



**HAL**  
open science

## Return times of hot and cold days via recurrences and extreme value theory

Davide Faranda, M Carmen Alvarez-Castro, Pascal Yiou

► **To cite this version:**

Davide Faranda, M Carmen Alvarez-Castro, Pascal Yiou. Return times of hot and cold days via recurrences and extreme value theory. *Climate Dynamics*, 2016, 47, pp.3803-3815. 10.1007/s00382-016-3042-6 . hal-01232444

**HAL Id: hal-01232444**

**<https://hal.science/hal-01232444>**

Submitted on 30 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Return times of hot and cold days via recurrences and extreme value**  
2 **theory**

3 **Davide Faranda · M. Carmen Alvarez-Castro · Pascal**  
4 **Yiou**

5  
6 Received: November 16, 2015/ Accepted:

7 **Abstract** In this paper we introduce a model evaluation and comparison metric based on the method-  
8 ology introduced in Faranda et al (2013) to assess biases and their potential origins in a historical model  
9 simulation against long-term reanalysis. The metric is constructed by exploiting recent results of dynam-  
10 ical systems theory linking rare recurrences to the classical statistical theories of extreme events for time  
11 series. We compute rare recurrences for 100 years daily mean temperatures data obtained in a model  
12 with historical greenhouse forcing (the Institut Pierre-Simon Laplace, IPSL-CM5 model) and compare  
13 them with the same quantities obtained from two datasets of reanalysis (20 Century Reanalysis and  
14 ERA 20C). The period chosen for the comparison is 1900-2000 and the focus is on the European region.  
15 We show that with respect to the traditional approaches, the recurrence technique is sensitive to the  
16 change in the size of the selection window of extremes due to the conditions imposed by the dynamics.  
17 Eventually, we study the regions which show robust biases with respect to all the techniques investigating  
18 the possible origins.

19 **Keywords** Climate · Dynamical Systems · Extreme events · Recurrences · Temperature

---

Davide Faranda  
LSCE, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France  
Tel.: +33-169081142  
Fax: +33-169087716  
E-mail: [davide.faranda@lsce.ipsl.fr](mailto:davide.faranda@lsce.ipsl.fr)

M. Carmen Alvarez Castro  
LSCE, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France

Pascal Yiou  
LSCE, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France

## 1 Introduction

Temperature extremes at one location are generally defined from the tail of the temperature probability distribution (Katz and Brown 1992; Coles 2001). This definition focuses the data analyses of the maxima/minima of time series or exceedances over high thresholds. This has justified the application of extreme value theory (EVT) in climate sciences, to assess the statistical properties of large values of temperature, precipitation or wind speed (Nikulin et al 2011).

EVT (Leadbetter et al 1983) has been widely used to study temperature extremes (Ghil et al 2011; Katz 2010). This statistical approach models the probability distribution of high values with Generalized Extreme Value distributions (GEV) or Generalized Pareto Distribution (GPD) (Gnedenko 1943; Pickands III 1975).

One of the outcomes of those statistical approaches is the computation of return periods for extreme events (Galambos 1980). Once the probability distribution of the extremes is modeled by a GEV or a GPD law, one is able to compute return level for events with a return period that is larger than the length of observations, i.e. a probability so small that they have not yet been observed in the available data.

A limitation of a straightforward application of EVT to climate variables is that it focuses on large events and does not treat events that are described by moderate values of a physical variable. One way of illustrating this limitation is to look at a chaotic dynamical system with (at least) one unstable fixed point near the center of the attractor, like in the system of Lorenz (1963). In such a system, two of the unstable fixed points are located near the center of the strange attractor and never reach extreme values in the phase space. The occurrence of such events escape an analysis with EVT, although they are rare because of their dynamical instability (Lucarini et al 2014).

We propose to extend the notion of return period of extreme events to rare events. This proposition is based on the application of EVT to the dynamical information obtained from the statistics of recurrences of events. The recurrence approach hails from dynamical systems theory and has been known since the work of Poincaré (Frisch 1956). However, only recently it has been proved that classical extreme value laws (GEV distributions) can be used to describe the recurrence properties around points of the phase space of a dynamical system (Freitas et al 2010; Faranda et al 2011). The recurrences of the dynamics around a point of the phase space allow to define *rare events*, i.e. regions of the phase space that are seldom visited. Extreme values obtained as recurrences around a point of the phase space have as limiting distribution a GEV if the system fulfills specific chaotic constraints that replace the independence condition for random variables (Freitas et al 2010).

The goal of this paper is to present a metric to evaluate the rare temperature events of climate model simulations and compare them with reanalysis data. This metric is compared to analyses of extremes of temperature with the EVT. Comparing the metrics and the datasets helps to assess regional biases in climate model simulations.

We focus on recurrence properties of European temperature data for the 20 Century Reanalysis (hereinafter 20CR) and a model run (IPSL-CM5) forced with historical greenhouse gases concentrations. We introduce the classical extreme value theory (EVT) metrics and the one based on the recurrence analysis in Section 2. We illustrate the recurrence metric on several temperature time series taken from reanalyses

65 and model simulation in Section 3. In Section 4 we compare the EVT and recurrence metrics on the  
 66 temperature data of the IPSL model and the two reanalysis product. Finally we provide a justification  
 67 of the results and summarize the main findings.

## 68 2 Methods

69 In this section, after recalling the EVT approaches to the study of extreme events as large (or small)  
 70 ones, we present the metric based on the recurrences analysis and focus on rare events. We also explain  
 71 the algorithmic procedure we use to obtain our results.

### 72 2.1 Extreme Value Theory: large deviations

73 The first way to model the extremes of a continuous random variable  $X$  is to determine the probability  
 74 distribution of the maxima of  $X$  over blocks of  $m$  observations. Under hypotheses of independence  
 75 and identical distribution (IID), the probability distribution of maxima  $M_m = \max\{X_0, X_1, \dots, X_{m-1}\}$   
 76 converge to generalized extreme value (GEV) law (Leadbetter et al 1983):

