A Generalized Spatio-Temporal Threshold Selection Method for Identification of Extreme Event Patterns

Vitaly Kholodovsky

A scholarly paper in partial fulfillment of the requirements for the degree of Master of Science

August 2018

Department of Atmospheric and Oceanic Science
University of Maryland College Park, Maryland

Advisor: Dr. Xin-Zhong Liang

Abstract

Extreme weather and climate events such as heavy precipitation, drought, heat waves and strong winds can cause extensive damage to society in terms of human lives and financial losses. As climate changes, it is important to understand how extreme weather events may change as a result. Climate and statistical models are often independently used to model extreme events. To better assess performance of the climate models, a variety of spatial forecast verification methods have been developed. However, in most cases, the spatial verification measures that are widely used to compare mean states do not have sufficient theoretical justification to benchmark extreme weather events. As part of an integrated modeling framework, we propose a new generalized spatio-temporal threshold selection method for the identification of extreme event episodes, which couples existing pattern recognition indices with high (or low) threshold choices. This integrated approach has four main steps: 1). Construction of essential climate quantities; 2). Dimension reduction; 3). Spatial domain mapping; and 4). Threshold clustering. We apply this approach to observed standardized precipitation rate anomalies over CONUS. The proposed method automates the threshold selection process and can be directly applicable in conjunction with modeling of extremes. Additionally, it allows for identification of synoptic scale spatial patterns that can be directly traced to individual extreme episodes, and it offers users the flexibility to select an extreme threshold that is linked to desired geometrical properties.

Table of Contents

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor Prof. Xin-Zhong Liang for the continuous support of my graduate study, for his patience, guidance and immense knowledge. In addition, I would like to thank Dr. Kenneth Kunkel for sharing with our group his precipitation dataset as well as David New and Xiaohui Zheng for processing it. I thank all faculty members and classmates at AOSC department as well as teammates at ESSIC for your time in helping me to expand my knowledge in atmospheric and oceanic science and for simulated discussions. My sincere thanks also goes to Jennifer Kennedy for reviewing this manuscript. and helpful suggestions. Last but not the least, I would like to thank my family for their support and understanding throughout writing this thesis.

This material is based upon work supported by the National Oceanic and Atmospheric Administration, Educational Partnership Program, U.S. Department of Commerce, under Agreement No. NA16SEC4810006. Its contents are solely the responsibility of the award recipient and do not necessarily represent the official views of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

List of Tables

List of Figures

List of Symbols

$Y(s,t)$ - spatio-temporal essential climate quantity

$\mathrm{Q}_{\mathrm{s}}^{(\eta)}(Y)$ - η-th quantile in space

$\mathrm{Q}_{\mathrm{t}}^{(\tau)}(Y)$ - τ-th quantile in time

$\mathrm{H}(s,t)$ - positive extreme field

$\mathrm{L}(s,t)$ - negative extreme field

$f(\theta)$ - conditional frequency

C_{index} - connectivity index

S_{index} - shape index

P_{min} - theoretical minimum perimeter

P - perimeter of the pixels above a given threshold

$\complement_{\mathrm{index}}$ - object complexity

A - area of the pixels

A_{convex} - area of the convex hull

A_{index} - area index

$\mathrm{I}[\cdot]$ - index operator for geometric index

θ - threshold

s - number of spatial pixels

t - number of temporal samples

d_{ACF} - dissimilarity metric based on autocorrelation

d_{PACF} - dissimilarity metric based on partial autocorrelation

$s(i)$ - silhouette coefficient object i

W - daily accumulation window

1 Introduction

Extreme events of essential climate variables (Bojinski et al. 2014) such as heavy rainfall, high temperatures, strong winds and their derivatives such as droughts and heat waves, are a major source of risk to society and the environment. The purpose of risk management is to design and implement a set of common procedures for identifying, measuring, and managing such high impact events. However, there is no uniformly accepted definition of extreme events. More often than not, the process of identifying of extreme events has been typically based on classifying them into groups with characteristics such as frequency of occurrence, intensity, temporal duration and timing (Stephenson 2008). In addition, the evaluation of multidimensional nature of extreme events, in most cases, has been confined to individual grid points, thus overlooking embedded spatial dependency. The Intergovernmental Panel on Climate Change (IPCC) defines climate extreme as "the occurrence of a value of a weather or climate variable above (or below) a threshold value near the upper (or lower) ends of the range of observed values of the variable" (IPCC 2012). Typically, choice of high (or low) threshold depends on the conventions of specific scientific disciplines. For example, the climate science community tends to choose location-specific thresholds based on either categories (e.g. Sun et al. 2007; Dai 2001; Dai et al. 2017), quantiles (e.g. Groisman et al. 2005; Lau et al. 2013; Pendergrass and Hartmann 2014), combinations (e.g. Kunkel et al. 2003, 2007, 2010, 2012) or indicies (e.g. Alexander et al. 2006; Zhang et al. 2011; Donat et al. 2013a,b). In the realms of statistical modeling, under appropriate conditions, excesses above (or below) high (or low) threshold are often represented as part of generalized Pareto distribution (GPD) (Balkema

and de Haan 1974; Pickands 1975), where requirements for a chosen threshold is to be high (or low) enough to satisfy chosen goodness of fit test. Although an interdisciplinary overlapping for the threshold choice is widely practiced, it is still hard to compare extreme value analysis results between different scientific studies if the threshold selection methodologies are differ. Even though, the same forecast verification strategy (in this case, a variety of traditional grid-by-grid point verifications methods (e.g Jolliffe and Stephenson 2003; Wilks 2011, Chapter 8) is used across studies, the choice between forecasting models could significantly diverge leading to different inference results. Thus a unified approach for threshold selection (i.e. universal extreme event definition) is important because it standardizes inference process for extreme events analysis, leading further to consistency and transparency when compare results between different scientific findings. Unification, standardization and transparency are the key pillars for disciplined risk management.

Although, many univariate methods have been proposed to automate the threshold choice (e.g. Fukutome et al. 2015; Scarrott and MacDonald 2012; Bader et al. 2017), there are no clear cut criteria, and in practice threshold choice tends to be determined by either exploration methods or by stability assessments of parameters estimates. Ultimately, the choice of threshold is always an interplay between bias and variance (Smith 1987). For high extreme thresholds, if the chosen threshold is not high enough, the GPD likely will not have a good fit to the excesses above the threshold, leading to approximation bias. Conversely, if the chosen threshold is too high, only a small number of exceedances will be generated and consequently, there will be high variance in the estimators (in the case of low extreme thresholds, these conditions are reversed). A more comprehensive evaluation of univariate

extreme value theory (EVT) is given in the excellent book by Coles (2001).

Analysis of extreme events must also incorporate the spatial nature of the key climate variables since many single quantities such as precipitation or temperature are measured at multiple sites. Implementation of spatial extreme value analysis has to be done using an integrative modeling approach (IMA), in which different scientific disciplines are combined into one holistic process, joined by a common modeling factor(s) subject to uniform assumptions. For example, in statistics, it may involve integration of multivariate EVT (e.g. Cooley et al. 2012), geostatistics (e.g. Banerjee et al. 2015), and spatial forecast verification (e.g. Friederichs and Thorarinsdottir 2012), and in atmospheric science, it would encourage use of coupled high resolution climate systems to resolve localized extreme events coupled with spatial threshold selection algorithm to model spacial dependencies. While myriad of other coupling combinations are possible, the challenging aspect of this integration lies in maintaining suitable and consistent statistical assumptions across all modeling blocks.

