### Title
Monitoring and understanding trends in extreme storms: State of knowledge




### Journal
Bulletin of the American Meteorological Society, 94(4)

### ISSN
0003-0007



### Authors
Kunkel, KE
Karl, TR
Brooks, H
et al.



### Publication Date
2013-04-01

### DOI
10.1175/BAMS-D-11-00262.1

Peer reviewed

The DOI for this manuscript is doi: 10.1175/BAMS-D-11-00262.1

# Monitoring and Understanding Trends in Extreme Storms: State of Knowledge

Kenneth E. Kunkel[1], Thomas R. Karl[2], Harold Brooks[3], James Kossin[2], Jay H. Lawrimore[2], Derek Arndt[2], Lance Bosart[4], David Changnon[5], Susan L. Cutter[6], Nolan Doesken[7], Kerry Emanuel[8], Pavel Ya. Groisman[2], Richard W. Katz[9], Thomas Knutson[10], James O'Brien[11], Christopher J. Paciorek[19], Thomas C. Peterson[2], Kelly Redmond[12], David Robinson[13], Jeff Trapp[14], Russell Vose[2], Scott Weaver[15], Michael Wehner[16], Klaus Wolter[17], Donald Wuebbles[18]

[1]NOAA Cooperative Institute for Climate and Satellites, National Climatic Data Center
[2]NOAA National Climatic Data Center
[3]NOAA National Severe Storms Laboratory
[4]State University of New York-Albany
[5]Northern Illinois University
[6]Hazards and Vulnerability Research Institute, Department of Geography, University of South Carolina, Columbia, South Carolina
[7]Colorado State University
[8]Massachusetts Institute of Technology
[9]National Center for Atmospheric Research
[10]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey
[11]Florida State University
[12]Desert Research Institute
[13]Rutgers University
[14]Purdue University
[15]NOAA Climate Prediction Center
[16]Lawrence Berkeley National Laboratory, Berkeley, California
[17]University of Colorado-CIRES Climate Diagnostics Center and NOAA-ESRL Physical Science Division, Boulder, Colorado
[18]University of Illinois at Urbana-Champaign
[19]Department of Statistics, University of California at Berkeley, Berkeley, California

**Abstract**

The state of knowledge regarding trends and an understanding of their causes is presented for a specific subset of extreme weather and climate types. For severe convective storms (tornadoes, hail storms, and severe thunderstorms), differences in time and space of practices of collecting reports of events make the use of the reporting database to detect trends extremely difficult. Overall, changes in the frequency of environments favorable for severe thunderstorms have not been statistically significant. For extreme precipitation, there is strong evidence for a nationally-averaged upward trend in the frequency and intensity of events. The causes of the observed trends have not been determined with certainty, although there is evidence that increasing atmospheric water vapor may be one factor. For hurricanes and typhoons, robust detection of trends in Atlantic and western North Pacific tropical cyclone (TC) activity is significantly constrained by data heterogeneity and deficient quantification of internal variability. Attribution of past TC changes is further challenged by a lack of consensus on the physical linkages between climate forcing and TC activity. As a result, attribution of trends to anthropogenic forcing remains controversial. For severe snowstorms and ice storms, the number of severe regional snowstorms that occurred since 1960 was more than twice that of the preceding 60 years. There are no significant multi-decadal trends in the areal percentage of the contiguous U.S. impacted by extreme seasonal snowfall amounts since 1900. There is no distinguishable trend in the frequency of ice storms for the U.S. as a whole since 1950.

**Capsule Summary**

The state of knowledge regarding trends and an understanding of their causes is presented for severe convective storms, extreme precipitation, hurricanes and typhoons, and severe snowstorms and ice storms.

## 1. Introduction

The record for the number of weather and climate disasters that exceeded $1 billion (U.S.) or more in losses was set in 2011 (http://www.ncdc.noaa.gov/oa/reports/billionz.html). Twelve of the fourteen events counted in this record were related to storms, including severe local weather (tornadoes), storm related excessive precipitation, snowstorms/blizzards, and hurricane/tropical storms[1]. There is broad recognition that our climate is non-stationary and changing (Global Climate Change Impacts in the US 2009), not only in mean conditions but in its extremes as well (Katz 2010). However, there is less certainty in our ability to detect multi-decadal changes in each of these phenomena, and to understand the causes for any changes we can detect. This motivates our interest in a status report on our ability to detect, analyze, and understand changes in the risk of weather and climate extremes. Due to the intense media coverage of and great public interest in the 2011 disasters, we suspect that many BAMS readers have received inquiries or have a personal interest about the nature of these events in the context of long-term trends and potential climate change. This paper is meant to present a clear record that can be used by meteorological professionals about what is known and unknown and why.

This paper examines a specific subset of extreme weather and climate types affecting the United States. For our purposes, storm-related extremes here refer to those short duration

[1] The observed changes in losses represent a combination of the effects of both physical climate and socio-economic variability (e.g., Pielke et al. 2008), and it is difficult to attribute any of these changes to climate (Bouwer 2011). Here we will concentrate on physical climate variability. The non-storm disasters were the Texas, Arizona, New Mexico wildfires and the Southern Plains/Southwest drought and heat wave.

events that have levels/types of wind and/or precipitation at local to regional scales that are uncommon for a particular place and time of year (Peterson et al. 2008). The categories of storms described herein were chosen because they often cause property damage and loss of life, but the identification of an extreme occurrence is based on meteorological properties, not on the destructiveness. Our primary purpose is to examine the scientific evidence for our capability to detect trends and understand their causes for the following weather types: (1) severe convective storms (tornadoes, hail storms, and severe thunderstorms), (2) extreme precipitation, (3) hurricanes and typhoons, and (4) severe snowstorms and ice storms. These storm categories are not independent. Extreme precipitation can occur in any of the other three. Categories 1 and 4 are both typically associated with extratropical cyclones and sometimes in the same one. Nevertheless, the particular impacts are distinct and thus a separate examination of each of these is warranted.

The reason society ultimately cares about variability and change in the above physical phenomena is that these translate into socio-economic and biophysical impacts (e.g. life, property, ecosystems). The assessment of changes in the physical phenomena is just the first step. It is essential that trends in the impacts also be assessed in a comprehensive manner. As will be addressed later, this second step is quite challenging.

## 2. Severe Convective Storms: Thunderstorms, Tornadoes, and Hail Storms

Severe thunderstorms (hail of at least 2.5 cm or wind gusts of more than 95 km/h) and tornadoes pose challenging problems in efforts to establish temporal trends. In general, reports of such events in the US are collected to verify weather warnings and, as such, changes in verification efforts and emphasis are likely to have led to most, if not all, of the reported changes in frequency. The problems have been discussed by Doswell et al. (2005) and Verbout et al. (2006). The occurrence of F1+ tornadoes shows no trend since 1954, the first year of near real-time data collection, with all of the increase in tornado reports resulting from an increase in the weakest tornadoes, F0 (Fig. 1). Stronger events may be more reliably reported than weaker events, but changes in tornado damage assessment procedures still lead to problems in trend identification (Doswell et al. 2009). Changnon and Changnon (2000) used reports from first-order station observers for the 20th century to assess severe weather conditions and found considerable regional variability in the incidence of hail—increasing trends in some areas, decreasing trends elsewhere. The change from human observers to automated stations beginning in the 1990s influences the comparability of observations from the past to the future. Due to the changing practices and the nature of rare events, we have little confidence in the accuracy of trends in the meteorological occurrence of severe thunderstorms (including hail storms) and tornadoes.

Since raw reports are fraught with difficulties, attention has focused on examining the environmental conditions associated with severe thunderstorms to estimate the frequency and distribution of events (Brooks et al. 2003). This is guided by our understanding of the ingredients for severe thunderstorm occurrence derived from studies of day-to-day weather forecasting (Rasmussen and Blanchard 1998). The quality of severe thunderstorm forecasts indicates that the understanding of the physical processes is relatively good (Moller 2001). For example, using measures of the potential energy available for storms and the organizing potential of tropospheric shear, discrimination between severe and non-severe thunderstorms is possible (Fig. 2). Severe thunderstorms occur in an environment with large values of potential energy and wind shear, and tornadoes, in particular, are favored in high shear environments. Moist enthalpy, combining temperature and moisture content, near the earth's surface has been increasing in recent decades (Peterson et al. 2011). By itself, this would lead to an increase in thunderstorms, but changes above the Earth's surface could reduce or counteract that effect with unknown impacts on the initiation of thunderstorms. Brooks and Dotzek (2008) found long-term changes in the overall occurrence of favorable conditions for severe thunderstorms, but the interannual variability in their study was so large as to make the results statistically insignificant. Trapp et al. (2009) used an ensemble of global climate model simulations for the second half of the 20$^{th}$ century and found qualitatively similar changes in the severe thunderstorm environments; however, the large observed interannual variability implies that statistical significance of trends may not be reached for several more decades. The use of high-resolution models to dynamically downscale such climate data has the potential of providing an alternative to the observation-based and storm-environment-based approaches mentioned above (Trapp et al. 2011).

## 3. Extreme Precipitation

The occurrence of extreme precipitation rates requires abundant atmospheric water vapor and strong upward motion. Upward motion arises from three principal mechanisms: dynamical forcing, release of convective instability, and orographic forcing. Depending on the situation, all of these mechanisms can make a significant contribution to a specific event. In the U.S., the principal meteorological phenomena associated with extreme precipitation events include extratropical cyclones (ETCs), tropical cyclones (TCs), mesoscale convective systems, and the North American Monsoon (Kunkel et al. 2011).

The U.S. observing network is better suited for the assessment of changes in very heavy precipitation than for any other class of extreme storm. For instance, the NWS Cooperative Observer Network (COOP) network has largely employed the same standard 8” nonrecording precipitation gauge throughout its history (Yang et al. 1998), minimizing time-dependent biases resulting from changes in instrumentation. Furthermore, the gauge itself exhibits only a minor wind-driven bias in measuring large amounts of liquid precipitation (Groisman and Legates 1994). In addition, field experiments (Sevruk 1982) and theoretical results (Folland 1988) show that gauge undercatch is not substantial in very heavy rainfall. From a spatial perspective, the U.S. COOP network is of sufficient density for the detection of changes in very heavy precipitation over most regions (Groisman et al. 2005), except for some high elevations in the west. The COOP data do not distinguish between convective and non-convective precipitation.

There are a variety of extreme precipitation metrics, analysis methods, observing stations sets, and time periods used in published trends studies, reflecting tradeoffs among these choices. Statistical methodological approaches tend to fall into two basic categories: purely empirically-based or those with a more theoretical basis. For the empirically-based methods, thresholds are defined in terms of the data distribution, statistics such as the frequency of threshold exceedance are calculated and aggregated across space, and trends fitted. For the theoretically-based methods, distributions from the statistical theory of extreme values (e.g., Coles 2001) are fitted to extreme statistics including seasonal or annual maxima and excesses over a high threshold, and with the provision for trends in the parameters of these extremal distributions. The advantages of the purely empirical approach include being automatically applied and relatively powerful in detecting any trends, and being relatively easy to explain to non-specialists; its disadvantages include providing information only in aggregate terms for large regions and only applicable to moderately extreme events.  The advantages of methods based on extreme value theory include providing information in a form useful to decision and policy makers (i.e., in terms of return levels that apply locally and to the most extreme events of greatest societal relevance); its disadvantages include difficulty in being routinely applied (e.g., requiring a choice of threshold for the statistical theory to be a reasonable approximation) and the lack of a straightforward way to account for the spatial dependence of extremes in trend analyses. The choice of metrics often involves a tradeoff between the desire to examine trends in the low probability events that are most societally-relevant and the need to minimize sampling uncertainty by including less extreme, but more frequent events. The time period is often chosen on the basis of the number of stations with relatively complete data. In this case, there is a tradeoff between the desire to examine as long of a period as possible, but longer periods reduce the number of stations, and the need to minimize sampling uncertainty by including a minimum number of stations. The different choices that can be made are represented in the following set of analyses that are described below.

