ORIGINAL PAPER

S. A. Sisson · L. R. Pericchi · S. G. Coles

A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean

Abstract There is an urgent need for the development and implementation of modern statistical methodology for long-term risk assessment of extreme hydrological hazards in the Caribbean. Notwithstanding the inevitable scarcity of data relating to extreme events, recent results and approaches call into question standard methods of estimation of the risks of environmental catastrophes that are currently adopted. Estimation of extreme hazards is often based on the Gumbel model and on crude methods for estimating predictive probabilities. In both cases the result is often a remarkable underestimation of the predicted probabilities for disasters of large magnitude. Simplifications do not stop here: assumptions of data homogeneity and temporal independence are usually made regardless of potential inconsistencies with genuine process behaviour and the fact that results may be sensitive to such mis-specifications. These issues are of particular relevance for the Caribbean, given its exposure to diverse meteorological climate conditions.

In this article we present an examination of predictive methodologies for the assessment of long-term risks of hydrological hazards, with particular focus on applications to rainfall and flooding, motivated by three data sets from the Caribbean region. Consideration is given to classical and Bayesian methods of inference for annual maxima and daily peaks-over-threshold models. We also examine situations where data non-homogeneity is compromised by an unknown seasonal structure, and the situation in which the process under examination has a physical upper limit. We highlight the fact that standard Gumbel analyses routinely assign near-zero probability to subsequently observed disasters, and that for San Juan, Puerto Rico, standard 100-year predicted rainfall estimates may be routinely underestimated by a factor of two.

Keywords Bayesian analysis · Caribbean · Extreme value theory · Hydrological hazards · Risk assessment

S. A. Sisson (✉)
Department of Statistics, School of Mathematics,
University of New South Wales, Sydney, 2052, Australia
E-mail: scott.sisson@unsw.edu.au

L. R. Pericchi
University of Puerto Rico, San Juan, PR, USA

S. G. Coles
University of Padova, Padova, Italy

1 Introduction

The Caribbean has a complex and often volatile meteorological system, which makes extreme value analysis an essential tool for planning purposes, but also difficult to implement in a non-superficial way. This is complicated further by the fact that, as in any statistical analysis, there is a dichotomy between making sampling errors as small as possible through efficient inference, but adopting a way of managing uncertainty in a way that does not diminish genuine and unavoidable effects due to random variation. There are many examples of extreme value analyses in which this dichotomy has been resolved by going to one extreme or the other: either by ignoring sampling error completely, or by publishing results whose estimation errors are so large as to be worthless in practice. Neither of these approaches is satisfactory, and we aim to show in this article, via application to a number of datasets drawn from the Caribbean, the advantages in taking a more balanced approach to handling statistical uncertainties in extreme value analyses.

The basis of the techniques we describe have been developed and proposed by Coles and Pericchi (2003) and Coles et al. (2003), although we extend this work here substantially. Taken as a whole the techniques offer a simple, yet powerful, tool with which to tackle the modelling of extremes. This is one motivation for the present article. Our other motivation, by pooling a number of analyses of datasets drawn from the Caribbean, is to emphasize the potential benefits to regional

planners in adopting the approach to extreme value modelling that we are advocating. The datasets we consider are water levels in Lago de Managua, Nicaragua, and daily rainfall levels at the international airports of both Maiquetia, Venezuela and San Juan, Puerto Rico. A common aspect of each of these datasets is that, in one way or another, what might be regarded as standard extreme value analyses prove unsatisfactory. In fact, for two of the series, standard analyses of historical data under predict by orders of magnitude the likelihood of events that subsequently occurred.

In many respects the techniques that we adopt are simply those of good statistical practice. We seek to use all relevant data, model components of non-homogeneity, and take proper account of uncertainty when making predictions. The novelty arises from the fact that the unique aspects of an extreme value analysis make each of these aspects more challenging to implement in practice. However, developments in the theory of extremes over the last 20 years or so have led to representations of extremal processes which afford the possibility to utilize more data than is possible with classical representations. Moreover, by moving from, say, annual to daily maxima of a process, the opportunity (not to say the need) arises of modelling within-year variations. These classes of extreme value model are integral to our methodology. The other main theme of our approach is a Bayesian re-interpretation of extreme value analyses. This offers a number of advantages including the possibility to add additional information in the form of a prior distribution. However, the main advantage is the ease and flexibility with which measures of uncertainty can be handled. In particular, we argue that a predictive version of the conventional return level plot is the most constructive way to present and interpret the results of an extreme value analysis. This type of analysis relies heavily on Markov chain Monte Carlo methods of computation (for example, Coles and Pericchi 2003 for a related analysis, or Robert and Casella 2004; Gilks et al. 1996 for more general MCMC methods), but these are found to be relatively simple to implement in our examples.

In Sect. 2 we discuss the various methods available for the analysis of extremes of environmental processes. In Sects. 3–5 respectively we consider three cases of interest in the Caribbean: water volume in Lago de Managua, the Vargas rainfall tragedy, and data from San Juan International Airport. Finally, we conclude with a discussion in Sect. 6.

