

North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends

Richard Grotjahn¹ · Robert Black² · Ruby Leung³ · Michael F. Wehner⁴ ·
Mathew Barlow⁵ · Mike Bosilovich⁶ · Alexander Gershunov⁷ ·
William J. Gutowski Jr.⁸ · John R. Gyakum⁹ · Richard W. Katz¹⁰ ·
Yun-Young Lee¹ · Young-Kwon Lim¹¹ · Prabhat⁴

Received: 7 August 2014 / Accepted: 4 May 2015 / Published online: 22 May 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The objective of this paper is to review statistical methods, dynamics, modeling efforts, and trends related to temperature extremes, with a focus upon extreme events of short duration that affect parts of North America. These events are associated with large scale meteorological patterns (LSMPs). The statistics, dynamics, and modeling sections of this paper are written to be autonomous and so can be read separately. Methods to define extreme events statistics and to identify and connect LSMPs to extreme temperature events are presented. Recent advances in statistical techniques connect LSMPs to extreme temperatures through appropriately defined covariates that supplement more straightforward analyses. Various LSMPs, ranging from synoptic to planetary scale structures, are associated with extreme temperature events. Current knowledge about the synoptics and the dynamical mechanisms leading to the associated LSMPs is incomplete. Systematic studies

of: the physics of LSMP life cycles, comprehensive model assessment of LSMP-extreme temperature event linkages, and LSMP properties are needed. Generally, climate models capture observed properties of heat waves and cold air outbreaks with some fidelity. However they overestimate warm wave frequency and underestimate cold air outbreak frequency, and underestimate the collective influence of low-frequency modes on temperature extremes. Modeling studies have identified the impact of large-scale circulation anomalies and land–atmosphere interactions on changes in extreme temperatures. However, few studies have examined changes in LSMPs to more specifically understand the role of LSMPs on past and future extreme temperature changes. Even though LSMPs are resolvable by global and regional climate models, they are not necessarily well simulated. The paper concludes with unresolved issues and research questions.

✉ Richard Grotjahn
grotjahn@ucdavis.edu

¹ Atmospheric Science Program, Department of L.A.W.R., University of California Davis, One Shields Ave., Davis, CA 95616, USA

² School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 First Drive, Atlanta, GA 30332-0340, USA

³ Pacific Northwest National Laboratory, Richland, WA 99352, USA

⁴ Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

⁵ University of Massachusetts Lowell, Lowell, MA 01854, USA

⁶ NASA GSFC Global Modeling and Assimilation Office, Greenbelt, MD 20771, USA

⁷ Climate, Atmospheric Science and Physical Oceanography (CASPO) Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA

⁸ Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA

⁹ Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, QC H3A 0B9, Canada

¹⁰ Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, CO 80307, USA

¹¹ NASA Goddard Space Flight Center, Global Modeling and Assimilation Office, Goddard Earth Sciences Technology and Research/I.M. Systems Group, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

Keywords Large scale meteorological patterns for temperature extremes · Heat waves · Hot spells · Cold air outbreaks · Cold spells · Statistics of temperature extremes · Dynamics of heat waves · Dynamics of cold air outbreaks · Dynamical modeling of temperature extremes · Statistical modeling of extremes · Trends in temperature extremes

1 Introduction to temperature extremes

Temperature extremes have large societal and economic consequences. While many heat waves are short-lived, longer events can have a large economic cost. Cold air outbreaks (CAOs) tend to be short-lived but carry large economic losses. Timing of the CAOs can be more important than the minimum temperatures of the freeze; during 4–10 April 2007 low temperatures across the South caused \$2B in agricultural losses since many crops were in bloom or had frost sensitive buds or nascent fruit (Gu et al. 2008). The event also exemplifies how monthly means can be misleading: April 2007 average temperatures were near normal. In short, both hot spells (HSs) and CAOs have great societal importance and they are short-term events that do not necessarily appear in monthly mean data.

This report focuses on short-term (5-day or less) extreme temperature events occurring in some part of North America. Temperature extremes considered in this paper include both short-term hottest days (warm season) and CAOs (winter and spring) as these have the greatest impacts. Such events, both observed and simulated, have received considerable attention (including research papers, e.g. Meehl and Tebaldi 2004; and active websites: <http://www.esrl.noaa.gov/psd/ipcc/extremes/>, <http://www.ncdc.noaa.gov/extremes/cei/>, and <http://gmao.gsfc.nasa.gov/research/subseasonal/atlas/Extremes.html>). However, the emphasis here is on the less well understood context for the extreme events. Our primary context is the large-scale meteorological patterns (LSMPs) that accompany these extreme events.

Temperature extreme events are usually linked to large displacements of air masses that create a large amplitude wave pattern (here called an LSMP). LSMPs have a spatial scale bigger than mesoscale systems but smaller than the near-global scale of some modes of climate variability. The LSMP often has some portion that is superficially similar to a blocking ridge, so a blocking index can be an LSMP indicator (Sillmann et al. 2011). Extreme events have also been linked to circulation indices like the North Atlantic Oscillation (NAO) (Downton and Miller 1993; Cellitti et al. 2006; Brown et al. 2008; Kenyon and Hegerl 2008; Guirguis et al. 2011) the Madden–Julian Oscillation (MJO) (Moon et al. 2011) and El Niño/Southern Oscillation (ENSO) (Downton and Miller 1993; Higgins et al. 2002;

Carrera et al. 2004; Meehl et al. 2007; Goto-Maeda et al. 2008; Kenyon and Hegerl 2008; Alexander et al. 2009; Lim and Schubert 2011). However, LSMPs are likely distinct from climate modes for several reasons. First, named climate modes such as the NAO are common modes of variability, whereas the LSMP is presumably as rare as the associated extreme event. Second, climate modes occur on a longer time scale than the LSMPs for the short-term events focused upon in this article. However, it is possible that an extreme event might occur when a climate mode has transient extreme magnitude or is amplified in association with another low frequency phenomenon. Third, tested LSMP patterns are not that similar to climate modes. In correlating eight NOAA teleconnection patterns (<http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml>) and California LSMPs and in assessing the PNA contribution to last winter's extreme cold in eastern North America (neither shown here), we do not find notable contribution from such modes. Several studies identified the LSMPs associated with specific extreme hot events (Grotjahn and Faure 2008; Loikith and Broccoli 2012) and CAO events (Konrad 1996; Carrera et al. 2004; Grotjahn and Faure 2008; Loikith and Broccoli 2012). Parts of these LSMPs tend to be uniquely associated with the corresponding extreme weather and those parts have some predictability (Grotjahn 2011).

The LSMPs for extreme events are not fully understood for different parts of North America. Also, local processes: topography, soil moisture, etc. play key roles but there is a knowledge gap in how well climate models simulate the LSMPs as well as these local processes and how the local and global modes interact with LSMPs. Bridging these knowledge gaps will reduce the uncertainty of future projections and drive model improvements.

Now is an opportune time to summarize critical issues and key gaps in understanding temperature extremes variability and trends because: (1) it is not known if current climate models used for future projections are producing extremes via the correct dynamical mechanisms, which directly impacts confidence in projections, and (2) knowledge of the LSMPs can improve downscaling (statistical or dynamical) by focusing attention on large scale patterns that are fundamental to the occurrence of the extreme event. Conversely, global models that do not reproduce the magnitude or duration of extreme temperature events accurately may still capture the correct LSMPs and facilitate downscaling. Finally, there is now sufficient preliminary work and growing interest to make a summary valuable.

From the LSMP context, the objectives of this review are surveys of: relevant statistical tools for extreme value analysis (Sect. 2), synoptic and dynamical interactions between LSMPs and other scales from local to global (Sect. 3), model simulation issues (Sect. 4), trends in these

Table 1 Some temperature related ETCCDI indices (Sillmann et al. 2013a)

ETCCDI index name	Semi-formal definition	Plain English
TX90p	The percentage of days when the high temperature is greater than 90 % of those in reference period	Hot days
TX10p	The percentage of days when the high temperature is less than 10 % of those in reference period	Cold days
TN90p	The percentage of days when the low temperature is greater than 90 % of those in reference period	Hot nights
TN10p	The percentage of days when the low temperature is less than 10 % of those in reference period	Cold nights
TXx	Monthly or seasonal maximum of daily maximum temperature	Hottest day
TXn	Monthly or seasonal minimum of daily maximum temperature	Coldest day
TNx	Monthly or seasonal maximum of daily minimum temperature	Hottest night
TNn	Monthly or seasonal minimum of daily minimum temperature	Coldest night
HWDI	Heat Wave Duration Index	Length of a heat wave
CWDI	Cold Wave Duration Index	Length of a cold spell
FD	Days below freezing	Frost days

For a complete list and formal definitions, see http://etccdi.pacificclimate.org/list_27_indices.shtml

temperature extremes (Sect. 5), and various open questions (summary section). Different readers may be interested in different surveys. Accordingly, the sections are autonomous allowing a reader to skip to a particular section(s) of interest.

2 Extreme statistics and associated large scale meteorological patterns (LSMPs)

2.1 Definitions of extreme events

The expert team on climate change detection and indices (ETCCDI) under the auspices of the World Meteorological Organization's CLIVAR program provide a useful, but somewhat incomplete starting point to explore the relationship between extreme temperatures and LSMPs. The ETCCDI indices as well as software to calculate them are available at <http://etccdi.pacificclimate.org> and are described in Alexander et al. (2006). The ETCCDI temperature indices are summarized in Table 1 and are designed for detecting and attributing human effects on extreme weather and do not necessarily represent particularly rare events. However, extreme value statistical methodologies can be applied to quantify the behavior of the tails of the distribution of certain ETCCDI indices and reveal insight into truly rare events (Brown et al. 2008; Peterson et al. 2013). Another common statistical measure is the return value (or level). Under a changing climate, the return value can be interpreted as an extreme quantile of the temperature

distribution that varies over time (e.g., the 20-year return value may be interpreted as that value that has a 5 % chance of being exceeded in a particular year). Trends in these measures of extreme temperature are detectable at the global scale (Brown et al. 2008) and have been attributed to human emissions of greenhouse gases (Christidis et al. 2005). At local scales, increases and decreases are observed reflecting the significant amount of natural variability in extreme temperatures.

Figure 1 shows observed trends over North America from 1950 to 2007 in 20-year return values of the hottest/coldest days and hottest/coldest nights. The temperature of very cold nights (Fig. 1d) exhibits pronounced warming over the entire continent as does the temperature of very cold days (Fig. 1b). The pattern of changes in the temperature of very hot days (Fig. 1a) and very hot nights (Fig. 1c) follows that of average temperature change with strong cooling in the southeastern United States. This “warming hole” has been linked to sea surface temperature patterns in the equatorial Pacific (Meehl et al. 2007) and changes in anthropogenic aerosols in the eastern US (Leibensperger et al. 2012). Changes are more pronounced at the higher latitudes, except in Quebec and Newfoundland. This analysis extends the work of Peterson et al. (2013) and uses annual anomalies to define the extreme indices. Although illustrative of the statistical techniques, analyses of annualized measures of extreme temperature cannot identify associated LSMPs that develop and dissipate on a much shorter time scale. Furthermore, events of high and low impact are better separated in seasonal analyses.

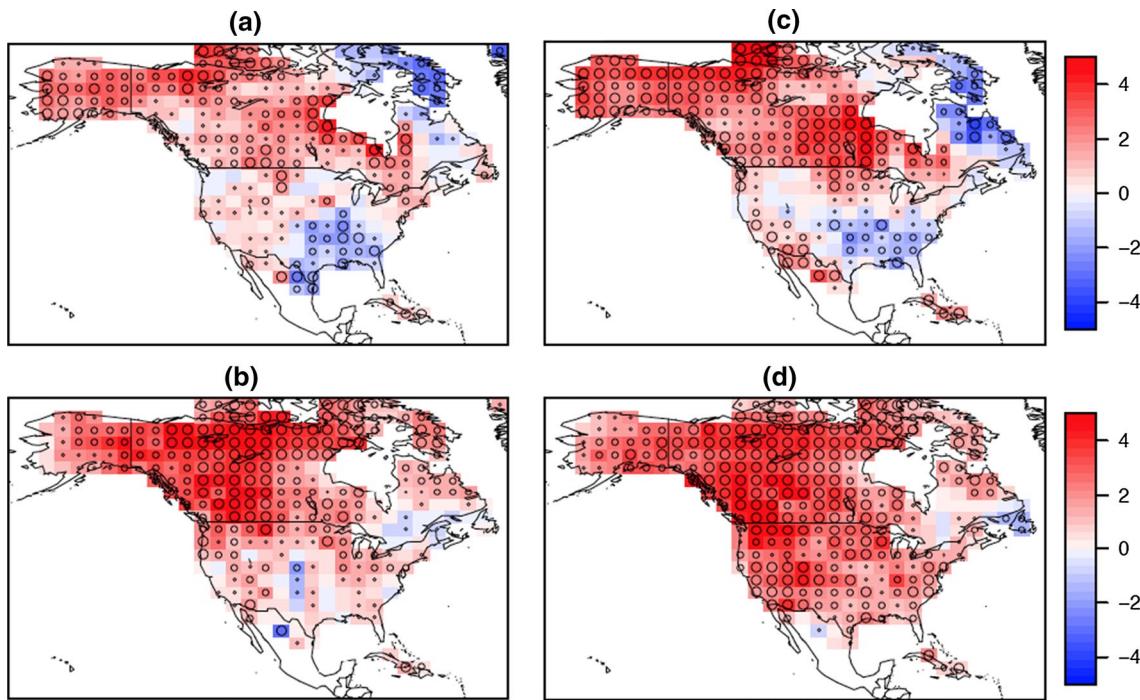


Fig. 1 Change over 1950–2007 in estimated 20-year annual return values ($^{\circ}\text{C}$) for **a** hot tail of daily maximum temperature (TXx), **b** cold tail of daily maximum temperature (TXn) **c** hot tail of daily minimum temperature (TNx) and **d** cold tail of daily minimum temperature (TNn). Results are based on fitting extreme value statistical models with a linear trend in the location parameter to exceedances of a location-specific threshold (greater than the 99th percentile for upper tail and less than the 1th percentile for lower tail). As this

analysis was based on anomalies with respect to average values for that time of year, hot minimum temperature values, for example, are just as likely to occur in winter as in summer. The *circles* indicate the z-score for the estimated change (estimate divided by its standard error), with absolute z-scores exceeding 1, 2, and 3 indicated by *open circles* of increasing size. Higher z-score indicates greater statistical significance

2.2 Application of extreme value statistical techniques

The observed changes in Fig. 1 are calculated using a time-dependent point process approach to fit “peaks over threshold” statistical models (Coles 2001). In this case, the extension of stationary extreme value methods to a time-dependent formalism used time as a “covariate” quantity to the ETCCDI indices (Kharin et al. 2013). A principal advantage of a fully time dependent formalism over a quasi-stationary approximation (Wehner 2004) is that the amount of data used to calculate extreme value parameters is substantially increased, resulting in higher quality fitted distributions and hence more accurate estimates of long period return values. Calculations involving climate model output gain additional statistical accuracy by using multiple realizations from ensembles of simulations, provided they are independent and identically distributed.

The “block maxima” and “peaks over threshold” methods to fit the tails of the distributions of random variables are asymptotic formalisms (Coles 2001). In this terminology, “block maxima” refers to use of only the maximum value during each “block” of time, usually a single season or year. The resulting generalized extreme value (GEV) or

Poisson and generalized Pareto distributions (GPDs) are both three parameter functions and can be transformed between each other. Hence provided that the data used to fit a distribution are in the “asymptotic regime”, i.e. far out in the tail of the distribution, the two methods are equivalent. Uncertainty in the estimate of long period temperature return values resulting from limited sample size can be appreciable and may be as large as that from unforced internal variability (Wehner 2010). However, variations in these estimates from multi-model datasets such as CMIP3/5 are generally significantly larger.

Confidence in the estimates of the statistical properties of the tail of the parent distribution of a random variable can be ascertained by exploring the sensitivity to the sample size used to fit the extreme value distribution. For block maxima methods, the length of the block is a season (or effectively so for temperature if the block length is a year). Lengthening the block makes the sample size smaller; shortening it makes the sample size larger since only one extreme value is drawn from each block. Such sensitivities are somewhat more straightforward to explore with peaks over threshold (POT) methods. Typical thresholds may be chosen between 80 and 99 % depending on

the size of the parent distribution that the extreme values are drawn from. However, standard POT methods may not discriminate between extreme values that occur at successive dates when the individual extreme values may not be truly independent. In these cases, declustering techniques (Coles 2001) are applied to avoid biased (low) estimates of the uncertainty. The trends in extreme temperature shown in Fig. 1 are calculated using such a declustering technique and a POT formalism with time as a linear covariate. An alternative approach retains possibly dependent consecutive extremes, adjusting the estimates of uncertainty through either resampling or more advanced techniques for quantifying extremal dependence (e.g., Fawcett and Walshaw 2012). Furthermore, LSMPs responsible for extreme events can be formed using only the dates of the onset of the event (e.g. Grotjahn and Faure 2008; Bumbaco et al. 2013) reducing the risk of autocorrelated extremes. In the cited studies, a 5 day gap was typically required between events. Low frequency factors (or climate modes), such as ENSO, are best treated using the previously mentioned covariate techniques.

Some of the advantages and challenges of applying statistical methods based on extreme value theory to analyze non-stationary climate extremes have been pointed out previously (e.g., Katz 2010), but are still not necessarily well appreciated by the climate science community. Conventional approaches tend to be either: (1) less informative (e.g., analyzing only the frequency of exceeding a high threshold, not the excess over the threshold and not measuring the intensity of the event); or (2) less realistic (e.g., based on assumed distributions such as the normal that may fit the overall data well, but not necessarily the tails). When relating extremes to LSMPs, standard regression approaches would not quantify the uncertainty in the relationship as realistically as using extremal distributions with covariates. Challenges in extreme value methods include specifying the dependence on LSMPs of the parameters of the extremal distributions in a manner consistent with our dynamical understanding. Moreover, heat waves and CAOs are relatively complex forms of extreme events, some of whose characteristics can be challenging to incorporate into the framework of extreme value statistics (Furrer et al. 2010). Finally, another advantage of the POT approach over the block maxima approach is being able to incorporate daily indices of LSMPs as covariates (not just monthly or seasonally aggregated indices).

2.3 Identification of LSMPs related to extreme temperatures

Several methods have been used to identify LSMPs that occur in association with extreme temperature events. These methods and their properties are summarized in

Table 2. The text summarizes each method, its advantages and its disadvantages.

2.3.1 Composites

Composite methods define the LSMP using a target ensemble average. The values at a grid point for a field on specified ‘target’ dates are averaged together. These dates might be when a temperature event begins (onset dates) defined as when some parameter(s) first meet some threshold and duration criteria. Table 3 illustrates various threshold and duration criteria that have been used to identify short-term extremely hot events. The different definitions yield somewhat different dates and results. A definition using a physiological hazard (Robinson 2001) might not be satisfied at night in the coastal and some inland regions of the West, even though the daytime temperature threshold is well exceeded. Lower thresholds (90th percentile) generate more events thus increasing the sample size that can improve the statistical fit, though the behavior of the highest 1 % may not be well fit. Similarly, the number of stations (or the size of the area over which those stations occur) impacts which dates are identified, even for regions that would seem to be meteorologically consistent; for example, which stations and how many are included from the Central Valley of California changes which dates exceed a threshold. Also, different definitions target different purposes: Grotjahn and Faure were interested in the LSMPs at the onset of the hottest events while Meehl and Tebaldi were interested in finding the longest duration events of some importance. The number of dates averaged equals the number of ensemble members.

