AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2015JD023586

Key Points:
- CMIP5 model performance is variable in resolving landfalling ARs
- In RCP 8.5, AR frequency broadens equatorward of peak historical frequency
- Equatorward increase in moisture flux dominated by dynamical response

Correspondence to:
A. E. Payne,
aepayne@uci.edu

Citation:
Payne, A. E., and G. Magnusdottir (2015), An evaluation of atmospheric rivers over the North Pacific in CMIP5 and their response to warming under RCP 8.5, *J. Geophys. Res. Atmos.*, *120*, 11,173–11,190, doi:10.1002/2015JD023586.

Received 27 APR 2015
Accepted 16 OCT 2015
Accepted article online 20 OCT 2015
Published online 13 NOV 2015

©2015. American Geophysical Union. All Rights Reserved.

An evaluation of atmospheric rivers over the North Pacific in CMIP5 and their response to warming under RCP 8.5

Ashley E. Payne[1] and Gudrun Magnusdottir[1]

[1]Department of Earth System Science, University of California, Irvine, California, USA

Abstract Landfalling atmospheric rivers over the North Pacific are evaluated in historical (1980–2005) simulations from 28 models participating in the fifth phase of the Coupled Model Intercomparison Project (CMIP5) and compared to the Modern-Era Retrospective Analysis for Research and Applications and ERA-Interim reanalyses. Landfalling dates are identified as having spatially elongated high wind and moisture features in physical proximity to the coastline. Model performance relative to reanalysis is found to be quite variable. The majority can resolve the spatial structure of landfalling events, but few correctly resolve the frequency distribution, interannual variability in number and amplitude of moisture flux, and median landfalling latitude. The response of a subset of high-performing models to projected warming at the end of the 21st century is investigated using Representative Concentration Pathway (RCP) 8.5 (2070–2100) projections. Selected models show a broadening of the frequency distribution, with the largest increase in frequency occurring equatorward of peak historical frequency. While there is a robust increase in AR-related moisture transport compared to the historical period, this increase is not linked to more extreme AR events within the context of the projected future climate. The occurrence of years with many AR dates is greater at the end of the 21st century than during the historical period. The equatorward increase in peak historical frequency is colocated with increases in the 850 and 250 hPa zonal winds. The response in moisture flux to warming is mostly thermodynamic, but equatorward of its peak distribution, it is dominated by a dynamic response.

1. Introduction

Despite their filamentary geometry, atmospheric rivers (ARs) play an important role in meridional moisture transport and in the global redistribution of heat from the tropics [*Zhu and Newell*, 1998; *Bao et al.*, 2006; *Newman et al.*, 2012]. They are characterized by high moisture content and a strong low-level jet [*Ralph et al.*, 2004]. ARs are closely associated with the warm conveyor belt within the warm sector of extratropical cyclones and are sustained by prefrontal moisture convergence [*Bao et al.*, 2006; *Stohl et al.*, 2008; *Sodemann and Stohl*, 2013; *Dacre et al.*, 2015]. When they cross over land (so-called landfall), they are a major source of cold-season precipitation [*Neiman et al.*, 2008; *Dettinger et al.*, 2011; *Dettinger*, 2013]. The extreme precipitation and flooding that sometimes accompany landfalling ARs over coastal areas with steep topography can have severe socioeconomic consequences.

Given the essential role ARs play in the hydrological cycle and extratropical moisture transport [*Newman et al.*, 2012], the influence of projected climate change on AR behavior must be considered. Lower tropospheric atmospheric moisture is expected to increase with warming, in-line with the Clausius-Clapeyron relation (see review in *Meehl et al.* [2007]). The increase in atmospheric moisture with warming has been found to result in a robust increase in the horizontal moisture transport in model projections, with little change in vertical mass fluxes [*Held and Soden*, 2006; *Lavers et al.*, 2015].

The association of ARs with extratropical cyclones and the midlatitude storm track must also be considered. In the Northern Hemisphere, amplified surface warming at high latitudes decreases the surface meridional temperature gradient, while the warming of the tropical upper troposphere increases the upper level temperature gradient. These changes are associated with an increase in upper tropospheric eddy kinetic energy and the vertical expansion of the climatological storm track [*Yin*, 2005; *Chang et al.*, 2012]. ARs over the eastern North Pacific are closely related to Rossby wave propagation and breaking [*Ryoo et al.*, 2013; *Payne and Magnusdottir*, 2014]. *Payne and Magnusdottir* [2014] showed a link between intense landfalling ARs and the development of anticyclonic Rossby wave breaking over the eastern North Pacific. The association of

Rossby wave breaking with the exit region of the midlatitude jet is, therefore, of importance in understanding the behavior of landfalling ARs. Over the North Pacific, *Barnes and Polvani* [2013] have identified an increased meridional wobbling of the jet at the end of the 21st century in Representative Concentration Pathway (RCP) 8.5 projections from the fifth phase of the Coupled Model Intercomparison Project (CMIP5). Increased meridional variability of the jet has implications for the location of breaking and is of interest in investigating the response of ARs to warming.

Much of the focus of recent research has been on the response of precipitation in warming projections. Few studies have focused on the explicit response of ARs, as intense moisture transport events, to warming. *Dettinger* [2011] was the first to investigate this topic using an ensemble of projections from seven global climate models used in the fourth Intergovernmental Panel on Climate Change assessment report. Comparing historical simulations to end of the century warming projections, he found an increase in moisture transport and an increase in the number of years with many ARs. Investigation was furthered by use of data from CMIP5 in *Warner et al.* [2015]. They identified ARs using the 99th percentile of moisture transport along a transect just off the coast of North America in 30 years of historical simulations and RCP 8.5 projections from 10 different models. Similar to the findings in *Dettinger* [2011], *Warner et al.* [2015] found a robust increase in moisture transport within ARs, but little change in lower level winds. While focused on the eastern North Atlantic, *Lavers et al.* [2013] found similar increases in AR intensity as *Dettinger* [2011] and *Warner et al.* [2015] using an ensemble of five models participating in CMIP5 and a different method to identify ARs.

A comprehensive overview of the performance of the CMIP5 models in resolving landfalling ARs over the North Pacific has yet to be undertaken. Here we evaluate the statistics of landfalling ARs in historical simulations of 28 different CMIP5 models. We add to the existing research by identifying a subset of high performing CMIP5 models for North Pacific ARs and characterize the AR response to warming using CMIP5 simulations under the RCP 8.5 forcing scenario. Our emphasis is on the large-scale dynamics, in association with Rossby wave dynamics, that generate the sometimes quite delicate AR spatial structures [*Payne and Magnusdottir*, 2014].

The purpose of this paper is twofold: (1) to evaluate the ability of CMIP5 models to simulate landfalling ARs over the eastern North Pacific and (2) to investigate the response of AR behavior to projected climate change, specifically focusing on their dynamical responses. The paper is organized as follows. In section 2, we outline the CMIP5 models and simulations we use, how landfalling ARs are identified in each data set and summarize our evaluation metrics. In section 3, we evaluate the ability of CMIP5 historical simulations to resolve AR frequency over the North Pacific as compared to two different reanalysis products. In section 4, we investigate changes in AR behavior in end of the century RCP 8.5 projections. We summarize our conclusions in section 5.

2. Data and Methods

2.1. Data Sets

In order to comprehensively evaluate CMIP5 performance, we use data from 28 models originating from 18 different modeling groups participating in CMIP5. From each model, we use two sets of simulations: historical (1980–2004) and RCP 8.5 projections (2070–2099). Historical simulations are forced by both anthropogenic and natural changes in atmospheric composition and include changes in land cover [*Taylor et al.*, 2012]. While data from historical simulations are available over the entire industrial period, we focus on the years that have overlap with our reanalysis data sets (satellite era) and are common among all models. We evaluate the response of ARs to warming in the most extreme scenario available, RCP 8.5, in which radiative forcing increases to a peak of 8.5 W m^{-2} by the end of the 21st century [*Taylor et al.*, 2012].

Only a handful of models output daily fields for multiple ensembles. Therefore, we chose to limit our study to that of a single ensemble member from each model (the first available). Specific humidity (q) and zonal and meridional winds (u, v) are downloaded on pressure levels 1000, 850, 500, and 250 hPa from a central repository, the Earth System Grid-Center for Enabling Technologies (ESG-CET) (http://pcmdi9.llnl.gov/esgf-web-fe/). Further information on each model is detailed in Table 1.

