JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD038321

Special Section:
Atmospheric Rivers: Intersection of Weather and Climate

Six Hundred Years of Reconstructed Atmospheric River Activity Along the US West Coast

Swatah Snigdha Borkotoky[1], **A. Park Williams**[2,3], **and Scott Steinschneider**[1]

[1]Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY, USA, [2]Department of Geography, UCLA, Los Angeles, CA, USA, [3]Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

Key Points:

- Precipitation & tree-ring based data in the Western US can be used to reconstruct atmospheric river (AR) activity along the US west coast
- Reconstructions show 2–8 and 10–20-year modes of low-frequency variability in AR landfalls at locations across the coastline
- There is also an undocumented, multi-decadal signature in the variability of AR landfall latitude along the US west coast

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. S. Borkotoky,
sb2499@cornell.edu

Citation:

Borkotoky, S. S., Williams, A. P., & Steinschneider, S. (2023). Six hundred years of reconstructed atmospheric river activity along the US west coast. *Journal of Geophysical Research: Atmospheres*, *128*, e2022JD038321. https://doi.org/10.1029/2022JD038321

Received 5 DEC 2022
Accepted 12 JUN 2023

Author Contributions:

Conceptualization: Swatah Snigdha Borkotoky, Scott Steinschneider
Data curation: A. Park Williams
Formal analysis: Swatah Snigdha Borkotoky
Funding acquisition: Scott Steinschneider
Investigation: Swatah Snigdha Borkotoky
Methodology: Swatah Snigdha Borkotoky
Project Administration: Scott Steinschneider
Resources: A. Park Williams, Scott Steinschneider

Abstract Atmospheric rivers (AR) are critically important to water resources management along the US west coast, driving variability in both droughts and floods across the region. Inter-annual variability of ARs is well documented in the instrumental record back to the mid-twentieth century, but long-term variations in the frequency and landfall location of ARs along the US west coast are poorly understood due to limited records. This limitation impedes the ability to contextualize emerging trends and projections of AR activity. Here we use station-based records of daily precipitation and tree-ring records to present novel, spatially explicit estimates of daily AR occurrences in the first half of the twentieth century and annual AR counts over the last 600 years. First, we use neural networks and daily precipitation across Western North America to classify the daily occurrence of AR landfalls in three regions along the US west coast during the cold season back to 1916 CE. Then, we reconstruct the annual frequency of AR landfalls in those same regions back to 1400 CE using a gridded, tree-ring based reconstruction of the standardized precipitation index and a Poisson regression framework. The skillful reconstruction of daily and annual AR occurrences provides previously unavailable estimates of AR landfall variability and highlights new peaks in AR activity and modes of low-frequency variability prior to the instrumental record. Our reconstructions suggest that the average latitude of AR landfall has varied considerably on multi-decadal scales over the last 600 years, but without any discernible trends beyond this quasi-oscillatory behavior.

Plain Language Summary The climate of the US west coast is highly variable, with prolonged droughts punctuated by extremely wet conditions. These extremes are driven largely by the frequency of one type of storm event, known as atmospheric rivers (ARs), and the location where these storms make landfall along the coastline. Although observational data are available to quantify the frequency of these storms over the second half of the twentieth century, it is not well understood how these storms varied prior to this period. This knowledge gap makes it difficult to interpret the significance of recent trends or climate model projections in AR activity. To address this gap, we develop a novel data set of daily AR occurrences in the cold season for different regions along the US west coast for the first half of the twentieth century, based on station-based precipitation data and machine learning. We then reconstruct the annual frequency of AR landfalls going back to 1400 CE using tree-ring based estimates of precipitation. The results show that AR landfalls along the west coast have varied considerably over the last 600 years at multi-decadal timescales, but without any trends that clearly reflect the impacts of anthropogenic climate change.

1. Introduction

Since the seminal work of Zhu and Newell (1998), atmospheric rivers (ARs) have been increasingly recognized as critical to the water resources of the US west coast, and California in particular. ARs are long, narrow and transient channels of intense integrated vapor transport (IVT; Ralph et al., 2018) in the lower troposphere (less than 3 km altitude; Guan & Waliser, 2017; Ralph et al., 2004). They are closely, but not exclusively, associated with extra-tropical cyclones (Sodeman & Stohl, 2013; Zhang et al., 2019). ARs provide much of the annual precipitation to the US west coast (Dettinger & Cayan, 2014), most of which is delivered in the form of heavy precipitation events (Ralph et al., 2019). As such, the frequency of ARs and their landfall locations are major drivers of variability in both droughts and floods across much of the western US coastline.

Despite their importance, variability in the annual frequency of AR landfalls along the US west coast is not well understood due to records of limited length. In the second half of the twentieth century, there is good agreement between AR landfall dates using different detection algorithms and reanalysis or remotely sensed data sources

Supervision: A. Park Williams, Scott Steinschneider
Validation: Swatah Snigdha Borkotoky, A. Park Williams, Scott Steinschneider
Visualization: Swatah Snigdha Borkotoky
Writing – original draft: Swatah Snigdha Borkotoky
Writing – review & editing: Swatah Snigdha Borkotoky, A. Park Williams, Scott Steinschneider

(Brands et al., 2017; Guan & Waliser, 2015; Jackson et al., 2016), with only modest differences arising from varying data resolution (Guan & Waliser, 2017) and with respect to marginal features (e.g., regions of modest IVT; Lora et al., 2020). For the first half of the twentieth century, however, the availability of observations for assimilation is limited and reanalysis-based estimates of US west coast AR landfall occurrences tend to disagree with the limited observations available during this period (NOAA-20C and ERA-20C; Brands et al., 2017; Compo et al., 2011, Poli et al., 2016). Thus, our current understanding of AR landfall variability is generally limited to the most recent 70 years (post 1950).

Furthermore, projected future changes to precipitation along the US west coast are highly uncertain, in part because of uncertainty around changes to the frequency of ARs (Dettinger, 2016). The amount of water vapor transported by ARs is projected to increase under future warming due to the Clausius-Clapeyron relation (Held & Soden, 2006; updated by Hwang et al., 2011), and correspondingly, contributions of total precipitation from ARs are expected to increase (Gershunov et al., 2019). However, precipitation intensity may not scale as fast as would be expected under Clausius-Clapeyron scaling due to a decrease in upslope wind (Dettinger et al., 2011) and because longer duration events driven by synoptic scale systems (like ARs) are impacted by both dynamics as well as thermodynamics (Hartmann et al., 2013). In particular, poleward Hadley cell expansion is expected under warming (Grise & Davis, 2020) and shifts in the Northern Annular Mode are possible (Screen et al., 2018). The associated effects of these dynamic climate changes on AR landfall location are unclear. For instance, Shields and Kiehl (2016) projected that landfalling ARs over the US west coast will move equatorward during winter, while Ma et al. (2020) identified a poleward movement of ARs in the southern hemisphere. The uncertainty in such estimates is significant (Guirguis et al., 2018) and highlights the need to better understand past variability of ARs, so that projections and emerging trends in AR frequency and landfall location can be contextualized within their range of natural variability. The primary goal of this study is to develop this understanding by reconstructing the frequency and landfall location of ARs along the US west coast over several centuries using a combination of station-based precipitation records (for the first half of the twentieth century) and tree-ring based moisture proxies (for pre-twentieth century).

Along the US west coast, chronologies of tree-ring widths are uniquely suited to record AR activity. Generally, tree growth is sensitive only to available soil moisture during the growing season, not the source of precipitation (Fritts, 1966; Meko & Woodhouse, 2010). But in the US west and especially the US southwest, tree growth is sensitive to cold-season precipitation extremes, even though sensitivity is generally stronger to dry extremes (Dannenberg et al., 2019; Stahle et al., 2020). There are a few reasons for this behavior. First, much of the annual precipitation in this region is delivered during the cold season by ARs (Dettinger et al., 2011), often in the form of heavy and extreme events (Ralph et al., 2019; Steinschneider et al., 2016). Therefore, these storms form a core source of moisture that fuels tree growth. Second, these events recharge the snowpack that is later released to soil moisture during the growing season, helping to bridge the difference in timing between moisture delivery and tree growth (Stahle et al., 2020). Finally, as shown recently in Howard et al. (2023), certain tree species like blue oak (*Quercus douglasii*) are especially sensitive to heavy cold season precipitation, more so than non-heavy precipitation. Howard et al. (2023) argue this is possibly because only heavy precipitation events in the cold season recharge deep soil moisture, especially at lower elevation and warmer locations of the Coastal Range and Sierra Nevada where blue oak are often found and precipitation falls more often as rain than snow. Conversely, non-heavy precipitation mostly wets the shallow soil, and so is more vulnerable to evaporative loss before the trees can access it for growth. For all these reasons, cold-season precipitation and the large storms that deliver it are uniquely important to tree growth across much of the US west coast.

Previous work has leveraged this unique response of tree growth to heavy, cold-season precipitation in the western US to reconstruct ARs at an annual scale over the southern California coastline (Steinschneider et al., 2018) and extreme precipitation (often related to ARs) in the Sacramento River basin (Borkotoky et al., 2021) and Sierra Nevada (Howard et al., 2023). These studies identified quasi-periodic, inter-decadal signals in their respective target variables back to the mid-1400s. In addition, Steinschneider et al. (2018) showed that chronologies in Arizona, New Mexico, Utah, Colorado and central Mexico were useful for reconstructing AR landfalls along the southern California coastline, and Borkotoky et al. (2021) highlighted the link between extreme precipitation events in the Sacramento River Watershed and tree-ring moisture proxies in the Pacific Northwest. These results indicated that tree-ring chronologies beyond the target region for reconstruction were able to provide useful information to support the reconstruction. ARs with high intensity often penetrate the Sierra Nevada, Cascades, and Peninsular Ranges through low-elevation corridors and deliver moisture further into the Intermountain West

(IMW) (Rutz & Steenberg, 2012). In so doing, precipitation patterns in the IMW can be indicative of the location of AR landfall along the coast (Alexander et al., 2015; Rutz et al., 2014, 2015; Swales et al., 2016). Similar results have been found for ARs striking the western Canadian coastline (Vallejo-Bernal et al., 2022). This provides a unique opportunity to capture the spatial pattern of landfalling and inland-penetrating ARs based on precipitation recorded in tree-ring chronologies along these moisture conducive pathways, especially from moisture-sensitive chronologies within the drier interior of western North America.

The work in Borkotoky et al. (2021) and Steinschneider et al. (2018) established the possibility of reconstructing ARs and related precipitation extremes within specific regions along the US west coast. However, those studies did not develop a pan-coastal-scale assessment of AR frequency or how the latitude of AR landfalls has varied over the last several centuries. In addition, those studies only focused on annual reconstructions of AR frequencies using tree-ring based records but did not attempt to use more finely resolved information (long daily precipitation records) to develop a more detailed reconstruction of AR behavior in the recent past. These are the tasks undertaken in this work. The present study aims to answer the following research questions:

1. How well can the daily and annual occurrence of AR landfalls be reconstructed using station data of daily precipitation (over the twentieth century) and tree-ring based estimates of annual precipitation anomalies (back to 1400 CE), respectively, for regions along the US west coast?
2. How has the frequency and landfall location of ARs across the US west coast varied over the last 600 years?

To answer these questions, we develop for the first time a reconstruction of daily, cold-season landfalling AR occurrences for three regions along the US Pacific coastline using a novel machine learning approach and station-based precipitation data that is available back to 1916 CE. This reconstruction reveals previously unreported AR occurrence variability at a daily scale for the first half of the twentieth century. We also develop a novel set of reconstructions of the frequency of annual cold-season AR precipitation events for those same regions using a recent tree-ring based reconstruction of cold-season precipitation totals that is available back to 1400 CE. This reconstruction can be used to understand long-term (decadal to centennial) variability of ARs along the US west coast. Together, these reconstructions compose the most comprehensive assessment of pre-instrumental AR landfall variability along the US west coast that is available to date.

2. Data

2.1. Atmospheric River Landfall Catalog

Daily occurrences of AR landfalls along the US west coast are taken from the Scripps Institute of Oceanography (SIO)-generated AR catalog (SIO-R1 catalog; Gershunov et al., 2017). This catalog reports individual AR events across western North America from 1 January 1948 to 31 December 2020, detected using IVT and integrated water vapor (IWV) from the National Center for Environmental Prediction-National Center for Atmospheric Research reanalysis data (Kalnay et al., 1996). ARs are defined as a 1500 km long structure with IVT and IWV in excess of 250 kg/m/s and 15 mm, respectively, sustained continuously for at least 18 hr. The grid cell corresponding to the maximum IVT along the coastline is considered to be the AR landfall location. The catalog was validated in three ways, including: (a) a comparison against the special sensor microwave/imager—integrated water vapor catalog by Neiman et al. (2008) (updated in 2015); (b) an assessment of the AR contributions to observed precipitation over the western US; and (c) a comparison to AR databases based on other reanalysis products.

For this study, we use the daily occurrence (binary value of 0 or 1) of cold-season (October–March) ARs landfalling at six locations along the US west coast. These locations are separated by 2.5° latitude (the native resolution of the SIO-R1 catalog), and are near the following metropolitan areas: Santa Barbara, California; San Francisco, California; Redding, California; Medford, Oregon; Portland, Oregon; and Seattle, Washington (see Figure 1). These daily occurrences are then aggregated spatially for three regions (south: Santa Barbara and San Francisco, central: Redding and Medford, north: Portland and Seattle), where a region experiences an AR landfall occurrence on a given day if either or both locations in that region experience an AR landfall on that day. This aggregation was used to reduce noise in the signal between landfall location and both the precipitation gauges and tree-ring based moisture proxies. AR landfalls in these three regions form the target variables for reconstruction in this work.