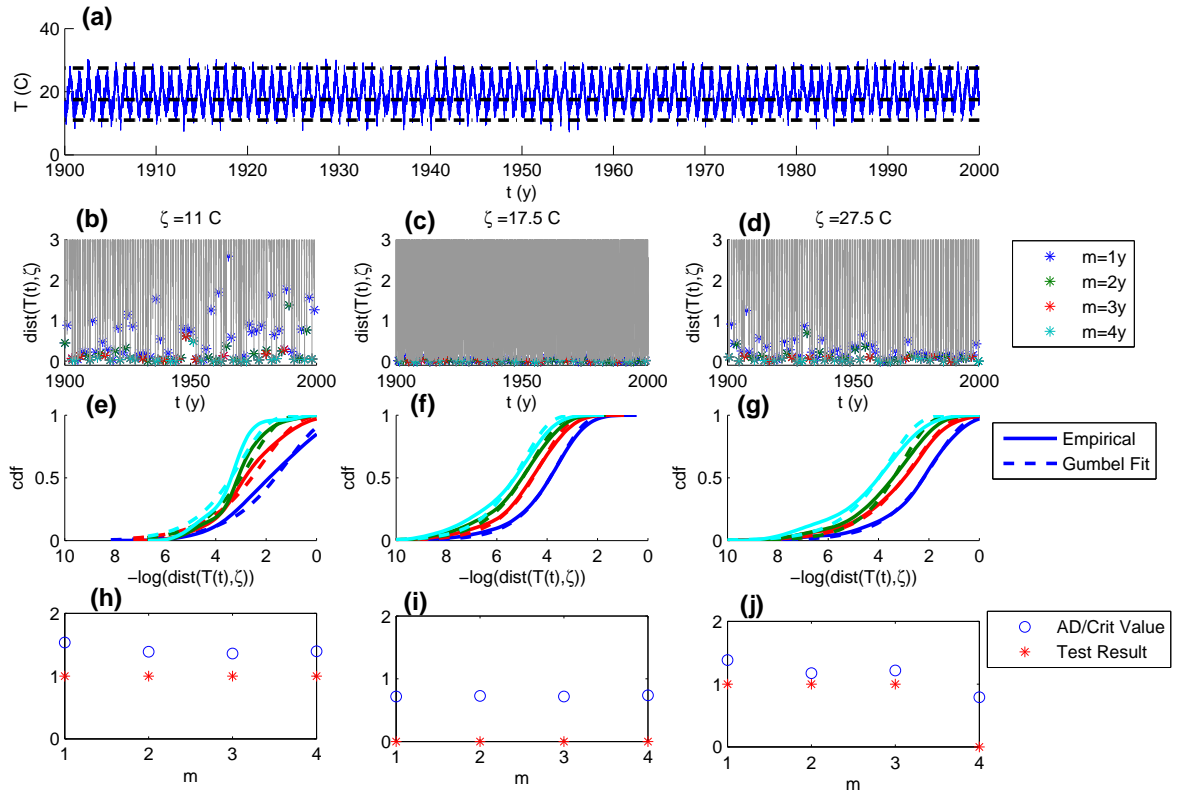
$$F_G(M_m < x) = \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu(m)}{\sigma(m)} \right) \right]^{-1/\xi} \right\} \quad (1)$$

77 with  $1 + \xi(x - \mu)/\sigma > 0$ . The location parameter  $\mu(m)$  and the scale parameter  $\sigma(m) > 0$  in Eq. 1  
 78 account for the normalization of the data. The sign of the shape parameter  $\xi$  discriminates the kind of  
 79 tail decay of the parent distribution. We call this approach Block Maxima (BM) and we determine the  
 80 GEV parameters from a maximum likelihood estimator.

81  
 82 The practical choices of optimal block lengths  $n$  stem from heuristic considerations on the length of  
 83 datasets (Coles 2001). The laws of extremes that are determined allow one to estimate return periods  
 84 associated to return levels of extreme events. By construction, extreme events are rare. Therefore their  
 85 return periods are large. The  $r$ -year return value is formally defined by setting Eq. 1 equal to  $1 - 1/r$ .  
 86 This computation is based on the assumption that, for large  $m$ , the shape parameter  $\xi$  of the GEV does  
 87 not depend anymore on  $m$  and therefore one can extrapolate over the length of the sample. Since we  
 88 work with limited data-sets,  $\xi$  is not independent on  $m$  and the previous formula give biased results.  
 89 In order to overcome this problem and compute reliable return levels for  $r$ -year return value, we have  
 90 to fit a different GEV for each bin length  $m$  and use as return level the estimate of  $\mu(m)$  which is, by  
 91 definition, the expected return level for return period  $r = m$  (Coles 2001).

### 92 2.2 Recurrence analysis: rare events

93 Here, we focus on rare events, i.e. which have a long return period. Such events are not necessarily  
 94 extreme in the sense outlined in the previous subsection. This complementary point of view is based on a  
 95 use of the dynamical information contained in the time series. In order to describe this idea, we consider  
 96 a discrete time dynamical system. This is a reasonable hypothesis for an atmospheric system (Lucarini  
 97 et al 2012; Faranda et al 2013). The dynamics is governed by a set of equations  $f$ , which controls the  
 98 variables of the systems  $x$  such that  $x_{t+1} = f(x_t)$ . We assume that, by starting from a random initial



**Fig. 1** Illustration of the recurrence technique. **a** Time series of temperature obtained from the 20CR analysis at Lon=30E, Lat=61N. **b-d** Series of the distances from the temperature of  $\zeta = 11$  °C, 17.5 °C, 27.5 °C respectively. Colored stars indicate the minima selected for different bin lengths, as in the legend. **e-f** Empirical cumulative distribution functions (cdf) for  $\zeta = 11$  °C, 17.5 °C, 27.5 °C respectively. Different colors correspond to different bin lengths. **h-j** Results of the Anderson Darling test in terms of the ratio between test parameters and critical values (circles) and test results (stars) for different bin lengths.

99 condition, the dynamics follows a chaotic trajectory. We fix an arbitrary point  $\zeta$  of the trajectory and  
 100 measure the time series of the distances between  $\zeta$  and the other observations in the trajectory:

$$g(t; \zeta) = -\log(d(x_t, \zeta)).$$

101  $d(x, y)$  is a distance function between two vectors: by definition it tends to 0 when the vectors  $x$  and  $y$  are  
 102 close to each other. We take the opposite of the logarithm ( $-\log$ ) in order to increase the discrimination  
 103 of small values of  $d(x, y)$  and treat high values of  $g$ . We are interested in the extremes of  $g(t; \zeta)$ , for all  $t$ .  
 104 Such extremes define the recurrences of the system for all  $\zeta$ . To select the extremes, we apply again the  
 105 BM approach this time by dividing the time series  $g(x_t)$ , in intervals of length  $m$ . Every  $m$  observations,  
 106 the closest recurrence to the point  $\zeta$  is taken. If  $n$  intervals of length  $m$  are available in the series, one  
 107 ends up with  $n$  closest recurrences.  
 108

109 If the system is chaotic, it has been shown that the distribution of the extremes of  $g(t; \zeta)$  converge to  
 110 a GEV of Gumbel type (Gumbel 1958). A detailed explanation for this can be found in Faranda and  
 111 Vaienti (2013). The reason is intuitive: the Gumbel law is in the form  $\exp(-\exp(-x))$ . Here one of the  
 112 exponential functions comes from the exponential recurrence statistics, the other by the inverse of the  
 113 logarithm function.

114  
 115 The distribution of extremes of  $g(t; \zeta)$  contains the dynamics of the systems in terms of rate of con-  
 116 vergence to the expected GEV: depending on the bin length  $m$ , one gets or not convergence to the  
 117 asymptotic extreme value laws. For this reason it is possible to define the shortest bin length when  
 118 appropriate statistical tests (such as the Lilliefors (1969) or the Anderson and Darling (1954) tests) fail  
 119 to reject the hypothesis that the extreme statistics of  $g(t; \zeta)$  is GEV distributed. This quantity gives  
 120 the shortest return time  $\tau$  for the point  $\zeta$ . Hence  $\zeta$  is normally recurrent for  $m > \tau$ , or rare for  $m < \tau$   
 121 (Faranda and Vaienti 2013).

