P2.15

ANALYZING SEASONAL TO INTERANNUAL EXTREME WEATHER AND CLIMATE VARIABILITY WITH THE EXTREMES TOOLKIT

Eric Gilleland*and Richard W. Katz

Research Applications Laboratory,
National Center for Atmospheric Research

This paper has been revised on February 6, 2006 from the originally submitted manuscript.

1 INTRODUCTION

Statistical analyses in weather and climate variability studies have often been concerned with averages of a random variable, such as mean precipitation or temperature. However, the *extremes* of random variables are important to consider, and have become increasingly studied in recent years (e.g., Wettstein and Mearns (2002) (hereafter, WM), Brown and Katz (1995), Zwiers and Kharin (1998), Kharin and Zwiers (2000,2005), Jagger *et al.* (2001), Ekström *et al.* (2005) and Fowler *et al.* (2005)). When studying changes in the average of a distribution, the Central Limit Theorem (CLT) indicates that the averages are asymptotically normally distributed, and therefore normality is an appropriate assumption for modeling and inference. For extremes, there is a similar theorem to the CLT called the Extremal Types Theorem, which gives asymptotic justification for assuming the extreme data (e.g., maxima, minima, etc.) follow one of three types of distributions: Gumbel, Fréchet or Weibull (see e.g., Beirlant *et al.* (2004), Coles (2001), Embrechts *et al.* (1997), Reiss and Thomas (2001) and Leadbetter *et al.* (1983)). Furthermore, these three distributions can be written in a single expression as a family of distributions referred to as the generalised extreme value (GEV) distribution. Figure 1 shows an example from having simulated means and maxima from 1,000 samples of standard normal distributions each of size 1,000. The resulting histograms for the means and maxima of these samples are displayed along with the best fit normal and GEV distributions. In each case, either distribution appears to be reasonable, but the normal distribution (solid line) clearly provides a better fit for the means and the GEV (dashed line) a better fit for the maxima as the theory suggests. Figure 2 is similar to Figure 1, but simulations are from a uniform distribution on the range of -1 to 1. It is much easier to see how the GEV distribution is a better fit in the case of maxima from the uniform distribution, though for the means the two are still similar to each other with the normal capturing the overall shape slightly better than the GEV.

Corresponding author address: Eric Gilleland, National Center for Atmospheric Research (NCAR), Research Applications Laboratory, 3450 Mitchell Ln, Boulder, CO 80301; email: ericg@ucar.edu

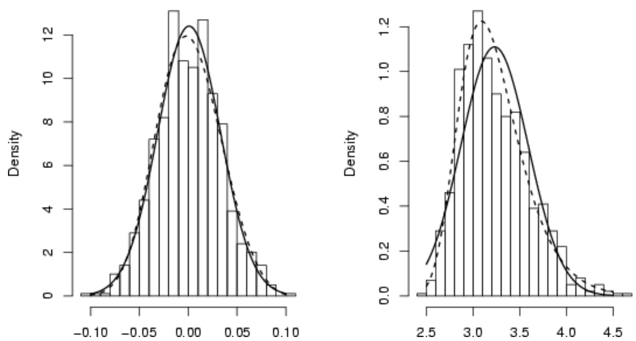


Figure 1: *One-thousand random samples each of size 1,000 simulated from a normal distribution with mean zero and unit standard deviation. The histograms shown here are for the means (left) and maxima (right) for each of these samples. The solid lines show the best fit normal pdf and the dashed lines show the best fit GEV (using maximum-likelihood estimation).*

In this paper, it will be demonstrated how extreme-value statistical analysis can be employed for studying extreme weather and climate variability incorporating seasonality and other covariates. Although several software packages exist for performing such analyses (Stephenson and Gilleland (2005)), the R (R Development Core Team (2004)) package `extRemes` (Gilleland and Katz (2005)) is used here because it is open source (as is R) and particularly well suited for weather and climate applications because of its extensive tutorial

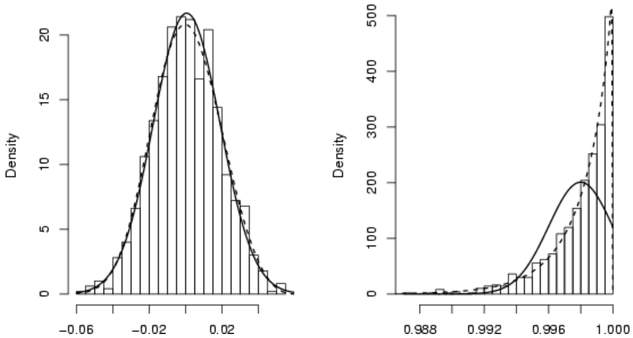


Figure 2: *One thousand random samples each of size 1,000 simulated from a uniform distribution over the range -1 to 1. The histograms shown here are for the means (left) and maxima (right) for each of these samples. The solid lines show the best fit normal pdf and the dashed lines show the best fit GEV (using maximum-likelihood estimation).*

aimed at such applications, and the ability to incorporate covariate information into parameter estimates.

The present paper analyzes the data of WM who also employ extreme-value analysis and relate the behavior of the extreme events to the mean and standard deviation of the quantities of interest. It is demonstrated here how such relations can be formalized directly in the extreme-value analysis using `extRemes`.

First, some background for extreme-value statistics is given in Section 2. Section 3 summarizes the extreme temperature application analyzed in WM. Section 4 provides analysis of these data. Some discussion is presented in Section 5, and finally, some tutorial instruction for using `extRemes` to perform the analysis described in Section 4 is given in the appendix.