2.2. Precipitation

The station-based Global Historical Climatology Network (GHCN; Menne, Durre, Vose, et al., 2012) is the primary source of daily precipitation data used in this study. We utilize a previous, gap-filled version of this data set that includes 470 stations across western North America with data that extend back to 1915 (Borkotoky et al., 2021; see Figure 1). For the gap-filling procedure, all missing daily precipitation totals are marked as zero if every other station within a specified radius (starting at 50 km but made larger if necessary) with valid precipitation measurements that day reported zero precipitation. The remaining missing values are then gap-filled based on quantile mapping to arrive at a continuous record of station-based daily precipitation. The final data set only includes those stations which have, prior to gap filling, at least 75% of daily values in at least half of all calendar months during at least three of the four 26-year sub-periods from 1 January 1915 to 31 March 2020 (with the final period being 17.25 years). As such, gap-filling is applied only to stations that already have good data coverage throughout 1915–2020. Borkotoky et al. (2021) provides additional detail regarding the gap-filling procedure.

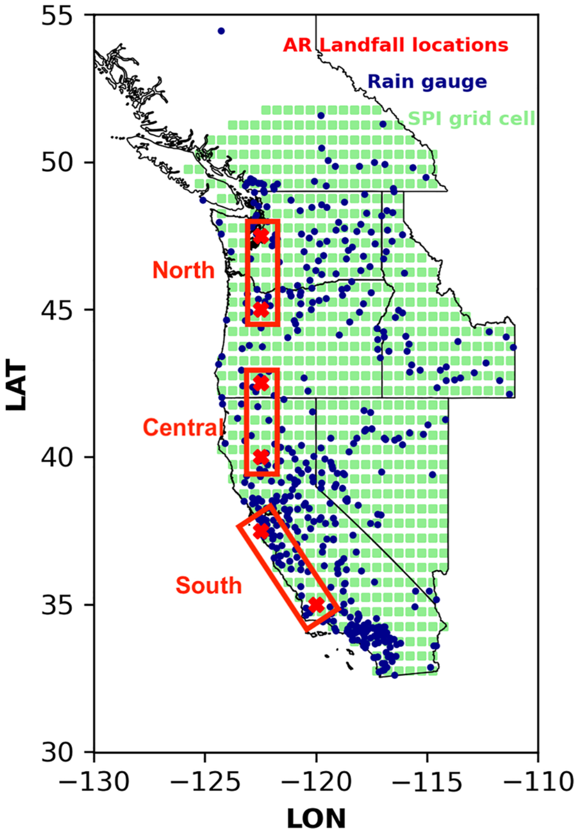

Figure 1. Atmospheric river (AR) landfall locations across the US west coast (red), the network of rainfall stations (blue), and the gridded standardized precipitation index (SPI) (green). The three rectangles designate the northern, central, and southern regions of AR landfall occurrences along the coast that are the targets for reconstruction.

2.3. Reconstructed SPI

We utilize a gridded, tree-ring based reconstruction (0.5° resolution) of the standardized precipitation index (SPI) for cumulative cold-season (November to March) precipitation total across western North America (Figure 1), described in detail in Borkotoky et al. (2021) and Williams et al. (2021). This data set, which extends from 1400 to 2020 CE, leverages an updated version of the chronology network used for previous hydroclimate reconstructions in western North America (e.g., Cook et al., 2010; Stahle et al., 2020; Williams et al., 2020). The network is composed of 1620 chronologies of tree-ring width index and spans from 24°N to 56°N, 102°W–124°W. The chronologies have continuous data coverage from at least 1800–1983 CE, and 1,458, 866, 483, and 332 chronologies extend back to 1700, 1600, 1500, and 1400 CE, respectively. Most of the chronologies, which are calculated from raw tree-ring width measurements obtained from the International Tree-ring Databank (hosted by NOAA), are detrended using the signal-free method (Cook et al., 2015; Melvin & Briffa, 2008). The chronologies are used in a point-by-point principal component (PC) regression to reconstruct cold-season SPI based on a merged precipitation product derived from the 1/24° NOAA Climgrid (USA; Vose et al., 2014b) and 0.5° CRU 4.04 (Canada and Mexico; Harris et al., 2014) monthly precipitation datasets. After 1983, the gridded product is based on observed SPI. Additional details about this gridded SPI data set are discussed in our previous study (Borkotoky et al., 2021; Williams et al., 2021).

3. Methods

We employ two predictive models to reconstruct the occurrence of AR landfalls in three regions along the US west coast (Figure 2). First, we develop independent neural networks to estimate daily cold-season AR landfall occurrences in each of the three regions in Figure 1 (i.e., one neural network per region), using the daily, gap-filled GHCN precipitation totals as predictors. We introduce methods for uncertainty propagation and model interpretability to understand better the reliability of these reconstructions. Using the fitted neural networks, we reconstruct the daily occurrence of AR landfalls back to 1916 CE. We then reconstruct annual counts of cold-season AR landfalls in the three regions using three separate Poisson regressions, each using PCs of the gridded SPI data as covariates. These regression models are used to reconstruct the annual landfall totals across the US west coast back to 1400 CE. We note that the neural network based daily reconstructions of AR landfalls are developed independently of the SPI based reconstructions of annual AR counts, and are used for two purposes: (a) as a useful stand-alone product that can help resolve uncertainty around AR landfall activity in the early twentieth century (see Brands et al., 2017); and (b) as an additional (and independent) data set against which to validate our SPI-based reconstructions of annual AR counts. Specifically, during the period of overlap between these two models when no AR observations are available (1916–1948), we use the neural networks' predictions, aggregated to an annual scale, as a secondary source of validation for the SPI-based Poisson regressions.

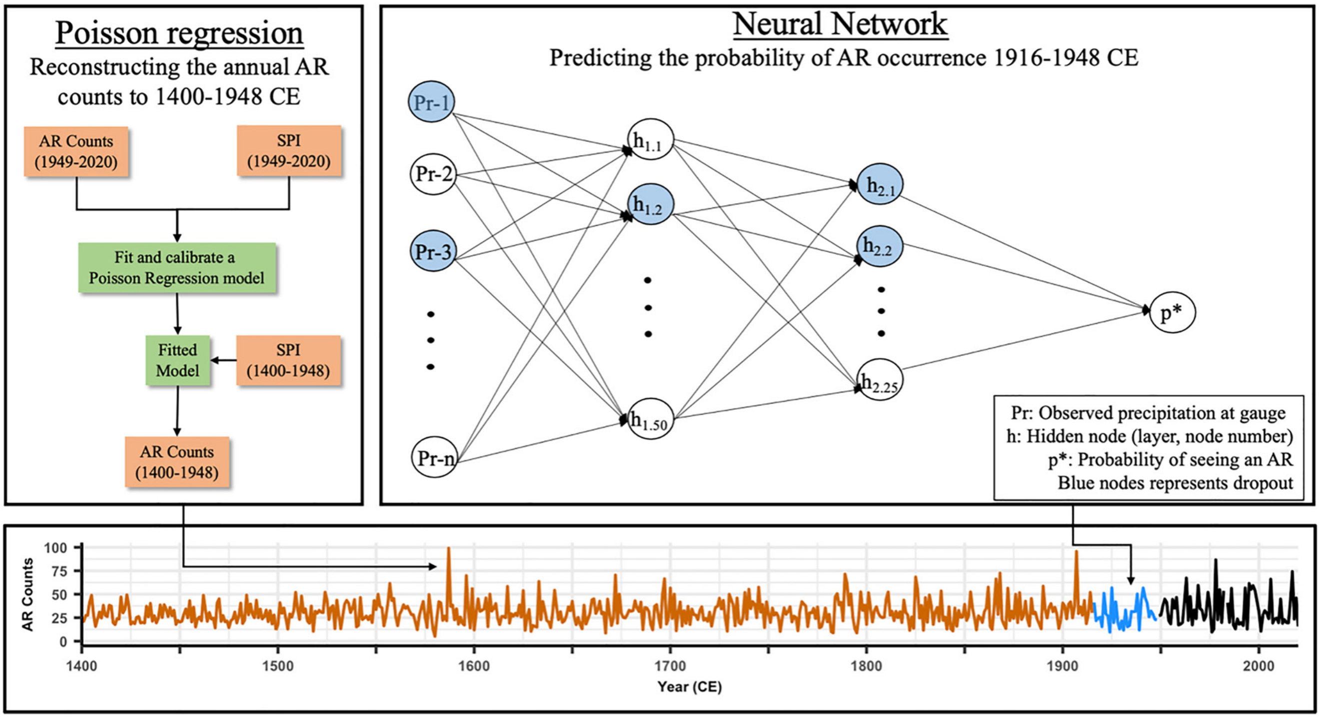

Figure 2. Workflow used to reconstruct atmospheric river (AR) counts across the US west coast. (Top right) The neural network is fitted using observed precipitation data (1949–2020) as input to predict daily occurrences of ARs back to 1916. These daily occurrences are then aggregated to an annual total. (Top left) The Poisson regression is fit based on reconstructed standardized precipitation index (SPI) data (1949–2020) and then used to reconstruct annual AR counts back to 1400 CE.

3.1. Reconstruction of Daily Cold-Season Atmospheric River Landfall Occurrences

3.1.1. Basic Neural Network Structure

Neural networks (NNs) are non-linear predictive models that learn associations between input data and an output (or target) variable of interest by processing the input data through a set of hidden layers (Hastie et al., 2013; Nielsen, 2015). The inter-connected hidden layers are composed of processing units known as neurons (or nodes). Every node in a hidden layer is derived from a weighted sum of the outputs of the previous layer. The associated weights for each node represent its relative importance in computing the corresponding output. The non-linearity of NNs comes from the activation function within a node, which allows NNs to approximate any function (Hornik et al., 1989). This function takes in a single value as an input and computes an output based on a pre-defined operation. Some of the common activation functions are logistic, tanh, ReLU, and SeLU. Every layer is also associated with a bias term that can shift the outputs of the activation function at each node in a layer by the same amount.

In our study, each of the three NNs is comprised of 470 input nodes (for 470 precipitation stations across western North America, see Figure 1), followed by a first hidden layer of 50 nodes, then a second hidden layer with 25 nodes, and lastly a single output node. The output values are the daily probability of occurrence of a landfalling AR in one of the three regions along the US west coast.

3.1.2. Training and Testing

Each NN is trained by minimizing the binary cross entropy loss function (Murphy, 2012). We use the Adam optimizer (Kingma & Ba, 2014) to tune the weights via backward propagation and employ the SeLU activation (Klambeur et al., 2017) across all nodes except the output node. In our application, the data from 1967 to 2020 was used for training and that from 1949 to 1966 was used for testing. This led to a 75%/25% split in the data for training and testing, a common ratio for such modeling efforts. We also note that the NN structure is well suited to adapt to any long-term biases in the precipitation data that might have arisen from the gap-filling procedure.

For each NN, we select a probability threshold above which an AR is considered to have occurred. This threshold is selected by maximizing the F1 score (Chincor, 1992) of the training data. For any arbitrary threshold, a sequence of AR occurrences (and non-occurrences) can be derived and compared to the observed sequence to produce four quantities: true positives (the model correctly predicts an AR occurrence), false positives (the model incorrectly predicts an AR occurrence), true negatives (the model correctly predicts no AR occurrence) and false negatives (the model fails to predict an AR). The F1 score is the ratio of true positives to the sum of true positives and the average of false positives and false negatives. The thresholds for each of the three models are selected separately; these thresholds are used to assess model performance on the test set.

We also assess model performance using Receivers Operation Characteristics (ROC) curves. The ROC curve is developed by considering a range of thresholds and plotting the true positives against the false positives. For an ideal classifier, the curve will be at the top left corner with all true positives and zero false positives. We quantify performance using the area under the ROC curve (AUC, Bradley, 1996), with the best performing model having an AUC equal to 1.

3.1.3. Dropout

NNs are very susceptible to overfitting, especially with limited training data. Dropout is a regularization technique wherein a fraction of nodes along with their connections are removed or “dropped out” during training (Srivastava et al., 2014). The rate at which nodes are removed at each layer (l) is given by a user-defined probability (p^l). Removing nodes randomly forces the NN to learn more rigorous weights and avoid features that are highly co-adapted (dependent) with each other, thereby ensuring that predictions are more generalizable out-of-sample. That is, it prevents the model from learning and fitting intricate structure of the training data that might not be present in the testing data (Hinton et al., 2012). Predictions in the testing set are made by scaling the weights of the retained nodes in layer l by its dropout rate p^l (multiplying by $1 - p^l$). This ensures that the output of the “thinned network” is maintained for the out-of-sample data. Another advantage of using dropout is that it gives an “ensemble” of thinned NNs rather than one overfitted and dense network, which helps in capturing model uncertainty. Dropout is mathematically equivalent to a Bayesian approximation (Gal & Ghahramani, 2016) and can provide a predictive posterior distribution by running multiple forward passes through a thinned network with dropout, a process known as *Monte Carlo dropout* (Gal & Ghahramani, 2016; Srivastava et al., 2014). In this study, a constant dropout rate of 50% was used across all layers.

3.1.4. Hyperparameter Tuning

Hyperparameters are parameters that are user-defined and not learned during the model training process. In this study, hyperparameters include the number of hidden layers, the number of nodes in each hidden layer, the learning rate, and the dropout rate. The number of hidden layers was selected based on cross validation performance of different independently run networks with one to five hidden layers, with 50, 25, 20, 10, and 5 hidden nodes in the first, second, third, fourth and fifth hidden layer, respectively. The optimal learning rate was selected separately for each model via a grid search. The performance of all models across all regions was found to be largely insensitive to the dropout rate and hence it was held constant at 50%.