This paper proposes a new generalized spatio-temporal threshold selection algorithm for extreme events within IMA framework. By the generalized, we mean the algorithm's applicability to the wide variety of essential climate variables and their derived quantities (hereafter, essential climate quantities (ECQs)). The algorithm detects large scale extreme spatial patterns that occupy fairly extensive geographical areas. Those patterns contain extreme fields (i.e. for high (or low) extreme thresholds, only fields from upper (or lower) quartile of the distribution are selected) in both temporal and spatial variations, and can be linked to the individual extreme episodes such as flash floods, hurricanes, droughts etc. The threshold selection process is conditioned on these extreme patterns and enables

one to detect latent spatial dependencies with a help of geometric indices from digital topology. The algorithm incorporates spatial and temporal dependence in one holistic modeling framework opening an opportunity for future analysis of statistical inference of extreme events for spatial data for either a single quantity (e.g. precipitation) or an index of multiple quantities (e.g. the Palmer Drought Severity Index), which is not possible with traditional grid-by-grid point methods. Moreover, the extreme threshold value that was estimated in the observational setting and the resulted extreme geometrical field can be further applied in spatial forecast verification and independently, in statistical modeling utilizing multivariate EVT within IMA framework. This ensures consistently in matching extreme threshold values and graphical fields across all stages of extreme value analysis.

Throughout, our analysis will describe the spatial pattern of conditional frequency of extreme ECQ of precipitation over land, when conditioned on positive extreme field (we can also condition on negative extreme field with minor algorithm modification). The concept is similar to the high field energy that has been described in (Gilleland et al. 2013) and defined as upper quartile in space of the product of wind speed of updrafts and wind shear greater than its 90th percentile over time, to identify severe storm environments. We adapted this approach to accommodate specificity of our precipitation dataset and apply to detect high (or low) extreme field for precipitation (or drought) as the 99th (or the 1st) percentile in space greater than its 95th (or 5th) percentile in time. The formal definition will be given in Section 4.2. Furthermore, to evaluate areas with the largest conditional frequency values, we apply a new generalized spatio-temporal threshold selection methodology that uses time series clustering procedure for a number of geometric indices. This facilitates the algorithm's ability to classify a spatial pattern of an

image (here, conditional frequency of extreme ECQ of rainfall or drought) by a number of geometrical constructs described in Table 1.

Pixel	Smallest unit of a digital image
Isolated Structure	Collection of adjacent pixels with value of one
Area	Number of non-zero pixels in the structure
Convex Hull	Smallest convex polygon that contains the structure
Perimeter	Length of the outside boundary of the structure

Table 1: Geometrical constructs used in geometric indices

As far as we aware, this is a first published IMA in atmospheric science that links threshold selection process conditioned on extreme patterns with clustering of geometric index series.

The reminder of the paper is structured as follows. In Section 2, the dataset is described. The conceptual framework and all methods are reviewed in Sections 3 and 4. The detailed desription of all steps for spatio-temporal threshold selection algorithm starts with Section 4.1, where ECQ constructions is clarified. In Section 4.2, the dimension reduction techniques are explained and important definitions are formalized. The spatial domain methodology of mapping an image to its corresponding geometric indices is formulated in Section 4.3. We integrate geometric index series with clustering analysis in Section 4.4 and clustering analysis with a threshold choice in Section 4.5. An application on daily precipitation dataset for conditional frequency of extreme ECQ and important findings are highlighted in Section 5. Our conclusions are presented in Section 6.

2 Data

We will consider station-based precipitation measurements in mm/day obtained from the Global Historical Climatology Network-Daily (GHCN-Daily; (Menne et al. 2012b)) collected over the period 01/01/1961 to 12/31/2016 (56 years). While the GHCN-Daily dataset is subject to a rigorous quality assurance procedure (Menne et al. 2012a), it is not adjusted for artificial discontinuities such as changes in station location, instrumentation and time of observation (Donat et al. 2013b). To deal with this issues, we only select stations with at least 40% data coverage, yielding 8516 observations shown in Figure 1.

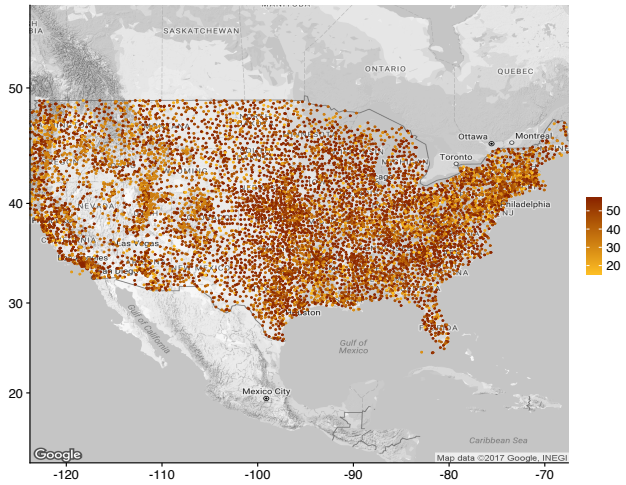

Figure 1: Spatial extent of GHCN-Daily dataset. The color shade indicates the length of the observation period (14-56 years).

In addition, as described in Liang et al. (2004), we perform topographical adjustments to account for the strong elevation dependence of the rain gauge sites

by using Parameter-Elevation Regression on Independent Slopes Model (PRISM) (Daly et al. 1994, 1997). We interpolate the GHCN-Daily dataset to a 30 km X 30 km grid covering the contiguous United States (CONUS), resulting over 20,000 days of spatial rainfall coverage.

3 Conceptual framework

The spatio-temporal threshold selection algorithm consist of four main steps:

1. ECQ construction is closely matched to the timeframe of interest.
2. Dimension reduction is based on statistical quantiles.
3. Spatial domain mapping is represented by geometric indices.
4. Threshold selection is linked to time series clustering.

For the purpose of this study, our focus is selecting a high threshold for frequency of extreme ECQ based on daily precipitation rates for daily (short term) rainfall events. Our algorithm is flexible enough to incorporate a low threshold choice and for longer lasting events as well. Next, we apply the dimension reduction step to identify extreme processes which are spread out over a fairly extended portion of spatial domain. These processes are used within the threshold selection framework and integrated with methods from digital topology. Values greater or equal than a chosen threshold are assigned to one, and values below to zero. This is called image digitizing, and is a widely used technique in computer imaging. The digitized image then can be mapped to a number of geometric indices that represent certain graphical quality of the image. As the threshold varies, so do the values of the geometric indices. As such, one can create threshold series that is

mapped to the corresponding geometric index series. The mapping process can be repeated for all geometric indices, creating a non-linear dependence between the threshold series and the desired geometrical properties. A time series clustering algorithm can be further used to select a representative threshold for each cluster.

Rather than considering extreme values at individual locations and their temporal dependence, we consider an overall spatial field that is conditioned on being extreme. In this case, it is possible to depict large-scale spatial extreme processes independent of whether or not individual grid cells in space are being extreme. The objective of the algorithm is to detect latent spatial dependencies for ECQ of interest within large scale extreme events and to automate threshold selection for extreme spatio-temporal processes. Ultimately, the identified spatial extremes can be linked to the individual weather patterns because the corresponding occurrences times of such extremes events are tracked during the identification process and the threshold choice for extreme events modeling can help to deepen our understanding of the underlying processes in forming those patterns that were possibly overlooked in conventional grid point by grid point methods.

