Chapter 8
Methods of Projecting Future Changes in Extremes

Michael Wehner

Abstract This chapter examines some selected methods of projecting changes in extreme weather and climate statistics. Indices of extreme temperature and precipitation provide measures of moderately rare weather events that are straightforward to calculate. Drought indices provide measures of both agricultural and hydrological drought that are especially suitable for constructing multi-model ensemble projections of future change. Extreme value statistical theories are surveyed and provide methodologies for projecting the changes in frequency and severity of very rare temperature and precipitation events.

Future changes in the average climate virtually guarantee that changes in extreme weather events will follow. Such rare events are best described statistically as it is difficult, but perhaps not impossible, to directly link individual disasters to human-induced climate change. Examples of extreme weather events with severe consequences to society that are amenable to projection include heat waves, cold spells, floods, droughts and tropical cyclones. Confidence in projections of future changes in the severity and frequency of such events is increased if the mechanisms of change can be identified and understood. Equally important, however, is the rigorous quantification of the uncertainties in these projections. These uncertainties include the inherent natural variability of the climate system as well as limitations in both the climate models' fidelity and the statistical methods used to analyze their output.

The discussions about future changes in extreme events in recent climate change assessment reports (including the IPCC 4th Assessment Report and the US national assessments) did not generally focus on sophisticated statistical analyses. Rather, extremes were presented in these documents by a series of "extreme indices".

M. Wehner (✉)
Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS50F, Berkeley, CA 94720, USA
e-mail: mfwehner@lbl.gov

A. AghaKouchak et al. (eds.), *Extremes in a Changing Climate*, Water Science and Technology Library 65, DOI 10.1007/978-94-007-4479-0_8,

Introduced first by Frich et al. (2002), they are often referred to as the Frich indices. While many of these represent significant departures from the mean climate, they are by no means descriptive of rare events or the far tails of the temperature or precipitation distributions. The fundamental difference between these index based treatments and formal Extreme Value Theory descriptions of rare events illustrates the difficulties in nomenclature when discussing climate and weather extremes. What constitutes "extreme" varies greatly in the literature and depends highly on the application of the final results. This chapter will survey some of these methods of projecting changes in climate and weather.

8.1 Extreme Indices

A set of extreme indices was part of the data output specifications for the Coupled Model Intercomparison Project (CMIP3, see www-pcmdi.llnl.gov). Table 8.1 lists these ten pre-calculated statistics that were specified to be calculated for each year of the simulations. Code was provided to the climate modeling groups to calculate these fields although they could also be replicated from the archived daily averaged surface air temperature and precipitation rates. Most of these indices are clearly motivated by their relevance to climate change impacts, e.g. the number of frost days, the growing season length and the number of consecutive dry days. However, for a more general interpretation of the effect of global climate change on

Table 8.1 The Frich indices saved as annualized quantities for the CMIP3 coordinated numerical experiment

Index name	Units	Description
fd	Day	Total number of frost days (days with absolute minimum temperature $< 0°$ C)
etr	Kelvin	Intra-annual extreme temperature range: difference between the highest temperature of any given calendar year (T_h) and the lowest temperature of the same calendar year (T_i)
gsl	Day	Growing season length: period between when $T_{day} > 5°$ C for > 5 days and $T_{day} < 5°$ C for > 5 days
hwdi	Day	Heat wave duration index: maximum period > 5 consecutive days with $T_{max} > 5°$ C above the 1961–1990 daily T_{max} normal
tn90	%	Fraction (expressed as a percentage) of time $T_{min} > 90$th percentile of daily minimum temperature, where percentiles are for the 1961–1990 base period
r10	Day	No. of days with precipitation greater than or equal to 10 mm day^{-1}
cdd	Day	Maximum number of consecutive dry days ($R_{day} < 1$ mm)
r5d	kg m^{-2}	Maximum 5 days precipitation total
sdii	kg m^{-2} s^{-1}	Simple daily intensity index: annual total/number of R_{day} greater than or equal to 1 mm day^{-1}
r95t	%	Fraction (expressed as a percentage) of annual total precipitation due to events exceeding the 1961–1990 95th percentile

extreme events, these three indices and the others based on fixed threshold values are somewhat less useful (Alexander et al. 2006; Tebaldi et al. 2006). For instance, every day is a frost day in the very high latitudes but none are in the deep tropics. Similarly, 10 mm of precipitation in a single event is fairly common in tropical regions but impossible in many desert regions. Of more utility in this context are the percentile-based indices such as tn90 (hot nights) and r95t (very wet days). These two indices define base states (the 1961–1990 period) from which departures can be calculated anywhere on the planet. The bottom panel of Fig. 8.1 shows a CMIP3 multi-model projection of the change in r95t over land regions at the end of the twenty-first century under the SRES A1B forcing scenario. This index is defined as the percentage of annual total precipitation due to events exceeding the 1961–1990 95th percentile. During the base state period, this field would be uniformly 5%. End of twenty-first century values over land in this figure range from a low of 9% to a high of 54%. This increase in the index may be interpreted in the following sense: what might be currently considered very wet days (i.e. the top 5%) will occur from two to ten times more frequently in this future scenario. This also suggests that the shape of the distribution of daily precipitation must change in this scenario because the mean precipitation is not projected to change in a similar manner. Note that care should be exercised in interpreting such exceedance rate changes as sampling errors may play a role (Zhang et al. 2005).

