DeGaetano Arthur (Orcid ID: 0000-0002-0183-1346)

Declining U.S. Regional and Continental Trends in Intra-annual and Interannual Extreme Temperature Swings

Arthur T. DeGaetano
Northeast Regional Climate Center
Department of Earth and Atmospheric Science
Cornell University

and

Li-Sha Lim
Centre for Climate Research Singapore (CCRS)
Meteorological Service Singapore

Submitted to:
International Journal of Climatology

February 2019

Keywords: Temperature Extremes, Intra-annual variability, time series, United States

Corresponding Author Address:
Dr. Art DeGaetano
1119 Bradfield Hall
Cornell University
Ithaca, NY 14853
atd2@cornell.edu

Abstract

The occurrence of regional temperature extremes on weekly to seasonal time scales has been a common climate impact in recent decades. Both instances of extreme warmth and extreme cold have been documented and analyzed in the literature. While these events have most often been analyzed independently, in this study the transition between temperature extremes is examined using station data. Five measures of extreme temperature change are examined.

At stations across the United States there has been a significant decrease in the temperature difference between the warmest and coldest percentile observed within each year based on 7-day, 30-day and 90-day temperature averages during both the 1900-2017 and 1950 to 2017 periods. The maximum difference between percentiles associated with adjacent 7-day, 30-day and 90-day periods in each year has also declined significantly. At the same time, the interval between the highest and lowest annual percentile occurrence has lengthened. On a decadal basis, the frequency of shifts from the sub-5th to over-95th temperature percentiles has also declined through time, while the average time period between temperature occurrences in opposite tails of the distribution has increased. In general these results are very consistent across the U.S., although some regional and duration-dependent differences are noted. For many of the extreme

temperature metrics, a high level of field significance is obtained in the Southwest, Great Plains and Midwest regions.

1. Introduction

In the most recent decade, an increase in extremely warm summer temperatures (e.g. Rahmstorf and Coumou, 2011; Coumou and Rahmstorf, 2012) has seemed to be accompanied by with several occurrences of much colder than normal winter temperatures (e.g. Tollefson, 2014; Cattiaux et al; 2010) across mid-latitude regions of the northern hemisphere. For instance the northeastern United States recorded its warmest February in 124 years of record in 2017, while only two years earlier experiencing its second coldest February. In 2014, the coldest July was recorded in parts of the Midwest and Southeast United States. More recently, while April 2018 was the coldest on record in the Midwest, May (2018) average temperatures were among the warmest. These seasonal or shorter-time-scale extremes have occurred over a backdrop of increasing average temperatures and warming daily extreme temperatures (Meehl et al., 2009; USGCRP, 2017). The longer term changes in warm and cold season average temperature and daily extremes are directly consistent with anthropogenic changes in radiative forcing (Allen et al. 2018).

Both warm and cold regional temperature extremes stem from the occurrence of high amplitude Rossby waves. Cattiaux et al. (2010) attribute the extremely cold winter of 2009-2010 across Europe to a quasi-stationary large scale circulation pattern, associated with an extremely negative phase of the North Atlantic Oscillation (NAO) (Hurrell, 1995). Wang et al. (2010) also attribute record-breaking-cold temperatures in Asia and North America to the extremely negative NAO values which resulted in northerly surface winds and the southward advection of Arctic air. Hoskins and Woollings (2015) offer a slightly different perspective whereby instead of viewing only spatially fixed flow patterns such as the NAO, they consider persistent anomalous flow patterns in general which may take the form of a classic block, or an amplified and stationary Rossby wave pattern.

Blocking patterns have also resulted in the occurrence and persistence of extreme warm temperature patterns in summer. The 2010 Russian heat wave was attributed to an “omega” blocking pattern (Dole et al. 2011). Near the center of the block, the presence of northwardly displaced subtropical air and sinking vertical air motions contributed to extremely warm surface temperature. This was amplified by feedbacks associated with severe drought conditions. Garcia-Herrera et al. (2010) attribute similar circulation features, the northward displacement of subtropical air from the North Atlantic and local soil moisture deficit feedbacks, with the 2003 European heat wave. Land surface feedbacks associated with soil moisture deficits have been widely shown to intensify and prolong heat waves (e.g. Hoerling et al., 2013; Miralles et al. 2014), but with important regional differences (Cheng et al., 2019)

These recent occurrences of temperature extremes have led to hypotheses regarding the occurrence and potential cause of long-term changes in the responsible atmospheric circulation features. Using self-organizing maps Horton et. al (2015) identified increasing trends in summer and autumn anticyclonic circulations that led to hot extremes in parts of Eurasia and North America. In winter, they found the an increased incidence of northerly flow led to winter cold extremes in central parts of Asia.   Blocking over Greenland, has increased significantly during summer over recent decades (Hanna *et al*. , 2016; McLeod and Mote, 2016) and has shown an increase in interannual variability during December (Hanna *et al., 2016).* In reanalysis data, Francis and Vavrus (2012) found a significant increase in block persistence as did Coumou *et al* (2015) during summer based on satellite data. Others such as Cattiaux *et al* (2016) found no evidence of increased block persistence using global climate models, suggesting that the detection and significance of such trends may be affected by differences in methodology. There is active debate in the literature regarding the mechanisms responsible for these circulation changes. Many studies point to Arctic amplification as a primary causal mechanism, while others argue that other factors contribute (e.g. Varvus, 2018; McCusker et al. 2016).

