**Nonlinear Processes
in Geophysics**

Estimating return levels from maxima of non-stationary random sequences using the Generalized PWM method

P. Ribereau[1], A. Guillou[2], and P. Naveau[3]

[1]Université Montpellier 2, Equipe Proba-Stat, CC 051, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France
[2]Université de Strasbourg, IRMA; 7, rue René Descartes, 67084 Strasbourg Cedex, France
[3]Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, Orme des Merisiers, 91191 Gif-sur-Yvette, France

Abstract. Since the pioneering work of Landwehr et al. (1979), Hosking et al. (1985) and their collaborators, the Probability Weighted Moments (PWM) method has been very popular, simple and efficient to estimate the parameters of the Generalized Extreme Value (GEV) distribution when modeling the distribution of maxima (e.g., annual maxima of precipitations) in the Identically and Independently Distributed (IID) context. When the IID assumption is not satisfied, a flexible alternative, the Maximum Likelihood Estimation (MLE) approach offers an elegant way to handle non-stationarities by letting the GEV parameters to be time dependent. Despite its qualities, the MLE applied to the GEV distribution does not always provide accurate return level estimates, especially for small sample sizes or heavy tails. These drawbacks are particularly true in some non-stationary situations. To reduce these negative effects, we propose to extend the PWM method to a more general framework that enables us to model temporal covariates and provide accurate GEV-based return levels. Theoretical properties of our estimators are discussed. Small and moderate sample sizes simulations in a non-stationary context are analyzed and two brief applications to annual maxima of CO_2 and seasonal maxima of cumulated daily precipitations are presented.

1 Introduction

Extreme value theory provides a solid mathematical foundation (e.g. Embrechts et al., 1997; Beirlant et al., 2004; de Haan and Ferreira, 2006) for studying maxima in fields like hydrology or climatology (e.g. Katz et al., 2002). In the Identically and Independently Distributed (IID) setup, this theory states that maxima should follow the Generalized Extreme Value (GEV) distribution whenever re-normalized maxima of a random sample converges to a non-degenerate random variable (e.g. Coles, 2001). From a statistical point of view, this means that the cumulative probability distribution function of maxima from IID samples is very likely to be correctly fitted by the following GEV distribution

$$G(x; \mu, \sigma, \gamma) = \exp\left(-\left\{1 + \gamma \frac{x - \mu}{\sigma}\right\}^{-1/\gamma}\right) \qquad (1)$$

where $\sigma > 0$, $\gamma \neq 0$ and $\mu \in \mathbb{R}$ are called the GEV scale, shape and location parameters, respectively, and with the constraint $1 + \gamma \frac{x-\mu}{\sigma} > 0$. If $\gamma \to 0$, then Eq. (1) corresponds to the Gumbel case and is equal to $\exp\left(-\exp\{-\frac{x-\mu}{\sigma}\}\right)$ with $x \in \mathbb{R}$.

To estimate the three parameters of a GEV distribution, several methods have been developed, studied and compared during the last twenty years. In 1979, Greenwood et al. (1979) and Landwehr et al. (1979) introduced the so-called Probability Weighted Moments (PWM). This method-of-moments approach (see Appendix A) has been very popular in hydrology (Hosking et al., 1985) and climatology because of its conceptual simplicity, its easy implementation and its good performance for most distributions encountered in geosciences. In 1985, Smith studied and implemented the Maximum Likelihood (ML) method for the GEV density (see Appendix C). According to Hosking et al. (1985), the PWM approach is superior to the MLE for small GEV distributed samples. Coles and Dixon (1999) argued that the PWM method assumes a priori on the shape parameter which is equivalent to assume a finite mean for the studied distribution. To integrate this condition in the ML approach, these authors proposed a penalized MLE scheme with the constraint $\gamma < 1$. If this condition is satisfied, then the ML approach is as competitive as the PWM one, even for small samples. Still, the debate over the advantages and drawbacks of both estimation methods is not closed. For example, the classical and penalized ML approaches impose a restriction on the lower values of γ, i.e. we need $\gamma > -0.5$ to have

Table 1. Bias of the estimated return levels obtained by the GPWM and ML method (for a return period $t=10\times n$).

