Impact of Dependence on Parameter Estimates of Autoregressive Process with Gumbel Distributed Innovation

[1]Bako Sunday Samuel*, **[2]Mohd Bakri Adam and [3]Anwar Fitrianto**

[1,2,3]Department of Mathematics, Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia

[2]Institute of Mathematical Research, Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia

[3]Kaduna State University
P.M.B 2339 Tafawa Balewa Way, Kaduna, Nigeria

*Corresponding author: bakosundays@gmail.com

Abstract Recent studies have shown that independent identical distributed Gaussian random variables is not suitable for modelling extreme values observed during extremal events. However, many real life data on extreme values are dependent and stationary rather than the conventional independent identically distributed data. We propose a stationary autoregressive (AR) process with Gumbel distributed innovation and characterise the short-term dependence among maxima of an (AR) process over a range of sample sizes with varying degrees of dependence. We estimate the maximum likelihood of the parameters of the Gumbel AR process and its residuals, and evaluate the performance of the parameter estimates. The AR process is fitted to the Gumbel-generalised Pareto (GPD) distribution and we evaluate the performance of the parameter estimates fitted to the cluster maxima and the original series. Ignoring the effect of dependence leads to overestimation of the location parameter of the Gumbel-AR (1) process. The estimate of the location parameter of the AR process using the residuals gives a better estimate. Estimate of the scale parameter perform marginally better for the original series than the residual estimate. The degree of clustering increases as dependence is enhance for the AR process. The Gumbel-AR(1) fitted to the threshold exceedances shows that the estimates of the scale and shape parameters fitted to the cluster maxima perform better as sample size increases, however, ignoring the effect of dependence lead to an underestimation of the parameter estimates of the scale parameter. The shape parameter of the original series gives a superior estimate compare to the threshold excesses fitted to the Gumbel distributed Generalised Pareto ditribution.

Keywords Autoregressive process; Gumbel distribution; generalised Pareto distribution; extremal index; Innovation.

Mathematics Subject Classification 37M10, 60G70.

1 Introduction

Extreme value theory as one of the most important statistical discipline has been applied in various fields, from atmospheric chemistry to environmental sciences to financial econometrics. The distinguishing feature of extreme value analysis is quantifying the stochastic behaviour of maxima (or minima) and the assessment of extremal behavior of random variables. Previous studies most time focus on independently and identically distributed random variables to consider the statistical property of their maxima (or minima) using parametric models [1]. It is straightforward under the independence assumptions to compute the likelihood function and estimate the parameters using the maximum likelihood estimators. However, due to temporal dependence, short-range dependence of extremes that leads to clusters of observations usually arise and the independence assumption of Extreme Value Theory no longer holds. Hence, an appropriate theoretical approach is needed to analyse the extremes of dependent series.

In recent years, evolving extreme value models have attracted considerable attention in literatures by investigating the degree of clustering cause by dependence of extreme values. The extension of extreme values to time series can be achieve by inducing dependence in the underlying state of the extreme value process. One approach to capture such dependence is to consider an autoregressive process for the model parameters of the extreme value distribution using a state space representation, see [2]. Through simulation study, Prado and Lopes [3] generated various Autoregressive state-space models using structured prior distributions to assess the performance of sequential filtering and parameter learning. Smith [4] also considered a class of max-stable processes approach and provides some application of the max-stable classes for the extreme of financial data. Toulemonde *at al.* [5] proposed a linear Gumbel distributed Autoregressive (AR) model using the logarithm of a positive α-stable random variable and the objective was to extend linear Autoregressive model to capture temporal dependence for Gumbel distributed maxima.

In this paper, we characterise the behaviour of the parameter estimates of an autoregressive process with Gumbel distributed innovation. Dependence is induce in the autoregressive process through the linear filtering of the Gumbel distributed innovation. We begin with the Generalised Extreme Value (GEV) model with time dependence which is induced by an autoregressive process having Gumbel innovations. The GEV distribution with location, scale and shape μ, σ, ξ respectively commonly use for the analysis of maxima of some large set of random variables is given by

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

where $\sigma > 0$ and $1 + \xi\frac{z-\mu}{\sigma} > 0$. The subset of Eq. (1) with $\xi = 0$, which is interpreted as the limit of $G(z)$ as $\xi \to 0$, is known as the Gumbel distribution with normalized maxima from GEV with light-tailed which is given as

$$G(x) = \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], \quad -\infty < x < \infty. \tag{2}$$

Let α_t be a stationary time series process with innovation following the Gumbel distribution given in Equation (2), therefore

$$x = \mu - \sigma\log(-\log\alpha_t). \tag{3}$$

If a time-dependence that follows a Gumbel distribution exists in a time series realisation, then Equation (3) provides a better approximation to the underlying process than the Gumbel distribution assuming time independence and Equation (3) is therefore useful in describing the sensitivity of the estimates of the parameters of the Gumbel distribution and of the GPD fitted to both the cluster maxima and to the original series.

If Gumbel distributed maxima are dependent, but this dependence is not recognised, what will happen to the resulting estimates of the Gumbel parameters? The main objective of this paper is to use a time-varying linear filtering operation to generate an autoregressive process with Gumbel innovations and to assess the performance of the resulting parameter estimates of the autoregressive process for different sample sizes with varying degrees of dependence. We also examine and compare the performance of the parameters of the GPD fitted to the cluster maxima and to the original series.

The rest of the paper is organised into the following sections: Sect. 2 which is divided into two subsections is the methodological aspects of the maximum likelihood estimates that shows the derivative of the parameters for the Gumbel distribution and the GPD under the maximum likelihood estimator. Section 3 presents the derivation of the model specification for the Gumbel-AR model through the linear filter operation. Section 4 presents the simulation study. Section 5 presents the results and discussion. A brief summary of conclusions from the study is given in Sect. 6.

