

Non-stationary extreme models and a climatic application

M. Nogaj¹, S. Parey¹, and D. Dacunha-Castelle²

¹EDF R&D, 6 Quai Watier, 78401 Chatou, France

²Department of Mathematics, University Paris-XI, Orsay, France

Received: 30 June 2006 – Revised: 7 December 2006 – Accepted: 5 June 2007 – Published: 25 June 2007

Abstract. In this paper, we study extreme values of non-stationary climatic phenomena. In the usually considered stationary case, the modelling of extremes is only based on the behaviour of the tails of the distribution of the remainder of the data set. In the non-stationary case though, it seems reasonable to assume that the temporal dynamics of the entire data set and that of extremes are closely related and thus all the available information about this link should be used in statistical studies of these events. We try to study how centered and normalized data which are closer to stationary data than the observation allows easier statistical analysis and to understand if we are very far from a hypothesis stating that the extreme events of centered and normed data follow a stationary distribution. The location and scale parameters used for this transformation (the central field), as well as extreme parameters obtained for the transformed data enable us to retrieve the trends in extreme events of the initial data set. Through non-parametric statistical methods, we thus compare a model directly built on the extreme events and a model reconstructed from estimations of the trends of the location and scale parameters of the entire data set and stationary extremes obtained from the centered and normed data set. In case of a correct reconstruction, we can clearly state that variations of the characteristics of extremes are well explained by the central field. Through these analyses we bring arguments to choose constant shape parameters of extreme distributions. We show that for the frequency of the moments of high threshold excesses (or for the mean of annual extremes), the general dynamics explains a large part of the trends on frequency of extreme events. The conclusion is less obvious for the amplitudes of threshold exceedances (or the variance of annual extremes) – especially for cold temperatures, partly justified by the statistical tools used, which require further analyses on the variability definition.

Correspondence to: M. Nogaj
(marta.nogaj@edf.fr)

1 Introduction

Studies of global warming and its potential cause – greenhouse gases, are raising fundamental questions and an increasing concern about the problem of the pertinent usage of extreme models in the field of climate (Beniston and Stephenson, 2004; Stott et al., 2003; Palmer and Räisänen, 2002). This work aims at analyzing the non-stationarity of climatic extreme events – their temporal variability beyond the seasonal effect, inherent to the system.

The classical analysis is based on the Extreme Value Theory (EVT) for the stationary case, obtaining constant parameters of the distribution of extremes. Those values are related in a complex manner to the distribution F of the data X_t (e.g. temperature time series) and are in general unobtainable from the usual estimation of F (Embrechts et al., 1997). Therefore, since the extreme parameters depend on subtle properties of the tail of the distribution, the whole data set is usually disregarded when modelling extremes.

The mathematical theory is divided into a probabilistic part (the approximation of the extremal distributions and of the dates of exceedances) mostly studied with stationary assumptions for distributions of observations with regular tails and a weak temporal dependence, and a statistical part difficult to assess because of small samples of extremal values. For the non-stationary case allowing for a temporal evolution of extremes, the main part of the probabilistic theory remains valid. The statistical properties are affected though primarily because of the lack of temporal asymptotics.

Our article aims at presenting a mathematical framework for non stationary extremes and in particular, at proposing a methodology to study the relations between the dynamics of the central field (characterised by a location parameter of the temperature - as the mean or the median, and a scale parameter - as the standard deviation or an inter-quantile range) and the extremes (characterised by quantities of the same type for very high or very low temperatures). We shall see later that

the main analysis are not affected by the choice of the location and scale parameters and thus, we choose the mean and the standard deviation for simplicity.

In the non-stationary case, a trend in the mean affecting the entire observation set must imply a trend of extreme events. The same is true for a trend in the variability of data. It is indeed physically reasonable to assume that the entire observation set can experience trends occurring in the mean and/or variance, contributing to trends in extremely rare values. The implicit question is the connection between the entities of the central field and extreme events: does a relation (simple enough to be useful) explain unequivocally a part or the whole non-stationarity of extremes?

It seems that the intuitive approach considers the behaviour of extremes to be relatively independent from the central field, and the tendency is to work only on extreme data. In this paper, we thus try to discuss this approach and present preliminary results intending to be a preamble to future works and development. We will examine to which extent are extremes explained by central trends. If the information given by the central field is insufficient, it could be interesting to investigate other reasons for the presence of trends in extremes. If F_t are time dependent distributions of X_t , we should examine what kinds of deformations F_t of some F can give the observed or estimated behaviour of extremes (of their parameters). Of course these deformations would be constrained to keep F_t at least with zero mean and unit variance but also possibly to keep time invariant a finite number of other parameters as moments, quantiles or the shape extreme parameter. It is quite impossible to choose one solution among all possibilities without a non mathematical a priori information and, until today, climatic physical considerations on temperature do not give such elements. The study of this problem is thus outside of the scope of this paper.

Although more of a qualitative than quantitative study, this type of general research offers an enriched comprehension of the climatic system and thus improves modelling of extremes.

Under the hypothesis quoted above, the distribution of M_j – maximum values of observations during year j , conveniently centered and scaled by quantities depending on the length n of the year, is approximated by a Generalised Extreme distribution (GEV). If F_t is considered annually constant but varying for each year, the approximation of the distribution will remain valid if for each year-block, the daily distribution is regular enough to apply the EVT. In this simple case we can theoretically extend the framework of the EVT to annual maxima M_j of year j modelled by independent variables from a Generalised Extreme distribution depending on parameters varying from year to year.

The same questions arise for observations above (below) a high (low) threshold instead of maxima per year-blocks. The asymptotic justification of the use of a POT.NS model (non stationary Peaks over Threshold models) is more complicated than for the maxima (GEV.NS). We will thus pro-

vide only the main results. We shall also discuss the case of a GEV model with variations inside a year-block and more generally what kind of non-stationarity can be allowed in these studies.

Our starting point, rather common, is the simplest data transformation defined by

$$Y_t = \frac{X_t - m_t}{s_t},$$

where m_t and s_t are a location and scale function respectively.

We denote by H the following hypothesis: $(H): Y_t = \frac{X_t - m_t}{s_t}$ follows a distribution G , with G invariant in time.

Y is the result of a simple transformation of data, but other transformations could be considered (i.e. logarithmic, powers, etc. ...)

This hypothesis (H) has been already discussed in the literature. Katz and Brown (1992) introduced “local properties” as the (relative) sensibility of the probability of extremes events to a change in mean or variance of the whole observation set. In a more global perspective, Ferro and Stephenson (2005) propose a procedure testing analogues of (H) using the estimates of a limited set of quantiles at two different times. One important reason to use these techniques is that popular methods for testing distributional equality between two samples (as the Kolmogorov-Smirnov test) are based on distances of the empirical functions not taking into account the extreme quantiles. Therefore, the mentioned paper uses, among others, very large quantiles whose behaviour is linked to that of extremes in order to test (H) on disjoint parts of original sample. Still, the tests performed are based on quantiles and the proposed bootstrap procedure remains perhaps difficult to apply in the context of extreme values. Our attempt is thus different and employs the EVT to address extreme distribution.

From the hereditary property of POT models (see Appendix A) or for GEV annual models, it is easy to compute the parameters of the non-stationary models (*POT.NS* and *GEV.NS*) of extremes of X_t from the properties of Y_t in our analysis and reciprocally. The advantage of these transformations is to alleviate the usage of statistical tools, as the transformation reduces the non stationarity of the data.

