An Intercomparison between Reanalysis and Dropsonde Observations of the Total Water Vapor Transport in Individual Atmospheric Rivers

BIN GUAN

Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, and Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

DUANE E. WALISER

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, and Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, California

F. MARTIN RALPH

Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

(Manuscript received 19 June 2017, in final form 22 November 2017)

ABSTRACT

A recent study presented nearly two decades of airborne atmospheric river (AR) observations and concluded that, on average, an individual AR transports $\sim 5 \times 10^8$ kg s^{-1} of water vapor. The study here compares those cases to ARs independently identified in reanalyses based on a refined algorithm that can detect less well-structured ARs, with the dual-purpose of validating reanalysis ARs against observations and evaluating dropsonde representativeness relative to reanalyses. The first comparison is based on 21 dropsonde-observed ARs in the northeastern Pacific and those closely matched, but not required to be exactly collocated, in ERA-Interim (MERRA-2), which indicates a mean error of −2% (−8%) in AR width and +3% (−1%) in total integrated water vapor transport (TIVT) and supports the effectiveness of the AR detection algorithm applied to the reanalyses. The second comparison is between the 21 dropsonde ARs and ~6000 ARs detected in ERA-Interim (MERRA-2) over the same domain, which indicates a mean difference of 5% (20%) in AR width and 5% (14%) in TIVT and suggests the limited number of dropsonde observations is a highly (reasonably) representative sampling of ARs in the northeastern Pacific. Sensitivities of the comparison to seasonal and geographical variations in AR width/TIVT are also examined. The results provide a case where dedicated observational efforts in specific regions corroborate with global reanalyses in better characterizing the geometry and strength of ARs regionally and globally. The results also illustrate that the reanalysis depiction of ARs can help inform the selection of locations for future observational and modeling efforts.

1. Introduction

Characterized by enhanced water vapor transport in long and narrow corridors in the lower troposphere, atmospheric rivers (ARs) play important roles in the global water cycle (Zhu and Newell 1998) and deliver precious freshwater to many arid/semiarid regions (Guan et al. 2010; Dettinger et al. 2011; Rutz and Steenburgh 2012), but they can also represent a significant hazard around the globe due to the associated extreme wind and precipitation (e.g., Waliser and Guan 2017). The phenomenology, characteristics, and processes of ARs have been the focus of a rapidly increasing volume of observing, modeling, theoretical, and application studies in the past decade (Ralph et al. 2017a). The development of AR science has led to the convergence of a formal definition of an AR recently included into the American Meteorological Society's Glossary of Meteorology (American Meteorological Society 2017).

Reflective of the overall definition but often emphasizing different aspects of the phenomenon, a number of

Corresponding author: Bin Guan, bin.guan@jpl.nasa.gov

DOI: 10.1175/JHM-D-17-0114.1

techniques have been previously developed for objective identification of ARs. For example, the technique based on the integrated water vapor (IWV) signature of ARs was developed associated with the availability of high-quality satellite retrievals of IWV over the northeastern Pacific (Ralph et al. 2004; Neiman et al. 2008; Wick et al. 2013). The technique based on point observations of IWV and surface wind was designed to best take advantage of high-resolution in situ observations from a limited number of coastal stations (Neiman et al. 2009; Ralph et al. 2013). More recent techniques often make use of integrated water vapor transport (IVT), a variable more directly related to orographic precipitation (Neiman et al. 2002; Neiman et al. 2009) and useful for understanding the role of ARs in the global water cycle (Zhu and Newell 1998).

A recent study by Ralph et al. (2017b) introduced the concept of total IVT (TIVT), which represents the total along-AR horizontal water vapor flux through a cross section perpendicular to an AR (analogous to the streamflow in a terrestrial river). The dataset used by Ralph et al. (2017b) had been collected using four types of research aircraft in a series of 37 research flights over nearly 20 years (Ralph et al. 2004, 2011, 2016; Neiman et al. 2014), of which 30 flights were conducted during the CalWater program of field studies. Ralph et al. (2017b) concluded that on average an individual AR transports horizontally roughly $5 \times 10^8\,\mathrm{kg\,s^{-1}}$ of water (as vapor), which is comparable to 27 times the discharge of water (as liquid) by the Mississippi River into the Gulf of Mexico. However, because the study had "only" 21 cases, a question remained as to how representative these cases are of ARs more globally.

The current study takes advantage of this unprecedented set of airborne observations by conducting an intercomparison of key AR characteristics calculated from the dropsonde-observed ARs and ARs in reanalysis products independently identified and measured based on different methodologies. For the dropsonde observations, the AR identification procedure is specific to the manner in which these observations were taken, that is, along transects across the ARs that attempt to sample the center part of ARs (Ralph et al. 2017b). For reanalyses, AR identification is based on applying a global AR detection algorithm that considers the 2D geometry of ARs (Guan and Waliser 2015, hereafter GW2015), providing a more complete spatiotemporal sampling than dropsondes but at the expense of some reliance on a model via data assimilation. The need for validating reanalysis-based AR characteristics against field observations and for understanding the representativeness of the available field observations motivates the current intercomparison study. Such two-way comparisons have not been conducted in previous AR studies. The two AR characteristics of interest here are AR width and TIVT across the AR width. Specifically, the objectives of the current study are 1) to validate the GW2015 global AR detection algorithm (with a few refinements to be detailed in section 2a) against dropsonde observations in terms of the above two AR characteristics; 2) to evaluate the representativeness of the two AR measures calculated from the limited number of dropsonde-observed ARs relative to all ARs detected by the GW2015 algorithm in reanalysis products; and 3) within this context, to examine seasonal and geographical variations in the two AR measures, first regionally, then globally.

2. Data and methodology

a. *Reanalyses and AR detection*

Global fields of specific humidity and vector winds are provided by two reanalysis products, namely, the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011) and the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. 2017). Comparison between the two products provides an estimate of the reanalysis uncertainty where appropriate. ERA-Interim data at 1.5° × 1.5° horizontal resolution are obtained at 17 pressure levels, and MERRA-2 data at 0.625° × 0.5° resolution are obtained at 21 pressure levels, between 1000 and 300 hPa. Six-hourly fields during the period of 1979–2016 (for ERA-Interim) or 1980–2016 (for MERRA-2) are used for the calculation of IVT and AR detection. Both ERA-Interim and the original version of MERRA compared well with aircraft observations of six northeastern Pacific ARs (Ralph et al. 2012) and 15 years of satellite observations of AR landfalls along the U.S. West Coast (Jackson et al. 2016) in terms of basic AR characteristics. Zonal and meridional components of IVT are calculated as

$$\mathrm{IVT}_x = -\frac{1}{g}\int uq\,dp \quad \text{and} \tag{1a}$$

$$\mathrm{IVT}_y = -\frac{1}{g}\int vq\,dp, \tag{1b}$$

where g is gravitational acceleration, u is zonal wind, v is meridional wind, q is specific humidity, p is pressure, and the integration is over all pressure levels between 1000 and 300 hPa.

An IVT-based AR detection algorithm was introduced in GW2015, which was designed for "tracking ARs globally as elongated targets" (tARget version 1).

The algorithm considers a combination of geometry and intensity thresholds in identifying the spatial distribution of ARs at any given time. The initial set of "objects" (i.e., contiguous area of connected pixels) is extracted based on the IVT magnitude threshold (the 85th percentile with a fixed lower limit of 100 kg m^{-1} s^{-1}) specific to each location and season. These objects are then filtered based on requirements on IVT direction (within 45° of the AR shape orientation and with an appreciable poleward component), length (>2000 km), and length/width ratio (>2), resulting in a defined set of ARs. Initial evaluation of the algorithm indicated reasonable agreement with Zhu and Newell (1998) in terms of AR fractional zonal circumference and fractional total meridional IVT and over ~90% agreement in detected AR landfall dates compared to other algorithms independently developed in three previous studies that focused on ARs in three regions with different climatologies.

A revised version of the algorithm, hereafter tARget version 2, is used in the current study, which contains refinements in three aspects. First, multiple, sequentially higher IVT percentile thresholds (i.e., 85th–95th percentiles), as opposed to only the 85th percentile, are used for AR detection. This is in view of the cases where an IVT object determined by the 85th percentile does not meet the AR requirements (e.g., too "fat" or otherwise less well structured), but a core region with higher IVT exists within the object that does meet the AR requirements by itself. A similar concept was used in Wick et al. (2013) in their AR detection algorithm based on IWV, where a series of IWV thresholds between 20 and 40 mm was used for AR detection. In implementing the multiple IVT thresholds, each map of IVT is first scanned with the 85th percentile threshold, and all the grid cells where ARs are detected are zeroed out from the map. The map is then rescanned with the 87.5th percentile threshold, and any new ARs detected are zeroed out from the current IVT map. The process is repeated five times until the 95th percentile threshold is applied. The use of multiple thresholds resulted in a 17% increase in the number of detected AR objects based on assessment with ERA-Interim. An example of a "new" AR, that is, detected here but not in GW2015, is shown in red in Fig. 1a.

Second, the ARs detected from the steps described above are further filtered for circular-shaped objects that are potentially tropical cyclones. The length/width ratio requirement (>2) implemented in the original version of the algorithm already filters out most of these round objects, but occasionally fails when an object is less regularly shaped. The new filter makes use of a robust characterization of the roundness of the object based on the "solidity" property (>0.8) combined with the ratio between the lengths of the minor and the major axes of a fitted ellipse (>0.8). This filter is applied to objects with centroids equatorward of 30°, where most tropical cyclones occur and where the distortion of an object shape due to the "lat–lon" map projection (commonly used by global geophysical datasets) is minimal. The filter leads to only a 0.2% reduction in the number of detected AR objects based on assessment with ERA-Interim. Note this roundness filter (and the length/width ratio requirement) does not aim to filter out all ARs potentially related to tropical cyclones, but only those well-shaped features that are most likely tropical cyclones themselves.

