AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2015JD023205

Key Points:
- Extreme precipitation is spatially coherent in the western United States
- Moisture sources and transport pathways vary regionally and seasonally
- ENSO affects the frequency of extreme events

Supporting Information:
- Figure S2
- Figure S1
- Text S1

Correspondence to:
C. Bracken,
cameron.bracken@colorado.edu

Citation:
Bracken, C., B. Rajagopalan, M. Alexander, and S. Gangopadhyay (2015), Spatial variability of seasonal extreme precipitation in the western United States, *J. Geophys. Res. Atmos.*, *120*, 4522–4533, doi:10.1002/2015JD023205.

Received 3 FEB 2015
Accepted 11 APR 2015
Accepted article online 22 APR 2015
Published online 19 May 2015

Spatial variability of seasonal extreme precipitation in the western United States

C. Bracken[1,2], B. Rajagopalan[1,3], M. Alexander[4], and S. Gangopadhyay[2]

[1]Department of Civil, Environmental and Architectural Engineering, University of Colorado Boulder, Boulder, Colorado, USA, [2]United States Bureau of Reclamation, Denver, Colorado, USA, [3]Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado, USA, [4]National Oceanic and Atmospheric Administration Earth System Research Laboratory, Boulder, Colorado, USA

Abstract We examine the characteristics of 3 day total extreme precipitation in the western United States. Coherent seasonal spatial patterns of timing and magnitude are evident in the data, motivating a seasonally based analysis. Using a clustering method that is consistent with extreme value theory, we identify coherent regions for extremes that vary seasonally. Based on storm back trajectory analysis, we demonstrate unique moisture sources and dominant moisture pathways for each spatial region. In the winter the Pacific Ocean is the dominant moisture source across the west, but in other seasons the Gulf of Mexico, the Gulf of California, and the land surface over the midwestern U.S. play an important role. We find the El Niño–Southern Oscillation (ENSO) to not have a strong impact on dominant moisture delivery pathways or moisture sources. The frequency of extremes under ENSO is spatially coherent and seasonally dependent with certain regions tending to have more (less) frequent extreme events in El Niño (La Niña) conditions.

1. Introduction

The nature of extreme precipitation varies widely across the western United States. For example, the Southwest is most likely to receive extreme precipitation in the summer, while the Pacific Northwest and California typically see the largest extreme events in the winter, and the intermountain west tends to see a more even distribution of occurrence across seasons [*Kunkel et al.*, 1999]. The mean occurrence day of annual maxima for all stations in the western U.S. clearly shows a peak in the winter months (December-January-February (DJF)), but in more than 50% of stations, annual maxima tend to occur between February and July (Figure 1). This motivates a careful examination of how extremes behave both seasonally and spatially. Past studies have classified the behavior of extreme events regionally throughout the western U.S. [*Groisman et al.*, 2001; *Arriaga Ramírez and Cavazos*, 2010; *Dos Santos et al.*, 2011; *Dulière et al.*, 2013; *Mullens et al.*, 2013; *Janssen et al.*, 2014] and for specific seasons [*Warner et al.*, 2012; *Pal et al.*, 2013; *Jiang et al.*, 2014], but few have taken a detailed look at both of these pieces. A seasonal analysis is especially important because while the most likely season for an extreme event may be identified for a specific region, extreme events can occur at any time of the year and the characteristics of these events vary both seasonally and spatially [*Kunkel et al.*, 1999].

Extreme precipitation takes many forms in the western U.S. due to complex topography and climatic interactions. Figure 2 shows the average magnitude within each season and the average occurrence day from the beginning of the season of 3 day extreme events for all stations in the western U.S. (see section 2 for a complete description of the data). A high degree of variability is present spatially and between seasons in both the average timing and magnitude of extreme events in the western U.S. and motivates further investigation. In terms of statistical distributions of extreme events, two weather stations close enough to identically observe all the same events would necessarily have the same extreme value distributions. Two stations farther apart might observe some of the same extreme events, but the events might differ in magnitude, perhaps causing the extreme value distributions to have the similar shape but have slightly different means. Two stations very far apart, not observing any of the same events and having different climatologies, would exhibit different statistical distributions, except by chance. How far need we go such that the extreme value distributions at two weather stations no longer resemble each other and is it possible to create groups of weather stations, or regions, such that the distributions looks sufficiently similar? This grouping, or clustering, is a straightforward way of identifying regions for which extreme values behave similarly [*Hsu and Li*, 2010].

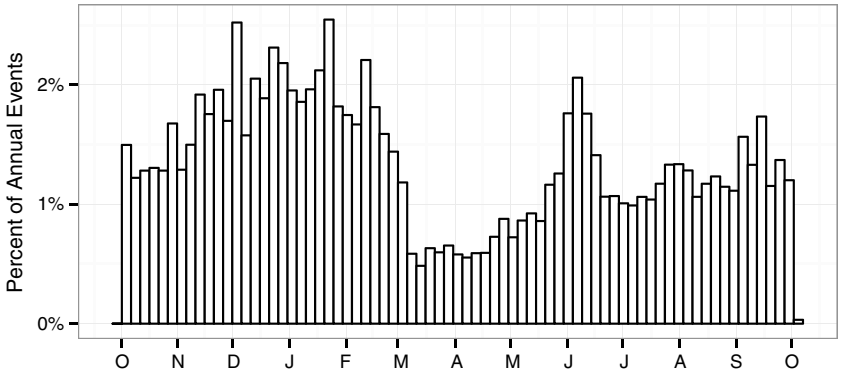

Figure 1. Occurrence day of annual 3 day precipitation maxima at ~14,000 stations in the western U.S. (5 day bins).

Past studies have grouped stations by climate divisions [*Kunkel*, 2003], though these regions are defined for climatological mean and are not appropriate for extreme values [*Jones et al.*, 2014]. Other studies have used grid-based groupings [*Kunkel et al.*, 2003], subjective [*Maraun et al.*, 2008; *Alexander et al.*, 2006], and objective [*Jones et al.*, 2014; *Wigley et al.*, 1984; *Dales and Reed*, 1989; *Neal and Phillips*, 2009] regions. *Jones et al.* [2014] defined regions in the UK using principal component analysis and clustering of a number of extreme value statistics. *DeGaetano* [1998] used a method for defining clusters based on similarity of extreme value cumulative distributions function and spatial proximity. *Bernard et al.* [2013] used a clustering method specifically tailored to the characteristics of extreme value distributions.