4 Methods

4 ECQ construction

The testing process for the algorithm will be based on standardized anomalies over daily accumulation window (W). The proposed ECQ derivation involves the following steps:

1. Remove seasonal patterns by subtracting climatological mean values over the entire time length from the raw data.

2. Calculate W - days rolling accumulated precipitation anomalies.

3. Standardize accumulated values by the corresponding 95th percentile.

In this study, we select $W = 0$, that is, our interest lies in short term (i.e.1-day) extreme rainfalls. These short-duration impactful precipitation events can produce a major natural hazard such as flash flood in some areas. For example, a major rainstorm on September 2010 in Minnesota and Wisconsin is shown in Figure 2, caused flash flooding and forced evacuations. The flexibility of this step enables us to capture others, longer term, impactful extremes, such as September 2013 Colorado floods ($W = 1$, Figure 3), October 2012 superstorm Sandy ($W = 4$, Figure 4) and May to October great floods of 1993 ($W = 180$, Figure 5)

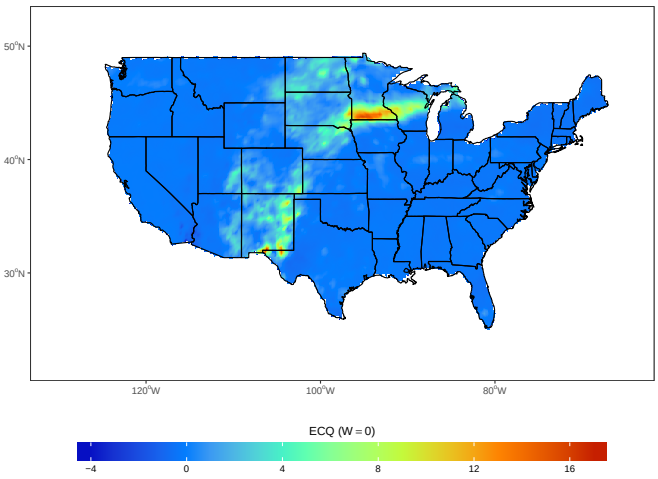

Figure 2: Minnesota and Wisconsin flash floods, September 2010 (9/22-9/23/10).

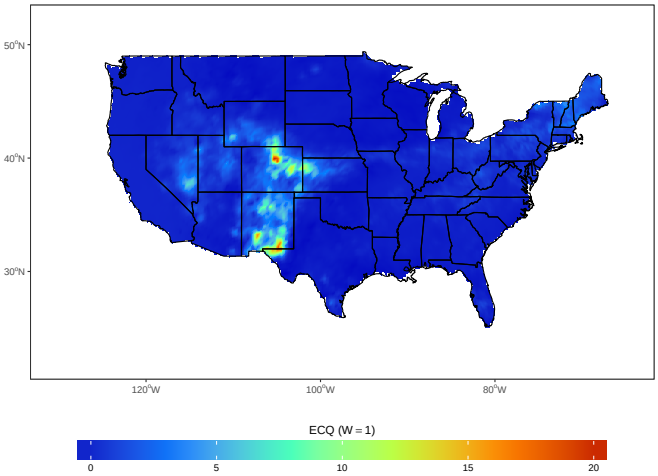

Figure 3: Colorado floods, September 2013 (9/11-9/12/13).

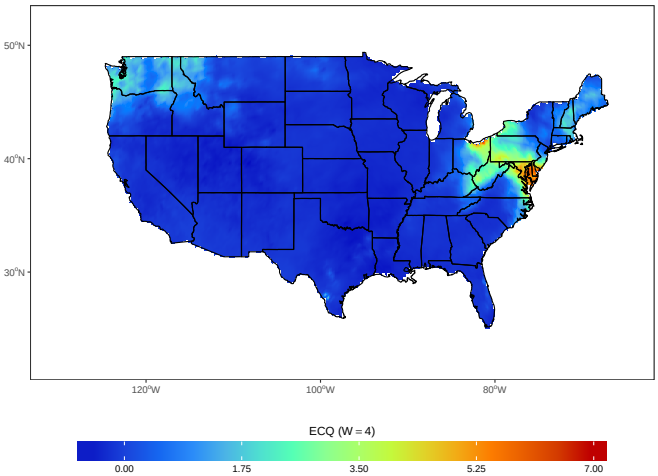

Figure 4: Superstorm Sandy, October 2012 (10/29-11/02/12).

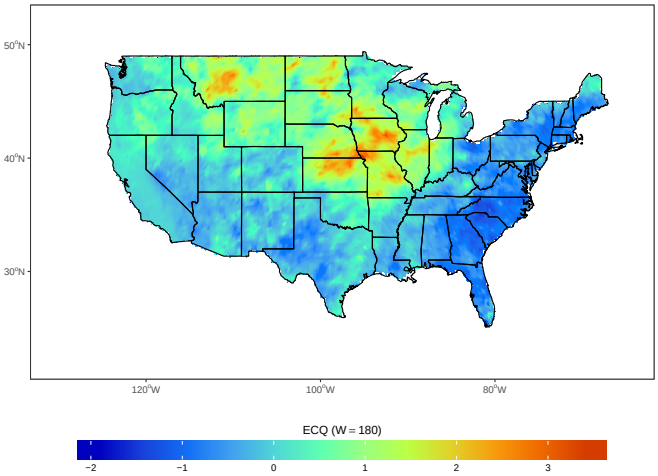

Figure 5: Great Flood of 1993, May - October 1993.

4 Dimension reduction

Traditional temporal dimension reduction methods such as principal component analysis (PCA) and linear discriminant analysis (LDA), which rely heavily on mean and covariance estimation, are not very informative for extremes. The block-maxima approach often used to analyze extremes potentially excludes relevant observations (Coles 2001). The challenge is therefore to subsample a relatively small number of unsmoothed observations to represent the tail of the distribution. Since our primary aim is to describe the spatial pattern of conditional frequency of extreme ECQ and its temporal evolution, we adapted a quantile-based dimension reduction methodology used by Gilleland et al. (2013) to analyze frequency of severe storm environments, and applied it for temporal and spatial quantile values more appropriate for the ECQ constructed from GHCN-Daily dataset. We further formalize this approach with the following definitions.

Definition 4.2.1. Let $Y(s,t)$ be a spatio-temporal ECQ, with number of spatial

pixels $s = [1,\ldots,S]$ and temporal samples $t = [1,\ldots,T]$, $\mathrm{Q}_{\mathrm{s}}^{(\eta)}(Y)$ is called as η-th quantile in space and $\mathrm{Q}_{\mathrm{t}}^{(\tau)}(Y)$ as τ-th quantile in time of ECQ Y, where $\eta, \tau \in [0,1]$.

Definition 4.2.2. Let $\mathrm{Q}_{\mathrm{s}}^{(\eta)}(Y)$ be a univariate time series representing η-th quantile in space of ECQ $Y(s,t)$, $\mathrm{Q}_{\mathrm{t}}^{(\tau)}[\mathrm{Q}_{\mathrm{s}}^{(\eta)}(Y)]$ is called as τ-th quantile in time of η-th quantile in space of ECQ Y.

Definition 4.2.3. Let $\mathrm{H}(s,t') \in Y$ be a spatio-temporal ECQ with $t' = [t'_1,\ldots,t'_{\mathrm{K}}]$, where $t' \in t$ and $K \leqslant T$, a process $\mathrm{H}(s,t')$ is called *Positive Extreme Field* (PEF) if $\mathrm{Q}_{\mathrm{s}}^{(\eta'_{\mathrm{h}})}(Y) > \mathrm{Q}_{\mathrm{t}}^{(\tau'_{\mathrm{h}})}[\mathrm{Q}_{\mathrm{s}}^{(\eta'_{\mathrm{h}})}(Y)]$ and $\eta'_{\mathrm{h}}, \tau'_{\mathrm{h}} \in [0.75, 1]$.