The third refinement to the algorithm is the identification of the AR transect that goes through the centroid of the AR in the direction perpendicular to the mean IVT averaged within the AR shape boundary (Fig. 1b, black). The AR transect width (hereafter, AR width for simplicity) and TIVT across the transect are also included in the output, with TIVT calculated as

$$\mathrm{TIVT} = \int \mathrm{IVT}_{\mathrm{transect}}\, dw, \tag{2}$$

where $\mathrm{IVT}_{\mathrm{transect}}$ is the component of IVT normal to the AR transect, and the integration is over all pixels that form the AR transect. The enhanced output facilitates comparisons with observations from dropsondes, which are often released along transects that go through the core section of ARs (Ralph et al. 2004, 2011; Neiman et al. 2014). Although not examined here, transects across other sectors of an AR (such as the landfall location) are also worthy of investigation to help understand variations along the AR axis.

b. *Dropsonde observations*

Dropsonde observations along 21 AR transects were reported in Ralph et al. (2017b). The observations were among those made during various campaigns in 1998–2016 that focused on ARs in the northeastern Pacific of importance to weather and hydrology in the western United States. Each AR was transected by at least nine dropsondes. The upper limit of the vertical profiles varies between dropsondes, which may affect the comparison of the calculated IVT values. However, the impact is expected to be small as the layer between the surface and 500 hPa, where the vast majority of water vapor is contained, is consistently sampled in all cases with no vertical gaps exceeding 50 hPa. To calculate the width and TIVT across the transects, the AR boundary is determined by the following conditions, which differ from those used in the tARget algorithm as described earlier: IVT magnitude greater than 250 kg m^{-1} s^{-1} at a minimum of three interior dropsonde locations and IVT magnitude less than 250 kg m^{-1} s^{-1} at both ends of the

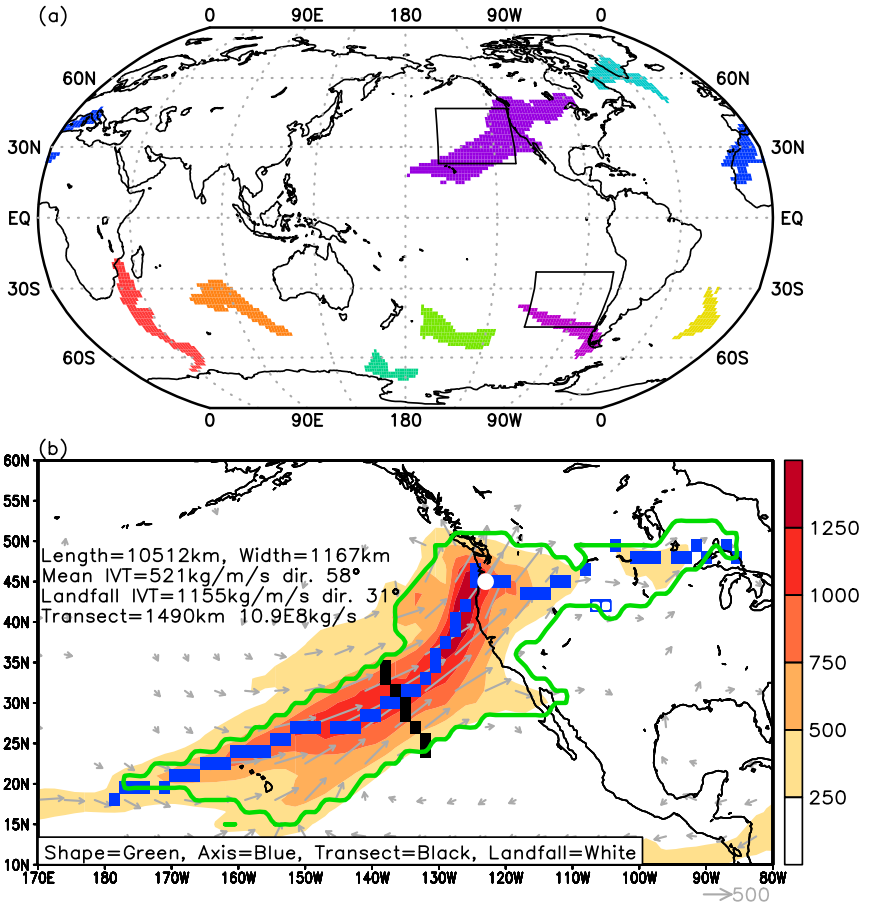

FIG. 1. (a) Example output from the AR detection algorithm, showing ARs detected at an arbitrary 6-h time step of the ERA-Interim reanalysis. Each color indicates a unique AR. The algorithm, tARget version 2, is a refinement of the original version introduced in GW2015; see text for details. The AR in red is "new" compared to those shown in Fig. 2 of GW2015 as a result of using multiple, sequentially higher IVT thresholds (i.e., 85th–95th percentiles) for AR detection. The two black boxes mark the domains used in subsequent analysis. (b) A specific AR from (a), showing the location of the AR transect (black), the width, and TIVT across the transect (numbers in the upper left of the panel), along with other key output from the algorithm (see legend). Also shown are IVT vectors ($kg\,m^{-1}\,s^{-1}$; gray arrows; not shown if the magnitude is smaller than $100\,kg\,m^{-1}\,s^{-1}$) and magnitudes (color shading) associated with the AR. The AR transect is defined as the great-circle arc that goes through the centroid of the AR in the direction perpendicular to the mean IVT averaged within the AR shape boundary (green).

transect. Some of the 21 cases do not strictly meet all of these criteria (usually because the transect did not go far enough equatorward to intersect the southern edge of the AR), but were nonetheless included in Ralph et al. (2017b) for the benefit of having more samples for the analysis, as are in the current study.

3. Results

a. Global AR frequency and precipitation

A detailed description of the global climatology of ARs, including frequency, geometry, intensity, seasonality, and climate modulations, was provided in GW2015. The discussion here focuses on the difference/similarity between the two versions of the tARget algorithm with respect to two quantities, namely, AR frequency and precipitation, which are among the most fundamental characteristics of ARs of relevance to their roles in the global water cycle. Daily precipitation from the 1° × 1° resolution Global Precipitation Climatology Project (GPCP) version 1.2 (Huffman et al. 2001), regridded to the ERA-Interim resolution, is used for this part of the analysis.

Shown in Fig. 2a is the global distribution of AR frequency during 1979–2016 based on ERA-Interim and tARget version 2. AR frequency is calculated at each grid cell as the percentage of time steps during which the

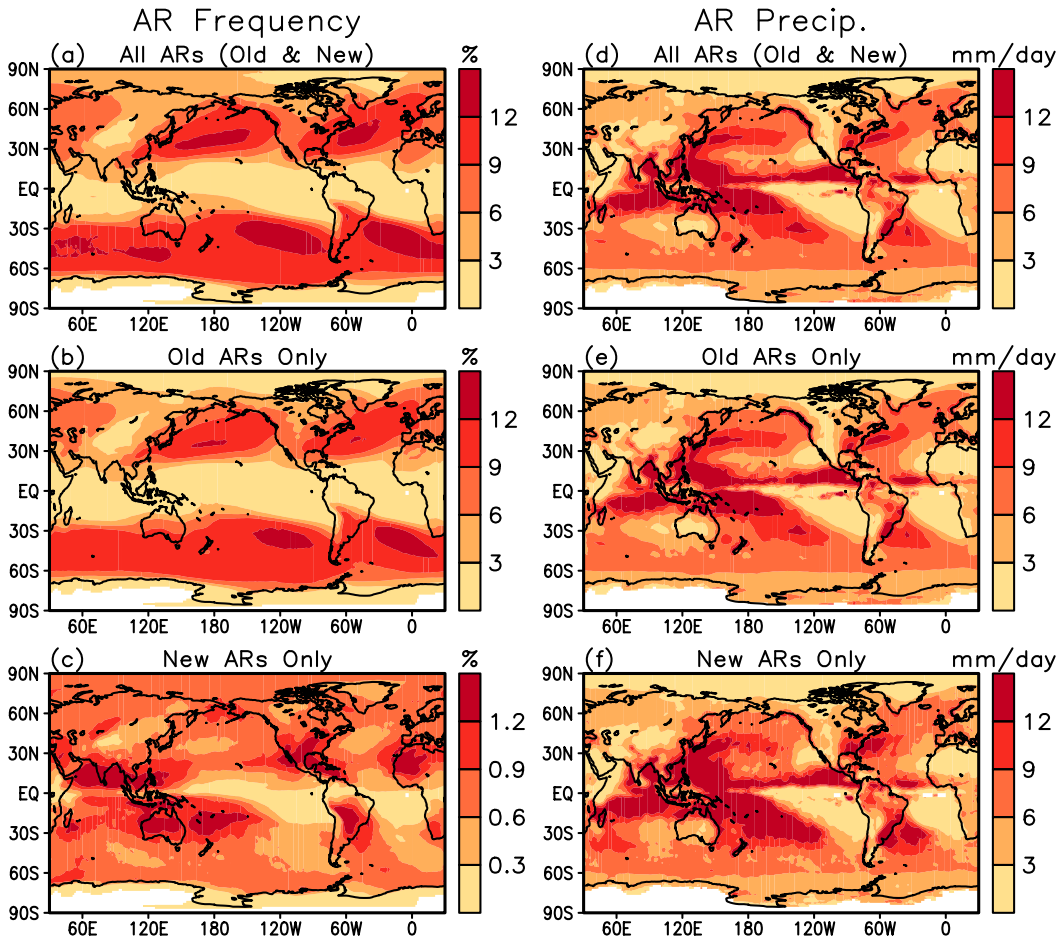

FIG. 2. (a)–(c) The percentage of AR time steps out of all time steps in ERA-Interim based on three groups: (a) all ARs detected using tARget version 2, (b) only those ARs detected using tARget version 1, and (c) only those ARs newly detected by tARget version 2. The percentage values in (b) and (c) add up to the percentage values in (a). Note the different scale in (c). (d)–(f) Mean AR precipitation intensity (mm day^{-1}) averaged over the three groups of ARs shown in (a)–(c). White shading indicates regions with no detected ARs during the analysis period. Precipitation data are from GPCP version 1.2.

grid cell is within any detected AR shape boundary. The overall pattern is characterized by enhanced AR frequencies in midlatitude ocean basins relative to inland and lower/higher-latitude regions, with multiple action centers located in the extratropical North Pacific/Atlantic, southeastern Pacific, South Atlantic, and southern Indian Ocean. The broad pattern is similar to Fig. 2b produced from tARget version 1, although AR frequency is notably increased based on the newer detection, especially around the various action centers. The increase in AR frequency is a result of using multiple IVT thresholds for AR detection, which overwhelms the tiny reduction in the number of ARs detected due to the filter for potential tropical cyclones, as also mentioned in section 2.

