STATE-SPACE MODELS FOR MAXIMA PRECIPITATION

Philippe NAVEAU [1], Paul PONCET [2]

ABSTRACT

A very active research field in atmospheric sciences is centered around the modeling of weather extremes. This is mainly due to the large economic and human impacts of such extreme events. In this paper, we focus on the statistical temporal modeling of precipitation maxima because daily and monthly maxima have been recorded for many decades and at various sites.

Our goal is to propose two new state-space models whose distributional foundations lie in Extreme Value Theory (EVT). Our first model takes advantage of max-stable processes, previously studied by Davis and Resnick (1989), among others. It can be viewed as a "translation" of the Gaussian linear Kalman filter into a Fréchet-type world in which the classical addition $a + b$ has been replaced by the max operator $a \vee b = \max(a, b)$ and the noise component is from a heavy-tailed distribution instead of being Gaussian. Our second state-space model is built from the mixture extremes framework proposed by Fougères *et al.*, (2006). Its strong points are its flexibility and richness with respect to applications. In addition to addressing the theoretical questions brought by our models, the main benefit of introducing them is to propose simple and powerful connections between EVT and data assimilation communities. The latter term "data assimilation" regroups statistical/dynamical techniques extensively used in climate studies. These procedures involve the combination of observational data with the underlying dynamical principles governing the physical system under observation. Hence, improving our knowledge about the representation of extremes in a state-space model framework is of strong interest from a data assimilation point of view.

Keywords : Data assimilation, Kalman filter, Extreme Value Theory, Generalized Extreme Value distribution, max-stable state-space model, GEV state-space model.

RÉSUMÉ

La modélisation des événements climatiques extrêmes est aujourd'hui un champ de recherches particulièrement actif, notamment de par l'importance de leurs impacts économiques et sociaux. Dans cet article nous portons notre attention sur la modélisation statistique des maxima de précipitations, car de telles données sont disponibles aux pas de temps journalier et mensuel sur plusieurs décennies et en de nombreux sites.

Notre but est de proposer deux nouveaux modèles à espace d'états dont les fondations probabilistes reposent sur la théorie des valeurs extrêmes (EVT en anglais). Notre premier modèle tire parti des processus max-stables, étudiés entre

1. Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, France ; naveau@lsce.cnrs-gif.fr

2. Université Pierre et Marie Curie, Paris, France ; paulponcet@yahoo.fr

autres par Davis and Resnick (1989). Il peut être vu comme la transposition du filtre de Kalman linéaire et gaussien à un monde de type Fréchet, où l'addition $a + b$ est remplacée par l'opération maximum $a \vee b = \max(a, b)$, et où les bruits sont à queue lourde au lieu d'être gaussiens. Notre second modèle se base sur le modèle de mélange pour les extrêmes proposé par Fougères *et al.* (2006). Sa flexibilité et sa richesse en termes d'applications en sont un atout essentiel. En plus des interrogations théoriques que suscitent nos modèles, leur principal intérêt est de créer des liens simples et puissants entre la EVT et le domaine de l'assimilation de données. Ce dernier regroupe des techniques statistiques et dynamiques abondamment utilisées dans les études climatiques. Ces procédures nécessitent de combiner d'une part des données issues d'observations et d'autre part les principes dynamiques sous-jacents qui gouvernent le système physique à l'œuvre. C'est pourquoi l'amélioration de notre connaissance des extrêmes et de leur représentation dans le cadre d'un modèle à espace d'états est d'un intérêt tout particulier du point de vue de l'assimilation de données.

Mots-clés : Assimilation de données, filtre de Kalman, théorie des valeurs extrêmes, distribution généralisée des valeurs extrêmes, modèle à espace d'états max-stable, modèle à espace d'états GEV.

1. Introduction

1.1. Statistical modeling of extreme events

Currently there is an increasing research activity in the area of climate extremes because they represent a key manifestation of complex systems and they have an enormous impact on economic and social human activities. Our understanding of the mean behavior of climate and its normal variability has been improving significantly during the last decades. In comparison, our knowledge of extreme events frequency and amplitude is much more incomplete and partial. Before motivating this work in detail, we first need to recall the basic principles of EVT.

