North American Extreme Temperature Events and Related Large Scale Meteorological Patterns: A Review of Statistical Methods, Dynamics, Modeling, and Trends

Richard Grotjahn[1], Robert Black[2], Ruby Leung[3], Michael F. Wehner[4], Mathew Barlow[5], Mike Bosilovich[6], Alexander Gershunov[7], William J. Gutowski Jr.[8], John R. Gyakum[9], Richard W. Katz[10], Yun-Young Lee[1], Young-Kwon Lim[11], Prabhat[4]

Submitted to Climate Dynamics: 7 August 2014, Revised: 13 March 2015

_____

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

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

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

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

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

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

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

[8] Dept. of Geological and Atmospheric Sciences, Iowa State University, Ames, IA USA 50011

[9] Dept. of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada H3A 0B9

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

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

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.

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 (five-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

83 climate variability. The LSMP often has some portion that is superficially similar to a blocking
84 ridge, so a blocking index can be an LSMP indicator (Sillmann et al. 2011). Extreme events
85 have also been linked to circulation indices like the North Atlantic Oscillation (NAO) (Downton
86 and Miller 1993; Cellitti et al. 2006; Brown et al. 2008; Kenyon and Hegerl 2008; Guirguis et al.
87 2011) the Madden-Julian Oscillation (MJO) (Moon et al. 2011) and El Niño / Southern
88 Oscillation (ENSO) (Downton and Miller 1993; Higgins et al. 2002; Carrera et al. 2004; Meehl et
89 al. 2007; Goto-Maeda et al. 2008; Kenyon and Hegerl 2008; Alexander et al. 2009; Lim and
90 Schubert 2011). However, LSMPs are likely distinct from climate modes for several reasons.
91 First, named climate modes such as the NAO are common modes of variability, whereas the
92 LSMP is presumably as rare as the associated extreme event. Second, climate modes occur on
93 a longer time scale than the LSMPs for the short-term events focused upon in this article.
94 However, it is possible that an extreme event might occur when a climate mode has transient
95 extreme magnitude or is amplified in association with another low frequency phenomenon.
96 Third, tested LSMP patterns are not that similar to climate modes. In correlating eight NOAA
97 teleconnection patterns (http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml) and
98 California LSMPs and in assessing the PNA contribution to last winter's extreme cold in eastern
99 North America (neither shown here), we do not find notable contribution from such modes.
100 Several studies identified the LSMPs associated with specific extreme hot events (Grotjahn and
101 Faure 2008; Loikith and Broccoli 2012) and CAO events (Konrad 1996; Carrera et al. 2004;
102 Grotjahn and Faure 2008; Loikith and Broccoli 2012). Parts of these LSMPs tend to
103 be uniquely associated with the corresponding extreme weather and those parts have
104 some predictability (Grotjahn 2011).

105 The LSMPs for extreme events are not fully understood for different parts of North
106 America. Also, local processes: topography, soil moisture, etc. play key roles but there is a
107 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 (section 2), synoptic and dynamical interactions between LSMPs and other scales from local to global (section 3), model simulation issues (section 4), trends in these temperature extremes (section 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 twenty-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 detectible 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 1950-2007 in twenty-year return values of the hottest/coldest days and hottest/coldest nights. The temperature of very cold nights (figure 1d) exhibits pronounced warming over the entire continent as does the temperature of very cold days (figure 1b). The pattern of changes in the temperature of very hot days (figure 1a) and very hot nights (figure 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 U.S. (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.

2.2 Application of extreme value statistical techniques

The observed changes in figure 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 (GPD) are both three parameter functions and can be

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

185 Confidence in the estimates of the statistical properties of the tail of the parent
186 distribution of a random variable can be ascertained by exploring the sensitivity to the sample
187 size used to fit the extreme value distribution. For block maxima methods, the length of the
188 block is a season (or effectively so for temperature if the block length is a year). Lengthening the
189 block makes the sample size smaller; shortening it makes the sample size larger since only one
190 extreme value is drawn from each block. Such sensitivities are somewhat more straightforward
191 to explore with peaks over threshold (POT) methods. Typical thresholds may be chosen
192 between 80% and 99% depending on the size of the parent distribution that the extreme values
193 are drawn from. However, standard POT methods may not discriminate between extreme
194 values that occur at successive dates when the individual extreme values may not be truly
195 independent. In these cases, declustering techniques (Coles 2001) are applied to avoid biased
196 (low) estimates of the uncertainty. The trends in extreme temperature shown in figure 1 are
197 calculated using such a declustering technique and a POT formalism with time as a linear
198 covariate. An alternative approach retains possibly dependent consecutive extremes, adjusting
199 the estimates of uncertainty through either resampling or more advanced techniques for
200 quantifying extremal dependence (e.g., Fawcett and Walshaw 2012). Furthermore, LSMPs
201 responsible for extreme events can be formed using only the dates of the onset of the event.
202 (e.g. Grotjahn and Faure 2008; Bumbaco et al. 2013) reducing the risk of autocorrelated
203 extremes. In the cited studies, a 5 day gap was typically required between events. Low

204 frequency factors (or climate modes), such as ENSO, are best treated using the previously
205 mentioned covariate techniques.

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

222

223 2.3 Identification of LSMPs related to extreme temperatures

224

225 Several methods have been used to identify LSMPs that occur in association with extreme
226 temperature events. These methods and their properties are summarized in Table 2. The text
227 summarizes each method, its advantages and its disadvantages.

228

229 2.3.1 Composites

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

247 Compositing has several advantages. One can track LSMP formation by compositing
248 fields with respect to the onset time of each event. Meteorologically relevant full fields (or
249 anomalies) are obtained and composite analyses are constructed to obtain information on the
250 synoptic and dynamical time evolution. The method is non-parametric in that it does not make
251 any assumption about the pattern or the event statistics. Unlike some other methods, criteria

252 can be applied (typically a minimum waiting period between events) to ensure events are
253 independent. Significance can be assessed using a bootstrap resampling procedure where the
254 target ensemble value at a grid point is compared to the distribution of values at that grid point
255 found from a large number of 'random ensembles' (each of which uses the same number but
256 randomly-chosen dates). Values above (or below) a threshold of the random ensemble
257 distribution imply significantly high (low) values at that grid point. For example, a target
258 ensemble value equal or higher than the top 10 of 1000 random ensemble values at that point is
259 significant at approximately the 99% level. Figure 2 shows the LSMPs for California Central
260 Valley cold air outbreaks and heat waves obtained by this method, including bootstrap
261 resampling significance.

262 Compositing has some disadvantages that can be addressed. The identification of the
263 extreme event dates must be done separately and before the compositing. The composite
264 produces one target ensemble average (for a specific field and level) for each set of target
265 dates. If more than one LSMP can produce the extreme event, then that must be identified
266 either with one or more additional criteria when choosing dates or identified by examining
267 (perhaps qualitatively) the maps for each individual member of the target ensemble. A
268 procedure like adding the number of positive and subtracting the number of negative anomaly
269 values at a grid point in the target ensemble members (called the 'sign count' in Grotjahn 2011)
270 can assist with identifying multiple patterns. (If the sign count equals the number of ensemble
271 members, then all members have the same sign of the anomaly field at that location.) Lee and
272 Grotjahn (2015) apply a cluster analysis to distinctly different parts of LSMPs prior to California
273 heat waves and identify two ways the onset LSMPs form.

274

275 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 section 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

348 squared Euclidean point-to-centroid distance in a group (Spath 1985; Seber 2008), where each
349 centroid is the mean of the patterns in its cluster.

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

363 A disadvantage of clustering analysis is that one pre-specifies the number of clusters
364 (e.g. k in k-means clustering). Determining the number k is subjective if one does not have
365 sufficient prior knowledge of related physical patterns. There are several statistics to check the
366 optimal number of k such as 'distance of dissimilarity' (Stephanon et al. 2012). Another
367 disadvantage is the ambiguity of cluster assignment for certain events. Another disadvantage is
368 the patterns of such events cannot be assigned clearly to one group over another group. Part of
369 a pattern may resemble Cluster #1, while another part may be more similar to Cluster #2. To
370 avoid ascribing some marginal events to specific clusters, probabilistic clustering methods (e.g.
371 Smyth 1998) are developed, which suggests the possibility (e.g. percentage) that an event
372 could be assigned to each cluster rather than assigning it only to one cluster. If one increases

373 the cluster number, one can expect to decrease such ambiguity in classification. However, the
374 point of clustering analysis is to give a physical insight with minimal groups and not to interpret
375 every single episode.

376

377 2.3.5. Self-Organizing Maps

378

379 Self-Organizing Maps (SOMs; Kohonen 1995) are two-dimensional arrays of maps that display
380 characteristic behavior patterns of a field (e.g., Cavazos 2000; Hewitson and Crane 2006;
381 Gutowski et al. 2004; Cassano et al. 2007). The SOM array is a discretization of the continuous
382 pattern space occupied by the field examined. Thus, in contrast to clustering analysis, SOMs do
383 not assume a clumping together of patterns, though such behavior can emerge if present in the
384 input data. Figure 3 gives an example of a SOM array of synoptic weather patterns in sea level
385 pressure over a region centered on Alaska. Individual maps in the array represent nodes in a
386 projection of this continuous space onto a two-dimensional surface, with the size of the array
387 determined by the degree of spatial discretization of the SOM space one "feels" is needed for
388 the analysis at hand. The two dimensions show the two primary pattern transitions for the field
389 examined. Although one could, in principle, use more than two dimensions, typical practice in
390 climatological work has used only two.

391 The input maps themselves determine the degree and types of pattern transitions, hence
392 the "self-organizing" nature of the resulting array. The SOM node array is trained on a sequence
393 of input maps through an artificial neural net technique. The SOM array does not necessarily
394 favor the largest scales in the input data, but rather the scales most relevant to the field for the
395 domain and resolution examined. Consequently, SOMs can extract nonlinear pattern changes in
396 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 pre-determined 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

422 (number of nodes) one applies to the SOM space. Fine discretization will allow apparent
423 detection of small differences in how different data sets occupy pattern space, but fine
424 discretization will also render very noisy frequency distribution functions of the fields in SOM
425 space, thus undermining detection of any significant differences. Coarse discretization limits
426 the ability of the SOM procedure to resolve features producing the extreme events, so a further
427 disadvantage is that an insufficient array size may obscure grouping that may be present in the
428 data. The training method also requires specification of parameters that govern the training
429 process. A well-trained SOM is insensitive to these choices, but care is needed to ensure such
430 a result. In addition, like some of the other methods described here, the extreme events are
431 defined separately from the SOM analysis.

432

433 2.3.6. Machine learning and other advanced techniques

434

435 Looking to the future, we note that substantial progress has been made in the field of machine
436 learning for extracting patterns from Big Data. Commercial organizations such as Google and
437 Facebook rely on sophisticated, scalable analytics techniques for mining web-scale datasets.
438 Both supervised and unsupervised machine learning tools could play an important role in
439 extracting spatio-temporal patterns from climate datasets. The technique Deep Belief Networks
440 (Salakhutdinov and Hinton 2012) has been applied with tremendous success to classifying
441 objects in digital images (Krizhevsky et al. 2012) and speech recognition (Hinton et al. 2012).
442 These methods have substantially outperformed existing techniques in the field with the same
443 underlying learning algorithm. While these techniques have not yet been adapted for a
444 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 Nino / Southern Oscillation (ENSO) does not exert much influence on extremely high temperature magnitudes or duration, the North Atlantic

469 Oscillation (NAO) and atmospheric blocking do. However to date, such covariate techniques
470 relating atmospheric blocking to extreme temperatures have not been applied to North America.
471 Many physically based covariate quantities potentially offer insight into the mechanisms behind
472 extreme temperature events and the response to changes in the average climate. Good North
473 American candidates for covariates include indices measuring modes of natural variability such
474 as those describing the ENSO, the Pacific Decadal Oscillation (PDO), the NAO, the North
475 Atlantic Subtropical High (NASH), and various blocking indices.

476 The ETCCDI indices were designed for climate change detection and attribution
477 purposes rather than for exploring the mechanisms causing extreme events. They are not ideal
478 for connecting extreme temperature events to LSMPs and they are not descriptive of particularly
479 rare events. However, ETCCDI indices are designed to be robust over the observational record
480 and have been calculated and described for CMIP5 models by Sillmann et al. (2013a) (see
481 Table 1). ETCCDI indices are intended to be applied globally and be meaningful in areas of
482 sparse observations. The relatively dense network of North American observations since the
483 beginning of the 20th century permits the construction of more specialized LSMP extreme
484 indices linked to specific extreme events. Grotjahn (2011) defines an index that is an
485 unnormalized projection of key parts of a target ensemble LSMP onto a daily map of the
486 corresponding variable. (He combined such projections onto 850 hPa temperature and 700 hPa
487 meridional wind to form his 'circulation index'.) His target ensemble members are from dates
488 satisfying criteria listed in Table 3. The key parts of the fields used are those where all the
489 extreme events in the training period were consistent, at least in having the same anomaly sign.
490 Grotjahn (2011) found that extreme values of such an index (based on upper air data) occurred
491 on many of the same dates as extreme values of surface stations in the California Central
492 Valley (CCV). Statistically significant relationships exist between extreme values of this
493 circulation index and both the rate of CCV daily maximum temperature exceeding a high

494 threshold and the distribution of the excess over the threshold (Katz and Grotjahn 2014).
495 Grotjahn (2013) used such an index to show that a particular climate model was notably under-
496 predicting the occurrence frequency (by half) of CCV hot spells. Grotjahn (2014) used such an
497 index to show how that same model compared with a 55 year historical record and what the
498 model implied for CCV hot spells during the last half of the 21st century under two
499 representative concentration pathways (RCPs) of greenhouse gases.

500

501 **3. Large Scale Meteorological Patterns related to Extreme Temperature Events**

502

503 Intraseasonal extreme temperature events (ETEs) are almost always associated with regional
504 air mass excursions induced by circulation anomalies that are part of large-scale meteorological
505 patterns (LSMPs). LSMPs can include synoptic features (e.g., midlatitude cyclones; Konrad
506 1996) that enhance the ETE and often with scales similar to a teleconnection pattern (though
507 the nodes may not align and ETE onset, itself, may impact the teleconnection pattern, Cellitti et
508 al. 2006). In some cases, the LSMP is interpreted as a juxtaposition of teleconnection patterns
509 that leads to ETE events (Lim and Kim 2013). Heuristically, the role of LSMPs in producing
510 ETEs could be considered the result of either (1) a direct contribution to the large-scale
511 circulation that facilitates the air mass excursion or (2) the indirect modulation of sub-scale
512 variability, such as regional modulation of storm track behavior by blocking patterns. Besides
513 such dynamically driven impacts, there exist possible local impacts related to the interaction of
514 the LSMP with local topography or coastline features, leading to possible local symmetries in
515 the response pattern (e.g., Loikith and Broccoli 2012). Current knowledge of the remote forcing,
516 dynamics and local forcing of LSMPs associated with ETEs is summarized next.