The net increase in AR frequency between the two versions of the detection is shown in Fig. 2c (the values are straight differences between Figs. 2a and 2b, not relative differences). In many regions outside the tropics, the increase is around 0.6%–0.9%, or 2–3 more AR days per year. Larger increases are seen in some tropical/subtropical regions, including southern North America into the Caribbean, central South America, northwest Africa, South/Southeast Asia, and Australia into French Polynesia. Most of these larger increases are located over land, suggesting an IVT threshold higher than the 85th percentile can be particularly effective for detecting ARs over land. No place shows a net decrease in AR frequency between the two versions.

Mean AR precipitation intensity averaged over each of the three groups of ARs in Figs. 2a–c is shown in Figs. 2d–f. At each grid cell, mean AR precipitation is independently calculated based on averaging over the time steps during which the grid cell is within the

detected AR shape boundary. Note that, unlike in the left column, the lower two panels in the right column are not additive as they represent intensities instead of summations. Largely similar to AR frequency, AR precipitation is enhanced in the extratropical ocean basins (Fig. 2d). In addition, large values are seen in the intertropical convergence zone, Indo-Pacific warm pool, and tropical North Atlantic, where ARs occur much less frequently compared to extratropical ocean basins (Fig. 2a). The structure of ARs in the tropical regions is only starting to be understood (GW2015; Yang et al. 2018), and AR precipitation in these regions could potentially include contributions from phenomena that have been differently named in the literature. For example, the three flood-producing ''tropical plume'' events in coastal northwest Africa analyzed in Knippertz and Martin (2005) are all detected as ARs in the current study, as they are in the regional AR archive independently developed by Brands et al. (2017; http://www.meteo.unican.es/es/atmospheric-rivers). Comparison between Figs. 2e and 2f suggests that the ''new'' ARs that were not detected by the original algorithm are as precipitating as the ''old'' ARs. The analysis in Fig. 2 suggests that the use of multiple IVT percentile thresholds, given the other requirements on geometry, is an effective way to detect additional ARs that have precipitation characteristics similar to other ARs but that can go undetected due to not meeting the geometry requirements when a single IVT percentile threshold is used.

b. *AR transects in northeastern Pacific*

The 21 dropsonde-observed AR transects (individually shown in Ralph et al. 2017b) correspond to 20 ERA-Interim reanalysis ARs based on searching the 6-h step of ERA-Interim that has an AR in the vicinity of the dropsonde transect and closest in time to the midpoint of the transect (Fig. 3). Two of the dropsonde transects (18 and 19) were close enough in time to correspond to the same reanalysis time step. In two cases, the observed AR is only detected by the revised algorithm (panels outlined in green), exemplifying the effectiveness of utilizing multiple IVT thresholds for AR detection. Various types of ARs are sampled, including landfalling versus offshore ARs, those centered in the subtropics versus midlatitudes, and long-and-narrow versus less-well-elongated ARs. Nineteen out of the 21 dropsonde transects have a matching reanalysis AR within ±3 h (i.e., within the temporal resolution of ERA-Interim). Two dropsonde transects have a matching reanalysis AR 12.6 and 9.7 h earlier, respectively. In 10 cases, the dropsonde and reanalysis transects are relatively close to each other in terms of location. The location differences in other cases are partly attributable to the different data sources and methods used for AR detection, and the timing difference between dropsonde and reanalysis, during which an AR can propagate in space and/or change in structure.

Comparison of AR widths based on dropsonde observations versus ERA-Interim reanalysis is shown in the scatterplot in Fig. 4a. Each circle represents one pair of dropsonde and reanalysis transects as shown in Fig. 3. Overall, the two sets of AR widths are well correlated, with a correlation coefficient of 0.60 ($p = 0.004$). The circles are reasonably distributed on both sides of the 1:1 line (black), indicating no systematic biases in the reanalysis-based results relative to dropsonde observations. The blue and red error bars indicate one standard deviation below/above the mean (where the two error bars cross). The reanalysis- and dropsonde-based results have very close mean values (see also the two dashed lines in Fig. 4c), although the former has a slightly larger standard deviation. A similar comparison for TIVT is shown in Fig. 4b. The agreement there is even better than in the case of AR width, with a stronger correlation coefficient of 0.80 ($p < 0.001$) and very close mean values and standard deviations. The better agreement in the case of TIVT is likely because the integrated value is dominated by stronger IVT values toward the center of the AR and therefore is not sensitive to small errors in the width of the identified AR transect. The three cases with the best space–time coincidence between dropsonde and reanalysis AR transects (green circles in Figs. 4a,b; numbers 5, 8, and 12 in Fig. 3) are among the cases, but not the only cases, with the best agreement in AR width and TIVT (i.e., circles located close to the 1:1 line), suggesting the comparison between the two sets of ARs is affected by, but not solely dependent on, the space–time coincidence. This is because the dropsonde observation and the reanalysis both independently aim to transect the core part of the AR based on what they respectively find to be where the core part is. In this sense, the manner of the comparison here is somewhat analogous to object-based forecast verification (Gilleland et al. 2009), where the lack of space–time coincidence between the forecast and target objects is not unduly penalized as long as the overall structure (e.g., shape, size) of the target system is reasonably captured. As such, the error statistics reported in this study should not be interpreted as reanalysis errors in the traditional sense, but more appropriately as errors in the AR ''objects'' tied to both the raw reanalysis data and the associated AR detection algorithm. Figures 4a and 4b indicate the AR detection algorithm, specifically the identification of AR transects, validates remarkably

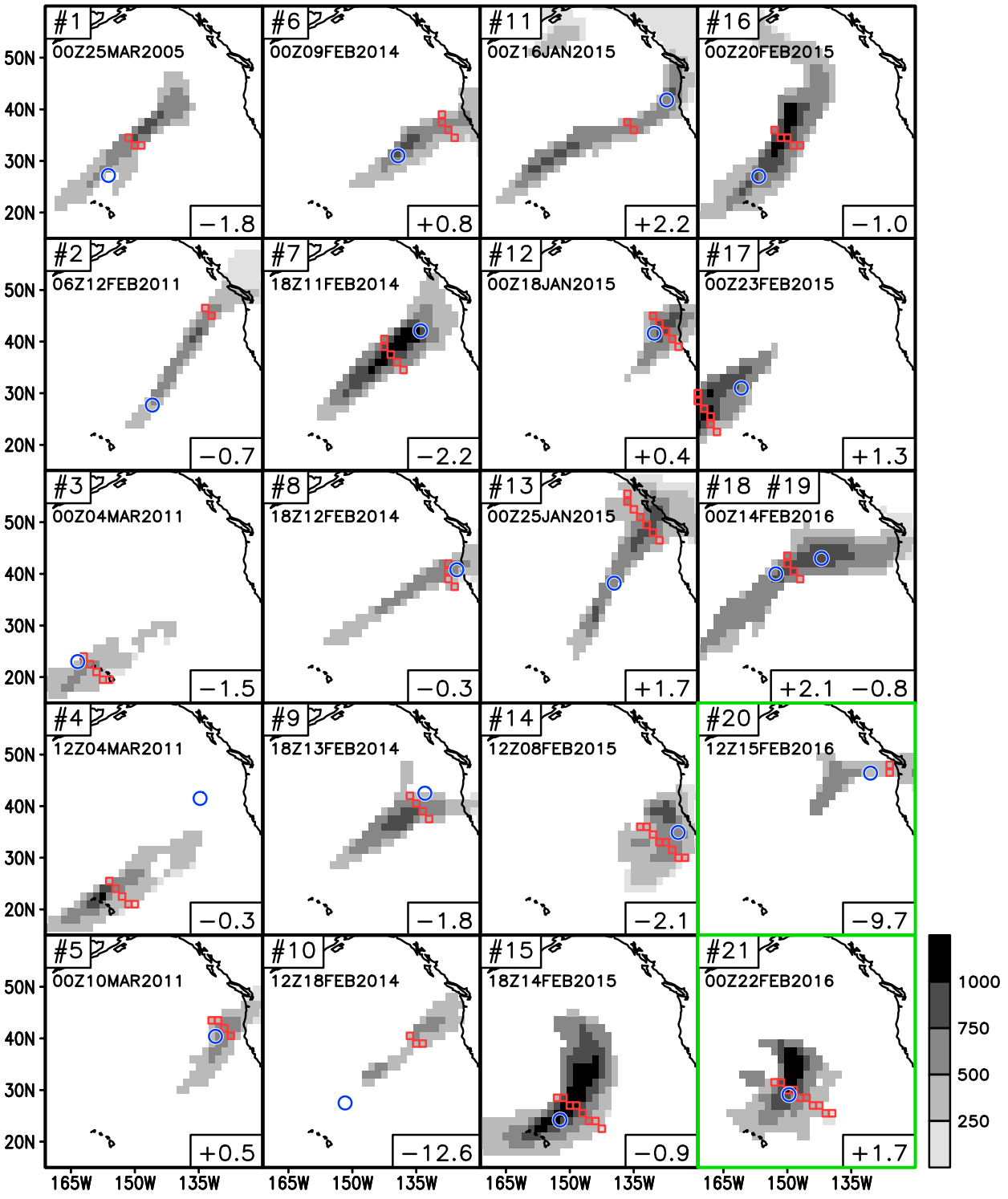

FIG. 3. The 20 ERA-Interim reanalysis ARs corresponding to the 21 dropsonde-observed AR transects. The numbering of the transect indicated in the top left of each panel is as in Ralph et al. (2017b). Shown in each case is the 6-h step (also indicated in the top left of each panel) of the reanalysis AR (IVT magnitude in shading; kg m^{-1} s^{-1}; transect in red) closest in time to the midpoint of the dropsonde transect (blue). Two of the 21 dropsonde transects (18 and 19) correspond to the same 6-h step of the reanalysis AR (third panel in the last column). Also indicated in each panel is the time difference (h) between the closest 6-h step of the reanalysis and the midpoint of the dropsonde transect (as a representation of the observation time), with positive/negative signs representing whether the reanalysis lags/leads the observation. Panels outlined in green show "new" ARs detected by tARget version 2 compared to the original version.