Many studies have found a statistically significant increase in the number and intensity of extreme precipitation events of durations ranging from hourly to a few days (Karl et al. 1996;

Karl and Knight 1998; Groisman et al. 2004; 2005; 2011; Kunkel et al. 2003, 2007; Global Climate Change Impacts in the United States 2009; Alexander et al. 2006). Given that trends in mean precipitation (+0.6% per decade; NOAA 2011) are less than extreme precipitation (2% per decade in top 1% of events; Kunkel et al. 2008), this apparently reflects a change in the tails of the distribution, rather than a shift in the entire distribution, over several decades compared to previous decades of the 20$^{th}$ century. The consistency of the results from these analyses reflects a degree of confidence in our ability to measure such changes in the U.S. For example, a set of precipitation-observing COOP stations with records extending back to around the turn of the 20$^{th}$ Century has been used to examine the temporal and spatial variations in number of extreme precipitation totals of 2-day duration exceeding a recurrence interval of 5 years. This duration was used to minimize instances of a single extreme precipitation event straddling the time of observation and the amount being split across the two days. Recurrence interval thresholds are used extensively in design of runoff control structure, which motivates their use as one component of a metric. Time series of station events were aggregated over decadal periods into 7 regions[2] of the coterminous U.S. and expressed as a spatially-averaged index (Fig. 3). There is considerable decadal-scale variability whose behavior often varies spatially (e.g. Mass et al. 2011). However, since 1991, all regions have experienced a greater than normal occurrence of extreme events. In the eastern regions, the recent numbers are the largest since reliable records begin (1895). For western regions, the recent decades are comparable to the early part of the historical record. Using the non-parametric Kendall's tau test for trends, the

[2] These are the regions being used for the 2013 National Climate Assessment Report

increase is statistically significant for the U.S. as a whole and the individual regions of the Midwest and Southeast (Table 1). Over the period 1957-2010, the Northeast region trend is also statistically significant. An analysis of another metric, the total amount of precipitation accumulated on days whose precipitation exceeds the 99$^{th}$ percentile for daily amounts, indicates a highly statistically significant upward trend for the period of 1957-2010 for the same set of regions (Midwest, Southeast, and Northeast) and the U.S. as a whole (Table 1); in this case, the results are robust to the choice of metric. No significant extreme precipitation trends are found in the western U.S. (see also Mass et al. 2011). Since the nature and magnitude of some impacts is sensitive to the duration of excessive precipitation, the sensitivity of results to the duration and return period has been studied (e.g. Kunkel et al. 2003, 2008) and qualitatively similar results have been found for durations of 1 to 90 days and return periods of 1 to 20 years in the definition of the metric.

The estimated change from 1948 to 2010 in the twenty year precipitation return value at individual stations based on daily accumulated precipitation station data (Fig. 4) from the Global Historical Climate Network-daily (Durre et al. 2008) were calculated using extreme value analysis (Tomassini and Jacob 2009; Cooley and Sain 2010). (see Supplemental Online Material for details). About 76% of all stations experience increases in extreme precipitation, with 15% showing a statistically significant increase based on station-specific hypothesis testing. From the central states to the north Atlantic these exhibit a high degree of spatial coherence. Regions with greater numbers of stations with decreases are of smaller spatial extent; the largest are in the northwest U.S. and the southern Appalachian Mountains. A field significance test was highly statistically significant. The choice of a 20-year return period in Fig. 4 is solely for illustrative purposes, with the estimated changes in return values for longer return periods being identical for this simplified form of extreme value analysis (see Supplementary Online Materials).

Figs. 3 and 4 and Table 1 display results for 3 different metrics. They are in best agreement over roughly the eastern half of the U.S., all indicating general upward trends. For the western half, the agreement is not as good; over the Great Plains and Southwest, the 20-yr return period threshold exhibits general upward trends in contrast to the lack of trends exhibited by the other two metrics.

Identification of the causes of long-term trends in extreme precipitation remains an area of active research, but some cogent work has already been completed. Globally, Min et al. (2009, 2011) have linked changes in extreme precipitation during the past several decades to human-caused changes in atmospheric composition.  Karl and Trenberth (2003) have empirically demonstrated that for the same annual or seasonal precipitation totals, warmer climates generate more extreme precipitation events compared to cooler climates.  This is consistent with water vapor being a critical limiting factor for the most extreme precipitation events. A number of analyses have documented significant positive trends in water vapor concentration and have linked these trends to human fingerprints in both changes of surface (Willet et al. 2007) and atmospheric moisture (Santer et al. 2007).

It is logical therefore to explore the connection.  The evidence in Table 2 from a pilot study (see Supplementary online material for details) depicts significant increases in the water vapor associated with extreme precipitation events, particularly east of the Rockies, and is suggestive that increases in water vapor in the environment of precipitation-producing systems may be a physical cause for the increase in intense precipitation events over the U.S.[3] In addition to the amount of water available for the generation of extreme precipitation events, dynamical factors must also be important. Even though there is no trend in U.S. landfalling Tropical Cyclones (TCs) (Global Climate Change Impacts in the United States, 2009), two studies found an upward trend in the number of extreme precipitation events associated with TCs (Knight and Davis 2009; Kunkel et al. 2010) while a third (Groisman et al. 2011) did not. There is also an upward trend in the number of extreme precipitation events in the vicinity of fronts associated with extra-tropical cyclones (Kunkel et al. 2011). However, there is no research indicating whether there has been a trend in the number and/or intensity of fronts. Gutowski et al. (2008) stated that the observed increases in extreme precipitation are "consistent with the observed increases in atmospheric water vapor, which have been associated with human-induced increases in greenhouse gases". While the role of water vapor as a primary cause for the increase in extreme precipitation events is compelling, the possibility of changes in the characteristics of meteorological systems cannot be ruled out. There may also be regional influences from the temporal redistribution of the number of El Nino events versus La Nina events and from land use changes such as the 20th Century increase in irrigation over the Great Plains and the post-World War II increase of corn and soybean acreage and planting density over the Midwest (DeAngelis et al. 2010; Groisman et al. 2011).

[3] In this analysis, each extreme precipitation event was assigned a precipitable water value, which was the maximum value from any radiosonde station within 300 km of the event location and within 24 hours of the observation time of the precipitation value.

## 4. Hurricanes and Typhoons

Detection of long-term changes in tropical cyclone (TC) activity has been hindered by a number of issues with the historical records. Heterogeneity introduced by changing technology and methodology is the major issue (e.g., Landsea et al. 2004). Data used to construct the historic “best track” archives are often initially collected and analyzed to support short-term forecasting needs using the best information, technology, and models of the day with no mandates in place to maintain heterogeneity. Improvements are generally implemented without any overlap or calibration against existing methods to document the impact of the changes on the longer-term climate record. The introduction of aircraft reconnaissance in some basins in the 1940s and satellite data in the 1960s had an important effect on our ability to identify and estimate the intensity of tropical cyclones, particularly those that never encountered land or a ship. The cessation in 1987 of regular aircraft reconnaissance into western North Pacific typhoons created a void in available *in situ* intensity measurements and our ability to calibrate satellite estimates against ground-truth, which adds further uncertainty to the records there. Efforts towards mitigation of these issues are ongoing, typically in the form of estimating storm frequency undercounts in the earlier parts of the Atlantic record (e.g., Vecchi and Knutson 2011), and using satellite data to construct less heterogeneous global records of storm intensity (e.g., Kossin et al. 2007). The latter efforts can be effective but at best are limited to the meteorological satellite era that began in the 1960’s, which limits their influence on trend detection on multi-decadal or longer time-scales. For example, comparisons between an index of tropical cyclone power dissipation (Emanuel 2005) derived from best track data versus a more homogeneous satellite reconstruction indicate high temporal consistency for the North Atlantic and somewhat less consistency for the western North Pacific since around 1980 (Fig. 5). The observed upward trend in the North Atlantic best track is robust to reanalysis, while the upward trend in the Pacific best track appears to be inflated by data heterogeneity issues.

Attempts to detect trends in intra-basin regions such as those defined by islands and archipelagos, or along coastlines are further constrained by the reduced data sample size associated with sub-setting the data. Intra-basin regional trend detection is also substantially challenged by variability in tropical cyclone tracks (e.g., Kossin et al. 2010; Holland 2007; Elsner 1998), which is driven largely by random fluctuations in atmospheric steering currents, but also is observed in response to more systematic climatic forcings such as El Niño / Southern Oscillation (ENSO). Landfalling tropical cyclone activity in the US, as well as East Asia, shows no significant long-term trends (e.g., Landsea 2005).

While data issues confound robust long-term (i.e., ~40-years or more) trend detection, trends in Atlantic TC frequency are robustly observed in the modern satellite period from around 1970 to present. In this case, the main challenge lies in attribution of these trends. A number of linkages between climate variability and TC activity have been well documented. In the tropical North Atlantic (tNA), observed climate variability and trends have been attributed using global climate models (e.g. Santer et al. 2006; Zhang 2007; Gillett et al. 2008; Ting et al. 2009; Zhang and Delworth 2009; Chang et al. 2011; Booth et al. 2012) or speculatively linked (e.g., Mann and Emanuel 2006; Evan et al. 2009) to a number of natural and anthropogenic factors. Natural multi-decadal internal variability of the North Atlantic is often referred to generically as the Atlantic Multi-decadal Oscillation (AMO) and has been linked, in modeling studies, to ocean thermohaline circulation variability (Delworth and Mann 2000). This variability is thought to contribute to the observed decadal variability of the tNA, but the robustness of evidence for this is presently a matter of debate. Natural tNA variability on shorter time-scales is also introduced by the North Atlantic Oscillation and remotely by ENSO via teleconnections. Uncertainties in the contribution of internal climate variability remain an important confounding factor (Hegerl et al. 2010) in the detection and attribution of climate trends in the tNA region. Owing to pronounced multidecadal variability evident in longer term records of Atlantic basin-wide or U.S. landfalling tropical cyclone frequency (e.g., Vecchi and Knutson 2011, see their Fig. 5), the period since around 1970 (e.g., Fig. 5) appears to be too short to draw confident inferences about longer term (e.g., century scale) trends in Atlantic tropical cyclone activity.

External forcing of the tropical climate can be natural or anthropogenic. Volcanoes are an important natural forcing agent, while greenhouse gas forcing has predominantly anthropogenic underpinnings. Attribution of forcing via aerosols is generally less clear. For example, sulfate aerosols occur naturally and are also a constituent of human-induced pollution. Sulfate aerosol concentration is associated with atmospheric dimming effects (e.g., Mann and Emanuel 2006) as well as changes in cloud albedo (e.g., Booth et al. 2012), both of which affect local external forcing. Concentrations of these and other aerosols have been reduced in the tNA subsequent to the US Clean Air Act amendments of the 1970s, but development in Asia has led to increased emissions in regions of the Indian and Pacific oceans, and one study has proposed a link between black carbon aerosol pollution and increased tropical cyclone intensity in the Arabian Sea (Evan et al. 2011). Mineral aerosols, such as dust transported westward over the tNA from the Sahara, are of natural origin, but may be at least partly modulated by human-induced land-use change. All of these forcings have been linked to tNA sea surface temperature (SST) variability, but significant questions remain about their relative contributions to the overall observed Atlantic hurricane variability. In terms of century-scale variability, only anthropogenic forcing has a *prima facie* expectation of introducing a significant trend on such time-scales, while inter-annual tropical variability can be largely attributed to natural fluctuations such as ENSO. Comparatively, attribution of the observed multi-decadal tNA variability is particularly uncertain and hypotheses span the range from mostly natural internal variability (e.g., Zhang and Delworth's (2009) attribution study for tNA vertical wind shear changes) to mostly external anthropogenic forcing (e.g., Mann and Emanuel 2006).