We can recover the extreme models of X_t from the knowledge we can get of the extreme model linked to Y_t and that of m_t, s_t . We call this approach the indirect model or reconstruction procedure.

Thus, we can compare the two ways of estimating the same model of extremes of X_t – the direct application of the non-stationary EVT to X_t ignoring the information of the whole observation set and the indirect approach consisting in estimating in a first step m_t and s_t from the whole data set, then fitting a (non)stationary model for Y_t and reconstructing the extreme parameters of X_t . We then discuss their differences and analyse the circumstances invalidating (H).

Under hypothesis (H), the trends in extreme events would be due to a translation/scaling effect of the central field. Thus the first question we attempt to answer concerns the complexity of the dependence of the evolution of extremes to that of the “central” climate.

If (H) is verified, we obtain a stationary reconstruction, i.e. all extreme parameters can (in a statistical meaning) be chosen as constant including the shape parameter ξ , and the same value of ξ can be chosen for both models. If (H) is not true, only ξ_t is the same for both models for every t . The previously mentioned hypothesis of time invariance of the ξ parameter of extreme distributions is a reasonable hypothesis and is justified by mathematical studies of the estimation of this parameter on different time periods of a long climatic series (Parey et al., 2006; Chavez-Demoulin, 1997). Moreover, considering the difficulties to estimate ξ , regular trends (e.g. linear trends) of this parameter will have minor statistical sense and will be in general refuted by goodness-of-fit tests (Nogaj et al., 2006). Furthermore, this parameter is the main parameter to describe the behaviour of the tail of the distribution (Embrechts et al., 1997), thus a varying ξ conveys the acceptance of a significant variation of the considered variable. If the distribution of M_j varies along with j , it seems more physically reasonable to consider that those variations occur with a constant ξ parameter, since this implies that the tail of the distribution of X_t remains substantially constant.

In the case of (partial) invalidation of (H), we could work with the centered and normed variable Y_t , and find possible deformations of the distribution G_t of Y_t leading to the observed trends of X_t . Indeed, we could show that many deformations occurring with a constant null mean and unit variance, constant ξ but also with a fixed finite sets of arbitrary quantiles or moments, can generate the same trends in extremes. As mentioned before, this analysis is out of the scope of this paper.

Seasonality driven non-stationarities are treated in different ways in literature (Coles, 2001) (e.g. for tidal extremes (Tawn et al., 1994) or for water vapors (Smith et al., 2000)). Some approaches solve these cases with deterministic temporally varying components, but theoretical and practical considerations are still not completely stabilised. In our study, we investigate trends of extreme events reaching beyond matters of seasonality.

Non stationarity is here expressed using the intuitive variable “time”, but for climatic time series, it could be any environmental covariate or index, as greenhouse gas emissions, atmospheric circulation indices (as the North Atlantic Oscillation index) or others. These indices, functions of time, would justify prediction studies based on extrapolations. The variation of extremes can thus be expressed in terms of covariates having their own dynamic and/or depending on time (Nogaj et al., 2006; Nogaj et al., 2007¹).

¹Nogaj, M., Yiou, P., and Dacunha-Castelle, D.: Study of time-

2 Models for non stationary extremes

We begin our study by investigating the simplest and most common problem, and we compare the behaviour of the temporal evolution of the whole observation data set with the trends observed in yearly most extreme data. Let F_t be the distribution of X_t .

In the simplest case of an annually constant F_t , the annual maxima M_j of year j are modelled by independent variables with a GEV distribution:

$$G(x) = \begin{cases} \exp \left\{ -1 + \xi \left(\frac{x - \mu(j)}{\sigma(j)} \right)^{-\frac{1}{\xi}} \right\}, & \text{if } \xi \neq 0, \\ \exp \left\{ -\exp \left(-\frac{x - \mu(j)}{\sigma(j)} \right) \right\}, & \text{if } \xi = 0. \end{cases} \quad (1)$$

Let $(\xi, \mu(j), \sigma(j), j=1, \dots, A)$ be the parameters of a non-stationary GEV model we denote by *GEV.NS*. We shall suppose that ξ does not depend on j as discussed previously.

An alternative approach consists in choosing a high threshold (often constant) $u(t)$ and considering observations X_t , such that $X_t > u(t)$. The constructed Peaks over Threshold model in this non-stationary context (*POT.NS*), consists of dates of exceedances following a non-stationary Poisson process with intensity $I(t)$, and of threshold excesses $x = X_t - u(t)$, with a Pareto distribution:

$$G(x) = \left(1 + \frac{\xi x}{\sigma(t)} \right)^{-1/\xi}.$$

The Poisson process and Pareto variables are independent. $(\xi, \sigma(t), I(t), u(t))$ are the parameters of the *POT.NS* model.

We have thus three different types of models.

1 – For X_t , we work with the following relation:

$$X_t = s_t Y_t + m_t$$

where m_t is a location parameter (mean or median) and s_t is a scale parameter (standard deviation or interquantile interval) and Y_t is a centered and normalized sequence of weakly dependent random variables

2 – For yearly records we define the direct annual model by the following relation, a being the year index:

$$M_a = \mu(a) + \sigma(a) G_\xi(a)$$

where $G_\xi(a)$ is a sequence of independent variables with a GEV distribution of location 0 and scale 1 (or another fixed value) and shape parameter ξ .

3 – We define the direct *POT.NS* model for a suitable threshold $u(t)$ as:

$$X_{t_i} = U(t_i) + \sigma(t_i) \pi_\xi(i)$$

varying extremes described by a POT model for temperature data – the concept of non-stationary return levels, in preparation, 2007.

where t_i are the successive events of a Poisson process of intensity $I(t)$ and $\pi_\xi(i)$ is a sequence of Pareto distributions with parameters $(\xi, 1)$.

The three relations are thus regression models with non constant variances and noises with specified distributions in the two last cases.

In a first step, for both GEV and POT models, the temporally dependent parameters are modelled by functions (called “non parametric” in statistical terminology). In order to be able to estimate these functions from only a finite data we need to suppose specific regularity conditions. In fact, these functions are estimated through non parametric statistical procedures such as penalised likelihoods providing cubic spline estimates (Green and Silverman, 1994; Wahba, 1990), kernel procedures (Silverman, 1986) or local polynomials (Davison and Ramesh, 2000).

To assess demands of predictability, and thus analytical extrapolations of the parameter functions we could, for instance, choose the class of polynomial functions of maximum degree K , defined by a relatively small number of parameters (“parametric models” in statistical terminology).

3 From temporal transformation of data to distribution deformations

As mentioned in the introduction, we denote by (H) the hypothesis: $Y_t = \frac{X_t - m_t}{s_t}$ follows a distribution G , with G invariant in time. We note that the (H) hypothesis, is a particular case of a general transformation of data of the form $Z_t = \Psi_t(X_t)$.

When the extreme models of X_t are recovered from the knowledge we can get of the extreme model linked to Y_t and that of m_t, s_t this model is said to be obtained by a “reconstruction procedure”.

If (H) is verified, we obtain a stationary model – extreme parameters of the transformed data are constant, and we note the reconstruction of the extreme model of X to be “constrained”.

3.1 Obtaining approximate stationarity after centering and scaling

3.1.1 GEV.NS models

We shall assume that the distribution of F_t is such that a GEV distribution with parameter ξ_1 successfully approximates the block maxima distribution, with m_t and s_t respectively denoting the location and scaling parameters.