Compositing has several advantages. One can track LSMP formation by compositing fields with respect to the onset time of each event. Meteorologically relevant full fields (or anomalies) are obtained and composite analyses are constructed to obtain information on the synoptic and dynamical time evolution. The method is non-parametric in that it does not make any assumption about the pattern or the event statistics. Unlike some other methods, criteria can be applied (typically a minimum waiting period between events) to ensure events are independent. Significance can be assessed using a bootstrap resampling procedure where the target ensemble value at a grid point is compared to the distribution of values at that grid point found from a large number of ‘random ensembles’ (each of which uses the same number but randomly-chosen dates). Values above (or below) a threshold of the random ensemble distribution imply significantly high (low) values at that grid point. For example, a target ensemble value equal or higher than the top 10 of 1000 random ensemble values at that point is significant at approximately the 99 % level. Figure 2 shows the LSMPs for California Central Valley cold air outbreaks

Table 2 A summary of various methodologies used to identify large scale meteorological patterns (LSMPs)

Method (example reference)	Approach	Attributes	Cautions	Significance testing
Composites (e.g. Grotjahn and Faure 2008)	Average together an ‘ensemble’ of maps on pre-identified dates As with some other procedures described here, it is important to assess how independent the dates are, e.g. consulting the autocorrelation	Ensemble average provides physical insights. Can be used to examine patterns leading up to and subsequent to the event onset by shifting dates used. Non-parametric in making no assumption about the pattern. The dates can be chosen to be independent	Must identify the event dates beforehand. Only finds one average unless individual members of ensemble examined and grouped when choosing which dates are in which average	Bootstrap resampling can compare ensemble average to many random ensemble averages (e.g. Grotjahn and Faure 2008). Need larger sample size or another method to assess consistency (e.g. sign counts; Grotjahn 2011)
Regression (e.g. Lau and Nath 2012)	Fit a regression line (polynomial in the predictand) to a time series of a predictor at each grid point	Pattern provides physical insights Parametric in assuming a specific regression line (polynomial). Can be used to examine patterns leading up to dates chosen event if event lasts for one time sample	Fit of regression line (polynomial) to extreme values may be notably altered by the choice of the polynomial assumed. Only finds one pattern. Does not incorporate time threshold criteria (e.g. event must last >2 days). Treats all dates as independent, which they might not be. Does not separate onset from other days during an event (e.g., mixing onset with dates during the event). However, low pass filtering sometimes used to aggregate the mixture of onset and during-event dates	Significance can be estimated by rejecting a null hypothesis about the coefficients, based on where the regression line values are higher (or lower) than a specified threshold of the predictand values
EOFs or PCs (Wu et al. 2012)	Calculate the eigenvectors of a space-weighted covariance (between values at different grid points) matrix. If only inputting extreme dates, then it is important to assess how independent the dates are, e.g. consulting the autocorrelation	Finds multiple patterns, each of which is orthogonal (not a subset) of another pattern. Can identify fraction of variability that is due to that EOF/PC. Most often used with filtered data to find leading low frequency structures	Not suitable when applied to all data as the leading EOFs will be common patterns not necessarily relevant to rare events. EOFs more useful when calculated only from data on event dates identified beforehand. Eigenvectors may not be sufficiently distinct. The patterns found may depend on domain chosen, though EOF ‘rotation’ may help. Each EOF explains only a fraction of the variance and so no single EOF might be an LSMP thereby limiting physical insight	No inherent significance test since the EOFs (or PCs) may not represent an actual weather pattern. The amount of variance associated with a particular EOF/PC is often used to indicate the importance of that EOF/PC

Table 2 continued

Method (example reference)	Approach	Attributes	Cautions	Significance testing
SOMs (e.g. Hewitson and Crane 2002)	Obtain the distribution of patterns covered by a set of input maps, using neural-net training	Resultant maps span the pattern space of input fields, represent nodes of a continuous pattern distribution, and can be related to a transformation of rotated EOFs back to physical space	Results can be sensitive to the domain used for pattern analysis. Care needed to ensure appropriate resolution of LSMP space. Low resolution can fail to resolve pattern features important for extreme events/Overly fine resolution undermines significance testing	Bootstrapping methods can compare frequency distributions of extreme-event days in LSMP space with random sampling of all days to indicate significance of event clustering in subregions of the full pattern space. (e.g. Cassano et al. 2007)
Clustering analysis (e.g. Stephanon et al. 2012)	Assigning events to K clusters so that it minimizes total sum of the generalized Euclidian distance between patterns in a cluster (Spath 1985; Seber 2008). If only inputting extreme dates, then it is important to assess how independent the dates are, e.g. consulting the autocorrelation	Objectively classify events and find relating LSMPs that provide physical insight No assumptions about the patterns such as orthogonality and symmetry	The number K of LSMP clusters can be somewhat arbitrary since K is pre-specified, but the 'dissimilarity index' can estimate the separation between clusters for different K and help optimize the value of K. Larger separation between clusters is preferred and generally declines as K increases. Optimal K can be unclear for some situations.	Significance of classification stability can be obtained from Monte Carlo test by rejecting the null hypothesis that the verification period temperature extreme cannot be classified in the clusters obtained from remaining periods. Lee and Grotjahn (2015) use bootstrap method (random ensembles same number of members as the cluster) to identify what parts of each cluster average are notable for the extreme event
Machine learning (Salakhutdinov and Hinton 2012)	Train a multi-layer neural network on a dataset. Training is done one layer at a time, with spatio-temporal patches at the bottom layer, and class labels being assigned at the top layer	Deep Belief Networks have proved to be powerful in capturing a range of patterns. It is likely that they will extract invariant patterns as well as anomalies	These techniques have produced state-of-the-art results in computer vision and speech recognition tasks. They have not yet been applied to climate datasets	Has not been conducted in the context of climate/LSMP applications. Classification performance of method has been conducted extensively for images by testing on held-out data

Table 3 Sampling of Various Criteria Used in Heat Wave and Hottest Day Definitions

Source	Definition
Robinson (2001)	A period of at least 48 h during which neither the overnight low nor the daytime heat index Hi falls below the NWS heat stress thresholds (80 and 105 °F). At stations where more than 1 % of both the high and low Hi observations exceed these thresholds, the 1 % values are used as the heat wave thresholds
Hajat et al. (2002)	3-day moving average temperatures >the 99th percentile of the whole record of temperature
Meehl and Tebaldi (2004)	The longest period of consecutive days satisfying the following 3 conditions 1. Daily maximum temperature $>T1$ for at least 3 days 2. Average daily maximum temperature $>T1$ for entire period 3. Daily maximum temperature $>T2$ for every day of entire period, where $T1$ (threshold 1) = 97.5th percentile of distribution of maximum temperatures in the observations and in simulated present day climate $T2$ = 81st percentile
Beniston (2004)	Maximum T exceeding the 90th quantile of summer temperature (30 °C) at a station (Basil, Switzerland)
Lipton et al. (2005)	Daily maximum high temperature remains 2 standard deviations above normal for at least 2 consecutive days
Gosling et al. (2007)	For 3 or more days the maximum T must be \geq 95th percentile of the maximum T in the summer climatology
Grotjahn and Faure (2008)	At least 3 consecutive days during which the daily maximum temperatures are above 100 °F (38 °C), and with at least one above 105 °F (40.5 °C)
Bachmann (2008)	Two combinations of criteria, I and II, were tested I. Must satisfy both conditions (where anomaly is relative to long term daily mean): 1. At least 3 consecutive days with daily anomaly maximum temperature ≥ 10 °C 2. At least 1 day must have maximum temperature anomaly ≥ 15 °C II. Must satisfy the 2 conditions above plus this additional condition: 3. The average maximum temperature for the event ≥ 100 F (38 °C)
Gershunov et al. (2009)	Individual stations exceeding the 99th percentile for 1, 2, or 3 dates in a row are aggregated, with the highest aggregation of values over the region including all of California and Nevada determining a ranking for an event. Daytime maximum and nighttime (highest) minimum treated separately
Lyon (2009)	Daily maximum temperature must exceed the 90th percentile for at least 3 consecutive days, where the percentile is based daily values from the 3-month summer season. Also tested, same temperature criterion over 5 consecutive days
Grotjahn (2011)	Daily maximum temperature anomaly (relative to long term daily mean) normalized by daily long term mean standard deviation at all three CV stations (KRBL, KFAT, KBFL) must all exceed 1.6 Note: this defines hottest days, not heat waves
Bumbaco et al. (2013)	Daily maximum temperature anomalies for stations in a region are averaged together. Heat wave when regional average daily anomaly exceeds 99th percentile for 3 or more consecutive days

and heat waves obtained by this method, including bootstrap resampling significance.

Compositing has some disadvantages that can be addressed. The identification of the extreme event dates must be done separately and before the compositing. The composite produces one target ensemble average (for a specific field and level) for each set of target dates. If more than one LSMP can produce the extreme event, then that must be identified either with one or more additional criteria when choosing dates or identified by

examining (perhaps qualitatively) the maps for each individual member of the target ensemble. A procedure like adding the number of positive and subtracting the number of negative anomaly values at a grid point in the target ensemble members (called the ‘sign count’ in Grotjahn 2011) can assist with identifying multiple patterns. (If the sign count equals the number of ensemble members, then all members have the same sign of the anomaly field at that location.) Lee and Grotjahn (2015) apply a cluster analysis to distinctly different parts of LSMPs

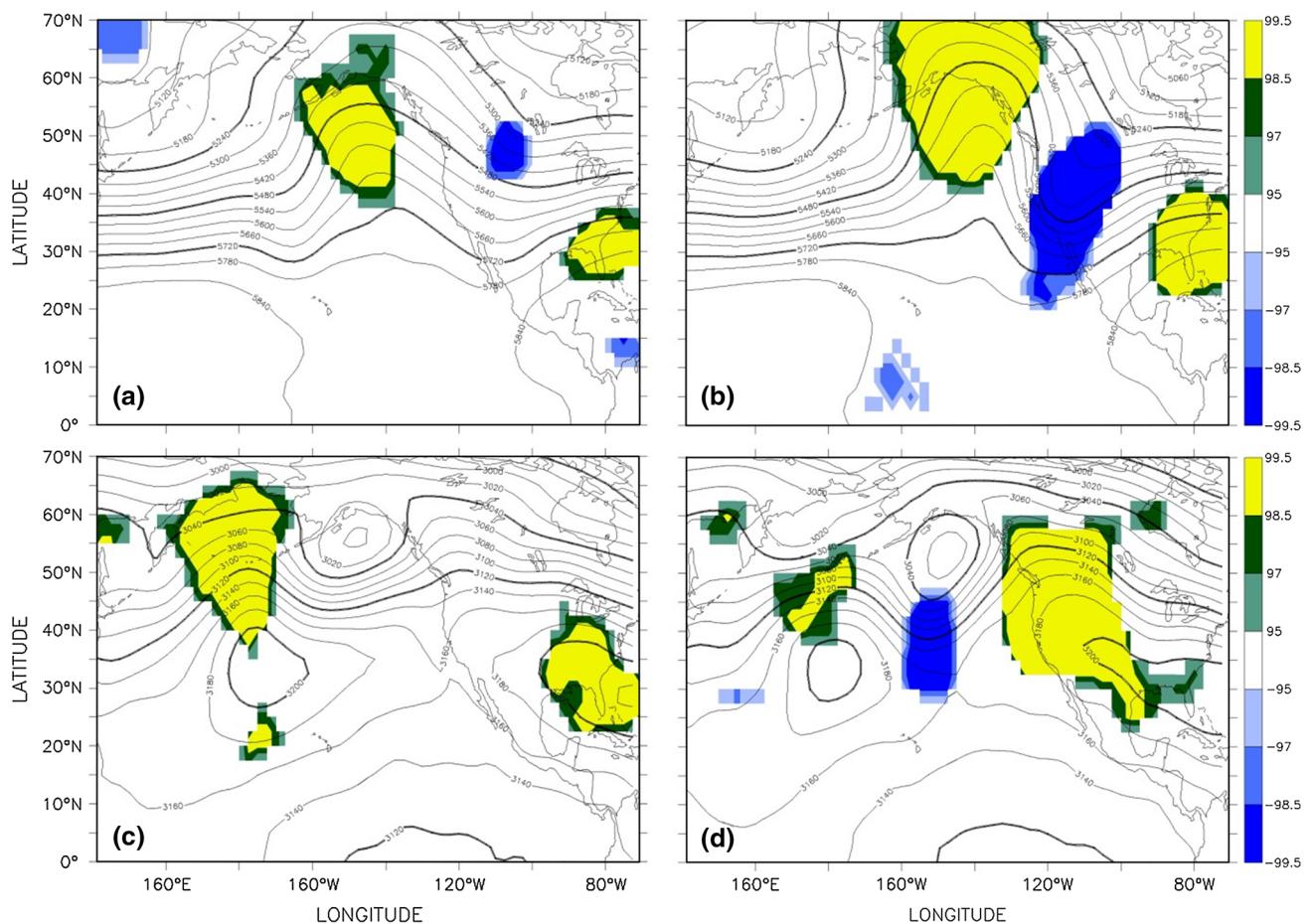


Fig. 2 Example large scale meteorological patterns (LSMPs) obtained as target ensemble mean composites for two types of California Central valley extreme events. Cold air outbreaks in winter (DJF) at **a** 72 h prior and **b** at onset of the events are shown in the 500 hPa geopotential height field. Heat waves during summer (JJAS),

c 36 h prior and **d** at the onset are shown in the 700 hPa geopotential height field. *Shading* indicates significance at the highest or lowest 5 % level, with the innermost shading significant at the 1.5 % level. Further discussion is in Grotjahn and Faure (2008)

prior to California heat waves and identify two ways the onset LSMPs form.

2.3.2 Regression

Regression estimates one quantity (the predictand) using a function of one or more other quantities (the predictors). The method is often parametric in assuming a specific function relates a predictor to the predictand (but nonparametric methods exist, too). An example predictor might be daily minimum surface temperature and the predictand might be 700 hPa level meridional wind. At each grid point the value of the predictand can be estimated using a polynomial function of the predictor, where the coefficients of that polynomial are calculated to minimize a squared difference between actual predictand values and polynomial values at that grid point. In general, the coefficients differ from grid point to grid point. To find the LSMP in this example of

extreme cold events, the polynomial can be used to construct the predictand at each grid point using a predictor value such as two standard deviations below normal.

The pattern obtained by regression can provide physical insights by directly linking the patterns in the predictand to extremes in the predictor. For example, lower tropospheric (700 hPa) winds prior to a California CAO flow from northern Alaska and northern western Canada to reach California without crossing over the Pacific Ocean. Regression can be used to examine patterns leading up to (or after) event onset dates by offsetting in time the predictor values from the predictand values when calculating the regression coefficients. The LSMP is again the resultant predictand when the predictor is at some specified value (e.g. predictor equals two standard deviations below normal). Significance can be estimated by rejecting a null hypothesis (e.g. that the regression coefficient is zero at the 1 % level using a student's *t* test).

One disadvantage of regression is that the assumption of a specific polynomial to represent how the predictand varies relative to the predictor. The fit of the regression line (polynomial) to extreme values may be notably altered by the order of the polynomial assumed. Regression, like composites, only finds one pattern. Regression does not incorporate event duration criteria (such as: the event must last at least 3 days). Regression treats all dates as independent, which they are not likely to be; however this can be somewhat mitigated by sub-sampling the data (e.g. only use every fifth day). Subsampling might be combined with low pass filtering to aggregate the mixture of onset and during-event dates. Another disadvantage is that a portion of the pattern may be highly significant but only have small correlation to the predictor.

2.3.3 Empirical orthogonal functions

Empirical orthogonal functions (EOFs) or principal components (PCs) can be used to identify LSMPs for extreme events. EOFs are the eigenfunctions of a matrix formed from the covariance between grid points on maps. EOFs from all maps in a time record will be ordered based on the eigenfunctions responsible for the largest amount of variance between time samples. Such eigenfunctions are the most common modes of variability and so not likely to be LSMPs of extreme events that are rare (except as mentioned in the introduction). However, EOFs can be formed only from maps selected in reference to an extreme event, such as maps only on the target dates of events onset. (EOFs of common low frequency modes may influence short-term extreme events focused on in this paper, as discussed in Sect. 3.1.) Weighting can be used for a variable grid spacing (such as occurs when using equal intervals of longitude across a range of latitudes).

An advantage of EOFs/PCs is the method finds multiple patterns, each of which is orthogonal (not a subset) of another pattern. The method can calculate the fraction of variability that is due to each EOF/PC. This approach is most often used with filtered data to find low frequency structures. This method is suitable for finding patterns leading up to and after the event by shifting the dates chosen by the event criteria (and only using those dates). In a study of California Central Valley hot spells, Grotjahn (2011) found the leading EOF based on dates satisfying the criteria in Table 3 to be very similar to the corresponding ensemble mean composite.

A disadvantage is that the patterns found may depend on the domain chosen, though EOF ‘rotation’ may help. Also, each EOF may explain only a small fraction of the variance and no single EOF might be an LSMP thereby limiting physical insight. Different leading EOFs/PCs might have structures influenced by the required orthogonality and

possibly not a pattern that occurs during an event. While the amount of variance associated with a given EOF is used to indicate the importance of the EOF, there is no inherent statistical significance test. Hence it can be unclear what portions of the EOF are significantly associated with the extreme event and what parts are not (and happen to reflect limited variation in the finite sample).

2.3.4 Clustering analysis

Clustering analysis is terminology indicating a widely used partitioning procedure that identifies separate groups of objects having common structural elements. Clustering analysis has been used to classify distinct sets of LSMPs associated with extreme events (Park et al. 2011; Stephanon et al. 2012). When we have 100 historical hot spell events over a given region, not all extreme events may have the same LSMP on or prior to their onset. Some events may have a wave-like height field, others may have a dipole pattern, and still others may have a third pattern. Although detailed grouping procedures vary for every clustering technique, the basic concept is to minimize the overall distance between patterns among events in resultant groups. For example, the k-means clustering technique applies an iterative algorithm in which events are moved from one group to another until there is no additional improvement in minimizing the squared Euclidean point-to-centroid distance in a group (Spath 1985; Seber 2008), where each centroid is the mean of the patterns in its cluster.

Output of clustering analysis is just the average field of events in each cluster, similar to the output from composite analysis. Unlike composite analysis in which members of clusters are pre-identified, the essential point of clustering analysis is to objectively classify events based on spatial pattern similarity. By applying cluster analysis to group similar onset patterns, one can isolate distinct dynamical origins of different extreme temperature events. Another advantage is that resultant clusters are based on physical maps without assumptions of orthogonality and symmetry such as in the mode separation by EOF/REOF. The robustness of a hot spell classification can be tested by a Monte Carlo test as follows. First one calculates a stability score that is the ratio of the number of verification period heat waves that are correctly attributed over the total number heat waves in the cluster. Second one estimates the probability density function (PDeF) of the null hypothesis that cluster assignment is purely random. Significance of each cluster can be estimated by rejecting a null hypothesis (e.g. stability score is located within highest 99 % of PDeF).

A disadvantage of clustering analysis is that one pre-specifies the number of clusters (e.g. k in k-means clustering). Determining the number k is subjective if one does not have sufficient prior knowledge of related physical

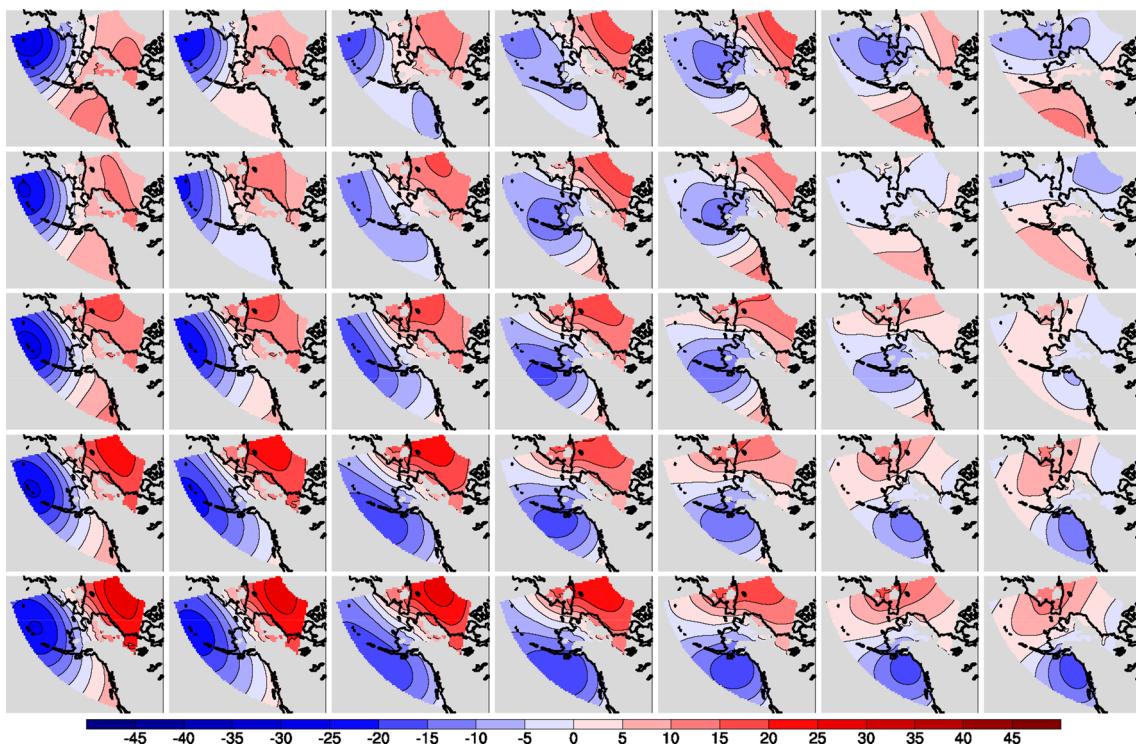


Fig. 3 Self-organizing map of synoptic weather patterns in a region focused on Alaska. The SOM array maps give the departure (in hPa) of sea level pressure (SLP) from the domain averaged sea level pressure. The SOM used daily December–January–February (DJF) SLP

for 1997–2007 from ERA-interim reanalyses and output from a regional climate model. Locations with elevation exceeding 500 m are not included in the maps to avoid using SLP in regions strongly influenced by methods used to extrapolate SLP from surface pressure

patterns. There are several statistics to check the optimal number of k such as ‘distance of dissimilarity’ (Stephanon et al. 2012). Another disadvantage is the ambiguity of cluster assignment for certain events. Another disadvantage is the patterns of such events cannot be assigned clearly to one group over another group. Part of a pattern may resemble Cluster #1, while another part may be more similar to Cluster #2. To avoid ascribing some marginal events to specific clusters, probabilistic clustering methods (e.g. Smyth et al. 1999) are developed, which suggests the possibility (e.g. percentage) that an event could be assigned to each cluster rather than assigning it only to one cluster. If one increases the cluster number, one can expect to decrease such ambiguity in classification. However, the point of clustering analysis is to give a physical insight with minimal groups and not to interpret every single episode.