In general, mean precipitation changes are a mix of increases and decreases and are smaller in magnitude as in the top panel of Fig. 8.1. In this multi-model example, an average projection is formed by equally weighting each climate model. Constructing weighted average projections based on model skill in replicating observed climate means or trends is a difficult task (Santer et al. 2009; Knutti et al. 2010a) and is presumably yet more difficult for extremes due to their less well characterized behavior. In Fig. 8.1, models with multiple realizations, if any, are ensemble averaged prior to inclusion into the multi-model result. Furthermore, the index is calculated on the models' native grids, then regridded to a common grid and masked prior to the multi-model averaging. These latter two points are the general practice in many climate change projection studies but have important implications for certain extremes, especially those related to precipitation, when evaluating model performance (Wehner et al. 2010).

Figure 8.2 shows a different way of representing the change in an extreme index. In this figure, a CMIP3 multi-model projection of the change in tn90 averaged over North American land regions under a variety of forcing scenarios is shown from the beginning of the twentieth century to the end of the twenty-first century. This index is defined as the percentage of time that daily minimum temperature exceeds the 90th percentile, of the 1961–1990 base period. This method of illustrating a projection, while lacking the spatial detail of the previous figure, allows the explicit depiction of projection uncertainty. The four major sources of projection uncertainty are the natural variability of the climate system, limited sample size (i.e. small ensembles and/or short time intervals), imperfect climate models (largely manifested by differences in climate model sensitivity to changes in greenhouse gas concentration but also realized in less well characterized ways for extremes), and

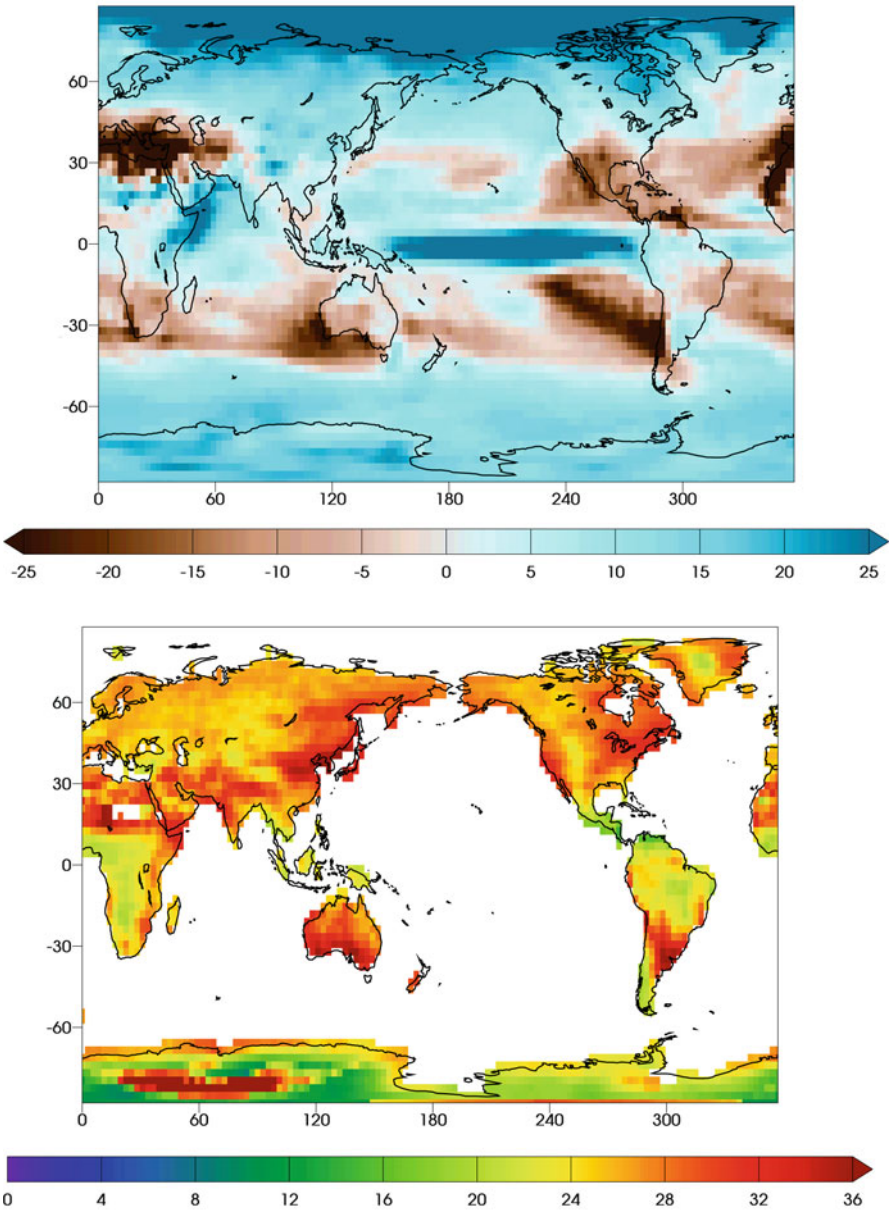

Fig. 8.1 A CMIP3 multi-model projection of changes in precipitation statistics at the end of the twenty-first century under the SRES A1B forcing scenario. (*Top panel*) Percent change of annual mean precipitation. (*Bottom panel*) Percentage of annual total precipitation due to events exceeding the 1961–1990 95th percentile (r95t). Ten different climate models were averaged with equal weighting in these projections (units: percent)

