Climate Dynamics manuscript No.
(will be inserted by the editor)

20th-Century Atmospheric River Activity along the West Coasts of Europe and North America: Algorithm Formulation, Reanalysis Uncertainty and Links to Atmospheric Circulation Patterns

S. Brands · J.M. Gutiérrez · D. San-Martín

Received: date / Accepted: date

Abstract A new atmospheric-river detection and tracking scheme based on the magnitude and direction of integrated water vapour transport is presented and applied separately over 13 regions located along the west coasts of Europe (including North Africa) and North America. Four distinct reanalyses are considered, two of which cover the entire 20th-century: NOAA-CIRES Twentieth Century Reanalysis v2 (NOAA-20C) and ECMWF ERA-20C. Calculations are done separately for the OND and JFM-season and, for comparison with previous studies, for the ONDJFM-season as a whole.

Comparing the AR-counts from NOAA-20C and ERA-20C with a running 31-year window looping through 1900-2010 reveals differences in the climatological mean and inter-annual variability which, at the start of the 20th-century, are much more pronounced in western North America than in Europe. Correlating European AR-counts with the North Atlantic Oscillation (NAO) reveals a pattern reminiscent of the well-know precipitation dipole which is stable throughout the entire century. A similar analysis linking western North American AR-counts to the North Pacific index (NPI) is hampered by the aforementioned poor reanalysis agreement at the start of the century. During the second half of the 20th-century, the strength of the NPI-link considerably varies with time in British Columbia and the Gulf of Alaska.

Considering the period 1950-2010, AR-counts are then associated with other relevant large-scale circulation indices such as the East Atlantic, Scandinavian, Pacific-North American and West Pacific patterns (EA, SCAND, PNA and WP). Along the Atlantic coastline of the Iberian Peninsula and France, the EA-link is stronger than the NAO-link if the OND season is considered and the SCAND-link found in northern Europe is significant during both seasons. Along the west coast of North America, teleconnections are generally stronger during JFM in which case the NPI-link is significant in any of the five considered subregions, the PNA-link

S. Brands, 1. MeteoGalicia - Consellería de Medio Ambiente, Territorio e Infraestructuras - Xunta de Galicia, Santiago de Compostela, Spain 2. Instituto de Física de Cantabria (CSIC-UC), Santander, Spain, E-mail: swen.brands@gmail.com
J. M. Gutiérrez, Instituto de Física de Cantabria (CSIC-UC), Santander, Spain
D. San-Martín, Predictia Intelligent Data Solutions, Santander, Spain

is significant in British Columbia and the Gulf of Alaska and the WP-link is so along the U.S. West Coast. During OND, these links are significant in the Gulf of Alaska only.

If AR-counts are calculated upon persistent- instead of instantaneous ARs, the link to the NAO weakens over the British Isles and western Iberia. For the experimental set-ups most closely mirroring those applied in Lavers et al (2012) and Ramos et al (2015), the NAO-links are completely or partly insignificant indicating that the inclusion of the persistence criterion notably alters the results. Visual support for the present study is provided by an exhaustive historical atmospheric river archive built at `http://www.meteo.unican.es/atmospheric-rivers`.

Keywords Atmospheric Rivers · Reanalysis Data · 20th century · Atmospheric Circulation · Europe · North America

1 Introduction

The poleward transport of water vapour in the atmosphere is not organized homogeneously in space and time. Rather, it is concentrated in narrow and elongated spatial structures of intense transport having a live-time of a few days at the utmost (Zhu and Newell, 1994, 1998). Due to their filamentary appearance reminiscent of a river's course seen from bird's-eye perspective, these structures have been originally referred to as "tropospheric rivers" (Newell et al, 1992), a term which later on developed to "atmospheric rivers" (ARs). Two processes contribute to the formation and maintenance of the water vapour constituting these structures. The first one is evapotranspiration in a remote source region followed by Lagrangian transport over thousands of kilometres, similar to the flow of a river, in which case evapotranspiration and condensation along the transport "route" play a minor role (Knippertz and Wernli, 2010; Gimeno et al, 2012; Sodemann and Stohl, 2013; Garaboa et al, 2015). The second process is small-scale moisture recycling (evapotranspiration, condensation and precipitation). In this case, water vapour is continuously lost and refreshed ahead of the cold front(s) of one or several extra-tropical cyclones, leading to a structure looking like a river but not sharing its transport properties (Bao et al, 2006). Recent studies point to the fact that, for most ARs, moisture recycling is more important than long-range transport (Newman et al, 2012; Dacre et al, 2015).[1]

ARs can be identified and tracked using either Eulerian or Langrangian methods (Newell et al, 1992; Bao et al, 2006; Knippertz and Wernli, 2010; Gimeno et al, 2012; Garaboa et al, 2015). The algorithms used within the Eulerian framework, which will be the focus of the present study, are capable to automatically detect and track AR-structures at a given point in time and usually operate on vertically integrated water vapour transport (Zhu and Newell, 1998; Lavers et al, 2012; Guan and Waliser, 2016). The corresponding data are ideally taken from dropsonde- or satellite observations which, however, have a limited spatial and temporal coverage (Zhu and Newell, 1998; Ralph et al, 2004; Lavers et al, 2011). This is why

[1] Author's comment: Since the two aforementioned studies rely on reanalysis data, the corresponding results might be sensitive to the physics and parametrization schemes of the global circulation model used for re-analysing. Note that the relative contribution of the two aforementioned factors might change if other models and/or parametrization schemes are used.

model-data from reanalyses, usually referred to as “quasi-observations” (Brands et al, 2012), are used if long time series and complete spatial coverage is required, e. g. for assessing the *climatological* aspects of atmospheric rivers (Higgins et al, 2000; Neiman et al, 2008; Knippertz et al, 2013; Dacre et al, 2015).

Recently, two long-term reanalyses covering the entire 20th century have become available: the NOAA-CIRES 20th century reanalysis [hereafter: NOAA-20C, Compo et al (2011), version 2 is used here], and the ECMWF ERA-20C reanalysis [ERA-20C, Poli et al (2013)]. In comparison to alternative reanalyses relying on surface-, upper-air- and satellite observations (Kalnay et al, 1996; Dee et al, 2011), only surface observations were considered in the data-assimilation procedure opted for in these projects. This was done to reduce the risk of artificial shifts (or inhomogeneities) in the time-series simulated by the Global Circulation Models used for re-analysing. These “observational shocks” (Ferguson and Villarini, 2012) are caused by sudden increases in the number of assimilated observations; the introduction of satellite data in the late 1970s being the most prominent example (Sturaro, 2003; Sterl, 2004).

On the regional scale, seasonal precipitation sums and extreme precipitation events have been associated with ARs and strong relationships were found for those regions characterized by specific topographic features such as mountain ranges near the coast or a coastline perpendicular to the main direction of the horizontal moisture flow (Neiman et al, 2004; Guan et al, 2012; Kim et al, 2013; Ramos et al, 2015; Eiras et al, 2016). Therefore, ARs are on the one hand beneficial for a region’s water supply but on the other are potentially harmful since they can trigger heavy flooding and landslide events [hereafter jointly referred to as “hydrological extreme events”, Lavers et al (2011)], especially in case they coincide with a previously accumulated thick snow pack and/or water-saturated soils (Leung and Qian, 2009; Ralph et al, 2013).

Whereas ARs triggering (extreme) precipitation have been studied extensively to-date, a spatially (and also temporally) complete picture on the large-scale atmospheric conditions triggering ARs is yet “under construction” [see Gimeno et al (2014) and references therein]. Bao et al (2006) found that enhanced IVT over the United States (U.S.) West Coast originating from the tropics is favoured by a weakened subtropical ridge in the central Pacific. Kim and Alexander (2015) found the strength and position of the Aleutian low to be key for the spatial pattern of IVT anomalies. If this low pressure system is anomalously deep, IVT is above normal in the northwestern U.S. and if it is displaced to the south moist conditions are exhibited by the southwestern U.S. and Mexico. Guan et al (2012) found that the exceptional AR-activity over California’s Sierra Nevada during the 2010/2011 snow season was linked to the negative phase of both the Pacific-North American pattern (PNA) and the Arctic Oscillation (Barnston and Livezey, 1987). The particular role of the PNA in “driving” AR-frequency counts in the region extending from the Canada-United States boarder to Alaska has been recently pointed out by Guan and Waliser (2016). Considering the 1979-2010 period, Jiang and Deng (2011) demonstrated that East-Asian cold surges increase the odds for ARs land-falling along the west coast of North America during the days following the peak of the cold-surge.

Since the above mentioned modes of atmospheric variability are known to be influenced by low-frequency modes originating in the tropics —which, in principle, are predictable on intra-seasonal to seasonal time-scales— attempts have been made to indirectly associate these tropical modes with IVT/AR-count anomalies along the west coast of North America (Bao et al, 2006; Guan et al, 2012; Payne and Magnusdottir, 2014; Kim and Alexander, 2015; Guan and Waliser, 2016).

Similar studies for Europe are sparse and partly contradictory. Lavers et al (2012) found the number of extended-winter season (October-to-March) AR landfalls over the British Isles to be inversely related to the Scandinavian Pattern (Barnston and Livezey, 1987). In a follow-up study conducted on continental scale (Lavers and Villarini, 2013), the sea-level pressure composite maps associated with AR-arrivals in northern and southern Europe were found to resemble the positive and negative phase of the North Atlantic Oscillation [NAO, (Hurrell et al, 2003)] respectively. Ramos et al (2015) focused on October-through-March AR-arrivals over the Iberian Peninsula and found them to be positively related to the Scandinavian pattern in first place. Unlike in Lavers and Villarini (2013), the "AR-NAO" link was found to be insignificant in Ramos et al (2015) which is perhaps somewhat counter-intuitive given that the NAO is known to describe a large fraction of variability of the wintertime precipitation totals in this region (Hurrell, 1995; Trigo et al, 2004). Reasons for this disagreement might be found in differences in the considered datasets, time-periods and season-definitions.

This study assesses the atmospheric river phenomenon along the west coasts of Europe and North America from a climatological point of view. Considering the October-through-December and January-through-March seasons (OND and JFM), a new AR-detection and tracking algorithm is proposed and applied to 6-hourly instantaneous data from four distinct reanalyses, two of which date back to the year 1900 (NOAA-20C and ERA-20C). This is done separately for 13 regions located along the west coasts of Europe and North America (see Figure 1). After pointing out the advantages of the new algorithm, the similarity between the year-to-year AR-count series from the two long-term reanalyses is assessed backwards in time to the early 20th-century. Similarity is measured 1) in terms of the climatological mean (represented by the $bias$) and 2) in terms of inter-annual variability (represented by the rank correlation coefficient, rs). Both measures are calculated for a sliding 31-year window moving forward by one year in a loop ranging from 1900 to 2010. In the absence of any "true" dataset dating back to 1900, and following the ideas of Sterl (2004), a comparison of two distinct reanalyses provides an estimate of their degree of realism. If similar results are obtained from the two for a given time-series aspect (i.e. inter-annual variability as documented by rs), this is likely due to a strong observational constrain, indicating that the result is a realistic estimation of the "truth". A large discrepancy, in contrast, indicates a loose observational constrain and an unreliable result. Encouraged by the finding that the inter-annual variability of the AR-counts for Europe is similar in the two reanalyses even at the start of the 20th century, the time-dependence of their association with the NAO is traced backwards until 1900 by means of a running correlation analysis. Results are then contrasted with a similar analysis relating AR-counts in western North America with the strength of the Aleutian low as described by the North Pacific index (Trenberth and Hurrell, 1994).

In a second working step, the search for atmospheric drivers of regional AR-activity is extended to other relevant indices describing the East Atlantic, Scandinavian, Pacific-North American and West Pacific patterns (Barnston and Livezey, 1987). In this case, we focus on the more reliable period 1950-2010. For ease of comparison with previous studies (Lavers et al, 2012; Ramos et al, 2015), seasonal

AR-counts are additionally derived from NCEP/NCAR reanalysis 1 (Kalnay et al, 1996) and ECMWF ERA-Interim (Dee et al, 2011) and the entire extended winter season (October-through-March), as well as the persistence criterion described in Lavers et al (2012) are considered in addition. It will be shown that the application of the persistence criterion can weaken the strength of the statistical relationships to the degree that the significant link to the NAO is lost for the experimental set-ups that most closely match those applied in Lavers et al (2012) and Ramos et al (2015).

Finally, an exhaustive historical archive of AR-events in the above mentioned 13 regions was built and made publicly available at `http://www.meteo.unican.es/atmospheric-rivers`. This archive will hereafter be referred to as the *"Atmospheric River Archive"*. It documents the behaviour of the proposed detection and tracking algorithm for thousands of cases and permits to openly discuss its advantages and disadvantages.

The remainder of this article is outlined as follows. The applied datasets are described in Section 2. The AR-detection and tracking scheme as well as the applied reanalysis similarity measures are described in Section 3. Results are presented in Section 4 and a discussion and some concluding remarks are provided in Section 5.

2 Data

For the purpose of AR detection and tracking, 6-hourly instantaneous data from the four reanalyses specified in Table 1 are used (the respective URLs are provided in the Acknowledgements).

The algorithm operates on the *magnitude* (IVT, in $kg\,m^{-1}\,s^{-1}$) and *direction* (D in degrees) of the vertically integrated water vapour transport which are calculated as follows:

$$IVT = \sqrt{{IVT_u}^2 + {IVT_v}^2} \tag{1}$$

$$D = atan2\left(\frac{IVT_u}{IVT}, \frac{IVT_v}{IVT}\right)\frac{180}{\pi} + 180 \tag{2}$$

where IVT_u and IVT_v are the vertical integrals of the zonal and meridional water vapour transport components respectively. The $atan2$ function returns the four-quadrant inverse tangent ranging in between $-\pi$ and π which is then transformed to degree values ranging in between 0° and 360°.

IVT_u and IVT_v were calculated from 2-dimensional pressure-level data between 1000 and 300 hPa (Lavers et al, 2012).

$$IVT_u = \frac{1}{g}\int_{1000}^{300} qu\,\mathrm{d}p \tag{3}$$

and

$$IVT_v = \frac{1}{g}\int_{1000}^{300} qv\,\mathrm{d}p \tag{4}$$

where q, u and v refer to specific humidity (in $kg\,kg^{-1}$), zonal and meridional wind (in $m\,s^{-1}$) at pressure level p, g to acceleration due to gravity and dp to the difference between adjacent pressure levels (in Pa).

