Autoregressive models for maxima and their applications to CH4 and N2O

BY GWLADYS TOULEMONDE

Institut de Mathématiques et de Modélisation de Montpellier

Université Montpellier II, CC051, 34095 Montpellier, France

gwladys.toulemonde@math.univ-montp2.fr

ARMELLE GUILLOU

Institut de Recherche Mathématique Avancée

Université de Strasbourg, 67084 Strasbourg Cedex, France

guillou@math.u-strasbg.fr

PHILIPPE NAVEAU

Laboratoire des Sciences du Climat et de l'Environnement

LSCE-IPSL-CNRS, 91191 Gif-sur-Yvette, France

philippe.naveau@lsce.ipsl.fr

MATHIEU VRAC

Laboratoire des Sciences du Climat et de l'Environnement

LSCE-IPSL-CNRS, 91191 Gif-sur-Yvette, France

mathieu.vrac@lsce.ipsl.fr

AND FREDERIC CHEVALLIER

Laboratoire des Sciences du Climat et de l'Environnement

LSCE-IPSL-CNRS, 91191 Gif-sur-Yvette, France

frederic.chevallier@lsce.ipsl.fr

SUMMARY

Recordings of daily, weekly or yearly maxima in environmental time series are classically fitted by the Generalized Extreme Value (GEV) distribution that originates from the well established Extreme Value Theory (EVT). One special case of such GEV distribution is the Gumbel family which corresponds to the modeling of maxima stemming from light-tailed distributions. To capture temporal dependencies, linear autoregressive (AR) processes offer a simple and elegant framework. Our objective is to extend linear AR models in such a way that they handle Gumbel distributed maxima. To reach this goal, we take advantage of the stability of Gumbel random variables when added to the logarithm of a positive α-stable random variable. This allows us to propose a linear Gumbel distributed AR model whose main theoretical properties are derived. For the atmospheric scientist, this link between linear AR processes and EVT widens the statistical treatment of extreme environmental recordings in which temporal dependencies are present. For example, our model is fitted to daily and weekly maxima of methane (CH4) and daily maxima of nitrous oxide (N2O) measured in Gif-sur-Yvette (France). Simulation results are also presented in order to assess the quality of our parameter estimations for finite samples.

Key words: Extreme Value Theory, Dependence, Gumbel Distribution, Autoregressive Model, Atmospheric Chemistry.

1 Autoregressive models and environmental sciences

The tight connection between environmental sciences and statistics can be exemplified by the key figure of Sir Gilbert Walker (1868-1958) whose name has been associated to both climatological and statistical concepts (Katz, 2002). For example, the Walker circulation characterizes a zonal atmospheric circulation at the equator and the Yule-Walker equations describe correlation relationships for autoregressive (AR) processes (e.g. Brockwell and Davis, 1987). These equations have been widely used in time series analysis, especially in climatology. Two reasons of the success of AR models are their conceptual simplicity and their flexibility for modeling quasi-periodic phenomena (e.g. sunspots time series) and short-term dependencies (e.g. day-to-day memories in weather systems). One drawback of current linear AR models is that they are unable to represent the distributional behavior of maxima. Classical Extreme Value Theory (EVT) (e.g. Embrechts *et al.*, 1997; de Haan and Ferreira, 2006; Coles, 2001; Beirlant *et al.*, 2004) dictates that correctly normalized maxima should follow (under various conditions) a Generalized Extreme Value (GEV) distribution. The key characteristic of the GEV is its stability for the max operator. The maximum of two independent and identically distributed (iid) GEV distributed random variables is still GEV distributed. But adding two GEV random variables does not generate a GEV distributed random variable. This explains why linear AR processes are not generally used to describe maxima behavior. For example, Davis and Resnick (1989) or Zhang and Smith (2008) defined and studied max AR (and not additive AR) models with Fréchet distributed marginals. In finance and reinsurance, two well studied EVT domains of applications in which heavy tailed distributions are prevalent, this issue may not be central because taking the maxima or the sum of two heavy tailed random variables is basically equivalent for the upper tail behavior, see chapter 2 of Embrechts *et al.* (1997). In contrast, light-tailed random variables are much more common in atmospheric sciences, e.g. temperature maxima. The link between max and sum for heavy tails is not valid anymore and other methods have to be developed to combine linear AR for light tails and maxima. It is well-known (Fisher and Tippett, 1928; Gnedenko, 1943) that correctly normalized maxima from such light-tailed distributions belong to the Gumbel domain. This means that maxima can be expected to be adequately fitted by the Gumbel distribution defined by

$$H_{\mu,\sigma}(x) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, \text{ with } -\infty < x < +\infty, \tag{1}$$

where μ and σ correspond to the so-called location and scale parameters, respectively. For example, the Quantile-quantile (QQ-) plots in Section 4 illustrates this point for maxima of methane (CH4) and of nitrous oxide (N2O).

At least two possible methods can be implemented to model the bivariate behavior between two consecutive maxima. The now popular copula approaches (e.g. Joe, 1997) allow construction of bivariate distributions under the assumption that all marginals can be identified. Another possibility is to take advantage of bivariate EVT, i.e. to choose and estimate a bivariate extremal dependence function (e.g. Naveau *et al.*, 2008). Although both aforementioned methods are flexible and practical, we prefer to opt for a new representation based on AR processes. This approach has the advantages that AR equations provide explicit relationships and can be directly used for prediction purposes. In addition, as we will see in the coming sections, parameter estimation is based on classical techniques and their interpretation is straightforward. All these elements are of importance for atmospheric scientists who, by training, are already well-versed in dynamical equations. For example, AR processes are routinely used in filtering schemes in atmospheric sciences (assimilation, kalman filtering, etc).

Before presenting and explaining in detail our models, our paper can be summarized as follows. In Section 2, the description and the main properties of our AR models are given. Simulations to assess the quality of our parameter estimators and prediction exercises on simulated data are presented in Section 3. Section 4 focuses on the analysis of CH4 and N2O maxima. As usual, conclusions and perspectives are given in the last section. Finally all the proofs are given in the Appendix.

2 Gumbel AR models

The building block of our models is an additive relationship between Gumbel and positive α-stable variables. Recall that a random variable S is said to be stable if for all non-negative real numbers c_1, c_2, there exists a positive real a and a real b such that $c_1 S_1 + c_2 S_2$ is equal in distribution to $aS + b$ where S_1, S_2 are iid copies of S.

If X is Gumbel distributed with parameters μ and σ and is independent of S which represents a positive α-stable variable with $\alpha \in (0,1)$ defined by its Laplace transform

$$\mathbb{E}(\exp(-uS)) = \exp(-u^{\alpha}), \text{ for all } u \geq 0, \tag{2}$$

then the sum $X + \sigma \log S$ is also Gumbel distributed with parameters μ and σ/α. Such an additive property has been recently studied by Fougères *et al.* (2008) in a mixture context. Crowder (1989), Hougaard (1986) and Tawn (1990) also worked with such distributions in survival analysis and the modeling of multivariate extremes. In time series analysis, the additive stability between Gumbel and positive $\alpha-$stable random variables allows us to propose a simple linear AR model that can be summarized by the following proposition.

PROPOSITION 1 *Let S_t be iid positive α-stable variables defined by (2) for any $t \in \mathbb{Z}$. Let $\{X_t, t \in \mathbb{Z}\}$ be a stochastic process defined by the recursive relationship*

$$X_t = \alpha\, X_{t-1} + \alpha\, \sigma\, \log S_t \tag{3}$$

where $\sigma > 0$. Equation (3) has a unique strictly stationary solution,

$$X_t = \sigma \sum_{j=0}^{\infty} \alpha^{j+1} \log S_{t-j} \tag{4}$$

and X_t follows a Gumbel distribution with parameters $(0, \sigma)$.

Although one can easily recognize the classical AR(1), $X_t = \alpha X_{t-1} + \sigma \epsilon_t$ with $\epsilon_t = \alpha \log S_t$ in (3), it is important to notice that the "noise" ϵ_t depends here on α. An advantage of the parameterization (3) is that X_t follows a Gumbel distribution whose parameters are independent of α. The covariance between X_t and X_{t-h} is increasing with α; more precisely, we have $\mathbb{C}ov(X_t, X_{t-h}) = \mathbb{V}ar(X_t)\alpha^{|h|}$. The first two moments of the Gumbel distribution can also be easily computed

$$\mu = \mathbb{E}(X_0) - \pi^{-1}\delta\sqrt{6\mathbb{V}ar(X_0)} \tag{5}$$

and

$$\sigma = \pi^{-1}\sqrt{6\mathbb{V}ar(X_0)} \tag{6}$$

where δ is the Euler's constant.

To simplify the statement of our proposition, the Gumbel location parameter μ was set equal to zero. In practice, μ can be different from zero. It suffices to add μ to X_t in (4) to have a Gumbel(μ, σ) distribution. Another possible extension is the following AR model

$$Z_t = \alpha_1\alpha_2 Z_{t-1} + \alpha_1\alpha_2\sigma \log S_t(\alpha_1) + \alpha_2\sigma \log S_{t-1}(\alpha_2) \tag{7}$$

where $S_t(\alpha_1)$ and $S_t(\alpha_2)$ are independent sequences of iid positive α_i-stable variables with $i = 1, 2$. If Z_0 is Gumbel distributed with parameters $(0, \sigma)$ then Z_t is also Gumbel distributed with parameters $(0, \sigma)$ for any $t > 0$. Equation (7) can be used to define a Gumbel ARMA(1,1) and, a generalization of the same idea could produce a Gumbel ARMA$(1, q)$. Instead of studying in detail such models, we prefer to investigate the properties of the simpler Gumbel AR(1) defined by (3). This latter model has fewer coefficients, whereas more complex models, although important for some applications, do not bring any new conceptual ideas in this paper.

