

# Multidimensional trends: the example of temperature

T.T.Huong Hoang<sup>1,2,a</sup>, S. Parey<sup>2,b</sup>, and D. Dacunha-Castelle<sup>1</sup>

<sup>1</sup> Laboratoire de Mathématiques, Université Paris 11, Orsay, France

<sup>2</sup> EDF/R&D, 6 quai Watier, 78 401 Chatou Cedex, France.

**Abstract.** Our goal is to discuss the notion of multidimensional trend in the climate change context. Temperature series over Europe are used to derive non parametric trends in mean, variance and extremes. A criterion is proposed to compare the different trends, and a general methodology is applied to test if the trends in extremes are only due to the trends in the central field, as mean and variance ones. Examples are given which show that the temporal evolution in mean and variance of the temperature time series are linked, and that generally, these evolutions explain a large part of the temporal evolution of the extremes.

## 1 Introduction

Climate change is generally presented and discussed in term of trends in the mean of climate variables over different spatial scales, from individual observation stations to spatial means over geographical areas or the entire globe. The study of temperature variability, confidence intervals and trends in extremes are also worth of interest. Generally, trends in mean are derived using ordinary least square regression methods [1,2]. On the other hand, trends are usually studied separately: papers are devoted either to trends in mean, or to trends in extreme events. The trends in extreme events are analysed using linear least squares fit on the series of so called extreme indices [3,4]. Mudelsee[5] uses kernel fitting to study flood risk in a non parametrical way.

In this paper, we would like to address the non parametric derivation of trends, as well as the links between trends in different quantities such as mean, variance and extremes, and then illustrate the point by considering temperature series in Europe [6].

Generally, a trend is computed (not defined) as a slow and thus regular component of a time series, superimposed on a series of quite stationary and less variable residuals. In other words, computing a trend consists in extracting some deterministic signal from noisy data. In most studies, as stated before, this is done using ordinary least squares fitting, which will be referred to as classical trends.

The point and main topic considered here can be formulated as follows: once the trend in mean has been estimated, are there other trends which could describe the temporal evolution of the series and how many are significant? Are they independent in some sense? If we leave out of this study, for simplicity, the usual seasonality and all periodic phenomena, our goal is then to look for multidimensional deterministic trends, which, when removed, leave an as stationary as possible residual series. In practice, as we will illustrate on temperature series, the search for

---

<sup>a</sup> e-mail: huong.hoang@math.u-psud.fr

<sup>b</sup> e-mail: sylvie.parey@edf.fr

stationarity has to be stopped when the physical meaning of the trends becomes too weak or when the trends are too strongly linked.

As our goal is to derive non parametric trends in the most objective and general possible way, we first mention the general qualitative properties, requirements and assumptions made to identify a trend. They can be listed as follows:

- **Choice of the time scale:** When observations are given over a time period  $T$  one has to decide which time scale (less than  $T$ ) is of interest to study the signal variation; to simplify we will call it the window length in this work. For example, if the topic of interest is the possible anthropogenic effect on climate change, if one has to consider data over the industrial period, the window length is more or less arbitrarily chosen as 30 years if two to three centuries of observations are available. Over the most recent period, the decade is often chosen as window length if the period length of observations is at most fifty years. The window length for a given observation period is indeed often chosen in a quite arbitrary way and this choice is very important for instance in the case of the increasing mean temperature. We will show in the paper that such a choice based on informal or subjective ideas (heuristics) can be improved when using more intrinsic (mathematical) considerations.

- **Almost invariance by data extension and localisation:** Once the length  $T$  of the observations has been fixed, if we add new observations whose period length is a significant fraction of  $T$  ( $T/2$  for example), then the modifications on the previously computed trend on any subinterval of the initial dataset have to be small. Of course, if we consider a trend estimated on a 50-year period of observations, and then include this period in a larger period of say 300 years, the trend is obviously modified and the intrinsic smoothness will change. Nevertheless, intrinsic choices allow to give an objective interpretation of this modification and can help to define changes or breaks in the trend behaviour.

- **Monotonicity:** In order to easily interpret the physical phenomena, the monotonicity can be imposed to a trend. For instance if trends are used in order to define an extrapolation necessary to compute return levels, monotonicity is required in general. Variants are in constraints on the number of intervals where the trend is monotone or convex. There are general tools to obtain constrained estimates as isotonic regression (see [7] for example).

Starting from these considerations, the paper is devoted to the identification of trends in mean, variance and extremes of a time series, together with their possible links. It is organised as follows: first, the statistical framework is given in section 2 before coming to the results for trends in mean, variance and extremes of temperature series in Europe in section 3. The link between the trends in mean and variance and the trends in extremes is discussed in section 4 and a methodology is proposed to test if the trends in extremes are due to trends in mean and variance, before concluding in section 5.

## 2 Statistical framework

### 2.1 Trend in mean

We first discuss the trend in mean, with the supposition that there is no seasonal trend (or any trend with an obvious shorter window, this point can be checked for example using a wavelet analysis).

The basic approach uses a moving average procedure, with a moving window length  $L$ . In this simplest case as for the more complicated forthcoming, we define the model, here :

$$X(t) = m(t) + e(t) \tag{1}$$

where  $X(t)$  is the observation series with unknown mean  $m(t)$  and  $e(t)$  is a centred process, which is expected in general to be stationary and uncorrelated.

This "signal+noise" approach is first improved in using non parametric statistics, and then often followed by the choice of a regression model in parametric statistics.

Non parametric methods satisfy the previously given requirements, but now the window length of interest is chosen in an intrinsic way and has a precise definition. The main lines are the following :

-We suppose that  $m(t)$  is in a space  $M$ , for instance that of the two continuously differentiable functions.