For NCEP/NCAR and NOAA-20C, 7 and 15 vertical pressure levels between 1000 and 300 hPa were available from the data providers respectively. Vertical integration is achieved by multiplying qu and qv at the pressure level p by a multiplier describing its contribution (as represented by the number of pressure levels in Pa) to the entire column extending from 1000 to 300 hPa (see Table 2), followed by summing up the resulting products. Since ECMWF's public server already provides IVT_u and IVT_v as vertical integrals between the pressure level at model surface and the top of the atmosphere (ECMWF, personal communication), it was not necessary to apply Equations 3 and 4 for ERA-20C and ERA-Interim. Note that q, u and v from NOAA-20C are ensemble-mean data.

In addition to the reanalysis datasets, monthly values of the large-scale atmospheric circulation indices relevant for the North Atlantic and North Pacific sectors were retrieved from the Climate Prediction Center (CPC) (Barnston and Livezey, 1987) and the University Cooperation for Atmospheric Research (UCAR) (Hurrell et al, 2003). A detailed description of these indices can be found in Table 3.

The considered time periods are as follows. A 31-year moving window running from 1900-2010 is used to assess time-variations in 1) the similarity between AR-counts from NOAA-20C and ERA-20C and 2) the strength of their link to the NAO or NP indices. The "full" association including indices others than the aforementioned two is conducted for the 1950 - 2010 period except for AR-counts from ERA-Interim, in which case 1979 - 2013 is used. Finally, 1979/80 - 2009/10 and 1950/51 - 2011/12 are considered for comparison with Lavers et al (2012) and Ramos et al (2015) respectively.

3 Methods

3.1 Atmospheric-River Detection and Tracking Algorithm

In the present study, ARs are detected separately in 8 regions ranging from Morocco to northern Norway and 5 regions ranging from southern California to the northern Gulf of Alaska respectively (see Fig. 1). Each detection region is defined as a "barrier" of grid-boxes approximately following the coastline. Due to distinct native horizontal resolutions, the exact coordinates of these barriers slightly differ from one dataset to another (the barriers shown in Fig. 1 refer to the ERA-20C dataset). Using the native resolution is preferable to interpolating to a common coarse grid, which would lead to a degradation of the higher-resolution datasets.

For a given detection region formed by a barrier of b grid-boxes, the following detection and tracking algorithm was applied every six hours (see also Figure 2).

1. The grid-box of maximum IVT along b is retained. This grid-box is hereafter referred to as the "targeted grid-box" e.
2. If the IVT value at e exceeds the predefined percentile threshold P_d (the detection percentile) the AR-tracking algorithm is activated, otherwise it proceeds to the next point in time.

3. Then, the direction (D) of the IVT-flow at e is calculated (see Equation 2) and discretized into the 8 cardinal directions: N, NE, E, SE, S, SW, W, NW. In the following example, we assume that D is from the W.
4. Out of the 8 possible neighbouring grid-boxes surrounding e, the algorithm considers the upstream grid-box s as well as the two grid-boxes neighbouring s (i. e., following the example, the 3 grid-boxes to the West, North-West and South-West of e). Among these 3 candidate grid-boxes the grid-box of maximum IVT is detected.
5. If this maximum IVT value exceeds the predefined percentile threshold P_t (the tracking percentile which not necessarily equals P_d, see also Table 4), the grid-box is retained as the new targeted grid box e. In this case, the algorithm proceeds to 3). Otherwise, it is stopped at this point in time and proceeds to the next point in time.
6. The algorithm continues until 5) is not met any more or until the detected IVT structure exceeds a length of l grid-boxes or in case a grid-box is detected twice, which can occur if the algorithm completely orbits a low pressure system. Note that l depends on the horizontal resolution of the dataset and equals 32, 40, 70 and 107 grid-boxes for NCEP/NCAR, NOAA-20C, ERA-20C and ERA-Interim respectively. For the ideal case of a purely meridional AR with no zonal displacement, this roughly corresponds to a longitude of 11000 km.
7. If the longitude of the detected IVT structure exceeds a threshold of 3000 km (spherical distance is considered), the detection region b is said to be affected by an AR at this point in time. If it is shorter than 3000 km, the structure is not considered an AR.

Considering the reference period 1979-2009[2], P_d and P_t were calculated separately for each grid box and month. Based on a comparison with the ARs detected in Neiman et al (2008) and Dettinger et al (2011), Lavers et al (2012) suggested the use of the 85^{th} percentile for P_d which, however, was replaced by other plausible values in some studies [e. g. Warner et al (2015)]. Thus, a secondary goal of the present study is to explore how sensitive the results are to variations not only in P_d but also in P_t. To this aim, our tracking algorithm was applied 6 times using 6 distinct combinations of the two parameters (see Table 4). The corresponding six values will hereafter be referred to as the "percentile sample". Its range describes the method-related sensitivity of the results. Additional sensitivity test were conducted 1) taking into account persistent and independent AR events only and/or 2) intentionally turning-off our algorithm's capability to track towards the N, NE, E and SE and/or 3) considering a length criterion of > 2000 instead of > 3000 km. An event is considered "persistent" if a given target region is continuously affected by an AR for at least 18 hours and if it is separated from other events by more than 24 hours (Lavers et al, 2012).

3.2 Reanalysis Comparison and Association with Circulation Indices

For each target region and season (OND or JFM), and each of the 6 AR definitions mentioned above, the seasonal AR-counts from the two long-term reanalyses

[2] common to all applied reanalysis datasets

(NOAA-20C and ERA-20C) are compared in terms of similarity in their *climatological mean* as expressed by the *bias*:

$$bias = \frac{\bar{y} - \bar{x}}{\bar{x}} \times 100 \quad (5)$$

where $\bar{x}$ and $\bar{y}$ are the climatological mean values of the seasonal AR-counts in the two reanalyses. Here, NOAA-20C is assumed to be the reference reanalysis x.

Similarity in terms of *inter-annual variability* is measured by correlating these counts with Spearman's rank correlation coefficient (rs). Prior to calculating rs, the year-to-year AR-count time series are optionally de-trended using Poisson regression with a log link function (Lavers et al, 2012).

To identify possible variations along the course of the entire study period (1900-2010), a 31-year window moving forward by one year from the start of the study period (1900-1930) till its end (1980-2010) is used and the above mentioned similarity measures, as well as the de-trending applied prior to calculating rs, are calculated separately for each sub-period. Apart from comparing the AR-counts from the two long-term reanalyses, rs is also used to associate these counts with the circulation indices listed in Table 3. Since the latter are continuous variables, they are optionally de-trended using ordinary least-squares regression instead of Poisson regression. The significance of rs is assessed with a two-tailed Student t-test conducted at the 5%-level, assuming temporal independence of the applied time series.

4 Results

4.1 AR-Detection and Tracking

Figure 3 provides an illustrative example of the algorithm's capability to detect and track AR structures. The figure shows an AR affecting southern Norway on 11 January 1971 OO UTC, as retrieved from NCEP/NCAR, NOAA-20C and ERA-20C (panels a, b and c respectively). Colour shadings and vector lengths are proportional to the magnitude of the vertically integrated water vapour flux. The direction of the flow is indicated by the orientation of the vectors and the cyan line represents the AR-track found by the algorithm. The initial landfall of this AR was detected earlier and this particular point in time is chosen to show the algorithm's capability to track towards the N, NE, E and SE (SE in this case, as described below) at any point along the AR track. This "eastward tracking" capability was not accounted for in the initial formulation of the Lavers et al (2012) algorithm, able to track towards the S, SW, W, NW only. Albeit this was corrected in the later versions of this algorithm (Lavers and Villarini, 2013, 2015), these do not do account for $\approx 180^\circ$ curves as those shown in Figure 3. Starting from a given detection barrier (e. g. 10°W for the case of western Europe), the Lavers and Villarini (2013) algorithm moves towards the West and tracks the maximum IVT threshold at each longitude. For the structure being an AR in Lavers and Villarini (2013), the tracked IVT values must exceed the assumed percentile threshold along a longitudinal distance of 20°. What is key for the understanding of our method is that the Lavers and Villarini (2013) algorithm only detects *one* grid-box per longitude. To perform a 180° turn, however, a second IVT value exceeding the threshold must be located at the *same* longitude further to the South (see Figure 3c) and this is not accounted for by Lavers and Villarini (2013), to the authors' knowledge. Telling the algorithm to move to the east, starting from the detection barrier, does not solve this problem either. Here, it will be shown that even though this limitation is of minor importance in Europe, it is detrimental to AR-detection in some regions along the west coast of North America (see below).

In spite of distinct native horizontal resolutions and applied data assimilation strategies, the three reanalyses produce virtually identical results for the AR event shown in Figure 3. Since the direction of the flow is scanned prior to searching the grid-box of maximum IVT, the algorithm correctly moves upstream after detecting the AR in southern Norway. The "curves" of the flow are captured well and so is the SE flow between the British Isles and the Iberian Peninsula. Finally, the algorithm stops in the central subtropical Atlantic because the allowed maximum of tracked grid-boxes (l) is exceeded. As an extension to this illustrative example, the *Atmospheric River Archive* available at `http://www.meteo.unican.es/atmospheric-rivers` documents all ARs detected in the 13 target regions displayed in Figure 1 during the period 1900-2010 (ERA-20C is compared to NOAA-20C) and 1979-2014 (only ERA-Interim is shown).

To draw some more general conclusions on the relevance of the "eastward tracking" capability, Figure 4 displays the fraction of ARs that are detected if this capability is intentionally turned off (F_{noeast}):

$$F_{noeast} = \frac{AR_{noeast}}{AR_{all}} \times 100 \tag{6}$$

where AR_{noeast} is the seasonal AR-count retrieved from an algorithm not capable to track towards the N, NE, E and SE, and AR_{all} is the respective count obtained from the fully capable algorithm as described above.

Figure 4 illustrates that eastward tracking is more relevant during OND than during JFM and more so in North America than in Europe. In the Gulf of Alaska, up to 70% of the ARs are "lost" if eastward tracking is not considered, which is due to the fact that ARs approaching this region from southerly directions frequently have a slight eastward component near landfall and turn to westerly directions when further tracked upstream. For an illustrative example of this phenomenon, the interested reader is referred to the AR-detections in December 2014 (see reanalysis: "ERA-Interim", continent: "western North America" and region: "northern Gulf of Alaska" at `http://www.meteo.unican.es/atmospheric-rivers`).

Finally, 90% of the AR-events documented in Dettinger et al (2011) (see their table 1) coincide with the ERA-Interim based AR-detections provided by the *Atmospheric River Archive* if the target day documented in Dettinger et al (2011) is "relaxed" by ± 18 hours. The "missing" 10% can largely be explained by the comparatively long AR-length criterion applied here (> 3000 km). If our algorithm is re-run with a shorter length criterion (> 2000 km) the coincidence rate rises to 97%. Interestingly, even though a longer length criterion is assumed, our archive contains more events than the Dettinger et al (2011) archive.

4.2 Temporal Variations in Reanalysis Similarity during 1900-2010

Figure 5 displays the year-to-year AR-count sequence obtained from NOAA-20C (blue) and ERA-20C (red) respectively; results are for the OND-season in this case. As above, the lines and shadings refer to the mean and range of the 6 seasonal AR-count values per year obtained from the 6 considered percentile combinations listed in Table 4. For the sake of completeness, AR-counts from NOAA-20C extend to 2012. Panels a to h refer to the results for Europe, panels i to h to the results for western North America. Figure 6 shows the respective results for the JFM-season. Note that the mean value of the 1900-2010 AR counts is displayed in the header of each panel for each of the two datasets (first number = NOAA-20C mean, second number = ERA-20C mean).

Results for the 31-year "running" *bias* in the AR-counts (ERA-20C minus NOAA-20C w.r.t to NOAA-20C, see Equation 5) are shown in Fig. 7. On the x-axis of each panel, the centre-year of a specific sub-period is displayed (e. g. "1920" refers to the time period 1905-1935). We will hereafter refer to this centre year instead of mentioning the entire sub-period. On the y-axis, the *bias* is displayed as the percentage deviation from the mean of the reference reanalysis for that sub-period, which is NOAA-20C. Again, the lines and shadings refer to the mean and range of the 6 *bias* values obtained from the percentile sample. To measure the stationarity of the *bias*, the standard deviation (std) of the 81 percentile-sample mean values for a given season and target region is displayed in the header of each panel (first number = OND std, second number = JFM std).

A visual inspection of the year-to-year time-series relevant for Europe (see panels a to h in Figures 5 and 6) reveals that up to at least the 1970s (1930s in northern Norway and 1940s in northern Iberia) NOAA-20C produces systematically more ARs than ERA-20C whereas the opposite is the case from approximately the 1980s onward. This translates into a change in the sign of the bias from negative values down to approximately -40% at the start of the 20th-century to positive values up to approximately $+25\%$ in the recent past (see panels a to h in Figure 7). As indicated by the standard deviation in the header of each panel, the non-stationarity of the bias is more pronounced in OND than in JFM, with the largest values obtained for the British Isles.

Contrary to what was found for Europe, ERA-20C produces up to twice as many ARs as NOAA-20C in western North America (exception: southern California, see panels i to m in 7). Such a large *bias* might be explained by the fact that the 56-member ensemble of NOAA-20C, during the "data-sparse" start of the 20th-century, suffers such a large spread that the percentile thresholds listed in Table 4 are exceeded by the ensemble-mean values far less often than during the later ("data-rich") period, leading to a reduction in AR detections for this reanalysis [see also Champion et al (2015)]. ERA-20C is a deterministic reanalysis and is therefore not affected by this issue. Nevertheless, due to the general lack of data, it cannot be expected to provide realistic AR-counts at the start of the century either. By approximately 1920s (with the exception of the northern Gulf of Alaska), the *bias* for western North America decreases to a magnitude comparable to the that found for Europe. As for Europe, temporal variations in the bias are more pronounced during the OND- than during the JFM-season, particularly over the southern and northern Gulf of Alaska.

Figure 8 displays the results of the running correlation analyses. On the y-axis, the rank correlation coefficient (rs), as well as the critical values for a two-tailed t-test applied at a test-level of 5% are shown (see dashed lines). Regarding the European regions, rs is systematically lower and its range (reflecting the method related uncertainty) systematically larger during the OND than during the JFM season. Values generally decrease as one moves backward in time. With rs exceeding +0.6 in nearly any case, the AR counts' inter-annual variability is roughly similar in both datasets even at the very beginning of the 20th century. From 1955 onwards, rs is greater than- or close to +0.8, indicating a close similarity during the last 7 decades of comparison. However, OND values in Norway —for unknown reasons— are smaller during the recent past than during the mid-20th-century (see panels g and h in Figure 8).