In addition to uncertainty about the relative contributions of the above forcings to the observed tNA variability, there is also uncertainty about how TCs respond to the ocean/atmosphere variability attributed to each individual forcing. Aerosol concentrations emanating from source regions are generally more spatially heterogeneous than greenhouse gas concentrations, and the AMO is generally associated with larger amplitude SST variations in the North Atlantic than in other basins. The nature of the forcing is important, because the response of tropical cyclone activity can be quite different for a given change in SST depending on the type of forcing. Thus, for example, reduced surface wind speeds will increase SSTs and also increase the thermodynamic potential for tropical cyclones, but the rate of increase in thermodynamic potential with SST will, in general, be much larger than if the same SST increase is brought about by increasing greenhouse gases (Emanuel 2007). This is because the degree of thermodynamic disequilibrium between the oceans and atmosphere depends directly on the net surface radiative flux, but inversely on surface wind speed. Thus SST is an imperfect proxy for the thermodynamic environment of tropical cyclones and it should not be used as the sole thermodynamic predictor of changing tropical cyclone activity. Nonetheless, analyses of potential intensity projections for the 21$^{st}$ century from CMIP3 climate models demonstrate that these modeled potential intensity changes are well correlated with changes in relative SST (i.e., the local SST relative to the tropical mean SST; Vecchi and Soden 2007).

In summary, robust detection of trends in Atlantic and western North Pacific TC activity is significantly constrained by data heterogeneity and deficient quantification of internal variability. Attribution of past TC changes is further challenged by a lack of consensus on the physical linkages between climate forcing and TC activity. As a result, attribution of any observed trends in TC activity in these basins to anthropogenic forcing remains controversial.

## 5. Severe Snow Storms and Ice Storms

Quantifying changes in the frequency, duration, and severity of winter storms requires the ability to accurately and consistently measure the amount of snow that falls and ice that accumulates during individual storms and throughout entire seasons. Changes in observing practices, reporting procedures, and observing technologies through time complicate these analyses. These include a transition from primarily afternoon to morning observation times, a gradual move to direct measurement from previous estimation of precipitation by “ten to one” snow to water ratio, and periodic changes in observer training practices. Although resulting artifacts in the climate record make analyses more difficult to accomplish, robust conclusions can be reached by selecting a subset of stations for which the snowfall record is of highest quality and which appear to have been minimally affected by non-climatic influences (Kunkel et al. 2009a, 2009b, 2009c). In addition, identification of extreme events such as severe regional snowstorms included here is likely less affected by changes in observing practices and procedures than the analysis of mean conditions.

The two most dominant factors that influence U.S. winter storm characteristics (trajectory, frequency, intensity) are the El Niño/Southern Oscillation (ENSO) and the North Atlantic Oscillation/Arctic Oscillation (N)AO phenomena. La Niña favors a more northerly storm track, bringing enhanced snow to the northern and central Rockies, while El Niño favors a more southerly storm track and potentially heavy precipitation in the southern states (e.g., Redmond and Koch 1991; Smith and O’Brien 2001). Over the last 110 years, ENSO behavior has varied greatly, with a period of low activity from the early 1930s into the late 1940s. During the most active periods, El Niño was favored early in the 20$^{th}$ century and from the mid-1970s to the late 1990s, while La Niña was most prominent from the 1950s to the mid-70s (Wolter and Timlin 2011).

The (N)AO, a dominant influence on eastern U.S. weather patterns also has undergone similar ‘regime changes’, favoring its positive phase in the early part and latter decades of the

20th century. More prominent spells of its negative phase occurred from the middle of the 20th century into the late 1960s. The last 15 to 20 years have seen a more even distribution of both phases, favoring the negative phase in the recent winters of 2009-2010 and 2010-2011 (Hurrell et al. 2003; Seager et al. 2010). Contributing factors to these regime changes are under investigation (e.g., L'Heureux et al. 2008; Allen and Zender 2011). The decadal scale variability of storm properties associated with each phenomenon can appear in observed records as a "trend," illustrating a need for caution before attribution to anthropogenic climate change.

The characteristics of what constitutes a severe winter storm vary regionally. Snowfall greater than 10 inches is common in many parts of the Northeast, and thus often only a short-term inconvenience. However, the same snowfall across the Southeast might cripple the region for a week or longer. A Regional Snowfall Index (RSI, Squires et al. 2009) has been formulated that takes into account the typical frequency and magnitude of snowstorms in each region of the eastern two-thirds of the U.S., providing perspectives on decadal changes in extreme snowstorms since 1900. An analysis based on the area receiving snowfall of various amounts shows there were more than twice the number of extreme regional snowstorms from 1961-2010 (21) as there were in the previous 60 years (9) (Figure 6). The greater number of extreme storms in recent decades is consistent with other findings of recent increases in heavier and more widespread snowstorms (Kocin and Uccellini 2004).

These extreme storms occurred more frequently in snow seasons that were colder and wetter than average (Fig. 6), but not exclusively. Approximately 35% of the snow seasons in which these events occurred were warmer than average and 30% drier than average. The implications are that even if temperatures continue to warm as they have over the past several decades, for the next few decades, at least, such record storms are possible as they have been observed during otherwise warmer- and drier-than-average seasons.

The impact of individual snowstorms is often immediate and dramatic, but the cumulative effects of all snowstorms in a season can also be costly and disruptive. Snowfall measured at approximately 425 high quality stations was used to assess variation and change in the percentage of the contiguous U.S. affected by extreme high or low seasonal snowfall since 1900 (Kunkel et al. 2009c). Observations do not show significant century-scale trends in either high or low seasonal totals. The areal percentage of the U.S. experiencing seasons with the heaviest accumulated snowfall (top 10%) was greatest in the 1910s, the 1960s and 1970s (Figure 7a). The areal percentage of the contiguous U.S. with unusually light seasonal snowfall totals (those in the lowest 10%) decreased from 1940 through the mid-1970s (Figure 7b). Areal coverage of extremely low seasonal snowfall has been steady or slightly increasing since that time.

It may appear contradictory that the number of extreme snowstorms could increase in the latter half of the 20$^{th}$ century (Fig. 6) without a coinciding decrease in areal coverage of extremely low seasonal snowfall totals (Fig. 7b). However, there should be no expectation that changes in the frequency of such extreme short-duration events, which can occur during otherwise unusually warm and snow-free seasons, would be correlated with trends in low seasonal snowfall totals. This is especially true in northern areas of the U.S. where seasonal snowfall totals can be lower than average even during years when an extreme snowstorm has occurred.

Severe winter conditions are not limited to heavy snowfall. Ice storms can disrupt transportation, and those exceeding certain threshold accumulations can cause catastrophic damage to ecosystems and infrastructure. Most freezing rain events occur east of the Rocky Mountains (Changnon and Creech 2003), and generally with less frequency than snow, particularly outside the South. Freezing rain climatologies typically begin in the mid-20$^{th}$ century, are generally limited to daily ("days with") values for a subset of stations, and at best only coarsely distinguish between different magnitudes. National and regional trends in the number of freezing rain days show no systematic trends since about 1960, after some regions experienced a relative maximum during the 1950s (Gay and Davis 1993; Changnon and Karl 2003).

Frozen precipitation and associated impacts will not disappear in a warmer world (Kodra et al. 2011), and means and extreme events may even increase, for example at elevations and latitudes where warmer conditions still remain below freezing.  Snow measurements are among the most challenging of all climate elements (Doesken and Judson 1996; Yang et al. 1998; Yang et al.2001), and climate analysis depends on a robust national system of reference stations, spanning all elevations, designed to track snow properties through time and to develop relations to other sensing technologies. Such a national system is especially important in measuring and assessing variations and trends in smaller amounts of snow and water content typical of low elevations (e.g., many cities and airports).

## 6. Discussion and Conclusions

The main conclusions of this scientific assessment are:

- Severe convective storms: thunderstorms, tornadoes, and hail storms- Differences in time and space of practices of collecting reports of events make the use of the reporting database to detect trends extremely difficult. Although some ingredients that are favorable for severe thunderstorms have increased over the years others have not, so that, overall, changes in the frequency of environments favorable for severe thunderstorms have not been statistically significant.
- Extreme precipitation-There is strong evidence for a nationally-averaged upward trend in the frequency and intensity of extreme precipitation events. The COOP network is considered adequate to detect such trends. The causes of the observed trends have not been determined with certainty, although there is evidence that increasing atmospheric water vapor may be one factor.
- Hurricanes and typhoons- Robust detection of trends in Atlantic and western North Pacific TC activity is significantly constrained by data heterogeneity and deficient quantification of internal variability. Attribution of past TC changes is further challenged by a lack of consensus on the physical linkages between climate forcing and TC activity. As a result, attribution of any observed trends in TC activity in these basins to anthropogenic forcing remains controversial.
- Severe snowstorms and ice storms-The number of severe regional snowstorms that occurred since 1960 was more than twice the number that occurred during the preceding 60 years. There are no significant multi-decadal trends in the areal percentage of the contiguous

U.S. impacted by extreme seasonal snowfall amounts since 1900. There is no distinguishable trend in the frequency of ice storms for the U.S. as a whole since 1950.

Figure 8 summarizes our scientific assessment of the current ability to detect multi-decadal changes and understand the causes of any changes, putting each phenomenon into one of three categories of knowledge from less to more. The position of each storm type was determined through extensive verbal discussion at a meeting of the author team to reach a group consensus. In terms of detection, the existing data for thunderstorm phenomena (hail, tornadoes, thunderstorm winds) are not considered adequate to detect trends with confidence. This is also the case with ice storms. The data adequacy for hurricanes and snow storms was judged to be of intermediate quality; although trends have been studied, there are a number of quality issues that add uncertainty to the results of such studies. The data adequacy for precipitation is of higher quality than the rest of the types, leading to higher confidence in the results of trend studies.

Knowledge of the potential physical causes of trends is higher for extreme precipitation than for other storm types while knowledge of causes for hail, tornadoes, hurricanes, and snow storms is intermediate among the types. The adequacy of knowledge is quite low for thunderstorm winds and ice storms.

Improving the status of the data and understanding can be advanced through the following steps:

- Severe convective storms- Consistent collection of severe thunderstorm and tornado reports that does not depend upon the severe weather warning process would be necessary to make the time series of reports useful for climate-scale purposes. Alternatively, development of objective remotely-sensed observations, most likely based upon radar, that serve as proxies for actual severe weather events could address issues, although challenges will exist as radar technology changes.

- Extreme precipitation-It is essential that the high quality data network be maintained so that future variations and trends can be detected. The role of water vapor trends as possible cause of extreme precipitation trends should be more thoroughly explored.
- Hurricanes and typhoons-Better understanding of factors controlling tropical cyclone variability will be realized through the development of improved theoretical frameworks, numerical and statistical modeling, and observations. Improved observations will most likely result from additional observing platforms, both in situ (e.g., expanded manned or unmanned aircraft reconnaissance and/or tethered blimps such as the Aeroclipper) and remote (e.g., better microwave and scatterometer coverage). Consistency of the data is essential, and calibration periods are needed when new instruments or protocols are introduced so that biases can be quantified and data heterogeneity can be minimized.
- Severe snow storms and ice storms-A high priority is reducing uncertainties in the historical record through the incorporation of new sources of data and development and application of techniques that properly account for changing technologies and observing practices that have occurred through time. This should be done while also creating a robust national system of observing stations with sufficient density spanning all elevations, integrating new

technologies, and employing well documented and consistent observing and reporting practices

The identification and understanding of trends in impacts shares many of the same difficulties, such as data quality and attribution of impacts, found for trends in the meteorological phenomena discussed here. For example, temporal and spatial changes in social vulnerability (Cutter and Finch 2008) make detection of robust trends on outcomes of small-scale meteorological events very challenging. As with the physical climate extreme data, changes in practices of economic loss reporting and attribution over time have occurred. Different datasets record information on different classes of events, not all parameters are collected, and the duration of the record is variable as well (Gall et al. 2009). Metrics that are recorded vary in precision and, in some cases, techniques attempting to adjust for population, wealth, mortality, or type of loss (insured/uninsured; direct/indirect) are inconsistent making cross-database comparisons very difficult.