2 BRIEF BACKGROUND FOR EXTREME-VALUE ANALYSIS

Because much has been written about extreme-value analysis, this section will be brief. For further reading on the subject, Coles (2001) is a good introductory text that is heavy on application, but still giving some theoretical development, and Beirlant *et al.* (2004) give a more thorough discussion. Stephenson and Tawn (2004) give a short, but particularly insightful introduction. For a more in-depth theoretical discussion, see Embrechts *et al.* (1997) and Leadbetter *et al.* (1983). For a basic introduction, see Gilleland and Katz (2005). Smith (2002) and Katz *et al.* (2002) give a more terse applied introduction with enough detail to satisfy a novice to extreme-value analysis.

There are two primary methods for analyzing extreme values statistically. The first is to fit data to a model using traditional statistical techniques, and then look at the extreme quantiles typically by simulating from the model (see e.g., Gilleland and Nychka (2005) and Chandler (2005)); the second method is to fit data to an extreme-value distribution. The former method will not be discussed further here. The latter method is carried out by two alternative approaches: block maxima and peaks over threshold (POT). The distributional theory is equivalent for either approach, though the distributional forms may, at first glance, appear to be different. The extreme-value distributions for each of these two approaches are introduced in Section 2.1, and return levels (quantiles) are discussed in Section 2.2 followed by an introduction to parameter estimation in Section 2.3. As is usual in the literature, only the upper tail extremes (e.g., maxima) are discussed because the lower tail extremes can be handled by taking the negative transformation and simply applying the same techniques as for maxima. Finally, a brief discussion is given on extending these univariate analyses to a spatial setting in Section 2.4.

2.1 Extreme-Value Distributions

When data are taken to be the maxima (or minima) over certain *blocks* of time (such as annual maximum precipitation, monthly maximum/minimum temperature), then it is appropriate to use the GEV distribution (1).

$$G(z;\mu,\sigma,\xi) = \exp\left[-\{1+\xi(z-\mu)/\sigma\}_{+}^{-1/\xi}\right], \quad (1)$$

where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ are the location, scale and shape parameters respectively, and $x_{+} = \max(x, 0)$. The three extremal types are determined by the sign of ξ arriving at the Weibull distribution for $\xi < 0$, the Gumbel distribution is obtained in the limit as $\xi \longrightarrow 0$, and the Fréchet distribution for $\xi > 0$. Each of the three types of distributions have distinct forms of behavior in the tails. The Weibull is bounded above, meaning that there is a finite value which the maximum cannot exceed. The Gumbel distribution yields a *light* tail, meaning that although the maximum can take on infinitely high values, the probability of obtaining such levels becomes small exponentially. The Fréchet, a *heavy* tailed distribution, decays polynomially so that higher values of the maximum are obtained with greater probability than would be the case with a lighter tail (see, e.g., Figure 3).

Note that some literature uses $\kappa = -\xi$ in (1). With such a parameterisation, positive and negative κ would yield the Weibull and Fréchet respectively.

The approach leading to distribution (1) assumes data are maxima from *blocks*, say of time. Arguably, for some problems, taking maxima from large

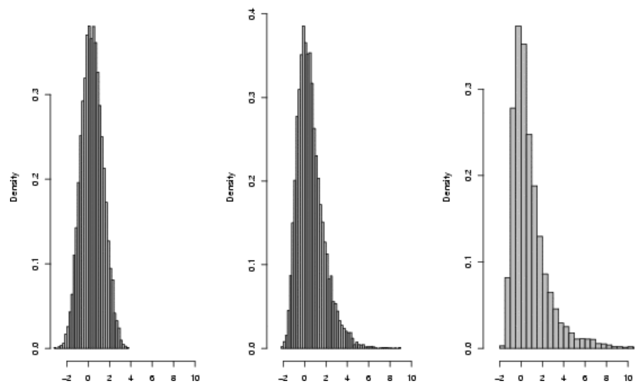


Figure 3: *Example histograms of data simulated from a GEV distribution with* $\xi < 0$ *(left),* $\xi = 0$ *(center) and* $\xi > 0$ *(right).*

blocks (e.g., annual maximum precipitation) discards too much data. The POT approach allows for more data to inform the analysis, but also increases the complexity of the problem.

For the POT approach, a threshold is first determined, and data above that threshold are fit to the generalised Pareto distribution (GPD).

$$G(x; \tilde{\sigma}, \xi, u) = 1 - \left[1 + \frac{\xi(x-u)}{\tilde{\sigma}}\right]^{-1/\xi}, \quad (2)$$

where $x - u > 0$, $1 + \frac{\xi(x-u)}{\tilde{\sigma}} > 0$ and $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

The GPD (2) gives the probability of a random variable exceeding a high value given that it already exceeds a *high* threshold, say u (i.e., $\Pr[X > x | X > u]$). Of course, the theoretical results that show the GPD to be the asymptotic distribution appropriate for exceedances over a high threshold requires these exceedances to be independent and identically distributed random variables.

Choice of threshold is critical to any POT analysis. Too high of a threshold could discard too much data leading to high variance of the estimate, but too low of a threshold can lead to bias because (2) is an asymptotic result requiring a high threshold. Practitioners are wont to using graphical tools in determining an appropriate threshold. One of the more popular approaches is to fit the GPD using a range of thresholds, and then graphing the parameter estimates along with their variability; an appropriate threshold being the lowest possible choice such that any higher threshold would result in similar estimates. It should be noted that the distribution (2) is equivalent to (1) under an appropriate transformation (see e.g., Katz *et al.* (2002) and Coles (2001)).

Apart from threshold selection, an important assumption for the GPD is that the threshold exceedances are independent. Such an assumption is often unreasonable for weather and climate data because high values of meteorological and climatological quantities are often succeeded by high quantities (e.g., a high temperature day is likely to be followed by another high temperature day). One measure of dependency that is frequently used is the extremal index, θ (e.g., Ferro and Segers (2003), Coles (2001)). The case of complete independence will yield a value of $\theta = 1$, but it is also possible to have dependent data where $\theta = 1$.