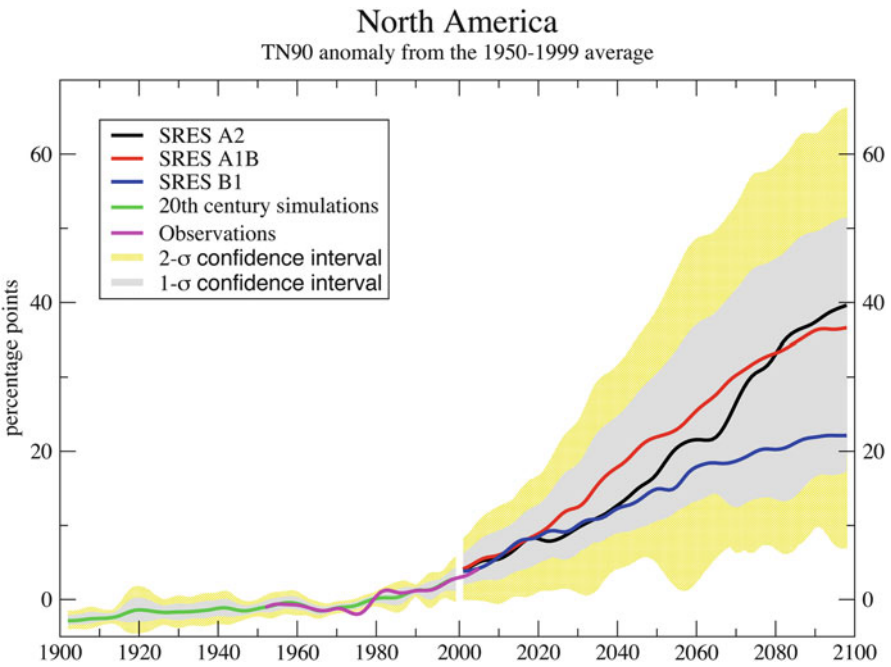

Fig. 8.2 A CMIP3 multi-model projection of the percentage of time the daily minimum temperature exceeds the 90th percentile of daily minimum temperature, calculated from 1961 to 1990 base period (tn90) at the end of the twenty-first century under the SRES A1B forcing scenario. Ten different climate models were averaged with equal weighting in this projection. A 13 point temporal filter is applied to all projections as in IPCC AR4 (units: percent)

the unpredictability of human behavior (i.e. the different scenarios). One method for quantifying the uncertainty from the imperfections of climate model is to calculate the variance in the projection across the ten climate models that provided this index to the CMIP3 database. In Fig. 8.2, one standard deviation across models is depicted by the gray shading and two standard deviations by the yellow shading. The envelopes plotted here are determined by the maximum spread across all three scenarios. One might also want to consider each scenario separately to base decisions on how significant the differences between the scenarios are. In this case, taken from the USGRCP report (Karl et al. 2009), this representation permitted usage of the "likelihood language" (Morgan et al. 2009). The gray shaded area represent the "likely" range of change (i.e. a 2 out of 3 chance of being a correct statement) while the yellow shaded bounds represent the "very likely" range of change (i.e. a 9 out of 10 chance of being a correct statement). However, given the limited set of available global models and that many of them are related, this measure of uncertainty does not completely sample the space of projection and underestimates to true uncertainty due to model deficiencies (Tebaldi and Knutti 2007; Knutti et al. 2010b).

As with projections of changes in mean quantities, the scenario uncertainty in the beginning of the twenty-first century is less than at the end (Hawkins and Sutton 2009; Yip et al. 2011). Comparison of Fig. 8.2 to similar figures for changes in mean temperature (Karl et al. 2009) reveals subtle differences in the timing of the separation of the low emissions scenario (B1, stabilizes at 550 ppm CO_2) from the high emissions scenarios (A2, business as usual). Furthermore, the relationship between the higher stabilization scenario (A1B, stabilizes at 720 ppm CO_2) from the business as usual scenario is quite different. In multi-model projections of the annual mean surface air temperature, the two scenarios are indistinguishable over most areas, including North America, until mid-century after which the business as usual scenario continues to increase and the stabilization scenario starts to stabilize. By the end of the twenty-first century, the differences between the scenarios are "likely" to be significant. In Fig. 8.2, warm nights increase in temperature over North America at the beginning of the twenty-first century at a greater rate in the stabilization scenario (A1B) than in the business as usual scenario (A2). This is followed by the A2 scenario catching up towards the end of the twenty-first century. There are enough differences in these forcing scenarios that one could hypothesize a plausible mechanism for why warm night temperatures might behave differently from annual mean temperatures. But the inter-model uncertainty in Fig. 8.2 is clearly large enough to prevent a conclusion that these differences are "likely" significant. Even a weaker statement about the significance of these differences is prevented by limitations in the sample size behind this index projection in comparison to that behind projection of mean temperature changes. For at the beginning of the twenty-first century, only about 10% of the daily temperature values are used in calculating the index as opposed to all of the values when calculating the annual mean. Although this fraction rises to about 50% towards the end of the century due to warming, the tn90 index remains a noisy quantity compared to annual mean temperatures. In order to ascertain, whether these tantalizing differences in the scenario behavior between the warm night index and the mean temperature are genuine, more realizations of each individual model are required. This will prove to be a recurring theme in ascertaining the significance of extreme changes. The exact details depend greatly on the variability of the quantity of interest and the magnitude of the differences (Wehner 2000).