In terms of the tangible surface weather that results from these circulation changes, Coumou et al. (2018) found that summer temperatures within the warmest percentiles are increasing at a faster rate than summer temperatures in the coldest percentiles. They argue that this is an indication that processes other than radiative forcing are influencing the change in summer extremes. Such a change increases the range between the warmest and coldest summer temperatures, but also contributes to the increase in intensity and frequency of summer heat waves in the mid-latitudes (Russo et al., 2014).

Francis et al. (2018) coined a new metric called long duration events (LDE) and found that there was an increase in LDEs, defined as 4-day periods with either all stations measuring ≤ 0.254 mm (dry LDE) or any station > 0.254 mm (wet LDE) of precipitation within a 2500 km^2 grid. They found both dry and wet spells have mainly increased since the mid-1990s. Conversely, Swain et al. (2018) examined what they termed the “precipitation whiplash” signal (a change from sub-twentieth-percentile to over 80th percentile precipitation in adjacent wet seasons). In climate model simulations for California, they found that this metric increased through the 21st century, with a larger increase in southern compared to northern California. Cohen (2016) referred to “weather whiplash” and used reanalysis data to examine trends in zonal wind speed, temperature and geopotential heights, to conclude that temperature variability has increased in mid-latitudes.

Here we develop and analyze a set of metrics referred to as “temperature swings”. Analogous to precipitation whiplash events, a swing is characterized either by the magnitude of the difference, or the period between, the highest and lowest percentile temperature anomalies within a year or as the difference and period between opposing extreme temperature anomalies (i.e. cold tail versus warm tail) across years. The temporal changes in the resulting temperature swing climatology are analyzed across the continental United States and by region, as well as over different time scales. As with precipitation whiplash events, changes in temperature swing occurrence impact a number of natural and manmade systems, for example avian social behavior (Jetz and Rubenstein, 2011), insect survival (Sholes, 2011), human health (Amiya, et al. 2009), agriculture (Wilhite et al. 2017), and electricity demand (Chang et al., 2016).

2. Methods

2.1 Data

Daily average temperature data for U.S. stations in the Global Historical Climatology Network (Menne, et al. 2012) for which 1981-2010 normals existed were obtained from the Applied Climate Information System (ACIS) (DeGaetano et al., 2015). Although GCHN data are not homogenized for non-climatic discontinuities, the values undergo extensive quality control (Menne, et al. 2012), assuring the data are of the highest quality possible. Two sets of stations were selected, those with records that began in 1950 or earlier and those with records that extended back to at least 1900. The set of stations with observations commencing in 1900 is a subset of the first. The daily data were grouped into overlapping 7-, 30-, and 90-day duration intervals and averaged. These intervals were allowed to cross calendar year boundaries. For example, the 7-day duration average assigned to 2 January included data from the preceding 27-31 December in addition to 1-2 January. The subsequent 7-day duration average for 3 January included data from 28 December -3 January. This choice of durations was intended to represent synoptic, sub-seasonal and seasonal time scales thereby reflecting conventional forecast periods. To facilitate spatial analysis the stations were aggregated into six National Climate Assessment regions (USGCRP, 2017).

Stations were excluded when more than 10% of the possible i-day-duration intervals contained missing data. Otherwise, missing data values were estimated using a procedure similar to that developed by DeGaetano et al. (1995). The method assumes that the departure of daily temperature from normal is homogenous over the region surrounding a station. For each missing temperature, the closest station with both non-missing daily temperature and a 1981-2010 normal was identified. Using these values, an estimate for the missing temperature, T_{miss}, was computed by

$$T_{miss} = (T_n - N_n) + N_{miss}\,, \tag{1}$$

where T_n and N_n are the daily temperature and official 1981-2010 daily temperature normal at the closest neighboring station and N_{miss} is the daily temperature normal at the station with missing data.

2.2 Swing Magnitude and Duration

Each of the 365 duration-interval averages in a specific year was transformed to a percentile. For example, a 100-year record includes one hundred 7-day duration averages ending on each specific date, *d*, within the year (e.g. March 2). The 100 values for each date were sorted and the position of the value for a specific year, *y*, (e.g. March 2, 1991) used to assign a percentile to each average (P_{dyi}), where *i* is the duration interval. Using these values, three metrics were computed for each year. Swing magnitude, *Sm*, was defined as:

$$Sm_{yi} = \max_d P_{dyi} - \min_d P_{dyi}\,. \tag{2}$$

Thus Sm_{yi} is the difference between the largest and smallest daily percentile for interval *i* in year *y*. Swing period, *Sp*, is the number of days between the date of occurrence of the maximum and minimum annual percentile,

$$Sp_{yi} = abs[\left\{d \middle| P_{dyi} = \max_d P_{dyi}\right\} - \left\{d \middle| P_{dyi} = \min_d P_{dyi}\right\}] . \qquad (3)$$

Adjacent swing magnitude *Sa* is similar to Sm_{yi} with the constraint that the two percentiles be associated with subsequent, non-overlapping duration intervals. Thus

$$Sa_{yi} = \max_d abs(P_{dyi} - P_{(d+i)yi}). \qquad (4)$$

These three metrics are illustrated graphically in Figure1.

Non-climatic discontinuities in daily GHCN data can potentially affect these swing metrics. However, since the metrics define relative changes within each year, the effect of these unavoidable artifacts is likely minimal. Moreover the regional grouping of stations in subsequent analyses further mitigates the influence of non-systematic discontinuities.