An approach frequently employed to handle such dependency is to decluster the data by identifying clusters and utilizing only a summary of each cluster (e.g., the maxima). Several such methods have been devised for determining clusters (see, e.g., Ferro and Segers (2003) and the references therein), and one of the simplest and most widely used is runs declustering. With runs declustering, a new cluster is formed once the value of the quantity of interest exceeds the threshold after having fallen below the threshold for a certain length of time, called the run length (denoted here by r).

There is also a point-process characterisation for the POT approach that allows for simulataneous fitting of the rate at which values exceed the threshold and the intensity of the exceedances (see, e.g., Coles (2001) Chapter 7 and Smith (2002)). Again, the same assumptions of a high threshold, common distribution, and independence are necessary for the model to be theoretically valid. Furthermore, the point-process model can be shown to be equivalent through appropriate transformations (see e.g., Katz *et al.* (2002) and Coles (2001) Section 7.4) to the GEV (1) and GPD (2), and unlike the GPD explicitly includes the location parameter, μ.

2.2 Return Levels (Quantiles)

Typically, when considering extreme values of a random variable, one is interested in the return level of an extreme event, defined as the value, z_p, such that there is a probability of p that z_p is exceeded in any given year, or alternatively, the level that is expected to be exceeded on average once every $1/p$ years ($1/p$ is often referred to as the return period). For example, if the 100-year return level for precipitation at a given location is found to be 1.5 cm, then the probability of precipitation exceeding 1.5 cm in any given year is $1/100 = 0.01$.

The return level is derived from the distribution (either (1) or (2)) by setting the cumulative distribution function equal to the desired probability/quantile, $1-p$; and then solving for the return level. For example, for the GEV distribution given in (1), the return level, z_p, is given by the following equation.

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - \{-\log(1-p)\}^{-\xi}], & \text{for } \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, & \text{for } \xi = 0 \end{cases}$$

2.3 Estimation

There are different methods available for performing parameter estimation including: Method of Moments Estimation (MME), Probability Weighted Moments (PWM) or equivalently L-Moments (LM), Maximum Likelihood Estimation (MLE), and Bayesian methods. For smaller sample sizes ($n < 50$), the MLE is unstable and can give unrealistic estimates for the shape parameter (e.g., Hosking and Wallis (1997), Coles and Dixon (1999), and Martins and Stedinger (2000,2001)). Madsen *et al* (1997) argue that the MME quantile estimators have smaller root mean square error when the true value of the shape parameter is within a narrow range around zero. For weather and climate applications, enough data are typically available to expect that MLE would be comparable in performance, especially when blocks smaller than years are used. Additionally, MLE allows one to easily incorporate covariate information into parameter estimates. Furthermore, it is more straightforward to obtain error bounds for parameter estimates with MLE compared with most alternative methods. Although work on Bayesian estimation with respect to extreme-value analysis has been sparse in the literature, good examples are available (see e.g., Stephenson and Tawn (2004) and the references therein, Coles (2001, Section 9.1), and Cooley *et al.* (2005a, 2005b)).

Obviously, one will never select the Gumbel when fitting data to a GEV because the Gumbel is reduced to a single point in a continuous parameter space. A common approach is to perform an initial hypothesis test to determine which of the three extremal types (e.g., the Gumbel) is appropriate, and then fit data only to that type. However, this approach does not account for the uncertainty of the choice of extremal type on the subsequent inference, which can be quite large. Stephenson and Tawn (2004) suggest a Bayesian approach to estimating these parameters that allows for the Gumbel to be achieved with positive probability; though results can be highly sensitive to choice of prior distributions.

2.4 Spatial Extensions

So far, only univariate data has been considered. In fact, incorporating spatial structure in the analysis of extremes is an area of active research in statistics (e.g., Schlather and Tawn (2002, 2003), Heffernan and Tawn (2004), Cooley *et al.* (2005a), Gilleland and Nychka (2005), Gilleland *et al.* (2006)). Work is currently underway to add spatial tools to `extRemes`, and the beta version already has the capability of fitting data to a GPD at several sites, and smoothing parameter estimates by way of a thin plate spline. For simplicity, the focus here is only on univariate data.

3 EXTREME TEMPERATURE DATA IN THE NORTHEAST UNITED STATES AND CANADA

Data to be analyzed here are a subset from the study carried out in WM; here summary description is given, but for more detail please refer to WM and the R help files for `SEPTsp` and `PORTw` included with the package `extRemes`. Here the focus centers on data from the two locations given special attention in WM: Port Jervis, New York and Sept-Iles, Québec. The Port Jervis data cover the winters from 1927 through 1995, and Sept-Iles cover the spring seasons 1945 through 1995 consisting of 68 and 51 monthly minima (i.e., block minima) derived from daily data respectively. The subsets are chosen not as much to compare findings with WM, but rather to demonstrate the statistical techniques available in using `extRemes` for performing such an analysis. Nevertheless, these two locations were treated with special attention in WM largely because results tended to be more significant in these two areas; and therefore make for a reasonable pedagogical example.

Each dataset contains measurements of monthly minimum (Figure 4) and maximum temperatures. Covariate information is also available, including: the associated Arctic Oscillation (AO) index, mean daily minima (maxima) over the one-month period, and standard deviations of daily minima (maxima) for each period.

4 ANALYSIS OF SEASONAL EXTREME TEMPERATURE

In WM, it was found that for Sept-Iles in the spring, lower mean minimum temperatures are coupled with significant increases in the standard deviation of daily minimum temperature, and that these minimum temperature extremes become more severe as the AO index increases. For Port Jervis in the winter, it was also found that increasing temperatures are associated with increases in the AO index. Here, these features are examined by modeling the extreme temperature data without any covariates, and then making comparisons with more complex extreme-value models that incorporate covariate information such as the AO index.