Given homogeneous clusters of extreme precipitation, dominant moisture pathways and sources of extreme precipitation can be examined regionally. Moisture pathway and delivery questions can be answered with back trajectory analysis. Back trajectory analysis calculates the pathway that an air parcel followed such that it arrives at the location and time of an observed extreme event. Trajectory analysis has been used

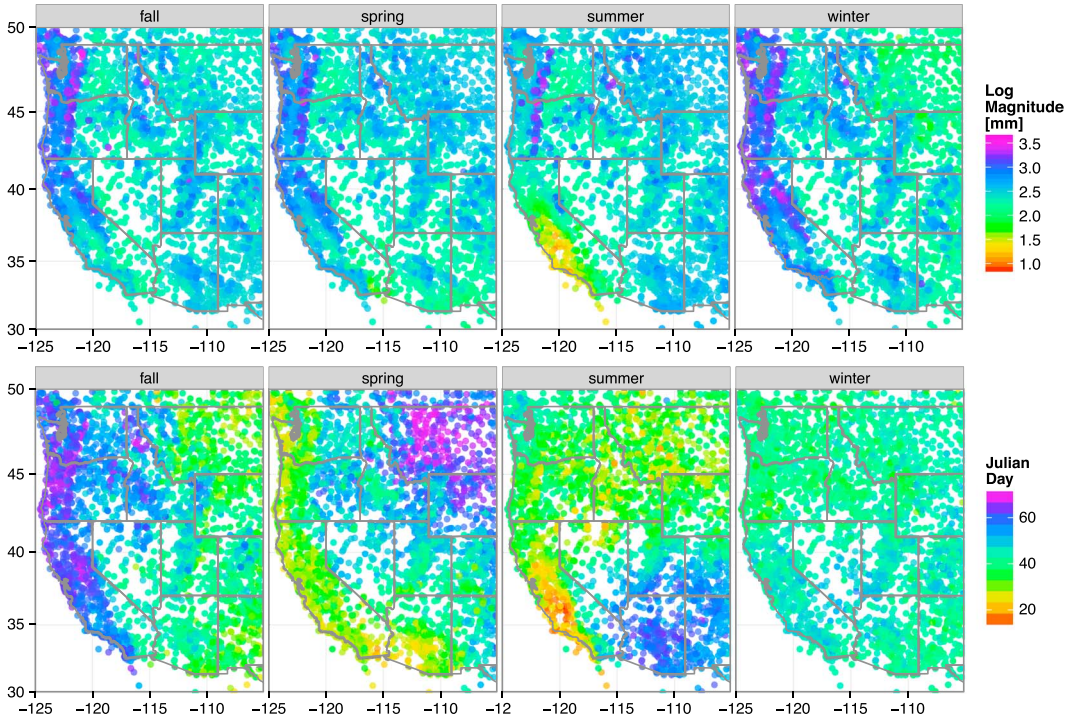

Figure 2. (top) Mean magnitude and (bottom) timing of extreme events in the western U.S. In Figure 2 (top), color indicates magnitude on a log scale. In Figure 2 (bottom), color indicates the average Julian day of occurrence of maxima at a station from the start of the season. For corresponding standard deviation point maps, please see the supporting information.

previously to identify moisture sources and pathways for extreme precipitation [*Gustafsson et al.*, 2010; *Massacand et al.*, 1998; *Reale et al.*, 2001; *Alexander et al.*, 2015]. In this study we investigate the statistically likely moisture sources and pathways for extreme events and how these vary between seasons and between regions [*DeGaetano*, 1998; *Jorba et al.*, 2004]. This is of importance in modeling, simulating, and predicting extremes in space and time and, consequently, for resource management.

In addition to a seasonal analysis, we also seek to understand the impact of the El Niño–Southern Oscillation (ENSO) on dominant moisture pathways and sources in the western U.S. The link between ENSO and extreme events has been previously explored [*Gershunov and Barnett*, 1998; *Cayan et al.*, 1999; *Higgins et al.*, 2010; *Feldl and Roe*, 2011; *DeFlorio et al.*, 2013]. In El Niño conditions, days with high daily precipitation are seen to be more frequent than average over the Southwest and less frequent over the Northwest. During La Niña conditions the signal is typically reversed, but the ENSO events are not all the same and the spatial patterns are complex.

We propose the following research questions related to extreme events in the western U.S.:

1. Can we objectively define coherent extreme value regions and how many regions are appropriate?
2. What are the dominant moisture sources and pathways for each season for these regions?
3. How do seasonal moisture sources pathways change under ENSO regimes?

This paper is structured as follows: We discuss data sources and preprocessing in section 2, methods in section 3, results in section 4, and discussion and conclusions in section 5.

2. Data and Preprocessing

2.1. Precipitation Data

We obtained daily precipitation data from the Global Historical Climatology Network (GHCN, http://www.ncdc.noaa.gov/oa/climate/ghcn-daily/) [*Durre et al.*, 2008, 2010; *Menne et al.*, 2012]. Of the approximately 13,000 stations falling in the study region, we included only those stations with greater than 75% data coverage (at least 50 years of data) from 1948 to 2013, leaving us with 1037 stations.

An extreme precipitation data set was developed using a seasonal block maxima approach. Seasons were defined as 3 month blocks: winter (DJF), spring (March-April-May), summer (June-July-April), and fall (September-October-November). For each station, seasonal 3 day maxima were computed for each year in the record. The extreme events were compiled into a database of approximately 600,000 events.

2.2. Back Trajectories

For each event, back trajectories were computed using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model [*Draxler*, 1999; *Draxler and Hess*, 1998, 1997]. A back trajectory is a three-dimensional pathway taken by a parcel of air such that it arrives at a particular location at the time of a known extreme event. Back trajectories are commonly used in air pollution modeling, but they have also been used for moisture source identification [*Lu et al.*, 2013; *Izquierdo et al.*, 2012; *Gustafsson et al.*, 2010; *Jorba et al.*, 2004; *Stohl and Trickl*, 1999]. HYSPLIT requires gridded climate data to calculate back trajectories, as well as a terminal time and location for the back trajectory. We use the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data [*Kalnay et al.*, 2011].

The upper limit on residence time of moisture in the atmosphere is approximately 8 days [*Trenberth*, 1998]. In addition, for trajectories computed from wind fields, errors of 20% the distance traveled are expected [*Stohl and Seibert*, 1998; *Stohl*, 1998]. Given this uncertainty and the limit on moisture retention, trajectory locations were computed hourly for 8 days preceding the extreme event (192 h). Trajectories were initialized for each event every 6 h during the event period and at 500 m vertical intervals from 0 m to 5000 m above ground level (totaling 151 trajectories per event). The results were compiled into a database of approximately 8,000,000 trajectories.

3. Methods

3.1. Identification of Extremes Regions

Clustering is an intuitive method for grouping multivariate data. In classical *k*-means clustering, clusters are determined so that they minimize the intracluster variance of a metric related to Euclidian distance. For example, if points in the plane are clustered based only on *x* and *y* position, the *k*-means algorithm creates *k* clusters that minimize the variance of the distance between each point in a cluster. Cluster centers are determined

by taking the mean of all cluster member points, and thus, the center of each cluster may not be located at a point in the cluster. The *k*-means method excels at identifying patterns related to mean behavior of Gaussian or Gaussian mixture data, but in the context of extreme values, *k*-means clustering is inappropriate for two reasons [*Bernard et al.*, 2013].

1. With *k*-means, clusters are developed based on the mean of observations within a cluster. With Gaussian data this poses no problem since the mean of Gaussian data remains Gaussian. When applying *k*-means to highly skewed extreme value data, *k*-means centroids are not interpretable, since the mean of generalized extreme value (GEV) data is Gaussian and not GEV. The solution to this issue is to use an alternative clustering algorithm called partitioning around medoids (PAM) [*Kaufman and Rousseeuw*, 1990]. This method creates clusters that are centered around an existing station and requires no averaging and therefore alleviates the interpretation issue.
2. The classical methods of assessing pairwise dependence between time series (i.e., Euclidian distance) are not in compliance with extreme value theory when applied to time series of maxima [*Bernard et al.*, 2013]. The solution to this issue is the use a measure of pairwise dependence called the *F*-madogram, which is explicitly tailored to extremal data. This measure of dependence can be used as the distance matrix in the PAM clustering algorithm, creating an theoretically sound extreme value clustering method.

