

MACHINE LEARNING

Earth



PAPER

OPEN ACCESS

RECEIVED
29 May 2026

REVISED
8 September 2026

ACCEPTED FOR PUBLICATION
22 September 2026

PUBLISHED
20 October 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Investigating western North America atmospheric rivers with machine learning-based detection

Katherine Dagon^{1,*} , Teagan King² , Christine A Shields¹ , John Truesdale¹ , Sol Kim³ ,
Kirsten J Mayer¹ , Andre Graubner⁴ , Annette Greiner⁵ , Lukas Kapp-Schwoerer⁷ ,
and Karthik Kashinath⁶

¹ NSF National Center for Atmospheric Research, Boulder, CO, United States of America

² Tomorrow.io, Golden, CO, United States of America

³ University of Chicago, Chicago, IL, United States of America

⁴ ETH Zurich, Zurich, Switzerland

⁵ NERSC, Lawrence Berkeley National Laboratory, Berkeley, CA, United States of America

⁶ NVIDIA, Santa Clara, CA, United States of America

⁷ Independent researcher.

* Author to whom any correspondence should be addressed.

E-mail: kdagon@ucar.edu

Keywords: atmospheric rivers, machine learning, Earth system models, detection and tracking, extreme precipitation, climate change

Supplementary material for this article is available [online](#)

Abstract

Atmospheric rivers (ARs) are long, narrow conduits transporting water vapor poleward. ARs are often associated with extreme precipitation and can have significant impacts on humans and the environment when they make landfall. Here we use a previously developed deep learning model, CGNet, to detect ARs in high resolution Community Earth System Model (CESM) historical and future simulations and evaluate the results using multiple reanalysis products. CGNet is a convolutional neural network trained on expert-labeled AR and tropical cyclone masks from the ClimateNet project. We probe the sensitivities of CGNet by testing how different characteristics of the input (variables, normalization values, spatial and temporal resolution) impact the resulting AR detection. Global and regional AR frequency magnitudes are strongly dependent on source data. To understand future changes, we compare features of the detected ARs across CESM simulations using historical and Representative Concentration Pathway (RCP) 8.5 forcing. Focusing on western North America, we find that ARs largely increase over California and decrease over the Pacific Northwest and western Canada in response to projected warming. We leverage built-in spatiotemporal tracking with CGNet to connect regional changes in ARs with shifts in lifecycle characteristics. We further investigate precipitation associated with ARs and find that AR precipitation increases in California are largely driven by changes in AR precipitation intensity, whereas AR frequency provides an important contribution to AR precipitation changes at more northerly latitudes. AR-associated extreme precipitation is projected to increase over much of the region, with important implications for impacts.

1. Introduction

Atmospheric rivers (ARs) are commonly defined as narrow conduits transporting water vapor poleward from tropical and/or extratropical moisture sources (American Meteorological Society 2026). ARs are frequently associated with extreme precipitation, especially in regions with coastal orography like western North America (Leung and Qian 2009). AR-driven extreme precipitation can lead to devastating impacts on humans and ecosystems through flooding and landslides (Ralph *et al* 2019a) that can impact infrastructure (Michaelis *et al* 2022), while also providing important long-term moisture sources such as mountain snowpack (Guan *et al* 2010). This interplay points to a need to

understand how ARs and their associated precipitation will evolve with regional and global climate changes.

Increased interest in understanding ARs and related impacts has aided the development of a wide range of detection methods for identifying ARs (Zhu and Newell 1998, Ralph *et al* 2004). The AR Tracking Method Intercomparison Project (ARTMIP) has been at the forefront of efforts to compare and quantify uncertainties in AR detection algorithms, which has led to enhanced understanding of ARs in current and future climates (Shields *et al* 2018, 2023) and the downstream effects of AR detection uncertainty on humans and the environment (Rutz *et al* 2019, Lora *et al* 2020, Reid *et al* 2020, Collow *et al* 2022). In particular, Henny and Kim (2025) applied global ARTMIP algorithms to reanalysis data and found that ARs are consistently becoming larger, more frequent, and associated with increased moisture content. In addition, Scholz and Lora (2024) found that ARs were associated with warmer winters and extreme heat events due to enhanced sensible heat transport, with implications for longer term changes in regional temperatures and global heat transport.

To better understand future changes, O'Brien *et al* (2022) applied global ARTMIP algorithms to Representative Concentration Pathway 8.5 (RCP8.5) and Shared Socioeconomic Pathways 5-8.5 (SSP5-8.5) projections using models from the fifth and sixth phases of the Coupled Model Intercomparison Project (CMIP5/6) and found increases in AR frequency, counts, and sizes that were particularly notable along western coastlines. Payne *et al* (2020) reviewed the characteristics of future ARs using a theoretical framework and found a global increase in AR frequency and intensity of AR precipitation driven by an increase in atmospheric moisture, though they point to the need for higher resolution Earth system model simulations to better represent AR processes. Shields *et al* (2023) used a higher resolution atmosphere model simulation that confirmed the global AR frequency and intensity increases with climate change, but also showed a large spread of AR precipitation responses, largely due to occurrence differences across detection algorithms. However, all methodologies robustly projected increases in extreme precipitation rates. Looking regionally, Payne and Magnusdottir (2015) evaluated North Pacific land-falling ARs in RCP8.5 projections using CMIP5 models and found a broadening of the historical distribution of AR frequency, consistent with an increase in atmospheric moisture. Using high resolution downscaling to study future extreme AR events in California, Huang *et al* (2020a) found a significant increase in associated total precipitation and precipitation intensity. Zhou *et al* (2024) analyzed AR clusters over the west coast of the United States in future climate simulations and found that more extreme ARs increased in frequency. Common throughout these studies is an emphasis that the choice of detection algorithm is an important factor in quantifying future changes in ARs, with careful consideration of detection methods needed based on scientific questions and/or region of interest (O'Brien *et al* 2022, Shields *et al* 2023).

Many of the existing detection algorithms use specific thresholds of physical quantities such as integrated water vapor, in a relative or absolute sense, to identify ARs. The use of thresholds can pose an issue in applying these detection algorithms to different climate states (Rush *et al* 2025). In recent years, machine learning (ML) or deep learning (DL) has emerged as an additional tool for automated detection of extreme weather features such as ARs. Employing a technique like semantic segmentation allows for threshold-free detection of extreme weather features like ARs, which can aid in the objectivity and computational expense of these algorithms. However, the use of supervised ML requires AR labels to provide a ground truth for training the algorithm. Previous work applying ML and DL to the detection of ARs has utilized heuristics such as the Toolkit for Extreme Climate Analysis (TECA, Prabhat *et al* 2012), to prepare a labeled dataset for training (e.g., Liu *et al* 2016, Mudigonda *et al* 2017, Racah *et al* 2017, Kurth *et al* 2018, Muszynski *et al* 2019). The resulting DL models were somewhat limited in performance due to the quality and reliability of the underlying training data. Galea *et al* (2024) used the TempestExtremes framework (Ullrich and Zarzycki 2017) to generate AR labels and train a DL algorithm with good performance on reanalysis data, though the high accuracy was likely due to the consistency and long record of ARs identified by TempestExtremes in reanalysis.

To improve the quality of training data and expand the applications of the ML models, Prabhat *et al* (2021) developed ClimateNet, an expert-labeled dataset of ARs and tropical cyclones (TCs). This hand labeled data was used to train a DL model to detect ARs and TCs via semantic segmentation, without imposing threshold conditions. Kapp-Schwoerer *et al* (2020) extended this work to produce a lightweight version of the segmentation model called CGNet. (More details on ClimateNet and CGNet will be discussed in section 2.) Higgins *et al* (2023) compared the performance of CGNet with threshold-based ARTMIP algorithms and found the approaches were generally consistent with each other with the exception of spatial area, demonstrating that hand labeled ARs have a tendency to be biased larger than heuristic-based methods. Tian *et al* (2023) used an ensemble approach to combine 20 segmentation

algorithms trained on the ClimateNet dataset, and found that they could increase AR detection accuracy with the ensemble relative to individual algorithms. Mahesh *et al* (2024) used a semi-supervised ML approach to identify ARs without the need for additional labeled data, leveraging the existing ARTMIP algorithms to produce an AR consensus index. However they did not apply their method to future climate simulations, noting that type of application is important to understanding how ML-detected ARs might change in the future.

In this paper, we use the CGNet DL model to identify and track ARs in high-resolution coupled Earth system model simulations. Distinguishing this work from previous ML efforts, here we use a future climate simulation with a warmer background base state to analyze its impact on global and regional AR frequency and other lifecycle characteristics. While CGNet is an existing method, its application to high resolution climate simulations under historical and future conditions, along with accompanying sensitivity tests, represent novel contributions. To address these gaps, this study investigates the following research questions:

1. What are the sensitivities of the CGNet algorithm to different input fields, normalization values, and spatial and temporal resolution of the input data?
2. How do Earth system model ARs produced by CGNet compare to existing ARTMIP catalogues and ML-detected ARs in reanalysis data?
3. How might AR characteristics such as frequency and lifecycle change in the future?
4. What are the drivers of changes in total and extreme precipitation from ARs over western North America?

2. Data and methods

2.1. ClimateNet project

The ClimateNet project was developed to use ML to detect and classify TCs and ARs (Prabhat *et al* 2021). An expert-labeled dataset was first assembled to train the ML model. A series of labeling campaigns were held in 2019 to collect these expert labels of TCs and ARs using a web-based ClimateContours tool. The ClimateContours tool presents labelers with a snapshot from a Community Atmosphere Model (CAM5.1) historical simulation at 25 km resolution (Wehner *et al* 2014, 2018). The CAM5.1 historical simulation used for the labeling covers the time period 1995–2015 with natural and anthropogenic forcings. The atmosphere model is forced by prescribed sea surface temperatures and sea ice extent following the Atmosphere Model Intercomparison Project (AMIP) protocol. Labelers use their cursor to trace the outline of an AR or TC by placing vertices of a polygon around the event. When the last vertex coincides with the first vertex, the labeler selects the event type (AR or TC). The ClimateContours tool provides users with a set of six variables they may switch between in order to inform their labeling of TCs and ARs: integrated vapor transport (IVT) ($\text{kg m}^{-1} \text{s}^{-1}$), integrated water vapor (mm), vorticity (s^{-1}), surface winds (m s^{-1}), 850 hPa winds (m s^{-1}), and sea level pressure (PSL) (hPa). About 80 scientists participated in these campaigns which collected 219 unique images. Each image was labeled by multiple experts, resulting in 459 labeled images over the time period from June 1996 to September 2013. The ClimateNet datasets are available at: <https://portal.nersc.gov/project/ClimateNet> and more details on quality assurance and quality control adjustments can be found in Prabhat *et al* (2021).

2.2. CGNet detection algorithm

The ClimateNet dataset was first used to train a deep neural network architecture DeepLabv3+ (Chen *et al* 2018). DeepLabv3+ is a segmentation model with an encoder-decoder framework which takes in atmospheric data as input, and produces segmentation masks as output with values 0 (background), 1 (TC), or 2 (AR) at each gridpoint. While this method is effective in segmenting extreme weather events (Prabhat *et al* 2021), it is also computationally expensive and can be difficult to train. To address these issues, Kapp-Schwoerer *et al* (2020) adopted CGNet, a more lightweight segmentation model than DeepLabv3+ (Wu *et al* 2021). CGNet uses a convolutional neural network approach with a core building block called the CGBlock which integrates local and surrounding features into a joint feature. This model uses 4 input variables from the ClimateNet dataset to detect TCs and ARs: total vertically integrated water vapor (TMQ), zonal wind at 850 hPa (U850), meridional wind at 850 hPa (V850), and PSL. The CGNet model code is available at: <https://github.com/andregraubner/ClimateNet>.

The original CGNet is trained on the 25 km resolution CAM5.1 historical simulations and the corresponding expert labels from ClimateNet. Here, we take the CGNet architecture and repeat the training process from Kapp-Schwoerer *et al* (2020) by splitting the 459 total samples into 398 samples (1996–2010) for training and 61 samples (2011–2013) for testing. The training and testing subsets are split using different blocks of years to account for temporal autocorrelation in the data (Mahesh *et al* 2024). We train with a batch size of 16 over 20 epochs using the Jaccard loss (related to the Intersection over Union, or IoU) (Rahman and Wang 2016) and the Adam optimizer (Kingma and Ba 2017). Radke *et al* (2025) trained CGNet with different batch sizes (4 and 10) and found very similar resulting IoU scores, noting that a larger batch size can reduce training time without negatively impacting skill. Following Kapp-Schwoerer *et al* (2020) we use a learning rate of 10^{-3} , which was confirmed by Tian *et al* (2023) as a suitable value to maintain accuracy with a reasonable training time.

We train using two V100 GPUs on the Casper analysis cluster at NSF NCAR. Each epoch takes about 1 min to run and the model reaches 60% mean training IoU and 56% mean testing IoU after 20 epochs (table S1). The Jaccard loss function for training CGNet was used to address data imbalance (Kapp-Schwoerer *et al* 2020). The training history as a function of Jaccard loss and mean training IoU over 20 epochs are shown in figure S1. Training data loss changes are subtle but trend downward (figure S1(a)), while the mean IoU trends upward throughout the training period (figure S1(b)). Figure S1(b) also shows the breakdown of training IoU across the output channels: background, TC, and AR, along with the testing data results. IoUs are greatest for background, then ARs, then TCs, due to the dominance of the background class. Testing IoUs are comparable to the performance of the last training epoch but slightly decreased (and more so for ARs). Prabhat *et al* (2021) note that the IoUs for ARs and TCs are strongly impacted by even minor differences in event boundaries, and, even when human experts agree reasonably well on labels, the mean IoU values can be depressed. They emphasize that due to this sensitivity to small disagreements between predicted masks and the ground truth, IoU values smaller than 1 are possible even for qualitatively correct predictions.

We also train a model with double the number of epochs (40), and the change in loss is negligible over this period (not shown). The IoU performance on testing data at 40 epochs degrades relative to the 20-epoch training, indicating the network at this point is likely overtrained. Tian *et al* (2023) also found that the optimal number of epochs for training an ensemble of DL-based AR detection algorithms (including CGNet) was between 10 and 30.

In addition to developing CGNet as a lightweight algorithm, Kapp-Schwoerer *et al* (2020) demonstrate spatiotemporal tracking as an extension of the detection work. Connected components are identified for both ARs and TCs, and these components are then assigned IDs and tracked over time. We use the same approach to track AR events, investigate genesis and termination locations, and produce summary statistics including lifetime, travel distance, and speed. We use the same minimum event criteria as Kapp-Schwoerer *et al* (2020) where event duration must be at least 12 h and size must be at least 250 pixels, consistent with other AR detection and tracking studies (e.g., Ralph *et al* 2019b, Inda-Díaz *et al* 2021, Kim *et al* 2025).

2.3. Earth system model simulations and reanalysis datasets

In this paper we apply the re-trained CGNet model to high resolution Earth system model simulations with the Community Earth System Model (CESM), version 1.3 (CESM1.3) (Meehl *et al* 2019). We use simulations of present-day and future climate change to detect ARs and understand associated impacts. The atmosphere model of CESM1.3 is CAM5 (Park *et al* 2014) with a spectral element dynamical core (Dennis *et al* 2012), a horizontal resolution of 0.25° , and 30 vertical atmospheric levels. The sea ice component is the Community Ice Code Version 4 (CICE4, Hunke and Lipscomb 2008), the ocean component is the Parallel Ocean Program Version 2 (POP2) with a horizontal resolution of 1° and higher resolution in the equatorial tropics (Smith *et al* 2010, Danabasoglu *et al* 2012), and the land component is the Community Land Model Version 4 (CLM4, Lawrence *et al* 2011), with the River Transport Model Version 1 (Branstetter and Famiglietti 1999).