517

518 3.1 Remote forcing of LSMPs and ETEs

519

520 3.1.1 Connection to low frequency modes of climate variability.

521

522 Numerous observational studies have ascertained that ETE behavior is modulated by recurring
523 large scale teleconnection patterns, particularly during winter. On intraseasonal time scales
524 there is a substantial modulation of North American ETEs during winter by the Pacific-North
525 American (PNA) pattern, North Atlantic (or Arctic) Oscillation (NAO or AO) and blocking patterns
526 (Walsh et al. 2001, Cellitti et al. 2006, Guirguis et al. 2011). On interannual and longer time
527 scales additional climate modes such as El Nino-Southern Oscillation (ENSO) and the Pacific
528 Decadal Oscillation (PDO) are also implicated (Westby et al. 2013). General relationships that
529 have emerged from these statistical analyses are illustrated in Fig. 4: The positive (negative)
530 phase of the NAO favors the occurrence of warm (cold) events over the eastern (southeastern)
531 United States. The positive (negative) phase of the PNA tends to favor cold events over the
532 southeastern (northwestern) US. These connections to climate modes are neither unique nor
533 independent. For example, the regional influence of the PNA pattern on ETEs largely mirrors
534 that of both the PDO and ENSO (Fig. 4) since the midlatitude atmospheric signatures of both
535 ENSO and the PDO project on the PNA pattern. Also, the prevalence of atmospheric blocking
536 patterns is intrinsically linked to particular climate mode phases (Renwick and Wallace 1996).

537 There have been pronounced episodes of climate modes influencing ETEs during recent
538 winters. Cold extremes over Europe and the southeastern United States during recent winters
539 (2009-10 and 2010-11) were primarily accounted for by the anomalous blocking associated with
540 persistent episodes of large amplitude negative phase of the NAO (Guirguis et al. 2011). There
541 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

566 principle, Arctic amplification may lead to regional alterations in the structure of storm tracks, jet
567 streams and planetary waves. In recent years, considerable attention has been focused on the
568 role of Arctic amplification-induced changes to the jet stream (Francis and Vavrus 2012, Liu et
569 al. 2012, Barnes 2013, Screen and Simmonds 2013). Francis and Vavrus (2012) found that a
570 weaker zonal flow (i.e., polar jet) from weakened meridional temperature gradient slows the
571 eastward Rossby wave progression and tends to create larger meridional excursions of height
572 contours and associated temperature displacements resulting in a higher probability of extreme
573 weather. In a similar vein, Liu et al. (2012) argue that the circulation change due to the decline
574 of Arctic sea ice leads to more frequent events of atmospheric blocking that cause severe cold
575 surges over large parts of northern continents. Francis et al. (2009), Overland and Wang (2010),
576 Jaiser et al. (2012) and Lim et al. (2012) found that there is a delayed atmospheric response to
577 the Arctic sea ice. Specifically, the Arctic sea ice extent in summer to fall influences the
578 atmospheric circulation in the following winter over the northern mid- to high-latitudes, affecting
579 the seasonal winter temperature and subseasonal warm/cold spells. Evidence presented in
580 more recent studies (Barnes 2013, Screen and Simmonds 2013), however, suggests that the
581 role of the mechanism put forth by Francis and Vavrus (2012) is uncertain at best. More
582 generally, Cohen et al. (2014) conclude that our understanding of the mechanistic link between
583 ongoing Arctic amplification and mid-latitude extreme weather is currently limited by
584 shortcomings in relevant data records, physical models and dynamical understanding, itself. As
585 such, the likely future impact of Arctic amplification upon extreme weather is highly uncertain.

586 Arctic amplification is also linked to long-term variability in high latitude snow cover. An
587 analogous linkage between autumnal variability in Eurasian snow cover and wintertime ETE
588 events over North America has been noted (Cohen and Jones 2011). In this case, autumnal
589 snow cover anomalies induce a subsequent weakening of the stratospheric polar vortex during
590 winter which, in turn, leads to a persistent negative phase episode of the tropospheric AO

591 favoring North American cold events (Cohen et al. 2007). In addition, there is evidence that the
592 changing Arctic sea ice extent may be linked to changes in the autumnal advance in Eurasian
593 snow cover (Cohen et al. 2013). As for Arctic amplification, however, there is considerable
594 uncertainty regarding the statistical robustness and physical nature of the Eurasian snow cover
595 influence upon mid-latitude extreme weather (Peings et al. 2013, Cohen et al. 2014).

596

597 3.1.3 Large-scale climate "markers" for climate model assessment

598

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

611 The Coupled Model Intercomparison Project (CMIP) provides an ideal resource for
612 assessing the ability of modern coupled climate models to represent the behavior of climate
613 modes. A recent analysis of CMIP5 models indicates that while most models studied perform
614 well in representing the basic aspects of the PNA pattern, a small subset of models have

615 difficulty qualitatively replicating the NAO pattern (Lee and Black 2013; Table 4). Otherwise
616 differences among model patterns consist of horizontal shifts or amplitude variations in the
617 circulation anomaly pattern features. CMIP5 models generally underestimate the regional
618 frequency of winter blocking events while summertime blocking events occurring over the high
619 latitude oceanic basins are typically overestimated (Masato et al. 2013). Conversely, Westby et
620 al (2013) found serious deficiencies in the representation of the PDO by CMIP5 models with
621 direct impacts upon the modulation of anomalous temperature regimes.

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

633 1) The tropical large-scale circulation response to tropical diabatic heating

634 2) Upper tropospheric meridional vorticity gradients in the subtropics

635 3) Barotropic deformation structures in the jet exit regions

636 4) The structure and intraseasonal variability of the Atlantic and Pacific storm tracks

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

640

641 3.2: Dynamics of LSMPs

642

643 3.2.1 Diagnostic tools to study dynamics of LSMP onset/decay

644

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

654 Most ETE events are associated with lateral air mass excursions, which are induced by
655 large-scale circulation anomalies in the lower troposphere forming the LSMP (e.g., Loikith and
656 Broccoli 2012). The anomalous circulation serves as a “dynamical trigger” for ETE events. The
657 delineation of LSMP dynamics is treated as a two-stage process: First it is of interest to assess
658 whether the main energy source(s) are local or remote. Once the energy source *location* is
659 determined, the second stage is to assess the specific *physical mechanism* providing the
660 proximate energy source in this location. An effective means for assessing the source location

661 regions for large-scale atmospheric waves is the application of wave activity flux analyses
662 (Plumb 1985, Takaya and Nakamura 2001). The wave activity flux is parallel to Rossby wave
663 group velocity and traces out a three-dimensional pathway between a wave source region (of
664 flux divergence) and a wave sink region (of flux convergence).

665 Possible primary physical mechanisms providing wave activity sources in this context
666 include large-scale barotropic growth, baroclinic growth/instability and nonlinear forcing by
667 synoptic-scale eddies (Evans and Black 2003). These mechanisms may be augmented via
668 secondary feedbacks related to internal diabatic processes or interactions of the LSMP with the
669 local topography or land surface. Past studies have introduced comprehensive dynamical
670 frameworks for studying the dynamical mechanisms leading to the growth and decay of large-
671 scale circulation anomalies (e.g., Feldstein 2002, 2003). These are based upon a local analysis
672 of tendencies in either the geopotential (Evans and Black 2003) or streamfunction (Feldstein
673 2002) fields. In these studies local tendencies are decomposed into separate forcing terms that
674 can be related to distinct physical processes. The two-stage process outlined above is a
675 generally useful means for the dynamical diagnosis of LSMP life cycles in both observations
676 and climate model simulations.

677

678 3.2.2 North American Arctic air mass formation

679

680 Both dynamics and thermodynamics have roles in the formation of extreme cold-air masses.
681 Wexler (1936, 1937) postulated that the cold air was formed at the surface by radiative cooling
682 from a snow-covered ground under clear, windless conditions, creating an intense temperature
683 inversion restricted to a very shallow layer at the surface. Finding that the Wexler model could
684 not adequately explain the depth of the cold layer observed in soundings, Gotaas and Benson

685 (1965) showed that the existence of suspended ice crystals were crucial to upper-level cooling.
686 Curry (1983) later modeled the effect with the introduction of condensate, particularly ice-
687 crystals, in the layer. In an experiment without moisture, the inversion that formed after 2 weeks
688 of radiative cooling was less than 1000 meters deep. Only with the addition of ice crystals did
689 the inversion rise above 900 hPa. More recently, Emanuel (2009) found that the rates and depth
690 of cooling in his model were sensitive to the amount of water vapor and clouds present.

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

702 Figure 5 (Turner et al. 2013), illustrating the LSMPs of a representative case of arctic air
703 mass formation in January-February 1979, shows surface ridging building southeastward along
704 the eastern slopes of the Rockies. The authors argue that diabatic cooling is needed to explain
705 the cooling observed. This cooling may consist of a combination of sublimational cooling from
706 precipitation falling into a dry layer, and/or radiational cooling from suspended ice crystals.

707

708 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 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 figure 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

734 the Rocky Mountains (Grotjahn and Faure 2008). As discussed in Grotjahn and Faure, the large
735 scale pattern is similar to outbreaks affecting the areas east of the Rockies however with a
736 primary difference being a small but statistically significant ridge over the southeastern US. (See
737 figure 2a,b.) This ridge precedes the CAO by several days, including the large ridge that
738 develops subsequently over Alaska. A cold air mass between these ridges is directed further
739 westward (a more northerly or northeasterly flow of the cold air) than for eastern CAOs. The
740 distortion of the flow by that southeastern ridge enables the cold air mass to cross the Rockies.
741 Loikith and Broccoli (2012) capture a portion of this pattern in geopotential heights at 500 hPa
742 using their 'grand composite' technique. More local to the area of the extreme, colder
743 temperatures are associated with an adjacent surface high (e.g. the conceptual model of Colle
744 and Mass 1995). Favre and Gershunov (2006) pursue this link to anticyclones affecting western
745 North America; and develop indices based on frequency and central pressure of transient
746 cyclones and anticyclones in the eastern North Pacific. Their 'CA' index is the difference,
747 cyclone minus anticyclone, in strength and frequency. Negative CA values are correlated with
748 the coldest 10% of winter minimum temperatures.

749 *Eastern US events (i.e., east of Rockies).* Several statistical and synoptic studies either
750 directly or indirectly relate to the topic of US CAOs occurring east of the Rockies (e.g., Konrad
751 and Colucci 1989, Colle and Mass 1995, Konrad 1996, Walsh et al 2001, Cellitti et al 2006,
752 Portis et al 2006). Common precursors to eastern CAO events include anomalously high
753 surface air pressure over western Canada (linked to a polar air mass), occurrences of the
754 negative (positive) phase of the NAO (PNA) teleconnection pattern and/or an anomalously weak
755 stratospheric polar vortex. CAO onset is characterized by the south/southeastward advection of
756 the polar air mass in association with one or more of the following synoptic LSMP features:
757 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.

781 3.2.4 Summer heat wave events

782

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

792 *West coast events.* Heat waves affecting the west coast are linked to an upper air LSMP
793 that has a local ridge. For California Central Valley heat waves, that ridge is typically aligned
794 with the west coast and the LSMP also has upstream features: a trough south of the Gulf of
795 Alaska and a ridge further west, south of the Aleutians (Grotjahn and Faure 2008). See figure
796 2d. The local ridge is easily understood as resulting from high thickness due to the anomalous
797 high temperatures through the depth of the troposphere generally centered along the west coast
798 of the US. (Plots can be found here:
799 http://grotjahn.ucdavis.edu/EWEs/heat_wave/heat_wave.htm) Grotjahn and Faure also show
800 that a statistically significant ridge in the northwest Pacific develops prior to the significant
801 intensification of that west coast ridge. These features are essentially equivalent-barotropic
802 through the depth of the troposphere. While Grotjahn and Faure found variation amongst the
803 events, the ensemble average consists of a significant temperature anomaly that in the lower
804 troposphere (850 and 700 hPa) is located just off the Oregon and northern California coast and
805 elongated meridionally. The narrow zonal with longer meridional scale is consistent with

Bachmann (2008). Bachmann showed that extreme surface temperature dates in Sacramento were more frequently matched by corresponding extreme dates at places far to the north (e.g. Seattle) than with places across the Sierra Nevada or Cascade ranges to the east (e.g. much closer Reno). 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, 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 figure 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 2014); 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, 2014), 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 (2014), 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 Tibaldi 2004, Lau and Nath 2012, Loikith and Broccoli 2012, and 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, Lokith 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

855 predictability (Teng et al. 2013). Teng et al. (2013) have noted the similarity of the circulation
856 pattern to a Rossby wave number 5 wave train in the jet stream waveguide. The commonality
857 of the wave train pattern structure and evolution among these studies is difficult to assess,
858 though.

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

867 *Eastern events.* While the large-scale meteorological patterns associated with heat
868 waves in the eastern US have not yet seen much study, they have been examined for the
869 Northeast and Central US Gulf Coast in Lau and Nath (2012) and for piedmont North Carolina
870 by Chen and Konrad (2006). For both the Northeast and Gulf Coast, Lau and Nath found a mid-
871 level ridge centered over the region as part of a wave train and precipitation deficits, broadly
872 similar to the results for the Midwestern regions. For piedmont North Carolina, Chen and
873 Konrad showed that, at upper levels, a strong ridge over or just upstream of the region was a
874 common feature along with, at lower levels, adiabatic warming associated with descent from the
875 Blue Ridge Mountains into the piedmont region.

876

877 **4. Modeling of temperature extremes and associated circulations**

878

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

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

899

900 4.1 Global and regional climate model skill in simulating temperature extremes

901

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

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

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 equation 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 Figure 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 figure 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 figure 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 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 twenty 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 20th 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 (figure 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,

949 in good agreement with Lim and Schubert (2011). Comparison of WWs and CAOs between the
950 CMIP5 models and observation indicated that the models tend to overestimate WW frequency,
951 but underestimate CAOs frequency.