A related metric is termed a tail swing. Unlike *Sp*, tail swings can span years and unlike *Sm* and *Sa*, tail swings are defined by a specific percentile. A tail swing commences when P_{dyi} crosses the 95th or 5th percentile and ends when P_{dyi} crosses the corresponding threshold in the other tail of the empirical average temperature distribution (Fig. 1).

For each 1900-2017 and 1950-2017 temperature time series, tail swings were tallied by assigning an occurrence to the date on which the change to the opposite percentile threshold first occurred. For example in a record that started in 1900, if P_{dyi} = 98 on 1 April 1900 and P_{dyi} = 4 on 15 June 1902, a tail swing occurrence would be assigned to 15 June 1902 even if the percentile for 16 June 1902 was also below the 5th percentile. The subsequent tail swing would correspond to the next date on which P_{dyi} > 95th percentile occurred. The period of each tail swing was defined as the length of time, in days, between tail swing occurrences. For the example above, the tail swing period is 806 days.

2.4 Detrended Swing Metrics

Since the computation of tail swings requires daily temperatures that span multiple years, longer term trends in the daily values can influence the frequency and period of tail swing occurrences. To identify whether long-term temperature trends have influenced tail swing characteristics, duration-interval average temperature series at each station were detrended by fitting a linear trend to each of the 365 duration interval average time series. The resulting time series of daily slopes were smoothed using a 30-day lowess filter from the sm.nonparametric.lowess routine within the StatsModels software library (https://www.statsmodels.org/). Figure S1 shows the resulting slopes by Assessment region.

For the 1900-2017 period, the regional and CONUS slopes for 7-day, 30-day and 90-day average temperature are predominantly positive, the exceptions being the Southeast regional series during all seasons, which is consistent with USGCRP (2017); the Midwest in autumn; and for the shorter durations the Northwest in late autumn (Fig. S1a, S1c, S1e). The autumn cooling in the Midwest and Northwest was also present for the 1950-2017 period (Fig. S1b, S1d, S1f). The Southeast, however, showed warming throughout most of the year during the 1950-2017 period except winter. In general, daily temperatures warmed at a rate of between 0.05 and 0.15

°C per decade in both periods with the greatest warming during winter and the least warming in autumn consistent with (Wang et al., 2009)

Residuals were calculated by subtracting the temperatures given by the smoothed regression slopes from the original temperatures. The residuals were then used to redefine the percentiles and ultimately recompute new tail swings and tail swing periods based on the residuals. Detrended daily temperature was also used to re-assess the S*m*, S*p*, and S*a* trends. However differences between the results using the original and detrended series were minimal.

In all regions, the number of positive trends exceeded the number of negative trends for both the coldest and warmest annual percentiles. In general, more than twice of the number of positive trends occurred compared to negative trends, with the exception of the Midwest region were a similar number of positive and negative trends existed (Table 1). Across the country the coldest percentiles on average warmed at a faster rate than the warmest percentiles. However over the period of record this difference was minimal (i.e. over a 67-year period the increase in the coldest percentile was on average <1 percentile greater than the increase in the warmest percentile (Table 1). Regionally, the annual warmest percentile averaged 99 for 7-day durations, 96 for 30-day durations and 88 for 90-day durations. The annual coldest percentiles averaged 2, 5 and 12 for these durations.

2.5 Swing metric trend assessment

The annual time series of *Sm*, *Sp* and *Sa* were evaluated for time dependent changes using both least-squares fit linear slope and the Kendall Tau statistic. In all cases, the null hypothesis Ho: slope= 0 was assessed relative to the alternative Ha: slope ≠ 0. This two-tailed test was conducted at the 95% level. The station-specific results were summarized by National Climate Assessment region and the contiguous United States by tallying the number of null hypothesis rejections and the number of positive and negative slopes (regardless of significance) in each region.

The regional groupings were used to examine the field significance of the test results. To quantify the significance of the regional results, the original chronology of years was randomized to destroy the time dependence that may have existed in the original 1900-2017 time series. A total of 1000 randomized sets of years were generated and each used to recompute the time series of *Sm*, *Sp* and *Sa* metrics, at all stations. The same set of randomized years was applied to all stations simultaneously to preserve any spatial relationships. Trend tests were then repeated on the reordered time series and new regional summaries of null hypothesis rejections and slope sign counts compiled. When repeated 1000 times, this provided an empirical distribution of regional test counts against which the summaries based on the original (proper) chronological order could be compared.

The significance of the observed changes in tail swing occurrence was tested in a similar fashion, albeit some modifications were required given the irregular (non-annual) spacing of tail swings events. Tail swing occurrences were aggregated into 10-year, non-overlapping time periods and for each block the total number of tail swing events and the average tail swing period were calculated. These decadal time series were evaluated for time dependent changes based on the least-squares fit linear slope as before.

The null hypothesis Ho: slope= 0 was assessed using an empirical distribution of 1000 bootstrap-resampled slopes. Each observed tail swing occurrence was randomly reassigned to a date within the available station record. No two swings were allowed to occur within a given duration interval. For example, when 90-day durations were considered, once a tail swing occurrence was assigned to a date *x*, a second tail swing event could not be assigned to any of the 89 days before or after *x*. This prevented the randomly assigned tail swing events from overlapping.

The resampled time series were then aggregated into non-overlapping 10-year periods and least-squares-fit slopes computed. These 1000 resampled slopes formed the empirical distribution against which the null hypothesis was evaluated. Ho was rejected when the original (non-randomized) values fell below the 2.5th or above the 97.5th percentile of the resampled distribution.