## 7. Acknowledgements

We thank Isaac Held for helpful comments on the Hurricane section. We also thank Imke Durre, Xungang Yin, Jared Rennie, Michael Palecki, and David Wuertz of NCDC for analytical assistance. The four reviewers provided valuable suggestions. This work was partially supported by National Oceanic and Atmospheric Administration Climate Program Office award NA07OAR4310063. Any opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of NOAA or the institutions for which they work. Christopher

Paciorek and Michael Wehner were supported by the *Office of Biological and Environmental Research* of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

## 8. References

Alexander, L.V., X. Zhang, T.C. Peterson, J. Caesar, B. Gleason, A.M.G. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, K. Rupa Kumar, J. Revadekar, G. Griffiths, L. Vincent, D.B. Stephenson, J. Burn, E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, and J.L. Vazquez-Aguirre, 2006: Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, **111**, D05109, doi:10.1029/2005JD006290.

Allen, R. J. and C. s. Zender, 2011: Forcing of the Arctic Oscillation by Eurasian Snow Cover. *J. Climate*, **24,** (in press).

Booth, B.B.B.,N. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, doi:10.1038/nature10946.

Bouwer, L.M., 2011: Have disaster losses increased due to anthropogenic climate change? *Bull. Amer. Meteor. Soc*., **92**, 39-46.

Brooks, H. E., and N. Dotzek, 2008: The spatial distribution of severe convective storms and an analysis of their secular changes. *Climate Extremes and Society*. H. F. Diaz and R. Murnane, Eds., Cambridge University Press, 340 pp, 35-53.

Brooks H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94.

Chang, C.-Y., J. C. H. Chiang, M. R. Wehner, A. F. Friedman, and R. Ruedy, 2011: Sulfate aerosol control of tropical Atlantic climate over the twentieth century. *J. Climate*, **24**, 2540-2555.

Changnon S. A., and D. Changnon, 2000: Long-term fluctuations in hail incidences in the United States. *J. Clim.,* **13**, 658–664.

Changnon, S.A. and T.G. Creech, 2003: Sources of data on freezing rain and resulting damages. *J. Appl. Meteor.*, 42, 1514–1518. doi: 10.1175/1520-0450(2003).

Changnon, S. A. and T. R. Karl. 2003. Temporal and spatial variations of freezing rain in the contiguous United States: 1948–2000. *J. Appl. Meteor.* 42:1302–1315.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*, Springer, London, 208 pp.

Cooley, D. and S.R. Sain, 2010: Spatial hierarchical modeling of precipitation extremes from a regional climate model. *Journal of Agricultural, Biological, and Environmental Statistics* **15**, 381-402.

Craven, J. P., and H. E. Brooks, 2004: Baseline climatology of sounding derived parameters associated with deep, moist convection. *Nat. Wea. Digest*, **28**,13-24.

Cutter, S. L., and C. Finch, 2008: Temporal and spatial changes in social vulnerability to natural hazards. *Proc. Nat. Acad. Sci.*, **104**, 2301-2306, doi:10.1073/pnas.0710375105

DeAngelis, A., F. Dominguez, Y. Fan, A. Robock, M. D. Kustu, and D. Robinson, 2010: Observational evidence of enhanced precipitation due to irrigation over the Great

Plains of the United States. *J. Geophys. Res.*, **115**, D15115, 14 pp., doi:10.1029/2010JD013892]

Delworth, T. L., and M E Mann, 2000: Observed and simulated multidecadal variability in the Northern Hemisphere. *Clim. Dyn*., **16,** 661-676.

Doesken, N.J., and A. Judson, 1996. The Snow Booklet. Colorado Climate Center, Atmospheric Science Department, Colorado State University Bulletin no. RC2, ISBN 0-9651056-2-8, 32 pp.

Doswell, C. A. III, H. E. Brooks, and M. P. Kay, 2005: Climatological estimates of daily local nontornadic severe thunderstorm probability for the United States. *Wea. Forecasting*, **20**, 577-595.

Doswell, C. A. III, H. E. Brooks, and N. Dotzek, 2009: On the implementation of the Enhanced Fujita Scale in the USA. *Atmos. Res.*, **93,** 554-563.

Durre, I., M. J. Menne, and R. S. Vose, 2008: Strategies for evaluating quality-control procedures. *J. Clim. Appl. Meteor.*, **47**, 1785-1791.

Elsner, J. B. (2003). Tracking hurricanes. *Bull. Amer. Meteor. Soc*., **84**, 353–6.

Emanuel, K. A., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686-688.

Emanuel, K., 2007: Environmental factors affecting tropical cyclone power dissipation. *J. Climate*, **20**, 5497-5509.

Evan, A. T., D. J. Vimont, A. K. Heidenger, J. P. Kossin, and R. Bennartz, 2009: The dominant role of aerosols in the evolution of tropical Atlantic Ocean temperatures. *Science*, DOI: 10.1126/science.1167404

Evan, A., T., J. P. Kossin, C. Chung, and V. Ramanathan, 2011: Arabian Sea tropical cyclones intensified by emissions of black carbon and other aerosols. Nature, 479, 94-97.

Folland, C. K., 1988: Numerical models of the rain-gauge exposure problem, field experiments, and an improved collector design. *Quart. J. Roy. Meteor. Soc*., **114**, 1485–1516.

Gall, M., K. A. Borden, and S. L. Cutter, 2009: When do losses count? Six fallacies of natural hazards loss data. *Bull. Amer. Meteor.,* **90**, 799-809.

Gay, D.A., and R.E. Davis, 1993: Freezing rain and sleet climatology of the southeastern USA. *Climate Research*, **3,** 209-220.

Gillett, N.P., P.A. Stott, and B.D. Santer, 2008a: Attribution of cyclogenesis region sea surface temperature change to anthropogenic influence. *Geophys. Res. Lett.*, **35**, L09707.

Global Climate Change Impacts in the United States, 2009: T.R.Karl, J.M. Melillo, and T.C. Peterson, (eds.) Cambridge University Press.

Groisman, P. Ya, and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc*., **75**, 215–227.

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. *J. Clim.*, **18**, 1326-1350.

Groisman, Pavel Ya, Richard W. Knight, Thomas R. Karl, David R. Easterling, Bomin Sun, Jay H. Lawrimore, 2004: Contemporary Changes of the Hydrological Cycle over the Contiguous

United States: Trends Derived from In Situ Observations. *J. Hydrometeorol.*, **5**, 64–85. doi: 10.1175/1525-7541(2004)005<0064:CCOTHC>2.0.CO;2

Groisman, P, Ya., R. W. Knight, and T.R.Karl, 2012: Changes in intense precipitation over the central US. *J. Hydrometeorol.*, 13, 47-66. doi: 10.1175/JHM-D-11-039-1.

Gutowski, W.J., G.C. Hegerl, G.J. Holland, T.R. Knutson, L.O. Mearns, R.J. Stouffer, P.J. Webster, M.F. Wehner, F.W. Zwiers, 2008: Causes of Observed Changes in Extremes and Projections of Future Changes in *Weather and Climate Extremes in a Changing Climate.Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*. T.R. Karl, G.A. Meehl, C.D. Miller, S.J. Hassol, A.M. Waple, and W.L. Murray (eds.). A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.

Hegerl, G.C., O. Hoegh-Guldberg, G. Casassa, M.P. Hoerling, R.S. Kovats, C. Parmesan,D.W. Pierce, P.A. Stott, 2010: Good Practice Guidance Paper on Detection and Attribution Related to Anthropogenic Climate Change. In: *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change* [Stocker, T.F., C.B. Field, D. Qin, V. Barros, G.-K. Plattner, M. Tignor, P.M. Midgley, and K.L. Ebi (eds.)]. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland.

Helsel, D.R., and R.M. Hirsch, 1993: *Statistical Methods in Water Resources*, Elsevier, Amsterdam, 522 pp.

Holland, G. J., 2007: Misuse of landfall as a proxy for Atlantic tropical cyclone activity. *EOS Trans. Amer. Geophys. Union*, **88**, 349-350.

Hollander, M., and D.A. Wolfe, 1973: *Nonparametric Statistical Methods*, John Wiley & Sons, New York, 503 pp.

Hurrell, J.W., Y. Kushnir, M. Visbeck, and G. Ottersen, 2003: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climate Significance and Environmental Impact,* J.W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck, Eds. Geophysical Monograph Series, **134**, pp. 1-35.

Karl, T. R, R.W. Knight, D.R. Easterling, and R.G. Quayle, 1996: Indices of climate change for the United States, *Bull. Amer. Meteor.,* **77**, 279-292.

Karl, T.R. and R.W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States., *Bull. Amer. Meteor.,* **79**, 231-241.

Karl, T.R. and K.E. Trenberth, 2003: Modern Climate Change. *Science 5 December 2003: Vol. 302 no. 5651 pp. 1719-1723. DOI: 10.1126/science.1090228*

Katz, R.W., 2010: Statistics of extremes in climate change. *Climatic Change*, **100**, 71-76.

Knight, D. B. and R. E. Davis, 2009: Contribution of tropical cyclones to extreme rainfall events in the southeastern United States. *J. Geophys. Res.* , **114** , D23102, doi:10.1029/2009JD012511.

Kocin, P.J., L.W. Uccellini, 2004: A snowfall impact scale derived from northeast storm snowfall distributions. *Bull. Amer. Meteor. Soc.*, **85**, 177–194. doi: 10.1175/BAMS-85-2-177

Kodra, E., K. Steinhaeuser, and A.R. Ganguly, 2011. Persisting cold extremes under 21st-century warming scenarios. *Geophys. Res. Lett.,* **38**, L08705. doi:10.1029/2011GL047103.

Kossin, J. P., K. R. Knapp, D. J. Vimont, R. J. Murnane, and B. A. Harper, 2007: A globally consistent reanalysis of hurricane variability and trends. *Geophys. Res. Lett.*, **34**, L04815, DOI:10.1029/2006GL028836.

Kossin, J.P., S.J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057-3076.

Kunkel, K.E., D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2010: Trends in U.S. heavy precipitation caused by tropical cyclones. *Geophys. Res. Lett.*, doi:10.1029/2010GL045164.

Kunkel, K. E., D.R. Easterling, K. Redmond, and K. Hubbard, 2003: Temporal variations of extreme precipitation events in the United States: 1895–2000, *Geophys. Res. Lett.*, **30**, 1900, 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. *J. Hydrometeor*., **8**, 1152-1160.

Kunkel, K.E., D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2011: Meteorological causes of the secular variations in observed extreme precipitation events for the conterminous United States. *J. Hydrometeor*., in press.

Kunkel, K.E., M. Palecki, L. Ensor, K.G. Hubbard, D. Robinson, K. Redmond, and D. Easterling, 2009a. Trends in 20th Century U.S. snowfall using a quality-controlled data set. *J. Atmos. Ocean. Tech.*, 26(1), 33-44. DOI: 10.1175/2008JTECHA1138.1

Kunkel, K.E., L. Ensor, M. Palecki, D. Easterling, D. Robinson, K. G. Hubbard, K. Redmond, 2009b. A new look at lake-effect snowfall trends in the Laurentian Great Lakes using a temporally homogeneous data set. *Journal of Great Lakes Research*, 35(1):23-29. doi: 10.1016/j.jglr.2008.11.003.

Kunkel, K.E., M.A. Palecki, L. Ensor, D. Easterling, K.G. Hubbard, D. Robinson, and K. Redmond, 2009c: Trends in Twentieth-Century U.S. Extreme Snowfall Seasons. *J. Climate*, **22**, 6204–6216. doi: 10.1175/2009JCLI2631.1

L'Heureux, Michelle L., R. Wayne Higgins, 2008: Boreal Winter Links between the Madden–Julian Oscillation and the Arctic Oscillation. *J. Climate*, **21**, 3040–3050.

Landsea, C. W., 2005: Hurricanes and global warming. Nature, 438, E11–E12, doi:10.1038/nature04477.