122

### 123 2.3 The algorithm

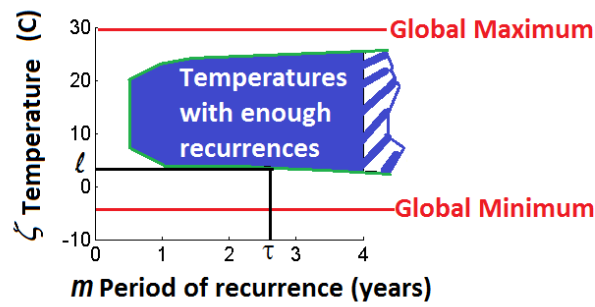
124 The recurrence approach can be adapted to the study of time series (rather than all the variables of a  
 125 dynamical systems). In the subsequent analysis  $\zeta$  is a reference temperature at a given location and  $x_t$   
 126 is a time series of daily mean temperatures.  $f$  represents the (unknown) atmospheric dynamics.

127

128 The analysis is illustrated step-by-step in Fig. 1. The daily temperature data at single grid points are  
 129 shown in Fig.1a. From the global minimum temperature, to the global maximum temperature, all the  
 130 values recorded in the series are checked alternatively as  $\zeta$ . The dotted lines correspond to three different  
 131 temperatures chosen to illustrate the technique.

132

133 Fig. 1b, c, d show the distances (grey curves) from the references temperatures  $T = 11$  °C, 17.5 °C, 27.5  
 134 °C respectively for all the 100 years analyzed. The minimum distances corresponding to a bin length  
 135 of  $m = 1, 2, 3, 4$  years are indicated in color. Since the scale is fixed to be the same for all the cases,  
 136 one remarks that for shortest bin length, when  $\zeta = 11$  °C or  $\zeta = 27.5$  °C, the minimum distances can  
 137 be relatively large. This does not happen for the temperature of 17.5 °C. In order to assess whether  
 138 the distribution of minimum distances are compatible with a Gumbel law, one needs first to take -log  
 139 and then compare the empirical cumulative density function (cdf) with the best fit to the Gumbel law.  
 140 Such comparison is shown in Fig.1e, f, g. As before, different colors correspond to different bin lengths.  
 141 By eye, it is difficult to assess whether a fit is good or not although qualitatively the fits for the cases  
 142  $\zeta = 17.5$  °C look more reliable. The quantitative answer to this question is given via the Anderson and  
 143 Darling (1954) test, whose results are displayed in Fig.1h, i, j. It measures the integral between the two  
 144 cdfs (empirical and fitted). If the value of this integral is small with respect to a tabulated critical value,  
 145 then the test does not reject the hypothesis that the data are Gumbel distributed. Fig.1h, i, j show the  
 146 results of the different Anderson-Darling tests in terms of the ratio between the test statistics and the  
 147 critical value (blue circles), and of the outcome of the test (red stars). When the ratio is smaller the one,  
 148 the Anderson Darling test succeed ( $h = 0$ ). This happens for all the bin length at  $\zeta = 17.5$  °C, and only  
 149 for  $m = 4$  years for  $\zeta = 27.5$  °C. The occurrence of this temperature can therefore be considered rare for  
 150 time scales smaller than 3 years and normal for time scales larger than 4.



**Fig. 2** Region of temperatures for which the Anderson-Darling test does not refuse the hypothesis that data are Gumbel distributed (blue area) for different period of recurrences  $m$  in years. Minimum return times  $\tau$  and corresponding return levels  $\ell$  are represented by the projections of the green line on the  $x$  and  $y$  axis respectively.

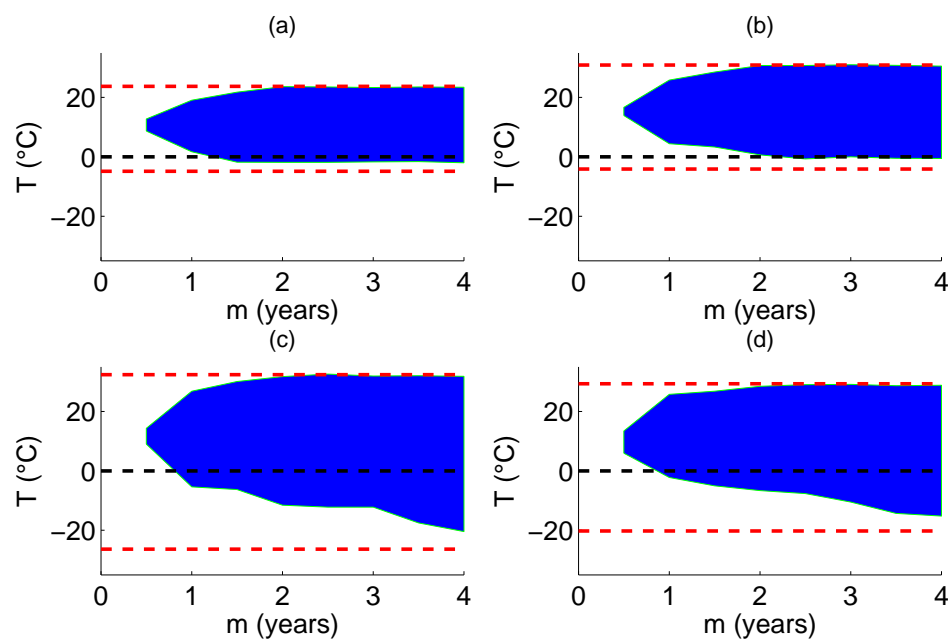
151 In Fig.2 we visualize the results of the Anderson-Darling test in the  $(m, \zeta)$  plane. The blue area represents  
 152 the region where the Anderson-Darling test does not reject the hypothesis that the data are Gumbel  
 153 distributed. Rare temperatures are located in the white area between the red lines (the absolute extreme  
 154 of the series, indicated as global maximum and global minimum) and the blue area. The minimum return  
 155 times  $\tau$  are the projection on the  $m$  axis of the green curve and the correspondent return levels  $\ell$  are the  
 156 projections on the temperatures axis.

### 157 3 Temperature recurrences of four European locations

158 The analysis is focused on the European region, between  $35^\circ N - 62^\circ N$  and  $12^\circ W - 32^\circ E$ . We com-  
 159 pare data from the 20 Century Reanalysis (20CR) (Compo et al 2011), the ERA 20C (Poli et al 2013)  
 160 and an historical run of the Institut Pierre-Simon Laplace (IPSL) model (Dufresne et al 2013). Here  
 161 we focus on surface temperature variations and relate them to the sea level pressure (SLP). The 20CR  
 162 data yields a horizontal resolution of  $1^\circ$ . They cover the period between 1871 and 2010. Our analysis is  
 163 based on the ensemble mean of 20CR. The ERA 20C assimilates observations of surface pressure and  
 164 surface marine winds only. The horizontal resolution is approximately 125 km (spectral truncation T159).  
 165

166 The IPSL-CM5, contributed to the 5<sup>th</sup> phase of the Coupled Model Intercomparison Project (CMIP5).  
 167 In addition to the physical atmosphere-land-ocean-sea ice model, it also includes a representation of  
 168 the carbon cycle, the stratospheric chemistry and the troposphere chemistry with aerosols. The version  
 169 used here is IPSL-CM5A medium resolution,  $1.25^\circ \times 2.5^\circ$  ( $143 \times 144L39$ ). The model is thoroughly  
 170 documented by Hourdin et al (2013a). The extremes were first investigated by Cattiaux et al (2013). In  
 171 order to compare the results, the reanalysis have been bilinearly re-interpolated on the same grid as the  
 172 model.

