EOS

EOS, TRANSACTIONS, AMERICAN GEOPHYSICAL UNION

New Software to Analyze How Extremes Change Over Time

PAGES 13–14

Average behavior is often studied, with well-developed techniques from the field of statistics allowing for inferences to be readily made. However, in many atmospheric, hydrologic, and other geophysical problems, extremes are of the greatest interest. The usual statistical methods for averages do not correctly inform scientists about extremes, but a more specialized area of statistical research focused on extremes can be applied instead. Where the normal distribution has theoretical support for use with averages, other forms of distribution have similar theoretical support for the extremes: the generalized extreme value (GEV) distribution for analyzing extreme statistics like annual maxima or minima, and the generalized Pareto (GP) distribution for excesses over a high threshold (or deficits below a low threshold) [e.g., *Coles,* 2001]. For more background on the statistics of geophysical extremes, especially those for weather and climate, see http://www.isse.ucar.edu/extremevalues/extreme.html.

With these same techniques it is now also possible to characterize extreme behavior changes over time by expressing one or more of the parameters of these distributions as a function of time. For example, given that global mean temperature is increasing over time, these techniques can be used to quantify the corresponding changes in extreme high-temperature events.

Engineering design in areas such as hydrology has long been making use of these distributions in the stationary case (i.e., under the assumption of an unchanging climate, in which the probability of extreme events such as flooding is treated as constant from one year to the next). Yet there is a consensus in the scientific community that climate should no longer be regarded as stationary (see, e.g., the AGU Position Statement "Human Impacts on Climate" at http://www.agu.org/sci_pol/positions/climate_change2008.shtml). In particular, through their provocative paper entitled "Stationarity is dead: Whither water management?," *Milly et al.* [2008] recently raised the issue that engineering design needs to take into account possible shifts in the probabilities of extreme hydrological events.

The Extremes Toolkit (extRemes) provides one avenue to address this issue. This software was designed to shorten the learning curve for making use of the statistical theory of extremes in analyzing geophysical events and their changes over time.

Using extRemes

One typical hydrologic engineering problem involves determining floodplain boundaries on the basis of the historical time series of annual peak flow. Annual peak flow is an extreme statistic, the maximum instantaneous discharge of a river at a gauging station over a given year. For example, an estimate is commonly needed of the "100-year flood," i.e., the value of annual peak flow with a 1% chance of being exceeded in a given year (technically, the 100-year "return level" or 0.99 quantile of the probability distribution of annual peak flow). In such a case, extRemes allows users to input data on the historical time series of annual peak flow not only to enable the estimation of this return level under stationary conditions but also to determine whether and how this level might have changed over the historical record.

The software for extRemes is freely available and utilizes graphical user interfaces (GUIs) for most of the functionality. It focuses on weather and climate applications, and an extensive manual accompanies it with support via its Web page (http://www.isse.ucar.edu/extremevalues/evtk.html; see Figure 1 for example output along with the corresponding GUIs). It is available as an R package, which is a widely used statistical language and environment that is also freely available (http://www.r-project.org). Other functionality has recently been added for which GUIs are not available, but extensive help files, as well as information on updates or new versions, are provided on the extRemes Web site.

Figure 2 provides an example of how it is possible to statistically model nonstationarity in observed extremes through the parameters of the GEV or GP distributions, one of the noteworthy capabilities of the extRemes
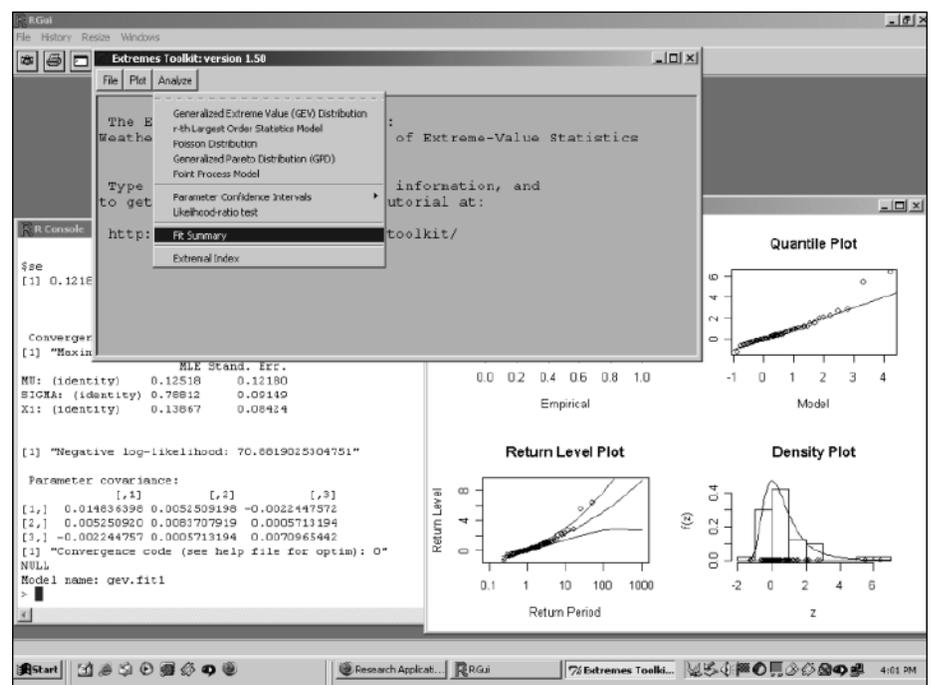