2 Methods

The central asymptotic theory of extremes is summarized by a series of results on weakly mixing stationary series (Leadbetter et al. 1983, for example). Denoting a stationary series by $X_1, X_2, ...$, which in our setting might represent daily rainfall measurements, for example, it is usual to characterize the behaviour of the extremes of this process by considering the limiting behaviour of a block maxima of size n

$$M_n = \max_{i=1,...,n} \{X_i\}. \tag{1}$$

Under some regularity conditions, normalizing sequences $\{a_n\}$ and $\{b_n\}$ can be found such that

$$\Pr\{(M_n - a_n)/b_n \leq z\} \to G(z) \tag{2}$$

as $n \to \infty$, for a non-degenerate distribution function G. In this case G must belong to the generalized extreme value (GEV) family of distributions, having distribution functions of the form:

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \tag{3}$$

defined on the set $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$, where μ and σ are location and scale parameters respectively. The shape parameter ξ determines the type of tail behaviour: the cases $\xi < 0$, $\xi > 0$ and $\xi = 0$ correspond to the Weibull, Fréchet and Gumbel sub-families of distributions respectively. In the latter case, the distribution function is interpreted as the limit as $\xi \to 0$ of (Eq. 3), leading to

$$G(z) = \exp\left[-\exp\left\{-\left(\frac{z - \mu}{\sigma}\right)\right\}\right], \quad -\infty < z < \infty. \tag{4}$$

The mathematical argument leading to the limits Eq. 3 and Eq. 4 provides a convenient strategy for modelling extremes. The population distribution is unknown, though its extremal characteristics determine the values of μ, σ and ξ in Eq. 2. Interpreting the limit as an approximation to the distribution of M_n for large, but finite, n in Eq. 1 and absorbing the unknown constants a_n and b_n into the parameters μ and σ, immediately gives rise to a working family for the distribution of block maxima M_n. Partly for historical reasons, and partly because the limit family for many well-known distributions is the Gumbel distribution, a tradition has developed in which the Gumbel distribution itself is used as the complete family with which to model block maxima rather than the full GEV family. We argue in subsequent sections that such a strategy runs very high risks unless there is external information which supports the Gumbel choice.

A typical application of model Eq. 3 involves fitting it to a series of annual maximum observations (effectively assuming $n = 365$ in Eq. 2 in the case of daily observations). Obviously, the choice of annual blocks is made for convenience, but experience also suggests it is often a reasonable choice in terms of validity of the asymptotic approximation. Many techniques are available for parameter estimation; our own preference is for likelihood-based methods. These enable modifications to the model to handle such features as non-stationarity, and also lead to a Bayesian formulation of extreme value problems which we will argue is the most natural for this type of analysis.

Regardless of inferential method, an unavoidable obstacle to obtaining an accurate inference is the limited amount of data that can be incorporated in the block maximum analysis. It is common to have series with as few as 10 years' of annual maxima, and rare to have more than 50 years. Often, models estimated with such short series have large sampling errors, especially on extrapolation. (As a side point, we note that many extreme value analyses resolve this problem by not reporting sampling errors. However, this simply hides rather than solves the difficulty.)

Efficient methods of model estimation may offer slight improvements in terms of reducing sampling error, but substantial changes can only be brought about by the inclusion into the model of additional information. One possibility, made feasible by the adoption of a Bayesian analysis, is the use of expert prior information, based perhaps on a knowledge of the underlying physical dynamics of a process. A complementary possibility is to try to use more of the available data than just the block maxima in drawing inference on extremal behaviour. It may well be, for example, that several extreme events arise in a given year, while only the largest of these would contribute to an analysis of annual maxima. This limitation of the block maxima approach has led to the development of modelling techniques based on an alternative threshold exceedance characterization of extremes.

Denoting the level of a process on (say) day j by X_j, one interpretation of the GEV limit Eq. 3, obtained by Taylor expansion of $n \log (F(x)) = \log G(x)$ about the point x (Coles 2001; Leadbetter et al. 1983, for example), is that

$$F(x) = P(X_j < x) = \exp\left[-\frac{1}{d}\left\{ 1 + \xi\left(\frac{x - \mu}{\sigma}\right) \right\}^{-1/\xi} \right], \quad (5)$$
$$x > u,$$

for a sufficiently large threshold u, where $d = 365.25$ corresponds to the (average) number of days in a year. Note we substitute d in place of n in this setting to highlight our use of "daily" data. This model is consistent with the GEV distribution for annual maxima in the sense that if Eq. 5 applies for the daily observations, then Eq. 3 is the distribution of the annual maxima. The advantage of this representation however is that all observations exceeding the threshold u contribute to inference on the extreme value parameters, for which we again favour likelihood-based methods. The choice of threshold u represents a trade-off between bias (due to failure of the asymptotic model at low thresholds) and variance (due to few exceedances of a high threshold). In practice there are simple diagnostic tools to assist with this choice (Davison and Smith 1990).

For most environmental applications the assumption of daily observations having a stationary distribution is not sustainable (Smith 1989; Walshaw 2000, for example). In this case model Eq. 5 suggests a natural generalization to

$$F_j(x) = P(X_j < x) = \exp\left[-\frac{1}{d}\left\{ 1 + \xi_j\left(\frac{x - \mu_j}{\sigma_j}\right) \right\}^{-1/\xi_j} \right],$$
$$x > u_j. \quad (6)$$

In this case parametric or seasonally-blocked models can be proposed for daily variations in extreme value behaviour as represented by the potential variations in the parameters μ_j, σ_j and ξ_j. It may also be appropriate to allow the threshold u_j to be time-dependent in this case. For such models likelihood-based inference represents the only feasible inferential method (see Ramesh and Davison 2002 for a local-likelihood approach).

Taking model Eq. 5 as a simple example, the likelihood function based on data $x_1,...,x_n$ takes the form

$$L(\mu, \sigma, \xi) = \prod_{i=1}^{n} g(x_i; \mu, \sigma, \xi) \quad (7)$$

where

$$g(x; \mu, \sigma, \xi) = \begin{cases} F(u) & \text{if } x \leq u \\ \frac{dF}{dx}(x) & \text{if } x > u. \end{cases} \quad (8)$$

That is, the likelihood contribution is of the form Eq. 6 for $x \in (u,\infty)$, or is censored at $F(u)$ for observations that fall below the threshold, u. Maximum likelihood estimation corresponds to the selection of parameters that maximize Eq. 7. Standard asymptotic theory of the likelihood function then yields immediate approximations to standard errors and confidence intervals.

