Regionalization of Hydroclimate Variables

Gregory J. Carbone (✉ greg.carbone@sc.edu)

University of South Carolina <https://orcid.org/0000-0001-5333-394X>

Peng Gao

University of North Carolina Wilmington

Junyu Lu

Arizona State University

Research Article

Keywords: Regionalization, hydroclimate, variables, Parameter-elevation Relationships, Independent Slopes Model (PRISM), planning

Posted Date: August 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-732999/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Regionalization of hydroclimate variables

Gregory J. Carbone^{*1}, Peng Gao², Junyu Lu^{3,4}

1. Department of Geography, University of South Carolina, Columbia, SC, 29208, United States; *greg.carbone@mailbox.sc.edu
2. Department of Earth and Ocean Sciences, University of North Carolina Wilmington, Wilmington, NC, 28403, United States; gaop@uncw.edu
3. School of Community Resources and Development, Arizona State University, 411 N. Central Avenue, Suite 550, Phoenix, AZ 85004, United States
4. The Hainan University-Arizona State University Joint International Tourism College, Hainan University, 58 Remin Road, Haikou, Hainan Province 570004, China; junyulu@asu.edu

* corresponding author: Gregory J. Carbone (greg.carbone@sc.edu)

ORCID Numbers:

Carbone: 0000-0001-5333-394X

Gao: 0000-0002-5104-2978

Lu: 0000-0002-0992-5662

Abstract

We apply a regionalization method to hydroclimate variables using the Parameter-elevation Relationships on Independent Slopes Model (PRISM) temperature and precipitation database. Resulting regions make intuitive sense from the perspective of driving influences on temperature and precipitation averages and anomalies, and are compatible with results from another empirically derived clustering scheme. Regions selected for individual variables show high similarity across different time frames. There is slightly less similarity when comparing regions created for different monthly or daily hydroclimate variables, and relatively low similarity between monthly vs. daily measures. It is unlikely that any one regionalization solution could summarize hydroclimate extremes given the wide range of variables used to describe them, but geographically sensitive datasets like PRISM and flexible algorithms provide useful methods for regionalization that can aid in drought monitoring and forecasting, and with impacts and planning associated with heavy precipitation.

Declarations

Funding (Not applicable)

Conflicts of interest/Competing interests (no conflicts)

Availability of data and material (The datasets generated during and/or analyzed during the current study are available from the corresponding author on request.)

Code availability (Software and code available from the corresponding author on request.)

1. Introduction

Constructing climate regions has both descriptive and applied purposes.

Regionalization can reveal new aspects of data, inform data stratification, or stimulate hypotheses about physical drivers influencing climatological observations (Wilks 2019). It can also provide a rational basis for managing resources influenced by climate variability. Given the threat of drought or heavy precipitation, understanding coherent regions derived from long-term datasets with high spatial resolution would improve both theoretical and applied aspects of hydroclimate extremes.

Early rule-based regionalization schemes include those by Köppen (1900), who devised a simple scheme based on monthly temperature and precipitation values to mirror existing global natural vegetation regions, and by Thornthwaite (1931, 1948), who created regions based on potential evapotranspiration for North America. More recent delineation of climate regions emphasizes more empirical or data-driven approaches (e.g., Fovell and Fovell 1993, Fovell 1997). The practical application of coherent, statistically-based climate regions has prompted several efforts to construct them at regional to globe scales. For example, Wolter and Allured (2007), recognizing how such units might benefit drought monitoring and seasonal climate forecasts, established an iterative clustering scheme to distill data from the US Cooperative Network into coherent patterns of temperature and precipitation anomalies. Their rationale was that the spatial extent and pattern of past droughts should inform operational drought monitoring. They further argued that, since seasonal forecasts rely on the response of mid-latitude circulation to tropical Pacific sea-surface temperature and pressure anomalies, defining regions on the basis of historical temperature and precipitation anomalies produces logical spatial units to verify forecasts. Their effort defined functional regions as an alternative to the two dominant schemes used for long-term regional records and for seasonal climate forecasting – the National Oceanographic and Atmospheric Administration’s 344 climate divisions and its 102 climate forecast divisions, respectively. The former were constructed using somewhat arbitrary criteria (Guttman and Quayle 1996); the latter, while less constrained by these criteria, overlap significantly with the climate divisions and are not based on statistical properties of the climate record (Wolter and Allured 2007). Bieniek et al. (2012), also recognizing the shortcomings of the established climate divisions, constructed empirical climate regions for Alaska that better reflected temperature and precipitation response to a suite of teleconnections.

Statistically-based climate regions provide an appropriate scale to measure climate change, to avoid shortcomings associated with estimates at individual stations (Wu et al. 2018), to identify trends (Karl et al. 1994a; Karl et al. 1994b; DeGaetano 2001; Bharath and Srinivas 2015), to assess the adequacy of observing networks (DeGaetano 2001;

Bharath and Srinivas 2015), or to attribute changes to specific causes (Fazel et al. 2018). Data-derived regions are appropriate for evaluating model output from numerical weather prediction models (e.g., Argüeso et al. 2011) or control runs of general circulation models (Belda et al. 2015), or for creating climate change scenarios via downscaling from coarse to more local spatial scale (Maraun 2010; Winkler et al. 2011; Carbone 2014; Perdinan and Winkler 2015).