173 For all the datasets, we consider 100 years (1900-1999) of daily temperature data. This allows for the  
 174 computation of return times between 6 months up to 4 years. This limit is imposed by the length of  
 175 the time series: the block-minima approach requires to keep a sufficient number of extremes as well as a  
 176 large set of observations in each bin for the selection of extremes. For recurrence windows longer than 4  
 177 years, we get less than 25 maxima and unreliable estimates of the GEV distributions. Conversely, events  
 178 chosen for bin length shorter than 6 month are not rare enough to be considered in the analysis.



**Fig. 3** Recurrence analysis at four specific grid points of the 20CR. Temperature for which the Anderson-Darling test does not refuse the hypothesis that data are Gumbel distributed, are located in the blue area for different period of recurrences  $m$  in years. Red dotted line represents global maximum and minimum of the time series, gray dotted line represents  $0\text{ }^{\circ}\text{C}$ . **a** Ireland, **b** Po Valley, **c** Vienna Region, **d** Paris Region.

179 In order to illustrate the recurrence technique, we compute the 20CR temperature recurrences at 4 grid  
 180 points with different climate characteristics. Results are displayed in Fig.3a for Ireland, Fig. 3b for the  
 181 Po Valley , Fig. 3c for Vienna Region and Fig. 3d for Paris Region.

182

183 We first describe the different shapes of the blue area of normal recurrences explainable in terms of the  
 184 climate characteristics of the grid points considered. The climate of Ireland is mild: extreme cooling or  
 185 warming is prevented by the presence of the Atlantic Ocean. The green curve is therefore very symmetric  
 186 and saturates at the values of  $0\text{ }^{\circ}\text{C}$  (for the minima) and  $18\text{ }^{\circ}\text{C}$  for the maxima.

187

188 The Po Valley climate is influenced by a Mediterranean component as well as by the local geographical  
 189 features. The presence of both the Alps and the Apennines creates, in combination with blocking events,  
 190 stable condition and large daily temperature excursions (Beniston 2006). Negative temperatures are gen-  
 191 erally associated to snowfalls which occur even at the sea level almost every year (Terzago et al 2010).  
 192 This feature is captured by the shape of the green curve for the minima: for a return period of 1 year,  
 193 recurrences of temperatures lower than 0 are rare, but for higher return times negative temperatures  
 194 recurrences become typical.

195

196 The region of Vienna features a continental climate and very cold extremes of temperature are produced  
 197 for 3 or 4 years return time. The Paris region is greatly influenced by the Atlantic Ocean, but blocking  
 198 conditions may trigger large temperature excursions associated to both cold spells and heat waves as



shown by the low values reached by the green curve for 3-4 years return period.

Over geographically uniform regions the variation between the recurrence plots shown in this example is uniform. However, at mountain ranges, or for grid points with land-sea sharp variations, great differences can appear even for contiguous grid points.

The asymmetry between the behavior of maxima and minima is coherent with the extreme value statistics computed with the GEV approach or the analysis of conditional variance (Yiou et al 2009). As described by Brown et al (2008), both the maxima and minima temperature extremes obey a Weibull law corresponding to a negative shape parameter of the GEV distribution, as generally expected for physical observables (Faranda et al 2014). Brown et al (2008) show that maxima are more bounded than minima i.e. the shape parameter is more negative for the maxima. In our analysis, this is reflected by the slower saturation of the  $\tau$  curve for the minima of temperature recurrences.

#### 4 Comparison between EVT and recurrence metrics

We compare the return levels obtained with the recurrence technique  $\ell$  to the ones computed by using the BM approach  $\mu$  by taking maxima and minima of temperatures over block lengths of  $m = 1, 2, 3, 4$  years. The biases are denoted as  $\Delta T_{max}$  and  $\Delta T_{min}$  and computed as  $\Delta T_{max}(m) = \ell_{20CR} - \ell_{ERA20C}$  for the recurrences, and  $\Delta T_{max}(m) = \mu_{20CR} - \mu_{ERA20C}$  for the GEV return levels. Analogous quantities are defined for the minima and other combinations of model/reanalysis.

The results are shown in Fig. 4a-f for the maxima and Fig. 4g-l for the minima. The box-plots provide information about the distribution of the biases: on each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. The  $x$ -axis show different bin lengths  $m$ . Fig. 4a-c and Fig. 4g-i show the bias obtained via the recurrence technique and Fig. 4d-f and Fig. 4j-l for the BM approach. The left hand-side column in Fig. 4 report the biases between 20CR and ERA 20C, the central report the box-plots the biases between the IPSL model and ERA 20C and the right hand-side the biases for the IPSL model and ERA 20C.

The two methodologies provide different results: the sign of the median can vary between recurrences and BM as well as the distribution of outliers. Moreover, the BM is substantially insensitive to the change of bin length  $m$ . For the BM approach, one can hardly spoil differences between the distributions obtained for  $m = 1$  or  $m = 4$ . In this case, the GEV fitting with short (e.g. 1 years) bin lengths is biased as the asymptotic regime, i.e. the limit  $m \gg 1$  is not reached. The procedure introduced to determine  $\ell$  prevents from short size effects because, when  $m$  is too short to contain good recurrences of  $\zeta$ , the Anderson Darling test simply refuses the hypothesis that the data are Gumbel distributed. This explains why the plots analysed in Fig. 3 have a conic shape (they would rather be rectangular for the BM approach) and show a sensible variation between 1 and 4 years estimations.

Before assessing the way the IPSL model represent temperature extremes, we compare the two reanalysis products to check the consistency between the two datasets (Fig. 4 a, d). For the maxima, the BM approach suggests that 20CR are warmer of about 2 °C than the ERA 20C. The median of the distribution

for the biases computed with the recurrence techniques is positive. Most of outliers are positive at  $m = 1$  year bin length and negative for  $m = 4$  years. For the minima the results are shown in Fig.4g, j. In this case the median biases are negative for both the methods. We remark that for  $m = 4$  years the box plots obtained with the recurrence technique and the BM approach are similar.