We use reanalysis data to serve as an "observational" comparison for quality control over the period 1980–2004. We use MERRA reanalysis [*Rienecker et al.*, 2011] at 1.25° × 1.25° horizontal resolution and ERA-Interim reanalysis [*Dee et al.*, 2011] at 0.75°×0.75° horizontal resolution. For consistency with CMIP5 data, daily averages are used. Both reanalysis data sets have previously been used to study ARs, and they perform

Table 1. Details For All 28 CMIP5 Models Used in Order of Resolution (Historical, 1980–2005 and RCP 8.5, 2070–2100)[a]

ID	Model	Resolution	Group
01	CMCC-CESM	3.7 × 3.7	Centro Euro-Mediterraneo per I Cambiamenti Climatici
02	FGOALS-g2	3.1 × 3.1	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS, Tsinghua University
03	BNU-ESM	2.8 × 2.8	College of Global Change and Earth System Science, Beijing Normal University
04	bcc-csm1-1	2.8 × 2.8	Beijing Climate Center, China Meteorological Administration
05	CanESM2	2.8 × 2.8	Canadian Centre for Climate Modelling and Analysis
06	MIROC-ESM-CHEM	2.8 × 2.8	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
07	MIROC-ESM	2.8 × 2.8	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
08	GFDL-ESM2G	2.0 × 2.0	NOAA Geophysical Fluid Dynamics Laboratory
09	GFDL-ESM2M	2.0 × 2.0	NOAA Geophysical Fluid Dynamics Laboratory
10	NorESM1-M	1.9 × 1.9	Norwegian Climate Centre
11	IPSL-CM5A-LR	1.9 × 1.9	Institut Pierre-Simon Laplace
12	IPSL-CM5B-LR	1.9 × 1.9	Institut Pierre-Simon Laplace
13	CMCC-CMS	1.9 × 1.9	Centro Euro-Mediterraneo per I Cambiamenti Climatici
14	MPI-ESM-LR	1.9 × 1.9	Max Planck Institute for Meteorology
15	CSIRO-Mk3-6-0	1.9 × 1.9	Commonweath Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence
16	inmcm4	1.5 × 1.5	Institute for Numerical Mathematics
17	MIROC5	1.4 × 1.4	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
18	CNRM-CM5	1.4 × 1.4	Centre National de Recherches Meteorologiques/Centre Europeen de Recherche et Formation Avancee en Calcul Scientifique
19	IPSL-CM5A-MR	1.3 × 1.3	Institut Pierre-Simon Laplace
20	ACCESS1-0	1.2 × 1.2	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia
21	ACCESS1-3	1.2 × 1.2	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia
22	HadGEM2-CC	1.2 × 1.2	Met Office Hadley Centre
23	EC-EARTH	1.1 × 1.1	EC-EARTH consortium
24	MRI-CGCM3	1.1 × 1.1	Meteorological Research Institute
25	MRI-ESM1	1.1 × 1.1	Meteorological Research Institute
26	bcc-csm1-1-m	1.1 × 1.1	Beijing Climate Center, China Meteorological Administration
27	CCSM4	0.9 × 0.9	National Center for Atmospheric Research
28	CMCC-CM	0.7 × 0.7	Centro Euro-Mediterraneo per I Cambiamenti Climatici

[a] A numerical identifier is shown in the first column, model name in the second column, spatial resolution in the third column, and modeling group in the fourth column.

well against observational case studies [e.g., *Jiang and Deng*, 2011; *Lavers et al.*, 2012; *Ryoo et al.*, 2013; *Rutz et al.*, 2014; *Payne and Magnusdottir*, 2014].

We focus on the extended winter (defined as October through March), which is considered the active season for ARs over the North Pacific [*Neiman et al.*, 2008]. Here years are referred to as starting in October and ending in March. As there is no consistent calendar type between the different data sets, we remove 29 February from each leap year resulting in a 365 day year. The exception is the Hadley model (HadGEM2-CC), which has a 360 day year.

2.2. AR Identification

We focus on ARs making landfall, which is defined by physical proximity to the coastline between 30°N and 60°N in the eastern North Pacific, using the methods described in *Payne and Magnusdottir* [2014] (see their

Figure 1). We note that the spatial resolution of the majority of the CMIP5 models is coarser than the fine-scale ARs. Therefore, in order to identify AR-like features in this data set, we look for events with high moisture transport and an elongated geometry, impacting the coastline. Since our analysis extends over different models and two climates, landfalling AR dates are defined according to data set-dependent percentiles based on the historical simulation for each model [e.g., *Lavers et al.*, 2012]. Our threshold on vertically integrated moisture flux *MF* (as defined in equation (1) below) is distribution based and fixed to the 85th percentile of MF over all dates for the extended winter (October–March) [*Lavers et al.*, 2013].

The following criteria are used to detect AR-like features: (1) MF exceeds the 85th percentile over the landfall region (verified in all models to be above the static 250 kg m^{-1} s^{-1} threshold commonly used) [e.g., *Rutz et al.*, 2014; *Warner et al.*, 2015], (2) precipitable water (PW, as defined in equation (2) below) exceeds 2 cm [*Ralph et al.*, 2004; *Neiman et al.*, 2008], (3) the 850 hPa zonal and meridional components of the wind are both positive and the wind speed exceeds 10 m s^{-1} [*Neiman et al.*, 2008; *Dettinger*, 2011; *Hagos et al.*, 2015]. We also require that the long axis of the identified feature extends more than 2000 km in the zonal direction [*Ralph et al.*, 2004]. We note that the combination of the data set-dependent percentile thresholds over a range of variables means the total number of landfalling AR dates varies between data sets. Using this set of criteria, MERRA captures 80% and ERA-Interim captures 77% of the dates listed in Tables 1 and 2 of *Neiman et al.* [2008]. Both reanalysis data sets capture the notable February 1986 and January 1997 landfalling AR events [*Leung and Qian*, 2009].

Our reanalysis data sets are at spatial resolutions approximately equal to or higher than the majority of the CMIP5 models investigated. For direct comparison of reanalysis with CMIP5 data, we interpolate all data sets to a common 2° × 2° grid using bilinear interpolation after case selection.

MF is calculated as the product of the horizontal wind speed and the specific humidity vertically integrated from 1000 hPa to 500 hPa:

$$\mathrm{MF}(\lambda,\phi,t) = g^{-1}\left|\int_{p_s}^{p_t} q(\lambda,\phi,p,t)\mathbf{V}(\lambda,\phi,p,t)\mathrm{d}p\right|, \tag{1}$$

where **V** is the horizontal velocity and q is specific humidity (both in SI units), g is 9.81 m s^{-2}, p_s is 1000 hPa, and p_t is 500 hPa. We calculate the climatology for each date in the extended winter, over all years, for each data set (1980–2004 and 2070–2099, respectively). For identified landfalling AR dates, the daily *MF* anomaly is the difference between the *MF* value and the *MF* climatology for that day of the season.

PW is calculated as specific humidity in units of equivalent depth of a liquid unit column of water, vertically integrated from 1000 hPa to 500 hPa:

$$\mathrm{PW}(\lambda,\phi,t) = \left(g\rho_w\right)^{-1}\int_{p_s}^{p_t} q(\lambda,\phi,p,t)\mathrm{d}p, \tag{2}$$

where ρ_w is 1000 kg m^{-3}.

The horizontal length of the feature is determined by connected grid points satisfying the criteria on MF, PW, and the 850 hPa horizontal winds (described above), trailing westward from the point of landfall (similar to that described in *Lavers et al.* [2013]). With the exception of seasonal and interannual analysis, we use composites over all landfalling dates to investigate differences between reanalysis products and historical simulations.

2.3. Statistical Metrics

We use several different measures to evaluate model performance in representing landfalling ARs compared to the two reanalysis data sets. All horizontal calculations are weighted by the normalized square root of the cosine of latitude. Model bias (*B*) is defined as the difference between the model historical data (*M*) and each reanalysis data set (*O*). Spatial correlations between each model and observations are quantified using the area-weighted Pearson correlation coefficient (*R*).

To describe the amplitude of the differences between the models and the observations, we calculate the area-weighted root-mean-square error (*E*):

$$E = \sqrt{\sum_{i=1}^{N}\left(M_i - O_i\right)^2}. \tag{3}$$

To allow for the comparison of the error in each model for a given reanalysis data set and field, we calculate the relative model error, E' [*Gleckler et al.*, 2008]:

$$E' = \frac{E - \tilde{E}}{\tilde{E}}, \tag{4}$$

where $\tilde{E}$ is the median error over all models for a given observational data set and field. This quantity allows for comparison of the relative amount of error in each model so that models with a low E' have less error than models with a high E'.

While we use criteria-based methods to rank models in sections 3.1 and 3.2, we summarize model performance against reanalysis by quantifying model rank using the Taylor score, S [*Taylor*, 2001]:

$$S = \frac{4(1+R)^4}{\left(\sigma_M/\sigma_O + \sigma_O/\sigma_M\right)^2 \left(1+R_o\right)^4}, \tag{5}$$

where R_o is the maximum correlation attainable, which is here assumed to be 1, and σ is the standard deviation of each model (σ_M) and of each reanalysis data set (σ_O). For each model we calculate the Taylor score for a range of variables and approximate model performance by averaging the Taylor scores over all variables. We weight the average model score by the standardized difference from the average number of AR dates in reanalysis.

3. Evaluation of Landfalling ARs in Historical CMIP5 Simulations

This section is broken into three parts; we focus on (1) the frequency of AR-like conditions over the basin during landfalling dates, (2) the total number of dates identified by each model, including their seasonal and interannual variability, and (3) the deviation in relevant fields from reanalysis (frequency of occurrence, moisture flux, specific humidity, and the zonal and meridional winds at 850 and 250 hPa). The purpose of this section is to identify those models that best simulate the statistics of extreme moisture transport over the eastern North Pacific. We focus our evaluation specifically on landfalling AR dates over the historical period so that the selected models may be used to evaluate their response to warming in RCP 8.5 simulations.

3.1. AR Distribution and Amplitude

3.1.1. AR Frequency

Figure 1 shows the frequency of AR occurrence (shaded) and model bias (contoured) for each of the CMIP5 models. Frequency is defined as the total number of times each grid point satisfies criteria detailed in section 2.2, divided by the total number of landfalling AR dates identified in each model over the entire time period (1980–2004, October–March). In order to consider the role of model native resolution in their performance, panels are ordered from lowest resolution in the top left (CMCC-CESM) to highest resolution in the bottom right (CMCC-CM). The two reanalysis data sets are shown in the bottom right.