3.1.5. Interpreting Neural Network Output

The results of a neural network are not easily interpretable due to its complex structure, but recent techniques have been introduced to help reduce the “black box” nature of these models. We use saliency maps (Simonyan et al., 2013), first introduced by Itti et al. (1998), to help understand how different precipitation gauges contribute to the prediction of AR occurrences for any given observation. Saliency is defined as the gradient of a backward pass at the input layer for a particular output. In our context, saliency helps to identify features (observations of precipitation) that have outsized influence on predictions of AR occurrence for specific days of the record. We average the total saliency for all the days in which an AR struck in a particular region and use it to identify which precipitation gauges have, on average, more influence in predicting the occurrence of landfalling ARs.

3.2. Reconstruction of Annual Cold-Season Atmospheric River Occurrences

3.2.1. Poisson Regression of Annual AR Counts Against the Principal Components of SPI

The daily cold-season reconstruction of AR occurrence over the US west coast extends only to 1916 CE. Using the gridded SPI data as covariates, we fit a Poisson regression to reconstruct the annual cold season counts of AR occurrences in each of the three regions along the US west coast back to 1400 CE. As a first step, we use principal component analysis to summarize variability in the entire gridded SPI data set (1400–2020 CE) within five western US states (California, Oregon, Washington, Idaho, and Nevada) and the Canadian province of British Columbia. These states and provinces were selected because they encompassed the areas with the highest

Spearman rank correlation of the gridded SPI against the annual AR counts in each of the three landfall regions (see Figure 1), as compared to areas further into the North American interior. Using varying numbers of PCs as predictands, we fit Poisson regression models to the annual cold-season AR counts:

$$P(\text{AR counts}_i = z) = \frac{{\lambda_i}^z e^{-\lambda_i}}{z!} \tag{1}$$

$$\log \lambda_i = \beta_0 + {x_i}^T \beta \tag{2}$$

Here, x_i is the vector of PCs for the ith year, λ_i is the mean estimate of the annual AR counts in year i, and β is a vector of regression coefficients. The relative performance based on Spearman rank correlation is used to select the optimal number of PCs needed to capture the inter-annual variability of AR counts.

3.2.2. Model Validation and AR Counts Reconstruction

After selecting the optimal number of PCs, the regression model is trained and validated using a *k*-fold (leave-10-year-out) cross-validation approach. Ten years of observed data are withheld, the model is then trained based on the rest of the data, and the trained model is used to estimate the withheld 10 years of AR counts. This is done for all three AR landfall regions separately, to explore the ability of the reconstructed SPI data and Poisson regression to capture annual variability of AR landfalls across the US west coast. We also evaluate the SPI-based reconstruction against the reconstruction of the NN models aggregated to an annual scale in the common period prior to available observations (1916–1948 CE).

After assessing the fit and out-of-sample prediction skill of the three Poisson regression models, we use the fitted models to reconstruct annual AR counts for each location back to 1400 CE. Using wavelet analysis, we examine the inter-annual to multi-decadal variability of AR counts in each of the three regions. We also assess the latitudinal variability of landfalling ARs across the entire US west coast using the following approach:

1. We normalize the annual counts of ARs in each of the three regions by taking the difference between the AR count in a year and the minimum AR count across all years under consideration, and then dividing by the range (maximum − minimum) of AR counts across those years for that region. This normalized count of AR landfalls ranges between 0 and 1.
2. For each year, we add the normalized metric of annual AR counts across all three regions, thus developing an index representative of the strength of AR activity over the entire US west coast.
3. In addition, we develop a weighted average of the AR landfall latitude across the three regions, using the normalized metric of annual AR counts from step 1 as a weight that gets applied to the latitude of each region. To do this, we multiply the normalized metric of annual AR counts from step 1 by the latitude of the region (i.e., the average latitude of the two sites within the region), add the resulting values across all three regions, and divide by the sum of the normalized counts across sites.
4. Finally, we center the weighted average of the AR landfall latitude from step 3 around its climatological mean, producing annual, latitudinal anomalies in AR landfall location.

Time series of latitudinal anomalies in AR landfall location (step 4) are examined across the entire reconstruction and are visualized to emphasize those years with a large absolute count of ARs (based on the index from step 2 above). The normalization approach helps to differentiate years with significant shifts in the AR landfalling latitude but few AR occurrences from those years with both higher AR activity and a considerable shift in the average landfalling location.

Finally, the reconstruction of latitudinal variability was compared to six climate indices, namely the Pacific Decadal Oscillation, PDO (1855–2021 CE); Pacific North American, PNA (1951–2021 CE); Atlantic Multidecadal Oscillation, AMO (1871–2016 CE); El Niño–Southern Oscillation [NINO] (1950–2021 CE); and Quasi-Biennial Oscillation, QBO at two levels: 70 hPa and 50 hPa (both 1954–2021 CE). These indices were gathered from National Oceanic and Atmospheric Association (NOAA) Physical Science Laboratory.

4. Results

4.1. Validation and Reconstruction of Daily AR Landfall Occurrences

For each of the three regions, 1000 simulations of the NN (with dropout) are used to develop a predictive ensemble of the daily probability of AR occurrence. We use the median of this ensemble to assess model performance.

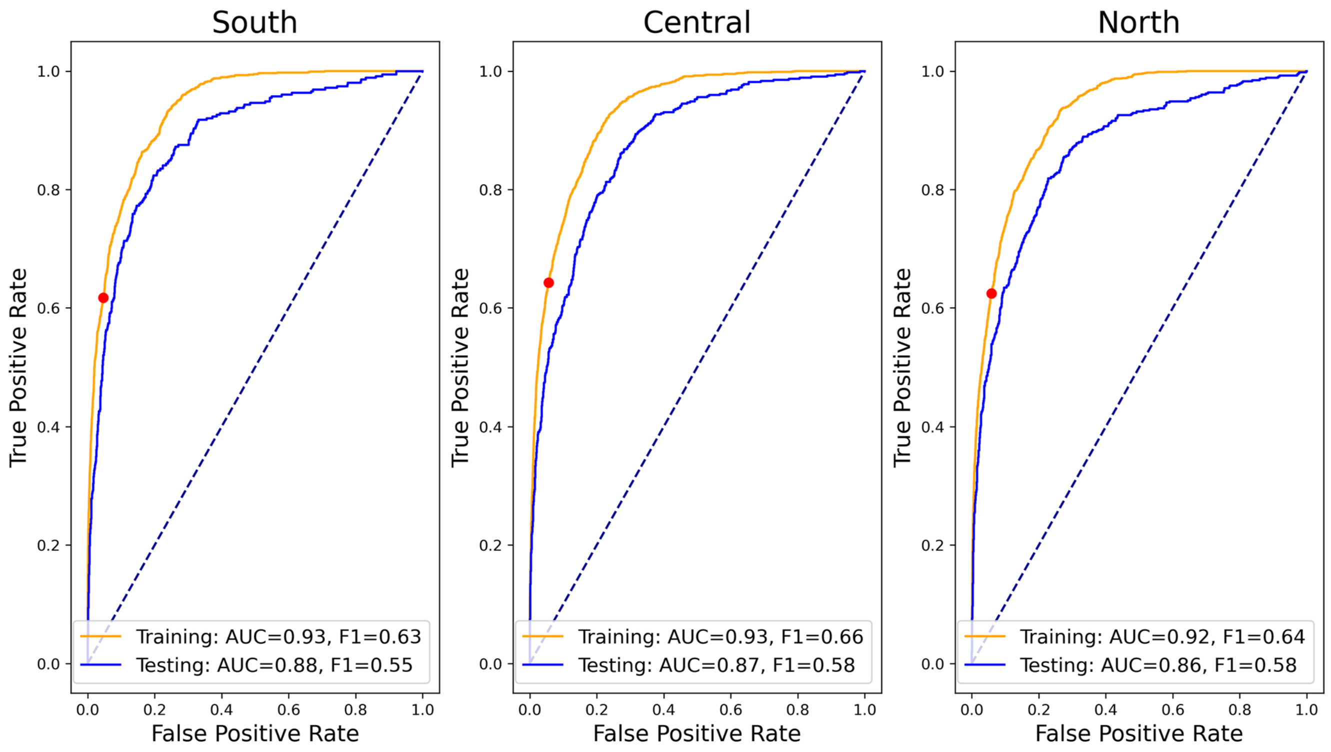

Figure 3. The receiver operating characteristic (ROC) curve for the neural networks in both the training and testing datasets for the three regions. The area under the curve (AUC) and the F1 score are shown on the bottom of each plot. The dashed diagonal line represents the performance of a "random guess" model. The red point shows model performance using the optimal probability threshold to distinguish atmospheric river occurrence from non-occurrence.

Figure 3 shows the ROC curves of the models based on a set of probability thresholds that ranges from 0 to 1. If the threshold of probability is 0, then the true positive and false positive will both be 1 (classify all days as AR days), and vice versa for a threshold of 1 (classify all days as non-AR days). The optimal probability threshold to distinguish AR occurrence from non-occurrence (i.e., the threshold that gives the optimal F1 score) is 0.3 for all regions over the training data (1967–2020). We highlight this threshold in Figure 3 (red point).

All models exhibit good skill, as evidenced by their ROC curves arching toward the top-left corner and AUC scores approaching 1. However, even with the optimal threshold, there are a non-trivial number of false positives and false negatives. Skill degrades modestly from the training to the testing set for all regions, with the smallest declines for the southern region and the largest decline for the northern region. Overall, though, performance is similar across regions.

The NN-based daily prediction ensembles of AR occurrence are aggregated to an annual scale to develop time series of annual cold-season counts of landfalling ARs across the three regions (i.e., 1000 time series of 72 years for each of the three regions). Figure 4 shows the performance of the three models on an annual basis for both training and testing periods, including the Spearman correlation and normalized root mean square error (nRMSE) for the observed period (1949–2020 CE). The model for the southern region was the best performing in terms of Spearman correlation (0.85), and within the predictive uncertainty bounds it was able to capture most years with high and low AR counts in both the training and testing sets. The performance for the models in the central region and the northern region were similar, with slightly lower values for the Spearman correlation coefficient (0.83 and 0.78, respectively) but also slightly better normalized RMSE than the southern model. Overall, these results indicate that the annual variability of AR counts is well captured by the NN models developed to reproduce the daily occurrence probability.

Importantly, the NNs identify previously undocumented (and sometimes record) peaks in annual AR activity in the first half of the twentieth century that align well with known flood and extreme precipitation events. For instance, AR landfalls in the northern region reached their highest levels in the 1933–1934 cold season, consistent with record precipitation totals during that year. Similarly, the NN reconstructions identify important peaks in AR frequency in the central region in the 1920s and in the late 1930s.

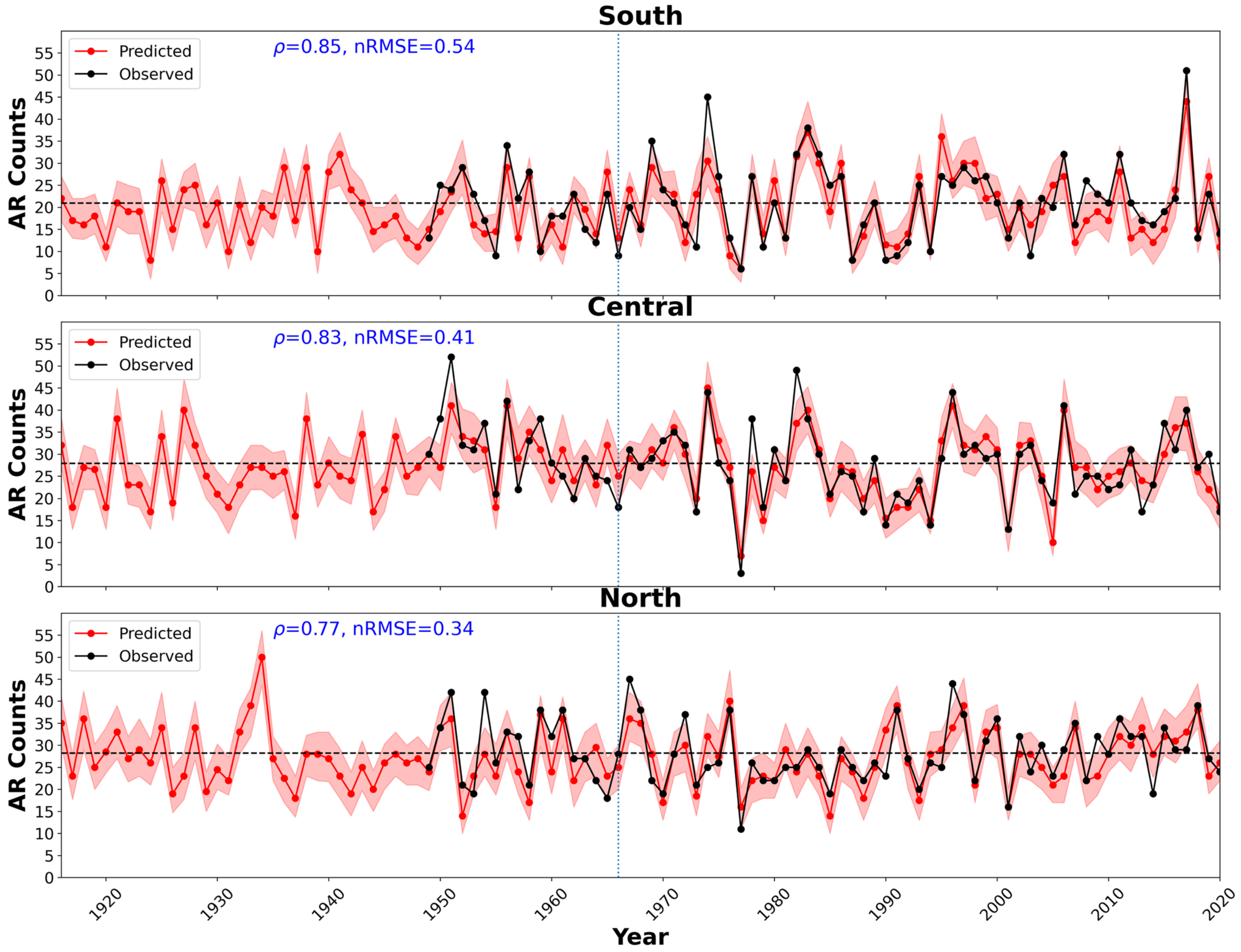

Figure 4. Time series of the observed (black) and NN-predicted (red) annual atmospheric river (AR) counts for the three regions. The red faded region represents the 95% predictive interval of the NN ensemble under dropout. The vertical line separates training and testing periods. The Spearman correlation and normalized root mean square error (i.e., RMSE divided by the climatological mean) for the observed period are also shown.