Parameterization (3) offers an explicit identification of the process X_t by its characteristic function as shown in the next proposition.

PROPOSITION 2 *Let $\{X_t, t \in \mathbb{Z}\}$ be defined as in Proposition 1. The characteristic function of any random vector $\mathbf{X_h} = (X_t, \ldots, X_{t-h})'$ with $h > 0$ can be written as*

$$\mathbb{E}\left(\exp[i\, u'\mathbf{X_h}]\right) = \Gamma\left(1 - i\sigma \sum_{j=0}^{h} u_j \alpha^{h-j}\right) \prod_{j=0}^{h-1} \frac{\Gamma\left(1 - i\sigma \sum_{k=0}^{j} u_k \alpha^{j-k}\right)}{\Gamma\left(1 - i\sigma \sum_{k=0}^{j} u_k \alpha^{j-k+1}\right)}.$$

A natural question connected to extreme events analysis is to know what kind of dependence is present in the upper tail. There are a variety of ways to answer such a question (e.g. Fougères, 2004). Coles (2001) or Coles *et al.* (1999) advocate the following two upper tail dependence coefficients

$$\chi = \lim_{x \to \infty} \frac{\mathbb{P}(X_{t-1} > x, X_t > x)}{\mathbb{P}(X_{t-1} > x)} \text{ and } \overline{\chi} = \lim_{x \to \infty} \frac{2 \log \mathbb{P}(X_{t-1} > x)}{\log \mathbb{P}(X_{t-1} > x, X_t > x)} - 1.$$

These quantities can be computed for our Gumbel AR model.

PROPOSITION 3 *Let $\{X_t, t \in \mathbb{Z}\}$ be defined as in Proposition 1. The parameter χ equals zero, while the dependence parameter $\overline{\chi}$ is equal to $\alpha/(2-\alpha) \in (0,1)$.*

As expected for light-tailed distributions, χ is null and this situation corresponds to the so-called asymptotic independence (e.g. Coles, 2001). Still, the coefficient $\overline{\chi}$ clearly indicates that the dependence strength in the upper tail increases almost proportionally to α.

To estimate our three parameters α, σ and μ in our Gumbel AR(1) model, we opt for a method of moments approach because of its simplicity of implementation and its good asymptotic properties. A possible alternative resides in a maximum likelihood procedure (e.g. Breidt and Davis, 1991; Andrews *et al.*, 2008). The expressions of the two parameters of the Gumbel distribution given in Equation (5) and Equation (6) provide the following estimators of μ and σ

$$\widehat{\mu} = \overline{X} - \pi^{-1}\delta\sqrt{6}s \text{ and } \widehat{\sigma} = \pi^{-1}\sqrt{6}s \tag{8}$$

where $\overline{X} = \sum_{t=1}^{T} X_t/T$ and $s^2 = \sum_{t=1}^{T}(X_t - \overline{X})^2/T$. Concerning the estimation of α, a least-square estimator can be introduced by writing

$$\arg\min_r \left\{ \sum_{t=1}^{T-1} \left([X_{t+1} - \mathbb{E}(X_0)] - r\,[X_t - \mathbb{E}(X_0)] \right)^2 \right\} = \frac{\sum_{t=1}^{T-1}(X_t - \mathbb{E}(X_0))(X_{t+1} - \mathbb{E}(X_0))}{\sum_{t=1}^{T-1}(X_t - \mathbb{E}(X_0))^2}.$$

This is similar to the classical Yule-Walker equation for AR(1) models. It follows that our estimator of α is simply

$$\widehat{\alpha} = \frac{1}{s^2 T} \sum_{t=1}^{T-1} (X_t - \overline{X})(X_{t+1} - \overline{X}). \tag{9}$$

The asymptotic properties of our triplet of estimators can be summarized by the following proposition.

PROPOSITION 4 *As T the sample size goes to infinity, the estimators of μ, σ and α defined by (8) and (9) are almost surely consistent and the vector $\sqrt{T}\,(\widehat{\mu}-\mu, \widehat{\sigma}-\sigma, \widehat{\alpha}-\alpha)'$ converges in distribution to a zero-mean Gaussian vector with covariance matrix*

$$\begin{pmatrix} \frac{\pi^2\sigma^2}{6}\frac{1+\alpha}{1-\alpha} - \frac{12\delta\sigma^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} + \frac{11\delta^2\sigma^2(1+\alpha^2)}{10(1-\alpha^2)} & \frac{6\sigma^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} - \frac{11\delta\sigma^2(1+\alpha^2)}{10(1-\alpha^2)} & -\alpha\sigma\delta \\ \frac{6\sigma^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} - \frac{11\delta\sigma^2(1+\alpha^2)}{10(1-\alpha^2)} & \frac{11\sigma^2(1+\alpha^2)}{10(1-\alpha^2)} & \alpha\sigma \\ -\alpha\sigma\delta & \alpha\sigma & 1-\alpha^2 \end{pmatrix} \tag{10}$$

where $\zeta(.)$ represents the Riemann function.

3 Simulation results

To study the finite sample size behavior of our estimators $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha}$, we generate 1000 samples from (3) with different values of n (sample size) and α. We have chosen $n = 50, 100, \ldots, 1000$, $\alpha \in \{0.2, 0.5, 0.8\}$, $\mu = 0$ and $\sigma = 2$. Figure 1 recapitulates the outputs of our simulations. The top, medium and bottom panels correspond respectively to the properties of $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha}$ with respect to different sample sizes (x-axis). The mean, first and third quartiles of our 1000 replica are represented by the dashed line and the two dotted-dashed lines, respectively. For small sample sizes, we observe some bias in the estimation of α which is asymmetric. As the dependence captured by α decreases, the estimation of the two Gumbel parameters μ and σ improves.

From a prediction point of view, it is interesting to quantify the error if one predicts with a classical Gaussian AR(1) model while the underlying true model is Gumbel AR(1) one. As discussed in the introduction, this could be the case if the variable of interest is a maximum obtained from a light-tailed distribution. To reach this goal, we remark that, under our model (3), we have

$$\mathbb{P}(X_{t+1} \leq y | X_t = x) = \mathbb{P}\left(\log S_{t+1} \leq \frac{y - \alpha x}{\alpha \sigma}\right). \tag{11}$$

This means that the predictive distribution of $[X_{t+1}|X_t = x]$ under (3) is simply the one of the log of a positive α-stable random variable. To visualize this one-step prediction density, we simulate a sample of 1000 observations from our AR(1) Gumbel model with $\alpha = 0.5$, $\mu = 0$ and $\sigma = 2$. After estimating μ, σ and α according to (8) and (9), 1000 values of $[\widehat{X}_{t+1}|X_t = x]$ are drawn from $\widehat{\alpha}$-stable positive realizations according to (11), $\widehat{\mu}$ and $\widehat{\sigma}$. The corresponding histogram and the true density (solid line) are superimposed in the left panel of Figure 2. The same exercise is repeated but under the wrong assumption that the model is Gaussian AR(1). The right panel clearly indicates a discrepancy between the true density