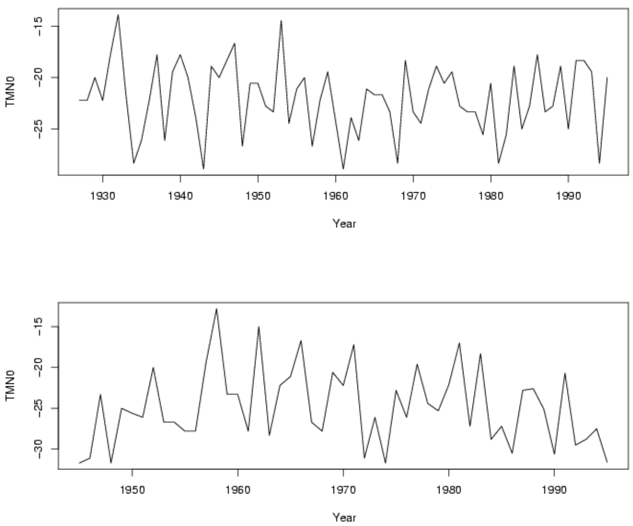


Figure 4: *Minimum temperatures (degrees celsius) at Port Jervis, New York (top) and Sept-Iles, Québec (bottom).*

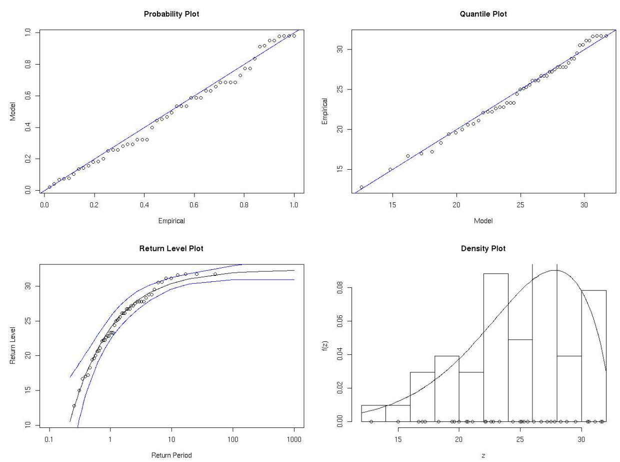


Figure 5: *Diagnostic plots from fitting the minimum monthly temperature data at Sept-Iles, Québec to a GEV distribution. Quantile and return-level graphs are for the negative transformed minima. From upper left to lower right: probability, quantile, return level, and histogram with fitted GEV density.*

To fit the monthly minimum temperature to a GEV distribution, the usual methods for maxima apply by realising that $\min\{X_1, \ldots, X_n\} = -\max\{-X_1, \ldots, -X_n\}$. That is, the negative transformation of the data ($Y_1 = -X_1, \ldots, Y_n = -X_n$) is fit to the GEV distribution. Here, results are presented in terms of the untransformed minima except where noted.

Maximum-likelihood fitted parameter values and other information are summarized in Table 1 (a). In WM, it is argued that the Gumbel ($\xi = 0$) distribution is reasonable, and that fitting the data to the other two possible extremal types would likely not be better. However, the estimates obtained from fitting all three extremal types simultaneously yields values of ξ that are far from zero. For Sept-Iles, Québec, $\xi \approx -0.6$, which is more than four standard errors away from zero. As noted in Section 2.3, fitting only to the Gumbel ignores the uncertainty associated with the choice of extremal types. For the 100-year event, the estimated return level here is about -32 degrees celsius with 95% confidence bounds estimated from the profile likelihood of about $(-34.1, -31.4)$ degrees celsius compared to an estimated return level of -14.87 estimated from the Gumbel fit, which is well below those estimated from the fit that accounts for the uncertainty of choice of extremal type. For the monthly minimum temperature data in the winter at Port Jervis, New York, $\xi \approx -0.3$ (Table 1 (b)). The 100-year return level for minimum temperature at this location is about -29.68 with 95% confidence bounds (profile-likelihood) of about $(-33.333, -28.245)$, whereas the

Table 1: *GEV parameter estimates from fitting monthly minimum temperatures at (a) Sept-Iles, Québec (spring) and (b) Port Jervis, New York (winter).*

	Parameter	Estimate	Std. Error
(a)	Location (μ)	-23.971	0.828
	Scale (σ)	5.255	0.719
	Shape (ξ)	-0.624	0.139
Negative log-likelihood: 147.38			
(b)	Location (μ)	-20.770	0.455
	Scale (σ)	3.346	0.324
	Shape (ξ)	-0.264	0.092
Negative log-likelihood: 179.56			

100-year return level assuming the Gumbel is estimated to be approximately -35.23, which is beyond the lower limit of the 95% confidence bounds for the bounded-tail Weibull.

The probability and quantile graphs for each data set (Figure 5 displays those for Sept-Iles, Québec) suggests that the underlying assumptions for the GEV distribution are reasonable for both datasets. The return level plot is shown in the lower left corner along with pointwise 95% confidence bounds estimated by the delta method (the default). The delta method assumes that the parameter estimates are symmetric; which is typically not the case for the shape parameter or extreme return levels. For example, Figure 6 shows the profile likelihoods for the 100-year return level and the shape parameter. In each case, there is clearly asymmetry–

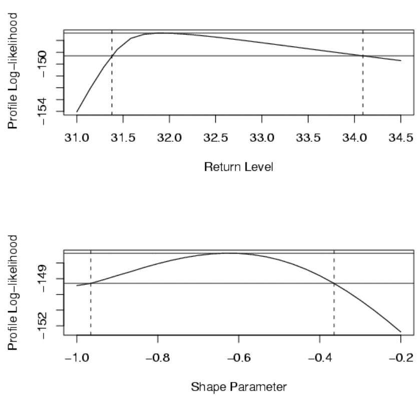


Figure 6: *Profile likelihoods for the negative transformed 100-year return level (top) and shape parameter (bottom) from having fit monthly minimum temperature at Sept-Iles, Québec (spring) to the GEV distribution.*

especially for the 100-year return level.