Regionalization has many hydroclimate applications, including assessing the spatial and temporal patterns and determining causes of drought. A variety of clustering algorithms have been used to identify the duration, spatial locus and extent, and intensity associated with significant droughts (Xu et al. 2015) or to define homogenous drought monitoring regions (Ali et al. 2019). Spatial clustering informs frequency analysis to uncover drought periodicity and to provide clues for a drought's causes (Vicente-Serrano 2006; Bordi et al. 2007; Santos et al. 2010; Gocic and Trajkovic 2014; Zhang et al. 2017; Manzano et al. 2019). Regionalization also can help improve existing methods for drought forecasts or for deriving future scenarios (Mishra and Singh 2011). Creating rational climate regions is equally relevant in analyzing heavy precipitation and flooding, as well as in storm-water and disaster and risk management (DeGaetano 2001; Du et al. 2014; Gao et al. 2018; Irwin et al. 2017). It is the basis for understanding heavy precipitation regions and is essential to stochastic storm transposition, wherein historic storms are superimposed on different regions as plausible heavy rain events. Resampling the intensity, duration, and spatial extent of a past event provides a complementary method for rainfall and flood frequency analysis (Yin et al. 2016; Wright et al. 2020) but demands coherency in the statistical properties used to construct intensity-duration-frequency (IDF) curves and other measures of probability and risk due to heavy precipitation (Gao et al. 2018). Hence, investigators have looked at spatial coherence in precipitation intensity, including diurnal rainfall patterns (Mooney et al. 2017), annual precipitation extremes (Yang et al. 2010), and variability in mountainous regions (Sugg and Konrad 2020).

While data-driven climate regions are often constructed from point data, gridded data sources abound. Derived from observational networks, reanalysis, or other model output, these data are an important part of modern climatological analysis. Gridded datasets often integrate in-situ measurements with areally-based data using multiple sources to maximize the spatial and temporal resolution. Examples include remotely sensed data, model output, and response variables such as vegetation (Rhee et al. 2008; Bieniek et al. 2012; Zhang et al. 2017). The resulting integrated products can be used to establish functional climate regions that respond to hydroclimate extremes.

In this paper, we apply a hierarchical clustering algorithm to statistical measures of the widely used Parameter-elevation Relationships on Independent Slopes Model (PRISM) temperature and precipitation database. We construct regions based on temperature and precipitation averages and anomalies, as well as the statistical parameters underlying the Standardized Precipitation Index (SPI), the Standardized Precipitation Evapotranspiration Index (SPEI), and 1-, 2-, 3-, and 4-day annual precipitation maxima. Our analysis offers an inherently areal perspective to regionalization with a gridded dataset widely employed in research and operations. The parameters upon which measures of drought or heavy precipitation are grounded both provide a rational basis for regionalization and assist in applications that depend on spatial coherence when applying probability distributions to regional data (Hosking and Wallis 1993; Bonnin et al. 2006). By clustering on the parameters of theoretical probability distributions, we consider a range of possible probability distributions beyond those provided by limited observed records (Wilks 2019). Clustering on these parameters emphasizes anomalies like those represented by the SPI and SPEI, return intervals associated with heavy precipitation, and operational products such as seasonal climate forecasts. Our analysis has three goals:

- to provide a regionalization scheme based on spatial coherence of the underlying data structure of the PRISM dataset;
- to compare our resulting regions with those of other data-driven analyses to measure similarity across varying algorithms and against established aggregation units (e.g., climate divisions, forecast regions);
- to measure how clustering solutions compare across hydroclimate variables and different periods of record; and
- to determine the feasibility of creating coherent regions that simultaneously capture multiple hydroclimate extremes.

2. Data and methods

We conduct all analyses using temperature and precipitation values from the PRISM dataset (Daly et al. 2008). PRISM data helps us address the challenge associated with the unit size of original datasets. Fovell and Fovell (1993) acknowledged that the climate divisions used to create their regions are more evenly sized in the eastern and southern USA, but more erratically sized in the west. We avoid this latent bias by clustering on the evenly-sized PRISM grid cells. These gridded data are interpolated from several different observation networks to a 4-km grid, incorporating properties of location, elevation, coastal proximity, slope orientation, vertical atmospheric profiles, topographic position, and orographic effectiveness. Like other gridded datasets, PRISM provides the advantage of complete records and spatial coverage. After 2001, PRISM precipitation data incorporated radar inputs. PRISM data have been used as the basis of

or to augment other regionalization studies (Abatzoglou et al. 2009; Bieniek et al. 2012; Sugg and Konrad 2020).

Our regionalization focuses on hydroclimate variables and includes independent analysis of the following:

- Monthly average precipitation (12 variables)
- Monthly average temperature and precipitation (24 variables)
- Combined temperature and precipitation anomalies (following Wolter and Allured 2007; Abatzoglou et al. 2009)
- Drought indices (parameters used to calculate SPI and SPEI)
 - 3-month SPI: two-parameter gamma distribution for each month (24 variables)
 - 3-month SPEI: three-parameter log-logistic distribution for each month (36 variables)
- Heavy precipitation: 1-, 2-, 3-, and 4-day annual maximum precipitation based on Anderson-Darling distance

We computed average monthly temperature and average total precipitation for each PRISM grid in each month for the period of record, 1895-2018. Monthly anomalies were calculated by subtracting the period-of-record average from each monthly observation. Our analysis for all monthly products includes three different time frames: 1895-2018, 1957-2018, and 1981-2018; in addition, we examined other time frames for specific comparison with previous studies. We do this not only because the parameters used to calculate SPI and SPEI are sensitive to record length (Wu et al. 2005; Vergni et al. 2017; Carbone et al. 2018), but also to consider changes that may be due to climate variability and change, or to changes in the number and distribution of stations used to produce the PRISM grids. While we cannot distinguish between these two factors, we can document how different time frames affect the coherency of regions inherent in this widely used dataset.

We computed SPI values following the method of McKee et al. (1993). Three-month running precipitation totals were calculated for each grid and month during the period of record. Probability distribution functions (PDFs) were fit independently for each month to derive gamma shape (alpha) and scale (beta) parameters using maximum likelihood estimation (Wilks 2019). We used the three-parameter log-logistic function to fit the probability distribution of the precipitation - potential evapotranspiration difference (Vicente-Serrano et al. 2010). We calculated SPEI using the R package SPEIcalc (URL: <http://sac.csic.es/spei>) developed by Vicente-Serrano et al. (2010). To regionalize the characteristics of heavy precipitation, we summed total rainfall for

consecutive 1-, 2-, 3-, and 4-day periods for all grid cells and each year from January 1, 1981, through December 31, 2018.

