Climatic Change (2010) 100:71–76
DOI 10.1007/s10584-010-9834-5

Statistics of extremes in climate change

Richard W. Katz

Received: 2 January 2010 / Accepted: 9 March 2010 / Published online: 19 May 2010

Abstract This editorial essay concerns the use (or lack thereof) of the statistics of extremes in climate change research. So far, the statistical theory of extreme values has been primarily applied to climate under the assumption of stationarity. How this theory can be applied in the context of climate change, including implications for the analysis of the economic impacts of extremes, is described. Future research challenges include the statistical modeling of complex extreme events, such as heat waves, and taking into account spatial dependence in the statistical modeling of extremes for fields of climate observations or of numerical model output. Addressing these challenges will require increased collaboration between climate scientists and statisticians.

> *In order to apply any theory we have to suppose that the data are homogeneous, i.e. that no systematical change of climate and no important change in the basin have occurred within the observation period and that no such changes will take place in the period for which extrapolations are made* (Gumbel 1941, p. 187).

1 Historical perspective

The above quote is from Emil Gumbel, a pioneer in the application of the statistics of extremes to a number of fields including climate and hydrology. Several decades ago, he was rightly concerned that the recently developed statistical theory of extreme values could apparently only be applied under the assumption of stationarity. Stimulated in large part by Gumbel (among other things, writing the first book on the statistics of extremes, Gumbel 1958), there have been many areas to which this theory has been successfully applied, such as in engineering design (e.g., for flood plain management), but nearly always under the paradigm of stationarity.

R. W. Katz (✉)
Institute for Mathematics Applied to Geosciences, NCAR, Boulder, CO 80307, USA
e-mail: rwk@ucar.edu

Since the time of the quote by Gumbel, the statistical theory of extreme values has undergone extensive further development, including devising approaches to incorporate non-stationarity (e.g., Coles 2001). Nevertheless, the statistics of extremes has only rarely been applied in the context of climate change until quite recently. Statistical techniques either designed for dealing with trends in averages (i.e., least squares regression analysis) or at least not specifically designed for extremes (i.e., nonparametric techniques) have been used instead. Perhaps a less obvious drawback is the common practice of rigidly defining extreme events (e.g., in terms of a threshold relevant for societal impacts) and then only examining their frequency of occurrence, in effect, ignoring any variations in their severity.

At least as early as Wigley (1985, 1988), it has been appreciated that trends in the frequency and severity of extreme climate events might well not closely resemble the corresponding ones for averages. Strictly speaking, Wigley's research made only quite limited use of extreme value theory, his goal being to provide a simple pedagogical example. He addressed the question of how much the return period of an extreme event, of the form exceeding a high threshold, would decrease as a function of a gradual increasing trend in the mean of the distribution. Other early work on this topic included Mearns et al. (1984), who treated more complex forms of extreme events than did Wigley, such as a run of temperatures above a high threshold as might be deleterious to crop yields, but not making any explicit use of extreme value theory. Katz and Brown (1992) extended the results of Wigley (1988), making more explicit use of extreme value theory, in showing that the frequency of extreme events, whether of the form of a single observation or of the maximum of a sequence of observations, exceeding a high threshold is more sensitive to changes in the standard deviation (or, more generally, the scale parameter) than to changes in the mean (or, more generally, the location parameter) of a distribution.

It is fitting that the Wigley (1988) paper, which originally appeared in a publication with limited distribution, has recently been reprinted in *Climatic Change*, along with a commentary by Cooley (2009) on the potential application of extreme value theory to climate change. Still Cooley focused rather narrowly on how Wigley's work might be extended, whereas the present editorial essay attempts to place such research within a broader historical perspective.