To help understand how the NNs use daily precipitation across the domain to estimate AR occurrence probabilities, Figures 5a–5c shows the percentage of seasonal precipitation delivered to each gauge during days of observed landfalling ARs for each of the three regions. These days are not mutually exclusive and may overlap across different regions. Almost 50% of seasonal precipitation was delivered across multiple precipitation gauges during days with AR landfalls across the southern and central regions. Interestingly, the gauges with a significant percentage of seasonal precipitation attributable to AR landfalls are not exclusively clustered around the AR landfall location. For instance, ARs making landfall in the southern and central regions (Figures 5a and 5b) often penetrate through low elevation conduits in the Sierra Nevada and Cascades and deliver precipitation further inland (Alexander et al., 2015), contributing a significant portion of seasonal precipitation to gauges in eastern Oregon, Washington, and central Idaho. Hence, landfalling ARs in the southern and central regions can contribute a substantial portion of the total cold season precipitation at locations further north and west. However, as the AR landfall location moves further north, the signal becomes weaker, with low seasonal contributions for AR landfalls in the northern region.

The contributions of ARs to total seasonal precipitation shown in Figures 5a–5c partially reflect how the NNs use precipitation information to generate predictions of AR occurrence. This is shown in Figures 5d–5f, which presents saliency maps for the three NN models. The saliency is analogous to the coefficients in a regression model. Positive (negative) saliency indicates a direct (inverse) association between the precipitation gauge and

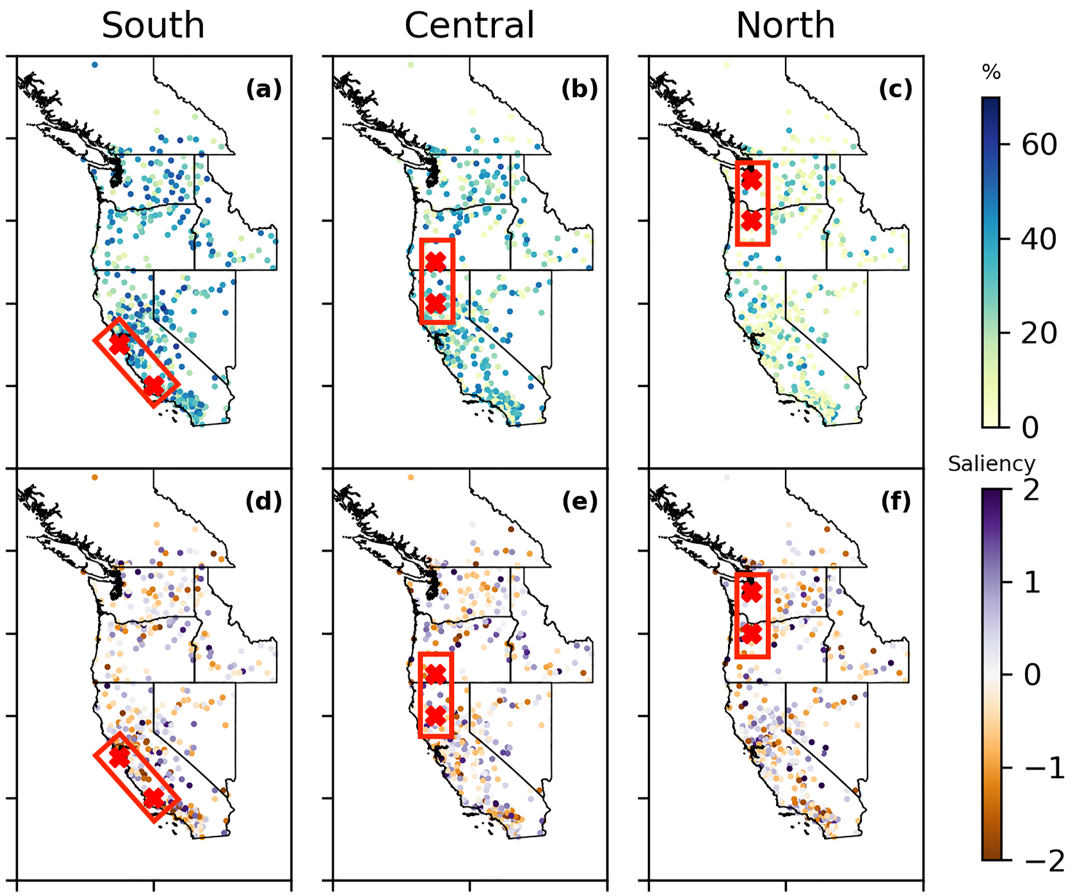

Figure 5. (Top, a–c) The percentage of cold season precipitation delivered during atmospheric river landfalling over the three regions. (Bottom, d–f) The average saliency of each precipitation gauge in each neural network, representing those gauges that most influence neural network predictions.

the corresponding prediction of AR occurrence. For all models, the gauges with large, positive saliency are similar to those that receive a large portion of seasonal precipitation during AR landfall days; for all regions, a regression of site-by-site average saliency (Figures 5d–5f) against percentage of seasonal precipitation during AR days (Figures 5a–5c) is significant at the 0.001 level. Sometimes, sites with high saliency are clustered near the sites of AR landfall, but not always. For instance, during AR landfalls in the northern region, the saliency map suggests that the model for this region uses precipitation signals far away from the location of AR landfall (e.g., in central Nevada). Further analysis (not shown) indicates that even though the area of most intense precipitation during AR landfall often occurs near the site of landfall, there is often lighter precipitation occurring over a much broader region of the US West. The neural network can predict AR landfalls well by identifying these complex spatial patterns in the precipitation gauging network.

Finally, we note that we also fit the NN models to AR occurrences at each of the original six locations in Figure 1, rather than the three regions. The performance for these models is slightly worse than that of the regional models (see: Text S1 and Figures S1 and S2 in Supporting Information S1).

4.2. Validation and Reconstruction of Annual AR Landfall Counts

We now consider predictions of annual AR counts based on the tree-ring based SPI reconstruction. As an initial diagnostic, we examine the correlation between the annual counts of landfalling ARs against the gridded SPI reconstruction during the years from 1949 to 2020 CE (Figures 6a–6c). For the southern and central regions, positive correlations are high in some locations (>0.7), and the gridded correlation values form a pattern that closely resembles that of an inland penetrating AR (Figures 6a and 6b). We also note that the areas of highest correlation

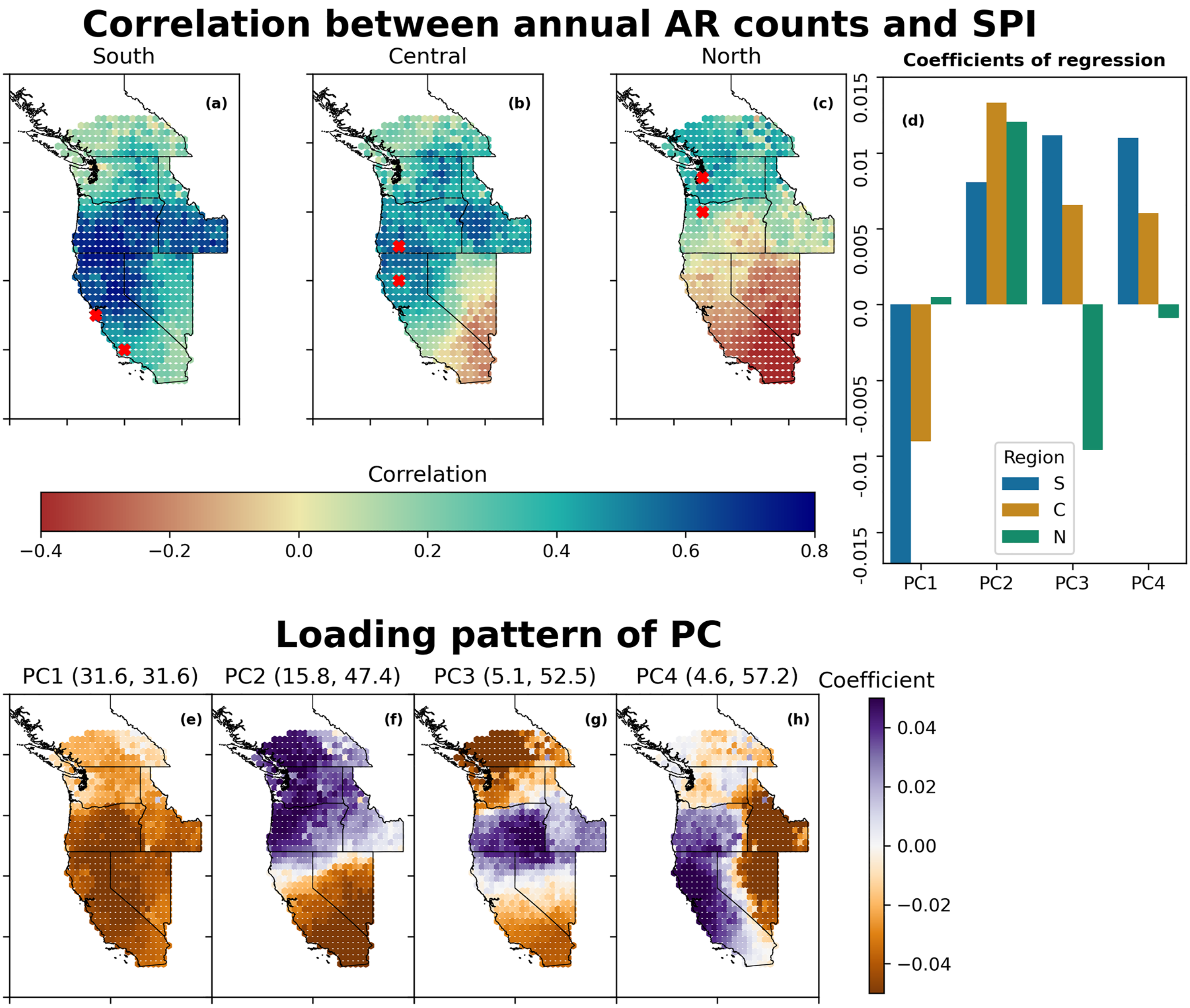

Figure 6. (Top left, a–c) The Spearman rank correlation between the observed annual atmospheric river counts in each of the three regions and the gridded standardized precipitation index (SPI) during the observed period (1949–2020). (Top right, d) The regression coefficients of each of the four principal components (PCs) used in the three Poisson regression models (south [S], central [C], and north [N]). (Bottom, e–h) The loading patterns of the four PCs used as predicands in the regression models, along with their percentage of variance explained and cumulative variance explained.

are often north of the area of AR landfall (Figures 6a–6c), consistent with previous findings that precipitation is often largest to the north of landfalls (Brodeur & Steinschneider, 2020). The areas of positive correlation between the SPI and annual AR counts weaken for the northern region. However, for this region, significant negative correlations emerge with the SPI in the Southwest, akin to a previously established dipole between precipitation in the Southwest and Pacific Northwest (Dettinger et al., 1998). We note that these correlation patterns are very consistent with a composite analysis of SPI anomalies during the 10 years with the greatest and lowest number of AR landfalls in each of the three regions (Text S2 and Figure S3 in Supporting Information S1).

Figure 6d shows the regression coefficients for the tree-ring based SPI PCs used by the regression models to reconstruct AR frequency, while Figures 6e–6h shows the loading patterns for those PCs. These four PCs explain nearly 60% of the total variance in the tree-ring based SPI, with PC1-PC4 explaining 31.6%, 15.8%, 5.1%, and 4.6%, respectively. Consistent with the results in Williams et al. (2021), PC1 exhibits a monopole across the entire western US (Figure 6e), particularly south of Washington State. The models for the southern and central regions rely heavily (and with negative association) on PC1. That is, if the SPI signal across most of the West is wet in a particular year, the models predict higher annual AR counts in all but the northern region. PC2 represents a north-south SPI dipole with centers over southern California and Nevada and the coastline of the Pacific Northwest (Figure 6f), and has a large, positive influence in all models (albeit slightly weaker for the southern model). PC3 emphasizes SPI variability in central and eastern Oregon, flanked by loadings of opposite sign in British Columbia and, to a lesser extent, southern California (Figure 6g). This PC has the most contrasting contribution to the regression models, with positive responses for models developed for the southern and central regions and a negative response for the northern region. PC4, with a zonally oriented dipole (Figure 6h), is most utilized for AR count predictions in the southern and central regions. Ultimately, the SPI signal that drives the AR reconstructions in each model depends on the regression coefficients (Figure 6d) and loading patterns (Figures 6e–6h) for all PCs simultaneously, and generally reflects the spatial patterns seen in the correlation maps in Figures 6a–6c.