$\alpha = 0.2$ $\alpha = 0.5$ $\alpha = 0.8$

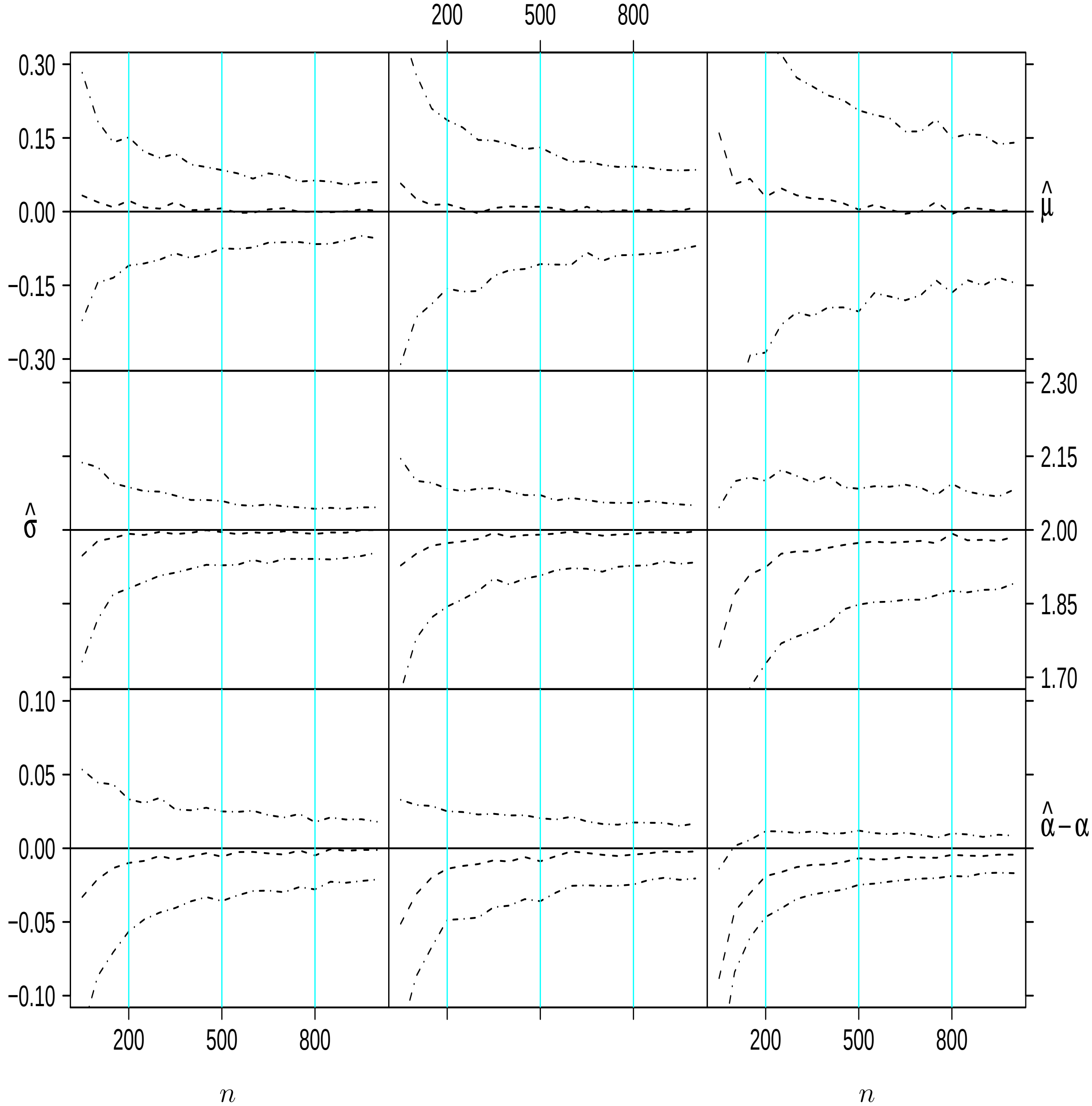

Figure 1: Simulation results for $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha}$ defined by (8) and (9). The mean, first and third quartiles obtained from 1000 replica with $\mu = 0$, $\sigma = 2$ and $\alpha \in \{0.2, 0.5, 0.8\}$ correspond to the dashed line and the two dotted-dashed lines, respectively. The x-axis represents different sample sizes. The top, medium and bottom panels corresponds to $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha} - \alpha$.

(solid line) and the estimated histogram obtained under the Gaussian setup. Of course, it is not surprising that the asymmetry present by construction in $[X_{t+1}|X_t = x]$ cannot be handled by the Gaussian model. This simply illustrates that the information contained in the type of random variable, here maxima, can help in the modeling of predictive densities whenever classical EVT can be applied.

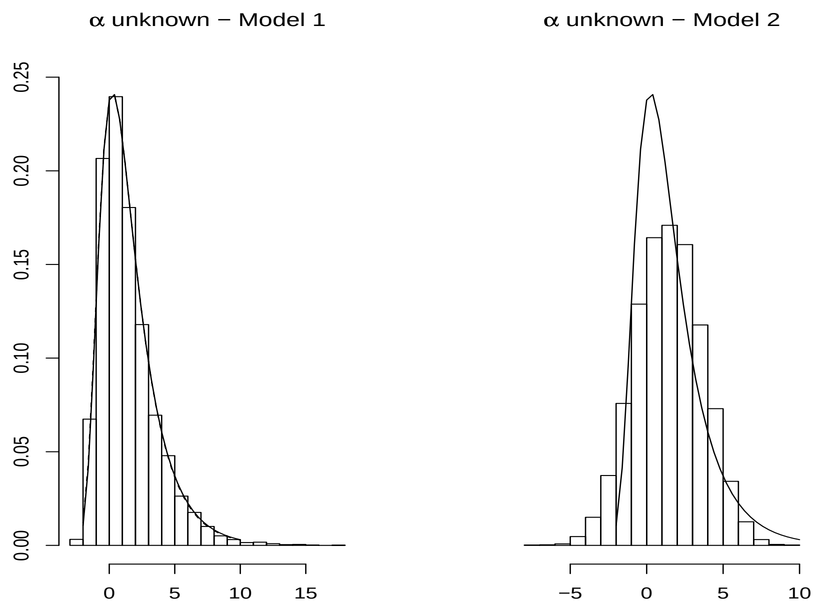

Figure 2: Conditional histograms of $\widehat{X}_{t+1}$ for the Gumbel model (Model 1) and for the Gaussian one (Model 2).

4 Analysis of maxima of CH4 and N2O

In practice the connection between light-tailed maxima and the Gumbel distribution can be illustrated by environmental variables. Our example choice is primarily motivated by atmospheric considerations. After water vapor, carbon dioxide, methane and nitrous oxide are the three most important greenhouse gases. They play a fundamental role in our understanding of the past, present and future state of the Earth atmosphere. In this context, identifying the temporal structure among the largest CH4 and N2O measurements is of primary interest for the atmospheric chemist because this can help to predict future maxima of CH4 and N2O at a specific location. For our site of Gif-sur-Yvette (France), the maxima block sizes of a day and a week are convenient with respect to the length and the resolution of our time series. The length of our records, five years, is too short to study yearly maxima, or even monthly maxima, at a climatic scale.

To illustrate the distributional behaviors of our chosen variables, a Gumbel distribution has been fitted to two time series of daily and weekly CH4 concentration recorded from 2002 to 2007. The data measured in parts per billion (ppb) are presented in Figure 3.

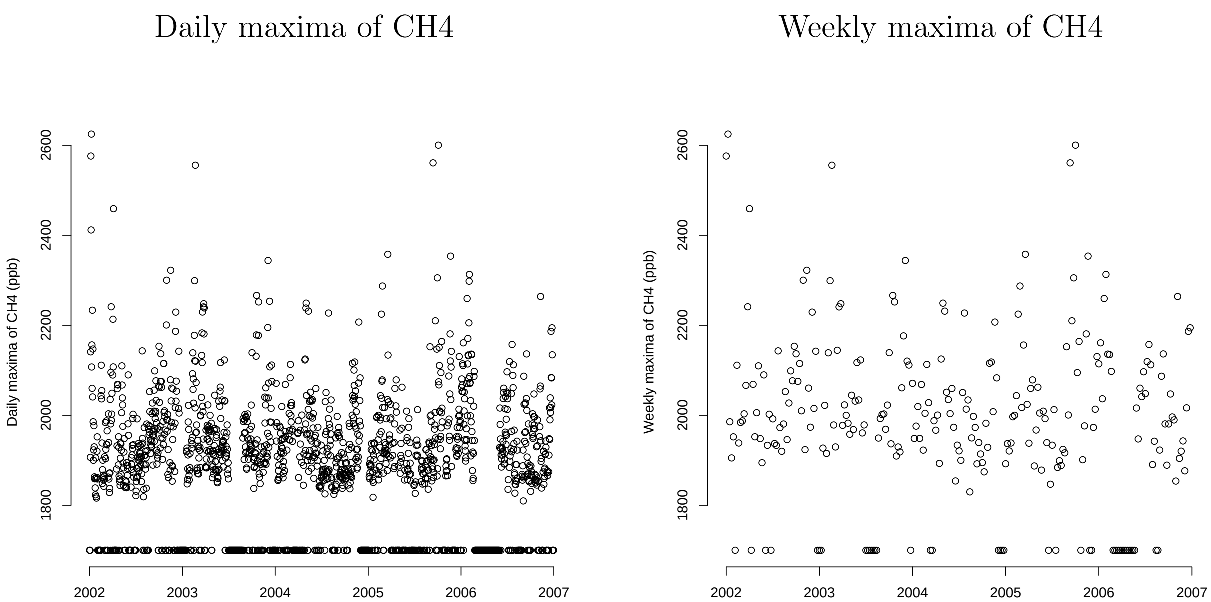

Figure 3: The y-axis corresponds to maxima of CH4 in ppb recorded at Gif-sur-Yvette (France) from 2002 to 2007 (x-axis). The zeros represent missing values.