2 Extreme value analysis under non-stationarity

If only analyzing whether there is a trend in the frequency of occurrence of an extreme event, not its severity, then the statistical technique of Poisson regression can be used. For example, Katz (2002) and Solow and Moore (2000) fitted trends in the frequency of hurricanes using this technique (e.g., assuming a linear trend in the logarithm of the rate parameter of the Poisson distribution). The Poisson distribution arises from the "law of small numbers," with Poisson regression being an extension in which the rate parameter is non-constant (e.g., can have a time trend). What is not nearly as well appreciated is that this sort of regression-like approach can be readily applied to modeling trends in the severity, or simultaneously in the frequency and severity, of extreme events.

The classical extreme value theory of Gumbel's era focused on block maxima (e.g., the single highest daily precipitation amount over an entire year or the highest temperature over a season). Under a wide range of conditions, block maxima should be approximately distributed as the generalized extreme value (GEV; e.g., Coles 2001). The GEV distribution is quite flexible with three parameters: (1) the *location parameter* specifies the center of the distribution; (2) the *scale parameter* determines the size of deviations about the location parameter; and (3) the *shape parameter* governs how rapidly the upper tail decays.

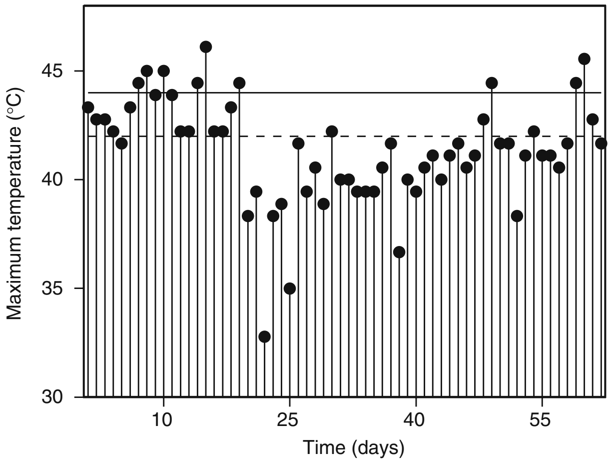
With the recent availability of statistical software for extreme value analysis (e.g., Stephenson and Gilleland 2006), it is straightforward to introduce trends into the parameters of the GEV (e.g., a linear trend in the location parameter and/or in the logarithm of the scale parameter). For examples of this approach, see Coles (2001) or Katz et al. (2002). Cooley (2009) applied this technique to a time series of annual maximum temperature data for central England, although Nadarajah (2005) had already demonstrated the use of the same technique to detect trends in temperature extremes in an article appearing in *Climatic Change* several years before.

A similar, but somewhat simpler to implement, analysis by Zwiers and Kharin (1998) compared the parameters of the GEV distribution fitted to extremes simulated by a general circulation model (GCM) under a control run to the corresponding ones under doubled CO_2. This paper served as a catalyst for making acceptable the application of the statistics of extremes to the output of numerical models of the climate system.

More modern extreme value theory, not fully developed until well after Gumbel's era, attempts to make use of more of the available information about the upper tail of a distribution than just the block maxima. For example, if the highest and second highest daily precipitation amounts over the entire historical record occurred during the same year, the second highest value would be ignored in the block maxima approach. The original idea is due to hydrologists motivated by the statistical modeling of floods. They devised a "peaks-over-threshold" approach consisting of two components: (1) a Poisson process governing the occurrence of an extreme event, in the form of exceeding a high threshold (i.e., consistent with a Poisson distribution for the number of exceedances within a given time interval); and (2) a generalized Pareto (GP) distribution for the "excess" over a high threshold. In extreme value theory, the GP distribution arises as an approximation for the excess over a high threshold, consistent with the GEV for block maxima (e.g., its shape parameter has the same interpretation, Coles 2001). As an example, Fig. 1 show the times series of daily maximum temperature during July–August 1948 at Phoenix, AZ. On day 1 (i.e., 1 July 1948), for example, the maximum temperature exceeded a threshold of 42°C with an excess of about 1.3°C.