It can be seen from both Table 2 and Figure 7 that the return levels for minimum temperature at Port Jervis, New York gradually decrease for longer return periods. Clearly, at least for these data, the delta method bounds are tighter for lower values of return levels whereas for return periods beyond about 10 years, the profile-likelihood has a much tighter lower bound, and slightly higher upper bound. Beyond about 100 years, both bounds are very wide reflecting the inherent uncertainty associated with making inferences far beyond the range of the data; but the profile-likelihood method gives a more accurate picture for such longer return periods because it accounts for the skew in the parameter distributions.

Inclusion of the AO index as a covariate in the location parameter (μ) of the GEV distribution yields a significant (at the 5% level) improvement over the fit without AO index (likelihood ratio test statistic is about 12, which is much greater than the associated $\chi^2_{1,1-0.05}$ critical value of about 4). Specifically, the

Table 2: *Estimated return levels and 95% confidence intervals for several return periods from having fit monthly minimum temperatures (degrees celsius) at Port Jervis, New York (winter) to the GEV distribution.*

Return period	Return level	Lower bound	Upper bound
5	-24.91	-25.95	-23.97
10	-26.45	-27.81	-25.47
15	-27.19	-28.86	-26.17
25	-27.00	-30.13	-26.90
50	-28.92	-31.77	-27.68
75	-29.38	-32.70	-28.03
100	-29.68	-33.33	-28.24
110	-29.77	-33.54	-28.31
125	-29.90	-33.82	-28.39
150	-30.06	-34.21	-28.50
200	-30.31	-34.81	-28.65
500	-30.98	-36.65	-29.02
1000	-31.40	-37.96	-29.21

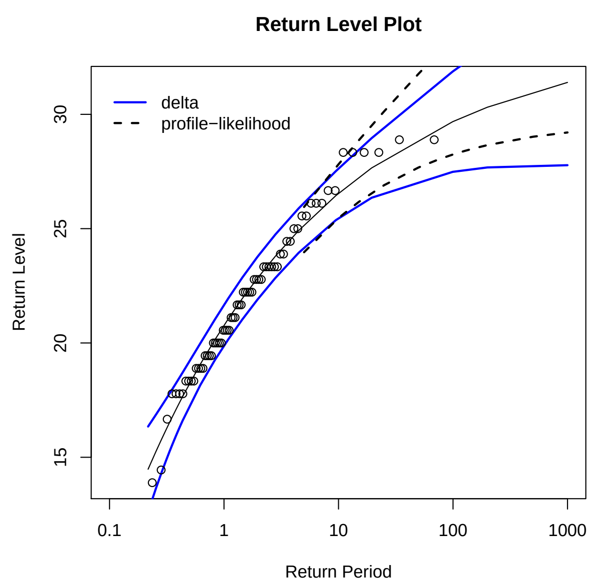


Figure 7: *Return-level graph of negative transformed monthly minimum temperature (degrees celsius) calculated from associated GEV distribution (solid line) with 95% confidence intervals calculated from the delta and profile-likelihood methods respectively for Port Jervis, New York (winter).*

Table 3: *GEV parameter estimates from fitting monthly minimum temperature (degrees celsius) recorded at Sept-Iles, Québec (spring) with AO index incorporated as a covariate in the location parameter as in Eq. (3).*

Parameter	Estimate	Stand. Error
Location (μ_0)	-23.42	0.668
Location (μ_1)	-2.09	0.686
Scale (σ)	4.19	0.502
Shape (ξ)	-0.368	0.117
Negative log-likelihood: 143.04		

model obtained is summarized in Table 3, where the location parameter is modeled as a linear regression of the following form.

$$\mu(x) = \mu_0 + \mu_1 x, \quad (3)$$

where x is the AO index.

Note that as the AO index increases, the values of the location parameter become more negative indicating that the minimum temperature extremes become more severe as the AO index increases. This result is consistent with the findings of WM. It might also be worth noting that the fitted distribution with AO index as a covariate now has a more strongly negative shape parameter (≈ -0.4) that is about three standard errors away from zero (Gumbel case).

It is also found in WM that this intensification of the extreme temperatures is coupled with increases in the standard deviation of daily minimum temperature for this location in spring. Performing a similar fit with a covariate in the location parameter as in Eq. (3), but with x the standard deviation of daily minimum temperatures, also results in a significant (5% level) improvement over the no-covariate model. Furthermore, adding the standard deviation to the model with AO index (i.e., $\mu(\mathbf{x}) = \mu_0 + \mu_1 x_1 + \mu_2 x_2$, where x_1 is the AO index and x_2 is the standard deviation) results in a significant improvement over the model with just AO index (likelihood ratio test statistic of about 62). On the other hand, the addition of AO index to the model with only the standard deviation as a covariate is not found to be significant. This result, however, should not be surprising because the standard deviation of daily minimum temperatures are derived from the same data as the block maxima. Additionally, it should be expected that data with greater variability should also have more intense extremes. For these reasons, using such a covariate in the model is misleading. Nevertheless, the AO index is not derived from the same data as the dependent variable (i.e., monthly minimum temperature), and is certainly an improvement over the model without a covariate.

It is also possible to incorporate covariates into the other parameters. In the case of the scale parameter, it is important to ensure that $\sigma > 0$ for all possible covariate values. This is easily attained by using the log link function instead of the identity–the two choices provided by `extRemes`. Results from trying such fits were not found to be as significant as for the location parameter for the Sept-Iles data. Typically, the difficulty in estimating the shape parameter (with wildly differing tail behavior for the three types of distributions) is enough to deter researchers from utilizing covariates in conjunction with this parameter, although such analysis is allowed with `extRemes`.

In WM, it is found that minimum temperature extremes at Port Jervis during winter are not significantly influenced by the AO index; and this is corroborated when performing an analogous analysis as for the Sept-Iles data.

