4-20-2015

A Seasonal Analysis of Extreme Precipitation Trends in the Contiguous United States

Rachael Balstad

Jaechoul Lee

A Seasonal Analysis of Extreme Precipitation Trends in the Contiguous United States

RACHAEL BALSTAD & JAECHOUL LEE, Ph.D. (FACULTY ADVISOR)

MATHEMATICS DEPARTMENT | BOISE STATE UNIVERSITY

Introduction

In recent years, instances of extreme weather have increased in both frequency and severity in comparison to years past, from record high and low temperatures to flooding and landslides. The two latter phenomena are often a byproduct of extreme precipitation. The National Oceanic and Atmospheric Administration (NOAA), a branch of the National Weather Service (NWS), estimates that, over the last 30 years, floods alone have caused the United States and its territories suffer a mean of 82 fatalities per year in addition to damages averaging $7.96 billion each year[1].

While many mathematicians and scientists have studied trends in average precipitation, extremes are of particular interest because of their direct (and often dire) consequences on both the physical and financial well-being of humans. Statistically, averages and extremes are independent of one another, meaning that a trend in averages will not necessarily indicate an analogous trend in extremes. Thus, it stands to reason that extremes be studied independently and in addition to average precipitation data.

In a 1998 study, Kunkel, Andsager, and Easterling published a study of extreme precipitation events (in durations of 1-7 days with 1-year recurrence intervals) across the United States from 1931-1996, noting that these events "are highly correlated with hydrologic flooding in some U.S. regions" [2]. They have concluded, using simple linear analysis, that the occurrence of these events is increasing in some regions at a highly statistically significant rate. For instance, the east north central region experienced an increase that is statistically significant at the 5% level, while in the Midwest and Great Lakes region, upward trends were locally significant. Another 2002 study by Kunkel noted that one potential cause of the increase of extreme precipitation events can be attributed to an increase in temperature, which in turn increases saturation water vapor pressure. As a result, it is likely that atmospheric water vapor content will increase. With more water vapor available, systems experience a greater likelihood of extreme precipitation events. However, Kunkel notes that while this increase in extreme precipitation is more or less global, these changes are not likely to occur homogeneously due to factors such as "moisture availability, thermodynamic instability,...[and] the frequency and intensity of precipitation-producing meteorological systems" [3].

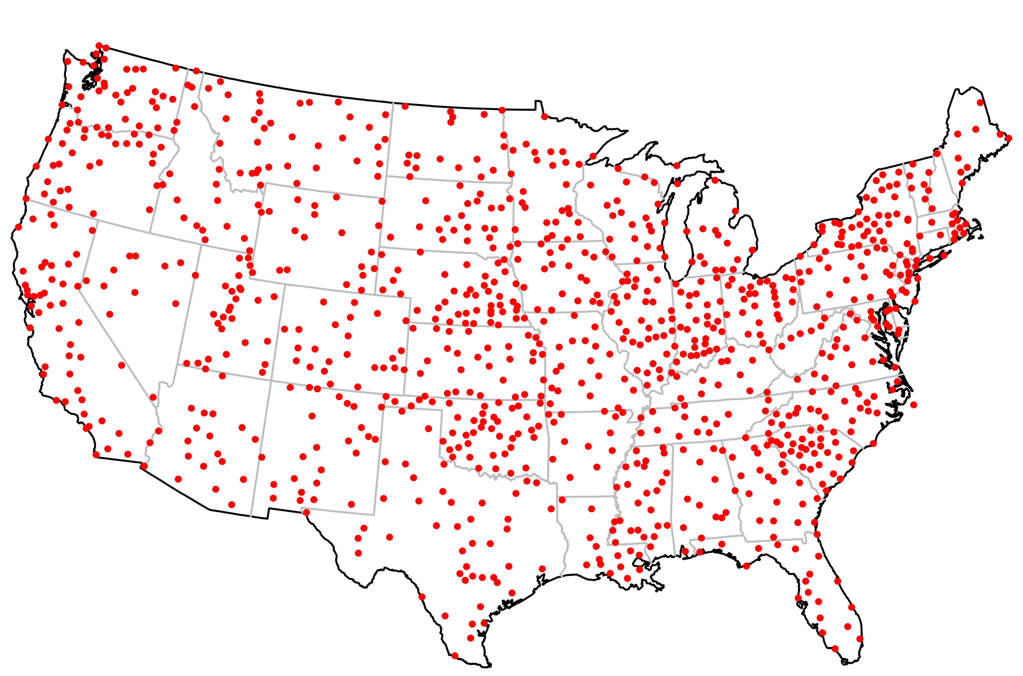

Figure 1: Location of the 1,026 stations used to examine seasonal extreme precipitation in this study