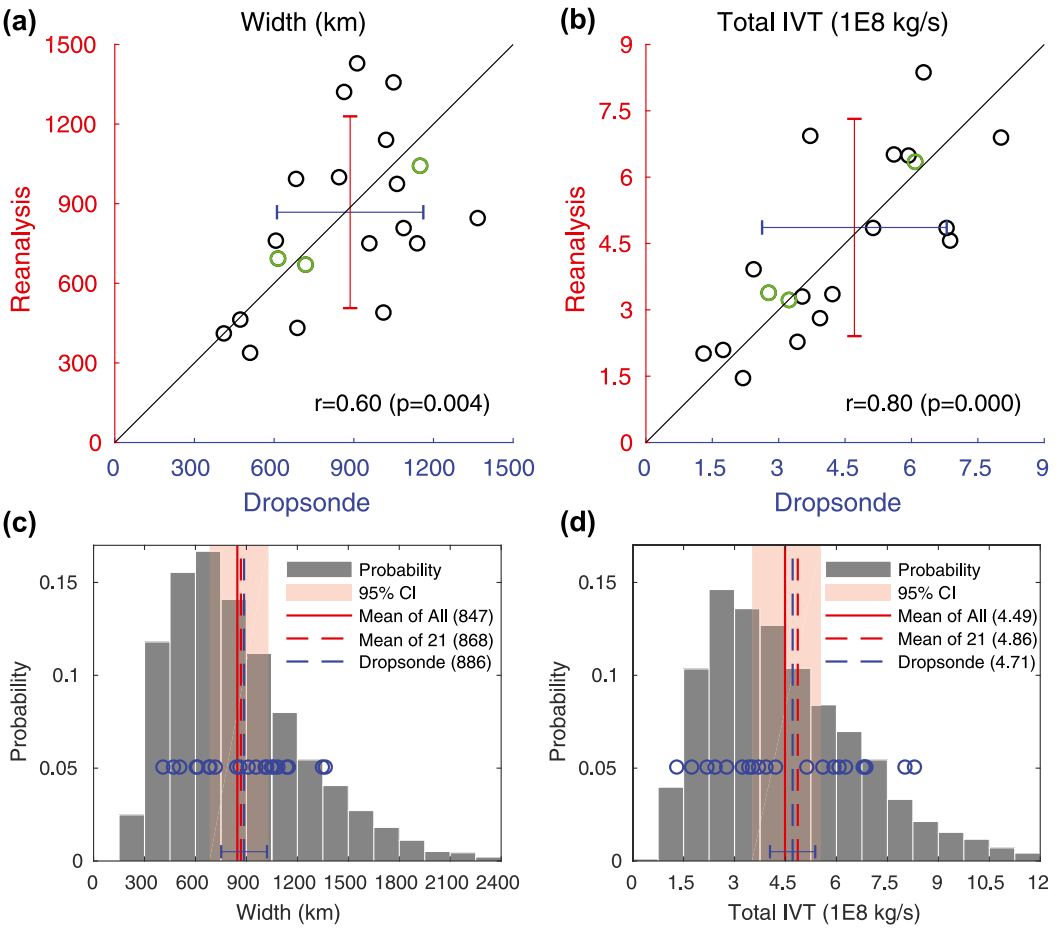

FIG. 4. (a) Scatterplot showing AR widths (km) based on dropsonde observations vs ERA-Interim reanalysis; see Fig. 3 for the time and location of each transect. The three cases with the best space–time coincidence between dropsonde and reanalysis AR transects (5, 8, and 12 in Fig. 3) are shown in green. The 1:1 line is in black. The correlation coefficient between the two sets of AR widths is shown, along with the p value. Blue and red error bars indicate one standard deviation below/above the mean (where the two error bars cross). (b) As in (a), but for TIVT (10^8 kg s^{-1}) across AR transects. (c) Histogram of AR widths based on all ARs detected in ERA-Interim over the northeastern Pacific (AR centroids within 23°–46.4°N, 163.4°–124.6°W) from 15 Jan to 25 Mar of 1979–2016 (gray bars). Also shown are the mean AR width (km) based on all reanalysis ARs that contributed to the histogram (red solid), the subset of the reanalysis ARs that corresponds to the 21 dropsonde transects (red dashed), and the observed value based on the 21 dropsonde transects as reported in Ralph et al. (2017b) (blue dashed line for the mean and blue circles for individual transects). The mean AR width value is also indicated in the figure legend for each sample. Red shading indicates the 95% confidence interval of the mean reanalysis AR width for a random 21-member sample drawn from the pool of reanalysis ARs based on 10 000 iterations. The error bar centered on the blue dashed line indicates the 95% confidence interval of the difference between the blue and red dashed lines based on a two-tailed, paired t-test. (d) As in (c), but for TIVT (10^8 kg s^{-1}) across AR transects.

well against dropsonde observations, with a relative error of only −2% (+3%) for mean width (TIVT).

The results above are put into context by further comparing statistics based on the 21 AR transects versus all ARs in ERA-Interim that were detected in the northeastern Pacific (AR centroids within 23°–46.4°N, 163.4°–124.6°W) from 15 January to 25 March of 1979–2016. The domain and the combination of days are selected to be just enough to encompass the location and time of the 21 AR transects analyzed earlier to facilitate comparisons. A total of 5636 reanalysis ARs are selected, based on which the probability distributions of AR widths and TIVT are obtained. Both AR width and TIVT are characteristic of a lognormal distribution with a longer right tail (Figs. 4c,d, gray). The mean values based on the entire population (red solid) compare well with the mean values based on only the subset of reanalysis ARs that correspond to 21 dropsonde ARs

(red dashed). The red shading shows the sampling variations in mean AR width and TIVT based on randomly selecting 21 ARs from the pool of all reanalysis ARs within the domain and time period, repeating for 10 000 times, and finding the 2.5nd and 97.5th percentiles based on the empirical distributions obtained for mean AR width and TIVT. The result that the red dashed lines are close to the red solid line and well within the red shading for both AR width and TIVT supports that the 21-member subset is a highly representative sampling of the entire population of ARs within the given domain and time period.

For AR width, the dropsonde-based mean value (Fig. 4c, blue dashed) is only slightly larger than the reanalysis counterpart (red dashed), as already seen in Fig. 4a. The error bar centered on the blue dashed line indicates the 95% confidence interval of the difference between the blue and red dashed lines for doing a two-tailed, paired *t* test. A paired *t* test is used since the two samples intend to characterize the same set of ARs. That the red dashed line is located well within the error bar suggests the difference between the two dashed lines is highly insignificant statistically. Moreover, the dropsonde-based mean value (blue dashed) well reflects the mean value based on a total of 5636 reanalysis AR transects (red solid), with a relative difference of 5%. Similar results can be seen for TIVT, also with a 5% relative difference. A reasonable range of AR width and TIVT values were sampled by individual dropsonde transects (blue circles) compared to the reanalysis-based full distributions (gray bars). Specifically, the minimum and maximum dropsonde-sampled AR widths (TIVT) are within the bottom 7% and top 11% (bottom 3% and top 6%) of the full distribution. Figures 4c and 4d suggest that the 21 dropsonde AR transects, although a relatively small sample, were able to characterize the mean width and TIVT very well compared to the much more spatiotemporally complete sampling by the ERA-Interim reanalysis.

Similar comparisons are done between dropsondes and MERRA-2. As in the case of ERA-Interim, the 21 dropsonde ARs are matched to 20 MERRA-2 ARs due to two temporally close dropsonde ARs matched to one reanalysis time step (Fig. 5). Nineteen of the MERRA-2 ARs are identified at the same 6-hourly time steps as the ERA-Interim ARs examined earlier. In the only one exception (number 4), the MERRA-2 AR is identified later than the corresponding ERA-Interim AR by 6 h. Eighteen out of the 21 dropsonde transects have a matching MERRA-2 AR within ±3 h (i.e., within the temporal resolution of MERRA-2). The general good agreement with dropsondes discussed earlier for ERA-Interim is also seen for MERRA-2 (Figs. 6a,b), with a correlation of 0.61 ($p = 0.003$) for AR width, 0.85 for TIVT ($p < 0.001$), and a relative error of −8% for mean width and −1% for mean TIVT.