-We suppose that  $e(t)$  is stationary. An important point must not be forgotten: the statistical tools used to identify  $m(t)$  depends on the properties of  $e(t)$ , for instance of its correlation or of its distribution.

-A criterion is chosen to define a "good estimate" of  $m(t)$ , for instance the minimization of the integrated quadratic error.

$$Err = \int_{t \in D} [m(t) - \hat{m}(t)]^2 \quad (2)$$

which requires the control of the bias and of the variance of the estimator (it exists of course the same but local criteria for fixed  $t$ ).

There are two families of methods: smoothers (splines, kernels, loess, lasso) and approximations (Fourier, orthonormal systems as Fourier series, wavelets families). See [7–11] for a general presentation of these basic statistical tools.

In all cases, the **intrinsic** window length is associated to a **tuning parameter** of the method (penalization parameter or kernel width in the first case, threshold for coefficients in the second one). This parameter is estimated from the data (adaptative estimation), normally with *independent data*, using methods such as the cross validation (CV) or the generalized cross validation (GCV) or other criteria as AIC (see [12–14]). Besides we have another parameter, the **degrees of freedom**, calculated by the trace of the smoother matrix  $df = tr(S)$  (see [13]) which is directly linked to the tuning parameter. The more df is large, the less the function is smooth. The degree of freedom is the only common parameter which allows to compare different smoothing methods such as loess and splines. It fixes the window length which becomes intrinsic because the degree of freedom only depends on the data without any a priori choice [7]. Similar notions can be inferred for approximation methods. Details are given in [30] about the technical use of CV and so the mathematical choice of the degrees of smoothness.

The methods are developed and deeply studied in the case  $e(t)$  is a sequence of independent and equidistributed variables. In the case of *dependent data*, they can be used with suitable and precise specifications and modifications depending on the hypothesis made on the process  $e(t)$ : for instance  $e(t)$  has a (short) memory (autocorrelation) or the variance of  $e(t)$  is not constant. Examples are given for temperatures in our previous work [15]. It remains to obtain more general results in mathematical statistics to improve the intrinsic character of the tools.

Confidence bands are obtained by bootstrap using also modifications (very close to that necessary for the choice of the tuning parameter) to take into account the properties of  $e(t)$  [16–18].

## 2.2 Trend in variance

To derive the trend in variance, we first choose a model, here :

$$X(t) = m(t) + s(t)\epsilon(t). \quad (3)$$

We shall use sequential plug in methods: once the trend  $m(t)$  for the mean has been estimated (as  $\hat{m}(t)$ ), it is then natural to look for a trend for the variance  $s^2(t)$  of the same observations with the same window length of interest.

The model is now:

$$Y(t) = X(t) - \hat{m}(t) = s(t)\epsilon(t) \quad (4)$$

where  $\epsilon(t)$  is a centred process with variance one, once again expected to be stationary and uncorrelated. The purpose then is to estimate  $s(t)$ .

For a large class of hypotheses on  $\epsilon(t)$  (like for  $m(t)$ ), we can proceed as for  $m(t)$  but on the data  $Y^2(t)$ . It is worth noticing here that the fact to plug  $\hat{m}(t)$  instead of  $m(t)$  in the data in order to obtain  $Y^2(t)$  does not modify, at least for large datasets, the efficiency of the usual non parametric methods [5]. Of course, here as for  $m(t)$ , statistical results are needed to extend to very general  $\epsilon(t)$  the use of the tools.

Periodic seasonality and global warming are not independent, global warming strongly depends on the season (even on the month) of the year(see [19]). Technically, this is a difficult point .What happens for seasonal trends if these trends in mean and variance exist? We do not discuss this problem here, as it is neither a problem of usual separation of frequencies nor a problem of space time decomposition usually solved using wavelets for instance. Here the main problems are that of the identification of parameters because seasonality is not invariant for the time scale defined by the mean and variance trends, and the mean and variance trends are highly seasonal. In a parametric context it is a simple problem but it needs tricky solutions in a non parametric context. We shall discuss it in a forthcoming paper.  $\epsilon(t)$ , as defined above, is not a stationary process because in general, and it is the case for temperatures, its dynamics depends on the season. Of course all the difficulties previously mentioned disappear, or at least are much attenuated, if one limits the study to a homogeneous part of the year, for instance the 2 hottest months.

### 2.3 Trend in extremes

It is suspected that if a warming effect is observed for temperature, more precisely increasing trends in mean and/or variance, then some linked phenomena has to be expected for the extremes. For the extremes, the statistical procedures are based on an asymptotic result coming from the probability theory [20], which requires a large amount of data (the convergence is slow) and some regularity conditions on the tails of distributions, in general quite impossible to check but accepted for their weakness. Two other conditions have to be verified in order to have a limit theory for the extremes:

- does the temporal structure of the process  $\epsilon(t)$  adjust to a theory of extremes for this kind of process?
- what is determinant for the extremes (what makes the extremes?): the deterministic trend  $m(t)$  or the random fluctuations of  $\epsilon(t)$ ?

In our applications, the effect of the deterministic trend  $m(t)$  during a short period (some months) can be neglected, with a statistical evidence, with respect to the random fluctuation  $\epsilon(t)$ , whose distribution has parameters which are the extreme trends. Now what happens when we consider a larger time scale?

