Trends and climate evolution: Statistical approach for very high temperatures in France

**Sylvie Parey · Farida Malek · Carine Laurent ·
Didier Dacunha-Castelle**

Abstract The existence of an increasing trend in average temperatures during the last 50 years is widely acknowledged. Furthermore, there is compelling evidence of the variability of extremes, and rapid strides are made in studies of these events. Indeed, by extending the results of the "extreme value theory" (EVT) to the non-stationary case, analyses can examine the presence of trends in extreme values of stochastic processes. Definition of extreme events, their statistical significance as well as their interpretations have to be handled with great care when used for environmental concerns and public safety. Thus, we will discuss the validity of the hypothesis allowing the use of mathematical theories for these problems. To answer safety requirements, respect installation norms and reduce public risk, return levels are a major operational goal, obtained with the EVT. In this paper, we give quantitative results for observations of high temperatures over the 1950–2003 period in 47 stations in France. We examined the validity of the non-stationary EVT and introduced the notion of return levels (RL) in a time-varying context. Our analysis puts particular accent on the difference between methods used to describe extremes, to perform advanced fits and tests (climatic science), and those estimating the probability of rare future events (security problems in an evolving climate).

After enouncing the method used for trend identification of extremes in term of easily interpretable parameters of distribution laws, we apply the procedure to long series of temperature measurements and check the influence of data length on trend estimation. We also address the problem of choosing the part of observations allowing appropriate extrapolation. In our analysis, we determined the influence of the 2003 heat wave on trend and return-level estimation comparing it to the RL in a stationary context. The application of the procedure to 47 stations spread over France is a first step for a refined spatial analysis. Working on the behavior of distribution parameters while assessing trend identification is a primary tool

S. Parey (✉) · C. Laurent
Electricité de France R&D, 6 quai Watier, 78401 Chatou Cedex, France
e-mail: sylvie.parey@edf.fr

F. Malek · D. Dacunha-Castelle
Mathématiques, Modélisation Stochastique et Statistique, Université Paris-Sud, Orsay, France

🖄 Springer

in order to classify climatic change with respect to the location of the station and open a systematic work using the same methodology for other variables and multivariate studies.

1. Introduction

Dimensioning of civil engineering works is based on the statistical evaluation of return levels for some defined return periods (20, 50 or 100 years depending on the required reliability). For example, air conditioning for some sensitive industrial sites is based on 100-year return levels of high local temperature. These methods, known as extreme evaluation methods, exist since mid 20th-century, and were primarily introduced by Gumbel (1958). Since then, much advancement has been made in numerous fields such as hydrology and oceanography. The greenhouse effect and its possible consequences drive increasing interest of the public and of scientists for these methods. Estimation of extreme temperature in a non-stationary context becomes crucial. We shall see that the best statistical fit to observed data in order to get a good description and understanding of a more or less recent behavior leads to different mathematical choices than the ones induced by extrapolation objective. An important part of this paper deals with descriptive statistics. Still, we approach and dispute pertinent conflicts in climate statistics between mathematical distribution fitting or description and prediction problems. Concerns on extreme events grow in strength because of their socio-economical impacts. Thus, their prediction becomes critical. A clear and unified definition of return levels in a non-stationary context is a fundamental basis for proper prediction. In addition, physical considerations are necessary for appropriate extrapolation. The foundation of the EVT relies on two general definitions of extreme events. Those can be considered as maxima of given blocks of time (e.g. a year or a month), described by the Generalized Extreme Value (GEV) distribution. Another specification consists of Peaks Over Threshold (POT), where extremes are retained values over a properly chosen high threshold. These exceedances, when independent and in sufficient quantity follow a Generalized Pareto Distribution (GPD). The series of occurrences of these events is the trajectory of a Poisson process.

The EVT has been primarily developed for the fields of oceanography and hydrology (Coles, 2001; Katz et al., 2002). In comparison with this last paper, we detail here the problematic of non stationarity in introducing new models and considering in more details the role of the observation length, the sensitivity to a large observation and the link between the general and extreme modeling.

Leadbetter et al. (1983) give the detailed probabilistic foundations of the theory. The paper of Smith (1989) is basic for environment examples and gives examples of non-stationary models.

An increased number of studies were since then applied to climate or environmental issues, but only a few investigate the potential non stationarity of extreme events. The validity of statistical methods depends on asymptotic properties, whereas the definition itself of extremes make them rare, thus implying the necessity to work with long data series. These are hard to obtain, which hinders considerably their statistical estimation. Some hypotheses can be generalized to data dependent in time, but there is no serious mathematical evidence of correctness of these extensions. Another difficulty when attempting to extrapolate rare events is related to the fractious inference of mathematical laws to a class of events never observed in the past. These objections can be resolved by statistical verifications of the quality of model choices. Furthermore, the non stationary context requires additional mathematical hypotheses in order to carry out extrapolation.

Although the methods are based on simple ideas, the theoretical foundation of the probability theory is not completely done until now. The underlying hypotheses one has to ascertain

to justify the statistical model are formulated in the EVT for the stationary case. Still, when dealing with non stationary spatially distributed time series, these processes are fastidious and often left behind in analyses. They often involve graphical tools which can only be applied for exploratory studies. Thus, the classical statistical tools need to be significantly improved.

Stationary series contain no cycle and no monotonous trend. Seasonal means of temperature present an increasing trend since the mid 20th century. This change justifies the question we will answer in this paper, that is: is there a similar change in behavior of extreme temperatures? For physical interpretation, it would be useful to get a better understanding of the link between trends for all observation and trends for extremes. We detail in appendix 4 why this is a very awkward problem without giving any clear solution.

Impact of climate change on extreme events remains a prevalent enquiry, which resolution is hindered by the lack of very long series and difficulties of climate models to efficiently represent rare events (see IPCC report 2001 or Moberg et al., 2004). Nevertheless, much work has been done or is in progress on this subject, especially within the framework of the European projects STARDEX (Statistical and Regional Dynamical downscaling of Extremes for European regions) and MICE (Modeling the Impacts of Climate Extremes). Work on extremes for simulated climate are also conducted in Semenov et al. (2002) and Huntingford et al. (2003). After comparing present simulated climate to observation regarding extreme events, evaluations are made with EVT in a stationary context for fixed periods in the future according to available simulations.

Some evidence on the evolution of the distribution of extreme and very high values can be found in papers of "Understanding changes in weather and climate extremes" (2000 in Bulletin of the American Society: Easterling et al., Meehl et al.). These papers also include a large bibliography on physical aspects of the problem and on data and studies for various areas in the world.

Many recent works focus on regions of different size, from parts of Europe to the entire world. They are concerned with the evolution of universally accepted indices proposed in order to represent high values of climate variables, such as temperature or precipitation. Based on particular standards (see standards of ECA for instance), these indices do not examine very high, rare events, but enable fair evaluations for very large scales (see Yan et al., (2002); see alsoYan et al., (2002)).

The definition of an extreme event, especially in a varying climate context, is not the only difficulty encountered when studying non stationary extremes. The choice of the model for trend, its statistical estimation, as well as confidence interval computation are core concerns which have to be treated by the theory. Only a few studies in probability theory are addressing the basic following problem: what level of non-stationarity keeps the results on climatological extremes valid? Thus, the EVT and the underlying hypotheses needs to be adapted.