An alternative to maximum likelihood, which offers several advantages, is a Bayesian analysis. In this case, starting with a prior density $\pi(\mu,\sigma,\xi)$ and a likelihood function $L(\mu,\sigma,\xi)$, the posterior density is obtained as

$$f(\mu, \sigma, \xi | x_1, \ldots, x_n) \propto L(\mu, \sigma, \xi)\pi(\mu, \sigma, \xi) \quad (9)$$

The main objections to a Bayesian analysis are the requirement to specify a prior distribution, and the computational difficulties involved in calculating the proportionality constant implied by Eq. 9. In our experience, however, posterior inferences are usually robust to prior choice provided the prior is reasonably flat, corresponding to an absence of genuine prior knowledge. On the other hand, if genuine prior knowledge is available, the opportunity to exploit such knowledge through the Bayesian paradigm is actually an advantage. For the second aspect, although computation is more demanding than maximum likelihood, modern simulation-based techniques such as Markov chain Monte Carlo provide simple procedures to effect Bayesian inferences.

The conventional way to summarize an extreme value analysis is by means of a return level plot. Based on a standard annual maximum analysis, the m-year return level z_m satisfies

$$P(Z \leq z_m) = 1 - \frac{1}{m}, \quad (10)$$

corresponding (loosely) to the level that is expected to occur once every m years. A return level plot is a plot of z_m against m, usually on a logarithmic scale, with parameters replaced by their maximum likelihood estimates.

The Bayesian analogue of the return level is given by the solution of

$$P(Z \leq z_m|\mathcal{H}) = 1 - \frac{1}{m}$$

where

$$P(Z \leq z|\mathcal{H}) = \int P(Z \leq z|\mu,\sigma,\xi)P(\mu,\sigma,\xi|\mathcal{H})\mathrm{d}(\mu,\sigma,\xi)$$

is the so-called predictive distribution of the annual maximum distribution given the historical data and prior information denoted by $\mathcal{H}$. Return levels can again be plotted as a function of m, with the advantage of providing a more conservative estimate having accounted for the uncertainty which derives from the parameter estimation.

3 Lago de Managua, Nicaragua

Nicaragua is a country that is accustomed to catastrophic events. It has suffered the destruction by earthquake of its capital, Managua City, twice in a 40-year period. Belonging to a tropical ecosystem, its climate produces great variations in precipitation through which the country alternates between cycles of drought and flooding. The tropical cyclones that traverse the Atlantic every year frequently cross Nicaragua due to its geographical location, thereby routinely provoking extensive flooding and landslides. In 1982, tropical storm Alleta caused severe flooding in the western part of the country, in 1990 flooding of the Bambana and

Prinzapolka rivers affected 100,000 people along its banks, and in 1998 the rains brought by Hurricane Mitch caused human and material damages that were without precedent in the history of climactic disasters in Nicaragua.

We focus on water levels of Lago de Managua, Tipitapa, determined by precipitation runoff in the vicinity. In particular, the cumulative effects of extreme storm events lead to extreme water levels in the lake. Figure 1 shows annual maxima of water levels in the lake for the period 1926–1998 excluding the years 1948–1955 for which the data are missing (it is perhaps worth noting that missing-data are quite naturally handled in a Bayesian analysis). The outstanding extreme level in 1998 is attributed to the effects of Hurricane Mitch. The volume of water produced by Mitch is in the order of 4,000 Hm^3, more than double the previously observed maximum volume induced by tropical storm Alleta.

These data were originally analyzed in an unpublished technical report by Córdova and Camachov (1999). Finding the Gumbel model to provide a poor fit to the entire series, their approach was to eliminate the smaller observations until a reasonable fit with the Gumbel distribution was obtained (Kite 1977). Such manipulation, however, leaves doubts about the validity of the Gumbel model for an arbitrary annual maximum, and it is our preferred approach to consider the broader GEV family as a candidate model.

A comparison of return level estimates based on the Gumbel and GEV models, and using both maximum likelihood and Bayesian methods of inference, is shown in Fig. 2. For the Bayesian analysis we assumed the noninformative prior $\pi(\mu,\sigma,\xi) \propto 1$ defined over the support of the parameters, simulating random draws from the posterior Eq. 9 via random-walk Metropolis-Hastings updates (Hastings 1970). The return level plots are obtained by solving Eq. 10, or the Bayesian analogue, for m. In Fig. 2 the Gumbel model is confirmed as being