Overall, the regression coefficients in Figure 6d, taken in conjunction with the loading patterns, show how the model for the northern region predicts the strongest increase in AR frequency when there are positive SPI anomalies in the far north and negative SPI anomalies in the far south of the domain. Due to the contrasting response to PC3, predictions from the northern model will deviate strongly from predictions from the other two models when SPI is positive in Washington State and further north, but relatively dry in Oregon, central Idaho, and further to the south. Predictions for the southern and central regions will be very correlated (given their similar regression coefficients), but dry conditions in the southernmost part of the domain will reduce predicted AR frequency more in the southern model compared to the central model.

The three independent Poisson regression models were evaluated using a leave-10-year-out cross validation. Figure 7 shows the out-of-sample predictions for all models. Additionally, the observations and annually aggregated AR counts from the three neural networks are overlaid to compare against the regression models' predictions. The models for the central region and especially the southern region estimate the interannual variability of AR counts very well, with Spearman rank correlations of 0.62 and 0.79, respectively (Figures 7a and 7b). In addition, the Poisson regressions show good agreement with the neural network estimates in the pre-observational period (1916–1948) for both the central and southern regions (rank correlations of 0.71 and 0.79). The model for the northern region exhibits somewhat lower skill, with Spearman rank correlation coefficients of 0.58 and 0.60 for the observed (1949–2020) and neural network predicted (1916–1948) AR counts (Figure 7c). Still, the northern model does capture some of the peaks and troughs in the observed series, and generally reflects observed variability on multi-year timescales. We also note that the northern model is able to distinguish between years with high AR frequencies in the south but low AR frequencies in the north using gradients in the SPI across those two regions (Text S3 and Figure S4 in Supporting Information S1).

Finally, we also refit the regression models for each of the original six locations and found some deterioration in performance for all locations when compared to their regional counterparts (Text S4 and Figure S5 in Supporting Information S1). This indicates that aggregating the six locations into the three representative regions improves the ability of the models to capture annual AR activity. Additionally, we performed a similar analysis using SPI data only through 1983 CE (a period during which the SPI data is based solely off tree-ring based estimates, rather than observed SPI post 1983). We found that the results were very similar to the models that use SPI data for the entire instrumental period (1916–2020 CE), including SPI PCs with extremely similar spatial loading patterns and explained variance within 0.5% of the original PCs, and regression coefficients of the same sign and very similar magnitude for all three models as compared to those in Figure 6d.

4.3. Inter-Annual Variability in AR Counts and Landfall Latitude

The fitted Poisson regression models were used to reconstruct the annual counts of landfalling ARs across the US west coast back to 1400 CE. These three reconstructions are shown in Figure 8. Overall, reconstructions for the southern and central regions are highly correlated (Pearson $r = 0.86$) and exhibit very similar low-frequency variability. However, the reconstructions between these two regions do occasionally deviate, for instance with the higher frequency of ARs in the south during the mid-1970s or occasional peaks in the south that are not present in the central region (e.g., just prior to 1650). The northern model deviates more from the central model and especially from the southern model, with correlation coefficients of 0.49 and 0.08, respectively. When the reconstructions are differenced, the northern and southern models exhibit the most notable, extended deviations in AR frequency during the 1470s, 1570s, 1840s, and 1950s, when AR frequency is higher in the north as compared to the south, and during the 1550s, 1740s, 1940s, and early 1980s when the opposite scenario emerges. These differences

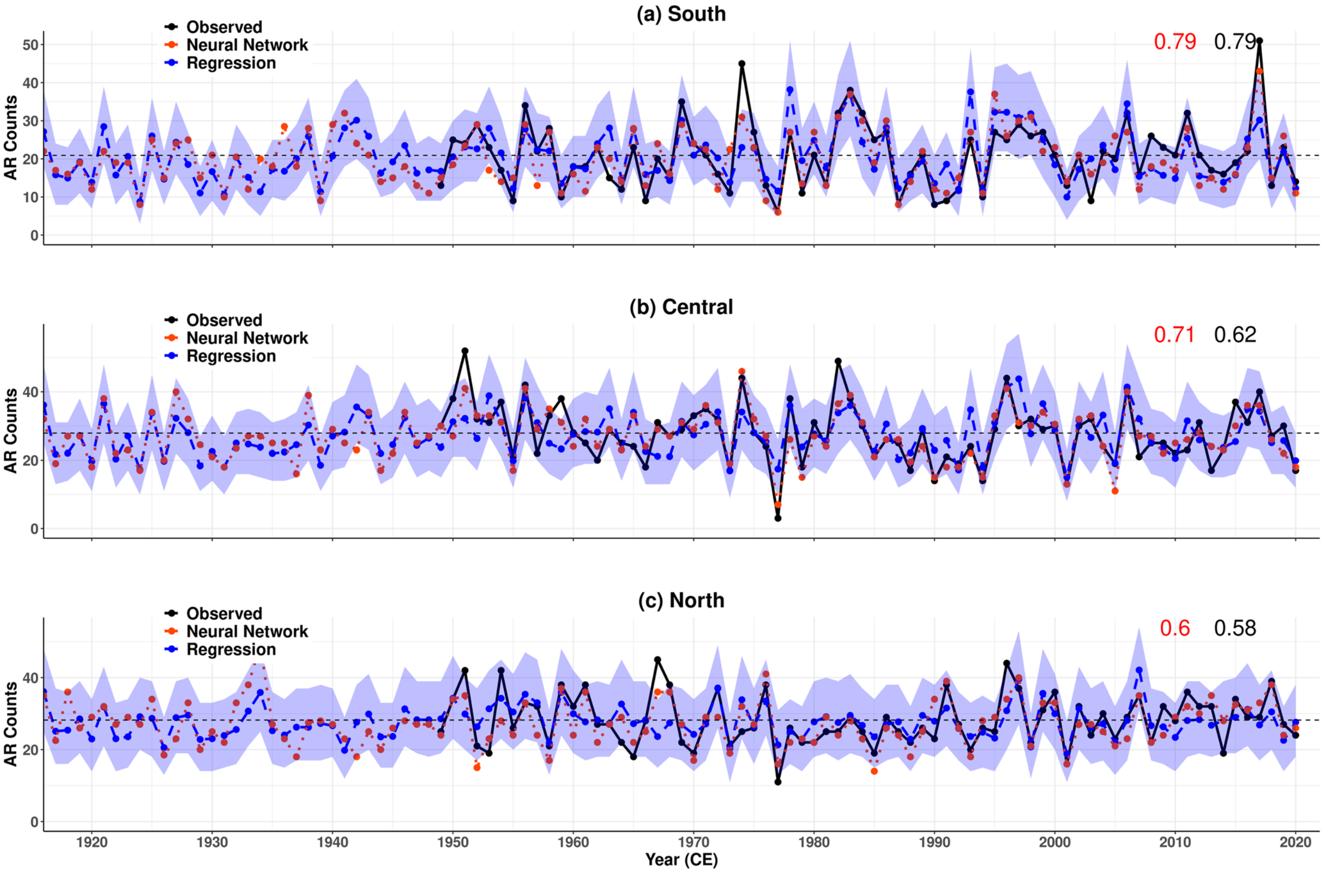

Figure 7. Time series of the observed (black) and leave-ten-years-out predicted (blue) annual atmospheric river (AR) counts for the three regression models. The red dashed line represents the neural network fitted (1949–2020) and reconstructed (1916–1948) annual AR counts for the three regions. The blue faded region represents the 95% predictive interval of the Poisson distribution under the regression models. The numbers in the top-right corner of each plot represent the Spearman rank correlation coefficient between the estimated AR counts from the regression model and the observations (in black) and neural network (in red).

emerge most prominently when there are significant spatial gradients in the SPI between the northern most part of the domain (Washington and southern Canada) and regions in Oregon, central Idaho, and further south (Figure S6 in Supporting Information S1). We also note that none of the three reconstructions suggest twentieth to twenty-first century trends in AR landfall frequency that exceed the variability experienced in previous centuries.

Figure 9 shows the results of a wavelet analysis applied to the three reconstructions, based on a Morlet wavelet (Torrence & Compo, 1998). In certain respects, the spectra are very similar across the three regions. For instance, they all exhibit a consistent periodic signal between 2 and 8 years that weakens prior to the 1500s. Another signal with 10–20 year periodicity is dominant in the late twentieth century everywhere (with diminished strength in the north), and appears intermittently throughout the three regions back to the mid-1500s. This signal is in agreement with the 13–15 years cyclicity in western US cold-season precipitation evaluated by Johnstone (2011), the 10–20 year periodicity of winter precipitation along the central Pacific coast (St. George & Ault, 2011), and an approximately 15-year periodic signal in AR landfalls in Northern California shown by Dettinger and Cayan (2014). Williams et al., 2021 also identified twentieth-century signals of similar periodicity for cool-season precipitation in the Sierra Nevada. An additional signal with a 30–70 year periodicity is present around the turn of the twentieth century for all regions and is seen throughout much of the reconstruction in the central region. This signal is also prominent in the north (with a break during the 1800s) but should be interpreted cautiously, as the model there exhibited somewhat weaker out-of-sample performance.

Figure 10 a shows the observed shift in latitude of the weighted-average AR landfall location across the US west coast during the instrumental period (using 1949–2020 CE for normalization; see Section 3.2.2), smoothed with a 10-year rolling average. Figures 10b and 10c show the same for the annual AR counts predicted by the NNs and

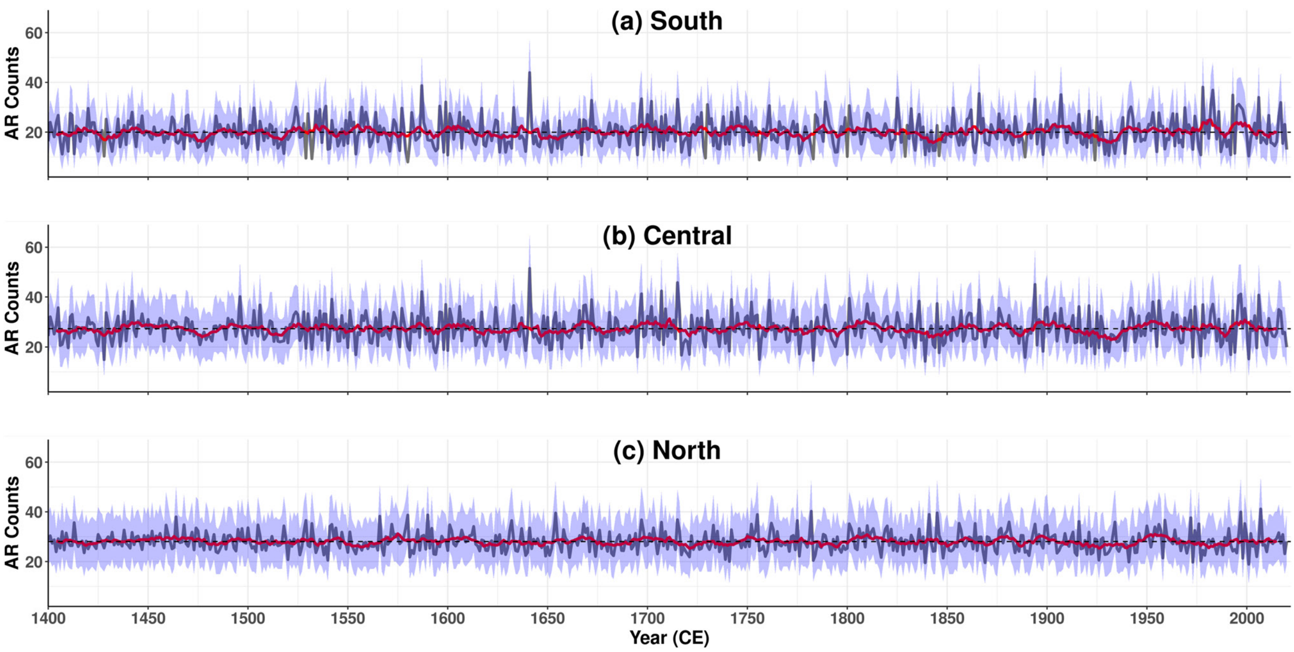

Figure 8. Time series of reconstructed annual atmospheric river (AR) counts (in black) for the three regions back to 1400 CE. The 10-year rolling average of the AR series is shown in red, while the 10-year rolling average of the 95% confidence interval is shown in blue bounds.

Poisson regression models, respectively. In the observations (Figure 10a), there was a significant poleward shift in AR landfall in the mid-1950s through the mid-1960s, followed by an equatorward shift that peaked in the early 1980s. The poleward shift from the mid-1950s to the mid-1960s is consistent with a shift in the transition zone of the winter precipitation dipole during that time (Wise, 2010). A weaker poleward shift of AR landfalls started in the late-1980s and lasted until the mid-1990s. Afterward, the 10-year rolling mean latitude of AR landfalls remained near the climatological average, with only a small tendency equatorward that appears to end around 2010.

The NNs (Figure 10b) generally replicate the same latitudinal shifts as seen in the observations, with a few exceptions. In particular, the NNs show smaller poleward shifts in landfall location in the mid-1950s and early 1960s that also extend to the early 1950s, a poleward shift in the 1970s that isn't present in the observations, a larger equatorward shift in the 1990s–2000s, and a stronger poleward shift after 2010. Arguably, the tree-ring based Poisson regression estimates (Figure 10c) do a better job in replicating the observed, decadal shifts in AR landfall location and frequency, although they predict a much more intense equatorward shift in the late 1990's and early 2000s. The Poisson regression models and NNs are also in close agreement for the period from 1916 to 1950, both in the frequency and direction of AR landfalls.