We adopted a hierarchical clustering method, Regionalization with Dynamically Constrained Agglomerative Clustering and Partitioning (REDCAP; Guo 2008), to identify clusters based on underlying parameters for the probability distributions used to calculate SPI (gamma), SPEI (Pearson-Type III), and 1- to 4-day precipitation (gamma). REDCAP is a widely used clustering algorithm with previous hydroclimate applications (Gao et al. 2018; Yang et al. 2020). REDCAP offers a flexible algorithm that overcomes limitations inherent in conventional clustering methods that do not consider spatial information, one of the key factors that shape climate regions. It uses the average linkage clustering method but directly enforces spatial contiguity during the clustering procedure forming spatially contiguous regions within which various analyses -- such as climate model evaluation and resampling of rare or extreme events -- could be conducted (Gao et al. 2015; Gao et al. 2018). By accommodating different variables derived from raw temperature and precipitation and various distance measurements (Table 1), REDCAP allows us to investigate multiple facets of hydroclimate and distill information from raw data.

REDCAP identifies groups of grid cells (i.e., clusters) from PRISM data that are spatially contiguous and have similar statistical properties defined by each variable (details below). The algorithm considers each grid cell as an individual cluster. Then, it iteratively merges pairs of clusters that are spatially contiguous and have the highest similarity (or the lowest dissimilarity/distance). This follows the same clustering procedure as conventional clustering methods (e.g., average linkage clustering) except that REDCAP requires that pairs of clusters merged in each iteration must be spatially contiguous. When all grid cells are merged, a spatially contiguous dendrogram is constructed. REDCAP then partitions the spatially contiguous dendrogram into the desired number of subtrees assigned by a user, each forming a cluster of spatially contiguous grid cells.

Dissimilarity between each pair of grid cells is defined using three distinct distance measures (Table 1). The first is Euclidian distance, applied to situations where each grid cell has multiple variables associated with that cell. The distance (D) between a pair of grid cells m and n each of which has k variables is defined by Equation 1.

$$D = \sqrt{\sum_{i=1}^n (m_i - n_i)^2} \quad (1)$$

The second distance measure follows that used by Wolter and Allured (2007) to compare our results with their statistically-derived climate divisions. Average temperatures and precipitation totals for every three-month season were computed for

each grid cell for the entire study period. The anomaly at each grid cell was then calculated by subtracting the average for the same season. Finally, the dissimilarity is determined as one minus the correlation of pairwise anomalies.

The third distance measure is Anderson–Darling (AD) distance measuring the similarity of distributions of annual precipitation maxima. The AD distance for two series of n year annual maxima at grid cell a and b is defined by Equation 2.

$$AD(a, b) = \frac{n}{2} \int_{-\infty}^{\infty} \frac{(F_a(x) - F_b(x))^2}{F_{ab}(x)(1 - F_{ab}(x))} dF_{ab}(x) \quad (2)$$

where $F_a(x)$, $F_b(x)$, and $F_{ab}(x)$ are the sample distribution functions of a , b , and a combined sample of a and b , respectively.

Hydroclimatological measure	Variable	Distance dissimilarity measure
Normality	Monthly average precipitation Monthly average temperature and precipitation	Euclidian
Anomaly	Temperature and precipitation anomaly	Wolter and Allured
Drought	Two-parameter Gamma distribution for SPI three-parameter Log-logistic distribution for SPEI	Euclidian
Heavy precipitation	1-, 2-, 3-, and 4-day annual max precipitation	Anderson-Darling

Table 1 Hydroclimate variables and dissimilarity used in regionalization

REDCAP produces any user-defined number of regions. We determined the number of clusters for each variable using two approaches: First, we selected the same number of regions used operationally or in prior studies. For example, we used 344 regions, the number of USA climate divisions used operationally and for which a 125-year historical record exists for many climate variables, including several important for hydroclimatology. We used 102 regions, the number of seasonal climate forecast regions used by NOAA’s Climate Prediction Center. We also selected 139 regions, the number established by Wolter and Allured (2007) when clustering three-month temperature and precipitation anomalies from 1978-2006 National Oceanographic and Atmospheric Administration Cooperative Observer Program (NOAA COOP) data.

Second, we used the “L method” to determine a statistically-optimal number of regions by analyzing within-region heterogeneity at each hierarchical level. Within-

region heterogeneity is defined as the sum of squared deviations in each region plotted against the number of clusters. The L method identifies points of maximum curvature in the plot, assuming that an appropriate number of clusters often coincides with rapid changes in within-region heterogeneity as clusters are merged. Visually, the L method locates the 'elbow' in the plot. Mathematically, the L method divides the plot into two parts, those with a lower or higher number of clusters (L_c and R_c , respectively) for each possible number of clusters (c). Separate lines were fitted for L_c and R_c , and total RMSE (Root Mean Squared Error) at c was defined as:

$$RMSE_c = \frac{c-1}{b-1} RMSE(L_c) + \frac{b-c}{b-1} RMSE(R_c) \quad (3)$$

where $RMSE(L_c)$ and $RMSE(R_c)$ are RMSEs of the lines in L_c and R_c respectively, and b is the largest possible number of clusters. The statistically optimal number of regions corresponds to the value of c that minimizes $RMSE_c$. When b is very large, the optimal number results in fine-grain clusters with little practical meaning. To solve this problem, the L method iteratively updates the plot by assigning b to be twice the optimal number of clusters in the last iteration. It keeps searching for the optimal number of regions in the updated plot, ultimately yielding the optimal number of regions at each iteration.