EVT is the branch of statistics which describes the behavior of the largest observations in a data set and it has a long history going back to 1928 (Fisher and Tippett, 1928). It has been applied to a variety of problems in finance (Embrechts *et al.*, 1997) and hydrology (Katz *et al.*, 2002). Surprisingly, its application to climate studies has been fairly recent, e.g. Kharin and Zwiers (2000). Similarly to many results in mathematics and physics, a stability property is the key element to understand EVT. One may wonder which types of distribution are closed for maxima (up to affine transformations), i.e. which family law allows that the observations and their maximum have the same type of distribution. For example, the Gaussian distribution does not satisfy this condition, since the maximum of Gaussian variables is not Gaussian (although the sum is and that's why the Gaussian law belongs to the family of stable laws). The solution of our question is called the group of *max-stable distributions* which is composed of three types of distribution: Fréchet, Weibull and Gumbel, see Embrechts *et al.* (1997) for more details on

EVT. These three types can be summarized by the GEV(μ, σ, ξ) distribution defined by

$$F(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]_{+}^{-1/\xi}\right\} \tag{1}$$

where $a_+ = a \vee 0$, i.e. $\max(a, 0)$. The parameters $\sigma > 0$ and μ represent scale and location, respectively. The shape parameter ξ describes the tail behavior of the distribution. If ξ is negative (Weibull type), the upper tail is bounded. If ξ is zero [1] (Gumbel type), all moments are finite. If ξ is positive (Fréchet type), the upper tail is still unbounded but higher moments eventually become infinite. These three cases are termed "bounded", "light-tailed", and "heavy-tailed", respectively. The justification for the GEV distribution arises from an asymptotic argument, for, as the sample size increases and under mild conditions, the distribution of the sample maximum asymptotically follows either a Fréchet, Weibull, or Gumbel distribution. Its flexibility to describe all three types of tail behavior makes it a universal tool for modeling block maxima. When considering the difference between the classical Gaussian and the GEV distributions, it is important to notice that the former does not encompass heavy tails and there is much evidence in the literature that the distributions of hydrological and climatological variables are heavy tailed, especially precipitation. Although it can be difficult to determine from one single site unless the record is relatively long, the distribution of maximum precipitation amount (e.g., daily) appears consistently to have a heavy tail (Katz *et al.*, 2002). Regional analyses of precipitation extremes, in which the shape parameter is constrained to be constant within the area, provide clear evidence of heavy tails (Buishand, 1991). To illustrate the heavy-tailed distribution, we comment on a simple but typical type of extremes, daily maxima precipitation. For example, we can look at the city of Nîmes in France in Figure 1. The x-axis corresponds to the days with positive precipitation (this explains that we have less than 365 points per year) and the y-axis represents the maximum of precipitation for these days. Such a time series exhibits the specific characteristics of heavy-tailed extremes. Its values can sometimes be very large, and consequently they cannot be well represented by Gaussian processes.

1.2 Statistical issues when assimilating extreme events

The fundamental problem of data assimilation may be simply stated as follows: given the state of atmospheric variables at one time, what is the state at a later time? To answer such a question, the basic laws and principles of physics, biology and chemistry are classically used to describe the evolution of the state variables forward in time. From a probabilistic point of view, understanding the present and future state variables is intimately linked to state-space modeling.

1. This case corresponds to the limiting case $\xi \to 0$ in (1), i.e. $F(x) = \exp\left\{-\exp(-\frac{x-\mu}{\sigma})\right\}$.