In contrast to the result for Europe, rs values along the west coast of North America are insignificant or even negative at the start of the century (note the distinct scale of the y-axes). Another distinction is that the rs values in OND are much closer to those obtained for JFM and actually are larger during the first decades of the 20th century. Following the running rs forward in time, a value of approximately +0.5 is at the latest reached around 1935 and a value of approximately $+0,8$ is so around 1965.

The method-related sensitivity of the results is small in comparison to the mean value (compare shadings with lines in Figures 5 to 7), which is generally also the case for the forthcoming results. Reducing the AR-length criterion to > 2000 km slightly improves the reanalysis agreement without, however, bringing the huge differences found over North America at the start of the century to an acceptable level (not shown).

4.3 Temporal Variations in the Link to the NAO and Aleutian Low during 1900-2010

Since the inter-annual variability of the AR-counts in Europe was found to be similar in the two long-term reanalyses even at the start of the 20-century, we proceed to assess their association with the seasonal-mean NAO (the station-based index is used here). To this, a running rank correlation analysis is applied in the aforementioned configuration, i. e. a 31-year moving window is used. The same is done for the AR-counts along the west coast of North America, having in mind that insignificant rs were obtained at the start of the 20th-century when comparing the two reanalysis there. The results for the OND and JFM seasons are displayed in Figures 9 and 10 respectively. Blue lines and shadings are for AR-counts from NOAA-20C and red ones are for AR-counts from ERA-20C. Also shown are the critical values for a significant rs at a test-level of 5% (see dashed lines).

Similar to the well-known correlation dipole for seasonal precipitation totals (Hurrell, 1995), AR-counts in southern Europe are inversely related to the NAO whereas in northern Europe a positive relationship is found, which is in agreement with the Lavers and Villarini (2013) results (see panels a to h in Figures 9 and 10). These relationships are generally weaker and less stationary (i.e. variable in time) during OND than during JFM. In the two southernmost and the two northernmost regions, rs in JFM is significant for any of the 81 considered sub-periods indicating a temporally robust link to the NAO during this season. In northern Iberia and western France, however, rs is significant from 1940 until the end of the 1970s only. Similarly, over the British Isles, rs is insignificant from approximately 1915 to 1921 and —for NOAA-20C— also from 1960 to 1970, indicating that the NAO-link in the three central regions of the European Atlantic seaboard is subject to non-negligible variations along the course of the century. During the OND season, rs is generally insignificant except for Morocco and southern Iberia from approximately 1915 to 1930 and from 1975 onwards, and for southern Norway from approximately 1940 to 1970. Albeit somewhat larger during OND than during JFM, dataset-induced differences are generally small for Europe.

As expected from the results of the reanalysis comparison, dataset-induced differences concerning the link between AR-counts in western North America and the Aleutian low can be larger than 0.5 correlation points at the start of the century (see panels i to m in Figures 9 and 10). In the two southernmost regions, rs is insignificant or prone to large dataset-differences along the entire study period, except during the JFM-season where significantly negative rs are obtained from 1945 to 1960 (in North California-Oregon-Washington only) and from 1990 onwards (in both regions).

In the 3 remaining regions, rs for JFM is significantly negative from approximately 1950 onwards, except for the northern Gulf of Alaska where insignificant values are obtained in the very recent past (from 1990 onwards). During OND, dataset-induced differences in the results are relatively large until at least 1955. Thereafter, these differences diminish, revealing sig. negative rs in British Columbia and the southern Gulf of Alaska, which, however, decrease when approaching the present, eventually becoming insignificant from 1970 / 1980 onwards. This decrease is most pronounced in British Columbia. The OND values for the northern Gulf of Alaska are constantly sig. negative from approximately 1955 onwards.

As can be seen from Figures S01 and S02 in the supplementary online material, similar results are obtained when the AR-count and index time series are de-trended (separately in each 31-year period of the running analyses) prior to calculating rs.

4.4 Relationship to the Large-Scale Atmospheric Circulation during 1950-2010

Figure 11 shows the rs between the seasonal AR counts in the eight considered European target regions and the seasonal-mean large-scale circulation indices relevant there. Unlike the running analyses conducted above, rs in this section is calculated once for the period 1950-2010[3], or 1979-2013 in case ARs from ERA-Interim are considered. As above, the bars and errorbars in a given panel refer to the mean and range of the percentile sample (see Table 4). The critical values obtained from a two-sided t-test conducted at a test-level of 5% are indicated by dashed lines. Along the rows, results for ARs retrieved from NCEP/NCAR, NOAA-20C, ERA-20C and ERA-Interim are displayed from the top to the bottom. The OND-, JFM- and ONDJFM results are provided in columns 1-3.

The three re-analyses covering the 1950-2010 period produce very similar results (see rows 1-3 in Figure 11). During both OND and JFM (see columns 1+2),

[3] note that the indices provided by the Climate Prediction Center are available from 1950 onwards only

relationships to the NAO are strongly negative in the southern European regions, weaker in the central regions and strongly positive in the northern regions, thereby depicting the well-known correlation dipole found for precipitation in earlier studies (Hurrell, 1995; Qian et al, 2000). Since r_{H-NAO} over the period 1950-2010 is significant for almost any region irrespective of the considered season and dataset and since the magnitude of rs is close to 0.8 in some cases, the NAO, and particularly the NAO based on SLP, is the most important circulation pattern influencing extended winter AR counts in Europe if the results are seen as a whole. Exceptions from this general finding are mainly found during the OND season in which case the AR-counts over western Iberia, northern Iberia and western France are more strongly linked to the EA than to the NAO (rs_{EA} lies in between +0.5 and +0.7) and those over the British Isles are more strongly linked to the SCAND ($rs_{SCAND} \approx -0.4$). Links to the NAO and SCAND are more pronounced during the JFM than during the OND season whereas the opposite is found for the links to the EA. During JFM, rs_{SCAND} is between -0.4 and-0.65 in the three northernmost regions and $\approx +0.4$ in the two southernmost ones. During OND, rs_{SCAND} is significant in the three northernmost regions only. Links to the EA/WR are significant in the latter three regions during OND and in northern Norway during JFM. Links to the POL are generally insignificant except during JFM in northern Iberia and western France. When considering the entire winter half-year (see column 3 in Figure 11), the strength of the teleconnections generally lies in between the values obtained for OND and JFM.

A series of additional sensitivity tests were conducted for the ONDJFM season and the respective results are displayed in Figure 12. The first column refers to solely considering persistent ARs, the second to "turning off" our algorithm's capability to track towards the N, NE, E and SE, and the third to using a length criterion of > 2000 instead of 3000 km (over the sphere). From these additional experiments, it becomes obvious that the inclusion of the persistence criterion weakens the link between the ONDJFM-AR counts and the NAO indices particularly over the British Isles (compare first column in Figure 12 with last column in Figure 11). This effect is most appreciable in case the experimental set-up considered in Lavers et al (2012)[4] is used in combination with a length criterion of > 3000 km, in which case $rs_{CPC-NAO}$ is consistently insignificant (see Table 5a). For the experimental set-up used in Ramos et al (2015)[5], the detrimental effect of the persistence criterion leads to insignificant $rs_{CPC-NAO}$ for two out of six percentile combinations irrespective of the applied length criterion (see Table 5). Finally, neither disabling the algorithm's capability to track towards the N, NE, E and SE nor applying the alternative length criterion does notably alter the results in this region of the world (compare columns 2 and 3 in Figure 12 with the last column in Figure 11).

The respective results for the west coast of North America and the circulation indices relevant there are shown in Figure 13 and Figure 14 respectively. Instantaneous AR counts along the Gulf of Alaska are positively correlated with the PNA and negatively correlated with the NP (see Figure 13). Yet significant in both seasons, these links are more pronounced during JFM than during OND (compare

[4] i. e. considering persistent ARs during ONDJFM 1979/80 - 2009/10 derived from ERA-Interim

[5] i. e. considering persistent ARs during ONDJFM 1950/51 - 2011/12 derived from NCEP/NCAR

first and second column). During JFM, AR-counts in SouthCal, NorthCal-OR-WA and British Columbia are also significantly associated with the NP index, with the exception of the AR-counts in SouthCal and NorthCal-OR-WA obtained from ERA-20C, in which case results are on the limit to significance (see panel h). It is during the JFM-season only when ARs over the two aforementioned regions are significantly related to the WP. Teleconnections involving the AR-counts in the Gulf of Alaska are systematically weaker during 1979-2013 than during 1950-2010 (compare last row to rows 1-3 in Figure 13). This finding is not dataset-dependent (see the "late" results of the running correlation analyses in Figures 9 and 10) and might be explained by the systematic strengthening of the wintertime Aleutian low after the Pacific Climate Shift in 1976/77 (Deser et al, 2004).

Unlike in Europe, the persistence criterion's effect on rs is not systematic along the west coast of North America, i. e. can lead to a slight increase or decrease in rs (compare first column in Figure 14 with last column in Figure 13). If the algorithms capability to track towards the N, NE, E and SE is disabled, teleconnections with the PNA an NP become insignificant in the northern and southern Gulf of Alaska (compare second column in Figure 14 with last column in Figure 13). Thus, the inclusion of this capability is key to properly capture the inter-annual variability of the AR-counts in these regions. As was the case for Europe, applying the alternative length criterion does not notably alter the results (compare last column in Figure 14 with last column in Figure 13).

5 Summary, Discussion and Concluding Remarks

On the basis of a new algorithm operating on the *magnitude* and *direction* of IVT, time series of year-to-year AR occurrence counts were calculated for 13 target regions along the west coasts of Europe (including North Africa) and North America. This was done separately for the OND and JFM-seasons using 6-hourly instantaneous data from 4 distinct reanalyses, two of which extend back to the early 20th century (1900). In principle, no AR-persistence criterion was considered.

A "running" comparison of the seasonal AR counts from the two long-term reanalyses over the period 1900-2010 revealed:

1. Biases which are especially pronounced in, but not limited to, the early 20th-century. With up to $> 100\%$, the biases during this early period are more severe in western North America than in Europe.
2. Along the west coast of Europe, the two reanalyses produce a similar inter-annual variability even at the start of the 20th-century (rank correlation $\geq +0.6$). This is in sharp contrast to the near-to-zero correlation found along the west coast of North America during the same period. In this region, rs steadily increases until approximately 1945-75 and thereafter remains constant at a level $\geq +0.8$.

Encouraged by finding 2), the stationarity of AR-NAO link was traced back to the early 20th-century using a 31-year running correlation analysis over the period 1900-2010. Albeit rs for individual target regions can vary along the time-axis, particularly during the OND-season, the dipole found on continental-scale (i.e. looking at the conjunction of target regions) is generally found in each sub-period, indicating that it is a robust feature along the course of the entire 20th century. Applying the same method for AR-counts along the west coast of North America and the strength of the Aleutian Low (as represented by the North Pacific Index) revealed larger variations in time which —during the early 20th century— are attributable to dataset uncertainties including uncertainties in the NP index itself (Trenberth and Hurrell, 1994). From the 1940-1970 climate period onwards, however, these uncertainties are small and the detected non-stationarities in the above link —which are most pronounced over British Columbia and the southern Gulf of Alaska— are likely to reflect real processes. A detailed assessment of the causes for this is recommended for the future.

For the reliable period 1950-2010, the search for atmospheric drivers of seasonal AR-occurrence counts was extended to circulation patterns others than the NAO and NP. For western Europe, the NAO was found to be the most important atmospheric driver of AR activity if the results are seen as a whole. In particular, the OND and ONDFJM AR-counts over the British Isles and western Iberia are significantly linked to the NAO if no persistence criterion is applied. Remarkably, if the Lavers et al (2012) persistence criterion is applied, rs values in the two aforementioned regions drop to insignificance (or near insignificance) for the experimental set-ups that most closely mirrow those applied in Lavers et al (2012) and Ramos et al (2015). However, despite conceptual similarities, the tracking algorithm applied here is not identical to that used in the above mentioned studies. Therefore, it cannot be ultimately demonstrated that the persistence criterion is the responsible for, e. g., the insignificant AR-NAO link found in Ramos et al (2015). During the OND-season, AR-counts along the Atlantic coast of Iberia and France were found to be more strongly linked to the East Atlantic pattern than to the North Atlantic Oscillation. As formerly pointed out in Lavers et al (2012), AR-counts over the British Isles were found to be significantly associated with the Scandinavian index. Here, it was shown that this index is a significant driver of the AR-activity in Norway.

Apart from the aforementioned links to the Aleutian low, the PNA was found to significantly alter the AR-counts in British Columbia and the Gulf of Alaska during the JFM-season, which is in agreement with the Guan and Waliser (2016). During the OND season these links are generally weaker, leading to insignificant results over British Columbia. It is during JFM only when AR-counts along the U.S. west coast are significantly related to the West Pacific pattern.

The above mentioned uncertainties in the “quasi-observed” climatological mean AR-counts should be taken into account when evaluating the bias of e. g. the CMIP5 Earth System Models (Taylor et al, 2012) against either of the two long-term reanalyses, particularly during the early 20th century. This type of uncertainty is also expected to hinder the association of specific AR events from either of the two reanalyses with hydrological extreme events documented by other sources. Finally, the close agreement on the seasonal AR-counts' inter-annual variability back to 1900-31 for Europe and 1920-51 for western North America permits to assess their variability (and predictability) with longer time series, as was shown here for their link to the NAO and NP indices. A logical future step is to relate AR-activity to sea-surface temperature variations on multiple time-scales (Zhang et al, 1997; Trenberth et al, 1998; Delworth and Mann, 2000; Broennimann, 2007).

Acknowledgements The authors would like to thank Jorge Eiras-Barca, Daniel Garaboa, Dr. Gonzalo Miguez-Macho, Dr. David Lavers and two anonymous referees for their constructive criticism and helpful advice. They acknowledge the use of the climate indices provided by UCAR https://climatedataguide.ucar.edu/climate-data/ and the Climate Prediction Center http://www.cpc.ncep.noaa.gov/, the ECMWF ERA-20C and ERA-Interim reanalyses http://apps.ecmwf.int/datasets, the NOAA-CIRES 20th Century Reanalysis version 2 http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.html and the NCEP/NCAR reanalysis 1 http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html. SB would like to thank the TRAGSA Group and the CSIC JAE-PREDOC programme for financial support.