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

965 Heat waves are associated with anomalous large-scale circulation patterns well resolved
966 by GCMs. Grotjahn (2013, 2014) describe how the LSMPs for California hot spells simulated by
967 a CMIP5 climate model compare with the observed LSMPs; while the model generates a similar
968 large scale pattern and included the needed large scale sinking, the model is unable to simulate
969 sea breezes so the effect of LSMP blocking of sea breezes during extreme heat episodes is
970 likely not captured by the global model. In general, heat wave intensity and duration can vary at
971 local to regional scales because surface temperature is influenced by processes such as land
972 surface fluxes, turbulence and winds, and clouds and radiation. For example, land-atmosphere
973 interactions were found to play an important role in European summer heat waves of the past

974 decades (Fischer et al. 2007). Using a regional model, Jaeger and Seneviratne (2011) and
975 Lorenz et al. (2010) highlighted the key role of soil moisture memory on heat wave intensity and
976 persistence, respectively. Using observations and a CMIP3 model, Gershunov and Douville
977 (2008) showed a strong negative relationship between preceding winter and spring precipitation
978 and summertime heat in the central, Midwestern, and eastern US. In regions with limited soil
979 moisture, surface temperature can respond more strongly to the large-scale circulation anomaly
980 that leads to clear sky and strong surface solar heating because surface cooling from
981 evapotranspiration is limited. As soil moisture is influenced by antecedent precipitation and
982 surface properties such as land cover, topography, and soil characteristics that vary at local to
983 regional scales, heat wave intensity and duration can vary similarly, within the context of large-
984 scale influence by a circulation anomaly. Thus the ability of climate models to simulate and
985 predict heat waves is also dependent on physics parameterizations and grid resolution.

986 Lau and Nath (2012) analyzed GCM simulations performed at 200km and 50km grid
987 resolutions and found considerable resemblance between the temperature anomaly patterns
988 from the coarse and fine resolution simulations, except for some local details generated by the
989 higher resolution model. Both models were able to capture realistic heat wave intensity,
990 duration, and frequency in various key regions of North America and the synoptic features
991 accompanying the warm episodes as revealed in the North American Regional Reanalysis.
992 Kunkel et al. (2010) compared regional climate simulations performed at 30km resolution over
993 North America driven by two GCM historical simulations with the respective GCM simulations
994 and observations. Although the two GCM simulations have opposite temperature biases,
995 downscaling by the RCM provided improved simulations of heat waves in both cases, with the
996 overall averaged biases reduced by a factor of two compared to the GCMs. Gao et al. (2012)
997 compared a regional simulation at 4km grid resolution over the eastern U.S. driven by a GCM
998 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 U.S. 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 Section 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 U.S. 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 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,

1094 especially maximum and minimum temperatures that are important to extreme temperature
1095 events. Figure 10 shows the U.S. summertime 2011 Warm Spell Duration Index (WSDI, e.g.
1096 Alexander et al. 2006) computed from the occurrences of daily mean temperature exceeding the
1097 90th percentile for 6 days or longer from MERRA. This diagnostic relies on continuous data to
1098 derive the seasonal index, but also identify the spatial extent. Figure 11 shows the US trends in
1099 WSDI from MERRA. Positive trends are apparent in many areas of the US with some regions
1100 having statistical significance (above 90% confidence).

1101 Walsh et al. (2001) used global reanalysis data to study cold air outbreaks and identified
1102 Midwest, Gulf Coast, and East Coast as regions influenced by CAOs. They did not identify any
1103 apparent trend in CAOs in North America between 1948 and1999, although analysis from a
1104 single station suggests that extreme outbreaks may have been more frequent in the late 1800s
1105 and early 1900s. Extending the study to include the recent decade from the global reanalysis
1106 data may provide further confidence in potential trends in CAOs in North America.

1107 The Twentieth Century Reanalysis (Campo et al. 2006; 2011) provides objectively
1108 analyzed four-dimensional data on a 2^{o} global grid. With assimilation of synoptic surface
1109 pressure available since the late 19th century, this dataset covers the period from 1871 – 2012
1110 and offers a unique opportunity for analysis of long-term trend in temperature extremes and the
1111 associated LSMP. Campo et al. (2011) showed encouraging comparison of this surface-based
1112 reanalysis product with other reanalyses that make use of upper-air and satellite data. Future
1113 work utilizing this dataset could potentially further advance understanding of variability and
1114 trends in LSMPs that influence temperature extremes.

1115

1116 5.2 Projected trends

1117

1118 The CMIP3 and CMIP5 multimodel ensembles have been used to investigate projected
1119 changes of temperature extremes for the next couple of decades (e.g., mid-21st century) (e.g.,
1120 Tebaldi et al. 2006; Orlowsky and Seneviratne 2012; Sillmann et al. 2013b; Wuebbles et al.
1121 2014). Climate change simulations in the CMIP5 multimodel ensembles showed greater
1122 changes in ETCCDI based on daily minimum temperatures than in ETCCDI based on daily
1123 maximum temperatures. Also, the strong emission scenario such as Representative
1124 Concentration Pathways (RCPs) 8.5 (Moss et al. 2010; van Vuuren et al. 2011) produced more
1125 definitive changes in temperature extreme statistics than weaker emission scenarios (e.g., RCP
1126 2.6 and 4.5). Furthermore, the changes in ETCCDI under RCP8.5 tend to be greater than
1127 changes under any of the scenarios in the Special Report on Emission Scenarios (SRES)
1128 (Nakicenovic et al. 2000) used in CMIP3. The spatial distributions of Sillmann et al. (2013b)
1129 showed that the northern high latitude regions have the strongest increase in the minimum daily
1130 minimum temperature, while changes in the maximum daily maximum temperatures tend to be
1131 evenly distributed globally. The study also noted that the percentile indices such as warm and
1132 cold nights, exhibit the highest increases in the tropical regions. This is a result of differences in
1133 interannual temperature variability between the two regions, which is relatively large in mid- to
1134 high latitudes but small in the low-latitudes.

1135 Figure 12 shows the projected ensemble mean increase in two of the ETCCDI indices
1136 assessed in figure 8 at the end of this century (2080-2100) from the reference period of 1985-
1137 2005 under the RCP8.5 forcing scenario of the 5th Assessment Report of the Intergovernmental
1138 Panel on Climate Change (IPCC AR5). Twenty-six models are included in this multi-model
1139 average. Models marked with an "*" in figure 8 are not included in this projection due to a lack of
1140 future simulations. The two indices, cold nights (Tnn) and hot days (Txx), are chosen to illustrate
1141 the seasonal differences of changes at the extremes of both ends of the distribution of daily
1142 surface air temperatures. The upper left panel shows that the wintertime increase in cold nights

1143 has large changes northward of the middle of the U.S. largely corresponding to present day
1144 winter snow covered areas. Kodra et al. (2011) studied cold extremes in the 21st century using
1145 nine GCMs and found that despite future warming, extreme cold events, both in terms of
1146 intensity and duration, can persist in the future. Cool nights in the summer, shown in the upper
1147 right panel, do not exhibit such a monotonic poleward behavior. Rather, the continental interior
1148 warms more than the edges for this measure of extreme temperature. Wintertime increases in
1149 warm days, shown in the lower left panel, are projected to be significantly less than wintertime
1150 cold nights; although some of the poleward gradient property is present. Summertime increases
1151 in hot days are projected slightly warmer than but generally similar to summertime increases in
1152 cool nights. One noticeable difference is that future extreme temperatures in the Mexican
1153 interior warm more in the summer than in the winter.

1154 Seasonality in the projected changes of these two indices is a reflection of the complex
1155 mechanisms affecting changes in extreme and mean temperatures in North America. Changes
1156 in snow cover certainly affect the future winter mean and extreme temperatures. Likewise,
1157 decreases in soil moisture affect the future summer mean and extreme temperatures. The
1158 additional possibility of biases and changes in LSMPs affecting extreme temperatures must be
1159 considered in light of these and other mechanisms. For example, van Oldenborgh et al. (2009)
1160 found that changes in the large-scale circulation, including a shift towards a more westerly
1161 circulation and the North Atlantic current need to be better simulated especially in winter and
1162 spring for more realistic simulations of warming over Western Europe in recent decades.
1163 Gershunov and Guirguis (2012) noted that three out of four GCMs analyzed did not adequately
1164 capture the synoptic causes of California heat waves. When sub-regional heat waves are
1165 defined relative to the changing local climate (a ‘non-stationary’ perspective), coastal heat
1166 waves have greater impact than those inland even though the summertime average warming is
1167 stronger inland. Sillmann and Croci-Maspoli (2009) found that European blocking events

1168 influence particularly the winter cold extremes in Europe. For future projection, decrease in
1169 blocking frequency with increase in maximum blocking duration was simulated under the A1B
1170 scenario. Also, the blocking pattern shifts northeastward, affecting a larger part of Europe by
1171 giving rise to anomalously cold winter months.

1172 To partly address the question of whether changes in heat waves in the future are
1173 associated with a shift in the daily maximum temperature or changes in LSMP, Lau and Nath
1174 (2012) noted that the probability distribution functions (PDFs) for the current and future climate
1175 over most regions that experience heat waves in the U.S. have similar shape but were shifted
1176 by the mean warming in the daily maximum temperature, and the magnitude of skewness of the
1177 PDFs exhibited only minor changes from the current to future climate. Further statistical
1178 manipulations of the model outputs suggest that the increase in heat wave intensity and
1179 frequency in the future is primarily associated with a shift in the daily maximum surface
1180 temperature, which suggests that changes in the characteristics of LSMPs associated with heat
1181 waves may be minor. Grotjahn (2014) shows trends in an LSMP index for California hot spells in
1182 reanalysis and both historical and future simulations by a climate model. Future changes in heat
1183 waves may also be modulated by changes in land-atmosphere interactions. For example, global
1184 warming may shift the climate regime northward and establish new transitional zones with
1185 strong land-atmosphere coupling strength (Seneviratne et al. 2006) and drying over the
1186 subtropics and mid-latitudes during summer in the future may increase land-atmosphere
1187 feedbacks in general (Dirmeyer et al. 2012) and result in amplified heat wave response to global
1188 warming.

1189 Regional climate models have also been used to investigate temperature extreme
1190 changes in the future. Kunkel et al. (2010) showed that the regional model has superior skill in
1191 simulating heat waves compared to the global models, but it largely inherited the climate
1192 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 U.S. 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 U.S. 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

1312 Research is sponsored by the National Science Foundation.

1313

1314 References

1315 Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M,
1316 Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths
1317 G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P,
1318 Rusticucci M, Vazquez-Aguirre JL (2006) Global observed changes in daily climate
1319 extremes of temperature and precipitation. J. Geophys. Res., 111, D05109,
1320 doi:10.1029/2005JD006290

1321 Alexander L, Uotila P, Nicholls N. (2009) Influence of sea surface temperature variability on
1322 global temperature and precipitation extremes, J Geophys Res 114:D18116

1323 Bachmann B (2008) The spatial extent of California heat waves. MS Thesis. Atmospheric
1324 Science Program, University of California Davis. March. 164 pp. (pdf available at:
1325 http://grotjahn.ucdavis.edu/EWEs/)

1326 Barnes EA (2013) Revisiting the evidence linking Arctic amplification to extreme weather in
1327 midlatitudes. Geophys. Res. Lett., 40, 4734–4739, doi:10.1002/grl.50880

1328 Benniston M (2004) The 2003 heat wave in Europe: a shape of things to come? An analysis
1329 based on Swiss climatological data and model simulations. Geophys Res Lett
1330 31:L02202. doi:10.1029/2003GL018857

1331 Black RX, Evans KJ (1998) The statistics and horizontal structure of anomalous weather
1332 regimes in the Community Climate Model, Mon.Weather Rev., 126(4), 841–859

1333 Beniston M (2004) The 2003 heat wave in Europe: A shape of things to come? An analysis
1334 based on Swiss climatological data and model simulations. Geophys. Res. Lett. 31:
1335 L02202, doi:10.1029/2003GL018857

1336 Bosilovich, MG (2013) Regional Climate and Variability of NASA MERRA and Recent
1337 Reanalyses: U.S. Summertime Precipitation and Temperature. *J. Appl. Meteor.*
1338 *Climatol.*, **52**, 1939–1951. doi: http://dx.doi.org/10.1175/JAMC-D-12-0291.1

1339 Brewer MC, Mass CF, Potter BE (2012) The West Coast Thermal Trough: Climatology and
1340 Synoptic Evolution. *Mon. Wea. Rev.*, 140: 3820–3843
1341 doi: http://dx.doi.org/10.1175/MWR-D-12-00078.1

1342 Brown S, Caesar J, Ferro C (2008), Global changes in extreme daily temperature since 1950, J
1343 Geophys Res 113: D05115, doi:10.1029/2006JD008091.

1344 Bumbaco KA, Dello KD, Bond NA (2013) History of Pacific Northwest Heat Waves: Synoptic
1345 Pattern and Trends. J. Appl. Meteor. Climatol., 52, 1618–1631
1346 doi: http://dx.doi.org/10.1175/JAMC-D-12-094.1

1347 Carrera M, Higgins R, Kousky V (2004) Downstream weather impacts associated with
1348 atmospheric blocking over the Northeast Pacific J Climate 17:4823-4839

1349 Cassano EN, Lynch AH, Cassano JJ, Koslow MR (2006) Classification of synoptic patterns in
1350 the western Arctic associated with extreme events at Barrow, Alaska. Clim. Res. 30: 83-
1351 97

1352 Cassano JJ, Uotila P, Lynch AH, Cassano EN (2007) Predicted changes in synoptic forcing of
1353 net precipitation in large Arctic river basins during the 21st century. J. Geophys. Res.
1354 112: G04S49 doi:10.1029/2006JG000332

1355 Cavazos T (2000) Using Self-Organizing Maps to investigate extreme climate events: An
1356 application to wintertime precipitation in the Balkans. J. Climate 13: 1718–1732

1357 Cellitti M, Walsh J, Rauber R, Portis D (2006), Extreme cold air outbreaks over the United
1358 States, the polar vortex, and the large-scale circulation, J. Geophys Res 111: D02114,
1359 doi:10.1029/2005JD006273

1360 Chang F-C, Wallace JM (1987) Meteorological conditions during heat waves and droughts in
1361 the United States Great Plains. Mon. Wea. Rev., 115, 1253–1269

1362 Chen F, CE Konrad II (2006) A Synoptic Climatology of Summertime Heat and Humidity in the
1363 Piedmont Region of North Carolina. J. Appl Meteor Climatol 45:5, 674-685

1364 Christidis N, Stott PA, Brown S, Hegerl GC, Caesar J (2005) Detection of changes in
1365 temperature extremes during the second half of the 20th century, Geophys. Res. Lett.,
1366 32, L20716, doi:10.1029/ 2005GL023885

1367 Cohen J, Jones J (2011) A new index for more accurate winter predictions. Geophysical
1368 Research Letters 38, L21701, http://dx.doi.org/10.1029/2011GL049626

1369 Cohen J, Barlow M, Kushner P, Saito K (2007) Stratosphere-troposphere coupling and links
1370 with Eurasian land-surface variability. Journal of Climate 20:5,335–5,343,
1371 http://dx.doi.org/10.1175/2007JCLI1725.1

1372 Cohen J, Jones J, Furtado JC, Tziperman E (2013) Warm Arctic, cold continents: A common
1373 pattern related to Arctic sea ice melt, snow advance, and extreme winter weather.
1374 Oceanography 26(4), http://dx.doi.org/10.5670/oceanog.2013.70

1375 Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K,
1376 Entekhabi D, Overland J, Jones J (2014) Recent Arctic amplification and extreme mid-
1377 latitude weather. Nature Geosci., 7, 627–637, doi:10.1038/ngeo2234

1378 Coles S (2001) An Introduction to Statistical Modeling of Extreme Values, Springer, London,
1379 208pp. ISBN: 978-1-84996-874-4 (Print) 978-1-4471-3675-0 (Online)

1380 Colle BA, Mass CF (1995) The Structure and Evolution of Cold Surges East of the Rocky
1381 Mountains. *Mon. Wea. Rev.*, **123**: 2577–2610. doi: http://dx.doi.org/10.1175/1520-
1382 0493(1995)123<2577:TSAEOC>2.0.CO;2

1383 Compo, GP, JS Whitaker, and PD Sardeshmukh, (2006) Feasibility of a 100 year reanalysis
1384 using only surface pressure data. Bull. Amer. Met. Soc., 87: 175-190.