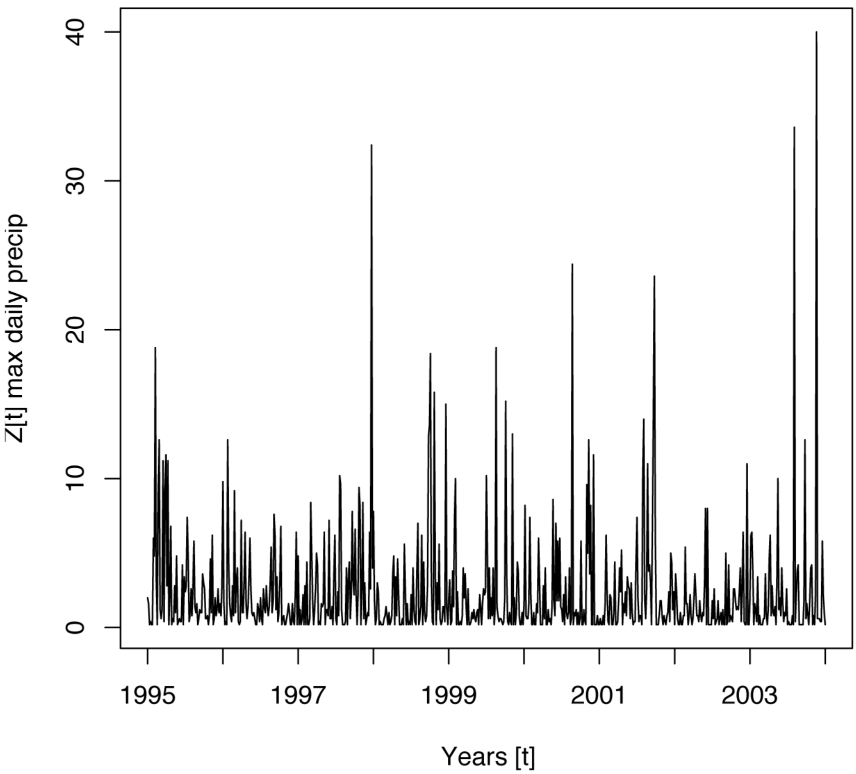

FIG 1. — Daily maxima of precipitation amounts recorded every three hours in Nîmes, France. The x-axis corresponds to the days with positive precipitation and the y-axis represents the maximum of precipitation for these days.

In the previous section, we recalled that a mathematical based theory states that maxima should be adequately represented by a GEV distribution and numerous studies have confirmed this approach, e.g. Katz *et al.* (2002); Kharin and Zwiers (2000); Koutsoyiannis and Baloutsos (2000); Hosking *et al.* (1985). Consequently, it is fundamental to assess if classical data assimilation techniques can integrate this distributional constraint within the available dynamical information, i.e., how to link GEV distributed observations with state-space dynamics. Classically, data assilimilation in geosciences relies on Gaussian distributions for observations and for prior information, e.g., Rodgers (2000). In order to preserve the Gaussian hypothesis, some studies filter or threshold data to remove large observations, e.g., Chevallier *et al.* (2004). When (limited) non-Gaussian features of the distributions are present, e.g., Dharssi *et al.* (1992) and Evensen (1994), they are not cycled in the Kalman filter, which reduces the optimality of the assimilation system. Our study aims at extending the data assimilation theoretical framework. To include heavy tailed distributions, we will propose simple and innovative models based on EVT that can better handle this complex topic of dealing with extreme values. More precisely, the fundamental question that we would like to address in this work is the following one. If observations, say Y_t (e.g. daily maxima precipitation amounts recorded every three hours, see Figure 1), follow a GEV distribution (as expected from the theory and justified by past studies), what are the dynamical models that can produce such GEV outputs?

To start answering this question, we have to address the issue of modeling temporal dependence that represents an essential component of data assimilation techniques. To illustrate the problem, we can look at one of the simplest temporal structures, the classical auto-regressive model of order one

$$Z_t = aZ_{t-1} + \text{noise}_t$$

where the sample noise is i.i.d. (independently and identically distributed). With this simple model, we can compare two cases of noise: either (a) Gaussian or (b) GEV with a heavy tail. In Figure 2, we fix $a = 0.35$ and we explore two cases of noise.

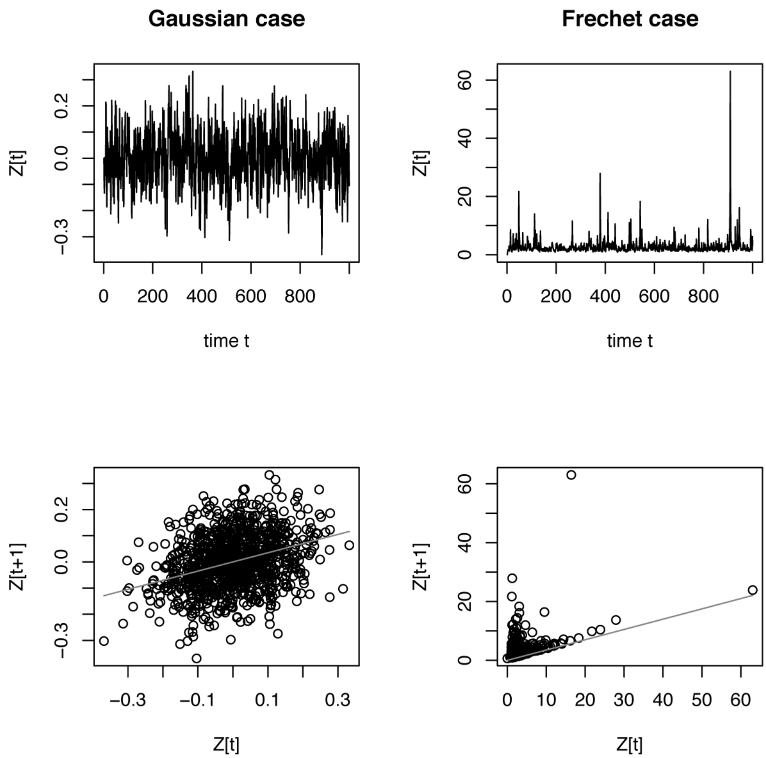

Fig 2. — Comparing AR(1) model with Gaussian (left panels) and Fréchet noises (right panels). The straight gray lines (lower panels) correspond to the slope of the AR model.