Other extreme indices than that developed by Frich et al. (2002) can be useful tools in analyzing future climate change projections. In particular, there are a number of drought indices in wide use by the agricultural and other water intensive industries. Table 8.2 shows five drought indices that are provided to the public at regular intervals by the US National Climatic Data Center (NCDC) on their website, http://www.drought.noaa.gov/. A recent paper (Wehner et al. 2011) examined the performance and projections for the Palmer Drought Severity Index in the CMIP3 models finding wide variations between the models. In that study, the models simulated the observed PDSI much better after a bias correction procedure. Bias corrections can take many forms and can be useful in enhancing confidence in projections. Bias correction assumes that errors in the mean state may not influence trends or changes to the same degree. In many instances, this assumption can be

Table 8.2 The NCDC drought indices (see http://www.drought.noaa.gov/)

Drought index name	Units	Description
PDSI	Palmer Drought Severity Index	Duration and intensity of long-term drought
PHDI	Palmer Hydrological Drought Index	Similar to PDSI except measures longer term hydrological effects relevant to reservoir levels, groundwater levels, etc.
Z-index	Palmer Z Index	Short-term drought on a monthly scale
CMI	Crop Moisture Index	Short-term drought on a weekly scale
SPI	Standardized Precipitation Index	A normalized precipitation only index that is reported on time scales ranging from weeks to years.

tested by applying the correction over one part of an observational record and testing against another part. In the PDSI study, the input (monthly averaged temperature and precipitation) to the drought index calculation was corrected by applying a monthly varying climatological factor that altered the models' long term temperature and precipitation means to the observations but kept each models' particular variability intact. The PDSI is constructed to measure excursions from a neutral base state. Since the models' variability was not corrected, performance in replicating observed PDSI statistics ranged greatly. The simple land surface model contained in the PDSI algorithm is particularly sensitive to temperature leading to large projected changes in the severity and spatial extent of future drought in North America. However, this large temperature sensitivity caused large inter-model differences in these projections at the end of the century because of the large differences in climate model sensitivities to changes in atmospheric greenhouse gases.

This source of projection uncertainty can be reduced in a certain sense by rephrasing how the climate change question is asked. Most climate change projection questions ask something like: "What will happen at the end of the century?" Instead consider if a question such as the following is asked: "What will happen if the global mean temperature rises by 2.5 K?" In the former case, the time period is fixed but the different models exhibit vastly different warmings. In the latter case, the question of timing is foregone but at least the model states bear some resemblance to each other. In fact, under the SRES A1B scenario, the date at which the running decadal average global mean surface air temperature reaches 2.5 K over its preindustrial value ranges from 2038 in the most sensitive model to 2110 in the least sensitive model. The average date over all models to reach this amount of warming is 2070. Figure 8.3 shows maps of future North American PDSI under SRES A1B forcing and the associated inter-model uncertainty relevant to these two ways of posing future climate change questions. The upper two panels (a and c) show decadal averaged PDSI values and represent what the climatological values of PDSI would be relative to the current climatology. For interpretation of PDSI, drought is classified into the following categories: incipient ($-0.5 \geq \text{PDSI} > -1.0$), mild ($-1.0 \geq \text{PDSI} > -2.0$), moderate

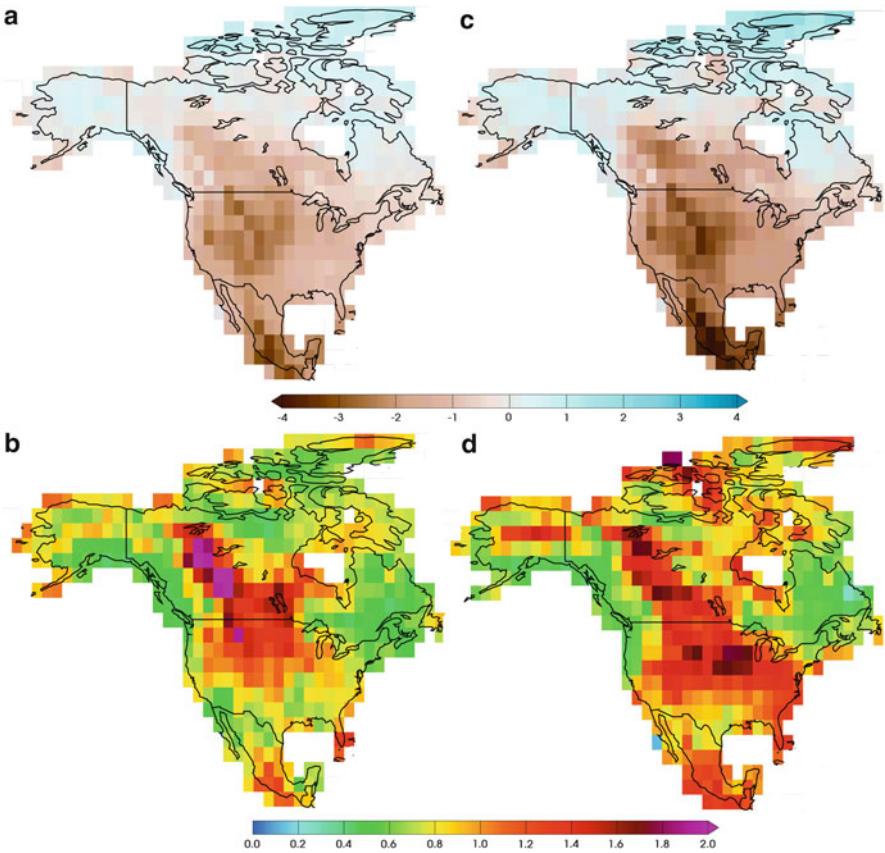

Fig. 8.3 (**a**) Multi-model average value of PDSI when the global average surface air temperature has increased 2.5 K over its 1900–1909 mean value (**b**) inter-model standard deviation of the values shown in panel (**a**, **c**) multi-model average value of PDSI for the decade centered at 2070 (**d**) inter-model standard deviation of the values shown in panel (**c**)