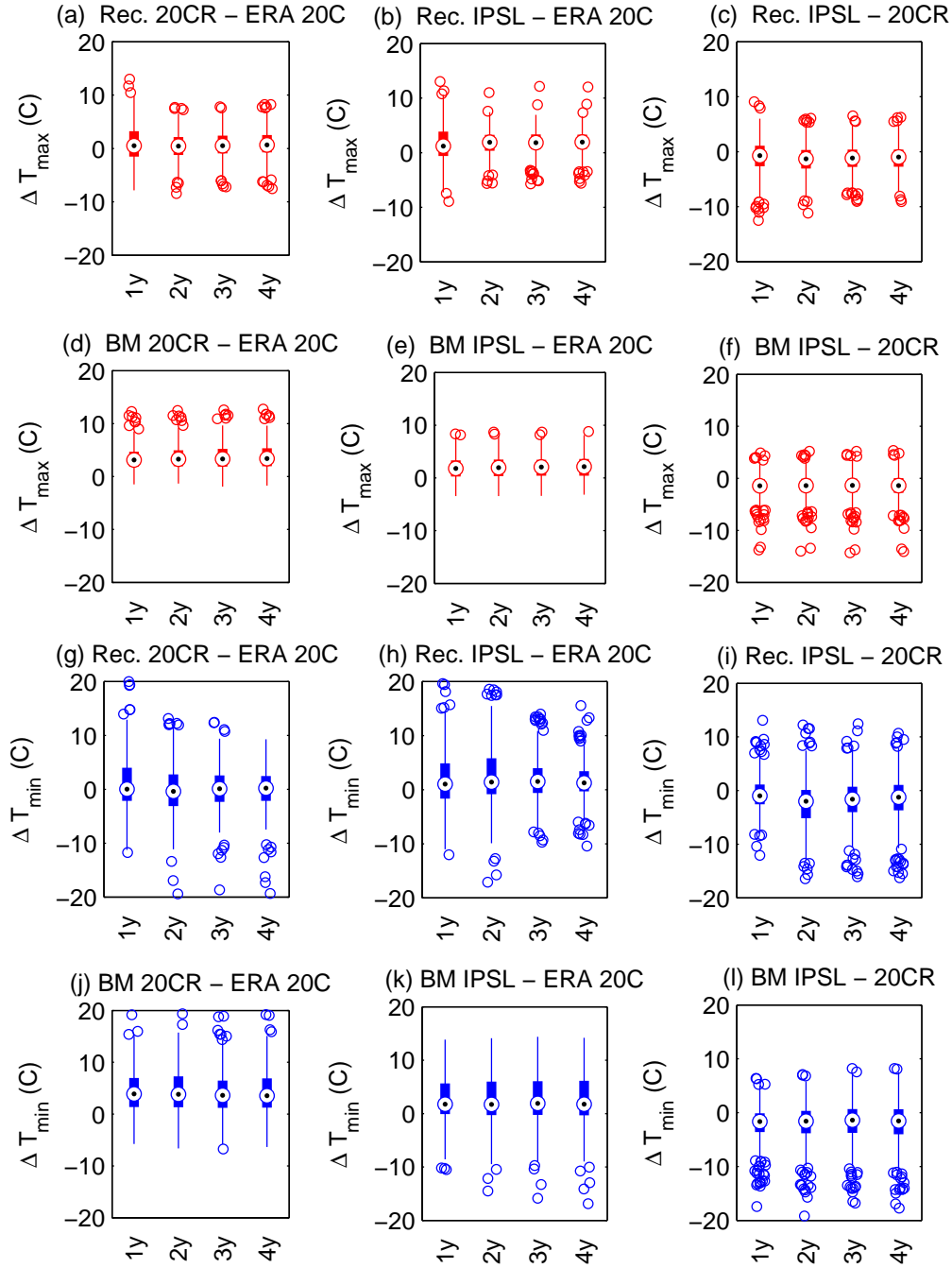
We have seen that for the reanalysis products analyzed the biases in the representation of both large (via the BM approach) and rare (via recurrences) extremes can reach up to 15 °C. It is not surprising that the comparison of the output of the IPSL historical simulation with the reanalysis suffer from the same problems. In addition to the box-plots, we represent the the spatial distributions of the biases between IPSL model and 20CR in Fig. 5a, b (maxima) and Fig. 6a, b (minima) and in Fig. 5c, d (maxima) and Fig. 6c, d (minima) for the ERA 20C. Maxima of the IPSL model seems to reproduce well those of the reanalysis over the Mediterranean basin. A positive bias over Great Britain and a negative bias for the Iberian peninsula appear as robust features with respect to the change of reanalysis data-set and the change of bin length. The results for the minima show less coherence, although one can identify a negative bias over the Scandinavian region and over England and a positive bias over continental Europe. Overall, biases for the minima are larger than for the maxima

## 5 Understanding the biases in extremes

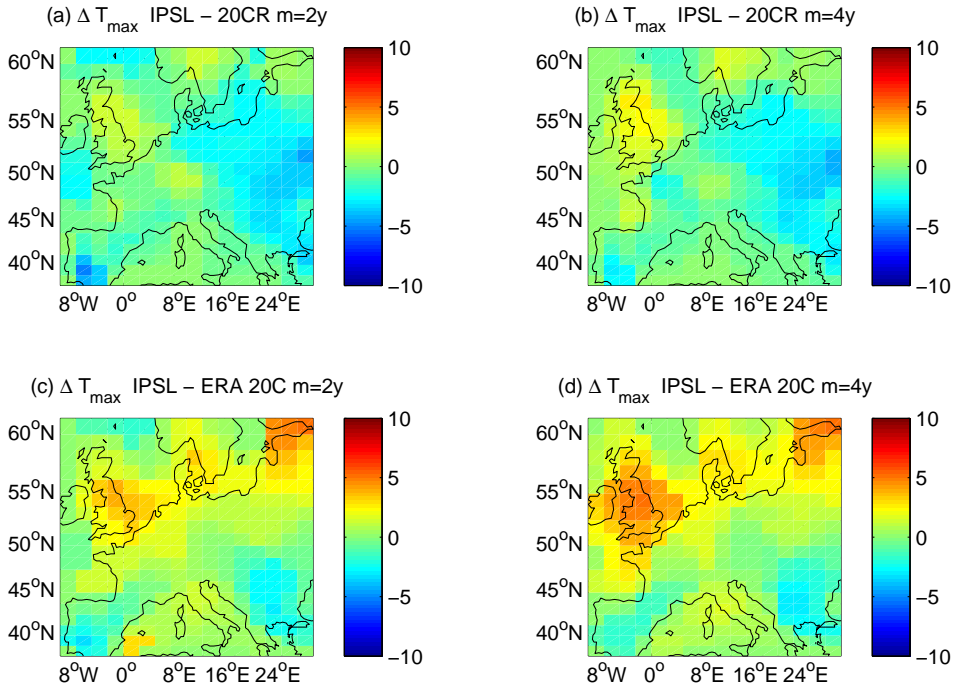
In order to explain the differences on both large and rare extremes, we first assess whether they originate by a poor representation of the mean and of the variability of the temperature distributions or whether they are related to finer statistical properties (higher moments of temperature distributions). This analysis tells whether the extreme value analysis add a real information with respect to the classical statistical characterization. We then focus on the Scandinavian region and the Iberian peninsula and try to understand the possible origin of some biases found with the previous analysis which seem robust with respect to the change in the reanalysis data-set and in the bin length  $m$ .

### 5.1 Analysis of Moments

We report in Fig. 7 the differences in the  $i$ th moment  $m_i$  between the two reanalysis products Fig.7a–c, the IPSL and the 20CR Fig.7d–f, the IPSL and the ERA 20C in Fig.7g–i. Differences in the mean  $m_1$  are plotted in Fig.7a, d, g, in the standard deviation  $m_2$  in Fig.7b, e, h and in the skewness  $m_3$  in Fig.7c, f, i. Differences in the mean affect the location parameter  $\mu$  of the GEV distribution. The effects of biases in the mean for the recurrence plots shown in Fig.3 mostly consist in shifting the diagrams up or down. From our analysis is evident that the biases found for the means are small and do not have a coherent spatial structure. Moreover, the difference in means are clearly not related to the difference in variance, as it should be in case of systematic bias according to the method proposed by Yiou et al (2009) for estimation of conditional variance. In addition, the deviations found for the standard deviation and skewness are also relatively small and locally compensate each other effects, e.g. over the Scandinavian region. They affect the recurrence diagrams by a change of the saturation values of  $\ell$ . It is evident that such statistical quantities (even combined) can not fully explain the biases found for rare recurrences. This means that the extremal behavior depends also on high order moments which are tricky to measure for short time series.