Definition 4.2.4. Let $\mathrm{L}(s,t') \in Y$ be a spatio-temporal ECQ with $t' = [t'_1,\ldots,t'_{\mathrm{K}}]$, where $t' \in t$ and $K \leqslant T$, a process $\mathrm{L}(s,t')$ is called *Negative Extreme Field* (NEF) if $\mathrm{Q}_{\mathrm{s}}^{(\eta'_{\mathrm{l}})}(Y) < \mathrm{Q}_{\mathrm{t}}^{(\tau'_{\mathrm{l}})}[\mathrm{Q}_{\mathrm{s}}^{(\eta'_{\mathrm{l}})}(Y)]$ and $\eta'_{\mathrm{l}}, \tau'_{\mathrm{l}} \in [0, 0.25]$.

Based on these definitions, we can formulate a conditional frequency for extreme ECQ as a function of high threshold θ_{h} as

$$f_{(\mathrm{Y|H})}(\theta_{\mathrm{h}}) = \frac{\sum_{i=1}^{S}\sum_{j=1}^{K}[\mathrm{H}(s_i,t'_j) > \theta_{\mathrm{h}}]}{K} \tag{1}$$

Alternatively, we can define similar quantity if a low threshold θ_{l} is of interest.

$$f_{(\mathrm{Y|L})}(\theta_{\mathrm{l}}) = \frac{\sum_{i=1}^{S}\sum_{j=1}^{K}[\mathrm{L}(s_i,t'_j) < \theta_{\mathrm{l}}]}{K} \tag{2}$$

For simplicity, we will omit the subscripts Y|H and Y|L when referring to the conditional frequency. In terms of additional notations, let $\theta = [\theta_1,\ldots,\theta_{\mathrm{n}}]$ be a set of initial thresholds and $\theta_1 < \theta_2 <,\ldots,< \theta_{\mathrm{n}}$. We can think of the series

$f(\theta_1),\ldots,f(\theta_n)$ as a procedure of partitioning a conditional frequency image into several overlapping homogeneous fields, where each field is mapped into the geometric indices described in the next section.

4 Spatial Domain Mapping

Geometric indices were first introduced by AghaKouchak et al. (2010) and applied to validate radar data against satellite precipitation estimates and weather prediction models. Gilleland (2017a) complemented geometric indices with mean-error and mean-square-error distances, to introduce new diagnostic plots within a spatial forecast verification framework. The diagnostics were applied to number of cases from the spatial forecast verification intercomparison project (ICP, `http://www.ral.ucar.edu/projects/icp`) (Ahijevych et al. 2009; Gilleland et al. 2009, 2010). The indices vary between zero and one and describe the connectivity, shape and area of the image pixels for a predefined threshold. The connectivity index is defined as

$$C_{\text{index}} = 1 - \frac{n-1}{\sqrt{m}+n}, \tag{3}$$

where n is the number of isolated structures, and m is the number of pixels with value of one (i.e. above a chosen threshold). The connectivity index shows how the structures within the image are interconnected. The higher (lower) the index the more connected (dispersed) the fields are within the image. The shape index is given by

$$S_{\text{index}} = \frac{P_{\min}}{P}, \tag{4}$$

where P is the perimeter of the pixels above a given threshold, and $P_{\min}$ is the theoretical minimum perimeter of an s-pixel pattern, which is attained if the

pattern of non-zero pixels was formed closest to a perfect circle. Mathematically, it is defined as

$$P_{\min} = \begin{cases} 4 \times \sqrt{s}, & \text{if } \lfloor \sqrt{s} \rfloor = \sqrt{s} \\ 2 \times (\lfloor 2 \times \sqrt{s} \rfloor + 1), & \text{otherwise} \end{cases},$$

where $\lfloor \cdot \rfloor$ is the floor function. The near one values of S_{index} imply approximately circular pattern of non-zero pixels. For more detailed discussion about these indices see Gilleland (2017a). Finally, to measure object complexity, we introduce another geometric index defined in Bullock et al. (2016) as follows

$$\mathsf{C}_{\text{index}} = 1 - \frac{A}{A_{\text{convex}}} = 1 - A_{\text{index}}, \tag{5}$$

where A is the area of the pixels (i.e. the number of non-zero pixels) above a given threshold, A_{convex} is the area of the convex hull around those pixels and A_{index} is defined as their ratio. The index values close to zero are representative of more structured image patterns, whereas fairly dispersive image patterns imply values near one.

As previously stated, past applications of geometric indices were performed primarily for model validations. However, we adapt this approach to the observational settings to evaluate geometric index valued for different thresholds. The aim is to determine specific geometrical properties that are relevant to the extreme ECQs, where high (or low) threshold values are derived via time series clustering procedure. That is, for every threshold θ_{i} from the set of initial thresholds θ, geometric indices are estimated and then grouped into the index series for time series clustering depicted in the following section. For brevity, we will represent

these index series as $\mathrm{I}_j[f(\theta)]$, where $j = [1,2,3]$ and $\mathrm{I}_j[\cdot]$ is an index operator for geometric index j of conditional frequency f at threshold θ. For example, $\mathrm{I}_1[\cdot] = C_{\text{index}}[\cdot]$, $\mathrm{I}_2[\cdot] = S_{\text{index}}[\cdot]$ and $\mathrm{I}_3[\cdot] = \complement_{\text{index}}[\cdot]$. An example of evolution of geometric indices as a function of a threshold is shown in Figure 6.

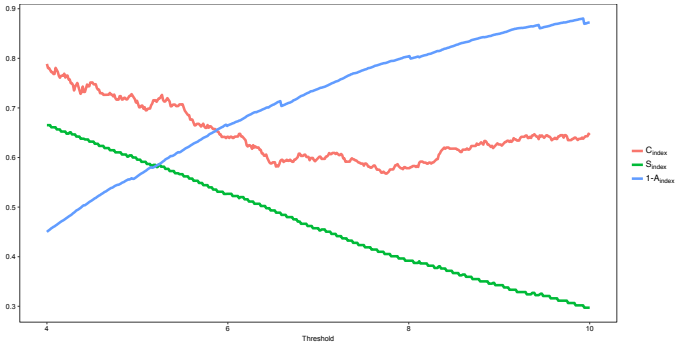

Figure 6: Geometric indices evolution as a function of threshold for summer ECQ (GHCN-Daily dataset).

Analyses are performed using the `SpatialVx` (Gilleland 2017b) package in R.

4 Time Series Clustering

Clustering is a Machine Learning process of grouping unlabeled data into homogeneous segments or clusters. Segmentation is performed in such a way that inter-group dissimilarity is maximized and intra-group similarity is minimized according to objective criterion. The degree of dissimilarity (or similarity) between the clustered objects is of major importance in cluster analysis. Partitioning in the clustering method entirely depends on a distance or dissimilarity metric, which measures how far away two objects are from each other. Common dissimilarity metrics such as Euclidean and Manhattan are not suitable for time series clustering because they ignore serial correlations within the time series. Our main aim

however is to investigate the spatio-temporal features of sequences of conditional frequency images that may exhibit high degrees of serial correlation at different lag times. Therefore, the first fundamental task in applying the clustering method is the selection of a dissimilarity metric that can capture these subtle characteristics.