5 DISCUSSION

Results found from analyzing two subsets of data from WM show agreement with the results in WM for one of the sets (minimum temperatures at Port Jervis, New York), but using a slightly more sophisticated analysis. However, return levels found from fitting the data at Sept-Iles, Québec were found to produce much lower return levels than those estimated using the approach in WM. The approach used here accounts for the uncertainty associated with the choice of extremal type, but under this paradigm the Gumbel case will be estimated with a probability of zero because it is a single point of a continuous parameter space.

A primary intent of the paper is to demonstrate how easily such an analysis can be perfomed using `extRemes`. It should be noted that `extRemes` has much more capability than is detailed here (see Gilleland and Katz (2005) for a more thorough description of the capabilities of `extRemes`). The inclusion of spatial tools is a much needed addition for a software package intended for weather and climate applications, and such tools will be available in the near future.

Naturally, a GUI-based software package will have limitations as far as analyzing cutting-edge research problems. The intention is for `extRemes` to serve as an aid in shortening the learning curve associated with using a possibly very new methodology.

APPENDIX

The Extremes Toolkit: Weather and Climate Applications of Extreme Value Statistics

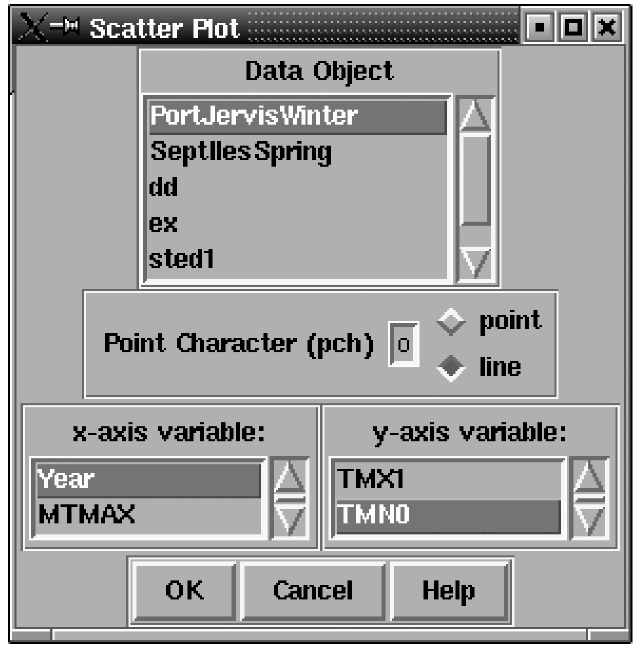
The extremes toolkit (`extRemes`) was developed at NCAR to assist scientists, especially scientists interested in trends in weather and climate extremes or in societal and ecological impacts of severe weather and climate change, not familiar with extreme value statistical techniques. The package is written in the open source[1] statistical computing language called R (R Development Core Team (2004)), but can be used without any knowledge of R. `extRemes` provides a graphical user interface (GUI) to another R package called `ismev`–itself an R port of the S-Plus package written by Stuart Coles (Coles (2001)). However, `extRemes` does provide some additional functionality (Stephenson and Gilleland (2005)). This section is only a short tutorial on using `extRemes` to perform the analyses described herein, but see Gilleland and Katz (2005) for a full tutorial on the package.

It is assumed here that `extRemes` (version 1.51 or greater) is already installed and loaded into an R session (see the toolkit's home page at `http://www.assessment.ucar.edu/toolkit` for installation and loading instructions). To load the datasets used here into `extRemes` one simply needs to click on **File** followed by **Read Data**, and then search for the files (one at a time) **PORTw.R** and **SEPTsp.R** (if these files are not found within the `extRemes` data directory, they can be obtained from the web at `http://www.isse.ucar.edu/extremevalues/data/`). After double-clicking (or selecting and clicking **Open**) these files, a new window will appear. Select **R source** and enter a name in the **Save As** field (e.g., the names **PortJervisWinter** and **SeptIlesSpring** are used here), then click **OK**. Summary information on the data should appear in the R console window, and the working directory is saved to memory.

Graphical Tools

It is always a good idea to graph data before analyzing it, and this can be readily performed from the toolkit dialogs. For example, to create the top graph in Figure 4, simply select **Plot** and then **Scatter Plot**, then make the selections as shown below.

[1]All of the packages available from the R-CRAN website are also open source, including `extRemes` and `ismev`.



Other graphs, such as AO index against monthly mean daily minima (maxima), can also be easily drawn using the `extRemes` GUI dialogs.[2]

Data Transformations

Several data transformations can be made using `extRemes`. Although most are easy to compute from the command line, it is preferable to use the toolkit dialog in order that the resulting transformations are placed where subsequent dialogs can include them.

To perform the negative transformation for the Sept-Iles springtime minimum temperature data, choose **Negative** from the **Transform Data** menu under **File**. Select **SeptIlesSpring** from the **Data Object** listbox, and then select the variable **TMN0** from the **Variables to Transform** field, and finally click **OK**. A message is displayed in the R console informing you that a new column has been added to the data with the heading **TMN0.neg**.