**Fig. 4** Box plots of the biases  $\Delta T_{max}$  (red) and  $\Delta T_{min}$  (blue) between 20CR and ERA 20C **a,d,g,j**, IPSL and ERA 20C **b,e,h,k**, IPSL and 20CR **c,f,i,l**. Central marks are the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually.  $x$  axis shows different bin lengths  $m$  from 1 to 4 years.



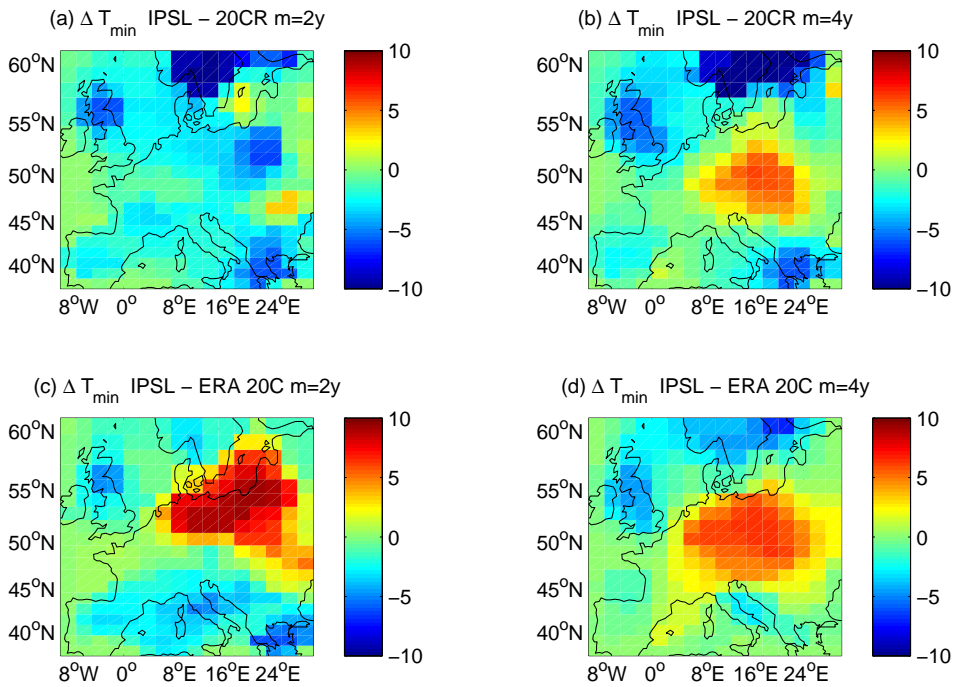
**Fig. 5** Analysis of the biases  $\Delta T_{max}$  obtained via the method of recurrences for the IPSL model and 20CR (a,b) and for the IPSL model and ERA 20C (c,d). From left to right the bin length  $m$  is increasing from 2 to 4 years. A spatial median filter on a window  $3 \times 3$  has been applied to the data.

## 282 5.2 Atmospheric circulation and biases

283 We now focus on the two regions (Scandinavia and south-west Iberian Peninsula) where the represen-  
 284 tation of rare recurrences highlights consistent differences between reanalyses and model historical run,  
 285 and we investigate the origins of such biases. The first hypothesis is that the biases are found for a  
 286 specific weather circulation pattern. In order to verify this claim, we construct an average pressure field  
 287 representing all the situation which are related to rare temperature in the models but not in the 20CR,  
 288 and vice-versa.

289  
 290 To obtain such pressure fields we follow this procedure: i) we select a period of recurrence  $m$  (e.g. two  
 291 years) ii) we compute the values of  $\ell$  located on the  $\tau$  line and therefore representing the boundaries  
 292 between normal and rare temperature recurrences (e.g.  $\ell = -35^\circ\text{C}$  for the minima of the IPSL-CM5  
 293 model and  $\ell = -21^\circ\text{C}$  over Scandinavia for  $m = 2$  years), iii) we determine the composite SLP field  
 294 corresponding to the temperature return levels  $\ell$  for the model and the correspondent values for the  
 295 reanalysis and average over them, iv) we repeat the procedure by changing the period of recurrence  $m$   
 296 and check the robustness of the results with respect to changes in  $m$ .

297



**Fig. 6** Analysis of the biases  $\Delta T_{min}$  obtained via the method of recurrences for the IPSL model and 20CR (a,b) and for the IPSL model and ERA 20C (c,d). From left to right the bin length  $m$  is increasing from 2 to 4 years. A spatial median filter on a window  $3 \times 3$  has been applied to the data.

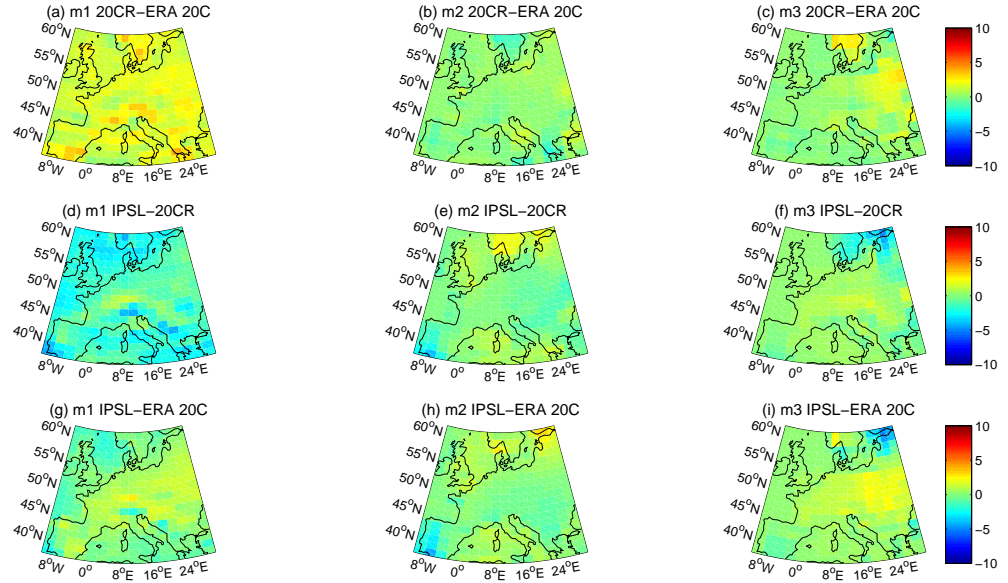
298 If a predominant pressure field appears from this analysis, then we can use this information to understand  
 299 the nature of the bias. In Fig. 8 we represent the average pressure field for the situation associated to  
 300 the bias in the Scandinavian region (left panel) and in the south-west Iberian Peninsula (right panel).  
 301 Some of the daily pressure fields which have been averaged to produce Fig. 8 is reported in Fig. 9 for  
 302 Scandinavia and in Fig. 10 for south-west Iberian Peninsula.