Further explanation of item 2 above is necessary. Using a variogram called the *F*-madogram, it is possible to construct a measure of pairwise dependence that is appropriate for extreme values [*Naveau et al.*, 2009; *Cooley et al.*, 2006]. Given T samples of bivariate maxima $(M_i^{(t)}, M_j^{(t)})^T$ from two locations i and j, the nonparametric estimator for the *F*-madogram is given by *Bernard et al.* [2013] as

$$\hat{d}_{ij} = \frac{1}{2T}\sum_{t=1}^{T}\left|\hat{F}_i(M_i^{(t)}) - \hat{F}_j(M_j^{(t)})\right|, \tag{1}$$

where

$$\hat{F}_i(u) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{\{M_i^{(t)}\le u\}}, \tag{2}$$

where $\mathbf{1}_{\{M_i^{(t)}\le u\}}$ is the indicator function for the event $\{M_i^{(t)} \le u\}$ which returns 1 if the statement is true or 0 otherwise. The entire function returns a proportion of the number of data points that are less than or equal to a given value u (the empirical cumulative distribution function). It is important to note that the *F*-madogram does not depend on the magnitude of extreme events, providing a dimensionless metric that compares the shape of the extreme value distributions between two stations. Coupled with the PAM algorithm, the *F*-madogram provides an efficient, nonparametric, and theoretically sound method for clustering maxima.

Bernard et al. [2013] applied this extreme value-oriented clustering method to station data covering France with favorable results. Since this method is based on only the shape of the extreme value distribution at a station and not on magnitude of event or geographic proximity, as physical extent of a region increases, so does the chance of misclassification of a station. In other words, extreme value distributions from two geographically disparate stations may look the same for different physical reasons. To address this issue, we propose an extension to the clustering algorithm that also incorporates physical proximity of stations. The extension involves computing a modified version of the madogram

$$\hat{\hat{d}}_{ij} = \hat{d}_{ij} + p_{ij}, \tag{3}$$

where

$$p_{ij} = \frac{q_{ij}}{\sum_{i=1}^{N} q_{ij}}\max_{ij}\hat{d}_{ij} \tag{4}$$

and

$$q_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{5}$$

The computation of p_{ij} is simply the Euclidian distance between the station locations scaled such that they will never exceed the largest value of the original *F*-madogram. Note that the Euclidian distance formula may be replaced with the Haversine distance formula if the original data are not projected.

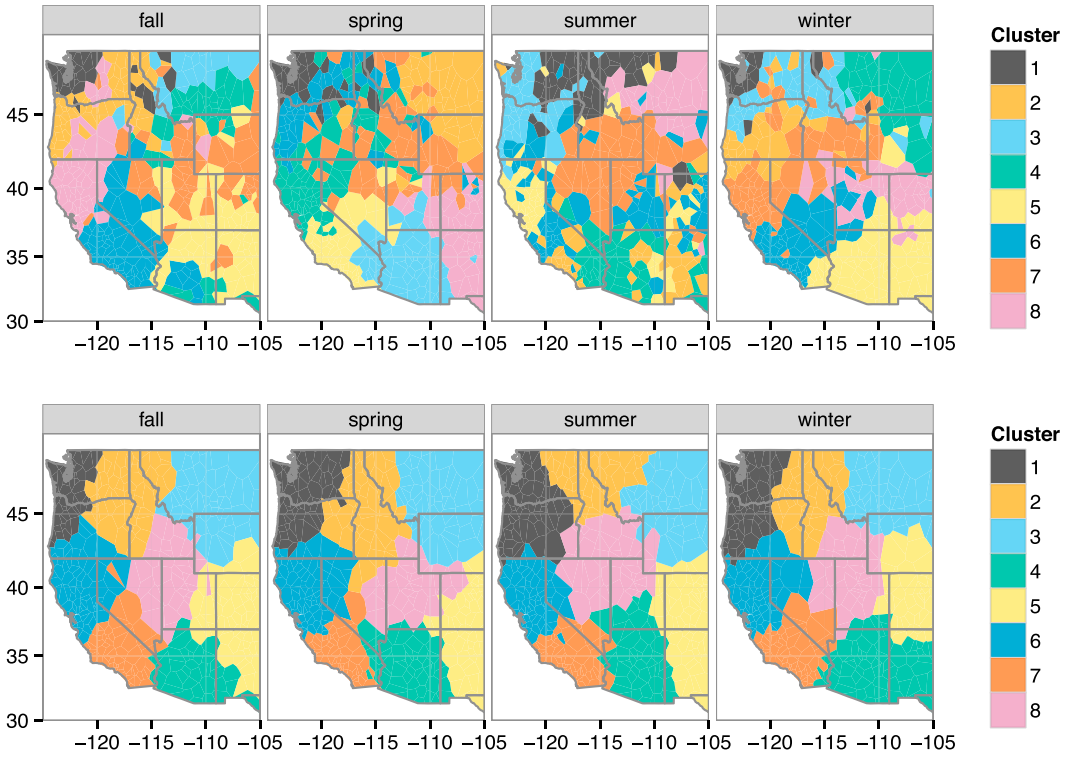

Figure 3. (top) Extreme regions defined by the clustering method of *Bernard et al.* [2013] and (bottom) regions defined with the modified extremes clustering method for large geographic regions.

We present a comparison of the original clustering algorithm of *Bernard et al.* [2013] (Figure 3, top) and the same algorithm with our improvements (Figure 3, bottom). Both graphics use only stations with near-complete record from 1950 to 2014 and eight clusters are specified. In the original algorithm, few clusters are physically contiguous. We still see similar clustering regions defined in both algorithms, but the inclusion of a weighted distance matrix into the clustering algorithm greatly improved the cluster coherence.

As a final step in determining extreme regions, we extended the clusters based on point data to cover the entire study region, so that any stations that did not have sufficient data to be included in the clustering can be associated with a cluster. This was achieved by using a Dirichlet tessellation (or Voronoi tessellation) of the points included in the clustering algorithm, attributing all the area closest to a station as part of that station's cluster. Note that a region need not be contiguous, though they typically are.

3.1.1. Choosing the Number of Clusters

For the PAM clustering method, the metric typically used for choosing a relevant number of clusters is called the silhouette coefficient [*Rousseeuw*, 1987]. Though several metrics exist for aiding in the determination of the number of clusters required in a nonhierarchical clustering algorithm, the choice always comes down to an application-based subjective decision [*DeGaetano*, 2001]. For our application we wanted to choose the number to be the same for all seasons for the purposes of comparing across seasons, near-optimal in terms of the silhouette coefficient, reasonably resolute, and manageable from a data analysis perspective. The choice of eight clusters satisfied these criteria, yet any number of clusters from 6 to 11 would have been reasonable. Clusters defined with this method exhibit self similarity in that the same clustering method could be applied again to a single region, further defining more resolute regions.

3.2. Moisture Source Identification

To identify moisture sources, we started by filtering trajectories by those that produced precipitation. We included 100 trajectories per station from each season that (1) produced the greatest decrease in specific humidity along a trajectory during the event period and (2) dropped below 1000 m at some point during the 8 day window. We required a height threshold so that we excluded trajectories that did not pick up moisture during the 8 day window. The loss of specific humidity coupled with the observed maxima provides a high

degree of confidence that the selected subset of trajectories produced precipitation at the surface. We refer to this subset as the "rain trajectories."