QQ-plots displayed in Figure 4 indicate that a Gumbel fit seems to be reasonable. Concerning the short-term temporal dependence, the scatter plot of two consecutive maxima of CH4 in Figure 5 shows a dependence, as one would expect, that seems stronger at the daily scale than at the weekly one. As the marginals can be approximated by a Gumbel density, classical

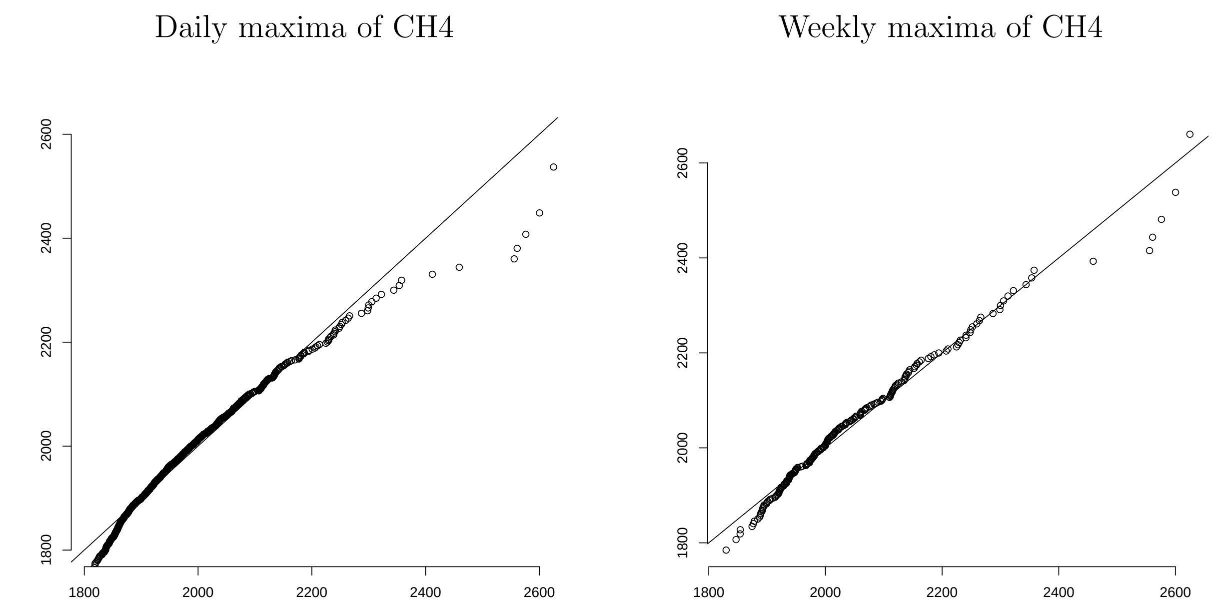

Figure 4: Gumbel QQ-plot of the daily and weekly maxima of CH4 displayed in Figure 3. The two Gumbel parameters of (1) are estimated by the method of moments.

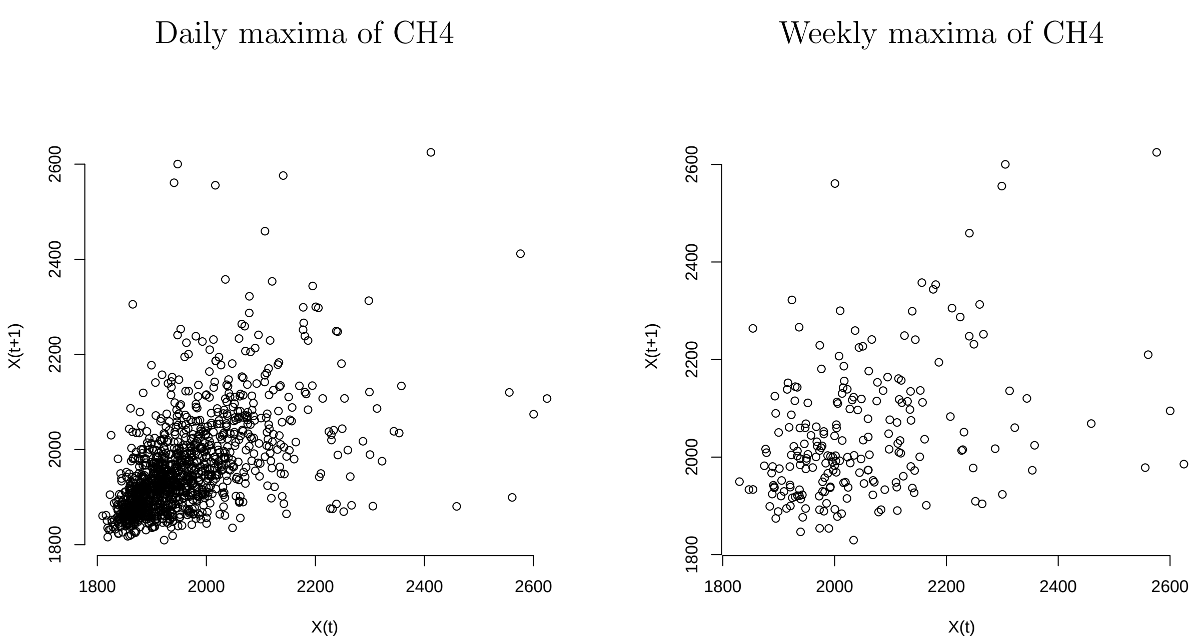

Figure 5: Scatter plots of consecutive maxima of CH4.

correlation measures are not appropriate to capture such dependencies among maxima. The same type of plots (QQ-plot and scatter plot) can be obtained for daily maxima of N2O, see Figure 6 and the same type of conclusions can also be made. Note that an increasing linear trend has been removed from the daily maxima of N2O in order to make them more stationary (see the upper panel of Figure 6).

Concerning our CH4 example, α's estimators are equal to $\widehat{\alpha}_{Day} = 0.54$ and $\widehat{\alpha}_{Week} = 0.35$. From Proposition 4, we obtain for α the following 95% confidence intervals: $CI_{95\%}(\alpha_{Day}) : [0.49, 0.59]$ and $CI_{95\%}(\alpha_{Week}) : [0.23, 0.47]$. As an example, for a visual check, we plot in Figure 7 the scatter plot of two consecutive maxima from a Gumbel AR(1) with parameters corresponding to the estimators from the series and with $\alpha = 0.35$. Moreover the sizes of the observed and simulated series are the same. Therefore comparison of Figure 5 for weekly maxima and Figure 7 shows a good visual agreement between the observed and simulated bivariate structure for successive maxima.

To assess the predictive power of our model, we estimate the three parameters on the first period, here from 2002 to the middle of 2006. For the second part of 2006, we draw 1000 $\widehat{X}_{t+1} = \widehat{\alpha} x_t + \widehat{\alpha}\widehat{\sigma} \log S_{t+1} + \widehat{\mu}(1-\widehat{\alpha})$ with x_t the observed value at time t and S_{t+1} a random positive $\widehat{\alpha}$-stable variable as defined in (2). Then the empirical quartiles of the distribution of $[\widehat{X}_{t+1} | X_t = x_t]$ are deduced. Figure 8 represents the prevision of daily maxima (on the left) and weekly maxima (on the right) of methane in Gif-sur-Yvette during the second period of 2006.

5 Conclusions and perspectives

Our Gumbel AR(1) process defined and studied in this paper offers a simple way to model short-term dependencies among maxima stemming from light-tailed distributions. Although

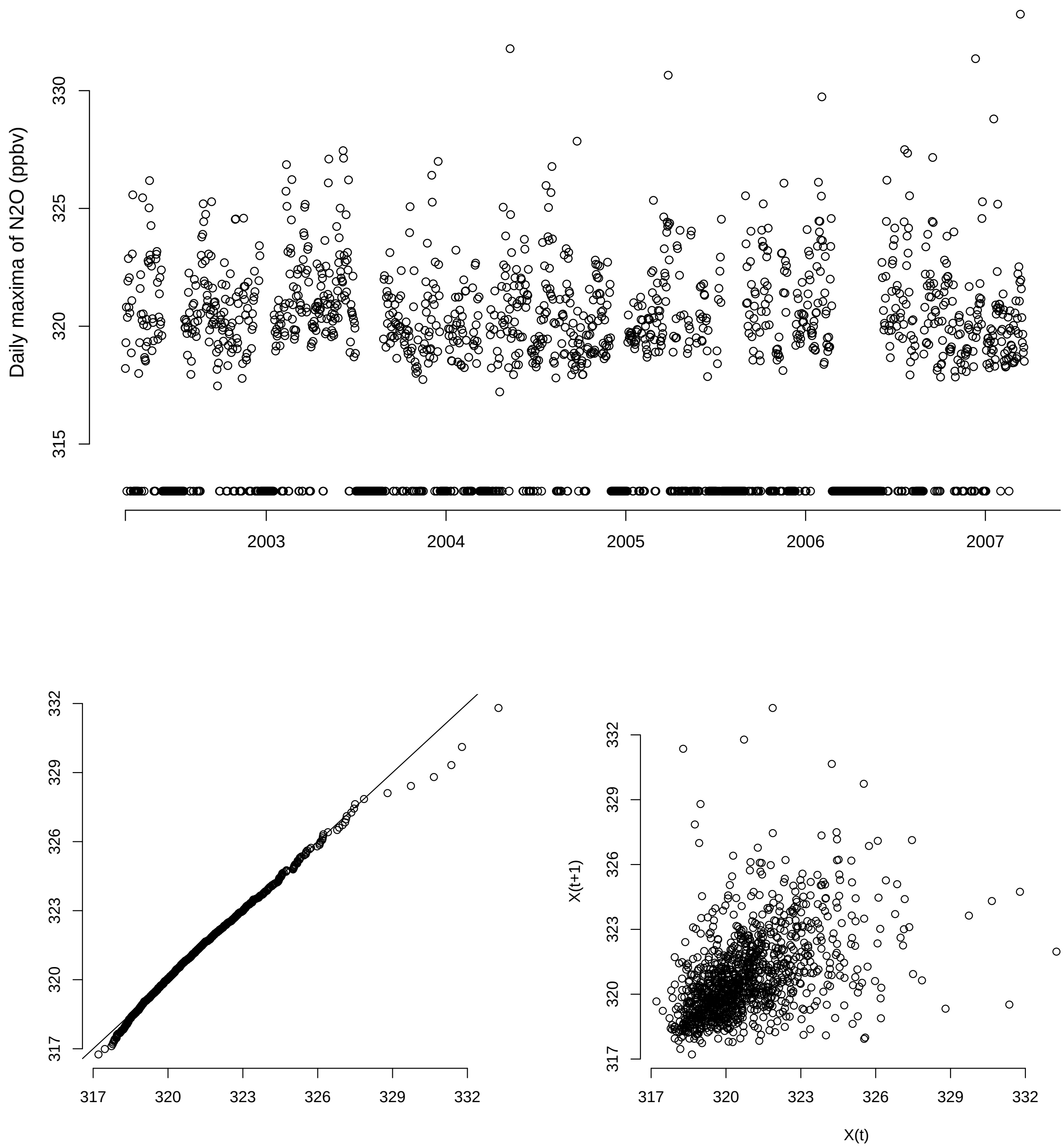

Figure 6: Daily maxima of N2O in parts per billion by volumn (ppbv) recorded at Gif-sur-Yvette (France) from 2002 to 2007. Upper panel: N2O time series after removing a linear trend (the "zeros" represent missing values). Left lower panel: Gumbel QQ-plot of the upper panel data. Right lower panel: scatter plot of consecutive values from the upper panel data.