We will consider annual models, i.e. m_t and s_t are constant in a block and are therefore denoted by m_a and s_a respectively, for $t \in a$. We will show further-on (see Sect. 3.1.2) how these hypotheses can be relaxed.

Let $M_a = \max X_t$. M_a has an approximate *GEV.NS* distribution $G(\mu_a, \sigma_a, \xi_1)$. If $M_a^* = \max_{t \in a} \frac{X_t - m_a}{s_a} = \max_{t \in a} Y_t$,

then

$$\mathbb{P}(M_a < x) = \mathbb{P}\left(M_a^* < \frac{x - m_a}{s_a}\right) \tag{2}$$

where M_a^* is a sequence with approximate GEV distribution $G(\mu_{0a}, \sigma_{0a}, \xi)$. If the hypothesis (H) is satisfied, μ_{0a} and σ_{0a} are independent from t . From Eq. (2), for every x and every a , we obtain

$$\xi_X = \xi_Y, \sigma_a = \sigma_{0a} s_a, \mu_a = m_a + \mu_{0a} s_a \tag{3}$$

3.1.2 POT.NS models

Let $u(t)$ and $w(t)$ be two thresholds, large enough for the model to be correctly approximated by a POT distribution. From $Y_t = \frac{X_t - m_t}{s_t}$, the threshold for reconstruction is given by $u(t) = s_t v(t) + m_t$ and $P(X_t > u(t)) = P(Y_t > v(t))$.

For convenience, we shall denote the trend in a parameter, say the location m , by m_t or $m(t)$ indistinguishably.

We find (see Appendix A) the following relations for parameters of a POT.NS model of X (with threshold $w(t), v(t)$ being the threshold for Y):

$$\xi_X = \xi_Y$$

$$\sigma_u(t) = \sigma_w(t) + \xi(s(t)v(t) + m(t) - w(t)) \tag{4}$$

$$J_u(t) = \left(1 + \frac{\xi}{\sigma_w} (s(t)v(t) + m(t) - w(t))\right)^{-\frac{1}{\xi}} I_w(t) \tag{5}$$

where $s_t v(t) + m_t - w(t) > 0$.

For the case where $s_t w(t) + m_t - u(t) > 0$, we can find analogue relations for $\sigma_{w,t}$ and $J_{w,t}$ by inverting X and Y and u and w .

Those relations are true no matter if the thresholds are constant or not and the hypothesis (H) is satisfied or not. If (H) is true though, σ_w and J_w are constant.

Similarly, those relations do not depend on the choice of the location $(m(t))$ and scale $(s(t))$ functions, and can thus be chosen as the mean and standard deviation.

4 On the use of these results

The procedure for both GEV and POT models consists of the following steps:

1. Direct estimation through non-parametric procedures of the model of extremes of X giving the parameters (μ_a, σ_a, ξ_X) for the GEV case and $(\xi_1, \sigma_{w,t}, I_{w,t})$ for the POT case, after the choice of w .
2. Estimation of m_t, s_t and computation of Y_t ; choosing a threshold v for the transformed data Y .

3. Estimation of $(\mu_{0,a}, \sigma_{0,a}, \xi_Y)$ and/or $(\sigma_{u,t}, I_{u,t}, \xi_Y)$ for both cases: without constraints and under stationarity hypothesis for Y extremes,
4. Application of the above formulas of reconstruction for X for the two cases; comparison and evaluation of the reconstructions

It seems possible to test directly the stationarity of Y_t , but in fact classical tests perform weakly in presence of a varying s_t .

The results presented in the following part aim at identifying statistical models for both methods acquired at a high threshold.

The arising question we address here concerns the statistical procedures allowing for the usage of those results. We use the reconstruction in two “directions”. The comparison of the trends in extremes obtained through the direct non-stationary procedure and the reconstructed model without constrains of Sect. 3.1.2 allows to test the quality of the statistical methods for the two results. Occurring dissimilarities are only due to statistical errors.

4.1 Verifying the “reconstruction” relations

We estimate in a non-parametric way, through procedures such as penalised weighted least squares, the mean m_t and the variance s_t of X_t . Similarly, smoothed $\mu_{w,t}$ and $\sigma_{w,t}$ and ξ can be obtained through penalised likelihood (Green and Silverman, 1994). We can estimate $I_{w,t}$ through kernel methods (Silverman, 1986). The same methods can be applied with (Y, v) instead of (X, w) .

As stated earlier, to verify the correctness of the (H) hypothesis, we estimate ξ, σ_u and J_u or μ_u through standard procedures with the constraint of stationarity of POT or GEV models, assuming the parameters of Y_t to be constant. We also estimate the parameters of Y_t without the stationarity constraint.

Non-parametric and qualitative plots allow us to evaluate the correctness of the equations relating X_t and Y_t – namely, the relations (4) and (5). The evaluation of the quality of the reconstructed trends of extreme parameters $\sigma_{u,t}$ and $I_{u,t}$ can be done through simulations to point the magnitudes of probabilities of deviations between the two models. A simpler method, consists in computing the absolute area difference between the reconstructed and estimated time-varying parameter normalised by the area under one of the functions, for instance the direct one. This value (denoted by Δ , see Sect. 5.1) can be considered as a useful adjustment criterion to evaluate the correctness of our methodology. We compute it both for the reconstruction based on the general POT model for Y_t and that considering the (H) hypothesis true. We can obtain distributions of this quantity through simulation analyses.

If the reconstructed trends $\sigma_{u,t}$ and $I_{u,t}$ are not convenient enough, the hypothesis is invalidated, and we can reasonably

state that m_t and s_t are not exhaustive explanations of present trends in the extreme values.

4.1.1 Time-varying threshold

Another issue raised by this analysis is the question of a time-varying threshold. Indeed, we can be interested in finding a non-stationary u_t so that the moments of exceedances would be a stationary Poisson process. Among all representations of extremes, the obtained model seems to be the easiest to interpret as u_t becomes the natural trend and the remaining term consists of a noise with varying variance.

Let X_{s_1}, \dots, X_{s_k} be the subsequence of X_{t_i} such that $X_{t_i} > u_t$, then X_{s_1}, \dots, X_{s_k} is the trajectory of X^* , a POT.NS model with parameters $(\xi, \sigma_{u,t}, I_{u,t})$ with $I_{u,t} = \left(1 + \frac{\xi}{\sigma_{w,t}s_t} (u_t - w s_t - m_t) \right)^{-\frac{1}{\xi}} J_{w,t}$. Therefore, we look for a constant I_u , the mean number of exceedances every year (arbitrary but large enough for statistical purposes) such that:

$$u_t = m_t + w s_t + \frac{\sigma_w s_t}{\xi} \left[\left(\frac{I_u}{J_{w,t}} \right)^{-\xi} - 1 \right] \tag{6}$$

The corresponding $\sigma_{u,t}$ is obtained through the following formula:

$$\sigma_{u,t} = \sigma_w \left(\frac{I_u}{J_{w,t}} \right)^{-\xi}.$$

4.1.2 Parametric estimations

A simple technique for the purpose of extrapolation, or to obtain a global visualisation (e.g. involving multiple grid points) requires to estimate s_t, m_t, σ_t, I_t with parametric models, like polynomials, to get confidence intervals. Still, the estimation of parameters as well as the choice of the best models, is done through procedures based on likelihoods, for which the statistical asymptotic theory might not be justified. Also, the problem of choosing a model remains complex for the purpose of our analysis. Still, an example of a procedure through polynomial estimation will be presented in Sect. 5.3.