The somewhat large (but statistically insignificant at the 95% level as indicated by the location of the red dashed line relative to the blue error bar in Fig. 6c) negative bias in MERRA-2 AR width relative to dropsondes (Fig. 6a) is consistent with the overall smaller AR width in the full distribution compared to ERA-Interim (comparing Fig. 6c to Fig. 4c, red solid). In general, MERRA-2 has a larger fraction of ARs with smaller width, and a smaller fraction of ARs with larger width (comparing Fig. 6c to Fig. 4c, gray). Specifically, MERRA-2 contains a considerable fraction of ARs with width smaller than 300 km, whereas such extremely narrow ARs are much rarer in ERA-Interim. This difference in the distributions may be contributed by the same observed ARs being detected in both products except with smaller widths in MERRA-2, but also by some of the narrow ARs that are only detected in MERRA-2. The overall smaller AR width in MERRA-2 relative to ERA-Interim has been found in Guan and Waliser (2017) on a global basis, where it was suggested to be associated with the native resolution of the dynamical models based on a systematic examination of 24 global weather/climate models and three reanalysis products—all with data archived on a common grid. In this regard, smaller AR width in MERRA-2 is consistent with its native resolution (0.625° × 0.5°) being slightly finer than that of ERA-Interim (T255, ~79 km). Unlike AR width, TIVT in MERRA-2 shows little bias relative to dropsondes (Fig. 6b), conceivably because of the same reason as discussed earlier, that TIVT is dominated by stronger IVT values toward the center of the AR and therefore is not sensitive to reasonable errors in the width of the identified AR transect.

The representativeness of the 21 dropsonde ARs relative to the total of 6077 MERRA-2 ARs from 15 January to 25 March of 1980–2016 is then assessed. Although the relative differences (20% for AR width and 14% for TIVT) are notably larger than in the case where ERA-Interim is used as the reference, the significance tests (that the blue dashed line is located within the red shading in Figs. 6c,d) still support the 21 dropsonde transects to be consistent with a reasonable—albeit not highly representative—sample from the entire population of MERRA-2 ARs.

c. *Sensitivity to analysis domain and period*

As mentioned earlier, the domain used for calculating the reanalysis-based AR statistics was intended to be representative of the location of the 21 dropsonde AR transects in the northeastern Pacific. Specifically, the

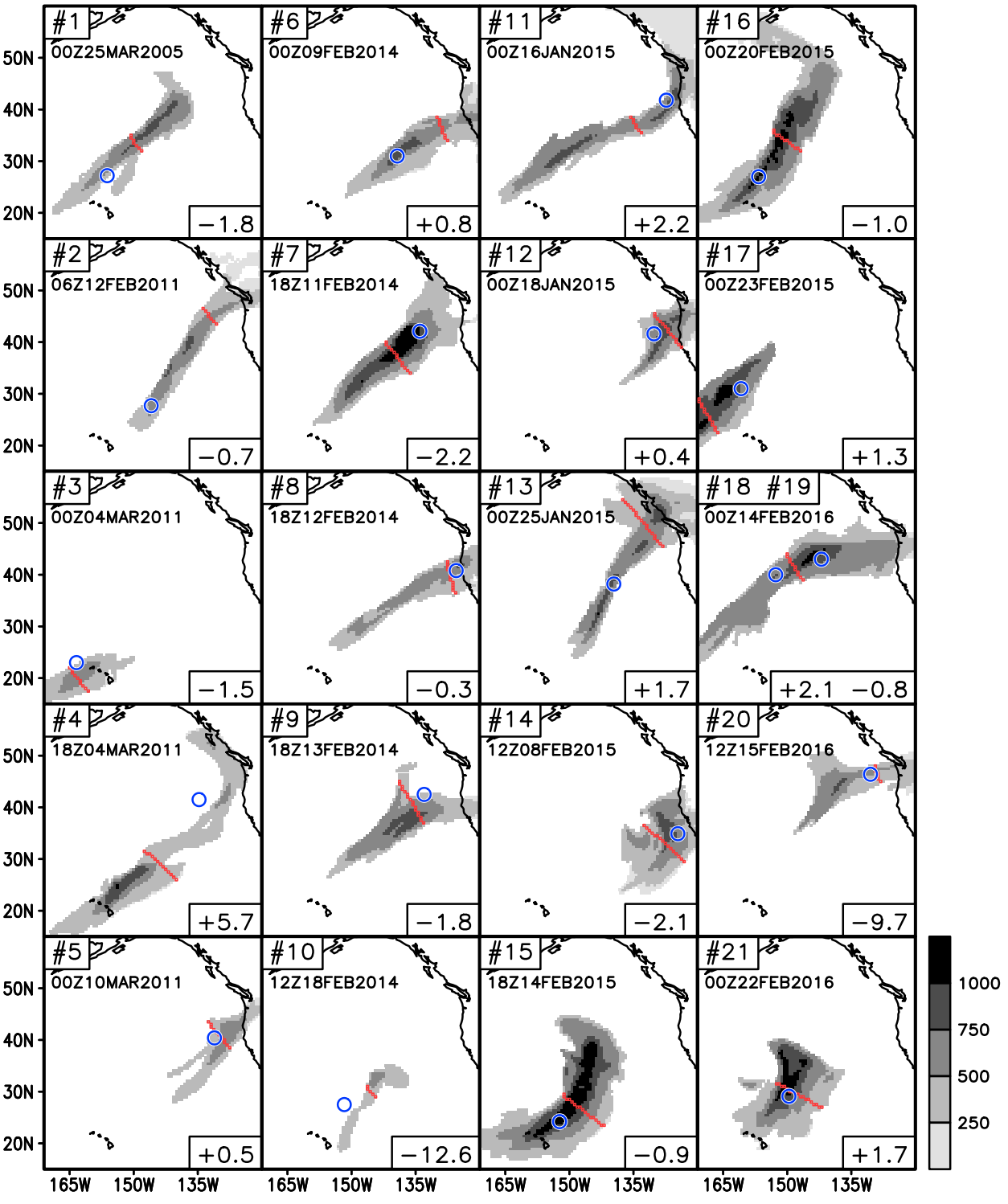

FIG. 5. As in Fig. 3, but with the reanalysis being MERRA-2.

domain used was of just enough size to encompass the centroids of the observed ARs. The location of the analysis domain is now perturbed to understand the sensitivity of the AR statistics to the exact location of the selected domain, hence the robustness of the reanalysis-based statistics presented earlier. In doing so, the domain used in the ''control'' analysis (23°–46.4°N, 163.4°–124.6°W) is shifted in each of the four directions

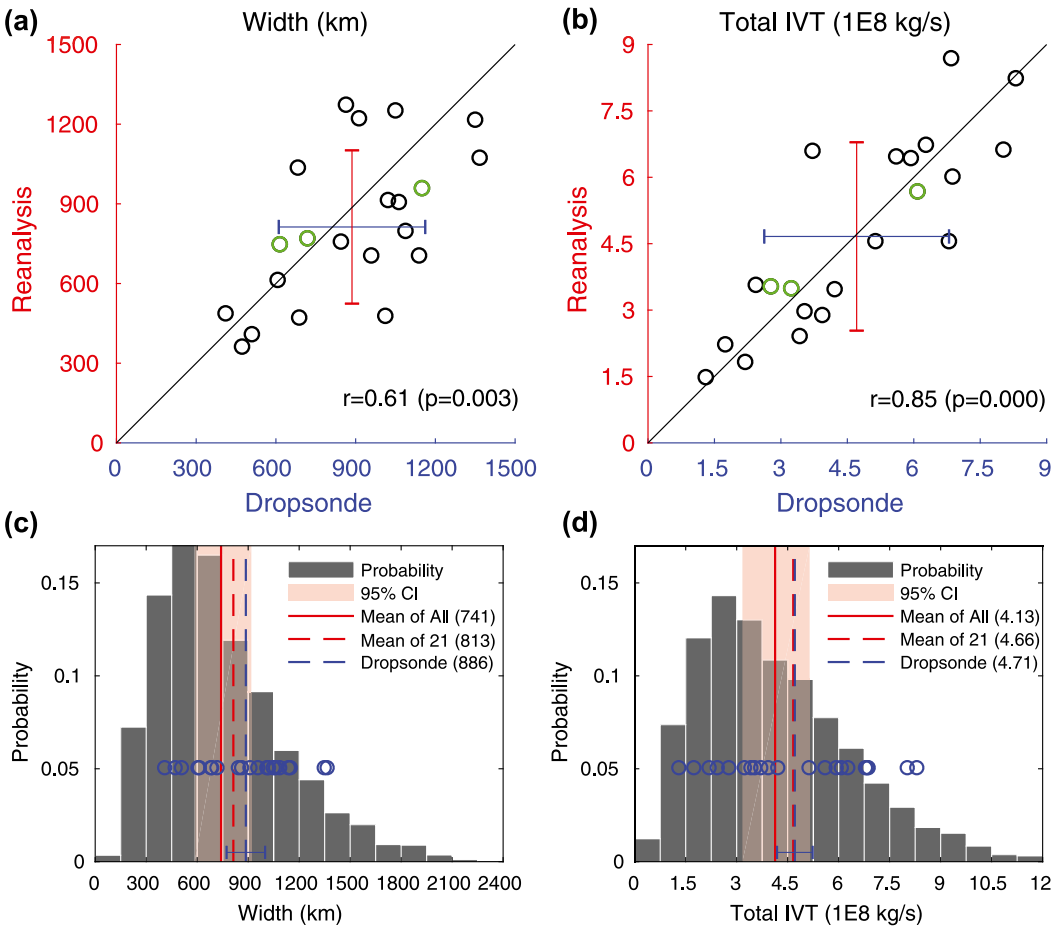

FIG. 6. As in Fig. 4, but with the reanalysis being MERRA-2.