($-2.0 \geq \text{PDSI} > -3.0$), severe ($-3.0 \geq \text{PDSI} > -4.0$), and extreme ($-4.0 \geq \text{PDSI}$). The upper right panel shows the PDSI averaged over all models for the decade centered around 2070 (with an average model global warming of 2.5 K). In this projection, conditions currently considered severe drought would become normal in the western US. In parts of Mexico, conditions currently considered extreme drought would become normal. However, uncertainty in this projection, shown as the inter-model standard deviation in the lower right panel (d), is large in these regions. By changing the climate change question to ask what the value of PDSI would be under a 2.5 K global warming (which would occur on average at 2070), this inter-model uncertainty is reduced in most areas as shown in the lower left panel (b). The actual projection of drought severity is also reduced as seen in the upper left panel (a), reflecting a nonlinear dependence of PDSI on temperature.

8.2 Extreme Value Theory Methods

Numerous studies in the literature as well as much of the contents of this book utilize sophisticated extreme value statistics to explore questions of climate change. These techniques differ from the index-based methods described in the previous section principally in their ability to quantify the statistical behavior of much rarer events. Rather than review the details of how and when extreme value theory methods may be applied to climate and weather datasets, this section discusses aspects of interpretation of results from these statistical formalisms in a context of climate change.

The parameters describing the generalized extreme value distribution and the generalized Pareto distribution can often offer interesting insight. However, these fields are not closely tied to observable quantities and are generally of limited utility to the users of climate change projections. Design engineers and other parties interested in climate change impacts are more concerned with how the limitations of their particular systems might be exceeded. Return value and/or return time often can provide the critical information necessary to make informed decisions about the impacts of rare weather and climate events. Whether the analysis takes a block maxima or threshold approach, these application relevant fields are readily calculated if the distribution parameters can be satisfactorily fit to the extreme data.

Extreme value theory (EVT) is often used to describe how extreme weather behaves in a changing climate by analyzing high frequency (i.e. daily) modeled or observed datasets. Return values from the fitted EVT distributions are defined over a fixed specified period, for instance, T with units in years. In a stationary climate, the return value can be interpreted as the value of the data that would be realized on average once every T years over a very long period of time. By introducing time as a covariate, EVT can be generalized to non-stationary datasets (Brown et al. 2008; Smith, Private communication, 2010). In a changing climate, this explanation loses meaning for a time dependent return value. Instead, a more appropriate alternative interpretation is that the return value at a given time represents the value that has a 1/T chance of occurring that year in the dataset.

Return time offers a slightly different way to express the same concepts. In a stationary climate, the return time is the average time between instances that the data take to reach or exceed a specified value over the course of a very long time. In a non-stationary climate, the return time for a fixed specified value would be a time dependent quantity. The inverse of the return time would be the chance that the specified value would be achieved in that year.

Uncertainty in return value and return time estimates depends on the magnitude of the time scale of interest in relation to the length of the datasets Wehner (2010). When this time scale is much less than the dataset length, the return values have likely been realized in the datasets and uncertainty is lower. When the time scales are much larger than the dataset lengths, the EVT estimates are extrapolations outside the datasets and uncertainty is higher. However, if the asymptotic assumptions of the EVT are valid, estimated return values and return times can be reasonable in extrapolated cases. The generalization of EVT to treat time dependent datasets can

help reduce uncertainty by allowing the consideration of longer datasets. However, care must be exercised as these generalizations assume specific time dependences of the EVT distribution parameters. These can be linear, quadratic or even higher order in the Smith (Private communication, 2010) formalism but it is not always clear how to generate the best fits. In fact, the observed climate change has not been particularly linear and future changes may not even be monotonic if drastic remediation procedures are taken.

The actual climate system is of course limited to the single world that actually exists. Climate model simulations have no such limitations as they are routinely integrated in statistically independent realizations to be combined into large ensembles by varying initial conditions. If one assumes quasi-stationarity over short periods of time, these independent realizations can be combined into much longer datasets and stationary EVT used to provide accurate estimates of the distribution parameters. The length of such a period depends greatly on the variable of interest as well as the rate of climate change.

Stationarity would be guaranteed if linear detrending is applied over these short periods. In the literature (for instance Kharin et al. 2007), it is not uncommon to assume a decade or two. Although Santer et al. (2011) showed that any individual decade in the last century might exhibit observed positive or negative temperature trends, they also showed that over a large sample of decades, a statistically significant positive trend can be found. This suggests that detrending is prudent when combining intervals over individual realizations to construct a larger stationary dataset for EVT analysis. Ensemble sizes in the CMIP3 database ranged between three and eight, if multiple realizations were performed at all. In the CMIP5 specifications, a minimum of ten realizations is called for in the "Tier 1" experiments (Taylor et al. 2009). This then affords the opportunity to concatenate detrended decadal segments to build quasi-stationary datasets of about 100 years in length representing any time period during the integration.