2.3.5 Self-organizing maps

Self-organizing maps (SOMs; Kohonen 1995) are two-dimensional arrays of maps that display characteristic behavior patterns of a field (e.g., Cavazos 2000; Hewitson and Crane 2006; Gutowski et al. 2004; Cassano et al. 2007). The SOM array is a discretization of the continuous

pattern space occupied by the field examined. Thus, in contrast to clustering analysis, SOMs do not assume a clumping together of patterns, though such behavior can emerge if present in the input data. Figure 3 gives an example of a SOM array of synoptic weather patterns in sea level pressure over a region centered on Alaska. Individual maps in the array represent nodes in a projection of this continuous space onto a two-dimensional surface, with the size of the array determined by the degree of spatial discretization of the SOM space one “feels” is needed for the analysis at hand. The two dimensions show the two primary pattern transitions for the field examined. Although one could, in principle, use more than two dimensions, typical practice in climatological work has used only two.

The input maps themselves determine the degree and types of pattern transitions, hence the “self-organizing” nature of the resulting array. The SOM node array is trained on a sequence of input maps through an artificial neural net technique. The SOM array does not necessarily favor the largest scales in the input data, but rather the scales most relevant to the field for the domain and resolution examined. Consequently, SOMs can extract nonlinear pattern changes in fields, such as shifts in strong gradients. In addition, the pattern at each node is essentially

a composite of input maps with similar spatial distribution for the field examined, so that patterns in the SOM array show archetypal patterns of the field examined and directly lend themselves to physical interpretation. Typically, the SOM array displays features having the highest temporal variance in the input data. From this perspective, the SOM array is roughly akin to a transformation of a rotated EOF from spectral space back to physical space.

An advantage of SOMs is that one can identify the nodes where extreme events occur frequently and thus the physical behavior yielding extremes. For example, extreme events may tend to cluster in a small portion of SOM space, thereby allowing identification of LSMPs yielding extreme events. A further advantage is that if more than one group emerges in the SOM-space frequency distribution, then the grouping provides a SOM-determined segregation of different types of extreme events. One can then focus analysis and composites of additional fields (e.g., precipitation, winds, temperature) on only events of the same type. For example, Cassano et al. (2006) used SOMs of sea-level pressure patterns over Alaska to determine which synoptic weather patterns were responsible for extreme wind and temperature events at Barrow, Alaska. They then found robust links between these large-scale synoptic weather patterns and local weather features (precipitation, winds, and temperature). One can construct estimates of the significance of differences in frequency distributions in SOM space through bootstrapping procedures to estimate the likelihood that frequency distributions are not simply the result of random, finite sampling of the pattern space. Thus one can compare frequency distributions between a present and projected climate to assess potential climate changes in LSMPs, or between observational and model climates to assess similarity of observed and simulated LSMPs yielding extremes.

A disadvantage of SOMs is that the array size is predetermined by the user, and there is no clear, objective guideline for selecting array size. There are, however, some factors that can affect the array-size choice. The issue of significance limits the degree of discretization (number of nodes) one applies to the SOM space. Fine discretization will allow apparent detection of small differences in how different data sets occupy pattern space, but fine discretization will also render very noisy frequency distribution functions of the fields in SOM space, thus undermining detection of any significant differences. Coarse discretization limits the ability of the SOM procedure to resolve features producing the extreme events, so a further disadvantage is that an insufficient array size may obscure grouping that may be present in the data. The training method also requires specification of parameters that govern the training process. A well-trained SOM is insensitive to these choices, but care is needed to ensure such a result. In addition, like

some of the other methods described here, the extreme events are defined separately from the SOM analysis.

2.3.6 Machine learning and other advanced techniques

Looking to the future, we note that substantial progress has been made in the field of machine learning for extracting patterns from Big Data. Commercial organizations such as Google and Facebook rely on sophisticated, scalable analytics techniques for mining web-scale datasets. Both supervised and unsupervised machine learning tools could play an important role in extracting spatio-temporal patterns from climate datasets. The technique Deep Belief Networks (Salakhutdinov and Hinton 2012) has been applied with tremendous success to classifying objects in digital images (Krizhevsky et al. 2012) and speech recognition (Hinton et al. 2012). These methods have substantially outperformed existing techniques in the field with the same underlying learning algorithm. While these techniques have not yet been adapted for a multivariate spatio-temporal dataset (such as in climate), research efforts are currently underway to evaluate the performance of such methods in extracting patterns as well as anomalies from datasets. It is too early to discern pros and cons fully for such methods.

2.4 Including large scale patterns in extreme statistics

Application of covariates in extreme value methods (termed “conditional extreme value analysis”) is relatively new to the climate science community, although it has been available to the larger statistics community for some time (Coles 2001). The basic idea of conditional extreme value analysis is to allow the extremal distribution to be dynamic; that is, shifting depending on the observed value of an index of a climate mode or LSMP (the index would be an example of a “covariate”). The book by Coles includes an example in which annual maximum sea level is related to a climate mode, the Southern Oscillation. In their study of changes in extreme daily temperatures, Brown et al. (2008) used the NAO as a covariate in addition to a trend component.

Such techniques have proven useful in connecting extreme temperatures to LSMPs. Sillmann and Croci-Maspoli (2009) and Sillmann et al. (2011) used a blocking index as a covariate for extremely cold European winter temperatures and found that extreme value distributions (based on block minima) were better fit and long period return values were somewhat colder. Furthermore, they concluded that projected future extremely cold events in Europe were less influenced by atmospheric blocking because of projected shifts in North Atlantic blocking patterns. Photiadou et al. (2014) used a similar technique (but based on the POT approach rather than block maxima) to connect blocking and other indices to European

high temperature events finding that while El Niño/Southern Oscillation (ENSO) does not exert much influence on extremely high temperature magnitudes or duration, the North Atlantic Oscillation (NAO) and atmospheric blocking do. However to date, such covariate techniques relating atmospheric blocking to extreme temperatures have not been applied to North America. Many physically based covariate quantities potentially offer insight into the mechanisms behind extreme temperature events and the response to changes in the average climate. Good North American candidates for covariates include indices measuring modes of natural variability such as those describing the ENSO, the Pacific Decadal Oscillation (PDO), the NAO, the North Atlantic Subtropical High (NASH), and various blocking indices.

The ETCCDI indices were designed for climate change detection and attribution purposes rather than for exploring the mechanisms causing extreme events. They are not ideal for connecting extreme temperature events to LSMPs and they are not descriptive of particularly rare events. However, ETCCDI indices are designed to be robust over the observational record and have been calculated and described for CMIP5 models by Sillmann et al. (2013a) (see Table 1). ETCCDI indices are intended to be applied globally and be meaningful in areas of sparse observations. The relatively dense network of North American observations since the beginning of the twentieth century permits the construction of more specialized LSMP extreme indices linked to specific extreme events. Grotjahn (2011) defines an index that is an unnormalized projection of key parts of a target ensemble LSMP onto a daily map of the corresponding variable. (He combined such projections onto 850 hPa temperature and 700 hPa meridional wind to form his ‘circulation index’.) His target ensemble members are from dates satisfying criteria listed in Table 3. The key parts of the fields used are those where all the extreme events in the training period were consistent, at least in having the same anomaly sign. Grotjahn (2011) found that extreme values of such an index (based on upper air data) occurred on many of the same dates as extreme values of surface stations in the California Central Valley (CCV). Statistically significant relationships exist between extreme values of this circulation index and both the rate of CCV daily maximum temperature exceeding a high threshold and the distribution of the excess over the threshold (Katz and Grotjahn 2014). Grotjahn (2013) used such an index to show that a particular climate model was notably under-predicting the occurrence frequency (by half) of CCV hot spells. Grotjahn (2015) used such an index to show how that same model compared with a 55 year historical record and what the model implied for CCV hot spells during the last half of the twentyfirst century under two representative concentration pathways (RCPs) of greenhouse gases.

3 Large scale meteorological patterns related to extreme temperature events

Intraseasonal extreme temperature events (ETEs) are almost always associated with regional air mass excursions induced by circulation anomalies that are part of large-scale meteorological patterns (LSMPs). LSMPs can include synoptic features (e.g., midlatitude cyclones; Konrad 1996) that enhance the ETE and often with scales similar to a teleconnection pattern (though the nodes may not align and ETE onset, itself, may impact the teleconnection pattern, Cellitti et al. 2006). In some cases, the LSMP is interpreted as a juxtaposition of teleconnection patterns that leads to ETE events (Lim and Kim 2013). Heuristically, the role of LSMPs in producing ETEs could be considered the result of either (1) a direct contribution to the large-scale circulation that facilitates the air mass excursion or (2) the indirect modulation of sub-scale variability, such as regional modulation of storm track behavior by blocking patterns. Besides such dynamically driven impacts, there exist possible local impacts related to the interaction of the LSMP with local topography or coastline features, leading to possible local symmetries in the response pattern (e.g., Loikith and Broccoli 2012). Current knowledge of the remote forcing, dynamics and local forcing of LSMPs associated with ETEs is summarized next.

3.1 Remote forcing of LSMPs and ETEs

3.1.1 Connection to low frequency modes of climate variability

Numerous observational studies have ascertained that ETE behavior is modulated by recurring large scale teleconnection patterns, particularly during winter. On intraseasonal time scales there is a substantial modulation of North American ETEs during winter by the Pacific-North American (PNA) pattern, North Atlantic (or Arctic) Oscillation (NAO or AO) and blocking patterns (Walsh et al. 2001; Cellitti et al. 2006; Guirgis et al. 2011). On interannual and longer time scales additional climate modes such as El Niño-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) are also implicated (Westby et al. 2013). General relationships that have emerged from these statistical analyses are illustrated in Fig. 4: The positive (negative) phase of the NAO favors the occurrence of warm (cold) events over the eastern (southeastern) United States. The positive (negative) phase of the PNA tends to favor cold events over the southeastern (northwestern) US. These connections to climate modes are neither unique nor independent. For example, the regional influence of the PNA pattern on ETEs largely mirrors that of both the PDO and ENSO (Fig. 4) since the midlatitude atmospheric signatures of

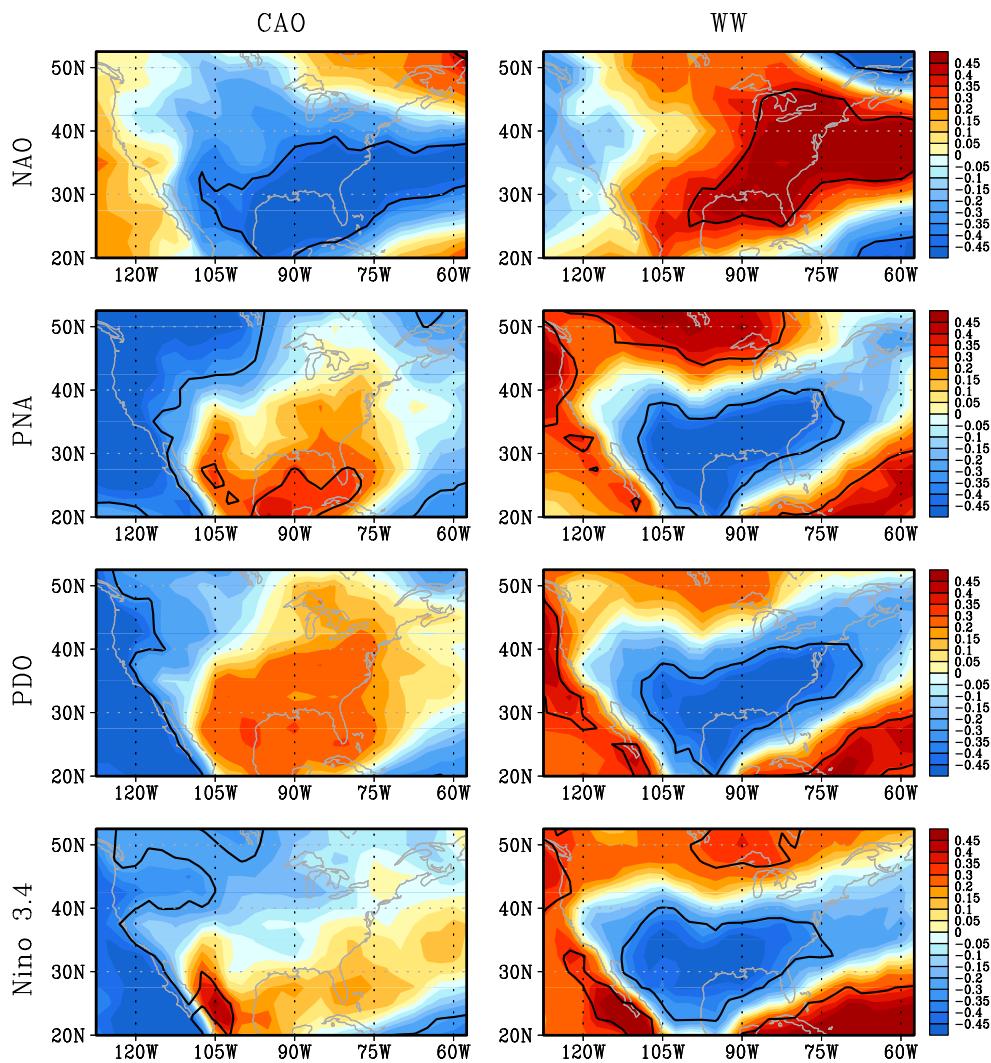


Fig. 4 Correlation between the local seasonal impact of cold days (left column) and warm days (right column) and the seasonal mean NAO (first row), PNA (second row), PDO (third row) and Niño 3.4

indices (fourth row) during winter, 1950–2011. The black contours encompass regions having correlations statistically significant at the 95 % confidence level (figure from Westby et al. 2013)

both ENSO and the PDO project on the PNA pattern. Also, the prevalence of atmospheric blocking patterns is intrinsically linked to particular climate mode phases (Renwick and Wallace 1996).

There have been pronounced episodes of climate modes influencing ETEs during recent winters. Cold extremes over Europe and the southeastern United States during recent winters (2009–2010 and 2010–2011) were primarily accounted for by the anomalous blocking associated with persistent episodes of large amplitude negative phase of the NAO (Guirguis et al. 2011). There is also evidence of an important role for stationary Rossby wave patterns in contributing to North American temperature extremes during summer (Schubert et al. 2011; Wu et al. 2012). These wave patterns appear to arise from internal forcing

associated with intraseasonal transient eddies (Schubert et al. 2011).

As discussed above, two more commonly recognized remote influences upon North American ETEs are associated with ENSO and the PDO (e.g., Westby et al. 2013), both involving local sea surface temperature anomalies and atmosphere–ocean coupling. These generally operate in conjunction with PNA-like teleconnection patterns that extend from the coupling region downstream into North America. Similar to the effect of climate modes, the impact of remote forcing upon warm season ETEs is partly limited by the relative inactivity and spatial extent of climate modes, which serve as horizontal pathways for Rossby wave energy between the remote forcing region and the local surface response (Schubert et al. 2011).

3.1.2 Connection to sea ice and snow cover

The atmospheric response to the Arctic sea ice reduction is thought to be Arctic warming and destabilization of the lower troposphere, increased cloudiness, and weakening of the poleward thickness gradient and polar jet stream (Francis et al. 2009; Outten and Esau 2012). As the Arctic warms faster than lower latitudes (so-called Arctic Amplification), the meridional temperature gradient at higher latitudes is likely to weaken altering the polar jet stream according to thermal wind balance. Changes in the high latitude jet stream in turn have the potential to impact weather conditions at middle and high latitudes. For example, during winter an enhanced westerly jet over the North Atlantic can help maintain relatively mild conditions over northwest Europe via heat transport from the Atlantic. Cohen et al. (2014) review three “pathways” by which Arctic amplification may impact extreme weather events in mid-latitudes. In principle, Arctic amplification may lead to regional alterations in the structure of storm tracks, jet streams and planetary waves. In recent years, considerable attention has been focused on the role of Arctic amplification-induced changes to the jet stream (Francis and Vavrus 2012; Liu et al. 2012; Barnes 2013; Screen and Simmonds 2013). Francis and Vavrus (2012) found that a weaker zonal flow (i.e., polar jet) from weakened meridional temperature gradient slows the eastward Rossby wave progression and tends to create larger meridional excursions of height contours and associated temperature displacements resulting in a higher probability of extreme weather. In a similar vein, Liu et al. (2012) argue that the circulation change due to the decline of Arctic sea ice leads to more frequent events of atmospheric blocking that cause severe cold surges over large parts of northern continents. Francis et al. (2009), Overland and Wang (2010), Jaiser et al. (2012) and Lim et al. (2012) found that there is a delayed atmospheric response to the Arctic sea ice. Specifically, the Arctic sea ice extent in summer to fall influences the atmospheric circulation in the following winter over the northern mid- to high-latitudes, affecting the seasonal winter temperature and subseasonal warm/cold spells. Evidence presented in more recent studies (Barnes 2013; Screen and Simmonds 2013), however, suggests that the role of the mechanism put forth by Francis and Vavrus (2012) is uncertain at best. More generally, Cohen et al. (2014) conclude that our understanding of the mechanistic link between ongoing Arctic amplification and mid-latitude extreme weather is currently limited by shortcomings in relevant data records, physical models and dynamical understanding, itself. As such, the likely future impact of Arctic amplification upon extreme weather is highly uncertain.

Arctic amplification is also linked to long-term variability in high latitude snow cover. An analogous linkage

between autumnal variability in Eurasian snow cover and wintertime ETE events over North America has been noted (Cohen and Jones 2011). In this case, autumnal snow cover anomalies induce a subsequent weakening of the stratospheric polar vortex during winter which, in turn, leads to a persistent negative phase episode of the tropospheric AO favoring North American cold events (Cohen et al. 2007). In addition, there is evidence that the changing Arctic sea ice extent may be linked to changes in the autumnal advance in Eurasian snow cover (Cohen et al. 2013). As for Arctic amplification, however, there is considerable uncertainty regarding the statistical robustness and physical nature of the Eurasian snow cover influence upon mid-latitude extreme weather (Peings et al. 2013; Cohen et al. 2014).

3.1.3 Large-scale climate “markers” for climate model assessment

Representation of fundamental climate modes One obvious essential minimum requirement for climate models to properly represent the modulation of ETEs by climate modes is the extent to which the models are able to represent the primary climate modes, themselves. Thus, fundamental markers for model assessment are metrics that measure the representation of key extratropical climate modes including those internally forced on intraseasonal time scales (PNA, AO/NAO and atmospheric blocking) and those externally forced on longer time scales (the extratropical response to ENSO and the atmospheric part of the PDO). Atmospheric models have had historical difficulty in representing some types of intraseasonal low frequency variability (Black and Evans 1998). A particular problem is an under-representation of atmospheric blocking activity (Scaife et al. 2010). In a similar vein, the representation of externally forced extratropical modes connected to ENSO and PDO depends on how well the coupled climate models simulate the associated oceanic phenomenological behavior.