2 Methodological Aspects

2.1 Generalized Extreme Value Distributions

The GPD is a distribution with scale, σ, and shape, ξ, parameters respectively and is use to model maxima above a given threshold, u. MLEs of the parameters are asymptotically normal and asymptotically efficient in many cases and therefore are preferred in the estimation of both σ and ξ. We proceed by stating the generalised Pareto distribution and show how generalised extreme value distribution and the GPD models are related before deriving the ML estimate of the GPD.

Cole [1] define a sequence $X_1, X_2, \cdots, X_n$ of independent and identically distributed random variables with common distribution function, F, that follows the GEV distribution in Equation (1), as a GPD distribution if

$$M_n = \max\{X_1, \cdots, X_n\}$$

so that

$$\Pr\{M_n\} \approx G_{(x)}$$

and for any reasonable large threshold u, we define the distribution function of $(X - u)$, conditional on $X > u$, as approximately

$$H(z) = 1 - \left(1 + \frac{\xi z}{\hat{\sigma}}\right)^{-\frac{1}{\xi}}, \tag{4}$$

which is defined on $\{z : z > 0\}$ and $(1 + \frac{\xi z}{\hat{\sigma}}) > 0$, where $\hat{\sigma} = \sigma + \xi(u - \mu)$ and z is the threshold excess. Equation (4) is call the GPD with scale and shape parameters σ and ξ respectively.

For the transformation from GEV to GPD, lets denote an arbitrary term in the sequence $X_1, X_2, \cdots, X_n$ by X. Then the GPD is given as

$$\Pr\{X > u + z | X > u\} = \left[1 + \frac{\xi z}{\tilde{\sigma}}\right]^{-\frac{1}{\xi}}, \tag{5}$$

where $\tilde{\sigma} = \sigma + \xi(u - \mu)$ and z is the excess above a given threshold.

An important parameter in extreme value theory is the extremal index. The extremal index measure the degree of clustering of extreme observations [6]. Let $(Y_j), -\infty < j < \infty$, be a strictly stationary sequence and $\theta \in (0, 1)$ a non-negative number. Assume that for every $\tau > 0$ there exists a sequence (u_n) such that

$$\lim_{n \to 0} n\overline{F}(u_n) = \tau \tag{6}$$

and

$$\lim_{n \to \infty} \Pr(M_n \leq u_n) = e^{-\theta\tau}. \tag{7}$$

Then θ is called the extremal index of the sequence (Y_j). Unless the extremal index, θ, is equal to one, the limit distributions for stationary and independent sequences are not the same. Clustering in observations signify the existence of dependence in the sequence of observations.

The derivation of the GPD from the GEV makes the dependence of the GPD parameters on the threshold, u, obvious. The parameters of the GPD of threshold excesses are determined by those of the corresponding GEV distribution and in particular, the shape parameter of the GPD is equal to that of the associated GEV. Let $x_1, x_2, \ldots, x_n$ be a sequence of observations of independent random variables which are the exceedances above a threshold,u. The log-likelihood function of the GPD is derived from taking logarithm of the joint density function. Let n_u be the number of observations above the thresholds u. Then the log-likelihood function of $x_1, x_2, \ldots, x_n$ of the i.i.d Generalised Pareto random variables is given by

$$\log L(\sigma_u, \xi | X) = -n_u \ln \sigma_u - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n_u} \ln \left[1 + \xi \left(\frac{x_i - u}{\sigma_u}\right)\right], \tag{8}$$

with $\left[1 + \xi \left(\frac{x_i - \mu}{\sigma_u}\right)\right] > 0$ and $\xi \neq 0$. The constraint implies $x < \mu - \sigma_u/\xi$. When $\xi = 0$, the log-likelihood function can be derived as

$$\log L(\sigma_u | X) = -n_u \ln \sigma_u - \sum_{i=1}^{n_u} \left(\frac{x_i - u}{\sigma_u}\right) \tag{9}$$

where $\xi = 0$.

3 Model Specification for the Gumbel-AR(1)

Define the autoregressive model with a first order AR process (Gumbel-AR(1)) as

$$\alpha_t = \phi\alpha_{t-1} + \eta_t, \eta_t \sim G(x), t = 1, 2, \ldots, n \tag{10}$$

where α_t is the output of the linear filter of the innovation η_t which is distributed as the Gumbel random variable in Eq. (2). For the Gumbel-AR(1) in Equation (10) to be stationary, the $|\phi| < 1$.

The error form (innovation) η_t in Eqquation (10) is suppose transformed to the process α_t by what is called a linear filter. The linear filtering operations simply takes a weighted sum of previous innovations η_t.

Hence, for the Gumbel-AR(1) in Equation (10), we have

$$
\begin{aligned}
\alpha_t &= \phi\alpha_{t-1} + \eta_t \\
&= \phi(\phi\alpha_{t-2} + \eta_{t-1}) + \eta_t \\
&= \phi^2\alpha_{t-2} + \phi\eta_{t-1} + \eta_t \\
&= \phi^3\alpha_{t-3} + \phi^2\eta_{t-2} + \phi\eta_{t-1} + \eta_t + \cdots.
\end{aligned}
$$

provided $|\phi| < 1$, the value of ϕ^k will get smaller, so that from Equation (10) we obtain

$$
\alpha_t = \eta_t + \phi\eta_{t-1} + \phi^2\eta_{t-2} + \phi^3\eta_{t-3} + \cdots. \tag{11}
$$

Thus, our linear filtering of the time-dependent Gumbel distribution is base on the fact that Equation (11) holds.