Figure 11 shows the 10-year rolling average latitudinal shift in landfalling ARs from the Poisson regression-based reconstruction (using 1400–2020 CE for normalization; see Section 3.2.2). The reconstruction suggests that there has been significant decadal periodicity in the average location of AR landfalls along the US west coast over the last several centuries. Throughout much of the 1400s, there was a significant poleward shift of ARs, only briefly interrupted by small equatorward shifts. Starting in the late 1400s and extending through until 1550, there were several short oscillations in landfall location. This was followed by a significant equatorward shift in ARs for about 30 years, and then from 1575 to 1675 there was a centennial poleward shift in AR occurrences interrupted only briefly by aperiodic breaks toward the equator. A long equatorward shift followed from 1675 to 1750, after which smaller multidecadal shifts in AR latitude continued until just before 1850. At the turn of the twentieth century through to present, there were several significant, equatorward shifts in AR occurrences, with only a few poleward shifts after 1950 and 2010. However, the full scope of the reconstruction suggests that while the latitude of AR landfalls has varied significantly on multi-decadal timescales over the last 600 years, there are no clear secular trends over the last several decades that suggest systematic shifts in the location of AR landfalls beyond what might be expected due to natural variability.

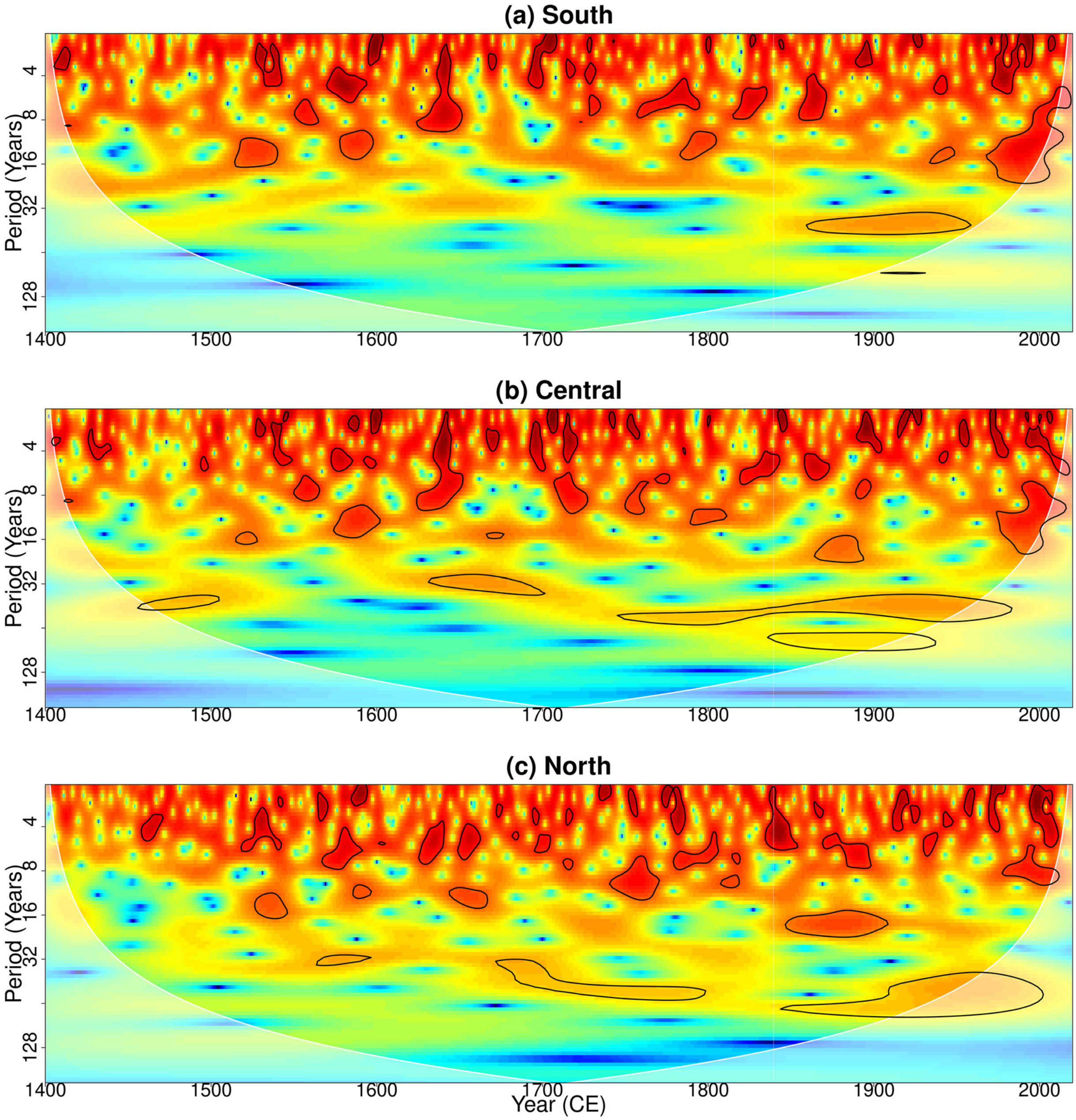

Figure 9. Wavelet power spectra for the annual atmospheric river count reconstructions across the three regions, based on the Poisson regression models. The areas within the shaded black boundary are significant at the 90% significance level.

Finally, Figures 12a–12e shows the 10-year rolling average of five climate indices (Pacific Decadal Oscillation [PDO], Pacific North American [PNA], Atlantic Multidecadal Oscillation [AMO], NINO3.4 and Quasi-Biennial Oscillation [QBO] at 70 hPa and 50 hPa) against the 10-year rolling average of the latitudinal variability of landfalling ARs (derived from the regression models). All of these indices are moderately correlated with AR latitudinal variability, with the PDO and PNA exhibiting negative correlations and AMO, NINO and QBO at 50 hPa and 70 hPa exhibiting positive correlations. The QBO at 70 hPa has the strongest correlation with AR latitudinal variability. This strong correlation is consistent with past work highlighting QBO controls over AR activity (Mundhenk et al., 2018). These six climate indices were used to estimate the latitudinal variability of landfalling ARs (1954–2016 CE) using a multivariate linear regression model. Figure 12f shows the observed versus the fitted AR latitudinal variability. It indicates that the combination of the six indices was able to capture the latitudinal variability of landfalling ARs quite well, with an $R^2 = 0.51$. Interestingly, if a similar model is

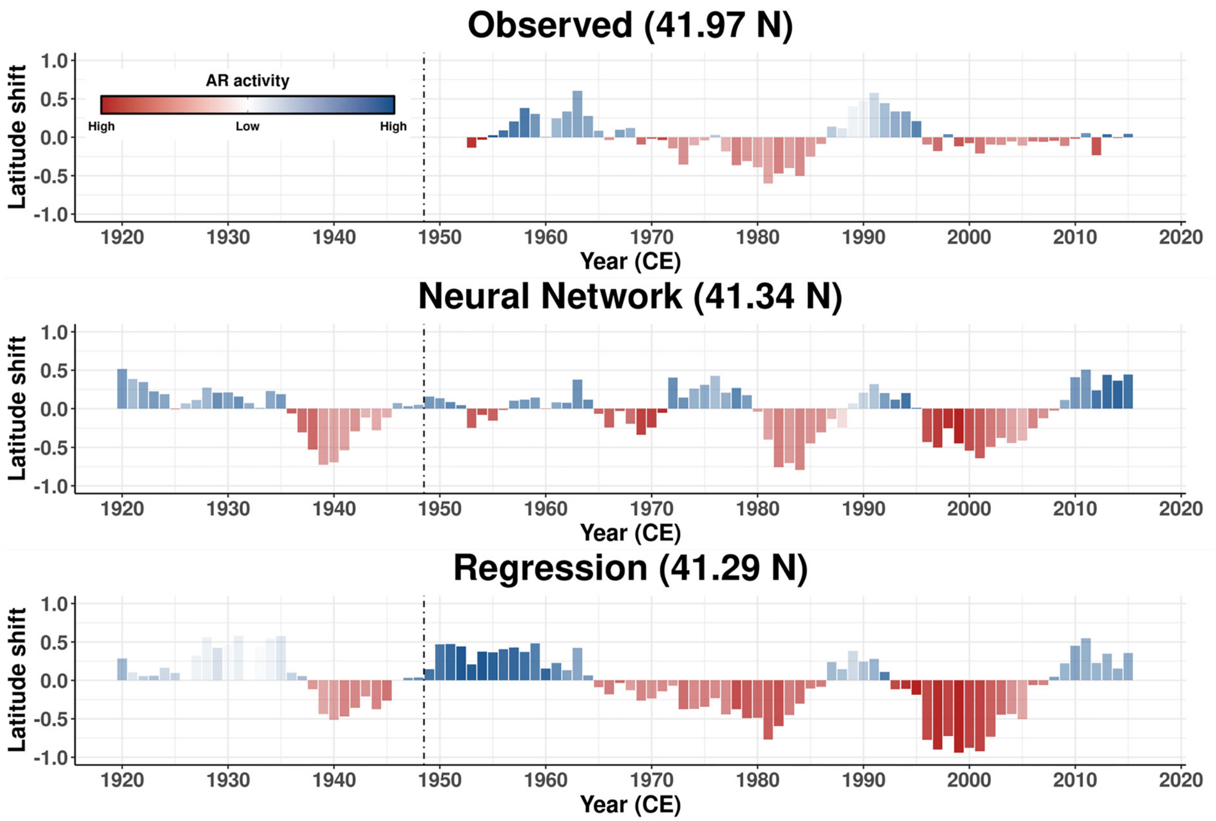

Figure 10. (a) The shift in weighted latitude (10-year rolling average) of landfalling atmospheric rivers (ARs) across the US west coast during the observed period (1949–2020 CE). The latitude within parentheses indicates the weighted-mean location of ARs during the instrumental period and the shading intensity represents a normalized score of the number of ARs landfalling during that year (darker shade means a greater number of ARs). A poleward (equatorward) shift is indicated by the blue (red) shading. The dashed vertical line demarcates the instrumental record from the reconstruction. (b) Represents the same weighted latitude shift of the landfalling ARs as estimated by the neural networks. (c) Denotes the same but for the tree-ring based Poisson regression estimates.

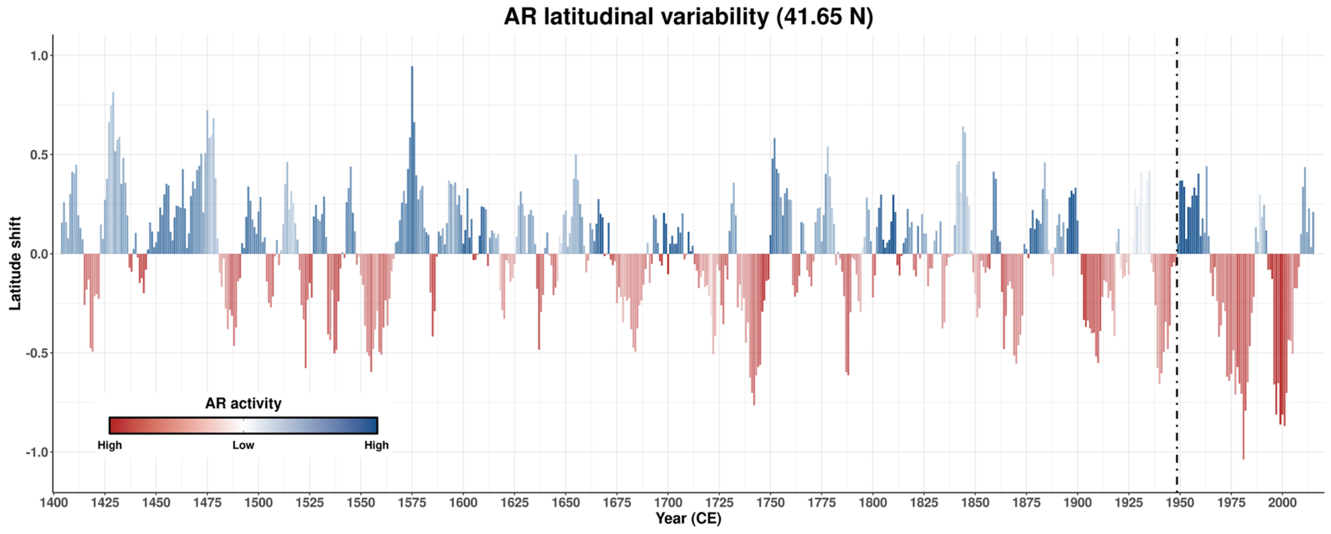

Figure 11. The 10-year rolling average shift in latitude of landfalling atmospheric rivers (ARs) across the US west coast over the last 621 years, based on reconstructions from the Poisson regression models. The latitude within parentheses indicates the weighted-mean location of ARs during the reconstruction period and the shading intensity represents a normalized score of the number of ARs landfalling during that year (darker shade means a greater number of ARs). A poleward (equatorward) shift is indicated by blue (red) shading.