We utilize years 2000–2005 from the historical simulation of CESM1.3, with time varying natural and anthropogenic forcings. We extend the historical simulation using a simulation with RCP 2.6 forcing from 2006 to 2014. The combined 2000–2014 time period will be referred to in the paper as the CESM historical simulation and serve as our present-day baseline. We utilize the years 2086–2100 from a simulation with RCP 8.5 to investigate the response to projected warming. We use CESM simulation output at 3-hourly frequency to match how the CGNet model is trained. For all simulation periods we use 3-hourly instantaneous U850, V850, and PSL. However, 3-hourly TMQ is not a default model output so for the historical (2000–2005) and RCP8.5 (2086–2100) simulation periods we calculate TMQ using 3-hourly average surface pressure (PS) and humidity (Q). Further, 3-hourly Q is not available for the

RCP2.6 simulation (2006–2014) so we use 6-hourly average TMQ which we interpolate to 3-hourly resolution. While previous work suggests that these approximations do not significantly impact the resulting feature detection (Dagon *et al* 2022a), interpolating a 6-hourly moisture field to a higher frequency will artificially smooth out peak values that are important to the dynamic nature of ARs, representing a potential limitation of this work. All CESM simulation output used in the analysis is available through the NCAR Geoscience Data Exchange (Dagon *et al* 2022b, 2026a) and software developed for this study is available through GitHub (Dagon *et al* 2026b).

We also use data from two reanalysis products to evaluate the detection of ARs with CGNet: the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2, Gelaro *et al* 2017), with a horizontal resolution of 0.5° latitude by 0.625° longitude, and the European Centre for Medium-Range Weather Forecasts' Reanalysis Version 5 (ERA5, Hersbach *et al* 2020), with a horizontal resolution of 0.25° latitude by 0.25° longitude. We leverage previously published ARTMIP catalogues of ARs detected in these reanalysis products (Shields *et al* 2022, Rutz *et al* 2024), as well as the associated source data, to understand differences across datasets.

2.4. Associating precipitation with ARs

To explore precipitation associated with the CGNet-detected ARs, we use the precipitation output from the aforementioned CESM1.3 simulations. Total precipitation is calculated by summing 3-hourly average precipitation rate on a seasonal basis for each gridpoint over time. Total precipitation associated with an AR is calculated by summing precipitation where there is a detected AR at the same gridpoint over time. AR precipitation is divided by total precipitation to get the fraction of total precipitation associated with an AR, on a seasonal basis at each gridpoint. We calculate changes in AR precipitation by comparing the 15 year CESM historical simulation (2000–2014) with the 15 year CESM RCP8.5 simulation (2086–2100).

We further separate the changes in AR precipitation into frequency and intensity contributions, using an approach outlined in Utsumi *et al* (2016), Pepler *et al* (2021), which was applied to changes in frontal precipitation in Dagon *et al* (2022a). The change in total AR precipitation ΔP_{AR} is calculated as the sum of three terms: a frequency term, an intensity term, and a covariance term (equation (1)). The frequency term is calculated by multiplying the change in AR frequency Δn_{AR} by the mean AR precipitation intensity in the historical simulation I_{AR} . The intensity term is calculated by multiplying the AR frequency in the historical simulation n_{AR} by the change in mean AR precipitation intensity ΔI_{AR} . The covariance term is the product of the frequency and intensity changes, and because it is small compared to the other two terms we limit our analysis to the frequency and intensity terms.

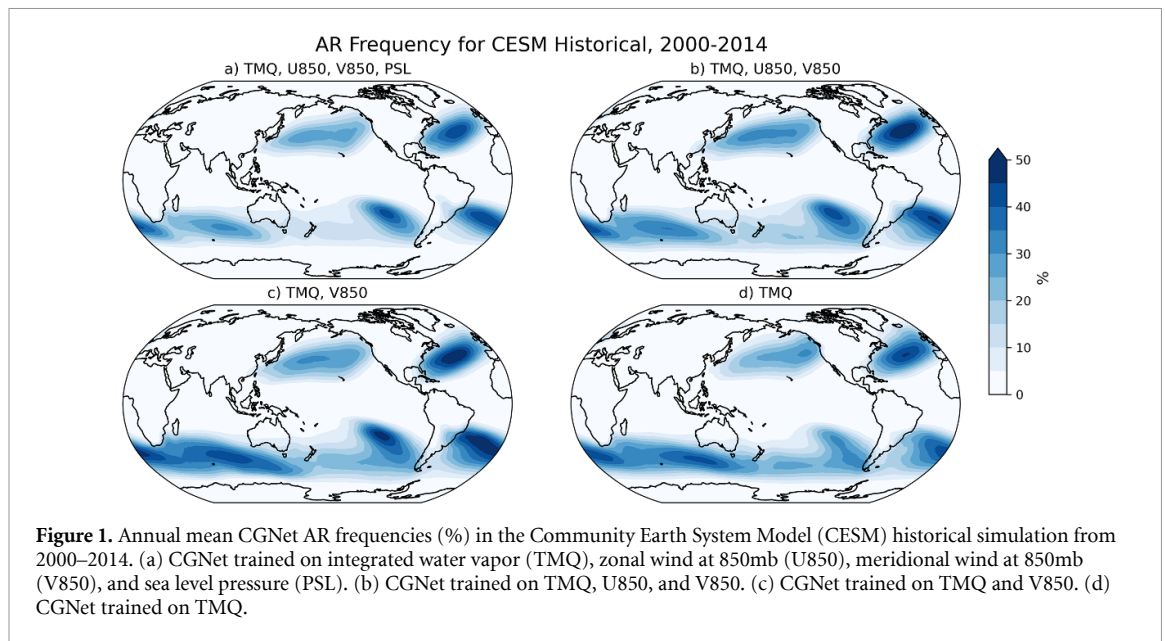
$$\Delta P_{AR} = \Delta n_{AR} I_{AR} + n_{AR} \Delta I_{AR} + \Delta n_{AR} \Delta I_{AR}. \quad (1)$$

Extreme precipitation is calculated using a condition of greater than 90th percentile precipitation over land gridpoints. We calculate the 90th percentile precipitation at each gridpoint for each season, and then subset the timepoints where precipitation exceeds that location- and season-specific threshold. Furthermore, we use the same baseline period from the CESM historical simulation (2000–2014) to calculate extreme precipitation in both the historical and RCP8.5 simulations (Tebaldi *et al* 2006, Dagon *et al* 2022a). AR-associated extreme precipitation is calculated the same as AR-associated total precipitation, by selecting gridpoints with a detected AR to subset the extreme precipitation. We calculate the probability ratio *PR* metric (Fischer and Knutti 2015, Dagon *et al* 2022a) to investigate changes in the probability of AR extreme precipitation (equation (2)). The *PR* compares the conditional probability of AR extreme precipitation $N_P|AR/N_{AR}$ with the climatological probability of extreme precipitation N_P/N . Here N is the number of time steps, N_P is the number of time steps with extreme precipitation, N_{AR} is the number of time steps with ARs, and $N_P|AR$ is the number of extreme precipitation time steps associated with ARs.

$$PR = \frac{N_P|AR/N_{AR}}{N_P/N}. \quad (2)$$

2.5. Statistical significance tests

Throughout the paper, statistical significance of differences is assessed using a *t*-test. For related groups such as the same dataset with different data processing, we use a two-sided paired *t*-test to compare the yearly means of dependent samples. For independent groups such as different datasets or future versus historical time periods, we use a two-sided Welch's *t*-test to compare the yearly means of independent samples.



3. CGNet sensitivity tests

3.1. Input fields

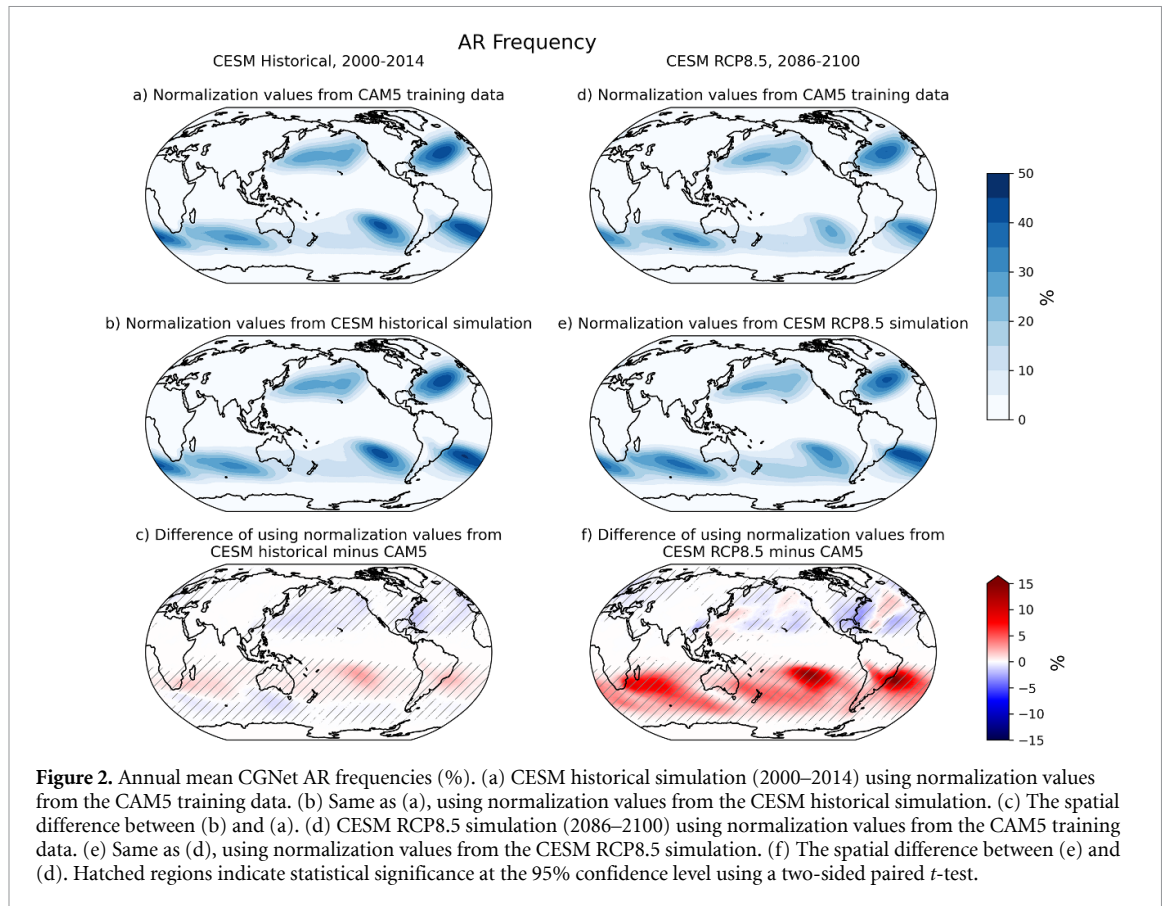
We repeat the training process described in section 2.2 with different combinations of atmospheric variables, to test the sensitivity of the CGNet algorithm to input fields. In addition to the 4-input-field version described above, we also train a version of the ML model with U850, V850, and TMQ (3-input model), a version with V850 and TMQ (2-input model), and a version with only TMQ (1-input model). These tests are to primarily understand the sensitivity of AR detection to different inputs, though we also include the IoU results for TC detection. The final training and testing mean IoUs along with the individual IoUs for each of the output categories (background, AR, TC) are shown in table S1. AR IoU values are consistently higher than TC values across the different trained models, confirming previous extreme weather segmentation results (Kurth *et al* 2018, Prabhat *et al* 2021). While AR training IoUs stay relatively constant across the trained model versions (~47%), AR testing IoUs are higher with the 4- and 3-input models (~40%) and decrease as the wind fields are removed from the input. However, the model trained on TMQ only is still able to detect ARs with a testing IoU of 37%, while TCs are unable to be detected in this version.

Figure 1 shows the global frequency maps for ARs detected in the CESM historical simulation across the different versions of CGNet. In general, the pattern of global AR detection frequencies are similar across the different versions. AR frequency goes up slightly when PSL and U850 are removed as inputs (figures 1(a)–(c)). Relative to the 3- and 2-input versions, AR frequency goes down when V850 is further removed but the overall frequency is slightly higher than the 4-input version of the model, demonstrating that TMQ alone is sufficient to globally detect ARs (figure 1(d)).

3.2. Normalization values

Because each input variable has different ranges of values and units, it is important to apply some normalization or standardization, such as z-score, to the input data before training (Kapp-Schwoerer *et al* 2020, Radke *et al* 2025). Typically the global means and standard deviations that are used for normalization are calculated across the training data and those values are also used when applying the trained model to unseen data, such as the test set or other out-of-sample applications (Géron 2019, Feng *et al* 2025). This is because the training data is assumed to be drawn from the same population as the data the model will encounter in deployment, and thus its mean and standard deviation serve as estimates of the true population statistics. However, this assumption may not hold when evaluating a future with a warmer background base state. Nevertheless, using a consistent set of normalization values can help the model retain the scale for AR detection and avoid losing its reference point.

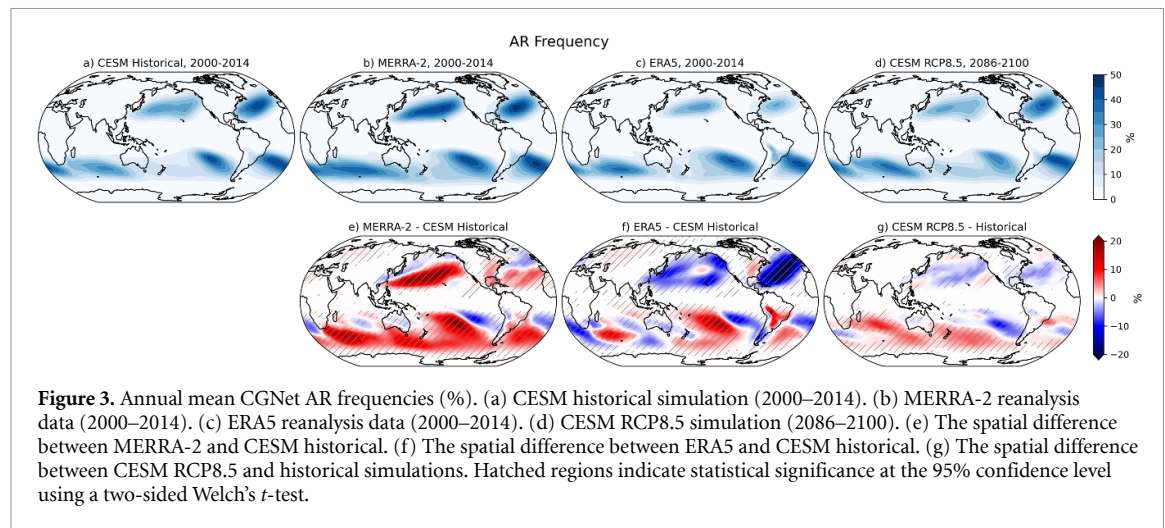
To better understand the impact of normalization values on the resulting detection of ARs, we vary the global mean and standard deviation values that are used in the z-score normalization for CGNet. Initially, we test the impact of using training data for normalization from an atmosphere-only versus a coupled historical simulation, before exploring the impact of normalization under a future climate using



a coupled simulation. Specifically, we analyze AR detection results when 1) the global mean and standard deviation values are calculated from the original CGNet training data (CAM5.1 historical simulation) compared to when 2) the normalization values are calculated from the unseen Earth system model data used for inference (CESM1.3 historical simulation). It is important to note that due to the sparse nature of the labeled data from CAM5.1, the global means and standard deviations were not calculated over a continuous time period but rather from the individual training data samples which unevenly sample 398 time points from 1996–2010. We show the resulting AR detection of option 1 in figure 2(a) compared to the AR detection from option 2 in figure 2(b).