The Coupled Model Intercomparison Project (CMIP) provides an ideal resource for assessing the ability of modern coupled climate models to represent the behavior of climate modes. A recent analysis of CMIP5 models indicates that while most models studied perform well in representing the basic aspects of the PNA pattern, a small subset of models have difficulty qualitatively replicating the NAO pattern (Lee and Black 2013; Table 4). Otherwise differences among model patterns consist of horizontal shifts or amplitude variations in the circulation anomaly pattern features. CMIP5 models generally underestimate the regional frequency of winter blocking events while summertime blocking events occurring over the high latitude oceanic basins are typically overestimated (Masato et al. 2013). Conversely, Westby et al. (2013) found serious deficiencies

Table 4 Correlation coefficients between the loading pattern of observation and that of each CMIP5 model for NAO-like and PNA-like modes

Model	NAO	PNA	Mean
GFDL-ESM2G	0.94	0.93	0.93
MPI-ESM-LR	0.83	0.92	0.88
HadCM3	0.90	0.82	0.86
CSIRO-Mk3-6-0	0.85	0.81	0.83
CCSM4	0.71	0.91	0.81
CanESM2	0.73	0.83	0.78
CNRM-CM5	0.78	0.77	0.78
MIROC-ESM-CHEM	0.73	0.78	0.75
inmcm4	0.72	0.76	0.74
NorESM1-M	0.56	0.91	0.73
IPSL-CM5A-MR	0.61	0.84	0.73
MIROC5	0.68	0.77	0.73
HadGEM2-CC	0.74	0.70	0.72
MIROC-ESM	0.72	0.70	0.71
IPSL-CM5A-LR	0.48	0.88	0.68
MRI-CGCM3	0.49	0.87	0.68
	0.72	0.83	0.77

Right-most column includes the averaged coefficients of two modes. Bold values denote high-top models and italic values denote low-top models (reproduced from Lee and Black 2013)

in the representation of the PDO by CMIP5 models with direct impacts upon the modulation of anomalous temperature regimes.

Regional flow parameters impacting remote dynamical communication The pathway between anomalous remote forcing and the regional circulation response involves several distinct factors. For example, in the case of the extratropical response to tropical heating anomalies, tropical divergent outflow interacts with subtropical vorticity gradients to produce a Rossby wave train that extends into the extratropics (Sardeshmukh and Hoskins 1988). The forced Rossby wave train can then dynamically interact with the background extratropical mean flow (such as barotropic deformation in the jet exit region) and the midlatitude storm tracks leading to a “net” large scale circulation pattern extending from the tropics into the midlatitude region of interest (e.g., Franzke et al. 2011). The ability of coupled climate models to accurately represent such pathways is dependent upon a concomitant representation of several regional atmospheric phenomena and circulation structures including:

1. The tropical large-scale circulation response to tropical diabatic heating
2. Upper tropospheric meridional vorticity gradients in the subtropics

3. Barotropic deformation structures in the jet exit regions
4. The structure and intraseasonal variability of the Atlantic and Pacific storm tracks

The authors recommend that model validation activities concentrate upon the above features in order to uncover likely sources for existing model deficiencies in leading teleconnection patterns.

3.2 Dynamics of LSMPs

3.2.1 Diagnostic tools to study dynamics of LSMP onset/decay

The onset and decay of LSMP structures linked to ETEs generally occur on relatively short time scales. As such, the range of dynamical forcing mechanisms that can directly account for LSMP time evolution is limited to internal atmospheric processes (given the relatively long time scales associated with boundary forcing). The atmospheric processes affecting LSMP evolution can be local or remote. For example, local synoptic-scale cyclogenesis can usher in a regional Arctic air outbreak while a transient episode of the PNA teleconnection pattern can provide a remote downstream influence on North American extremes. In some cases it is an optimal juxtaposition of local and remote influences operating on multiple time scales that is required to produce an extreme event (e.g., Dole et al. 2014).

Most ETE events are associated with lateral air mass excursions, which are induced by large-scale circulation anomalies in the lower troposphere forming the LSMP (e.g., Loikith and Broccoli 2012). The anomalous circulation serves as a “dynamical trigger” for ETE events. The delineation of LSMP dynamics is treated as a two-stage process: First it is of interest to assess whether the main energy source(s) are local or remote. Once the energy source *location* is determined, the second stage is to assess the specific *physical mechanism* providing the proximate energy source in this location. An effective means for assessing the source location regions for large-scale atmospheric waves is the application of wave activity flux analyses (Plumb 1985; Takaya and Nakamura 2001). The wave activity flux is parallel to Rossby wave group velocity and traces out a three-dimensional pathway between a wave source region (of flux divergence) and a wave sink region (of flux convergence).

Possible primary physical mechanisms providing wave activity sources in this context include large-scale barotropic growth, baroclinic growth/instability and nonlinear forcing by synoptic-scale eddies (Evans and Black 2003). These mechanisms may be augmented via secondary feedbacks related to internal diabatic processes or interactions of the LSMP with the local topography or land surface. Past studies have introduced comprehensive dynamical frameworks

for studying the dynamical mechanisms leading to the growth and decay of large-scale circulation anomalies (e.g., Feldstein 2002, 2003). These are based upon a local analysis of tendencies in either the geopotential (Evans and Black 2003) or streamfunction (Feldstein 2002) fields. In these studies local tendencies are decomposed into separate forcing terms that can be related to distinct physical processes. The two-stage process outlined above is a generally useful means for the dynamical diagnosis of LSMP life cycles in both observations and climate model simulations.

3.2.2 North American Arctic air mass formation

Both dynamics and thermodynamics have roles in the formation of extreme cold-air masses. Wexler (1936, 1937) postulated that the cold air was formed at the surface by radiative cooling from a snow-covered ground under clear, windless conditions, creating an intense temperature inversion restricted to a very shallow layer at the surface. Finding that the Wexler model could not adequately explain the depth of the cold layer observed in soundings, Gotaas and Benson (1965) showed that the existence of suspended ice crystals were crucial to upper-level cooling. Curry (1983) later modeled the effect with the introduction of condensate, particularly ice-crystals, in the layer. In an experiment without moisture, the inversion that formed after 2 weeks of radiative cooling was less than 1000 meters deep. Only with the addition of ice crystals did the inversion rise above 900 hPa. More recently, Emanuel (2009) found that the rates and depth of cooling in his model were sensitive to the amount of water vapor and clouds present.

Turner and Gyakum (2011), in their composite study of 93 Arctic air mass formations in Northwest Canada, found that the cold air mass lifecycle is a multi-stage set of processes. During the first stage, snow falls into a layer of unsaturated air in the lee of the Rockies, causing moisture increases in the sub-cloud layer. Simultaneously, the mid-troposphere is cooled by cloud-top radiation. On the second day, snowfall abates, the air column dries and clear-sky surface radiational cooling prevails, augmented by the high emissivity of fresh snow cover. The surface temperature falls very rapidly, as quickly as $18\text{ }^{\circ}\text{C day}^{-1}$. On the third day, after near-surface temperatures fall below their frost point, ice crystals and, nearer the surface, ice fog form. At the end of formation, there is cold-air damming (Forbes et al. 1987; Fritsch et al. 1992), with a cold pool and anticyclone in the lee of the Rockies, lower pressure in the Gulf of Alaska and an intense baroclinic zone oriented northwest to southeast along the mountains.

Figure 5 shows an event having extremely cold surface temperatures and unusual duration (17 days, compared with Turner et al., 2013, 93-case mean of 5 days). SLP anomalies tend to be higher near the coldest air, but are

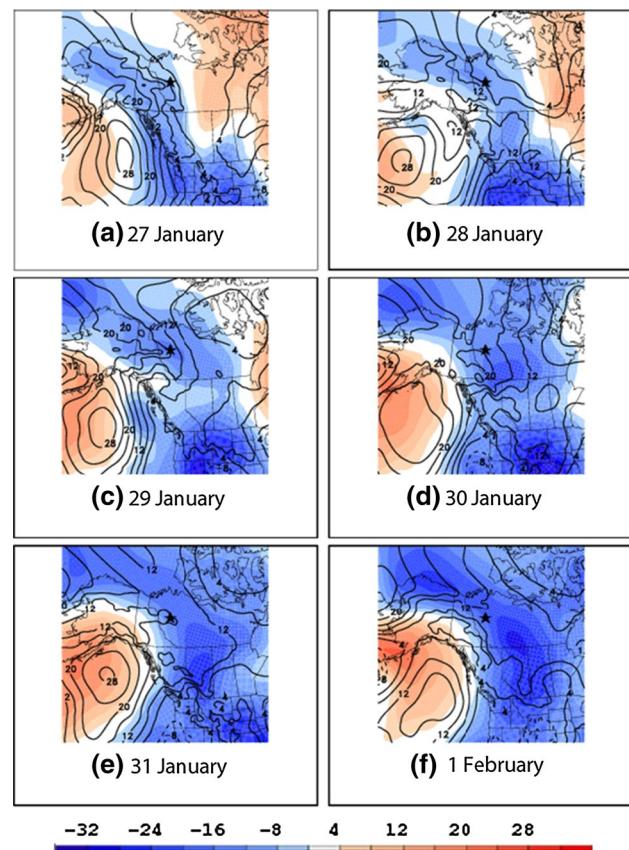


Fig. 5 Sea-level pressure anomalies (*solid/dashed* contour interval of 4 hPa indicating positive/negative values) and 1000–500 hPa thickness anomalies (*shaded*, dam) at 1200 UTC during the arctic air mass formation period of 27 January 1979 through 1 February 1979. The Norman Wells, NWT rawinsonde station is indicated by the star; by 8 February the temperature there reaches $-48.1\text{ }^{\circ}\text{C}$. The climatological reference period is 1970–2000

not extreme during this 5-day formation period. More to the point, the rapid 1000–500 hPa cooling of $\geq 10\text{ }^{\circ}\text{C}$ (≥ 20 dam) over northwestern Canada is noteworthy. Surface ridging that builds southeastward along the eastern slopes of the Rockies, particularly during 31 January and 1 February, facilitates cold-air damming and areal expansion of anomalously cold air southward to the US-Canada border. Turner et al. (2013) invoke diabatic cooling to explain the observed cooling during this formation period. This cooling may consist of sublimational cooling from precipitation falling into a dry layer and radiational cooling from suspended ice crystals. Turner et al. cite empirical evidence that much of the surface weather in this region and period includes snowfall, ice crystals, and ice fog.

3.2.3 CONUS wintertime cold air outbreaks

Cold air outbreaks over North America typically consist of a two-stage process: the first stage is the formation of

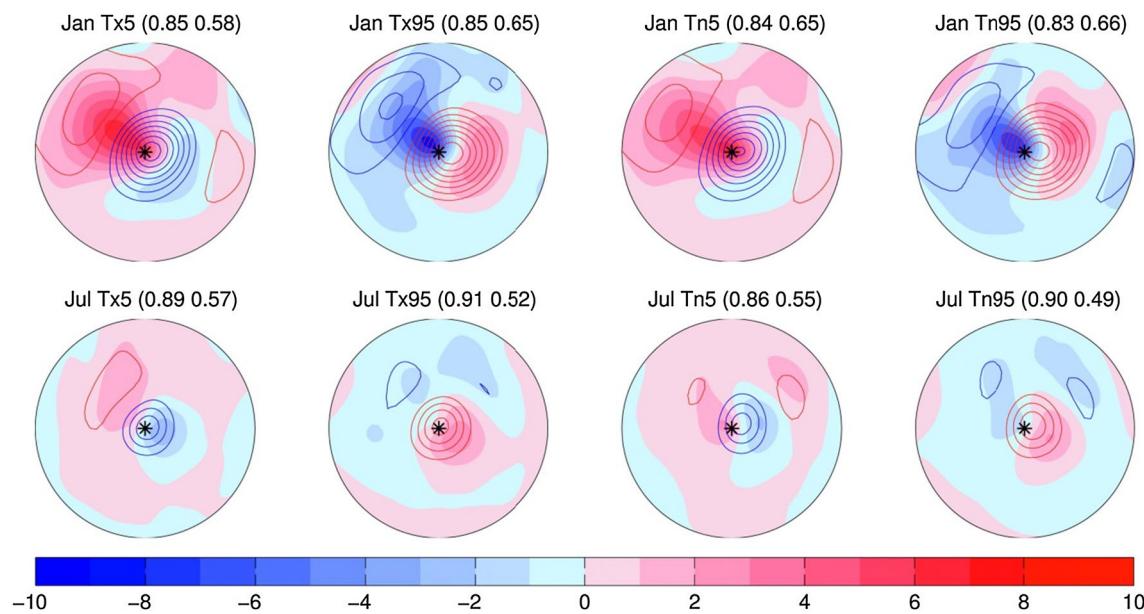


Fig. 6 Grand composites of anomalies associated with temperature regimes over North America: SLP anomalies (hPa) are shaded and Z_{500} height anomalies are contoured every 20 m; red (blue) contours are positive (negative) Z_{500} anomalies. Grand composites are shown for (*top*) January and (*bottom*) July extreme (from *left to right*) cold maximum (Tx5), warm maximum (Tx95), cold minimum (Tn5),

and warm minimum (Tn95) temperatures. The Grand composite combines information from extreme temperature events occurring at multiple disparate geographical locations over North America. The anomaly fields are displayed for a circular domain with a radius of 4500 km. From Loikith and Broccoli (2012); see text for further details

an Arctic air mass surface anticyclone at higher latitudes over Canada. This is followed by the rapid horizontal transport of the air mass to lower latitudes. The latter stage is enabled by the lower-tropospheric circulation anomaly embedded within LSMPs. North American cold air outbreaks are typically associated with a 500 hPa geopotential height anomaly pattern consisting of a broad region of negative anomalies over the CAO region, itself, along with a positive anomaly feature located to the northwest. This is illustrated in Fig. 6 (panels Jan Tx5 and Jan Tn5), which displays horizontally phase-shifted composites of the tropospheric circulation anomalies associated with local temperature extremes over North America (noting that the so-called “grand composite” patterns encompass a circular domain with a radius of 4500 km). This 500 hPa dipole structure is linked to a positive sea level pressure anomaly feature that extends between the upper level features along with a considerably weaker negative SLP anomaly located to the southeast. Thus, the CAO events are linked to a near-surface northerly flow embedded within a northwestward tilting anticyclonic circulation anomaly. Loikith and Broccoli observe that although these circulation anomaly patterns resemble large-scale teleconnection patterns, the inherent spatial length scale is closer to synoptic scale. This illustrates the importance of local synoptic features in the life cycles of ETE events. On the other hand, recent papers by Westby et al. (2013) and Loikith and Broccoli (2014)

demonstrate an additional statistically significant modulation of ETEs by larger-scale teleconnection patterns.

West coast events (e.g., CA Central Valley events) Compared with cold air outbreaks affecting the eastern and central parts of the US, less has been written about outbreaks affecting the western US. Extreme cold over California is related to a large scale pattern that brings cold air from the Arctic and northern Canada without crossing the Pacific, and hence over the Rocky Mountains (Grotjahn and Faure 2008). As discussed in Grotjahn and Faure, the large scale pattern is similar to outbreaks affecting the areas east of the Rockies however with a primary difference being a small but statistically significant ridge over the southeastern US (see Fig. 2a, b). This ridge precedes the CAO by several days, including the large ridge that develops subsequently over Alaska. A cold air mass between these ridges is directed further westward (a more northerly or northeasterly flow of the cold air) than for eastern CAOs. The distortion of the flow by that southeastern ridge enables the cold air mass to cross the Rockies. Loikith and Broccoli (2012) capture a portion of this pattern in geopotential heights at 500 hPa using their ‘grand composite’ technique. More local to the area of the extreme, colder temperatures are associated with an adjacent surface high (e.g. the conceptual model of Colle and Mass 1995). Favre and Gershunov (2006) pursue this link to anticyclones affecting western North America; and develop indices based on frequency

and central pressure of transient cyclones and anticyclones in the eastern North Pacific. Their ‘CA’ index is the difference, cyclone minus anticyclone, in strength and frequency. Negative CA values are correlated with the coldest 10 % of winter minimum temperatures.

Eastern US events (i.e., east of Rockies) Several statistical and synoptic studies either directly or indirectly relate to the topic of US CAOs occurring east of the Rockies (e.g., Konrad and Colucci 1989; Colle and Mass 1995; Konrad 1996; Walsh et al. 2001; Cellitti et al. 2006; Portis et al. 2006). Common precursors to eastern CAO events include anomalously high surface air pressure over western Canada (linked to a polar air mass), occurrences of the negative (positive) phase of the NAO (PNA) teleconnection pattern and/or an anomalously weak stratospheric polar vortex. CAO onset is characterized by the south/southeastward advection of the polar air mass in association with one or more of the following synoptic LSMP features: southward extension or propagation of the surface high pressure from Canada, surface cyclogenesis over the eastern US, anomalously low 500 hPa geopotential heights over the Great Lake region and the southeastward movement of an upper level shortwave from Canada.

Distinguishing between east coast (EC) and Midwest (MW; just east of the Rockies) CAO events, EC events result from the geostrophic cold advection enhanced by southeastward *propagation* of the surface high pressure system (with winds often amplified by surface cyclogenesis further east) leading to detachment of the polar air mass from its source (Konrad and Colucci 1989; Walsh et al. 2001). MW events are linked to the southeastward meridional *extension* of surface high pressure from Canada leading to unusually lengthy cold air transport (Walsh et al. 2001). For MW events occurring just east of the Rockies, the cold advection is related to low-level northerly ageostrophic flow within a region of cold-air damming against the topography (Colle and Mass 1995). In all cases, the salient LSMP features are synoptic-scale in nature and, thus, governed by generally well-understood dynamical physical principles considered to be well-represented by weather and climate models. The CAO process, itself, is largely due to radiative cooling (during polar air mass formation and early stages of CAO onset) and southward air mass transport (by the horizontal wind) partly offset by adiabatic warming induced by large-scale subsidence (Konrad and Colucci 1989; Walsh et al. 2001; Portis et al. 2006). Although the proximate physics associated with the CAO process itself is reasonably well characterized, our knowledge of the physics of the LSMP patterns is implicit and depends on subjective associations (between advective circulation features and synoptic LSMPs) that have been made in earlier studies. As such, there is an existing need to (a) more objectively link critical advective CAO circulation

features to dynamical entities (e.g., potential vorticity) and (b) subsequently assess the physical origin of these dynamical entities.

3.2.4 Summer heat wave events

Hot spells over North America have an intense upper level ridge as one expects hydrostatically. This ridge is seen in Figs. 2 and 6. There can also be a shallow layer of relatively lower pressure, sometimes called a ‘thermal low’. Hot spells occur from a combination of factors: a hot air mass displaced from its normal location (displaced from the southwestern desert), strong subsidence (causing adiabatic warming), and solar heating. Solar heating is more effective when surface latent heat fluxes are constrained (by drought, for example). The first two factors are enhanced and help define the LSMP for a hot spell; there may be an upstream trough (e.g. Figs. 2, 6) that enhances southwestward advection of warmer air. Other factors influence how the LSMP develops. What follows are details for different regions of North America.