3.1 Bootstrapping the Gumbel-AR(1) Model

Bootstrap is a computer-based method for assigning measures of accuracy to statistical estimates. Let α_t be a stationary time series realisation at time t with n observations that is distributed as the Gumbel-AR(1) in Equation (10). The bootstrap, like the Jackknife is a resampling of the data points in α_t. While the Jackknife considers n new data sets, each contains all the original data points less 1 observation, the bootstrap uses n_{boot} data sets, each containing n observations which are obtained by random sampling of the original set of n observations. The random sampling is done with replacement, and the probability that an observation is picked is $1/n$ [7].

Efron [8] pioneered the bootstrap geared towards independent data. Smith [9] recognised that if one applies independent identical distributed bootstrap to data that are dependent, inconsistency follows. Because of the total data scrambling induced by i.i.d bootstrap, all dependence information will be lost. For the Gumbel-AR(1) model in Equation (10), the standard bootstrap method cannot be applied because our objective is to retain the dependence structure of the model. [10] first introduced the block bootstrap to be use for dependent data. The block bootstrap method propose by [11] is use in this study. The basic idea of the block bootstrap is closely related to the i.i.d non-parametric bootstrap. The two procedures are based on drawing observations with replacement. In the block bootstrap however, instead of single observations, blocks of consecutive observations are drawn. This is done to ensure the dependence structure of neighbourhood observations are captured.

4 Simulation Study

In this section, simulation study is done to examine the behaviour of the estimators of the Gumbel-AR(1) model. We also examine the estimates of the parameters of the Gumbel

GPD fitted to the cluster maxima and the original series. Simulation will be perform for $n \in \{300, 500, 1000, 3000, 5000\}$ and $\phi \in \{0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, and without loss of generality we set $\mu = 0$ and $\sigma = 1$.

For each n and ϕ, 10,000 bootstrap samples will be generated from the Gumbel-AR(1) model. As the value of ϕ increases, dependence is enhance. The value of ϕ is steadily increase to assess the sensitivity and consistency of the respective estimate of the model parameters to changes in the degree of dependence for a reasonable range of sample sizes. For each combination of n and ϕ, Augmented Dickey Fuller (ADF) will be perform on the respective models to check if they are indeed stationary. Furthermore, for each sample size n and ϕ, we estimate μ and σ for the Gumbel-AR(1) and its residuals using the maximum likelihood estimate.

Having establish the presence of dependence in the autoregressive, we proceed to decluster the data and then fit the Gumbel Generalised Pareto distribution to the threshold excesses. An important property of the extremal index estimator use for threshold selection is the threshold stability plot. The threshold stability plot examines varying threshold for invariance of the extremal index to change in threshold. For every given threshold, the extremal index is computed. In other to decluster the Gumbel-AR(1) model, the declustering scheme propose by [12] will be use. The method involve estimating the extremal index of the sequence of excesses above a given threshold of the respective models before the identification of independent clusters of exceedances above the threshold. A suitable threshold for the declustering is indicated by interexceedance times which correspond to inter-cluster times lying close to a diagonal line indicated by a quantile-quantile plot.

After declustering the data, the exceedances of the thresholds are then use to estimate the extremal index and hence, the parameters of the GPD_{gumbel}, ξ and σ, can then be estimated for both the cluster maxima and for the original series.

5 Results and Discussion

In this section, results and discussion of the simulation study are presented in Table 1 to Table 10.

Table 1: Gumbel-AR(1) Lag 1 Autocorrelation

n	ϕ									
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
300	0.036	0.091	0.189	0.296	0.398	0.495	0.588	0.697	0.791	0.896
500	0.034	0.090	0.186	0.293	0.391	0.493	0.573	0.696	0.788	0.887
1000	0.020	0.088	0.170	0.290	0.390	0.490	0.569	0.694	0.784	0.884
3000	0.011	0.064	0.160	0.281	0.385	0.482	0.561	0.690	0.778	0.875
5000	0.007	0.054	0.153	0.258	0.379	0.465	0.551	0.689	0.755	0.872

Table 1 give the lag 1 autocorrelation of the Gumbel autoregressive process. As the value of ϕ increases, the autocorrelation increases for a fixed sample size. For constant ϕ value with increasing sample size, the autocorrelation diminishes.

Table 2: Gumbel-AR(1) Estimate of μ with Different Combination of n and ϕ with $\mu = 0$

	ϕ									
n	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
300	0.6907	0.7513	0.7986	0.9570	1.1151	1.2822	1.3558	1.9687	3.1117	6.1317
500	0.6763	0.7173	0.7678	0.8996	1.1046	1.1312	1.3403	1.9120	3.0227	5.9336
1000	0.5262	0.6875	0.7332	0.8697	1.0048	1.1105	1.3272	1.9005	3.0202	5.8030
3000	0.5195	0.6300	0.6821	0.8037	0.9198	1.0873	1.3229	1.9001	3.0023	5.6065
5000	0.5183	0.5828	0.6459	0.7845	0.9138	1.0809	1.1445	1.8074	2.9751	5.5123

Table 3: Residual of Gumbel-AR(1) Estimate of μ with Different Combination of n and ϕ with $\mu = 0$

	ϕ									
n	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
300	0.0048	0.0049	0.0050	0.0051	0.0054	0.0056	0.0057	0.0060	0.0061	0.0066
500	0.0046	0.0047	0.0048	0.0050	0.0053	0.0055	0.0056	0.0059	0.0060	0.0064
1000	0.0045	0.0046	0.0047	0.0049	0.0051	0.0052	0.0054	0.0056	0.0058	0.0060
3000	0.0039	0.0040	0.0041	0.0045	0.0046	0.0047	0.0049	0.0050	0.0053	0.0058
5000	0.0037	0.0039	0.0040	0.0044	0.0045	0.0046	0.0047	0.0049	0.0051	0.0057

Table 2 and Table 3 presents the estimate of the location parameter μ for the Gumbel-AR(1) and its residual respectively. As dependence is increase, the estimate of the location parameter of the autoregressive process and its residual increases for a fixed sample size. For a fixed ϕ value, the estimate of the location parameter perform better as n increases. However, a simultaneous increase in n and ϕ lead to a corresponding increase in the estimate of the location parameter of the autoregressive process.