The following dissimilarity measures have been adapted from Galeano and Peña (2000) and Caiado et al. (2006). Given $f(\theta)$ mapping to $\mathrm{I}_j[f(\theta)]$, where $j = [1,2,3]$, we let $\mathrm{I}_l[f(\theta)]$ and $\mathrm{I}_m[f(\theta)]$ to be two geometric index series, that is $\mathrm{I}_l[f(\theta)] = \mathrm{I}_l[f(\theta_1)],\ldots,\mathrm{I}_l[f(\theta_n)]$ and $\mathrm{I}_m[f(\theta)] = \mathrm{I}_m[f(\theta_1)],\ldots,\mathrm{I}_m[f(\theta_n)]$ for $l \neq m$. We denote estimated autocorrelation vectors of I_l and I_m by $\widehat{\rho}_{I_\mathrm{l}} = (\widehat{\rho}_{1,I_\mathrm{l}} \ldots \widehat{\rho}_{R,I_\mathrm{l}})$ and $\widehat{\rho}_{I_\mathrm{m}} = (\widehat{\rho}_{1,I_\mathrm{m}} \ldots \widehat{\rho}_{R,I_\mathrm{m}})$ for some R such as $\widehat{\rho}_{i,I_\mathrm{l}} \approx 0$ and $\widehat{\rho}_{i,I_\mathrm{m}} \approx 0$ for $i > R$. Dissimilarity between I_l and I_m is measured by the following Euclidean distance:

$$d_{ACF}(I_\mathrm{l}, I_\mathrm{m}) = \sqrt{\sum_{i=1}^{R} (\widehat{\rho}_{i,I_\mathrm{l}} - \widehat{\rho}_{i,I_\mathrm{m}})^2} \tag{6}$$

Analogously, we define $\widehat{\phi}_{ii,I_\mathrm{l}}$ and $\widehat{\phi}_{ii,I_\mathrm{m}}$ as the estimated partial autocorrelations of I_l and I_m. A corresponding dissimilarity metric is given by

$$d_{PACF}(I_\mathrm{l}, I_\mathrm{m}) = \sqrt{\sum_{i=1}^{R} (\widehat{\phi}_{ii,I_\mathrm{l}} - \widehat{\phi}_{ii,I_\mathrm{m}})^2} \tag{7}$$

We use the `TSclust` package in R (Montero and Vilar 2014) to calculate both distances. Choice of the number of clusters K is another important task in the clustering algorithm. It typically involves an extensive search for number of clusters greater than one, which is further evaluated by the a variety of criterion measures. For example, the NbClust() function in the `NBClust` package in R (Charrad et al. 2014) gives user a choice between 30 different indices to judge a quality of clustering solution. Generally, cluster selections is still an unresolved

problem requiring multiple iterations. This step, however, is somewhat simpler in our case. Given the initial threshold series θ, which is constructed in the ascending order, the clustering choice will be limited to the three clusters only ($K = 3$), we call them: *low extreme*, *penultimate extreme* and *extreme* with particular emphasis on the *penultimate extreme* (here, second) cluster. This decision is motivated by the tradeoff between bias and variance. After the clustering solution is found, the first (*low extreme*) cluster should have threshold values that are too low for the extreme values selection and the third (*extreme*) cluster should have threshold values that are too high. If the threshold choice is made from the first cluster, bias can occur and the choice from the third cluster could lead to increased variance in the estimators. We therefore focus on the second cluster in order to minimize both bias and variance impact. Having decided on the number of clusters and the dissimilarity measures, makes it possible to describe clustering algorithm. The most popular partitioning clustering approaches are k-means and partitioning around medoids (PAM) (Kaufman and Rousseeuw 1987). K-means creates cluster centers by averaging points within the cluster. The averaging process can be sensitive to outliers and it also breaks the max-stability property, so that mean of two maxima is no longer a maximum, which is an important assumption in EVT. PAM, implemented in `cluster` package in R (Maechler et al. 2017), provides a more robust outcome, where each cluster is identified by its most representative object called a medoid. Additionally, it preserves observational features of the dataset, which is important for clustering extremes. The algorithm has two phases, BUILD and SWAP:

1. In the first phase, k representative objects are selected to form initial clustering set.

2. In the second phase, an attempt is made to improve clustering by exchanging selected and unselected objects.

The objective of PAM is for all selected clusters to minimize average dissimilarity between their centrally located representative object and any other object in the same cluster. Further details can be found in Kaufman and Rousseeuw (1990). The quality of the resulting clusters and the choice of number of clusters can be assessed using the so-called "silhouette coefficient" (Rousseeuw 1987). A silhouette coefficient is a useful statistic for determining how similar an object is to its own cluster (cohesion) compared to other clusters (separation). For an object i it is defined as follows,

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{8}$$

where $a(i)$ is the average dissimilarity of object i to all other objects in its cluster and $b(i)$ is the minimum average distance of object i to all other objects in the given cluster not containing i. The value of the silhouette coefficient $s(i)$ ranges between -1 and 1, where a negative value is undesirable. If $s(i) \approx 1$, a strong structure has been found, which means that inter-cluster distance is much larger than intra-cluster distance. Conversely, if $s(i) < 0.25$, no substantial structure has been found. The average silhouette coefficient $\bar{s}(K)$ can be used to evaluate a quality of segregation into K clusters, such as, for any $K = [1, \ldots, n]$ find $\max_K \bar{s}(K)$.

4 Threshold Selection

It is worth mentioning that there is a considerable amount of literature on threshold selection for image segmentation (e.g. Sezgin et al. 2004; Sahoo et al. 1988), a

process of image partitioning into meaningful regions, which is somewhat similar to our procedure. This may include such computer vision tasks as finding a single threshold to separate an object in an image from its corresponding background, or segregation of light and dark regions, or in our case, it involves identifying distinct extreme episodes that spread out over a large portion of spatial domain. Though offering invaluable information, these methods are not informative about spatial patterns (Kaur and Kaur 2014) and are thus not appropriate for complex images (i.e. color images where multiband thresholding may be necessary) (Yogamangalam and Karthikeyan 2013).

Our threshold choice is directly linked to the outcome of the time series clustering step. The clustering solution produces results for the three clusters and their medoids, where every member is linked to the threshold and corresponding values of three geometric indices. As mentioned in previous section, we address bias/variance trade-off by selecting members from the second cluster only. The threshold selection is implemented as follows. First, we select maximum average silhouette coefficient between $\bar{S}_{\text{ACF}}$ and $\bar{S}_{\text{PACF}}$ and call it $\bar{S}_{\text{best}}$. This is necessary to select a best clustering solution between two distance measures. Next, the medoid from the second cluster is matched by the $\bar{S}_{\text{best}}$. This will correspond to our threshold choice. Finally, the medoid index in the overall time series is used to determining corresponding geometric indices. Equipped with both, the high spatio-temporal threshold value and the spatial domain snapshot represented by geometric indices, the user can make more informative choice on the threshold selection for the extremes and its applications.

5 Results

We applied the threshold selection algorithm to the daily ECQ for precipitation over CONUS. The heavy rainfall is seasonally dependent and varies in space and time. Thus, in order to understand the complete evolution of $f(\theta)$ and its spatial extent, we stratified the dataset by seasons defined as winter (DJF), spring (MAM), summer (JJA), and fall (SON). For every season, we automatically determined a high threshold value using the geometric index series clustering procedure. Each cluster strength was analyzed with the average silhouette coefficient; values near one suggest strong clustering structures. Every threshold value was mapped to a corresponding set of geometric indices, revealing various geometrical properties of the spatial image. In addition, to evaluate performance of the algorithm on how well three clusters are separated and the resultant geometrical properties of the conditional frequency of extreme ECQ, we characterized the results for GPD fit in the following manner:

Cluster 1. Low extreme. Sampling from this cluster will most likely result in approximation bias.

Cluster 2. Penultimate extreme. The threshold values from this cluster can be used for GPD analysis.

Cluster 3. Extreme. Selection from this cluster may lead to high variance in the estimators because of lack of data points.

The outcome of this process for all three clusters is displayed in Tables 2 - 4.