The source region identification method is somewhat imprecise, technically providing the end of the source region. *Gustafsson et al.* [2010] identified the beginning of a source region as the first time the specific humidity of a trajectory drops below its ending specific humidity. We found this criteria difficult to apply since in many cases our ending specific humidity was close to zero. Given the inaccuracy of a single trajectory, using the point where maximum specific humidity occurred was a sufficiently good indicator of source region location.

From the set of rain trajectories, we identified a moisture source location by the point at which a trajectory achieved its maximum specific humidity. Minimum physical height would have also been a reasonable indicator; initial exploration indicated that these two indicators gave nearly identical results. Computing moisture source locations for each rain trajectory and binning the resulting locations on a 1° grid, we get a picture of moisture source regions for a given cluster and season.

3.3. Trajectory Clustering

Trajectory clustering is a well-known statistical method for analyzing moisture transport patterns from trajectory data [*Moody and Galloway*, 1988; *Moody*, 1986; *Stohl*, 1998; *Dorling et al.*, 1992]. The method is an application of the *k*-means clustering algorithm to longitudinal data [*Genolini and Falissard*, 2010] and is implemented in the KmL package [*Genolini and Falissard*, 2011]. We used the Haversine distance formula for calculating intracluster root-mean-squared error.

We fixed the number of clusters in each region and season to 10—the aim was to explore groups produced with a fixed number of clusters. For each group we determined an average trajectory as the mean of all the trajectory locations at a given time.

To investigate the ENSO connection, we identified events based on the multivariate ENSO index (MEI) [*Wolter and Timlin*, 1998, 1993] (http://www.esrl.noaa.gov/psd/enso/mei/). A season was identified as having strong El Niño or La Niña conditions if the MEI value was greater than 1 or less than −1, respectively. Trajectories were composited based on this ENSO classification to identify the teleconnections.

4. Results

4.1. Cluster Descriptions

Clusters were similar between seasons but not identical—the boundaries shift seasonally indicating differences in extreme distributions across seasons. Remarkably, the centroids of each region fell near the same location for each cluster (Figure 3, bottom). This is coincidental since each season is clustered independently, but we leveraged this to describe the regions across seasons. Each region is prefixed with a label that will be used to refer to the region. (1) (PNW) Western Oregon and Washington roughly extending east to the Cascade Range; (2) (PNE) Eastern Oregon and Washington and Western and Northern Idaho and far Western Montana; (3) (MT) Eastern Montana and Northern Wyoming; (4) (AZ) Arizona and in winter, New Mexico; (5) (CB) Western Colorado, Eastern Utah, Southern/Central Wyoming, and New Mexico (in fall, spring, and summer only); (6) (NC) Northern California, Southern Oregon, and Western Nevada; (7) (SC) Southern California and Southern Nevada; and (8) (GB) Great Basin watershed including Western Utah, Eastern Nevada, and Southern Idaho.

4.2. Moisture Sources

Figure 4 shows moisture source regions for each cluster and season. In winter, not surprisingly, more than 98% of moisture sources for extremes are in the Central to Eastern Pacific Ocean (including the Gulf of California). In coastal clusters (PNW, NC, and SC) moisture tends to be picked up farther from the coast than inland clusters (PNE, MT, AZ, CB, and GB).

In summer a few patterns stand out. Moisture is generated in a band of the northwest Pacific that can extend to the far western Pacific (PNW, PNE, MT, NC, SC, and GB). This pattern can be seen in all clusters except for the southwest clusters (AZ and CB). In southwest clusters (AZ and CB), the near-coast eastern Pacific, the Gulf of California, and the Gulf of Mexico appear as dominant moisture sources. A somewhat surprising result is the dominance of the land as a moisture source for inland clusters (PNE and MT), likely due to a combination of moisture transport and moisture recycling [*Bosilovich and Schubert*, 2001; *Dirmeyer et al.*, 2014]. The inland cluster GB exhibits a mixture of all the summer patterns.

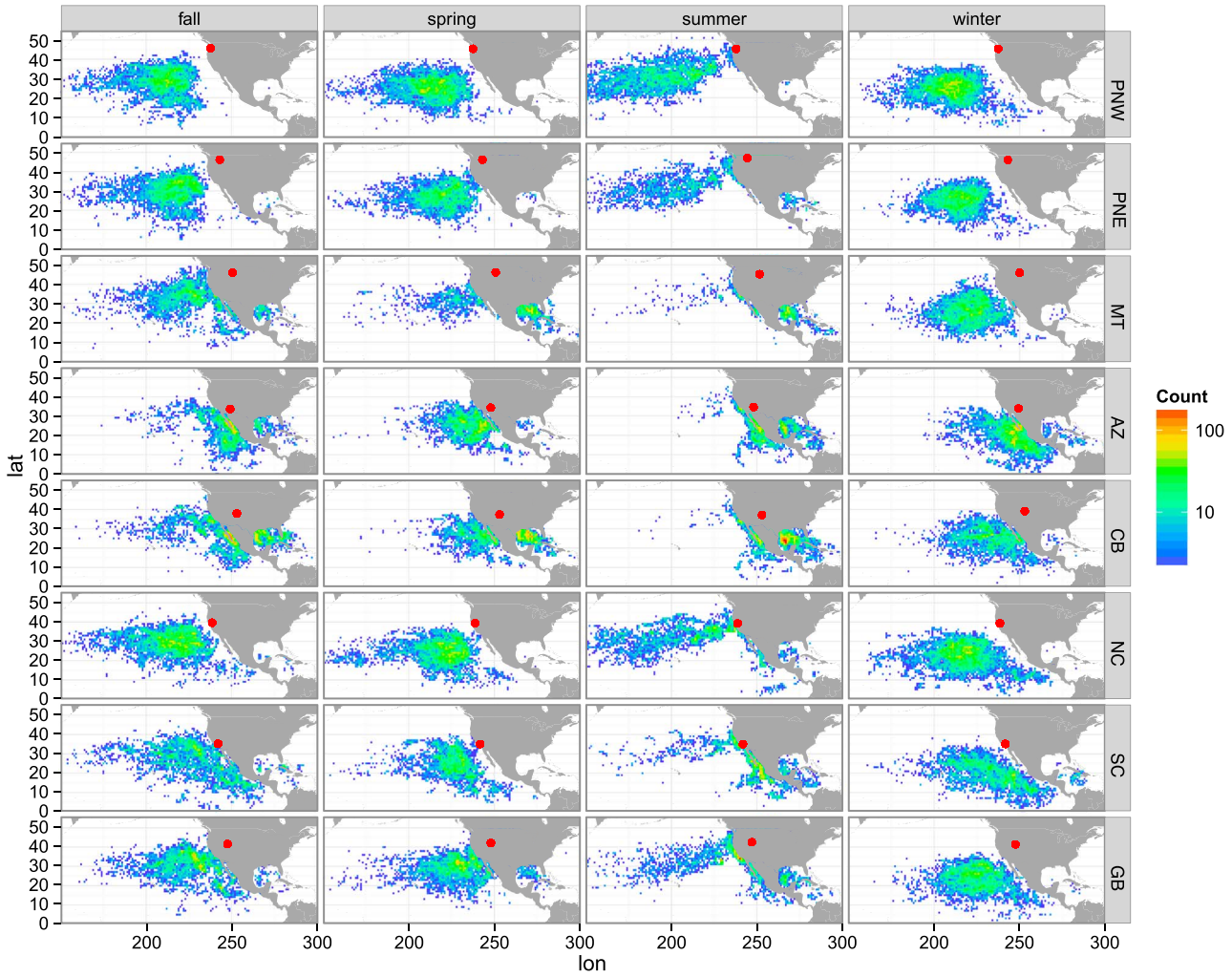

Figure 4. Moisture source location counts for each season and region, binned on a 1° grid. Red dots indicate the center of each extreme region.