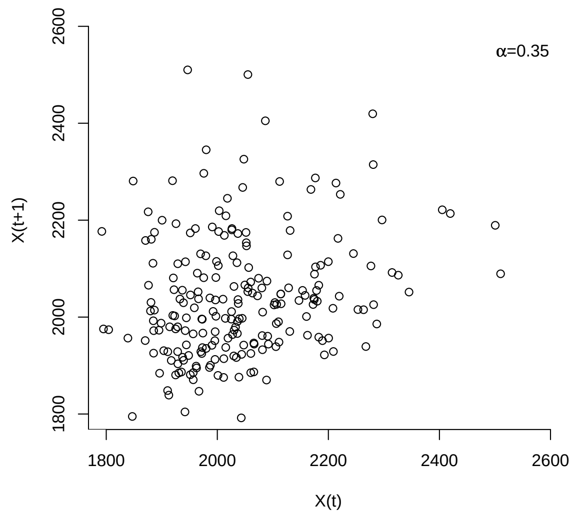

Figure 7: Scatter plot from a Gumbel AR(1) with $\alpha = 0.35$.

it is possible to extend it to ARMA$(1, q)$, we do not yet know how to preserve the Gumbel characteristic for AR(p) for $p \geq 2$. In addition, we impose that maxima have to follow a Gumbel distribution. As noticed in Section 1, this is reasonable for a lot of variables in atmospheric sciences. Still, some variables like precipitation records at some locations and specific temporal scales may not be light tail distributed but rather slightly heavy tailed. Hence, for such phenomena, it would be of interest to propose a more general AR model like a GEV AR(1) process. This is possible if we define S_t, $t \in \mathbb{Z}$, as in Proposition 1 and $\{X_t, t \in \mathbb{Z}\}$ by the recurrence equation

$$X_t + \frac{\sigma}{\xi} = \left(X_{t-1} + \frac{\sigma}{\xi} \right)^{\alpha} \times S_t^{\alpha\xi} \times \left(\frac{\sigma}{\xi} \right)^{1-\alpha} \tag{12}$$

where $\sigma > 0$ and $\xi \in \mathbb{R}^*$. It is possible to demonstrate as in Proposition 1 that Equation (12) has a unique strictly stationary solution given by

$$X_t + \frac{\sigma}{\xi} = \frac{\sigma}{\xi} \prod_{j=0}^{\infty} (S_{t-j})^{\xi\alpha^{j+1}} \tag{13}$$

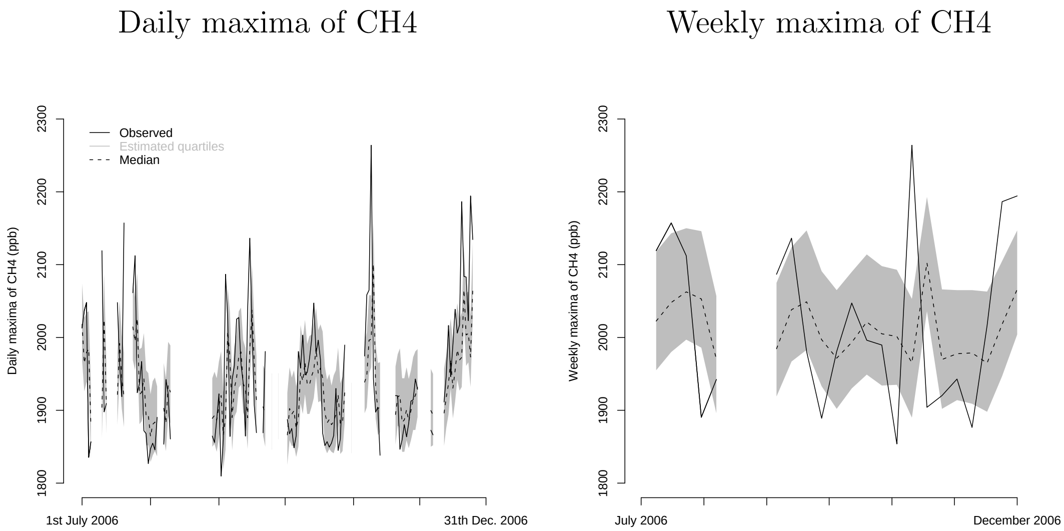

Figure 8: One-step prevision of methane daily maxima (on the left) and methane weekly maxima (on the right) on the second part of the year 2006.

and that X_t follows a $GEV(0, \sigma, \xi)$ distribution. This model could be used in practice. Nevertheless, this extension leads to a non additive model. This complexity could diminish its application in atmospheric sciences. Maybe, a more promising road would be to develop and study a state-space model based on our Gumbel AR(1) process. This could lead to important applications in filtering schemes like data assimilation, the latter being routinely used by geoscientists.

Acknowledgment

The CH4 and N2O measurements have been provided by the RAMCES Observatory for Research on the Environment. This work was supported by the CS of the University of Strasbourg, by the AssimilEx project of the french ‘Agence Nationale de la Recherche’ and by the Weather and Climate Impact Assessment Science Initiative at the National Center for Atmospheric Research (USA). We would like also to thank Anne-Laure Fougères for valuable discussions about our Gumbel AR(1) model. Finally, the insightful suggestions and remarks from two anonymous reviewers and the editor have greatly improved the scope and

the clarity of this work.

6 Appendix

Proof of Proposition 1. With $Y_t = X_t - \delta\sigma$ and $\varepsilon_t = \alpha\sigma \log S_t - \delta\sigma(1-\alpha)$ where δ is the Euler's constant, (3) may be rewritten as

$$Y_t = \alpha Y_{t-1} + \varepsilon_t. \tag{14}$$

This is a well-known model, an AR(1) process where the random variables ε_t are iid with null expectation and variance equal to $\alpha^2\sigma^2\sigma_\varepsilon^2$. Indeed, according to Zolotarev (1986, Section 3.6), $\mathbb{E}(\log S) = \delta\,(1/\alpha - 1)$ and $\mathbb{V}ar(\log S) = (\pi^2/6) \times (1/\alpha^2 - 1) =: \sigma_\varepsilon^2$. According to classical results concerning the AR process of order one detailed in Brockwell and Davis (1987, Section 3.1), we have in case $|\alpha| < 1$ and if $\{Y_t\}$ is stationary that

$$\lim_{n\to\infty} \mathbb{E}\left(Y_t - \sum_{j=0}^{n} \alpha^j \varepsilon_{t-j}\right)^2 = 0 \tag{15}$$

and $\sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j}$ is mean-square convergent. Consequently the process $\{Y_t\}$ has a unique second-order stationary solution

$$Y_t = \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j}. \tag{16}$$

Since it is obvious that (16) is equivalent to (4), the process $\{X_t\}$ defined in (4) is the unique second-order stationary solution of (3).