References

Bao J, Michelson S, Neiman P, Ralph F, Wilczak J (2006) Interpretation of enhanced integrated water vapor bands associated with extratropical cyclones: Their formation and connection to tropical moisture. Mon Weather Rev 134(4):1063–1080, DOI {10.1175/MWR3123.1}

Barnston A, Livezey R (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon Weather Rev 115(6):1083–1126, DOI {10.1175/1520-0493(1987)115⟨1083:CSAPOL⟩2.0.CO;2}

Brands S, Gutierrez JM, Herrera S, Cofino AS (2012) On the use of reanalysis data for downscaling. J Clim 25(7):2517–2526, DOI {10.1175/JCLI-D-11-00251.1}

Broennimann S (2007) Impact of El Nino Southern Oscillation on European climate. Rev Geophys 45(2), DOI {10.1029/2006RG000199}

Champion AJ, Allan RP, Lavers DA (2015) Atmospheric rivers do ot explain UK summer extreme rainfall. J Geophys Res Atmos DOI {10.1002/2014JD022863}

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE Jr, Vose RS, Rutledge G, Bessemoulin P, Broennimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli O, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ (2011) The Twentieth Century Reanalysis Project. Q J R Meteorol Soc 137(654, A):1–28, DOI {10.1002/qj.776}

Dacre HF, Clark PA, Martinez-Alvarado O, Stringer MA, Lavers DA (2015) How do atmospheric rivers form? Bull Amer Meteorol Soc 96(8):1243–1255, DOI {10.1175/BAMS-D-14-00031.1}

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kallberg P, Koehler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656, Part a):553–597, DOI {10.1002/qj.828}

Delworth T, Mann M (2000) Observed and simulated multidecadal variability in the Northern Hemisphere. Clim Dyn 16(9):661–676, DOI {10.1007/s003820000075}

Deser C, Phillips A, Hurrell J (2004) Pacific interdecadal climate variability: Linkages between the tropics and the North Pacific during boreal winter since 1900. J Clim 17(16):3109–3124, DOI {10.1175/1520-0442(2004)017⟨3109:PICVLB⟩2.0.CO;2}

Dettinger MD, Ralph FM, Das T, Neiman PJ, Cayan DR (2011) Atmospheric Rivers, Floods and the Water Resources of California. Water 3(2):445–478, DOI {10.3390/w3020445}

Eiras-Barca J, Brands S, Miguez-Macho G (2016) Seasonal variations in North Atlantic atmospheric river activity and associations with anomalous precipitation over the Iberian Atlantic Margin. J Geophys Res:in press, DOI {10.1002/2015JD023379}

Ferguson CR, Villarini G (2012) Detecting inhomogeneities in the Twentieth Century Reanalysis over the central United States. J Geophys Res Atmos 117, DOI {10.1029/2011JD016988}

Garaboa D, Eiras-Barca J, Huhn F, nuzuri V PM (2015) Langrangian coherent structures along atmospheric rivers. Chaos

Gimeno L, Nieto R, Vázquez M, Lavers D (2014) Atmospheric rivers: a mini-review. Front Earth Sci, doi: 10.3389/feart.2014.00002

Gimeno L, Stohl A, Trigo RM, Dominguez F, Yoshimura K, Yu L, Drumond A, Maria Duran-Quesada A, Nieto R (2012) Oceanic and terrestrial sources of continental precipitation. Rev Geophys 50, DOI {10.1029/2012RG000389}

Guan B, Waliser DE, Molotch NP, Fetzer EJ, Neiman PJ (2012) Does the Madden-Julian Oscillation influence wintertime atmospheric rivers and snowpack in the Sierra Nevada? Mon Weather Rev 140(2):325–342, DOI {10.1175/MWR-D-11-00087.1}

Guan B, Waliser DE (2016) Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. J Geophys Res 120:in press, DOI {10.1002/2015JD024257}

Higgins R, Schemm J, Shi W, Leetmaa A (2000) Extreme precipitation events in the western United States related to tropical forcing. J Clim 13(4):793–820, DOI {10.1175/1520-0442(2000)013⟨0793:EPEITW⟩2.0.CO;2}

Hurrell J (1995) Decadal trends in the North-Atlantic Oscillation - Regional temperatures and precipitation. Science 269(5224):676–679, DOI {10.1126/science.269.5224.676}

Hurrell J, Kushnir Y, Ottersen G, Visbeck M (2003) The North Atlantic Oscillation: Climate significance and environmental Impact, Geophysical Monograph Series, vol 134. AGU, Washington, D. C.

Jiang T, Deng Y (2011) Downstream modulation of North Pacific atmospheric river activity by East Asian cold surges. Geophys Res Lett 38, DOI {10.1029/2011GL049462}

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo K, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Amer Meteorol Soc 77(3):437–471

Kim HM, Alexander MA (2015) ENSO's modulation of water vapor transport over the Pacific-North American region. J Clim 28(9):3846–3856, DOI {10.1175/JCLI-D-14-00725.1}

Kim J, Waliser DE, Neiman PJ, Guan B, Ryoo JM, Wick GA (2013) Effects of atmospheric river landfalls on the cold season precipitation in California. Clim Dyn 40(1-2):465–474, DOI {10.1007/s00382-012-1322-3}

Knippertz P, Wernli H (2010) A Lagrangian climatology of tropical moisture exports to the Northern Hemispheric extratropics. J Clim 23(4):987–1003, DOI

{10.1175/2009JCLI3333.1}

Knippertz P, Wernli H, Glaeser G (2013) A global climatology of tropical moisture exports. J Clim 26(10):3031–3045, DOI {10.1175/JCLI-D-12-00401.1}

Lavers DA, Villarini G (2013) The nexus between atmospheric rivers and extreme precipitation across Europe. Geophys Res Lett 40(12):3259–3264, DOI {10.1002/grl.50636}

Lavers DA, Villarini G (2015) The contribution of atmospheric rivers to precipitation in Europe and the United States. Journal of Hydrology 522:382–390, DOI {10.1016/j.jhydrol.2014.12.010}

Lavers DA, Allan RP, Wood EF, Villarini G, Brayshaw DJ, Wade AJ (2011) Winter floods in Britain are connected to atmospheric rivers. Geophys Res Lett 38, DOI {10.1029/2011GL049783}

Lavers DA, Villarini G, Allan RP, Wood EF, Wade AJ (2012) The detection of atmospheric rivers in atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation. J Geophys Res Atmos 117, DOI {10.1029/2012JD018027}

Leung LR, Qian Y (2009) Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model. Geophys Res Lett 36, DOI {10.1029/2008GL036445}

Neiman P, Persson P, Ralph F, Jorgensen D, White A, Kingsmill D (2004) Modification of fronts and precipitation by coastal blocking during an intense landfalling winter storm in southern California: Observations during CALJET. Mon Weather Rev 132(1):242–273, DOI {10.1175/1520-0493(2004)132⟨0242:MOFAPB⟩2.0.CO;2}

Neiman PJ, Ralph FM, Wick GA, Lundquist JD, Dettinger MD (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 Hydrometeorol 9(1):22–47, DOI {10.1175/2007JHM855.1}

Newell R, Newell N, Zhu Y, Scott C (1992) Tropospheric rivers - A pilot-study. Geophys Res Lett 19(24):2401–2404, DOI {10.1029/92GL02916}

Newman M, Kiladis GN, Weickmann KM, Ralph FM, Sardeshmukh PD (2012) Relative contributions of synoptic and low-frequency eddies to time-mean atmospheric moisture transport, including the role of atmospheric rivers. J Clim (25):73417361, DOI {http://dx.doi.org/10.1175/JCLI-D-11-00665.1}

Payne AE, Magnusdottir G (2014) Dynamics of landfalling atmospheric rivers over the North Pacific in 30 Years of MERRA reanalysis. J Clim 27(18):7133–7150, DOI {10.1175/JCLI-D-14-00034.1}

Poli P, Hersbach H, Tan D, Dee D, Thépaut JN, Simmons A, Peubey C, Laloyaux P, Komori T, Berrisford P, Dragani R, Trémolet Y, Hólm E, Bonavita M, Isaksen L, Fisher M (2013) The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (era-20c). Tech. rep., ERA Report Series 14

Qian B, Corte-Real J, Xu H (2000) Is the North Atlantic Oscillation the most important atmospheric pattern for precipitation in Europe? J Geophys Res Atmos 105(D9):11,901–11,910, DOI {10.1029/2000JD900102}

Ralph F, Neiman P, Wick G (2004) Satellite and CALJET aircraft observations of atmospheric rivers over the eastern north pacific ocean during the winter of 1997/98. Mon Weather Rev 132(7):1721–1745, DOI {10.1175/1520-0493(2004)132⟨1721:SACAOO⟩2.0.CO;2}

Ralph FM, Coleman T, Neiman PJ, Zamora RJ, Dettinger MD (2013) Observed impacts of duration and seasonality of atmospheric-river landfalls on soil moisture and runoff in coastal Northern California. J Hydrometeorol 14(2):443–459, DOI {10.1175/JHM-D-12-076.1}

Ramos AM, Trigo RM, Liberato MLR, Tome R (2015) Daily precipitation extreme events in the Iberian Peninsula and its association with atmospheric Rivers. J Hydrometeorol 16(2):579–597, DOI {10.1175/JHM-D-14-0103.1}

Sodemann H, Stohl A (2013) Moisture origin and meridional transport in atmospheric rivers and their association with multiple cyclones. Mon Weather Rev 141(8):2850–2868, DOI {10.1175/MWR-D-12-00256.1}

Sterl A (2004) On the (in)homogeneity of reanalysis products. J Clim 17(19):3866–3873, DOI {10.1175/1520-0442(2004)017⟨3866:OTIORP⟩2.0.CO;2}

Sturaro G (2003) A closer look at the climatological discontinuities present in the NCEP/NCAR reanalysis temperature due to the introduction of satellite data. Clim Dyn 21(3-4):309–316, DOI {10.1007/s00382-003-0334-4}

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Amer Meteorol Soc 93(4):485–498, DOI {10.1175/BAMS-D-11-00094.1}

Trenberth K, Hurrell J (1994) Decadal atmosphere-ocean variations in the Pacific. Clim Dyn 9(6):303–319, DOI {10.1007/BF00204745}

Trigo RM, Pozo-Vázquez D, Osborn TJ, Castro-Díez Y, Gámiz-Fortis S, Esteban-Parra MJ (2004) North Atlantic oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula. Int J Climatol 24(8):925–944, DOI {10.1002/joc.1048}

Trenberth K, Branstator G, Karoly D, Kumar A, Lau N, Ropelewski C (1998) Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. J Geohphys Res Atmos 103(C7):14,291–14,324, DOI {10.1029/97JC01444}

Warner MD, Mass CF, Salathé Jr EP (1998) Changes in winter atmospheric rivers along the North American west coast in CMIP5 climate models. J Hydrometeorol 16:118-128, DOI {http://dx.doi.org/10.1175/JHM-D-14-0080.1}

Zhang Y, Wallace J, Battisti D (1997) ENSO-like interdecadal variability: 1900-93. J Clim 10(5):1004–1020, DOI {10.1175/1520-0442(1997)010⟨1004:ELIV⟩2.0.CO;2}

Zhu Y, Newell R (1994) Atmospheric rivers and bombs. Geophys Res Lett 21(18):1999–2002, DOI {10.1029/94GL01710}

Zhu Y, Newell R (1998) A proposed algorithm for moisture fluxes from atmospheric rivers. Mon Weather Rev 126(3):725–735, DOI {10.1175/1520-0493(1998)126⟨0725:APAFMF⟩2.0.CO;2}

Table 1 Considered reanalysis datasets, 6h-instantaneous values are applied in any case. Listed are the acronyms used throughout the study, the full names, horizontal resolutions (lat. × lon.), reference publications and the number of runs conducted for each reanalysis. The ensemble-mean data from NOAA-20C are used in the present study.

Acronym	Full Name	Resolution	Reference	Nr. runs
NCEP/NCAR	NCEP/NCAR Reanalysis 1	$2.5° \times 2.5°$	Kalnay et al (1996)	1 run
NOAA-20C	NOAA CIRES 20th-Century Reanalysis v2	$2° \times 2°$	Compo et al (2011)	56-member ensemble
ERA-20C	ECMWF ERA-20C Reanalysis	$1.125° \times 1.125°$	Poli et al (2013)	1 run
ERA-Interim	ECMWF ERA-Interim Reanalysis	$0.75° \times 0.75°$	Dee et al (2011)	1 run

Table 2 List of multipliers used for multiplication with qu and qv at a given pressure level p. Horizontal bars indicate that the data at the corresponding pressure level were not available from the data provider. See text for more details.

p	$NCEP/NCAR$	$NOAA-20C$
300	5000	2500
350	-	5000
400	10000	5000
450	-	5000
500	10000	5000
550	-	5000
600	10000	5000
650	-	5000
700	12500	5000
750	-	5000
800	-	5000
850	11250	5000
900	-	5000
925	7500	-
950	-	5000
1000	3750	2500

Table 3 Considered large-scale atmopsheric circulation indices

Used Acronym	Description	Provider	Citation
H-NAO	J. Hurrell's NAO index based on PCA and SLP fields	UCAR	Hurrell et al (2003)
station-NAO	J. Hurrell's NAO index based on SLP station values	"	"
CPC-NAO	NAO index based on rotated PCA and geopotential height fields	CPC	Barnston and Livezey (1987)
EA	East Atlantic Pattern index	"	"
SCAND	Scandinavian Pattern index	"	"
EA/WR	East Atlantic / Western Russia index	"	"
PNA	Pacific-North American pattern index	"	"
WP	West Pacific Index	"	"
NP	North Pacific / Aleutian Low index	UCAR	Trenberth and Hurrell (1994)

Table 4 The 6 percentile combinations used for AR detection and tracking. P_d is the percentile threshold used for detection at the region of AR-arrival and P_t is the percentile threshold used along the track of the AR.