We will consider here the classical block maxima approach [20]. The time period of observations is divided into  $n$  consecutive blocks of days of same length. Then, we select into each block of time the maximum value of the observations  $X(t)$  obtaining a sequence  $M_1, M_2, , M_k, , M_n$  of the maxima. We suppose that the independence or at least a weak dependence of the values is insured if the length of the blocks is sufficiently large. In this case, the probability theory

allows to suppose that there exist sequences  $\mu_k$  and  $\sigma_k > 0$  such that the distribution of  $M_k$  is approximated by one of  $\mu_k + \sigma_k Z_{\xi_k}$  where  $Z_{\xi_k}$  is a sequence of independent variables with a distribution  $G_{\xi_k}$ .

$$\begin{aligned} &\text{if } \xi_k \neq 0, G_{\xi_k}(z) = \exp -(1 + \xi_k z)^{-1/\xi} \text{ when } 1 + \xi_k z > 0; = 1 \text{ if not} \\ &\text{if } \xi_k = 0, G_{\xi_k}(z) = \exp (-\exp (-z)). \end{aligned} \quad (5)$$

Then the discrete model of extremes is:  $M_k = \mu_k + \sigma_k Z_{\xi_k}$  and the trends for the extremes are  $\mu_k$  and  $\sigma_k$  ( $k$  being the block index) which will be extended into continuous functions of time  $t$ .

The expectation and variance of the distribution are expressed as:

$$E(Z) = \frac{1}{\xi_k}(\Gamma(1 - \xi_k) - 1); V(Z) = \frac{1}{\xi_k^2}(\Gamma(1 - 2\xi_k) - \Gamma^2(1 - \xi_k)) \quad (6)$$

So if  $M_k$  is the maximum in a block  $k$  with distribution  $G_{\xi_k}$ , we have:

$$E(M_k) = \mu_k + \sigma_k E(Z); V(M_k) = \sigma_k^2 V(Z) \quad (7)$$

where  $\mu_k$ ,  $\sigma_k$  and  $\xi_k$  are functions of the block index  $k$  and respectively location parameter, scale parameter and shape parameter of  $G_{\xi_k}$ . The location parameter of the GEV models is not a natural parameter, it mixes mean and variance. Thus we often consider in the following the mean and the variance of extremes, which are more easily interpretable.

We first supposed that  $\xi$  does not depend on time (see the discussion on this hypothesis in [21]).  $\mu_k$  and  $\sigma_k > 0$  depending on the block index  $k$  are then the only trends. To be compared with the trends in mean and variance, they must be defined for every time  $t$ . Thus, we now have to estimate the trends  $\mu_X(t)$  and  $\sigma_X(t)$  (subscript  $X$  referring to the observation series  $X(t)$ ) and also  $\xi_X$  in using a non parametric method based in any case on the likelihood, and not on least squares like for the mean trend (see [22, 23]). We use cross validations to choose the smoothing parameters for  $\mu_k$  and  $\sigma_k$ . These problems of optimization and choice of the tuning parameters are quite difficult because of the correlation between the parameters (see [22]).

## 2.4 Measure of the similarity of trends

The different trends considered here, for instance trends in the temperature mean and variance, show some similarities and are supposed to be both affected by climate change. Thus, for physical reasons, they have to be considered on the same time scale (or with the same window length).

Furthermore, the dependence between the different trends can be measured, as for random series, by a formal correlation defined as the scalar product of their reduced form, respectively called  $nf$  and  $ng$ , centred around their mean  $\bar{f} = \int_{t \in D} f(t) dt$  and normalized by

$$\begin{aligned} &\sqrt{\int_{t \in D} (f(t) - \bar{f})^2}: \\ &r(f, g) = \int_{t \in D} \langle nf(t), ng(t) \rangle dt \end{aligned} \quad (8)$$

$D$  is the ensemble of the dates in the total time period  $T$ .

This correlation is a measure of the linear dependence of  $f$  and  $g$  and of the quality of the approximation of  $g(t)$  by a function of the form  $a + bf(t)$ . To analyse more accurately the links between trends, we can also calculate the correlation coefficient between the first derivatives of  $f$  and  $g$  which is a measure of the quality of the approximation once the best linear trend has

been removed. Another indicator of the similarity of trends is the number and the distances between the locations of local extremes of the curves. Other quantities can be used in order to measure the likeliness between functions  $f$  and  $g$  including non linear transformations, for instance one can estimate  $\theta$  such that the correlation between  $f$  and  $\theta(g)$  is maximal.