Fig. 1 Lago de Managua, Nicaragua, annual maximum monthly volume. Scale is measured in Hm^3. *Open circles* denote the 1982 tropical storm Alleta, and the clear outlier Hurricane Mitch in 1998
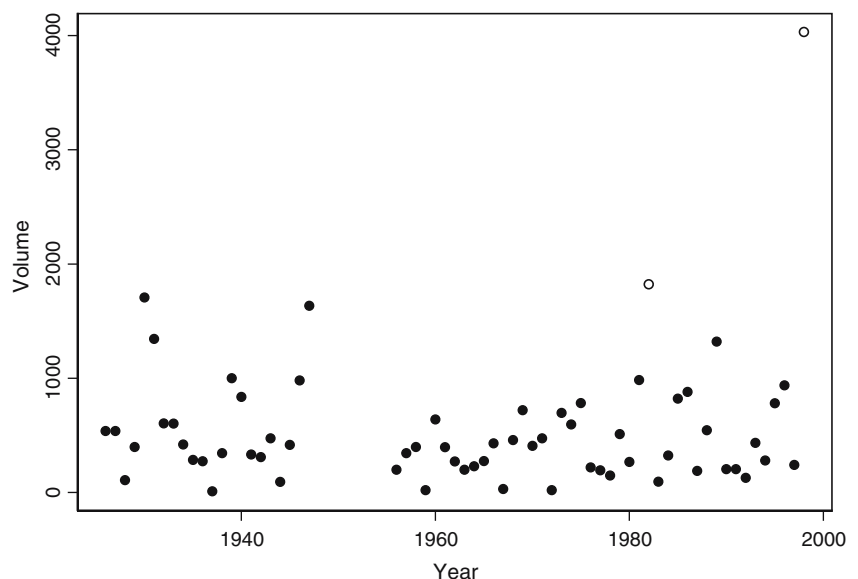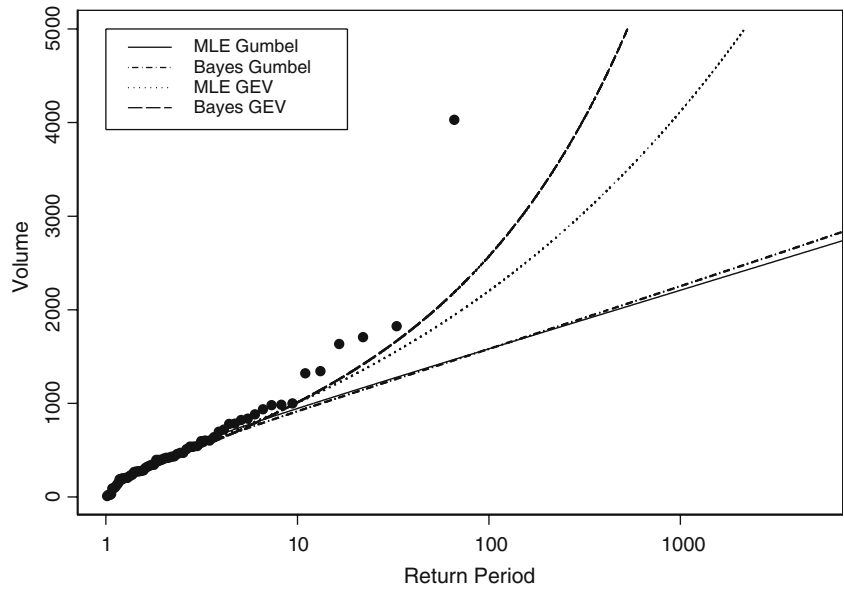
Fig. 2 Return levels for expected annual maximum volume of Lago de Managua, excluding Hurricane Mitch, using both Gumbel and GEV models. Both Bayesian and MLE estimates are shown

woefully inadequate: according to this model the return level of the 1998 event is around 800,000 years. In contrast, the GEV model fares much better: the estimated shape parameter of $\hat{\xi} = 0.21$ enables a curvature in the return level plot that reflects much better the empirical information. The standard error of 0.11 for this estimate implies that at usual levels of significance the Gumbel hypothesis of $\xi = 0$ should be rejected.

The correspondence between data and model is even clearer from the Bayesian analysis, reflecting the importance of using a model that does not impose a strong form of tail behaviour and an inferential paradigm in which uncertainties are properly managed and illustrated. A numerical summary of various return period estimates is given in Table 1.

Despite the improvements that we believe accrue from both the retention of the full GEV family and the use of a Bayesian inference, there remain some caveats about the accuracy of the final analysis. First, our model assumes a stationarity in process behaviour. Though this is supported by the data for the period of observation (for example, through rejection of models with time-dependent parameters (μ_t, σ_t, ξ_t), or with oscillation indices as predictors), slight changes in general climate may result in dramatically different behaviour for the regional hurricane climate, and consequent changes in

extremal characteristics (Goldenberg et al. 2001). For the Venezuelan data in the next section, assumptions of time-constant parameters are not reasonable for the analysis. Secondly, we have assumed the validity of the GEV model for annual maxima of monthly observations. In this case we are effectively adopting the GEV limit with just $n = 12$ in Eq. 1. Though the empirical evidence is that the model fits reasonably well, the strength of the asymptotic basis for this model should be treated with some caution.

4 Vargas, Venezuela (Maiquetia International Airport)

As stressed in Sect. 2, it is essential in an extreme value analysis to exploit all relevant information. In the previous example, there was precious little data, and so little advantage in digressing from a conventional annual maximum type of analysis (apart from the importance of using a GEV family and the preference for a Bayesian analysis). In this section we consider a different example based on measurements of rainfall recorded at Maiquetia International Airport, Venezuela, for the period 1951–1999 (see Fig. 3). Daily observations are available from 1961, making a threshold-type analysis available. There is particular interest in this series since the event

Table 1 Predicted volumes for Lago de Managua at 50, 100, 500 and 1000 year return levels in Hm³, under various annual maxima models. Final columns denote the expected return period (years) for an event the size of Mitch, and the negative log likelihood evaluated at the maximum likelihood estimate

Model	Return period (years)				Return period of Mitch (4,031 Hm³)	$-\log(L)$
	50	100	500	1000		
MLE Gumbel	1,394	1,583	2,021	2,209	823,570	462.94
Bayes Gumbel	1,384	1,584	2,051	2,255	343,042	
MLE GEV	1,786	2,204	3,445	4,121	916	460.74
Bayes GEV	1,967	2,574	4,887	6,532	311	

Fig. 3 Time series and realizations against day of year for the rainfall values recorded in Venezuela. Rainfall is measured in mm. Annual maxima are represented by *open circles*. Lower plot focuses on extremes by containing only events above 10 mm
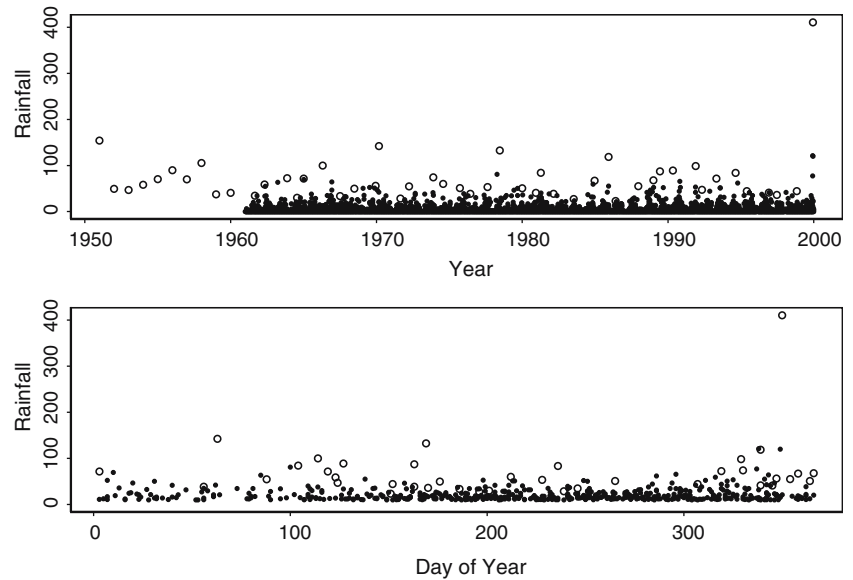
which occurred on 15th December 1999, at over 410 mm, was almost three times greater than any previously recorded and was regarded as virtually impossible by previously published models for the process.