Number	P_d	P_t
1	85	75
2	85	80
3	85	85
4	90	75
5	90	80
6	90	85

Table 5 Rank correlation coefficient (rounded to the next integer $\times 100$) measuring the link between AR-counts and Climate Prediction Centre's NAO index during the ONDJFM-season, with and without considering the Lavers et al (2012) persistence criterion. Results are for the experimental set-ups most closely refelecting Lavers et al (2012) (setup 1) and Ramos et al (2015) (setup 2); see text for more details. Significant results ($\alpha = 0.05$, two-sided t-test) are printed in bold. a) results for an AR length criterion of $>$ 3000 km, b) results for $>$ 2000 km.

a) 3000 km	Setup 1	Setup 2
Without persistence criterion	**48**, **49**, **46**, **45**, **45**, **40**	**-37**, **-38**, **-37**, **-39**, **-40**, **-41**
With persistence criterion	17, 15, 14, 24, 19, 26	**-26**, **-35**, -21, **-32**, -23, **-34**
b) 2000 km		
Without persistence criterion	**52**, **54**, **55**, **53**, **53**, **51**	**-33**, **-33**, **-32**, **33**, **-35**, **-36**
With persistence criterion	**44**, **40**, **37**, 34, **42**, **48**	-18, **-31**, -20, **-31**, -25, **-32**

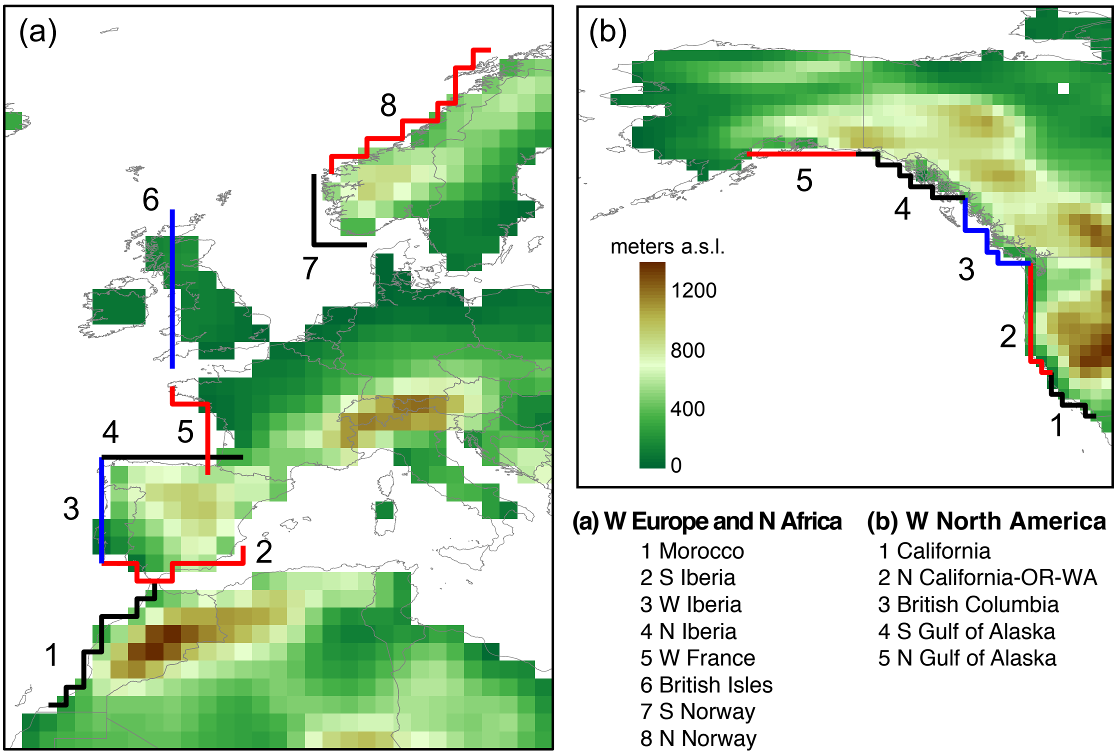

Fig. 1 Target regions used for AR-detection and tracking for the case of ERA-20C. Also shown is the corresponding orography. The detection "barriers" used for the 3 remaining reanalyses are in the direct vicinity of those shown here.

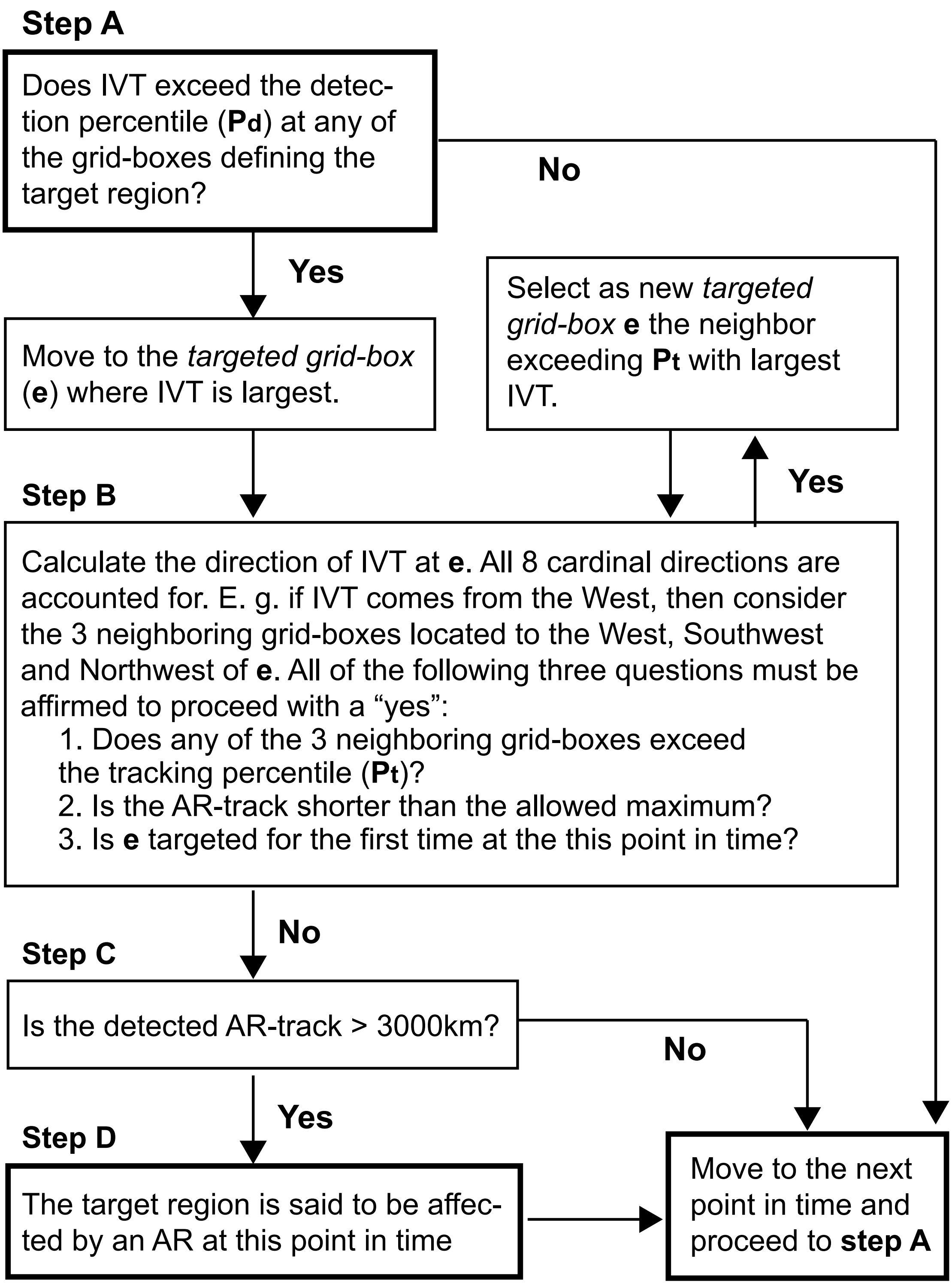

Fig. 2 Schematic overview of the proposed AR detection and tracking algorithm

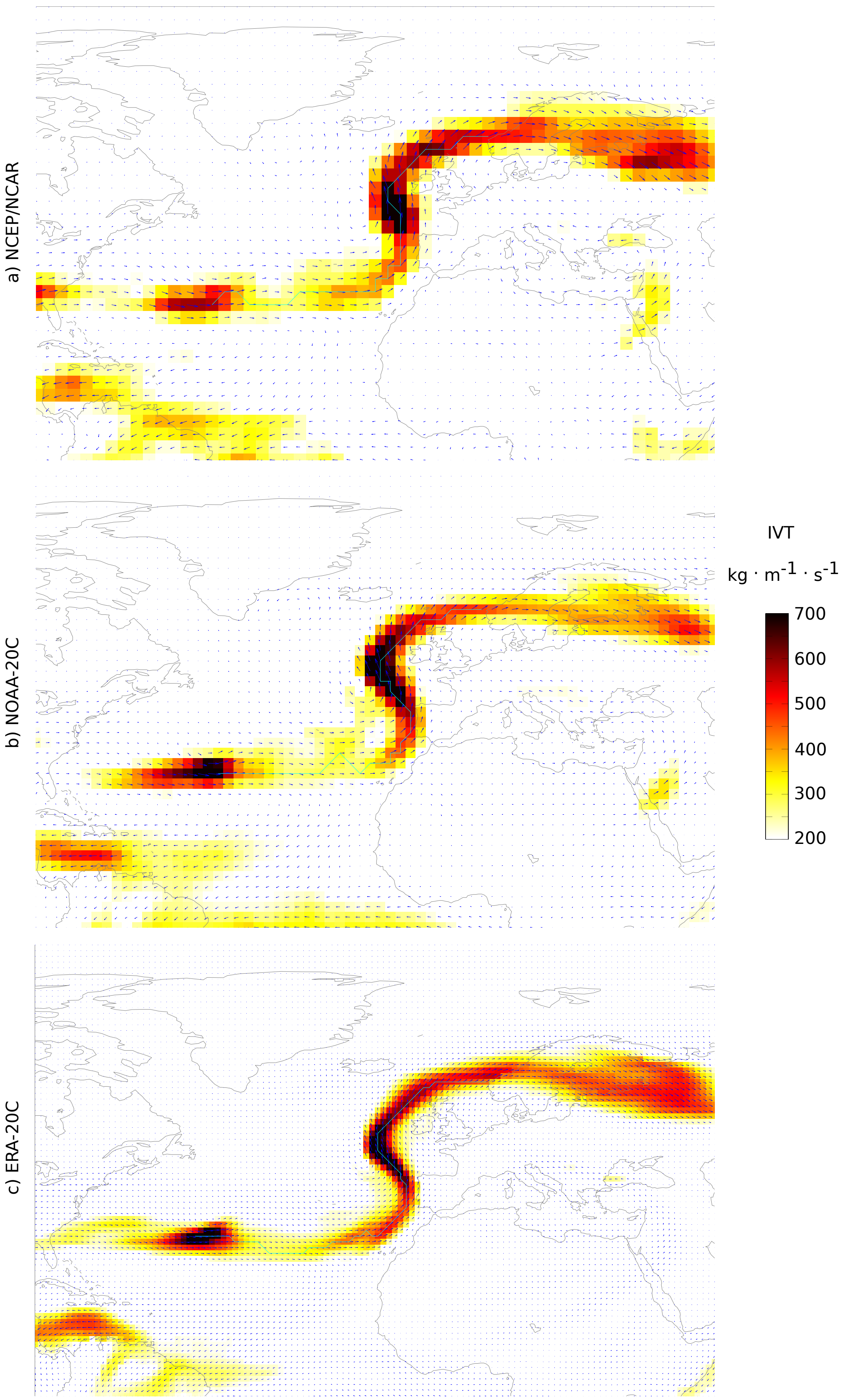

Fig. 3 Illustrative example for an AR affecting southern Norway on 11 January 1971 OO UTC for a) NCEP/NCAR, b) NOAA-20C and c) ERA-20C. Colour shadings and vector lengths are proportional to the strength of the vertically integrated water vapour flux. The direction of the flow is indicated by the orientation of the vectors. The cyan line represents the AR-track found by the algorithm.

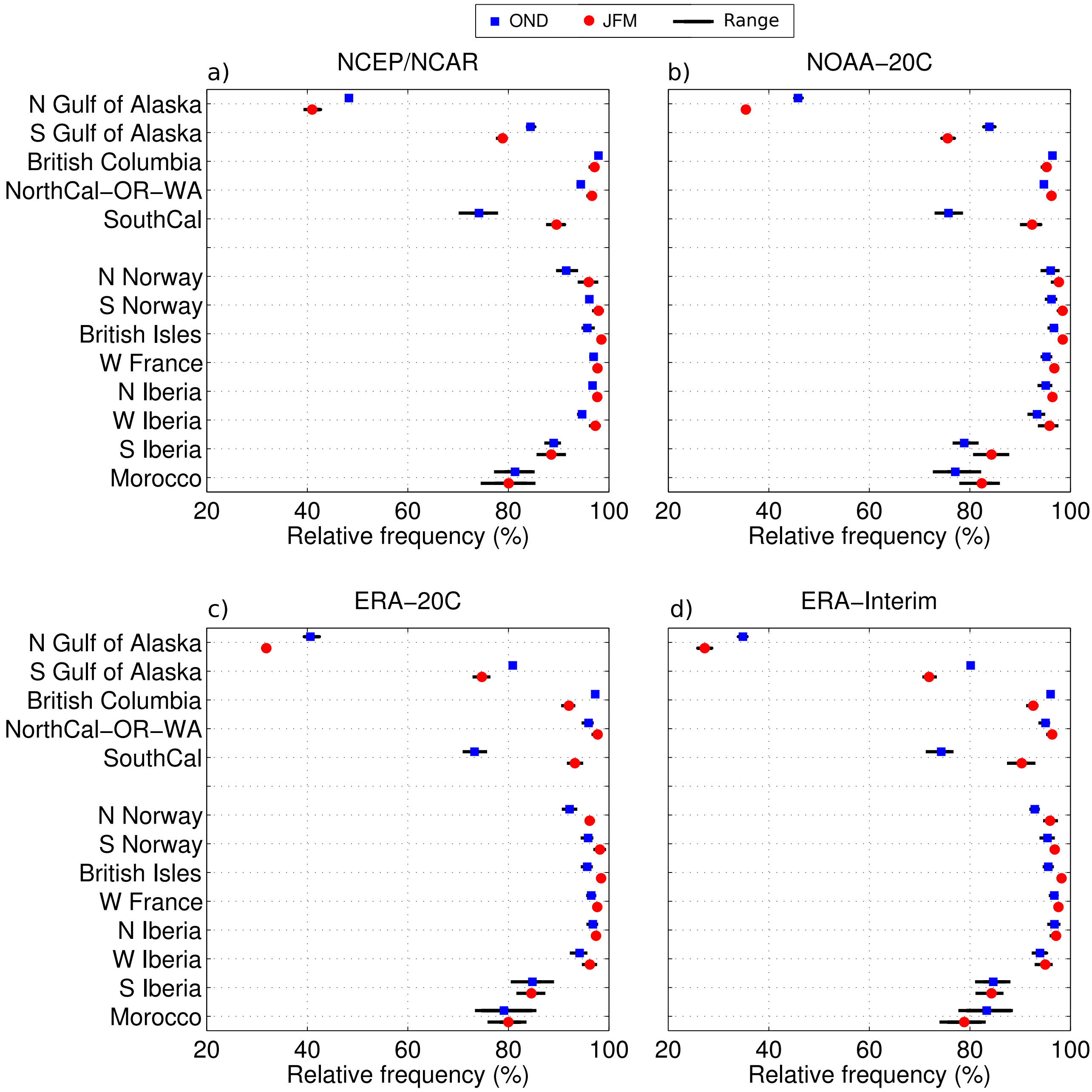

Fig. 4 Fraction of ARs that are detected if the capability to track towards the north, north-east, east or south-east is intentionally disabled (see Equation 6) for a) NCEP/NCAR, b) NOAA-20C, c) ERA-20C and d) ERA-Interim. Results are for the October-to-December (OND) and January-to-March (JFM) seasons, considering the time period 1979-2010. Squares / circles and errorbars refer to the mean and range of the 6 results obtained from the 6 considered percentile-threshold combinations listed in Table 4, i. e. refer to the method-related uncertainty of the results.