303

304 Fig. 8a illustrates the SLP pattern linked to the Scandinavian bias. The average pressure field shows a  
 305 maximum over Scandinavia and the daily pressure fields reported in Fig. 9 show comparable weather  
 306 patterns. In cold nights with snow or ice cover, the presence of stable condition (high pressure) with clear  
 307 skies can produce a significant cooling. The over-cooling observed in the model can be therefore due to  
 308 an insufficient representation of the compensation mechanisms, like the parametrization of the low level  
 309 clouds and of the snow/ice graying that affect the surface albedo (Hourdin et al 2013a,b).

310

311 On the other hand, by looking at Fig. 8b and the daily pressure fields of Fig. 10, we cannot conclude  
 312 that there is a predominant weather pattern associated with the temperature bias in the south-west  
 313 Iberian Peninsula. Moreover, the analysis of daily pressure fields reported in Fig. 10 show a large variety  
 314 of patterns which all have different signatures on surface temperature (Yiou and Nogaj 2004). Contrary



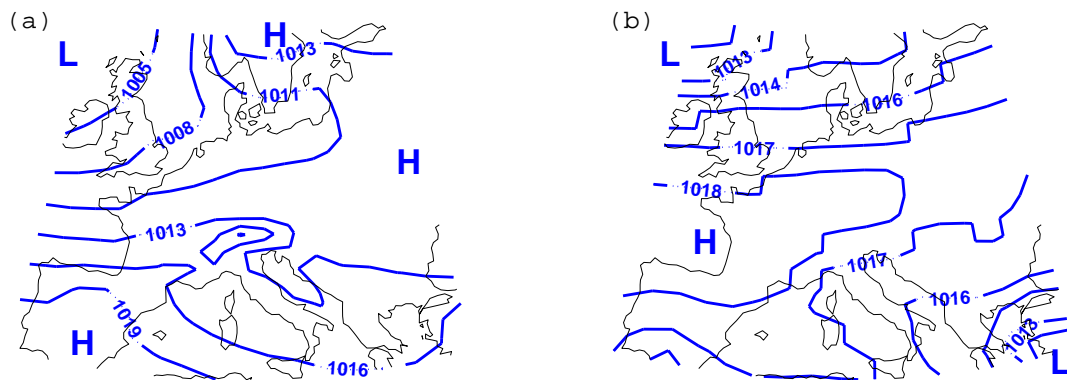
**Fig. 7** Differences in the statistics of the first three central moments of the temperature distribution  $T$ . **a-c** show the biases between 20CR and ERA 20C, **d-f** the biases between IPSL model and 20CR, **g-i** the ones between IPSL model and ERA 20C. From left to right first (mean m1), second (standard deviation m2) and third order (m3) moments respectively. A spatial median filter on a window  $2 \times 2$  has been applied to the data.

315 to the results found for Scandinavia, we conclude that the bias over the Iberian peninsula is independent  
 316 on the weather pattern considered. The origin of the bias are instead linked to the geography: the model  
 317 represents south-west Iberian Peninsula not as a continental grid point, but as a mixture of land and  
 318 sea. This area is characterized by the presence of the Guadalquivir valley, which is located between the  
 319 mountain ranges Sierra Morena and Betic System, creating a local continental climate over the region  
 320 where high temperatures can be easily attained in summer (Rivas-Martínez et al 1997; Hernández-  
 321 Ceballos et al 2013).

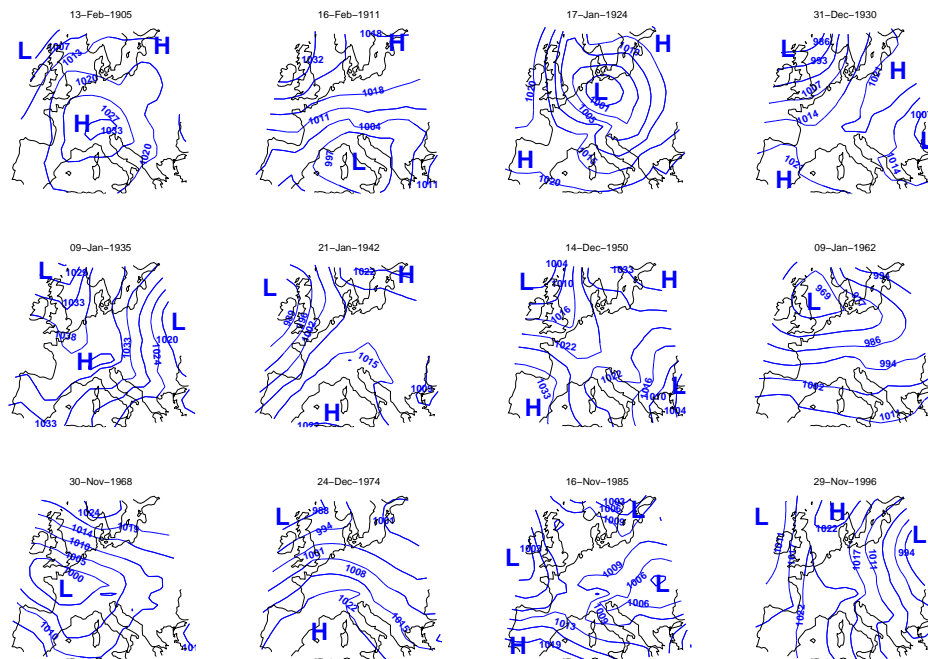
## 322 6 Discussion

323 In this paper we have used the recurrence analysis proposed in Faranda and Vaienti (2013) to study  
 324 the dynamics of temperature extremes over the European region. The GEV fitting, based on the block-  
 325 maxima approach, considers extreme events as large (or small) observations. The recurrence technique  
 326 is focused on the analysis of events that are rare independently on the magnitude of the observables  
 327 considered. We have shown that the metrics used to evaluate extremes over the European region for  
 328 two reanalysis products and an historical simulation provide different results and that the recurrence  
 329 technique is more sensitive to the change in bin-length with respect to the BM approach.