Regionalization of different climatic variables yielded different sets of regions. We determined spatial coherence of these regions using Strehl and Ghosh's (2002) index which measures the mutual spatial information shared by two random variables. Among a variety of cluster similarity indices, the Strehl and Ghosh index has several advantages. It is independent of the number of clusters or the number of elements, and it does not rely on distribution assumptions for hypothesis testing. The index is based on mutual information between two sets of clusters C and C' ($I(C, C')$) and is defined as:

$$I(C, C') = \sum_{i=1}^k \sum_{j=1}^l P(i, j) \log_2 \frac{P(i, j)}{P(i)P(j)} \quad (4)$$

where $P(i, j)$ is the probability that an element belongs to cluster C_i in C and to cluster C'_j in C' (equation 5). The Strehl and Ghosh index is the mutual information between clusters C and C' standardized by the geometric mean of the entropy of cluster C ($H(C)$) and C' ($H(C')$) defined by equation 6, where $P(i)$ is the probability that an element is in cluster $C_i \in C$ which has n elements (equation 7).

$$P(i, j) = \frac{|C_i \cap C'_j|}{n} \quad (5)$$

$$H(C) = - \sum_{i=1}^k P(i) \log_2 P(i) \quad (6)$$

$$P(i) = \frac{|c_i|}{n} (7)$$

We apply the Strehl and Ghosh index for two purposes: 1) to compare spatial similarity of clusters derived for two different variables, and 2) to compare spatial similarity of clusters derived from the same variable across different periods of record.

3. Results

Monthly temperature and precipitation averages

Two distinct “elbows” define statistically optimal solutions for 64 and 19 clusters using monthly temperature and precipitation averages (Table 2). The 64-region solution shows familiar patterns in the eastern two-thirds of the contiguous USA (Fig. 1): a north-south gradient from the Gulf of Mexico to the Great Lakes driven by solar insolation and temperature, an east-west gradient in the central USA/Great Plains driven by precipitation, and SW-NE oriented clusters reflecting the Appalachian Mountains or proximity to the Atlantic. In the West, the clusters are considerably smaller and highlight the influences of elevation and coastal proximity.

Fig. 1 Sixty-four regions based on monthly temperature and precipitation normals from PRISM data, 1981-2018

In order to compare REDCAP clusters with a prior regionalization for monthly temperature and precipitation averages, we use a 14-cluster solution (close to a statistically optimal breakpoint). The 14-cluster REDCAP solution (Fig. 2a) shows general similarities to the well-known Fovell and Fovell (1993) clusters (Fig. 2b). In particular, large latitudinal bands characterize regions in the eastern USA. This north-south temperature gradient, combined with elevation effects and the east-west precipitation gradient dominate the regional patterns in the western USA. The most striking differences produced by REDCAP are associated with a more refined precipitation gradient across the central Plains and more detailed regions in the western USA reflecting topographic influences on temperature and precipitation. These differences stem from the use of PRISM data that incorporates elevation and slope into its gridding algorithm. By contrast, the climatological division data used by Fovell and Fovell (1993) are aggregated at the coarser-scale climate divisions and without explicit consideration for elevation. Moreover, the climatological division data in mountainous areas are disproportionately constructed from stations at lower elevations.

Also noteworthy is that when forced to 14 clusters, REDCAP produces different clusters for different periods of the PRISM record, e.g., 1931-1981 vs. 1981-2018. While similar *general* patterns emerge from the two periods, there are distinct differences in the central Plains and western USA (Fig. 2c). From this analysis, it is impossible to distinguish the degree to which such differences represent changes in temperature and precipitation after 1981 as opposed to different stations used to construct the PRISM dataset. One clear difference regarding the latter period is that it overlaps considerably with the period from which PRISM incorporates Snow Telemetry (SnoTel) data in the mountainous west and radar-based precipitation estimates across the USA.

Fig. 2 REDCAP clustering into 14 regions based on monthly temperature and precipitation normals from PRISM data, 1931-1981 (a), and 1981-2018 (c); Digital version of Fovell and Fovell's (1993) 14 regions using climate division data, 1931-1981(b).

Closer inspection across different periods of record reveals that similarity varies with the number of clusters selected and the record length and/or overlap between periods. Strehl and Ghosh index values, quantifying the mutual information shared by clusters, show that, with forty or more clusters, spatial similarity is consistently high (approximately 0.8) irrespective of record length or specific time periods (Fig. 3). This suggests that, for much of the country, the use of statistical properties from coherent regions formed on the basis of monthly temperature and precipitation averages is relatively insensitive to the period of record used to construct the regions when the total number of regions is sufficiently large. With fewer clusters, mutual shared information decreases rapidly. For example, a ten-cluster solution derived from 1981-2018 average temperature and precipitation data has a relatively low similarity index (approximately 0.65) to ten-cluster solutions based on 1895-1956, 1895-2018, 1957-2018, or 1931-1981 data. By contrast, increasing period length and/or overlap—1895-2018 vs. 1957-2018—results in similarity index values above 0.8 with as few as five clusters.

Fig. 3 Similarity of regions based on monthly temperature and precipitation averages across four different periods: a) 1957-2018 vs 1981-2018, b) 1895-2018 vs. 1981-2018, c) 1895-2018 vs. 1957-2018, and d) 1931-1981 vs. 1981-2018.

Since we report the Strehl and Ghosh similarity index values throughout this paper, it is helpful to calibrate these values to a series of maps. The similarity index between the two time periods used in the 14-cluster solution above (Figs. 2a vs. 2c), for example, is 0.69. At 60 clusters, the similarity for the same two time periods (1981-2018 and 1931-1981) is 0.81 (Fig. 4).

Fig. 4 Sixty-four regions based on monthly temperature and precipitation normals from PRISM data, 1981-2018 (a) and 1931-1981 (b)

REDCAP's solution for 102 clusters based on monthly temperature and precipitation normals produces regions quite different than NOAA's standard forecast regions (Fig. 5). REDCAP regions vary more in size and are more elongated and irregular. The REDCAP clusters reveal familiar climate controls. In the eastern USA, the elongated clusters amplify the role of latitude and proximity to the Gulf of Mexico or Atlantic coasts on temperature and precipitation anomalies. But the most dramatic differences between REDCAP's clusters and NOAA's forecast regions is in the West where properties of elevation and slope inherent in the PRISM dataset matter more. These results are not surprising as the current forecast regions -- a result of National Weather Service reorganization in the 1990s -- retain remnants of state borders with only limited exceptions for physical features. Many forecast region boundaries are determined by either climate divisions or political boundaries rather than by data-driven regions.