In order to obtain the distribution of X_t we are interested in the characteristic function of $\log S$. According to Zolotarev (1986, p. 117), it is possible to establish the characteristic function of $\log S$ as follows:

$$\mathbb{E}\left(\exp(iu\log S)\right) = \frac{\Gamma\left(1-\frac{iu}{\alpha}\right)}{\Gamma(1-iu)}. \tag{17}$$

The characteristic function of X_t could now be computed. According to (15) we have

$$\lim_{n\to\infty} \mathbb{E}\left(X_t - \sigma\sum_{j=0}^{n} \alpha^{j+1}\log S_{t-j} - \delta\sigma\alpha^{n+1}\right)^2 = 0 \tag{18}$$

which implies

$$\mathbb{E}\left(e^{iuX_t}\right) = \lim_{n\to\infty} \mathbb{E}\left(e^{iu\left(\sigma\sum_{j=0}^{n}\alpha^{j+1}\log S_{t-j}+\delta\sigma\alpha^{n+1}\right)}\right) = \lim_{n\to\infty} e^{iu\delta\sigma\alpha^{n+1}}\mathbb{E}\left(\prod_{j=0}^{n} e^{iu\sigma\alpha^{j+1}\log S_{t-j}}\right).$$

Since the variables S_t, $t\in\mathbb{Z}$, are independent, it is possible to show that $\mathbb{E}\left(e^{iuX_t}\right) = \Gamma(1-iu\sigma)$ which exactly corresponds to the characteristic function of a Gumbel$(0,\sigma)$ distribution. The same result holds for a Gumbel(μ,σ) with μ not necessarily equal to zero. It suffices to add μ to X_t. Moreover, since S_t, for all integers t, are iid, the process X_t is not only second-order stationary but strictly stationary. □

Proof of Proposition 2. For any $h>0$, let

$$\mathbf{X_h} = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-k} \\ \vdots \\ X_{t-h} \end{pmatrix} = \begin{pmatrix} \alpha^h X_{t-h} + \sigma\sum_{j=0}^{h-1}\alpha^{j+1}\log S_{t-j} \\ \alpha^{h-1}X_{t-h} + \sigma\sum_{j=1}^{h-1}\alpha^{j}\log S_{t-j} \\ \vdots \\ \alpha^{h-k}X_{t-h} + \sigma\sum_{j=k}^{h-1}\alpha^{j-k+1}\log S_{t-j} \\ \vdots \\ X_{t-h} \end{pmatrix} \text{ and } u = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_k \\ \vdots \\ u_h \end{pmatrix}.$$

Then

$$
\begin{aligned}
\mathbb{E}\left(\exp[i\ u^{'}\mathbf{X_h}]\right) =&\mathbb{E}\left(e^{iu_0X_t+iu_1X_{t-1}+\dots+iu_kX_{t-k}+\dots+iu_hX_{t-h}}\right)\\
=&\mathbb{E}\left(e^{iu_0\left(\alpha^hX_{t-h}+\sigma\sum_{j=0}^{h-1}\alpha^{j+1}\log S_{t-j}\right)+\dots+iu_k\left(\alpha^{h-k}X_{t-h}+\sigma\sum_{j=k}^{h-1}\alpha^{j-k+1}\log S_{t-j}\right)+\dots+iu_hX_{t-h}}\right)\\
=&\mathbb{E}\left(e^{iX_{t-h}\sum_{j=0}^{h}u_j\alpha^{h-j}}\right)\ \mathbb{E}\left(e^{i\sigma\sum_{j=0}^{h-1}\left(\sum_{k=0}^{j}\alpha^{j-k+1}u_k\right)\log S_{t-j}}\right)\\
=&\Gamma\left(1-i\sigma\sum_{j=0}^{h}u_j\alpha^{h-j}\right)\prod_{j=0}^{h-1}\mathbb{E}\left(e^{i\sigma\left(\sum_{k=0}^{j}\alpha^{j-k+1}u_k\right)\log S_{t-j}}\right)\\
=&\Gamma\left(1-i\sigma\sum_{j=0}^{h}u_j\alpha^{h-j}\right)\prod_{j=0}^{h-1}\frac{\Gamma\left(1-i\sigma\sum_{k=0}^{j}u_k\alpha^{j-k}\right)}{\Gamma\left(1-i\sigma\sum_{k=0}^{j}u_k\alpha^{j-k+1}\right)}. \qquad \square
\end{aligned}
$$

Proof of Proposition 3. The following quantity is essential in the computation of the two dependence parameters χ and $\overline{\chi}$

$$
\begin{aligned}
\mathbb{P}(X_{t-1}>x,X_t>x) &= \mathbb{P}(X_{t-1}>x,\alpha X_{t-1}+\alpha\sigma\log S_t>x)\\
&= \int_x^\infty \mathbb{P}\left(\log S_t>\frac{1}{\alpha\sigma}(x-\alpha y)\mid X_{t-1}=y\right)dH_{0,\sigma}(y)\\
&= \int_x^\infty \mathbb{P}\left(S_t>\exp\left(\frac{1}{\alpha\sigma}(x-\alpha y)\right)\right)dH_{0,\sigma}(y)
\end{aligned}
$$

where S_t is a positive $\alpha-$stable variable and X_t a Gumbel $(0,\sigma)-$distributed random variable. Consequently

$$
\begin{aligned}
\mathbb{P}(X_{t-1}>x,X_t>x) &= \int_0^{\exp(-\frac{1}{\sigma}x)}\mathbb{P}\left(S_t>z\exp\left(\frac{x}{\alpha\sigma}\right)\right)\exp(-z)dz\\
&= \exp\left(-\frac{x}{\alpha\sigma}\right)\int_0^{\exp(-\frac{x}{\alpha\sigma}(\alpha-1))}\exp\left(-u\exp\left(-\frac{x}{\alpha\sigma}\right)\right)\mathbb{P}(S_t>u)du\\
&= \exp\left(-\frac{x}{\alpha\sigma}\right)\int_0^\infty\exp\left(-u\exp\left(-\frac{x}{\alpha\sigma}\right)\right)1_{0\le u\le e^{\frac{x}{\alpha\sigma}(1-\alpha)}}\mathbb{P}(S_t>u)du.
\end{aligned}
$$

Note that $\frac{1}{\mathbb{P}(X_{t-1}>x)} \overset{x\to\infty}{\sim} \exp\left(\frac{x}{\sigma}\right)$. The dependence parameter χ can be written as

$$\begin{aligned}\chi &= \lim_{x\to\infty} \frac{\mathbb{P}(X_{t-1} > x, X_t > x)}{\mathbb{P}(X_{t-1} > x)} \\ &= \lim_{x\to\infty} \exp\left(\frac{x}{\alpha\sigma}(\alpha-1)\right) \int_0^\infty \exp\left(-u \exp\left(-\frac{x}{\alpha\sigma}\right)\right) 1_{0\le u\le e^{\frac{x}{\alpha\sigma}(1-\alpha)}} \mathbb{P}(S_t > u) du \\ &= \lim_{x\to\infty} \int_0^1 \exp\left(-\omega \exp\left(-\frac{x}{\sigma}\right)\right) \mathbb{P}\left(S_t > \omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right) d\omega.\end{aligned}$$

Since S_t is a positive $\alpha-$stable variable, we have $\mathbb{P}(S_t > x) = x^{-\alpha}L(x)$ where L is a slowly varying function (see Bingham *et al.*, 1987, Section 8.3.5), which implies that

$$\chi = \lim_{x\to\infty} \int_0^1 \exp\left(-\omega \exp\left(-\frac{x}{\sigma}\right)\right) \omega^{-\alpha} L\left(\omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right) d\omega\, e^{-\frac{x}{\sigma}(1-\alpha)}.$$

By Taylor expansion, we have $\exp(-\omega\exp(-x/\sigma)) = 1 - \omega e^{-\omega\kappa} e^{-\frac{x}{\sigma}}$ with $\kappa \in \left(0, e^{-\frac{x}{\sigma}}\right)$, and therefore

$$\begin{aligned}\chi &= (1+o(1)) \lim_{x\to\infty} \int_0^1 \omega^{-\alpha} L\left(\omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right) d\omega\, e^{-\frac{x}{\sigma}(1-\alpha)} \\ &\sim \frac{1}{1-\alpha} \lim_{x\to\infty} \frac{L\left(e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right)}{e^{\frac{1}{\sigma}x(1-\alpha)}} \\ &= 0,\end{aligned}$$

since $0 < \alpha < 1$, the approximation coming from Karamata's theorem.

Now, we are interested in the dependence parameter $\overline{\chi}$:

$$
\begin{aligned}
\overline{\chi} &= \lim_{x\to\infty} \frac{2\log \mathbb{P}(X_{t-1} > x)}{\log \mathbb{P}(X_{t-1} > x, X_t > x)} - 1 \\
&= \lim_{x\to\infty} \frac{-2x/\sigma}{-\frac{x}{\alpha\sigma} + \log\left(\int_0^\infty \exp\left(-u\exp\left(-\frac{x}{\alpha\sigma}\right)\right) 1_{0\le u\le e^{\frac{x}{\alpha\sigma}(1-\alpha)}}\, \mathbb{P}(S_t > u)du\right)} - 1 \\
&= \lim_{x\to\infty} \left[\frac{1}{2\alpha} - \frac{\sigma}{2x}\log\left(\int_0^1 \exp\left(-\omega\exp\left(-\frac{x}{\sigma}\right)\right)\mathbb{P}\Big(S_t > \omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\Big)\, d\omega\, e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right)\right]^{-1} - 1 \\
&= \lim_{x\to\infty} \left[\frac{1}{2} - \frac{\sigma}{2x}\log\left(\int_0^1 \exp\left(-\omega\exp\left(-\frac{x}{\sigma}\right)\right)\mathbb{P}\Big(S_t > \omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\Big)\, d\omega\right)\right]^{-1} - 1.
\end{aligned}
$$

As previously, we have $\mathbb{P}(S_t > x) = x^{-\alpha}L(x)$ where L is a slowly varying function. Consequently, we obtain

$$
\begin{aligned}
\overline{\chi} &= \lim_{x\to\infty} \left[\frac{1}{2} - \frac{\sigma}{2x}\log\left(\int_0^1 \exp\left(-\omega\exp\left(-\frac{x}{\sigma}\right)\right)\omega^{-\alpha}L\left(\omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right) d\omega\, e^{-\frac{x}{\sigma}(1-\alpha)}\right)\right]^{-1} - 1 \\
&= \lim_{x\to\infty} \left[1 - \frac{\alpha}{2} - \frac{\sigma}{2x}\log\left(\int_0^1 \exp\left(-\omega\exp\left(-\frac{x}{\sigma}\right)\right)\omega^{-\alpha}L\left(\omega e^{\frac{x}{\alpha\sigma}(1-\alpha)}\right) d\omega\right)\right]^{-1} - 1 \\
&= \left[1 - \frac{\alpha}{2}\right]^{-1} - 1 \\
&= \frac{\alpha}{2-\alpha}. \qquad \square
\end{aligned}
$$