Spring and fall can be viewed as transitionary seasons between summer and winter. All of the previously described patterns are present in varying degrees. In coastal clusters (PNW and NC), due to the direct influence of the Pacific Ocean, we found the transition between summer and winter to be much less strong than in inland clusters (MT, for example).

4.3. Moisture Pathways

Figure 5 shows the dominant pathways for extremes in the western U.S., and Table 1 shows the percentage of trajectories included in each trajectory group. Winter pathways typically run from the winter moisture sources, i.e., the central to eastern Pacific, directly onto land for all clusters. Many winter pathways for coastal regions (PNW, NC, and SC) follow well-known atmospheric river (AR) patterns [*Newell et al.*, 1992; *Zhu and Newell*, 1998; *Ralph and Dettinger*, 2011; *Neiman et al.*, 2011; *Villarini et al.*, 2013; *Alexander et al.*, 2015]. With few exceptions, pathways for the coastal regions (PNW, NC, and SC) mimic winter, while the more inland regions (PNE, MT, AZ, CB, and GB) display pathways from the Gulf of Mexico, north over land, and from the Gulf of California. Summer pathways display the most diversity; for example, regions MT and CB display numerous distinct pathways. Fall and spring pathways appear as a transition from summer to winter, as with the moisture sources.

4.4. ENSO Connections to Extremes

From a back trajectory perspective, we examined the effects of ENSO on the frequency of extremes. Figure 6 shows the ratio of trajectories occurring in La Niña versus El Niño years. A value greater than 1 (blue) indicates that a station is Niña dominated (experiences relatively more strong events in La Niña conditions), and a value less than 1 (red) indicates that a station is Niño dominated (experiences more events in El Niño conditions).

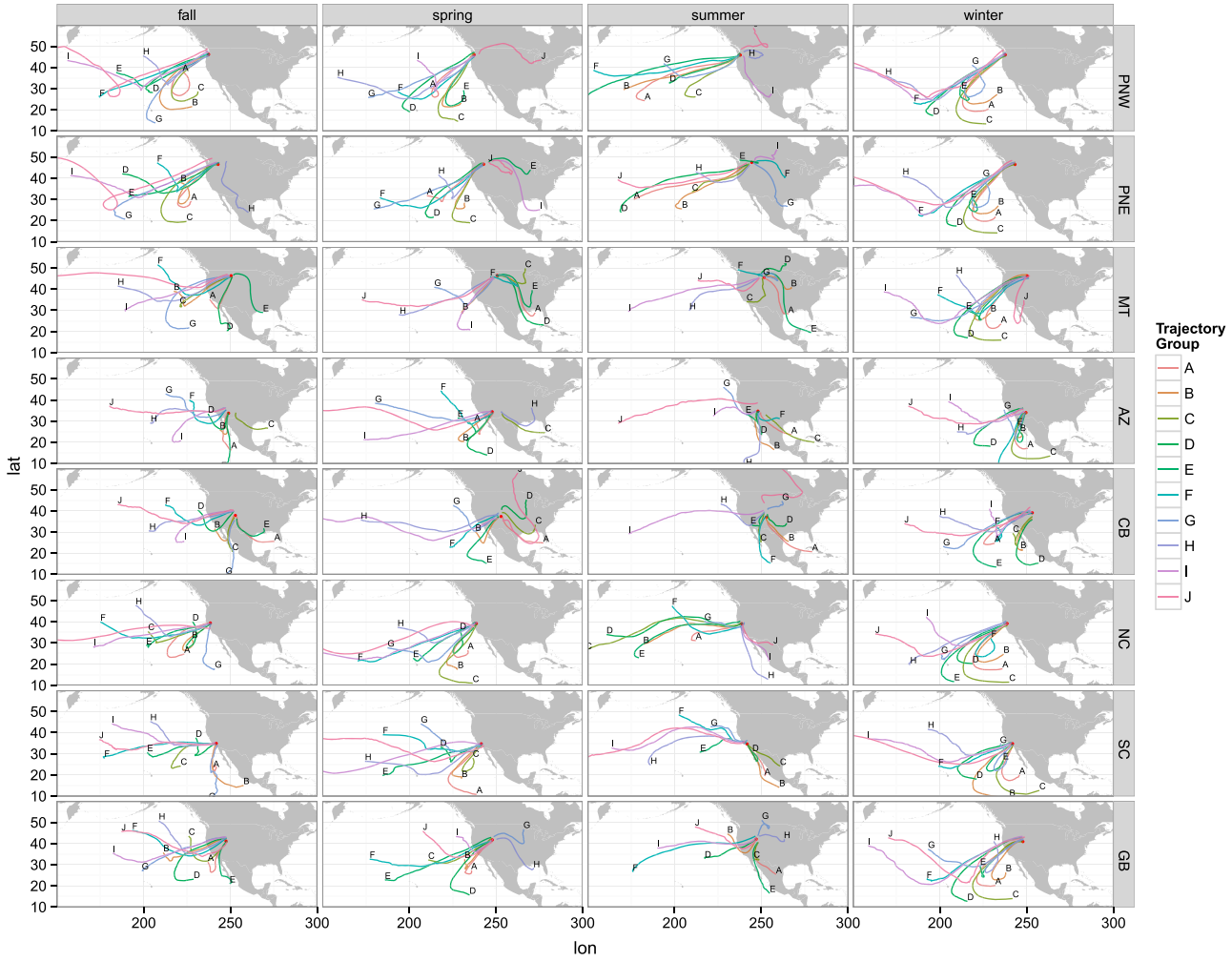

Figure 5. Dominant pathways (determined by trajectory clustering) by region and season.

A north/south divide can be seen in both summer and winter, the north being predominately Niña dominated and the south being predominantly Niño dominated. The signal is more mixed in the fall and spring. Southern California is always Niño dominated, and western Oregon is always Niña dominated. Similar north-south dipoles have been documented in *Cayan et al.* [1999] and *Feldl and Roe* [2011] although a direct comparison was not possible because of dissimilar time windows.

We did not find ENSO to have a strong effect on moisture source locations. We found ENSO to affect trajectory pathways but typically only prior to a moisture source location, indicative of ENSO's influence on larger atmospheric circulation patterns. Moisture source locations and delivery pathways from moisture sources to event locations remained essentially unchanged. See supporting information for additional graphics.

5. Discussion and Conclusions

We have presented a detailed analysis of 3 day total extreme precipitation throughout the western United States. Using an improved extreme value clustering method, we defined spatially coherent regions for extremes that vary seasonally and made physical sense based on topography. Based on storm back trajectory analysis, we were able to demonstrate unique moisture sources and dominant moisture pathways for each spatial region.