In the left panels, we assume that the noise follows a standardized Gaussian distribution. In contrast, the right panels are obtained by plugging in a heavy-tailed noise (here a Fréchet random variable with shape parameter equal to 1.9). The upper panels show a realization for the two cases. While it is possible to simulate very large extremes with the Fréchet noise, the Gaussian model is very limited in generating values greater than twice the standard deviation. The lower panels also indicate another strong difference between the two models. For the Gaussian noise, the temporal dependence obtained by plotting the values of Z_t at time t versus its values at time $t+1$ is rather linear (with the gray slope equal to a). For the Fréchet noise, the lower right panel tells us a different story. Very large values occur in a burst and they dominate the temporal structure for a short period. It also means that the classical statistical reasoning based on the covariance function may not be adapted for capturing the temporal dependence in maxima behavior. Overall, comparing Figures 1 and 2 indicates that observed temporal precipitation patterns could be better mimicked with the Fréchet model than the Gaussian one. Although this toy example can generate interesting non-Gaussian responses with the right choice of noise, it has many limitations. First, its simplicity does not allow the practitioner to blend observations with larger scale atmospheric information, and this constraint is fundamental for data assimilation. Secondly, it is rather difficult to derive the statistical properties of the process Z_t because the GEV distribution is not stable for the sum, i.e. Z_t does not follow a GEV even when the noise is GEV distributed.

2. Data-assimilation for maxima

2.1. The classical state-space model

Kalman filtering and state-space modeling have been at the core of past statistical data assimilation procedures. Given the state of the atmosphere at time t, the question of interest for climatologists is to determine the state of the atmosphere at time $t+1$ given that observations are available at time t. To perform such a task, it is classically assumed that the underlying dynamical principles governing the system are known and the link between the observations and the state are also given. Mathematically, this corresponds to a state-space formulation in which the state equation drives the dynamics of the system and the observational equation integrates the measurements with the state variables. *State-space models* have been widely studied in statistics, e.g., West and Harrison (1997) and Shepard (1994). These models have become a powerful tool for modeling and forecasting dynamical systems and they have been used in a wide range of disciplines such as biology, economics, engineerings, climatology and statistics (Guo *et al.*, 1999; Naveau *et al.*, 2005). The basic idea of a state-space model is that the dimensional vector of observations Y_t at time t is generated by two equations, the *observational* and the *system* equations. The first equation describes how the observations vary as a function of the unobserved state vector X_t:

$$Y_t = F_t(X_t, \epsilon_t),$$

where ϵ_t represent a noise and F_t is a function that is assumed to be known. The temporal dynamical structure is incorporated via the system equation:

$$X_t = G_t(X_{t-1}, \eta_t),$$

where η_t represents a noise and G_t is a function that is assumed to be known. In practice, such observational and state equations are too complex for real applications and extra assumptions are classically added. The three most common ones are: (a) independence between and within observation and state noises, (b) Gaussian noises, and (c) linearity for the observational and state equations, i.e.

$$Y_t = F_t X_t + \epsilon_t \text{ and } X_t = G_t X_{t-1} + \eta_t. \tag{2}$$

If the sequence $\{Y_t\}$ represents the observation of precipitation maxima, and if one assumes that Y_t follows a GEV distribution, then it is, by design, impossible to work with (2) under both assumptions (a) and (b). Of course, models that have the advantage of obeying the distributional GEV constraint have to break one of the three assumptions. Basically, assumption (a) is needed to insure simple and manageable estimation procedures. Hence, we keep this assumption in this paper and propose alternatives to (b) and (c) in the next section.

Note that, to simplify the exposition and the computations, we assume in this paper that the dimension of all vectors is equal to one and that the sampling is equally spaced, $t = 1, ..., T$. While it is straightforward to remove the latter assumption if needed, going from the univariate to the multivariate case is much more challenging and further research is needed to resolve this issue.