Fig. 5 NOAA's 102 forecast regions (a); 102 regions based on normal T and P from 1981 to 2018 created by REDCAP (b).

Monthly temperature and precipitation anomalies

Forcing REDCAP to produce 139 clusters to compare against the 139 clusters created by Wolter and Allured (2007) resulted in remarkably similar regional patterns (Fig. 6). While both schemes use the same measure of seasonal temperature and precipitation anomalies and both the same time periods, they differ in constraints on spatial contiguity (REDCAP demands spatial contiguity, Wolter and Allured do not) and operate on different datasets (PRISM grids vs. NWS Coop stations). Yet, both schemes produce clusters that reflect elevation, orientation, and proximity to the coast. In the eastern USA, both schemes produce larger regions than the well-established climatic divisions. In the western USA, the derived clusters are often smaller than the standard climate divisions used in mountainous states. When considering more than twenty-five clusters, REDCAP produces similar seasonal temperature and precipitation anomalies (Strehl and Ghosh similarity index of approximately 0.8), irrespective of the period of record. For 139 clusters, the similarity index is approximately 0.83 whether comparing clusters derived from 1895-1956, 1895-2018, 1957-2018, or 1981-2018 data.

Fig. 6 Comparison of 139 clusters of monthly, October 1978 - September, 2006 temperature and precipitation anomalies derived by Wolter and Allured (2007; colored dots) vs. REDCAP (black boundary lines).

Drought indices

The L method produces breakpoints for monthly temperature-precipitation anomalies, SPEI, and SPI clusters at 74, 56, and 50 respectively (Table 2).

	break 1	break 2	break 3	break 4
Normal precipitation and temperature	6	19	64	496
Anomaly precipitation and temperature	7	10	17	33
SPI	8	15	50	210
SPEI	4	10	26	56
1-day ann. max. precipitation	6	11	37	153
2-day ann. max. precipitation	6	10	27	182
3-day ann. max. precipitation	6	11	26	158
4-day ann. max. precipitation	6	13	45	187

Table 2. “Elbow” breakpoints for each variable.

Here, we show that for a mid-range value of 60 clusters, temperature-precipitation anomalies regions are remarkably uniform in size across the USA (Fig. 7a). Clusters

based on statistical parameters used to construct SPI or SPEI values vary more in size and more clearly highlight elevation and slope characteristics captured by the PRISM dataset (Figs. 7b and 7c). Yet, the statistical similarity for all possible pairs of these three variables is high -- the Strehl and Ghosh similarity index is greater than 0.75 for 60 or more clusters.

Fig. 7 Sixty-cluster solutions for: a) temperature-precipitation anomalies, b) SPEI parameters, and c) SPI parameters.

In addition, clustering results are robust across a range of time frames for each of these variables. With more than thirty clusters, the Strehl and Ghosh similarity index is greater than 0.75 when comparing clusters derived from 1895-1956, 1895-2018, 1957-2018, or 1981-2018 data for temperature-precipitation anomalies, as well as for SPI and SPEI parameters.

The Strehl and Ghosh similarity index for SPEI vs. monthly temperature and precipitation averages ranges from 0.7 to 0.78 when creating more than 60 regions. The similarity index for SPI vs. monthly temperature and precipitation averages and temperature and precipitation anomaly vs. monthly temperature and precipitation averages is in the same approximate range -- 0.7 - 0.76 for more than 60 regions. Both are lower than the similarity between the drought indices -- SPI vs. SPEI -- or the similarity of a single anomaly or drought measure across different time frames.

Heavy Precipitation

We examine clusters forming 11 and 30 heavy precipitation regions based on average breakpoint elbows found in 1- to 4-day annual maxima precipitation totals (Table 2). The resulting regional patterns reflect three general influences: proximity to the Gulf, Atlantic, or Pacific, the east-west moisture gradient in the central USA, and orographic effects (Figs. 8 and 9).

Fig. 8 REDCAP 11-cluster solution optimizing similarity in annual maxima for a) 1-day, b) 2-day, c) 3-day, and d) 4-day precipitation.

Fig. 9 As in Figure 8, for the REDCAP 30-cluster solution.

The regions show a moderate degree of similarity across different durations (1- to 4-days) when eleven or more clusters are selected -- Strehl and Ghosh similarity index values range from approximately 0.70 to 0.78 from 11 to 30 regions. Strehl and Ghosh

similarity index values are also in this range for a single precipitation variable (i.e., 1- to 4-day annual maximum) when compared across different record lengths (e.g., 1981-1999, 2000-2018, 1981-2018). By contrast, similarity index values between any of the daily precipitation intensity measures and any of the monthly variables range are considerably lower, ranging from 0.5 to 0.58. Here, we illustrate that similarity values between monthly mean precipitation and SPI are considerably higher than either variable compared with 1- to 4-day annual maximum precipitation (Fig. 10).

Fig. 10 Similarity index for sample pairs of monthly average precipitation (p), monthly SPI, and 1- to 4-day annual maximum precipitation.

The clusters created by such analysis could form the basis for regional frequency analysis whereby precipitation data are pooled to calculate intensity-duration-frequency curves (Hosking and Wallis 1993; Irwin et al. 2017; Yang et al. 2010; Burn 2014). Whether such pooling is done independently for each duration period (1- to n-days) depends on the level of similarity required. Clustering of heavy precipitation statistics also reduces interannual variability associated with station measurements (Kunkel et al. 2020).

4. Discussion and Summary

Our method creates empirically derived regions exploiting the underlying statistical structure of important hydroclimate variables. By using PRISM, we define regions with a dataset that is widely employed in research and operations and also incorporates the role of topography, proximity to the coast, and other geographic features. These organic regions are often similar to those derived with other clustering algorithms and often quite different than many commonly-used spatial units (e.g., climate divisions, forecast regions). Of course, such units will remain a well-established standard and have practical value in many circumstances. State, county, or other political boundaries are important for management and policy decisions. However, it is important to recognize that there are drawbacks to using these when selecting individual stations to represent a region, as an aggregated coherent spatial unit, or for downscaling or aggregating data. While the boundaries of data-driven regions do not stop at political borders, this should not preclude their use. Resource managers concerned with a political unit can use them within their jurisdictions, not worrying about how large, small, or differently sized they are.