In practice the normalization values do not vary substantially (see table S2), but they do impact the resulting AR detection on the order of 2%–3% significant differences in AR frequency (figure 2(c)). When comparing the AR detection with normalization values taken from option 2 (the CESM climatology) versus option 1 (the CAM climatology), we find that AR detection decreases in the Northern Hemisphere, increases in the Southern midlatitudes, and decreases in the Southern Ocean. It is possible that this sensitivity test is revealing structural model differences between the AMIP-style simulation and fully coupled CESM and/or the shift in years, number of data points, and discontinuity of the data used to calculate the means and standard deviations. For consistency, in the remainder of the paper we use means and standard deviations calculated from the original training data for AR detection in the historical simulation.

Since we found minimal differences between CAM and CESM historical climate normalization, we next consider the choice of normalization values when applying the trained CGNet model to a future coupled climate simulation. Again we use the variable global means and standard deviations calculated from the original CGNet training data (figure 2(d)) and compare the resulting AR detection in the CESM RCP8.5 simulation to using normalization values from the RCP8.5 simulation itself (figure 2(e)). When using the values from the future climate simulation, the means and standard deviations are roughly similar to the historical values, with the exception of TMQ which increases in mean and standard deviation (table S2). This increase in TMQ is expected in a warming world with more moisture (Held and Soden 2006). When we consider the difference in using historical versus future normalization values, we see that the changes in AR detection are greater than comparing different historical values, especially in the Southern Hemisphere where there are statistically significant increases of greater than 15% (figure 2(f)). We also see that the spatial patterns of the two difference plots are similar, again



potentially highlighting structural model differences that have imprinted onto AR detection through the choice of normalization values.

Previous research has found that updating the normalization values can aid in the ability of ML to generalize to unseen climates (Dagon *et al* 2022a, Beucler *et al* 2024). One reason for improved generalization could be because normalizing with unseen climate statistics would likely shift the raw values toward the original training dataset distribution. In this case, by normalizing with a future, warmer climate, the changes due to thermodynamics (e.g. the Clausius–Clapeyron (C–C) relationship) are essentially removed, tailoring the analysis to dynamical changes. Here we are interested in the relative changes to the dynamic nature of ARs under climate change (e.g. AR genesis and termination, associated precipitation and impacts), so in the remainder of the paper we use results based on updating the normalization values for detection under RCP8.5.

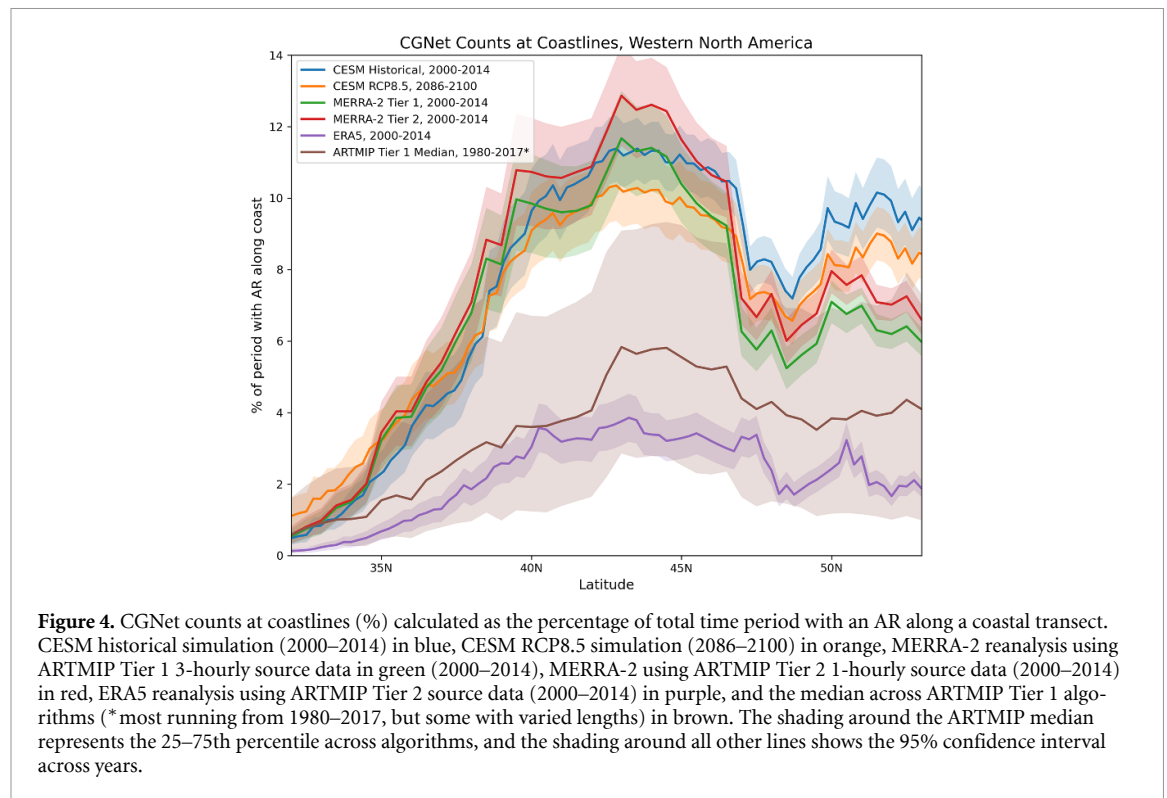
4. Evaluation of detected ARs

4.1. CGNet in the context of ARTMIP

To evaluate the CGNet-detected ARs in CESM, we first compare the results to ARTMIP, which was designed to understand and quantify differences in detection and tracking methodologies for ARs. Characteristics of ARs such as frequency and seasonality are strongly dependent on the detection method (Rutz *et al* 2019), and this translates to impacts of ARs such as associated precipitation (Collow *et al* 2022). It is useful and important to understand how the CGNet algorithm compares to other ARTMIP methods.

The first phase of ARTMIP (Tier 1) launched in 2016 and applied 20 different global and regional AR detection tools (ARDTs) to MERRA-2 data from January 1980 to June 2017 with the experimental design described in Shields *et al* (2018). Since then, ARTMIP has expanded to over 40 ARDTs with results remaining consistent with Rutz *et al* (2019). Lora *et al* (2020) used a subset of Tier 1 methods to focus on those with global coverage. In particular their figure 1a showed the ensemble average AR frequency from 1980–2016 across 10 methods. There are several AR frequency hotspots over midlatitude oceanic regions, where the ensemble mean frequencies reach 20%–25%. Comparing this result to the global map of annual mean AR frequencies using CGNet applied to the CESM historical simulation (figure 3(a)), we see that while the spatial patterns and ocean hotspot regions are similar, the frequency of detection is much greater in CESM with CGNet (peaking around ~45%).

A regional metric we use for further comparison to ARTMIP is ‘counts at coastlines’, which selects points from a coastal transect that have fractional land/sea cover between a given latitude band (Rutz *et al* 2019). The AR counts are averaged across longitudes (typically there are only a few points selected at each latitude), and the counts at coastlines metric is then calculated as a percentage of the total time period with an AR along the coast. A comparison of ARTMIP Tier 1 algorithms applied to MERRA-2 data from 1980–2017 for a coastal transect in western North America (between 32–55°N and 130–115°W) is shown in figure S3 of Shields *et al* (2023). We repeat this metric with the CGNet results from the CESM historical simulation over 2000–2014 (figure 4, blue line) and MERRA-2 data as part of ARTMIP Tier 1 using a published catalogue of CGNet ARs from Rutz *et al* (2024), also over 2000–2014



(figure 4, green line). The CESM results are similar to CGNet in ARTMIP, with a primary peak in counts between 40–45°N and a secondary peak between 50–55° N. We also calculate the median across ARTMIP Tier 1 algorithms from Shields *et al* (2023) (figure 4, brown line), confirming the higher frequency of AR detection in CGNet relative to other ARTMIP methods.

4.2. CGNet ARs in reanalysis products

We further compare the CESM historical ARs with ARs detected in MERRA-2 and ERA5 over the same time period (2000–2014). Again we utilize published ARTMIP catalogues of ARs detected by CGNet in reanalysis (Shields *et al* 2022, Rutz *et al* 2024). Figures 3(b) and (c) show the global frequency maps for annual mean MERRA-2 and ERA5 ARs from CGNet. The locations of AR hotspots compare well across CESM and the reanalysis products, though the frequency magnitudes differ. The peak magnitudes are generally higher in MERRA-2 relative to the CESM historical simulation (figure 3(e)). The difference plot indicates there is likely a southward shift in peak locations, including the Pacific and Atlantic Ocean basins in both hemispheres. ERA5 ARs are generally less frequent than ARs in CESM, especially in the Northern Hemisphere (figure 3(f)). In the Southern Hemisphere, the southward shift in peak locations which was evident in the MERRA-2 difference plot is also seen with ERA5. Given the consistency in the Southern Hemisphere differences, this could point to a difference in how CESM represents winds and water vapor relative to the reanalysis products, which then translates to a shift in detected AR location with CGNet. In contrast, the varying Northern Hemisphere AR frequencies across the Earth system model and reanalysis products points to inconsistencies in AR detection with CGNet, and may be related to differences in spatial and temporal resolution of the datasets (which will be discussed further in section 4.3). To check that the reanalysis differences with CESM are not driven by the particular years sampled, we also plot global maps of MERRA-2 and ERA5 CGNet AR frequencies for the full time periods available from ARTMIP (MERRA-2: 1980–June 2017, ERA5: 2000–2019) and do not find any significant differences (not shown).

To better understand the differences in AR detection across CESM and the reanalysis products, we compare the annual mean climatologies of the primary input variables used to detect ARs (integrated water vapor, zonal wind at 850mb, and meridional wind at 850mb) (figure S2). We find that water vapor is generally lower in MERRA-2 and ERA5 versus CESM, with the exception of some land regions and the tropical Pacific (figures S2(a) and (d)), the latter likely related to the double intertropical convergence zone issue in Earth system models (Zhou *et al* 2022). The differences in wind velocities are more regionally variable (figures S2(b), (c), (e) and (f)). The lower integrated water vapor in ERA5 relative to MERRA-2 (figure S2(g)) could be contributing to the lower AR frequencies in ERA5 relative to

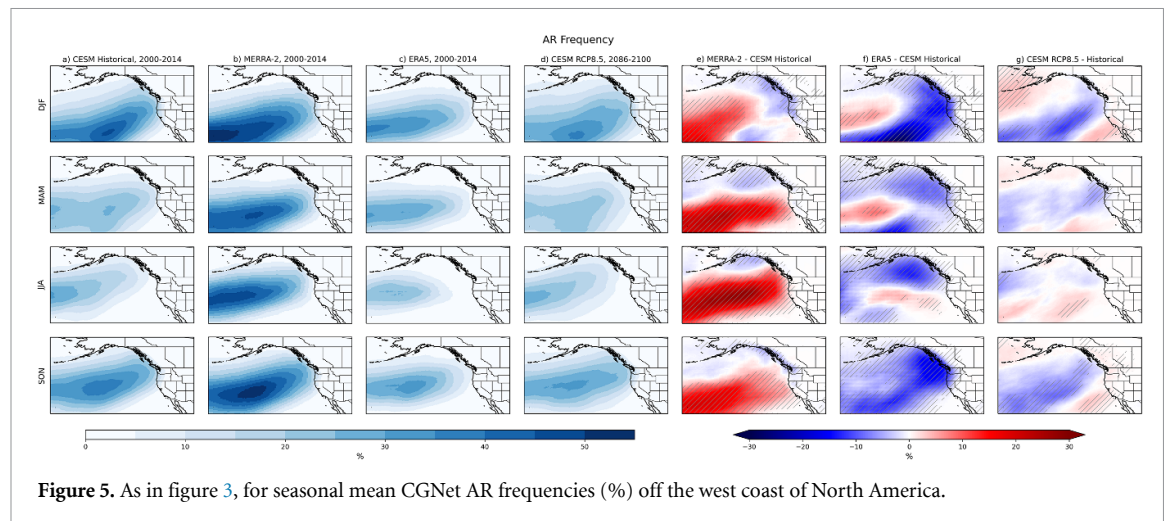


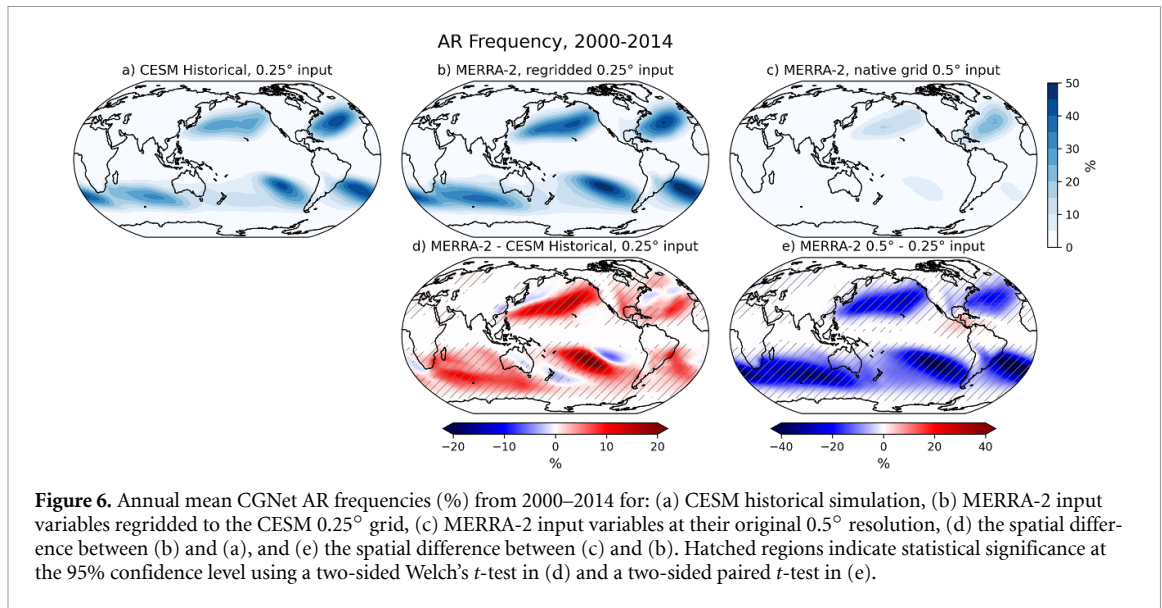
Figure 5. As in figure 3, for seasonal mean CGNet AR frequencies (%) off the west coast of North America.

CESM (figure 3(f)). In general the underlying input fields from the reanalysis products are more similar to each other than to CESM, indicating that the differences in AR detection across reanalysis products and CESM (figure 3) are unlikely to be fully explained by the differences in the base states of winds and water vapor.

Looking regionally, we next compare AR frequencies off the west coast of North America (between 25–65°N and 180–105°W) (figure 5). Similar to the global picture, this region shows an overall higher magnitude of ARs detected in MERRA-2 versus CESM in this region (figure 5(e)). This statistically significant increase is also consistent across seasons, where seasonal AR frequencies are calculated over four 3-month periods: December–January–February (DJF), March–April–May (MAM), June–July–August (JJA), and September–October–November (SON). All seasons do show a localized lower magnitude but statistically significant decrease in MERRA-2 ARs relative to CESM, focused in the northern coastal latitudes. Regional and seasonal AR frequencies in ERA5 also confirm the overall lower magnitude of ARs relative to the CESM historical simulation (figure 5(f)). Some seasons show lower magnitude increases in ARs in ERA5 relative to CESM, though these increases are generally not located near the coastline.