Hence, there are two EVT methods that can be used to make projections of future changes in extreme weather event statistics. The first method is to fit non-stationary datasets with time dependent EVT distributions. The advantage in this approach is that the single realization of the observed climate system can be treated without any ad hoc assumptions of stationarity. The length of the record should be chosen carefully such that the trend is well fit by the specified time dependence. For multiple realizations of a single climate model, each realization should be treated separately in this method and ensemble mean return values and/or return times calculated. A continuous picture of change including trends is provided by this method. Additionally, a measure of the models' internal variability can be obtained by calculating the inter-realization variance to provide insight into this source of projection uncertainty. The second method is applicable to climate models with multiple realizations. In this approach, short intervals from each realization are concatenated to form a larger dataset. Detrending of the segments prior to concatenation is often desirable. Fitting a stationary EVT distribution to concatenated datasets formed at different times permits changes in return values and/or return times to be directly calculated. In both methods, the uncertainty from

estimation of the fitted distribution parameters can be estimated by the scheme outlined by Hosking and Wallis (1997). This technique involves first estimating the distribution parameters for the actual dataset then generating random datasets distributed by the EVT distribution defined by those parameters. To estimate uncertainty, each of the random distributions is fit to an EVT with appropriate parameters and return fields estimated along with their variances. This forms a measure of uncertainty associated with the finiteness of the data. For the large ensembles promised with CMIP5, this uncertainty should be reduced.

Presentation of the results from an EVT analysis of climate change poses serious challenges as the concept of return value and return time may not be intuitive to the non-specialist. Changes in return value (for a fixed return period) can be expressed in similar forms to widely published changes in mean values. For instance, Fig. 8.4 shows a multi-model CMIP3 projection of the end of century changes under a business as usual scenario (SRES A1B) of the 20 year return value of the annual maximum daily average surface air temperature (upper panel) and the annual mean surface air temperature (lower panel). Exhibiting projected changes in pairs of figures such as these allows discussion of the differences between them. In this case, large changes in the 20-year return value are confined to land masses and are generally larger than changes in the annual mean. Not shown in these figures are the seasonal behaviors of projected changes that can reveal mechanistic insights. This seasonal aspect is particularly important in the interpretation of changes in precipitation extremes. Also not shown are changes in extremes of minima temperatures that exhibit different behaviors than the changes in maxima temperatures, again providing opportunities for understanding physical mechanisms of change.

Changes in severity of rare weather events is only a part of an EVT analysis. Changes in frequency may be yet more important. Figure 8.5 attempts to illustrate this point in the upper panel by showing the return time in the future for daily surface air temperatures exceeding the present day 20-year return value. In this case, the future return time is projected to become less than 20 years over most of the globe. Alternatively, the lower panel of Fig. 8.5 shows the number of times in a future 20 year period that this same temperature threshold can be expected to be exceeded on average. If the climate does not change, this number would be one. However, for annual maximum daily temperatures, the value is much greater than that over most of the globe. The EVT data used in both Figs. 8.4, 8.5 come from the same analysis. These figures reveal that warm weather events currently considered rare (once every 20 years) are projected to become relatively commonplace and that warm events of a fixed rarity are projected to become more severe.

8.3 Multi-Variate Climate and Weather Extremes

The literature of multi-variate extreme value statistics is well developed (see Chap. 7). However, it has not seen significant application to climate change projections or historical analyses despite an urgent need. For instance, consider