3. Results

3.1 Seasonality

Figure 2 shows the seasonal cycle of the occurrence of the annual warmest and coldest percentiles. These two values define S*m* and S*p*. There is little regional variation to these national patterns. For 7- and 30-day duration events (Fig. 2a-b, 2d-e) a strong seasonal cycle is lacking. For the 90-day duration, however, both the annual maximum and minimum percentiles show a tendency toward winter occurrences (Fig. 2g-h). In some cases, nearly twice as many occurrences are evident in winter months as in summer months.

The seasonality of adjacent swings (S*a* events) is more pronounced (Fig. 2c, 2f, 2i). For 7-day duration swings, S*a* events are generally more common in the transition seasons (Fig. 2c). This tendency transitions to a winter maximum and summer minimum as duration increases. The between-month differences in S*a* occurrence are larger than those for the S*m* event percentiles. Ninety-day S_a events are twice as likely to occur in January compared to July (Fig. 2i).

3.2 Trends

Swing Magnitude

Trends in S*m* are distinctly negative from 1950-2017, indicating that within a particular year the difference between the highest and lowest percentile occurrence has declined (Fig. 3a, 3d, 3g). The decline in this metric is more pronounced for durations greater than seven days. For 7-day durations (Fig. 3a) about half the trends are positive, but few reach the level of statistical significance. More of the negative trends are significant, with many of these located in California, the Gulf Coast and the Northeast. For 30-day durations (Fig. 3d), the number of positive S*m* trends declines, with only a few reaching the level of statistical significance. In contrast, the number of significant negative S*m* trends increases, particularly in the Great Plains and Midwest. For 90-day durations (Fig. 3g), the number of significant negative trends increases further, particularly in the intermountain west. Positive trends also increase in number from those observed for 30-day duration S*m*. These positive trends, which are generally not statistically significant, tend to be concentrated in the Ohio Valley, Southeast, and Mid Atlantic regions.

Over the longer 1900-2017 period, the geographic pattern of S*m* trends is similar to that for trends starting in 1950 (Fig. 4a, 4d, 4g). Although there are markedly fewer available stations and most are concentrated in the Mississippi Valley, 7-day duration trends are fairly evenly divided between increases and decreases, with few trends reaching the level of significance. However 30- and 90-day duration S*m* trends are predominately negative (Fig. 4d, 4g) with a concentration of significant trends in the northern Great Plains and Midwest. Like the 1950-2017 trends, a cluster of positive trends exists in the Midwest, especially for the 90-day durations.

Swing Period

Trends in S*p* are distinctly positive from 1950-2017 (Fig. 3b, 3e, 3h) indicating that within a particular year the period between the highest and lowest percentile occurrence has lengthened. For 7-day durations, the majority S*p* trends are positive (Fig. 3b). Most of the significant S*p* trends are concentrated in the Midwest. Negative trends in S*p* dominate the Southeast. S*p* trends for 30-day durations are also primarily positive (Fig. 3e). The number of significant (positive) trends also increases particularly in the Midwest and Northern Great Plains. Positive trends occur more frequently in the Southeast compared to those for the shorter duration. For 90-day durations, a concentrated area of significant positive 90-day duration S*p* trends extends from western parts of the Northeast, through the Great Lakes and Midwest, into the Northern Great Plains (Fig. 3h). This is the general area that experienced significant decreasing S*m* trends. The number of negative S*p* trends also increases for 90-day durations, particularly in the Intermountain, Northwest and central regions of the U.S (Fig. 3h).

For the 1900-2017 period, the geographic patterns of S*p* trends and significant S*p* trends are similar to those for the 1950-2017 period (Fig. 4). The most pronounced difference occurs with the significant 90-day duration trends. Over the longer time period, these trends extend farther south ranging from the Northern Great Plains to the Gulf Coast and remain largely positive (Fig. 4h).

S*p*, by definition, is not influenced by the magnitude of the percentiles. Therefore, the changes observed in Figures 3 and 4 are unlikely to be related to differential warming of the cold versus warm extremes. Rather other factors, potentially related to the persistence and/or magnitude of the atmospheric circulation patterns that cause the extremes must contribute to the observed S*p* changes. Alternatively, temperatures in the warm tail of the distribution (highest percentiles) increasing at a slower rate than temperatures in the cold tail, could result in the decreases in S*m* shown in Figures 3 and 4. Table 1 shows that this pattern of change (greater warming of the lowest percentile relative to the warmest) on average occurs in all Assessment

Regions, however the magnitude of this difference is small. Detrending the multi-day temperature averages prior to computing S*m*, S*p* and S*a* trends had minimal effect on the results, since this did not address the differential warming experienced by different parts of the distribution.

Adjacent Swing Magnitude

Like trends in S*m*, S*a* trends are predominately negative (Fig. 3c, 3f, 3i). For 7-day durations, there is a similar number of negative and positive trends, most of which are not significant statistically (Fig. 3c). Negative trends are concentrated in the Mid-Atlantic and Southeast regions. As duration increases, with the exception of locations in the extreme northern Midwest (e.g. Wisconsin), trends are predominately negative and largely significant (Fig. 3f). The geographic pattern of 90-day duration trends is similar to that for the 30-day values. The exception being the concentrated area of positive trends, which are mostly non-significant, is located farther south across Ohio and Indiana and extending into the Southeast region.