The modern point process approach combines these two component of the peaks-over-threshold technique into a single two-dimensional process (Davison and Smith 1990). In Fig. 1, all the points above the threshold of 42°C can be viewed as distributed in a two-dimensional space, with time being one dimension and the excess in temperature above the threshold being the other. If the threshold is high enough, then this point process is approximately a two-dimensional Poisson process, with non-constant (or non-homogenous) rate parameter because higher excesses should be less likely than smaller excesses. Because the two-dimensional Poisson process is already non-homogenous, it is straightforward to allow any of its parameters to depend on time as well. Although fitting such point process models is challenging numerically, software is available (e.g., the extRemes package in the open source

Fig. 1 Time series of daily maximum temperature during July–August 1948 at Phoenix, AZ, with two thresholds, one at 42°C (*dashed line*) and another at 44°C (*solid line*), to motivate point process approach to extreme value analysis



statistical programming language R, cran.r-project.org/web/packages/extRemes/, also see Stephenson and Gilleland 2006).

This two-dimensional point process can be thought of as applying to any sufficiently high threshold, not just the actual threshold used in fitting the model. For instance, parameters corresponding to the higher threshold of 44°C (also shown in Fig. 1) can be derived indirectly from those for the lower threshold of 42°C (in particular, the shape parameter remains the same). As such, there is no need for a rigid focus on a single threshold (e.g., as based on impact considerations).

One complication is that exceedances over a high threshold tend to occur in clusters, especially for persistent climate variables such as temperature. We postpone addressing this issue until the treatment of more complex extreme events, such as hot spells/heat waves, in Section 4.

3 Economic impact of extremes

The statistical theory of extreme values also has more direct implications for the economic impact of extreme climate events. Near the end of the 19th century, Vilfredo Pareto first postulated the distribution, now bearing his name, as a model for high income or wealth. So it would be natural to anticipate that the distribution of economic damage from extreme events could have a heavy tail. In fact, the Pareto distribution (i.e., the heavy-tailed type of the GP, in which the upper tail decays at such a slow rate that higher-order moments are infinite) has been used to model the economic damage from hurricanes (Jagger et al. 2008; Katz 2002) and, at least indirectly, that from high wind storms (Dorland et al. 1999). Given much evidence that stream flow itself has a heavy tail, it is an immediate consequence that the economic damage from floods should be heavy-tailed as well (Katz et al. 2002).

The possibility that the distribution of economic damage from climate extremes can be heavy-tailed would seem relevant to the debate about whether conventional cost-benefit analysis can adequately deal with the economics of climate change. Yet heavy-tailed distributions (or a broader class of “fat”-tailed distributions; e.g., the lognormal is fat-tailed but not heavy-tailed) have entered into this debate only by an indirect mechanism, involving contentious concepts such as "deep" uncertainty and "irreversible" climate change (e.g., Weitzman 2009). For instance, the distribution of economic damage from climate extremes being heavy-tailed (strictly speaking, only fat-tailed is required) has a fundamental effect on the probability of eventual insolvency for insurance companies in the business of providing coverage for catastrophic events, even under an unchanging climate (e.g., Embrechts et al. 1997).

4 Future challenges

To be sure, resistance to the use of the statistics of extremes remains within the climate change research community. Among other things, it continues to be argued that shifts in extremes can be more reliably derived indirectly from changes in the overall probability distribution of a climate variable (e.g., shifts in the mean and standard deviation) than through direct statistical modeling of extremes (e.g., Ballester et al. 2010; Solow 1999). At first glance, this argument sounds consistent with the method originally used by Wigley (1988). Yet, as already mentioned, his purpose was clearly pedagogical and would not necessarily justify the use of this approach in practice. This issue remains unresolved because of the conflict between: (1) the physical appeal of a consistent shift in the entire distribution, including the extremes, of a climate variable; and (2) the practical concern that trends in extremes may well be of a different nature than those driven primarily by the center of the distribution.