Comparison of the frequency distribution of the two reanalysis data sets shows excellent agreement over the ocean ($R = 0.988$, $p \leq 0.01$). Differences between the two data sets are concentrated over land (not shown). Overall, models capture the general shape of AR frequency. Frequency peaks in the Pacific Northwest and Northern California regions and drops off toward lower latitudes along the coastline. From the coastline, this peak in frequency trails westward and equatorward over the basin, consistent with observations [*Neiman et al.*, 2008]. Closer examination of the frequency distribution for each data set, however, shows large variation. Correlation coefficients (relative to MERRA and ERA-Interim, respectively) are shown in the lower left of each panel and show variable performance, with correlations ranging from a low of 0.69 (IPSL-CM5B-LR) to a high of 0.98 (EC-EARTH) (Figure 1).

For a comprehensive overview of the distribution of AR frequency over the domain shown in Figure 1, we show the standardized frequency distribution for each data set (the frequency distribution divided by its standard deviation). Shading indicates values between the 25th and 75th percentiles, and the solid black line refers to the median frequency. Again, agreement between the two reanalysis data sets is excellent, with consistent median (MERRA: 0.71, ERA-Interim: 0.67) and quantile values (MERRA: 0.23–1.50, ERA-Interim: 0.15–1.43). With the exception of MIROC-ESM-CHEM and MIROC-ESM, model frequency distribution is generally consistent with reanalysis. Models tend to underestimate median frequency ($\sim$0.1%/σ lower than average reanalysis) and high extremes ($\sim$0.07%/σ lower than average reanalysis).

Figure 1. For historical simulations (1980–2004), AR frequency (shading, %) for all models (lowest resolution, top left, to highest resolution, bottom right) compared against MERRA and ERA-Interim reanalysis in the lower right corner. Model bias (MERRA, dark, and ERA-Interim, light) is shown in contours (intervals of 10%, from ±10%, where blue is negative and red is positive). The correlation coefficient is shown in the lower left corner of each panel (MERRA, ERA-Interim). Only statistically significant bias shown (Student's *t* test, 95% level).

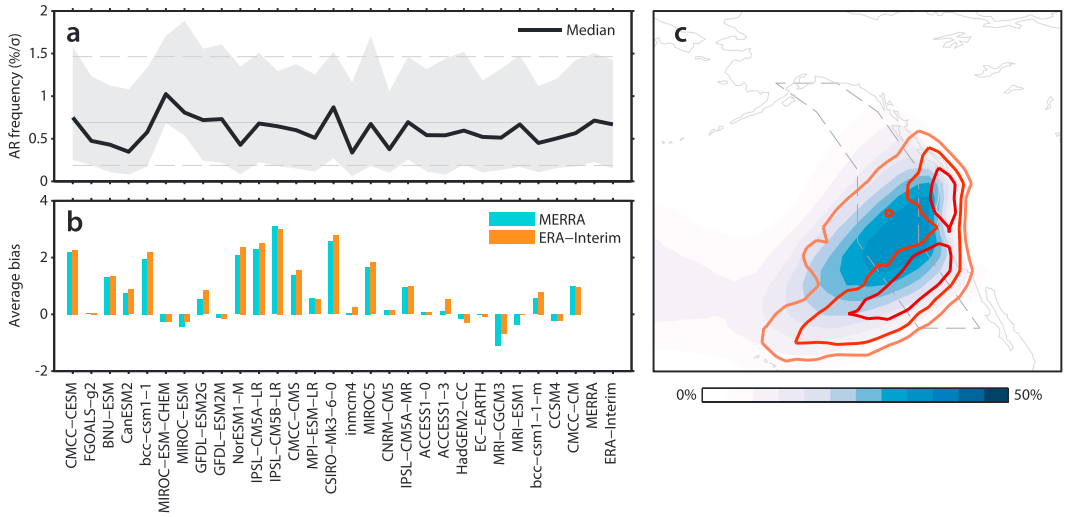

Figure 2. (a) The standardized distribution of nonzero AR frequency for each model in order of resolution (left to right) and MERRA and ERA-Interim reanalysis (far right), as represented by their median (black) and 25th (bottom edge) and 75th (top edge) percentiles. For reference, the average median (solid grey) and 75th and 25th percentiles (dashed grey, top and bottom, respectively) for both reanalysis are shown as horizontal lines. (b) For statistically significant bias (Student's *t* test, 95% level), the average bias for each model compared to (blue) MERRA and (orange) ERA-Interim reanalysis. (c) For AR frequency, the standard deviation across all models (contour intervals of 2σ, from 6σ) and the multimodel mean (shading, intervals of 10%, from 10%). The outlined region (25°N–60°N) is used to calculate anomalies for Figure 3.

3.1.2. Model Bias

Returning to Figure 1, we show the spatial distribution of model bias calculated as the difference in frequency between each model and each reanalysis data set. Positive bias is indicated by red contours, negative bias is indicated by blue contours, and each reanalysis data set is distinguished by contour shading (MERRA, dark colored contours and ERA-Interim, light colored contours). At each grid point, we calculate significance using a two-sided Student's *t* test in which the interannual variation in AR frequency is used for sample variance, following the example in *Anstey et al.* [2013]. Only statistically significant bias (at the 95% level) is reported.

Comparison of the bias relative to each reanalysis data set shows very good agreement both in magnitude and in location (compare dark and light contours in Figure 1). To provide a quantitative overview of the bias in each model, Figure 2b shows the average magnitude of bias over all points for all models. While a comparison of bias between the lowest-resolution models (top left) and the highest-resolution models (bottom right) shows a general decrease with increasing resolution, bias does not scale linearly with resolution. Changes in the resolution of the common grid do not have any impact on the results reported here. Average bias is generally positive indicating a spatial overestimation of AR frequency compared to reanalysis. We confirm this in Figure 2c where we show the multimodel composite of AR frequency (shading) and its standard deviation calculated over all 28 models (contouring). The figure illustrates that the largest disagreement between models is concentrated equatorward of the peak frequency, indicating disagreement in the average latitudinal position of AR landfall. We note that this peak in standard deviation is not an artifact of high moisture.

3.1.3. MF Distribution and Amplitude

We further our investigation of the spatial variability of ARs in the different models in Figure 3, which shows Hövmoller diagrams of the seasonal variability over all years for each model compared to reanalysis (in the lower right corner). Each seasonal composite is composed of the sector zonal average (over the dashed box in Figure 2c) standardized MF anomalies for all landfalling AR dates. Positive anomalies of MF, as areas of anomalously high moisture transport, are used as a proxy for AR frequency. We compare the median latitude of the peak in positive anomalies between each model (black) and the average over reanalysis (grey). As we do not require an equal number of AR dates to be identified in each model (discussed in detail in section 3.2), years with no identified dates are shown in white. It is apparent from Figure 3 that the performance of models in resolving AR latitude compared to reanalysis varies greatly. While several models show an equatorward shift in median latitude relative to reanalysis, this shift varies in magnitude and is not consistent with changes in model resolution.

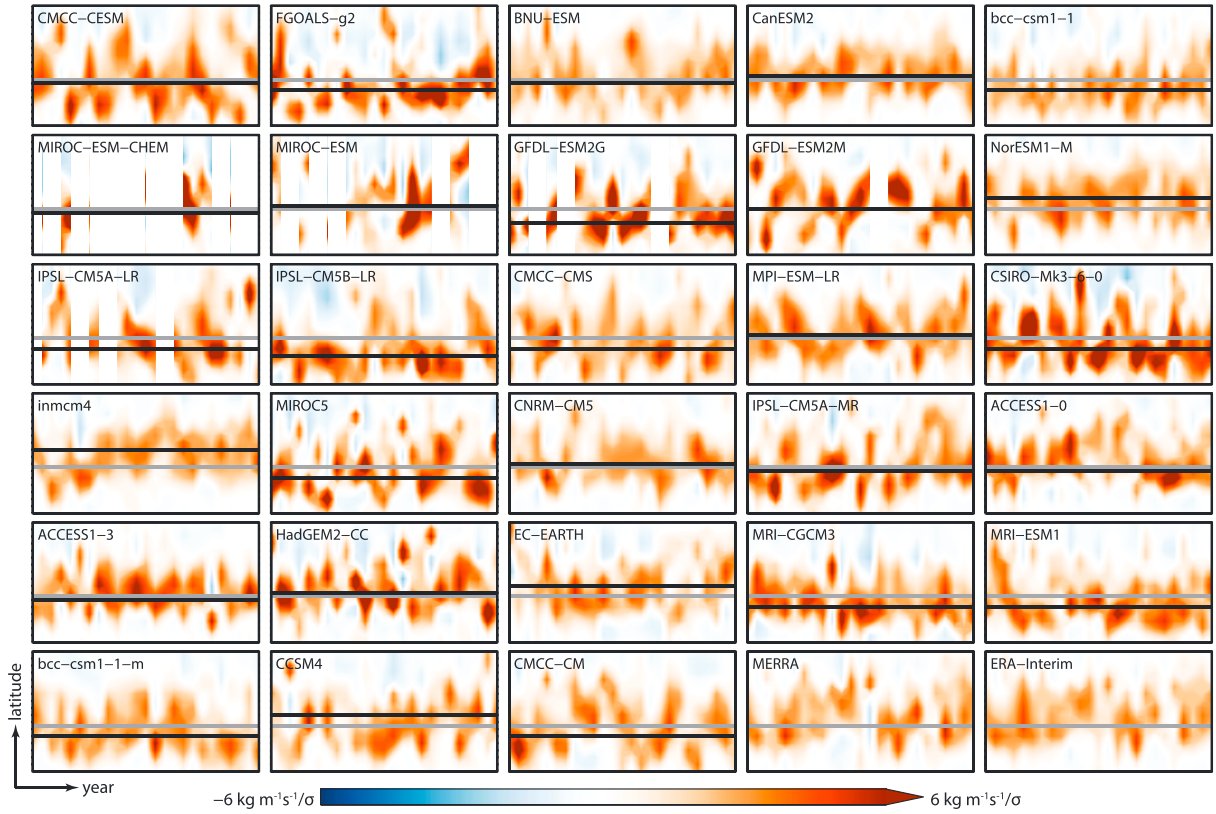

Figure 3. For historical simulations (1980–2004), a Hövmoller diagram of the interannual variability for the extended winter (*x* axis) of the sector zonal average (calculated over dashed region in Figure 2c) of daily anomalies of MF for landfalling AR dates. The median latitude of positive MF anomalies in each model (black) is compared to that in the average of both reanalysis (grey).