Months	θ	$\bar{s}(K)$	C_{index}	S_{index}	$\complement_{\text{index}}$	A
DJF	7.07	0.51	0.69	0.44	0.75	2218
MAM	9.46	0.49	0.61	0.40	0.77	1796
JJA	9.42	0.51	0.58	0.40	0.80	1953
SON	9.58	0.51	0.63	0.45	0.73	2594

Table 2: Results of spatio-temporal threshold selection algorithm for second (penultimate extreme) cluster.

Months	θ	$\bar{s}(K)$	C_{index}	S_{index}	$\complement_{\text{index}}$	A
DJF	4.05	0.79	0.80	0.65	0.48	5787
MAM	5.62	0.79	0.74	0.62	0.49	5180
JJA	6.67	0.72	0.69	0.56	0.62	3970
SON	6.76	0.74	0.74	0.62	0.50	5423

Table 3: Results of spatio-temporal threshold selection algorithm for first (non-extreme) cluster.

Months	θ	$\bar{s}(K)$	C_{index}	S_{index}	$\complement_{\text{index}}$	A
DJF	9.09	0.56	0.71	0.37	0.79	1173
MAM	10.35	0.56	0.64	0.31	0.86	989
JJA	11.52	0.58	0.64	0.32	0.86	1209
SON	11.27	0.58	0.63	0.36	0.82	1597

Table 4: Results of spatio-temporal threshold selection algorithm for third (extreme) cluster.

Since our main interest lies in the spatial distribution of frequency for ex-

treme ECQ conditioned on PEF, we performed inter-cluster comparison to further demonstrate that the three clusters are well separated. The first, *low extreme*, cluster depicted in Table 3 has higher connectivity and shape index, in addition to a larger areal extent and smaller complexity than the other two cluster. Utility of this cluster is to sample relatively small threshold values. not appropriate for extreme value analysis. Comparing Tables 2 and 4 (Cluster 2 and Cluster 3) reveals that the frequency of extreme ECQ in the *penultimate extreme* cluster has the following characteristics: lower connectivity (all numbers have been rounded to two digits) and higher shape index, with larger areal extent and smaller complexity. The same information can be visually validated in Figure 7.

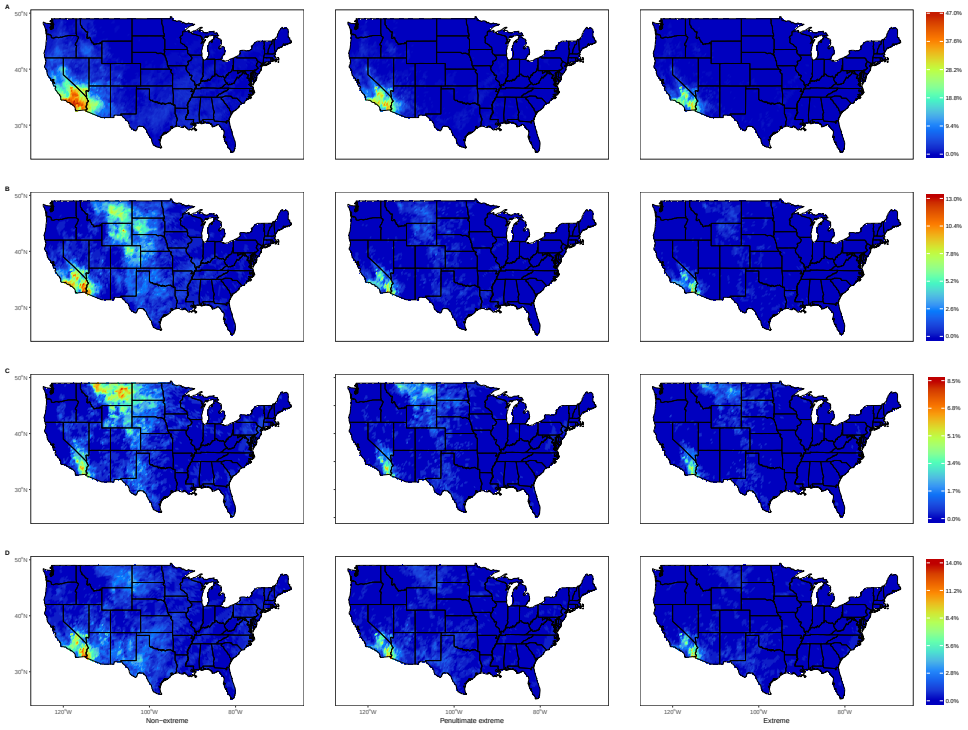

Figure 7: Frequency of ECQ conditioned on PEF with $W = 0$ for GHCN-Daily dataset. Left column - *low extreme* cluster with threshold values from Table 3; middle column - *penultimate extreme* cluster with threshold values from Table 2, and right column - *extreme cluster* with threshold values from Table 3, for every season: Winter (top row), Spring (second row), Summer (third row), and Fall (bottom row).

The main advantage of having larger areal extent and lower connectivity in spatial forecast verification for extremes is that it provides a more coherent representation of extremes in space and, as a result, allows for a more informative decision about the causes, severity and impact of the extremes under climate change, assuming climate models can simulate these patterns reasonable well. These graphical qualities make the *penultimate extreme* cluster the best choice for higher threshold selection for extreme value analysis.

In addition, this threshold selection approach illustrates a fundamental difference between the methodology adopted here and other approaches for defining extreme event frequency. It is a common practice for threshold choice in univariate EVT to carry out an exploratory analysis or stability assessment of estimated parameters in temporal domain without considering any spatial dependencies. In our case, having frequency for extreme ECQ conditioned on PEF algorithmically determined, allows for the analysis of temporal patterns within a fairly large spatial field, notwithstanding that many of the individual grid cells are not necessarily extreme. Each box plot in Figure 8 summarizes one of the geometric indices and the areal extent distribution when the dataset was stratified by year.

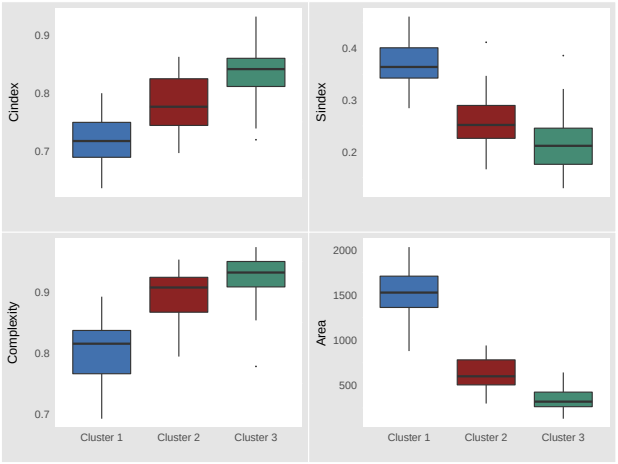

Figure 8: Time series of C_{index} (top left), S_{index} (top right), $\mathsf{C}_{\text{index}}$ (buttom left), and A (bottom right) for annual ECQ conditional frequency for three clusters in the threshold selection algorithm.

Unsurprisingly, for the *penultimate extreme* cluster, the spatial and temporal dependence of the geometric indices and the area of frequency for extreme ECQ conditioned on PEF determined by the threshold selection algorithm is different from two other clusters. Of course, one important question to be addressed from these analyses is whether the difference is statistically significant. We compared the second cluster to remanding two clusters by performing a two-sided Wilcoxon rank sum test. The results were all significant at the 99th percentile confidence level. Clearly, our clustering process characterized clustering structure for *low extreme*, *penultimate extreme* and *extreme* clusters well, which is also confirmed by the relatively high average silhouette coefficients in the corresponding tables. Therefore, a threshold selection from three different clusters for extreme value analysis could lead to different outcomes in spatial forecast verification, and ultimately to different statistical inferences regarding factors that explain extreme weather and climate events.