Landsea, C.W., C. Anderson, N. Charles, G. Clark, J. Dunion, J. Fernandez-Partagas, P. Hungerford, C. Neumann, and M. Zimmer, 2004: The Atlantic hurricane database re-analysis project: Documentation for the 1851-1910 alterations and additions to the HURDAT database. In: Hurricanes and Typhoons: Past, Present and Future [R.J. Murnane, and K.B. Liu (eds.)]. Columbia University Press, New York, pp. 177-221.

Mann, M. E., and K. Emanuel, 2006: Atlantic hurricane trends linked to climate change. *Eos, Trans. Amer. Geophys. Union*, **87**, 233–244.

Mass, C., A. Skalenakis, and M. Warner, 2011: Extreme precipitation over the West Coast of North America: Is there a trend? *J. Hydrometeor*., 10.1175/2010JHM1341.1

Min, S.-K., X. B. Zhang, F.W. Zwiers, P. Friederichs, and A. Hense, 2009: Signal detectability in extreme precipitation changes assessed from twentieth century climate simulations. *Clim. Dyn.*, **32**, 95–111.

Min S.K., X. Zhang, F.W. Zwiers, G.C. Hegerl., 2011; Human contribution to more-intense precipitation extremes. *Nature,* 470, 378-381.

Moller, A. R., 2001: Severe local storms forecasting. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 433–480.

NOAA National Climatic Data Center, State of the Climate: National Overview for Annual 2011, published online December 2011, retrieved on April 27, 2012 from http://www.ncdc.noaa.gov/sotc/national/2011/13.

Peterson, T.C., D.M. Anderson, S.J. Cohen, M. Cortez-Vázquez, R.J. Murnane, C. Parmesan, D. Phillips, R.S. Pulwarty, J.M.R. Stone, 2008: Why Weather and Climate Extremes Matter in *Weather and Climate Extremes in a Changing Climate. Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*. T.R. Karl, G.A. Meehl, C.D. Miller, S.J. Hassol, A.M. Waple, and W.L. Murray (eds.). A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.

Peterson, Thomas C., Katharine M. Willett, and Peter W. Thorne, 2011: Observed changes in surface atmospheric energy over land. *Geophys. Res. Lett.*, **38**, L16707, doi:10.1029/2011GL048442.

Pielke, R.A., Jr., C.W. Gratz, C.W. Landsea, D. Collins, M.A. Saunders, and R. Musulin, 2008: Normalized hurricane damage in the United States: 1900-2005. *Natural Hazards Review*, **9**, 29-42.

Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148-1164.

Redmond, K.T., and R.W. Koch, 1991. Surface Climate and Streamflow Variability in the Western United States and Their Relationship to Large Scale Circulation Indices. *Water Resour. Res.,* **27,** 2381-2399.

Santer, B.D., T.M.L. Wigley, P.J. Glecker, C. Bonfils, M.F. Wehner, K. AchutaRoa, T.P. Barnett, J.S. Boyle, W. Brüggemann, M. Fiorino, N. Gillet, J.E. Hansen, P.D. Jones, S.A. Klein, G.A. Meehl, S.C.B. Raper, R.W. Reynolds, K.E. Taylor, and W.M. Washington, 2006: Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions. *Proceedings of the National Academy of Sciences*, 103(38), 13905-13910.

Santer, B.D., C. Mears, F.J. Wentz, K.E. Taylor, P. J. Gleckler, T,.M.L. Wigley, T. P. Barneett, J. S. Boyle, W. Bruggermann, N. P. Gillett, S. A. Klein, G. A. Meehl, T. Nozawa, D. W. Pierce, P. A. Stott, W. M. Washington, and M. F. Wehner, 2007: Identification of human-induced changes in atmospheric moisture content. *Proceedings of the National Academy of Science*, 104(39), 15248-15253.

Seager, R., Y. Kushnir, J. Nakamura, M. Ting, and N. Naik (2010), Northern Hemisphere winter snow anomalies: ENSO, NAO and the winter of 2009/10, *Geophys. Res. Lett.*, *37*, L14703, doi:10.1029/2010GL043830.

Sevruk, B., 1982: Methods of correction for systematic error in point precipitation measurement for operational use. *Operational Hydrology Report No. 21*, World Meteorol. Organ., 91 pp.

Smith, S. R., and J. J. O'Brien, 2001: Regional snowfall distributions associated with ENSO: Implications for seasonal forecasting. *Bull. Amer. Meteor.*, **82**, 1179-1191.

Squires, M.F., J.H. Lawrimore, R.R. Heim, D.A. Robinson, M. Gerbush, T. Estilow, C. Tabor, and A. Wilson, 2009: Development of new snowstorm indices and databases at the National Climatic Data Center. *Eos, Transactions of the AGU*, 90(52), Fall Meeting Supplement, Abstract IN13A-1076 Poster.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and natural 20th century SST trends in the North Atlantic. *J. Climate*, **22**, 1469–1481.

Tomassini, L. and D. Jacob, 2009: Spatial analysis of trends in extreme precipitation events in high-resolution climate model results and observations for Germany. *J. Geophys. Res.*,**114**, D12113.

Trapp R. J., N. S. Diffenbaugh, and A. Gluhovsky, 2009: Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophys. Res. Lett.*, **36**, L01703.

Trapp, R.J., E.D. Robinson, M.E. Baldwin, N.S. Diffenbaugh, and B.R.J. Schwedler, 2011: Regional climate of hazardous convective weather through high-resolution dynamical downscaling. *Clim. Dynamics*, 10.1007/s00382-010-0826-y

Vecchi, G. A., and T. R Knutson, 2011: Estimating annual numbers of Atlantic hurricanes missing from the HURDAT database (1878-1965) using ship track density. *J. Climate*, **24**, doi:10.1175/2010JCLI3810.1.

Vecchi, G. A., and B. J. Soden, 2007: Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature*, 450(7172), DOI:doi:10.1038/nature06423.

Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the US tornado database: 1954-2003. *Wea. Forecasting*, **21**, 86-93.

Villarini, G., F. Serinaldi, J.A. Smith, and W.F. Frajewski, 2009: On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, **45**, W08417, doi:10.1029/2008WR007645.

Willett, K.M., N.P. Gillett, P.D. Jones, and P.W. Thorne, 2007: Attribution of observed surface humidity changes to human influence. *Nature*, **449**, 710-712. doi:10.1038/nature06207.

Wolter, K., and M.S. Timlin, 2011: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *International J. Climatology*, **31**, 1074-1087

Yang, D, B.E. Goodison, J.R. Metcalfe, V.S. Golubev, R. Bates, T. Pangburn, and C. Hanson, 1998. Accuracy of NWS 8" standard nonrecording precipitation gauge: Results and application of WMO Intercomparison. *J. Atmos. Ocean. Tech*., **15**, 54-68.

Yang, D., B.E. Goodison, J. Metcalfe, P. Louie, E. Elomaa, C. Hanson, V. Golubev, T. Gunther, J. Milkovic, and M. Lapin, 2001. Compatibility evaluation of national precipitation gage measurements. *J. Geophys. Res*., **106**, D2, 1481-1491.

Zhang, R. 2007: Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic. *Geophys. Res. Lett.*, 34, L12713, doi:10.1029/2007GL030225.

Zhang, R., and T. L. Delworth, 2009: A new method for attributing climate variations over the Atlantic Hurricane Basin's main development region. *Geophys. Res. Lett.*, **36**, L06701, DOI:10.1029/2009GL037260.

**Figure Captions**

Figure 1. Reported tornadoes in NWS database from 1950-2011. Blue line is F0 tornadoes, red dots are F1 and stronger tornadoes.

Figure 2. Convective parameters from 0000 UTC soundings in US 1997-9. “Wmax” is vertical velocity based on parcel theory estimate of updraft associated with convective available potential energy. “Shear” is magnitude of vector wind difference between surface winds and 6 km winds. Red dots are associated with F2 and stronger tornadoes, blue dots are associated with non-tornadic significant severe thunderstorms (hail of at least 5 cm diameter and/or winds of hurricane force), grey dots are non-severe thunderstorms. Data from Craven and Brooks (2004).

Figure 3. Time series of decadal values of an index (standardized to 1) of the number of 2-day precipitation totals exceeding a threshold for a 1 in 5-yr occurrence for 7 regions and the U.S. as a whole. This was based on an individual analysis of 930 long-term stations. Station time series of the annual number of events were gridded and then regional annual values were determined by averaging grid points within the region. Finally, the results were averaged over decadal periods.

Figure 4. Changes in observed twenty year return value of the daily accumulated precipitation from 1948 to 2010. Units: inches. Only locations for which data from at least 2/3 of the days in the 1948-2010 period were recorded are included in this analysis. The change in the return value at each station is shown by a circle whose relative size portrays its statistical significance: the large circles indicate the z-score (estimated change in the return value divided by its

standard error) is greater than two in magnitude, medium circles indicate the z-score is between one and two in magnitude, and the small circles indicate the z-score is less than one in magnitude.

Figure 5. Comparisons of tropical cyclone Power Dissipation Index (PDI; defined in Emanuel 2005) in the North Atlantic and western North Pacific. The red curves show the annual values derived from the best track data and the blue curves show annual values derived from the more homogeneous satellite-based intensity reconstructions. Thin lines show the raw values, thick lines show the smoothed time series, and least-squares linear trend lines calculated from the raw series are shown. The data are updated and adapted from Kossin et al. (2007).

Figure 6. Number of extreme snowstorms (upper 10 percentile) occurring each decade within the six U.S. climate regions in the eastern two-thirds of the contiguous U.S. (Based on an analysis of the 50 strongest storms for each of the six climate regions from October 1900-April 2010). The inset map shows the boundaries of each climate region. These regions were selected for consistency with NOAA’s monthly to annual operational climate monitoring activities. The map includes standardized temperature anomalies and precipitation departures from the $20^{th}$ century mean calculated across all snow seasons in which each storm occurred. The snow season is defined as December-March for the South and Southeast regions and November-April for the other four regions.

Figure 7. (a). Area weighted annual percentage of U.S. homogenous snowfall stations exceeding their own $90^{th}$ percentile seasonal totals, 1900-01 to 2010-11. Reference period is 1937-38 to 2006-07. Adapted from Kunkel et al. (2009c). Thick blue line: 11-year running mean of the

percentages. Dashed line: Number of grid cells with active stations each year. (b) as (a) but for the percentage of the contiguous U.S. snowfall data below the 10th percentile.

Figure 8. Authors' assessments of the adequacy of data and physical understanding to detect and attribute trends. Phenomena are put into one of three categories of knowledge from less to more. The dashed lines on the top and right sides denote that knowledge about phenomena in the top category is not complete.

Table 1. Nonparametric test for trend in extreme precipitation based on Kendall's $\tau$ for the number of occurrences of 2-day precipitation exceeding a threshold for a 1-in-5yr return period over the period of 1895-2010 and over the period of 1957-2010, as well as the total precipitation exceeding the 99 percentile for daily amounts over the period of 1957-2010.

| Region | Kendall's $\tau$ (2-dy,5-yr) 1895-2010 | Kendall's $\tau$ (2-dy,5-yr) 1957-2010 | Kendall's $\tau$ (99%ile) 1957-2010 |
|---|---|---|---|
| United States | 0.240*** | 0.388*** | 0.340*** |
| Northeast | 0.065 | 0.266*** | 0.360*** |
| Southeast | 0.242*** | 0.192** | 0.188** |
| Midwest | 0.206*** | 0.224** | 0.301*** |
| N. Great Plains | 0.032 | 0.146 | 0.085# |
| S. Great Plains | 0.097 | 0.053 | ------- |
| Northwest | -0.006 | 0.063 | 0.062 |
| Southwest | 0.012 | 0.121 | 0.048 |

*Significant at 0.10 level

**Significant at 0.05 level

***Significant at 0.01 level

#Results for combined Northern and Southern Great Plains

*Notes on Table 1*: Kendall's $\tau$ can be used to perform a nonparametric test for trend (Chapter 8, Hollander and Wolfe 1973). The statistic $\tau$ is a measure of association between the variable and time, ranging between −1 and 1 like an ordinary correlation coefficient. The *P*-value is based on the null hypothesis of no trend (i.e., the time series is uncorrelated with time). Positive values of $\tau$ indicate indices increasing with time, but not necessarily linearly. Kendall's $\tau$ is commonly used to test for trends in hydrologic time series (Chapter 8, Helsel and Hirsch 1993: Villarini et al. 2009).