Differences in the amplitude of positive MF anomalies are less consistent with changes in resolution. Figure 4a shows the yearly amplitude of positive MF anomalies (shown in Figure 3) across all models compared to the average amplitude from reanalysis (MERRA, blue and ERA-Interim, orange). For models, we note that we are focusing not only on the year-to-year variability but rather on the overall amplitude of variability. While the median amplitude of the standardized positive MF anomalies is consistent with reanalysis, the total range of values over all models is quite large. In particular, model extremes are largely overestimated. This may be due to differences in the total number of dates identified in each model, which will be addressed in section 3.2.

3.1.4. Overview

In this subsection we evaluated the distribution of landfalling AR dates in historical model simulations and investigated differences in the amplitude of identified events using MF anomalies. Based our analysis, we can characterize high-performing models as those having: (1) high spatial correlation of AR frequency relative to reanalysis ($R \geq .93$), (2) low average bias ($\bar{B} \leq 1$), and (3) a median latitude of positive MF anomalies close to the average of reanalysis ($\leq 3°$). By this criterion, an initial subset of high-performing models include CanESM2, MPI-ESM-LR, CNRM-CM5, IPSL-CM5A-MR, ACCESS1-0, ACCESS1-3, HadGEM2-CC, EC-EARTH, and CCSM4. To test the performance of this subset against all models considered, we return to Figure 4a, where we show their range (shading) in positive MF anomalies. These models show a marked improvement of the multimodel ensemble and generally capture the range of the interannual variability of MF amplitude in reanalysis. Many of the extreme outliers present in the original 28 models are excluded in this subset.

3.2. Seasonal Evolution and Interannual Variability

3.2.1. Total Counts and Seasonal Evolution

Figure 5 shows the total number of landfalling dates identified in each data set and the breakdown of that total into the number identified in each month. Resolution increases from left to right with the two reanalysis data sets shown at the far right side of the figure. There is rather good agreement between the two reanalysis data sets, with more AR dates identified in ERA-Interim (322) than in MERRA (258). Compared to reanalysis, the number of AR dates identified in models varies widely. The highest number of landfalling dates

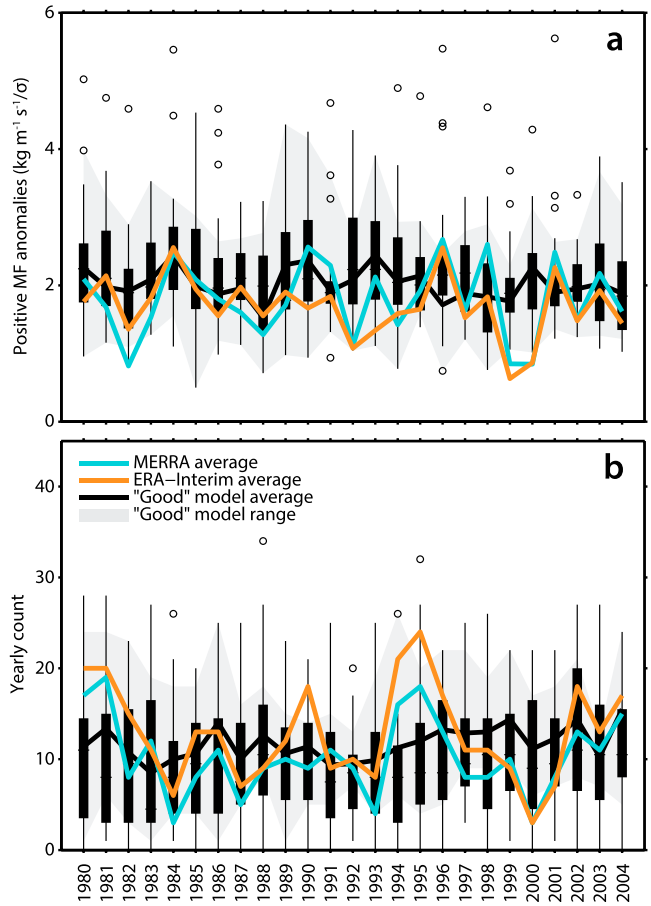

Figure 4. For all models, a box-and-whisker plot showing (a) the range of positive MF anomalies for landfalling AR dates identified in each extended winter year (October–March) and (b) yearly counts of landfalling AR dates. The solid boxes show the model spread between the 25th and the 75th quantiles, with the median shown as a black line. The circles represent outliers. The averages for MERRA (blue) and ERA-Interim (orange) are shown in each panel. The range of the high performing subset of models is shown in grey shading, with their average (black) in reference to the evaluation in (a) sections 3.1 and (b) 3.2.

is identified in inmcm4 (501), and the fewest are identified in MIROC-ESM-CHEM (26). Generally, increases in model resolution are associated with an increased consistency in the number of identified AR dates compared with reanalysis.

From observations we know that the highest number of landfalling events (over the entire coastline) occurs early in the season [*Neiman et al.*, 2008]. Models generally perform well against observations, identifying more AR dates early in the season. Exceptions to this seasonality include NorESM1-M, MIROC5, and CSIRO-Mk3-6-0.

3.2.2. Interannual Variability in Total Counts

Figure 4b shows the interannual variability of identified landfalling dates for each year in the historical period as compared to the two reanalysis data sets for all models. Again, the two reanalysis data sets are consistent with each other, with notable exceptions in 1982, 1990, and 1995. However, given the difference in the total number of dates identified by each data set, this is to be expected (there are 64 more dates identified in ERA-Interim). While the number of identified dates in all models varies widely each year, their average amplitude is generally consistent with observations.

3.2.3. Overview

In this subsection we evaluated the number of AR dates identified and their seasonal and interannual variation. We add to our criteria of high-performing models from section 3.1, to include (1) total counts consistent with observations (within 100 of the average of reanalysis) and (2) a higher number of identified dates early in the season.

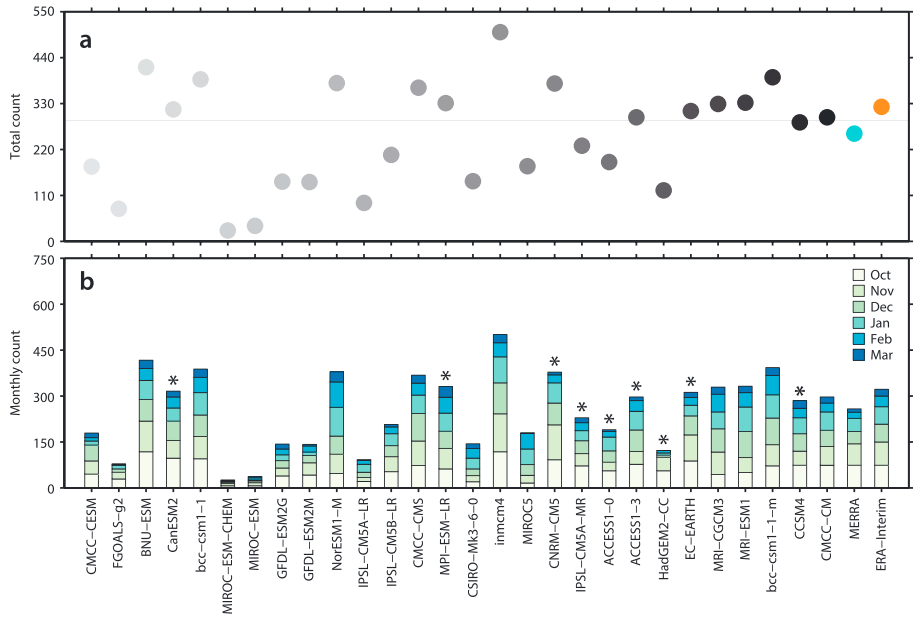

Figure 5. (a) The total number of landfalling dates selected by each model (lowest resolution, lightest grey and highest resolution, darkest grey) compared to reanalysis (MERRA, blue and ERA-Interim, orange). (b) The separation of the total number of landfalling dates into monthly counts. The subset of high-performing models identified in section 3.1 is marked with an asterisk.

For a closer investigation of their seasonal variability, we indicate the subset of models identified in section 3.1 in Figure 5b with an asterisk. All models in this subset correctly resolve the AR seasonality, with the highest number of AR dates in the earliest 3 months (October–December) and the fewest in the last 3 months (January–March). Likewise, the majority of the identified models have total counts consistent with observations. An exception is HadGEM2-CC, which we now exclude from our subset.