When $\phi \geq 0.7$, the estimate of the location parameter is quite large and unreasonable therefore, when fitting data to autoregressive processes with Gumbel innovation without taking the effect of dependence into consideration, the location parameter is overestimated.

Table 4: Gumbel-AR(1) Estimate of σ with Different Combination of n and ϕ with $\sigma = 1$

	ϕ									
n	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
300	0.0120	0.0122	0.0126	0.0127	0.0132	0.0136	0.0143	0.0160	0.0176	0.0250
500	0.0126	0.0127	0.0130	0.0139	0.0140	0.0148	0.0152	0.0164	0.0198	0.0255
1000	0.0127	0.0129	0.0135	0.0140	0.0141	0.0153	0.0154	0.0168	0.0207	0.0286
3000	0.0128	0.0132	0.0136	0.0142	0.0144	0.0155	0.0156	0.0170	0.0248	0.0307
5000	0.0129	0.0133	0.0139	0.0145	0.0147	0.0156	0.0162	0.0174	0.0260	0.0309

Table 4 and Table 5 give the estimate of the scale parameter for the Gumbel-AR(1) and

Table 5: Residual of Gumbel-AR(1) estimate of σ with different combination of n and ϕ with $\sigma = 1$

n	ϕ									
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
300	0.0118	0.0121	0.0123	0.0125	0.0127	0.0129	0.0131	0.0132	0.0137	0.0143
500	0.0120	0.0122	0.0124	0.0126	0.0128	0.0130	0.0132	0.0134	0.0139	0.0145
1000	0.0123	0.0124	0.0125	0.0127	0.0129	0.0131	0.0133	0.0136	0.0141	0.0146
3000	0.0124	0.0125	0.0126	0.0128	0.0130	0.0133	0.0134	0.0138	0.0145	0.0149
5000	0.0125	0.0126	0.0128	0.0130	0.0132	0.0134	0.0136	0.0140	0.0147	0.0152

its residual respectively. The estimate of the scale parameter increases for a fixed sample size with varying ϕ and a fixed ϕ with varying sample size respectively for both the original series and its residual.

From Table 2, as dependence is increase, the estimate of the location parameter becomes larger. However, for a fixed ϕ value with increasing sample size, the estimate of the location parameter decreases. From Table 7 and Table 8, when ϕ is increase for a constant sample size, the estimate of the scale parameter increases. Same situation is observe as the sample size increases for a fixed ϕ value. A simultaneous increase in n and ϕ increases the estimate of the scale parameter.

From result presented in Table 2, ignoring the effect of dependence in the model increases the estimate of the location parameter. For a fixed dependence level with an increasing sample size, the estimate of the location parameter improves; but an increase in both n and ϕ results in over estimating the location parameter. Results shown in Table 3 suggest that estimating the location parameter using the residuals of the Gumbel-AR(1) gives a superior estimate.

Table 6: Threshold Used to Decluster the Gumbel-AR(1) for Sample Size, n, with ϕ Degree of Dependence

n	ϕ									
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1000	0.562	0.585	0.749	0.787	0.920	1.093	1.347	1.706	2.827	5.515
3000	0.566	0.610	0.753	0.795	0.945	1.219	1.534	1.817	2.856	5.700
5000	0.575	0.653	0.764	0.830	0.982	1.220	1.539	1.961	2.962	5.818

For the estimate of the scale parameter presented in Table 4, as ϕ increases, the estimate perform better. Similarly, at a constant ϕ value with increasing sample size, the estimate of the scale parameter perform better. From result presented in Table 5, we see that estimating the scale parameter using the original series gives a marginally better estimate than using the residual of the Gumbel-AR(1).

Table 6 shows the threshold use to decluster the Gumbel-AR(1) model. After declustering the data using the automatic declustering scheme proposed by [12], the exceedances of the thresholds are then used to estimate the extremal index and hence, the parameters of the

GPD can then be estimated from the resulting declustered series. A suitable threshold for the declustering is indicated by interexceedance times which correspond to inter-cluster times lying close to a diagonal line which is indicated by a quantile-quantile plot. From Table 6, for a fixed sample size and an increasing ϕ value, the threshold show an increasing trend. As the sample size increases for a fixed ϕ value, the threshold stability increases; also, increasing n and ϕ simultaneously increases the threshold use for declustering. Thus, a high dependence in the model implies that a high threshold will be use in declustering the data and an increase in sample size leads to a corresponding increase in the threshold use for declustering.

Table 7: Estimates of the Extremal Index for Sample Size, n, with ϕ Degree of Dependence

n	$\phi = 0$	$\phi = 0.1$	$\phi = 0.2$	$\phi = 0.3$	$\phi = 0.4$	$\phi = 0.5$	$\phi = 0.6$	$\phi = 0.7$	$\phi = 0.8$	$\phi = 0.9$
1000	1.0000	0.9449	0.9176	0.7513	0.7454	0.7337	0.5243	0.4813	0.3176	0.2896
3000	1.0000	0.9784	0.9773	0.8853	0.8298	0.7339	0.5999	0.5096	0.4743	0.2943
5000	1.0000	0.9879	0.9833	0.9043	0.8428	0.7903	0.6225	0.5209	0.4839	0.3467

Table 8: Independent Clusters for Sample Size, n

n	ϕ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1000	164	63	141	94	94	101	65	63	19
3000	337	212	161	170	194	215	89	154	51
5000	390	448	329	349	220	259	175	147	91