(east/west/north/south) by 6°. The number of degrees to shift is selected to be just large enough to illustrate the sensitivity, but not so large as to change the overall geographical characteristic of the domain. The probability distributions and the means of AR width and TIVT are then recalculated, shown in Figs. 7a–d. The following discussion applies to both ERA-Interim and MERRA-2, except that mean AR width and TIVT are consistently smaller in MERRA-2 as noted earlier. The analysis suggests that AR width has little sensitivity to small variations in the east/west location of the analysis domain, but some sensitivity when the domain is shifted in the north/south direction (Figs. 7a,c). Specifically, northward (southward) shifting of the domain leads to a shifting of the probability distribution toward the left (right), hence a reduction (increase) in the mean AR width. Unlike the directional contrast in the case of width, TIVT exhibits comparable, and relatively small, sensitivities to both north/south and east/west shifting of the analysis domain (Figs. 7b,d). Similar sensitivity analysis is done based on breaking the period used in the control (1979–2016) into two subperiods of equal length, with the latter subperiod assimilating substantially more satellite observations in the reanalyses (Dee et al. 2011; Gelaro et al. 2017). Mean AR width is somewhat smaller in the second half-period than in the first half-period, whereas little sensitivity to analysis period is seen in the case of TIVT (Figs. 7e,f). The sensitivity analysis here indicates that, out of the six cases considered, the largest sensitivity is associated with AR width with respect to latitudinal location of the analysis domain. In all cases examined, the sensitivity is not unduly large to prevent meaningful comparison of the AR width and TIVT statistics presented in the previous subsection.

d. Seasonal and geographical variations

The main analysis above focused on ARs in the northeastern Pacific and during the winter season. Seasonal and geographical dependence of the results are now examined for each of the four seasons and for two representative domains. The northeastern Pacific domain is the same one used in the main analysis, and the

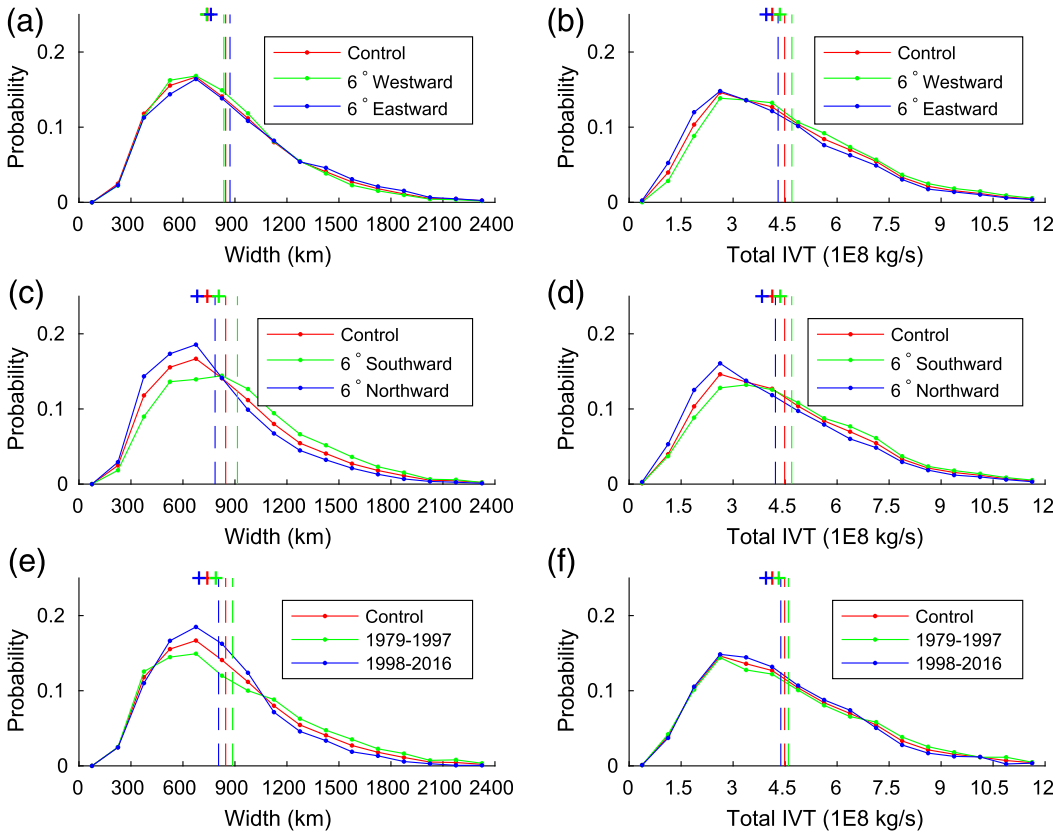

FIG. 7. (a),(b) Sensitivity of (left) AR width (km) and (right) TIVT (10^8 kg s^{-1}) to zonal location of the analysis domain based on shifting the domain used in Figs. 4 and 6, the control, by 6° to the east/west. Curves and vertical lines show the histograms and the means, respectively, based on ERA-Interim. The plus signs indicate the means based on MERRA-2. (c),(d) As in (a) and (b), but for sensitivity to meridional location. (e),(f) As in (a) and (b), but for sensitivity to analysis period based on breaking the period used in the control into two subperiods with equal length.

southeastern Pacific domain is of the same size and is placed offshore of South America, where ARs frequently make landfalls, to mimic its northern counterpart (see the two boxes in Fig. 1a). ARs in both regions have been shown to be hydrologically important, with enhanced activities during their respective winter seasons (e.g., Dettinger et al. 2011; Viale and Nuñez 2011).

In both regions, AR width and TIVT are characteristic of a lognormal distribution with a longer right tail, with seasonal and geographical variations in the shape of the distribution and the associated mean value (Fig. 8). For the northeastern Pacific, AR width tends to have the largest (smallest) values in winter (summer), and intermediate values in spring and fall (Fig. 8a). Seasonal contrast in TIVT is largely between the two extended cold and warm seasons, with notably larger TIVT values in fall and winter (Fig. 8b). The contrasts are consistent with the seasonal variations in AR frequency and intensity in this region (GW2015). Seasonal variations are in general smaller in the southeastern Pacific for both AR width and TIVT (Figs. 8c,d). Relatively weak seasonality in AR width and TIVT in the southeastern Pacific is consistent with weaker seasonality in AR frequency and intensity in this region compared to the northeastern Pacific (GW2015). The discussion above applies to both ERA-Interim and MERRA-2, except that mean AR width and TIVT are consistently smaller in MERRA-2 in the two regions examined, as noted earlier for northeastern Pacific.

A similar examination of the seasonal and geographical variations in AR width and TIVT is done globally and shown in Fig. 9 (ERA-Interim) and Fig. 10 (MERRA-2). At each grid cell, the value represents an average over the ARs that have centroids located at that grid cell, calculated separately for the four seasons. To facilitate comparison, a 1.5° × 1.5° grid is used for both ERA-Interim and MERRA-2 for binning the AR centroids. The following discussion applies to both products, except that mean AR width and TIVT are consistently smaller in MERRA-2 compared to ERA-Interim across

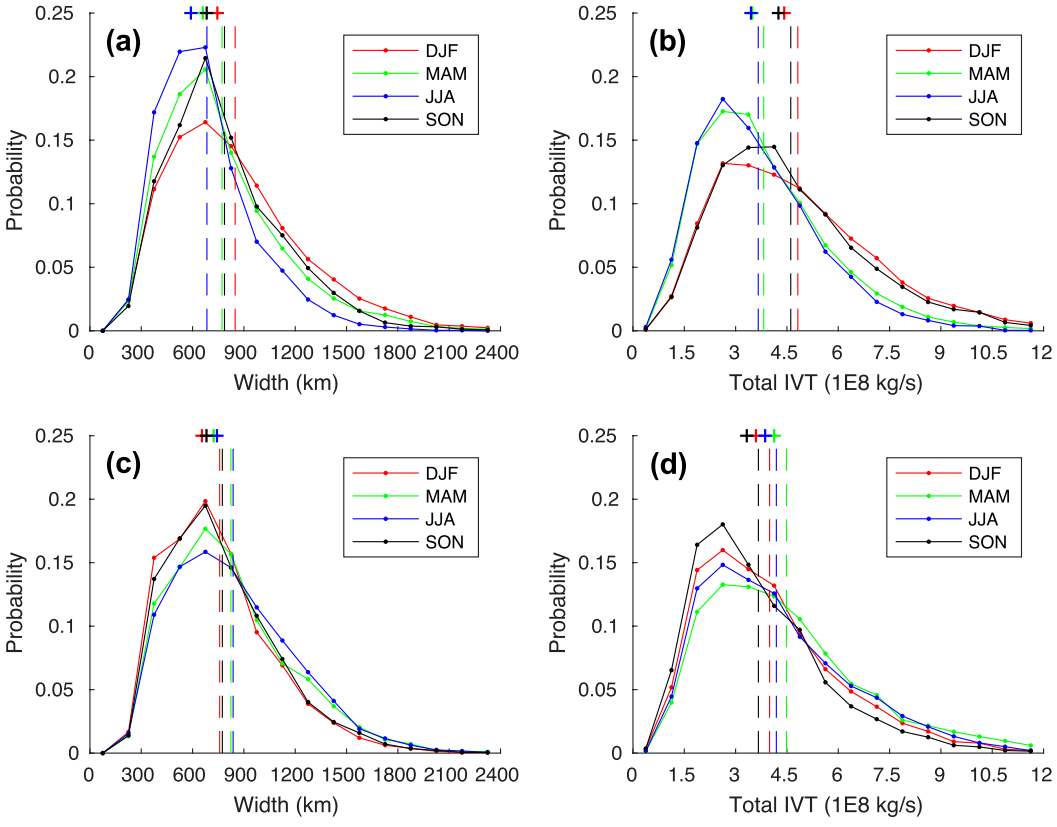

FIG. 8. Seasonal (see legend) and geographical variations in AR width (km) and TIVT (10^8 kg s^{-1}): (a), (b) northeastern Pacific and (c),(d) southeastern Pacific. The two domains are outlined in Fig. 1a. Curves and vertical lines show the histograms and the means, respectively, based on ERA-Interim. The plus signs indicate the means based on MERRA-2.

the globe, as noted earlier for two eastern Pacific regions. AR centroids are rarely located in the tropical and polar areas and some continental interiors, as indicated by the white shading. In general, the largest AR width tends to occur in the subtropical ocean basins. Large AR width also occurs in some land areas during specific seasons, including in southwestern North America, northwestern Africa, and Australia during December–February, and in South Asia during July–August. The largest TIVT also tends to occur in the subtropics, but in regions somewhat different from the case of AR width. Specifically, the largest TIVTs are seen in the western/central North Pacific and western North/South Atlantic in December–February and near the south/east coast of Asia and in the South Pacific during July–August.