Considering the seasonal differences in winds and water vapor in this region, we again find that water vapor is generally lower in both reanalysis products relative to CESM across seasons with some exceptions off the coast in MAM and JJA (figures S3(a) and S4(a)) which could be contributing to localized higher AR frequencies in those seasons (figures 5(e) and (f)). Water vapor is lower in ERA5 than MERRA-2 in this region (figure S5(a)), again potentially contributing to lower AR frequencies in ERA5 relative to CESM (figure 5(f)). Wind velocities off the west coast of the U.S. are generally higher in the reanalysis products than in CESM (figures S3(b),(c) and S4(b),(c)). One notable exception is meridional wind in DJF, which is lower in MERRA-2 and ERA5 than in CESM. Again the differences in underlying input fields between the reanalysis products and CESM are greater than the differences in underlying input fields between the reanalysis products themselves (figure S5), indicating that other factors are contributing to the regional and seasonal differences in AR detection across the model and reanalysis datasets (figure 5).

To compare landfalling ARs across CESM and the reanalysis products, we can use ARTMIP's 'counts at coastlines metric' calculated over the same time period as the CESM historical simulation (2000–2014). The CESM historical results compare well with MERRA-2 at the peak latitudes (~40–45°N), but the CESM counts are significantly higher than MERRA-2 counts at higher latitudes (north of 45°N), and slightly lower at lower latitudes (south of 40°N) but within the 95% confidence interval across years. While AR frequencies over the oceanic AR hotspots are generally higher in MERRA-2 than in CESM, including in the Eastern Pacific (figures 3(e) and 5(e)), differences in western North America coastal ARs can vary depending on the subregion. Looking at ERA5, we see a much smaller number of coastal ARs in western North America compared to MERRA-2 and the CESM historical simulation (figure 4, purple line). Unlike with MERRA-2 where the global and regional differences showed higher AR frequencies than CESM but the coastal ARs were still very comparable, here we see that the lower magnitude of detected ARs in ERA5 versus CESM (figures 3(f) and 5(f)) does indeed translate to much lower land-falling ARs in this region. This result also confirms what Collow *et al* (2022) found looking at 10 different AR detection algorithms across the U.S. West Coast, where ERA5 had overall fewer ARs detected than MERRA-2.



To investigate the differences in seasonality of landfalling ARs, we can further separate the counts at coastlines over western North America by season (figure S6). The overall shapes of the plots are similar, though the distribution of seasons differs. The prominent DJF peak in coastal ARs in CESM across 40–45°N (figure S6(a)) is less present in MERRA-2 Tier 1 where DJF and SON have similar magnitudes (figure S6(b)), though a lower magnitude DJF peak is evident in ERA5 (figure S6(d)). The peak season north of 45°N is SON across MERRA-2 and CESM, consistent with Payne and Magnusdottir (2014). SON and DJF show similar AR counts in this region in ERA5 albeit much lower than MERRA-2 and CESM. Differences in MAM are minimal across CESM and MERRA-2, though CESM has greater high latitude coastal ARs than MERRA-2 (figure S6(f)), and ERA5 has far fewer across the entire region (figure S6(h)). In JJA, MERRA-2 has higher overall frequencies than CESM (figure S6(f)), whereas there are essentially no ERA5 landfalling ARs in JJA (figure S6(d)).

4.3. Impact of spatial and temporal resolution

The reanalysis products are both on different grids than the CESM simulations, and MERRA-2 in particular has a lower horizontal resolution (0.5° latitude by 0.625° longitude). ERA5 is on a 0.25° latitude by 0.25° longitude grid which is very similar to, though not exactly the same, as the CESM grid. Because CGNet was trained on atmospheric model simulations with a 0.25° grid, changes to the resolution of the input data will likely impact the ability of the algorithm to detect ARs. The MERRA-2 and ERA5 CGNet-detected ARs from ARTMIP (Shields *et al* 2022, Rutz *et al* 2024) both match their underlying data product's native resolution, however we were unable to determine whether there was any regridding performed during the analysis. If regridding was not performed, this implies that the MERRA-2 data was input into the algorithm at a lower resolution than how CGNet was trained. CGNet would have seen a smaller number of pixels in the input relative to data at higher spatial resolution like ERA5 and CESM, which could help explain the differences in AR frequencies seen in figure 3.

To better understand the impact of spatial resolution, we test AR detection with CGNet on the MERRA-2 input variables regridded via bilinear interpolation to match the CESM horizontal resolution of 0.25° (figure 6). We compare the regridded MERRA-2 results with the CESM historical AR frequencies, and calculate the spatial difference (figure 6(d)). Overall the differences are very similar to (though not exactly the same as) what we observe in figure 3(e), which is based on the existing CGNet catalogue from ARTMIP. To further probe the spatial resolution sensitivity, we use the MERRA-2 input fields at their original 0.5° resolution and run them through CGNet to replicate the ARTMIP results. We find that the AR detection is significantly impacted by using this input spatial resolution (figure 6(c)). We are able to detect some ARs but they are much less frequent than using MERRA-2 input variables that have been regridded to 0.25° (figure 6(e)). This result shows a significant impact of spatial resolution on AR detection in our implementation, and we are unable to produce reasonable results without regridding the input data to the resolution that the ML model was trained on.

We also test differences in AR detection with varying temporal frequency. While the ARTMIP Tier 1 experiments used 3-hourly data from MERRA-2, the Tier 2 experiments used 1-hourly data from both MERRA-2 and ERA5. Previous work has shown that using 1-hourly versus 3-hourly IVT from MERRA-2

results in negligible differences in the climatology of IVT over western North America with the largest differences of about 10% over the ocean (Collow *et al* 2022, figure S1). We compare the resulting CGNet global AR frequencies using ARTMIP Tier 1 and Tier 2 source data in figure S7 (Shields *et al* 2022, Rutz *et al* 2024). The comparison between MERRA-2 data at different temporal frequencies is shown in figure S7(a) (3-hourly) and figure S7(b) (1-hourly). We see that the difference in temporal resolution minimally impacts global AR frequencies on the order of 2%–4% though the changes are statistically significant (figure S7(d)). When examining landfalling ARs in western North America, we see that MERRA-2 Tier 2 data results in more ARs than in Tier 1 across the whole domain (figure 4, red line). This increase in coastal ARs also translates to higher SON and DJF peaks in the seasonal counts at coastlines, though overall the characteristics of the seasonal counts do not change very much (figure S6(c)).

Comparing MERRA-2 and ERA5 1-hourly data across ARTMIP Tier 2 (figures S7(b) and (c)), we see that the differences are much more pronounced than the differences between MERRA-2 Tier 1 and Tier 2 (figure S7(d)) and in general ERA5 detects fewer ARs than MERRA-2 (figure S7(e)) despite having a higher native spatial resolution. One notable exception is over southern South America, where ERA5 detects up to 20% more ARs than MERRA-2 in that location, consistent with Collow *et al* (2022) where they note that the increase in ERA5 ARs observed over interior South America is related to the greater moisture transport associated with the South American Low Level Jet.

5. Future changes in ARs and regional climate

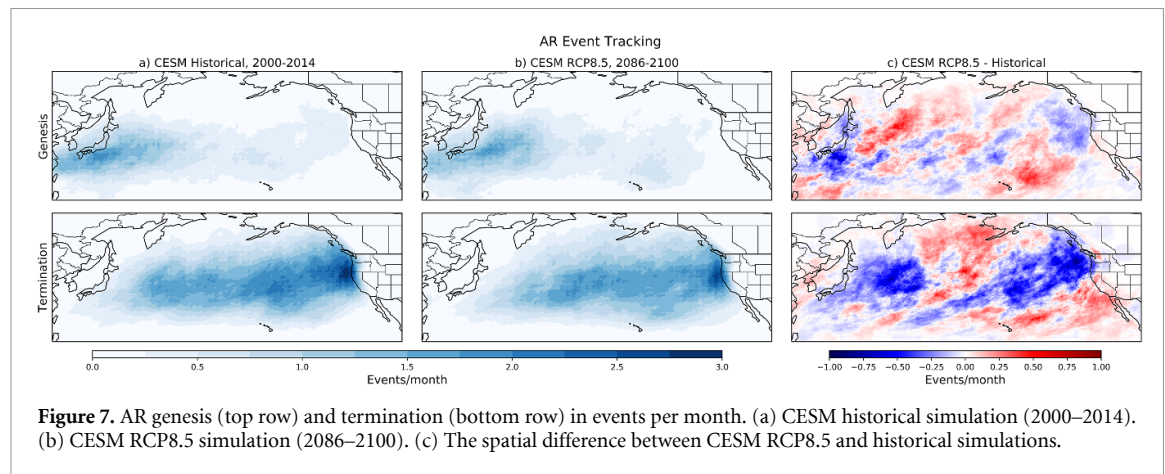
5.1. AR frequency and event changes

To explore changes in ARs with a warming climate, we apply the CGNet detection algorithm to a future climate simulation with CESM using RCP8.5 forcing. The resulting global frequency map is shown in figure 3(d). On an annual mean basis, AR frequency largely decreases in the Northern Hemisphere and increases in the Southern Hemisphere with the exception of the west coast of South America where there are statistically significant decreases in ARs under RCP8.5 (figure 3(g)). Regional and seasonal differences between AR frequencies in the CESM RCP8.5 and historical simulations are shown in figure 5(g). For landfalling ARs, there is a statistically significant decrease over the Pacific Northwest in MAM and an statistically significant increase over central and southern California in DJF and SON. There is also an increase in Alaskan ARs across all seasons, with significant changes in DJF. The regional changes in landfalling ARs are also reflected in the plot of AR counts at coastlines (figure 4, orange line). Relative to the CESM historical simulation, the CESM RCP8.5 simulation lowers the annual AR counts at all latitudes except the southernmost portion of the region (~ 32 – 38°N). The seasonality of AR frequency also changes in a future climate (figure S6(e)), with a potential shift in AR activity to earlier months. The difference plot between RCP8.5 and historical simulations (figure S6(i)) highlights how there are also increases in DJF AR counts at 35°N and decreases in MAM AR counts across the region.

To understand potential mechanisms behind changes in AR frequency, we also analyze changes in the key input fields used to detect ARs. Global plots of annual mean differences in TMQ, U850, and V850 between the CESM RCP8.5 and historical simulations are shown in figure S8. Unsurprisingly, TMQ increases globally as a warmer atmosphere holds more water vapor (figure S8(a)). Because water vapor increases everywhere, it is difficult to discern its impact on specific changes in AR frequency, though our sensitivity tests from section 3.1 show that TMQ on its own is enough for CGNet to reasonably detect ARs. The impact of water vapor changes could also be mitigated by the adjustment we make to the input normalization values for AR detection in the RCP8.5 simulation (see section 3.2). The mean and standard deviation of TMQ are the most significant changes in input normalization values from historical to RCP8.5 climates (table S2). By accounting for the higher mean and standard deviations of integrated water vapor, the resulting changes in detected ARs may be less influenced by changes in TMQ.

Changes in winds are more regionally variable, and while they are still important for the overall detection of ARs in CGNet (figure 1), winds are potentially less impacted by changing the input normalization values for detection in RCP8.5 (table S2). Changes in U850 (figure S8(b)) are similar to AR frequency changes (figure 3(g)) demonstrating a potential connection between changes in zonal winds and detected ARs. Changes in V850 in the North Pacific also mirror changes in AR frequency in that region, though the V850 changes in the Southern Hemisphere are less consistent with AR frequency changes (figure S8(c)). Looking at regional and seasonal changes in 850mb zonal and meridional winds (figure S9) again emphasizes the connection with regional and seasonal AR frequency changes (figure 5(g)).

Using the spatiotemporal event tracking tools provided by Kapp-Schwoerer *et al* (2020), we track connected AR events across the CESM historical and RCP8.5 simulations. For this analysis we increase



the spatial domain of interest beyond western North America to better capture the lifecycle of AR events in the North Pacific (between 15–65°N and 240–105°W). Once events are tracked, we calculate genesis and termination frequencies at each grid point (figure 7). In the CESM historical simulation, AR genesis peaks in the Northwest Pacific off the coast of Japan, with a secondary genesis location north of the Hawaiian Islands (figure 7(a)). These origin locations are consistent with the Kuroshio Extension and Pineapple Express AR classifications identified in previous work (e.g., Wang *et al* 2022). Termination in the CESM historical simulation is spread across most of the Central and Northeastern Pacific with the peak at the west coast of the U.S. These results are also consistent with AR formation colocated with the midlatitude storm track, propagation eastward across the North Pacific, and termination due to land-fall in western North America (Guan and Waliser 2019, Zhou *et al* 2021). Plotting genesis minus termination (figure S10) highlights the regions where genesis is more common (red colors) or termination is more common (blue colors). The area of peak genesis in the western Pacific shifts northeastward in RCP8.5 relative to the historical simulation, with the secondary area of genesis in the central Pacific shifting southeastward (figures 7(b) and (c)). The termination peak shifts southeastward along the U.S. coast to southern California and Baja California, consistent with the overall AR frequency changes across seasons (figure 5(g)).

We further calculate the lifetime and travel distance of the AR events and plot histograms to compare the statistics across the CESM historical and RCP8.5 simulations in figure 8. Most of the AR events are between 12 and 24 h, with the number of events sharply and then gradually dropping off with each increasing 12 h increment as is typical of AR behavior in this region (Zhou *et al* 2018, Guan and Waliser 2019) (figure 8(a)). Most AR events travel between 1500 and 3000 km, with a gradual decrease in the number of events with increasing distance traveled (figure 8(b)). Lifetime and distance traveled will be heavily influenced by genesis location, as ARs typically terminate quickly after reaching land (Zhou and Kim 2019). Speed (calculated as distance over lifetime), also shows a peak in the number of events with low-medium speeds (between 20 and 80 km/h) then a decrease in number of events with increasing speed (figure 8(c)), consistent with other AR tracking results (Zhou *et al* 2021). When comparing changes between the CESM historical and RCP8.5 simulations, the shortest (between 12–24 h) and longest (over 144 h or 6 days) duration events decrease, while all other event lifetimes show an increase in number of events. The number of events in each distance bin mostly increases, though the peak distance bin (1500–3000 km) shows a decrease in the number of AR events. Accordingly, the number of events with low-medium speeds (20–60 km/h) decreases while the number of higher speed events increases. Despite fewer ARs overall in most of this region with climate change (figure 3(g)), the event tracking results indicate these ARs could last longer and travel further at higher speeds. This result is likely related to the shifts in genesis and termination locations (figure 7) that could increase the distance traveled across the Pacific, and thus the overall event lifetime.

5.2. Regional AR precipitation changes

Figure 9 shows seasonal plots of the fraction of total precipitation associated with ARs in the same western North American region shown in figure 5. Most of the AR precipitation is located off the coast in all seasons. Up to 80% of the SON and DJF total precipitation is associated with ARs in this region of the Northeast Pacific. The overall spatial patterns are similar in the CESM historical and RCP8.5 simulations (figures 9(a) and (b)), though there are some notable seasonal changes (figure 9(c)). AR precipitation over the ocean shifts northward in most seasons, especially SON and DJF. All seasons also show some

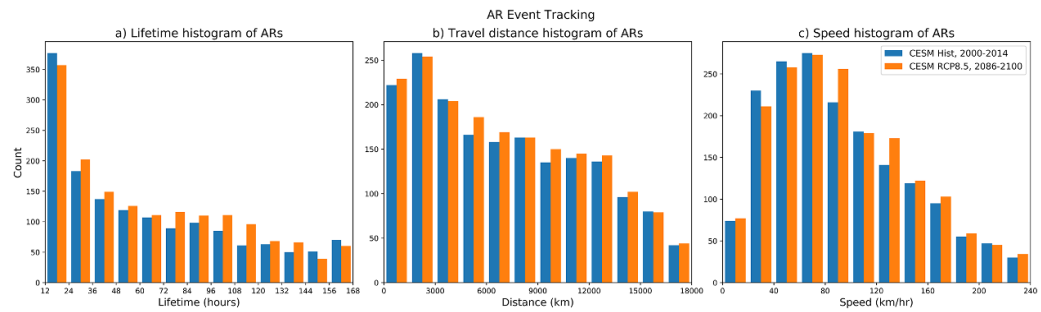


Figure 8. Histograms of AR lifecycle characteristics for CESM historical simulation (2000–2014) in blue and CESM RCP8.5 simulation (2086–2100) in orange. (a) Lifetime duration (hours), (b) travel distance (km), and (c) speed (km/h).