5 Examples

The presented examples will show the application of our methodology to different types of data. To check the validity of the (H) hypothesis, we obtain the mean m_t and the variance s_t as well as the stationary extreme parameters. With those data, and with Eqs. (4) and (5), we reconstruct the non-stationary parameters of the POT.NS model $\sigma_{u,t}$ and $I_{u,t}$.

5.1 Simulated example and statistical criteria

To assess the correctness of our theoretical results as well as the precision of the obtained reconstructions, we will work

Table 1. Table representing the main features (mean, standard deviation (“SD”), and the 25th and 75th percentiles) of the distribution of Δ – the absolute area difference, obtained for 1000 simulations of Gaussian data with varying mean and standard deviation. The adjustment criterion was computed to compare estimated and true mean and variance ($\Delta(m(t), \hat{m}(t)), \Delta(s(t), \hat{s}(t))$), as well as between different reconstructions, direct estimations or theoretical Poisson intensities with notations specified in Sect. 5.1.

Δ	Mean	SD	25%	75%
$\Delta(m(t), \hat{m}(t))$	0.0012	0.0004	0.0009	0.0015
$\Delta(s(t), \hat{s}(t))$	0.0139	0.0052	0.0103	0.0174
$\Delta^t(I_X, I_X^{r,s})$	0.0705	0.0335	0.0463	0.0868
$\Delta^t(I_X, I_X^k)$	0.1224	0.0473	0.0877	0.1503
$\Delta^t(I_X, I_X^l)$	0.01792	0.0802	0.1223	0.2281
$\Delta^e(I_X^s, I_X^r)$	0.0912	0.0329	0.0656	0.1126
$\Delta^e(I_Y^s, I_Y^r)$	0.1062	0.0378	0.0784	0.1311
$\Delta^t(I_Y^s, I_Y^r)$	0.0901	0.0293	0.0685	0.1091

in the first place on simulated data. Indeed, the aim is to examine how well our methodology determines trends e.g. for the Poisson intensity when comparing blind reconstructions with known initial variations. For this purpose, we perform 1000 simulations of gaussian (other distributions are possible) data X_t , with varying mean and constant variance and try to reconstruct the trend in the frequency of extremes of $Y(t) = \frac{X_t - m_t}{s_t}$. We will proceed similarly for the case where the mean is constant but the variance increases linearly with time and where those both entities vary in time. Thus, we will consider three cases of gaussian data of length $T=5400$. The length of the observation series corresponds roughly to 60 seasons of 90 days. This is motivated by the climatological time series we will work with later on. First (case 1) we simulate data X_t with a linear trend of the mean $m_t = 27.5 + 9.95 \times 10^{-5}t$ and with a constant variance $s=1$. The imposed variation of the mean corresponds to 0.536°C during the simulated period, which is a realistic assumption. Then (case 2), we simulate data with a constant mean $m_t=0$ and a linear trend of the variance $s_t = 1.2 + 6 \times 10^{-5}t$, corresponding to a variation of 0.324°C . Finally (case 3), we shall consider both the mean and the variance as varying in time. Both functions can be estimated through different methods – moving averages, spline functions or local polynomials. We present the results for the moving average estimation. Through kernel methods, we estimate $I_u(t)$ and $J_w(t)$ – with no assumption of stationarity of the obtained distribution of extreme parameters of Y_t . Simulations on the σ parameter of the GPD are also possible, and the estimation would be done by spline functions or local polynomials.

For each case, we compute the adjustment criterion we defined earlier. We will use the following notation for this Δ index, computed as the absolute area difference between two functions f and $g - \Delta(f, g)$. We will denote by $\Delta^e(f, g)$

the case where we estimate m and s to obtain f and g , and $\Delta^t(f, g)$ the case where we use the true functions m and s . Functions f, g can be the Poisson intensity for data $X - I_X^k$ (obtained directly through kernel methods), I_X^r (reconstructed), $I_X^{r,s}$ (reconstructed with stationarity assumption of data Y), and for data $Y - I_Y^k$ (kernel method on the transformed data) and I_Y^s (under stationarity constraint). Similar notation can be adopted for the scale parameter of extremes (σ), estimated by spline functions or local polynomials (σ_X^l). To assess the errors of statistical procedures, we also compute the Δ for the difference between the true and the estimated mean m ($\Delta(m(t), \hat{m}(t))$) and standard deviation s ($\Delta(s(t), \hat{s}(t))$).

An illustration of case (1) and (3) is represented on panels (a) and (b) of Fig. 1, respectively. It depicts the reconstruction of the Poisson intensity under hypothesis (H) (plain line with stars), the reconstruction of the Poisson intensity based on the assumption of non-stationarity of Y (dashed line), the kernel estimation with bandwidth $h=629$ (plain line) and the “theoretical” curve of $I(t)$ obtained through substitution of the true m_t and s_t for the centered and normed data. The reconstructions are qualitatively good, $\Delta^t(I_X, I_X^{r,s})=0.06$ and 0.09 , respectively. The depicted trend in extremes is due to the variation of the mean of the whole data set. As stated earlier, deviations from the theoretical value 0 of Δ indicate imprecision in the reconstructed parameter due to errors in the estimations of m_t and s_t and that of the POT parameters (see Table 1). Also, we infer from the comparison of panels (a) and (b) of Fig. 1, that the introduction of temporal variability in the variance hinders the reconstruction procedure.

As we have introduced it, Δ denotes the difference between each time-varying parameter computed with the direct model for X_t (“theoretical” if m_t and s_t are known) and the reconstructed model with or without the (H) hypothesis. It measures the precision of statistical procedures. In theory, as we showed above, Δ should be equal to 0. The distances between the reconstructed model with stationarity constraint and the unrestricted procedure measures the plausibility of (H). The distribution of Δ has been studied through simulations and can be characterised by some of its main features represented in the Table 1. This table of distributions can be a good basis for decisions when trying to debate on the validity of hypothesis (H), although conclusions should be drawn with care.

5.2 Warm observed temperatures example – the (H) hypothesis

The aim of our study is to be able to obtain trends in extreme events based on the information we get from the whole observation set, and more precisely the mean m_t and the variance s_t . Encouraged by our previous results performed on simulated data, we apply our procedure to temperature observations over a predefined station over Paris and over Prague.