2.2. Max-stable state-space models

Before defining our max-stable state-space models, we need to recall basic properties and ideas concerning max-stable laws. We can start with the simple max-stable auto-regressive model $Z_t = (aZ_{t-1}) \vee \epsilon_t$ where $a \geqslant 0$, Z_0 and ϵ_t are Fréchet distributed and independent. Such an auto-regressive model is max-stable because any combination of the type $\bigvee_i a_i Z_i$ with $a_i \geqslant 0$ remains in the same distributional class. Such a concept can be generalized (Davis and Resnick, 1989) and max-stable processes have been used in a variety of applications. For example, Helland and Nilsen (1976) considered a max-autoregressive model with a random coefficient to model deep water exchanges in a sill fjord. In their model, Z_t represents the annual density of resident water in the fjord basin and ϵ_t corresponds to the annual density of coastal water adjacent to the fjord. Using the operator $\vee$ allowed them to model a non-linear exchange of water. Max-autoregressive processes have also been taken advantage of for solving utility problems. Suppose that Z_t is the utility of the computer model currently held and ϵ_t is the utility of the new model. If ϵ_t is much larger than Z_t, then a switch is made, i.e. $Z_{t+1} = \epsilon_t$, otherwise the current utility falls by a certain percentage due to aging, i.e. $Z_{t+1} = aZ_t$. Because of its simplicity to model non-linear behaviors with intermittent bursts, such models have also been implemented to a variety of others applications, e.g. queueing and storage theory with abrupt changes in job loads.

But, to our knowledge, no one has yet developed a Kalman filter method for max-stable state-space models. In order to start filling this methodological gap, we propose the following max-stable state-space model. The observation equation becomes

$$Y_t = F_t X_t \vee \epsilon_t, \tag{3}$$

where ϵ_t represents an i.i.d. noise with a Fréchet margin, and $F_t > 0$. We also replace the sum operator by the max operator in the system equation:

$$X_t = G_t X_{t-1} \vee \eta_t, \tag{4}$$

where $G_t > 0$ and η_t represent an i.i.d. noise with a Fréchet margin. By design, if Y_t records a precipitation maximum that follows a Fréchet distribution, then (3) and (4) ensure that this distributional constraint is satisfied while introducing a dynamical temporal structure. Concerning the interpretation of (3), the variable ϵ_t should not be viewed as a measurement error, but rather as a source of strong variability. It could be interpreted as an unobservable physical variable that has the power to highly influence observed precipitation maxima. For example, one may think of the vertical velocity (denoted W in meteorology). The latter is very likely to be heavy-tailed, and strong vertical wind bursts in a column of moisture (advection) could create the condition for heavy rainfall (Wilson and Toumi, 2005). In other words, whenever a strong and intermittent wind burst occurs a switch is made (i.e. $Y_t = \epsilon_t$), otherwise the current state decreases proportionally to the quantity F_t. Of course, such an interpretation remains, at this stage, hypothetical and more collaboration with the atmospheric community is required to validate such a scheme. To illustrate the stochastic behavior of this system, Figure 3 shows the trajectories of X_t (vertical gray lines) and Y_t (solid line) in the case $G_t = 0.7$ and $F_t = 0.9$. As in a Kalman filter context, one of the main statistical questions is to determine how to estimate X_t given the past observations $(Y_1, \ldots, Y_t)$. One difficult issue to solve before answering such a question is that classical minimization schemes based on the mean and variance cannot be implemented because the assumption of Fréchet margins does not guarantee that these moments are finite. While it is beyond the scope of this paper to solve this complex issue, it is easy to derive bounds that provide valuable information on the almost sure trajectory of X_t. Indeed, a natural upper bound for X_t can be defined by

$$\overline{X}_t = \frac{Y_t}{F_t}, \text{ for all } t \geqslant 0. \tag{5}$$

One can also derive an almost sure lower bound. To do this, we first need to introduce the random times $T_0 = 0$ and

$$T_{j+1} = \inf \left\{ t > T_j : \frac{Y_t}{Y_{t-1}} \overset{\text{a.s.}}{=} \frac{F_t G_t}{F_{t-1}} \right\}, \text{ for } j = 0, 1, 2, \ldots \tag{6}$$

and then, we define sequentially $\underline{X}_0 = 0$ and for $t \geqslant 1$

$$\underline{X}_t = \begin{cases} \dfrac{Y_t}{F_t} & \text{, if } t = T_j \text{ for some } j, \\ G_t \underline{X}_{t-1} & \text{, otherwise.} \end{cases} \tag{7}$$

Note that the events $\epsilon_t/\epsilon_{t-1} = F_t G_t / F_{t-1}$ have a null probability of occuring. Figure 4 displays the upper and lower bounds defined by (5) and (7), respectively the solid and dotted lines. As in Figure 3, the vertical grey lines represent the trajectory of X_t. The vertical dotted lines represent the random times defined by (6). Note that the upper and lower bounds intersect at these random times. This means that the value of X_t is completely determined at these occurrences. Overall, this figure clearly indicates that these bounds provide useful feature about the hidden values of (X_t).