γ	Method	$n=15$	$n=25$	$n=50$	$n=100$
1	z_t^*	148.4	248.4	502.4	1002.4
1	GPWM	−98.9	−163.8	29.4	14.3
1	ML	190.6	402.5	628.9	181.4
0.6	z_t^*	8.6	44.0	71.6	107.4
0.6	GPWM	−18.6	−22.4	−3.8	1.0
0.6	ML	42.5	122.5	77.2	36.8
0.4	z_t^*	6.7	20.2	31.5	41.1
0.4	GPWM	−13.3	−8.0	−2.6	−0.7
0.4	ML	19.1	31.1	11.2	7.8
0.2	z_t^*	5.5	10.0	16.3	18.9
0.2	GPWM	−2.9	−2.9	−1.3	−0.81
0.2	ML	11.4	7.6	2.3	1.5
0	z_t^*	5.0	5.5	10.2	10.9
0	GPWM	−1.6	−1.0	−0.67	−0.46
0	ML	3.7	1.0	0.28	0.15
−0.2	z_t^*	3.1	3.3	7.5	7.7
−0.2	GPWM	−0.4	−0.37	−0.33	−0.24
−0.2	ML	0.05	0.01	−0.09	−0.07
−0.4	z_t^*	2.1	2.2	6.2	6.3
−0.4	GPWM	−0.26	−0.10	−0.14	−0.09
−0.4	ML	−0.09	−0.14	−0.14	−0.07
−0.6	z_t^*	3.1	3.3	5.6	5.6
−0.6	GPWM	−0.15	0.01	−0.04	−0.03
−0.6	ML	−0.11	−0.15	−0.09	−0.05
−1	z_t^*	0.99	0.99	4.99	4.99
−1	GPWM	0.15	0.08	0.03	0.00
−1	ML	0.42	−0.05	−0.01	−0.00

regularity of the ML based estimators. Although it is rare to work with bounded upper tails, they can be encountered in geophysics. For example, atmospheric scientists can be interested in relative humidity maxima, a bounded random variable. One problem with the PWM method is the range of validity ($\gamma<1/2$) to derive the asymptotic properties of the PWM estimators. This constraint may be too restrictive for some applications in hydrology and climatology. Recently, Diebolt et al. (2008) introduced the concept of the General-

ized Probability Weighted Moments (GPWM) for the GEV. It broadens the domain of validity of the PWM approach, allowing heavier tails to be fitted. Despite this advantage, the ML method has kept a strong advantage over a PWM approach: its inherent flexibility in a non-stationary context. When studying climatological and hydrological data, it is not always possible to assume that the distribution of the maxima remains unchanged in time. For example, trends can be present in extreme values of different hydroclimatological series (e.g., Kharin et al., 2007; IPCC Report, 2007). The MLE can easily integrate temporal covariates within the GEV parameters (e.g., Katz et al., 2002; Coles, 2001; El Adlouni et al., 2007) and conceptually, the MLE procedure remains the same if the GEV parameters vary in time (see El Adlouni and Ouarda (2008) for a comparison study of different methods for non-stationary GEV models). In practice, numerical problems quickly arise and estimated ML return levels can be misleading in some non-stationary situations (see Tables 1 and 2), especially for strong heavy tails ($\gamma\geq0.4$). With these limitations in mind, our aim is to propose and to study a novel GPWM procedure that can handle temporal covariates. Our main motivation is to keep the interesting GPWM properties identified in the IID case while adding the needed flexibility to handle non-stationarities that are often present in real case studies. This will provide a valuable alternative to the MLE for non-stationary GEV distributed data.

Before closing this introduction, we would like to emphasize that, beyond the three aforementioned estimation methods (MLE, PWM and GPWM), there exists other variants and extensions. For example, Hosking (1990) proposed and studied the L-moments, Zhang (2007) recently derived a new and interesting method-of-moments, and Bayesian estimation has also generated a lot of interest in extreme value analysis (e.g. Lye et al., 1993; Coles, 2001; Cooley et al., 2007). But we will neither compare nor discuss these alternatives with respect to our proposed method for the following reasons. The objective of this work is not to write a review paper that compares all existing estimation methods. Instead our aim is to extend the applicability of a well-known approach (PWM) by working with a larger class of estimators (GPWM) and by proposing an algorithm that encompasses temporal dependences. We focus on comparing our approach with the classical MLE because the later may be the most popular in statistical climatology and hydrology for non-stationary time series analysis.