The estimate of the extremal index, θ, for the Gumbel-AR(1) is given in Table 7. The extremal index is an important parameter use in measuring the degree of clustering of extremes in a stationary process and $\theta \in [0, 1]$, where $\theta = 1$ indicates that exceedances occur singly and $\theta < 1$ implies that exceedances occurs in clusters. The case $\theta = 0$ is rare, although not impossible and it entails that sample maxima of the Gumbel-AR(1) are of smaller order than sample maxima of the associated independent Gumbel distribution. Clustering is an indication of dependence in the sequence of observations. From Table 7, as dependence is enhance for a fixed sample size, the degree of clustering increases. However, for a fixed dependence level with increasing sample size, the degree of clustering decreases and the model tend towards an i.i.d model. Consequently, the number of independent clusters increases as the sample size increases. When $\phi = 0$, the extremal index is one indicating that the Gumbel-AR(1) is indeed i.i.d. Enhancing the dependence level for an increasing sample size leads to a corresponding increase in the degree of clustering.

After declustering the Gumbel-AR(1), the GPD-AR(1) is assumed to be independent and identically distributed, and varying ϕ have no impact whatsoever on the declustered data. Therefore, our interest is now to focus on consistency for an increasing n only after declustering the Gumbel-AR(1).

Table 8 gives the number of observations that exceeds the respective threshold use in estimating the extremal index. Out of the original series n, we have identified serial dependence

in the data (with the exception of $\phi = 0$ with extremal index 1), and therefore, the threshold excesses are not independent and in fact correspond to the number of independent clusters for the different sample sizes shown in Table 8

Having decluster the Gumbel-AR(1), we proceed to fit the generalised Pareto distribution and estimate the parameters, ξ and $\sigma_{GPD_{gumbel}}$, of the GPD-AR(1) use to describe the conditional distribution of the cluster maxima given that it exceeds the threshold use for declustering. We compare the parameter estimates of the GPD-AR(1) model fitted to the cluster maxima with those obtained by fitting GPD_{gumbel} to the Gumbel-AR(1) before declustering.

The MSE of the parameter estimates of the shape parameter, ξ, fitted to the cluster maxima and the original series is shown in Table 9. As the sample sizes increases, ξ performs better for both the GPD-AR(1) and Gumbel-AR(1), however, for any given sample size, the MSE of the Gumbel-AR(1) is smaller than that of the corresponding GPD-AR(1).

Table 9: MSE for Estimate of ξ Fitted to GPD-AR(1) and Gumbel-AR(1) for n and ϕ

		ϕ								
	n	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
GPD-AR(1)	1000	0.0829	0.1526	0.0905	0.0854	0.0849	0.0642	0.0639	0.1555	0.0833
	3000	0.0668	0.0405	0.0547	0.0451	0.0737	0.0593	0.0576	0.0349	0.0556
	5000	0.0378	0.0187	0.0041	0.0336	0.0431	0.0492	0.0228	0.0344	0.0008
Gumbel-AR(1)	1000	0.0536	0.0627	0.0417	0.0492	0.0662	0.0492	0.0524	0.0442	0.0352
	3000	0.0408	0.0301	0.0327	0.0257	0.0427	0.0447	0.0437	0.0301	0.0124
	5000	0.0293	0.0138	0.0010	0.0169	0.0134	0.0378	0.0137	0.0220	0.0002

Table 10: Estimate of $\sigma_{GPD_{gumbel}}$ fitted to GPD-AR(1) and Gumbel-AR(1) for n and ϕ

		ϕ								
	n	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
GPD-AR(1)	1000	0.0079	0.0068	0.0069	0.0079	0.0081	0.0102	0.0094	0.0114	0.0124
	3000	0.0092	0.0082	0.0089	0.0084	0.0083	0.0116	0.0124	0.0124	0.0136
	5000	0.0096	0.0088	0.0122	0.0106	0.0095	0.0130	0.0129	0.0183	0.0190
Gumbel-AR(1)	1000	0.0074	0.0063	0.0063	0.0069	0.0073	0.0091	0.0076	0.0097	0.0119
	3000	0.0081	0.0076	0.0079	0.0073	0.0078	0.0095	0.0108	0.0110	0.0129
	5000	0.0082	0.0087	0.0099	0.0091	0.0079	0.0113	0.0138	0.0137	0.0152

The estimate of $\sigma_{GPD_{gumbel}}$ fitted to the cluster maxima and the original series is shown in Table 10. After declustering the data and the Gumbel-generalised Pareto distribution is fitted to the independent cluster maxima and to the Gumbel-AR(1), the result show that as the sample size increases, the parameter estimate, $\sigma_{GPD_{gumbel}}$, fitted to both the cluster maxima and to

the original series perform better. A comparison of the estimate of $\sigma_{GPD_{gumbel}}$ between the corresponding GPD$_{gumbel}$-AR(1) and the Gumbel-AR(1) indicate that the estimate, $\sigma_{GPD_{gumbel}}$, of the GPD-AR(1) are larger than the corresponding $\sigma_{GPD_{gumbel}}$ estimate of the Gumbel-AR(1).