1385 Compo, GP, JS Whitaker, PD Sardeshmukh, N Matsui, RJ Allan, X Yin, BE Gleason, RS Vose,
1386 G Rutledge, P Bessemoulin, S Brönnimann, M Brunet, RI Crouthamel, AN Grant, PY
1387 Groisman, PD Jones, MC Kruk, AC Kruger, GJ Marshall, M Maugeri, HY Mok, Ø Nordli,
1388 TF Ross, RM Trigo, XL Wang, SD Woodruff, and SJ Worley (2011) The Twentieth
1389 Century Reanalysis Project. Quart. J. Roy. Meteorol. Soc., 137: 1–28.
1390 doi: 10.1002/qj.776

1391 Curry J (1983) On the Formation of Continental Polar Air. J. Atmos. Sci., 40, 2278-2292

1392 Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda
1393 MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann
1394 N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H,
1395 Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM,
1396 Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F
1397 (2011), The ERA-Interim reanalysis: configuration and performance of the data
1398 assimilation system, Quart. J. Roy. Meteor.Soc., 137, 553–597

1399 Diffenbaugh, NS, Ashfaq M (2010) Intensification of Hot Extremes in the United States.
1400 Geophys. Res. Lett., 37(15), DOI: 10.1029/2010GL043888

1401 Dirmeyer PA, Cash BA, Kinter JL III, Stan C, Jung T, Marx L, Towers P, Wedi N, Adams JM,
1402 Altshuler EL, Huang B, Jin EK, Manganello J (2012) Evidence for Enhanced Land-
1403 Atmsophere Feedback in a Warming Climate. J. Hydrometeor., 13(3), 981-995

1404 Dole R, Hoerling M, Kumar A, Eischeid J, Perlwitz J, Quan X-W, Kiladis G, Webb R, Murray D,
1405 Chen M, Wolter K, and Zhang T (2014) The Making of an Extreme Event: Putting the
1406 Pieces Together. Bull. Amer. Meteor. Soc., 95, 427–440. DOI:
1407 http://dx.doi.org/10.1175/BAMS-D-12-00069.1

1408

1409 Downton M, Miller K (1993) The freeze risk to Florida citrus. Part II: Temperature variability and
1410 circulation patterns, J. Clim 6:364–372

1411 Easterling D (1999) Development of regional climate scenarios using a downscaling approach.
1412 *Climatic Change*, **41**, 615-634

1413 Emanuel K (2009) Back to Norway: An Essay. *Synoptic–Dynamic Meteorology and Weather*
1414 *Analysis Forecasting,* Meteor Monogr, No. 55, Amer. Meteor. Soc., 87–96

1415 Evans KJ, Black RX (2003) Piecewise Tendency Diagnosis of Weather Regime Transitions. J.
1416 Atmos. Sci., 60, 1941–1959

1417 Favre A, Gershunov A (2006) Extra-tropical cyclonic/anticyclonic activity in North-Eastern
1418 Pacific and air temperature extremes in Western North America. Clim. Dyn. 26: 617–
1419 629. DOI 10.1007/s00382-005-0101-9

1420 Fawcett L, Walshaw D (2012) Estimating return levels from serially dependent extremes.
1421 Environmetrics 23:272-283

1422 Feldstein SB (2000) Teleconnections and ENSO: The timescale, power spectra, and climate
1423 noise properties. J Clim:13, 4430-4440

1424 Feldstein SB (2002) Fundamental mechanisms of PNA teleconnection pattern growth and
1425 decay. Quart. J. Roy. Meteorol. Soc., 128, 775-796

1426 Feldstein SB (2003) The dynamics of NAO teleconnection pattern growth and decay. Quart. J.
1427 Roy. Meteorol. Soc., 129, 901-924

1428 Fischer EM, Seneviratne SI, Lüthi D, Schär C (2007) Contribution of Land-Atmosphere Coupling
1429 to Recent European Summer Heat Waves, Geophys. Res. Lett., 34, L06707,
1430 doi:10.1029/2006GL029068

1431 Forbes GS, Anthes RA, Thomson DW (1987) Synoptic and Mesoscale Aspects of an
1432 Appalachian Ice Storm Associated with Cold-Air Damming. Mon. Wea. Rev., 115**:** 564-
1433 591

1434 Francis JA, Chan W, Leathers DJ, Miller JR, Veron DE (2009) Winter northern hemisphere
1435 weather patterns remember summer Arctic sea-ice extent. Geophys Res Lett
1436 36:L07505. doi: 10.1029/2009GL037274

1437 Francis JA, Vavrus SJ (2012) Evidence linking Arctic amplification to extreme weather in mid-
1438 latitudes. Geophys. Res. Lett., 39(6), doi: 10.1029/2012GL051000

1439 Franzke C, Feldstein SB, Lee S (2011) Synoptic analysis of the Pacific–North American
1440 teleconnection pattern. Q. J. R. Meteorol. Soc. 137: 329–346. DOI:10.1002/qj.768

1441 Fritsch JM, Kapolka J, Hirschberg PA (1992) The Effects of Subcloud-Layer Diabatic Processes
1442 on Cold Air Damming. *J. Atmos. Sci.*, 49**,** 49-70

1443 Furrer EM, Katz RW, Walter MD, Furrer R (2010) Statistical modeling of hot spells and heat
1444 waves. Climate Research 43:191-205

1445 Gao Y, JS Fu, JB Drake, Y Liu, and J-F Lamarque (2012) Projected Changes of Extreme
1446 Weather Events in the Eastern United States Based on a High Resolution Climate
1447 Modeling System. Environ. Res. Lett., doi:10.1088/1748-9326/7/4/044025

1448 Gao, Y, Leung LR, Lu J, Masato G (2015) Persistent cold air outbreaks over North America in a
1449 warming climate. Environ. Res. Lett., accepted.

1450 Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models,
1451 J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972

1452 Gershunov A, Cayan DR, Iacobellis SF (2009) The great 2006 heat wave over California and
1453 Nevada: signal of an increasing trend. J. Climate 22: 6181-6203

1454 Gershunov A, Douville H (2008) Extensive summer hot and cold extremes under current and
1455 possible future climatic conditions: Europe and North America. In: H. F. Diaz and R. J.
1456 Murnane (Eds.), *Climate Extremes and Society.* Cambridge University Press, New York,
1457 348pp, 74-98

1458 Gershunov A, Guirguis K (2012) California heat waves in the present and future. *Geophysical*
1459 *Research Letters*, 39, L18710, doi:10.1029/2012GL052979

1460 Gleason KL, Lawrimore JH, Levinson DH, Karl TR, Karoly DJ (2008) A Revised U.S. Climate
1461 Extremes Index. J. Clim, 21: 2124–2137. doi: http://dx.doi.org/10.1175/2007JCLI1883.1

1462 Gordon HB, O'Farrell SP (1997) Transient climate change in the CSIRO coupled model with
1463 dynamical sea-ice. Mon Weather Rev 125:875-907

1464 Gosling SN, McGregor GR, Paldy A (2007) Climate change and heat-related mortality in six
1465 cities Part 1: Model construction and validation. Int. J. Biometeorol. 51: 525-540

1466 Gotaas, Y, Benson CS (1965) The Effect of Suspended Ice Crystals on Radiative Cooling. *J.*
1467 *Appl. Meteor.*, **4,** 446-453

1468 Goto-Maeda Y, Shin D, O'Brien J (2008), Freeze probability of Florida in a regional climate
1469 model and climate indices, Geophys Res Lett 35: L11703, doi:10.1029/2008GL033720.

1470 Grotjahn R (2011) Identifying extreme hottest days from large scale upper air data: a pilot
1471 scheme to find California central valley summertime maximum surface temperatures.
1472 Clim Dyn 37: 587–604. doi:10.1007/s00382-011-0999-z

1473 Grotjahn R (2013) Ability of CCSM4 to simulate California extreme heat conditions from
1474 evaluating simulations of the associated large scale upper air pattern. Clim Dyn 41:
1475 1187-1197. DOI 10.1007/s00382-013-1668-1

1476 Grotjahn R (2014) Western North American extreme heat, associated large scale synoptic-
1477 dynamics, and performance by a climate model. In: *Dynamics and Predictability of*
1478 *Large-Scale, High-Impact Weather and Climate Events*. Eds: J Li, R Swinbank, H
1479 Volkert, R Grotjahn, Cambridge Univ Press

1480 Grotjahn R, Faure G (2008) Composite predictor maps of extraordinary weather events in the
1481 Sacramento California region. Weather Forecast 23:313–335.
1482 doi:10.1175/2007WAF2006055.1

1483 Gu L, Hanson P, Post W, Kaiser D, Yang B, Nemani R, Pallardy S, Meyers T (2008) The 2007
1484 eastern U.S. spring freezes: Increased cold damage in a warming world? BioScience
1485 58:253–262

1486 Guirguis K, Gershunov A, Schwartz R, Bennett S (2011) Recent warm and cold daily winter
1487 temperature extremes in the Northern Hemisphere. Geophys Res Lett 38:L17701
1488 doi:10.1029/2011GL048762

1489 Gutowski, WJ, Otieno F, Arritt RW, Takle ES, Pan Z (2004) Diagnosis and attribution of a
1490 seasonal precipitation deficit in a U.S. regional climate simulation. J. Hydrometeor. 5:
1491 230-242 [DOI: 10.1175/1525-7541(2004)005<0230:DAAOAS>2.0.CO;2]

1492 Hajat S, Kovats RS, Atkinson RW, Haines A (2002) Impact of hot temperatures on death in
1493 London: A time series approach. J. Epidemiol. Comm. Health 56: 367-372

1494 Hewitson, BC, Crane RG, (2002) Self-organizing maps: applications to synoptic climatology.
1495 Clim. Res., 18, 41-57

1496 Hewitson BC, Crane RG (2006) Consensus between GCM climate change projections with
1497 empirical downscaling: precipitation downscaling over South Africa. Int. J. of Climatology
1498 26: 1315–1337. DOI: 10.1002/joc.1314

1499 Higgins R, Leetmaa A, Kousky V (2002), Relationships between climate variability and winter
1500 temperature extremes in the United States, J. Clim., 15, 1555–1572

1501 Hinton, G, Deng L, Yu D, Dahl G, Mohamed A-R, Jaitly N, Senior A, Vanhoucke V, Nguyen P,
1502 Sainath T, Kingsbury B (2012) "Deep neural networks for acoustic modeling in speech
1503 recognition: The shared views of four research groups." *Signal Processing Magazine,*
1504 *IEEE* 29.6 : 82-97.

1505 Ho CK, Stephenson DB, Collins M, Ferro CAT, Brown SJ (2012) Calibration strategies: A
1506 source of additional uncertainty in climate change projections. *Bulletin of the American*
1507 *Meteorological Society*, **93**, 21-26

1508 Hohenegger C, Schär C (2007) Atmospheric predictability at synoptic versus cloud-resolving
1509 scales. Bull Amer Meteor Soc 88:1783-1793.

1510 Jaeger EB, Seneviratne SI (2011) Impact of Soil Moisture-Atmosphere Coupling on European
1511 Climate Extremes and Trends in a Regional Climate Model. *Climate Dynamics*, 36 (9-
1512 10), 1919-1939, doi: 10.1007/s00382-010-0780-8

1513 Jaiser R, Dethloff K, Handorf D, Rinke A, Cohen J (2012) Impact of sea ice cover changes on
1514 the northern hemisphere atmospheric winter circulation. Tellus A 64:11595.
1515 doi:10.3402/tellusa.v64i0.11595

1516 Kallache M, Vrac M, Naveau P, Michelangeli P-A (2011) Nonstationary probabilistic
1517 downscaling of extreme precipitation. *Journal of Geophysical Research*, 116: D05113,
1518 doi: 10.1029/2010JD014892

1519 Karl TR, Knight RW, Easterling DR, Quayle RG (1996) Indices of Climate Change for the United
1520 States. Bull. Amer. Meteor. Soc., 77: 279–292. doi: http://dx.doi.org/10.1175/1520-
1521 0477(1996)077<0279:IOCCFT>2.0.CO;2

1522 Karl TR, Quayle RG (1981) The 1980 Summer Heat Wave and Drought in Historical
1523 Perspective. Mon Wea Rev, 109: 2055–2073. doi: http://dx.doi.org/10.1175/1520-
1524 0493(1981)109<2055:TSHWAD>2.0.CO;2

1525 Katz RW (2010) Statistics of extremes in climate change. Clim. Change 100:71-76

1526 Katz RW, Grotjahn R (2014) Statistical methods for relating temperature extremes to Large-
1527 Scale Meteorological Patterns. US CLIVAR Var. 12 (1):4-7
1528 http://www.usclivar.org/sites/default/files/documents/2014/Variations-winter-2014-v3.pdf

1529 Kenyon J, Hegerl G (2008) Influence of Modes of Climate Variability on Global Temperature
1530 Extremes J Clim 21:3872–3889. doi: http://dx.doi.org/10.1175/2008JCLI2125.1

1531 Kharin, VV, Zwiers FW, Zhang X, Wehner M (2013) Changes in temperature and precipitation
1532 extremes in the CMIP5 ensemble, Climatic Change 119: 345-347

1533 Klein W (1952) The weather and circulation of June 1952. Mon Wea Rev, 80: 99-104

1534 Kodra E, Steinhaeuser K, Ganguly AR (2011) Persisting Cold Extremes Under 21st-Century
1535 Warming Scenarios. Geophys. Res. Lett., 38, L08705, doi:10.1029/2011GL047103

Kohonen T (1995) Self-Organizing Maps. 2d ed. Springer-Verlag, 426 pp

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 Wea Rev 124:1067-1083 doi: http://dx.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 Wea Rev, 117: 2687–2700

Krizhevsky A, Sutskever I, Hinton G (2012) "Imagenet classification with deep convolutional neural networks." Advances in Neural Information Processing Systems 25 (Proceedings from NIPS 2012) Available at: 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. Amer. Meteor. Soc., 77: 1507–1518

Kunkel, KE, Liang X-Z, Zhu J (2010) Regional Climate Model Projections and Uncertainties of U.S. 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 Climate, 25: 4761–4784. doi: http://dx.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

Lee, Y-Y, Grotjahn R (2015) California Central Valley summer heat waves form two ways. Submitted to Climate Dyn.