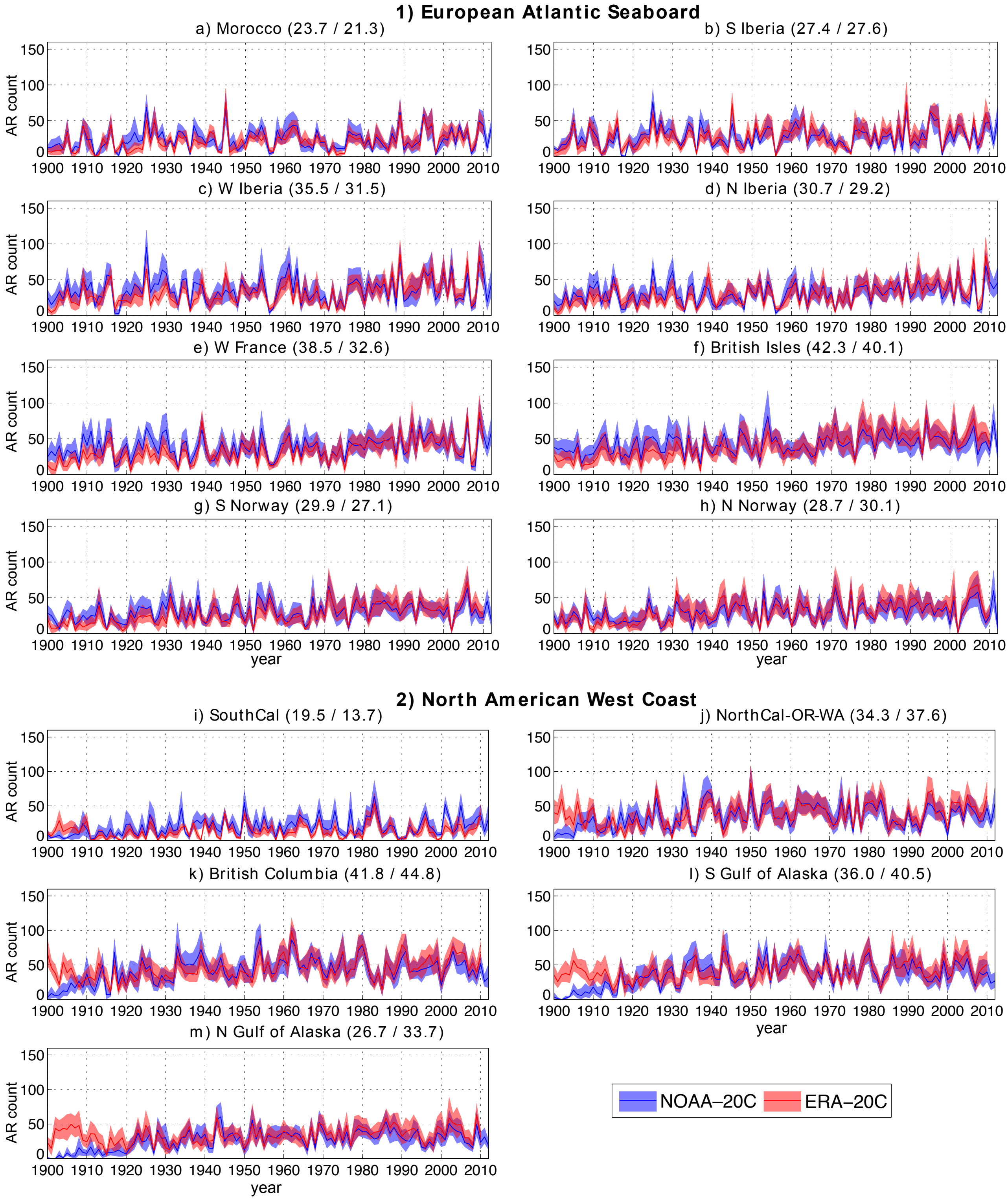

Fig. 5 Year-to-year sequence of seasonal AR-occurrence counts during the OND season for NOAA-20C (blue) and ERA-20C (red). The lines and shadings refer to the mean and range of the percentile sample (see Table 4), i. e. refer to the method-related uncertainty of the results. Displayed are 1900-2012 time series for NOAA-20C and 1900-2010 time series for ERA-20C.

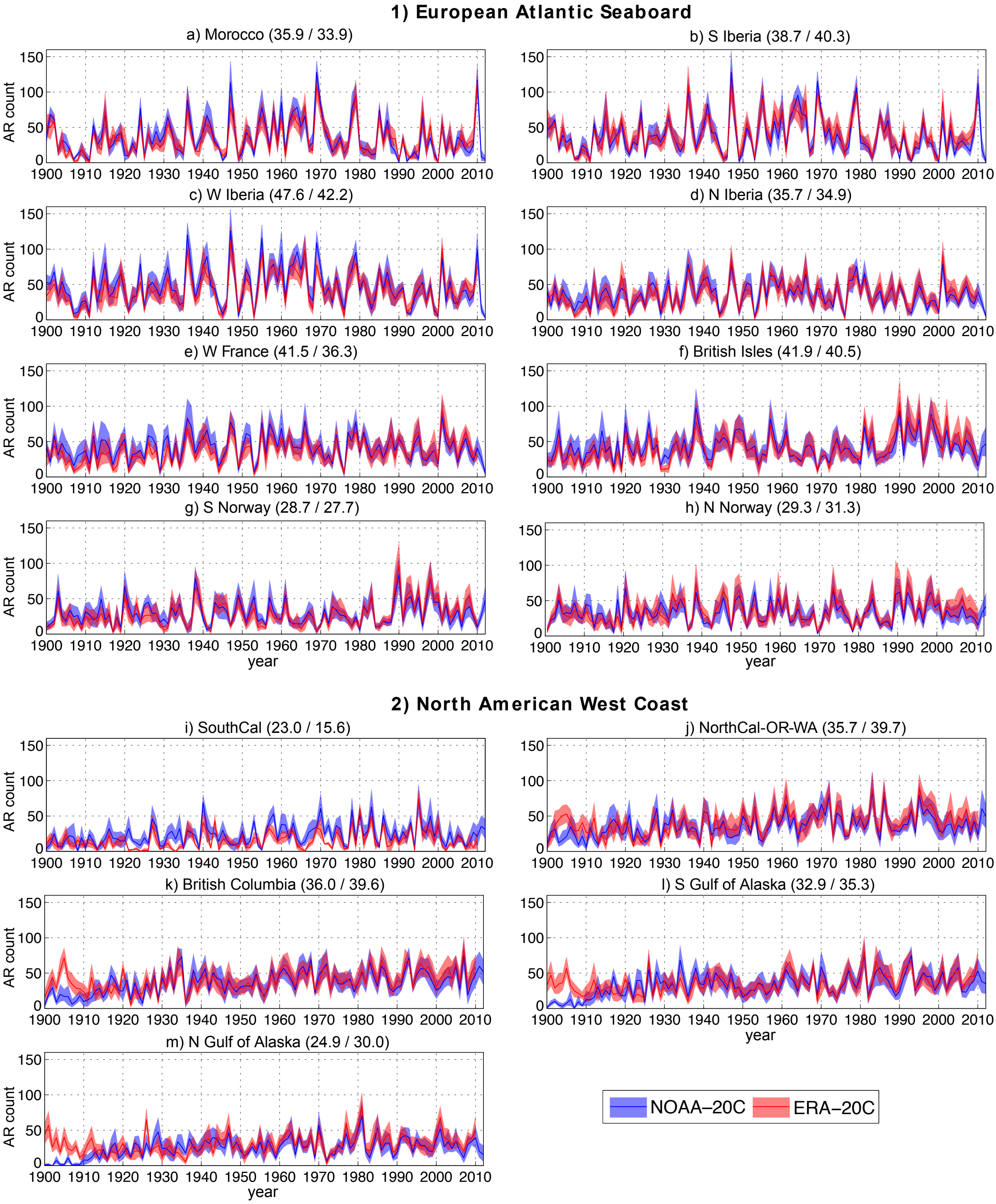

Fig. 6 As Figure reff.f04, but for the JFM-season.

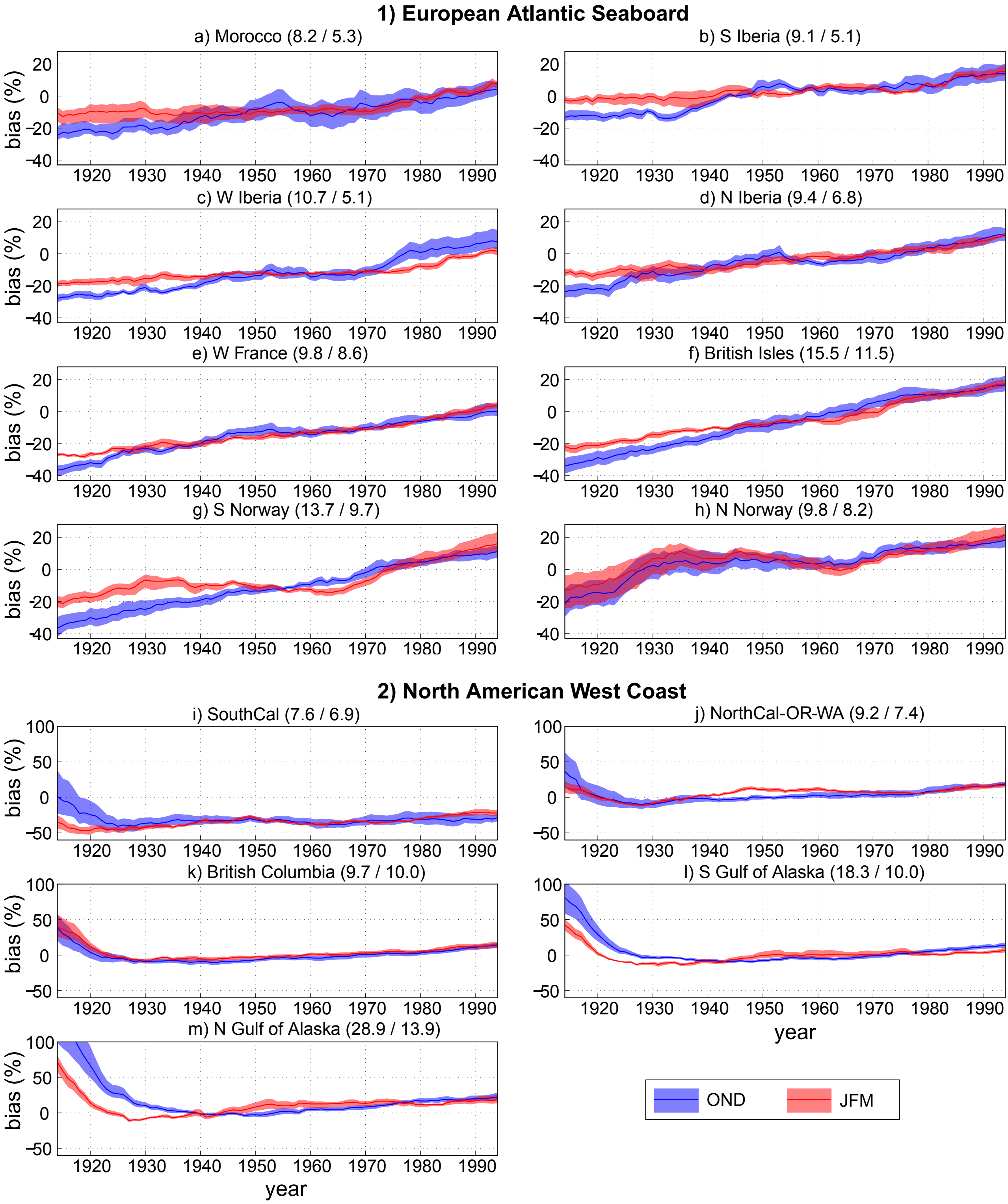

Fig. 7 Relative difference in the climatological mean AR-occurrence counts (NOAA-20C minus ERA-20C with respect to NOAA-20, in %, see Equation 4) along the course of the 20th century, obtained by applying a 31-year running window starting in 1900-1931 and ending in 1900-2010. Along the x-axis of each panel, the centre year of each sub-period is displayed. Lines and shadings refer to the mean and range of the percentile sample (see Table 4), i. e. refer to the method-related uncertainty of the results; blue = OND season, red = JFM season. To measure the stationarity of the the bias, the standard deviation (std) of the 81 mean bias values (as depicted by the lines) is displayed in the header of each panel. The first number refers to std for OND, the second to std for JFM. Note the distinct scale of the y-axes for Europe/North Africa and western North America.