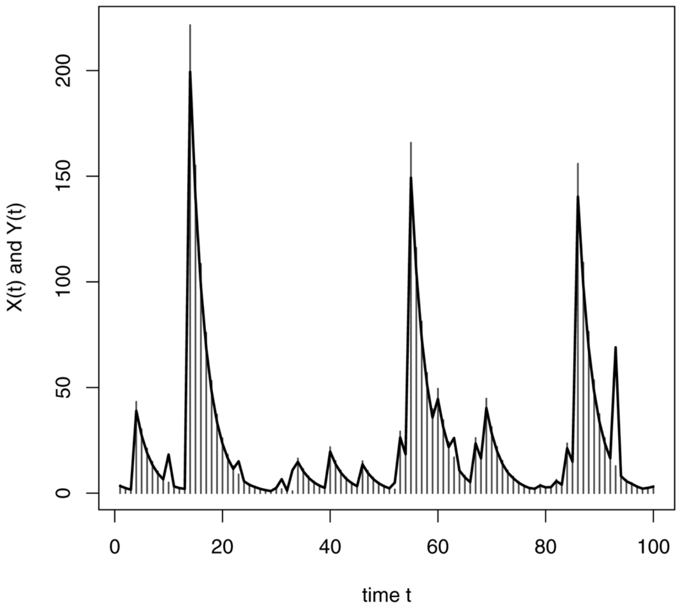

FIG 3. — One realization of X_t (vertical grey lines) and Y_t (solid line) obtained from the system (3) and (4) for which $G_t = 0.7$ and $F_t = 0.9$. The x-axis represents the time $t = 0, \ldots, 100$ and the y-axis corresponds to the values of X_t and Y_t.

In our examples and for the model defined by (3) and (4) we opted to work with Fréchet margins. This choice was made since precipitation maxima are heavy-tailed, and it is also convenient from a theoretical point of view, although our model could be easily extended to GEV margins.

Despite the fact that the system defined by (3) and (4) has the advantage of being simple to define and of introducing an interesting temporal dependence, it has a substantial limitation. Some researchers in climatology may have

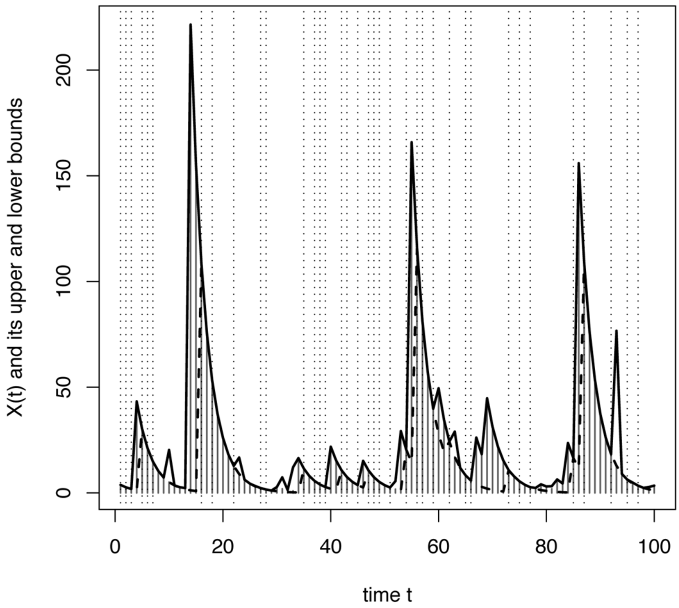

FIG 4. — Upper and lower bounds defined by (5) and (7), respectively the solid and dotted lines, of the realization of X_t (grey vertical lines) from Figure 3. The vertical dotted lines correspond to the random times defined by (6).

concerns about working with the max operator $a \vee b = \max(a, b)$ and they would much prefer to deal with an additive model. To resolve this issue and to provide alternatives to max-stable processes, we propose in the next section an additive-type model based on the Gumbel distribution which also belongs to the GEV family.