Returning to Figure 4b, we again test the performance of our revised subset and show their range of yearly counts in shading (EC-EARTH, CNRM-CM5, MPI-ESM-LR, CCSM4, ACCESS1-0, ACCESS1-3, IPSL-CM5A-MR, and CanESM2). While totals remain consistent with reanalysis, comparison of the model average of the subset with reanalysis still shows some overestimation and underestimation of total counts, though now within the range of what is seen in reanalysis.

3.3. Model Performance Summary

In this subsection we evaluate the performance of all models against relevant fields: (1) AR frequency, (2) MF, (3) 850 hPa specific humidity, (4) 850 hPa horizontal winds, and (5) 250 hPa horizontal winds. We investigate the performance of each model against upper level winds because of the association of intense ARs with the extratropical jet and Rossby wave dynamics [*Payne and Magnusdottir*, 2014]. Figure 6a shows the relative error, E′ for each variable (*y* axis) from each model (*x* axis). Each grid box is divided along the diagonal to show calculations relative to MERRA reanalysis (bottom triangle) and ERA-Interim reanalysis (top triangle). The shading shows where models have high error (positive, brown) and low error (negative, teal) relative to the entire set of models being evaluated. The revised subset of models identified in the previous subsection is marked by an asterisk at the top of Figure 6a.

Results show that during landfalling AR dates, relative error tends to be consistent across different metrics and there is a clear separation between high- and low-performing models. The subset of models previously identified all show relatively low error. The exception is IPSL-CM5A-MR, which is inconsistent with reanalysis for both upper and lower level winds and shows marginal improvement in performance over the remaining metrics. CCSM4 shows poor performance in lower level moisture but otherwise performs reasonably well.

Figure 6b shows a Taylor diagram display (see *Taylor* [2001] for details) of the average of standardized metrics for each variable shown in Figure 6b, where models are indicated by their numerical identifier from Table 1. High-performing models are those falling closest to reanalysis (black star). Radial distance of each point from the origin shows standard deviation. Radial distance of each point from reanalysis shows error, and the

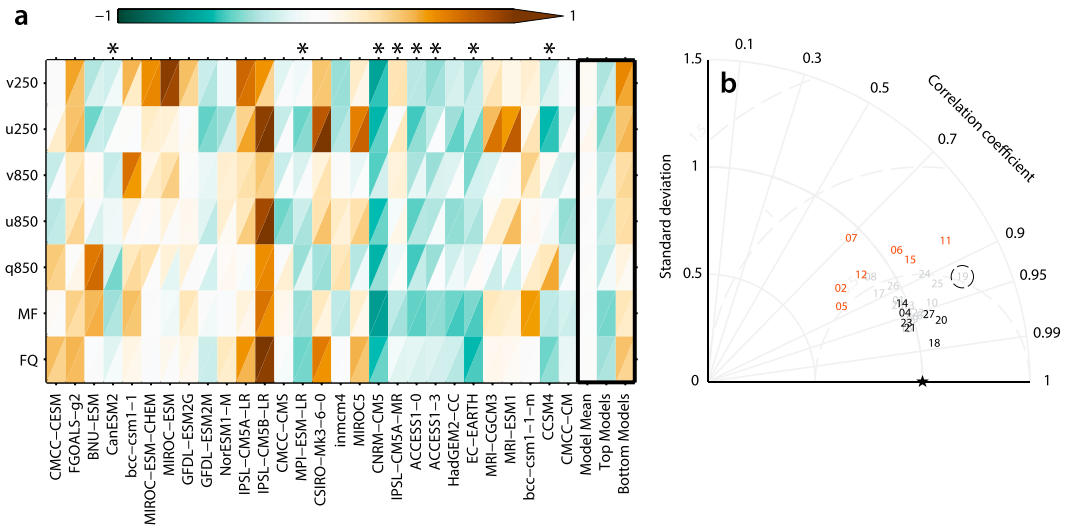

Figure 6. (a) A portrait diagram display of the relative error for each model (*x* axis) against each observational data set (MERRA, top triangle and ERA-Interim, bottom triangle) over (*y* axis, from top) 250 hPa meridional wind (v250), 250 hPa zonal wind (u250), 850 hPa meridional wind (v850), 850 hPa zonal wind (u850), 850 hPa specific humidity (q850), MF and AR frequency (FQ). (b) Taylor diagram display of the average model performance for variables (each standardized by reanalysis) in Figure 6a. Models are identified by their ID in Table 1, where both reanalyses are represented by a black star. Points are positioned according to their standard deviation (radial distance from the origin), root-mean-square error (radial distance from reanalysis) and correlation (azimuthal position). The identifiers for the final list of the subset of high-performing models are in black bold, and for the seven lowest-performing models are in red bold.

azimuthal position of each point shows the correlation. The identifier for each high-performing model in our subset is in bold (black). The poor performance of IPSL-CM5A-MR (dashed circle in Figure 6b) is reflected in its distance from the other models in the subset. For this reason, we exclude IPSL-CM5A-MR from our subset.

Models are ranked according to the average of the Taylor scores and standardized difference from the average number of AR dates in reanalysis. In Figure 6a (black outline), the relative error for the model mean is compared to our subset of high-performing models (EC-EARTH, CNRM-CM5, MPI-ESM-LR, CCSM4, ACCESS1-0, ACCESS1-3, and CanESM2) and to the lowest seven performing models (FGOALS-g2, bcc-csm1-1, MIROC-ESM-CHEM, MIROC-ESM, IPSL-CM5A-LR, IPSL-CM5B-LR, and CSIRO-Mk3-6-0), identified in bold (red) in Figure 6b. Error is reduced in the high-performing models and increased in the low-performing models.

We note that our evaluation is rather robust to changes in the AR identification scheme. Increasing the threshold on MF from the 85th percentile to the 95th percentile (close to the threshold used in *Warner et al.* [2015]) changed only one of the seven high-performing models we identified (CMCC-CM replaced CanESM2).

4. Future Changes in Landfalling AR Dates in RCP 8.5 Simulations

In this section, we (1) characterize the change in AR behavior with warming by comparing end-of-century RCP 8.5 simulations to historical simulations and (2) decompose the moisture flux equation into the thermodynamic and dynamic components to investigate the influence on the change in AR frequency with warming. The purpose of this section is to characterize the large-scale response of the AR, as an atmospheric feature, to warming.

4.1. Response of AR Behavior to Warming

4.1.1. AR Frequency

Figure 7 shows the difference in the frequency of AR occurrence between RCP 8.5 and historical simulations (shading, with historical frequency contoured) for the seven high-performing models identified by our evaluation in section 3. As in section 3.1, frequency is defined as the total number of times each grid point satisfies criteria in section 2.2, divided by the total number of landfalling AR dates identified. The multimodel average is shown in the lower right corner. Comparison to the superimposed historical frequency distribution (contours) shows a consistent pattern of increased frequency equatorward in RCP 8.5 projections. Frequency increases (of varying magnitudes) are also apparent poleward of historical peak frequency in several models

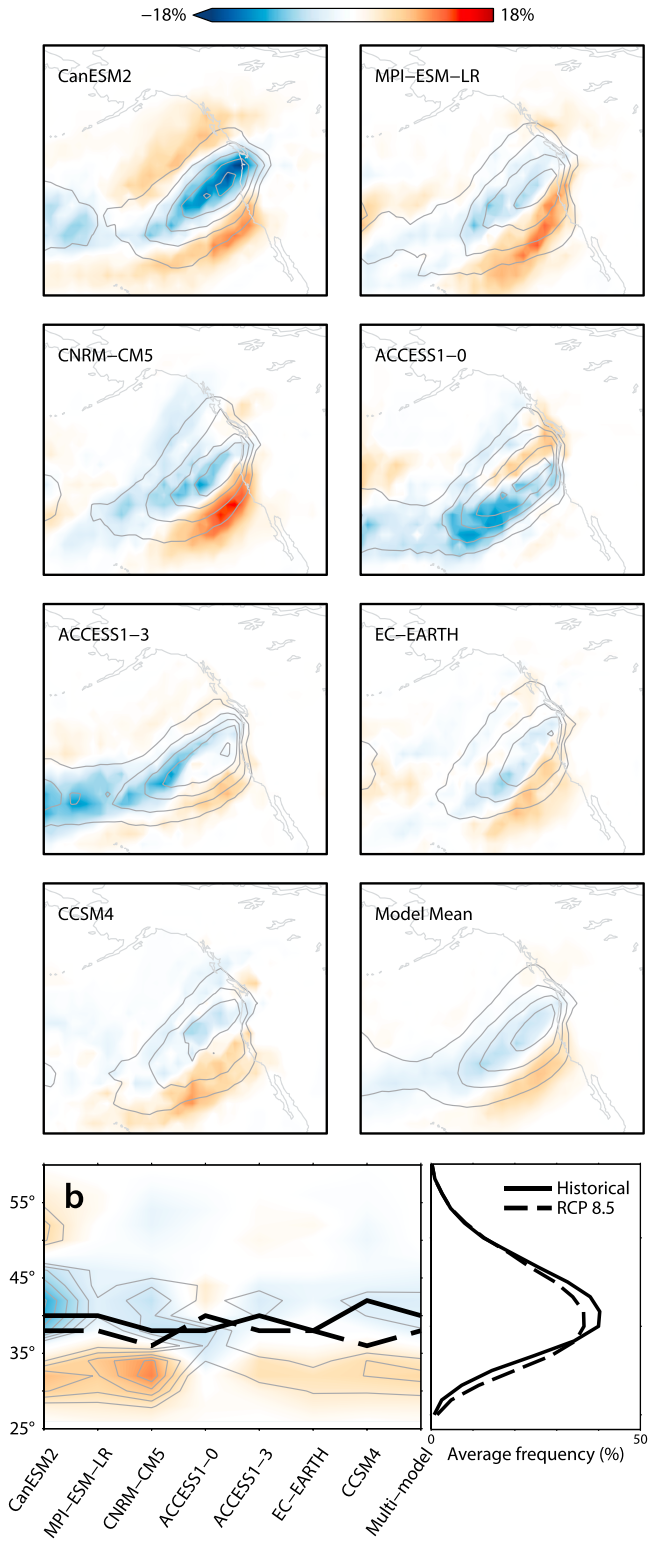

Figure 7. For selected models and multimodel mean, (top eight panels) the AR frequency difference (shading, %, RCP 8.5—historical), with the historical distribution shown in contours (intervals of 10%, starting from 10%) and (b) the sector zonal average (over the dashed box in Figure 2c) of the difference field, with the latitude of peak frequency for the historical (solid black) and RCP 8.5 (dashed black) data sets is shown to the left. The standard deviation over each data set is shown in contour (intervals of 1σ, from 2.5σ). The multimodel zonal average frequency for each period is shown in Figure 7b (right).