2   A GEV regression model

One of the most simples, most frequently used and most studied models in statistics is the classical regression

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \tag{2}$$

where $\mathbf{Y}=(Y_1, \ldots, Y_n)^t$ represents an observational vector of length n, $\mathbf{X}$ is a $n \times p$ known matrix of explanatory variables and β a vector of unknown parameters of length p that characterizes the relationship between observations and explanatory variables. For example, in a time series context where a cycle of length T can be present, a model like $Y_i = \beta_0 + \beta_1 \cos(2\pi i / T) + \epsilon_i$ corresponds to a matrix

$$\mathbf{X} = \begin{pmatrix} 1 & \cos(\frac{2\pi}{T}) \\ \vdots & \vdots \\ 1 & \cos(\frac{2\pi n}{T}) \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}. \quad (3)$$

A regression like Eq. (2) can also model linear and even polynomial trends and consequently, it can handle various types of non-stationarities in time. Classically, the vector ϵ in Eq. (2) is assumed to be a zero-mean Gaussian vector, but this hypothesis is not reasonable whenever the observations can be considered as maxima. As already mentioned in the Introduction, Extreme Value Theory tells us that a more adequate fit for maxima should be the GEV distribution defined by Eq. (1). For this reason, we assume in this paper that the vector ϵ in Eq. (2) consists of IID random variables from a GEV distribution defined by Eq. (1) and with parameters $(0, \sigma, \gamma)$. Note that μ is set to zero here because β_0 in the vector β usually plays the role of the location parameter. Overall, this is equivalent to state that the sequence of maximum Y_i represents a sequence of independent GEV distributed random variables with the same scale parameter σ, the same shape parameter γ but with a varying location parameter μ_i that depends on the covariate $\mathbf{X}$, i.e. $\mu_i = (\mathbf{X}\beta)_i$.

Classically, most estimation methods for estimating β in a regression model assume that the "noise" ϵ is zero-mean. In our case, we have imposed that ϵ_i follows a GEV distribution with $\mu=0$. But this does not imply that the mean ϵ_i is null because the GEV density is not always symmetric around μ. To be in accordance with the zero-mean constraint, we just have to re-parametrize our model

$$\mathbf{Y} = \mathbf{X}\beta^* + \epsilon^* \quad (4)$$

where

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{np-1} \end{pmatrix}, \quad \epsilon^* = \begin{pmatrix} \epsilon_1 - \mathbf{E}(\epsilon_1) \\ \vdots \\ \epsilon_n - \mathbf{E}(\epsilon_1) \end{pmatrix}, \quad (5)$$

$$\beta^* = \left(\beta_0 + \mathbf{E}(\epsilon_1), \beta_1, \ldots, \beta_{p-1}\right)^t,$$

and $\mathbf{E}(\epsilon_1)$ corresponds to the mean value of ϵ_1.

To estimate return levels, we propose to implement the following three-steps algorithm:

Step a: Implement an existing regression method on Eq. (4) to find regression estimates that we call $(\widehat{\beta}_1, \ldots, \widehat{\beta}_{p-1})$.

Table 2. Standard deviation of the estimated return levels obtained by the GPWM and ML method (for a return period $t=10\times n$).

γ	Method	$n=15$	$n=25$	$n=50$	$n=100$
1	GPWM	238.1	374.5	1163.6	2304.0
1	ML	10 616.4	28 206.3	4670.08	2671.2
0.6	GPWM	54.76	30.4	69.8	83.7
0.6	ML	846.6	627.3	268.0	108.4
0.4	GPWM	12.0	11.7	20.1	23.1
0.4	ML	126.0	226.13	37.8	27.1
0.2	GPWM	6.2	5.0	7.0	6.9
0.2	ML	11.5	38.9	9.1	6.7
0	GPWM	2.2	2.4	2.9	2.5
0	ML	19.5	14.9	2.5	1.8
−0.2	GPWM	1.51	1.4	1.4	1.1
−0.2	ML	4.3	2.8	0.81	0.52
−0.4	GPWM	0.42	0.89	0.85	0.62
−0.4	ML	2.5	2.2	0.36	0.23
−0.6	GPWM	0.33	0.64	0.56	0.41
−0.6	ML	0.70	0.34	0.18	0.11
−1	GPWM	0.64	0.44	0.28	0.18
−1	ML	0.48	0.23	0.10	0.08

Step b: Apply the GPWM approach described in Appendix B to estimate the GEV parameters of the "pseudo-residuals"

$$\widehat{\varepsilon}_i = Y_i - \sum_{j=1}^{p-1} \widehat{\beta}_j x_{ij}, \text{ for } i=1, \ldots, n. \quad (6)$$

Step c: From the GEV estimates obtained from Step b, compute the desired return levels.