The main specificity of this paper is to give a methodology that allows formulating empirical laws on distributions and on the evolution of temperature extremes with time. Trend is usually defined as a function or set of functions of time, modeling the variations of some parameters to roughly describe the evolution of data. In this paper, the notion of trend is split into two complementary concepts to explain distributions of extreme temperatures and their occurrences in time. For extrapolation purposes, trend has to be a clearly monotonous or a periodic, well-identified function, which in our analysis is not always the case. When best fitting models having the largest likelihood are not monotonous, we need to modify our choice, to obtain simple return level evolutions. However, this approach can lead to inadmissible physical or climatic considerations we will try to explain. A better approach is to consider time as an ordinary explicative (clearly deterministic) variable (in statistical language a covariate). Mathematical theory does not mind a specific covariable, we can replace it by any

other one as for instance an index of greenhouse gas emission. Both variables can also be taken into account. We do not detail this approach here for the following reasons: for a recent past, it can be shown that in order to built a non stationary model, the choice of either of these covariates is almost equivalent. But from a statistical point of view it is not correct to keep both variables in the model, they are redundant. For the long period we want to study, time seems clearly more adequate. For prediction we then choose to extrapolate recent temporal trends, which is equivalent to extrapolate past greenhouse gas evolution.

In this paper, we propose a detailed methodology to identify and estimate trends in series of very high temperatures and to perform the required tests on the statistical model validity. Two kinds of statistical problems are considered: the specific effect of the chosen period and the sensitivity to one particular observation. We make also some remarks on the effect of clustering. These results are used to do extrapolations for the evaluation of return levels.

We have done computation for the two approaches, GEV and POT, only the POT method (and results) will be described here, as outcomes are similar in both cases.

Identifying and evaluating the trend will be presented in part 2, with details and results in part 3. Then, part 4 is dedicated to the effect of the very high-observed temperature in 2003, and part 5 will give some results on the use of these statistical models for return level evaluations, before concluding in part 6. Part 7 is an introduction to spatial variation. Part 8 and following are mathematical appendices.

2. Data description and methods

Provided by Météo-France, series of temperature are 3-hourly daily measurements at fixed intervals (0 h, 3 h, 6 h, 12 h, 15 h, 18 h, 21 h) with an instrumental precision of 1/10 degree. Data is spread over 47 stations in France, which locations are shown in Figure 1. For each station, we retain the daily maximum during the A years of observations as the maximum of the 8 3-hourly measurements. In our examples A varies from 19 to 55 (except for Paris-Montsouris, where $A = 130$), in the period 1949–2003 (1873–2002 for Paris-Montsouris, the series being considered as relatively homogeneous by Météo-France only from 1915).

2.1. Seasonal effects

Temperature data contains 1-year seasonality, even if this period effect is less apparent for high values. Suppressing this yearly effect would leave us with irrelevant seasons, as winter, where high temperatures are not inclined to appear. Thus, we will only keep observations for hot seasons. This period is determined empirically from daily and weekly histograms for each station constrained by the homogeneousness rule. A hot season of 100 days from the 14th of June to the 21st of September is finally retained as a compromise between the need of informative data and the required homogeneity.

2.2. POT model

In this paper, we chose to discuss the Peaks Over Threshold (POT) model, as opposed to the Generalized Extreme Value (GEV) model, easier to handle, but less informative. In a first step, we define the statistical model used, we recall the hypotheses and main principles of the method.

From each season (every year) of length $L = 100$ days, we extract extreme values. Let $Y(t)$ be the maximal temperature observed at date t. Let u be a high level threshold defined
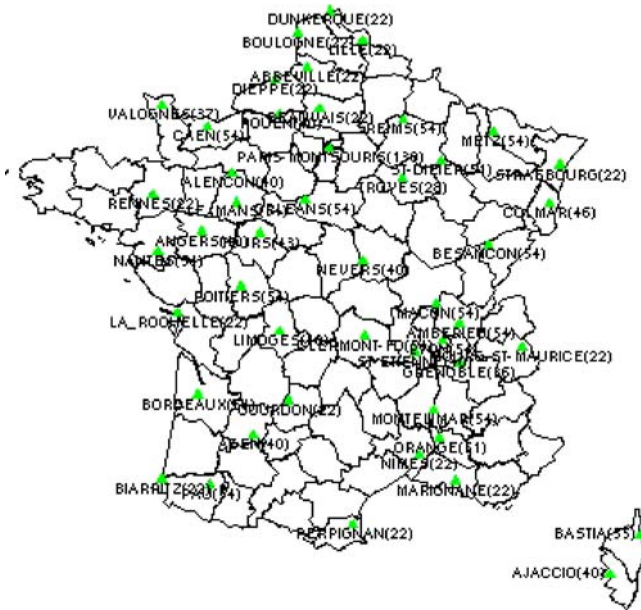
Fig. 1 Geographical location of the 47 Météo-France stations used in the study, with series length in brackets

as the number of days t such that $Y(t) \geq u$ is equal to N with $\frac{N}{AL} = \alpha$, α fixed, A being the number of observed years. Then $U = \{t / Y_t \geq u\}$ is the set of dates corresponding to events exceeding the threshold. The declustering procedure we preferred consists in extracting the highest value of an aggregation of several consecutive peaks, considered as a "cluster". Then $(X_1, \ldots, X_N)$ and $t_1, \ldots, t_N$ are the data after declusterization, X_i being the magnitude of the exceedance at date t_i.

The data $(X_1, \ldots, X_N)$ and $t_1, \ldots, t_N$ can be considered as the sample of a POT model if the following hypotheses are verified:

(K_1) $(t_1, \ldots, t_N)$ is a sample from a non-homogeneous Poisson process of intensity $I(t)$ (see Leadbetter et al., 1983).

(K_2) $X_1, \ldots, X_N$ and $t_1, \ldots, t_N$ are independent sequences of random variables.

(K_3) $X_1 - u, \ldots, X_N - u$ are independent variables with a Pareto distribution depending on time with σ, the scale parameter and ξ, the shape parameter.

This means that there exist functions $\sigma(t)$ and $\xi(t)$ such that for every $t_1, \ldots, t_N$ $P(X_j > x | X_j > u) = H(x, \xi, \sigma)$ with $\xi = \xi(t_j)$ and $\sigma = \sigma(t_j)$

$$H(x, \xi, \sigma) = \begin{cases} \left(1 + \dfrac{\xi}{\sigma}(x - u)\right)^{-\frac{1}{\xi}} & \text{with } \dfrac{x-u}{\sigma} < -\dfrac{1}{\xi} \quad \text{for } \xi < 0 \\[2ex] \left(1 + \dfrac{\xi}{\sigma}(x - u)\right)^{-\frac{1}{\xi}} & \text{with } \dfrac{x-u}{\sigma} > -\dfrac{1}{\xi} \quad \text{for } \xi > 0 \\[2ex] \exp\left(-\dfrac{(x-u)}{\sigma}\right) & \text{for } \xi = 0 \end{cases}$$

To apply the POT theory, a few pre-processing steps need to be applied (details are presented in Appendix A).

A proper threshold choice is the first concern, followed by the declustering procedure. In this way, we can obtain the clusterization coefficient θ, defined as the inverse value of the (random) length of clusters. It is estimated by the ratio of the number of clusters divided by the number of points over the threshold.

A first non-parametric analysis with splines or wavelets (Green et al., 1994) indicates the class of models we should consider in an effort to estimate the temporal evolution of the $\sigma(t)$, $\xi(t)$ and $I(t)$ parameters. The description of these parameters will give us an overview of the general variation of extreme events. Indeed, the $I(t)$ function will describe the change in frequencies of extremes, whereas $\sigma(t)$ will depict the evolution of their strength. This descriptive method suggested the choice of polynomials and the class of continuous piecewise linear (CPL) functions. To choose the optimal model among one class we opted for the objective likelihood test. Finally, goodness of fit tests allow us to check the fit of the obtained distribution of exceedances to the Generalized Pareto Distribution (GPD), the belonging of the set of dates to a Poisson process, and validate the hypotheses of independence.

2.3. Trends

The details for trend identification are given in appendix 1 and discussed hereafter.

As stated earlier, the notion of trend in our analysis is subtle and we will now discuss its definition and interpretation.