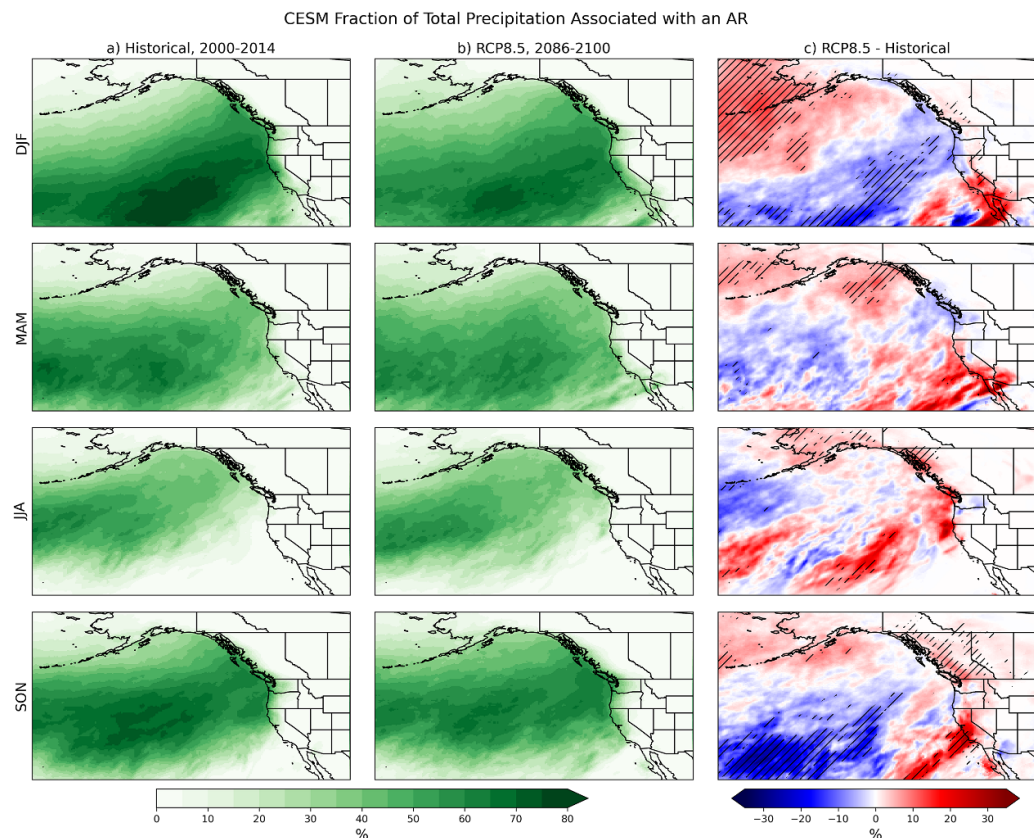
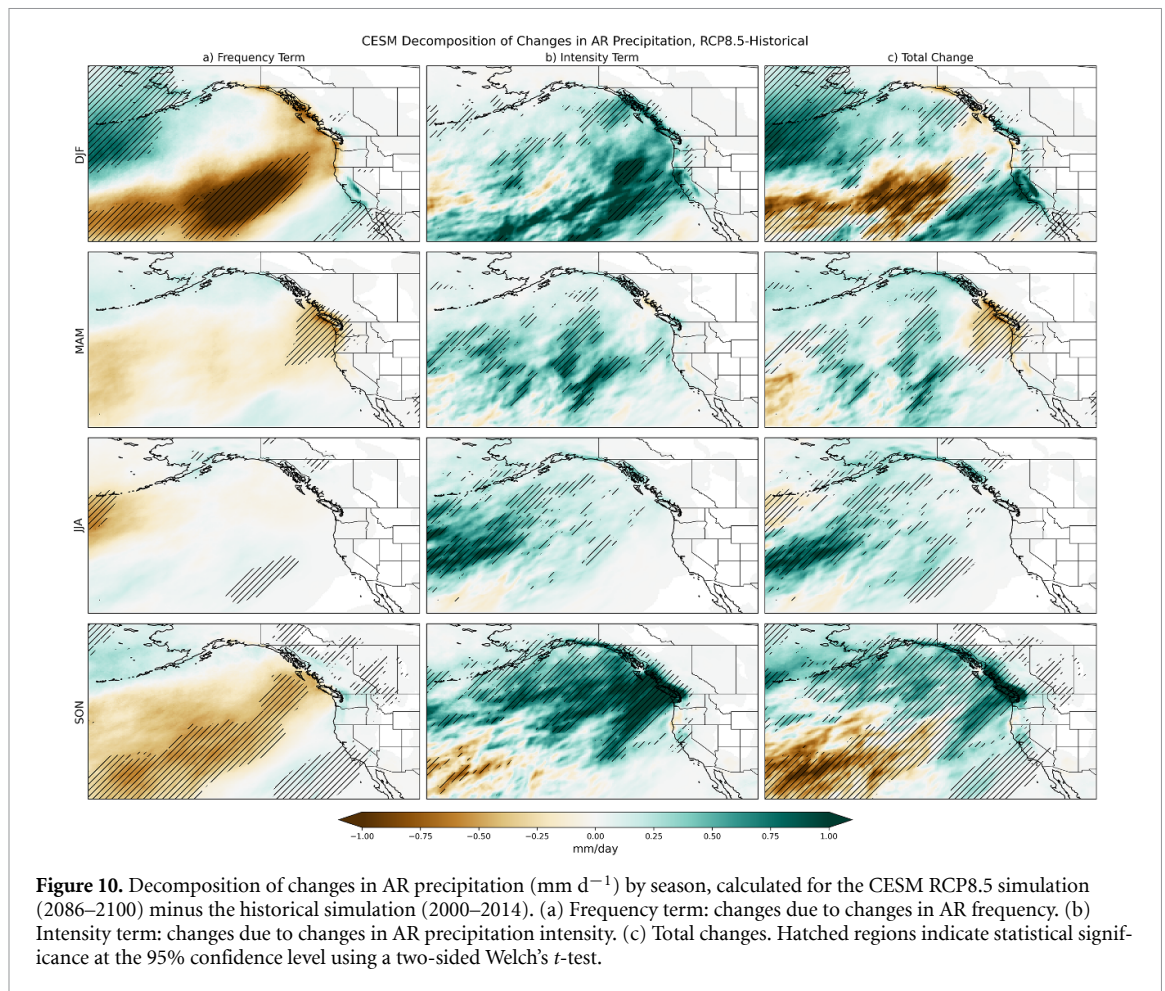


Figure 9. Fraction of total precipitation (%) associated with an AR by season. (a) CESM historical simulation (2000–2014). (b) CESM RCP8.5 simulation (2086–2100). (c) The spatial difference between CESM RCP8.5 and historical simulations. Hatched regions indicate statistical significance at the 95% confidence level using a two-sided Welch's *t*-test.

southeastward shifts in AR precipitation with increases in southern California and Baja California in all seasons except JJA where there is very little change. AR precipitation in the Pacific Northwest and western Canada shows mixed responses across seasons.

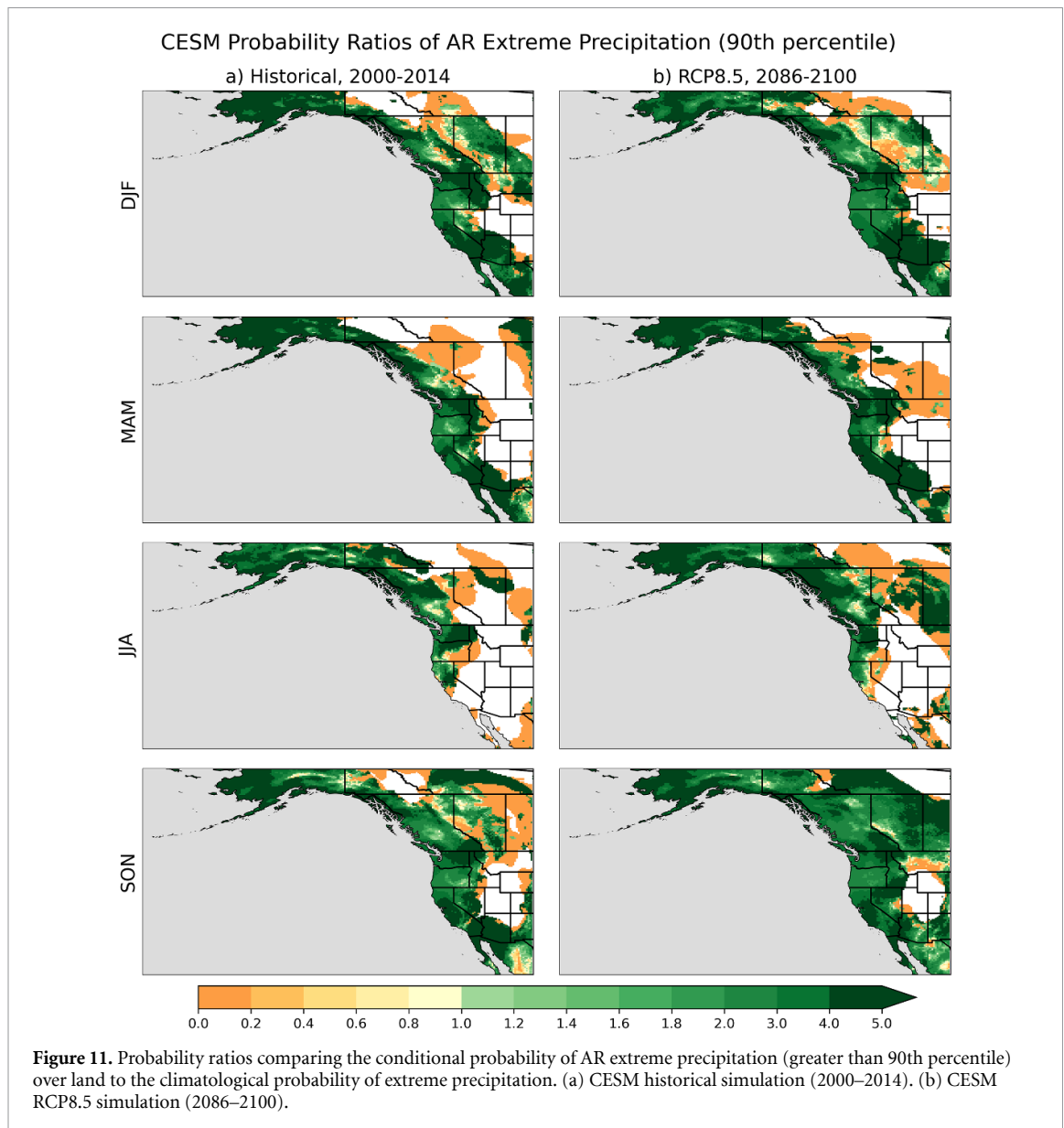
To focus on landfalling AR precipitation, we calculate the zonal mean fraction of total precipitation associated with an AR for the same coastal transect evaluated in section 4.1 (figure S11). Similar to AR frequencies, DJF and SON have the highest fractions of coastal AR precipitation across western North America in the CESM historical simulation (figure S11(a)), confirming prior analysis with satellite observations that linked landfalling AR precipitation in winter to stronger synoptic storms than summer (Neiman *et al* 2008). The peak in DJF AR precipitation is around 38–40° N, where more than 60% of total precipitation is associated with ARs. AR precipitation in SON is also around 60% of the total slightly north of the DJF peak, between 43–47° N. AR precipitation in MAM is between 10%–35% of total precipitation across the region. JJA has very little AR precipitation at low latitudes and around 10%–20% of the total in higher latitudes of the region. These fractions are consistent with prior work using historical data to estimate the contribution of ARs to precipitation in this region (e.g.,



Dettinger *et al* 2011, Lamjiri *et al* 2017, Arabzadeh *et al* 2020, Collow *et al* 2022) Comparing the historical simulation to the RCP8.5 simulation (figure S11(b)), we find that all seasons show increases and decreases in the fraction of precipitation associated with an AR along the coast (figure S11(c)). SON, DJF, and MAM show increases at lower latitudes with SON in particular increasing up to 30%. The increases in JJA AR precipitation are between 40–45° N and around 50° N. SON and DJF show decreases between 40–45° N and MAM AR precipitation decreases at higher latitudes. These results mirror the changes in coastal AR frequency in this region (figure S6(i)), though some seasons show stronger AR precipitation increases not necessarily reflected in the AR frequency changes (e.g. MAM in southern California and JJA in the Pacific Northwest). While these plots focus on landfalling AR precipitation, we note that changes in inland AR precipitation seen in figure 9(c) are also important for potential impacts.

The contributions of frequency and intensity to overall changes in AR precipitation are shown in figure 10. The frequency term (figure 10(a)) shows mostly decreases across seasons, mirroring the results of figure 5(g) while highlighting the importance of DJF to the frequency changes. Conversely, the intensity term (figure 10(b)) shows mostly increases across seasons, demonstrating that AR precipitation intensity is projected to increase in this region with climate change. Comparing the individual frequency and intensity terms with the total change (figure 10(c)), we see that the intensity term is contributing most of the increases in total AR precipitation changes across the region. However in certain seasons the increases in the intensity term are somewhat canceled out by decreases in the frequency term (e.g. MAM and SON in the Pacific Northwest). There are also regions where both terms contribute positively to the total increase in AR precipitation (e.g. central California in DJF).

Figure 11 shows the probability ratios (PR) of extreme precipitation (greater than 90th percentile) for the CESM historical and RCP8.5 simulations. In both simulations and across all seasons there are inland areas without data due to a lack of any detected ARs in those locations. In the historical simulation during all seasons except JJA we find that the entire coastline of the region is an area where extreme precipitation is more likely to be associated with an AR ($\text{PR} > 1$, green colors; figure 11(a)). In JJA this is true to a lesser extent, except in the lower latitudes including southern California where



there are not any detected ARs to calculate this metric. All seasons also show inland areas where extreme precipitation is less likely to be associated with an AR ($PR < 1$, orange colors). These areas tend to be at the edges of mountainous regions where the extreme precipitation is likely due to other sources like frontal systems and extratropical cyclones (Dagon *et al* 2022a). Examining the spatial patterns of the RCP8.5 simulation, we see an increase in areas more likely to have extreme precipitation associated with an AR across all seasons. In DJF and MAM, this expansion is focused in the U.S. southwest, whereas in JJA and SON the increase is seen in western and central Canada (figure 11(b)). While ARs overall are not as frequent in these inland regions (figure 5), our results show that even small changes in the locations of detected ARs can impact the likelihood of extreme precipitation originating from an AR.

We also calculate the probability ratios based on defining extreme precipitation as greater than 95th and 99th percentiles, and find that our results generally hold. The expansion of areas more likely to have AR-associated extreme precipitation under the RCP8.5 simulation, such as the U.S. southwest in DJF, is consistent using the 95th percentile definition (figure S12). This result also holds true using the 99th percentile definition (figure S13), though there is an overall increase in areas where ARs are less likely to be associated with extreme precipitation due to the more extreme definition. Notably, the increase in AR-associated extreme precipitation in western and central Canada during SON is greater when using the 99th percentile threshold, relative to the other definitions of extreme precipitation.