Although each of these three steps appears to be simple, they deserve careful examinations. In order to apply Step b, the pseudo-residuals $(\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_n)$ should be IID and GEV distributed. This is not true because β_j has to be estimated in Step a (it would be true if we knew β_j). This issue is less relevant if the sample size n gets larger. For example, if the classical least squares regression[1] is used for Step a, our estimation of β_j asymptotically becomes better and our IID GEV assumption for our residuals truer. More precisely, a

---

[1] $\widehat{\beta}^*=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ is obtained by minimizing the sum of the squared errors $(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$ whenever $\mathbf{X}'\mathbf{X}$ is invertible.

general result (Azais and Bardet, 2005), about least squares regression tells us that, under the assumptions

$$\begin{cases} \lim_{n \to \infty} n^{-a} \mathbf{X}'\mathbf{X} \text{ equals a positive definite matrix for } a > 0 \\ \lim_{n \to \infty} \max_{1 \le i \le n} \left(\mathbf{X} \left(\mathbf{X}'\mathbf{X} \right)^{-1} \mathbf{X}' \right)_{(i,i)} = 0 \qquad (7) \\ \text{the variance of all } \epsilon_i \text{ is finite (i.e. } \gamma < \frac{1}{2}) \end{cases}$$

the pseudo-residuals $\widehat{\varepsilon}_i$ can be asymptotically viewed as independent random variables distributed according to a GEV$(\beta_0, \sigma, \gamma)$ distribution. This result is general because it does not impose a specific type of distribution for the ϵ in Eq. (2) but only a finite variance condition. Still extreme values like maxima do not necessary obey the condition $\gamma < \frac{1}{2}$, it may be more appropriate to implement a robust/resistant regression in Step a.

Such a remark was confirmed by our simulation study. In Sect. 3 we take advantage of the Least Trimmed Squares (LTS) regression (Venables and Ripley, 2002, p. 156–163). Basically, the influence of very high values that could mislead the estimation of the β_j's is trimmed by minimizing an error/cost function that is only based on the core data. This LTS method usually gives accurate estimators even in the presence of large values but it needs a long computation time. This is not an important issue for maxima because of the rarity of the observations. Hence, for Step a, we advice to apply such a resistant/robust method for heavy tail distributions.

Concerning Step c, the return level z_t for the fixed time period t can be easily defined in the IID case. More precisely, the return z_t corresponds to the $1-1/t$ quantile, i.e. it is the level that, in average, is crossed one time during the time period t. For the GEV(μ, σ, γ) distribution, this means that

$$z_t = \mu + \frac{\sigma}{\gamma} \left[-1 + \left(-\ln\left(1 - \frac{1}{t} \right) \right)^{-\gamma} \right]. \qquad (8)$$

In a non-stationary context, no unique definition exists for z_t because extreme values recorded during the time $1, \ldots, n$ have a different distribution than the extremes occurring during any translated time $1+C, \ldots, n+C$ for any $C \neq 0$. With respect to our regression model Eq. (4), the IID return level defined by Eq. (5) can be adapted in the following way

$$z_t^* = (\mathbf{X}\beta)_t + \frac{\sigma}{\gamma} \left[-1 + \left(-\ln\left(1 - \frac{1}{t} \right) \right)^{-\gamma} \right]. \qquad (9)$$

This means the trend at time t plus an IID return level based on the residual of the regression.

To estimate beyond the range of observations, e.g. to compute the centennial return level with only 50 years of data, one needs to extrapolate the trends in Eq. (9) captured by $(\mathbf{X}\beta)_t = \mu_t$. As it is well know in statistics, it is very dangerous to extrapolate a linear regression outside the sample support. Hence, extra caution is required when interpreting z_t^* in real applications because it means that the trend will remain the same until time t.

3 Analysis of non-stationary maxima

3.1 Simulations

Simulations are performed for four sample sizes $n = 15, 25, 50$ and 100 and for nine shape parameters $\gamma = -1, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6$ and 1. Without loss of generality, we can set σ to one. As covariates, we fix

$$\mathbf{X} = \begin{pmatrix} 1 & \cos(\pi/2 \times 1) \\ 1 & \cos(\pi/2 \times 2) \\ \vdots & \vdots \\ 1 & \cos(\pi/2 \times n) \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}. \qquad (10)$$

For each combination of n and γ, $10\,000$ random samples are generated from a GEV$(\beta_0 + \beta_1 \times \cos(\pi/2 \times i), 1, \gamma)$ distribution.