Analyses of these data were previously performed in an unpublished technical report by González and Córdova (2000). Inference based on fitting the Gumbel model to annual maxima prior to 1999 suggested a return period of some 17 million years for an event of the magnitude which subsequently occurred in 1999. Coles and Pericchi (2003) then suggested a number of modifications which, in combination, led to an analysis which gives much greater credibility to the 1999 event—see Table 2. The fundamental changes were: replacement of the Gumbel with the GEV model; adoption of a Bayesian inference; use of the daily data; handling of seasonality. The first two of these aspects correspond to the changes we made in the analysis in the previous section, but the extra information now available in the daily data enabled the analysis to be taken further.

A particular feature of the analysis by Coles and Pericchi (2003) was the specification of a two-season structure to the rainfall process, but with non-specified seasonal changepoints. Specifically, the model assumed two seasons $\mathscr{I}_1$ and $\mathscr{I}_2$, within each of which the tail parameters of the model Eq. 6 were assumed constant, with the intervals themselves being treated as unknown quantities. Based on the pre-1999 data, it emerged that whilst an event of the magnitude of that of December 1999 was indeed exceptional, its level was by no means

unforeseeable—the predictive probability of such an event at some point during a 50-year observation period was around 1/3.

This analysis is open to criticism of the assumption of a two-season structure, despite accordance to the meteorology of the region consisting of both 'wet' and 'dry' seasons. If the modelling of seasonal-structure is crucial in properly accounting for fluctuations in the process, sensitivity of any conclusions to imposing a two-season structure should be examined.

Here we re-evaluate the analysis of Coles and Pericchi (2003) by lifting the restriction on a known and fixed number of seasons. Adopting a Bayesian model averaging framework we analysed a model containing r seasons, where r is now a random variable upon which we require to make inference. While requiring a more technical computational implementation to handle the non-constant dimensionality of the problem (Green 1995), we are able to effectively integrate over our uncertainty in the number of seasons.

Following Eq. 7, the log-likelihood for an r-season model is given by

$$l(\mu, \sigma, \xi) \propto -\sum_{j=1}^{r} \left\{ \sum_{i \in \mathscr{I}_r} \left(\log \sigma_j + (1 + 1/\xi_j) \right. \right.$$
$$\left. \log \left[1 + \xi_j \left(\frac{x_i - \mu_j}{\sigma_j} \right) \right] \right) + \frac{n_y |\mathscr{I}_r|}{n} \left[1 + \xi_j \left(\frac{u - \mu_j}{\sigma_j} \right) \right]^{-1/\xi_j} \right\},$$

where $\mu = (\mu_1, \ldots, \mu_r)$, $\sigma = (\sigma_1, \ldots \sigma_r)$ and $\xi = (\xi_1, \ldots, \xi_r)$ are now vectors of (random) length r, $|\mathscr{I}_r|$ is the number of days in season $\mathscr{I}_r$, and n_y is the total number of years of observation. The prior distribution, which we specify as $\pi(\mu, \sigma, \xi, r) \propto \pi(r) \prod_{j=1}^{r} \pi_j(\mu_j, \sigma_j, \xi_j)$, is similarly defined across all models. Apart from the non-constant dimensionality, and a prior distribution of $\pi(r-1) \sim$ Poisson(1)

Table 2 Expected return period (years) for an event the size realized on 15th December, 1999, in Vargas, Venezuela

	MLE	Bayes
Gumbel	17,600,000	2,170,000
GEV	4,280	660

Table 3 Prior and posterior model probabilities for number of seasonal components

Model	Number of seasons					
	1	2	3	4	5	6
Prior	0.368	0.368	0.184	0.061	0.015	0.003
Posterior excluding 1999 event	0.000	0.873	0.123	0.004	0.000	0.000
Posterior including 1999 event	0.000	0.584	0.381	0.034	0.002	0.000

Table 4 Return period estimates (years) for Vargas, Venezuela, of 410.4 mm using models with differing seasonal assumptions, both including and excluding the 1999 event

Model	Predictive return period of 410.4 mm	
	1999 Data excluded	1999 Data included
Homogeneous model	260.0	116.0
2 Season model	130.5	69.8
r Season model	129.6	69.6

on the unknown number of seasons, which has the advantage of a prior mean value coincident with the meteorological belief in a two-season structure, all other details are similar to those adopted in the original analysis. We note that adopting an improper prior $\pi_j(\mu_j,\sigma_j,\xi_j)P\ 1$ becomes problematic in this multi-model setting, as unknown normalizing constants then define the posterior model probabilities. Instead we specify $\pi_j(\mu_j,\sigma_j,\xi_j) = N(0,\tau_\mu)\times\log\ \ N(v,\tau_\sigma)\times N(0,\tau_\xi)$, and adopt sensible values for τ_μ, τ_σ, τ_ξ and v for which repeated simulations indicate an insensitivity in the posterior to their specification.