### 3 Results for the European temperature series

#### 3.1 Mean and variance trends

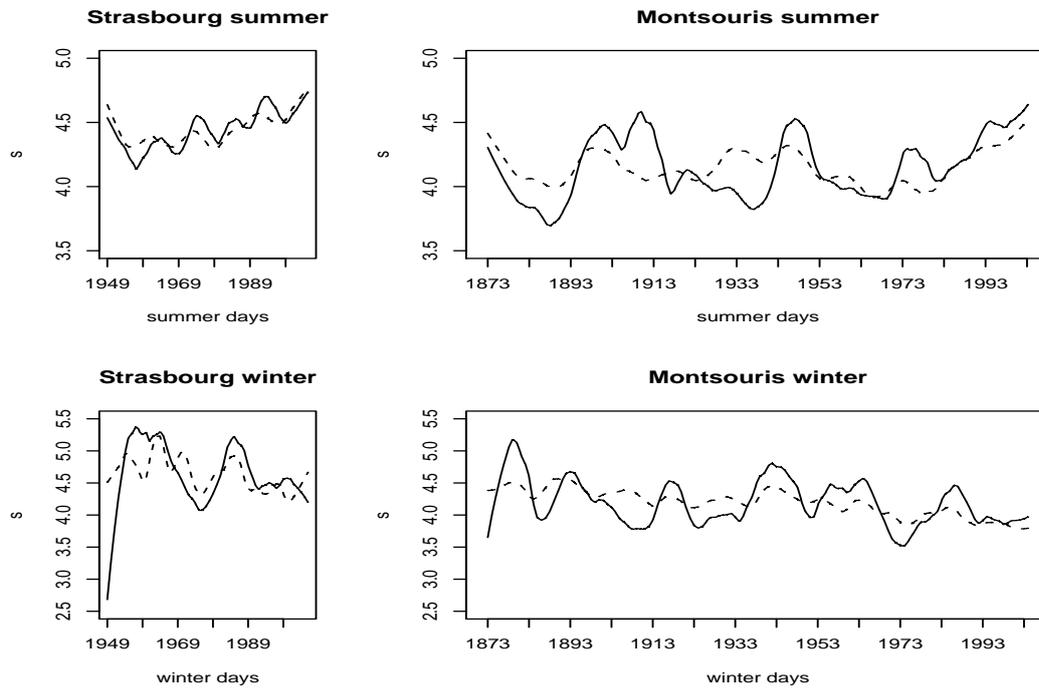
First we computed the trends in mean and variance of daily minimum temperature in winter and daily maximum temperature in summer for 2 observation series in France: Paris-Montsouris over the 1873-2003 period (131 years) and Strasbourg over the 1949-2005 period (57 years). This avoided to deal with the seasonal effect, as previously discussed in section 2.2. The summer is defined as the 100 days between the 14th of June and the 21st of September and the winter as the 90 days between the 1st of December and the 28th of February. These periods have been chosen because the extremes mainly occur during these times. In our applications to temperature series in Europe [15,24], we found degrees of freedom corresponding to a local window length of 20% of the total length, it means 11 to 15 years depending on the station. Therefore, the trends are computed here with loess and a smoothing parameter corresponding to an 11-year window length for Strasbourg and a 26-year window length for Paris-Montsouris. The steps are then the following:

- selection of the days of the studied season over the total period length (this leads to a series of  $n_{year}100$  days for summer and  $(n_{year}-1)90$  days for winter,  $n_{year}$  being the number of years in the total period length)
  - computation of  $\hat{m}(t)$  from  $X(t)$  using loess
  - computation of  $[X(t) - \hat{m}(t)]^2$
  - computation of  $\hat{s}^2(t)$  from  $[X(t) - \hat{m}(t)]^2$  using loess.
- where the hat notation corresponds to estimates.

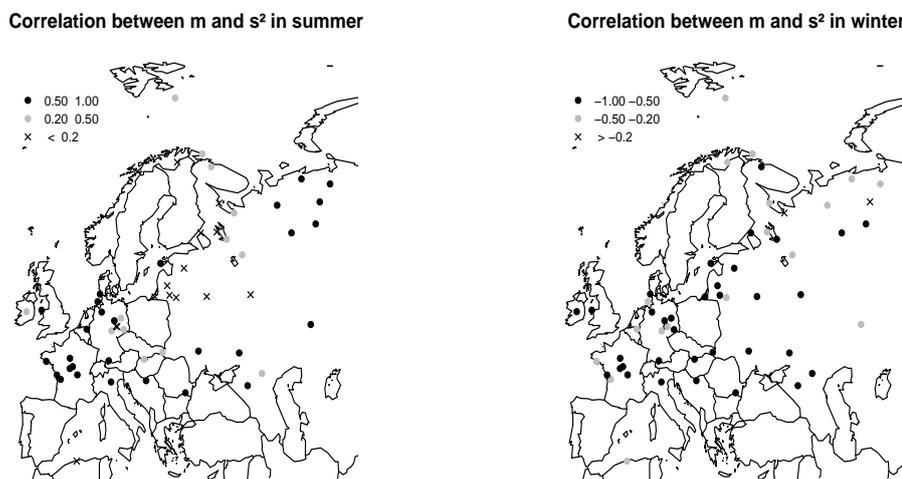
When computing the trends in mean and variance, it came out that their evolutions seemed linked: variance increasing when mean increases in summer and when mean decreases in winter. In order to verify this visual link, the standard deviation evolution is super-imposed with that obtained from a linear regression regarding the mean evolution. The results are shown in figure 1 (the length of the plotting area is proportional to the period length of each series). Although the linear reconstruction is far from perfect, it globally shows a similar evolution. In summer, the correlation is positive: when mean increases, variance increases; whereas in winter, the correlation is negative: when mean increases, variance decreases.

In particular, in summer, the linear relationship seems stronger for the last 50 years in Paris-Montsouris than for the beginning of the period, coherently with the result found for Strasbourg for which only the last period is available. While in winter, this relationship is similar for whole times.

In order to verify if this is a common behaviour, the correlation coefficients between the 2 functions of time  $\hat{m}(t)$  and  $\hat{s}(t)$  have been computed for a set of 55 observation series in Europe provided by the European Climate Assessment and Dataset (ECA&D) project [6], with period lengths between 40 and 100 years. As mentioned in section 2.4, the correlation coefficients are used as a measure of the likeliness of the curves. The results are plotted in figure 2, and show that this is a quite systematic behaviour in winter (correlation coefficients generally lower than -0.5, represented by black dots), while it is also true in summer, although less systematically (most points show a correlation coefficient larger than 0.5, represented by black dots). A more detailed study of this link can be found in [15].



**Fig. 1.** Standard deviation function  $\hat{s}(t)$  (solid line) and the estimate of  $\hat{s}(t)$  as a linear function of  $\hat{m}(t)$  :  $\hat{s}(t) = a\hat{m}(t) + b$  (dashed line) for summer (top panels) and winter (bottom panels) and for Strasbourg (left) and Paris-Montsouris (right). Note that for the summer  $a > 0$  and for the winter  $a < 0$ .