2.3. The GEV state-space model

In this section we recall the main result obtained by Fougères *et al.* (2006) about stable linear processes driven by Gumbel dynamics. It stems from Tawn (1990) and it is based on the properties of stable laws [2]. In climate studies, it is often overlooked that the Gaussian law is *only one* member of this stable family. Some heavy tailed distributions like the Cauchy law can also be closed under summation. In our case, we work with a sub-class of stable variables. Let S be a positive α-stable random variable whose Laplace transform is such that $\mathbb{E}[\exp(-xS)] = \exp(-x^\alpha)$, for all $x > 0$ and for some $\alpha \in (0, 1)$. This

2. 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 $c_1S_1 + c_2S_2 \stackrel{d}{=} aS + b$ where S_1, S_2 are iid copies of S and where $\stackrel{d}{=}$ denotes equality in distribution.

definition allows us to recall a proposition found in Fougères *et al.* (2006). If

$$Y_t = F_t \log \left(\sum_{a \in A} c_{t,a} S_a \right) + \epsilon_t, \text{ with } t = 0, \ldots, T, \tag{8}$$

where $\{c_{t,a}\}$ are non-negative constants, $\{S_a, a \in A\}$ are independent positive α-stable variables, ϵ_t follows an i.i.d. GEV$(\mu_t, F_t, 0)$ distribution (that is a Gumbel distribution), and all variables are mutually independent, then we have

$$\mathbb{P}(Y_t \leqslant x_t, t = 0, \ldots, T) = \prod_{a \in A} \exp \left(-(\sum_{t=0}^{T} c_{t,a} e^{-\frac{x_t - \mu_t}{F_t}})^{\alpha} \right). \tag{9}$$

Although rather complex at first sight, this result tells us that Equation (8) can generate observations that follow a Gumbel distribution, and even more importantly, the multivariate distribution of the vector $(Y_0, \ldots, Y_T)$ can be explicitly written, see Equation (9). Hence, Equation (8) provides a solid foundation to construct state-space models for extremes, as we shall see.

From Equation (8) and the structure of the classical state-space model described in Section 2.1, it is natural to propose the following Gumbel state-space model. The observation equation becomes

$$Y_t = F_t \ \log X_t + \epsilon_t, \tag{10}$$

where ϵ_t represents an i.i.d. Gumbel noise and F_t is a non-negative scalar. The system equation has now the following form

$$X_t = G_t X_{t-1} + S_t, \tag{11}$$

where G_t is non-negative scalar and S_t represents an i.i.d. positive α-stable noise. Because of its stability, the vector X_t also follows a positive α-stable distribution. More precisely, we can write that

$$X_t = \sum_{i=0}^{t} c_{t,i} S_i,$$

where $c_{t,t} = 1$ and $c_{t,i} = \prod_{k=i+1}^{t} G_k$ for $i < t$. This implies

$$Y_t = F_t \ \log \left(\sum_{i=0}^{t} c_{t,i} S_i \right) + \epsilon_t.$$

This latter form of Y_t corresponds to Equation (8). Consequently, we know exactly the distribution of $(Y_0, \ldots, Y_T)$ through Equation (9), as well as the distribution of $(X_0, \ldots, X_T)$. Hence, the system of equations (10) and (11) provides a state-space model in which the observations and the state vector can be expressed in a closed Gumbel form. For the first time (to our knowledge), such a structure offers an additive and flexible way to perform data assimilation on maxima within EVT.

The estimation of X_t given the observations is a problem that we have not solved yet. A possible strategy is to extend the work of Stuck (1977), who found the Kalman filter steps for symmetric α-stable laws, within our Gumbel state-space model context. More precisely, suppose that we have the additive model defined by (2) where ϵ_t and η_t are i.i.d symmetric α-stable noises. For this case, Stuck derived the best linear estimate of X_t with respect to a distance called dispersion. To do so, he minimized $\gamma_t := \gamma(X_t - \hat{X}_t)$, where $X_t - \hat{X}_t$ denotes the error and $\gamma(Z)$ is the positive coefficient such that $E[\exp(ixZ)] = \exp(-\gamma(Z)|x|^\alpha)$, $x \in \mathbb{R}$, for any symmetric α-stable variable Z. His results generalized the classical Kalman filter in the Gaussian case ($\alpha = 2$). However, extending Stuck's approach to our model is not straightforward and more research is needed in this direction.

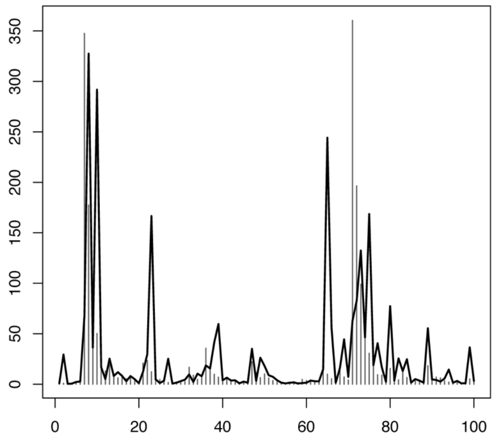

FIG 5. — One realization of X_t (vertical grey lines) and $\exp(Y_t)$ (solid line) obtained from the system (10) and (11) for which $\alpha = 0.6$, $F = 0.8$ and $G = 0.5$. The x-axis represents the time $t = 0, \ldots, 100$ and the y-axis corresponds to the values of X_t and Y_t.