Data

The data for this study were cultivated out of the United States Historical Climatology Network (USHCN), a detailed collaboration of 1,218 stations across the contiguous United States. This data set contains both daily and monthly records which include maximum and minimum temperature, precipitation, snowfall, and snow depth. The data are found on the Carbon Dioxide Information Analysis Center (CDIAC) website[4]. The 1,218 stations that are a part of the USHCN have been chosen in observation of record longevity, coverage of the contiguous United States, percentage of missing data points, and number of significant station changes (such as new equipment or new location). The data were put through quality control procedures which checked for internal consistency, outliers, frequent values, and spatial consistency.

For the purposes of this study, we used the daily precipitation data, which was first treated to remove obvious erroneous entries by ensuring no data in the set was above the record high for daily precipitation, and that precipitation values were all nonnegative. Seasonal maxima were selected as the data was cropped into seasonal blocks (spring: March, April, May; summer: June, July, August; fall: September, October, November; and winter: December, January, February). Next, stations that did not meet one of the following criteria were precluded from the set:

- A record of at least 75 years with up to 40% missing data
- A record of 45 to 75 years with up to 10% missing data

A season was considered to be "missing" if one or more of the daily observations from that season were missing.

Figure 1 depicts spatial location of the 1,026 stations used in our analysis. This covers most of the contiguous United States, with somewhat sparse coverage in South Nevada and the southeastern Oregon.

Methods

There are several possible approaches that can be used for analyzing extreme values. In this study, data are assumed to follow a generalized extreme value (GEV) distribution. Maximum likelihood methods are used to estimate the parameters of the GEV. We supposed that random variables were independent and identically distributed (IID), noting that for IID variables observed from a population with common cumulative distribution function $F_Y(\cdot)$, the Fisher-Tippett-Gnedenko Theorem says that given a sequence of constants $\{a_m\}$ and $\{b_m > 0\}$, as $m \to \infty$

$$P\left(\frac{\max\{Y_1,...,Y_m\} - a_m}{b_m} \le z\right) \to F_{Max}(z),$$

provided $F_{Max}(\cdot)$ is a nondegenerate cumulative distribution function. From this, we know that $F_{Max}(\cdot)$ belongs to either a Gumbel, Fréchet, or Weibull distribution. The three types can be represented simultaneously through the GEV cumulative distribution function:

$$F_{Max}(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\}$$

with $y_+ = \max\{y,0\}$, $-\infty < \mu < \infty$, and $-\infty < \xi < \infty$. The three parameters, μ, σ, and ξ are respectively location, scale, and shape parameters, with varying values of ξ to distinguish between the three.

Taking $\{X_t\}_{t=1}^{N}$ as the seasonal maximum precipitation series with seasonal block size $m \approx 90$, we have assumed a nonstationary GEV (μ_t, σ_t, ξ_t) distribution with varying seasonal parameters to account for variation in seasonal characteristics. The distribution with parameters varying over seasons are as follows:

$$\mu_{nS+\upsilon} = \alpha_\upsilon + \beta_\mu (nS+\upsilon)/(100S)$$

$$\ln(\sigma_{nS+\upsilon} = \tau_\upsilon + \beta_\sigma (nS+\upsilon)/(100S)$$

The seasonal period $S = 4$ and $t = nS + \upsilon$ gives a periodic expression for time, with t expressed in terms of n, the number of full calendar years up to time t, with $\upsilon = 0,1,2,3$ representing winter, spring, fall, and summer, respectively. The location parameter each season is denoted α_υ, and σ_υ is the seasonal scale parameter. With some degree of simplification, we note that the trend term in location parameter β_μ can be expressed as

$$E(X_{(n+100)S+\upsilon}) - E(X_{nS+\upsilon}) = \beta_\mu$$

provided that β_σ is 0. This implies that β_μ models the expected change in extremes over a century, provided that there is no trend in the scale parameter σ_t (i.e. that $\beta_\sigma = 0$).