West coast events Heat waves affecting the west coast are linked to an upper air LSMP that has a local ridge. For California Central Valley heat waves, that ridge is typically aligned with the west coast and the LSMP also has upstream features: a trough south of the Gulf of Alaska and a ridge further west, south of the Aleutians (Grotjahn and Faure 2008) (see Fig. 2d). The local ridge is easily understood as resulting from high thickness due to the anomalous high temperatures through the depth of the troposphere generally centered along the west coast of the US. (Plots can be found here: http://grotjahn.ucdavis.edu/EWEs/heat_wave/heat_wave.htm) Grotjahn and Faure also show that a statistically significant ridge in the northwest Pacific develops prior to the significant intensification of that west coast ridge. These features are essentially equivalent-barotropic through the depth of the troposphere. While Grotjahn and Faure found variation amongst the events, the ensemble average consists of a significant temperature anomaly that in the lower troposphere (850 and 700 hPa) is located just off the Oregon and northern California coast and elongated meridionally. The narrow zonal with longer meridional scale is consistent with station data analyzed by Bachmann (2008). Bachmann found that extreme surface temperature dates in Sacramento were more frequently matched by corresponding extreme dates at many stations located west of the Sierra Nevada and Cascades mountain ranges than closer stations east of the mountains. Grotjahn (2011) shows that the temperature anomaly leads to a thermal low being at the coast that sets up a low level pressure gradient that opposes penetration by a cooling sea breeze. The lower and mid-tropospheric flow has anomalous significant easterlies that are also downslope over some regions,

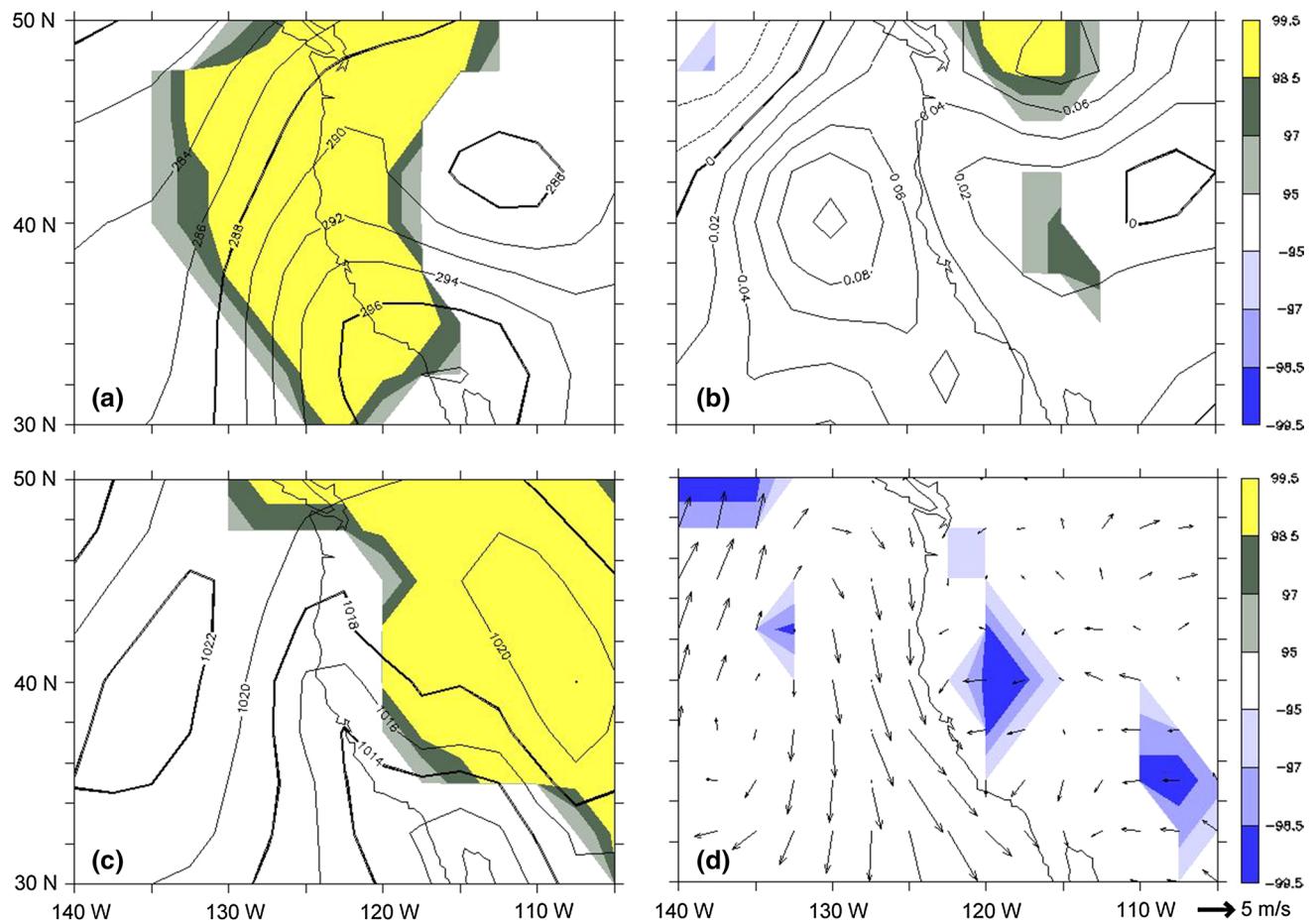


Fig. 7 Composite synoptic weather patterns at the onset of the 14 Sacramento California summer (JJAS) heat waves studied by Grotjahn and Faure (2008). **a** Temperature at 850 hPa with a 2 K interval. **b** 700 hPa level pressure velocity with 2 Pa/s interval and where positive values mean sinking motion. **c** Sea level pressure with 2 hPa interval. **d** Surface wind vectors with shading applicable to the zonal

component. Areas with yellow (lighter inside dark) shading are positive (above normal) anomalies that are large enough to occur only 1.5 % of the time by chance in a same-sized composite; areas that are blue (darker inside light) shading are negative anomalies occurring only 1.5 % of the time

most notably the Sierra Nevada mountains. Finally, there is notable sinking that lowers the climatological subsidence inversion and hence sunlight rapidly heats up the shallow layer beneath. (These factors are seen in Fig. 7).

Gershunov et al. (2009, hereafter GCI) make the novel distinction of considering separately heat waves that have high daytime maximum and events with high nighttime minimums. In so doing they uncover an LSMP with unusually high values of precipitable water (PW) over their region of interest occurring during their ‘nighttime’ heat waves (but not their ‘daytime’ events). GCI emphasize a trend of increasing occurrence of ‘nighttime’ events. Gershunov and Guirguis (2012) also find that trend in all their sub-regions of California and also show another trend: increasing longitudinal extent of Central Valley heat waves. GCI show maps of anomalous geopotential heights which also show the west coast ridge (height anomaly centered

over Washington State) with upstream (and downstream) troughs, consistent with Grotjahn and Faure. GCI show the LSMP for SLP (high over the Great Plains and low off the west coast) and remark that both daytime and nighttime events have a corresponding general southeasterly flow of air out of the desert southwest. This flow occurs throughout the lower troposphere with associated high frequency (<7 day period) heat fluxes that are prominent just off the coast (Grotjahn 2015); and consistent with setting up the offshore pressure gradient (to oppose a cooling sea breeze).

For heat waves affecting the western areas of Washington and Oregon, Bumbaco et al. (2013, hereafter BDB) show an upper level (500 hPa) ridge over the west coast and trough upstream over the Gulf of Alaska. That trough and ridge pattern is generally similar to the pattern for California heat waves, including the tropospheric height anomaly centered at the coast (Grotjahn 2011). Using a regression

approach Lau and Nath (2012) find a ridge and trough further upstream (over the western North Pacific) that are stronger several days prior to the event onset; both the location and timing of the ridge and trough are similar to those shown in Grotjahn and Faure. Similar to Grotjahn (2011, 2013, 2015), BDB find the lower tropospheric temperature anomaly (850 hPa) to be centered at the west coast of North America. Hence, BDB find negative values of sea level pressure anomaly centered offshore that sets up an offshore low level pressure gradient and offshore (northeasterly) flow (again similar to California heat waves). Consistent with the heat fluxes shown in Grotjahn (2015), the near-coast thermal trough appears to migrate from the southwestern deserts across the length of California to finally reach western Oregon and Washington (Brewer et al. 2012).

Midwest events Heat waves affecting the Midwest have anticyclonic flow at mid-levels, either as a closed anticyclone or as a strong ridge (Klein 1952; Karl and Quayle 1981; Namias 1982; Chang and Wallace 1987; Lyon and Dole 1995; Kunkel et al. 1996; Livezey and Tinker 1996; Palecki et al. 2001; Meehl and Tebaldi 2004; Lau and Nath 2012; Loikith and Broccoli 2012; Teng et al. 2013), with associated clear skies allowing maximum solar heating of the surface as well as adiabatic warming from subsidence.

For at least some events, the continental anticyclonic flow at mid-levels is part of a larger pattern of anomalies with remote centers over the Pacific and Atlantic (Namias 1982; Chang and Wallace 1987; Livezey and Tinker 1996; Lau and Nath 2012; Loikith and Broccoli 2012; Teng et al. 2013), with some suggestion that the anomalies are generated in the Pacific (Namias 1982; Lyon and Dole 1995; Livezey and Tinker 1996; Teng et al. 2013), with possible predictability (Teng et al. 2013). Teng et al. (2013) have noted the similarity of the circulation pattern to a Rossby wave number 5 wave train in the jet stream waveguide. The commonality of the wave train pattern structure and evolution among these studies is difficult to assess, though.

Studies of the surface flow have been less common, but negative SLP anomalies have been noted (Chang and Wallace 1987; Lau and Nath 2012; Loikith and Broccoli 2012), and Lau and Nath (2012) have shown anomalous southerly (or southwesterly) flow at the surface, consistent with strong horizontal warm advection. The regional-scale flow also interacts with local mechanisms, particularly the Urban Heat Island (UHI) effect (Kunkel et al. 1996; Palecki et al. 2001). Another common element identified in several studies is the presence of drought (Chang and Wallace 1987; Karl and Quayle 1981; Namias 1982; Lyon and Dole 1995) as well as simultaneous precipitation deficits (Lau and Nath 2012).

Eastern events While the large-scale meteorological patterns associated with heat waves in the eastern US have not yet seen much study, they have been examined for the Northeast

and Central US Gulf Coast in Lau and Nath (2012) and for piedmont North Carolina by Chen and Konrad (2006). For both the Northeast and Gulf Coast, Lau and Nath found a mid-level ridge centered over the region as part of a wave train and precipitation deficits, broadly similar to the results for the Midwestern regions. For piedmont North Carolina, Chen and Konrad showed that, at upper levels, a strong ridge over or just upstream of the region was a common feature along with, at lower levels, adiabatic warming associated with descent from the Blue Ridge Mountains into the piedmont region.

4 Modeling of temperature extremes and associated circulations

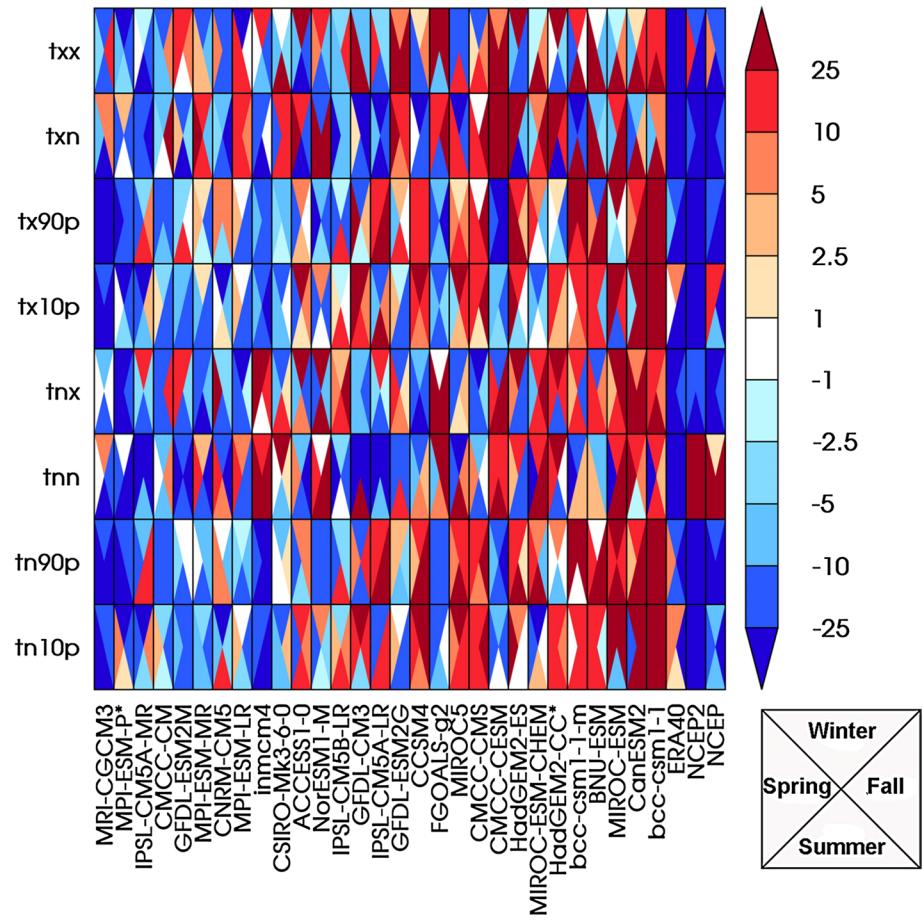
Temperature extremes have some predictability at weather and climate time scales that can be exploited by models for short term predictions and long term projections. At weather time scales the predictability of North American extreme temperature events (ETEs) is largely dependent upon the nature of the LSMPs that help organize their occurrence. The greatest predictability is expected to occur during the boreal cool season when ETEs are, at least in part, influenced by low frequency modes (PNA, NAO and blocking events) (Westby et al. 2013) with intrinsic time scales of several days to weeks (Feldstein 2000). Loikith and Broccoli (2012) illustrated that the local LSMPs linked to ETEs generally exhibit synoptic spatial scales rather than planetary scales, though a wave train can be long (e.g. Grotjahn and Faure, 2008; Fig. 2). Given the essential role of synoptic-scale disturbances such as east coast cyclones and southward moving polar anticyclones over the Midwest (Konrad and Colucci 1989; Walsh et al. 2001) in ETEs, pointwise predictability of ETEs is ultimately limited by our ability to forecast the details of synoptic-scale phenomena several days in advance (Hohenegger and Schär 2007).

Hence a significant challenge for models to predict temperature extremes is their ability to predict or simulate synoptic scale phenomena, low frequency modes that provide a large-scale meteorological context, and small scale atmospheric processes and land surface processes that influence the surface heat fluxes. Both dynamical and statistical models have been developed and used to simulate and project changes in temperature extremes. This section briefly summarizes the methods and skills of the models, while analysis of observed trends and projections of future trends are summarized in Sect. 5.

4.1 Global and regional climate model skill in simulating temperature extremes

Both global and regional climate models have been used to elucidate processes contributing to ETEs and evaluate

Fig. 8 Performance portrait of the CMIP5 models' ability to represent the temperature based ETCCDI indices over North American land. The colors represent normalized root mean square errors of seasonal indices compared to the ERA Interim reanalysis. Blue colors represent errors lower than the median error, while red colors represent errors larger than the median error. Seasons are denoted by triangles within each square. Models marked with “*” are not included in the RCP8.5 projections. Root mean square errors normalized by the model median RMSE for 3 other reanalyses are shown in the rightmost columns for comparison



model skills. The ability of global models to reproduce the observed temperature extreme statistics has been assessed by Sillmann et al. (2013a, b), Westby et al. (2013) and Wuebbles et al. (2014) using CMIP5 multi-model ensembles. Figure 8 shows a performance portrait of the normalized root mean square errors over the North American land area from the 28 available CMIP5 “historical” models for eight temperature based ETCCDI indices over 1979–2005 compared to the ERA Interim reanalysis (Dee et al. 2011). As in Gleckler et al. (2008), to plot errors from multiple variables on the same scale, they are normalized by the median error of the CMIP5 models using the formula

$$E_j^R = 100 \times \frac{(E_j - E_{median})}{E_{median}}. \quad (1)$$

Here, E_{median} is the median root mean square error (RMSE) of the CMIP5 models, E_j is the RMSE of the j th model and E_j^R is that model's “relative RMSE” and is plotted for seasonal means of the indices. In this analysis, the model median RMSE in Eq. 1 is calculated for each variable over all seasons and then applied to normalize each season in order to assess the relative seasonal performance. Blue colors represent errors lower than the median error, while red colors represent errors larger than the median error. Seasons are

denoted by triangles within each square. The different models are arranged in order of increasing average relative error with the models with the lowest average relative error on the left. The average relative error is positive for all 8 indices as Fig. 8 has more deep red than deep blue colors even though the number of positive and negative relative errors is equal. Spring has significantly lower relative error (3.8 %) averaged across all models and variables, compared to winter (5.0 %), summer (5.2 %) and fall (5.8 %).

To get a sense of model errors relative to the observational uncertainty, the 3 rightmost columns of Fig. 8, show the error of three other reanalyses considered in Sillmann et al. (2013a) relative to the model median RMSE. The two best models (leftmost in Fig. 8) are competitive with these other reanalyses but most models' RMSE relative to ERA Interim are significantly larger than the disagreement among reanalyses for many fields. Reanalysis disagreements in the cold night metrics (tnn, tn10p) and a cold day metric (tx10p) are notable and a larger number of models perform better in these fields relative to this crude measure of observational uncertainty.

Sillmann et al. (2013a) found that the spread amongst CMIP5 temperature extreme indices tends to be smaller than that of the CMIP3 models (earlier version of the

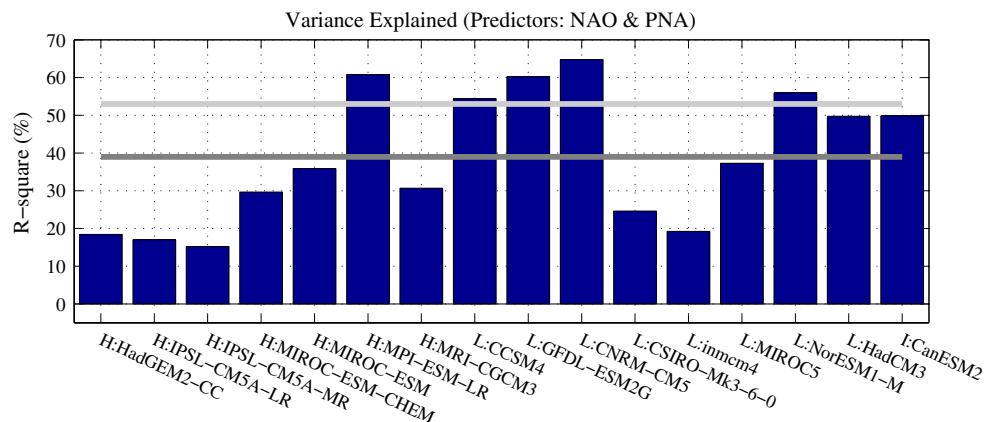


Fig. 9 The variance explained in an annual metric of the impact of wintertime Warm Waves on the southeast United States from a multiple linear regression using the NAO and PNA indices as predictors. Results are displayed for individual CMIP5 model simulations (blue

bars) while the light and dark gray lines denote variance values for observations and the model mean, respectively (from Westby et al. 2013)

CMIP5), indicating reduced uncertainty. They identified an increase in the monthly maximum daily maximum temperature and a decrease in the monthly minimum daily minimum temperature by the CMIP5 models over the high northern latitudes, compared to CMIP3. Analysis of 20 year return values in the CMIP3 models by Zwiers et al. (2011) found that models tend to warm the minimum temperature less than observed and that they warm the maximum temperature more than observed over the latter half of the twentieth century.

Westby et al. (2013) analyzed reanalysis data and CMIP5 historical simulations to study the statistics and low frequency mode modulation of wintertime CAOs and warm waves (WWs) occurring over the continental United States during 1949–2011. The observational results indicate (a) a lack of significant long-term trends in ETE frequency and (b) a seasonal modulation of CAOs (WWs) by the NAO and PNA (NAO, PNA, ENSO and PDO) patterns (Fig. 4). Similar behavior is found in the CMIP5 models, showing CAOs are mainly modulated by NAO in the southeastern US and by PNA in the northwestern US, while WW frequency is modulated by NAO over the eastern US and by a combination of PNA, PDO and ENSO over the southern US. The study also found modest influence of ENSO on WWs over the southern US, in good agreement with Lim and Schubert (2011). Comparison of WWs and CAOs between the CMIP5 models and observation indicated that the models tend to overestimate WW frequency, but underestimate CAOs frequency, although using a larger ensemble of CMIP5 models, Gao et al. (2015) found a better agreement between CAO frequency from CMIP5 models and reanalysis datasets.

Overall, CMIP5 models properly represent many of the significant associations between ETEs and low frequency

modes, particularly the modulation by the NAO and PNA patterns. Similar to Lee and Black (2013) who noted model deficiencies in representing low-frequency variability, Westby et al. (2013) found that the CMIP5 models underestimated the collective influence of low-frequency modes on temperature extremes (e.g., Fig. 9). One notable model failure is the virtual absence of a seasonal modulation of ETEs by the PDO because models do not adequately represent the nature and physics of the PDO. On the other hand, CMIP5 models do a considerably better job in representing the general behavior of the NAO and PNA patterns (Lee and Black 2013). Westby et al. conclude that predictions of future ETE behavior are ultimately limited by the ability of state-of-the-art coupled climate models to properly represent the (evolving) behavior of prominent low frequency climate modes. Nonetheless, consistent with observations, little evidence of significant trends (in either WW or CAO frequency) is found in the model simulations over the continental US from 1949 to 2011.