A summary of results in terms of inference on seasonal structure, that is the posterior $\pi(r-1|x)$ having seen the data, x, is given in Table 3. By any measure the evidence against a homogeneous structure across the entire year is overwhelming, as is the evidence against a structure more complicated than three-seasonal. It

would seem then that the data entertains only the possibility of two or three seasons per year. Furthermore, measured in terms of posterior probabilities of the number of seasons, the support for a three-seasonal structure is considerably weaker than that of a two-seasonal structure, although this is less the case when the 1999 event is included as specific parameter combinations are required to account for the magnitude of the event. Moreover, even if uncertainty in the number of seasons is maintained, results are largely unchanged: see Table 4 and Fig. 4. This suggests that a two-seasonal structure, as proposed in the meteorological literature and adopted by Coles and Pericchi (2003), actually has considerable empirical support within the historical data, strengthening the validity of the original analysis.

5 San Juan, Puerto Rico (Luis Muñoz Marin International Airport)

Given the apparent failure of the Gumbel model in each of the two previous examples, its widespread use as the basic model for annual maxima is disconcerting. Another example drawn from the Caribbean relates to rainfall levels in Puerto Rico. The current estimation of hazards in Puerto Rico (U.S. Department of Commerce 1961, Technical Paper No. 42 and its updates) still mainly rests on the Gumbel model. However, on this basis and using data prior to 1961, the 100-year return

Fig. 4 Predictive return level plots for the 2-seasonal and r-seasonal models for Vargas, Venezuela, both with and without the 1999 event. The 2- and r-seasonal return levels coincide almost exactly—the two obvious curves correspond to without (*solid line*) and with the 1999 event respectively. Points correspond to empirical estimates
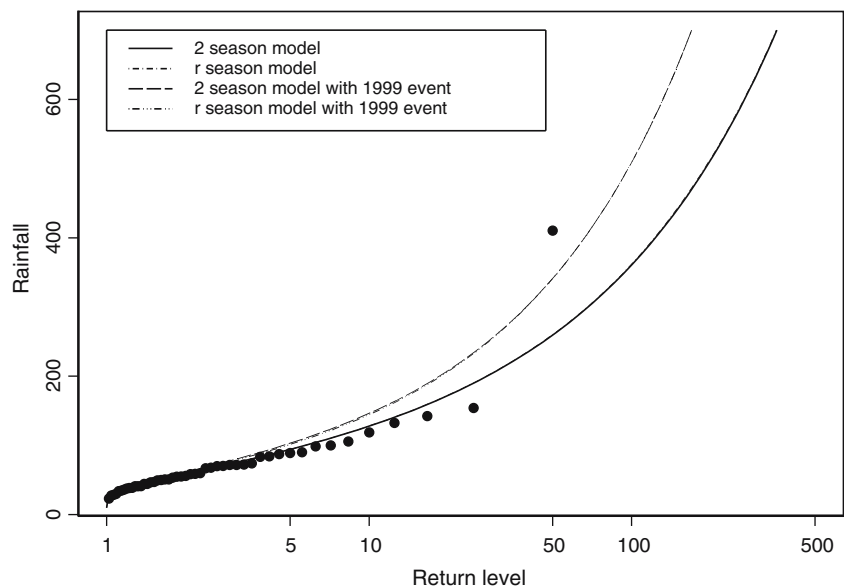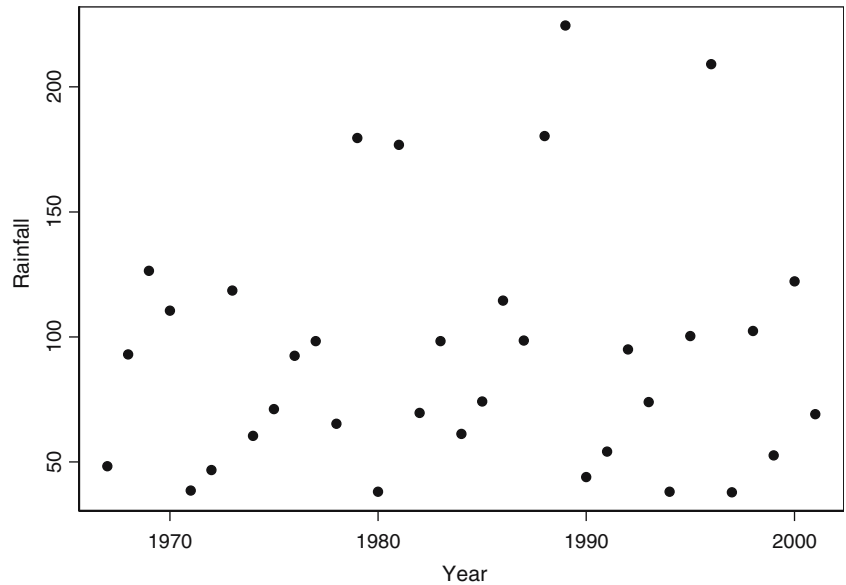
Fig. 5 San Juan annual maxima of daily rainfall. Rainfall is measured in mm

level maximum 24 h rainfall near the southern town of Ponce, for example, was estimated in the range 12–14 in. This falls excessively short of levels realized during the rainstorms that caused the Barrio Mameyes landslide disaster just northeast of the city. In particular, on 6–7 October 1985 over 22 in. of rainfall were recorded in one 24 h period. Just as with our previous examples, such levels were regarded as virtually impossible by the Gumbel analysis.