1558 Leibensperger EM, Mickley LJ, Jacob DJ, Chen W-T, Seinfeld JH, Nenes A, Adams PJ, Streets
1559 DG, Kumar N, Rind D (2012) Climatic effects of 1950-2050 changes in US
1560 anthropogenic aerosols – Part 2: Climate response, Atmos. Chem. Phys., 12: 3349–
1561 3362, doi:10.5194/acp-12-3349-2012.

1562 Lim YK, Ham YG, Jeong JH, Kug JS (2012) Improvement in simulation of Eurasian winter
1563 climate variability with a realistic Arctic sea ice condition in an atmospheric GCM.
1564 Environ Res Lett 7:044041. doi: 10.1088/1748-9326/7/4/044041

1565 Lim YK, Kim HD (2013) Impact of the dominant large-scale teleconnections on winter
1566 temperature variability over East Asia. J Geophys Res 118:7835–7848.
1567 doi:10.1002/jgrd.50462

1568 Lim Y-K, Schubert S (2011), The impact of ENSO and the Arctic Oscillation on winter
1569 temperature extremes in the southeast United States, Geophys Res Lett 38: L15706,
1570 doi:10.1029/2011GL048283

1571 Lipton K, Grumm R, Holmes R (2005) Forecasting Heat Waves using Climatic Anomalies. 21st
1572 Conference on Weather Analysis/17th Conference on Numerical Weather Prediciton. 1-5
1573 August 2005. Available at: http://ams.confex.com/ams/pdfpapers/94498.pdf

1574 Liu J, Curry JA, Wang H, Song M, Horton RM (2012) Impact of declining Arctic sea ice on winter
1575 snowfall PNAS. 109:4074–4079. doi:10.1073/pnas.1114910109

1576 Livezey RE, Tinker R (1996) Some Meteorological, Climatological, and Microclimatological
1577 Considerations of the Severe U.S. Heat Wave of Mid-July 1995. Bull. Amer. Meteor.
1578 Soc., 77: 2043–2054

1579 Loikith P, Broccoli A (2012) Characteristics of Observed Atmospheric Circulation Patterns
1580 Associated with Temperature Extremes over North America J Clim 25:7266–7281 doi:
1581 http://dx.doi.org/10.1175/JCLI-D-11-00709.1

1582 Loikith PC, Broccoli AJ (2014) The Influence of Recurrent Modes of Climate Variability on the
1583 Occurrence of Winter and Summer Extreme Temperatures over North America. J Clim
1584 27: 1600–1618.

1585 Lorenz R, Jaeger EB, Seneviratne SI (2010) Persistence of Heat Waves and its Link to Soil
1586 Moisture Memory. Geophys. Res. Lett, 37:, L09703, doi:10.1029/2010GL042764.

1587 Lyon B (2009) Southern Africa Summer Drought and Heat Waves: Observations and Coupled
1588 Model Behavior. J. Climate, 22: 6033–6046.
1589 doi: http://dx.doi.org/10.1175/2009JCLI3101.1

1590 Lyon B, Dole RM (1995) A Diagnostic Comparison of the 1980 and 1988 U.S. Summer Heat
1591 Wave-Droughts. J Clim, 8: 1658–1675

1592 Maraun, D (2013) Bias correction, quantile mapping, and downscaling: Revisiting the inflation
1593 issue. J Clim, 26: 2137-2143

1594 Masato G, Hoskins BJ, Woollings T (2013) Winter and Summer Northern Hemisphere Blocking
1595 in CMIP5 Models. J. Clim, 26: 7044–7059

1596 Meehl G, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the
1597 twenty-first century. Science 305:994–997

1598 Meehl G, Tebaldi C, Teng H, Peterson T (2007) Current and future U.S. weather extremes and
1599 El Niño Geophys Res Lett 34: L20704, doi:10.1029/2007GL031027

1600 Moon J-Y, Wang B, Ha K (2011) ENSO regulation of MJO teleconnection. Clim Dyn 37: 1133–
1601 1149

1602 Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori
1603 S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ,
1604 Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of
1605 scenarios for climate change research and assessment. Nature, 463: 747-756 (11
1606 February 2010) | doi:10.1038/nature08823

1607 Nakicenovic N, Alcamo J, Davis G, de Vries HJM, Fenhann J, Gaffin S, Gregory K, Grubler A,
1608 Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Papper W, Pitcher H,
1609 Price L, Riahi K, Roehrl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S,
1610 Swart R, van Rooijen S, Victor N, Dadi Z (2000) Special Report on Emissions Scenarios.
1611 Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge.

1612 Namias J (1982) Anatomy of Great Plains protracted heat waves (especially the 1980 U.S.
1613 summer drought). Mon Wea Rev, 110: 824-838

1614 Orlowsky B, Seneviratne SI (2012) Global changes in extreme events: Regional and seasonal
1615 dimension. *Climatic Change*, 110: 669-696, doi: 10.1007/s10584-011-0122-9

1616 Outten SD, Esau I (2012) A link between Arctic sea ice and recent cooling trends over Eurasia.
1617 Clim Change 110:1069–1075. doi:10.1007/s10584-011-0334-z

1618 Overland JE, Wang M (2010) Large-scale atmospheric circulation changes associated with the
1619 recent loss of Arctic sea ice. Tellus A 62:1–9. doi:10.1111/j.1600-0870.2009.00421.x

1620 Palecki MA, Changnon SA, Kunkel KE (2001) The nature and impacts of the July 1999 heat
1621 wave in the midwestern United States: Learning from the lessons of 1995. Bull. Amer.
1622 Meteor. Soc., 82: 1353–1367

1623 Park T-W, Ho C-H, Yang S (2011) Relationship between the Arctic Oscillation and cold surges
1624 over East Asia. J. Clim, 24: 68-83

1625 Peings Y, Brun E, Mauvais V, Douville H (2013) How stationary is the relationship between
1626 Siberian snow and Arctic Oscillation over the 20th century? Geophys. Res. Lett., 40:
1627 183–188, doi:10.1029/2012GL054083

1628 Peterson TC, Heim R, Hirsch R, Kaiser D, Brooks H, Diffenbaugh NS, Dole R, Giovannettone J,
1629 Guiguis K, Karl TR, Katz RW, Kunkel K, Lettenmaier D, McCabe GJ, Paciorek CJ,
1630 Ryberg K, Schubert S, Silva VBS, Stewart B, Vecchia AV, Villarini G, Vose RS, Walsh
1631 J,Wehner M, Wolock D, Wolter K, Woodhouse CA, Wuebbles D (2013) Monitoring and
1632 Understanding Changes in Heat waves, Cold waves, Floods and Droughts in the United
1633 States: State of Knowledge, Bull Amer Meteorol Soc, 94: 821-834, DOI: 10.1175/BAMS-
1634 D-12-00066.1, Supplement DOI: 10.1175/BAMS-D-12-00066.2

1635 Photiadou C, Jones MR, Keelings D, Dewes CF (2014) Modelling European Hot Spells Using
1636 Extreme Value Analysis. Climate Res, 58: 193-207 doi:10.3354/cr01191

1637 Plumb RA (1985) On the three-dimensional propagation of stationary waves. J Atmos Sci, 42:
1638 217–229

1639 Portis DH, Cellitti MP, Chapman WL, Walsh JE (2006) Low-Frequency Variability and Evolution
1640 of North American Cold Air Outbreaks. Mon Wea Rev 134:579-597

1641 Renwick JA, Wallace JM (1996) Relationships between north Pacific wintertime blocking, El
1642 Nin˜o, and the PNA pattern. Mon Wea Rev, 124: 2071–2076

1643 Rienecker M, and Coauthors (2007) The GEOS-5 data assimilation system—Documentation of
1644 versions 5.0.1 and 5.1.0.NASA GSFC Tech. Rep. Series on Global Modeling and Data
1645 Assimilation, NASA/TM-2007-104606, Vol. 27, 92 pp.

1646 Robinson P (2001) On the Definition of a Heat Wave. J. Applied Meteor. 40:762-775

1647 Saha S, and Coauthors (2010) The NCEP Climate Forecast System Reanalysis. Bull. Amer.

1648 Meteor. Soc., 91: 1015–1057

1649 Salakhutdinov R, Hinton G (2012) An efficient learning procedure for deep Boltzmann
1650 machines. Neural Computation, 24(8):1967–2006

1651 Sardeshmukh PD, Hoskins BJ (1988) The Generation of Global Rotational Flow by Steady
1652 Idealized Tropical Divergence. J. Atmos. Sci., 45: 1228–1251

1653 Scaife AA, Woollings T, Knight J, Martin G, Hinton T (2010) Atmospheric blocking and mean
1654 biases in climate models, J. Clim, 23(23): 6143–6152

1655 Schubert S, Wang H, Suarez M (2011) Warm season subseasonal variability and climate
1656 extremes in the Northern Hemisphere: The role of stationary Rossby waves. J Clim,
1657 24:4773–4792

1658 Screen JA, Simmonds I (2013) Exploring links between Arctic amplification and midlatitude
1659 weather. Geophys. Res. Lett., 40, 959–964, doi:10.1002/GRL.50174

1660 Seber GAF (2008) Multivariate Distributions, Multivariate Observations, Wiley, 17-58, ISBN 0-
1661 471-69121-6

1662 Seneviratne SI, Luthi D, Litschi M, Schar C (2006) Land-Atmosphere Coupling and Climate
1663 Change in Europe. Nature, 443: 205-209 doi:10.1038/nature05095

1664 Sillmann J, Croci-Maspoli M (2009) Present and future atmospheric blocking and its impact on
1665 European mean and extreme climate. Geophys Res Lett , 36: L10702

1666 Sillmann J, Croci-Maspoli M, Kallache M, Katz RW (2011) Extreme Cold Winter Temperatures
1667 in Europe under the Influence of North Atlantic Atmospheric Blocking. J. Clim, 24: 5899–
1668 5913. doi: http://dx.doi.org/10.1175/2011JCLI4075.1

1669 Sillmann J, Kharin VV, Zwiers FW, Zhang X, Bronaugh D (2013a) Climate extremes indices in
1670 the CMIP5 multi model ensemble: Part 1. Model evaluation in the present climate. J
1671 Geophys Res, 118: 1716–1733

1672 Sillmann J, Kharin VV, Zwiers FW, Zhang X, Bronaugh D (2013b) Climate extremes indices in
1673 the CMIP5 multi model ensemble: Part 2. Future climate projections. J Geophys Res,
1674 118: 2473–2493

1675 Simmons AJ, Willett KM, Jones PD, Thorne PW, Dee DP (2010) Low-frequency variations in
1676 surface atmospheric humidity, temperature, and precipitation: Inferences from
1677 reanalyses and monthly gridded observational data sets, J. Geophys. Res., 115:
1678 D01110, doi:10.1029/2009JD012442

1679 Smyth P, Ghil M, Ide K (1999) Multiple regimes in Northern Hemisphere height fields via mixture
1680 model clustering. J. Atmos. Sci. 56: 3704–3723

1681 Spaete P, Johnson DR, Schaack TK (1994) Stratospheric–Tropospheric Mass Exchange during
1682 the Presidents' Day Storm. Mon Wea Rev, 122: 424–439

1683 Spath H (1985) Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples.
1684 Horwood, 226 pp.

1685 Stephanon M, D'Andrea F, Drobinski P (2012) Heatwave classification over Europe and the
1686 Mediterranean region. Environ Res Lett 7: 014023 (9pp) http://dx.doi.org/10.1088/1748-
1687 9326/7/1/014023

1688 Takaya K, Nakamura H (2001) A formulation of a phase-independent wave-activity flux for
1689 stationary and migratory quasigeostrophic eddies on a zonally varying basic flow. J
1690 Atmos Sci, 58: 608–627

1691 Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA (2006) Going to the extremes; An
1692 intercomparison of model-simulated historical and future changes in extreme events,
1693 Climatic Change, 79: 185-211

1694 Teng H, Branstator G, Wang H, Meehl GA, Washington WM (2013) Probability of US heat
1695 waves affected by a subseasonal planetary wave pattern. Nature Geosci, 6: 1056–1061
1696 doi:10.1038/ngeo1988

1697 Thrasher B, Maurer EP, McKellar C, Duffy PB (2012) Bias correcting climate model simulated
1698 daily temperature extremes with quantile mapping. Hydrol Earth Sys Sci, 16: 3309-3314

1699 Turner JK, Gyakum JR (2011) The development of Arctic air masses in Northwest Canada and
1700 their behavior in a warming climate. J. Clim, 24: 4618-4633

1701 Turner JK, Gyakum JR, Milrad SM (2013) A thermodynamic analysis of an intense North
1702 American arctic air mass. Mon Wea Rev, 141: 166-181

1703 van Oldenborgh GJ, Drijfhout S, van Ulden A, Haarsma R, Sterl A, Severijns C, Hazeleger W,
1704 Dijkstra H (2009) Western Europe is warming much faster than expected, Clim. Past, 5:
1705 1-12.

1706 van Vuuren, DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T,
1707 Krey V, Lamarque J-F, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK
1708 (2011) The representative concentration pathways: an overview. Climatic Change,
1709 109:5-31 10.1007/s10584-011-0148-z

1710 Vavrus SJ, Walsh JE, Chapman WL, Portis D (2006) The behavior of extreme cold air
1711 outbreaks under greenhouse warming. International Journal of Climatology, 26:1133.