Over the longer 1900-2017 time period, the pattern and magnitude of S*a* trends is very similar to that for the 1950-2017 period (Fig. 4). The most notable exception is that the concentration of positive S*a* trends located in the upper Midwest for 7- and 30-day durations is characterized by mainly negative trends over the longer time interval. Given the close correspondence between the S_a and S_m results, subsequently only the S_m results are presented.

Field Significance

The concentration of significant trends in Figures 3-4 is potentially an artifact of the generally high spatial correlation among the S*m*, and S*p* time series from adjacent stations. While this is the case in some Assessment regions, in others, the number of statistically significant and negative S_m and positive S_p trends was greater than would be expected by chance (Fig. 5). Over the 1900-2017 period, the numbers of positive and statistically significant S*p* trends was greater than expected by chance ($\alpha \leq 0.05$) across the U.S. (Fig. 5). Regionally, the number of positive and statistically significant S*p* trends in the Great Plains and Southwest regions exhibited field significance most consistently.

For S*m* trends, the number of significant trends across the entire U.S. from 1900-2017 was field significant at the $\alpha = 0.10$ level or less. The concentration of significant S*m* trends was particularly noteworthy in the SW region ($\alpha \leq 0.05$) for all durations.

Over the shorter and more recent 1950-2017 period, the field significance of the trends was higher, especially in terms of the number of statistically significant trends (Fig. 5). The number of statistically significant S*m*, and S*p* trends was greater than that expected by chance (α <=0.5 for all but one duration across the U.S.). Regionally the highest field significance levels were obtained in the Southwest and in most other regions for durations > 7 days, with the exception of the Midwest, which failed to attain field significance. The number of negative S*m* trends (regardless of significance) was significant in the Southwest and the Great Plains (with the exception of 7-day duration trends). Across the U.S. field significance for duration of ≥15 days was at least at the $\alpha = 0.10$ level. For positive S*p* trends, field significance was not obtained nationally and occurred only for some regions and durations (Fig. 5).

Regional Time Series

The station-specific S*m*, and S*p* values within each U.S. National Climate Assessment Region were averaged to produce an aggregate time series. A contiguous United States (CONUS) time series was also computed based on an average of the station data. As expected from Figure 5, S*m* declines through time in the CONUS, with the 1950-2017 linear decline greater than that experienced from 1900-2017 (Fig. 6). In all regions (the Great Plains is shown as a representative example) the subset of stations with the longest records (beginning in at least 1900) reflect the regional trends and year-to-year variations of the larger set of stations with records beginning in 1950 (Fig. 6). The CONUS average time series for S*a* mirror the S*m* values, with consistent decreasing trends for all durations and a slightly faster rate of decline since 1950 relative to 1900 (not shown).

Time series of the interval between annual percentile swings, S*p*, generally increase in all regions and the CONUS over the longer-term 1900-2017 period regardless of duration (Fig. 7). This indicates within a given year the relative temperature extremes tend to be separated by longer time periods. Over the more recent time period, 1950-2017, the regional and CONUS S*p* time series show a mix of increases and decreases, depending on region and duration (Fig. 7).

This corresponds well with the lack of field significance for S*p* trends over the 1950-2017 in Figure 5.

Tail Swings

Like S*m* and S*p*, tail swing occurrence decreased through time while the period between tail swings increased. For the 1900-2017 period, this was especially true in the northern Great Plains, Midwest and Southeast (Fig. 8) where significant trends were common for 7-day and 30-day durations. Tail swings for data averaged over 90-day periods also exhibited this behavior (decreases in decadal tail swing occurrence and increases in the time between tail swings), but the number of statistically significant trends was substantially lower (Fig. 8e-f). Tail swing trends based on non-detrended daily data (not shown), were nearly identical to those based on the detrended data used for Figure 8. The geographic pattern of decreasing tail swing magnitude and increasing tail swing period is also reflected based on the 1950-2017 period of record (Fig. 9).

Figure 10 focuses on the Midwest region, given the prominence of tail swings in this area. In the early 20th century, stations in the Midwest experienced on the order of 40 7-day tail swings, 13 30-day tail swings and five, 90-day tail swings, generally consistent with the number of potential opportunities for tail swings (the number of 7 day periods is 12 times greater than 90 day periods). The decreasing trend in tail swing occurrence with time that characterized the region in Figure 8 and 9 for 7-day and 30-day durations is readily apparent in Figure 10a and 10d. Likewise, the increase in the period between tail swings is also apparent in Figures 10b and

10e. For both 7- and 30-day duration tail swings the average period between events increases by approximately 20%. In the earliest three decades 7-day duration tail swings were separated by 92 days on average, while in the most recent three decades the time between events increased to 109 days (Fig. 10b). The average time between 30-day duration tail swings is longer and increases from on average 254 days in the early part of the record to 332 days in more recent decades (Fig. 10e).   When the maximum time period between tail swing events (at any station) in each decade is considered (Fig. 10c and 10f) , the most recent decade (2008-2017) experienced the longest interval between successive tail swings for both 7-day (512 days between events) and 30-day durations (1324 days between events). On average for the entire period of record, the decadal maximum time intervals between 7- and 30-day tail swings were 433 and 1028 days respectively. Recent 90-day tail swings were not noteworthy in recent decades.