**Fig. 2.** Correlation coefficients between the evolution of the mean and the variance of daily minimum temperature in summer (left) and daily maximum temperature in winter (right) for the 55 ECA&D series.

### 3.2 The residuals

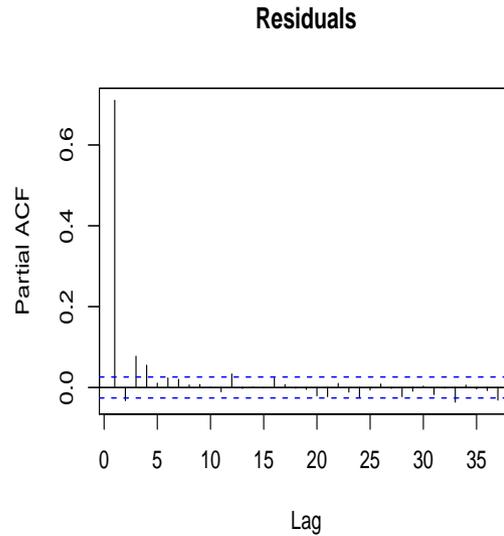
Now we consider that the data have been reduced, since the seasonal effect is avoided by selecting the days of the hot and cold seasons respectively and the mean and variance trends have been removed; it remains a process  $\epsilon(t)$  defined as:

$$\epsilon(t) = \frac{X(t) - \hat{m}(t)}{\hat{s}(t)} \quad (9)$$

where  $\epsilon(t)$  as random process can have a dynamic which is not stationary.

Once  $\epsilon(t)$  has been estimated as residuals, it is necessary to check the hypothesis done on  $e(t)$  or  $\epsilon(t)$  which allows the application of the detrending procedures. So one has to study the stationarity of  $\epsilon(t)$ , its autocorrelation, the boundness and the stationarity of its moments of order larger than 2, and then, if possible, to select a parametric model for  $\epsilon(t)$ , but this is not at all a requirement in order to apply the detrending procedures. An amount of non stationarity is possible in order to use the mathematical basis of the method of estimation of trends.

For temperature,  $\epsilon(t)$  is often modelled in a quite convenient way, by considering it as stationary, with an autocorrelation of slow order  $p$ ,  $AR(p)$ , chosen by AIC criterion. As an example see figure 3 representing partial-autocorrelation in the case of the Strasbourg temperature series in summer, for which  $p=4$ .



**Fig. 3.** Partial- autocorrelation of the residuals  $\epsilon(t)$  for the temperature series of Strasbourg in summer, the lag being expressed in days

Thus, one can fit for instance a model which has the form:

$$X(t) = \sum_{k=1}^p X(t-k) + a(X(t))\eta(t) \quad (10)$$

where  $a$  is an estimated non linear function bounded and uniformly lipschitz, which is zero for  $X$  large enough in the summer, and  $\eta$  a sequence of independent centred random variables. But in general this modelling cannot take completely into account the behaviour of the extremes

even if it explains the behaviour of the very large (or low) percentiles (see [25]).

### 3.3 Trend in extremes

Applying the cross-validation with a specific algorithm to the time evolution of the parameters of the GEV model (5) for the European temperature time series, we find degrees of freedom corresponding to the local (intrinsic) window length of 15 to 20 years. It is larger than the window length of the central field ( $m(t)$  and  $s(t)$ ). This can be explained by the fact that the extremes are more variable.

We then compute the trends for the mean of the series and of the extremes ( $m(t)$  and  $m_X(t)$ ), where  $m_X$  refers to the mean of the extremes, estimated from the GEV parameters as stated in formulas (6) and (7) and for their variances ( $s^2(t)$  and  $s_X^2(t)$ ),  $s_X^2$  being as for the variance of the extremes, expressed from the GEV parameters in (6) and (7). Here  $X(t)$  are the summer daily maximum temperatures or the winter daily minimum temperatures observed in Paris-Montsouris over the 1873-2003 period and in Strasbourg over the 1949-2005 period, defined in the same way as previously. In computing these trends, we obtained very parallel temporal evolutions. Thus here again, we chose to plot the evolution of the mean or standard deviation function of the extremes together with their estimation via a linear regression with the mean or standard deviation of the whole series. The results are shown in figure 4.

We can see that the evolutions of the mean of the whole observations and that of the extremes are strongly linked, whereas the same is true, although in a less systematic way, for the standard deviations. In the same way as previously, we verified the generality of the link in computing the correlation coefficients between the curves for  $m(t)$  and  $m_X(t)$  for the 55 ECA&D series over Europe. The results shown in figure 5 confirm that the two means are strongly linked, as well in winter as in summer, with correlation coefficients generally larger than 0.5 (black dots).