Many clustering algorithms draw on raw data or an underlying statistical structure. While we do not provide a comprehensive evaluation of different clustering schemes here, our method shows high similarity with that of Wolter and Allured (2007) who use

station data and employ a different algorithm. This suggests that a number of different data-based clustering solutions can offer a sound approach for aggregation, one that has practical applications for monitoring temperature and precipitation anomalies, and for producing seasonal temperature or precipitation forecasts or assessing their accuracy.

Our analysis also sought to assess similarity of regions based on different periods of record and different hydroclimate variables. The monthly hydroclimate variables considered here show high similarity across different time periods when the total number of regions was sufficiently large -- typically more than 25 clusters across the conterminous United States. Similarity indices for monthly temperature and precipitation means and for temperature and precipitation anomalies exceeded 0.8 when comparing first and second halves of the record; SPI and SPEI were only slightly lower (~0.75). Even heavy precipitation had relatively high similarity (~0.7) across different time periods, despite analysis using shorter record lengths. Similarity values across different hydroclimate variables were typically lower than a single variable across different time periods. Combinations of monthly average temperature, temperature-precipitation anomalies, SPI, and SPEI, and combinations of 1- to 4-day annual maximum precipitation values ranged from 0.7 to 0.78. Similarity between any single monthly variable and any single daily precipitation variable was considerably lower (0.5-0.58). This is not surprising given the short-term nature of intense precipitation events and differences between the drivers of these events versus the controls on monthly temperature and precipitation anomalies.

Certainly, no single regionalization accommodates the variables commonly used to define hydroclimate extremes like drought or heavy precipitation, making the creation of consistent "hydroclimate extreme" regions challenging. Our findings suggest that no one solution suits our particular selection of hydroclimate extremes. Nor would one solution satisfy the diverse number of research or operational needs for which climate regions might be appropriate (Abatzoglou et al. 2009). Since the definition of clusters depends on the variable of interest and specific application, flexible algorithms that allow user-defined parameters (e.g., REDCAP) offer a practical tool for decision makers.

Declarations

Funding: This research was supported by supported by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office (grant no. NA16OAR4310163).

Conflicts of interest/Competing interests: (The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.)

Author's Contribution: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed largely by Peng Gao and Junyu Lu. All authors contributed to the first draft and editing of the manuscript. All authors read and approved the final manuscript.

Availability of data and material: (The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.)

Code availability (Software and code used in the study is available from the corresponding author upon reasonable request.)

Ethics approval (Not applicable)

Consent to participate (Not applicable)

Consent for publication (Not applicable)

References

- Abatzoglou JT, Redmond KT, Edwards LM (2009) Classification of Regional Climate Variability in the State of California. *Journal of Applied Meteorology and Climatology* 48:1527-1541. <https://doi.org/10.1175/2009Jamc2062.1>
- Ali Z, Hussain I, Faisal M, Shoukry AM, Gani S, Ahmad I (2019) A framework to identify homogeneous drought characterization regions. *Theoretical and Applied Climatology* 137:3161–3172. <https://doi.org/10.1007/s00704-019-02797-w>
- Andreadis KM, Clark EA, Wood AW, Hamlet AF, Lettenmaier DP (2005) Twentieth-Century Drought in the Conterminous United States. *Journal of Hydrometeorology* 6(6): 985-1001. <https://doi.org/10.1175/JHM450.1>
- Argüeso D, Hidalgo-Muñoz JM, Gámiz Fortis SR, Esteban-Parra MJ, Dudhia J, Castro-Díez Y (2011) Evaluation of WRF parameterizations for climate studies over southern Spain using a multistep regionalization. *Journal of Climate* 24: 5633-5651. <https://doi.org/10.1175/JCLI-D-11-00073.1>
- Belda M, Holtanová E, Halenka T, Kalvová J, Hlávka Z (2015) Evaluation of CMIP5 present climate simulations using the Köppen-Trewartha climate classification. *Climate Research* 64:201-212. <https://doi.org/10.3354/cr01316>
- Bharath R, Srinivas VV (2015) Delineation of homogeneous hydrometeorological regions using wavelet-based global fuzzy cluster analysis. *International Journal of Climatology* 31(15):4707-4727. <https://doi.org/10.1002/joc.4318>
- Bieniek PA, Bhatt US, Thoman RL, Angeloff H, Pertain J, Papineau J, Fritsch F, Holloway E, Walsh JE, Daly C, Shulski M, Hufford G, Hill DF, Calos S, Gens R (2012) Climate divisions for Alaska based on objective methods. *Journal of Applied Meteorology and Climatology* 51(7):1276-1289. <https://doi.org/10.1175/JAMC-D-11-0168.1>
- Bonnin GM, Martin D, Lin B, Parzybok T, Yekta M, Riley D (2006) *Precipitation-Frequency Atlas of the United States*, NOAA Atlas 14, Volume 2, Version 3.0, NOAA, National Weather Service, Silver Spring, Maryland.
- Bordi I, Sutera A (2007) Drought monitoring and forecasting at large scale. In: Rossi G, Vega T, Bonaccorso B (eds) *Methods and tools for drought analysis management*. Springer, Dordrecht, pp 3-27

Burn DH (2014) A framework for regional estimation of intensity–duration–frequency (IDF) curves. *Hydrological Processes* 28:4209–4218. <https://doi.org/10.1002/hyp.10231>

Carbone GJ (2014) Managing climate change scenarios for societal impact studies. *Physical Geography* 35(1):22-49. <https://doi.org/10.1080/02723646.2013.869714>

Carbone GJ, Lu J, Brunetti M (2018) Estimating uncertainty associated with the standard precipitation index. *International Journal of Climatology* 38(S1):e607-e616. <https://doi.org/10.1002/joc.5393>