Heat waves are associated with anomalous large-scale circulation patterns well resolved by GCMs. Grotjahn (2013, 2015) describe how the LSMPs for California hot spells simulated by a CMIP5 climate model compare with the observed LSMPs; while the model generates a similar large scale pattern and included the needed large scale sinking, the model is unable to simulate sea breezes so the effect of LSMP blocking of sea breezes during extreme heat episodes is likely not captured by the global model. In general, heat wave intensity and duration can vary at local to regional scales because surface temperature is influenced by processes such as land surface fluxes, turbulence and winds, and clouds and radiation. For example, land-atmosphere interactions were found to play an important role in European summer heat waves of the past decades (Fischer

et al. 2007). Using a regional model, Jaeger and Seneviratne (2011) and Lorenz et al. (2010) highlighted the key role of soil moisture memory on heat wave intensity and persistence, respectively. Using observations and a CMIP3 model, Gershunov and Douville (2008) showed a strong negative relationship between preceding winter and spring precipitation and summertime heat in the central, Midwestern, and eastern US. In regions with limited soil moisture, surface temperature can respond more strongly to the large-scale circulation anomaly that leads to clear sky and strong surface solar heating because surface cooling from evapotranspiration is limited. As soil moisture is influenced by antecedent precipitation and surface properties such as land cover, topography, and soil characteristics that vary at local to regional scales, heat wave intensity and duration can vary similarly, within the context of large-scale influence by a circulation anomaly. Thus the ability of climate models to simulate and predict heat waves is also dependent on physics parameterizations and grid resolution.

Lau and Nath (2012) analyzed GCM simulations performed at 200 km and 50 km grid resolutions and found considerable resemblance between the temperature anomaly patterns from the coarse and fine resolution simulations, except for some local details generated by the higher resolution model. Both models were able to capture realistic heat wave intensity, duration, and frequency in various key regions of North America and the synoptic features accompanying the warm episodes as revealed in the North American Regional Reanalysis. Kunkel et al. (2010) compared regional climate simulations performed at 30 km resolution over North America driven by two GCM historical simulations with the respective GCM simulations and observations. Although the two GCM simulations have opposite temperature biases, downscaling by the RCM provided improved simulations of heat waves in both cases, with the overall averaged biases reduced by a factor of two compared to the GCMs. Gao et al. (2012) compared a regional simulation at 4 km grid resolution over the eastern US driven by a GCM with observations and the GCM simulation. They found statistically significant improvements in heat wave intensity and duration in the regional simulation compared to the global simulation in 14 out of 16 states in the eastern US and attributed the improvements to high-resolution topography and land use information that impact local surface temperature.

In summary, global climate models have demonstrated useful skill in simulating heat waves and cold air outbreaks over North America and their linkages to low frequency variability such as NAO, PNA, and ENSO. However, models generally underestimate the combined influence of low frequency modes on temperature extremes, with PDO and its modulations on temperature extremes being most notably deficient. Although heat waves and their LSMP are

well resolved by global models, some studies have demonstrated improved skill in simulating heat waves using regional models that may better capture regional processes such as land–atmosphere interactions influenced by surface heterogeneity.

4.2 Statistical modeling of temperature extremes

Most statistical methods for downscaling of temperature can be thought of as extensions of the model output statistics and related approaches to statistical weather forecasting (e.g., Wilks 2006). The most popular method is multiple regression analysis (e.g., Easterling 1999) but it cannot be expected to represent extremes well, as only the conditional mean of temperature given the values of the predictor variables (such as indices of LSMPs) is statistically modeled. In particular, the error term in the regression equation is usually assumed normally distributed, which is not necessarily realistic for temperature extremes.

Recently, bias correction techniques involving quantile mapping (Thrasher et al. 2012), or quantile matching (Ho et al. 2012), have become a popular way to adjust the entire probability distribution of temperature. Because of the mismatch in spatial (and sometimes temporal) scales involved in downscaling, the downscaled temperatures typically do not possess high enough variance. This issue is crucial for extremes, with it being preferable to increase the variance through randomization rather than direct “inflation” (Maraun 2013; von Storch 1999). Ideally, quantile mapping for the extreme tails of the temperature distribution should use statistical methods based on extreme value theory (e.g., Coles 2001). Kallache et al. (2011) proposed such a method, in which the quantile mapping involves fitting the Generalized Pareto distribution (GPD) to the upper (or lower) tail of the distribution. Although their application was to precipitation extremes, the technique should apply equally well to temperature extremes.

The statistical methods for relating extreme temperatures to LSMPs and climate modes, described earlier in Sect. 2, could also be applied in the context of statistical downscaling. In this approach, the extremal distribution (GEV or GPD) is conditioned on indices of LSMPs or climate modes, as well as possibly on other covariates in the context of statistical downscaling. Brown et al. (2008) relate daily maximum and minimum temperatures to the NAO index; Sillmann et al. (2011) relate European CAOs to a blocking index; Katz and Grotjahn (2014) relate California hot spells to an associated LSMP index. The chief limitation is that, given the effective reduction in data when only considering extremes, not as many predictors necessarily could be included as in more conventional statistical downscaling.

5 Observed and projected trends of temperature extremes

5.1 Observed trends and variability

Traditional methods in climate research for detecting trends in temperature extremes tend to focus on the relative frequency of exceeding a high threshold or of falling below a low threshold, as well as indices derived from such rates (Karl et al. 1996; Gleason et al. 2008). Such approaches generally fail to give a complete picture of trends in temperature extremes as they neglect the intensity of the event. It should be noted that other forms of indices are also commonly used to monitor changes in temperature extremes (Zhang et al. 2011).

Extreme value statistics (Coles 2001) formally account for both the rate of exceeding a high threshold and in the excess over a high threshold so they can be applied to detect trends in each of these two components (Katz 2010). A fundamental result in extreme value theory is that block maxima (e.g., highest daily maximum temperature each summer), suitably normalized, have an approximate generalized extreme value (GEV) distribution. In practice, it is more efficient and can be more informative to fit the GEV indirectly using the peaks over threshold (POT) approach, which involves a Poisson distribution to approximate the rate of exceeding a high threshold; and a generalized Pareto (GP) distribution to approximate the excess over the high threshold. Possible trends in extremes can be introduced into these extreme value distributions with time as a covariate. That is, the location, scale, and shape parameters of the GEV (or, equivalently, the rate parameter of the Poisson distribution and the scale and shape parameters of the GP distribution) can shift with time (Coles 2001). See Brown et al. (2008) for examples of the application of the POT approach to trend analysis of temperature extremes.

Peterson et al. (2013) applied the POT approach with a threshold corresponding to the 99th percentile for the upper tail (or 1st percentile for the lower tail) specific to each grid point in the US. To avoid temporal dependence of extremes, only the single highest (or lowest) daily temperature within a run of consecutive days exceeding the threshold was retained (termed “runs declustering”). Using the equivalent parameterization in terms of the GEV distribution, a linear trend in the location parameter of the GEV was fitted by the technique of maximum likelihood (Coles 2001), assuming that the scale and shape parameters were constant over time. To adjust for seasonality, the analysis was performed in terms of daily anomalies.

Figure 1 shows the estimated change in 20-yr return levels between 1950 and 2012, along with an indication of statistical significance. The broadest region of warming occurs for the cold tail of minimum temperature (panel

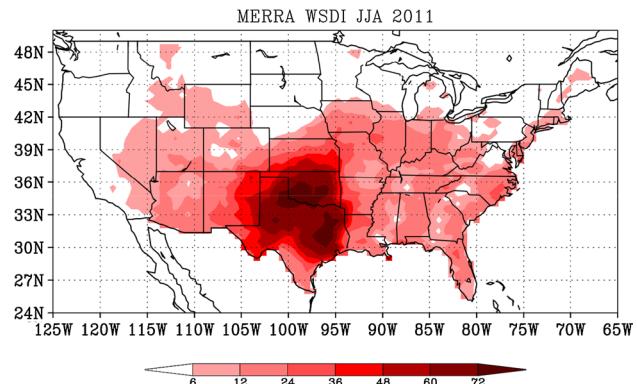


Fig. 10 Warm Spell Duration Index from MERRA for 2011 summer over the United States. WSDI is computed from the 90th percentile of daily mean temperatures for the summer season

d). Cooling occurs in the upper tail of both daily maximum (panel a) and minimum (panel c) temperature in the southeastern US. Although these results are not explicitly in terms of heat waves or cold air outbreaks, they do reflect changes in extremes that are often part of such multi-day events (Furrer et al. 2010).

Atmospheric reanalyses have provided a significant source of data for weather and climate research because they assimilate multitudes of available observing systems and provide continuous fields for regions or variables with sparse or no direct observations. Utilizing reanalyses in the calculation of extreme indices should be approached with a strong validation strategy (as discussed in Zhang et al. 2011). In the case of temperature, processes near the surface can influence the distribution of temperature. However, Simmons et al. (2010) show that ERA Interim reanalysis surface temperature reproduces that of HadCRUTv very closely, owing to the correction of soil water by the analysis of near-surface atmospheric water and temperature. Further, Bosilovich (2013) shows that even without surface analysis, reanalyses can provide robust interannual variability of seasonal surface temperature at regional scales. Furthermore, reanalysis temperature, augmented with observations, can provide a substantially improved representation of the diurnal cycle of temperature (maximum and minimum) consistently over the global land mass (Wang and Zeng 2013).

Two reanalyses, the modern-era retrospective-analysis for research and applications (MERRA) (Rienecker et al. 2011) and the climate forecast system reanalysis (CFSR) (Saha et al. 2010) have stored data at hourly intervals that allow evaluation of the diurnal cycle, especially maximum and minimum temperatures that are important to extreme temperature events. Figure 10 shows the US summertime 2011 Warm Spell Duration Index (WSDI, e.g. Alexander et al. 2006) computed from the occurrences of daily mean

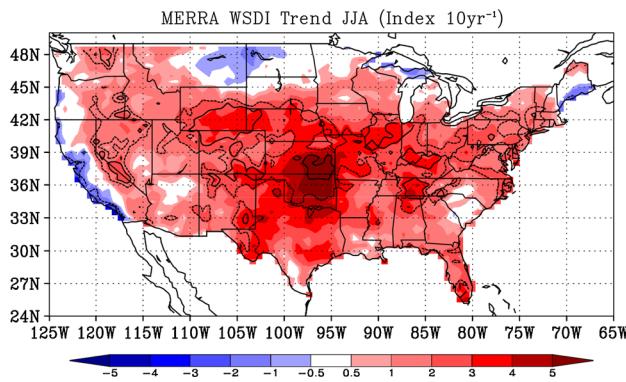


Fig. 11 Trend in WSDI as determined from MERRA for US summertime temperatures. Dashed and solid black contours indicate statistical significance (at 90 and 95 % confidence respectively)

temperature exceeding the 90th percentile for 6 days or longer from MERRA. This diagnostic relies on continuous data to derive the seasonal index, but also identify the spatial extent. Figure 11 shows the US trends in WSDI from MERRA. Positive trends are apparent in many areas of the US with some regions having statistical significance (above 90 % confidence).

Walsh et al. (2001) used global reanalysis data to study cold air outbreaks and identified Midwest, Gulf Coast, and East Coast as regions influenced by CAOs. They did not identify any apparent trend in CAOs in North America between 1948 and 1999, although analysis from a single station suggests that extreme outbreaks may have been more frequent in the late 1800s and early 1900s. Extending the study to include the recent decade from the global reanalysis data may provide further confidence in potential trends in CAOs in North America.

The twentieth century reanalysis (Compo et al. 2006, 2011) provides objectively analyzed four-dimensional data on a 2° global grid. With assimilation of synoptic surface pressure available since the late nineteenth century, this dataset covers the period from 1871 to 2012 and offers a unique opportunity for analysis of long-term trend in temperature extremes and the associated LSMP. Compo et al. (2011) showed encouraging comparison of this surface-based reanalysis product with other reanalyses that make use of upper-air and satellite data. Future work utilizing this dataset could potentially further advance understanding of variability and trends in LSMPs that influence temperature extremes.

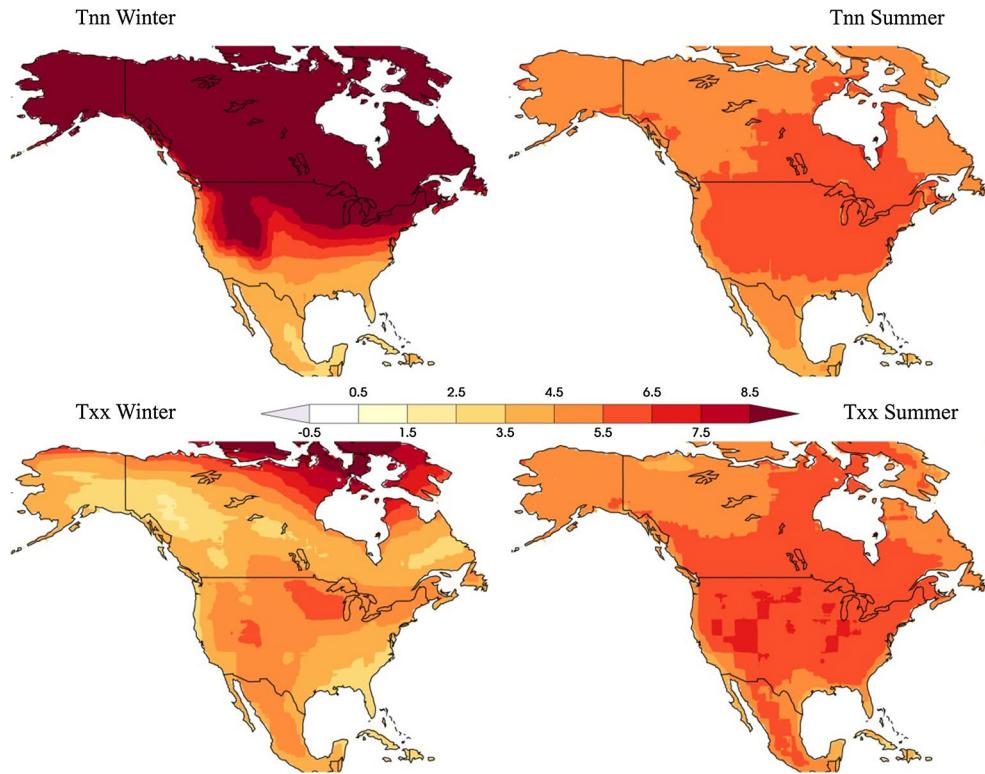
5.2 Projected trends

The CMIP3 and CMIP5 multimodel ensembles have been used to investigate projected changes of temperature extremes for the next couple of decades (e.g.,

mid-twentyfirst century) (e.g., Tebaldi et al. 2006; Orlowsky and Seneviratne 2012; Sillmann et al. 2013b; Wuebbles et al. 2014). Climate change simulations in the CMIP5 multimodel ensembles showed greater changes in ETCCDI based on daily minimum temperatures than in ETCCDI based on daily maximum temperatures. Also, the strong emission scenario such as representative concentration pathways (RCPs) 8.5 (Moss et al. 2010; van Vuuren et al. 2011) produced more definitive changes in temperature extreme statistics than weaker emission scenarios (e.g., RCP 2.6 and 4.5). Furthermore, the changes in ETCCDI under RCP8.5 tend to be greater than changes under any of the scenarios in the special report on emission scenarios (SRES) (Nakicenovic et al. 2000) used in CMIP3. The spatial distributions of Sillmann et al. (2013b) showed that the northern high latitude regions have the strongest increase in the minimum daily minimum temperature, while changes in the maximum daily maximum temperatures tend to be evenly distributed globally. The study also noted that the percentile indices such as warm and cold nights, exhibit the highest increases in the tropical regions. This is a result of differences in interannual temperature variability between the two regions, which is relatively large in mid- to high latitudes but small in the low-latitudes.

Figure 12 shows the projected ensemble mean increase in two of the ETCCDI indices assessed in Fig. 8 at the end of this century (2080–2100) from the reference period of 1985–2005 under the RCP8.5 forcing scenario of the 5th assessment report of the intergovernmental panel on climate change (IPCC AR5). Twenty-six models are included in this multi-model average. Models marked with an “*” in Fig. 8 are not included in this projection due to a lack of future simulations. The two indices, cold nights (Tnn) and hot days (Txx), are chosen to illustrate the seasonal differences of changes at the extremes of both ends of the distribution of daily surface air temperatures. The upper left panel shows that the wintertime increase in cold nights has large changes northward of the middle of the US largely corresponding to present day winter snow covered areas. Kodra et al. (2011) studied cold extremes in the twentyfirst century using nine GCMs and found that despite future warming, extreme cold events, both in terms of intensity and duration, can persist in the future. Cool nights in the summer, shown in the upper right panel, do not exhibit such a monotonic poleward behavior. Rather, the continental interior warms more than the edges for this measure of extreme temperature. Wintertime increases in warm days, shown in the lower left panel, are projected to be significantly less than wintertime cold nights; although some of the poleward gradient property is present. Summertime increases in hot days are projected slightly warmer than but generally similar to summertime increases in cool nights. One noticeable difference is that future extreme

Fig. 12 Projected seasonal changes in North American extreme temperatures from the CMIP5 multi-model at the end of this century under the RCP8.5 forcing scenario. The reference period is 1985–2005 while the future period is 2080–2100. Winter changes are shown on the *left* while summer changes are shown on the *right*. The *top* figures represent changes in cold nights (Tnn) while the *lower* figures represent changes in hot days (Txx). Units: Kelvins



temperatures in the Mexican interior warm more in the summer than in the winter.

Seasonality in the projected changes of these two indices is a reflection of the complex mechanisms affecting changes in extreme and mean temperatures in North America. Changes in snow cover certainly affect the future winter mean and extreme temperatures. Likewise, decreases in soil moisture affect the future summer mean and extreme temperatures. The additional possibility of biases and changes in LSMPs affecting extreme temperatures must be considered in light of these and other mechanisms. For example, van Oldenborgh et al. (2009) found that changes in the large-scale circulation, including a shift towards a more westerly circulation and the North Atlantic current need to be better simulated especially in winter and spring for more realistic simulations of warming over Western Europe in recent decades. Gershunov and Guirguis (2012) noted that three out of four GCMs analyzed did not adequately capture the synoptic causes of California heat waves. When sub-regional heat waves are defined relative to the changing local climate (a ‘non-stationary’ perspective), coastal heat waves have greater impact than those inland even though the summertime average warming is stronger inland. Sillmann and Croci-Maspoli (2009) found that European blocking events influence particularly the winter cold extremes in Europe. For future projection, decrease in blocking frequency with increase in maximum blocking duration was simulated under the A1B scenario.

Also, the blocking pattern shifts northeastward, affecting a larger part of Europe by giving rise to anomalously cold winter months.

To partly address the question of whether changes in heat waves in the future are associated with a shift in the daily maximum temperature or changes in LSMP, Lau and Nath (2012) noted that the probability distribution functions (PDFs) for the current and future climate over most regions that experience heat waves in the US have similar shape but were shifted by the mean warming in the daily maximum temperature, and the magnitude of skewness of the PDFs exhibited only minor changes from the current to future climate. Further statistical manipulations of the model outputs suggest that the increase in heat wave intensity and frequency in the future is primarily associated with a shift in the daily maximum surface temperature, which suggests that changes in the characteristics of LSMPs associated with heat waves may be minor. Grotjahn (2015) shows trends in an LSMP index for California hot spells in reanalysis and both historical and future simulations by a climate model. Future changes in heat waves may also be modulated by changes in land-atmosphere interactions. For example, global warming may shift the climate regime northward and establish new transitional zones with strong land-atmosphere coupling strength (Seneviratne et al. 2006) and drying over the subtropics and mid-latitudes during summer in the future may increase land-atmosphere feedbacks in general (Dirmeyer et al. 2012) and

result in amplified heat wave response to global warming. For CAOs, Schneider et al. (2015) and Gao et al. (2015) both demonstrated that the mean warming and the reduced temperature variance due to polar amplification account for most of the decrease in CAOs in a warmer climate. Gao et al. (2015) further found that changes in temperature skewness associated with changes in blocking frequency and thermodynamical modulation by melting of snow and sea ice along the 0 °C isotherm contribute additionally to regional differences in CAO changes.

Regional climate models have also been used to investigate temperature extreme changes in the future. Kunkel et al. (2010) showed that the regional model has superior skill in simulating heat waves compared to the global models, but it largely inherited the climate sensitivity from the GCMs. They found considerable regional variations in future heat wave characteristics depending on the emission scenarios and GCMs used. Using a suite of regional climate simulations driven by 5 ensemble members of a GCM, Diffenbaugh and Ashfaq (2010) found that hot extremes are intensified over much of the US associated with a summer anticyclonic circulation anomaly in the future, which reduces precipitation and soil moisture, and amplifies severe hot and dry conditions. The regional simulations showed a summer anticyclonic anomaly that is more widespread than in the GCM.