6. Discussion

In this study, we use a DL model (CGNet), trained on crowd-sourced human labels of extreme weather, to investigate ARs and their associated impacts across Earth system model simulations and reanalysis datasets. We gain a better understanding of the algorithm's sensitivities by varying input fields (table S1, figure 1) and input normalization values (table S2, figure 2). We evaluate the performance of CGNet on a CESM historical simulation by comparing its performance with MERRA-2 and ERA5 reanalysis datasets globally (figure 3) and regionally (figures 4 and 5). We find that AR detection varies widely across datasets, driven by factors such as the climatology of input variables (figures S2–S5) and the spatial and temporal resolution of input datasets (figures 6 and S7). ARs detected in a future climate simulation show regionally and seasonally varying increases and decreases in frequency (figures 3–5 and S6), associated with changes in circulation (figures S8 and S9) and shifts in AR genesis and termination locations (figures 7 and S10). These changes alter the overall characteristics of ARs including lifetime and distance traveled (figure 8), which further map onto regional impacts such as associated precipitation (figures 9–11 and S11).

6.1. CGNet sensitivities

Our tests with different input fields for CGNet show that integrated water vapor is a key variable for AR detection, and using that variable alone results in comparable performance for AR detection to the original model trained on 4 input variables including upper level winds and PSL (table S1, figure 1). Including wind variables may also add a small amount of skill, though PSL does not appear to be particularly important to AR detection, supported by results from Galea *et al* (2024), Radke *et al* (2025). The importance of water vapor and winds in AR detection also reflects the threshold-based algorithms from ARTMIP where integrated water vapor and IVT are almost universally used while PSL is not typically used in algorithms (Shields *et al* 2018). Reducing the number of input fields could decrease the computational expense of training the algorithm (Tian *et al* 2023), thus it is a useful exercise to consider carefully what input variables are most relevant to the detection. Future work could consider further modifications to the input variables of CGNet, depending on data availability.

Our tests varying input normalization values show small but zonally consistent and statistically significant differences in AR detection using input variable means and standard deviations calculated from the fully coupled CESM historical simulation (continuous time points from 2000–2014) versus values calculated from the CAM5 training data (398 time points from 1996–2010) (figure 2(c)). There are consistent decreases in detected ARs in the Northern Hemisphere and the Southern Ocean and increases in detected ARs across the southern midlatitudes. It is possible that this is related to climatological differences between the models run in fully coupled versus AMIP mode, and we do see that the CESM historical simulation has lower (higher) integrated water vapor in the Northern (Southern) Hemisphere than the CAM5 training data (not shown), loosely matching the pattern of AR frequency changes in figure 2(c). The two datasets also have differences in the specific time period and sampling frequency from which the normalization values are calculated, which could have small impacts on the results.

Changing normalization values has a larger impact on AR detection in the future climate simulation. Using input variable means and standard deviations calculated from the CESM RCP8.5 simulation (continuous time points from 2086–2100) versus values calculated from the CAM5 training data results in statistically significant AR frequency increases of greater than 15% in parts of the Southern Hemisphere (figure 2(f)). Northern Hemisphere changes in AR frequency are more muted, and include a mix of statistically significant increases and decreases. Some of these changes could be explained by the differences in model structure and sampling frequency as noted above, though there is likely also an impact of shifting the input variables' baseline states (table S2). In particular, increases in water vapor with warming will have different impacts on AR detection methods depending on how they are defined (Zhao 2020, Quagrainne *et al* 2025). In this study we choose to update the normalization values using the future climate simulation to improve the ability of the ML model to generalize (Dagon *et al* 2022a, Beucler *et al* 2024). However, we have shown that changing the values does have an impact on the resulting detection. Going forward, it is likely that the choice of normalization values will depend on the science question and specific application (Rasp *et al* 2018, Molina *et al* 2021, Radke *et al* 2025), and care should be taken to understand the potential differences that could be introduced when changing or updating values used for data standardization.

While the ClimateNet project represented a unique contribution to the field of AR detection through the development of an expert-labeled dataset, the ability to train high quality machine and DL models is limited by the number of labeled images. The labeling campaigns collected a total of 459 expert-labeled images using a 20 year climate model simulation as source data (Prabhat *et al* 2021). While these global images each include multiple labeled features, the overall number of independent synoptic conditions that were sampled from remains small, representing a limitation of using CGNet in this work. The use of professional annotators, though potentially costly, represents a powerful alternative to generating large numbers of labels that can then be used to train or refine supervised learning models (Kim *et al* 2025). Class imbalance is a perennial issue when a large percentage of the input pixels are classified as background, and future work around data enhancement could help address this issue (Tian *et al* 2023). Unsupervised learning, for example using local causal states, may also provide a viable approach to AR identification that is robust across changes in spatial resolution (Rupe *et al* 2025). Furthermore, investigating alternate ML and DL architectures (e.g. vision transformers, backbone architectures, and regional modeling) is a worthwhile pursuit (Tian *et al* 2023, Galea and Ma 2025, Radke *et al* 2025).

6.2. AR detection across datasets and resolutions

Evaluating CGNet in the context of ARTMIP shows that while CGNet captures similar spatial patterns in global AR frequency in the CESM historical simulation (figure 3(a)), the magnitude of detection is greater than the ARTMIP Tier 1 ensemble mean across 10 algorithms (Lora *et al* 2020). The difference in AR frequency is likely due to the significant spread in global AR detection methods which depresses the mean across an ensemble of detection methods, versus the frequencies from an individual method. CGNet is also known to be a ‘permissive’ algorithm, as it is more liberal with its detection and classification of ARs (Collow *et al* 2022, Higgins *et al* 2023), which is also confirmed in our comparison of CGNet landfalling ARs to the median across ARTMIP methods (figure 4). Furthermore, an algorithm trained on human labels can show a higher frequency compared to traditional methods (Shields *et al* 2018, Kapp-Schwoerer *et al* 2020, Kim *et al* 2025, Crespo-Otero *et al* 2026).

Another key difference between our results in figure 3(a) and the ARTMIP findings is that we compare AR detection with CESM and multiple reanalysis datasets from MERRA-2 and ERA5. The resulting global AR frequencies are quite different across the Earth system model and reanalysis products (figures 3(a)–(c)). While MERRA-2 AR frequencies are generally higher than the CESM historical simulation, ERA5 AR frequencies are generally lower, especially in the Northern Hemisphere (figures 3(e) and (f)). Baek and Lora (2021) found a similar pattern of differences between MERRA-2 and a CESM1 historical simulation, though they quantified ARs using IVT thresholds. Tian *et al* (2023) found that CESM2 historical AR frequencies were higher than ERA5 AR frequencies, though they used model output from a simulation run at a lower native spatial resolution. The higher AR frequency that we find off the coast of western North America in MERRA-2 relative to the CESM historical simulation (figure 5(e)) is also evident in Shields *et al* (2023) where they found a greater frequency of spring and summer ARs in MERRA-2 relative to a CAM5 historical simulation. Our finding that the CESM historical simulation has more landfalling ARs over this region than the reanalysis products (figure 4) echoes Rhoades *et al* (2020) where they found fewer landfalling ARs across the west coast of the U.S. in ERA5 compared to a variable resolution configuration of CESM2. Consistent with our results on seasonal landfalling ARs (figure S6), Kamnani *et al* (2025) found that DJF was the peak AR season for California across multiple ARTMIP Tier 1 detection methods, including ClimateNet. They also found that SON was the peak season for more northerly latitudes in this region, reflecting what we see in both MERRA-2 and the CESM historical simulation (figures S6(a)–(c)).

Consistent differences between ARs in the reanalysis products and CESM, for example the pattern of increases and decreases in AR frequency in the Southern Hemisphere (figures 3(e) and (f)), are likely due to differences in water vapor and winds (figures S2(a)–(f)). However, given that the underlying input fields from the reanalysis products are in general more similar to each other than to CESM (figures S2(g)–(i)), differences in these quantities are unlikely to fully explain the variation in resulting AR detection. Furthermore, our results show a clear impact of spatial resolution on AR detection, and we cannot adequately detect ARs with MERRA-2 fields unless the input data is regridded to the same resolution that CGNet was trained on (figure 6). This result is echoed in Swenson *et al* (2018), Reid *et al* (2020), where both studies found that AR detection increased as the spatial resolution of the input data increased. However the opposite result was found in Hagos *et al* (2015) where AR frequency in aquaplanet simulations decreased with increasing spatial resolution. Furthermore, Payne and Magnusdottir (2015) found that while the bias between AR detection in Earth system models and reanalysis products did decrease when resolution increased, the change in bias was not linear with changing resolution.

Underlying the comparison of our results with other studies is the definition of ARs themselves. Our results show that even with a threshold-free method, the source and horizontal resolution of the input data have large impacts on the resulting AR detection. Future work with ML detection algorithms could investigate alternative loss functions that avoid pixel-wise comparisons, along with semi-supervised methods such as image style transfer, to help generalize across input datasets and resolutions (Mahesh *et al* 2024).

6.3. Changes in AR frequency and lifecycle

Our global results show a mix of increases and decreases in ARs with climate change (figure 3(g)), consistent with previous studies using the Community Climate System Model version 4 (CCSM4, a predecessor to CESM) that showed ARs tracking with changes in the low level jet (Shields and Kiehl 2016a, 2016b). Consistent with our results, Shields *et al* (2023), Tian *et al* (2023) both found global increases in AR frequency in the future relative to historical simulations, despite different AR detection methods, Earth system models, and/or spatial resolution. Similarly, Tian *et al* (2023) also found a decrease in ARs off the west coast of southern South America. Nevertheless, while our results capture higher spatial scales and coupled atmosphere-ocean interactions known to be important for ARs (Liu *et al* 2021, 2022, Shields *et al* 2024), they are necessarily contingent on the sensitivities of the CGNet algorithm itself. Focusing on extreme ARs (Higgins *et al* 2025) or on AR width (Baek *et al* 2025), or weighting Earth system models by skill and independence (Radić *et al* 2015, Massoud *et al* 2019) could help isolate the climate change signal and avoid uncertainties in AR detection and future projections.

Our regional analysis shows increases in western North America coastal ARs over more southerly latitudes like California, and decreases further north in the Pacific Northwest and western Canada (figure 4), with notable changes across all seasons except JJA (figures 5(g) and S6(i)). The changes in DJF could be related to the poleward movement of the jet and expansion of the subtropical jet providing extra moisture to the southern California region (Shields and Kiehl 2016b, Baek and Lora 2021, Huang and Swain 2022). Consistent with our California results, Shields *et al* (2023) found that less restrictive AR detection methods (the category which CGNet fits) detected more ARs over this region under RCP8.5 relative to a CAM5 historical simulation. Higgins *et al* (2025) also found an increase in U.S. west coast extreme ARs under different warming scenarios with CGNet. Regarding the decrease in Pacific Northwest and western Canada ARs found in our results, some insight can be gained by comparing changes in input variables used to detect ARs with changes in AR frequency. This comparison indicates that while water vapor is universally increasing, zonal wind changes are likely driving the pattern of changes in AR frequency (figures S8 and S9). This interplay between moisture and wind contributions to regional AR changes was discussed in Payne and Magnusdottir (2015), where they also found consistent increases in North Pacific atmospheric moisture with decreases in 850 hPa zonal winds in CMIP5 models under RCP8.5, indicating that dynamical responses in conjunction with thermodynamic changes play an important role in AR projections. Furthermore, Gao *et al* (2015) found a seasonal dependence on the dynamical contribution to changes in western North America ARs (also in CMIP5 models under RCP8.5), noting that in certain regions and seasons the dynamical component was a significant contribution to total AR changes. However, uncertainties in representing dynamical responses in Earth system models limit the ability to project AR changes (Payne *et al* 2020), highlighting an opportunity for improvement in model process representation.

Using event tracking tools, we find that the areas of peak AR genesis shift northeastward in the western Pacific and southeastward in the central Pacific (figure 7). These two origin locations align with categorizations from Zhou and Kim (2019), Wang *et al* (2022). Both studies found that western Pacific originating (or Kuroshio Extension) ARs tended to make landfall in the Pacific Northwest, whereas eastern Pacific originating (or Pineapple Express) ARs made landfall at locations spread across western North America. Our results show a decrease in ARs in the Pacific Northwest, consistent with changes to western Pacific originating ARs where the origin location shifts further north. This higher latitude genesis also has implications for ARs forming in regions with less relative moisture than before, where previously AR formation in this area was found to be associated with water vapor from TCs that is transported poleward towards the North Pacific jet (Cordeira *et al* 2013). The increase in ARs in southern California and off the west coast of Mexico that we observe is consistent with a southward shift in the origin location of eastern Pacific originating ARs, possibly driven by enhanced westerly winds (Gao *et al* 2015). These results also imply that AR tracks over the North Pacific may become more zonal, instead of curving poleward as previously observed (Zhou *et al* 2018). We also find a general increase in AR lifetimes, travel distances, and speeds with the exception of the shortest and longest-lived events which decrease in frequency (figure 8). These changes in AR lifecycle characteristics are consistent with a stronger westerly jet and deepened Aleutian low (Kim and

Chiang 2022), also reflected in the modeled changes in zonal winds (figure S8(b)). While potential future increases in AR length and width have been shown (e.g., Espinoza *et al* 2018), changes in size characteristics are strongly dependent on the definition of ARs in a warming climate, and in particular whether or not historical thresholds are applied to future data (Zhao 2020). More broadly, large-scale circulation drivers are known to modulate AR changes in the North Pacific. Long-term changes in North Pacific ARs have been linked to the intensity of the mid-latitude westerly jet and the strength of low-level monsoon flows (Liang *et al* 2023, Liang and Haywood 2023). Park *et al* (2021) demonstrated the relationship between ARs and upstream quasi-stationary moisture source regions such as the East Asian summer monsoon and migrating extratropical cyclones. As with other results summarized here, the choice of tracking method will heavily influence resulting AR lifecycle characteristics (Guan and Waliser 2019, Zhou *et al* 2021). Comparing projected AR event changes using other tracking methods, Earth system models, and warming scenarios should be a priority for future work.

6.4. Changes in AR precipitation and associated mechanisms

We find that AR-associated precipitation increases in California across most seasons (figures 9 and S11) are driven by changes in AR precipitation intensity (figure 10). A shift to higher AR rainfall rates in this region under climate change scenarios is consistent with previous studies (e.g., Shields and Kiehl 2016b), Gershunov *et al* 2019), Prabhat *et al* 2021), Shields *et al* 2023). In particular Huang *et al* (2020a) found increases in AR total precipitation and AR precipitation intensity over California using high resolution modeling under RCP8.5, driven primarily by thermodynamic changes in water vapor with some smaller positive contributions from zonal wind increases. Dynamical changes such as an expansion and strengthening of the Aleutian low have been linked to increases in AR intensity in this region (Polade *et al* 2017). The presence of orography can act to further enhance AR precipitation (Payne *et al* 2020), pointing to a need for high resolution modeling to capture the topographic detail that is important to simulating precipitation associated with ARs (Huang *et al* 202b).

AR precipitation changes across seasons are more inconsistent in the Pacific Northwest and western Canada, and have a greater contribution from changes in AR frequency. Radić *et al* (2015) also found important contributions from AR frequency changes to changes in AR precipitation over British Columbia in CMIP5 models under multiple warming scenarios. Negative frequency contributions in combination with positive intensity contributions were also found in Konstali *et al* (2024) where they linked future precipitation changes in CESM2 under SSP3-7.0 to multiple weather features, including ARs. Gershunov *et al* (2019) noted that the wetter and less variable background state in the Pacific Northwest results in less extreme changes in AR precipitation relative to drier regions like California.