The POT model clearly shows the problem that we have to face. We have two stochastically independent objects: the Poisson process and the sequence of exceedances over the threshold. If we suppose the shape parameter ξ of the GPD constant, we have two parameters characterizing extreme climate. First, the intensity $I(t)$ of the Poisson process of the dates of exceedances accounts for the frequency of extreme events. Then, from the GPD, the mean $\sigma(t)/(1-\xi)$ of these peaks, describes their magnitude (let us note that σ is also proportional to the square root of the variance of the Pareto distribution which makes the interpretations more difficult). From the stochastic independence, the evolution of these two functions has no reason to be linked, for instance $I(t)$ may increase when $\sigma(t)$ decreases, increases or remains constant. If $\sigma(t)$ decreases while $I(t)$ increases, the occurrence of rare events changes and becomes more frequent, but their strength is less intense.

A very natural question concerns the relation between the change in the distribution of the whole data set and the evolution of extremes. In the case of the total observation period, trends will describe the alteration of the mean and sometimes of the variance of some distribution F, often specified as Gaussian. For example, Schar and Jendeitsky (2004) attempt to explain the 2003 very high temperature in Western Europe by the increase in the variance of a Gaussian model. In their analyses, trend in extremes is implicitly explained by the trend in the central part of the distribution but for mean values which are of course "more Gaussian". We try in Appendix D to show how trends depend on the distribution of the observations, especially when the shape parameter is fixed and strictly negative for high temperatures thus depicting non-gaussian distributions.

In fact, there do not exist any mathematical theory to deduce trends for extremes from trends for the whole distribution. In Appendix D we try to explain why. The main reason is the following: the knowledge of this transfer function implies a precise knowledge of the distribution of the observations, and the knowledge of trends in mean and variance and of the shape parameter, which characterizes the distribution of extremes, is not sufficient. However, even in the stationary case, extreme theory is often used without any attempt to extract all

the information contained in the observations. Working without this extreme value theory, in making direct and explicit computations on distributions, implies the estimation of the probability distribution with a large precision. This is impossible in the stationary case, then even more in a non stationary context.

3. Results

3.1. Period of observation: Effect on trend choice

It is clear that the estimated trend in the statistical model will be highly dependent on the period of observation. But does the ξ parameter depend on the period or is it a physical characteristic of extreme temperatures at a given station? The long Paris-Montsouris time series can help us examine the question and explore the parameter in detail.

In order to analyze it, we consider 6 nested periods (noted as P) as follows:

P10: [1873, 2002] (the whole series) P21: [1873, 1937] (the first half)
P22: [1938, 2002] (the second half) P31: [1873, 1915] (the first third)
P32: [1916, 1958] (the second third) P33: [1959, 2002] (the last third)

When examining the trend, we obtain the best polynomial degrees of fitted models for the two considered parameters: the Poisson intensity $I(t)$ and the Pareto parameter $\sigma(t)$ (see Table 1). Presented results are the same for the two confidence levels 0.1 and 0.05, these tests are performed using two methods detailed in Appendix A.

The optimal polynomial fitted depends strongly on the observation period. We can thus conclude that our analyzed data is highly non stationary. The degree 5 of the $I(t)$ function for the whole period is not surprising, it allows to capture smooth oscillations of the Poisson intensity already detected by a non parametric analysis (splines or wavelets, Green and Silverman, 1994) without any use of arbitrary choice of polynomial model.

Figure 2 shows the evolution of $I(t)$, the intensity of the Poisson process, with time, for the whole period and for the optimal models found. The optimal CPL model obtained consists of four linear fragments, thus depending on 8 parameters against 6 for the optimal polynomial. A direct comparison with the optimal polynomial model is, as we said previously, difficult. For shorter periods the degree for amplitudes of exceedances is smaller. The quartic observed in the $\sigma(t)$ parameter Pareto distribution of period P31 has to be considered as an expression of an important variability (σ is proportional to the quadratic mean) fitted on a short period rather than an evidence of a trend on values above the threshold.

A notable result is that there is no significant dependence of the ξ parameter on the period, nor on the family of models chosen (for the optimal model). Thus, this parameter seems invariant in time, at least when the time series is long enough to obtain an appropriate precision for the estimation, i.e. around 40 years. Thus, we can do some physically coherent classification with interesting physical features of the stations using this parameter.

Table 1 The best degrees of polynomial models for $I(t)$ and $\sigma(t)$ for each sub-period of time P

Period	P10	P21	P22	P31	P32	P33
$I(t)$ degree	5	0	2	0	0	2
$\sigma(t)$ degree	1	0	0	0	4	0
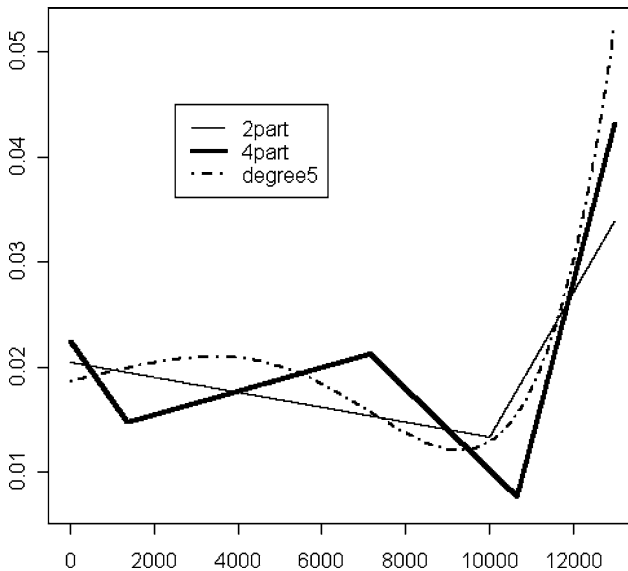
Fig. 2 Paris-Montsouris station (130 years, 13000 days in the hot seasons, 260 points over the threshold): Poisson intensity models: optimal polynomial degree 5, optimal CPL 4 fragments, and a non optimal model with 2 fragments. The *x*-axis represents the concatenated time (days in the hot seasons)

3.2. Results for the 47 stations

We now give results for the 47 stations in France mapped on Figure 1. For all these stations, we have at least 22 and at most 55 years of observations. Each individual threshold is chosen to issue 70 exceedances after declusterization for the 19 first years (until 2000) of the 22-year series. This threshold is chosen removing the years 2001–2003 because they add too many points, which leads, for short observation periods (22 years), to an under-representation of the beginning of the period. Then we add 3 points for each supplementary year. This empirical procedure has been compared to the use of plots which are difficult to handle in a non stationary context. It shows to give better results for tests of fit of models.

3.2.1. Threshold and clusters

Figure 3 shows the geographical repartition of the threshold values. France is divided into three parts: a hot south part (south of 45 °) with a threshold about 31 to 32 °C, an intermediate part about 30 °C and a North part with a level between 23 °C and 28 °C.

To apply the theory of POT models without corrections, we have to be sure that the observations are provided by a weakly dependent stochastic process (see Leadbetter et al., 1983 for precisions on "weakly"). The first step in order to discuss this point is to estimate the clustering coefficient θ. We use the simplest estimator already defined; it gives here $\theta = 0.5 \pm 0.1$. This is roughly equivalent to say that the mean cluster length is 2 days. When θ is clearly less than 1 (in practice less than 0.2) then it can happen that the temporal dependence is so strong that one has to modify the theoretical results used here. This is not the case for hot air temperatures. There is an important loss of information for 2003 and similar episodes. In this cases, the clusters are very long (this is due to the high difference between

Fig. 3 Threshold value (°C) for
each station on its whole period
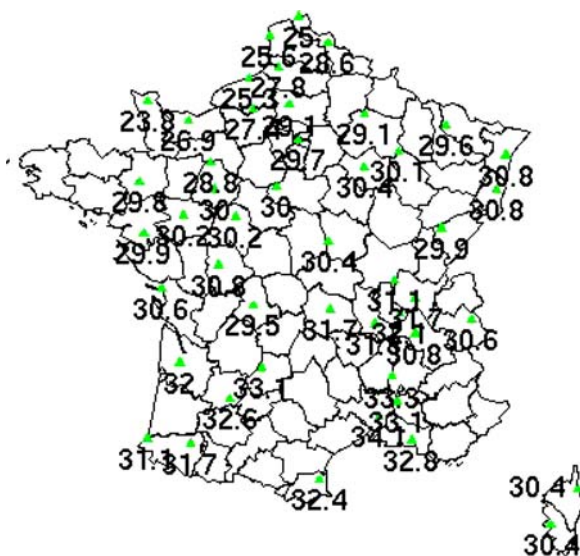
the maximal temperature inside the cluster and the threshold temperature). We shall discuss
this point in a forthcoming paper. In fact, declusterization forces the independence between
exceedances. However, it has the deep drawback not to take into account another kind of non
stationarity: the increase of the cluster length. A non stationary threshold could be thought
as an alternative method, but this is not true because a suitable threshold varies too slowly to
solve this problem.