Table 2. Differences between two periods (1990-2009 minus 1971-1989) for daily, 1-in-5yr extreme events and maximum precipitable water values measured in the spatial vicinity of the extreme event location and within 24 hours of the event time.

| Region | Extreme Precipitation Frequency index Difference (%) | Precipitable Water Difference (%) |
|---|---|---|
| Northeast | +55** | +2 |
| Southeast | +11* | +9*** |
| Midwest | +21** | +6** |
| North Great Plains | +18* | +16*** |
| South Great Plains | +15 | +8*** |
| Northwest | +36* | +4 |
| Southwest | +36* | -4 |

*Significant at 0.10 level

**Significant at 0.05 level

***Significant at 0.01 level

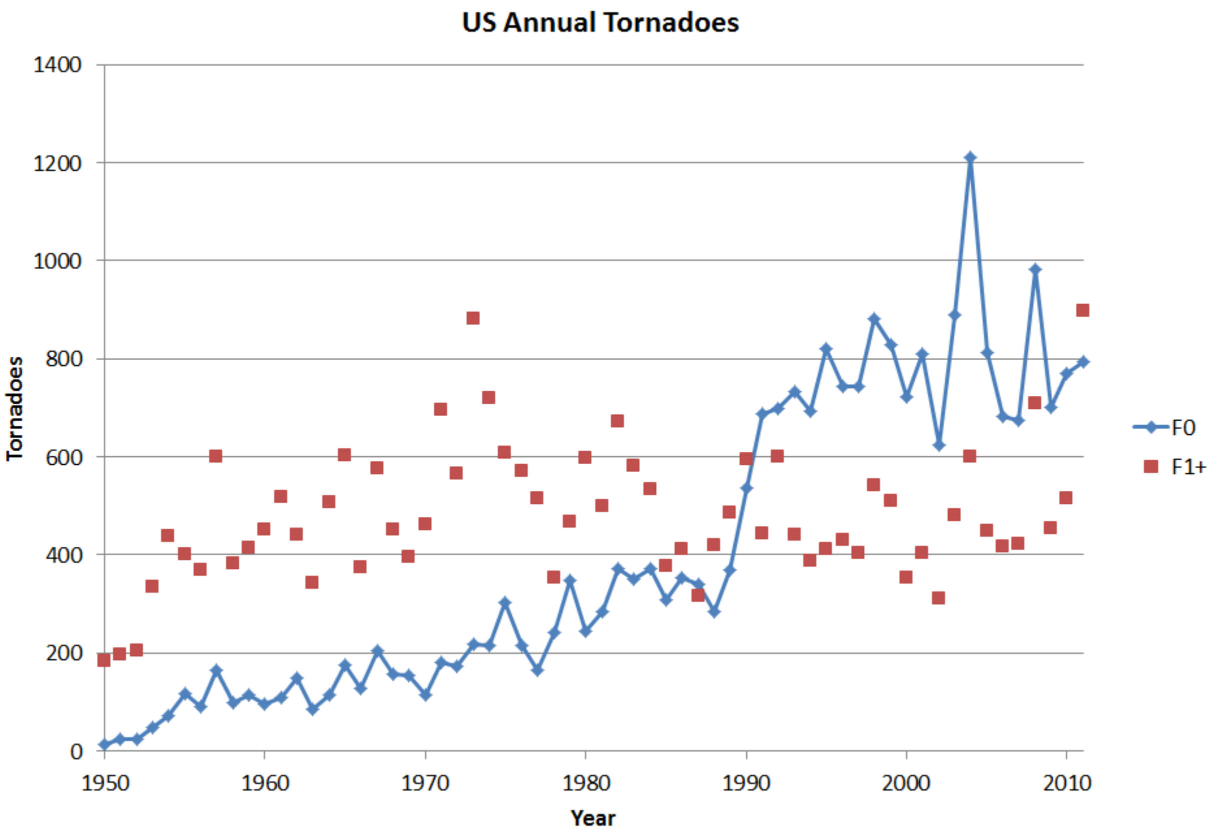



Figure 1. Reported tornadoes in NWS database from 1950-2011. Blue line is F0 tornadoes, red dots are F1 and stronger tornadoes.

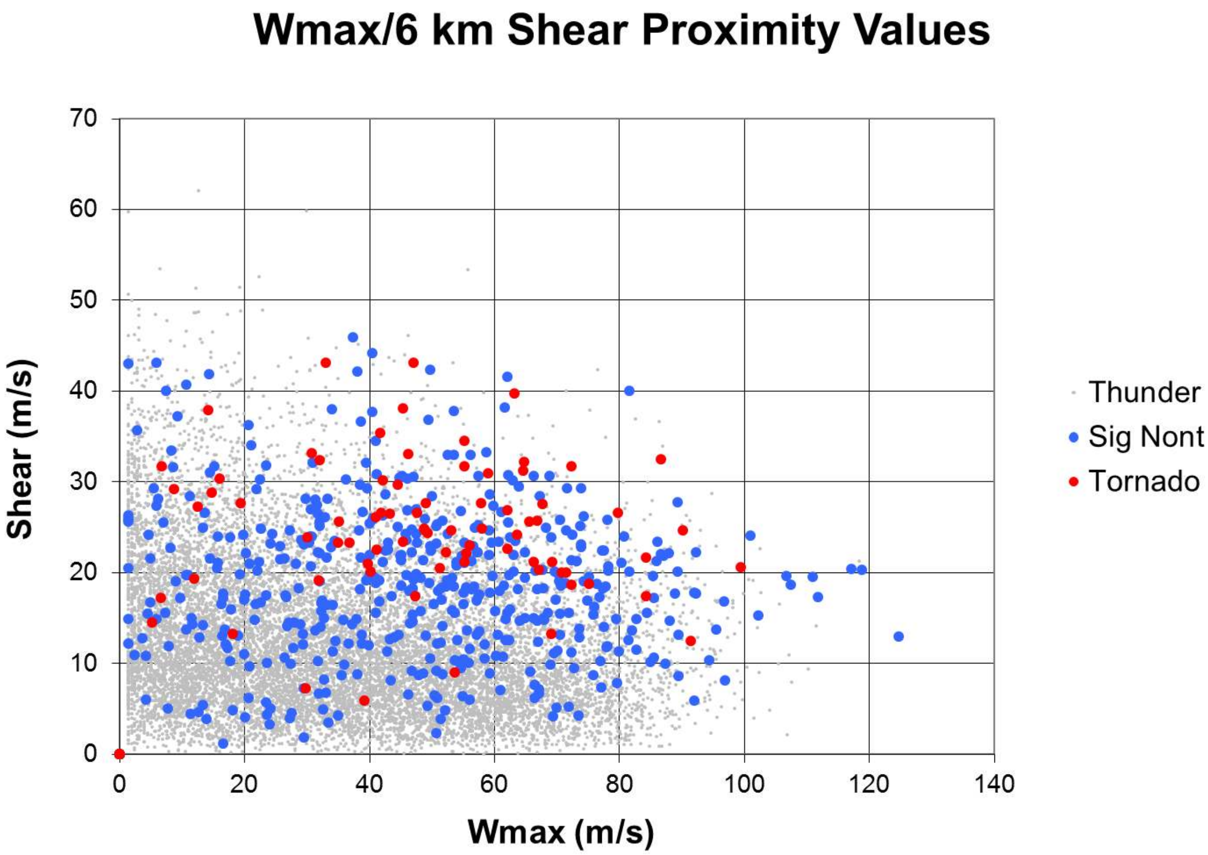



Figure 2. Convective parameters from 0000 UTC soundings in US 1997-9. "Wmax "is vertical velocity based on parcel theory estimate of updraft associated with convective available potential energy. "Shear" is magnitude of vector wind difference between surface winds and 6 km winds. Red dots are associated with F2 and stronger tornadoes, blue dots are associated with non-tornadic significant severe thunderstorms (hail of at least 5 cm diameter and/or winds of hurricane force), grey dots are non-severe thunderstorms. Data from Craven and Brooks (2004).

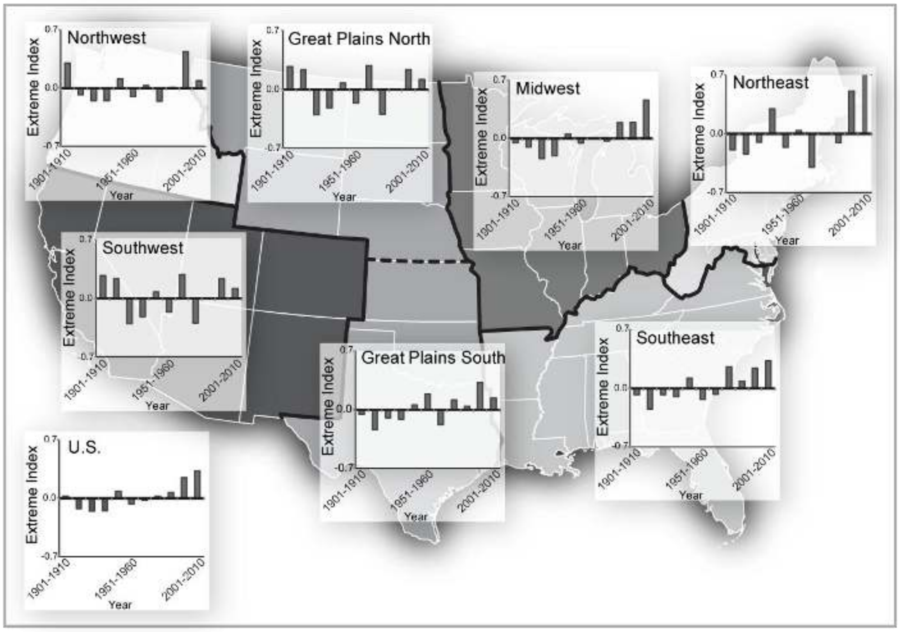



Figure 3. Time series of decadal values of an index (standardized to 1) of the number of 2-day precipitation totals exceeding a threshold for a 1 in 5-yr occurrence for 7 regions and the U.S. as a whole. This was based on an individual analysis of 930 long-term stations. Station time series of the annual number of events were gridded and then regional annual values were determined by averaging grid points within the region. Finally, the results were averaged over decadal periods.