Real Data Problem

We now use a real data example to identify and classify the temporal structure of extreme values in the Gumbel AR(1) model. Furthermore, we show how clusters as a result of dependence influence the parameter estimates of the autoregressive process and the Gumbel-Generalised Pareto distribution.
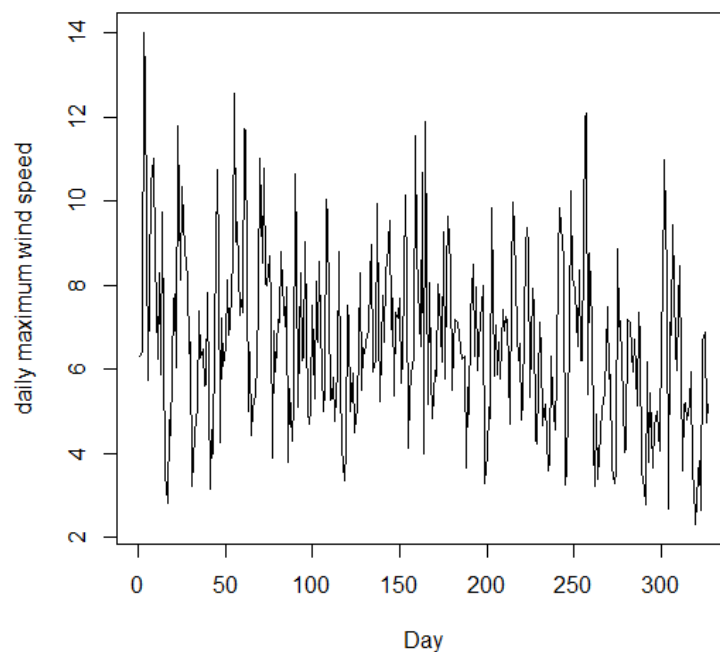
Figure 1: Scatter Plot of Daily Maximum Wind Speeds

In this example, we use an environmental dataset to show the link between light tailed maxima stemming from an extreme value distribution with Gumbel distributed maxima. The dataset consist of 327 daily maximum wind speeds (m/s) recorded in a specific location in Canada in 2012.

The catastrophe precipitated by Hurricanes Katrina and Rita in August and September 2005, the extraordinary hurricane season of 2004 in which five hurricanes made landfall in Florida, and the May 1999 outbreak of damaging tornadoes in Oklahoma highlight the significant and growing risks to society of extreme wind speed hazards. Extreme wind storms can cause severe damage leading to huge economic losses. Consequently, understanding the behaviour of the parameter estimates of maximum wind speeds is imperative in order to make predictions of wind storm and proffer ways of mitigating its consequence.

A scatter plot of the daily maximum wind speeds is displayed in Figure 1. It seems to show a linear decrease of maximum wind speeds. The autocorrelation for the wind speed data is
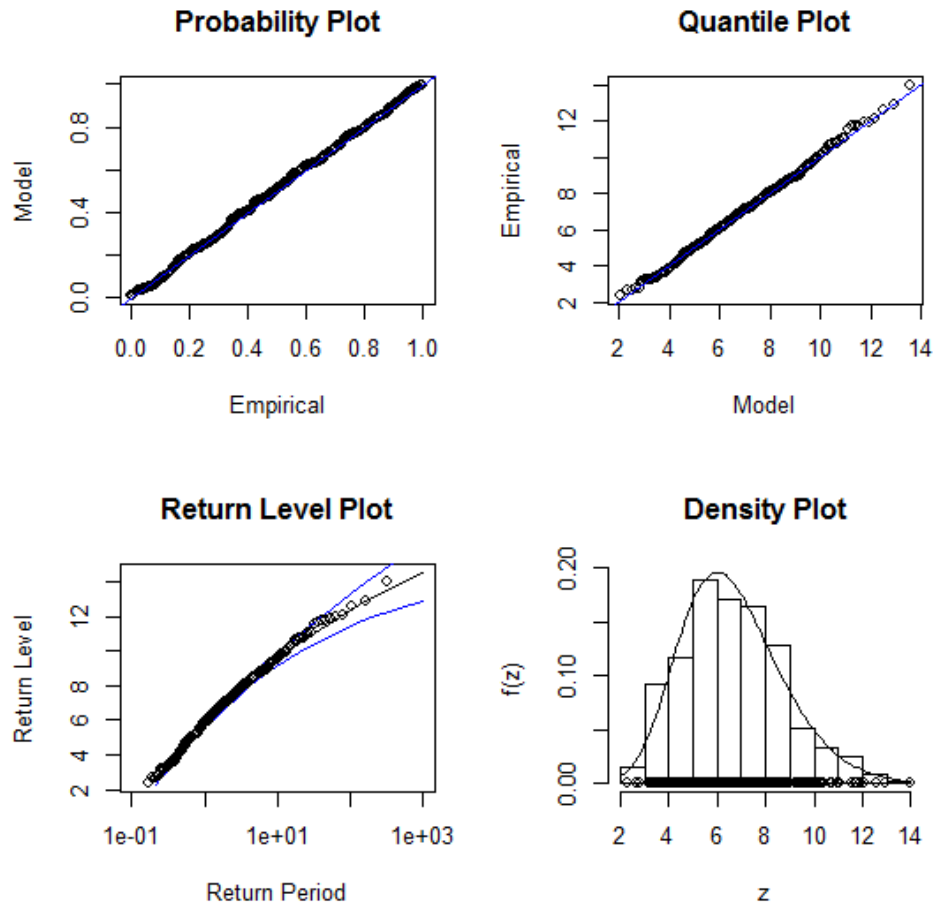
Figure 2: Goodness of Fit Test for Wind Speed Dataset to the Gumbel distribution

0.388 showing a positive dependence. The Augmented Dickey Fuller gives a probability value of 0.01 signifying that the wind speeds dataset is stationary.

Diagnostic plots for assessing the goodness of fit of the Gumbel model fitted to the data are shown in Figures 2. The plots show a linear relationship, which indicates that a Gumbel fit seems to be reasonable. To get more information on the dependence structure and fit the observed data to a time series model, the sample autocorrelation function and sample partial autocorrelation function are plotted in Figure 3 and 4 respectively. We see from Figure 3 and 4 that the sample autocorrelation function decays exponentially and the sample partial autocorrelation function mostly cuts of at lag 1 suggesting an autoregressive model of order 1 for the maximum daily wind speeds dataset.