Cheng L, AghaKouchak A (2014) Nonstationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate. *Scientific Reports* 4:7093. <https://doi.org/10.1038/srep07093>

Daly C, Halbleib M, Smith JL, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology* 28:2031-2048. <https://doi.org/10.1002/joc.1688>

DeGaetano AT (2001) Spatial grouping of United States climate stations using a hybrid clustering approach. *International Journal of Climatology* 21: 791-807. <https://doi.org/10.1002/joc.645>

Du H, Xia J, Zeng S (2014) Regional frequency analysis of extreme precipitation and its spatio-temporal characteristics in the Huai River Basin, China. *Natural Hazards* 70: 195–215. <https://doi.org/10.1007/s11069-013-0808-6>

Fazel N, Berndtsson R, Uvo CB, Madani K, Kløve B (2018) Regionalization of precipitation characteristics in Iran's Lake Urmia basin. *Theoretical and Applied Climatology* 132(1-2): 363-373. <https://doi.org/10.1007/s00704-017-2090-0>

Fovell RG (1997) Consensus clustering of US temperature and precipitation data. *Journal of Climate* 10(6):1405-1427. [https://doi.org/10.1175/1520-0442\(1997\)010<1405:CCOUST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<1405:CCOUST>2.0.CO;2)

Fovell RG, Fovell M-YC (1993) Climate zones of the conterminous United States defined using cluster analysis. *Journal of Climate* 6:2103-2135. [https://doi.org/10.1175/1520-0442\(1993\)006<2103:CZOTCU>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<2103:CZOTCU>2.0.CO;2)

- Gao P, Carbone GJ, Guo D (2015) Assessment of NARCCAP model in simulating rainfall extremes using a spatially constrained regionalization method. *International Journal of Climatology* 36(5):2368-2378. <https://doi.org/10.1002/joc.4500>
- Gao P, Carbone GJ, Lu J, Guo D (2018) An area-based approach for estimating extreme precipitation probability. *Geographical Analysis* 50(3):314-333. <https://doi.org/10.1111/gean.12148>
- Gogic M, Trajkovic S (2014) Spatiotemporal characteristics of drought in Serbia. *Journal of Hydrology* 510:110-123. <https://doi.org/10.1016/j.jhydrol.2013.12.030>
- Guo D (2008) Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographic Information Science* 22(7):801–823. <https://doi.org/10.1080/13658810701674970>
- Guttman NB, Quayle RG (1996) A historical perspective of U.S. Bulletin of the American Meteorological Society 77(2):293-304. [https://doi.org/10.1175/1520-0477\(1996\)077<0293:AHPOUC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2)
- Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. *Water Resources Research* 29(2): 271-281. <https://doi.org/10.1029/92WR01980>
- Irwin S, Srivastav RK, Simonovic SP, Burn DH (2017) Delineation of precipitation regions using location and atmospheric variables in two Canadian climate regions: the role of attribute selection. *Hydrological Sciences Journal* 62(2):191-204. <https://doi.org/10.1080/02626667.2016.1183776>
- Karl TR, Easterling DR, Knight RW, Hughes PY. 1994a. US, national and regional temperature anomalies. In: Boden TA, Kaiser DP, Sepanski RJ, Stoss FW (eds) *Trends '93: A Compendium of Data on Global Climate Change*. ORNL/CDIAC-65. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory: Oak Ridge, TN, pp. 686–736
- Karl TR, Easterling DR, Groisman PY 1994b. United States historical climatology network – national and regional estimates of monthly and annual precipitation. In: Boden TA, Kaiser DP, Sepanski RJ, Stoss FW (eds) *Trends '93: A Compendium of Data on Global Climate Change*. ORNL/CDIAC-65. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory: Oak Ridge, TN, pp. 830–905

Köppen W (1900) Versuch einer Klassifikation der Klimate, Vorzugsweise nach ihren Beziehungen zur Pflanzenwelt [Attempted climate classification in relation to plant distributions]. *Geographische Zeitschrift* 6:593-611, 657-679.

Kunkel KE, Karl TR, Squires MF, Yin X, Stegall ST, Easterling DR 2020. Precipitation extremes: Trends and relationships with average precipitation and precipitable water in the contiguous United States. *Journal of Applied Meteorology and Climatology* 59(1):125-142. <http://dx.doi.org/10.1175/JAMC-D-19-0185.1>

Manzano A, Clemente MA, Morata A, Luna MY, Begueria S, Vicente-Serrano SM, Martin ML (2019) Analysis of the atmospheric circulation pattern effects over SPEI drought index in Spain. *Atmospheric Research* 230, UNSP 104630. <https://doi.org/10.1016/j.atmosres.2019.104630>

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, et al. (2010) Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics* 48(3), RG3003. <https://doi.org/10.1029/2009RG000314>

McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. Eighth Conference on Applied Climatology, 17-22 January 1993, Anaheim, California. <https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf>

Mishra AK, Singh VP (2011) Drought modeling --- A review. *Journal of Hydrology* 403(1-2):157-175. <https://doi.org/10.1016/j.jhydrol.2011.03.049>

Mooney PA, Broderick C, Bruyere CL, Mulligan FJ, Prein AF (2017) Clustering of observed diurnal cycles of precipitation over the United States for evaluation of a WRF Multiphysics regional climate ensemble. *Journal of Climate* 30(22):9267-9286. <https://doi.org/10.1175/JCLI-D-16-0851.1>

Oliver MA, Webster R, (1989) A geostatistical basis for spatial weighting in multivariate classification. *Mathematical Geology* 21:15–35. <https://doi.org/10.1007/BF00897238>

Perdinan, Winkler JA (2015) Selection of climate information for regional climate change assessments using regionalization techniques: an example for the Upper Great

Lakes Region, USA. *International Journal of Climatology* 35(6):1027-1040.
<https://doi.org/10.1002/joc.4036>

Santos JF, Pulido-Calvo I, Portella MM (2010) Spatial and temporal variability of droughts in Portugal. *Water Resources Research* 46(3) W03503.
<https://doi.org/10.1029/2009WR008071>

Strehl A, Ghosh J (2002) Cluster Ensembles – A Knowledge Reuse Framework for Combining Multiple Partitions. *Journal of Machine Learning Research* 3: 583-617.
<https://doi.org/10.1162/153244303321897735>

Sugg JW, Konrad CE (2020) Defining hydroclimatic regions using daily rainfall characteristics in the southern Appalachian Mountains. *International Journal of Climatology* 13(7):785-802. <https://doi.org/10.1080/17538947.2019.1576785>

Thornthwaite CW (1931) The climates of North America, according to a new classification. *Geographical Review* 21:633-655.