Winter and summer extremes display distinct patterns of moisture delivery, timing, and magnitude throughout the western U.S. In these terms, winter extremes behave mostly uniform across the west, while summer

Table 1. Percentage of Trajectories Included in Each Group by Region and Season

		Trajectory Group									
Season	Region	A	B	C	D	E	F	G	H	I	J
Fall	PNW	20.53	16.26	14.15	9.70	9.59	9.25	7.67	7.39	5.29	0.18
Fall	PNE	15.95	15.25	14.27	13.41	10.36	7.87	5.93	5.69	5.63	5.62
Fall	MT	19.17	16.34	12.93	10.30	10.19	9.19	7.05	5.84	4.89	4.10
Fall	AZ	21.23	18.33	15.77	13.96	9.31	7.24	6.70	5.44	1.85	0.16
Fall	CB	23.43	18.25	15.57	12.43	12.23	7.62	4.45	4.24	1.77	0.01
Fall	NC	22.21	12.45	12.42	10.57	10.38	8.43	7.56	7.37	5.19	3.41
Fall	SC	28.72	15.63	11.68	11.01	10.95	7.89	5.85	5.06	3.19	0.00
Fall	GB	21.81	15.55	15.03	10.09	9.07	8.06	7.67	5.29	4.31	3.12
Spring	PNW	16.75	16.74	11.99	11.72	10.75	9.86	9.59	5.71	4.86	2.02
Spring	PNE	18.07	17.08	14.87	10.43	8.60	8.50	8.26	5.35	5.11	3.73
Spring	MT	16.55	14.99	13.91	10.42	10.24	9.56	9.28	9.01	4.10	1.94
Spring	AZ	27.27	21.96	17.72	9.55	6.08	4.82	4.64	3.41	2.91	1.63
Spring	CB	20.31	15.58	13.51	12.89	10.16	9.64	6.32	5.06	3.77	2.75
Spring	NC	20.75	18.83	12.50	11.14	11.02	10.93	6.99	5.06	2.77	0.00
Spring	SC	20.27	15.97	14.14	9.97	9.20	8.16	7.99	7.72	4.46	2.12
Spring	GB	23.07	16.63	13.23	10.77	8.95	7.76	7.66	5.35	3.56	3.02
Summer	PNW	26.50	16.25	15.53	10.21	8.94	8.68	7.74	2.65	2.44	1.05
Summer	PNE	25.28	17.05	14.03	9.53	8.64	7.40	6.34	6.12	3.30	2.32
Summer	MT	25.31	18.82	16.11	13.57	8.92	5.64	4.72	4.52	1.53	0.88
Summer	AZ	19.23	17.33	15.12	14.87	12.77	9.18	5.95	3.42	1.75	0.39
Summer	CB	23.14	14.54	14.24	12.29	10.77	9.65	7.36	6.07	1.91	0.00
Summer	NC	15.47	15.38	13.01	12.09	9.55	9.45	8.50	6.86	5.18	4.50
Summer	SC	21.38	20.65	14.62	11.04	9.30	6.76	5.48	4.09	3.58	3.11
Summer	GB	22.12	19.38	15.53	10.38	10.04	9.98	6.98	5.34	0.00	0.00
Winter	PNW	21.90	13.95	12.86	11.65	11.39	11.27	8.07	6.54	2.17	0.19
Winter	PNE	21.53	18.51	12.54	11.81	9.06	8.52	8.01	4.67	3.62	1.72
Winter	MT	20.12	16.71	13.66	11.94	9.33	8.69	5.78	4.85	4.81	4.10
Winter	AZ	21.96	17.69	16.06	15.49	11.20	8.98	2.95	2.87	2.79	0.01
Winter	CB	23.33	15.76	15.45	13.84	12.95	6.68	6.50	2.48	2.43	0.58
Winter	NC	18.81	16.69	16.48	14.49	9.34	8.74	6.63	4.51	3.11	1.20
Winter	SC	17.86	17.24	16.08	15.36	10.92	7.38	6.40	4.61	2.89	1.25
Winter	GB	25.45	20.81	18.04	11.40	7.61	6.29	4.30	3.81	1.96	0.33

extremes display a much higher degree of spatial variability. Fall and spring manifest as transition periods between these two regimes. In addition to the Pacific Ocean (including the Gulf of California), the Gulf of Mexico appears as a moisture source predominantly for inland clusters in fall, spring, and summer. For northeastern regions mostly covering Wyoming and Montana, the land surface is a dominant moisture source. Coastal regions including Northern California, Oregon, and Washington source nearly all moisture directly from the Pacific Ocean. In the summer, southwest regions including Southern California, Arizona, New Mexico, and western Colorado tend to have multiple dominant moisture pathways from both the Gulf of Mexico, Gulf of California, and the Pacific Ocean.

ENSO had little effect on moisture sources; moisture pathways were typically only affected prior to moisture source locations. The occurrence frequency of extremes under ENSO conditions is spatially coherent and seasonally dependent with certain regions tending to have more or less extreme events depending on El Niño or La Niña conditions.

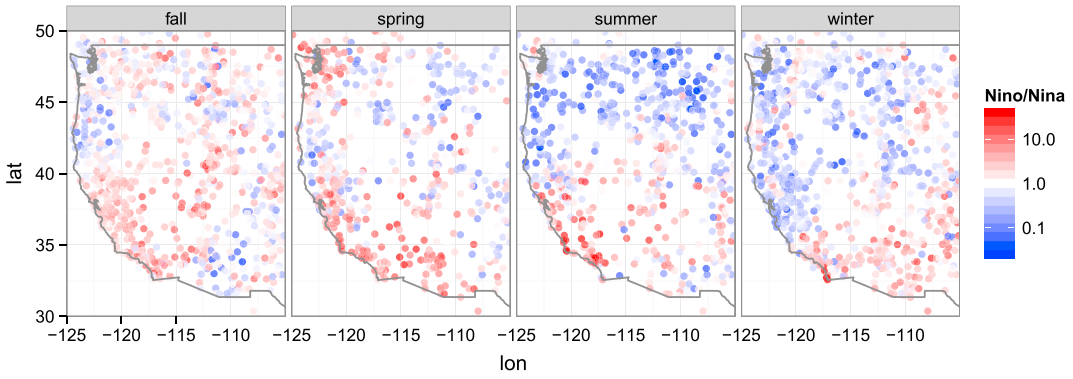

Figure 6. Ratio of the number of rain trajectories occurring in La Niña versus El Niño years. For points in red, more rain trajectories occurred in El Niño seasons, and for points in blue, more rain trajectories occurred in La Niña seasons.

To more precisely determine moisture source locations, the method for determining source region locations could be improved, although using higher resolution reanalysis data would likely be a better fist step. We were not able to determine the proportion of moisture that is gained from land versus open water, an apportionment which could be useful in more detailed studies. In addition, when multiple moisture sources are present for a single region, a more detailed analysis would be required to determine the degree to which certain moisture sources are active in ENSO conditions.

By only including trajectories with moisture sources during the 8 day back trajectory, we excluded trajectories that started the 8 day period with high specific humidity. There are likely extreme events produced by air that gains moisture prior to the 8 day window, but we did not have the means to identify those sources.

The NCEP reanalysis has a very coarse (2.5°) resolution and thus details of the pathways, such as the specific paths through the mountains, are not well resolved. For this study, the level of detail was sufficient, but we could identify higher resolution features such as specific pathway through mountains [*Alexander et al.*, 2015]. The use of newer (and shorter duration) reanalysis such as the ERA-Interim [*Dee et al.*, 2011], the Climate Forecast System Reanalysis [*Saha et al.*, 2010], the Modern-era Retrospective Analysis for Research [*Bosilovich et al.*, 2012], or the North American Regional Reanalysis [*Mesinger et al.*, 2006] would allow us to examine pathways in more detail.