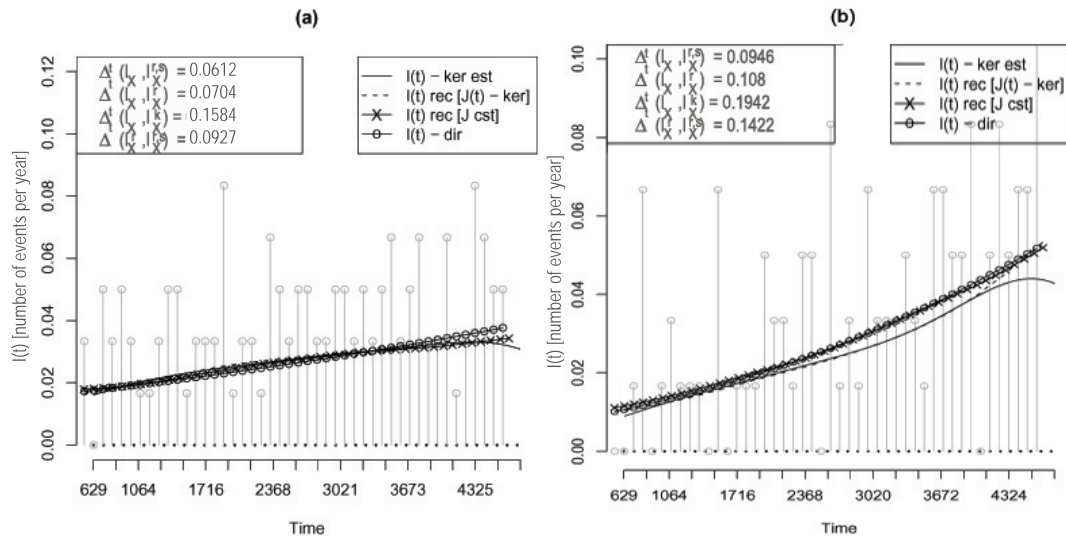


Fig. 1. Reconstruction of the trend in the Poisson intensity from simulated data with an increasing linear trend in the mean ($m_t=27.5+9.95\times 10^{-5}$) and a constant variance ($s=1$) (panel a) and a varying mean and varying variance ($s_t=1.2+9.5\times 10^{-5}$) (panel b). The graphs represent an approximation of the histogram of occurrences of extremes (vertical lines), the reconstruction of I_t (plain line with stars), the reconstruction of $I(t)$ based on kernel estimations (dashed line) with no stationary assumption of Y , the kernel estimation of I_t with bandwidth =629 (plain line) and the direct function I_t obtained by substituting the real values of m_t, s_t . The results presented concern the reconstructions based on $Y(t)=X_t-m_t$. The adjustment criterion is also provided for the evaluation of the correctness of the reconstruction (for notations see Sect. 5.1).

We make use of a long time series over Montsouris (Paris, France) of 131 years (1873 to 2003) taken from the data set produced by the European Climate Assessment project (Klein Tank et al., 2002), where we define the summer season to start the 14 June and end the 21 September. We first estimate the trend of the mean and variance. We consider those values to be constant during a season and to vary on a yearly basis. We first obtain m_a and s_a for X_t , where a corresponds to each year, m is the mean of the summer observations and s is the standard deviation. The same analysis has been performed choosing the median and the inter-quantile range, obtaining similar results to those illustrated later-on. We then apply a cubic spline procedure to smooth m_a and s_a , which are then replicated for each year, to obtain m_t and s_t . The resulting trends are used to compute $Y_t = \frac{X_t - m_t}{s_t}$. We assume Y_t satisfies the previously introduced hypothesis (H), and we apply the POT procedure to compute the extreme parameters of Y_t . Thus, we get $\xi = -0.314$, $u = 2.12$, $\sigma_u = 0.66$. The threshold for X_t is $w = 32.05$. Without any complex procedures to detect trends in the extremes of X_t , the plain line on Fig. 2 shows the reconstructed $\sigma(t)$ and $I(t)$ of the GPD (panel a) and Poisson process (panel b) respectively.

As a way of validation, we estimate $\sigma(t)$ through cubic splines (dashed line), which we represent on the Fig. 2a. Similar results were obtained with local polynomial estimations. The reconstruction based on Eq. (4) follows the spline curve rather closely ($\Delta(\sigma^{r,s}, \sigma^l) = 0.33$), which is a good indication of the performance of our statistical tools, although

the reconstruction without the stationarity assumption of Y_t is better ($\Delta(\sigma^r, \sigma^l) = 0.175$). Similarly, the application of Eq. (5) gives a proper reconstruction of $I(t)$. Indeed, the computed $I(t)$ (plain line on panel b) follows closely the kernel estimation (dashed line) ($\Delta(I^{r,s}, I^k) = 0.14$). The kernel estimator is subject to overfitting and depends on the size of the bandwidth used. Thus, the beginning and ending section corresponding to the width of the window were not represented.

We notice that the difference between $\Delta(I^{r,s}, I^k) = 0.14$ – providing the adjustment criterion between the reconstructed $I(t)$ (under the (H) hypothesis) and the kernel estimate, and the $\Delta(I^r, I^k) = 0.11$ between the kernel estimate of $I(t)$ and the reconstructed Poisson intensity discarding the (H) hypothesis (no stationary assumption of the centered and normed Y_t variables) are small. As they are relatively close to 0, and close to the mean of the Δ value obtained through simulation, they show that both reconstructions are correct and that deviations from 0 might be the result of the statistical procedures used. The immediate conclusion we get from the analysis of these results is that the (H) hypothesis seems to be validated – the mean and the variance of the whole observation set are closely linked to the trends in the occurrence of extreme events in the sense that the hypothesis “ $Y_t = \frac{X_t - m_t}{s_t}$ are stationary observations” can be accepted in a statistical framework.

A closer analysis shows that the general trends are correctly reconstructed for intensities despite visible difficulties at turning points. These problems are mainly due to

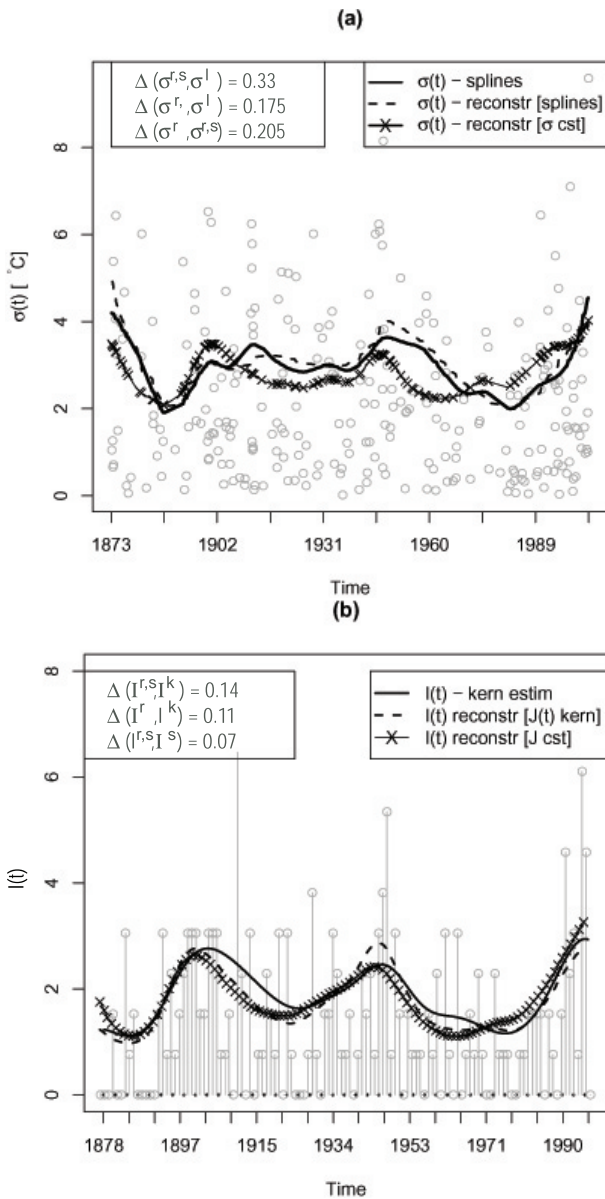


Fig. 2. Extreme trend analysis over the Montsouris station of warm summer temperatures. Panel (a): Exceedances over the threshold $w=32.045$ (dots) and $\sigma(t)$ obtained through different procedures: penalised likelihood (plain line), reconstruction based on Eq. (4) (plain line with stars), reconstruction discarding the (H) hypothesis of stationarity of Y_t (dashed line). Panel (b): Approximation of the Poisson intensity (bars) and $I(t)$ obtained through different procedure - kernel methods (plain line), reconstruction based on Eq. (5) (plain line with stars) and reconstruction without the stationarity hypothesis of Y_t (dashed line).