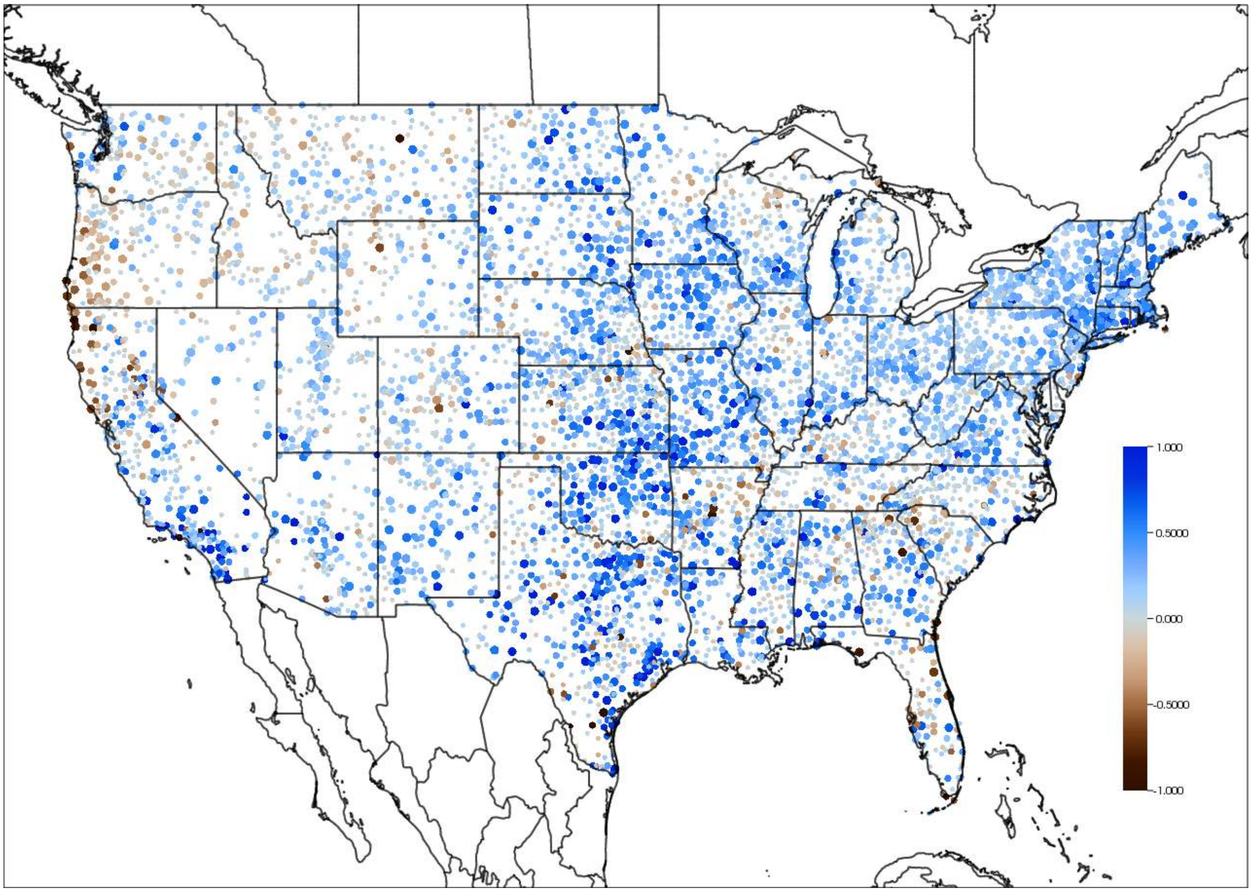



Figure 4. Changes in observed twenty year return value of the daily accumulated precipitation from 1948 to 2010. Units: inches. Only locations for which data from at least 2/3 of the days in the 1948-2010 period were recorded are included in this analysis. The change in the return value at each station is shown by a circle whose relative size portrays its statistical significance: the large circles indicate the z-score (estimated change in the return value divided by its standard error) is greater than two in magnitude, medium circles indicate the z-score is between one and two in magnitude, and the small circles indicate the z-score is less than one in magnitude.

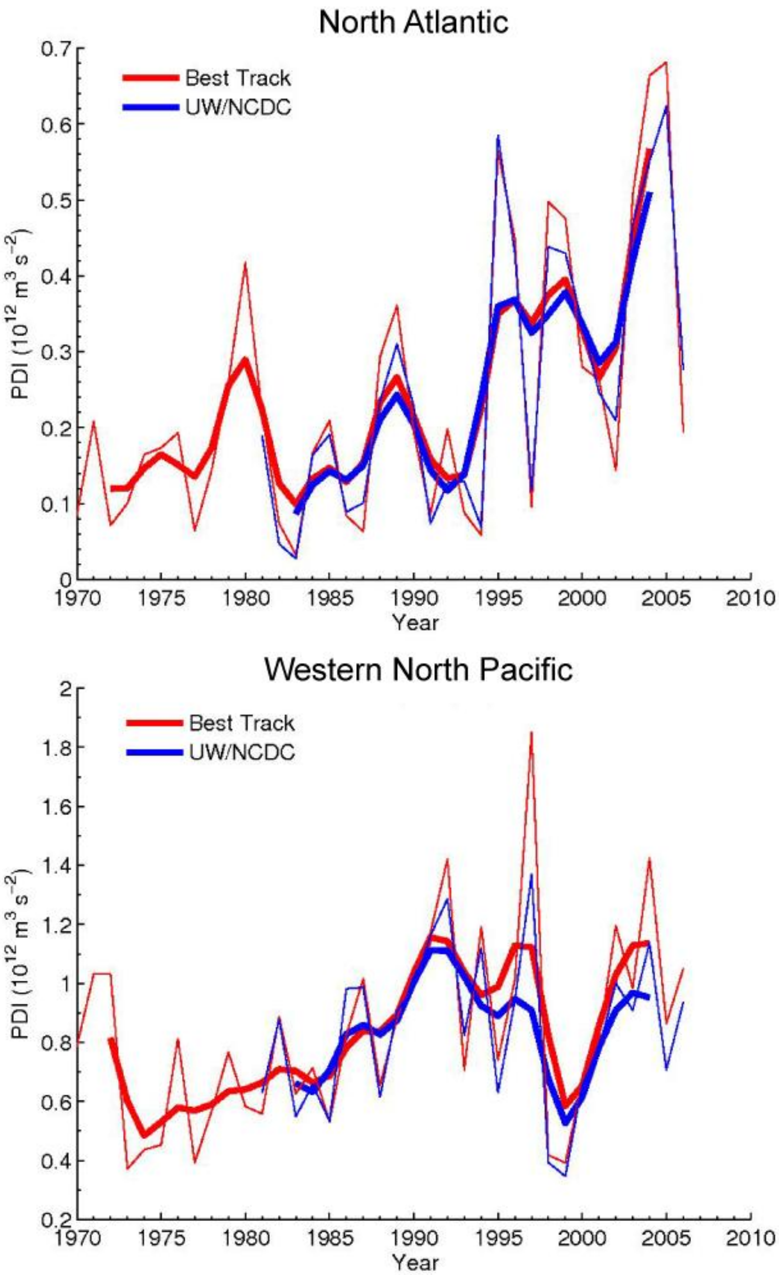



Figure 5. Comparisons of tropical cyclone Power Dissipation Index (PDI; defined in Emanuel 2005) in the North Atlantic and western North Pacific. The red curves show the annual values derived from the best track data and the blue curves show annual values derived from the more homogeneous satellite-based intensity reconstructions. Thin lines show the raw values, thick lines show the smoothed time series, and least-squares linear trend lines calculated from the raw series are shown. The data are updated and adapted from Kossin et al. (2007).

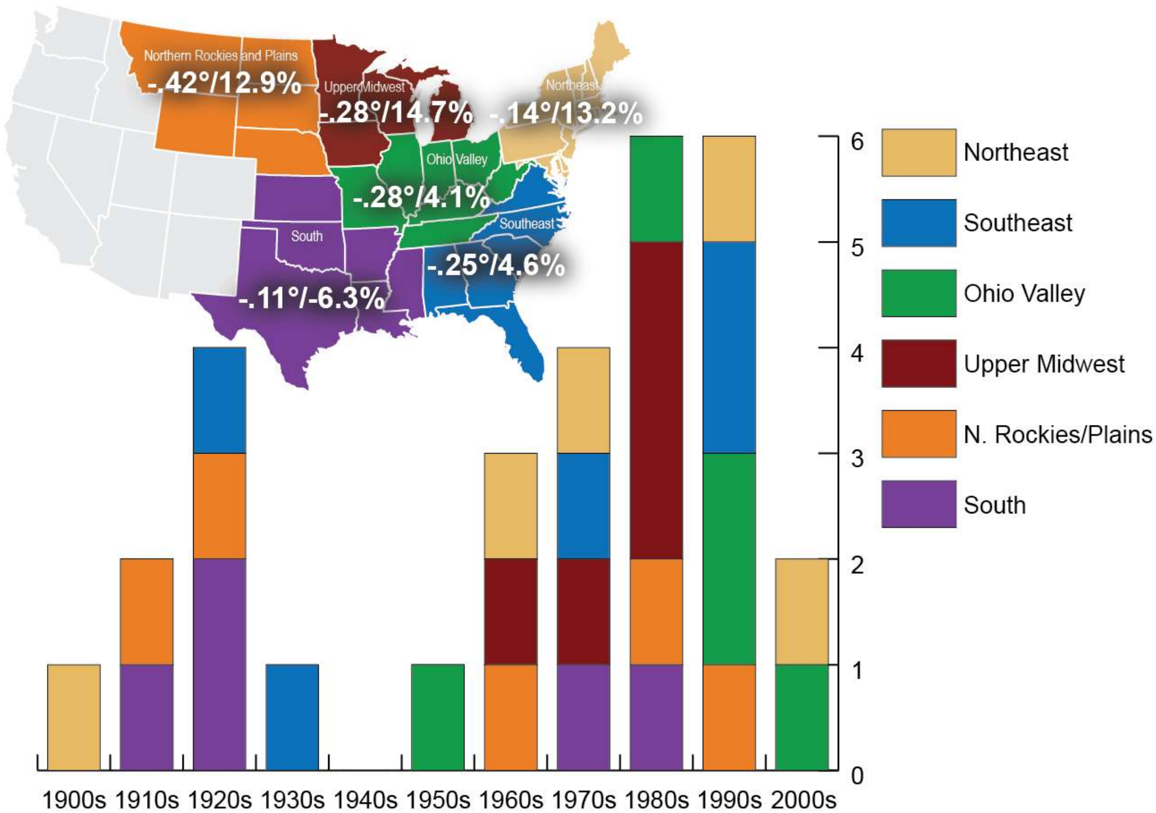
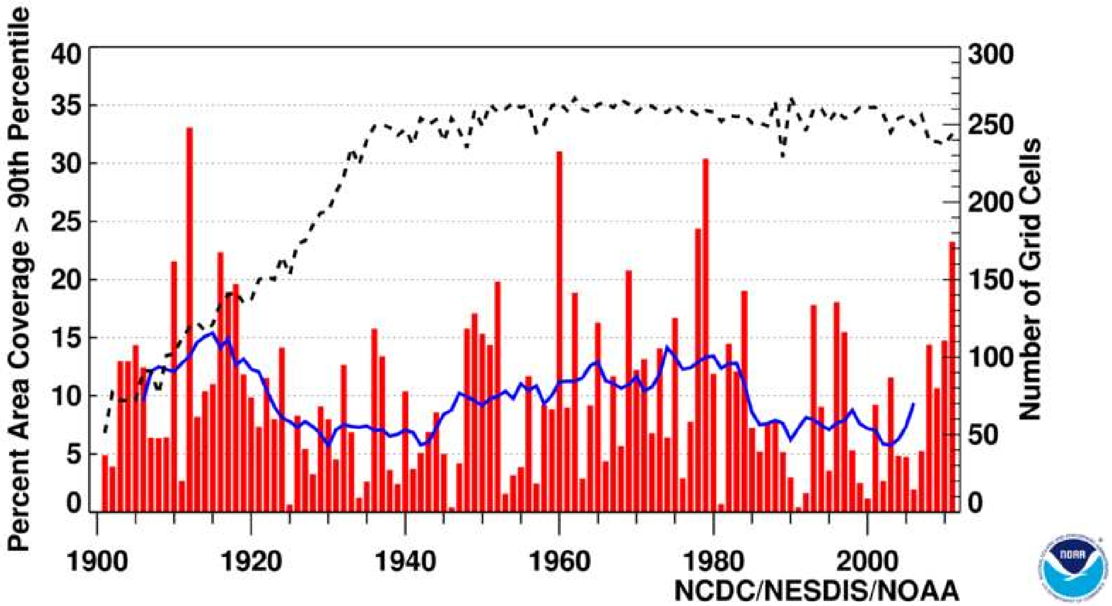



Figure 6. Number of extreme snowstorms (upper 10 percentile) occurring each decade within the six U.S. climate regions in the eastern two-thirds of the contiguous U.S. (Based on an analysis of the 50 strongest storms for each of the six climate regions from October 1900-April 2010). The inset map shows the boundaries of each climate region. These regions were selected for consistency with NOAA's monthly to annual operational climate monitoring activities. The map includes standardized temperature anomalies and precipitation departures from the 20th century mean calculated across all snow seasons in which each storm occurred. The snow season is defined as December-March for the South and Southeast regions and November-April for the other four regions.