3.2.2. Shape values and extreme values amplitude trend

We identify an optimal polynomial model for the Poisson process intensities associated with
the days of occurrence of temperature over the threshold and for the parameter $\sigma(t)$ of the
Pareto distribution. We also describe an optimal CPL (continuous piecewise linear) model.
This has been done for the whole observation periods for each station, and the main results
are described in this section.

As it is well known, the shape parameter ξ is difficult to estimate, which is confirmed with
our obtained confidence intervals. First result: the shape coefficient is negative in France with
the following values:

– between -0.1 and -0.5 for periods shorter than 25 years with a standard error (SE)
  between 0.1 and 0.2
– between -0.20 and -0.35 for observation periods of more than 40 years with a SE of
  0.08.

This result can be checked on non stationary Pareto models as well as on stationary ones.
So the distribution of temperatures is not a Gaussian one for which $\xi = 0$. Of course a highly
non stationary Gaussian process could break this result (see Appendix A and D) but we
dispose of a sufficient number of POT stationary models (for a large number of stations) to
reject this possibility. This result is crucial since the "central part" of the distribution is often
considered to be Gaussian! In fact, the distribution of extremes always has a finite upper
bound B given by $B = u - \sigma(t)/\xi$ and for the observed values of ξ, the density of the Pareto
distribution has a zero derivative at the upper bound. We give in the following table (Table 2)

Table 2 Upper bounds of the distribution

Upper bound	Rouen	Bordeaux	Marignane
Stationary	41.97	44.10	46.21
Optimal			
Min	40.21	41.93	42.94
Max	50.37	46.44	61.24

examples of such bounds for stationary and non stationary optimal models. In the second case, the optimal maximum bound (max) will be obtained for the last date of observation (given by a 0.95 confidence interval), whereas the minimal one refers to the first date (min).

Estimated values of ξ such as $\xi > -0.05$ or $\xi < -0.4$ have to be looked with suspicion on the data quality or on the statistical model fit. The absolute value of the parameter ξ generally increases with the degree of the Pareto model. Furthermore, the estimators of the parameters ξ and σ are not asymptotically independent, so the precision of ξ estimation diminishes strongly when the degree of σ increases because there are more parameters to estimate. There doesn't seem to be any evident spatial or climatological classification of the ξ parameter.

We have investigated optimality for two classes of models: polynomial and CPL functions. Results are very similar. The main advantage of CPL models is to avoid unstable choices for Pareto distributions when the degree is 3 or more. Then the dependence between the estimators of the different parameters leads to very large variances. For instance in the case of Dunkerque with only 22 years of observation the matrix of variance-covariance of the estimated parameters is almost infinite and the result on optimality has no statistical meaning. High Pareto degrees with $\xi < -0.4$ lead to very large confidence intervals and cannot be considered as reasonable results. This only happens for observations of length less than 25 years.

Figure 4 shows the spatial distribution for polynomial degree of trend of the mean of the extreme values which is proportional to $\sigma(t)$. The previously described phenomenon is observed in two stations in extreme North but for very short observation length and is thus doubtful.

Fig. 4 Degree of the polynomial trend for $\sigma(t)$ or for the mean of exceedances for the maximal length of observation

We obtain degrees larger than 3 mainly in the south east part of France, specifically in the Alps, the Massif Central and the Rhodanian corridor and in the Alsatian region, which are hot summer regions, either mountains or valleys between two mountain blocks. For short observation periods in other parts of France the parameter remains constant, for longer than 40 years it is linear or quadratic. So the mean of exceedances is highly unstable in hot stations for short periods; this can be seen as a "year effect" and this result disappears for a longer observation for almost all stations. Hot stations often show a quadratic σ, and the other ones a constant σ. It is worth noticing that the degree of the $\sigma(t)$ function divides the country in significant regions – different for Poisson intensity and for the threshold.

3.2.3. Poisson models and spatio-temporal analysis

Figure 5 shows the optimal model degree for the intensity $I(t)$ of the Poisson process which gives the description of the trend for the frequency of exceedances over the threshold.

For more than 40 years of observation, the intensity is very often increasing in a quadratic way, causing an increase in occurrences of extreme events. An exception is the Mediterranean coast with a cubic evolution. Between 40 and 25 years of observation the degree falls to a linear or null dependence in the north near the Channel.

In summary, trends classify climate variation of extremely high temperatures in the following way: the South and the Alps have strong increase of both probability of exceedances and mean over the threshold; Central part has only a quadratic trend for amplitudes of extremes; Channel and Atlantic coasts have a weaker increasing trend. A strong confidence on these results requires at least 30 years (these is an order of magnitude) of observations.

3.2.4. Polynomials versus Continuous Piecewise Linear (CPL) functions

Let us now discuss the choice of the model class between polynomials or CPL models. The advantage of CPL models is to capture more details. Statistical tests allow more parameters for these models (we will not give mathematical reasons in this paper). For instance for long periods, instead of a quadratic trend (3 parameters) for the Poisson intensity we generally

Fig. 5 Optimal degree of the polynomial trend for the Poisson process intensity $I(t)$

find a model with three pieces (6 parameters) which is much more informative. Furthermore, the problems related to polynomials of higher degrees estimated on short observation periods fade in some cases but remains if exceptional data are in an exceptional position as it is the case for 2003 (end of the observation). Confidence intervals of 3-piece linear models are more acute than for a quartic polynomial. Still, statistical optimization is harder for CPL functions than for polynomials and the mathematical theory for model choice and estimation is more difficult (see Dacunha-Castelle et al., 1997 for details on this point). It thus seems a practical advantage to benefit from both classes of functions and to avoid statistical artifacts.

4. Discussion on period length and 2003 effects

The effect of the period length is highly relevant in the case of non stationary process. Similarly, the year 2003 will have an impact on our estimations, considering its critical characteristic: it is the strongest and the last value of the time series. The most important results are given here on selected stations with 54 years of observation.

   We can see that the year 2003 changes dramatically the highest extreme. Table 3 shows the record for the 7 longest series until 2002, and then until 2003.

   Table 4 shows the ten largest values for the whole period for Poitiers (after declusterization) (1950–2003).

   The 2003 record is larger than the previous one (when the ten largest values occur before 2002) by more than 1 degree or even by 2 degrees. If we look at the ten largest values until 2002, they are approximately separated by 0.1–0.4 degree. Such a jump in the record has a negligible probability if the distribution is the non stationary Pareto fitted on data until 2002. So we have to conclude that 2003 temperatures had a very weak probability to be "explained" by the trend in the last 30 years (perhaps using 3 centuries of observation the situation would not be the same). In Poitiers the observed value in 2003 is larger than the theoretical upper bound of the distribution estimated for 2002!

   Adding years 2001 to 2003 to the sample leads to much more threshold exceedances, but more are lost during declusterization, because these exceedances occur in a consecutive way. Table 5 illustrates this behaviour for the same series as previously but considering a 25 years period (1976–2000) and then the same period with years 2001, 2002 and 2003.