statistical subtleties of estimations and smoothing. For σ_t the reconstruction with hypothesis (H) does not seem to fit its estimated analogue everywhere. It seems that additionally to a statistical problem linked to the approximation of s_t , there

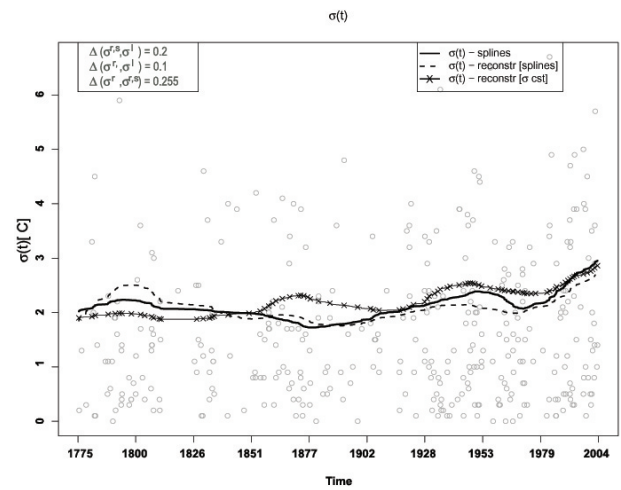


Fig. 3. Same as Fig. 2, but for the Prague station.

remains a part of σ not taken into account by the stationary model. The conclusion remains unclear.

Figure 3 illustrates the same analysis, performed on the Prague station for σ_t and Fig. 4a shows the results for I_t . The results on I_t present a spectacular accuracy of our methodology. Again, general trends are correctly estimated and small discrepancies are found over turning points, still producing a relatively small adjustment criterion $\Delta(\sigma^{r,s}, \sigma^l)=0.2$. Comparing Figs. 3 and 4a shows that the reconstruction performed on the Poisson intensity is better than for σ_t . Since the reconstruction of the scale parameter of the GPD is not very accurate (but better than that of Montsouris) both for the case considering the H hypothesis true and the one with no constraints on the stationarity of the Y_t model, we conclude that the inexactitudes are probably not due (in this case) to the validity of the H hypothesis. It might be a statistical error connected with estimation problems of s_t , σ_t or of the shape parameter ξ of the extreme distribution.

5.2.1 Warm temperatures – POT vs. GEV

A remarkable feature appears when analysing the POT models versus the maxima of the series. Figures 4a, b present the results of our methodology on a long time series of temperature observations over Prague (1997–2004). Panel (b) of Fig. 4 presents in parallel the sequence of annual maxima, and a smoothing produced through a moving average procedure of a 5 years window combined with an affinity to accentuate the turning points around local extremes. We establish a parallel between the variations of the annual maxima and the frequency of occurrence of extreme events. In general, a higher sequence of annual maxima will produce in response as accentuated succession of extreme exceedances. Adversely, lower than average maxima are related to fewer threshold excesses. This result directly relates block maxima to the POT methodology in an empirical way.

5.2.2 Extreme warm temperatures with a time-varying threshold

As presented in Sect. 4, the presented analysis gives mathematical foundations for an analysis of POT models with time-varying thresholds. For the presented time series of daily maximum temperatures over Prague (X_t), we compute the extreme parameters of the POT model ($\xi = -0.25$, $\sigma_w = 2.048$) for a threshold $w = 31.1$. The Poisson intensity I_t of X_t is approximated by kernel estimators. We then apply Eq. (6) for $J = 0.016$, which corresponds to the stationary intensity. Those values are chosen based on previously mentioned statistical properties and a fixed chosen intensity corresponding to an average of 1 event per year.

The obtained time-varying threshold $u(t)$ is used to select new exceedances X^* , for which we look for a POT model with a stationary intensity of the Poisson process. We thus estimate a new Poisson intensity $I^*(t)$ through kernel approximations. Figure 4c illustrates the resulting time-varying threshold producing stationary Poisson excess dates.

The superposition of the Poisson intensities of threshold exceedances of X_t (panel a), of the annual maxima of X_t (panel b) and of the time-varying threshold u_t (panel c) shows the impressive parallel between those features. The same analysis has been done for a time varying intensity of the dates of events of the Y data set. The resulting threshold differs slightly from the one presented on Fig. 4c, a dissimilarity mainly caused by the statistical constraints applied during the procedure.

As penalised likelihood methods are time-consuming, constructing $\sigma(t)$ with the information we get from X_t as well as the stationary GPD model of Y_t seems to be an efficient alternative, which might prove to be more precise than polynomial fits and even spline functions, as more data is available to model the temporal evolution of m_t and s_t than $\sigma(t)$, thus resulting in more accurate estimations.

5.3 North Atlantic example

We applied the described methodology to the National Center for Environmental Prediction (NCEP) reanalysis data over an extended region of the North Atlantic (NA) (80° W–40° W, 30° N–70° N), with a grid of 2.5° by 2.5° degrees. Warm seasons were selected from the 1948–2005 scope of daily data, thus providing 58 summer (June to August (JJA)) seasons of 90 days each. Although other variables can be used, we focused on very warm temperatures.

The advantage of this data set is that it provides gridded temperature data over the whole NA region. On the other hand, since we are in the presence of many gridpoints visual evaluation plots as presented earlier are not feasible. By proceeding in a point-wise manner, we determine of a high threshold (the 95th percentile of the data X_t) and apply the POT methodology presented in Nogaj et al. (2006). We will use the simple polynomial family as the function for the time

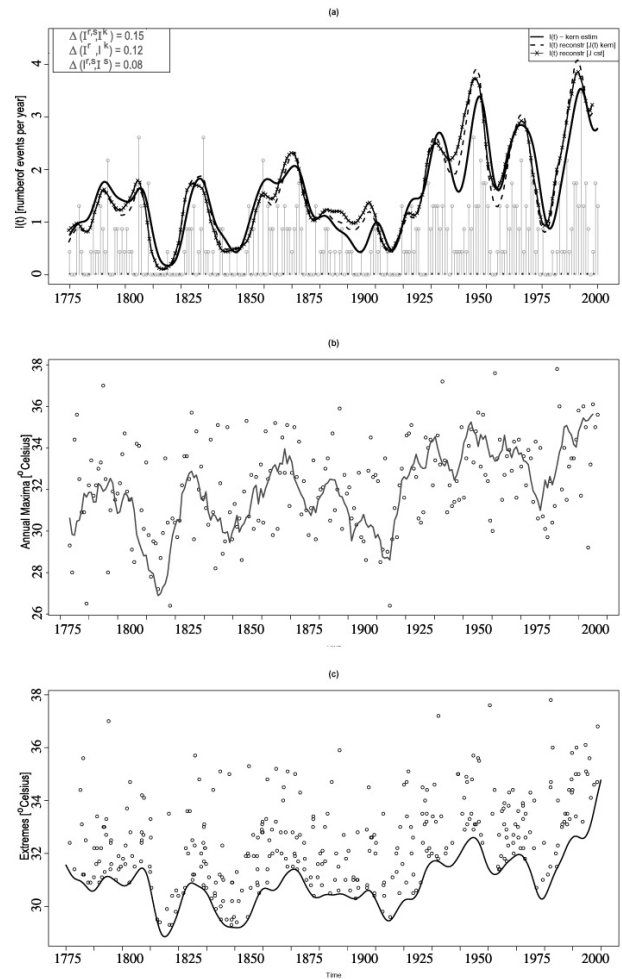


Fig. 4. The upper panel (a) illustrates a similar result to that presented in Fig. 2 but for observations over Prague. Panel (b) draws the annual maxima and smoothed annual maxima through a moving average procedure with a 5 years window. An affinity function is applied to the smoothed maxima to enhance local fluctuations. Panel (c) plots the exceedances of a time-varying threshold (plain line) obtained through Eq. (6).