4. Conclusions

A unique intercomparison is conducted between dropsonde-observed ARs and ARs in reanalysis products independently identified and measured based on different methodologies. The dropsonde observations are along 21 AR transects over the northeastern Pacific (Ralph et al. 2017b). The two reanalysis products considered (ERA-Interim and MERRA-2) each contain ~6000 ARs detected during 1979–2016 (for ERA-Interim) or 1980–2016 (for MERRA-2) over the domain and a combination of days that encompass the dropsonde observations. AR identification is based on a fixed IVT threshold of 250 kg m^{-1} s^{-1} for the dropsonde observations and seasonally and geographically dependent IVT thresholds, that is, the 85th–95th percentiles, for the reanalyses. The global AR detection algorithm applied to the reanalyses, tARget version 2, is a refinement of the original version introduced in GW2015. The new version implements multiple IVT percentile thresholds for more effective AR detection (e.g., Fig. 3, panels with green outline), and a roundness filter for potential tropical cyclones, which respectively increased (decreased) the number of detected AR objects by 17% (0.2%) based on assessment with ERA-Interim.

The characteristics compared here are the AR width and TIVT, a parameter introduced recently (Ralph et al. 2017b)

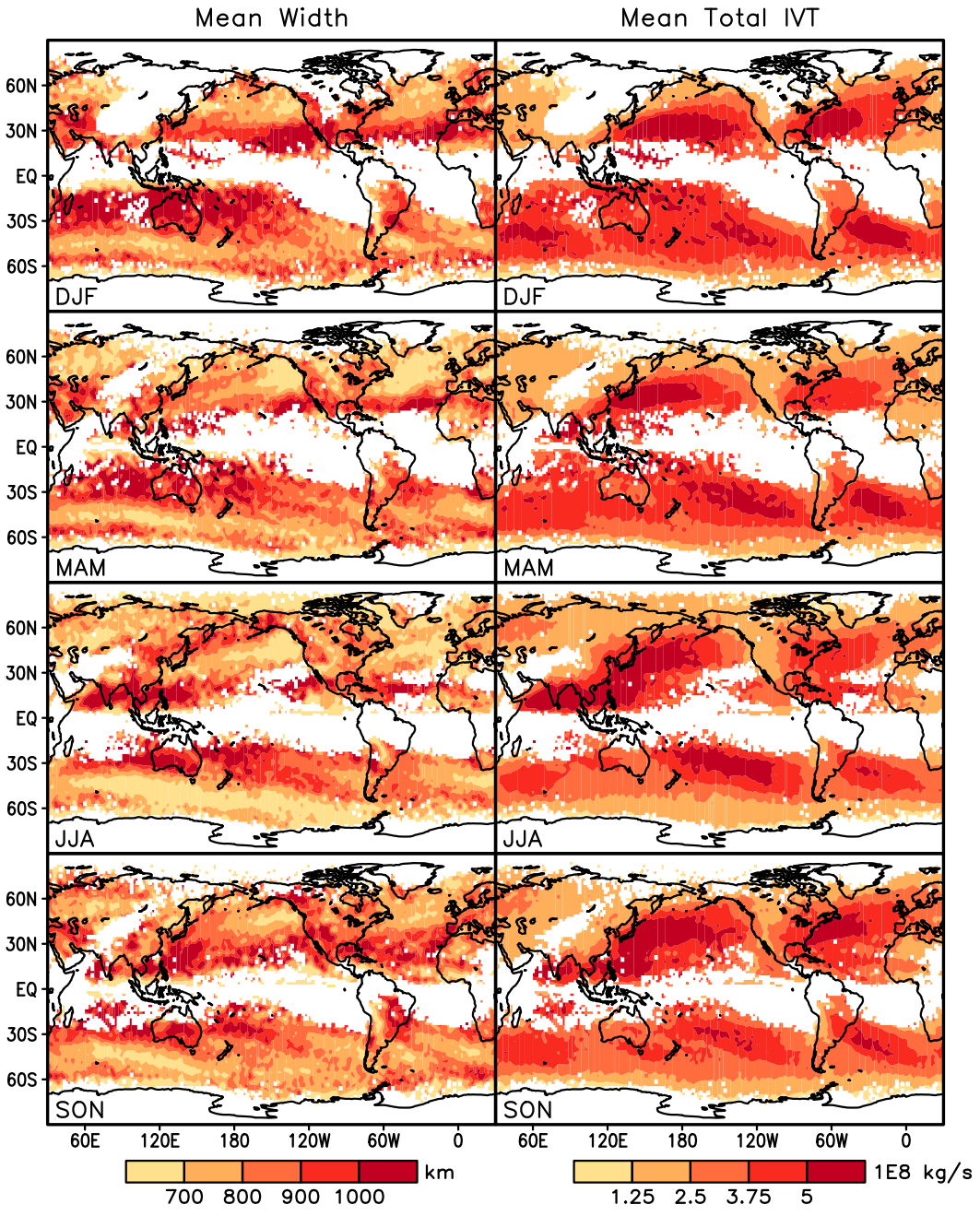

FIG. 9. Global distribution of mean AR width (km) and TIVT (10^8 kg s^{-1}) for each of the four seasons based on ERA-Interim. At each grid cell, the value is based on averaging over the ARs with centroids located at that grid cell.

representing the total horizontal transport of water vapor along an AR through a cross section normal to the AR. These parameters are calculated from the dropsonde-observed and reanalysis ARs and are intercompared. The first comparison is based solely on the 21 dropsonde ARs and 20 reanalysis ARs that match the overall time and location of the dropsonde transects (two of the dropsonde ARs correspond to the same reanalysis AR), with the purpose of validating ARs detected in the reanalyses against field observations. Of the 21 dropsonde-observed ARs, 19 (18) have a corresponding AR detected in ERA-Interim (MERRA-2) within ±3 h (Figs. 3, 5). Comparison of ERA-Interim (MERRA-2) ARs relative to dropsonde ARs indicates a

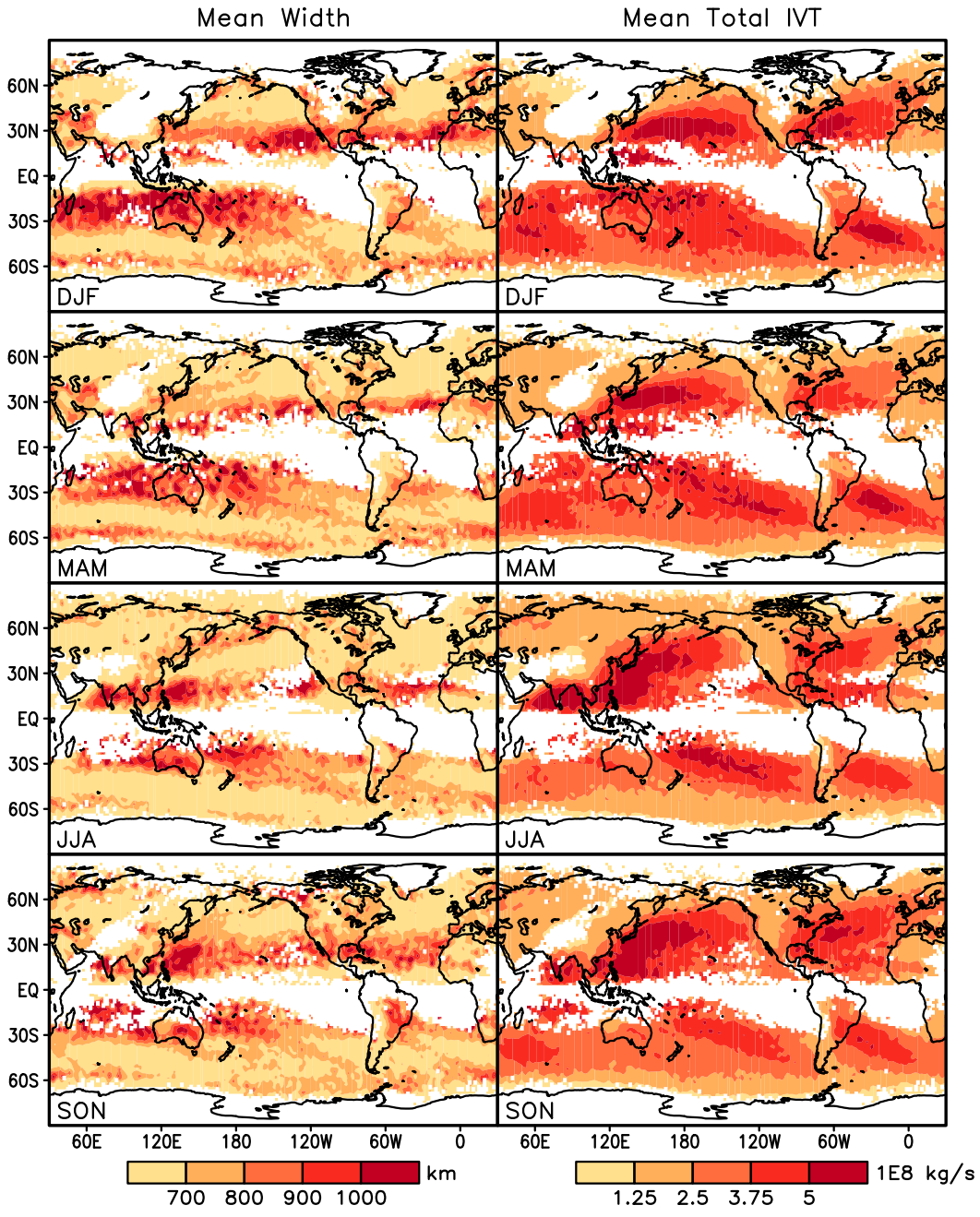

FIG. 10. As in Fig. 9, but based on MERRA-2. The coarser ERA-Interim grid is used for binning the AR centroids to facilitate comparison.

relative error of −2% (−8%) for mean AR width, and +3% (−1%) for mean TIVT (Figs. 4a, 4b, 6a, 6b), which supports the effectiveness of the global AR detection algorithm in detecting winter ARs over the northeastern Pacific and the reanalyses for their representation of the AR width and TIVT values.