The estimator of ϕ is $\hat{\phi} = 0.3872$ which further confirms that the dependence does exist among daily maxima wind speeds in the observed datasets. The maximization of the Gumbel log- likelihood for this dataset gives the estimates of the location and scale parameters as
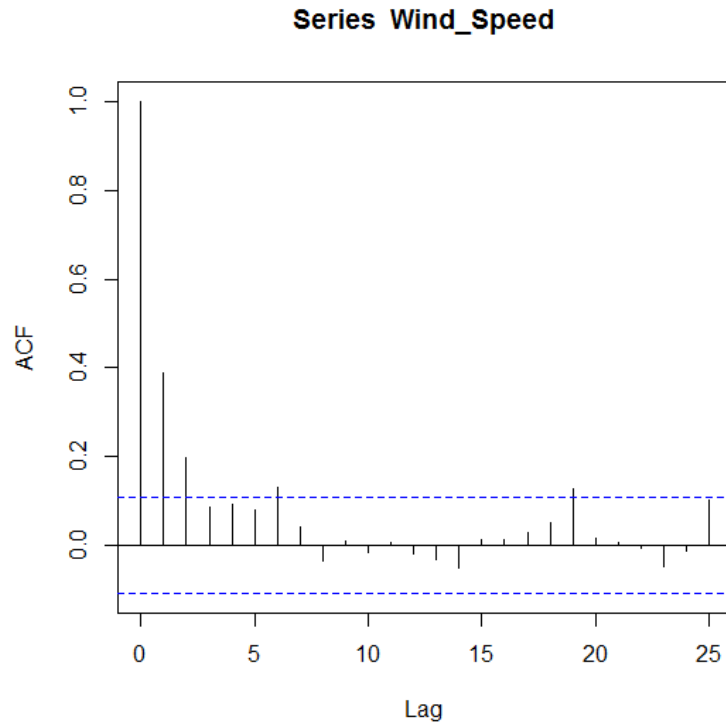
Series Wind_Speed

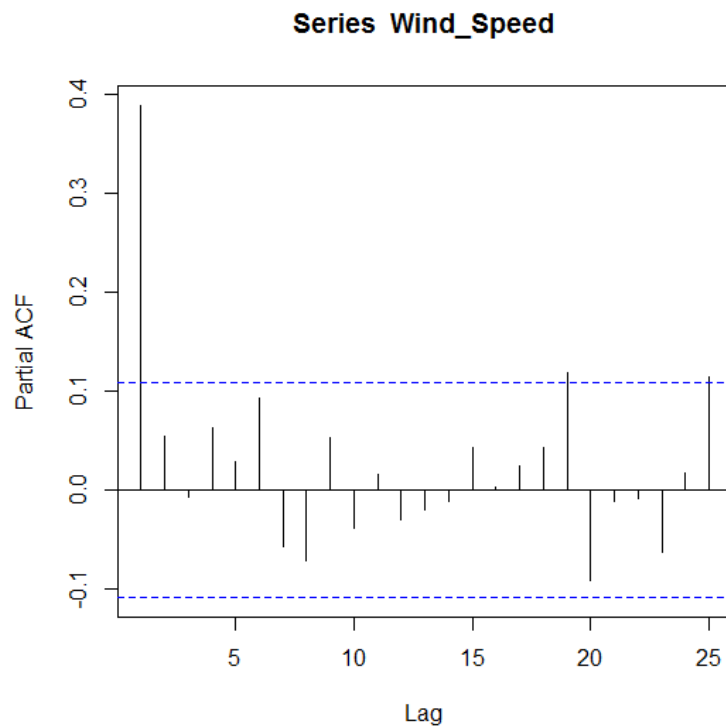Figure 3: Sample ACF of Daily Maximum wind Speed Series

Series Wind_Speed

Figure 4: Sample PACF of Daily Maximum Wind Speed Series

$(\hat{\mu}; \hat{\sigma}) = (5.6423; 1.8319)$. The estimates of the location and scale parameter for the residuals of the wind speeds data are $(\hat{\mu}; \hat{\sigma}) = (-0.9379; 1.7591)$.
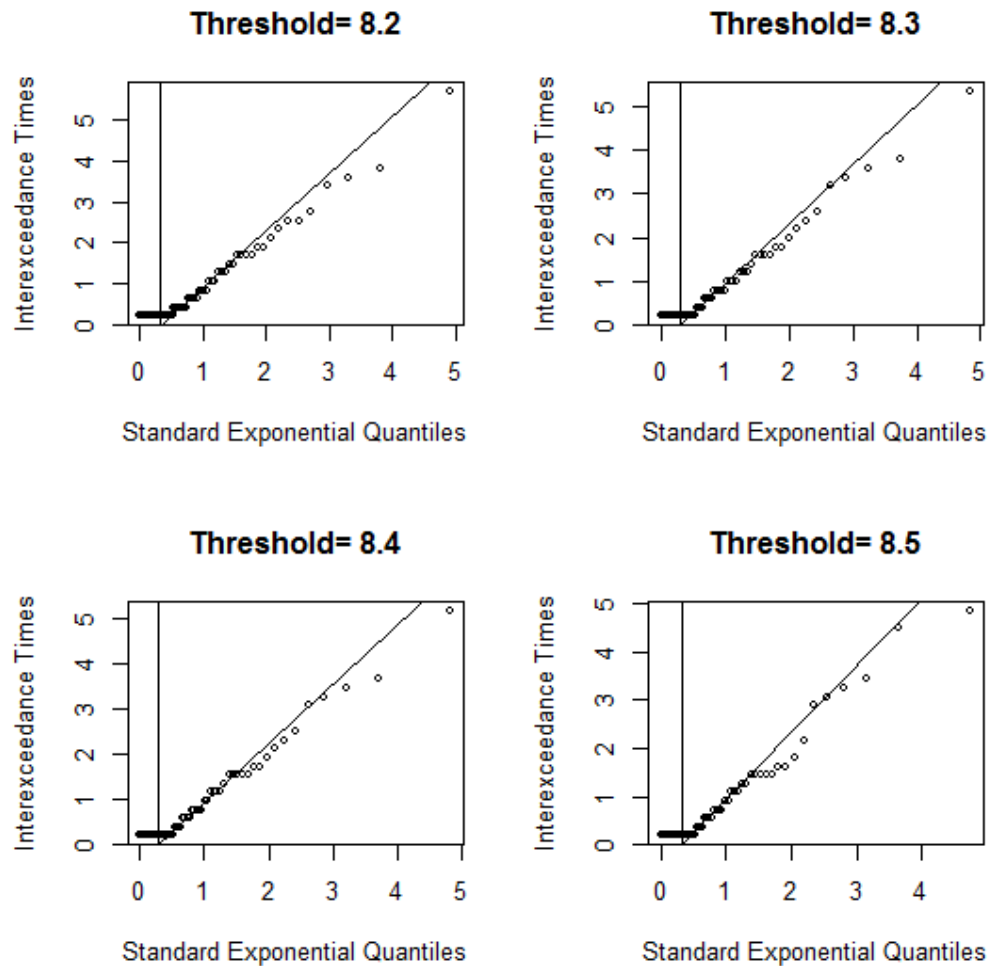
Figure 5: Quantile-Quantile Plot for Wind Speed Data Set