To estimate all parameters and return levels with our approach the same procedure is followed. Step a is implemented by regressing with a LTS method. Then Step b is applied. Finally, Step c is used to estimate the return level z_t^* defined by Eq. (9) for $t = 10 \times n$.

To obtain the return level estimate from the ML approach, we directly obtain estimation of β_0, β_1, γ and σ by maximizing the log-likelihood function (see Appendix C). Abnormal ML estimates have been eliminated by removing all samples that provide ML estimates of γ greater than 2.5[2].

For both estimation methods, the bias between true and estimated return values is displayed in Table 1 and the standard deviation in Table 2. These tables clearly showed that, for this specific example, the MLE does not provide adequate estimates whenever the sample size is small ($n < 50$) and the shape parameter is large, especially in terms of standard deviation. In contrast, our GPWM regression method does a better job at handling small samples and large γ (except for $\gamma = 1$ where both methods fail), especially in terms of standard deviation. Other simulations have been done, in particular a comparison between the two methods in a linear dependence case. Since the conclusions are the same, we do not include them.

3.2 Annual maxima of CO$_2$ concentrations

We consider the annual concentrations of CO$_2$ at Amsterdam in the Indian Ocean from 1981 to 2001. These observations are available on the website of RAMCES (http://www.ipsl.jussieu.fr/services/Observations/fr/RAMCES.htm) where an extensive bibliography can be found on this subject, for example in Gros et al. (1999). Each of the 21 measures is assumed to follow a GEV$(\beta_0 + \beta_1 \times \text{year}, \sigma, \gamma)$ distribution. With the GPWM and ML approaches, we have estimated γ, σ, β_0 and β_1 and also the return level for a period of 200 years. The results are listed in Table 3. We observe that

---

[2]Without this thresholding, the ML standard deviations would have been even larger for γ's greater than 0.4.

the two methods lead to very similar results for all the parameters estimated. This corroborates our simulations analysis. For light tailed random variables such as CO_2 concentration maxima, the GEV parameters estimates are fairly similar for values of γ near zero, see Table 1.

3.3 Seasonal maximum of cumulated daily precipitations

We treat here seasonal maxima of cumulated daily precipitations recorded in the Orgeval basin (France) during 31 years. We suppose that each of the 124 measures $(Y_{i,j})_{i=1,\ldots,31;\, j=1,\ldots,4}$ follows a $\mathrm{GEV}\left(\beta_0+\beta_1\times\cos\left(\frac{\pi}{2}\times j\right)+\beta_2\times i,\, \sigma,\, \gamma\right)$ distribution. The indice i corresponds to the year while the indice j corresponds to the season. The parameter β_2 indicates the trend of the series. Table 4 provides the estimates of γ, σ, β_0, β_1 and β_2 and the 100-year return level obtained from the GPWM and ML approaches.

In contrast to our previous example summarized by Table 3, the MLE and GPWM procedures give very different estimates of the 100-year return level estimates in Table 4. To better understand this difference, we re-estimate the GEV parameters but with only the first 20 years. Table 5 summarizes our findings. As observed in Tables 1 and 2, the properties of a GPWM estimated shape parameter around 0–0.2 does not vary greatly when the sample size changes. This is not the case for the ML approach where the estimates of γ can double by changing the sample size, see Table 5. This also impacts the estimates of the return levels.

4 Conclusions

In this paper, we propose an extension of the PWM method which can be viewed as an alternative to the MLE method that enable us to model temporal covariates and provide accurate return levels. We illustrate our approach by a simulation study and by applying our approach to two time series of maxima. Compared to the ML method, the GPWM method performs better and is computationally easy, but we cannot obtain confidence intervals and, up to now, we can only have a non-stationary location parameter, whereas with the MLE method the three parameters of the GEV distribution can be non-stationary. Concerning the estimation of return levels in such non-stationary cases, we would like to conclude with a word of caution. Extrapolating the trend beyond the range of observations is always a delicate and sometimes dangerous operation (since we assume that the trend will remain the same in the future), especially when dealing with extremes. Hence high return levels in a non-stationary context must be interpreted with extreme care.

Table 3. Annual maxima of CO_2 concentrations analysis with a $\mathrm{GEV}(\beta_0+\beta_1\times\mathrm{year},\, \sigma,\, \gamma)$.