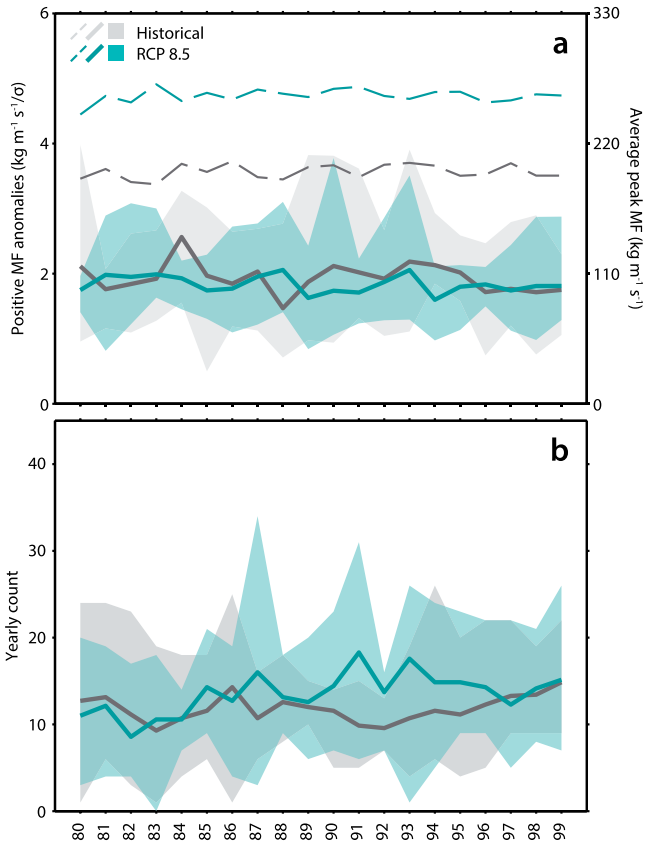

Figure 8. For two 20 year periods (historical: 1980–1999 and RCP 8.5: 2080–2099), for identified landfalling AR dates, (a) the range of positive MF anomalies (shading) in each extended winter year (October–March) and (b) the range of yearly counts (shading). The full range of historical data is shown in grey shading, and the full range of RCP 8.5 data is shown in teal shading. The average for each data set in both panels is shown as a dark grey and dark teal solid line, respectively. In Figure 8a the average peak MF in each extended winter (October–March) is shown as a dark grey and dark teal dashed line, respectively.

(CanESM2, MPI-ESM-LR, ACCESS1-0, ACCESS1-3, and EC-EARTH). Frequency decreases are approximately colocated with the historical peak frequency and are apparent in all models.

In order to determine whether the dominant increase in equatorward frequency is due to an absolute shift in landfalling latitude or due to a broadening of the peak frequency region, we show the sector zonal average (over the dashed box in Figure 2c) of the frequency difference field in Figure 7b. We compare the latitude of the frequency peak for the historical (solid black) and RCP 8.5 (dashed black) data sets to the centers of maximum standard deviation in the frequency difference field (grey solid). While there is a small difference in latitude (~2°), the centers of variability are generally concentrated equatorward and poleward of this latitudinal shift. These results do not change when the full ensemble of models evaluated in the previous section are considered, with several exceptions (bcc-csm1-1, MIROC-ESM-CHEM, MIROC-ESM, GFDL-ESM2M, IPSL-CM5A-LR, MIROC5, IPSL-CM5A-MR, HadGEM2-CC, and CCSM4). The rather static latitudinal peak in frequency suggests that a portion of the change in the distribution in Figure 7b is due to a broadening of the region of peak frequency. This means that the end of the century landfall is likely to occur over a wider latitude range than for the current climate. Broadening of the frequency distribution is confirmed in Figure 7b (right), which shows the multimodel zonal average frequency for each period.

4.1.2. Interannual Variability

In order to focus on the direct response of AR-like features to warming, we examine the change in MF at landfall. Consistent with previous research [*Lavers et al.*, 2013; *Warner et al.*, 2015], we find an absolute increase in MF within ARs in projections. Our results show between a 23% and 35% increase in the number of AR dates identified with warming. However, ARs are extreme events by definition. Therefore, we focus not only on absolute changes in MF but on changes in daily positive MF anomalies (as shown in Figure 3) between the

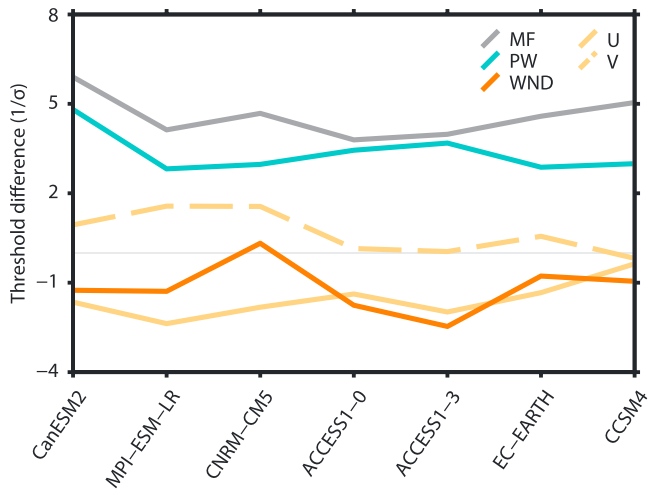

Figure 9. For selected models, the standardized difference in thresholds for MF (grey solid), PW (blue solid), wind speed (dark orange solid), meridional wind (light orange, dashed), and zonal wind (light orange, solid).

two climates. For a direct comparison of the interannual variability between the two data sets, we use only 20 years from each (October–March, historical: 1980–1999 and RCP 8.5: 2080–2099).

We first address the question, do ARs at the end of the century have increased intensity compared to ARs at the end of the previous century? Another way to phrase this question: Is there a shift in the distribution of moisture transport extremes? Figure 8a shows a comparison of the full range of positive MF anomalies between historical simulations (grey shading) and RCP 8.5 projections (teal shading). The mean of the distribution of each data set is shown as a grey and a teal line, respectively. Positive MF anomalies in late century RCP 8.5 projections show similar amplitudes to historical simulations, with extreme values falling within the total range of historical anomalies. This is also apparent when comparing the average of the two data sets, which are statistically indistinguishable (Student's *t* test at 95% level). We do find an absolute increase in the MF of AR dates with warming, seen as dashed lines in Figure 8a. This increase in MF does translate to more intense landfalling events, which is in agreement with previous research [*Dettinger*, 2011; *Lavers et al.*, 2013; *Warner et al.*, 2015]. However, within the context of the projected climate at the end of the 21st century, we do not find an increased occurrence of extreme events measured as percentiles.

In Figure 8b, we compare the interannual variability of the number of AR dates identified for the same two 20 year periods as in Figure 8a. All models show an increase in the total number of landfalling dates in RCP 8.5 projections (between 30 and 100 days, with an average of a 62 day increase). Results generally show similar interannual variability in both data sets, with years of many and few AR dates. However, years with an extreme number of identified dates in RCP 8.5 fall well outside the historical range. This is most apparent in comparison of the years 1987/2087 and 1991/2091, in which high extremes in RCP 8.5 projections greatly exceed historical extremes. An analysis of the clustering of AR dates (multiday landfalling events) within each data set shows an increase in their number, but little to no change in their average length. This suggests that the increase in the total number of landfalling dates is related to an increased number of multiday landfalling events.