Proof of Proposition 4. Since X_t is Gumbel(μ, σ) distributed, its first and second moments are known and finite. By the ergodic theorem, $\overline{X} = \frac{1}{T}\sum_{t=1}^T X_t$ converges almost surely to $\mathbb{E}(X_0)$ and $s^2 = \frac{1}{T}\sum_{t=1}^T (X_t - \overline{X})^2$ to $\mathbb{V}ar(X_0)$. Then by continuity it follows that $\widehat{\mu}$ converges almost surely to μ and $\widehat{\sigma}$ to σ. Concerning $\widehat{\alpha}$, as $\frac{1}{T}\sum_{t=1}^{T-1}(X_t - \overline{X})(X_{t+1} - \overline{X})$ converges almost surely to $\mathbb{C}ov(X_0, X_1) = \alpha\mathbb{V}ar(X_0)$, it follows that $\widehat{\alpha} = \frac{1}{Ts^2}\sum_{t=1}^{T-1}(X_t - \overline{X})(X_{t+1} - \overline{X})$ converges almost surely to α.

Now let us introduce $\{Y_t\}$ the two-sided moving average defined by $Y_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}$ where $\psi_j = \sigma\alpha^{j+1}$ for $j \geq 0$ and 0 otherwise with $\alpha \in (0, 1)$. We note $\varepsilon_t = \log S_t - \frac{\delta}{\alpha}(1-\alpha)$ with S_t defined as in Proposition 1. Therefore the random variables ε_t are iid with null expectation and variance equal to $\sigma_\varepsilon^2 = (\pi^2/6) \times (1/\alpha^2 - 1)$.

The proof of the remaining part of the proposition can be divided into two parts. First, we establish the following lemma and then we combine it with the delta method in order to conclude.

LEMMA 5 *Let $\{Y_t\}$ defined as previously. The random vector*

$$\sqrt{T}\begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} Y_t \\ \frac{1}{T}\sum_{t=1}^{T}(Y_t^2 - \mathbb{E}Y_t^2) \\ \frac{1}{T}\sum_{t=1}^{T}(Y_tY_{t+1} - \mathbb{E}Y_tY_{t+1}) \end{pmatrix} \tag{19}$$

converges to a normal distribution with mean vector equal to $\mathbf{0}$ *and covariance matrix* $\mathbf{\Sigma_Y}$ *equal to*

$$\begin{pmatrix} \gamma_0\frac{1+\alpha}{1-\alpha} & \frac{2\sigma^3\zeta(3)(1+\alpha+\alpha^2)}{1-\alpha^2} & \frac{2\sigma^3\zeta(3)\alpha(1+\alpha+\alpha^2)}{1-\alpha^2} \\ \frac{2\sigma^3\zeta(3)(1+\alpha+\alpha^2)}{1-\alpha^2} & \frac{22}{5}\gamma_0^2\frac{1+\alpha^2}{1-\alpha^2} & \frac{4\alpha\gamma_0^2}{1-\alpha^2}(1+\frac{3}{5}(1+\alpha^2)) \\ \frac{2\sigma^3\zeta(3)\alpha(1+\alpha+\alpha^2)}{1-\alpha^2} & \frac{4\alpha\gamma_0^2}{1-\alpha^2}(1+\frac{3}{5}(1+\alpha^2)) & \frac{\gamma_0^2}{1-\alpha^2}[(1+\alpha^2)(\frac{12}{5}\alpha^2+1)+\alpha^2(3-\alpha^2)] \end{pmatrix} \tag{20}$$

where $\zeta(.)$ is the Riemann function and $\gamma_k = \gamma(k)$ with $\gamma(.)$ the autocovariance function of $\{Y_t\}$.

We are interested in the behavior of $\widehat{\theta} = \sqrt{T}\,(\widehat{\mu}-\mu, \widehat{\sigma}-\sigma, \widehat{\alpha}-\alpha)'$ with $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha}$ defined in (8) and (9). First we are going to study $\widetilde{\theta} = \sqrt{T}\,(\widehat{\mu}-\mu, \widehat{\sigma}-\sigma, \widetilde{\alpha}-\alpha)'$ where

$$\widetilde{\alpha} = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(Y_tY_{t+1} - \overline{Y}^2\right)}{\frac{1}{T}\sum_{t=1}^{T}(Y_t - \overline{Y})^2}.$$

We combine Lemma 5 with the delta method applied to the functions

$$\phi_1(\mathbf{z}) := z_1 - \frac{\delta\sqrt{6}}{\pi}\sqrt{z_2 - z_1^2},\quad \phi_2(\mathbf{z}) := \frac{\sqrt{6}}{\pi}\sqrt{z_2 - z_1^2} \text{ and } \quad \phi_3(\mathbf{z}) := \frac{z_3 - z_1^2}{z_2 - z_1^2}$$

where $\mathbf{z} = (z_1, z_2, z_3)'$. The partial derivatives of ϕ_1, ϕ_2 and ϕ_3 exist, they are continuous and $H(\mu, \sigma, \alpha)$ defined as

$$H(\mu,\sigma,\alpha) = \frac{\partial \phi_\mu}{\partial z_\nu}(\mathbf{z})\bigg|_{(z_1,z_2,z_3)=(0,\mathbb{E}(Y^2),\mathbb{E}(Y_0Y_1))}, \ \mu,\ \nu = 1,2,3$$

is equal to $(2\gamma_0)^{-1}\begin{pmatrix} 2\gamma_0 & -\sigma\delta & 0 \\ 0 & \sigma & 0 \\ 0 & -2\alpha & 2 \end{pmatrix}$ and has rank 3. Therefore we obtain by the delta method the asymptotic normality of $\widetilde{\theta}$ with a mean vector equals to $\mathbf{0}$ and a covariance matrix equals to $H(\mu,\sigma,\alpha)\boldsymbol{\Sigma}_{\mathbf{Y}}H(\mu,\sigma,\alpha)'$ where $\boldsymbol{\Sigma}_{\mathbf{Y}}$ is defined in Lemma 5. After computations we obtain the matrix given in (10). Since it is possible to show that $\sqrt{T}(\widehat{\alpha} - \widetilde{\alpha}) = o_{\mathbb{P}}(1)$ with $\widehat{\alpha}$ defined in (9), we conclude that the vector $\widehat{\theta}$ converges to the same limiting distribution.

It remains to give the proof of Lemma 5.

Proof of Lemma 5. In a first part we compute the elements of $\boldsymbol{\Sigma}_{\mathbf{Y}}$. Since it is possible to show that $\mathbb{E}(\varepsilon_t^4) = \eta\sigma_\varepsilon^4 < \infty$ with $\eta = 3 + \frac{12}{5}\frac{1+\alpha^2}{1-\alpha^2}$ and since $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$ we can directly deduce the asymptotic covariance matrix of the vector $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Y_t^2,\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}Y_tY_{t+1}\right)'$ using Proposition 7.3.1 in Brockwell and Davis (1987). Using similar techniques, we obtain

$$\lim_{T\to\infty} T\mathbb{V}ar\begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T}Y_t \\ \frac{1}{T}\sum_{t=1}^{T}Y_t^2 \\ \frac{1}{T}\sum_{t=1}^{T}Y_tY_{t+1} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=-\infty}^{\infty}\gamma_k & \sum_{k=-\infty}^{\infty}\mathbb{E}(Y_0Y_k^2) & \sum_{k=-\infty}^{\infty}\mathbb{E}(Y_0Y_kY_{k+1}) \\ \sum_{k=-\infty}^{\infty}\mathbb{E}(Y_0Y_k^2) & (\eta-3)\gamma_0^2 + 2\sum_{k=-\infty}^{\infty}\gamma_k^2 & (\eta-3)\gamma_0\gamma_1 + 2\sum_{k=-\infty}^{\infty}\gamma_k\gamma_{k+1} \\ \sum_{k=-\infty}^{\infty}\mathbb{E}(Y_0Y_kY_{k+1}) & (\eta-3)\gamma_0\gamma_1 + 2\sum_{k=-\infty}^{\infty}\gamma_k\gamma_{k+1} & (\eta-3)\gamma_1^2 + \sum_{k=-\infty}^{\infty}(\gamma_k^2 + \gamma_{k+1}\gamma_{k-1}) \end{pmatrix}.$$

Since $\mathbb{E}(\varepsilon_t^3) = 2\zeta(3)(1/\alpha^3 - 1)$ and $\gamma_k = \gamma_0 \alpha^{|k|}$, the expression for $\mathbf{\Sigma_Y}$ given in (20) follows.