	GPWM	ML
γ	-0.1102	0.044
σ	0.52	0.52
β_0	-2528.96	-2528.94
β_1	1.44	1.44
z_t^*	441.89	439.53

Table 4. Seasonal maxima of cumulated daily precipitations analysis with a $\mathrm{GEV}\left(\beta_0+\beta_1\times\cos\left(\frac{\pi}{2}\times j\right)+\beta_2\times i,\, \sigma,\, \gamma\right)$ distribution fitted to 31 years of data.

	GPWM	ML
γ	0.068	0.152
σ	7.44	7.51
β_0	16.42	17.60
β_1	-3.49	-2.35
β_2	0.113	0.038
z_t^*	86.35	97.51

Table 5. Seasonal maxima of cumulated daily precipitations analysis with a $\mathrm{GEV}\left(\beta_0+\beta_1\times\cos\left(\frac{\pi}{2}\times j\right)+\beta_2\times i,\, \sigma,\, \gamma\right)$ distribution but only with the first 20 years of data.

	GPWM	ML
γ	0.072	0.271
σ	8.07	7.58
β_0	18.18	17.82
β_1	-4.61	-1.59
β_2	0.085	0.042
z_t^*	91.74	134.60

Appendix A

Probability Weighted Moments (PWM)

The PWMs of a random variable Z with distribution function F are the quantities (see Greenwood et al., 1979)

$$M_{p,r,s} = \mathbb{E}\Big[Z^p (F(Z))^r (1 - F(Z))^s \Big]$$

where p, r and s are real numbers. When the distribution F equals to the GEV distribution defined by Eq. (1), a subclass of PWM ($p=1$, $r=0, 1, 2, ...$ and $s=0$) can be explicitly obtained (see Hosking et al., 1985)

$$M_{1,r,0} = \frac{1}{r+1}\left\{\mu - \frac{\sigma}{\gamma}\Big[1 - (r+1)^\gamma \Gamma(1-\gamma)\Big]\right\}, \quad \text{(A1)}$$

for $\gamma < 1$ and $\gamma \neq 0$. This provides a system of three equations with three unknown parameters (μ, σ, γ)

$$
\begin{cases}
M_{1,0,0} = \mu - \dfrac{\sigma}{\gamma}\Big(1 - \Gamma(1-\gamma)\Big) \\[2mm]
2M_{1,1,0} - M_{1,0,0} = \dfrac{\sigma}{\gamma}\Gamma(1-\gamma)(2^\gamma - 1) \\[2mm]
\dfrac{3M_{1,2,0} - M_{1,0,0}}{2M_{1,1,0} - M_{1,0,0}} = \dfrac{3^\gamma - 1}{2^\gamma - 1}
\end{cases}
\quad \text{(A2)}
$$

In the IID case, the PWM $M_{1,r,0}$ can be estimated by the unbiased estimator (see Landwehr et al., 1979)

$$\widehat{M}_{1,r,0} = \frac{1}{n}\sum_{j=1}^{n}\Big(\prod_{\ell=1}^{r}\frac{j-\ell}{n-\ell}\Big)Z_{j,n}$$

or by the asymptotically equivalent consistent estimator

$$\widetilde{M}_{1,r,0} = \frac{1}{n}\sum_{j=1}^{n} p_{j,n}^r Z_{j,n},$$

where $(Z_1, \ldots, Z_n)$ is an IID sample, $Z_{1,n} \leq \ldots \leq Z_{n,n}$ the ordered sample and $p_{j,n}$ a plotting position, i.e. a distribution-free estimate of $F(Z_{j,n})$. Hence, the PWM estimators $(\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma})$ of (μ, σ, γ) are the solutions of the system Eq. (A2). In particular, we have

$$\widehat{\mu} = \widehat{M}_{1,0,0} + \frac{\widehat{\sigma}}{\widehat{\gamma}}\Big(1 - \Gamma(1-\widehat{\gamma})\Big) \text{ and } \widehat{\sigma} = \frac{\widehat{\gamma}\Big[2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0}\Big]}{\Gamma(1-\widehat{\gamma})(2^{\widehat{\gamma}} - 1)}.$$

To obtain $\widehat{\gamma}$, the last equation of (A2) has been solved numerically. In the specific case where $-1/2 < \gamma < 1/2$, γ can be estimated by :

$$\widehat{\gamma} = -7.859c + 2.9554c^2 \text{ with } c = \frac{2M_{1,1,0} - M_{1,0,0}}{3M_{1,2,0} - M_{1,0,0}} - \frac{\log 2}{\log 3}.$$

Under the condition $\gamma < 0.5$, Hosking et al. (1985) established the asymptotic normality of the vector $\theta = (\mu, \sigma, \gamma)'$, i.e., as the sample size n increases, $\sqrt{n}(\widehat{\theta} - \theta)$ converges in distribution to a trivariate zero-mean Gaussian vector whose covariance structure is given in Hosking et al. (1985).