   For a fixed desired number of independent extreme values, this quantity is multiplied by 1.2–2.5 times after the addition of the year 2001–2003. Still, recent observed years lead to increase clusterization and episode length. The clusterization coefficient θ is estimated at 0.45 until 2000 (with a SE of 0.05) and at 0.26 until 2003.

   The effect of a record year as 2003 is notable, especially for short observation periods: details will be given later on, but let us remark here that the estimation of ξ can highly depend

Table 3 Record for the series until 2002, and then until 2003

	Threshold 3% points	Records for period 1950–2002	Records for period 1950–2003
Bordeaux	31.99	08/09/98 38.30	08/04/03 40.01
Lyon	31.96	07/22/83 39.30	08/13/03 40.49
Metz	29.48	08/11/98 37.86	08/07/03 38.79
Montelimar	33.21	07/06/82 39.49	08/05/03 40.49
Orleans	29.98	09/04/52 37.85	08/06/03 39.30
Poitiers	30.69	07/22/90 37.09	08/05/03 39.10
Reims	29.11	08/11/98 37.07	08/12/03 38.90

Table 4 10 largest values for
Poitiers between 1950 and 2003

Poitiers	Temperature
08/05/03	**39.1**
06/22/03	**37.7**
07/22/90	37.1
08/04/90	36.9
08/10/98	36.8
08/04/75	36.8
08/29/61	36.4
06/30/52	36.4
08/15/74	36.0
08/25/01	36.0

Table 5 Number of threshold exceedances and number of values lost during declusterization for 2 periods: 1976–2000 and 1976–2003 in 7 stations

Station	Period 1976–2000 3.4% exceedances		Period 1976–2003 3.4% exceedances	
	Total number of values > threshold	Number of values lost during declusterization	Total number of values > threshold	Number of values lost during declusterization
Bordeaux	162	77	284	190
Lyon	197	112	355	261
Metz	199	114	372	278
Montelimar	206	121	374	280
Orleans	167	82	326	232
Poitiers	177	92	313	219
Reims	205	120	353	259

on a fraction of values in the sample. This point emphasizes our previous remark: ξ does not depend on the period if its length is sufficient. In general, 2003 changes dramatically the absolute record, and the model fit is very sensitive to a single very high value. In particular, it increases the degree of the polynomial giving the optimal model. Pareto models depend on the observation period but the position in the sample of a record year is also very important. If we permute years 2003 and 1995 (located in the middle part of the observed period) the effect of the record year on trends is now less meaningful.

2003 effect is strong even for a long period as illustrated for Caen in Figure 6 below.

Regarding the Poisson process intensity trend, we can see that for short periods, we generally do not capture the increasing trend if year 2003 is not included. Otherwise, a linear temporal dependence is the prevailing result with two important exceptions: in the south of France or near the Alps, where the degree is at least 3, and on the Atlantic Ocean or Channel coasts where the frequency of extremes remains constant, as shown on Figure 7 below.

5. Return levels

5.1. A suitable definition for the non stationary case

In the stationary case, the return level z_a for a years is the level for which the probability of exceedance every year is equal to $1/a$. In the non-stationary case, knowing the identified

trend, we need to re-define z_a as the unique level such that the expectation of the number of exceedances over z_a in the next a years will be 1. We shall define the return level z_a for a years starting from the date t_0, so the return level has to be thought as a function of the initial date and of the number of years taken into account to make the prediction.

Let $D(t_0, a)$ be the set of all days that belong to $[t_0, t_0 + 365a]$. For instance $D(t_0, a)$ are all the hot season days from 14 June to 21 September. Then, return level z_a is such that:

$$\sum_{t \in D(t_0, a)} \left(1 + \frac{\xi}{\sigma(t)}(z_a - u)\right)^{-\frac{1}{\xi}} I(t) = 1$$

$I(t)$ is the mathematical expectation of the number of exceedances at date t and the Pareto term $(1 + \frac{\xi}{\sigma(t)}(z_a - u))^{-\frac{1}{\xi}}$ is the mathematical expectation that the exceedance is larger than the level z_a.

Fig. 6 Optimal degree for polynomial trend of Pareto $\sigma(t)$ parameter for station Caen on periods 1982–2002, 1982–2003 and 1950–2003. (the abscissa scale is not the same for all graphs which corresponds to different period lengths)
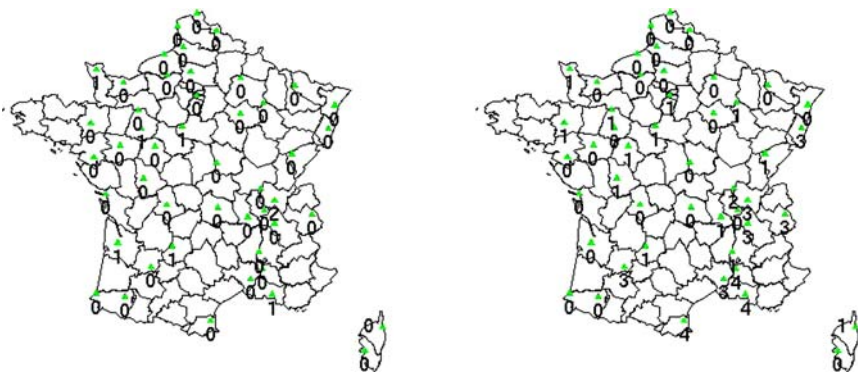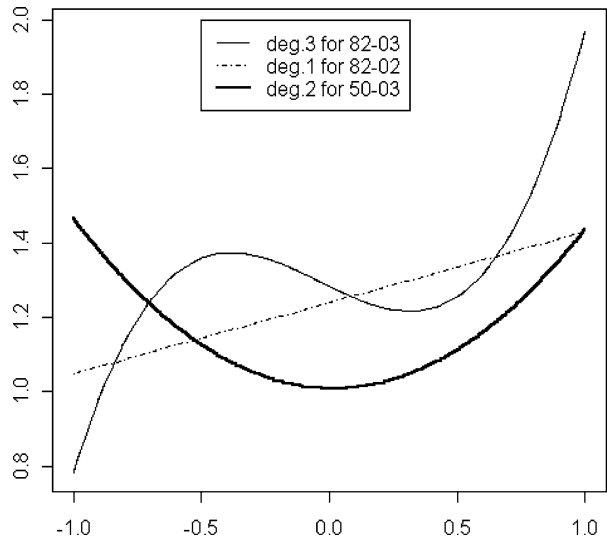
Fig. 7 Geographical repartition of optimal degree of polynomial trend for Poisson process intensity $I(t)$ on period 1982–2002 (left) and 1982–2003 (right)

The definition for the stationary case is a particular case of this one.

5.2. Influence of non monotonous model choice

This extension of the return levels to the non stationary case is a very important point for the use of statistical tools to make predictions on extreme values. In this paper, the prediction is done using an extrapolation of the trends of the chosen model in order to estimate a return level for the future in a changing climate.

Prediction in extremes as well as in the whole observation distribution cannot only be based on statistical tools if extra information can be given by physical simulation models or by scenario based not only on physical considerations but also on sociological and political ones, associated with energy consumption. Our concern here was to evaluate return levels valid for the near future (the next 20–30 years) taking climate evolution into account. Climate simulations do not yet provide spatially refined information at that time scale, thus, we choose to use statistical extrapolations.

It is reminded that the optimal model has been identified using statistical tests in order to fit a model to the data. We apply the so-called principle of parsimony that chooses the model with the largest number of parameters allowed by the mathematical theory. What about this choice for extrapolation? If the estimated parameter function is neither monotonous nor periodic, there is no natural way to define this extrapolation (of course this point remains a problem every time the explicative covariate is not given by extra statistical information). Therefore, even if the model choice leads to a non-monotonous polynomial of high degree, we will have (for instance, but this is a quite arbitrary decision) to decrease the degree to obtain a monotonous polynomial. Moreover, we have to check that extrapolation will always provide positive Poisson intensity $I(t)$ and $\sigma(t)$ functions, so decreasing functions are not admissible estimates for long return levels computation.