In the second part of the proof of Lemma 5, we have to show the convergence to a normal distribution. To this aim, we define the truncated sequence as follows $\mathbf{W_{m,t}} = (Y_{m,t}\,,\, Y_{m,t}^2 - \mathbb{E}Y_{m,t}^2\,,\, Y_{m,t}Y_{m,t+1} - \mathbb{E}Y_{m,t}Y_{m,t+1})^{'}$ where $Y_{m,t} = \sum_{j=-m}^{m} \psi_j \varepsilon_{t-j}$ and ψ_j as previously. The idea is to prove first the asymptotic normality of $T^{-1/2}\sum_{t=1}^{T} \mathbf{W_{m,t}}$ and then to let m tend to infinity. To this aim, we have to show that any linear combination of the three components of $T^{-1/2}\sum_{t=1}^{T} \mathbf{W_{m,t}}$ is Gaussian. Note that, for any $\lambda \in \mathbb{R}^3$, the sequence $\{\lambda \mathbf{W_{m,t}}\}$ is strictly stationary $(2m+1)$-dependent. Moreover $\lim_{T\to\infty} T^{-1}\mathbb{V}ar(\sum_{t=1}^{T} \lambda^{'}\mathbf{W_{m,t}}) = \lambda^{'}\mathbf{\Sigma_{Y_m}}\lambda$ where $\mathbf{\Sigma_{Y_m}}$ is defined as

$$\begin{pmatrix} \sum_{k=-\infty}^{\infty} \gamma_{m,k} & \sum_{k=-\infty}^{\infty} \mathbb{E}(Y_{m,0}Y_{m,k}^2) & \sum_{k=-\infty}^{\infty} \mathbb{E}(Y_{m,0}Y_{m,k}Y_{m,k+1}) \\ \sum_{k=-\infty}^{\infty} \mathbb{E}(Y_{m,0}Y_{m,k}^2) & (\eta-3)\gamma_{m,0}^2 + 2\sum_{k=-\infty}^{\infty} \gamma_{m,k}^2 & (\eta-3)\gamma_{m,0}\gamma_{m,1} + 2\sum_{k=-\infty}^{\infty} \gamma_{m,k}\gamma_{m,k+1} \\ \sum_{k=-\infty}^{\infty} \mathbb{E}(Y_{m,0}Y_{m,k}Y_{m,k+1}) & (\eta-3)\gamma_{m,0}\gamma_{m,1} + 2\sum_{k=-\infty}^{\infty} \gamma_{m,k}\gamma_{m,k+1} & (\eta-3)\gamma_{m,1}^2 + \sum_{k=-\infty}^{\infty}(\gamma_{m,k}^2 + \gamma_{m,k+1}\gamma_{m,k-1}) \end{pmatrix}$$

with $\gamma_{m,k} = \gamma_m(k)$ where $\gamma_m(.)$ is the autocovariance function of $\{Y_{m,t}\}$. Therefore we can directly apply Theorem 6.4.2 in Brockwell and Davis (1987) and we obtain that

$$T^{-1/2}\sum_{t=1}^{T} \lambda^{'}\mathbf{W_{m,t}} \xrightarrow{d} \Theta_m \text{ with } \Theta_m \sim \mathcal{N}\left(\mathbf{0}, \lambda^{'}\mathbf{\Sigma_{Y_m}}\lambda\right)$$

for all vectors $\lambda \in \mathbb{R}^3$ such that $\lambda^{'}\mathbf{\Sigma_{Y_m}}\lambda > 0$. Consequently

$$T^{-1/2}\sum_{t=1}^{T} \mathbf{W_{m,t}} \xrightarrow{d} \Omega_m \text{ with } \Omega_m \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_{Y_m}}). \tag{21}$$

The last step in the proof of this lemma is to show that the asymptotic normality described in (21) remains true if $\mathbf{W_{m,t}}$ is replaced by $\mathbf{W_t} = (Y_t\,,\, Y_t^2 - \mathbb{E}Y_t^2\,,\, Y_tY_{t+1} - \mathbb{E}Y_tY_{t+1})^{'}$.

The idea is to derive the result for $\mathbf{W_t}$ by letting $m \to \infty$. Using mainly the convergence dominated theorem, it is easy to show that $\mathbf{\Sigma_{Y_m}}$ converges to $\mathbf{\Sigma_Y}$ as m tends to infinity, which entails that $\Omega_m \xrightarrow{d} \Omega$ as m tends to infinity with $\Omega \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_Y})$.

The proof can now be completed by an application of Proposition 6.3.9 in Brockwell and Davis (1987) as in the proof of Proposition 7.3.3 of the same book. To this aim, we have to check one condition. It has already been proved in Proposition 7.3.3 that for $p = 0, 1$, the following limit

$$\lim_{m\to\infty} \limsup_{T\to\infty} \mathbb{P}\left(\sqrt{T}\left|\frac{1}{T}\sum_{t=1}^{T}(Y_{m,t}Y_{m,t+p} - \mathbb{E}(Y_{m,t}Y_{m,t+p})) - \frac{1}{T}\sum_{t=1}^{T}(Y_tY_{t+p} + \mathbb{E}(Y_tY_{t+p}))\right| > \varepsilon\right)$$

is equal to 0. Similarly, we have to show that

$$\lim_{m\to\infty} \limsup_{T\to\infty} \mathbb{P}\left(\sqrt{T}\left|\frac{1}{T}\sum_{t=1}^{T} Y_{m,t} - \frac{1}{T}\sum_{t=1}^{T} Y_t\right| > \varepsilon\right) = 0. \tag{22}$$

Using Chebychev's inequality

$$\begin{aligned}
&\mathbb{P}\left(T^{1/2}\left|T^{-1}\sum_{t=1}^{T} Y_{m,t} - T^{-1}\sum_{t=1}^{T} Y_t\right| > \varepsilon\right) \\
&\le \varepsilon^{-2} T \mathbb{V}ar\left(T^{-1}\sum_{t=1}^{T} Y_{m,t} - T^{-1}\sum_{t=1}^{T} Y_t\right) \\
&= \varepsilon^{-2} T\Bigg[\mathbb{V}ar\left(T^{-1}\sum_{t=1}^{T} Y_{m,t}\right) + \mathbb{V}ar\left(T^{-1}\sum_{t=1}^{T} Y_t\right) \\
&\qquad\qquad - 2\mathbb{C}ov\left(T^{-1}\sum_{t=1}^{T} Y_{m,t}, T^{-1}\sum_{t=1}^{T} Y_t\right)\Bigg].
\end{aligned}$$

The two variances and the covariance involved in this bound can easily be computed and we obtain

$$\lim_{m\to\infty} \limsup_{T\to\infty} \varepsilon^{-2} T \mathbb{V}ar\left(T^{-1}\sum_{t=1}^{T} Y_{m,t} - T^{-1}\sum_{t=1}^{T} Y_t\right) = 0.$$

This establishes (22), achieves the proof of Lemma 5 and also the one of Proposition 4. $\square$

References

Andrews B, Calder M, Davis RA. 2008. Maximum likelihood estimation for α-stable autoregressive processes. *Submitted*.

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

Bingham NH, Goldie CM, Teugels JL. 1987. *Regular variation*; Cambridge University Press, Cambridge.

Breidt FJ, Davis RA. 1991. Maximum likelihood estimation for noncausal autoregressive processes. *J. Multivar. Anal.*, **36**:175-198.

Brockwell PJ, Davis RA. 1987. *Time series: theory and methods*; Springer Series in Statistics, Springer-Verlag, New York.

Coles S. 2001. *An introduction to statistical modeling of extreme values*; Springer Series in Statistics, Springer-Verlag, London.

Coles S, Heffernan J, Tawn J. 1999. Dependence measures for extreme value analyses. *Extremes*, **2**:339-65.

Crowder MJ. 1989. A multivariate distribution with Weibull components. *J. Roy. Statist. Soc. B*, **51**:93-108.

Davis RA, Resnick SI. 1989. Basic Properties and Prediction of Max-Arma Processes. *Adv. Appl. Probab.*, **21**:781-803.

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

Fisher RA, Tippett LHC. 1928. Limiting forms of the frequency distribution in the largest particle size and smallest member of a sample. *Proc. Camb. Phil. Soc.*, **24**:180-190.

Fougères A-L. 2004. *Multivariate extremes.* In Extreme Values in Finance, Telecommunications, and the Environment, Finkenstadt B and Rootzén H (eds); Boca Raton, CRC, Boca Raton, Florida; 373-388.

Fougères A-L, Nolan JP, Rootzén H. 2008. Models for dependent extremes using stable mixtures. *Scand. J. Statist.*, to appear.

Gnedenko B. 1943. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.*, **44**:423-453.

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

Hougaard P. 1986. A class of multivariate failure time distributions. *Biometrika*, **73**:671-678.

Joe H. 1997. *Multivariate models and dependence concepts*; Monographs on Statistics and Applied Probability, 73; Chapman & Hall, London.

Katz RW. 2002. Sir Gilbert Walker and a Connection Between El Nino and Statistics. *Statistical Science*, **17**:97-112.

Naveau P, Guillou A, Cooley D, Diebolt J. 2008. Modeling Pairwise Dependence of Maxima in Space. *Biometrika* (in press).

Tawn JA. 1990. Modelling multivariate extreme value distributions. *Biometrika*, **77**:245-253.

Zhang Z, Smith RL. 2008. On the estimation and application of max-stable processes. *Submitted.*

Zolotarev VM. 1986. *One-Dimensional Stable Distributions.* American Mathematical Society Translations of Mathematical Monographs; Vol. 65. American Mathematical Society, Providence, RI. Translation of the original 1983 (in Russian).