## 4 Trends in extremes and trends in mean and variance

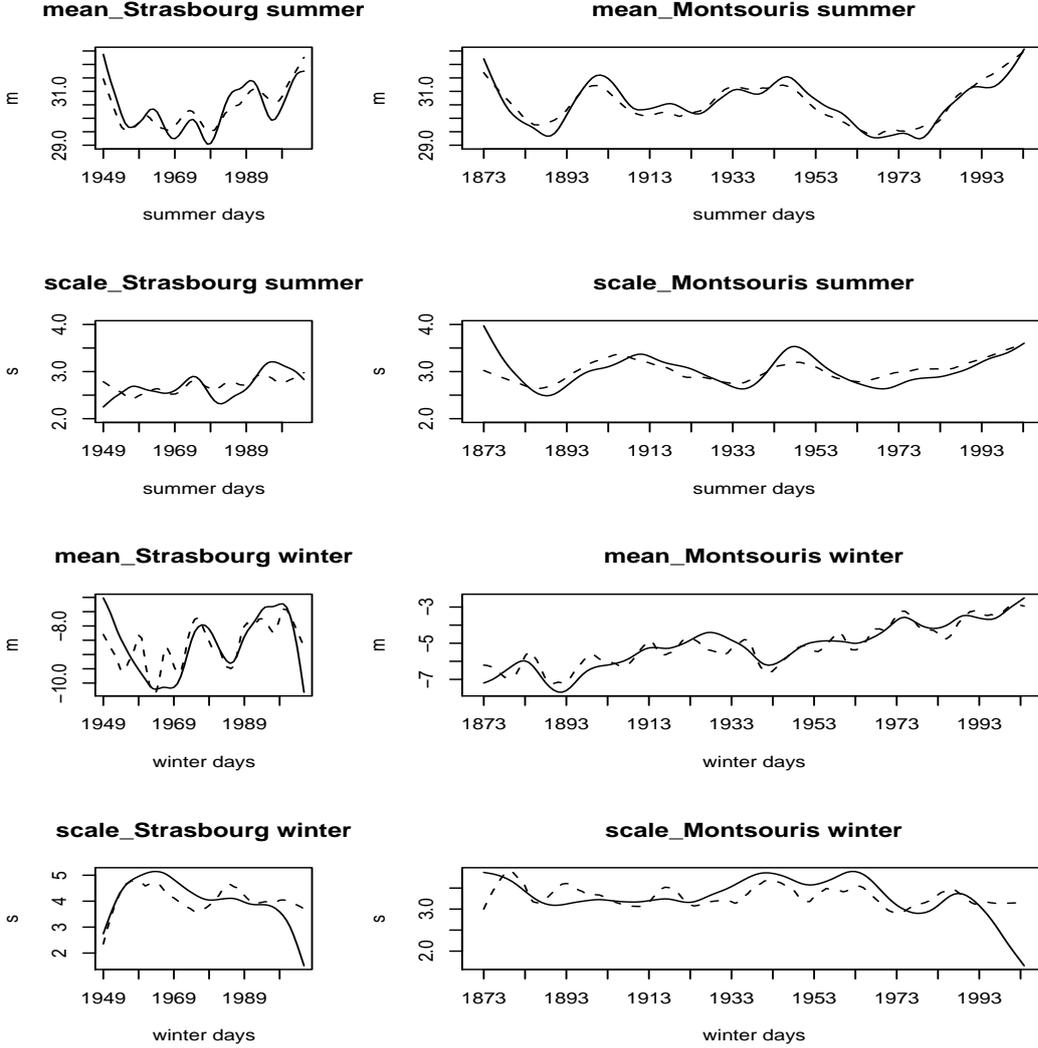
### 4.1 The K hypothesis

As can be seen from the previous results, trends in total mean and in the mean of the extremes, as well as trends in total variance (or standard deviation) and in the variance (or standard deviation) of the extremes look very similar. The question is then: is there some systematic link between these trends, and how can this be tested? Consider now  $Y(t)$  the reduced data:  $Y(t) = \frac{X(t)-m(t)}{s(t)}$ . Some works [26–28] address the so called H hypothesis:  $Y(t)$  is a stationary sequence. It is, as we shall see, a too strong hypothesis and so we choose to study the weaker K hypothesis: the tails of  $Y(t)$  are stationary, which is equivalent to say that the parameters of  $ext_Y$ :  $\mu_Y$ ,  $\sigma_Y$ ,  $\xi_Y$  are constant, where  $ext_X$  ( $\mu_X$ ,  $\sigma_X$ ,  $\xi_X$ ) and  $ext_Y$  ( $\mu_Y$ ,  $\sigma_Y$ ,  $\xi_Y$ ) are respectively the GEV models for  $X(t)$  and  $Y(t)$ .

Their parameters are linked in the following way:

$$\begin{aligned}\xi_X(k) &= \xi_Y(k) = \xi \\ \sigma_X(k) &= \sigma_Y(k)s(k) \\ \mu_X(k) &= m(k) + \mu_Y(k)s(k)\end{aligned}\tag{11}$$

$k$  being the block index.  $\sigma_Y(k)$  appears as the multiplier allowing to pass from the standard-deviation  $s(t)$  of the global sample at time  $k$  of the maximum of the temperature in block  $k$ , to the scale parameter of extremes. In fact, as  $s(t)$  is a smooth function,  $s(k)$  is the standard

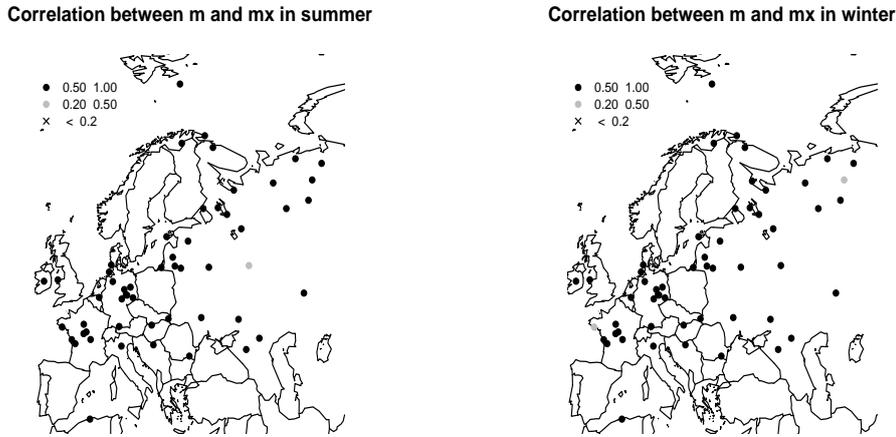


**Fig. 4.** Estimates of mean function and standard deviation function of the extremes and their linear estimates in function of the mean function, standard deviation function of the whole observations  $\hat{m}_X(t) = a\hat{m}(t) + b$ ,  $\hat{s}_X(t) = a\hat{s}(t) + b$  (dashed line) for summer (top) and winter (bottom) seasons in Strasbourg (left) and Paris-Montsouris (right)

deviation for the days of block  $k$ . The same role for  $\mu_X(k)$  is played by  $\mu_Y(k)$  but here again multiplied by  $s(k)$ .