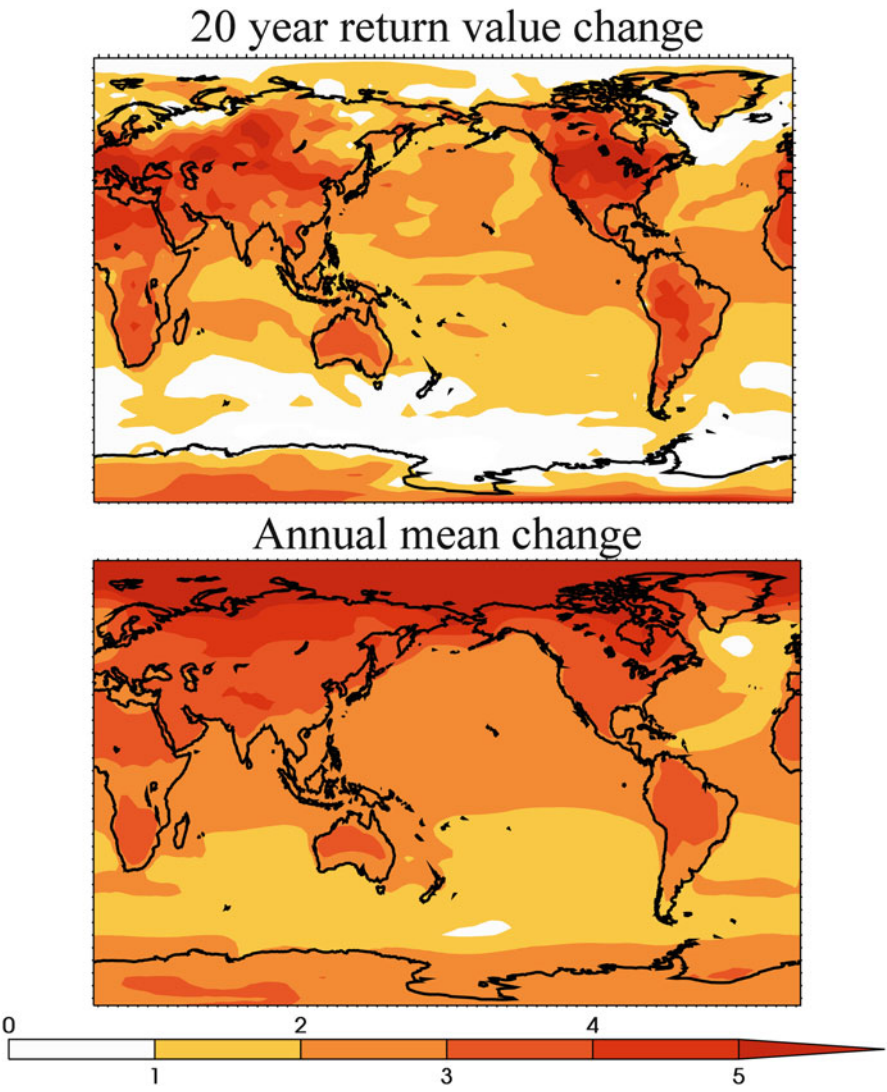

Fig. 8.4 Changes in the end of twenty-first century surface air temperature properties relative to the end of the twentieth century under SRES A1B forcing from the CMIP3 models. *Upper panel*: Change in 20-year return value of the annual maximum daily averaged temperature. *Lower panel*: Change in annual mean temperature (units: kelvin)

hot, dry and windy events versus hot, moist and stagnant events. The impacts of such events are very different. The former may lead to increased risk of fires while the latter may lead to increased human mortality through heatstroke or air quality issues. In both cases, at least one of the salient variables is not extreme in itself. In fact, it is often the combination of multiple events, each common in isolation, that is considered rare and/or dangerous.

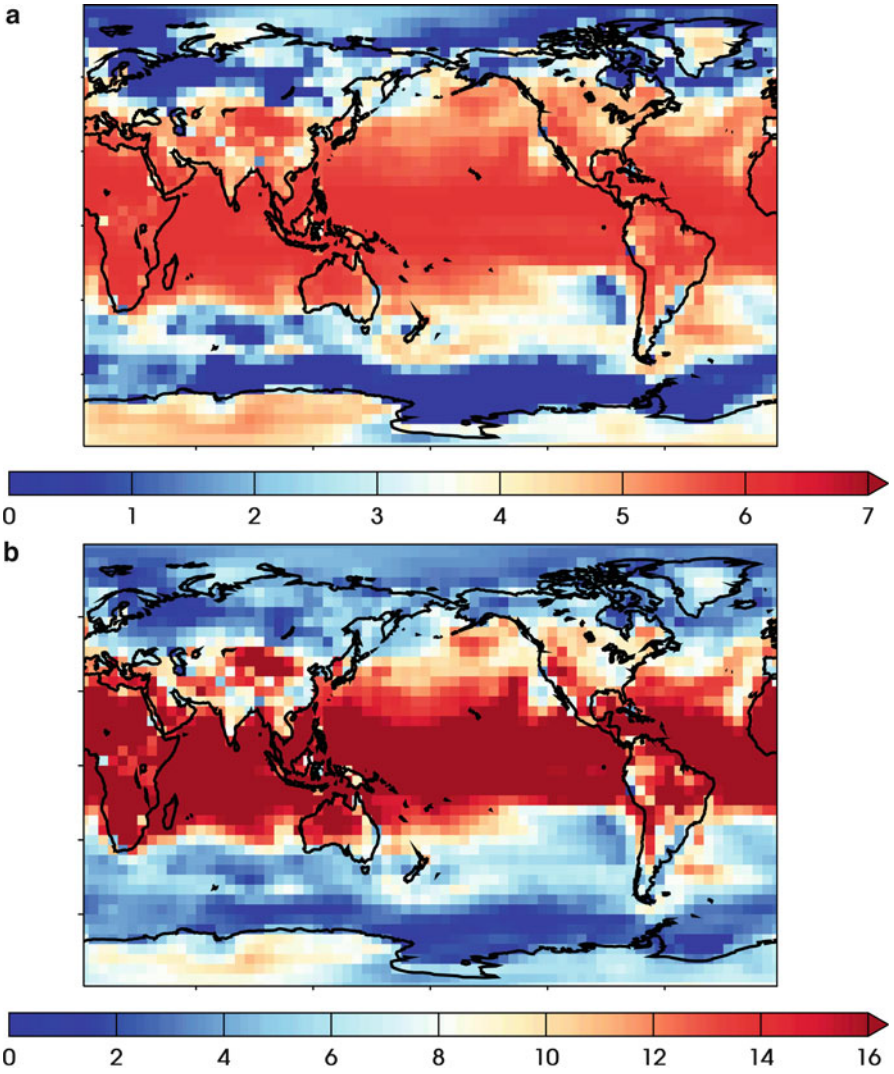

Fig. 8.5 Upper panel (**a**) The projected return time at the end of the twenty-first century under SRES A1B forcing associated with daily temperature threshold defined by the end of the twentieth century 20-year return value of the annual maximum daily averaged surface air temperature (units: years). Lower panel (**b**) The number of occurrences per 20 year period at the end of the twenty-first century when the daily averaged surface air temperature exceeds that same threshold. If the climate had not changed, this number would be one (units: dimensionless)

Two multivariate indices are in common usage in weather forecasting. Similar to the drought indices discussed above, they can be used to define the frequency and severity of extreme events in climate change projections. The first of these is the "Heat Index" (HI) that combines air temperature and relative humidity (Steadman

1979a, b). The second of these is the "Wind Chill" index combining temperature and wind speed (Osczevski and Bluestein 2005). Both of these indices, expressed in degrees, are used to estimate effects on the human body and are often said to describe how hot or cold "it feels". The derivations of both indices are rather involved and often implemented via tabular lookups or fitted polynomials. Delworth et al. (1999) projected that patterns of future increases in HI are largely dependent on temperature increases but are amplified by changes in moisture, illustrating a complex interplay between variables.