Appendix B

Generalized Probability Weighted Moments (GPWM)

Introduced and studied by Diebolt et al. (2008) for the GEV, the Generalized Probability Weighted Moments (GPWM) can be viewed as an extension of the PWMs. They are defined as

$$\nu_\omega = \int_{-\infty}^{\infty} x\omega(G(x))dG(x),$$

where G represents the GEV distribution defined by Eq. (1) and $\omega(.)$ is any suitable continuous function. One advantage of this definition is that it is easy to propose estimators of ν_ω. By rewriting the GPWMs as

$$\nu_\omega = \int_0^1 G^{-1}(u)\omega(u)du,$$

the following estimator can be easily computed for any given $\omega(.)$

$$\widehat{\nu}_{\omega,n} = \int_0^1 \mathbb{F}_n^{-1}(u)\omega(u)du \quad \text{(B1)}$$

where $\mathbb{F}_n^{-1}$ denotes the inverse of the classical empirical distribution function of a IID GEV distributed sample. To make the link between PWMs and GPWMs, one can choose $\omega(u) = u^r$. Then ν_ω corresponds to the classical PWM captured by Eq. (A1). But this choice imposes a strong restriction on the GEV shape parameter $\gamma < 1/2$. To handle heavier tails (i.e. larger γ), it is preferable to work with $\omega(u) = u^a(-\log u)^b$. For such a type of function, Diebolt et al. (2008) showed that

$$\nu_\omega = \frac{\sigma}{\gamma}\frac{\Gamma\Big(b - \gamma + 1\Big)}{(a+1)^{b-\gamma+1}} - \Big(\frac{\sigma}{\gamma} - \mu\Big)\frac{\Gamma\Big(b+1\Big)}{(a+1)^{b+1}}.$$

If b is chosen such that $\gamma < b+1$, then ν_ω exists. In this case, the asymptotic normality for $\widehat{\nu}_{\omega,n}$ holds when $\gamma < \frac{1}{2} + b$ (Diebolt et al., 2008) and confidence intervals can be provided for bounded, light and heavy tails. To simplify notations, ν_ω is now called $\nu_{a,b}$ when $\omega(u) = u^a(-\log u)^b$.

To estimate the three GEV parameters from $\nu_{a,b}$, we need to solve a system of three equations. In practice, we fix three pairs for (a, b): $(1, 1)$, $(1, 2)$ and $(2, 1)$. This simple choice provides an easy system to solve

$$\frac{\widehat{\gamma}}{1 - (\frac{3}{2})^{\widehat{\gamma}}} = \frac{2[\widehat{\nu}_{11} - \widehat{\nu}_{12}]}{\widehat{\nu}_{11} - \frac{9}{4}\widehat{\nu}_{21}},$$

$$\widehat{\sigma} = 2^{3-\widehat{\gamma}}\frac{\widehat{\nu}_{11} - \widehat{\nu}_{12}}{\Gamma(2-\widehat{\gamma})} \quad \text{and} \quad \widehat{\mu} = \frac{\widehat{\sigma}}{\widehat{\gamma}} - \frac{\widehat{\sigma}}{\widehat{\gamma}}2^{\widehat{\gamma}}\Gamma(2-\widehat{\gamma}) + 4\widehat{\nu}_{11}.$$

Compared to the PWM approach, the concept remains the same (method-of-moments) but the system of equations slightly differs (see Eq. A2) and provides a wider range of possible values for γ.