This analysis opens the door to many new applications. By sampling the trajectory data conditioned on features such as current location of trajectories and sea surface temperatures, short-term projections of spatially coherent extremes can be made. This resampling is akin to generating hurricane tracks [for example, *Yonekura and Hall*, 2011]. The trajectory projections can be coupled with spatial extremes model [for example, *Cooley et al.*, 2007] to produce return level maps of extreme precipitation, and consequently hydrologic extremes, important for resource management. Furthermore, these trajectories can be resampled to produce physically based simulations of extreme precipitation. We can also investigate the climatic conditions in moisture source regions that favor the production of extremes, potentially to develop statistical forecast models. Finally, given the evidence of the land surface as a moisture source for some regions, we would like to investigate the relative contribution of the land surface versus open water bodies as a moisture source for extreme precipitation.

Acknowledgments
Funding for this research by a Science and Technology grant from Bureau of Reclamation is gratefully acknowledged. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver, and the National Center for Atmospheric Research. The authors are thankful for support from Janus supercomputer staff at the University of Colorado. Analysis was conducted using the R language [*R Core Team*, 2014]. GHCN daily data were obtained from the NCDC website (https://www.ncdc.noaa.gov/oa/climate/ghcn-daily/). The NCEP/NCAR reanalysis used in this study is available in HYSPLIT format from the NOAA ARL ftp server (ftp://arlftp.arlhq.noaa.gov/pub/archives/reanalysis). Observed data and computed trajectories are available at http://bechtel.colorado.edu/~bracken/western-us-extremes-data/.

References

Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, *111*, D05109, doi:10.1029/2005JD006290.

Alexander, M. A., J. D. Scott, D. Swales, M. Hughes, K. Mahoney, and C. A. Smith (2015), Moisture pathways into the US Intermountain West Associated with heavy winter precipitation events, *J. Hydrometeorol.*, doi:10.1175/JHM-D-14-0139.1.

Arriaga Ramírez, S., and T. Cavazos (2010), Regional trends of daily precipitation indices in northwest Mexico and southwest United States, *J. Geophys. Res.*, *115*, D14111, doi:10.1029/2009JD013248.

Bernard, E., P. Naveau, M. Vrac, and O. Mestre (2013), Clustering of maxima: Spatial dependencies among heavy rainfall in France, *J. Clim.*, *26*, 7929–7937, doi:10.1175/JCLI-D-12-00836.1.

Bosilovich, M. G., and S. D. Schubert (2001), Precipitation recycling over the central United States diagnosed from the GEOS-1 data assimilation system, *J. Hydrometeorol.*, *2*, 26–35, doi:10.1175/1525-7541(2001)002<0026:PROTCU>2.0.CO;2.

Bosilovich, M. G., M. Rixen, P. van Oevelen, G. Asrar, G. Compo, A. Simmons, and K. Trenberth (2012), *Report of the 4th World Climate Research Programme International Conference on Reanalyses*, NASA Global Modeling and Assimilation Office, Silver Spring, Md.

Cayan, D. R., K. T. Redmond, and L. G. Riddle (1999), ENSO and hydrologic extremes in the western United States*, *J. Clim.*, *12*(9), 2881–2893.

Cooley, D., P. Naveau, and P. Poncet (2006), Variograms for spatial max-stable random fields, in *Dependence in Probability and Statistics*, vol. 187, pp. 373–390, Springer, New York.
Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels, *J. Am. Stat. Assoc.*, *102*(479), 824–840.
Dales, M. Y., and D. W. Reed (1989), *Regional Flood and Storm Hazard Assessment*, NERC Open Research Archive, Wallingford, Oxfordshire.
Dee, D. P., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597.
DeFlorio, M. J., D. W. Pierce, D. R. Cayan, and A. J. Miller (2013), Western U.S. extreme precipitation events and their relation to ENSO and PDO in CCSM4, *J. Clim.*, *26*(12), 4231–4243.
DeGaetano, A. T. (1998), A Smirnov test-based clustering algorithm with application to extreme precipitation data, *Water Resour. Res.*, *34*, 169–176, doi:10.1029/97WR03133.
DeGaetano, A. T. (2001), Spatial grouping of United States climate stations using a hybrid clustering approach, *Int. J. Climatol.*, *21*, 791–807.
Dirmeyer, P. A., J. Wei, and M. G. Bosilovich (2014), Comparing evaporative sources of terrestrial precipitation and their extremes in MERRA using relative entropy, *J. Hydrometeorol.*, *15*, 102–116.
Dorling, S. R., T. D. Davies, and C. E. Pierce (1992), Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—Method and applications, *Atmos. Environ. Part A*, *26*(14), 2575–2581.
Dos Santos, C., C. Neale, and T. Rao (2011), Trends in indices for extremes in daily temperature and precipitation over Utah, USA, *Int. J. Climatol.*, *31*, 1813–1822, doi:10.1002/joc.2205.
Draxler, R. R., and G. D. Hess (1997), Description of the HYSPLIT-4 modeling system, *NOAA Tech. Mem. ERL ARL-224*, Natl. Oceanic and Atmos. Admin., Silver Spring, Md.
Draxler, R. R., and G. D. Hess (1998), An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition, *Aust. Meteorol. Mag.*, *47*, 295–308.
Draxler, R. R. (1999), HYSPLIT4 user's guide, *NOAA Tech. Memo. ERL ARL-230*, NOAA Air Resources Laboratory, Silver Spring, Md.
Dulière, V., Y. Zhang, and E. P. Salathé Jr. (2013), Changes in twentieth-century extreme temperature and precipitation over the western United States based on observations and regional climate model simulations*, *J. Clim.*, *26*(21), 8556–8575.
Durre, I., M. J. Menne, and R. S. Vose (2008), Strategies for evaluating quality assurance procedures, *J. Appl. Meteorol. Climatol.*, *47*(6), 1785–1791.
Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose (2010), Comprehensive automated quality assurance of daily surface observations, *J. Appl. Meteorol. Climatol.*, *49*(8), 1615–1633.
Feldl, N., and G. H. Roe (2011), Climate variability and the shape of daily precipitation: A case study of ENSO and the American West, *J. Clim.*, *24*(10), 2483–2499.
Genolini, C., and B. Falissard (2010), KmL: K-means for longitudinal data, *Comput. Stat.*, *25*(2), 317–328.
Genolini, C., and B. Falissard (2011), Kml: A package to cluster longitudinal data, *Comput. Methods Programs Biomedicine*, *104*(3), e112–e121.
Gershunov, A., and T. P. Barnett (1998), ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Observations and model results, *J. Clim.*, *11*(7), 1575–1586.
Groisman, P. Y., R. W. Knight, and T. R. Karl (2001), Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century, *Bull. Am. Meteorol. Soc.*, *82*(2), 219–246.
Gustafsson, M., D. Rayner, and D. Chen (2010), Extreme rainfall events in southern Sweden: Where does the moisture come from?, *Tellus A*, *62*(5), 605–616.
Higgins, R. W., J. -K. E. Schemm, W. Shi, and A. Leetmaa (2010), Extreme precipitation events in the western United States related to tropical forcing, *J. Clim.*, *13*(4), 793–820.
Hsu, K. C., and S. T. Li (2010), Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network, *Adv. Water Resour.*, *33*(2), 190–200.
Izquierdo, R., A. Avila, and M. Alarcón (2012), Trajectory statistical analysis of atmospheric transport patterns and trends in precipitation chemistry of a rural site in NE Spain in 1984–2009, *Atmos. Environ.*, *61*, 400–408.
Janssen, E., D. J. Wuebbles, K. E. Kunkel, S. C. Olsen, and A. Goodman (2014), Observational- and model-based trends and projections of extreme precipitation over the contiguous United States, *Earth's Future*, *2*(2), 99–113.
Jiang, T., K. J. Evans, Y. Deng, and X. Dong (2014), Intermediate frequency atmospheric disturbances: A dynamical bridge connecting western U.S. extreme precipitation with East Asian cold surges, *J. Geophys. Res. Atmos.*, *119*, 3723–3735, doi:10.1002/2013JD021209.
Jones, M. R., S. Blenkinsop, and H. J. Fowler (2014), Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall, *Int. J. Climatol.*, *34*, 751–765.
Jorba, O., C. Pérez, F. Rocadenbosch, and J. M. Baldasano (2004), Cluster analysis of 4-day back trajectories arriving in the Barcelona area, Spain, from 1997 to 2002, *J. Appl. Meteorol.*, *43*(6), 887–901.
Kalnay, E., et al. (2011), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*(3), 437–471.
Kaufman, L., and P. Rousseeuw (1990), *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York.
Kunkel, K. E. (2003), North American trends in extreme precipitation, *Nat. Hazards*, *29*(2), 291–305.
Kunkel, K. E., K. Andsager, and D. R. Easterling (1999), Long-term trends in extreme precipitation events over the conterminous United States and Canada, *J. Clim.*, *12*(8), 2515–2527.
Kunkel, K. E., D. R. Easterling, K. Redmond, and K. Hubbard (2003), Temporal variations of extreme precipitation events in the United States: 1895–2000, *Geophys. Res. Lett.*, *30*(17), 1900, doi:10.1029/2003GL018052.
Lu, M., U. Lall, A. Schwartz, and H. Kwon (2013), Precipitation predictability associated with tropical moisture exports and circulation patterns for a major flood in France in 1995, *Water Resour. Res.*, *49*(10), 6381–6392, doi:10.1002/wrcr.20512.
Maraun, D., T. J. Osborn, and N. P. Gillett (2008), United Kingdom daily precipitation intensity: Improved early data, error estimates and an update from 2000 to 2006, *Int. J. Climatol.*, *28*(6), 833–842, doi:10.1002/joc.1672.
Massacand, A. C., H. Wernli, and H. C. Davies (1998), Heavy precipitation on the Alpine southside: An upper-level precursor, *Geophys. Res. Lett.*, *25*, 1435–1438.
Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An overview of the global historical climatology network-daily database, *J. Atmos. Oceanic Technol.*, *29*(7), 897–910.
Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360.
Moody, J. L. (1986), The influence of meteorology on precipitation chemistry at selected sites in the Eastern United States, PhD thesis, The Univ. of Michigan, Ann Arbor.
Moody, J. L., and J. N. Galloway (1988), Quantifying the relationship between atmospheric transport and the chemical composition of precipitation on Bermuda, *Tellus B*, *40*(5), 463–479.