Other obstacles to the use of the statistics of extremes include its difficulty in being applied to more complex forms of events, such as hot spells or heat waves (Meehl and Tebaldi 2004). "De-clustering," a technique to deal with the temporal dependence of extremes (Coles 2001), could be readily extended to model extreme spells, on the one hand, consistent with extreme value theory, on the other hand, simple enough to be applied under climate change (i.e., to detect trends in the spell frequency, duration, and intensity).

A challenging problem concerns the statistical modeling of climate extremes across space. Beside being important for its own sake, such spatial dependence needs to be taken into account when modeling extremes for fields of climate observations or of numerical model output. Even if the purpose is limited to providing local information about trends in extremes, it still would be beneficial to simultaneously model the same extremes over an adjacent region ("borrowing strength" in statistical terminology or akin to regional analysis for flood estimation in hydrology).

Finally, the question remains of the terms in which to most effectively convey the risk of extreme events under a changing climate. The familiar, if often misunderstood, concepts of return period and return level (e.g., the proverbial "100-year flood"), strictly speaking, no longer apply in a non-stationary climate. The most appropriate way to extend these concepts is an open question.

Satisfactorily addressing these challenges and obstacles will require an interdisciplinary approach, calling for increased collaboration between climate scientists and statisticians.

Acknowledgements Research supported by funds from the NCAR Weather and Climate Impact Assessment Science Program. NCAR is sponsored by the National Science Foundation.

References

Ballester J, Giorgi F, Rodó J (2010) Changes in European temperature extremes can be predicted from changes in PDF central statistics: a letter. Clim Change 98:277–284

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London

Cooley D (2009) Extreme value analysis and the study of climate change: a commentary on Wigley 1988. Clim Change 97:77–83

Davison AC, Smith RL (1990) Models for exceedances over high thresholds. J R Stat Soc, B 52: 393–442

Dorland C, Tol RSJ, Palutikof JP (1999) Vulnerability of the Netherlands and Northwest Europe to storm damage under climate change. Clim Change 43:513–535

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events for insurance and finance. Springer, Berlin

Gumbel EJ (1941) The return period of flood flows. Ann Math Stat 12:163–190

Gumbel EJ (1958) Statistics of extremes. Columbia University Press, New York

Jagger TH, Elsner JB, Saunders MA (2008) Forecasting US insured hurricane losses. In: Diaz HF, Murnane RJ (eds) Climate extremes and society. Cambridge University Press, Cambridge, pp 189–208

Katz RW (2002) Stochastic modeling of hurricane damage. J Appl Meteorol 41:754–762

Katz RW, Brown BG (1992) Extreme events in a changing climate: variability is more important than averages. Clim Change 21:289–302

Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. Adv Water Resour 25:1287–1304

Mearns LO, Katz RW, Schneider SH (1984) Extreme high-temperature events: changes in their probabilities with changes in mean temperature. J Clim Appl Meteorol 23:1601–1613

Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305:994–997

Nadarajah S (2005) Extremes of daily rainfall in West Central Florida. Clim Change 69:325–342

Solow AR (1999) On testing for change in extreme events. Clim Change 42:341–349

Solow AR, Moore L (2000) Testing for trend in a partially incomplete hurricane record. J Clim 13:3696–3699

Stephenson A, Gilleland E (2006) Software for the analysis of extreme events: the current state and future directions. Extremes 8:87–109

Weitzman ML (2009) On modeling and interpreting the economics of catastrophic climate change. Rev Econ Stat 91:1–19

Wigley TML (1985) Impact of extreme events. Nature 316:106–107

Wigley TML (1988) The effect of changing climate on the frequency of absolute extreme events. Clim Monit 17:44–55 (reprinted in Climatic Change (2009) 97:67–76)

Zwiers FW, Kharin VV (1998) Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling. J Clim 11:2200–2222