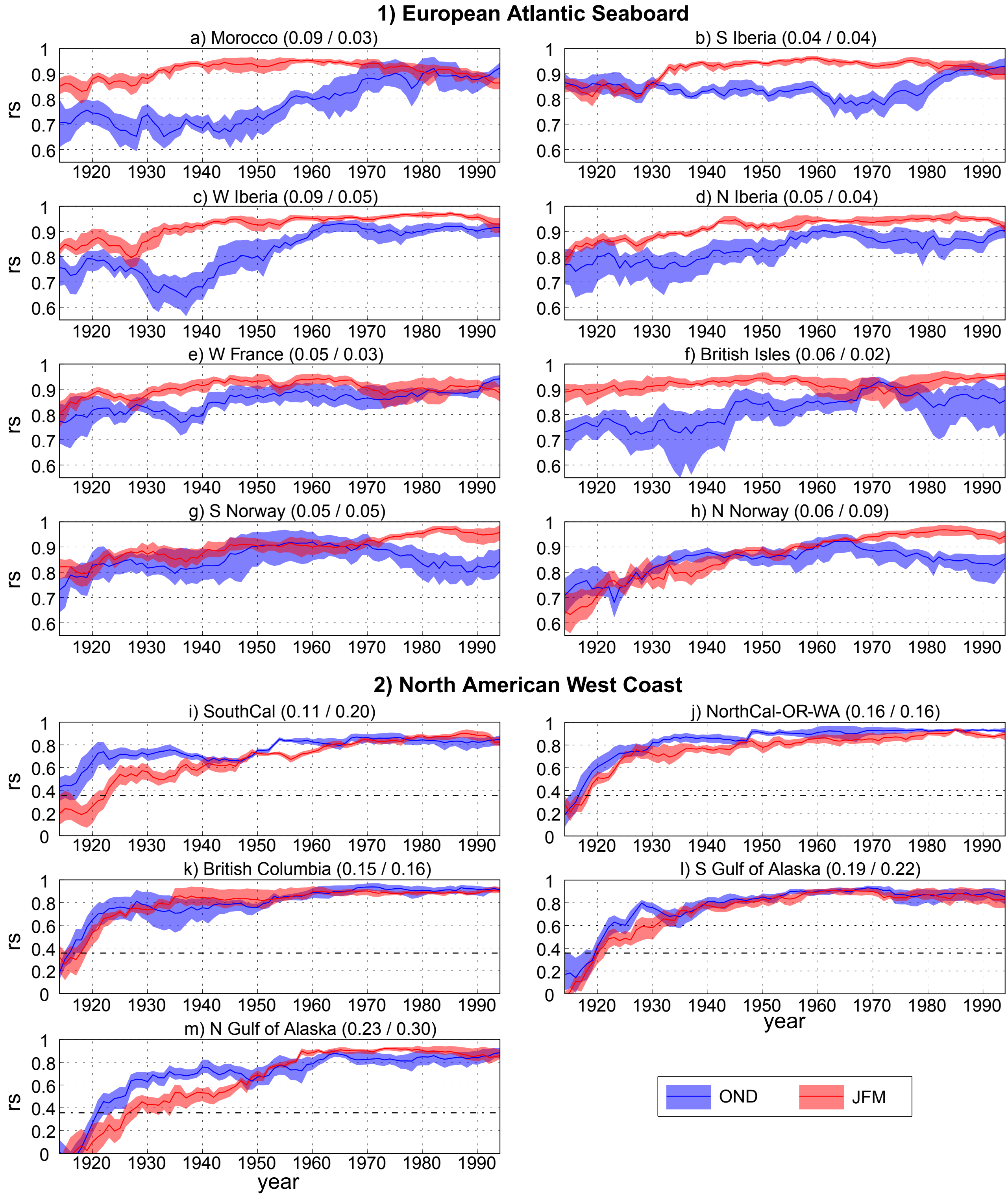

Fig. 8 As Figure 8 but for the rank correlation coefficient (rs) between the seasonal AR-occurrence counts from NOAA-20C and ERA-20C. Dashed horizontal lines mark the critical values below / above which rs is significant at a test-level of 5%. Note the distinct scale of the y-axes for Europe/North Africa and western North America.

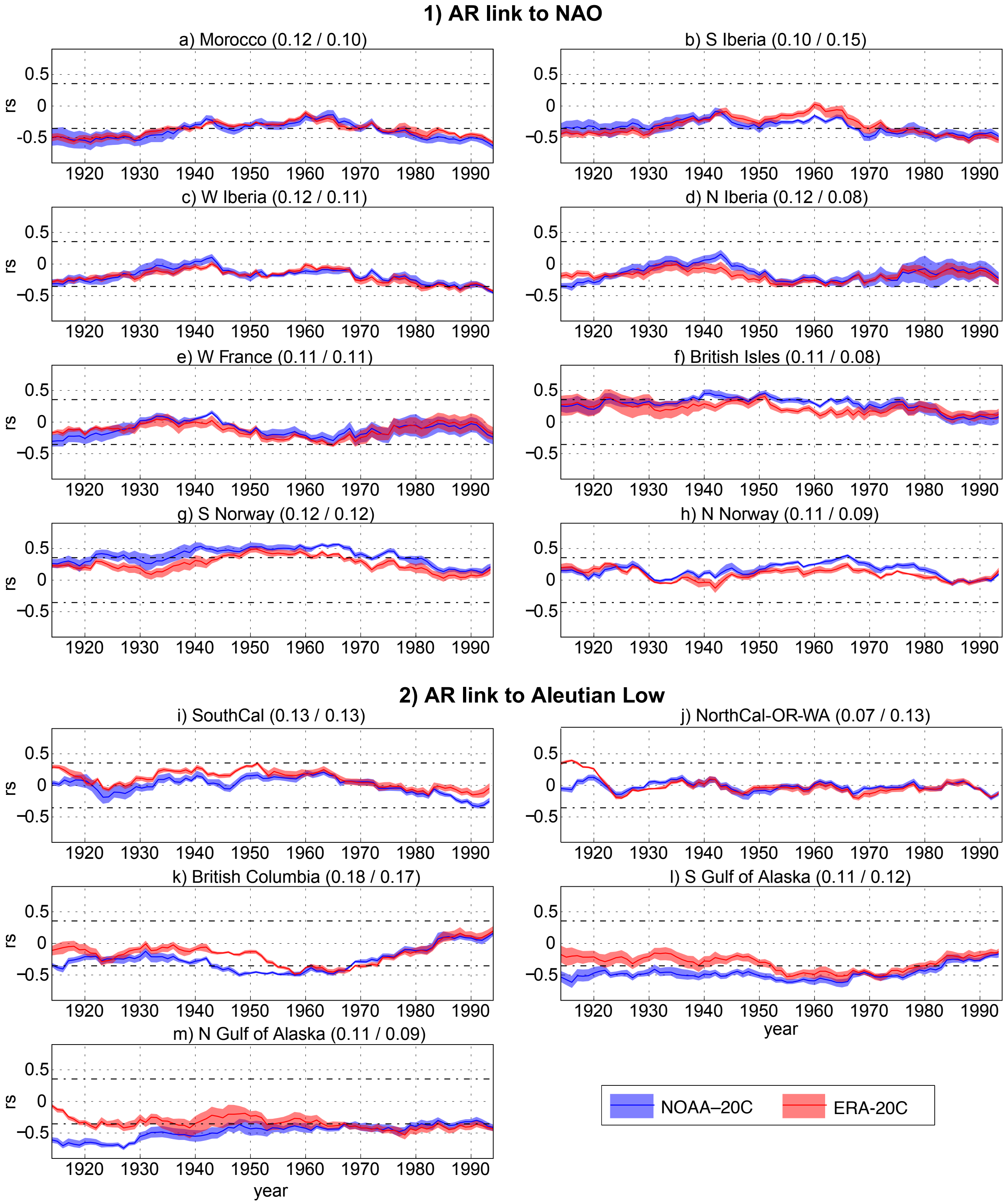

Fig. 9 As Figure 8 but for the rank correlation coefficient (rs) between the OND AR-occurrence counts in NOAA-20C (blue) or ERA-20C (read) and the station-based NAO or North Pacific index. Dashed horizontal lines mark the critical values below / above which rs is significant at a test-level of 5%.

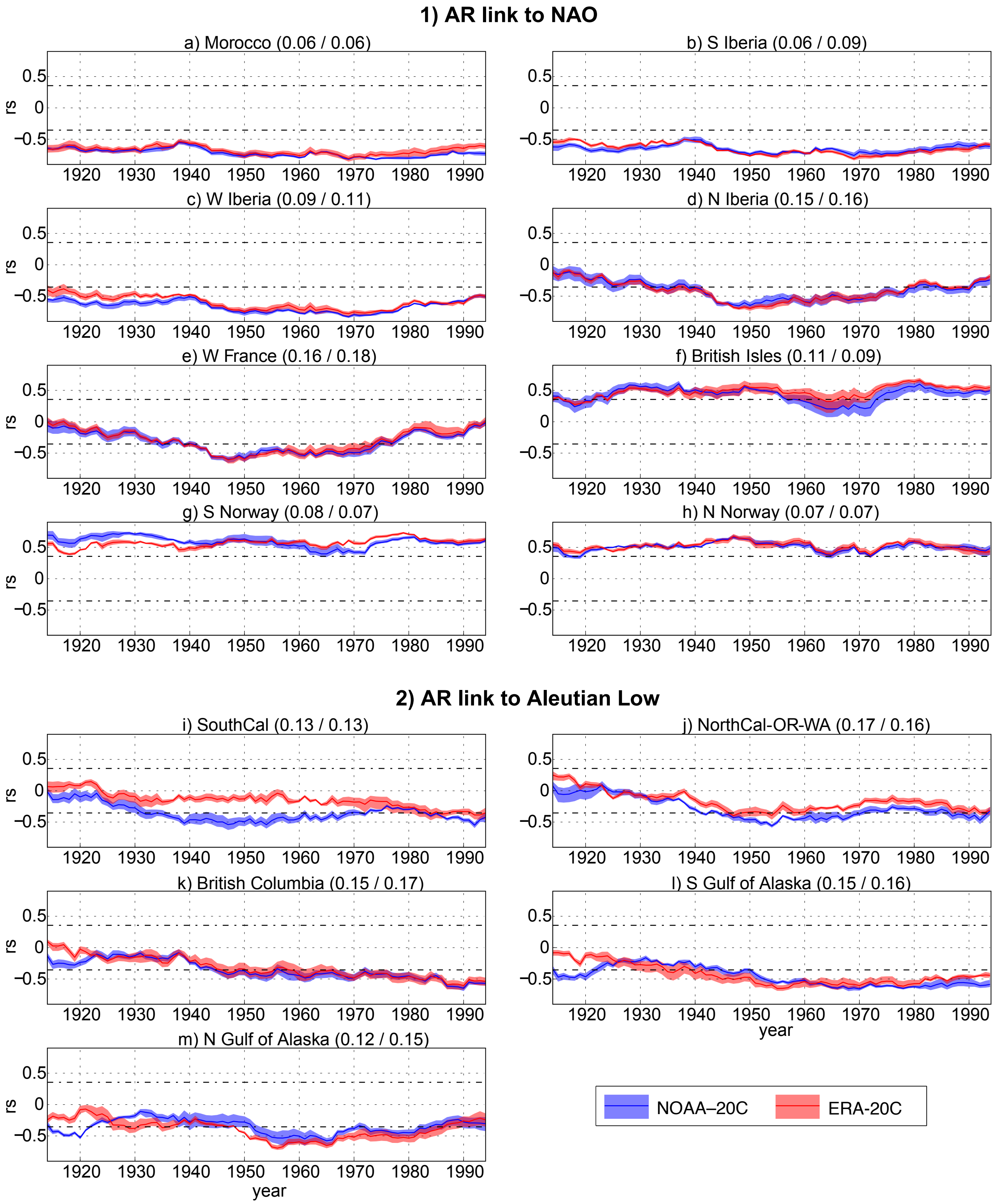

Fig. 10 As Figure 9, but for the JFM-season.