These parameters $\underline{\theta} = (\underline{\alpha}^{\mathrm{T}}, \underline{\tau}^{\mathrm{T}}, \underline{\beta}^{\mathrm{T}}, \xi)^{\mathrm{T}}$, with $\underline{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)^{\mathrm{T}}$, $\underline{\tau}^{\mathrm{T}} = (\tau_0, \tau_1, \tau_2, \tau_3)^{\mathrm{T}}$, and $\underline{\beta}^{\mathrm{T}} = (\beta_\mu, \beta_\sigma)^{\mathrm{T}}$, can be estimated using the maximum likelihood method. This likelihood function for $\underline{\theta}$ can be determined from $F_{Max}(\cdot)$, the GEV distribution function. Lastly, because there is no closed form expression for the maximum likelihood estimators, numerical optimization methods must be used to estimate the value of these parameters.

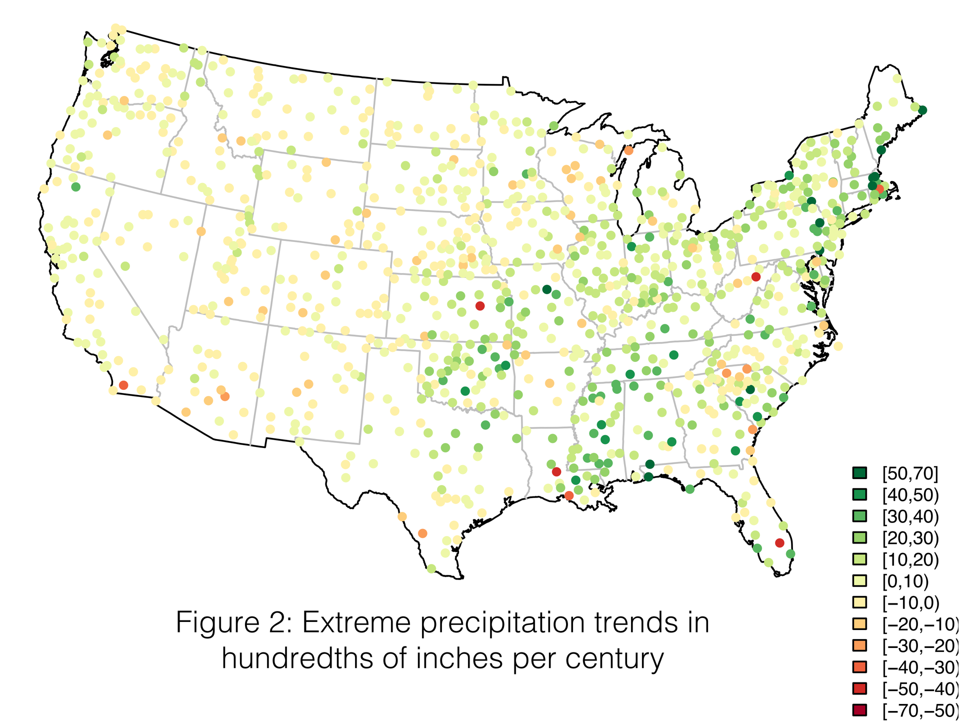

Figure 2: Extreme precipitation trends in hundredths of inches per century

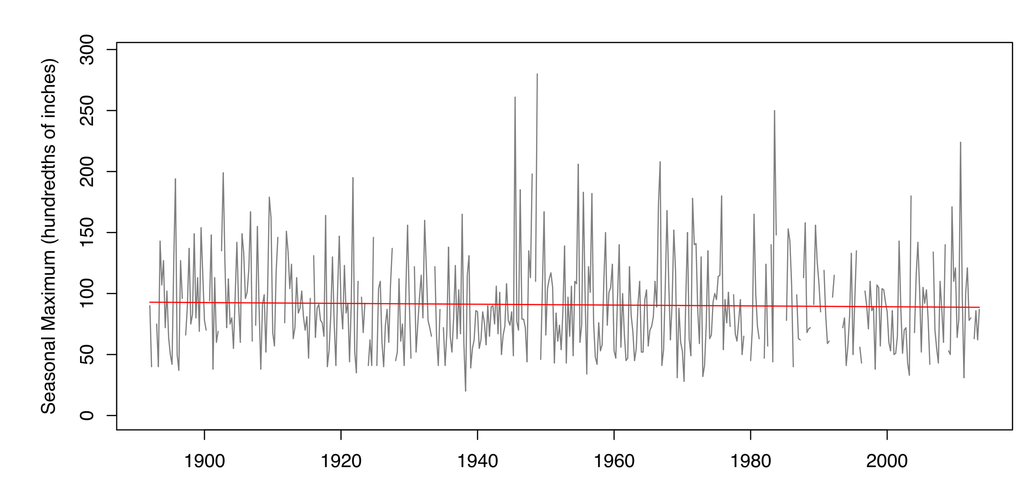

Figure 3: Scale parameter trends for station 456096 in Olga, WA

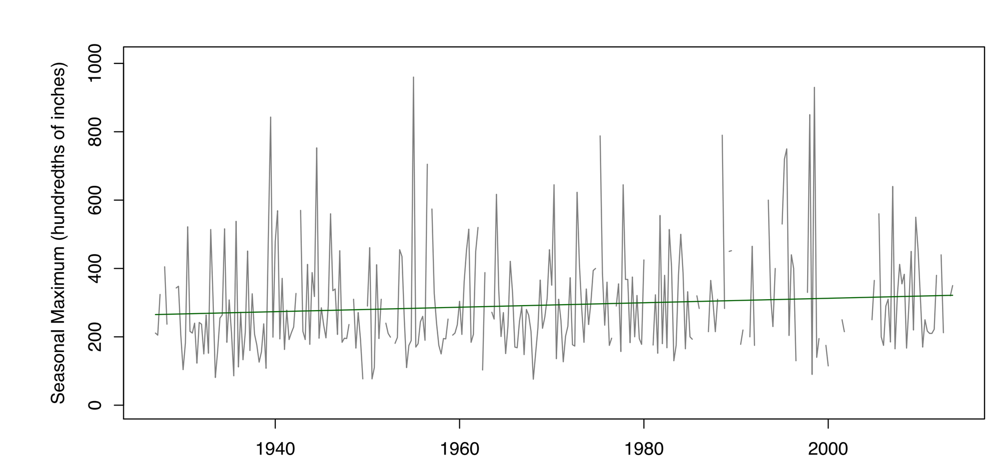

Figure 4: Scale parameter trends for station 011084 in Brewton, AL

Results

Figure 2 shows a geographic interpretation of changes in maximum precipitation in hundredths of inches over the past 100 years around the continental United States. Stations shown in dark red experienced a strong decrease in maximum precipitation observations over this time period, stations marked with dark green showed a strong increase, and stations shown in yellow observed little or no change in extreme precipitation. Trend estimates for individual stations are shown in Figures 3 and 4. These graphs show a trend line based on the values of β_μ, which was estimated by the maximum likelihood function. Figure 3 illustrates that, while the change is small, there is an estimated decrease overall in maximum precipitation values over the last century in Olga, Washington. However, notably, even a decrease of 0.3" over an area of one square mile implicates significantly less water. Conversely, Figure 4 shows an overall increase in maximum precipitation observations for the station located in Brewton, Alabama. In particular, we note that the severity of the increase (that is, the absolute value of β_μ) is larger for the station in Brewton than for the station in Olga.

Conclusions