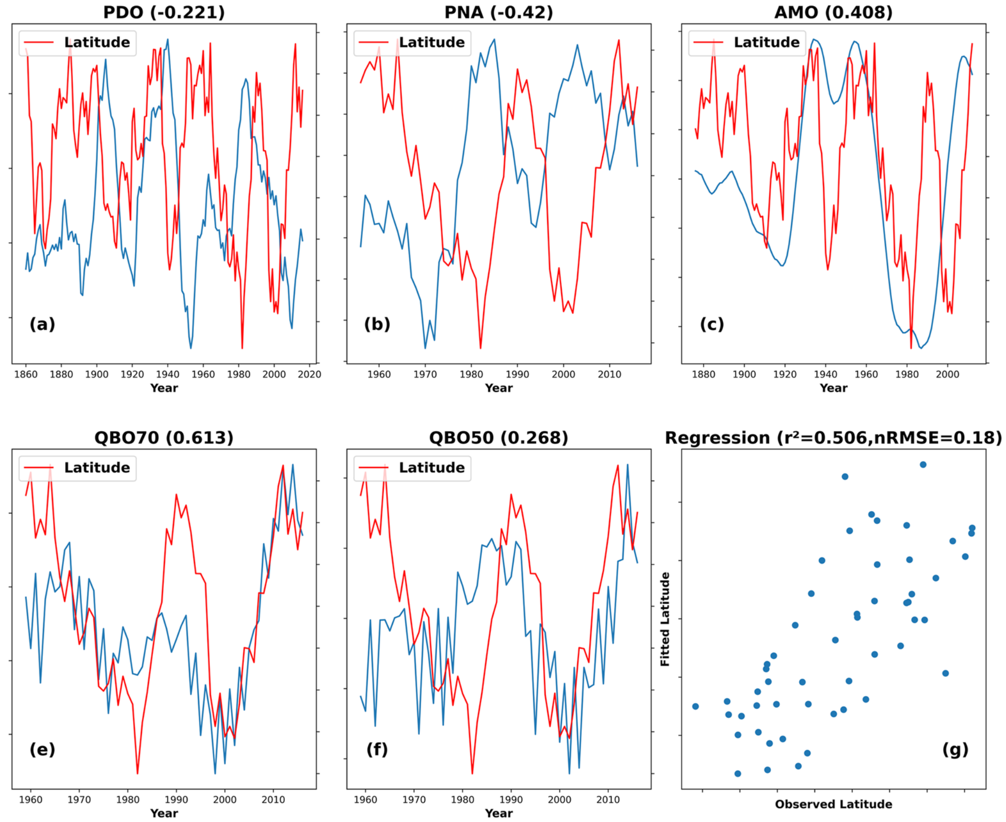

Figure 12. (a–f) The shift in landfalling atmospheric river (AR) latitude derived from the regression models (red) along the US west coast plotted against six climate indices (Pacific Decadal Oscillation [PDO], Pacific North American [PNA], Atlantic Multidecadal Oscillation [AMO], El Niño–Southern Oscillation [NINO] and Quasi-Biennial Oscillation [QBO] at 70 and 50 hPa shown in blue). The number within parentheses indicate the Spearman rank correlation between the two time series. (g) The observed shift in landfalling AR latitude against the AR latitude fitted using a linear regression model with the six climate indices. The number within the parentheses indicates the R^2 and the normalized root mean square error between the predicted and observed AR latitudinal variability.

developed for any of the individual AR frequency reconstructions shown in Figure 8, the R^2 is significantly lower and always less than 0.15 (see Figures S7–S9 in Supporting Information S1), suggesting that the climate indices are better related to the latitudinal variability of AR landfalls across the three regions rather than the frequency of AR landfalls in any one region.

5. Conclusions

This study established, for the first time, a comprehensive reconstruction of AR frequency and landfall location along the US west coast that extends, at daily timescales, back to 1916 and, at annual timescales, back to 1400 CE. The neural networks skillfully reproduced the daily occurrence of ARs for all coastal landfall regions and identified important (and sometimes record) peaks in AR frequency in the first half of the twentieth century that align with well-known extreme precipitation and flood events from that period. The spatial pattern of saliency for precipitation gauges varied by landfall location, but generally reflected those gauges that receive a large fraction of their precipitation from AR events landfalling at specific locations. For all landfall regions, the saliency highlighted precipitation gauges near the landfall location but also some very far away. This highlights how complex spatial patterns of precipitation across the western US can help capture the dynamics of moisture delivery as ARs make landfall along the coastline.

The NN-based AR reconstructions helped validate annual reconstructions of AR landfalls derived from tree-ring based cold-season precipitation data. Similar to the daily analysis, correlations between observed annual AR counts and the tree-ring based SPI suggested that different spatial signatures in the SPI data could help reconstruct AR landfalls in the Pacific Northwest and California. This was further supported by the Poisson regression coefficients assigned to different SPI PCs in AR reconstructions for each of the three regions. The annual AR reconstructions highlighted significant, low-frequency variability for all regions over the last 600 years. The most prominent signals included consistent periodicity in the 2–8 years band and a more intermittent 10–20 years signal. However, the most important result of the annual reconstruction was a previously undocumented, multi-decadal signature in the variability of AR landfall latitude along the entire western US coastline. Recent (i.e., post-1950) variations in landfall latitude from this reconstruction appear well within the range and pattern of natural variability predicted over the last 600 years, suggesting that any effects of climate change on AR latitude, if present, are not yet detectable.

Both the daily and annual reconstructions are novel datasets that can be used to understand AR variability and to assess the fidelity of AR detection algorithms applied to reanalysis products that extend prior to 1950 (see Brands et al., 2017). However, these reconstructions do have limitations that require discussion. First, daily AR occurrence is modeled using gauged precipitation that is influenced by other sources besides ARs. While the NNs combine the effects across multiple gauges to help isolate AR-related spatial patterns of moisture delivery, precipitation from non-AR sources certainly obscures the signal to a degree. Second, our reconstructions do not provide estimates of the strength of individual AR events, but rather only their frequency. Further work could explore whether alternative model structures could distinguish between weak and strong AR landfalls based on the spatial pattern and intensity of precipitation.

For the tree-ring based reconstructions, the southern and central regions were able to capture much of the variability in AR frequency, but skill is lower (albeit still notable) in the northern region. There are two reasons for this discrepancy in model performance. First and foremost, the skill of the SPI reconstruction is limited in most of Washington, Oregon, and Idaho (see Williams et al., 2021), which likely degrades the reconstruction of AR frequencies. This highlights that cold-season precipitation totals in the Pacific Northwest and northern IMW, which otherwise can be a good proxy for the frequency of AR landfalls, may not be well-suited for this purpose if first reconstructed using tree-ring chronologies. Additional work should focus on improving the reconstruction of cold-season precipitation in these regions, which will likely in turn help to resolve AR landfalls along the US Pacific Northwest coastline. This effort will have to contend with the larger warm season precipitation signal in the Pacific Northwest that can make reconstructions of cold-season precipitation challenging (St. George & Ault, 2014).

Second, ARs are not the only type of storm that can deliver precipitation, increase growing season soil moisture, and eventually be recorded as tree growth. This is particularly true in the Pacific Northwest, where extra-tropical cyclones unassociated with ARs generally deliver a greater proportion of cold season precipitation (Slinskey et al., 2020). More generally, the reconstructions of annual AR frequencies are based on proxies developed from moisture-sensitive tree-ring chronologies, which can at best only reflect changes in the likelihood of annual AR counts and cannot be used to identify individual events. In addition, the availability of tree-ring chronologies decreases substantively prior to 1550 CE, increasing the uncertainty in reconstructions prior to this date (Williams et al., 2021). Finally, as in most reconstructions, we have assumed stationarity in the relationship between AR landfalls and the SPI moisture proxies, which if incorrect could significantly reduce the credibility of our reconstructions.

Despite these limitations, the reconstructions of AR landfalls produced in this work are quite skillful and provide a state-of-the-art data set that can be used to better understand better past AR activity over the US west coast. This study builds on a growing literature trying to understand past shifts in the location of storm tracks (Wise & Dannenberg, 2014, 2017) and large-scale patterns of atmospheric flow (Gupta et al., 2022), so that emerging trends in atmospheric circulation can be put into historical context. Future work should explore the extension of the methods presented here to other regions of the world where ARs contribute substantially to total annual precipitation and networks of precipitation gauges and tree-ring chronologies are available to record such events, for example, central Chile and certain portions of the Iberian Peninsula (Eiras-Barca et al., 2018; Gallego et al., 2006; Morales et al., 2020; Natalini et al., 2016; Ramos et al., 2015; Viale et al., 2018).

Data Availability Statement

The data used in this analysis are publicly available, including daily station data gathered from the GHCN gauging network (Menne, Durre, Korzeniewski, et al., 2012), the NOAA Climgrid data set (Vose et al., 2014a), and the CRU 4.04 data set (Harris et al., 2020). The AR data set (SIO-R1) is derived from Gershunov et al. (2017). The cold-season SPI reconstruction and the RWI chronologies used to produce the SPI reconstruction are available at Borkotoky (2023). The AR reconstruction data can be accessed at Borkotoky (2022).

Acknowledgments

The authors acknowledge support from NSF grants AGS1702273, AGS1703029, and AGS1702184, NOAA award MAPP NA19OAR4310278, a grant from California's Department of Water Resources, and the many tree-ring scientists over the past decades who have collected, processed, and made available their tree-ring data for studies such as this. Special thanks to Jeremy Littell who collected, measured, and shared unpublished tree-ring datasets from the Northwest US.

References

Alexander, M. A., Scott, J. D., Swales, D., Hughes, M., Mahoney, K., & Smith, C. A. (2015). Moisture pathways into the U.S. intermountain west associated with heavy winter precipitation events. *Journal of Hydrometeorology*, *16*(3), 1184–1206. https://doi.org/10.1175/JHM-D-14-0139.1

Borkotoky, S. S. (2022). Six hundred years of reconstructed atmospheric river activity on the US west coast [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7384272

Borkotoky, S. S. (2023). Ring Width Index (RWI) and Standard Precipitation Index (SPI) data for atmospheric river reconstruction [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7893863

Borkotoky, S. S., Williams, A. P., Cook, E. R., & Steinschneider, S. (2021). Reconstructing extreme precipitation in the Sacramento River watershed using tree-ring based proxies of cold-season precipitation. *Water Resources Research*, *57*(4), e2020WR028824. https://doi.org/10.1029/2020WR028824

Bradley, A. P. (1996). ROC curves and the X^2 test. *Pattern Recognition Letters*, *17*(3), 287–294. https://doi.org/10.1016/0167-8655(95)00121-2

Brands, S., Gutierrez, J. M., & San-Martin, D. (2017). Twentieth-century atmospheric river activity along the west coasts of Europe and North America: Algorithm formulation, reanalysis uncertainty and links to atmospheric circulation patterns. *Climate Dynamics*, *48*(9–10), 2771–2795. https://doi.org/10.1007/s00382-016-3095-6

Brodeur, Z., & Steinschneider, S. (2020). Spatial bias in medium-range forecasts of heavy precipitation in the Sacramento River basin: Implications for water management. *Journal of Hydrometeorology*, *21*(7), 1405–1423. https://doi.org/10.1175/JHM-D-19-0226.1

Chincor, N. (1992). MUC-4 evaluation metrics. In *Proceedings of the 4th Conf. on message understanding* (pp. 22–29). https://doi.org/10.3115/1072064.1072067

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Cook, E. R., Seager, R., Heim, R. R., Jr., Vose, R. S., Herweijer, C., & Woodhouse, C. (2010). Megadroughts in North America: Placing IPCC projections of hydroclimatic change in a long-term paleoclimate context. *Journal of Quaternary Science*, *25*(1), 48–61. https://doi.org/10.1002/jqs.1303

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., et al. (2015). Old world droughts and pluvials during the Common Era. *Science Advances*, *1*(10), e1500561. https://doi.org/10.1126/sciadv.1500561

Dannenberg, M. P., Wise, E. K., & Smith, W. K. (2019). Reduced tree growth in the semiarid United States due to asymmetric responses to intensifying precipitation extremes. *Science Advances*, *5*(10), eaaw0667. https://doi.org/10.1126/sciadv.aaw0667

Dettinger, M. D. (2016). Historical and future relations between large storms and droughts in California. *San Francisco Estuary and Watershed Science*, *14*(2). https://doi.org/10.15447/sfews.2016v14iss2art1

Dettinger, M. D., & Cayan, D. R. (2014). Drought and the California delta-A matter of extremes. *San Francisco Estuary and Watershed Science*, *12*(2). https://doi.org/10.15447/sfews.2014v12iss2art4

Dettinger, M. D., Cayan, D. R., Diaz, H. F., & Meko, D. M. (1998). North-South precipitation patterns in Western North America on interdecadal-to-decadal timescales. *Journal of Climate*, *11*(12), 3095–3111. https://doi.org/10.1175/1520-0442(1998)011%3C3095:NSPPIW%3E2.0.CO;2

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

Eiras-Barca, J., Lorenzo, N., Taboada, J., Robles, A., & Miguez-Macho, G. (2018). On the relationship between atmospheric rivers, weather types and floods in Galicia (NW Spain). *Natural Hazards and Earth System Sciences*, *18*(6), 1633–1645. https://doi.org/10.5194/nhess-18-1633-2018

Fritts, H. C. (1966). Growth-rings of trees: Their correlation with climate. *Science*, *154*(3752), 973–979. https://doi.org/10.1126/science.154.3752.973

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd international conference on machine learning, New York, NY, USA*. Retrieved from http://proceedings.mlr.press/v48/gal16.pdf

Gallego, M. C., Garcia, J. A., Vaquero, J. M., & Mateos, V. L. (2006). Changes in frequency and intensity of daily precipitation over the Iberian Peninsula. *Journal of Geophysical Research*, *111*(D24), D24105. https://doi.org/10.1029/2006JD007280

Gershunov, A., Shulgina, T., Clemesha, R. E. S., Guirguis, K., Pierce, D. W., Dettinger, M. D., et al. (2019). Precipitation regime change in Western North America: The role of atmospheric rivers. *Scientific Reports*, *9*(9944), 9944. https://doi.org/10.1038/s41598-019-46169-w

Gershunov, A., Sulghina, T., Ralph, F. M., Lavers, D. A., & Rutz, J. J. (2017). Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophysical Research Letters*, *44*(15), 7900–7908. https://doi.org/10.1002/2017GL074175

Grise, K. M., & Davis, S. M. (2020). Hadley cell expansion in CMIP6 models. *Atmospheric Chemistry and Physics*, *20*(9), 5249–5268. https://doi.org/10.5194/acp-20-5249-2020

Guan, B., & Waliser, D. E. (2015). Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, *120*(24), 12514–12535. https://doi.org/10.1002/2015JD024257

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

Guirguis, K., Gershunov, A., Clemesha, R. E. S., Shulgina, T., Subramanian, A. C., & Ralph, F. M. (2018). Circulation drivers of atmospheric rivers at the North American West coast. *Geophysical Research Letters*, *45*(22), 12576–12584. https://doi.org/10.1029/2018GL079249

Gupta, R., Steinschneider, S., & Reed, P. M. (2022). A multi-objective paleo-informed reconstruction of western U.S. weather regimes over the past 600 years. *Climate Dynamics*, *60*(1–2), 339–358. https://doi.org/10.1007/s00382-022-06302-4

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - The CRU TS3.10 dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Harris, I., Osborne, T. J., Jones, P. D., & Lister, D. (2020). Climate research unit gridded time series [Dataset]. Retrieved from https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.04/

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., et al. (2013). Observations: Atmosphere and surface. In T. F. Stocker, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2013). *The elements of statistical learning* (p. 737). Springer.