Thornthwaite CW (1948) An approach towards a rational classification of climate. *Geographical Review* 38(1):55-94.

Vergni L, Di Lena B, Todisco F, Mannocchi F (2017) Uncertainty in drought monitoring by the Standardized Precipitation Index: the case study of the Abruzzo region (central Italy). *Theoretical and Applied Climatology* 128:13–26. <https://doi.org/10.1007/s00704-015-1685-6>

Vicente-Serrano SM (2006) Spatial and temporal analysis of droughts in the Iberian Peninsula (1910–2000). *Hydrological Sciences Journal* 51(1):83-97.
<https://doi.org/10.1623/hysj.51.1.83>

Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate* 23(7):1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>

Wilks D (2019) *Statistical Methods in the Atmospheric Sciences*, 4th edn. Elsevier, Cambridge

Winkler JA, Guentchev GS, Perdinan, Tan P-N, Zhong S, Liszewska M, Abraham Z, Niedźwiedź T, Ustrnul Z (2011) Climate scenario development and applications for local/regional climate change impact assessments: An overview for the non-climate

scientist, Part I: Scenario development using downscaling methods *Geography Compass* 5(6):301-328. <https://doi.org/10.1111/j.1749-8198.2011.00425.x>

Wolter K, Allured D (2007) New climate divisions for monitoring and predicting climate in the U.S. *Intermountain West Climate Summary* 3(5):2-6. <https://climas.arizona.edu/sites/default/files/pdf2007julnewclimatedivisions.pdf>. Accessed 16 July 2021.

Wright DB, Guo Y, England JF (2020) Six decades of rainfall and flood frequency analysis using stochastic storm transposition: Review, progress, and prospects. *Journal of Hydrology* 585. 124816 <https://doi.org/10.1016/j.jhydrol.2020.124816>

Wu FF, Yang XH, Shen ZY (2018) A three-stage hybrid model for regionalization, trends and sensitivity analyses of temperature anomalies in China from 1966 to 2015. *Atmospheric Research* 205:80-92. <https://doi.org/10.1016/j.atmosres.2018.02.008>

Wu H, Hayes MJ, Wilhite DA, Svoboda MD (2005) The effect of the length of record on the standardized precipitation index calculation. *International Journal of Climatology* 25:505– 520. <https://doi.org/10.1002/joc.1142>

Xu K, Yang DW, Yang HB, Li Z, Qin Y, Shen Y (2015) Spatio-temporal variation of drought in China during 1961-2012: A climatic perspective. *Journal of Hydrology* 526:253-264. <http://doi.org/10.1016/j.jhydrol.2014.09.047>

Yang T, Shao, QX, Hao ZC, Chen X, Zhang ZX, Xu CY, Sun LM (2010) Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. *Journal of Hydrology* 380(3-4):386-405. <http://doi.org/10.1016/j.jhydrol.2009.11.013>

Yang W, Deng M, Tang J, Jin R (2020) On the use of Markov chain models for drought class transition analysis while considering spatial effects. *Natural Hazards* 103:2945– 2959. <https://doi.org/10.1007/s11069-020-04113-6>

Yin Y, Chen H, Xu C, Xu W, Chen C, Sun S (2016) Spatio-temporal characteristics of the extreme precipitation by L-moment-based index-flood method in the Yangtze River Delta region, China. *Theoretical and Applied Climatology* 124:1005–1022. <https://doi.org/10.1007/s00704-015-1478-y>

Zhang Q, Kong DD, Singh VP, Shi PJ (2017) Response of vegetation to different time-scales drought across China: Spatiotemporal patterns, causes and implications. *Global and Planetary Change* 152:1-11. <https://doi.org/10.1016/j.gloplacha.2017.02.008>

Figures

Figure 1

Sixty-four regions based on monthly temperature and precipitation normals from PRISM data, 1981-2018

Figure 2

REDCAP clustering into 14 regions based on monthly temperature and precipitation normals from PRISM data, 1931-1981 (a), and 1981-2018 (c); Digital version of Fovell and Fovell's (1993) 14 regions using climate division data, 1931-1981(b).

Figure 3

Similarity of regions based on monthly temperature and precipitation averages across four different periods: a) 1957-2018 vs 1981-2018, b) 1895-2018 vs. 1981-2018, c) 1895-2018 vs. 1957-2018, and d) 1931-1981 vs. 1981-2018.

Figure 4

Sixty-four regions based on monthly temperature and precipitation normals from PRISM data, 1981-2018 (a) and 1931-1981 (b)

Figure 5

NOAA's 102 forecast regions (a); 102 regions based on normal T and P from 1981 to 2018 created by REDCAP (b).

Figure 6

Comparison of 139 clusters of monthly, October 1978 - September, 2006 temperature and precipitation anomalies derived by Wolter and Allured (2007; colored dots) vs. REDCAP (black boundary lines).

Figure 7

Sixty-cluster solutions for: a) temperature-precipitation anomalies, b) SPEI parameters, and c) SPI parameters.

Figure 8

REDCAP 11-cluster solution optimizing similarity in annual maxima for a) 1-day, b) 2-day, c) 3-day, and d) 4-day precipitation.

Figure 9

As in Figure 8, for the REDCAP 30-cluster solution.

Figure 10

Similarity index for sample pairs of monthly average precipitation (p), monthly SPI, and 1- to 4-day annual maximum precipitation.