The other comparison is between the 21 dropsonde ARs and the total of ~6000 ARs in each reanalysis product that represent a more complete spatiotemporal sampling of winter ARs over the northeastern Pacific, with the purpose of evaluating the representativeness of AR statistics derived from the limited number of dropsonde observations relative to the reanalyses. The comparison indicates a 5% relative difference for dropsonde-observed mean AR width and TIVT over the 21 transects relative to the total of ~6000 ERA-Interim

ARs (Figs. 4c,d), suggesting that the limited number of dropsonde observations over the northeastern Pacific is a highly representative sampling of the entire population of ARs depicted by ERA-Interim. The relative difference is larger (20% for AR width and 14% for TIVT) but statistically insignificant when comparing the 21 dropsonde ARs to the ~6000 MERRA-2 ARs, due to the presence of considerably more ARs with smaller width and TIVT in MERRA-2 compared to ERA-Interim. Sensitivity analysis supports the robustness of the above comparisons, especially for TIVT, which exhibits smaller sensitivity than AR width to perturbations in analysis domain and period.

AR width and TIVT have considerable seasonal and geographical variations. For both variables, the largest values tend to occur in the subtropics, although focused in somewhat different locations. Large AR width can also occur in some land areas during specific seasons—not the case for TIVT, for which the largest values seldom occur over land. Close examination of two offshore boxes over the northeastern Pacific and southeastern Pacific suggests that ARs tend to have larger width and TIVT during the region's respective cold season, and that between the two regions, larger seasonal contrasts exist in the northeastern Pacific.

The results presented in this paper provide a case where dedicated observational efforts in specific regions corroborate with global reanalysis products in better characterizing the geometry and strength of ARs regionally and over the globe, highlighting the value of dedicated observational efforts for benchmarking AR characteristics derived from more spatiotemporally complete data products—albeit with their own shortcomings, such as those from satellites, reanalysis, and model simulations. The global distributions of AR width and TIVT could better inform the selection of the locations for dedicated observational and modeling efforts in the future.

Acknowledgments. Dropsonde data used here were collected during several field campaigns over nearly 20 years involving many scientists, engineers, air crews, project managers, program managers, and others. These include individuals from NOAA, NASA, the Air Force, and elsewhere. Without their efforts, these data would not be available for this study. The WISPAR Global Hawk dropsonde data were quality controlled by J. Wang and colleagues at the NCAR Earth Observing Laboratory, where the data are maintained. NCAR is sponsored by NSF. The ERA-Interim reanalysis is obtained online from http://apps.ecmwf.int/datasets/data/interim-full-daily/, and MERRA-2 from https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/. This research was supported by the NASA Energy and Water cycle Study (NEWS) program, the California Department of Water Resources, and the Office of Naval Research. D.E.W.'s contribution to this study was carried out on behalf of the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

REFERENCES

American Meteorological Society, 2017: Atmospheric river. Glossary of Meteorology, http://glossary.ametsoc.org/wiki/Atmospheric_river.

Brands, S., J. M. Gutiérrez, and D. San-Martín, 2017: Twentieth-century atmospheric river activity along the west coasts of Europe and North America: Algorithm formulation, reanalysis uncertainty and links to atmospheric circulation patterns. *Climate Dyn.*, **48**, 2771–2795, https://doi.org/10.1007/s00382-016-3095-6.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

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

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1.

Guan, B., and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *J. Geophys. Res. Atmos.*, **120**, 12 514–12 535, https://doi.org/10.1002/2015JD024257.

——, and ——, 2017: Atmospheric rivers in 20 year weather and climate simulations: A multimodel, global evaluation. *J. Geophys. Res. Atmos.*, **122**, 5556–5581, https://doi.org/10.1002/2016JD026174.

——, N. P. Molotch, D. E. Waliser, E. J. Fetzer, and P. J. Neiman, 2010: Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophys. Res. Lett.*, **37**, L20401, https://doi.org/10.1029/2010GL044696.

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.*, **2**, 36–50, https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

Jackson, D. L., M. Hughes, and G. A. Wick, 2016: Evaluation of landfalling atmospheric rivers along the U.S. West Coast in reanalysis data sets. *J. Geophys. Res. Atmos.*, **121**, 2705–2718, https://doi.org/10.1002/2015JD024412.

Knippertz, P., and J. E. Martin, 2005: Tropical plumes and extreme precipitation in subtropical and tropical West Africa. *Quart. J. Roy. Meteor. Soc.*, **131**, 2337–2365, https://doi.org/10.1256/qj.04.148.

Neiman, P. J., F. M. Ralph, A. B. White, D. E. Kingsmill, and P. O. G. Persson, 2002: The statistical relationship between upslope flow and rainfall in California's coastal mountains: Observations during CALJET. *Mon. Wea. Rev.*, **130**, 1468–1492, https://doi.org/10.1175/1520-0493(2002)130<1468:TSRBUF>2.0.CO;2.

——, ——, G. A. Wick, J. D. Lundquist, and M. D. Dettinger, 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. *J. Hydrometeor.*, **9**, 22–47, https://doi.org/10.1175/2007JHM855.1.

——, A. B. White, F. M. Ralph, D. J. Gottas, and S. I. Gutman, 2009: A water vapor flux tool for precipitation forecasting. *Water Manage.*, **162**, 83–94, https://doi.org/10.1680/wama.2009.162.2.83.

——, G. A. Wick, B. J. Moore, F. M. Ralph, J. R. Spackman, and B. Ward, 2014: An airborne study of an atmospheric river over the subtropical Pacific during WISPAR: Dropsonde budget-box diagnostics and precipitation impacts in Hawaii. *Mon. Wea. Rev.*, **142**, 3199–3223, https://doi.org/10.1175/MWR-D-13-00383.1.

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

——, ——, G. N. Kiladis, K. Weickmann, and D. W. Reynolds, 2011: A multiscale observational case study of a Pacific atmospheric river exhibiting tropical–extratropical connections and a mesoscale frontal Wave. *Mon. Wea. Rev.*, **139**, 1169–1189, https://doi.org/10.1175/2010MWR3596.1.

——, G. A. Wick, P. J. Neiman, B. J. Moore, J. R. Spackman, M. Hughes, F. Yong, and T. Hock, 2012: Atmospheric rivers in reanalysis products: A six-event comparison with aircraft observations of water vapor transport. *WCRP Reanalysis Conf.*, Silver Spring, MD, World Climate Research Programme, 1 p., https://www.wcrp-climate.org/ICR4/posters/Hughes_AT-20.pdf.

——, T. Coleman, P. J. Neiman, R. J. Zamora, and M. D. Dettinger, 2013: Observed impacts of duration and seasonality of atmospheric-river landfalls on soil moisture and runoff in coastal northern California. *J. Hydrometeor.*, **14**, 443–459, https://doi.org/10.1175/JHM-D-12-076.1.

——, and Coauthors, 2016: CalWater field studies designed to quantify the roles of atmospheric rivers and aerosols in modulating U.S. West Coast precipitation in a changing climate. *Bull. Amer. Meteor. Soc.*, **97**, 1209–1228, https://doi.org/10.1175/BAMS-D-14-00043.1.

——, and Coauthors, 2017a: Atmospheric rivers emerge as a global science and applications focus. *Bull. Amer. Meteor. Soc.*, **98**, 1969–1973, https://doi.org/10.1175/BAMS-D-16-0262.1.

——, and Coauthors, 2017b: Dropsonde observations of total integrated water vapor transport within North Pacific atmospheric rivers. *J. Hydrometeor.*, **18**, 2577–2596, https://doi.org/10.1175/JHM-D-17-0036.

Rutz, J. J., and W. J. Steenburgh, 2012: Quantifying the role of atmospheric rivers in the interior western United States. *Atmos. Sci. Lett.*, **13**, 257–261, https://doi.org/10.1002/asl.392.

Viale, M., and M. N. Nuñez, 2011: Climatology of winter orographic precipitation over the subtropical central Andes and associated synoptic and regional characteristics. *J. Hydrometeor.*, **12**, 481–507, https://doi.org/10.1175/2010JHM1284.1.

Waliser, D. E., and B. Guan, 2017: Extreme winds and precipitation during landfall of atmospheric rivers. *Nat. Geosci.*, **10**, 179–183, https://doi.org/10.1038/ngeo2894.

Wick, G., P. J. Neiman, and F. M. Ralph, 2013: Description and validation of an automated objective technique for identification and characterization of the integrated water vapor signature of atmospheric rivers. *IEEE Trans. Geosci. Remote Sens.*, **51**, 2166–2176, https://doi.org/10.1109/TGRS.2012.2211024.

Yang, Y., T. Zhao, G. Ni, and T. Sun, 2018: Atmospheric rivers over the Bay of Bengal lead to northern Indian extreme rainfall. *Int. J. Climatol.*, https://doi.org/10.1002/joc.5229, in press.

Zhu, Y., and R. E. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.