Since we are interested in finding the clustering tendencies of the extreme values, the threshold used for declustering the dataset is chosen such that the extremal index, the shape and scale parameters of the generalised Pareto distribution are invariant to a changes in threshold use for declustering. This leads to a threshold of 8.4. The Q-Q plot in Figure 5 is used to check the goodness of fit of our model for cluster occurrence. By these plots, we can be reassured that the chosen threshold of 8.4 gives a good fit of the model to the wind speeds data. The estimate of the extremal index for this threshold is 0.7578 and the average cluster size is $1/0.7578 = 1.3196$. This implies that wind tends to be heavy on consecutive days but very strong wind tend not to last longer than two consecutive days.

We can now decluster the wind speeds dataset. So of the observed dataset which is of length 327, there are 63 independent exceedances of the threshold 8.4. We now proceed to estimate

the parameters of the Gumbel Generalised Pareto Distribution used to describe the conditional distribution of a cluster maxima given that it exceeds the threshold used to decluster the dataset is given as $(\hat{\sigma}_{GPD_{gumbel}}; \hat{\xi}) = (2.3501; -0.3542)$. The parameter estimates of GPD_{gumbel}-AR(1) fitted to the observed wind speeds dataset is given as $(\hat{\sigma}_{GPD_{gumbel}}; \hat{\xi}) = (1.8107; -0.2207)$. The key difference between the fits is that the Generalised Pareto distribution is fitted to the cluster maxima after declustering the dataset in the former, while the Generalised Pareto distribution is fitted to the original dataset without declustering in the latter.

6   Conclusions

The study used Gumbel-AR(1) to characterize short-term dependence among maxima stemming from light-tailed distributions over a range of sample sizes with varying degrees of dependence. A simulation study done examine and compare the performance of the estimates of Gumbel-AR(1) and GPD_{gumbel}-AR(1) for the different sample sizes with varying degrees of dependence.

The study found that estimating the parameters of the Gumbel autoregressive process of order one, and ignoring the effect of dependence, the location parameter is overestimated. The estimate of the location parameter of the autoregressive process using the residuals gives a better estimate. The estimate of the scale parameter of the original series gives a marginally better estimate than the residuals of Gumbel-AR(1). The degree of clustering increases as dependence is enhanced in the autoregressive process. The number of independent observations increases as sample size is increased for at a constant dependence level. When the effect of extreme value innovation is ignored, we risk underestimating the scale parameter of the Gumbel generalised Pareto distribution. Comparing the estimate of the shape parameter of the Gumbel-Generalised Pareto distribution before and after declustering shows that the original series gives a better estimate of the shape parameter for the autoregressive processes.

References

[1] Coles, S. *An Introduction to Statistical Modeling of Extreme Values.* vol. 208. London, Britain: Springer. 2001.

[2] Nakajima, J., Kunihama, T., Omori, Y. and Frühwirth-Schnatter, S. Generalized extreme value distribution with time-dependence using the ar and ma models in state space form. *Computational Statistics & Data Analysis.* 2012. 56(11): 3241–3259.

[3] Prado, R. and Lopes, H. F. Sequential parameter learning and filtering in structured autoregressive state-space models. *Statistics and Computing.* 2013. 23(1): 43–57.

[4] Smith, R. L. Statistics of extremes, with applications in environment, insurance and finance, chap 1. *Statistical analysis of extreme values: with applications to insurance, finance, hydrology, and other fields. Birkhuser, Basel.* 2003.

[5] Toulemonde, G., Guillou, A., Naveau, P., Vrac, M. and Chevallier, F. Autoregressive models for maxima and their applications to CH4 and N2O. *Environmetrics.* 2010. 21(2): 189–207.

[6] Reiss, R.-D. and Thomas, M. *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and other Fields*. Basel: Birkhäuser. 2007.

[7] Young, P. Jackknife and bootstrap resampling methods in statistical analysis to correct for bias. *Statistical Science*. 1996. 11: 189–228.

[8] Efron, B. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*. 1979. 7(1): 1–26.

[9] Singh, K. On the asymptotic accuracy of efron's bootstrap. *The Annals of Statistics*. 1981. 9(6): 1187–1195.

[10] Kunsch, H. R. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*. 1989. 17(3): 1217–1241.

[11] Politis, D. N. and Romano, J. P. The stationary bootstrap. *Journal of the American Statistical Association*. 1994. 89(428): 1303–1313.

[12] Ferro, C. A. and Segers, J. Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2003. 65(2): 545–556.