Overall, some robust changes in temperature extremes have emerged from analysis of the CMIP3 and CMIP5 multimodel ensembles as well as targeted global and regional climate modeling efforts. Model projections generally indicate increased frequency, intensity, and duration of heat waves, and despite global warming, cold extremes may persist in some regions due to changes in blocking events (Vavrus et al. 2006; Gao et al. 2015). More detailed analyses using multimodel ensembles may provide further insights on the mechanisms of heat wave changes associated with land–atmosphere coupling and potential changes in LSMPs affecting both heat waves and cold air outbreaks. Additionally, well-designed climate modeling experiments combined with observational analyses that target specific mechanisms and hypotheses may help isolate different factors and provide constraints for more robust projections of temperature extreme changes.

6 Summary

This review paper summarizes our current knowledge of, and context for developing new understandings about, extreme hot and cold temperature events affecting regions of North America. The topic of extremes encompasses many scientific issues and a breadth of time and space scales. Understanding extreme events ranges from how

events are defined and measured, to how extreme events are studied statistically, theoretically, and with models. To reduce the breadth of scales, the paper focuses upon the large scale meteorological patterns (LSMPs) that accompany extreme temperature events (ETEs) of short duration. However, even with this narrower focus, one easily sees that much further progress is needed to understand the properties of extreme temperature events and their supporting LSMP processes. Thus, the primary goal of this report is to provide guidance to researchers interested in studying extreme temperature events, especially in relation to their associated LSMPs. As such, a variety of techniques are described for identifying the LSMPs, including the relative merits of each approach. A variety of analysis tools are identified which highlight the linkage of LSMPs to both remote low frequency phenomena and local factors. Also highlighted is the information gleaned from multiple climate models including simulation issues and projected trends.

Section 2 presents methods of defining extreme temperature events statistics using both simple indices and extreme value statistical techniques. Also surveyed are methods used to identify and connect LSMPs to extreme temperatures. Recent advances in statistical techniques, involving conditional extreme value analysis, offer an opportunity to connect LSMPs to extreme temperatures through appropriately defined covariates (i.e., indices of LSMPs) that supplement more straightforward analyses.

Section 3 surveys our current knowledge of LSMPs related to ETEs. Although phenomena ranging from synoptic-scale waves to planetary-scale climate modes are implicated as contributors to ETEs, existing information on (a) the physical nature of these contributions and (b) the dynamical mechanisms responsible for the implicated LSMPs is uneven and incomplete. A diagnostic formalism is put forth for systematically isolating the underlying physics of LSMP life cycles with an ancillary goal of identifying essential large-scale circulation “markers” for climate model validation purposes.

Section 4 summarizes the approaches used to model extreme temperatures, including dynamical methods using global and regional climate models and statistical models. Although climate models generally capture heat waves and cold air outbreaks with some fidelity compared with observations, they overestimate warm wave frequency and underestimate CAO frequency. Furthermore, while CMIP5 models properly represent many of the significant associations between ETEs and low frequency modes, they underestimate the collective influence of low-frequency modes on temperature extremes, particularly related to the Pacific Decadal Oscillation. Statistical methods used to relate LSMPs with extremes are limited mainly by small sample sizes of extreme events, so that only a small number

of predictors can be included to ensure robust statistical relationships.

Section 5 surveys observed and projected trends in extreme temperatures. The studies that form the basis of the IPCC AR5 reports and the 3rd US National Climate Assessment do not consider the role of LSMPs in the magnitude and trends of extreme temperatures. However, a survey of the limited literature exploring these roles in North America and Europe suggests that future assessments, particularly at the regional scale, must include the connection between LSMPs and extreme temperature events.

In preparing this report it became apparent that much work remains to understand better past and future occurrences of ETEs and their underlying physical causes. Below are key questions relating to LSMPs that may guide future research.

Data and statistical methods Can statistical significance tests be developed for all LSMP detection methods? What are the effects of spatial resolution and gridding of station data on extreme value temperature analysis? How shall multivariate extreme statistical tools be optimally employed? What extreme values are most useful to various application sectors? Can the uncertainty be quantified in statistical quantities for given lengths of the observational or model data?

Synoptic behavior and physical mechanisms What physical/dynamical mechanisms form and maintain these rare LSMPs? Is more than one type of LSMP pattern responsible for a given regional ETE? Do specific LSMPs arise by different processes? (e.g. Lee and Grotjahn 2015) Can one quantify the extent to which low frequency phenomena such as blocking, MJO and ENSO contribute to each ETE? (e.g. Dole et al. 2014) Can one quantify the relative contributions of local factors such as topography, land surface states, and land use/land cover in ETE duration and severity? How do LSMPs vary with season and regional location? How sensitive is LSMP characterization to how the associated ETE is defined?

Model behavior and trends How well do global and regional climate models simulate the synoptic-dynamic behavior of relevant LSMPs and their variability and trends compared to observations/reanalyses? Can deficiencies in model dynamics or physical parameterizations be linked to model deficiencies in LSMP properties? How well must LSMP properties be simulated by global climate models to provide sufficiently accurate lateral boundary conditions for regional climate models to accurately simulate the associated ETE? Can metrics be designed to discern and understand differences in model skill in simulating an ETE and the associated LSMP?

Future properties What are the relative roles of dynamical and thermodynamical changes to future changes in LSMPs associated with ETEs? Considering current limitations of

climate models, how can one optimally assess the uncertainty in model projections of future changes in relevant LSMPs and their ETEs? Beyond ensemble modeling, what type of coordinated experiments can be performed to systematically evaluate models and infer the sources of model differences in LSMPs and ETEs? How can such model evaluations improve our interpretation of model projections of future changes in extremes and LSMPs?

While many gaps exist, investigators now have a large variety of tools and a useful LSMP framework by which to pursue better understanding of both historical and future extreme temperature events.

Acknowledgments The authors thank Dr. Christopher J. Paciorek for his assistance in preparing this article. The authors also thank Dr. Steven Vavrus for his comments. Most of the authors are members of the US CLIVAR Extremes working group who greatly appreciate the support provided by the US CLIVAR office. This report was enhanced by discussions held at the 2013 workshop on Analyses, Dynamics, and Modeling of Large Scale Meteorological Patterns Associated with Extreme Temperature and Precipitation Events held at the Lawrence Berkeley National Laboratory and also funded by US CLIVAR (<https://usclivar.org/meetings/extremes-workshop-agenda>). Research by Grotjahn was funded in part by NSF Grant 1236681 and also supported by the USDA National Institute of Food and Agriculture, Hatch project CA-D-LAW-4264-H. Leung and Wehner were supported by the Regional and Global Climate Modeling Program of the Office of Biological and Environmental Research in the Department of Energy Office of Science. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the DOE under contract DE-AC05-76RL01830 and Lawrence Berkeley National Laboratory is under contract DE-AC02-05CH11231. Black was supported by the US Department of Energy, Office of Biological and Environmental Research, awards DE-SC0004942 and DE-SC0012554, and the National Science Foundation Grant ARC-1107384. Gutowski was supported by National Science Foundation Grant ARC1023369 and Department of Energy Grant DESC0006643. Gyakum was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant, and by an International Polar Year Grant. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL (2006) Global observed changes in daily climate extremes of temperature and precipitation. *J Geophys Res* 111:D05109. doi:[10.1029/2005JD006290](https://doi.org/10.1029/2005JD006290)
- Alexander L, Uotila P, Nicholls N (2009) Influence of sea surface temperature variability on global temperature and precipitation extremes. *J Geophys Res* 114:D18116

- Bachmann B (2008) The spatial extent of California heat waves. MS Thesis, Atmospheric Science Program, University of California Davis. March. 164 pp. <http://grotjahn.ucdavis.edu/EWES/>
- Barnes EA (2013) Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes. *Geophys Res Lett* 40:4734–4739. doi:[10.1002/grl.50880](https://doi.org/10.1002/grl.50880)
- Beniston M (2004) The 2003 heat wave in Europe: a shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophys Res Lett* 31:L02202. doi:[10.1029/2003GL018857](https://doi.org/10.1029/2003GL018857)
- Black RX, Evans KJ (1998) The statistics and horizontal structure of anomalous weather regimes in the community climate model. *Mon Weather Rev* 126(4):841–859
- Bosilovich MG (2013) Regional climate and variability of NASA MERRA and recent reanalyses: US summertime precipitation and temperature. *J Appl Meteorol Climatol* 52:1939–1951. doi:[10.1175/JAMC-D-12-0291.1](https://doi.org/10.1175/JAMC-D-12-0291.1)
- Brewer MC, Mass CF, Potter BE (2012) The west coast thermal trough: climatology and synoptic evolution. *Mon Weather Rev* 140:3820–3843. doi:[10.1175/MWR-D-12-00078.1](https://doi.org/10.1175/MWR-D-12-00078.1)
- Brown S, Caesar J, Ferro C (2008) Global changes in extreme daily temperature since 1950. *J Geophys Res* 113:D05115. doi:[10.1029/2006JD008091](https://doi.org/10.1029/2006JD008091)
- Bumbaco KA, Dello KD, Bond NA (2013) History of Pacific Northwest heat waves: synoptic pattern and trends. *J Appl Meteorol Climatol* 52:1618–1631. doi:[10.1175/JAMC-D-12-094.1](https://doi.org/10.1175/JAMC-D-12-094.1)
- Carrera M, Higgins R, Kousky V (2004) Downstream weather impacts associated with atmospheric blocking over the Northeast Pacific. *J Clim* 17:4823–4839
- Cassano EN, Lynch AH, Cassano JJ, Koslow MR (2006) Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska. *Clim Res* 30:83–97
- Cassano JJ, Uotila P, Lynch AH, Cassano EN (2007) Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21st century. *J Geophys Res* 112:G04S49. doi:[10.1029/2006JG000332](https://doi.org/10.1029/2006JG000332)
- Cavazos T (2000) Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans. *J Clim* 13:1718–1732
- Cellitti M, Walsh J, Rauber R, Portis D (2006) Extreme cold air outbreaks over the United States, the polar vortex, and the large-scale circulation. *J Geophys Res* 111:D02114. doi:[10.1029/2005JD006273](https://doi.org/10.1029/2005JD006273)
- Chang F-C, Wallace JM (1987) Meteorological conditions during heat waves and droughts in the United States great plains. *Mon Weather Rev* 115:1253–1269
- Chen F, Konrad CE II (2006) A synoptic climatology of summertime heat and humidity in the piedmont region of North Carolina. *J Appl Meteorol Climatol* 45(5):674–685
- Christidis N, Stott PA, Brown S, Hegerl GC, Caesar J (2005) Detection of changes in temperature extremes during the second half of the 20th century. *Geophys Res Lett* 32:L20716. doi:[10.1029/2005GL023885](https://doi.org/10.1029/2005GL023885)
- Cohen J, Jones J (2011) A new index for more accurate winter predictions. *Geophys Res Lett* 38:L21701. doi:[10.1029/2011GL049626](https://doi.org/10.1029/2011GL049626)
- Cohen J, Barlow M, Kushner P, Saito K (2007) Stratosphere–troposphere coupling and links with Eurasian land-surface variability. *J Clim* 20:5335–5343. doi:[10.1175/2007JCLI1725.1](https://doi.org/10.1175/2007JCLI1725.1)
- Cohen J, Jones J, Furtado JC, Tziperman E (2013) Warm Arctic, cold continents: a common pattern related to Arctic sea ice melt, snow advance, and extreme winter weather. *Oceanography*. doi:[10.5670/oceanog.2013.70](https://doi.org/10.5670/oceanog.2013.70)
- Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J, Jones J (2014) Recent Arctic amplification and extreme mid-latitude weather. *Nat Geosci* 7:627–637. doi:[10.1038/ngeo2234](https://doi.org/10.1038/ngeo2234)
- Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London
- Colle BA, Mass CF (1995) The structure and evolution of cold surges east of the Rocky Mountains. *Mon Weather Rev* 123:2577–2610. doi:[10.1175/1520-0493\(1995\)123<2577:TSAEOC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<2577:TSAEOC>2.0.CO;2)
- Compo GP, Whitaker JS, Sardeshmukh PD (2006) Feasibility of a 100 year reanalysis using only surface pressure data. *Bull Am Meteorol Soc* 87:175–190
- Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Krul MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ (2011) The twentieth century reanalysis project. *Q J R Meteorol Soc* 137:1–28. doi:[10.1002/qj.776](https://doi.org/10.1002/qj.776)
- Curry J (1983) On the formation of continental polar air. *J Atmos Sci* 40:2278–2292
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597
- Diffenbaugh NS, Ashfaq M (2010) Intensification of hot extremes in the United States. *Geophys Res Lett*. doi:[10.1029/2010GL043888](https://doi.org/10.1029/2010GL043888)
- Dirmeyer PA, Cash BA, Kinter JL III, Stan C, Jung T, Marx L, Towers P, Wedi N, Adams JM, Altshuler EL, Huang B, Jin EK, Manganelli J (2012) Evidence for enhanced land-atmosphere feedback in a warming climate. *J Hydrometeorol* 13(3):981–995
- Dole R, Hoerling M, Kumar A, Eischeid J, Perlitz J, Quan X-W, Kiladis G, Webb R, Murray D, Chen M, Wolter K, Zhang T (2014) The making of an extreme event: putting the pieces together. *Bull Am Meteorol Soc* 95:427–440. doi:[10.1175/BAMS-D-12-00069.1](https://doi.org/10.1175/BAMS-D-12-00069.1)
- Downton M, Miller K (1993) The freeze risk to Florida citrus. Part II: temperature variability and circulation patterns. *J Clim* 6:364–372
- Easterling D (1999) Development of regional climate scenarios using a downscaling approach. *Clim Change* 41:615–634
- Emanuel K (2009) Back to Norway: an essay. Synoptic–dynamic meteorology and weather analysis forecasting. *Meteorol Monogr* 55, Am Meteorol Soc 87–96
- Evans KJ, Black RX (2003) Piecewise tendency diagnosis of weather regime transitions. *J Atmos Sci* 60:1941–1959
- Favre A, Gershunov A (2006) Extra-tropical cyclonic/anticyclonic activity in North-Eastern Pacific and air temperature extremes in Western North America. *Clim Dyn* 26:617–629. doi:[10.1007/s00382-005-0101-9](https://doi.org/10.1007/s00382-005-0101-9)
- Fawcett L, Walshaw D (2012) Estimating return levels from serially dependent extremes. *Environmetrics* 23:272–283
- Feldstein SB (2000) Teleconnections and ENSO: the timescale, power spectra, and climate noise properties. *J Clim* 13:4430–4440
- Feldstein SB (2002) Fundamental mechanisms of PNA teleconnection pattern growth and decay. *Q J R Meteorol Soc* 128:775–796
- Feldstein SB (2003) The dynamics of NAO teleconnection pattern growth and decay. *Q J R Meteorol Soc* 129:901–924
- Fischer EM, Seneviratne SI, Lüthi D, Schär C (2007) Contribution of land-atmosphere coupling to recent European summer

- heat waves. *Geophys Res Lett* 34:L06707. doi:[10.1029/2006GL029068](https://doi.org/10.1029/2006GL029068)
- Forbes GS, Anthes RA, Thomson DW (1987) Synoptic and mesoscale aspects of an Appalachian ice storm associated with cold-air damming. *Mon Weather Rev* 115:564–591
- Francis JA, Vavrus SJ (2012) Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys Res Lett*. doi:[10.1029/2012GL051000](https://doi.org/10.1029/2012GL051000)
- Francis JA, Chan W, Leathers DJ, Miller JR, Veron DE (2009) Winter northern hemisphere weather patterns remember summer Arctic sea-ice extent. *Geophys Res Lett* 36:L07505. doi:[10.1029/2009GL037274](https://doi.org/10.1029/2009GL037274)
- Franzke C, Feldstein SB, Lee S (2011) Synoptic analysis of the Pacific-North American teleconnection pattern. *Q J R Meteorol Soc* 137:329–346. doi:[10.1002/qj.768](https://doi.org/10.1002/qj.768)
- Fritsch JM, Kapolka J, Hirschberg PA (1992) The effects of subcloud-layer diabatic processes on cold air damming. *J Atmos Sci* 49:49–70
- Furrer EM, Katz RW, Walter MD, Furrer R (2010) Statistical modeling of hot spells and heat waves. *Clim Res* 43:191–205
- Gao Y, Fu JS, Drake JB, Liu Y, Lamarque J-F (2012) Projected changes of extreme weather events in the eastern United States based on a high resolution climate modeling system. *Environ Res Lett*. doi:[10.1088/1748-9326/7/4/044025](https://doi.org/10.1088/1748-9326/7/4/044025)
- Gao Y, Leung LR, Lu J, Masato G (2015) Persistent cold air outbreaks over North America in a warming climate. *Environ Res Lett*. doi:[10.1088/1748-9326/10/4/044001](https://doi.org/10.1088/1748-9326/10/4/044001)
- Gershunov A, Douville H (2008) Extensive summer hot and cold extremes under current and possible future climatic conditions: Europe and North America. In: Diaz HF, Murnane RJ (eds) Climate extremes and society. Cambridge University Press, New York, pp 74–98
- Gershunov A, Guirguis K (2012) California heat waves in the present and future. *Geophys Res Lett* 39:L18710. doi:[10.1029/2012GL052979](https://doi.org/10.1029/2012GL052979)
- Gershunov A, Cayan DR, Iacobellis SF (2009) The great 2006 heat wave over California and Nevada: signal of an increasing trend. *J Clim* 22:6181–6203
- Gleason KL, Lawrimore JH, Levinson DH, Karl TR, Karoly DJ (2008) A revised US climate extremes index. *J Clim* 21:2124–2137. doi:[10.1175/2007JCLI1883.1](https://doi.org/10.1175/2007JCLI1883.1)
- Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. *J Geophys Res* 113:D06104. doi:[10.1029/2007JD008972](https://doi.org/10.1029/2007JD008972)
- Gosling SN, McGregor GR, Paldy A (2007) Climate change and heat-related mortality in six cities Part 1: model construction and validation. *Int J Biometeorol* 51:525–540
- Gotaas Y, Benson CS (1965) The effect of suspended ice crystals on radiative cooling. *J Appl Meteorol* 4:446–453
- Goto-Maeda Y, Shin D, O'Brien J (2008) Freeze probability of Florida in a regional climate model and climate indices. *Geophys Res Lett* 35:L11703. doi:[10.1029/2008GL033720](https://doi.org/10.1029/2008GL033720)
- Grotjahn R (2011) Identifying extreme hottest days from large scale upper air data: a pilot scheme to find California central valley summertime maximum surface temperatures. *Clim Dyn* 37:587–604. doi:[10.1007/s00382-011-0999-z](https://doi.org/10.1007/s00382-011-0999-z)
- Grotjahn R (2013) Ability of CCSM4 to simulate California extreme heat conditions from evaluating simulations of the associated large scale upper air pattern. *Clim Dyn* 41:1187–1197. doi:[10.1007/s00382-013-1668-1](https://doi.org/10.1007/s00382-013-1668-1)
- Grotjahn R (2015) Western North American extreme heat, associated large scale synoptic-dynamics, and performance by a climate model. In: Li J, Swinbank R, Volkert H, Grotjahn R (eds) Dynamics and predictability of large-scale, high-impact weather and climate events. Cambridge University Press, Cambridge
- Grotjahn R, Faure G (2008) Composite predictor maps of extraordinary weather events in the Sacramento California region. *Weather Forecast* 23:313–335. doi:[10.1175/2007WAF2006055.1](https://doi.org/10.1175/2007WAF2006055.1)
- Gu L, Hanson P, Post W, Kaiser D, Yang B, Nemani R, Pallardy S, Meyers T (2008) The 2007 eastern US spring freezes: increased cold damage in a warming world? *Bioscience* 58:253–262
- Guirguis K, Gershunov A, Schwartz R, Bennett S (2011) Recent warm and cold daily winter temperature extremes in the Northern Hemisphere. *Geophys Res Lett* 38:L17701. doi:[10.1029/2011GL048762](https://doi.org/10.1029/2011GL048762)
- Gutowski WJ, Otieno F, Arritt RW, Takle ES, Pan Z (2004) Diagnosis and attribution of a seasonal precipitation deficit in a US regional climate simulation. *J Hydrometeorol* 5:230–242. doi:[10.1175/1525-7541\(2004\)005<0230:DAAOAS>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0230:DAAOAS>2.0.CO;2)
- Hajat S, Kovats RS, Atkinson RW, Haines A (2002) Impact of hot temperatures on death in London: a time series approach. *J Epidemiol Community Health* 56:367–372
- Hewitson BC, Crane RG (2002) Self-organizing maps: applications to synoptic climatology. *Clim Res* 18:41–57
- Hewitson BC, Crane RG (2006) Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int J Climatol* 26:1315–1337. doi:[10.1002/joc.1314](https://doi.org/10.1002/joc.1314)
- Higgins R, Leetmaa A, Kousky V (2002) Relationships between climate variability and winter temperature extremes in the United States. *J Clim* 15:1555–1572
- Hinton G, Deng L, Yu D, Dahl G, Mohamed A-R, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T, Kingsbury B (2012) Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *Sig Process Mag IEEE* 29(6):82–97
- Ho CK, Stephenson DB, Collins M, Ferro CAT, Brown SJ (2012) Calibration strategies: a source of additional uncertainty in climate change projections. *Bull Am Meteorol Soc* 93:21–26
- Hohenegger C, Schär C (2007) Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull Am Meteorol Soc* 88:1783–1793
- Jaeger EB, Seneviratne SI (2011) Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model. *Clim Dyn* 36(9–10):1919–1939. doi:[10.1007/s00382-010-0780-8](https://doi.org/10.1007/s00382-010-0780-8)
- Jaiser R, Dethloff K, Handorf D, Rinke A, Cohen J (2012) Impact of sea ice cover changes on the northern hemisphere atmospheric winter circulation. *Tellus A* 64:11595. doi:[10.3402/tellusa.v64i0.11595](https://doi.org/10.3402/tellusa.v64i0.11595)
- Kallache M, Vrac M, Naveau P, Michelangeli P-A (2011) Nonstationary probabilistic downscaling of extreme precipitation. *J Geophys Res* 116:D05113. doi:[10.1029/2010JD014892](https://doi.org/10.1029/2010JD014892)
- Karl TR, Quayle RG (1981) The 1980 summer heat wave and drought in historical perspective. *Mon Weather Rev* 109:2055–2073. doi:[10.1175/1520-0493\(1981\)109<2055:TSHWAD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<2055:TSHWAD>2.0.CO;2)
- Karl TR, Knight RW, Easterling DR, Quayle RG (1996) Indices of climate change for the United States. *Bull Am Meteorol Soc* 77:279–292. doi:[10.1175/1520-0477\(1996\)077<0279:IOCCFT>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0279:IOCCFT>2.0.CO;2)
- Katz RW (2010) Statistics of extremes in climate change. *Clim Change* 100:71–76
- Katz RW, Grotjahn R (2014) Statistical methods for relating temperature extremes to large-scale meteorological patterns. *US CLIVAR Var* 12(1):4–7. <http://www.usclivar.org/sites/default/files/documents/2014/Variations-winter-2014-v3.pdf>
- Kenyon J, Hegerl G (2008) Influence of modes of climate variability on global temperature extremes. *J Clim* 21:3872–3889. doi:[10.1175/2008JCLI2125.1](https://doi.org/10.1175/2008JCLI2125.1)