covariate, to detect trends in the scale parameter σ_t and in I_t . We estimate the extreme parameters through maximum likelihood procedures and determine of the best polynomial fit through likelihood ratio tests. Then, we center and normalise the data X_t , and apply the procedure to the obtained data $Y_t = \frac{X_t - m_t}{s_t}$. If the trend in extreme event over the NA region is uniquely described by the changes occurring in the mean and the variance of the whole data set, then the procedure should reject trends after centering and normalising the data.

Figure 5 shows the obtained degrees of the polynomial fits for the trends of the σ_t (panel a) of the GPD and I_t (panel b) of the Poisson process of the POT.NS model. 0 denotes stationary data, 1 and 2 linear and quadratic trends respectively.

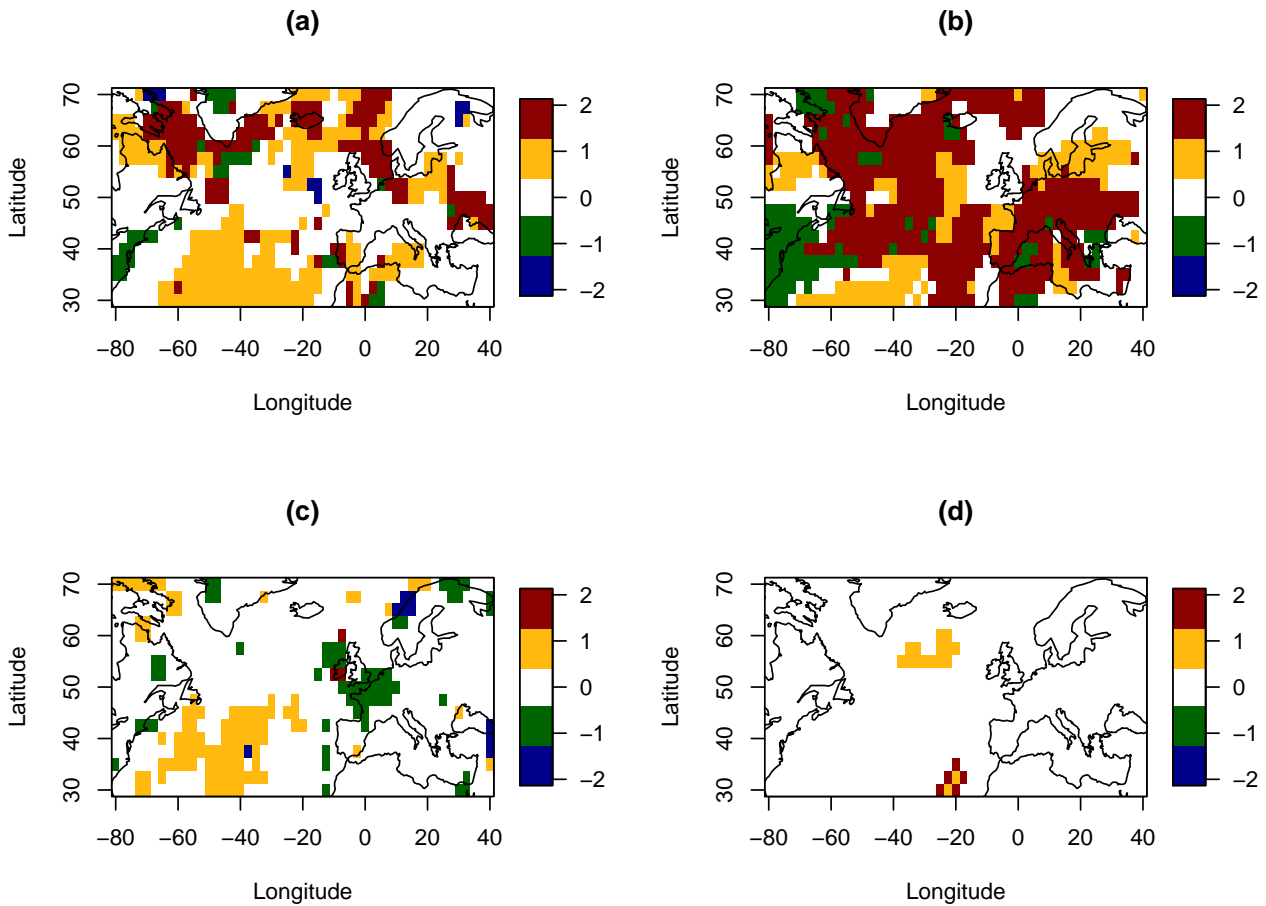


Fig. 5. Detected trends in the extreme parameters of the non-stationary POT model over the North Atlantic. Panel (a) represents the degree of the best polynomial fit for the scale parameter of the GPD. Positive and negative values correspond to the degrees of the polynomials and depict increasing and decreasing trends, respectively. Similarly as for panel (a), panel (b) shows the dynamics of the frequency of the occurrences of extremes. Panels (c–d) are analogues of panels (b–d) but for centered and normed data.

Positive and negative values denote increasing and decreasing trends, respectively. Panels (c–d) represent the detected trends applied on Y_t . Although norming and centering seem to “whiten” the graphs, which is translated by data becoming stationary after the removal of the trends in the mean and variance of the whole data set, we observe some regions where some trends in the σ parameter appears to remain. The south-west of the NA is still affected by increasing linear trends. Surprisingly, some parts of western Europe (mainly northern France and Ireland) present decreasing trends of the scale parameter of extremes after the centering and norming whereas no trend was detected on unmodified data. As seen for the 5.2 results, we have more difficulties in the reconstruction of σ_t than of I_t .

6 Conclusion

The analysed models for the central field and the extremes are seen as regression models. The POT model is applied to exceedances, which occur at moments following a Poisson process. The fundamental method presented in this paper is the introduction of a reconstruction procedure, which recovers the information we have about the extreme parameters characterising the dynamics of the extreme models. We showed, that from the information we have of the mean and the standard deviation of the whole dataset and that on the behaviour of the extremes of a process closer to stationarity, we retrieve the behaviour of the parameters of the GEV or POT models applied to block maxima or threshold excesses respectively.

This procedure is based on the introduction of a centered and normed variable $Y_t = \frac{X_t - m_t}{s_t}$, where m_t and s_t^2 are the mean and the variance of the observations X_t or any other parameters of location and scale. We formulated the hypothesis

(*H*), stating that the extremal model of Y_t is stationary. We focused our analysis on a verification of the mathematical foundation of this hypothesis through simulation procedures. We compare three types of trends to assess the quality of the retrieved information. First, we perform non-parametric estimations of the extreme parameters, which we then compare to the reconstructed trends under hypothesis (*H*) (with constant parameters of the extreme models of the Y_t variable), and the trends obtained while discarding the (*H*) hypothesis, and considering the threshold excesses of Y_t as non-stationary.