8.4 Summary

Changes in climate and weather extremes can be projected by a wide variety of methods. Indices and thresholds defined by their relevance to climate change impacts can be particularly useful. Changes in truly rare events, often associated with dire consequences, are well described by return value or return time changes using extreme value theories. Projecting changes in multi-variate climate and weather extremes is still a developing skill. The description of changes in rare compound events via multi-variate extreme value theory would be an important advance in the field.

Acknowledgement This work was supported by the Regional and Global Climate Modeling and the Earth System Modeling Programs of the Office of Biological and Environmental Research in the Department of Energy Office of Science under contract number DE-AC02-05CH11231.

References

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

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

Delworth TL, Mahlman JD, Knutson TR (1999) Changes in heat index associated with CO_2-induced global warming. Clim Chang 43:369–386

Frich P, Alexander LV, Della-Manta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. Clim Res 19:193–212

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc 90(8):1095–1107

Hosking JRM, Wallis JR (1997) Regional frequency analysis, an approach based on L-moments. Cambridge University Press, Cambridge/New York

Karl TR, Mellilo JM, Peterson TC (eds) (2009) Global climate change impacts in the United States: a state of knowledge report Cambridge [England]. Cambridge University Press, Cambridge/New York. Available at www.globalchange.gov

Kharin VV, Zwiers FW, Zhang X, Hegerl GC (2007) Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. J Clim 20:1419–1444

Knutti R, Abramowitz G, Collins M, Eyring V, Gleckler PJ, Hewitson B, Mearns L (2010a) Good practice guidance paper on assessing and combining multi model climate projections. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Midgley PM (eds) Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland

Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl G (2010b) Challenges in combining projections from multiple climate models. J Clim 23:2739–2758

Morgan MG, Dowlatabadi H, Henrion M, Keith D, Lempert R, McBride S, Small M, Wilbanks T (eds) (2009) Best practice approaches for characterizing, communicating, and incorporating scientific uncertainty in decisionmaking. A report by the Climate Change Science Program and the Subcommittee on Global Change Research. National Oceanic and Atmospheric Administration, Washington, DC, 96pp

Osczevski R, Bluestein M (2005) The new wind chill equivalent temperature chart. Bull Am Meteorol Soc 86:1453–1458

Santer BD, Taylor KE, Gleckler PJ, Bonfils C, Barnett TP, Pierce DW, Wigley TML, Mears C, Wentz FJ, Brueggemann W, Gillett NP, Klein SA, Solomon S, Stott PA, Wehner MF (2009) Incorporating model quality information in climate change detection and attribution studies. Proc Natl Acad Sci. doi:10.1073/pnas.0901736106

Santer BD, Mears C, Doutriaux C, Gleckler P, Wigley T, Gillett N, Ivanova D, Karl T, Lanzante J, Meehl G, Stott P, Taylor K, Thorne P, Wehner M, Wentz F (2011) Separating signal and noise in atmospheric temperature changes: the importance of timescale. J Geophys Res Atm 116:D22105. doi:10.1029/2011JD016263

Steadman RG (1979a) The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. J Appl Meteorol 18:861–873

Steadman RG (1979b) The assessment of sultriness. Part II: Effects of wind, extra radiation and barometric pressure on apparent temperature. J Appl Meteorol 18:874–885

Taylor KE, Stouffer RJ, Meehl GA (2009) A summary of the CMIP5 experiment design. http://www-pcmdi.llnl.gov/

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Phil Trans R Soc MathPhys Eng Sci 365:2053–2075. doi:2010.1098/rsta.2007.2076

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA (2006) Going to the extremes: an intercomparison of model-simulated historical and future changes in extreme events. Clim Chang 79:185–211. doi:10.1007/s10584-006-9051-4

Wehner MF (2000) Determination of the sampling size of AGCM ensemble simulations. Clim Dyn 16:321–331

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

Wehner MF, Smith R, Duffy P, Bala G (2010) The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. Clim Dyn 32:241–247. doi:10.1007/s00382-009-0656-y

Wehner MF, Easterling DR, Lawrimore JH, Heim RR Jr, Vose RS, Santer B (2011) Projections of future drought in the continental United States and Mexico. J Hydrometeorol 12:1359–1377. doi:10.1175/2011JHM1351.1

Yip S, Ferro CAT, Stephenson DB, Hawkins E (2011) A simple, coherent framework for partitioning uncertainty in climate predictions, to appear in. J Clim. doi:10.1175/2011JCLI4085.1

Zhang X, Hegerl G, Zwiers FW, Kenyon J (2005) Avoiding inhomogeneity in percentile-based indices of temperature extremes. J Clim 18:1641–1651