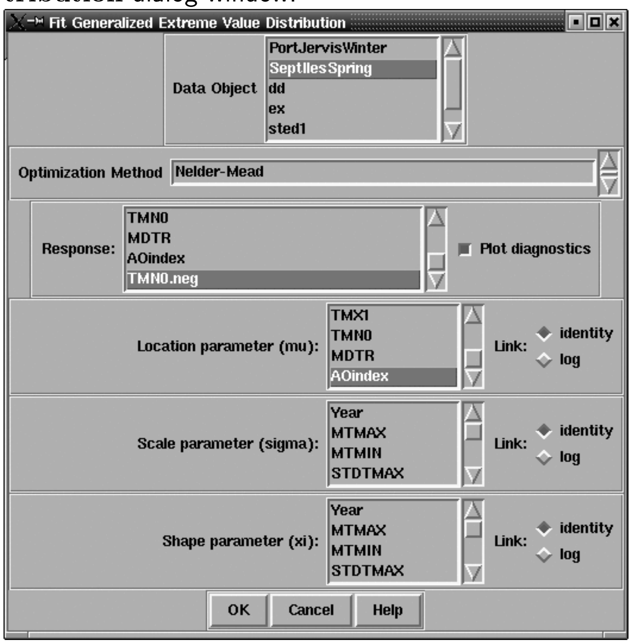
Parameter Estimation

After taking the negative transformation, the monthly minima can be fit to a GEV distribution by selecting **Generalized Extreme Value (GEV) Distribution** under **Analyze**. Again, select **SeptIlesSpring** from the **Data Object** field followed by **TMN0.neg** in the **Response** field. Check the **Plot diagnostics**

[2]For more advanced graphs, a good start might be an inspection of the code executed by the GUI windows found in the `extRemes.log` file located in the path of the current R working directory (use the R command `getwd()` to find this path) and the R help files for the function `plot` (i.e., use `help( plot)`) and associated parameters (use `help( par)`).

checkbutton to also obtain the graphs of Figure 5, but do not make any other selections (yet)–simply click **OK**. For the example below, the same steps are taken, but **AOindex** is selected from the **Location parameter (mu)** field.

The estimation for Table 3 can be performed in `extRemes` by making the selections (shown below) from the **Fit Generalized Extreme Value Distribution** dialog window.

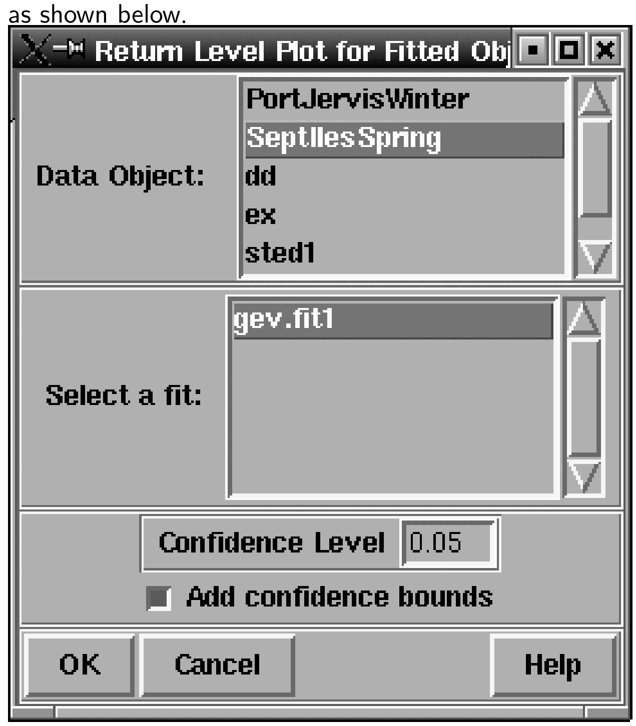


There are several choices available for the numerical optimization method. See the help file for the R function `optim` and the references therein for more information. Press *et al.* (1989) provide Fortran algorithms for all of these methods.

Profile Likelihood

Figure 6 is graphed using `extRemes` by selecting **GEV fit** from the **Parameter Confidence Intevals** under **Analyze**. In the window that opens, select **SeptIlesSpring** from the **Data Object** field followed by **gev.fit1** in the **Select a fit** field. Enter 31 and 35 in the **Lower** and **Upper** limit fields for the **Return Level Search Range** respectively; and -1 to -0.2 for the **Shape parameter (xi) Search Range**.[3] You may wish to change the **nint** entry to a lower value for this return level and these data (25 was used here). Finally, check the **Plot profile likelihoods** checkbutton, and then **OK**. Note that the parameter estimate along with its confidence intervals are reported to the R console. Here, the 100-year return level is estimated to be about 31.92 degrees celsius (recall that the model is fit to the negative minimum temperatures so that the -31.92 reported actually refers to the negative transformed minima) with a 95% confidence interval of about (28.24, 33.33).

Although somewhat tedious, it is possible to use `extRemes` to create a graph similar to that of Figure 5 using results from the profile-likelihood estimation. First, obtain confidence bounds for each of several return levels using the profile-likelihood method as done for the 100-year return level above.[4]. Next, graph the return levels by selecting **Return Level Plot** from the **Plot** menu, and make the selections as shown below.



To add the profile-likelihood derived confidence bounds to the graph requires use of the command line. First, for convenience, set up a 13×3 matrix of the return year and negative transformed values from the second and third columns of Table 2 by using, for example, the following commands (here, the resulting matrix is assigned to the object `ci`).

```
ci <- rbind( c( 5, 23.97, 25.95),
        c( 10, 25.47, 27.81),
        c( 15, 26.17, 28.86),
        c( 25, 26.90, 30.13),
        c( 50, 27.68, 31.77),
        c( 75, 28.03, 32.70),
        c( 100, 28.24, 33.33),
```

[3] Choice of search ranges and number of intervals is a matter of trial and error. Different ranges are possible so long as the range includes the points where the profile likelihood crosses the lower horizontal line (see Coles (2001) Section 2.7 for the meaning of this line).

[4] The delta method is a reasonable approximation for the shorter return levels, and therefore these are not estimated via the more time consuming profile-likelihood method

```
        c( 110, 28.31, 33.54),
        c( 125, 28.39, 33.82),
        c( 150, 28.50, 34.21),
        c( 200, 28.65, 34.81),
        c( 500, 29.02, 36.65),
        c( 1000, 29.21, 37.96))
```

Finally, add the values to the graph using the `lines` command as follows (coloring (`col`) and line type (`lty`) are optional to the user's preference).

```
lines( ci[,1], ci[,2], lty=2, col="red")
lines( ci[,1], ci[,3], lty=2, col="red")
```

Likelihood-Ratio Test

The likelihood-ratio test is performed with `extRemes` by selecting **Likelihood-ratio test** from the **Analyze** menu. Simply select the data object (in this case **SeptIlesSpring**), and then choose the model fit associated with the base model (**M0**) and the more complicated model (**M1**) being sure that **M0** is nested in **M1** (`extRemes` will automatically switch the two if **M0** has more parameter estimates than **M1**), and finally click **OK**.