Fig. 1. Example output from Extremes Toolkit (extRemes).

package. The figure shows the time series of annual peak flow at Mercer Creek in the state of Washington (http://pubs.usgs.gov/fs/FS-229-96). A small drainage basin, Mercer Creek experienced abrupt changes in the statistical characteristics of floods because of a period of rapid urban development starting about 1970. Using extRemes, a GEV distribution was fitted to the annual peak flow time series. The conditional median and 0.9 and 0.95 quantiles for the fitted non-stationary GEV are also shown in the figure, with abrupt increases in the midst of the record—corresponding to the influence of urban development—being readily apparent. Despite this trend in annual peak flow resulting from urban changes, no trend is apparent in annual mean flow at Mercer Creek, which remains about 15–30 cubic feet per second.

Other capabilities of extRemes include fitting the GEV and GP distributions with annual or diurnal cycles in the parameters, or dependence on other variables such as El Niño events. Not only can return levels be estimated, but also confidence intervals can be attached to these estimates to reflect their uncertainty. Finally, extRemes includes a more powerful approach (an extension of the "peaks over threshold" or "partial duration series" technique in hydrology), in which the occurrence of an extreme event (e.g., exceeding a high threshold) and the intensity of an event (e.g., excess over a high threshold) are simultaneously modeled.

New Potentials in Extreme Value Analysis

There are several other packages in R that focus on more advanced aspects of extreme value analysis not included in extRemes. For more information on these packages (and other non-R packages), see http://www.ral.ucar.edu/staff/ericg/softextreme.php. Unique features of extRemes are (1) its ability to account for nonstationarity in extreme behavior of one variable, whether this be a gradual trend over time or shifts from year to year in association with another variable
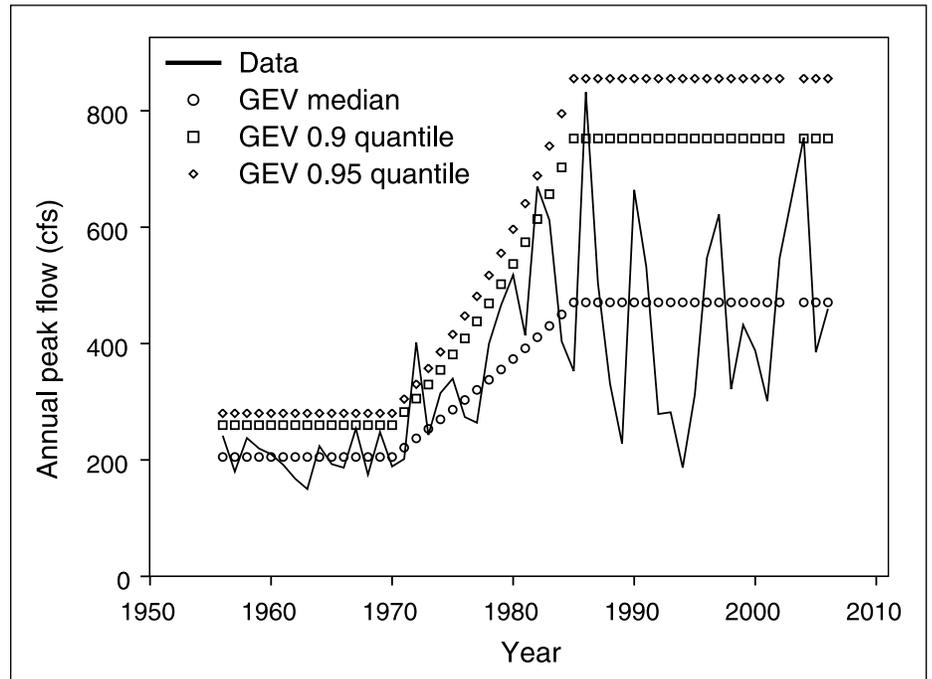

Fig. 2. Conditional quantiles of a nonstationary generalized extreme value (GEV) distribution fit to annual peak flows, measured in cubic feet per second (cfs) at Mercer Creek in the state of Washington.

(e.g., El Niño events); and (2) the incorporation of GUI windows for ease of use. Besides floods in hydrology, extRemes can be applied to a wide variety of other geophysical phenomena, including precipitation, temperature, and wind extremes in climate; extreme wave heights and sea level maxima in oceanography; earthquake magnitudes in seismology; and high sediment yields in paleoclimatology.

References

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

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: Whither water management?, *Science, 319*(5863), 573–574, doi:10.1126/science.1151915.

—Eric Gilleland and Richard W. Katz, National Center for Atmospheric Research, Boulder, Colo.; E-mail: ericg@ucar.edu