1712 Vavrus S, Holland MM, Jahn A, Bailey DA, Blazey BA (2012) 21st-century Arctic climate change
1713 in CCSM4. J Clim 25:2696-2710. doi:10.1175/JCLI-D-11-00220.1

1714 Von Storch, H (1999) On the use of “inflation” in statistical downscaling. Journal of Clim, 12:
1715 3505-3506.

1716 Walsh, JE, Philips AS, Portis DH, Chapman WL (2001) Extreme Cold Outbreaks in the United
1717 States and Europe, 1948-99. J. Clim., 14: 2642-2658

1718 Wang A, Zeng X (2013) Development of Global Hourly 0.5° Land Surface Air Temperature
1719 Datasets. J. Clim, 26: 7676–7691. doi: http://dx.doi.org/10.1175/JCLI-D-12-00682.1

1720 Wehner MF (2004) Predicted 21st century changes in seasonal extreme precipitation events in
1721 the Parallel Climate Model, J. Clim, 17: 4281-4290

1722 Wehner MF (2010) Sources of uncertainty in the extreme value statistics of climate data.
1723 Extremes 13: 205-217, doi 10.1007/s10687-010-0105-7

1724 Westby RM, Lee YY, Black RX (2013) Anomalous Temperature Regimes during the Cool
1725 Season: Long-Term Trends, Low-Frequency Mode Modulation, and Representation in
1726 CMIP5 Simulations. J Clim 26:9061–9076.

1727 Wexler H (1936) Cooling in the lower atmosphere and the structure of polar continental air. Mon
1728 Wea Rev, 64**:** 122-136

1729 Wexler H (1937) Formation of Polar Anticyclones. Mon Wea Rev, 65**:** 229-236

1730 Wilks DS (2006) *Statistical Methods in the Atmospheric Sciences*. Second Edition, Academic
1731 Press.

1732 Wu Z, Lin HL, Li J, Jiang Z, Ma T (2012) Heat wave frequency variability over North America:
1733 Two distinct leading modes. J Geophys Res 117: D02102.
1734 http://dx.doi.org/10.1029/2011JD016908

1735 Wuebbles D, Meehl G, Hayhoe K, Karl TR, Kunkel K, Santer B, Wehner M, Colle B, Fischer EM,
1736 Fu R, Goodman A, Janssen E, Lee H, Li W, Long LN, Olsen S, Seth A, Sheffield J, Sun

L (2013) CMIP5 Climate Model Analyses: Climate Extremes in the United States, Bull Amer Meteorol Soc. 95: 571-583 doi: http://dx.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

Zhang XD, Walsh JE (2006) Toward a seasonally ice-covered Arctic ocean: Scenarios from the IPCC AR4 model simulations. J Clim 19:1730–1747. doi:10.1175/JCLI3767.1

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: http://dx.doi.org/10.1175/2010JCLI3908.1

1751 Figure Captions

1752

1753 Fig. 1 Change over 1950-2007 in estimated 20-year annual return values (^{o}C) for a) hot tail of
1754 daily maximum temperature (TXx), b) cold tail of daily maximum temperature, (TXn) c) hot tail of
1755 daily minimum temperature, (TNx) and d) cold tail of daily minimum temperature (TNn). Results
1756 are based on fitting extreme value statistical models with a linear trend in the location parameter
1757 to exceedances of a location-specific threshold (greater than the 99th percentile for upper tail
1758 and less than the 1th percentile for lower tail). As this analysis was based on anomalies with
1759 respect to average values for that time of year, hot minimum temperature values, for example,
1760 are just as likely to occur in winter as in summer. The circles indicate the z-score for the
1761 estimated change (estimate divided by its standard error), with absolute z-scores exceeding 1,
1762 2, and 3 indicated by open circles of increasing size. Higher z-score indicates greater statistical
1763 significance.

1764

1765 Fig. 2 Example large scale meteorological patterns (LSMPs) obtained as target ensemble mean
1766 composites for two types of California Central valley extreme events. Cold air outbreaks in
1767 winter (DJF) at (a) 72 hours prior and (b) at onset of the events are shown in the 500 hPa
1768 geopotential height field. Heat waves during summer (JJAS) (c) 36 hours prior and (d) at the
1769 onset are shown in the 700 hPa geopotential height field. Shading indicates significance at the
1770 highest or lowest 5% level, with the innermost shading significant at the 1.5% level. Further
1771 discussion is in Grotjahn and Faure (2008).

1772

1773 Fig. 3 Self-organizing map of synoptic weather patterns in a region focused on Alaska. The
1774 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.

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 statistical significant at the 95% confidence level (Figure from Westby et al. 2013).

Fig. 5 (adapted from Turner et al. 2013): Sea level pressure (contours, every 4 hPa), 700-hPa winds (barbs, m s^{-1}), and 1000–500-hPa thicknesses (shaded, every 6 dam) at 1200 UTC daily during the formation period.

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).

1799

1800 Fig. 7 Composite synoptic weather patterns at the onset of the 14 Sacramento California
1801 summer (JJAS) heat waves studied by Grotjahn and Faure (2008). a) Temperature at 850 hPa
1802 with a 2 K interval. b) 700 hPa level pressure velocity with 2 Pa/s interval and where positive
1803 values mean sinking motion. c) Sea level pressure with 2 hPa interval, d) Surface wind vectors
1804 with shading applicable to the zonal component. Areas with yellow (lighter inside dark)) shading
1805 are positive (above normal) anomalies that are large enough to occur only 1.5% of the time by
1806 chance in a same-sized composite; areas that are blue (darker inside light) shading are
1807 negative anomalies occurring only 1.5% of the time.

1808

1809 Fig. 8 Performance portrait of the CMIP5 models' ability to represent the temperature based
1810 ETCCDI indices over North American land. The colors represent normalized root mean square
1811 errors of seasonal indices compared to the ERA Interim reanalysis. Blue colors represent errors
1812 lower than the median error, while red colors represent errors larger than the median error.
1813 Seasons are denoted by triangles within each square. Models marked with "*" are not included
1814 in the RCP8.5 projections. Root mean square errors normalized by the model median RMSE for
1815 3 other reanalyses are shown in the rightmost columns for comparison.

1816

1817 Fig. 9 The variance explained in an annual metric of the impact of wintertime Warm Waves on
1818 the southeast United States from a multiple linear regression using the NAO and PNA indices
1819 as predictors. Results are displayed for individual CMIP5 model simulations (blue bars) while
1820 the light and dark gray lines denote variance values for observations and the model mean,
1821 respectively (From Westby et al. 2013).

1822

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.

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).

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.

1837

1838

Table 1 Some temperature related ETCCDI indices (Sillmann et al. 2013a). For a complete list and formal definitions, see http://etccdi.pacificclimate.org/list_27_indices.shtml

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

1839

1840

1841

1842

1843

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	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

			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.	
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	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.

			fraction of the variance and so no single EOF might be an LSMP thereby limiting physical insight.	
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	Objectively classify events and find relating LSMPs that provide physical insight. No assumptions about the patterns such as orthogonality and symmetry.	The number of LSMPs can be somewhat arbitrary since number of cluster is pre-specified, although larger separation of clusters is acquired a little from the 'dissimilarity index'. The cluster assignment can be vague for	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. In theory,

	important to assess how independent the dates are, e.g. consulting the autocorrelation		some events.	one could 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.	DBNs 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.

1844

1845

1846

1847

1848

Table 3 Sampling of Various Criteria Used in Heat Wave and Hottest Day Definitions	
Source	**Definition**
Robinson (2001)	A period of at least 48 hours during which neither the overnight low nor the daytime heat index Hi falls below the NWS heat stress thresholds (80ºF and 105ºF). 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 (30C) 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 ≥ 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 100ºF (38ºC), and with at least one above 105ºF (40.5ºC)
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 ≥ 10C 2). At least 1 day must have maximum temperature anomaly ≥ 15C II. Must satisfy the 2 conditions above plus this additional condition: 3.) The average maximum temperature for the event ≥ 100 F (38C)
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

1849
1850

1851

1852

Table 4 Correlation coefficients between the loading pattern of observation and that of each CMIP5 model for NAO-like and PNA-like modes. Right-most column includes the averaged coefficients of two modes. Dark (light) grey shadings denote high-top (low-top) models (Reproduced from Lee and Black 2013).

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**

1853

1854

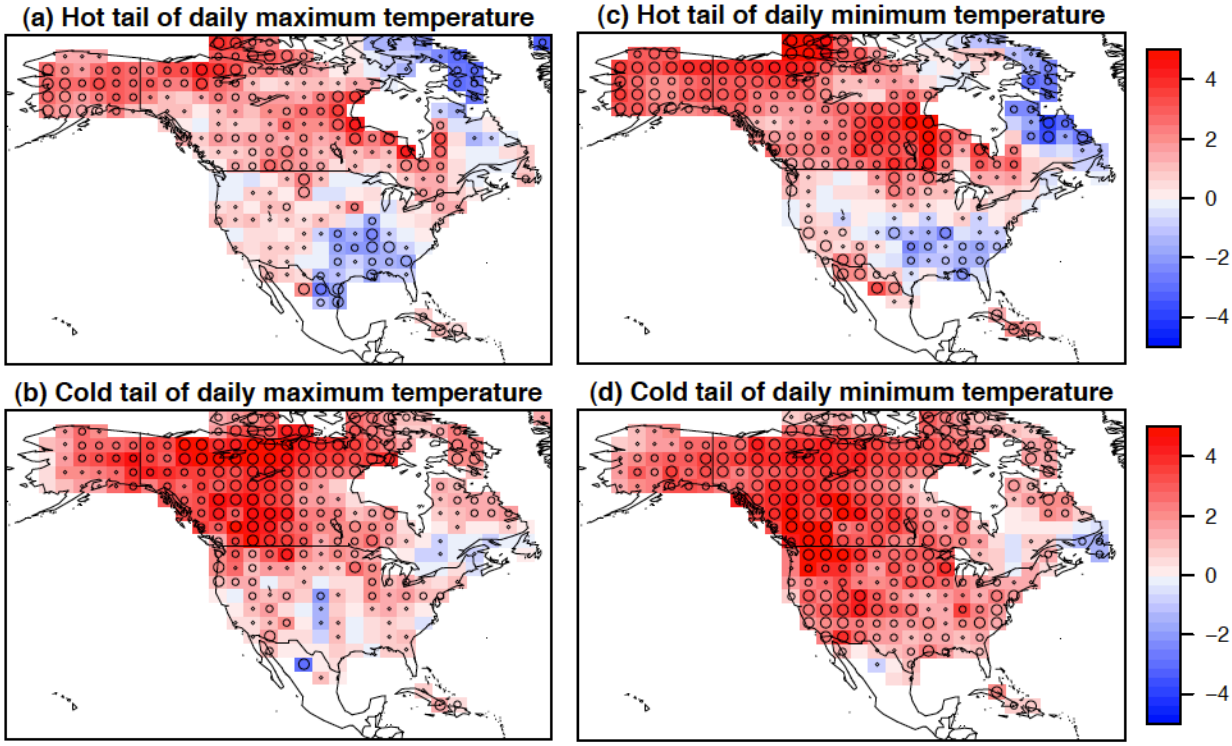


Fig. 1 Change over 1950-2007 in estimated 20-year annual return values (^{o}C) for a) hot tail of daily maximum temperature (TXx), b) cold tail of daily maximum temperature, (TXn) c) hot tail of daily minimum temperature, (TNx) and d) cold tail of daily minimum temperature (TNn). Results are based on 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.

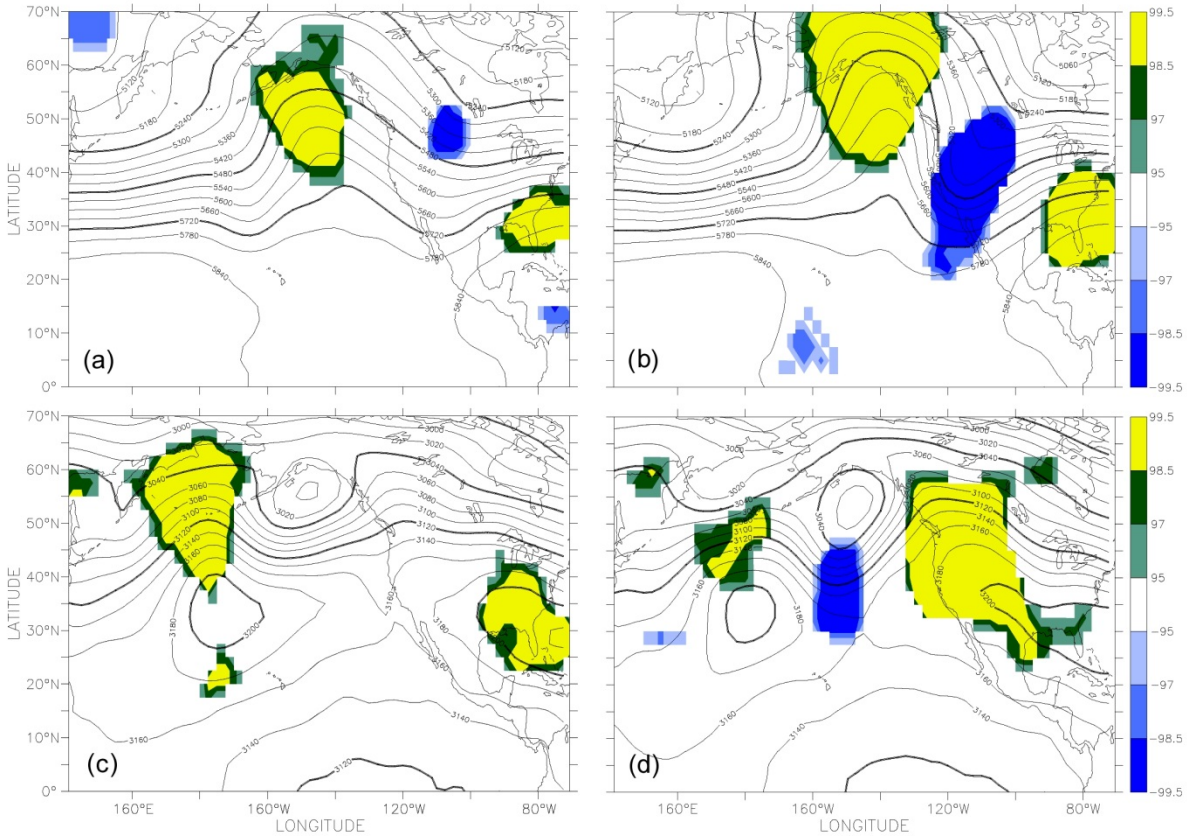


Fig. 2 Example large scale meteorological patterns (LSMPs) obtained as target ensemble mean composites for two types of California Central valley extreme events. Cold air outbreaks in winter (DJF) at (a) 72 hours prior and (b) at onset of the events are shown in the 500 hPa geopotential height field. Heat waves during summer (JJAS) (c) 36 hours 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).