The magnitude of AR precipitation increases in this region are aligned with what was found in Tian *et al* (2023) using ensembles of DL methods and Earth system models to detect ARs, though we also see localized decreases in AR precipitation when the changes are broken down by season. The ‘permissive’ nature of the CGNet algorithm, as discussed above, could be playing a role in some of the differences we see relative to previous studies. Quagrainie *et al* (2025) found that permissive AR detection methods were associated with less AR precipitation relative to restrictive methods, due to permissive methods capturing a broader spectrum of events including lightly precipitating ARs.

Even within the same region, AR precipitation can be driven by different thermodynamic responses to climate change that may be linked to the orientation of the event relative to topography (Hecht and Cordeira 2017, Hu *et al* 2017, Collow *et al* 2022, Michaelis *et al* 2022). ARs that are more wet versus more windy can produce different precipitation responses that also vary with orography (Gonzales *et al* 2020). Categorizing ARs by severity could further illuminate the changes in AR precipitation and how they vary with region and season (Ralph *et al* 2019a, Rhoades *et al* 2020). Including time lags in the precipitation response to ARs may also provide a more holistic picture of the localized impacts and how they vary across the region (Vallejo-Bernal *et al* 2023, Crespo-Otero *et al* 2026). Where and when AR precipitation will cause the most impact in this region can be traced back to the lifecycle of AR events and where they originate (Zhou and Kim 2019), providing important context for forecasting ARs on subseasonal timescales (Baggett *et al* 2017, Mundhenk *et al* 2018).

Our results show several inland regions across western North America have potential for increases in AR-associated extreme precipitation (figure 11), consistent with Gao *et al* (2015), Hagos *et al* (2016), and with important implications for flooding (Konrad and Dettinger 2017, Curry *et al* 2019), rain-on-snow events (Guan *et al* 2016, Chen *et al* 2019), and water resource management (Rhoades *et al* 2018, Payne *et al* 2020). The potential increase of ARs in more inland regions has been linked to interior penetration

of moisture with warming (Mahoney *et al* 2018) and a weakened rain-shadow effect (Huang *et al* 2020a). Using the more complex encoder-decoder architecture trained from ClimateNet expert labels, Prabhat *et al* (2021) showed an increase in AR extreme precipitation in California that exceeded the C–C scaling. The fact that changes in AR extreme precipitation can be explained by C–C scaling (and even super-C–C scaling) confirms previous work projecting AR changes in this region (e.g. Gao *et al* 2015, Warner *et al* 2015, Mahoney *et al* 2018, Ma *et al* 2020).

7. Conclusion

In summary, our study explores how ML-based detection of ARs varies across reanalysis datasets and Earth system model output, and shows how these variations can be traced back to algorithmic sensitivities and properties of the input data. Specifically, we find that water vapor is a key variable for AR detection, and that the choice of normalization values (i.e. variable means and standard deviations) will impact AR detection frequency, especially in a future climate scenario. We also find that the spatial resolution of the input data matters to the detection and that regridding input data to align with what the trained model expects is an essential pre-processing step. These results have important implications for future work in ML for detection of features like ARs. Even within a single ML-based framework, there are still large uncertainties in detection. Furthermore, using a threshold-free approach like ML to define features in a future climate needs careful consideration of data pre-processing aspects such as the input normalization values and spatial resolution, based on the science question of interest. Lastly it is important to note that CGNet was trained entirely on AMIP protocol simulations with prescribed sea surface temperatures. Applying this version of CGNet to fully prognostic, coupled CESM simulations may introduce uncertainties as the dynamic ocean/ice/atmosphere feedbacks could alter the presentation of ARs in a way that CGNet was not trained to detect. Future work should prioritize testing the use of coupled model data for training ML algorithms. With those considerations in mind, CGNet provides a convenient and computationally efficient method of detecting ARs across Earth system datasets that also has built-in spatiotemporal tracking, allowing for analysis of AR lifecycle characteristics which has previously been understudied in the climate change context.

We leverage this automated detection and tracking to investigate potential changes in AR frequency and event characteristics, with particular attention to the North Pacific and western North America. In more southerly latitudes like California, AR frequency increases are linked to circulation changes and a southward shift in the origin location of eastern Pacific ARs. AR precipitation also increases in this region due to increases in AR precipitation intensity. In contrast, AR frequency decreases in the Pacific Northwest and western Canada with an important dynamical contribution to these changes that is also connected to a northward shift in the origin location of western Pacific ARs. The decrease in AR frequency in this region also contributes to the seasonally mixed responses of AR precipitation. Increases in extreme precipitation from ARs are not limited to coastal regions, demonstrating the importance of understanding the potential impacts across a larger domain.

Acknowledgments

Portions of this study were supported by the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program of the U.S. Department of Energy's (DOE) Office of Biological & Environmental Research (BER) under Lawrence Livermore National Lab subaward DE-AC52-07NA27344, Lawrence Berkeley National Lab subaward DE-AC02-05CH11231, and Pacific Northwest National Lab subaward DE-AC05-76RL01830. This work was also supported by the National Science Foundation (NSF) National Center for Atmospheric Research (NCAR), which is a major facility sponsored by NSF under Cooperative Agreement No. 1852977. We would like to acknowledge computing support from the Casper system (<https://ncar.pub/casper>) provided by NSF NCAR. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This research was also part of the Blue Waters sustained-petascale computing project, supported by NSF (Awards OCI-0725070 and ACI-1238993) and the state of Illinois. This work was also part of the 'High Resolution Earth System Modeling Using Blue Waters Capabilities' PRAC allocation support by NSF (Award ACI-1516624), and used resources of the National Energy Research Scientific Computing Center (NERSC),

a U.S. DOE Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. We thank Nan Rosenbloom, Gary Strand, and Sasha Glanville for assistance with managing CESM data. We thank the editor and two anonymous reviewers for feedback and suggestions that improved the paper.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.5065/4g3t-bn97> Dagon *et al* (2026a).

Supporting Information available at: <https://doi.org/10.1088/3049-4753/aeaac7/data1>.

ORCID iDs

Katherine Dagon  0000-0002-4518-8225
Teagan King  0000-0001-8031-8606
Christine A Shields  0000-0003-4481-1377
John Truesdale  0009-0007-0292-0205
Sol Kim  0000-0002-2641-8929
Kirsten J Mayer  0000-0003-3582-8337
Andre Graubner  0000-0002-3345-2690
Annette Greiner  0000-0001-6465-7456
Lukas Kapp-Schwoerer  0009-0000-7685-5548
Karthik Kashinath  0000-0002-9311-5215

References

- American Meteorological Society 2026 Atmospheric river (available at: <https://glossary.ametsoc.org/wiki/atmospheric-river/>)
- Arabzadeh A, Ehsani M R, Guan B, Heflin S and Behrangi A 2020 Global intercomparison of atmospheric rivers precipitation in remote sensing and reanalysis products *J. Geophys. Res.: Atmos.* **125** e2020JD033021
- Baek S H and Lora J M 2021 Counterbalancing influences of aerosols and greenhouse gases on atmospheric rivers *Nat. Clim. Change* **11** 958–65
- Baek S H, McClenny E E, Ullrich P A, Quagrainie K T and Collins W D 2025 Response of atmospheric river width and intensity to aquaplanet warming: a detection algorithm- and background moisture-independent approach *J. Geophys. Res.: Atmos.* **130** e2025JD043367
- Baggett C F, Barnes E A, Maloney E D and Mundhenk B D 2017 Advancing atmospheric river forecasts into subseasonal-to-seasonal time scales *Geophys. Res. Lett.* **44** 7528–36
- Beucler T *et al* 2024 Climate-invariant machine learning *Sci. Adv.* **10** eadj7250
- Branstetter M L and Famiglietti J S 1999 Testing the sensitivity of GCM-simulated runoff to climate model resolution using a parallel river transport algorithm *14th Conf. on Hydrology* (American Meteorological Society) pp 391–2 (available at: <https://ams.confex.com/ams/older/99annual/abstracts/851.htm>)
- Chen L-C, Zhu Y, Papandreou G, Schroff F and Adam H 2018 Encoder-decoder with atrous separable convolution for semantic image segmentation (arXiv:1802.02611)
- Chen X, Leung L R, Wigmosta M and Richmond M 2019 Impact of atmospheric rivers on surface hydrological processes in western U.S. watersheds *J. Geophys. Res.: Atmos.* **124** 8896–916
- Collow A B M *et al* 2022 An overview of ARTMIP's tier 2 reanalysis intercomparison: uncertainty in the detection of atmospheric rivers and their associated precipitation *J. Geophys. Res.: Atmos.* **127** e2021JD036155
- Cordeira J M, Ralph F M and Moore B J 2013 The development and evolution of two atmospheric rivers in proximity to western north pacific tropical cyclones in October 2010 *Mon. Weather Rev.* **141** 4234–55
- Crespo-Otero A, Insua-Costa D, Deman V M H, Hernández-García E, López C and Míguez-Macho G 2026 Global asymmetries in the moisture origins of atmospheric river precipitation *npj Clim. Atmos. Sci.* **9** 137
- Curry C L, Islam S U, Zwiers F W and Déry S J 2019 Atmospheric rivers increase future flood risk in Western Canada's largest pacific river *Geophys. Res. Lett.* **46** 1651–61
- Dagon K, King T and Truesdale J 2026b katedagon/ML-extremes: CGNet ARs manuscript revisions code release (<https://doi.org/10.5281/zenodo.7126838>)
- Dagon K, King T, Truesdale J, Rosenbloom N A and Bates S 2026a Machine learning-based detection of atmospheric rivers in CESM1.3 UCAR/NCAR-GDEX (available at: <https://doi.org/10.5065/4g3t-bn97>)
- Dagon K, Truesdale J, Biard J C, Kunkel K E, Meehl G A and Molina M J 2022a Machine learning-based detection of weather fronts and associated extreme precipitation in historical and future climates *J. Geophys. Res.: Atmos.* **127** e2022JD037038
- Dagon K, Truesdale J, Rosenbloom N and Bates S 2022b Machine learning-based detection of weather fronts and associated extreme precipitation in CESM1.3 UCAR/NCAR GDEX (<https://doi.org/10.5065/q6t7-ta06>)
- Danabasoglu G, Bates S C, Briegleb B P, Jayne S R, Jochum M, Large W G, Peacock S and Yeager S G 2012 The CCSM4 ocean component *J. Clim.* **25** 1361–89
- Dennis J M *et al* 2012 CAM-SE: a scalable spectral element dynamical core for the community atmosphere model *Int. J. High Perform. Comput. Appl.* **26** 74–89

- Dettinger M D, Ralph F M, Das T, Neiman P J and Cayan D R 2011 Atmospheric rivers, floods and the water resources of California *Water* **3** 445–78
- Espinoza V, Waliser D E, Guan B, Lavers D A and Ralph F M 2018 Global analysis of climate change projection effects on atmospheric rivers *Geophys. Res. Lett.* **45** 4299–308
- Feng Z et al 2025 Mesoscale convective systems tracking method intercomparison (MCSMIP): application to DYAMOND global km-scale simulations *J. Geophys. Res.: Atmos.* **130** e2024JD042204
- Fischer E M and Knutti R 2015 Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes *Nat. Clim. Change* **5** 560–4
- Galea D and Ma H-Y 2025 Intercomparison of deep learning model architectures for atmospheric river prediction *Artif. Intell. Earth Syst.* **4** 240057
- Galea D, Ma H-Y, Wu W-Y and Kobayashi D 2024 Deep learning image segmentation for atmospheric rivers *Artif. Intell. Earth Syst.* **3** 230048
- Gao Y, Lu J, Leung L R, Yang Q, Hagos S and Qian Y 2015 Dynamical and thermodynamical modulations on future changes of landfalling atmospheric rivers over western North America *Geophys. Res. Lett.* **42** 7179–86
- Gelaro R et al 2017 The modern-era retrospective analysis for research and applications, version 2 (MERRA-2) *J. Clim.* **30** 5419–54
- Géron A 2019 *Hands-On Machine Learning With Scikit-Learn, Keras and Tensorflow* 2nd edn (O'Reilly Media)
- Gershunov A et al 2019 Precipitation regime change in Western North America: the role of atmospheric rivers *Sci. Rep.* **9** 9944
- Gonzales K R, Swain D L, Barnes E A and Diffenbaugh N S 2020 Moisture- versus wind-dominated flavors of atmospheric rivers *Geophys. Res. Lett.* **47** e2020GL090042
- Guan B, Molotch N P, Waliser D E, Fetzer E J and Neiman P J 2010 Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements *Geophys. Res. Lett.* **37** L20401
- Guan B and Waliser D E 2019 Tracking atmospheric rivers globally: spatial distributions and temporal evolution of life cycle characteristics *J. Geophys. Res.: Atmos.* **124** 12523–52
- Guan B, Waliser D E, Ralph F M, Fetzer E J and Neiman P J 2016 Hydrometeorological characteristics of rain-on-snow events associated with atmospheric rivers *Geophys. Res. Lett.* **43** 2964–73
- Hagos S M, Leung L R, Yoon J-H, Lu J and Gao Y 2016 A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the large ensemble CESM simulations *Geophys. Res. Lett.* **43** 1357–63
- Hagos S, Leung L R, Yang Q, Zhao C and Lu J 2015 Resolution and dynamical core dependence of atmospheric river frequency in global model simulations *J. Clim.* **28** 2764–76
- Hecht C W and Cordeira J M 2017 Characterizing the influence of atmospheric river orientation and intensity on precipitation distributions over North Coastal California *Geophys. Res. Lett.* **44** 9048–58
- Held I M and Soden B J 2006 Robust responses of the hydrological cycle to global warming *J. Clim.* **19** 5686–99
- Henny L and Kim K-M 2025 The changing nature of atmospheric rivers *J. Clim.* **38** 1435–56
- Hersbach H et al 2020 The ERA5 global reanalysis *Q. J. R. Meteorol. Soc.* **146** 1999–2049
- Higgins T B, Subramanian A C, Graubner A, Kapp-Schwoerer L, Watson P A G, Sparrow S, Kashinath K, Kim S, Monache L D and Chapman W 2023 Using deep learning for an analysis of atmospheric rivers in a high-resolution large ensemble climate data set *J. Adv. Model. Earth Syst.* **15** e2022MS003495
- Higgins T B, Subramanian A C, Watson P A G and Sparrow S 2025 Changes to atmospheric river related extremes over the united states west coast under anthropogenic warming *Geophys. Res. Lett.* **52** e2024GL112237
- Hu H, Dominguez F, Wang Z, Lavers D A, Zhang G and Ralph F M 2017 Linking atmospheric river hydrological impacts on the U.S. west coast to rossby wave breaking *J. Clim.* **30** 3381–99
- Huang X and Swain D L 2022 Climate change is increasing the risk of a California megaflood *Sci. Adv.* **8** eabq0995
- Huang X, Swain D L and Hall A D 2020a Future precipitation increase from very high resolution ensemble downscaling of extreme atmospheric river storms in California *Sci. Adv.* **6** eaba1323
- Huang X, Swain D L, Walton D B, Stevenson S and Hall A D 2020b Simulating and evaluating atmospheric river-induced precipitation extremes along the U.S. pacific coast: case studies from *J. Geophys. Res.: Atmos.* **125** e2019JD031554
- Hunke E C and Lipscomb W H 2008 CICE: the Los Alamos sea ice model documentation and software user's manual, version 4 *Technical Report No. LA-CC-06-012* (Los Alamos National Laboratory)
- Inda-Díaz H A, O'Brien T A, Zhou Y and Collins W D 2021 Constraining and characterizing the size of atmospheric rivers: a perspective independent from the detection algorithm *J. Geophys. Res.: Atmos.* **126** e2020JD033746
- Kamnani D, O'Brien T A, Smith S, Staten P W and Shields C A 2025 Regional and temporal variability of atmospheric river seasonality: influences of detection algorithms and moisture transport dynamics *J. Geophys. Res.: Atmos.* **130** e2024JD043032
- Kapp-Schwoerer L, Graubner A, Kim S and Kashinath K 2020 Spatio-temporal segmentation and tracking of weather patterns with light-weight Neural Networks *NeurIPS 2020 Workshop: AI for Earth Sciences* (available at: https://ai4earthscience.github.io/neurips-2020-workshop/papers/ai4earth_neurips_2020_55.pdf)
- Kim S and Chiang J C H 2022 Atmospheric river lifecycle characteristics shaped by synoptic conditions at genesis *Int. J. Climatol.* **42** 521–38
- Kim S, Graubner A, Kapp-Schwoerer L, Kashinath K and Schindler K 2025 A large-scale dataset for training deep learning segmentation and tracking of extreme weather *Sci. Data* **12** 1151
- Kingma D P and Ba J 2017 Adam: a method for stochastic optimization (arXiv:1412.6980)
- Konrad C P and Dettinger M D 2017 Flood runoff in relation to water vapor transport by atmospheric rivers over the western united states *Geophys. Res. Lett.* **44** 11456–62
- Konstali K, Spengler T, Spensberger C and Sorteberg A 2024 Linking future precipitation changes to weather features in CESM2-LE *J. Geophys. Res.: Atmos.* **129** e2024JD041190
- Kurth T et al 2018 Exascale deep learning for climate analytics (arXiv:1810.01993)
- Lamjiri M A, Dettinger M D, Ralph F M and Guan B 2017 Hourly storm characteristics along the U.S. West Coast: role of atmospheric rivers in extreme precipitation *Geophys. Res. Lett.* **44** 7020–8
- Lawrence D M et al 2011 Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model *J. Adv. Model. Earth Syst.* **3** M03001
- Leung L R and Qian Y 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
- Liang J and Haywood J 2023 Future changes in atmospheric rivers over East Asia under stratospheric aerosol intervention *Atmos. Chem. Phys.* **23** 1687–703