Even so, extrapolation using high-degree polynomial functions is risky. For instance it is clear that if the increasing trend for Poisson intensity or for the Pareto mean relies on polynomial models of cubic degree or higher, extrapolation over 50 years will be reduced to the cubic term and will have neither reasonable physical nor mathematical significance. Thus, the model choice is physically not realistic for the Pareto distributions as soon as the degree is larger than 2 and the period larger than the observation one.

To solve this problem, we can work with sub-optimal models with an upper limit of quadratic polynomials, or perform computations with CPL models, which will avoid the artificial effect of the high degree of polynomials.

This constraint is still not sufficient. The weakness of CPL models is that prediction is done extrapolating the last linear piece. The slope of this linear fragment may be too dependant on the distribution of the last observed values in the last period, for instance on the "2003 effect" if the prediction is done in 2004! We define a specific statistical procedure in order to control these edge effects (see Appendix C).

The detected change in temperatures since the seventies (most certainly due to anthropogenic influence) implies a careful choice of the length of data considered to compute return levels.

5.3. Results

We present the results of computed return levels for maximum observation period featured for each station (as described in Section 2), with comparison between stationary and sub-optimal polynomial model (i.e. with a quadratic limit to the Pareto $\sigma(t)$ parameter function)
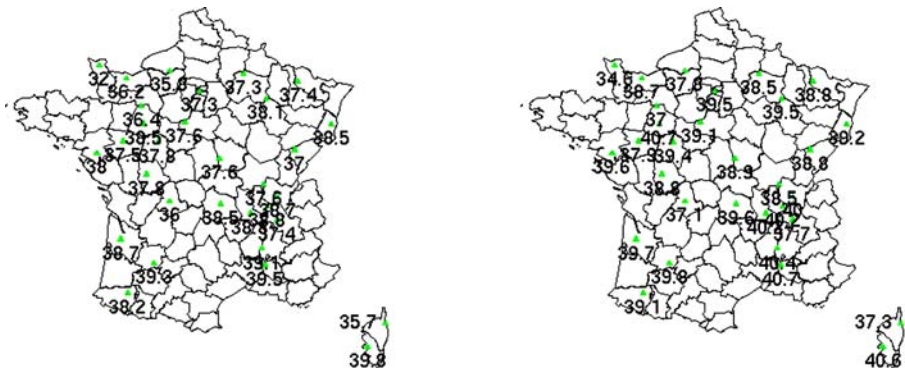
Fig. 8 30-year return levels considering the series as stationary (left) or taking trends into account (right) for series longer than 40 years

for 30-, 50- and 100-year return periods. Figure 8 shows results for 30-year return levels for the stations with more than 40 years of available observations.

For a 30-year return period, optimal (or sub-optimal) models generally increase the return level by 1 to 2 degrees (1.5 is exceeded near the Atlantic and Channel coasts). For shorter observation periods (for instance 22 years) this difference is less marked, with only a 0.5–1 degree increase in RLs.

When the optimal Pareto model is of higher degree than 3, considering CPLs doesn't improve the results.

To assess statistical significance of obtained return levels, we need to compute their confidence intervals. We noticed that more than 90% of the variance is connected to errors in the fit of the Pareto distribution. Thus, the contribution of the Poisson error is insignificant. In particular, for high optimal degrees of the Pareto model, the mean square error on the RL reaches 0.5 degrees for a 30-year return period and 50 years of available data. Since asymptotic hypotheses are not valid for almost all non stationary models we have considered (see Malek, 2004), we need to estimate confidence intervals through simulations. Of course, the simulation is a heavy procedure; we give in Appendix A some methodological procedures.

6. Conclusion

We studied very high and very low (not shown) temperatures for a large set of stations over France. Depending on the location, different observation length were available, from 22 to 55 years (year 2003 inclusive), and 130 years for 1 station.

Complex statistical analysis allows us to draw several relevant conclusions. First, the shape parameter ξ of the Pareto distribution is negative ($-0.4 < \xi < -0.1$) thus implying a non-Gaussian and short-tailed distribution of temperatures with physically sensible upper bounds. The value of ξ can vary but only slightly with the observation period, its variation can be neglected in first approximation allowing to restrict non stationarity characterization to the other parameters, more easy to interpret physically. We then choose the order of polynomial models of parameters to look for a physically sensible classification of stations; this was only possible from this property of ξ. The shape parameter will change significantly for an erroneous choice of the $\sigma(t)$ function since the estimation of these coefficients are linked, so we always work with statistically optimal models.

The intensity $I(t)$ of Poisson processes gives informative results on the frequency of exceedances over a high threshold. It allows a clear classification of stations (and it can be closely related to the mean parameter of a GEV model). We determined two kinds of trend of $I(t)$: a general, soft aperiodic oscillation with duration of 20 to 40 years for long observation, and a clear increasing drift during the last decades of the 20th century. From the analysis of the longest time series of 130 years, we can infer that this last increase lies in the last 25 ± 10 years. This trend is mainly quadratic for polynomial models and presents 3 or 2 fragments when considering CPL functions. Probably, this method gives intuitive partitions of the observation periods. These two families of models give similar results.

The $\sigma(t)$ function describes the non-stationarity of the magnitude but also the variability of extreme events and its behavior is station-dependent. Still, the interpretation of this parameter is difficult since it accounts for both the mean and the variance of extremes. $\sigma(t)$ tends to increase in hottest and elevated regions in France. Additional studies are needed in order to analyze events as the 2003 heat wave. We discussed the sensitivity of a very hot year placed at the end of the observation period – the year 2003. For short observation periods (<25 years), this particular data modifies trends significantly. From a purely statistical point of view, 2003 cannot be considered as a "product" of the recent trend. Even taking into account the increasing trends for temperature, including a possible increase of variability, the observed value is predicted with a very small probability, as an extreme between extremes. The jump in records sequence is between twice and four times the observed and regular jumps of more or less 0.4 degrees in past observations. 2003 observation could also be explained with a not so small probability by an acceleration of the increasing trend. However, even using sophisticated CPL models, there is not enough observations to prove or invalidate this hypothesis.

Estimation of ξ is one of the main technical difficulties for the computation of return levels. Results are more given in this part in order to illustrate a methodology than for their intrinsic value. A 30-year return level for very hot temperatures (above the 99th percentile for the whole year, and above the 97th percentile for the hot periods), will be more realistic for the CPL models, where extrapolation is more justified for long observation periods. For polynomial optimal non-stationary models, the increase for the 30-year return level varies from 1 to 2 degrees, with a 0.5 degree precision. This result is similar to the mean trend of the observation period.

This complex but exploratory analysis is meant to be exploited in advanced studies of extreme events or for different meteorological variables. It can also be used with other varying variables than time, alone or jointly with time.

7. Perspective: Spatial qualitative aspects

Till now, we applied our procedure independently for a set of S stations. To avoid local fluctuations, we could try to compute a global trend for the whole set or sub-set of stations. This problem, for flood predictions, was brought by Katz et al. (2002), and is well-known in classical statistics with Gaussian models. In the framework of rare events, correlations between stations are insufficient to summarize dependencies. Furthermore, a multivariate model is not available for extremes as is it in the Gaussian case to study average fields. The accessible copula theory of multivariate extremes is an efficient tool in probability theory to analyze spatial dependencies, but these complex models require a large amount of data. Pairwise exploratory studies confirm that correlation is not a good dependency measure and that the agreement of extreme events is strongly dependent on the geographical proximity.