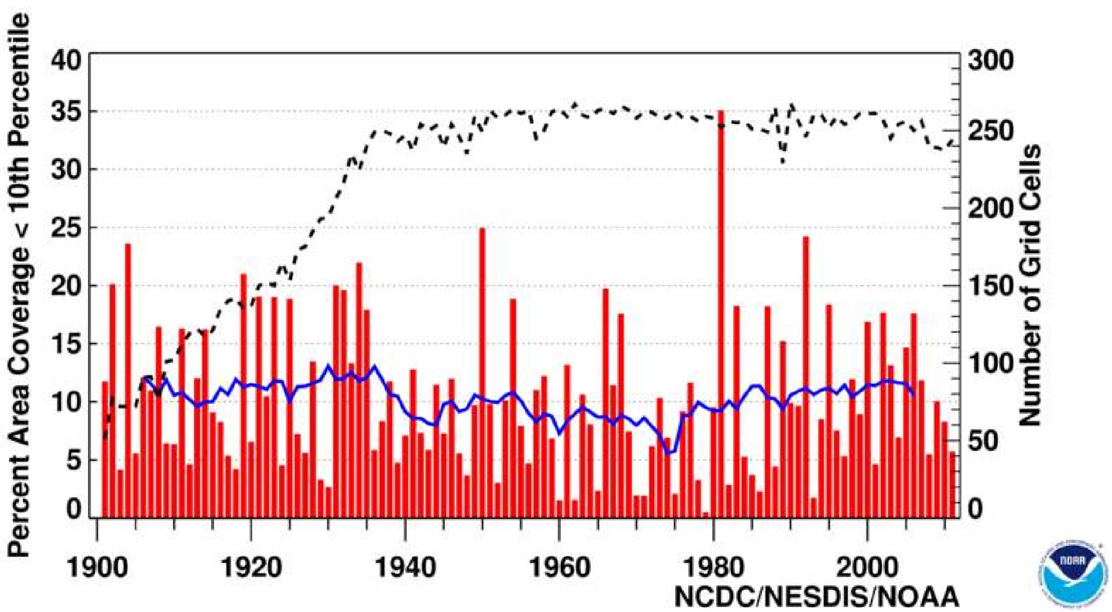




Figure 7. (a). Area weighted annual percentage of U.S. homogenous snowfall stations exceeding their own 90th percentile seasonal totals, 1900-01 to 2010-11. Reference period is 1937-38 to 2006-07. Adapted from Kunkel et al. (2009c). Thick blue line: 11-year running mean of the percentages. Dashed line: Number of grid cells with active stations each year. (b) as (a) but for the percentage of the contiguous U.S. snowfall data below the 10th percentile.

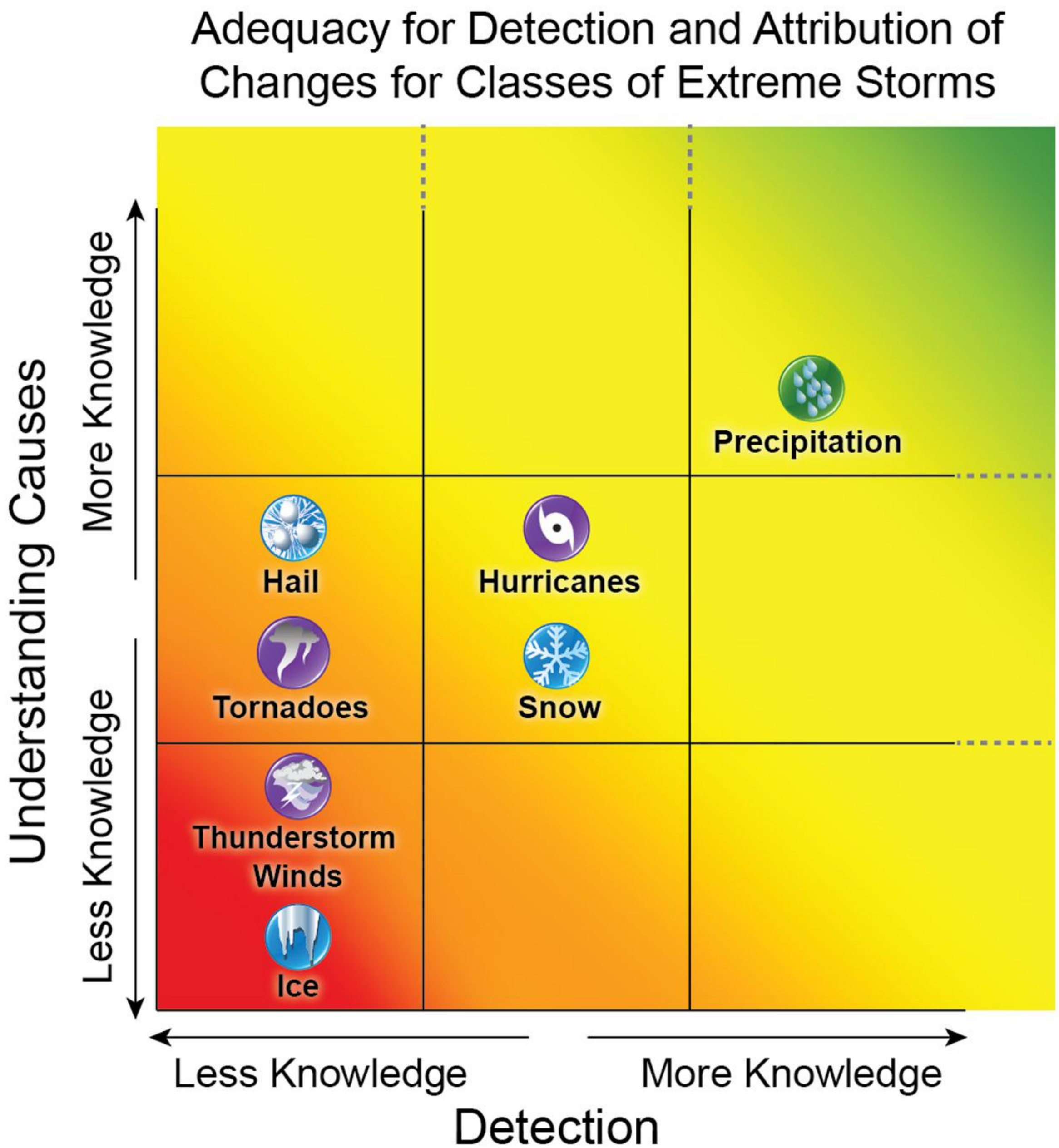



Figure 8. Authors' assessments of the adequacy of data and physical understanding to detect and attribute trends. Phenomena are put into one of three categories of knowledge from less to more. The dashed lines on the top and right sides denote that knowledge about phenomena in the top category is not complete.

## Supplementary Online Material

1. Time dependent peaks over threshold methodology

To produce figure 4, we used data for 1948-2010 from the Global Historical Climate Network-daily dataset for stations in the contiguous United States including only stations providing data for at least 2/3 of the days in that period. At each station, we found the station-specific 97$^{th}$ percentile of daily precipitation based on the entire time period, using only days with at least 1 mm of precipitation. We then fit a station-specific time-varying statistical extreme value model (Coles 2001) to daily exceedances of the 97$^{th}$ percentile. We used only the maximum daily value when consecutive days exceeded the threshold to avoid temporal dependence from multi-day storms (i.e., runs declustering with parameter $r = 1$, Coles 2001). We used a point process model for exceedances over a high threshold (or peaks over threshold), as in Tomassini and Jacob (2009) and Cooley and Sain (2010). The model is equivalent to a generalized Pareto distribution for excesses over a threshold combined with a Poisson process for the occurrence of threshold exceedances and is consistent with a generalized extreme value (GEV) distribution for block maxima. The basic parameters of the point process model can be expressed in terms of those of a GEV, namely location, scale, and shape. The shape parameter determines the heaviness of the tail of the distribution, encompassing the Weibull (bounded tail), Fréchet (heavy tail), and Gumbel (light tail) distributions. We allowed the location parameter to vary linearly in time, while assuming the shape and scale parameters were constant over time. To minimize complexity, any seasonality in these parameters was ignored. As a result of this parameterization, the change over time in the return level (for any return period) is linear with the same slope as that for the location parameter (Coles 2001). An additional consequence is that the change is not a function of the return period considered - that is the 1948-2010 change in the 20-year return level is the same as the 1948-2010 change in the X-year return level for any X. Note that by fitting a separate shape parameter value at each location, we allowed for the possibility that the heaviness of the tail differs by location. Uncertainty estimates were based on the Hessian of the point process likelihood according to standard maximum likelihood theory, with the standard error for the return level depending on not only the standard error for the linear trend parameter, but on the standard errors of the other parameters of the GEV as well. Standard diagnostics for extreme value distributions (Coles 2001) indicated no obvious lack of fit, and analysis with thresholds based on percentiles other than the 97$^{th}$ (90, 95, 98, 99, 99.5) indicated the results did not change substantially apart from the expected bias-variance tradeoff as the percentile increased. The station-specific results are noisy because of the uncertainty in estimating the behavior of extremes from short time series. Statistical approaches that smooth over the noise are feasible, but standard techniques have not been developed, so we show the station-specific results without smoothing. The results are not sensitive to the available data criterion. We repeated the analysis for stations with 90% and 95% available data. We found that the stations excluded by these criterion levels exhibited the same spatial patterns as the stations with more complete data.

To account for multiple testing, we carried out a field significance analysis. Each of 1000 simulations consisted of 63 years of synthetic data resampled with replacement from the 63 years of observations comprising 1948-2010. Each resampled year included all the data from all locations for that year, thereby preserving the spatial dependence and within-year temporal structure, but breaking the between-year dependence. This produced simulated datasets under the null hypothesis of no temporal trend across years. For each of the 1000 simulated datasets, we carried out the point process model analysis, calculating the field significance P-value based on the number of locations with z-score (change in return level divided by its standard error) exceeding 1, 1.64, and 1.96. In all three cases, none of the simulations had as high a proportion of stations with z-scores exceeding the value as the proportion of stations in the original analysis, giving $P < 0.001$.

## 2. Extreme Precipitation Water Vapor Analysis

A set of extreme precipitation events (daily, 1-in-5yr recurrence) used in Kunkel et al. (2011) was the basis for this analysis. For each station event, radiosonde data from the Integrated Global Radiosonde Archive were used to find the highest precipitable water value occurring within 3 degrees latitude and longitude and on the day before or the day of the event. This was assumed to be the best representation of the water vapor environment available to the precipitation-producing system.

For each of the NCA regions, we averaged these precipitable water values for two periods: 1971-1989 and 1990-2009. We also averaged the values of the extreme precipitation index. The statistical significance of the differences was tested using the two-sample t-test. These two periods were compared because they span a period of sizeable changes in extreme precipitation occurrences and the data from the Integrated Global Radionsonde Archive (Durre et al. 2006) are most complete after 1970.

## References

Cooley, D. and S.R. Sain. 2010. Spatial hierarchical modeling of precipitation extremes from a regional climate model. Journal of Agricultural, Biological, and Environmental Statistics 15: 381-402.

Coles, S. 2001. An Introduction to Statistical Modeling of Extreme Values . Springer, London.

Durre, I., R.S. Vose, and D.B. Wuertz. 2006: Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*, 19, 53-68.

Kunkel, K.E., D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2011: Meteorological causes of the secular variations in observed extreme precipitation events for the conterminous United States. *J. Hydrometeor*., in press.

Tomassini, L. and D. Jacob. 2009. Spatial analysis of trends in extreme precipitation events in high-resolution climate model results and observations for Germany. Journal of Geophysical Research 114, D12113.