4. Conclusions

Long term changes in temperature variability are examined from the prospective of a metric termed temperature swings. On an annual basis, a swing is based on the highest and lowest observed temperature percentiles and is defined by its magnitude (the absolute difference of the two percentiles) and period (number of days between the percentile extremes). Related metrics specify the maximum swing magnitude between consecutive (period=0) swings, and the frequency of and period between occurrences of values below and above specific extreme percentiles in opposite tails of the temperature distribution, for example below the 5th and above the 95th percentile.

Across the U.S., there is a general tendency for swing magnitude to decrease from 1900 (and 1950) to the present. This occurs regardless of whether the opposing extremes occur anytime within the same year or during consecutive time intervals. In addition, the period of time between extremes has consistently increased through these time periods. This occurs in terms of the warmest and coldest extremes within each year and also between the occurrence of fixed extremes in opposite tails of the temperature distribution which typically span multiple years. The decadal frequency of swings from one tail to the opposite tail has declined through time.

These metrics provide a novel perspective on changes in intra- and inter-annual extreme temperature variability through time. Decreases in intra-annual variation tend to be more pronounced and widespread for longer duration temperature averages (i.e. 30- and 90-day duration) as opposed to shorter 7-day aggregations. However on an interannual basis, it is the shorter duration events (i.e. 7- and 30-day durations) that exhibit the most significant changes. The results are generally consistent with the studies that have examined the differential warming rates between cold and warm temperature extremes (Meehl et al., 2009) and changes in atmospheric circulation patterns, such as the tendency for increased blocking and more persistent patterns.

Meehl et al. (2009) found that warm extremes have occurred twice as frequently as cold extremes. Potentially this could lead to cold extremes becoming more uncommon, thus decreasing swing frequency. This explanation is most plausible for 7-day duration tail swings as the short duration and use of fixed 5th and 95th percentiles in the definition of tail swings is most analogous to the daily records analyzed by Meehl et al. (2009). Detrending the mean time series would not address the differential warming between opposing extremes, potentially explaining the correspondence between the trended and detrended results.

For longer durations, and especially for the S_m, S_p and S_a metrics, differential warming of the warm and cold extremes is less likely to be the only factor responsible for the decline in swings. This is because these metrics are defined by relative, rather than fixed percentiles. The S_m value is the same in a year regardless of whether the annual percentile extremes are defined by the 1st and 50th percentiles or the 50th and 99th percentiles. Also, as duration increases the data become less similar to daily extremes.

Rather, especially for S_p, the results are consistent with an increase in persistence. If the circulation pattern associated with the warmest (or coldest) percentile anomaly within a year is more persistent, S_p will increase. This is because the length of the persistent pattern is included in the interval leading up to the start of the alternate percentile anomaly. Increased persistence could also play a role in the decline in S_m and S_a as this would tend to lower variability and hence limit the range of percentile anomalies experienced within any particular year, particularly on synoptic time scales (Schneider et al., 2015). Given that the strongest declines in S_m and S_a

are found in the Midwest and Great Plains, it is plausible to consider land-atmosphere feedbacks as having a contributing role, given the influence of soil moisture deficits on the intensity and persistence of heat waves (e.g. Hoerling et al., 2013; Miralles et al. 2014).

Collectively, the S_m, S_p and S_a results are tangentially consistent with a similar study using precipitation. Francis et al. (2018) found an increase in the duration of persistent wet and dry periods and concluded that weather patterns across North America were becoming more persistent. As in this study, Francis et al. (2018) found the most pronounced changes in and adjacent to the midwestern U.S.

More analogous to the tail swings examined here is the "precipitation whiplash" signal analyzed by Swain et al. (2018). However, they found an increase in year-to-year change from sub-20th to over-80th percentile precipitation, as opposed to the decrease identified in this study using a similar metric for opposing temperature extremes. Arguably the decreases in temperature swings were not spatially consistent along the west coast, which was the regional focus of Swain et al. (2018). Furthermore precipitation is not affected by potentially different rates of warming at low and high percentiles that likely influence our temperature tail swing results. Similarly, although Cohen's (2016) observation that zonally averaged temperature standard deviation has increased in mid-latitudes seems counter to these results, our focus on the most extreme temperatures and longer time scales complicates a direct comparison.

The use of station-based temperature data complements previous studies examining changes in atmospheric circulation patterns. The swing metrics analyzed here can serve as proxies for circulation patterns related to the warmest and coldest temperatures experienced in different geographic areas over different time scales from weekly to seasonal. Collectively these metrics indicate a slowing of the transition from circulation regimes producing the warmest and coldest temperatures. In addition, when these transitions occur the magnitude of the change (i.e. the difference between the temperatures associated with warm and cold extremes) decreases. This pattern has been generally consistent through time. Spatially, although the most significant changes are concentrated in the Southwest and Great Plains, there is a general consistency in the temperature swing trends across the U.S.

Having identified these characteristics of extreme temperature transitions, future research is necessary to determine the relationship between transitions and the ambient circulation features. For instance, understanding whether similar circulation features result in transitions in different regions of the country or during different times is necessary to ascribe a mechanism to the observed changes in extreme temperature variation and potentially explain the geographic patterns of change, which may be related large scale factors such as preferred wave patterns or local features such as regional soil moisture trends (e.g. Ardilouze, 2017).

Acknowledgments

This work was supported by NOAA Contract AB-133E-16-CQ-0025. The work is an extension of a senior research honors thesis submitted by Ms. Lim as a student in the College of Agriculture and Life Sciences at Cornell University.