Thus, for simulated data with a trend in the mean and the variance, the centered and normed variable was obtained according to the estimated \hat{m}_t and \hat{s}_t , and we analysed to what extend the trends occurring in extremes are driven by m_t and s_t . We quantified the statistical error inherent to the procedure (due to estimation errors), based on the computation of the error between the reconstructed trends and the theoretical ones.

We then applied this reconstruction procedure to climatic data, to appraise the validity of the (*H*) hypothesis on temperature observations. Our results, obtained for several time series, clearly justify the (*H*) hypothesis for the intensity of occurrences of threshold exceedances. This inference is probably false in the case of amplitudes of those excesses, probably due in part to some important inaccuracies of the statistical procedures (both cubic splines and local polynomials were tested) but also of the classical modelling tools. Still, this procedure does not depend on the functions chosen to depict the location and scale parameters (m and s), as both couples, mean/standard deviation and median/inter-quantile range were used to perform the analyses with very similar results.

As more data is available for the computation of the explanatory parameters of the whole data set, the estimation of m_t and s_t and therefore the computation of trends in extreme events as we showed in this paper seems to be more accurate than that of a direct search for time-varying parameters of extremes. Still, we acknowledge that the estimation of m_t and s_t is by itself a first source of statistical error specifically in the estimation of s_t . In fact, in climatic applications, it is not very satisfactory to consider s_t as a smooth version of the intra-annual variability. There appears to exist an inter-annual variability, probably not negligible. An improved model of s_t should thus include both the intra- and inter-annual variability, also adding a smooth deterministic parameter to a random effect depending on the observed temperature. This last point is relevant in order to have a model with random effects. But estimations of such a model remain to be performed.

Appendix A

Mathematical tools

Justification of Eqs. (4) and (5)

Let us recall the following hereditary property for *POT.NS* models. Let a *POT.NS* model with threshold $w(t)$, intensity $J_w(t)$ and parameters $(\xi, \sigma_w(t))$. Also, let t_i the dates of exceedance of the threshold for the observation X_t . Let a new threshold $u(t) > w(t)$, and X_{s_1}, \dots, X_{s_k} be the subsequence of X_{t_i} such that $X_{t_i} > u(t)$, then X_{s_1}, \dots, X_{s_k} is the trajectory of X^* , *POT.NS* model with parameters $(\xi, \sigma_u(t), I_u(t))$ with

$$I_u(t) = \left(1 + \frac{\xi}{\sigma_w}(u(t) - w)\right)^{-\frac{1}{\xi}} J_w(t)$$

$$\sigma_u(t) = \sigma_w(t) + \xi(u(t) - w(t))$$

Suppose now we have estimated a model *POT.NS* for X_t for the threshold $w(t)$ with parameters $(\xi_1, \sigma_w(t), I(t))$ and a *POT.NS* for $Y_t = \frac{X_t - m_t}{s_t}$ with parameters $(\xi, \sigma_v(t), J_v(t))$. If $(X_t > u(t)) = (Y_t > v(t))$, then $u(t) = s(t)v(t) + m(t)$. Applying the previous result for hereditary property for $u(t) > w(t)$, we obtain:

$$\sigma_u(t) = \sigma_w(t) + \xi(s(t)v(t) + m_t - w(t)) \tag{A1}$$

$$J_u(t) = \left(1 + \frac{\xi}{\sigma_w}(s(t)v(t) + m_t - w(t))\right)^{-\frac{1}{\xi}} I_w(t) \tag{A2}$$

A similar result can be obtained for $u(t) < w(t)$ considering first the POT model of Y_t .

If the distribution is bounded, the computation is straightforward: let x_{F_t} and x_G be the upper bounds of X_t and Y_t respectively, we thus obtain:

$$x_{F_t} = s_t x_G + m_t = s_t \left(v - \frac{\sigma_u}{\xi}\right) + m_t$$

From a statistical point of view, the situation is often more delicate: we cannot always find a value w satisfying $w > l(u, t)$ for every $l(u, t) = s_t u + m_t$ and such that there remains enough points above the threshold w in order to perform a correct estimation of the parameters. To overpass this situation we notice that we can prove the relation (A2) (if $w < l(u, t)$), proceeding from the variable Y to X). In this case, the limiting factor is the Poisson approximation, which can be invalidated for an inappropriate number of data points. In the unfrquent case where both approaches cannot solve the proper choice of the threshold u , a possible solution would consist in splitting the time series and applying two thresholds: $w > l(u, t)$ for one part of the data, and $w < l(u, t)$ for the other one.

Edited by: H. A. Dijkstra

Reviewed by: two anonymous referees

References

- Beniston, M. and Stephenson, D.: Extreme climatic events and their evolution under changing climatic conditions, *Global Planet. Change*, 44, 1–9, 2004.
- Chavez-Demoulin, V.: Ajustement de vraisemblance selon Mccullagh et Tibshirani, Internal Report 97.1, Department of Mathematics, EPFL, 1997.
- Coles, S.: *An Introduction to Statistical Modeling of Extreme Values*, Springer, New York, 2001.
- Davison, A. and Ramesh, N.: Local likelihood smoothing of sample extremes, *J. Roy. Statist. Soc.*, 62, 191–208, 2000.
- Embrechts, P., Kluppelberg, C., and Mikosch, T.: *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin, 1997.
- Ferro, C. A., Hannachi, A., and Stephenson, D.: Simple non-parametric techniques for exploring changing probability distributions of weather, *J. Climate*, 18(21), 4344–4354, 2005.
- Green, P. and Silverman, B.: *Nonparametric Regression and Generalized Linear Models*, Chapman and Hall, 1994.
- Katz, R. and Brown, B.: Extreme events in a changing climate: variability is more important than averages, *Clim. Change*, 21, 289–302, 1992.
- Klein Tank, A., Wijngaard, J., Können, G., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Kern-Hansen, L. H. C., Heino, R., Bessemoulin, P., Tzanakou, G. M.-W. M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Aberfeld, A. B. R., van Engelen, A., Forland, E., Mielus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Pachaliuk, A. C. O., Alexander, L., and Petrovic, P.: Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment, *Int. J. Climatol.*, 22, 1441–1453, Data and metadata available at <http://eca.knmi.nl>, 2002.
- Nogaj, M., Yiou, P., Parey, S., Malek, F., and Naveau, P.: Amplitude and frequency of temperature extremes over the North Atlantic region, *Geophys. Res. Lett.*, 33, L10801, doi:10.1029/2005GL024251, 2006.
- Palmer, T. N. and Räisänen, J.: Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature*, 415, 512–514, 2002.
- Parey, S., F. M., Laurent, C., and Dacunha-Castelle, D.: Trends and climate evolution: statistical approach for very high temperatures in france, *J. Climate*, 81(3–4), 331, 2006.
- Silverman, B.: *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, 1986.
- Smith, C., Toumi, R., and Haigh, J.: Seasonal trends in stratospheric water vapor, *Geophys. Res. Lett.*, 27, 1687–1690, 2000.
- Stott, P. A., Stone, D. A., and Allen, M. R.: Human contribution to the European heatwave of 2003, *Nature*, 432, 610–614, 2003.
- Tawn, J., Dixon, M., and Woodworth, P.: *Statistics for the Environment 2: Water related issues*, chapter Trends in sea-level, pages 147–181, Chichester, Wiley, 1994.
- Wahba, G.: *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.