4.1.3. Moisture and Wind Response

We break the change in MF due to warming into the response in the moisture field and the response in the wind field. Returning to the criteria we use to define landfalling AR dates, we first look at the change in our percentile-based thresholds, noting that we are now focused on the region hugging the coastline, eastward of the box in Figure 2c. Since the thresholds we use to identify ARs are distribution based, the difference in the threshold used for historical simulations and RCP 8.5 can give insight into the response of each of the fields to warming. Figure 9 shows the standardized change in the threshold between the two climates. The change in MF thresholds over all models shows a large positive shift, consistent with the robust increase in MF discussed previously. The similar magnitude of this shift over all models is confirmation of our evaluation in section 3, as it indicates similarities in the MF distribution of the selected models. Consistent with an increase in atmospheric moisture, we also find a large shift in the PW threshold with warming. The threshold for wind speed generally shows a decrease in RCP 8.5 projections. We investigate the change in the 850 hPa meridional

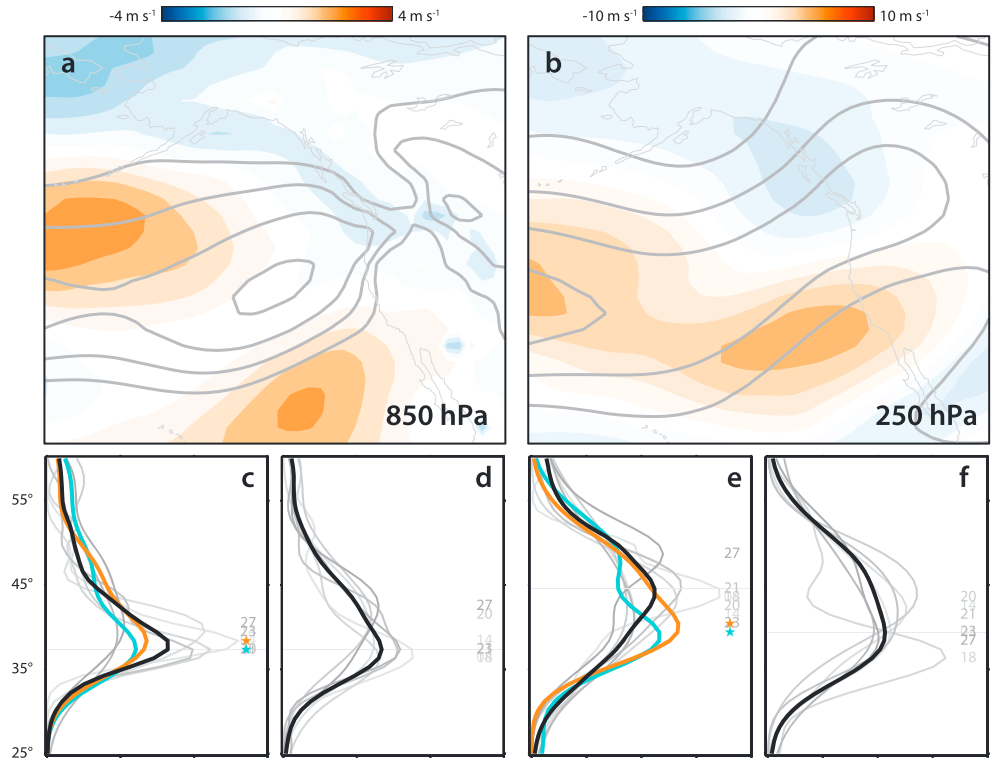

Figure 10. (a, b) Difference in the multimodel composite zonal winds at 850 and 250 hPa (shading, RCP 8.5—historical). For reference, the multimodel zonal wind composite for the historical period is shown in contour (at 850 hPa: intervals of 3 m s^{-1} starting from 3 m s^{-1} and at 250 hPa: intervals of 10 m s^{-1} starting from 10 m s^{-1}). Probability density functions showing the latitudinal peak in zonal wind anomalies at (c, d) 850 hPa and at (e, f) 250 hPa for (Figures 10c and 10e) historical simulations and (Figures 10d and 10f) RCP 8.5 projections for the seven selected models (grey lines). The multimodel average is shown in black. For reference, the latitudinal peak in MERRA (blue) and ERA-Interim (orange) are shown in Figures 10a and 10c. The peak of the PDF distribution for each model is marked by the numerical identifier (from Table 1).

and zonal components of the wind using 10 m s^{-1} as a baseline for the historical period. The change in the 850 hPa zonal and meridional baselines are shown in solid and dashed light orange lines, respectively, in Figure 9. All selected models show a decrease in the zonal wind baseline and some degree of increase in the meridional wind baseline in projections.

Beginning with a broad overview of the response of the zonal winds to warming, a multimodel composite analysis of the difference between RCP 8.5 projections and historical simulations (shading in Figure 10a) shows a weakening of the 850 hPa zonal wind speed in the jet exit region and a broadening of the jet maximum compared to the historical mean position (contour). A similar pattern is seen in the response of the upper level jet, where there is an eastward extension of the jet maximum over the central North Pacific and an increase equatorward of the jet core in the eastern North Pacific (Figure 10b). These results are consistent with findings in *Neelin et al.* [2013], who show an eastward extension of the climatological wintertime jet with CMIP5 projections.

We next focus on the variability in the latitude of the jet maximum at 850 and 250 hPa for each data set (Figures 10c–10f). The latitudinal distribution of the zonal winds is shown as probability density functions over all models (grey), compared to the multimodel mean (black). The distributions of MERRA (blue) and ERA-Interim (orange) are shown in the historical panels for comparison. The multimodel average in RCP 8.5 shows a widening of the latitudinal range of peak wind, consistent with the broadening of the lower level jet maximum seen in Figure 10a. While there is increased variability in the latitudinal position of peak lower level wind corresponding to AR events between the models in RCP 8.5 compared to historical simulations, there is little change in the mean latitude between the two climates (horizontal grey lines in Figures 10c and 10d). Similar to the results for 850 hPa, Figures 10e and 10f show an increase in the meridional variability of the upper level jet associated with ARs. Our results are consistent with the pattern of increased meridional vari-

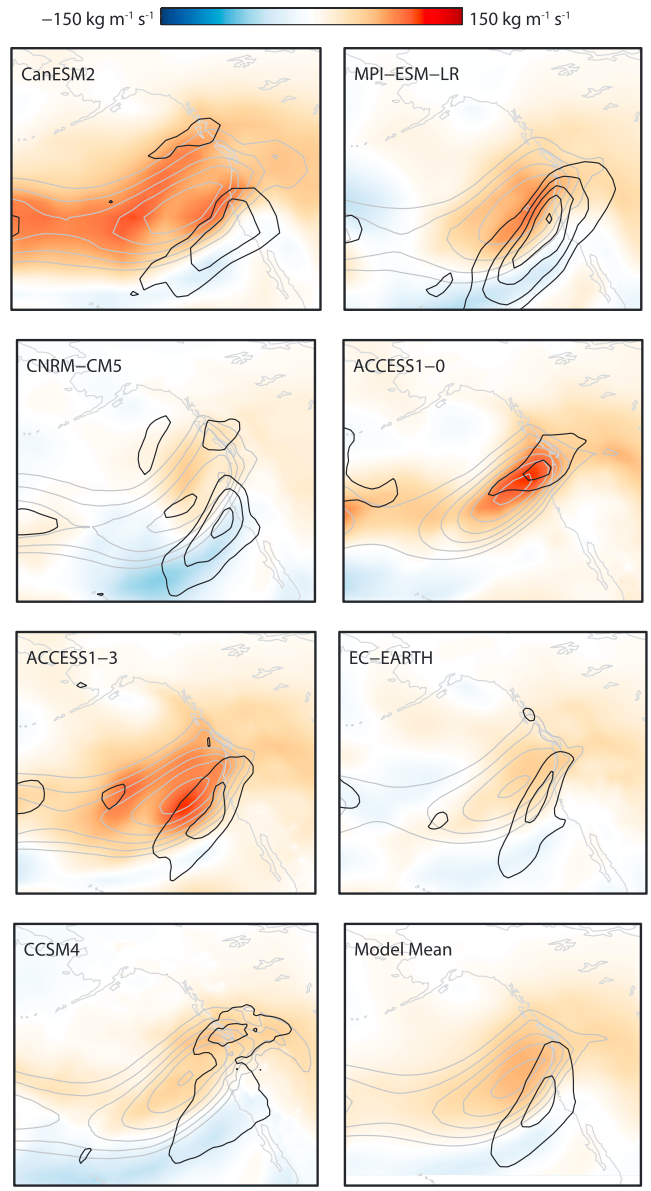

Figure 11. The thermodynamic response of MF to warming minus the dynamic response of MF to warming (kg m^{-2} s^{-1}) in shading for each of the selected models (lowest resolution, top left to highest resolution, bottom right). The multimodel mean difference is in the bottom right. The transient eddy response is shown in black contours (intervals of 10 kg m^{-2} s^{-1} starting from 15 kg m^{-2} s^{-1}), and the late century RCP 8.5 composite MF for landfalling dates is shown in grey contours (intervals of 50 kg m^{-2} s^{-1} starting from 200 kg m^{-2} s^{-1}).

ability in the midlatitude jet identified in *Barnes and Polvani* [2013]. We suggest that the pattern of frequency change between the two climates (seen in Figure 7), where we show a broadening of the distribution, is a consequence of a meridionally variable mean state. This is explored in more detail in section 4.2.

4.2. Dynamical Response and Decomposition of the Moisture Flux

In order to investigate the spatial response of MF in projections from previous subsection, we approximate the change in MF as [*Trenberth and Guillemot*, 1995; *Seager et al.*, 2010]

$$\delta \mathrm{MF} \approx g^{-1} \left| \int_{p_s}^{p_t} \left(\overline{\mathbf{V}}\, \delta \overline{q} + \delta \overline{\mathbf{V}}\, \overline{q} + \delta \left(\overline{\mathbf{V}' q'} \right) \right) \mathrm{d}p \right|, \tag{6}$$

where

$$\delta(\cdot) = (\cdot)_{\mathrm{RCP8.5}} - (\cdot)_{\mathrm{historical}} \tag{7}$$

and where overbars represent averages over all landfalling dates and primes indicate daily anomalies. In equation (6), the first term describes the thermodynamic response, the second term describes the dynamic response and the third term describes the transient eddy response of MF to warming. In Figure 11, we show the thermodynamic response minus the dynamic response in shading and the transient eddy response in black contours for the selected models. Positive values indicate regions where the thermodynamic component dominates and negative values indicate regions where the dynamic component dominates. The late century RCP 8.5 composite MF for landfalling dates is shown in grey contours for each panel. The multimodel ensemble mean is shown in the lower right of the figure.