Appendix A: Trend model choice

A.1. Threshold selection

Many methods can be used to select a threshold. One of the most popular is to use properties of Pareto distribution. Let's consider the threshold u such that exceedances over u have a Pareto distribution with parameter ξ and, the dates are the trajectory of a non stationary Poisson process. For another threshold $v > u$, exceedances over v will have the same properties. Furthermore, the mean of the exceedances over v is a linear function of v Graphical translations of these properties (see Coles, 2001) could be used for control purposes once adapted to non-stationary cases but they are hard to apply when dealing with a large set of stations. Our method consists in selecting an adaptive threshold u according to a fixed, desired number of exceedances N (e.g. $N = 150$). In fact, we choose a number depending on the length of the observation period following the empirical law $N = 10 + 3A$ where A is the observation length in years. We always check this rule verifying first the stability: the model parameters have to be stable when the number of declusterized exceedances is modified by a number of 10.

   Then we have to check the model adequacy. To perform classical tests of fitting, we have (it is a fastidious but obliged way) to transform the models into stationary models, by change of clock for the Poisson process and change of scale for the Pareto one.

   Therefore, we are now able to use the classical graphical rules with a correct mathematical basis.

   We have made this verification for all stations in order to verify that our threshold is correct for the stationary process obtained after transformation. It has always been the case except for 4 stations.

A.2. Rough trend research, choice of model into a given class and confidence intervals

To deal with a possible trend, we use the POT model with $I(t)$ (intensity parameter of the Poisson process) and $\sigma(t)$ (of the Pareto distribution) as functions of time t. The independence of those two parameters allows separate partial likelihoods. The parameter ξ gives the shape of the Pareto density and is the most difficult to estimate. As we already detailed, we have verified that a generally admitted use of time-independent ξ is reasonable. A first, non parametric study suggests a proper choice of classes of model functions for the parameters $I(t)$ and $\sigma(t)$. This is done using the method of penalized likelihood and cross validation giving cubic splines estimates or a wavelet method (Green and Silverman, 1994). The behavior of $I(t)$ and $\sigma(t)$ provided by the non parametric study allows choosing a parameterized class of functions. Thus, a polynomial of degree at most 5 is a reasonable choice for time dependence of those functions. Alternatively, piece-wise linear functions have been applied and provide interesting results. We address the problem of the choice of the best polynomial function, the estimation of its coefficients and the evaluation of the variance of the estimators.

   The detailed procedure concerns the simplest case of the Poisson intensity $I(t)$ or the $\sigma(t)$ function, but can be applied to any parametric nested model (like CPL functions).

   For every integer d ($d \geq 1$), we compute the set of coefficients which maximizes the log-likelihood for data X_i with a Poisson intensity I depending on time as a polynomial of degree d with these coefficients (see Appendix B for the expression of log-likelihood).

To choose the optimal model for $I(t)$, we test the best model for d against the best model for $d + 1$ using a likelihood test (see for instance Coles, 2001) and an associated level of confidence of the test, for instance $\alpha = 0.05$ or 0.10. We decide to choose degree $d + 1$ if the test rejects degree d at level α.

Confidence intervals (CI) for the estimators are critical to determine parameter significance. The classical methods, i.e the global asymptotic method, with variances of estimators estimated from the observed information matrix, and the profiled likelihood method (see Coles, 2001 for Pareto models and Dacunha-Castelle (1986) for general results) are based on asymptotic hypotheses.

But in our case, asymptotic hypotheses are not always satisfied. For Poisson non stationary process estimation, asymptotic theory can be irrelevant and very inappropriate in some cases (see Malek, 2004). Therefore, to be cautious, we prefer to control classical methods by simulation procedures to perform tests (for example, we do not want to apply without control chi-square approximations to the test of likelihood ratios) and to compute confidence intervals (reasonable confidence level is chosen as 0.05 or 0.1). We start from the optimal estimated model. We simulate 10000 trajectories of the POT model. On every trajectory, we estimate the optimal model and his parameters. Thus, we obtain an estimate of their distribution, a confidence interval and confidence volumes. Now, we can check the coherence of the initial choice with the value of the percentage of trajectories which give this choice. For the return level, we compute its value for every trajectory and then we compute the confidence interval from the estimated distribution of the return level. The agreement between classical methods modified as previously explained and the simulation results are good enough. As confidence intervals, we choose the longer of the two estimates.

A.3. Model validation

Following the above procedure, the optimal model in the class of all polynomial or CPL models is chosen, and validated through Kolmogorov or chi-square goodness of fit tests. To check the Poisson character, we test if the distribution of intervals between dates of events is exponential. In this validation, non stationarity implies to perform a change of clock and adapt the test. Similarly, we test for adjustment of the Pareto distribution and the independence between the two processes. These tests are performed on the "best candidate" in the class.

Appendix B: Log-likelihood expression

Let u be the chosen threshold, a the number of years of observation, L the hot season length and N the number of observed values $X_1, \ldots, X_N$ over the threshold at dates $t_1, \ldots, t_N$ (after declusterization) among the aL observed values. The log-likelihood is given by $L_N = L_N^1 + L_N^2$, where L_N is the sum of two independent terms:

$$L_N^1 = \sum_{n=1}^{N} h_\xi \left(\frac{X_n - u}{\sigma(t_n)} \right) \quad \text{the log-likelihood of the exceedances (GPD),}$$

with h_ξ defined as

$$\begin{cases} \dfrac{1}{\sigma(t)} \left(1 + \dfrac{\xi}{\sigma(t)}(z - u) \right)^{-1 - \frac{1}{\xi}} & \text{for} \quad z < u - \dfrac{\sigma(t)}{\xi} \\ 0 & \text{in others cases} \end{cases}$$

$$L_N^2 = \sum_{n=1}^{N} \log I(t_n) - \int_0^{AL} I(t)dt$$

the log-likelihood of the Poisson process of dates whose intensity is $I(t)$.

Appendix C: Technical problems with CPL models

The main difficulty of CPL functions lies on their non identifiability, i.e the same model can be defined by different sets of parameters. The classical theory of likelihood is not valid anymore and one has to use a specific theory similar to that given for mixture models by Dacunha-Castelle and Gassiat (1996). This issue results in problems of estimation of parameters that can have no consistent estimators. This can be avoided under the constraint of separation of angular nodes, e.g. by 10 years.

Appendix D: Trend on all observations and trend on extremes

We don't try to expose a complete mathematical theory (which does not exist as we explain). We just give some ideas of what can happen for climate variable, considering simpler situations than the one considered in this paper.

First, suppose we have to deal with gaussian variables. Daily maximum temperature is then T_t with distribution $N(m(t), \sigma(t))$. What is a trend in this case? For σ constant, the trend is the function $m(t)$. It can be estimated by some smoothing procedures as moving averages, kernels or wavelets.

If σ is not constant, the notion of trend is more sophisticated and perhaps rather arbitrary because a variation of the mean (thought physically as a signal) is only significant if it is large enough compared to σ (level of noise). Even if we have a good estimate of m and σ, there can exist significant trends on some time intervals and not on other ones.

Let us look now at extremes. Suppose $\sigma = 1$ and $m(t) = \alpha + \beta t$ with $\beta > 0$. Let $M_N = \max(T_1, \ldots, T_N)$. Then, M_N is equivalent to βN! There isn't any random effect for the behaviour.

This can be seen as follows: if $Y_t = T_t - m(t)$, it can be proved that $Z_N = \max(Y_1, \ldots, Y_N)$ increases very slowly and that $\frac{Z_N}{N} \to 0$. Z_N is equivalent to $\sqrt{\log(N)}G + \sqrt{\log(N) + a\log_2(N)}$ for large N, where a is some constant and G a Gumbel centered random variable with variance 1. For large t, $\mathrm{Prob}(G < \alpha + \beta t) < ce^{-\beta t}$ for some $c > 0$. Then from elementary probability result, for all N large enough (see later what "enough" means), $Z_N < \alpha + \beta N$.

Now if we take K disjoint blocks of large size N, say $N = K = 200$, and look at the behaviour of the maximum into the K blocks of size N, there is not a GEV model and maxima in the block J; $J = 1, \ldots, K$, is given by the general mean trend and so equivalent to βNJ! Block size is only an amplifier.