3. Conclusion

Although Gaussian linear state-space models have been very successful in classical time series analysis, they are not adapted to represent temporal changes of maxima. In this context, our main motivation was to introduce two state-space models that are in compliance with EVT. Such models have strong potential but still, a lot of work remains to understand their properties, drawbacks and qualities. More precisely, we identify at least two questions for future research:

1. Given the observations Y_t and the dynamics of our model (i.e. F_t and G_t), what is the optimal procedure to estimate X_t in equations (3) and (10) at each time step? This question amounts to finding the Kalman filter steps for our models.

2. What is the error when applying a classical Gaussian Kalman filter to maxima?

Acknowledgments: We would like to thank an anonymous reviewer for his/her insightful comments that have greatly improved the clarity of this work. This research was supported by the E2C2 European project, the NSF-GMC (ATM-0327936) grant and by the ANR-AssimilEx project.

References

Buishand T.A. (1991) Extreme rainfall estimation by combining data from several sites. *Hydrolog. Sci. J.* **36**(4):345-365.

Chevallier F., Lopez P., Tompkins A.M., Janiskov M., and Moreau E. (2004) The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Quart. J. Roy. Meteor. Soc.* **130**:917-932.

Coles S.G. (2001) *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics. Springer-Verlag London Ltd., London.

Cooley D., Naveau P., Jomelli V., Rabatel A., and Grancher D. (2005) A bayesian hierarchical extreme value model for lichenometry. *Environmetrics* **16**:1-20.

Davis R.A. and Resnick S.I. (1989) Basic Properties and Prediction of Max-Arma Processes. *Adv. Appl. Probab.* **21**:781-803.

Dharssi I., Lorenc A.C., and Ingleby N.B. (1992) Treatment of gross errors using maximum probability theory. *Quart. J. Roy. Meteor. Soc.* **118**(507), Part B:1017-1036(20).

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

Evensen G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **5**:10143-10162.

Fisher R.A. and Tippett L.H.C. (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge. Philos. Soc.* **24**:180-190.

Fougères A.L., Nolan J.P., and Rootzén H. (2006) Mixture Models for Extremes. *Submitted.*.

Guo W., Wang Y., and Brown M. (1999) A signal extraction approach to modeling hormone time series with pulses and a changing baseline. *J. Amer. Stat. Assoc.* **94**:746-756.

Helland I.S. and Nilsen T.S. (1976) On a General Random Exchange Model. *J. Appl. Probab.* **13**(4):781-790.

Hosking J.R.M., Wallis J.R., and Wood E.F. (1985) Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. *Technometrics*, **27**(3):251-261.

Katz R.W., Parlange M.B., and Naveau P. (2002) Statistics of extremes in hydrology. *Adv. Water Resour.* **25**:1287-1304.

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

Koutsoyiannis D. and Baloutsos G. (2000) Analysis of a Long Record of Annual Maximum Rainfall in Athens, Greece, and Design Rainfall Inferences. *Natural Hazards* **22**(1):31-51.

Naveau P., Genton M.G., and Shen X. (2005) A skewed Kalman filter. *J. Multiv. Anal.* **94**(2):382-400.

Rodgers C.D. (2000) Inverse methods for atmospheric sounding. Theory and practice. *Series on Atmospheric, Oceanic and Planetery Physics*, Volume 2. World Scientific. Singapore-New Jersey-London-Hong-Kong.

Shephard N. (1994) Partially Non-Gaussian State-space Models. *Biometrika* **81**:115-131.

Stuck B.W. (1977) Minimum Error Dispersion Linear Filtering of Scalar Symmetric Stable Processes. *IEEE Trans. Automat. Contr.* **AC-23**(3):507-509.

Tawn J. (1990) Modelling multivariate extreme value distributions. *Biometrika* **77**:245-253.

West M. and Harrison J. (1997) *Bayesian forecasting and dynamic models.* Springer, New York.

Wilson P.S. and Toumi R. (2005) A fundamental probability distribution for heavy rainfall. *Geophys. Res. Lett.* **32**(14).