ACKNOWLEDGEMENTS

This work is funded by the NCAR Weather and Climate Impacts Assessment Science (WCIAS) Program. NCAR is sponsored by the National Science Foundation (NSF). The authors would like to thank everyone from the "Statistical Analysis of EXTREMES in GEOPHYSICS" (http://www.ral.ucar.edu/staff/ericg/readinggroup.html) reading group who helped to edit this paper.

REFERENCES

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J., 2004: *Statistics of Extremes*, Wiley, Chichester, England.

Brown, B.G. and Katz, R.W., 1995: Regional analysis of temperature extremes: Spatial analog for climate change?, *J. of Climate* **8**:108–119.

Chandler, R.E., 2005: On the use of generalized linear models for interpreting climate variability, *Environmetrics* **16**(7):699–715.

Coles, S.G., 2001: *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.

Coles, S.G. and Dixon, M.J., 1999: Likelihood-based inference for extreme value models, *Extremes* **2**(1):5–23.

Cooley D., Nychka D., Naveau P., 2005a: Bayesian Spatial Modeling of Extreme Precipitation Return Levels, (submitted).

Cooley D., Naveau P., and Jomelli, V., 2005b: A Bayesian Hierarchical Extreme Value Model for Lichenometry. *Environmetrics*, (in press).

Ekström, M., Fowler, H.J., Kilsby, C.G., and Jones, P.D., 2005: New Estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. *J. Hydrology* **300**:234–251.

Embrechts, P., Klüppelberg, C., and Mikosch, T., 1997: *Modelling Extremal Events*, Springer-Verlag, New York.

Ferro, C.A.T. and Segers, J., 2003: Inference for clusters of extreme values, *J. R. Stat. Society B* **65**:545–556.

Fowler, H.J., Ekström, M., Kilsby, C.G., and Jones, P.D., 2005: New Estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *J. Hydrology* **300**:212–233.

Gilleland, E., Nychka, D., and Schneider, U., April 27, 2006: Spatial models for the distribution of extremes, *Computational Statistics: Hierarchical Bayes and MCMC Methods in the Environmental Sciences*, Edited by J.S. Clark and A. Gelfand. Oxford University Press. (*in press*)

Gilleland, E. and Katz R.W., 2005: *Tutorial for The Extremes Toolkit: Weather and Climate Applications of Extreme Value Statistics*, `http://www.assessment.ucar.edu/toolkit`.

Gilleland, E. and Nychka, D., 2005: Statistical models for monitoring and regulating ground-level ozone, *Environmetrics* **16**:535–546.

Heffernan, J.E. and Tawn, J.A., 2004: A conditional approach for multivariate extreme values, *J. R. Stat. Society B* **66**(3):497-530(34).

Hosking, J.R.M. and Wallis, J.R.,1997: *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, New York.

Jagger, T., Elsner, J.B., and Xufeng, N., 2001: A dynamic model of hurricane winds in coastal counties of the United States, *J. Appl. Meteor.* **40**(5):853–863.

Katz, R.W., Parlange, M.B., and Naveau, P., 2002: Statistics of extremes in hydrology, *Advances in Water Resources*, 25:1287–1304.

Kharin, V.V. and Zwiers, F.W., 2005: Estimating extremes in transient climate change simulations, *J. Climate* **18**:1156–1173.

Kharin, V.V. and Zwiers F.W., 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM, *J. Climate* **13**:3760–3788.

Leadbetter, M.R., Lindgren, G., and Rootzén, H., 1983: *Extremes and Related Properties of Random Sequences and Series*, Springer-Verlag, New York.

Madsen, H., Rasmussen, P.F., and Rosbjerg, D., 1997: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, 1, At-site modeling, *Water Resour. Res.* **33**(4):747–758.

Martins, E.S. and Stedinger, J.R., 2000: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, *Water Resour. Res.* **36**(3):737–744.

Martins, E.S. and Stedinger, J.R., 2001: Generalized maximum likelihood Pareto-Poisson estimators for partial duration series, *Water Resour. Res.* **37**(10):2551–2557.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., 1989: *Numerical Recipes (Fortran)*, Cambridge University Press, New York.

Reiss, R.D. and Thomas, M., 2001: *Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and Other Fields*, Birkhauser, New York.

R Development Core Team, 2004: *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, `http://www.R-project.org`.

Schlather, M. and Tawn, J.A., 2003: A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90**(1):139–156.

Schlather, M. and Tawn, J.A., 2002: Inequalities for the extremeal coefficients of multivariate extreme value distributions. *Extremes* **5**(1):87–102.

Smith, R.L., 2002: Statistics of extremes with applications in environment, insurance and finance, `http://www.stat.unc.edu/postscript/rs/semstatrls.pdf`

Stephenson, A. and Gilleland, E., 2005: Software for the analysis of extreme events: the current state and future directions, *Extremes* (submitted).

Stephenson, A. and Tawn, J.A., 2004: Bayesian inference for extremes: Accounting for the three extremal types, *Extremes* **7**:291–307.

Wettstein, J.J. and Mearns, L.O., December 2002: The influence of the North Atlantic-Arctic Oscillation on mean, variance and extremes of temperature in the northeastern United States and Canada, *J. of Climate* **15**:3586–3600.

Zwiers, F.W. and Kharin, V.V., 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling, *J. Climate* **11**:2200–2222.