Much of the change in MF is dominated by the thermodynamic response to warming (warm colors in Figure 11), consistent with findings in *Lavers et al.* [2013] and *Warner et al.* [2015]. However, in the region dominated by frequency increases in Figure 7, the change in MF is largely due to dynamic and transient responses to warming (cool colors and black contours in Figure 11, respectively). The colocation of these two regions supports our conclusions in the previous section that equatorward increases in the lower level zonal wind play a role in the change in AR distribution in Figure 11. Within the context of the broadening of the frequency distribution from Figure 7 and the variable upper level jet from Figure 8, we suggest that while small, the dynamical atmospheric response to warming may be important to understanding the response of AR behavior, particularly in determining the latitudinal distribution of landfalling events.

5. Conclusions

We present a comprehensive evaluation of CMIP5 model performance in resolving landfalling ARs over the North Pacific basin. While the majority of models evaluated correctly resolved the spatial structure of landfalling ARs, their performance in identifying landfalling dates and in resolving interannual variability in terms of number, amplitude, and median latitude varied greatly. Through an evaluation of these characteristics relative to reanalysis, we identify a subset of high-performing models for analysis of the response of ARs to warming: CanESM2, MPI-ESM-LR, CNRM-CM5, ACCESS1-0, ACCESS1-3, EC-EARTH, and CCSM4.

We extend our investigation to the response of AR behavior to projections under the RCP 8.5 scenario in the seven high performing models. Comparison of the historical distribution of AR frequency to RCP 8.5 projections shows a broadening of the distribution, particularly apparent equatorward of peak historical frequency. Further investigation of the response of MF to warming shows a robust increase in atmospheric moisture, but a decrease in the lower level zonal winds. Composite analysis shows a broadening of both the 850 and 250 hPa zonal winds equatorward of the historical jet maxima and a slight weakening in the jet exit region, apparent in the multimodel composite of 850 hPa zonal winds. Spatial decomposition of the response of MF shows that the thermodynamic response dominates in the peak region of AR frequency, consistent with increases in the atmospheric moisture content. However, we also find that the dynamical response dominates equatorward of the peak in the multimodel composite of MF in RCP 8.5.

Our results in section 4 support many of the findings in *Dettinger* [2011]; *Warner et al.* [2015] about changes in AR landfalling behavior in future climate projections. However, the increased variability of the upper level jet and broadening of the region of peak AR frequency calls for further analysis into the dynamical response to warming in order to fully characterize potential changes in ARs in late century projections.

Acknowledgments
We thank Wenchang Yang, Graham Simpkins, and the three anonymous reviewers for their helpful comments on the manuscript. This research was supported by NSF grant AGS-1206120 and NSF grant DGE-1321846. All data used in this research are described in *Rienecker et al.* [2011] and *Dee et al.* [2011].

References

Anstey, J. A., P. Davini, L. J. Gray, T. J. Woollings, N. Butchart, C. Cagnazzo, B. Christiansen, S. C. Hardiman, S. M. Osprey, and S. Yang (2013), Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution, *J. Geophys. Res. Atmos.*, *118*, 3956–3971, doi:10.1002/jgrd.50231.

Bao, J.-W., S. A. Michelson, P. J. Neiman, F. M. Ralph, and J. M. Wilczak (2006), Interpretation of enhanced integrated water vapor bands associated with extratropical cyclones: Their formation and connection to tropical moisture, *Mon. Weather Rev.*, *134*, 1063–1080.

Barnes, E. A., and L. Polvani (2013), Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models, *J. Clim.*, *26*, 7117–7135.

Chang, E. K., Y. Guo, and X. Xia (2012), CMIP5 multimodel ensemble projection of storm track change under global warming, *J. Geophys. Res.*, *117*, D23118, doi:10.1029/2012JD018578.

Dacre, H. F., P. A. Clark, O. Martinez-Alvarado, and M. A. Stringer (2015), How do atmospheric rivers form?, *Bull. Am. Meteorol. Soc.*, *96*, 1243–1255, doi:10.1175/BAMS-D-14-00031.1.

Dee, D., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*, 553–597.

Dettinger, M. (2011), Climate change, atmospheric rivers, and floods in California—A multimodel analysis of storm frequency and magnitude changes, *J. Am. Water Resour. Assoc.*, *47*(3), 514–523.

Dettinger, M., 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.

Dettinger, M. D. (2013), Atmospheric rivers as drought busters on the U.S. West Coast, *J. Hydrometeorol.*, *14*, 1721–1732.

Gleckler, P., K. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *J. Geophys. Res.*, *113*, D06104, doi:10.1029/2007JD008972.

Hagos, S., L. R. Leung, Q. Yang, C. Zhao, and J. Lu (2015), Resolution and dynamical core dependence of atmospheric river frequency in global model simulations, *J. Clim.*, *28*, 2764–2776.

Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Clim.*, *19*, 5686–5699.

Jiang, T., and Y. Deng (2011), Downstream modulation of North Pacific atmospheric river activity by East Asian cold surges, *Geophys. Res. Lett.*, *38*, L20807, doi:10.1029/2011GL049462.

Lavers, D. A., G. Villarini, R. P. Allan, E. F. Wood, and A. J. Wade (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.*, *117*, D20106, doi:10.1029/2012JD018027.

Lavers, D. A., R. P. Allan, G. Villarini, B. Lloyd-Hughes, D. J. Brayshaw, and A. J. Wade (2013), Future changes in atmospheric rivers and their implications for winter flooding in Britain, *Environ. Res. Lett.*, *8*, 034010, doi:10.1088/1748-9326/8/3/034010.

Lavers, D. A., F. M. Ralph, D. E. Waliser, A. Gershunov, and M. D. Dettinger (2015), Climate change intensification of horizontal water vapor transport in CMIP5, *Geophys. Res. Lett.*, *42*, 5617–5625, doi:10.1002/2015GL064672.

Leung, L. R, and Y. Qian (2009), Atmospheric rivers induced heavy precipitation and flooding in the Western U.S. simulated by the WRF regional climate model, *Geophys. Res. Lett.*, *36*, L03820, doi:10.1029/2008GL036445.

Meehl, G. A., et al. (2007), Global climate projections, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 747–845, Cambridge Univ. Press, Cambridge, U. K., and New York.

Neelin, J. D., B. Langenbrunner, J. E. Meyerson, A. Hall, and N. Berg (2013), California winter precipitation change under global warming in the Coupled Model Intercomparison Project Phase 5 Ensemble, *J. Clim.*, *26*, 6238–6256, doi:10.1175/JCLI-D-12-00514.1.

Neiman, P. J., F. M. Ralph, 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. Hydrometeorol.*, *9*, 22–47.

Newman, M., G. N. Kiladis, K. M. Weickmann, F. M. Ralph, and P. D. Sardeshmukh (2012), Relative contributions of synoptic and low-frequency eddies to time-mean atmospheric moisture transport, including the role of atmospheric rivers, *J. Clim.*, *25*, 7341–7361.

Payne, A. E., and G. Magnusdottir (2014), Dynamics of landfalling atmospheric rivers over the North Pacific in 30 years of MERRA reanalysis, *J. Clim.*, *27*, 7133–7150.

Ralph, F., P. Neiman, and G. Wick (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*, 1721–1745.

Rienecker, M. M., et al. (2011), MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, *24*, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Rutz, J. J., W. J. Steenburgh, and F. M. Ralph (2014), Climatological character of atmospheric rivers and their inland penetration over the Western United States, *Mon. Weather Rev.*, *142*, 905–921.

Ryoo, J., Y. Kaspi, D. Waugh, G. Kiladis, D. Waliser, E. Fetzer, and J. Kim (2013), Impact of Rossby wave breaking on U.S. West Coast winter precipitation during ENSO events, *J. Clim.*, *26*, 6360–6382, doi:10.1175/JCLI-D-12-00297.1.

Seager, R., N. Naik, and G. A. Vecchi (2010), Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming, *J. Clim.*, *23*, 4651–4668.

Sodemann, H., and A. Stohl (2013), Moisture origin and meridional transport in atmospheric rivers and their association with multiple cyclones, *Mon. Weather Rev.*, *141*, 2850–2868.

Stohl, A., C. Forster and H. Sodemann (2008), Remote sources of water vapor forming precipitation on the Norwegian west coast at 60°N—A tale of hurricanes and an atmospheric river, *J. Geophys. Res.*, *113*, D05102, doi:10.1029/2007JD009006.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*, 485–498.

Trenberth, K. E. and C. J. Guillemot (1995), Evaluation of the global atmospheric moisture budget as seen from analyses, *J. Clim.*, *8*, 2255–2272.

Warner, M. E., C. F. Mass, and E. P. Salathé Jr. (2015), Changes in winter atmospheric rivers along the North American west coast in CMIP5 climate models, *J. Hydrometeorol.*, *16*, 118–128.

Yin, J. H. (2005), A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophys. Res. Lett.*, *32*, L18701, doi:10.1029/2005GL023684.

Zhu, Y. and R. Newell (1998), A proposed algorithm for moisture fluxes from atmospheric rivers, *Mon. Weather Rev.*, *126*, 725–735.