In a more illustrated example, Fig. 5 shows 35 years of annual maxima of daily rainfall recorded at Luis Muñoz Marin International Airport, San Juan, Puerto Rico. These data form a subset of a much larger archive of rainfall figures recorded at a number of locations within the island by the local National Weather Service. Although no exceptionally extreme rainfall or flooding

events are apparent in the airport location data set, the island's geographical location places it within a high risk area for such events. Analyses similar to those of the Lago de Managua data result in fitted return level plots (Fig. 6) that again display wide variation in predictive estimates of future extreme events. Modelling with the generalized extreme value distribution in a Bayesian framework again offers the shortest estimates for recurrence of events of fixed magnitudes. Table 5 enumerates a number of predicted maximum rainfall estimates for a range of return periods, for all models. Even for return periods of 50 and 100 years, commonly used as design parameters in engineering applications, the predicted extreme rainfall levels vary considerably. Of particular interest are estimates of the 100-year return period, which is itself a standard infrastructure design

Fig. 6 Return levels for San Juan annual maxima data using both Gumbel and GEV distributions. Both Bayesian and maximum likelihood estimates are shown
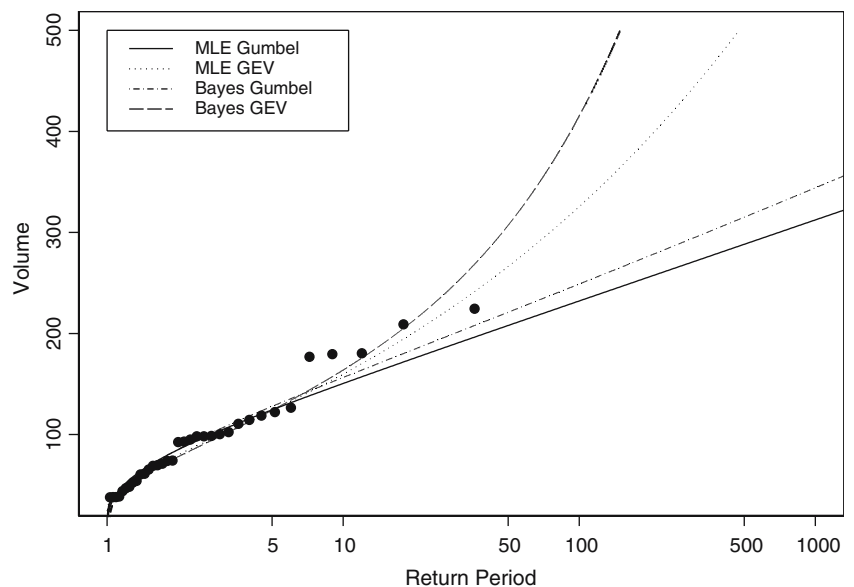
Table 5 Expected rainfall (mm) for San Juan at 50, 100, 500 and 1000 year return levels for both Gumbel and GEV models, using maximum likelihood and Bayesian frameworks

Model	Return period (years)			
	50	100	500	1000
Gumbel MLE	208	232	288	312
Gumbel Bayes	221	249	315	343
GEV MLE	266	326	507	609
GEV Bayes	307	416	938	1,404

Table 6 Maximum likelihood estimates (MLE) and Bayesian posterior means with standard errors for model parameters in standard and truncated GEV models, for annual San Juan data

	μ	σ	ξ
MLE	68.06 (6.40)	30.83 (5.30)	0.24 (0.20)
Truncated MLE	68.05	30.82	0.24
Bayes	68.67 (6.64)	34.26 (6.32)	0.27 (0.22)
Truncated Bayes	68.60 (6.77)	34.38 (6.60)	0.30 (0.25)

parameter. In this case the Gumbel maximum likelihood estimate is roughly half that of the Bayesian predictive estimate based on the GEV.

It must be admitted that it is rather easy to adopt our position of advocating the use of a Bayesian GEV model, since this will generally tend to a more conservative estimate than either maximum likelihood or Gumbel model approaches. Practitioners with design responsibility cannot always afford this luxury since the cost of designing structures to overly conservative levels may be unaffordable in practice. For example, the predictive 1000-year return level of around 1,700 mm for daily rainfall might well be regarded as completely unrealistic on physical grounds, and predicting beyond this level may not make sense. This argument has been strongly expressed to us by hydrologists (personal communication). A comprehensive approach would be a formal cost-based decision analysis, but this is beyond the scope of this article. Instead, we explore the consequences of imposing a predetermined physical limit on the model.

One method is to restrict the GEV family to negative values of ξ. However, whilst this would result in bounded distributions, the upper tail of these distributions is restrictively light. Moreover, this option would be rather in opposite spirit to the approaches we have adopted and advocated so far. An alternative, that we introduce here is the use of a GEV model truncated at some maximum process bound, m. That is, we propose a distribution function of the form

$$G^*(z) = \begin{cases} \frac{G(z;\mu,\sigma,\xi)}{G(m;\mu,\sigma,\xi)} & z < m \\ 1 & \text{else,} \end{cases} \quad (11)$$

where $G(\cdot;\mu,\sigma,\xi)$ is the GEV distribution defined by Eq. 3. The dependence of the numerator and the normalizing constant in the denominator upon the model parameters is highlighted.

The advantage is that while bounded, the tails of the distribution can be relatively heavy up to the upper limit. Despite failure of the asymptotic support such a departure from the GEV family implies, this model appears to provide a reasonable fit to the data, particularly for truncation at levels well in excess of the realized data observations. Parameter values for model Eq. 11 obtained by maximum likelihood and from posterior marginal means are displayed in Table 6, and the resulting return levels in Fig. 7. Nonetheless, we maintain that this type of modelling should be used only with extreme caution, when the physical limit of a process is absolutely bounded by some known level. In situations where it is believed that a process may be bounded, but the level of such a bound is unknown, it is almost certainly better to work with the full GEV family which