Appendix C

Maximum Likelihood Estimation (MLE)

Let $(Y_1, ..., Y_n)$ be a sequence of independent random variables such that Y_i follows a $\text{GEV}((\mathbf{X}\beta)_i, \sigma, \gamma)$ distribution. Let θ be the vector of unknown parameters to be estimated (in our case, $\theta = (\beta, \sigma, \gamma)$). In the case $\gamma \neq 0$, the log likelihood function is given by :

$$\log L(\beta, \sigma, \gamma) = -n \log \sigma$$
$$-\left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{n} \log\left(1 + \gamma \frac{Y_i - (\mathbf{X}\beta)_i}{\sigma}\right)$$
$$-\sum_{i=1}^{n} \left(1 + \gamma \frac{Y_i - (\mathbf{X}\beta)_i}{\sigma}\right)^{-\frac{1}{\gamma}} \quad \text{(C1)}$$

provided $1 + \gamma \frac{Y_i - (\mathbf{X}\beta)_i}{\sigma} > 0$ for all $i = 1, ..., n$. The Maximum Likelihood estimates $(\widehat{\beta}, \widehat{\sigma}, \widehat{\gamma})$ are obtained by maximizing Eq. (C1). In the case $\gamma > -\frac{1}{2}$, the usual properties of consistency, asymptotic efficiency and asymptotic normality hold. Since there is no explicit formula from the maximization, we obtain the estimates by numerically optimizing Eq. (C1), e.g. by using a Newton Raphson algorithm.

References

Azais, J. M. and Bardet, J. M.: Le modèle linéaire par l'exemple, régression, analyse de la variance et plans d'expériences illustrés avec R, SAS et Splus, Dunod, 2005.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J.: Statistics of Extremes: Theory and Applications, Wiley Series in Probability and Statistics, 2004.

Coles, S. G.: An introduction to statistical modeling of extreme values, Springer, 2001.

Coles, S. G. and Dixon, M. J.: Likelihood-based inference of extreme value models, Extremes, 2, 5–23, 1999.

Cooley, D., Nychka, D., and Naveau, P.: Bayesian Spatial Modeling of Extreme Precipitation Return Levels, J. Amer. Statist. Assoc., 102–479, 824–840, 2007.

de Haan, L. and Ferreira A.: Extreme Value Theory: An Introduction, Springer Series in Operations Research, 2006.

Diebolt, J., Guillou, A., Naveau, P., and Ribereau, P.: Improving probability-weighted moment methods for the generalized extreme value distribution, REVSTAT – Statistical Journal, 6–1, 33–50, 2008.

El Adlouni, S. and Ouarda, T. B. M. J.: Comparaison des méthodes d'estimation des paramètres du modèle GEV non-stationnaire, Revue des Sciences de l'Eau, 21(1), 35–50, 2008.

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R., and Bobée, B.: Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, Water Resour. Res., 43, W03410, doi:10.1029/2005WR004545, 2007.

Embrechts, P., Klüppelberg, C., and Mikosch, T.: Modelling Extremal Events for Insurance and Finance, vol. 33, Applications of Mathematics, Springer-Verlag, Berlin, 1997.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: Probability-weighted moments: definition and relation to parameters of several distributions expressable in inverse form, Water Resour. Res., 15, 1049–1054, 1979.

Gros, V., Bonsang, B., Martin, D., Novelli, P. C., and Kazan, V.: Short term carbon monoxide measurements at Amsterdam Island: estimations of biomass burning emissions rates, Chemosphere, 1, 163–172, doi:10.1016/S1465-9972(99)00009-4, 1999.

Hosking, J. R. M.: L-moments: analysis and estimation of distributions using linear combinations of order statistics, J. R. Stat. Soc., 52, 105–124, 1990.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F.: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, Technometrics, 27, 251–261, 1985.

IPCC Climate Change: Synthesis Report, http://www.ipcc.ch/ipccreports/, 2007.

Katz, R., Parlange, M., and Naveau, P.: Extremes in hydrology, Adv. Water Resour., 25, 1287–1304, doi:10.1016/S0309-1708(02)00056-8, 2002.

Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, J. Climate, 20, 1419–1444, 2007.

Landwehr, J., Matalas, N., and Wallis, J. R.: Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, Water Resour. Res., 15, 1055–1064, 1979.

Lye, L. M., Hapuarachchi, K. P., and Ryan., S.: Bayes estimation of the extreme-value reliability function, IEEE Trans. Reliab., 42, 641–644, doi:10.1109/24.273598, 1993.

Smith, R. L.: Maximum likelihood estimation in a class of nonregular cases, Biometrika, 72, 69–90, 1985.

Venables, W. N. and Ripley, B. D.: Modern Applied Statistics With S, Springer, 2002.

Zhang, J.: Likelihood Moment Estimation for the Generalized Pareto Distribution, Aust. NZ. J. Stat., 49, 69–77, doi:10.1111/j.1467-842X.2006.00464.x, 2007.