Of course this analysis is only valid in an asymptotic situation and it can be extended to POT models.

For temperature data, there is a trend on mean more or less linear, (in fact in some cases there exists a significant quadratic term). We observe a linear trend (or a quadratic one) for the mean parameter of GEV. It is in general stronger than the previous one. This can be a consequence of many kinds of changes in the distributions. For instance, if the tail moves slowly from a gaussian behaviour e^{x^2} to an exponential behaviour e^{-ax}, then we can observe

the previous qualitative phenomena without any change for the popular shape coefficient ξ which always remains equal to 0.

Before giving a practical and perhaps methodological conclusion, let us go to another simple example.

Let now T_t have an exponential distribution depending on the sample size N given by its density $\lambda_N \exp{-\lambda_N(x - \theta_N)}$ for $x > \theta_N$ and 0 for $x < \theta_N$.

Let M_N be defined as the maximum of observations as previously and let us look at the behaviour of M_N.

As usually, we look to some sequences of real numbers A_N and B_N with evident physical meaning such that $(M_N - B_N)/A_N$ has a non zero limit distribution.

From $P(M_N < A_N x + B_N) = (1 - e^{\lambda_N(A_N x + B_N - \theta_N)})^N$ we get immediately that the unique solution up to constants is given by: $A_N = \frac{1}{\lambda_N}$, $B_N = \frac{\log(N)}{\lambda_N} - \theta_N$ and a Gumbel limit distribution.

Suppose now that we have as previously K blocks of size N, in block J the parameters are λ_J and θ_J. Then trends are as follows:

means: observation $\dfrac{1}{\lambda_J} + \theta_J$ and extremes $\dfrac{1}{\lambda_J} \log(N) + \theta_J$

variance: observation and extremes $(\dfrac{1}{\lambda_J})^2$.

From these formulas, one can see that for the mean trend, the λ variation effect is multiplied by $\log(N)$ and not the θ effect. Of course for other distributions the factor $\log(N)$ changes drastically.

Practically, the interpretation depends on the size N by the way of the distribution: the transfer function associated to the following correspondence *trends of all sample* → *trends of extremes* is quite complicated.

One has to be convinced that in order to be able to understand the previous arrow, it should be necessary to estimate with an extreme precision the distribution of the observed temperatures including the tail and not only the shape parameter. This is quite impossible and in fact it is the reason why even in the stationary case, the use of GEV-POT theory is very general, precisely to avoid the previous, quite impossible, estimation.

References

Brabson BB, Lister DH, Palutikof, Jones PD (2002) Characteristics of extreme temperature in the 21st century. Preprint

Buishand TA (1991) Extreme rainfall estimation by combining data from several sites. Hydrol Sci J 36:345–365

Chavez-Demoulin V (1999) Two problems in environmental statistics: capture-recapture models and smooth extremal models. PhD thesis, EPFL, Lausanne, Switzerland

Coles S (2001) An introduction to statistical modeling of extreme values, Springer Series in Statistics. Springer Verlag

Dacunha-Castelle D, Duflo M (1986) Probability and statistics. T.2. Springer New-York

Dacunha-Castelle D, Gassiat E (1997) Testing the order of a model using locally parameterization: population mixtures and stationary ARMA processes, Annals of statistics no 4

Davison AC (1984) Modeling excesses over high thresholds, with an application. In: Tiago de Oliveira J (ed) Statistical extremes and application. Reidel, Dordrecht, pp 461–482

Davison AC, Smith RL (1990) Models for exceedances over high thresholds (with discussion). J Royal Statistical Soc B 52:393–442

Easterling DR, Evans JL, Groisman P Ya, Karl TR, Kundel KE, Ambernje P (2000) Observed variability and trends in extreme climate events: a brief review. Bull of the AMS

Fougeres P, Genest AL (1997) Nonparametric estimation procedure for bivariate extreme value copulas. Biometrika 84:567–577

Frei C, Schar C (2001) Detection probability of trnds in rare events: theory and application to heavy precipitation in Alpine region. J Climate 14:1564–1584

Green PJ, Silverman BW (1994) Nonparametric regression and generalized linear models: a roughness penalty approach. Chapman and Hall

Gumbel EJ (1958) Statistics of extremes. Columbia University Press, New York, p 375

Huntingford C, Jones RG, Prudhomme C, Lamb R, Gash HHC, Jones DA (2003) Regional climate-model predictions in extreme rainfall for a changing climate. Q J Roy Meteorol Soc 129:1607–1621

Jenkinson AF (1969) Statistics of extremes, In: Estimation of maximum flood. Technical Note 98. World Meteorological Organization, Geneva, pp 183–227

Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. Advances in water resources 25:1287–1304

Keim BM, Cruise JF (1997) A technique to measure trends in the frequency of discrete random events. Climate 11:848–855

Kiktev D, Sexton DMH, Alexander L, Folland CK (2003) Comparison of modeled and observed trends in indices of daily climate extremes. J Climate 16:3560–3571

Klein Tank AMG, Können GP (2003) Trends in indices of daily temperature and precipitation extremes in Europe 1946–1999. J Climate 3665–3680

Leadbetter MR, Lindgren G, Rootzen H (1983) Extremes and related properties of random sequences and series. Springer Verlag, New York

Leadbetter MR, Weissman I, De Haan L, Rootzen H (1989) On clustering of high levels in statistically stationary series. In: Sansom J (ed) Proceedings of the 4th International Meeting on Statistical Climatology. New Zealand Meteorological Service, Wellington

Malek F (2004) Asymptotique des processus de Poisson non stationnaire et application au niveau de retour pour des séries de très hautes températures. Preprint Univ. Paris-Sud

Mearns LO, Katz RW, Schneider SH (1984) Extreme high-temperature events: changes in their probabilities with changes in mean temperature. Journal of Climate and Applied Meteorology 23:1601–1613

Meehl GA, Karl T, Easterling DR, Changnon S, Pielke R, Changnon D, Evans J, Groisman PY, Knutson TT, Kunkel KE, Mearns LO, Parmesan C, Pulwarty R, Root T, Sylves RT, Whetton P, Zwiers F (2000) An introduction to trends in extreme weather and climate projections: observations, socioeconomic impacts and model projections. Bull of the AMS 413–416

Meehl GA, Zwiers F, Evans J, Knutson T, Mearns L, Whetton P (2000) Trends in extreme weather and climate events:issues related to modeling extremes in projection of future climate change projections. Bull of the AMS 427–436

Moberg A, Jones PD (2004) Regional climate model simulations of daily maximum and minimum near-surface temperature across Europe compared with observed station data 1961–1990. Climate Dyn 23:695–715

Pauli F, Coles SG (2001) Penalized likelihood inference in extreme value analyses. Journal of Applied Statistics 28:547–560

Rootzen H, Tajvidi N (1997) Extreme value statistics and wind storm losses: a case study. Scand Actuar J 70–94

Schar C, Jendritsky G (2004) Hot news from summer 2003. Nature 432:559–569

Semenov VA, Bengtsson L (2002) Secular trend in daily precipitation characteristics/greenhouse gas simulation with a coupled AOGCM. Climate Dyn 19:123–140

Smith RL (1989) Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Statist Sci 4:367–393

Smith RL, Shively TS (1995) A point process approach to modeling trends in tropospheric ozone. Atmos Environ 29:3489–3499

Yan Z, Jones PD, Davies TD, Moberg A, Bergström H, Camuffo D, Cocheo C, Maugeri M, Demarée GR, Verhoeve T, Thoen E, Barriendos M, Rodríguez R, Martín-Vide J, Yang C (2002) Trends in extreme temperatures in Europe and China based on daily observations. Climatic Change 52:355–392

Zhang X, Hogg WD, Bonsal BR, WDA (2001) Cautionary note on the use of seasonally varying thresholds to assess temperature extremes. Climatic Change 50:505–507