References

Allen, M., Dube, O. P., Solecki, W., Aragón–Durand, F., Cramer, W., Humphreys, S., Kainuma, M., Kala, J., Mahowald, N., Mulugetta, Y., Perez,R., Wairiu, M. and Zickfeld, K. (2018) Framing and Context. In: *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* [V. Masson-Delmotte, Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors,S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock,T., Tignor, M. and Waterfield, T. (eds.)].

Amiya, S., Nuruki, N., Tanaka, Y., Tofuku, K., Fukuoka, Y., Sata, N., Kashima, K. and Tsubouchi, H., (2009) Relationship between weather and onset of acute myocardial infarction: can days of frequent onset be predicted?. *Journal of cardiology*, *54*(2), 231-237.

Ardilouze, C., Batté, L., Bunzel, F., Decremer, D., Déqué, M., Doblas-Reyes, F.J., Douville, H., Fereday, D., Guemas, V., MacLachlan, C. and Müller, W., (2017) Multi-model assessment of the impact of soil moisture initialization on mid-latitude summer predictability. *Climate Dynamics*, *49*(11-12), 3959-3974.

Cattiaux, J., Peings, Y., Saint-Martin, D., Trou-Kechout, N. and Vavrus, S.J. (2016) Sinuosity of midlatitude atmospheric flow in a warming world. *Geophysical Research Letters*, *43*(15), 8259-8268.

Cattiaux, J., Vautard, R., Cassou, C., Yiou, P., Masson-Delmotte, V. and Codron, F. (2010) Winter 2010 in Europe: A cold extreme in a warming climate. *Geophysical Research Letters*, *37*(20).

Chang, Y., Kim, C.S., Miller, J.I., Park, J.Y. and Park, S. (2016) A new approach to modeling the effects of temperature fluctuations on monthly electricity demand. *Energy Economics*, *60*, 206-216.

Cheng, L., Hoerling, M., Liu, Z. and Eischeid, J. (2019) Physical understanding of human-induced changes in US hot droughts using equilibrium climate simulations. *Journal of Climate*, *32*(14), 4431-4443.

Cohen, J. (2016) An observational analysis: Tropical relative to Arctic influence on midlatitude weather in the era of Arctic amplification. *Geophysical Research Letters*, *43*(10), 5287-5294.

Coumou, D., Di Capua, G., Vavrus, S., Wang, L. and Wang, S. (2018) The influence of Arctic amplification on mid-latitude summer circulation. *Nature communications*, *9*(1), 2959.

Coumou, D., Lehmann, J. and Beckmann, J. (2015) The weakening summer circulation in the Northern Hemisphere mid-latitudes. *Science* **348**, 324–327.

Coumou, D. and Rahmstorf, S. (2012) A decade of weather extremes. *Nature climate change*, *2*(7), 491.

DeGaetano, A.T., Eggleston, K.L. and Knapp, W.W. (1995) A method to estimate missing daily maximum and minimum temperature observations, *J. Appl. Meteor.*, **34**, 371-380.

DeGaetano, A.T., Noon, W. and Eggleston, K.L. (2015) Efficient access to climate products using ACIS web services. *Bulletin of the American Meteorological Society*, *96*(2), 173-180.

Dole, R., Hoerling, M., Perlwitz, J., Eischeid, J., Pegion, P., Zhang, T., Quan, X.W., Xu, T. and Murray, D. (2011) Was there a basis for anticipating the 2010 Russian heat wave?. *Geophysical Research Letters*, *38*(6).

Francis, J.A., Skific, N. and Vavrus, S.J. (2018) North American weather regimes are becoming more persistent: Is Arctic amplification a factor?. *Geophysical Research Letters*, *45*(20), 11-414.

Francis, J.A. and Vavrus, S.J. (2012) Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters*, *39*(6).

García-Herrera, R., Díaz, J., Trigo, R.M., Luterbacher, J. and Fischer, E.M. (2010) A review of the European summer heat wave of 2003. *Critical Reviews in Environmental Science and Technology*, *40*(4), 267-306.

Hanna, E., Cropper, T.E., Hall, R.J. and Cappelen, J. (2016) Greenland blocking index 1851–2015: a regional climate change signal. International Journal of Climatology, 36, 4847–4861.

Hoerling, M., Kumar, A., Dole, R., Nielsen-Gammon, J.W., Eischeid, J., Perlwitz, J., Quan, X.W., Zhang, T., Pegion, P. and Chen, M. (2013) Anatomy of an extreme event. *Journal of Climate*, *26*(9), 2811-2832.

Horton, D.E., Johnson, N.C., Singh, D., Swain, D.L., Rajaratnam, B. and Diffenbaugh, N.S. (2015) Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature*, *522*(7557), 465.

Hoskins, B. and Woollings, T. (2015) Persistent extratropical regimes and climate extremes. *Current Climate Change Reports*, *1*(3), 115-124.

Hurrell, J.W. (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science*, *269*(5224), 676-679.

Jetz, W. and Rubenstein, D.R. (2011) Environmental uncertainty and the global biogeography of cooperative breeding in birds. *Current Biology*, *21*(1), 72-78.

McCusker, K.E., Fyfe, J.C. and Sigmond, M. (2016) Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss. *Nature Geoscience*, *9*(11), 838.

McLeod, J.T. and Mote, T.L. (2016) Linking interannual variability in extreme Greenland blocking episodes to the recent increase in summer melting across the Greenland ice sheet. International Journal of Climatology, 36, 1484–1499.