- Liang J, Yong Y and Hawcroft M K 2023 Long-term trends in atmospheric rivers over East Asia *Clim. Dyn.* **60** 643–66
- Liu X, Chang P, Dan F, Saravanan R, Wang H, Rosenbloom N, Zhang S and Wu L 2022 Improved simulations of atmospheric river climatology and variability in high-resolution CESM *J. Adv. Model. Earth Syst.* **14** e2022MS003081
- Liu X, Xiaohui M, Chang P, Jia Y, Dan F, Guangzhi X, Wu L, Saravanan R and Patricola-DiRosario C M 2021 Ocean fronts and eddies force atmospheric rivers and heavy precipitation in western North America *Nat. Commun.* **12** 1268
- Liu Y, Racah E, Prabhat, Correa J, Khosrowshahi A, Lavers D, Kunkel K, Wehner M and Collins W 2016 Application of deep convolutional neural networks for detecting extreme weather in climate datasets (arXiv:1605.01156)
- Lora J M, Shields C A and Rutz J J 2020 Consensus and disagreement in atmospheric river detection: ARTMIP global catalogues *Geophys. Res. Lett.* **47** e2020GL089302
- Ma W, Norris J and Chen G 2020 Projected changes to extreme precipitation along North American west coast from the CESM large ensemble *Geophys. Res. Lett.* **47** e2019GL086038
- Mahesh A, O'Brien T A, Loring B, Elbashandy A, Boos W and Collins W D 2024 Identifying atmospheric rivers and their poleward latent heat transport with generalizable neural networks: ARCNNv1 *Geosci. Model Dev.* **17** 3533–57
- Mahoney K, Swales D, Mueller M J, Alexander M, Abel M R and Malloy K 2018 An examination of an inland-penetrating atmospheric river flood event under potential future thermodynamic conditions *J. Clim.* **31** 6281–97
- Massoud E C, Espinoza V, Guan B and Waliser D E 2019 Global climate model ensemble approaches for future projections of atmospheric rivers *Earth's Future* **7** 1136–51
- Meehl G A et al 2019 Effects of model resolution, physics and coupling on southern hemisphere storm tracks in CESM1.3 *Geophys. Res. Lett.* **46** 12408–16
- Michaelis A C, Gershunov A, Weyant A, Fish M A, Shulgina T and Ralph F M 2022 Atmospheric river precipitation enhanced by climate change: a case study of the storm that contributed to California's Oroville dam crisis *Earth's Future* **10** e2021EF002537
- Molina M J, Gagne D J and Prein A F 2021 A benchmark to test generalization capabilities of deep learning methods to classify severe convective storms in a changing climate *Earth Space Sci.* **8** e2020EA001490
- Mudigonda M, Kim S, Mahesh A, Kahou S, Kashinath K, Williams D, Michalski V, O'Brien T and Prabhat 2017 Segmenting and tracking extreme climate events using neural networks *31st Conf. on Neural Information Processing Systems (NIPS 2017)* (available at: https://dl4physicsciences.github.io/files/nips_dlps_2017_20.pdf)
- Mundhenk B D, Barnes E A, Maloney E D and Baggett C F 2018 Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden–Julian oscillation and quasi-biennial oscillation *npj Clim. Atmos. Sci.* **1** 20177
- Muszynski G, Kashinath K, Kurlin V, Wehner M and Prabhat 2019 Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets *Geosci. Model Dev.* **12** 613–28
- Neiman P J, Ralph F M, Wick G A, Lundquist J D and Dettinger M D 2008 Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the West Coast of North America based on eight years of SSM/I satellite observations *J. Hydrometeorol.* **9** 22–47
- O'Brien T A et al 2022 Increases in future AR count and size: overview of the ARTMIP Tier 2 CMIP5/6 experiment *J. Geophys. Res.: Atmos.* **127** e2021JD036013
- Park C, Son S-W and Kim H 2021 Distinct features of atmospheric rivers in the early versus late east asian summer monsoon and their impacts on monsoon rainfall *J. Geophys. Res.: Atmos.* **126** e2020JD033537
- Park S, Bretherton C S and Rasch P J 2014 Integrating cloud processes in the community atmosphere model, version 5 *J. Clim.* **27** 6821–56
- Payne A E, Demory M-E, Leung L R, Ramos A M, Shields C A, Rutz J J, Siler N, Villarini G, Hall A and Ralph F M 2020 Responses and impacts of atmospheric rivers to climate change *Nat. Rev. Earth Environ.* **1** 143–57
- Payne A E and Magnusdottir G 2014 Dynamics of landfalling atmospheric rivers over the north pacific in 30 Years of MERRA reanalysis *J. Clim.* **27** 7133–50
- Payne A E and Magnusdottir G 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** 11173–90
- Pepler A S, Dowdy A J and Hope P 2021 The differing role of weather systems in southern Australian rainfall between 1979–1996 and 1997–2015 *Clim. Dyn.* **56** 2289–302
- Polade S D, Gershunov A, Cayan D R, Dettinger M D and Pierce D W 2017 Precipitation in a warming world: assessing projected hydro-climate changes in California and other Mediterranean climate regions *Sci. Rep.* **7** 10783
- Prabhat et al 2021 ClimateNet: an expert-labeled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather *Geosci. Model Dev.* **14** 107–24
- Prabhat, Rübel O, Byna S, Wu K, Li F, Wehner M and Bethel W 2012 TECA: a parallel toolkit for extreme climate analysis *Proc. Comput. Sci.* **9** 866–76
- Quagrainie K T, O'Brien T A and Zhou Y 2025 Similarities in meteorological composites among different atmospheric river detection tools during landfall over western coastal North America *J. Geophys. Res.: Atmos.* **130** e2024JD042163
- Racah E, Beckham C, Maharaj T, Kahou S E, Prabhat and Pal C 2017 ExtremeWeather: a large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events (arXiv:1612.02095)
- Radić V, Cannon A J, Menounos B and Gi N 2015 Future changes in autumn atmospheric river events in British Columbia, Canada, as projected by CMIP5 global climate models *J. Geophys. Res.: Atmos.* **120** 9279–302
- Radke T, Fuchs S, Wilms C, Polkova I and Rautenhaus M 2025 Explaining neural networks for detection of tropical cyclones and atmospheric rivers in gridded atmospheric simulation data *Geosci. Model Dev.* **18** 1017–39
- Rahman Md A, Wang Y et al 2016 Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation *Advances in Visual Computing (ISVC 2016). Lecture Notes in Computer Science* (ed) G Bebis et al (Springer) pp 234–44
- Ralph F M et al 2019b ARTMIP-early start comparison of atmospheric river detection tools: how many atmospheric rivers hit northern California's Russian river watershed? *Clim. Dyn.* **52** 4973–94
- Ralph F M, Neiman P J and Wick G A 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** 1721–45
- Ralph F M, Rutz J J, Cordeira J M, Dettinger M, Anderson M, Reynolds D, Schick L J and Smallcomb C 2019a A scale to characterize the strength and impacts of atmospheric rivers *Bull. Am. Meteorol. Soc.* **100** 269–89
- Rasp S, Pritchard M S and Gentile P 2018 Deep learning to represent subgrid processes in climate models *Proc. Natl Acad. Sci.* **115** 9684–9
- Reid K J, King A D, Lane T P and Short E 2020 The sensitivity of atmospheric river identification to integrated water vapor transport threshold, resolution and regridding method *J. Geophys. Res.: Atmos.* **125** e2020JD032897

- Rhoades A M, Jones A D, Srivastava A, Huang H, O'Brien T A, Patricola C M, Ullrich P A, Wehner M and Zhou Y 2020 The shifting scales of western U.S. landfalling atmospheric rivers under climate change *Geophys. Res. Lett.* **47** e2020GL089096
- Rhoades A M, Jones A D and Ullrich P A 2018 The changing character of the California Sierra Nevada as a natural reservoir *Geophys. Res. Lett.* **45** 13008–19
- Rupe A, Kashinath K, Kumar N and Crutchfield J P 2025 Unsupervised discovery of extreme weather events using universal representations of emergent organization *Chaos* **35** 083110
- Rush W D et al 2025 Atmospheric river detection under changing seasonality and mean-state climate: ARTMIP Tier 2 paleoclimate experiments *J. Geophys. Res.: Atmos.* **130** e2024JD042222
- Rutz J J et al 2019 The atmospheric river tracking method intercomparison project (ARTMIP): quantifying uncertainties in atmospheric river climatology *J. Geophys. Res.: Atmos.* **124** 13777–802
- Rutz J J et al 2024 Atmospheric river tracking method intercomparison project tier 1 source data and catalogues UCAR/NCAR GDEX (<https://doi.org/10.5065/D62R3QFS>)
- Scholz S R and Lora J M 2024 Atmospheric rivers cause warm winters and extreme heat events *Nature* **636** 640–6
- Shields C A et al 2018 Atmospheric river tracking method intercomparison project (ARTMIP): project goals and experimental design *Geosci. Model Dev.* **11** 2455–74
- Shields C A et al 2022 Atmospheric river tracking method intercomparison project tier 2 reanalysis source data and catalogues UCAR/NCAR GDEX (<https://doi.org/10.26024/rawv-yx53>)
- Shields C A et al 2023 Future atmospheric rivers and impacts on precipitation: overview of the ARTMIP tier 2 High-resolution global warming experiment *Geophys. Res. Lett.* **50** e2022GL102091
- Shields C A and Kiehl J T 2016a Simulating the pineapple express in the half degree community climate system model, CCSM4 *Geophys. Res. Lett.* **43** 7767–73
- Shields C A and Kiehl J T 2016b Atmospheric river landfall-latitude changes in future climate simulations *Geophys. Res. Lett.* **43** 8775–82
- Shields C A, Li H, Castruccio F S, Fu D, Nardi K, Liu X and Zarzycki C 2024 Response of the upper ocean to northeast Pacific atmospheric rivers under climate change *Commun. Earth Environ.* **5** 603
- Smith R D et al 2010 The Parallel Ocean Program (POP) Reference Manual *Technical Report No. LAUR-10-01853* (Los Alamos National Laboratory)
- Swenson E T, Lu J and Straus D M 2018 Resolution dependence and rossby wave modulation of atmospheric rivers in an aquaplanet model *J. Geophys. Res.: Atmos.* **123** 6297–311
- Tebaldi C, Hayhoe K, Arblaster J M and Meehl G A 2006 Going to the extremes *Clim. Change* **79** 185–211
- Tian Y, Zhao Y, Son S-W, Luo J-J, Oh S-G and Wang Y 2023 A deep-learning ensemble method to detect atmospheric rivers and its application to projected changes in precipitation regime *J. Geophys. Res.: Atmos.* **128** e2022JD037041
- Ullrich P A and Zarzycki C M 2017 TempestExtremes: a framework for scale-insensitive pointwise feature tracking on unstructured grids *Geosci. Model Dev.* **10** 1069–90
- Utsumi N, Kim H, Kanae S and Oki T 2016 Which weather systems are projected to cause future changes in mean and extreme precipitation in CMIP5 simulations? *J. Geophys. Res.: Atmos.* **121** 10522–37
- Vallejo-Bernal S M, Wolf F, Boers N, Traxl D, Marwan N and Kurths J 2023 The role of atmospheric rivers in the distribution of heavy precipitation events over North America *Hydrol. Earth Syst. Sci.* **27** 2645–60
- Wang S, Xiaohui M, Guangzhi X, Zhou S, Chang P and Wu L 2022 A Comparison between the Kuroshio extension and pineapple express atmospheric rivers affecting the west coast of North America *J. Clim.* **35** 3905–25
- Warner M D, Mass C F and Salathé E P 2015 Changes in winter atmospheric rivers along the North American West Coast in CMIP5 climate models *J. Hydrometeorol.* **16** 118–28
- Wehner M F et al 2014 The effect of horizontal resolution on simulation quality in the community atmospheric model, CAM5.1 *J. Adv. Model. Earth Syst.* **6** 980–97
- Wehner M F, Reed K A, Loring B, Stone D and Krishnan H 2018 Changes in tropical cyclones under stabilized 1.5 and 2.0°C global warming scenarios as simulated by the community atmospheric model under the HAPPI protocols *Earth Syst. Dyn.* **9** 187–95
- Wu T, Tang S, Zhang R, Cao J and Zhang Y 2021 CGNet: a light-weight context guided network for semantic segmentation *IEEE Trans. Image Process.* **30** 1169–79
- Zhao M 2020 Simulations of atmospheric rivers, their variability and response to global warming using GFDL's new high-resolution general circulation model *J. Clim.* **33** 10287–303
- Zhou W, Leung L R and Lu J 2022 Linking large-scale double-ITCZ bias to local-scale drizzling bias in climate models *J. Clim.* **35** 7965–79
- Zhou Y and Kim H 2019 Impact of distinct origin locations on the life cycles of landfalling atmospheric rivers over the U.S. West Coast *J. Geophys. Res.: Atmos.* **124** 11897–909
- Zhou Y, Kim H and Guan B 2018 Life cycle of atmospheric rivers: identification and climatological characteristics *J. Geophys. Res.: Atmos.* **123** 12715–25
- Zhou Y, O'Brien T A, Ullrich P A, Collins W D, Patricola C M and Rhoades A M 2021 Uncertainties in atmospheric river lifecycles by detection algorithms: climatology and variability *J. Geophys. Res.: Atmos.* **126** e2020JD033711
- Zhou Y, Wehner M and Collins W 2024 Back-to-back high category atmospheric river landfalls occur more often on the west coast of the United States *Commun. Earth Environ.* **5** 187
- Zhu Y and Newell R E 1998 A proposed algorithm for moisture fluxes from atmospheric rivers *Mon. Weather Rev.* **126** 725–35