Held, I., & Soden, B. J. (2006). Robust response of the hydrological cycle to global warming. *Journal of Climate*, *19*(21), 5686–5699. https://doi.org/10.1175/JCLI3990.1

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutidinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. Retrieved from https://arxiv.org/pdf/1207.0580.pdf

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Howard, I. M., Stahle, D. W., Dettinger, M. D., Poulsen, C., Ralph, F. M., Torbenson, M. C. A., & Gershunov, A. (2023). A 440-year reconstruction of heavy precipitation in California from Blue Oak Tree Rings. *Journal of Hydrometeorology*, *24*(3), 463–477. https://doi.org/10.1175/JHM-D-22-0062.1

Hwang, Y., Frierson, D. M. W., Soden, B. J., & Held, I. (2011). Corrigendum for Held and Soden (2006). *Journal of Climate*, *24*(5), 1559–1560. https://doi.org/10.1175/2010JCLI4045.1

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Learning*, *20*(11), 1254–1259. https://doi.org/10.1109/34.730558

Jackson, D. L., Hughes, M., & Wick, G. A. (2016). Evaluation of landfalling atmospheric rivers along the U.S. West coast in reanalysis data sets. *Journal of Geophysical Research: Atmospheres*, *121*(6), 2705–2718. https://doi.org/10.1002/2015JD024412

Johnstone, J. A. (2011). A quasi-biennial signal in western US hydroclimate and its global teleconnections. *Climate Dynamics*, *36*(3–4), 663–680. https://doi.org/10.1007/s00382-010-0755-9

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–472. https://doi.org/10.1175/1520-0477(1996)077%3C0437:TNYRP%3E2.0.CO;2

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint. https://doi.org/10.48550/arXiv.1412.6980

Klambeur, G., Unterthiner, T., Mayr, A., & Hochreiter, S. (2017). Self-normalizing neural networks. In *31st Conf. on advances in neural information processing systems, Long Beach, CA, USA*. Retrieved from https://arxiv.org/pdf/1706.02515

Lora, J. M., Shields, C. A., & Rutz, J. J. (2020). Consensus and disagreement in atmospheric river detection: ARTMIP global catalogues. *Geophysical Research Letters*, *47*(20), e2020GL089302. https://doi.org/10.1029/2020GL089302

Ma, W., Chen, G., & Guan, B. (2020). Poleward shift of atmospheric rivers in the southern hemisphere in recent decades. *Geophysical Research Letters*, *47*(21), e2020GL089934. https://doi.org/10.1029/2020GL089934

Meko, D. M., & Woodhouse, C. A. (2010). Application of streamflow reconstruction to water resources management. In M. Hughes, T. Swetnam, & H. Diaz (Eds.), *Dendroclimatology. Development in paleoenvironmental research* (pp. 231–261). Springer. https://doi.org/10.1007/978-1-4020-5725-0_8

Melvin, T. M., & Briffa, K. R. (2008). A "signal" free approach to dendroclimatic standardization. *Dendrochronologia*, *26*(2), 71–86. https://doi.org/10.1016/j.dendro.2007.12.001

Menne, M. J., Durre, I., Korzeniewski, B., McNeill, S., Thomas, K., Yin, X., et al. (2012). Global historical climatology network - Daily (GHCN-Daily). Version 3 [Dataset]. NOAA National Climatic Data Center. https://doi.org/10.7289/V5D21VHZ

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, *29*(7), 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1

Morales, M. S., Cook, E. R., Barichivich, J., Christie, D. A., Villalba, R., LeQuesne, C., et al. (2020). Six hundred years of South American tree rings reveal an increase in severe hydroclimatic events since mid-20th century. *Proceedings of the National Academy of Sciences*, *117*(29), 16816–16823. https://doi.org/10.1073/pnas.2002411117

Mundhenk, B. D., Barnes, E. A., Maloney, E. D., & Baggett, C. F. (2018). Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden–Julian oscillation and quasi-biennial oscillation. *NPJ Climate and Atmospheric Science*, *1*, 20177. https://doi.org/10.1038/s41612-017-0008-2

Murphy, K. (2012). *Machine learning: A probabilistic perspective* (p. 1104). MIT press. ISBN-13: 978-0262018029. Retrieved from http://noiselab.ucsd.edu/ECE228/Murphy_Machine_Learning.pdf

Natalini, F., Alejano, R., Vázquez-Piqué, J., Pardos, M., Calama, R., & Büntgen, U. (2016). Spatiotemporal variability of stone pine (*Pinus pinea L.*) growth response to climate across the Iberian Peninsula. *Dendrochronologia*, *40*, 72–84. https://doi.org/10.1016/j.dendro.2016.07.001

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

Nielsen, M. A. (2015). Neural networks and deep learning. Retrieved from http://neuralnetworksanddeeplearning.com/

Poli, P., Hersbatch, H., Dee, D. P., Barrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Ralph, F. M., Dettinger, M., Cairns, M. M., Galarneau, T. J., & Eylander, J. (2018). Defining "Atmospheric River": How the *Glossary of Meteorology* helped resolve a debate. *Bulletin of the American Meteorological Society*, *99*(4), 837–839. https://doi.org/10.1175/BAMS-D-17-0157.1

Ralph, F. M., Neiman, P. J., & Wick, G. A. (2004). Satellite and CALJET aircraft observations of atmospheric rivers over the eastern North Pacific Ocean during the winter of 1997/98. *Monthly Weather Review*, *132*(7), 1721–1745. https://doi.org/10.1175/1520-0493(2004)132%3C1721:SACAOO%3E2.0.CO;2

Ralph, F. M., Wilson, A. M., Shulgina, T., Kawzenuk, B., Sellars, S., Rutz, J. J., et al. (2019). ARTMIP-early start comparison of atmospheric river detection tools: How many atmospheric rivers hit northern California's Russian River Watershed? *Climate Dynamics*, *52*(7–8), 4973–4994. https://doi.org/10.1007/s00382-018-4427-5

Ramos, A. M., Trigo, R. M., Liberato, M. L. R., & Tome, R. (2015). Daily precipitation extreme events in the Iberian Peninsula and its association with atmospheric rivers. *Journal of Hydrometeorology*, *16*(2), 579–597. https://doi.org/10.1175/JHM-D-14-0103.1

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

Rutz, J. J., Steenberg, W. J., & Ralph, F. M. (2014). Climatological characteristics of atmospheric rivers and their inland penetration over the western United States. *Monthly Weather Review*, *142*(2), 905–921. https://doi.org/10.1175/MWR-D-13-00168.1

Rutz, J. J., Steenberg, W. J., & Ralph, F. M. (2015). The inland penetration of atmospheric rivers over the western North America: A Lagrangian analysis. *Monthly Weather Review*, *143*(5), 1924–1944. https://doi.org/10.1175/MWR-D-14-00288.1

Screen, J. A., Bracegirdle, T. J., & Simmonds, I. (2018). Polar climate change as manifest in atmospheric circulation. *Current Climate Change Reports*, *4*, 383–395. https://doi.org/10.1007/s40641-018-0111-4

Shields, C. A., & Kiehl, J. T. (2016). Atmospheric river landfall-latitude changes in future climate simulations. *Geophysical Research Letters*, *43*(16), 8775–8782. https://doi.org/10.1002/2016GL070470

Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional neural networks: Visualizing image classification models and saliency maps. Retrieved from https://arxiv.org/pdf/1312.6034.pdf

Slinskey, E. A., Loikith, P. C., Waliser, D. E., Guan, B., & Martin, A. (2020). A climatology of atmospheric rivers and associated precipitation for the seven U.S. National Climate Assessment Regions. *Journal of Hydrometeorology*, *21*(11), 2439–2456. https://doi.org/10.1175/JHM-D-20-0039.1

Sodeman, H., & Stohl, A. (2013). Moisture origin and meridional transport in atmospheric rivers and their association with multiple cyclones. *Monthly Weather Review*, *141*(8), 2850–2868. https://doi.org/10.1175/MWR-D-12-00256.1

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958. Retrieved from http://jmlr.org/papers/v15/srivastava14a.html

Stahle, D. W., Cook, E. R., Burnette, D. J., Torbenson, M. C. A., Howard, I. M., Griffin, D., et al. (2020). Dynamics, variability, and change in seasonal precipitation reconstructions for North America. *Journal of Climate*, *33*(8), 3173–3195. https://doi.org/10.1175/JCLI-D-19-0270.1

Steinschneider, S., Ho, M., Cook, E. R., & Lall, U. (2016). Can PDSI inform extreme precipitation? An exploration with a 500 year long paleoclimate reconstruction over the U.S. *Water Resources Research*, *52*(5), 3866–3880. https://doi.org/10.1002/2016WR018712

Steinschneider, S., Ho, M., Williams, A. P., Cook, E. R., & Lall, U. (2018). A 500-year tree-ring based reconstruction of extreme cold-season precipitation and number of atmospheric river landfalls across the Southwestern United States. *Geophysical Research Letters*, *45*(11), 5672–5680. https://doi.org/10.1029/2018GL078089

St. George, S., & Ault, T. R. (2011). Is energetic decadal variability a stable feature of the central Pacific Coast's winter climate? *Journal of Geophysical Research*, *116*(D12), D12102. https://doi.org/10.1029/2010JD015325

St. George, S., & Ault, T. R. (2014). The imprint of climate within Northern Hemisphere trees. *Quaternary Science Review*, *89*, 1–4. https://doi.org/10.1016/j.quascirev.2014.01.007

Swales, D., Alexander, M., & Hughes, M. (2016). Examining moisture pathways and extreme precipitation in the U.S. Intermountain West using self-organizing maps. *Geophysical Research Letters*, *43*(4), 1727–1735. https://doi.org/10.1002/2015GL067478

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079%3C0061:APGTWA%3E2.0.CO;2

Vallejo-Bernal, S. M., Wolf, F., Boers, N., Traxl, D., Marwan, N., & Kurths, J. (2022). Spatio-temporal synchronization of heavy rainfall events triggered by atmospheric rivers in North America. EGUsphere [preprint]. https://doi.org/10.5194/egusphere-2022-530

Viale, M., Valenzuela, R., Garreaud, R. D., & Ralph, F. M. (2018). Impacts of atmospheric rivers on precipitation in southern South America. *Journal of Hydrometeorology*, *19*(10), 1671–1687. https://doi.org/10.1175/JHM-D-18-0006.1

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams, C. N., Jr., et al. (2014a). NOAA monthly U.S. climate gridded dataset (NClimGrid), Version 1 [Dataset]. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5SX6B56

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams, C. N., Jr., et al. (2014b). Improved hydrological temperature and precipitation time series for U.S. climate divisions. *Journal of Applied Meteorology and Climatology*, *53*(5), 1232–1251. https://doi.org/10.1175/JAMC-D-13-0248.1

Williams, A. P., Anchukaitis, K. J., Woodhouse, C. A., Meko, D. M., Cook, B. I., Bolles, K., & Cook, E. R. (2021). Tree rings and observations suggests no stable cycles in Sierra Nevada cool-season precipitation. *Water Resources Research*, *57*(3), e2020WR028599. https://doi.org/10.1029/2020WR028599

Williams, A. P., Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles, K., et al. (2020). Large contribution from anthropogenic warming to an emerging North American megadrought. *Science*, *368*(6488), 314–218. https://doi.org/10.1126/science.aaz9600

Wise, E. (2010). Spatiotemporal variability of the precipitation dipole transition zone in the western United States. *Geophysical Research Letters*, *37*(7), L07706. https://doi.org/10.1029/2009GL042193

Wise, E., & Dannenberg, M. P. (2014). Persistence of pressure patterns over North America and the North Pacific since AD 1500. *Nature Communications*, *5*(4912), e1602263. https://doi.org/10.1038/ncomms5912

Wise, E., & Dannenberg, M. P. (2017). Reconstructed storm tracks reveal three centuries of changing moisture delivery to North America. *Science Advances*, *3*(6). https://doi.org/10.1126/sciadv.1602263

Zhang, Z., Ralph, F. M., & Zheng, M. (2019). The relationship between extratropical cyclone strength and atmospheric river intensity and position. *Geophysical Research Letters*, *46*(3), 1814–1823. https://doi.org/10.1029/2018GL079071

Zhu, Y., & Newell, R. (1998). A proposed algorithm for moisture fluxes from atmospheric rivers. *Monthly Weather Review*, *126*(3), 725–735. https://doi.org/10.1175/1520-0493(1998)126%3C0725:APAFMF%3E2.0.CO;2