Meehl, G.A., Tebaldi, C., Walton, G., Easterling, D. and McDaniel, L. (2009) Relative increase of record high maximum temperatures compared to record low minimum temperatures in the US. *Geophysical Research Letters*, *36*(23).

Menne, M.J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R.S., Gleason, B.E. and Houston, T.G. (2012) Global historical climatology network-daily (GHCN-Daily), Version 3. *NOAA National Climatic Data Center*, *10*, p.V5D21VHZ.

Miralles, D.G., Teuling, A.J., Van Heerwaarden, C.C. and De Arellano, J.V.G. (2014) Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature geoscience*, *7*(5), 345.

Rahmstorf, S. and Coumou, D. (2011) Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, *108*(44), 17905-17909.

Russo, S., Dosio, A., Graversen, R.G., Sillmann, J., Carrao, H., Dunbar, M.B., Singleton, A., Montagna, P., Barbola, P. and Vogt, J.V. (2014) Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres*, *119*(22), 12-500.

Schneider, T., Bischoff, T. and Płotka, H., 2015. Physics of changes in synoptic midlatitude temperature variability. *Journal of Climate*, *28*(6), pp.2312-2331.

Sholes, O.D. (2011) Consequences of narrow temperature tolerance for a pinyon pine sawfly, Neodiprion edulicolus. *Environmental entomology*, *40*(4), 777-781.

Swain, D.L., Langenbrunner, B., Neelin, J.D. and Hall, A. (2018) Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, *8*(5), 427.

Tollefson, J. (2014) US cold snap fuels climate debate. *Nature News*.

USGCRP, (2017) Climate Science Special Report: Fourth National Climate Assessment, Volume I [Wuebbles, D.J., Fahey, D.W., Hibbard, K.A., Dokken, D.J., Stewart, B.C. and Maycock, T.K. (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 470 pp., doi: 10.7930/J0J964J6.

Vavrus, S.J. (2018) The influence of Arctic amplification on mid-latitude weather and climate. *Current Climate Change Reports*, *4*(3), 238-249.

Wang, C., Liu, H. and Lee, S. K. (2010) The record-breaking cold temperatures during the winter of 2009/2010 in the Northern Hemisphere, *Atmos. Sci. Lett.*, 11(3), 161–168.

Wang, H., Schubert, S., Suarez, M., Chen, J., Hoerling, M., Kumar, A. and Pegion, P. (2009) Attribution of the seasonality and regionality in climate trends over the United States during 1950–2000. *Journal of Climate*, *22*(10), 2571-2590.

Wilhite, D.A., Morrow, K.C. and Shulski, M. (2017) Changes in Arctic Climate and Central US Weather Patterns: Is There a Link?. *Bulletin of the American Meteorological Society*, *98*(1), ES1-ES2.

Figure Captions

Figure 1. Illustration of S_m, S_p, S_a and tail swing metrics using artificially generated data. In each year, S_m is depicted by the height of the light gray rectangles and S_p is given by the rectangle's width. The quasi-vertical dotted lines denote S_a, which may be contained within the interval defined by S_p. The two complete tail swings in the data series are shown by the dark gray rectangles at the top (above the horizontal line marking the 95th percentile) and bottom (below the horizontal line marking the 5th percentile) of the graph.

Figure 2 Monthly frequency of occurrence of the highest annual percentile (a, d and g); lowest annual percentile (b, e and h) and the largest percentile change between adjacent periods (c, f and i) for 7-day (a-c), 30-day (d-f) and 90-day (g-i) periods ending in the given month for stations in the contiguous U.S.

Figure 3 Trends in Sm (a, d and g), Sp (b, e and h), and Sa (c, f, and i) for 7-day (a-c), 30-day (d-f) and 90-day (g-i) duration temperatures during the period 1950-2017 . Blue (red) symbols indicate negative (positive) trends, with filled circles indicating significance at the $\alpha = 0.05$ level.

Figure 4 As in Figure 3, but for the 1900-2017 period .

Figure 5 Field significance of the number of negative S*m and* positive S_p trends and statistically significant S*m* and S*p* trends for different durations by Assessment region during 1900-2017 (left panels) and 1950-2017 (right panels) . Dark and light blue squares indicate field significance at the α = 0.05 and 0.10 level, respectively. Combinations without shading indicate a lack of regional field significance.

Figure 6 Time series of annual S*m* for 7- (a and b), 30- (c and d) and 90- (e and f) day temperature durations. The leftmost panels are averaged over stations in the contiguous U.S., the rightmost panels are averages for the Great Plains Climate Assessment region. The blue lines are based on stations with available data in the 1900-2017 period and the green lines based on stations with data in the 1950-2017 period. The red lines are the linear least squares slopes associated with each period.

Figure 7 As in Figure 6, but for annual S*p.*

Figure 8 Trends in decadal tail swing frequency (a, c and e) and decadal average period between tail swings (b, d, and f) for 7-day (a-b), 30-day (c-d) and 90-day (e-f) duration temperatures during the period 1900-2017 . Blue (red) symbols indicate significant (α = 0.05 level) negative (positive) trends.

Figure 9 As in Figure 8, but for 1950-2017.

Figure 10 Average number of tail swing occurrences (a, d and g); average time between tail swings (b, e, and h) and maximum time between tail swings (c, f, and i) by decade based on 7-day (a-c), 30-day (d-f) and 90-day (g-i) average temperatures. Dark bars represent stations with data available in the 1900-2017 period, lighter bars show stations with data in the 1950-2017 period. The Midwest Climate Assessment region is shown.