- Kharin VV, Zwiers FW, Zhang X, Wehner M (2013) Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Clim Change* 119:345–347
- Klein W (1952) The weather and circulation of June 1952. *Mon Weather Rev* 80:99–104
- Kodra E, Steinhäuser K, Ganguly AR (2011) Persisting cold extremes under 21st-century warming scenarios. *Geophys Res Lett* 38:L08705. doi:[10.1029/2011GL047103](https://doi.org/10.1029/2011GL047103)
- Kohonen T (1995) Self-organizing maps, 2nd edn. Springer, Berlin
- Konrad C II (1996) Relationships between the intensity of cold-air outbreaks and the evolution of synoptic and planetary-scale features over north America. *Mon Weather Rev* 124:1067–1083. doi:[10.1175/1520-0493\(1996\)124<1067:RBTIOC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1067:RBTIOC>2.0.CO;2)
- Konrad CE II, Colucci SJ (1989) An examination of extreme cold air outbreaks over eastern North America. *Mon Weather Rev* 117:2687–2700
- Krizhevsky A, Sutskever I, Hinton G (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems 25 (proceedings from NIPS 2012). <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- Kunkel KE, Changnon SA, Reinke BC, Arritt RW (1996) The July 1995 heat wave in the Midwest: a climatic perspective and critical weather factors. *Bull Am Meteorol Soc* 77:1507–1518
- Kunkel KE, Liang X-Z, Zhu J (2010) Regional climate model projections and uncertainties of US Summer heat waves. *J Clim* 23:4447–4458
- Lau N-C, Nath MJ (2012) A model study of heat waves over North America: meteorological aspects and projections for the twenty-first century. *J Clim* 25:4761–4784. doi:[10.1175/JCLI-D-11-00575.1](https://doi.org/10.1175/JCLI-D-11-00575.1)
- Lee Y-Y, Black RX (2013) Boreal winter low-frequency variability in CMIP5 models. *J Geophys Res Atmos* 118:6891–6904. doi:[10.1002/jgrd.50493](https://doi.org/10.1002/jgrd.50493)
- Lee Y-Y, Grotjahn R (2015) California Central Valley summer heat waves form two ways. *J Clim* (under review)
- Leibensperger EM, Mickley LJ, Jacob DJ, Chen W-T, Seinfeld JH, Nenes A, Adams PJ, Streets DG, Kumar N, Rind D (2012) Climatic effects of 1950–2050 changes in US anthropogenic aerosols—Part 2: climate response. *Atmos Chem Phys* 12:3349–3362. doi:[10.5194/acp-12-3349-2012](https://doi.org/10.5194/acp-12-3349-2012)
- Lim YK, Kim HD (2013) Impact of the dominant large-scale teleconnections on winter temperature variability over East Asia. *J Geophys Res* 118:7835–7848. doi:[10.1002/jgrd.50462](https://doi.org/10.1002/jgrd.50462)
- Lim Y-K, Schubert S (2011) The impact of ENSO and the Arctic Oscillation on winter temperature extremes in the southeast United States. *Geophys Res Lett* 38:L15706. doi:[10.1029/2011GL048283](https://doi.org/10.1029/2011GL048283)
- Lim YK, Ham YG, Jeong JH, Kug JS (2012) Improvement in simulation of Eurasian winter climate variability with a realistic Arctic sea ice condition in an atmospheric GCM. *Environ Res Lett* 7:044041. doi:[10.1088/1748-9326/7/4/044041](https://doi.org/10.1088/1748-9326/7/4/044041)
- Lipton K, Grumm R, Holmes R (2005) Forecasting heat waves using climatic anomalies. In: 21st conference on weather analysis/17th conference on numerical weather prediction. 1–5 August 2005. <http://ams.confex.com/ams/pdfpapers/94498.pdf>
- Liu J, Curry JA, Wang H, Song M, Horton RM (2012) Impact of declining Arctic sea ice on winter snowfall PNAS. *Proc Natl Acad Sci* 109:4074–4079. doi:[10.1073/pnas.1114910109](https://doi.org/10.1073/pnas.1114910109)
- Livezey RE, Tinker R (1996) Some meteorological, climatological, and microclimatological considerations of the severe US heat wave of mid-July 1995. *Bull Am Meteorol Soc* 77:2043–2054
- Loikith P, Broccoli A (2012) Characteristics of observed atmospheric circulation patterns associated with temperature extremes over North America. *J Clim* 25:7266–7281. doi:[10.1175/JCLI-D-11-00709.1](https://doi.org/10.1175/JCLI-D-11-00709.1)
- Loikith PC, Broccoli AJ (2014) The influence of recurrent modes of climate variability on the occurrence of winter and summer extreme temperatures over North America. *J Clim* 27:1600–1618
- Lorenz R, Jaeger EB, Seneviratne SI (2010) Persistence of heat waves and its link to soil moisture memory. *Geophys Res Lett* 37:L09703. doi:[10.1029/2010GL042764](https://doi.org/10.1029/2010GL042764)
- Lyon B (2009) Southern Africa summer drought and heat waves: observations and coupled model behavior. *J Clim* 22:6033–6046. doi:[10.1175/2009JCLI3101.1](https://doi.org/10.1175/2009JCLI3101.1)
- Lyon B, Dole RM (1995) A diagnostic comparison of the 1980 and 1988 US summer heat wave-droughts. *J Clim* 8:1658–1675
- Maraun D (2013) Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J Clim* 26:2137–2143
- Masato G, Hoskins BJ, Woollings T (2013) Winter and summer northern hemisphere blocking in CMIP5 models. *J Clim* 26:7044–7059
- Meehl G, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the twenty-first century. *Science* 305:994–997
- Meehl G, Tebaldi C, Teng H, Peterson T (2007) Current and future US weather extremes and El Niño. *Geophys Res Lett* 34:L20704. doi:[10.1029/2007GL031027](https://doi.org/10.1029/2007GL031027)
- Moon J-Y, Wang B, Ha K (2011) ENSO regulation of MJO teleconnection. *Clim Dyn* 37:1133–1149
- Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463:747–756. doi:[10.1038/nature08823](https://doi.org/10.1038/nature08823)
- Nakicenovic N, Alcamo J, Davis G, de Vries HJM, Fenner J, Gaffin S, Gregory K, Grubler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Papper W, Pitcher H, Price L, Riahi K, Roechl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) Special report on emissions scenarios. Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge
- Namias J (1982) Anatomy of great plains protracted heat waves (especially the 1980 US summer drought). *Mon Weather Rev* 110:824–838
- Orlowsky B, Seneviratne SI (2012) Global changes in extreme events: regional and seasonal dimension. *Clim Change* 110:669–696. doi:[10.1007/s10584-011-0122-9](https://doi.org/10.1007/s10584-011-0122-9)
- Outten SD, Esau I (2012) A link between Arctic sea ice and recent cooling trends over Eurasia. *Clim Change* 110:1069–1075. doi:[10.1007/s10584-011-0334-z](https://doi.org/10.1007/s10584-011-0334-z)
- Overland JE, Wang M (2010) Large-scale atmospheric circulation changes associated with the recent loss of Arctic sea ice. *Tellus A* 62:1–9. doi:[10.1111/j.1600-0870.2009.00421.x](https://doi.org/10.1111/j.1600-0870.2009.00421.x)
- Palecki MA, Changnon SA, Kunkel KE (2001) The nature and impacts of the July 1999 heat wave in the midwestern United States: learning from the lessons of 1995. *Bull Am Meteorol Soc* 82:1353–1367
- Park T-W, Ho C-H, Yang S (2011) Relationship between the Arctic Oscillation and cold surges over East Asia. *J Clim* 24:68–83
- Peings Y, Brun E, Mauvais V, Douville H (2013) How stationary is the relationship between Siberian snow and Arctic Oscillation over the 20th century? *Geophys Res Lett* 40:183–188. doi:[10.1029/2012GL054083](https://doi.org/10.1029/2012GL054083)
- Peterson TC, Heim R, Hirsch R, Kaiser D, Brooks H, Diffenbaugh NS, Dole R, Giovannettone J, Guiguis K, Karl TR, Katz RW, Kunkel K, Lettenmaier D, McCabe GJ, Paciorek CJ, Ryberg K, Schubert S, Silva VBS, Stewart B, Vecchia AV, Villarini G, Vose RS, Walsh J, Wehner M, Wolock D, Wolter K, Woodhouse

- CA, Wuebbles D (2013) Monitoring and understanding changes in heat waves, cold waves, floods and droughts in the United States: state of knowledge. *Bull Am Meteorol Soc* 94:821–834. doi:[10.1175/BAMS-D-12-00066.1](https://doi.org/10.1175/BAMS-D-12-00066.1), Supplement doi:[10.1175/BAMS-D-12-00066.2](https://doi.org/10.1175/BAMS-D-12-00066.2)
- Photiadou C, Jones MR, Keelings D, Dewes CF (2014) Modelling European hot spells using extreme value analysis. *Clim Res* 58:193–207. doi:[10.3354/cr01191](https://doi.org/10.3354/cr01191)
- Plumb RA (1985) On the three-dimensional propagation of stationary waves. *J Atmos Sci* 42:217–229
- Portis DH, Cellitti MP, Chapman WL, Walsh JE (2006) Low-frequency variability and evolution of north american cold air outbreaks. *Mon Weather Rev* 134:579–597
- Renwick JA, Wallace JM (1996) Relationships between north Pacific wintertime blocking, El Niño, and the PNA pattern. *Mon Weather Rev* 124:2071–2076
- Rienecker M et al (2007) The GEOS-5 data assimilation system—documentation of versions 5.0.1 and 5.1.0. NASA GSFC technical report series on global modeling and data assimilation, NASA/TM-2007-104606, vol 27, 92 pp
- Robinson P (2001) On the definition of a heat wave. *J Appl Meteorol* 40:762–775
- Saha S et al (2010) The NCEP climate forecast system reanalysis. *Bull Am Meteorol Soc* 91:1015–1057
- Salakhutdinov R, Hinton G (2012) An efficient learning procedure for deep Boltzmann machines. *Neural Comput* 24(8):1967–2006
- Sardeshmukh PD, Hoskins BJ (1988) The generation of global rotational flow by steady idealized tropical divergence. *J Atmos Sci* 45:1228–1251
- Scaife AA, Woollings T, Knight J, Martin G, Hinton T (2010) Atmospheric blocking and mean biases in climate models. *J Clim* 23(23):6143–6152
- Schneider T, Bischoff T, Plotka H (2015) Physics of changes in synoptic midlatitude temperature variability. *J Clim* 28:2312–2331
- Schubert S, Wang H, Suarez M (2011) Warm season subseasonal variability and climate extremes in the Northern Hemisphere: the role of stationary Rossby waves. *J Clim* 24:4773–4792
- Screen JA, Simmonds I (2013) Exploring links between Arctic amplification and midlatitude weather. *Geophys Res Lett* 40:959–964. doi:[10.1002/GRL.50174](https://doi.org/10.1002/GRL.50174)
- Seber GAF (2008) Multivariate distributions, multivariate observations, Wiley, pp 17–58, ISBN 0-471-69121-6
- Seneviratne SI, Luthi D, Litschi M, Schar C (2006) Land-atmosphere coupling and climate change in Europe. *Nature* 443:205–209. doi:[10.1038/nature05095](https://doi.org/10.1038/nature05095)
- Sillmann J, Croci-Maspoli M (2009) Present and future atmospheric blocking and its impact on European mean and extreme climate. *Geophys Res Lett* 36:L10702
- Sillmann J, Croci-Maspoli M, Kallache M, Katz RW (2011) Extreme cold winter temperatures in europe under the influence of North Atlantic atmospheric blocking. *J Clim* 24:5899–5913. doi:[10.1175/2011JCLI4075.1](https://doi.org/10.1175/2011JCLI4075.1)
- Sillmann J, Kharin VV, Zwiers FW, Zhang X, Branaugh D (2013a) Climate extremes indices in the CMIP5 multi model ensemble: Part 1. Model evaluation in the present climate. *J Geophys Res* 118:1716–1733
- Sillmann J, Kharin VV, Zwiers FW, Zhang X, Branaugh D (2013b) Climate extremes indices in the CMIP5 multi model ensemble: Part 2. Future climate projections. *J Geophys Res* 118:2473–2493
- Simmons AJ, Willett KM, Jones PD, Thorne PW, Dee DP (2010) Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets. *J Geophys Res* 115:D01110. doi:[10.1029/2009JD012442](https://doi.org/10.1029/2009JD012442)
- Smyth P, Ghil M, Ide K (1999) Multiple regimes in Northern Hemisphere height fields via mixture model clustering. *J Atmos Sci* 56:3704–3723
- Spath H (1985) Cluster dissection and analysis: theory, FORTRAN programs, examples. Horwood, 226 pp
- Stephanon M, D'Andrea F, Drobinski P (2012) Heatwave classification over Europe and the Mediterranean region. *Environ Res Lett* 7:014023. doi:[10.1088/1748-9326/7/1/014023](https://doi.org/10.1088/1748-9326/7/1/014023)
- Takaya K, Nakamura H (2001) A formulation of a phase-independent wave-activity flux for stationary and migratory quasigeostrophic eddies on a zonally varying basic flow. *J Atmos Sci* 58:608–627
- Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA (2006) Going to the extremes: an intercomparison of model-simulated historical and future changes in extreme events. *Clim Change* 79:185–211
- Teng H, Branstator G, Wang H, Meehl GA, Washington WM (2013) Probability of US heat waves affected by a subseasonal planetary wave pattern. *Nat Geosci* 6:1056–1061. doi:[10.1038/ngeo1988](https://doi.org/10.1038/ngeo1988)
- Thrasher B, Maurer EP, McKellar C, Duffy PB (2012) Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydroclim Earth Syst Sci* 16:3309–3314
- Turner JK, Gyakum JR (2011) The development of Arctic air masses in Northwest Canada and their behavior in a warming climate. *J Clim* 24:4618–4633
- Turner JK, Gyakum JR, Milrad SM (2013) A thermodynamic analysis of an intense North American arctic air mass. *Mon Weather Rev* 141:166–181
- van Oldenborgh GJ, Drijfhout S, van Ulden A, Haarsma R, Sterl A, Severijns C, Hazleger W, Dijkstra H (2009) Western Europe is warming much faster than expected. *Clim Past* 5:1–12
- van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK (2011) The representative concentration pathways: an overview. *Clim Change* 109:5–31. doi:[10.1007/s10584-011-0148-z](https://doi.org/10.1007/s10584-011-0148-z)
- Vavrus SJ, Walsh JE, Chapman WL, Portis D (2006) The behavior of extreme cold air outbreaks under greenhouse warming. *Int J Climatol* 26:1133
- Von Storch H (1999) On the use of “inflation” in statistical downscaling. *J Clim* 12:3505–3506
- Walsh JE, Philips AS, Portis DH, Chapman WL (2001) Extreme cold outbreaks in the United States and Europe, 1948–99. *J Clim* 14:2642–2658
- Wang A, Zeng X (2013) Development of global hourly 0.5° land surface air temperature datasets. *J Clim* 26:7676–7691. doi:[10.1175/JCLI-D-12-00682.1](https://doi.org/10.1175/JCLI-D-12-00682.1)
- Wehner MF (2004) Predicted 21st century changes in seasonal extreme precipitation events in the parallel climate model. *J Clim* 17:4281–4290
- Wehner MF (2010) Sources of uncertainty in the extreme value statistics of climate data. *Extremes* 13:205–217. doi:[10.1007/s10687-010-0105-7](https://doi.org/10.1007/s10687-010-0105-7)
- Westby RM, Lee YY, Black RX (2013) Anomalous temperature regimes during the cool season: long-term trends, low-frequency mode modulation, and representation in CMIP5 simulations. *J Clim* 26:9061–9076
- Wexler H (1936) Cooling in the lower atmosphere and the structure of polar continental air. *Mon Weather Rev* 64:122–136
- Wexler H (1937) Formation of polar anticyclones. *Mon Weather Rev* 65:229–236
- Wilks DS (2006) Statistical methods in the atmospheric sciences, 2nd edn. Academic Press, Waltham
- Wu Z, Lin HL, Li J, Jiang Z, Ma T (2012) Heat wave frequency variability over North America: two distinct leading modes. *J Geophys Res* 117:D02102. doi:[10.1029/2011JD016908](https://doi.org/10.1029/2011JD016908)

- Wuebbles D, Meehl G, Hayhoe K, Karl TR, Kunkel K, Santer B, Wehner M, Colle B, Fischer EM, Fu R, Goodman A, Janssen E, Lee H, Li W, Long LN, Olsen S, Seth A, Sheffield J, Sun L (2014) CMIP5 climate model analyses: climate extremes in the United States. *Bull Am Meteorol Soc* 95:571–583. doi:[10.1175/BAMS-D-12-00172.1](https://doi.org/10.1175/BAMS-D-12-00172.1)
- Zhang X, Alexander LV, Hegerl GC, Klein Tank A, Peterson TC, Trewin B, Zwiers FW (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip Rev Clim Change* 2:851–870. doi:[10.1002/wcc.147](https://doi.org/10.1002/wcc.147)
- Zwiers FW, Zhang X, Feng Y (2011) Anthropogenic Influence on long return period daily temperature extremes at regional scales. *J Clim* 24:881–892. doi:[10.1175/2010JCLI3908.1](https://doi.org/10.1175/2010JCLI3908.1)