If  $\mu_Y(k)$  and  $\sigma_Y(k)$  can be considered as constant, which corresponds to the validity of the K hypothesis, we can see that the difference of behaviour between the trends of the whole sample and those of the extremes is completely due to the mean and variance trends. And then, according to (6) and (7), we have:

$$\begin{aligned} m_X(k) &= m(k) + s(k)(\mu_Y(k) + \sigma_Y(k)E(Z)) \\ s_X(k) &= s(k)\sigma_Y(k)/\xi.(V(Z))^{-1/2} \end{aligned} \tag{12}$$



**Fig. 5.** Correlation coefficients of the mean of the central field against the one of the extremes in summer (left) and in winter (right) for the 55 European ECA&D temperature series.

If  $K$  is true, these formulas show that the strong almost affine relation found between  $m(t)$  and  $s^2(t)$  implies a strong relation between the mean  $m(t)$  and the mean function  $m_X(t)$  of the extremes.

#### 4.2 How to test $K$ ?

This is an example of a general method for testing the absence of trends.

The set of possible evolutions of  $ext_Y$  parameters is very large. We have previously more or less excluded a possible evolution of  $\xi$  [23], but there are many subtle deformations of the distribution of  $Y(t)$ . So any test suffers from the absence of natural alternatives. It is why we prefer the use of a distance  $\Delta$  between two functions of time, defined as:

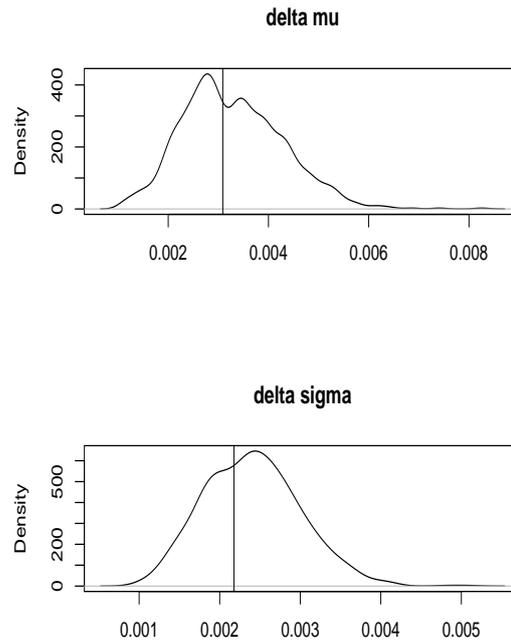
$$\Delta(f, g) = \left( \int_{t \in D} (f(t) - g(t))^2 \right)^{1/2} \quad (13)$$

to test the  $K$  hypothesis. If we estimate a function of time  $f$  by  $g$ ,  $\Delta(f, g)$  is a measure of the quality of  $g$  as an estimate of  $f$ . Here, we have two estimates of the parameters of  $ext_Y$ :  $\tilde{f}$ , the non parametric estimate, and  $\hat{f}$  the constant estimate obtained if  $K$  is true, that means if  $ext_Y$  is stationary. Now, in any case,  $K$  being true or false, one can prove that  $\tilde{f}$  converges to  $f$  when the sample size  $T$  tends to infinity, the speed of convergence depends on the supposed smoothness of  $f$  (see [29]). The situation is of course different for  $\hat{f}$ , if  $K$  is true it converges to  $f$  with a speed of the order of  $\sqrt{T}$  and in this case  $\Delta(\hat{f}, \tilde{f})$  is, for a large sample, very close to  $\Delta(f, \tilde{f})$ . On the contrary if  $K$  is false,  $\hat{f}$  converges in general to a constant (this is a theoretical result, [28,30]) which is of course different from  $f$  and even if it does not converge;  $\Delta(\hat{f}, f)$  does not tend to zero and remains larger than some  $A > 0$  and the same is true for  $\Delta(\tilde{f}, \hat{f})$ . The intuitive reason is that we try to find  $f$  in a set of functions far away from  $f$  if  $K$  is false. These ideas could be translated in an asymptotic result. We prefer the use of a more numerical

approach based on simulation. See [30] for details.

Our proposed solution is then to construct by simulation or bootstrap the distribution of  $\Delta(\hat{f}, \tilde{f})$  under the K hypothesis, that is the distribution of the distances between the non parametric estimates and the constants in case we know that the parameters are constant. To do this, we simulate 1000 samples of the stationary GEV  $(\mu_Y, \sigma_Y, \xi_Y)$  distribution with the same size as the series of the maxima of  $Y(t)$  (an alternative will be to use bootstrap of the residuals for the model of extremes with very similar results). From each sample, we estimate the extreme parameters in two ways: first, in considering them as constant; secondly, in considering them as functions of time. Then we calculate the distances between these two estimations of the parameters and so we obtain a distribution of the statistical error of estimation under the K hypothesis. If our true distances are found lower than the 90th percentile, the K hypothesis is considered as satisfied: the distances are only due to statistical errors.