Mullens, E. D., M. Shafer, and J. Hocker (2013), Trends in heavy precipitation in the southern USA, *Weather*, *68*(12), 311–316.
Naveau, P., A. Guillou, D. Cooley, and J. Diebolt (2009), Modelling pairwise dependence of maxima in space, *Biometrika*, *96*(1), 1–17.
Neal, R. A., and I. D. Phillips (2009), Summer daily precipitation variability over the East Anglian region of Great Britain, *Int. J. Climatol.*, *29*, 1661–1679, doi:10.1002/joc.1826.
Neiman, P. J., L. J. Schick, F. M. Ralph, M. Hughes, and G. A. Wick (2011), Flooding in Western Washington: The connection to atmospheric rivers*, *J. Hydrometeorol.*, *12*(6), 1337–1358.
Newell, R. E., N. E. Newell, Y. Zhu, and C. Scott (1992), Tropospheric rivers? A pilot study, *Geophys. Res. Lett.*, *19*(24), 2401–2404.
Pal, I., B. T. Anderson, G. D. Salvucci, and D. J. Gianotti (2013), Shifting seasonality and increasing frequency of precipitation in wet and dry seasons across the U.S., *Geophys. Res. Lett.*, *40*(15), 4030–4035, doi:10.1002/grl.50760.
R Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna.
Ralph, F. M., and M. D. Dettinger (2011), Storms, floods, and the science of atmospheric rivers, *Eos Trans. AGU*, *92*(32), 265–266.
Reale, O., L. Feudale, and B. Turato (2001), Evaporative moisture sources during a sequence of floods in the Mediterranean region, *Geophys. Res. Lett.*, *28*(10), 2085–2088.
Rousseeuw, P. J. (1987), Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.*, *20*, 53–65.
Saha, S., et al. (2010), The NCEP climate forecast system reanalysis, *J. Am. Stat. Assoc.*, *91*(8), 1015–1057.
Stohl, A. (1998), Computation, accuracy and applications of trajectories—A review and bibliography, *Atmos. Environ.*, *32*(6), 947–966.
Stohl, A., and P. Seibert (1998), Accuracy of trajectories as determined from the conservation of meteorological tracers, *Q. J. R. Meteorol. Soc.*, *124*(549), 1465–1484.
Stohl, A., and T. Trickl (1999), A textbook example of long-range transport: Simultaneous observation of ozone maxima of stratospheric and North American origin in the free troposphere over Europe, *J. Geophys. Res.*, *104*(D23), 30,445–30,462.
Trenberth, K. E. (1998), Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change, *Clim. Change*, *39*(4), 667–694.
Villarini, G., J. A. Smith, and G. A. Vecchi (2013), Changing frequency of heavy rainfall over the central United States, *J. Clim.*, *26*(1), 351–357.
Warner, M. D., C. F. Mass, and E. P. Salathé Jr. (2012), Wintertime extreme precipitation events along the Pacific Northwest Coast: Climatology and synoptic evolution, *Mon. Weather Rev.*, *140*(7), 2021–2043, doi:10.1175/MWR-D-11-00197.1.
Wigley, T., J. M. Lough, and P. D. Jones (1984), Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series, *J. Climatol.*, *4*, 1–25.
Wolter, K., and M. S. Timlin (1993), Monitoring ENSO in COADS with a seasonally adjusted principal component index, in *Proceedings of the 17th Climate Diagnostics*, pp. 52–57, NOAA/NMC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, Norman.
Wolter, K., and M. S. Timlin (1998), Measuring the strength of ENSO events: How does 1997/98 rank?, *Weather*, *53*(9), 315–324.
Yonekura, E., and T. M. Hall (2011), A statistical model of tropical cyclone tracks in the western North Pacific with ENSO-dependent cyclogenesis, *J. Appl. Meteorol. Climatol.*, *50*(8), 1725–1739.
Zhu, Y., and R. E. Newell (1998), A proposed algorithm for moisture fluxes from atmospheric rivers, *Mon. Weather Rev.*, *126*(3), 725–735.