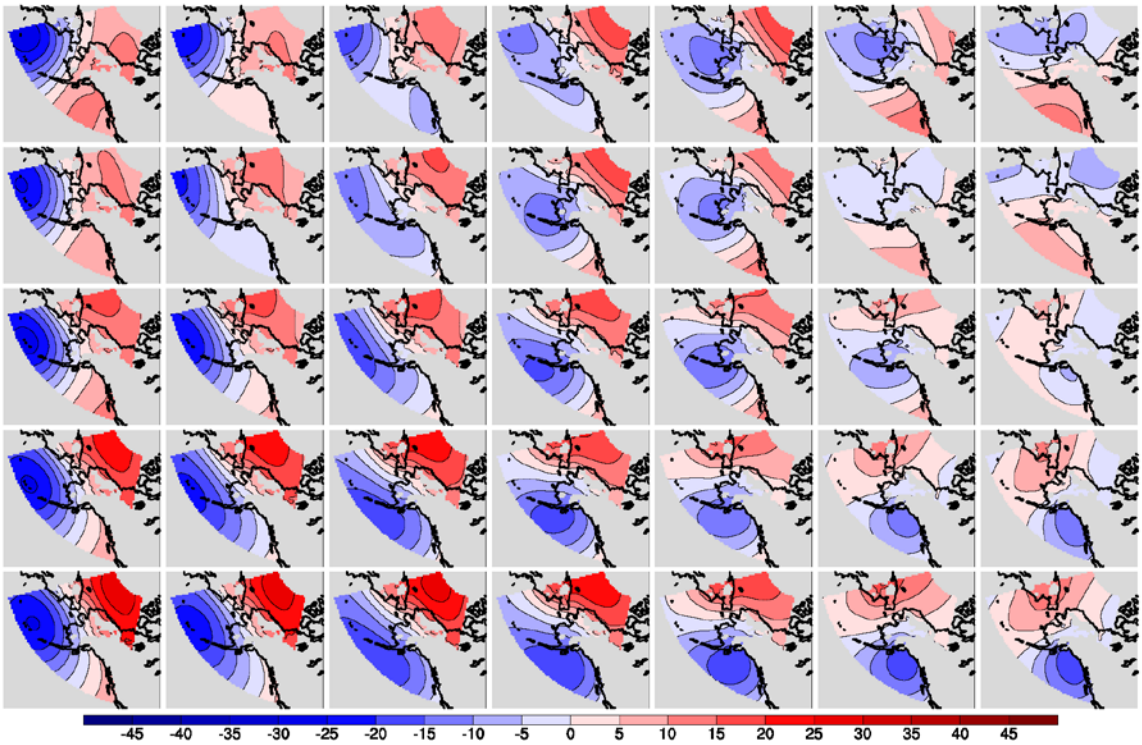


Fig. 3 Self-organizing map of synoptic weather patterns in a region focused on Alaska. The SOM array maps give the departure (in hPa) of sea level pressure (SLP) from the domain averaged sea level pressure. The SOM used daily December-January-February (DJF) SLP for 1997-2007 from ERA-Interim reanalyses 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.

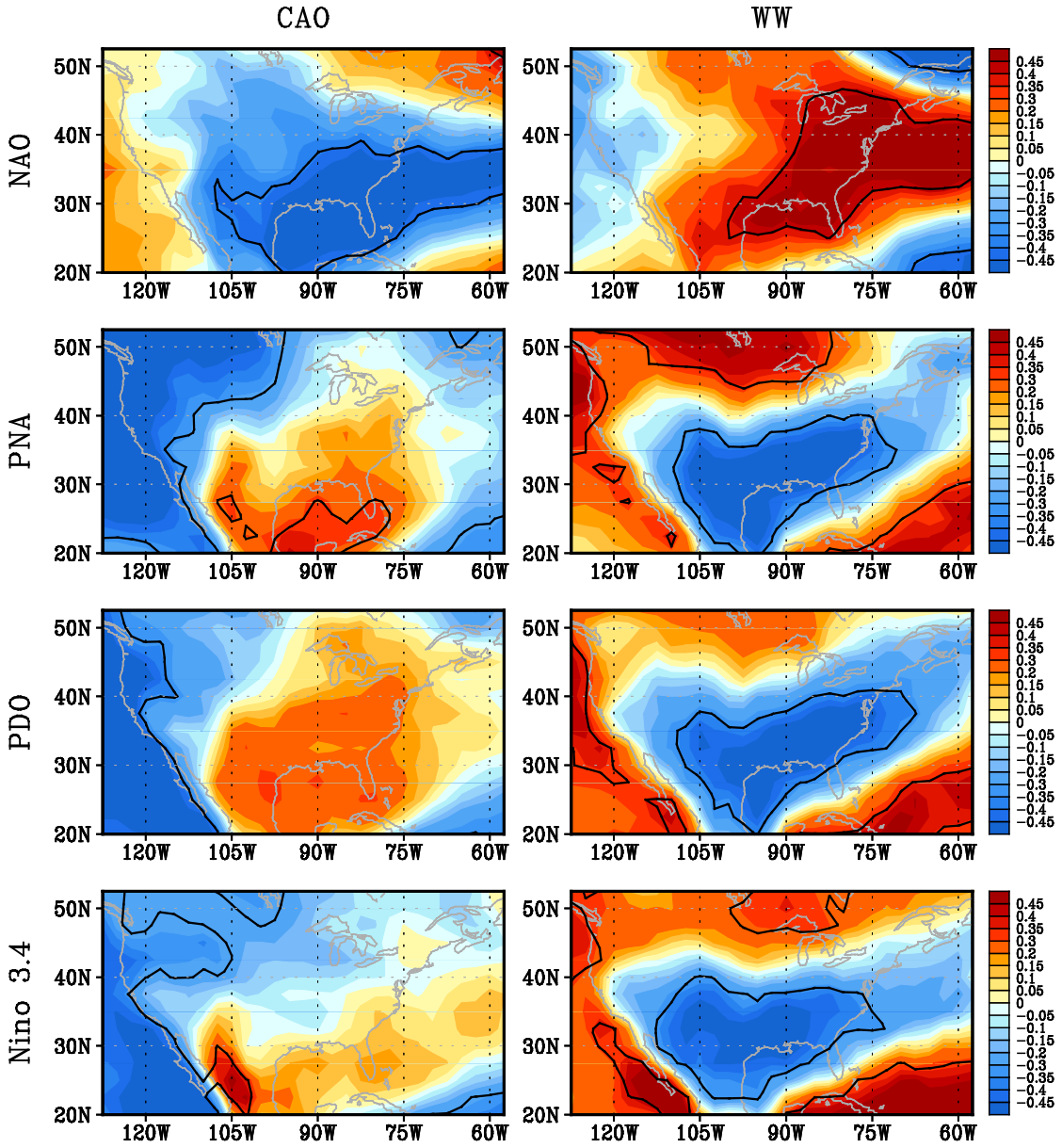


Fig. 4 Correlation between the local seasonal impact of cold days (left column) and warm days (right column) and the seasonal mean NAO (first row), PNA (second row), PDO (third row) and Niño 3.4 Indices (fourth row) during winter, 1950-2011. The black contours encompass regions having correlations statistical significant at the 95% confidence level (Figure from Westby et al. 2013).

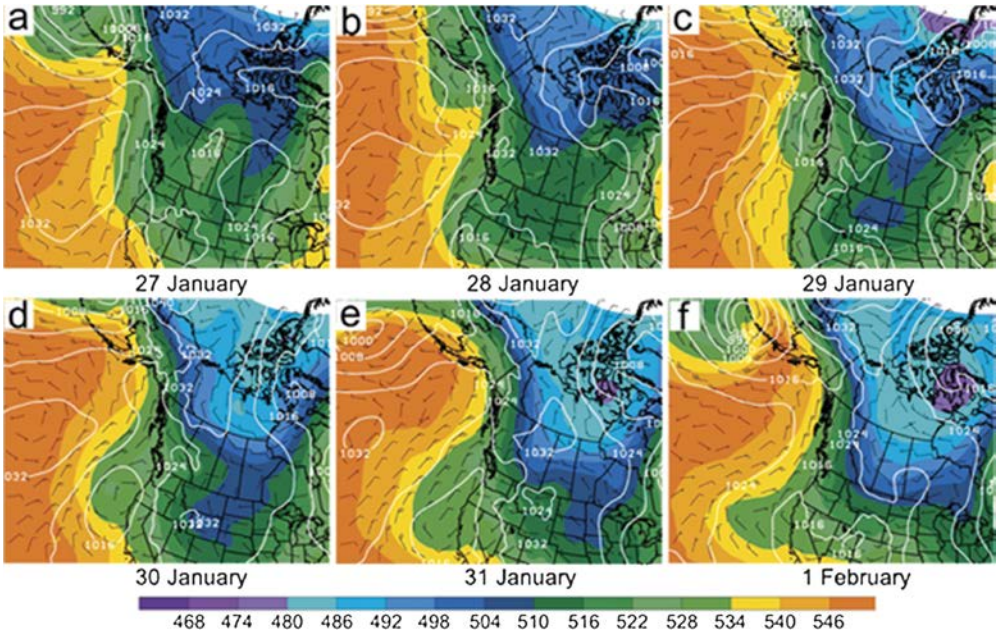
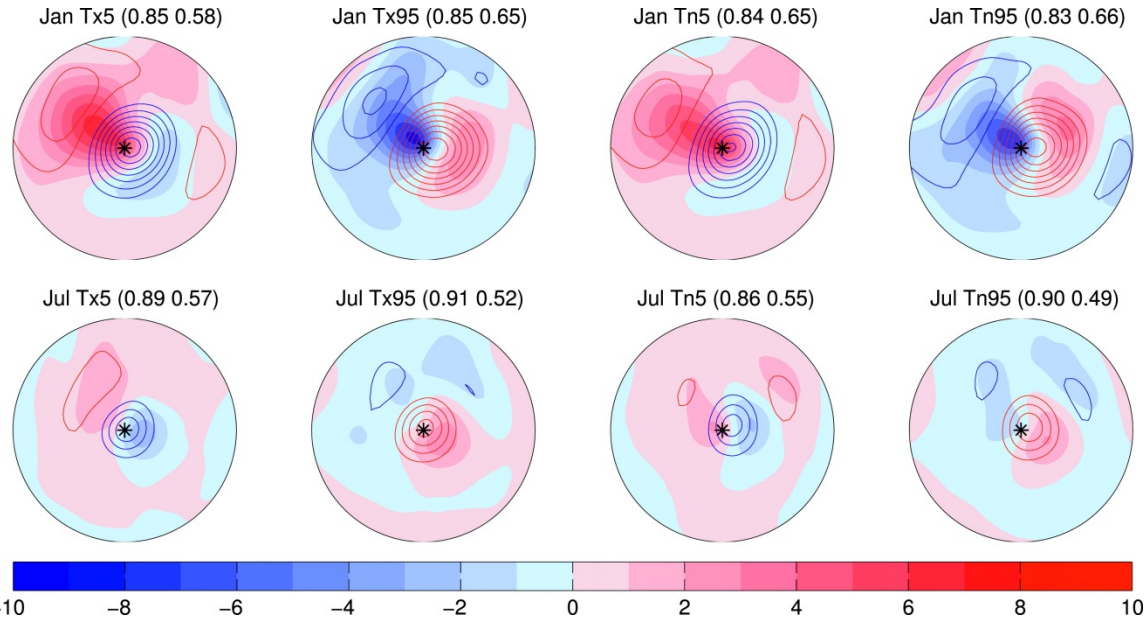


Fig. 5 (adapted from Turner et al. 2013): Sea level pressure (contours, every 4 hPa), 700-hPa winds (barbs, m s^{-1}), and 1000–500-hPa thicknesses (shaded, every 6 dam) at 1200 UTC daily during the formation period.



1926

1927

1928 **Fig. 6** Grand composites of anomalies associated with temperature regimes over North
1929 America: SLP anomalies (hPa) are shaded and Z_{500} height anomalies are contoured every 20
1930 m; red (blue) contours are positive (negative) Z_{500} anomalies. Grand composites are shown for
1931 (top) January and (bottom) July extreme (from left to right) cold maximum (Tx5), warm
1932 maximum (Tx95), cold minimum (Tn5), and warm minimum (Tn95) temperatures. The Grand
1933 composite combines information from extreme temperature events occurring at multiple
1934 disparate geographical locations over North America. The anomaly fields are displayed for a
1935 circular domain with a radius of 4500 km. From Loikith and Broccoli (2012; see text for further
1936 details).

1937

1938

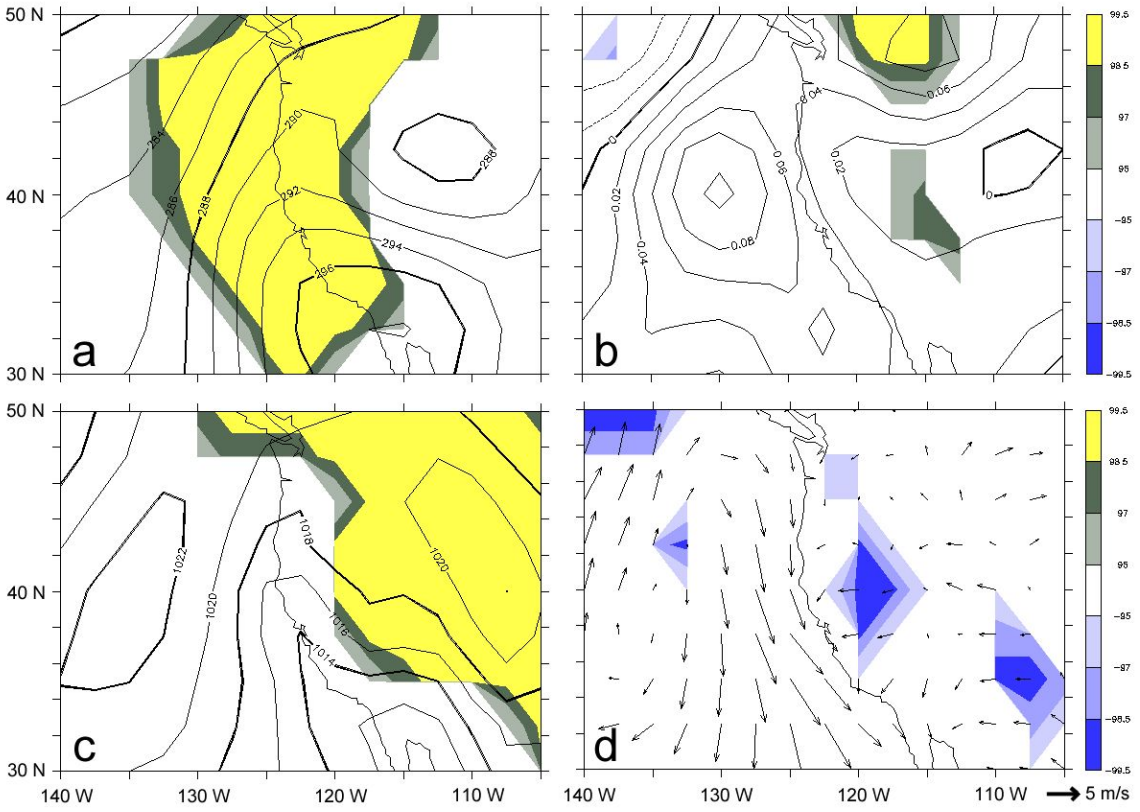


Fig. 7 Composite synoptic weather patterns at the onset of the 14 Sacramento California summer (JJAS) heat waves studied by Grotjahn and Faure (2008). a) Temperature at 850 hPa with a 2 K interval. b) 700 hPa level pressure velocity with 2 Pa/s interval and where positive values mean sinking motion. 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. Figure reproduced from Grotjahn (2011; DOI 10.1007/s00382-011-0999-z)