6 Conclusions

Lack of universal definition of what we mean by the word "extreme" creates a difficulty when comparing different scientific studies of extreme events. This makes statistical inference much harder and obfuscates transparency and risk management process. The current study is an attempt to formalize this important concept. We introduced a new spatio-temporal threshold selection algorithm for extreme events within IMA framework. With the help of three clusters, we demonstrated that the results for the *penultimate extreme* cluster are statistically significant when compared to two other, *low extreme*, and *extreme*, clusters.

The main advantages of this algorithm are the following. It automates the threshold selection process in the objective way and can ultimately be used in conjunction with spatial forecast verification and modeling of extreme events. It is adaptable to model extremes with both high and low threshold choices. It incorporates spatial and temporal dependence in one holistic modeling framework thus opening an opportunity for future analysis of statistical inference of extreme events for univariate and possibly multivariate ECQ, which is not possible with traditional grid point by grid point methods. It also links a threshold choice with desired geometrical properties therefore offering a user a more informative and flexible choices in selecting extreme events. The algorithm is relatively fast and could be embedded within weather modeling systems to identified synoptic scale spatial patterns that can be linked to the individual extreme episodes.

Furthermore, an interesting perspective for this work is to determine whether or not, from the ensemble of different climate models and multiple statistical simulations, one can capture dynamic and thermodynamic properties of the spatio-

temproal structure of $f(\theta)$. Another perspective is to investigate how the observed spatio-temporal patterns for frequency of extreme ECQ have changed and the causes of those changes, utilizing the most recent detection and attribution framework (e.g. Knutson et al. 2017; Ribes et al. 2017).

References

AghaKouchak, A., N. Nasrollahi, J. Li, B. Imam, and S. Sorooshian, 2010: Geometrical characterization of precipitation patterns. *J. Hydrometeor*, **12 (2)**, 274–285, doi:10.1175/2010JHM1298.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2010JHM1298.1.

Ahijevych, D., E. Gilleland, B. G. Brown, and E. E. Ebert, 2009: Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Weather and Forecasting*, **24 (6)**, 1485–1497, doi:10.1175/2009WAF2222298.1, URL http://journals.ametsoc.org/doi/full/10.1175/2009WAF2222298.1.

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research*, **111 (D5)**, doi:10.1029/2005JD006290, URL http://doi.wiley.com/10.1029/2005JD006290.

Bader, B., J. Yan, and X. Zhang, 2017: Automated selection of r for the r largest order statistics approach with adjustment for sequential testing. *Statistics and Computing*, **27 (6)**, 1435–1451, doi:10.1007/s11222-016-9697-3, URL http://link.springer.com/10.1007/s11222-016-9697-3.

Balkema, A. A., and L. de Haan, 1974: Residual Life Time at Great Age. *The Annals of Probability*, **2 (5)**, 792–804, doi:10.1214/aop/1176996548, URL http://projecteuclid.org/euclid.aop/1176996548.

Banerjee, S., B. P. Carlin, and A. E. Gelfand, 2015: *Hierarchical modeling and*

analysis for spatial data. Second edition ed., No. 135, Monographs on statistics and applied probability, CRC Press, Taylor & Francis Group, Boca Raton.

Bojinski, S., M. Verstraete, T. C. Peterson, C. Richter, A. Simmons, and M. Zemp, 2014: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. *Bulletin of the American Meteorological Society*, **95 (9)**, 1431–1443, doi:10.1175/BAMS-D-13-00047.1, URL http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-13-00047.1.

Bullock, R., B. Brown, and T. Fowler, 2016: Method for Object-Based Diagnostic Evaluation. Tech. rep., University Corporation For Atmospheric Research (UCAR):National Center for Atmospheric Research (NCAR):NCAR Library (NCARLIB). URL https://opensky.ucar.edu/islandora/object/technotes:546, dOI: 10.5065/D61V5CBS.

Caiado, J., N. Crato, and D. Peña, 2006: A periodogram-based metric for time series classification. *Computational Statistics & Data Analysis*, **50 (10)**, 2668–2684, doi:10.1016/j.csda.2005.04.012, URL http://linkinghub.elsevier.com/retrieve/pii/S0167947305000770.

Charrad, M., N. Ghazzali, V. Boiteau, and A. Niknafs, 2014: NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, **61 (6)**, 1–36, URL http://www.jstatsoft.org/v61/i06/.

Coles, S., 2001: *An introduction to statistical modeling of extreme values*. Springer series in statistics, Springer, London ; New York.

Cooley, D., J. Cisewski, R. J. Erhardt, S. Jeon, E. Mannshardt, B. O. Omolo, and

Y. Sun, 2012: A survey of spatial extremes: Measuring spatial dependence and modeling spatial effects. *Revstat*, **10 (1)**, 135–165.

Dai, A., 2001: Global Precipitation and Thunderstorm Frequencies. Part I: Seasonal and Interannual Variations. *Journal of Climate*, **14 (6)**, 1092–1111, doi: 10.1175/1520-0442(2001)014<1092:GPATFP>2.0.CO;2.

Dai, A., R. M. Rasmussen, C. Liu, K. Ikeda, and A. F. Prein, 2017: A new mechanism for warm-season precipitation response to global warming based on convection-permitting simulations. *Climate Dynamics*, doi: 10.1007/s00382-017-3787-6.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, **33 (2)**, 140–158, doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Daly, C., G. Taylor, and W. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. Preprints. 10th AMS Conf. on Applied Climatology, Reno, NV. *American Meteorological Society*, 10–12.

Donat, M., L. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar, 2013a: Global Land-Based Datasets for Monitoring Climatic Extremes. *Bulletin of the American Meteorological Society*, **94 (7)**, 997–1006, doi:10.1175/BAMS-D-12-00109.1.

Donat, M. G., and Coauthors, 2013b: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset: HADEX2-GLOBAL GRIDDED CLIMATE EXTREMES.

Journal of Geophysical Research: Atmospheres, **118 (5)**, 2098–2118, doi:10.1002/jgrd.50150, URL http://doi.wiley.com/10.1002/jgrd.50150.

Friederichs, P., and T. L. Thorarinsdottir, 2012: Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction: VERIFICATION FOR EXTREME VALUE DISTRIBUTIONS. *Environmetrics*, **23 (7)**, 579–594, doi:10.1002/env.2176, URL http://doi.wiley.com/10.1002/env.2176.

Fukutome, S., M. A. Liniger, and M. Süveges, 2015: Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in switzerland. *Theoretical and Applied Climatology*, **120 (3)**, 403–416, doi:10.1007/s00704-014-1180-5, URL https://doi.org/10.1007/s00704-014-1180-5.

Galeano, P., and D. P. Peña, 2000: Multivariate analysis in vector time series. *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo*, **4 (4)**, 383–403.

Gilleland, E., 2017a: A New Characterization within the Spatial Verification Framework for False Alarms, Misses, and Overall Patterns. *Weather and Forecasting*, **32 (1)**, 187–198, doi:10.1175/WAF-D-16-0134.1, URL http://journals.ametsoc.org/doi/10.1175/WAF-D-16-0134.1.

Gilleland, E., 2017b: *SpatialVx: Spatial Forecast Verification*. URL https://CRAN.R-project.org/package=SpatialVx, r package version 0.6-1.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of Spatial Forecast Verification Methods. *Weather and Fore-*

casting, **24 (5)**, 1416–1430, doi:10.1175/2009WAF2222269.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009WAF2222269.1.