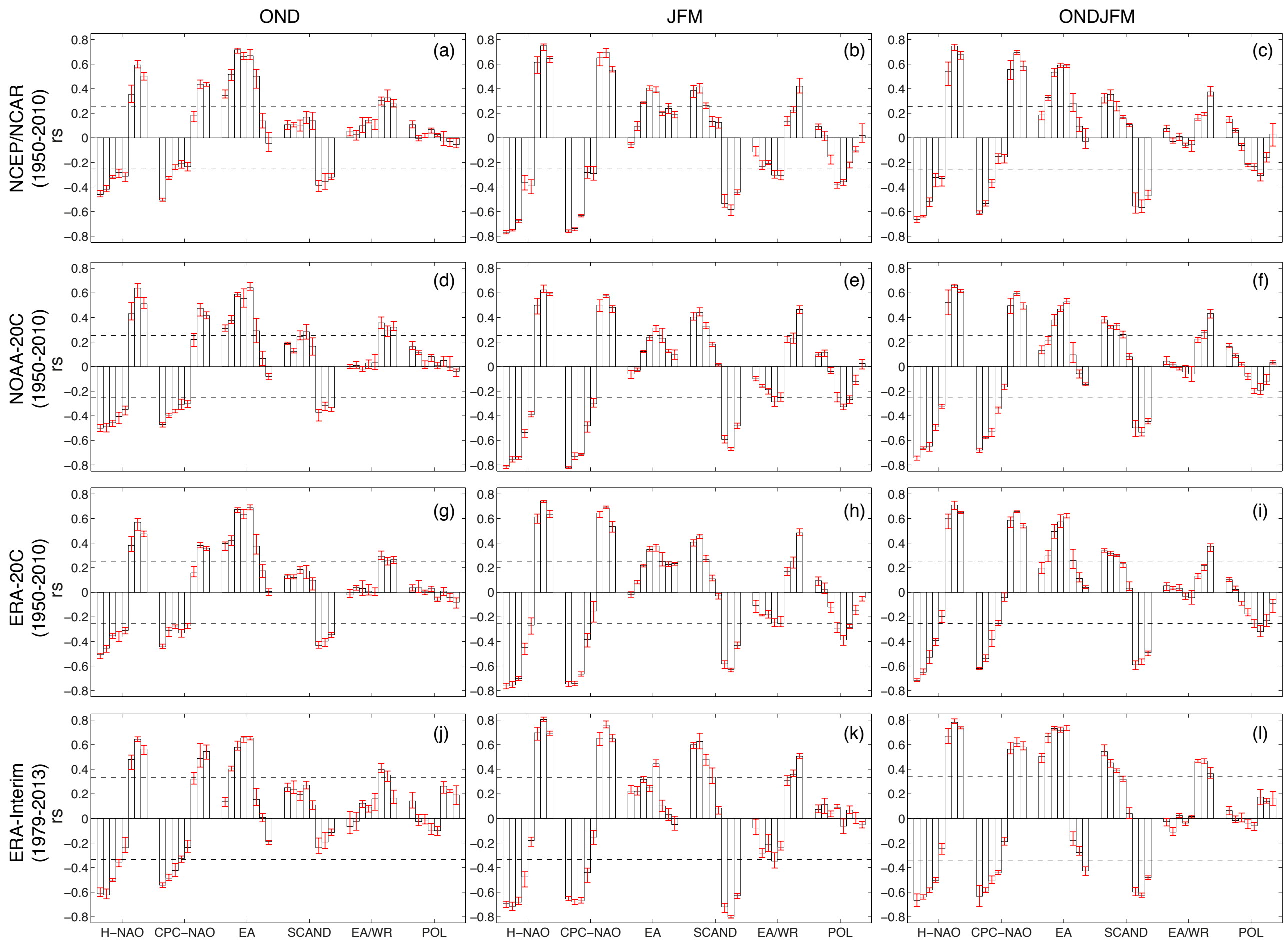

Fig. 11 Rank correlation coefficient (rs) between seasonal AR-occurrence counts and seasonal-mean atmospheric circulation indices for the 8 considered target regions in Europe/North Africa ordered from the South to the North, i.e. the first bar or each group of bars refers to Morocco and the last to northern Norway respectively. Results are for NCEP/NCAR, NOAA-20C, ERA-20C and ERA-Interim (each row corresponds to a dataset) and for OND, JFM, ONDJFM and ONDFJM considering the persistence criterion (each column corresponds to a season-definition). Bars and errorbars refer to the mean and range of the 6 results obtained from the 6 considered percentile-threshold combinations, i. e. refer to the method-related uncertainty of the results. Dashed horizontal lines mark the critical values below / above which r is significant at a test-level of 5%. Results are for 1950-2010 except for ERA-Interim in which case they are for 1979-2013.

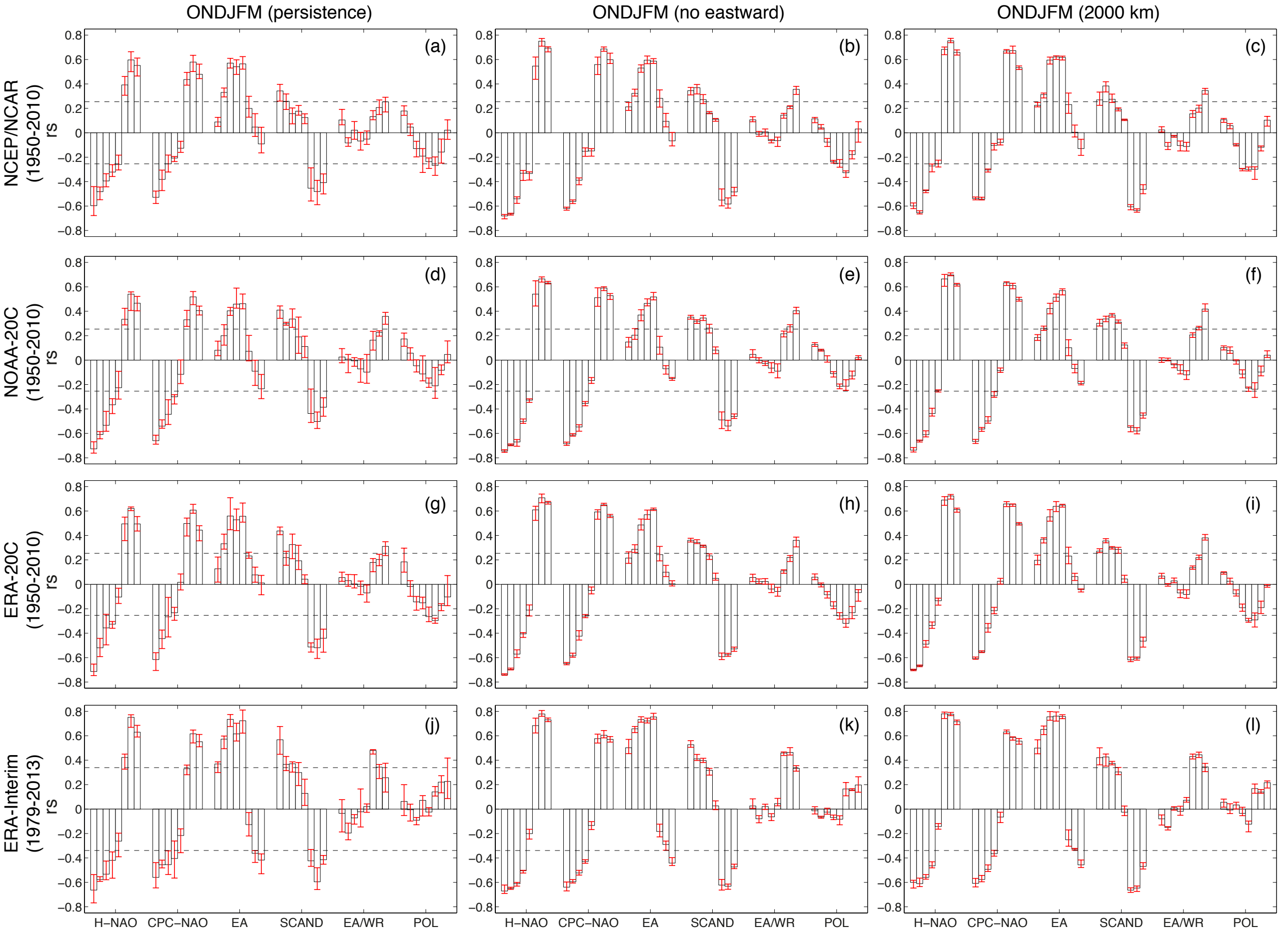

Fig. 12 As Figure 11, but for AR-counts (first column) calculated upon persistent events only, (second column) obtained without eastward tracking capability and (third column) obtained with a length criterion of > 2000 km.

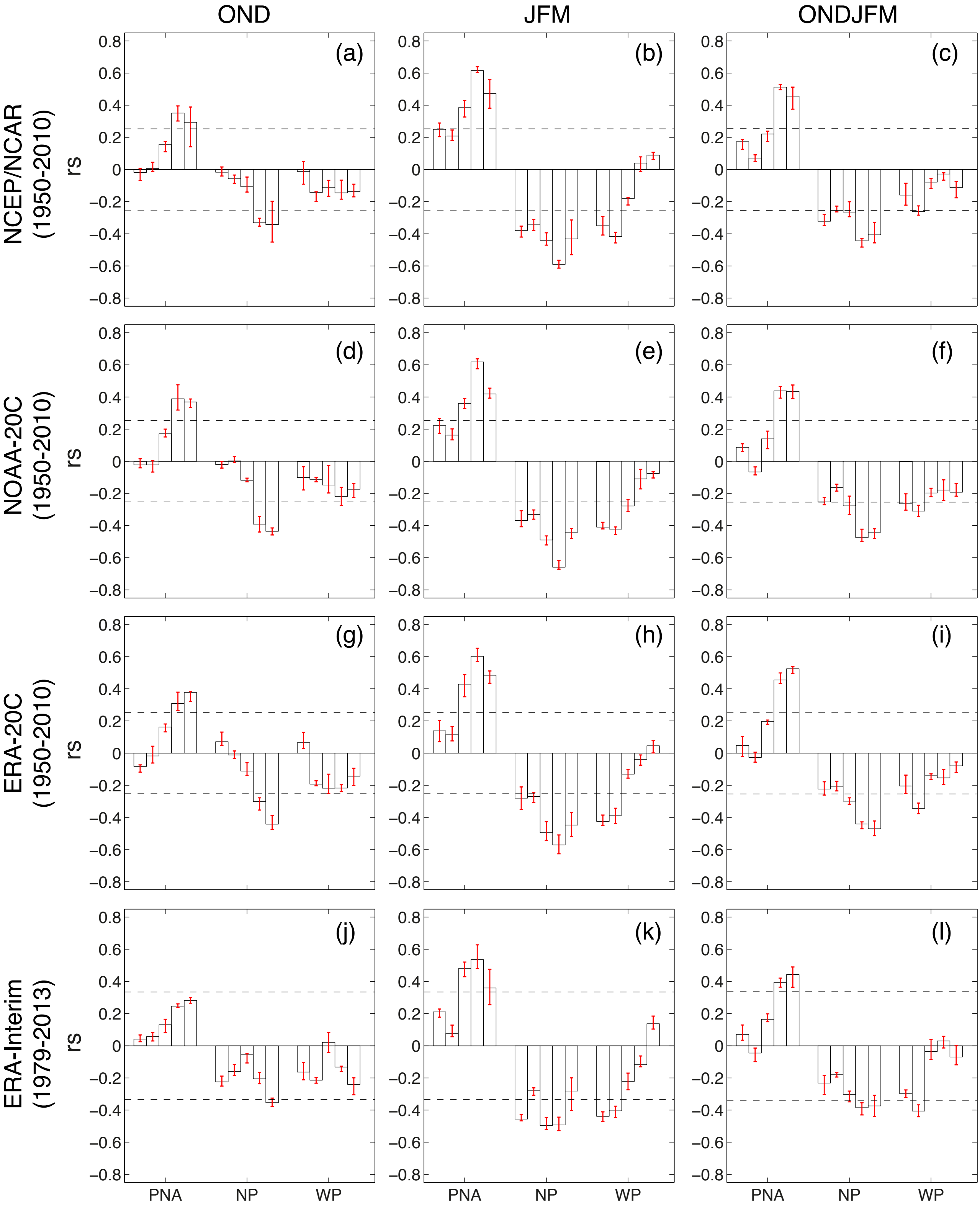

Fig. 13 As Figure 11, but for rs between seasonal AR-occurrence counts in the 5 regions along the west coast of North America and the circulation coefficients relevant there, PNA = Pacific-North American, NP = North Pacific, WP = West Pacific. The first bar or each group of bars refers to southern California, the last one to the northern Gulf of Alaska.

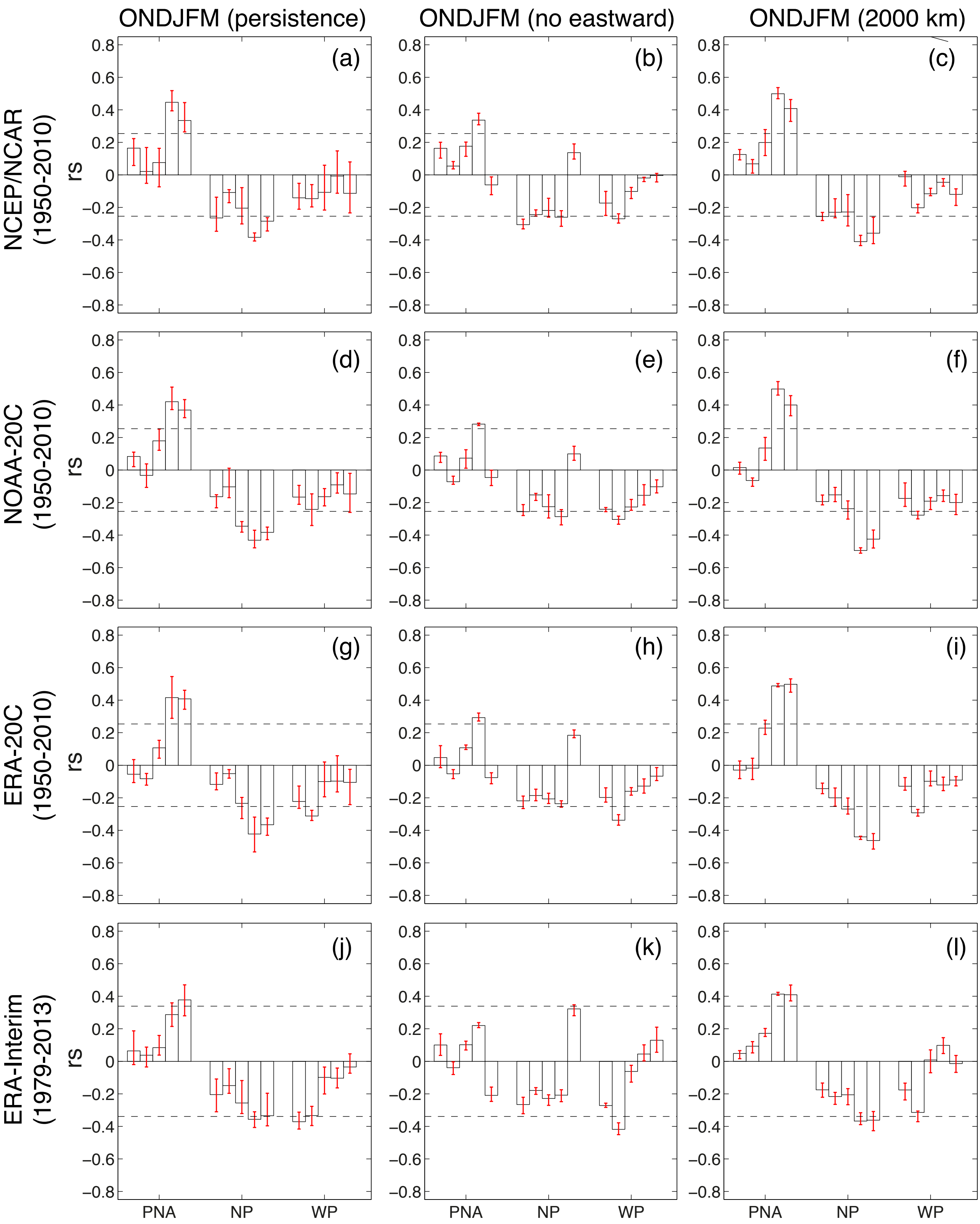

Fig. 14 As Figure 13, but for AR-counts (first column) calculated upon persistent events only, (second column) obtained without eastward tracking capability and (third column) obtained with a length criterion of > 2000 km.