What happens if  $ext_Y$  is not stationary (from the test)? This is a very difficult question, probably without any precise answer with these data. For example, when  $\mu_Y(k)$  can be considered as constant and  $\sigma_Y(k)$  as a function of  $k$ , so the upper bound of the distribution of  $Y(k)$  which is the same as that of  $X(k)$  (taken at time  $t$  which is the date of the maximum in block  $k$ ) is  $\mu_Y(k) - \sigma_Y(k)/\xi$ . The deformation of the tail of the distribution can be described by this trend on the upper bound (of course to keep a mass one to the probability this is compensated by a slight modification of the central part of the distribution). We tested the K hypothesis



**Fig. 6.** The distribution of distances under the K hypothesis with the real values of the same distance for the observations (vertical lines) in summer for the series of Deols and for the location parameter  $\mu$ (top panel) and the scale parameter  $\sigma$ (bottom panel).

using this method for the series in Paris-Montsouris, Strasbourg and all European temperature series. An example is shown in figure 6 for the Deols temperature series in summer (ECA&D

series over the period 1901-2006). The distances computed from the observed series (for both the location  $\mu$  and scale  $\sigma$  parameters) lie inside the distribution of the distances when K is satisfied, thus in this case, the K hypothesis can be accepted. When considering the 55 series over Europe, the hypothesis can be accepted in about 80% of the cases, but differences appear between summer and winter which need to be further analysed.

## 5 Conclusion

In this paper, we discussed the requirements and assumptions currently associated to the identification of trends in data series and applied non parametric statistical methods to derive trends in European temperature series. We could then show that the trends in mean and variance are linked, and that this link induces a strong relation between trends in the whole dataset and trends in extremes. We developed a testing procedure to verify if the extremes of the centred and normalized series are stationary. If the hypothesis is verified, then the trends in extremes are only due to trends in mean and variance of the whole dataset. This analysis has been conducted for an important number of European temperature series, and showed that this is the case for the majority of the series in Europe. Further detailed analyses have then to be conducted in order to analyse the different behaviours and to identify some physical explanations.

## References

1. Angell J.K., J. Clim. **16**,2288-2295 (2003).
2. Percival D.B., Rothrock D.A., J. Clim. **18**, 886-891 (2005)
3. Kiktev D., Sexton D.M.H. Alexander L., Folland C.K., J. Clim. **16**, 3560-3571 (2003)
4. Klein Tank A.M.G., Knnen G.P., J. Clim **16**, 3665-3680 (2003)
5. Mudelsee M., Deutsch M., Brngen M., Tetzlaff G., Hydr. Sciences J. **51**, 833 (2006)
6. Klein Tank A.M.G. *et al.* Int. J. of Clim., **22**, 1441-1453 (2002)
7. Z. Wu, N.E. Huang, S.R. Long, C.K. Peng, proc. Nat. Acad. Sci. USA September **18**, 104(38) 148896164894 (2007)
8. J. Fan and I. Gijbels, *Local polynomial modelling and its applications*. Chapman & Hall (1996)
9. P. Green and B. Sliverman, *Nonparametric regression and generalized linear models*. Chapman & Hall (1994)
10. D. Ruppert, M.P. Wand, U. Holst and O. Hossjer, Technometrics **39**, 262-273 (1997)
11. J.S. Simonoff, *Smoothing methods in statistics*. Springer Series in Statistics (1998)
12. C. Loader, *Local Regression and Likelihood*. Springer (1999)
13. Hastie, T.J. and Tibshirani, R.J., *Generalized Additive Models*. Chapman and Hall (1990)
14. M.H. Neumann, Ann. Statist. **23**, 1937-1959 (1995)
15. S. Parey, D. Dacunha-Castelle, T.T.H. Hoang, submitted to Clim. Dynam. (2008)
16. P. Hall, Ann. Statist. **20**, 695-711 (1992)
17. S.N. Lahiri, *Resampling methods for dependent data*. Springer (2003)
18. M.H. Neumann, Statistics **29**, 1-36 (1997)
19. Schönwise, C.D, Rapp, J., Fuchs, T. and Denhard, M., Meteorologische Zeitschrift **3**, 22-28 (1994).
20. S. Coles, *An Introduction to Statistical Modelling of Extreme Values*. Springer (2001)
21. M. Nogaj, S. Parey, D. Dacunha Castelle, Nonlin. Processes. Geophys. **14**, 305316 (2007)
22. V. Chavez-Demoulin, A.C. Davison, J. Appl. Statist. **54**, 207-222 (2005)
23. A.C. Davison, N.L. Ramesh, J.R. Statist. Soc. B. **19** (1999)
24. P. Yiou, D. Dacunha-Castelle, S. Parey, T.T.H. Hoang, Geophys. Res. Lett. (2009)
25. G. Benmenzer, D. Dacunha-Castelle, T.T.H. Hoang, submitted to Envirometrics (2008)
26. C. A. Ferro, A. Hannachi, D. Stephenson, J. Clim. **18**, 43444354 (2005)
27. R. Katz, B. Brown, Clim. Change **21**, 289302 (1992)
28. C. Schär, P.L. Vidale, D. Lthi, C. Frei, C. Hberli, M.A. Liniger and C. Appenzeller, Nature **427**, 332-336 (2004)
29. D. Dacunha-Castelle and M. Duflo, *Probabilits et Statistiques vol 2*. Masson (1983) and Springer (1984)
30. T.T.Huong Hoang, S. Parey, D. Dacunha-Castelle, preprint, Department of Mathematics Paris-Sud Orsay (2009)