Gilleland, E., D. A. Ahijevych, B. G. Brown, and E. E. Ebert, 2010: Verifying Forecasts Spatially. *Bulletin of the American Meteorological Society*, **91 (10)**, 1365–1373, doi:10.1175/2010BAMS2819.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2010BAMS2819.1.

Gilleland, E., B. Brown, and C. Ammann, 2013: Spatial extreme value analysis to project extremes of large-scale indicators for severe weather: SPATIAL EXTREMES OF LARGE-SCALE PROCESSES. *Environmetrics*, **24 (6)**, 418–432, doi:10.1002/env.2234, URL http://doi.wiley.com/10.1002/env.2234.

Groisman, P. Y., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvaev, 2005: Trends in Intense Precipitation in the Climate Record. *Journal of Climate*, **18 (9)**, 1326–1350, doi:10.1175/JCLI3339.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI3339.1.

IPCC, 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp.

Jolliffe, I. T., and D. B. Stephenson, Eds., 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ.

Kaufman, L., and P. Rousseeuw, 1987: Clustering by means of medoids, *Statistical data analysis based on the L_1-norm and related methods*. North-Holland; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co, Amsterdam ; New York : New York, N.Y., U.S.A, 405–416 pp.

Kaufman, L., and P. J. Rousseeuw, Eds., 1990: *Partitioning Around Medoids (Program PAM)*. John Wiley & Sons, Inc., Hoboken, NJ, USA, 68–125 pp., URL http://doi.wiley.com/10.1002/9780470316801.ch2, dOI: 10.1002/9780470316801.ch2.

Kaur, D., and Y. Kaur, 2014: Various image segmentation techniques: A review. *International Journal of Computer Science and Mobile Computing*, **3 (5)**, 809–814.

Knutson, T., and Coauthors, 2017: Ch. 3: Detection and Attribution of Climate Change. Climate Science Special Report: Fourth National Climate Assessment, Volume I. Tech. rep., U.S. Global Change Research Program, – pp. DOI: 10.7930/J01834ND.

Kunkel, K. E., D. R. Easterling, D. A. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2010: Recent increases in U.S. heavy precipitation associated with tropical cyclones: HEAVY PRECIPITATION AND TROPICAL CYCLONES. *Geophysical Research Letters*, **37 (24)**, n/a–n/a, doi:10.1029/2010GL045164, URL http://doi.wiley.com/10.1029/2010GL045164.

Kunkel, K. E., D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2012: Meteorological Causes of the Secular Variations in Observed Extreme Precipitation Events for the Conterminous United States. *Journal of*

Hydrometeorology, **13 (3)**, 1131–1141, doi:10.1175/JHM-D-11-0108.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-0108.1.

Kunkel, K. E., D. R. Easterling, K. Redmond, and K. Hubbard, 2003: Temporal variations of extreme precipitation events in the United States: 1895-2000: TEMPORAL VARIATIONS IN THE U.S. *Geophysical Research Letters*, **30 (17)**, n/a–n/a, doi:10.1029/2003GL018052, URL http://doi.wiley.com/10.1029/2003GL018052.

Kunkel, K. E., T. R. Karl, and D. R. Easterling, 2007: A Monte Carlo Assessment of Uncertainties in Heavy Precipitation Frequency Variations. *Journal of Hydrometeorology*, **8 (5)**, 1152–1160, doi:10.1175/JHM632.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JHM632.1.

Lau, W. K.-M., H.-T. Wu, and K.-M. Kim, 2013: A canonical response of precipitation characteristics to global warming from CMIP5 models: PRECIPITATION AND GLOBAL WARMING. *Geophysical Research Letters*, **40 (12)**, 3163–3169, doi:10.1002/grl.50420, URL http://doi.wiley.com/10.1002/grl.50420.

Liang, X.-Z., L. Li, K. E. Kunkel, M. Ting, and J. X. Wang, 2004: Regional climate model simulation of US precipitation during 1982–2002. Part I: Annual cycle. *Journal of Climate*, **17 (18)**, 3510–3529, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(2004)017%3C3510%3ARCMSOU%3E2.0.CO%3B2.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik, 2017: *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.6 — For new features, see the 'Changelog' file (in the package source).

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston,

2012a: An Overview of the Global Historical Climatology Network-Daily Database. *Journal of Atmospheric and Oceanic Technology*, **29 (7)**, 897–910, doi:10.1175/JTECH-D-11-00103.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-11-00103.1.

Menne, M. J., and Coauthors, 2012b: Global Historical Climatology Network - Daily (GHCN-Daily), Version 3. NOAA National Climatic Data Center, URL http://www.nodc.noaa.gov/archivesearch/catalog/search/resource/details.page?uuid=gov.noaa.ncdc:C00861, dOI: 10.7289/v5d21vhz.

Montero, P., and J. A. Vilar, 2014: TSclust: An R package for time series clustering. *Journal of Statistical Software*, **62 (1)**, 1–43, URL http://www.jstatsoft.org/v62/i01/.

Pendergrass, A. G., and D. L. Hartmann, 2014: Changes in the Distribution of Rain Frequency and Intensity in Response to Global Warming*. *Journal of Climate*, **27 (22)**, 8372–8383, doi:10.1175/JCLI-D-14-00183.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-14-00183.1.

Pickands, J., 1975: Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, **3 (1)**, 119–131, doi:10.1214/aos/1176343003, URL http://projecteuclid.org/euclid.aos/1176343003.

Ribes, A., F. W. Zwiers, J.-M. Azaïs, and P. Naveau, 2017: A new statistical approach to climate change detection and attribution. *Climate Dynamics*, **48 (1-2)**, 367–386, doi:10.1007/s00382-016-3079-6, URL http://link.springer.com/10.1007/s00382-016-3079-6.

Rousseeuw, P. J., 1987: Silhouettes: A graphical aid to the interpretation and

validation of cluster analysis. *Journal of Computational and Applied Mathematics*, **20**, 53–65, doi:10.1016/0377-0427(87)90125-7, URL http://linkinghub.elsevier.com/retrieve/pii/0377042787901257.

Sahoo, P. K., S. Soltani, and A. K. Wong, 1988: A survey of thresholding techniques. *Computer vision, graphics, and image processing*, **41 (2)**, 233–260.

Scarrott, C., and A. MacDonald, 2012: A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Statistical Journal*, **10 (1)**, 33–60, URL https://www.ine.pt/revstat/pdf/rs120102.pdf.

Sezgin, M., and Coauthors, 2004: Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, **13 (1)**, 146–168.

Smith, R. L., 1987: Estimating Tails of Probability Distributions. *The Annals of Statistics*, **15 (3)**, 1174–1207, doi:10.1214/aos/1176350499, URL http://projecteuclid.org/euclid.aos/1176350499.

Stephenson, D., 2008: Definition, diagnosis, and origin of extreme weather and climate events. In *Climate Extremes and Society*, Diaz, Henry F. and Murnane, Richard J. (eds.). Cambridge University Press, Cambridge, New York, 348.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2007: How Often Will It Rain? *Journal of Climate*, **20 (19)**, 4801–4818, doi:10.1175/JCLI4263.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI4263.1.

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*. 3rd ed., No. v. 100, International geophysics series, Elsevier/Academic Press, Amsterdam ; Boston.

Yogamangalam, R., and B. Karthikeyan, 2013: Segmentation techniques comparison in image processing. *International Journal of Engineering and Technology (IJET)*, **5 (1)**, 307–313.

Zhang, X., L. Alexander, G. C. Hegerl, P. Jones, A. K. Tank, T. C. Peterson, B. Trewin, and F. W. Zwiers, 2011: Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change*, **2 (6)**, 851–870, doi:10.1002/wcc.147, URL http://doi.wiley.com/10.1002/wcc.147.