In observation of the distribution of colors in Figure 2, we conclude that extreme precipitation is, by and large, increasing in the North East, South East, and South Central United States, with more severe increase shown in South Central states such as Mississippi, Alabama, Oklahoma, and Tennessee. By contrast, in the western United States, extreme precipitation trends show a slight decrease, becoming more severe in the Southwest and Central Western states. The coastal stations along the west appear to have experienced slight increase, while land-locked states and stations further from the coast have, overall, experienced slight decrease.

Future Study

In the near future, we expect to apply techniques to identify the times and locations of unknown changepoints. A changepoint occurs when a station undergoes a change that could significantly affect data observations, such as a change in location or acquisition of new equipment. However, these changepoints are not always recorded when stations undergo them. If gone undetected, sudden shifts in station data may be erroneously attributed to other factors. Thus, we intend to use changepoint techniques to identify where and when changepoints have occurred so that we can fit trend lines on the intervals between changepoints, rather than global trend lines that span across the record. This will allow us to more precisely determine trend estimates for individual stations. We will use the results of this study to help us determine possible locations for unaccounted changepoints. For instance, we note that Louisiana contains several stations in close proximity with vastly different trends. Such a stark difference within a small geographical radius leads us to believe this may be the location of an unknown changepoint. Thus, we can identify possible changepoints by sight before looking into specific station data to analytically verify whether a changepoint is likely to have occurred there.

References

1 http://www.nws.noaa.gov/hic/
2 Andsager, K., Kunkel, K., & Easterling, R. (1998, 7 August). Long-Term Trends in Extreme Precipitation Events in the Conterminous United States and Canada. *Journal of Climate, 12 (8)*, 2515-2527.
3 Kunkel, K. (2002, April 28). North American Trends in Extreme Precipitation. *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards, 29 (2)*, 291-305.
4 http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn.html