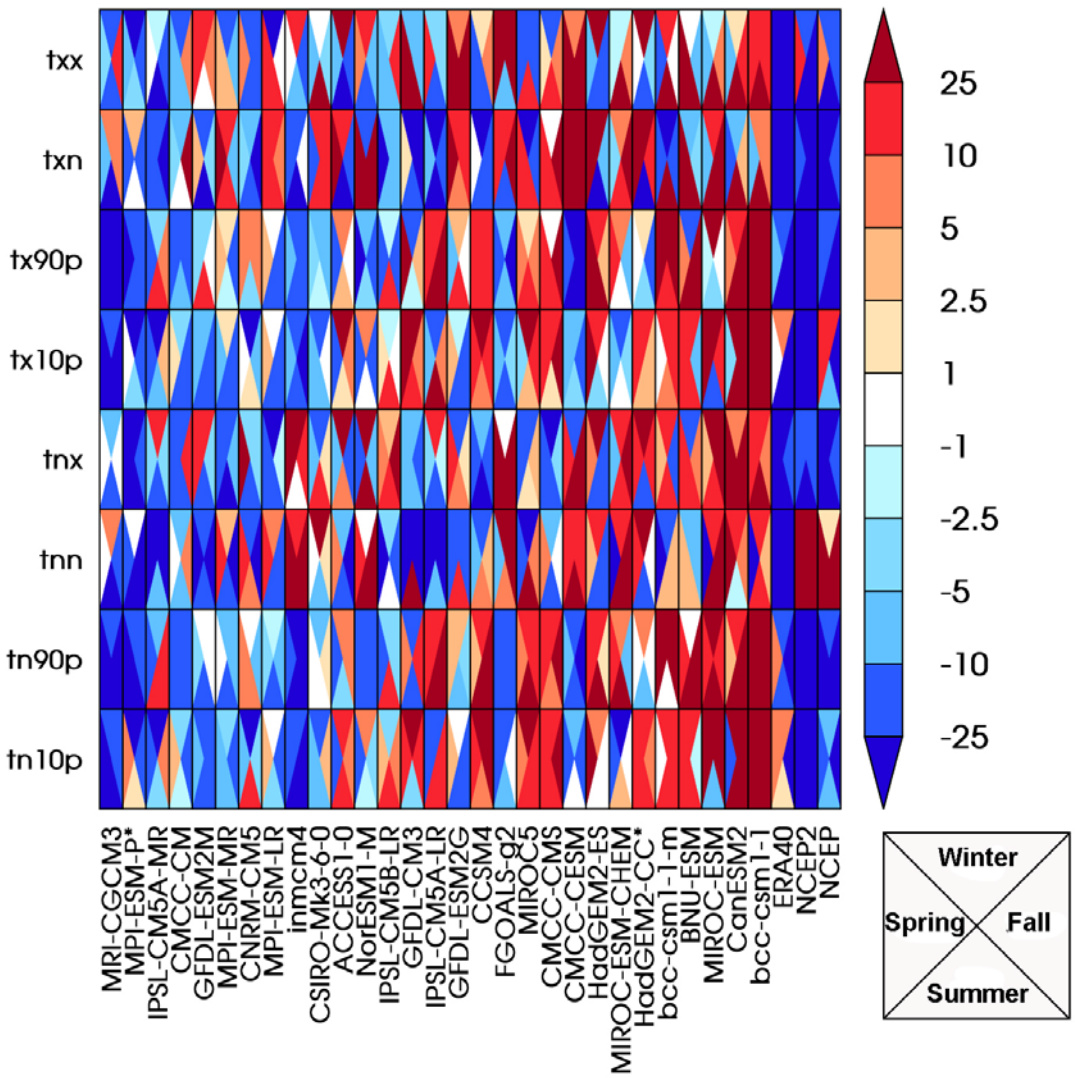


Fig. 8 Performance portrait of the CMIP5 models' ability to represent the temperature based ETCCDI indices over North American land. The colors represent normalized root mean square errors (RMSE) of seasonal indices compared to the ERA Interim reanalysis. Blue colors represent errors lower than the median 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.

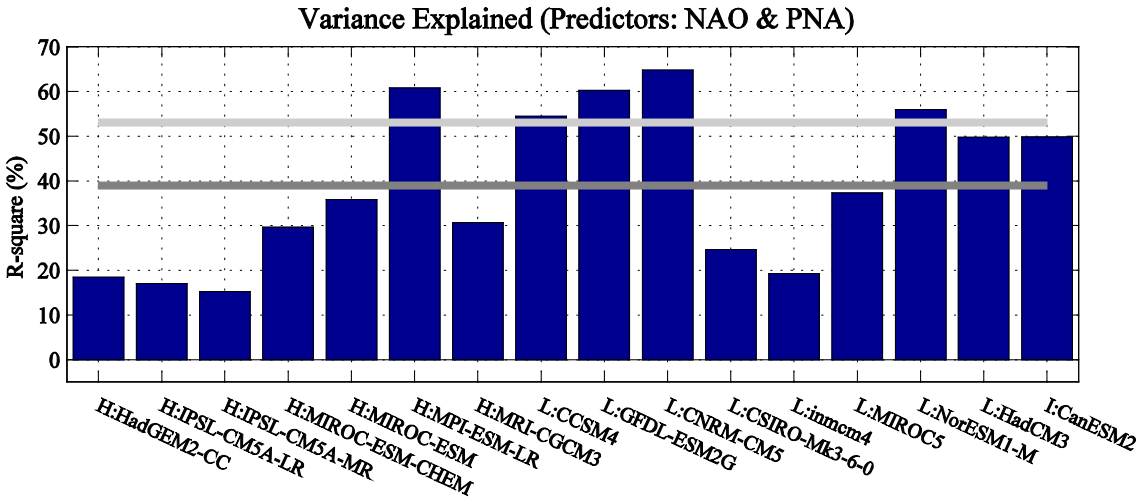


Fig. 9 The variance explained in an annual metric of the impact of wintertime Warm Waves on the southeast United States from a multiple linear regression using the NAO and PNA indices as predictors. Results are displayed for individual CMIP5 model simulations (blue bars) while the light and dark gray lines denote variance values for observations and the model mean, respectively (From Westby et al. 2013).

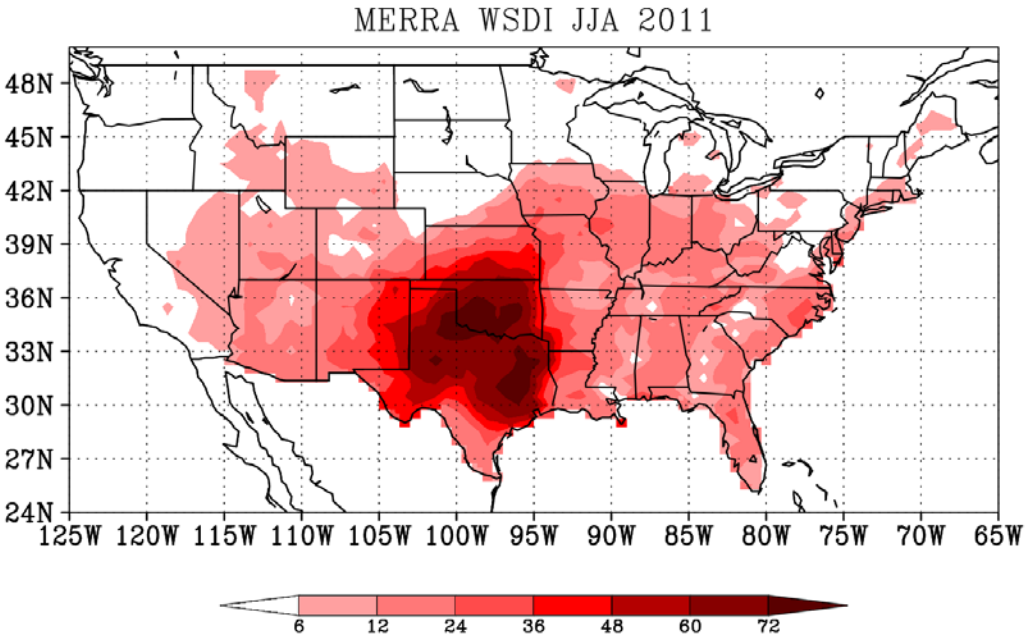


Fig. 10 Warm Spell Duration Index from MERRA for 2011 summer over the United States. WSDI is computed from the 90th percentile of daily mean temperatures for the summer season.

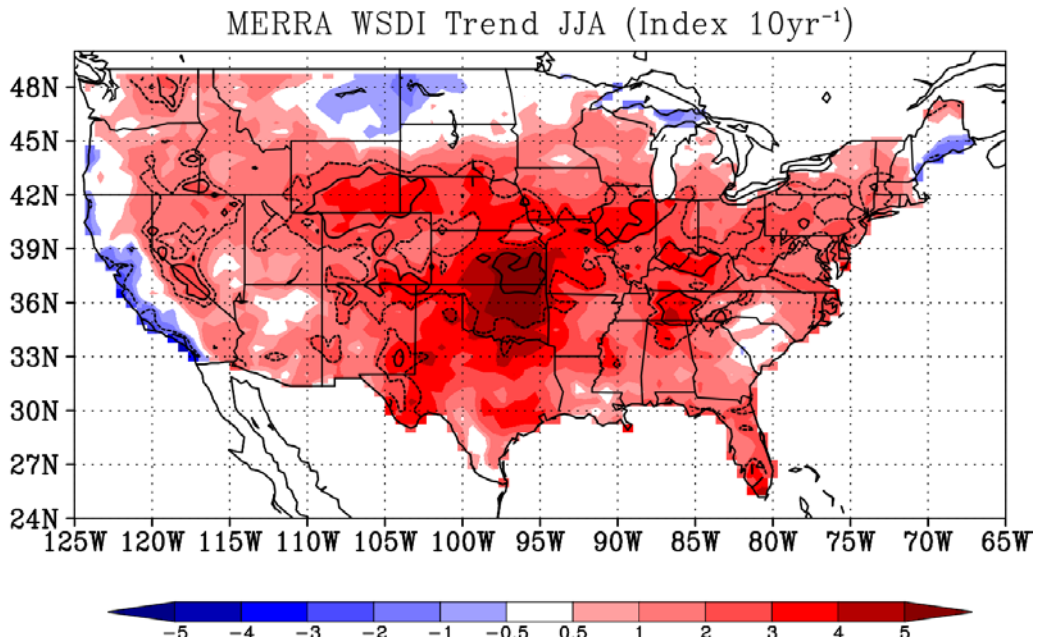


Fig. 11 Trend in WSDI as determined from MERRA for US Summertime temperatures. Dashed and solid black contours indicate statistical significance (at 90% and 95% confidence respectively).

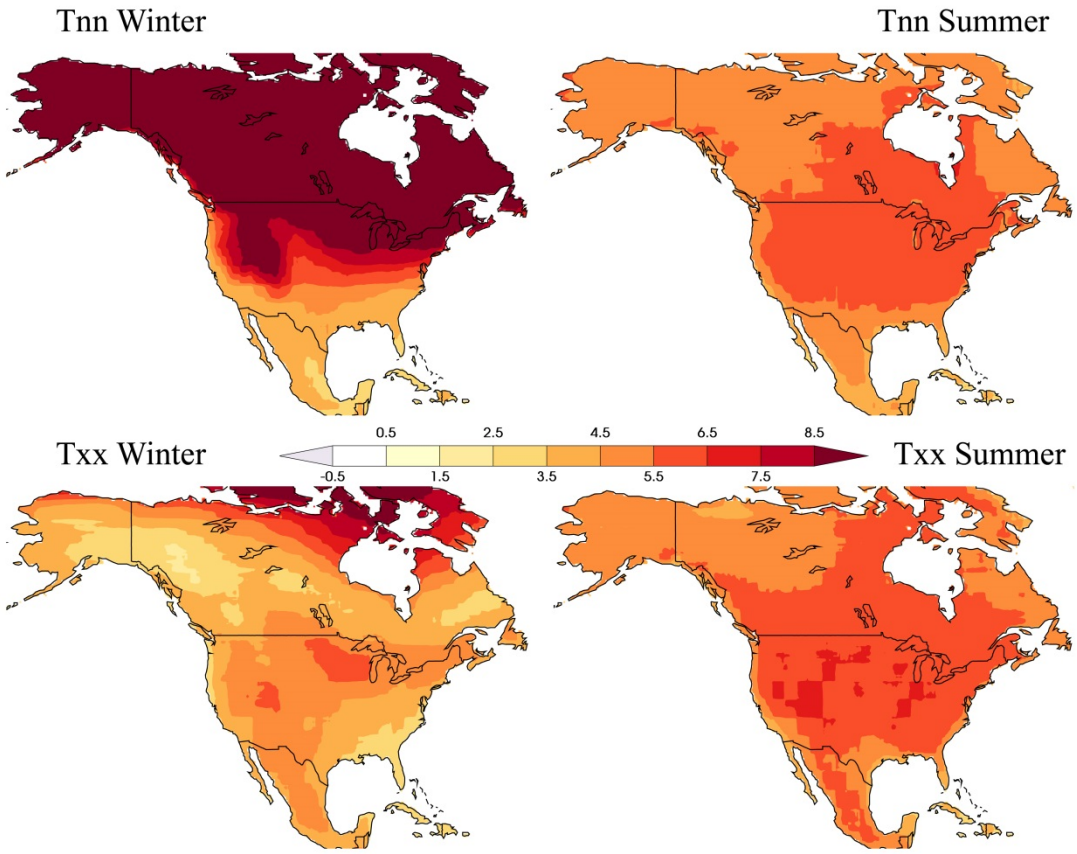


Fig. 12 Projected seasonal changes in North American extreme temperatures from the CMIP5 multi-model at the end of this century under the RCP8.5 forcing scenario. The reference period is 1985-2005 while the future period is 2080-2100. Winter changes are shown on the left while summer changes are shown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