Fig. 7 Return levels for San Juan annual maxima data using both standard and truncated GEV distributions. Both Bayesian and maximum likelihood estimates are shown
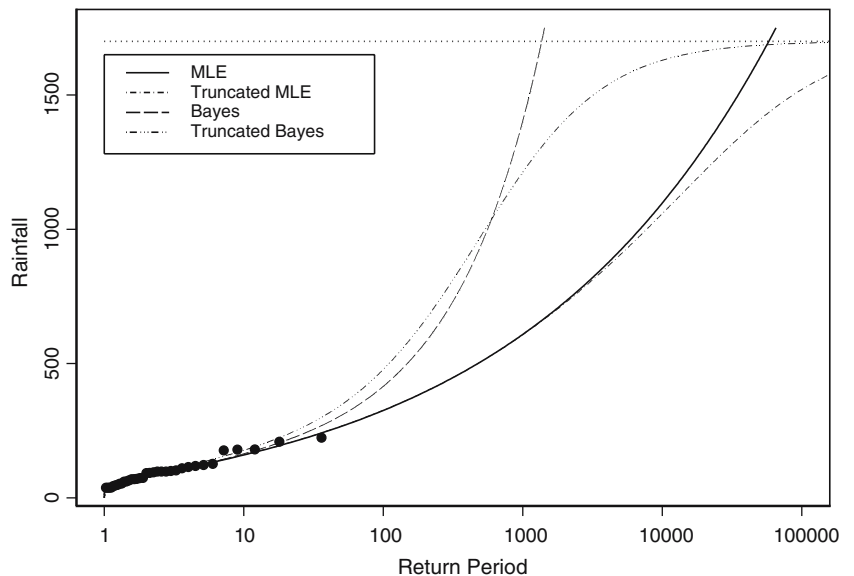
admits both the possibility that there is no such bound and the uncertainty in the value of a bound even if it exists.

6 Discussion

This article discusses models and inference methodologies for long-term risk assessment of hydrological hazards in the Caribbean. We have presented analyses of three sites in the Caribbean region, two of which have witnessed major catastrophes in the form of Hurricane Mitch in Nicaragua, and the Vargas Tragedy in Venezuela, each regarded as virtually impossible events under a standard analysis. Both of these environmental and humanitarian disasters emphasize the need for a critical examination of the standard methods used to generate statistics upon which policy decisions are made. The third site, Puerto Rico, experiences a similar meteorology as the other locations, but in recent history has not suffered a truly devastating hydrological event. It is therefore of great concern that in this case the Bayesian GEV model predicts precipitation events almost twice that of the maximum likelihood Gumbel model at the 100-year return period level.

An essential part of our analysis is the adoption of a Bayesian framework—a natural statistical approach that enables the array of uncertainties involved in predictive inference to be probabilistically incorporated into the analysis, as well as enabling the application of models that are intractable under classical frameworks. Previous quantitative precipitation studies in Puerto Rico have proposed support for a Bayesian approach, for example Caster et al. (2001), although in this case the focus was not on extremes. Additionally, Krzysztofowicz (1983) and Olson et al. (1995) both demonstrate that the Bayesian approach can substantially outperform classical predictive methods in rainfall analyses. Moreover, the bias that can result from extreme events causing missing data is most easily handled within a Bayesian analysis (Carter et al. 2001). This phenomenon actually arose in the Venezuelan case study where one extreme event broke the rainfall recording equipment, and as a consequence the extreme event of the following day was unrecorded (Coles et al. 2003).

We have also suggested that the reduction of model complexity based on often unwarranted assumptions can have profound effects on inference, especially in the prediction of long-term risks. This situation arises both with the adoption of the standard Gumbel distribution over the generalized extreme value distribution, and the case where an overly restrictive a priori seasonal structure is assumed in the case of non-stationary (e.g. daily) data. Thus there is a clear challenge in the case of systems such as Puerto Rico and the Virgin Islands, where the complex local meteorology may prohibit detailed expert opinion on seasonal structure.

To summarize, in this article we hope to have highlighted via example serious shortcomings inherent in procedures currently adopted for risk assessment of extreme hydrological hazards in the Caribbean. We have advocated more flexible inferential methods for the analysis of the extremes of environmental processes, which due to enhanced generality are less susceptible to flawed inferences. However, there clearly remains an urgent need for a detailed systematic and exhaustive review of such standard procedures, particularly for implementation in the Caribbean region.

References

Carter MM, Elsner JB, Bennett SP (2001) A quantitative precipitation forecast experiment for Puerto Rico. J Hydrology 239:162–178

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

Coles SG, Pericchi LR (2003) Anticipating catastrophies through extreme value modelling. Appl Stat 52:405–416

Coles SG, Pericchi LR, Sisson SA (2003) A fully probabilistic approach to extreme rainfall modelling. J Hydrology 273:35–50

Córdova JR, Camachov R (1999) Análisis del períod de retorno del volumen de agua aportado por el huracán Mitch al lago de Managua. Unpublished technical report

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

Gilks WR, Richardson S, Spiegelhalter DJ (1996) Markov chain Monte Carlo in practice. Chapman & Hall, CRC

Goldenberg SB, Landsea CW, Mestas-Nuñez AM, Gray WM (2001) The recent increase in atlantic hurricane activity: causes and implications. Science 293:474–479

González M, Córdova JR (2000) Consideraciones sobre la probabilidad de ocurrencia de lluvias máximas en la zona litoral de norte de Venezuela. Memorias del Seminario International 'Los Aludes Torrenciales de Diciembre 1999 en Venezuela', Instituto de Mecánica de los Fluidos, Universidad Central de Venezuela, Science

Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82:711–732

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109

Kite GW (1977) Frequency and risk analyses in hydrology. Water Resources Publications, Fort Collins

Krzysztofowicz RA (1983) Why should a forecaster use Bayes theorem? Water Resources Research 19:327–336

Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and series. Springer, Berlin Heidelberg New York

Olson DA, Junker NW, Korty B (1995) Evaluation of 33 years of quantitative precipitation forecasting at the NMC. Weather Forecasting 10:498–511

Ramesh NI, Davison AC (2002) Local models for exploratory analysis of hydrological extremes. J Hydrology 256:106–119

Robert C, Casella G (2004) Monte Carlo statistical methods, 2nd edn. Springer, Berlin Heidelberg New York

Smith RL (1989) Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). Stat Sci 4:367–393

Walshaw D (2000) Modelling extreme wind speeds in regions prone to hurricanes. Appl Stat 49:51–62

U.S. Department of Commerce (1961) Generalized estimates of probable maximum precipitation and rainfall frequency data from Puerto Rico and the Virgin Islands (Technical Paper 42). Precipitation frequency study for Puerto Rico and the U.S. Virgin Islands (5th revision)