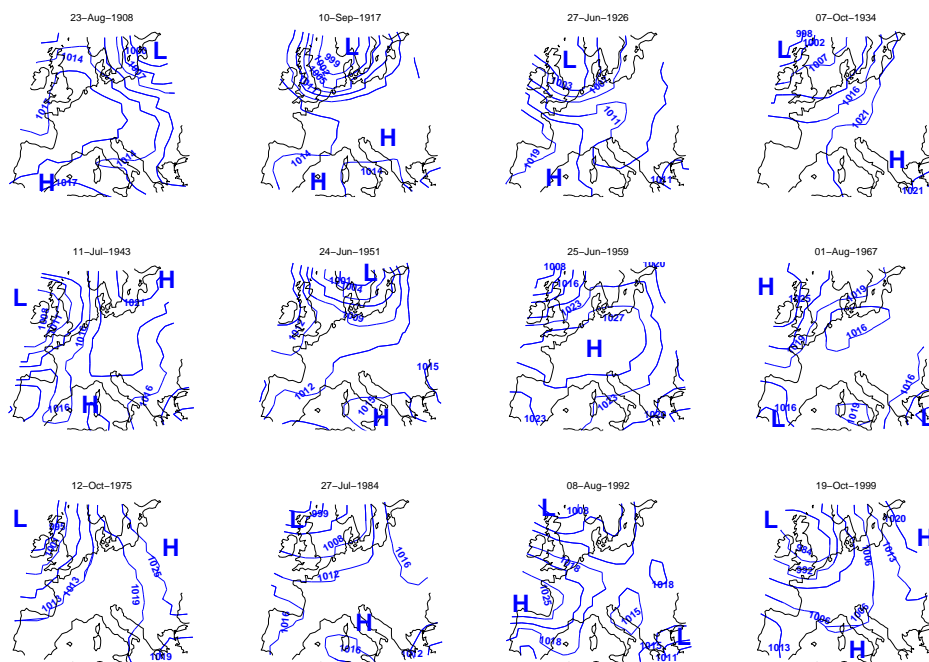
330



**Fig. 8** a: average pressure field (in hPa) obtained by averaging all the pressure fields of the days where the IPSL-CM5 model simulated temperatures between  $-20\text{ }^{\circ}\text{C}$  and  $-35\text{ }^{\circ}\text{C}$  in the Scandinavian region. b: average sea level pressure field (in hPa) obtained by averaging all the pressure fields of the days where the 20CR registered temperatures between  $25$  and  $35\text{ }^{\circ}\text{C}$  in the south-west Iberian Peninsula .



**Fig. 9** Daily sea level pressure fields (in hPa) for some of the situations where the IPSL-CM5 model simulated temperatures between  $-20\text{ }^{\circ}\text{C}$  and  $-35\text{ }^{\circ}\text{C}$  in the Scandinavian region.



**Fig. 10** Daily sea level pressure fields (in hPa) for some of the situations where the 20CR registered temperatures between  $25^{\circ}\text{C}$  and  $35^{\circ}\text{C}$  in the south-west Iberian Peninsula.

331 For the maxima we found that the representation of warm days given by the IPSL model is coherent  
 332 with the reanalysis data-sets over most of the European region. Differences are localized over specific  
 333 areas, as the south-west Iberian peninsula and Great Britain. For the minima biases are larger and there  
 334 is some inconsistencies with respect to the change of bin length over some regions. Negative biases over  
 335 England and Scandinavia appear as robust features with respect to the change in reanalysis data-set and  
 336 bin length.

337

338 We then investigated the biases obtained with the recurrence analysis against the first order moments  
 339 of the temperature biases distribution and the atmospheric circulation. By analysing the pressure fields  
 340 associated to cold and warm days misrepresented in the IPSL model, we have associated the bias for  
 341 south-west Iberian Peninsula to the sea/land mask of the model which represents these grid points with  
 342 a fraction of sea, whereas this region is affected by a marked local continental climate. The over-cooling  
 343 observed in Scandinavia have been associated to the representation of atmospheric circulation in the  
 344 model because a precise weather pattern consisting of an high pressure over this region appears for most  
 345 of the events considered for the conditional average. The misrepresentation is then attributed to the  
 346 radiative effects which are possibly not compensated by model parameterizations. These and the other  
 347 biases can hardly be explained by just looking at the first order moments of the temperature distribution.

348



Besides the results obtained for the IPSL model examined in this paper, the technique can be easily adapted for the analysis of different models and variables and to combine the extremes of multiple variables simultaneously. The computation of robust minimum return times for extremes of temperature data avoids the use of asymptotic statistical theories which can provide return times on longer - even infinite - periods without considering the underlying dynamics.

Future applications of these methods include the study of climate change for temperature recurrences of models and reanalysis data. The signature of climate change will correspond to modifications of the blue area and the green curve of the recurrence diagrams when two period with different greenhouse forcing are considered.

## Acknowledgments

D. Faranda and P. Yiou were supported by ERC grant No. 338965-A2C2 and M. Carmen Alvarez-Castro was supported by the Swedish Research Council grant: "Euro-Atlantic climate variability during the last millennium: atmospheric circulation and extreme events (MILEX)".

## References

- Anderson TW, Darling DA (1954) A test of goodness of fit. *Journal of the American Statistical Association* 49(268):765–769
- Beniston M (2006) Mountain weather and climate: A general overview and a focus on climatic change in the alps. *Hydrobiologia* 1(562):3–16
- Brown S, Caesar J, Ferro CA (2008) Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research: Atmospheres* (1984–2012) 113(D5)
- Cattiaux J, Quesada B, Araklian A, Codron F, Vautard R, Yiou P (2013) North-atlantic dynamics and european temperature extremes in the ipsl model: sensitivity to atmospheric resolution. *Climate Dynamics* 40(9-10):2293–2310
- Coles S (2001) *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, New York.
- Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose R, Rutledge G, Bessemoulin P, et al (2011) The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society* 137(654):1–28
- Dufresne JL, Foujols MA, Denvil S, Caubel A, Marti O, Aumont O, Balkanski Y, Bekki S, Bellenger H, Benschila R, et al (2013) Climate change projections using the ipsl-cm5 earth system model: from cmip3 to cmip5. *Climate Dynamics* 40(9-10):2123–2165
- Faranda D, Vaienti S (2013) A recurrence-based technique for detecting genuine extremes in instrumental temperature records. *Geophysical Research Letters* 40(21):5782–5786
- Faranda D, Lucarini V, Turchetti G, Vaienti S (2011) Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *Journal of Statistical Physics* 145(5):1156–1180
- Faranda D, Freitas JM, Lucarini V, Turchetti G, Vaienti S (2013) Extreme value statistics for dynamical systems with noise. *Nonlinearity* 26(9):2597
- Faranda D, Bourgoïn M, Miralles S, Odier P, Pinton JF, Plihon N, Daviaud F, Dubrulle B (2014) Robust estimate of dynamo thresholds in the von kármán sodium experiment using the extreme value theory. *New Journal of Physics* 16(8):083,001
- Freitas ACM, Freitas JM, Todd M (2010) Hitting time statistics and extreme value theory. *Probability Theory and Related Fields* 147(3-4):675–710
- Frisch HL (1956) Poincaré recurrences. *Physical Review* 104(1):1
- Galambos J (1980) *The asymptotic theory of extreme order statistics*
- Ghil M, Yiou P, Hallegatte S, Malamud B, Naveau P, Soloviev A, Friederichs P, Keilis-Borok V, Kondrashov D, Kossobokov V, et al (2011) Extreme events: dynamics, statistics and prediction. *Nonlinear Processes in Geophysics* 18(3):295–350
- Gnedenko B (1943) Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics* pp 423–453
- Gumbel EJ (1958) *Statistics of Extremes*. Columbia University Press, New York

- 395 Hernández-Ceballos J MA and Adame, Bolívar J, De la Morena B (2013) A mesoscale simulation of coastal circulation in  
396 the Guadalquivir valley (southwestern Iberian peninsula) using the WRF-ARW model. *Atmospheric Research* 124:1–20
- 397 Hourdin F, Foujols MA, Codron F, Guemas V, Dufresne JL, Bony S, Denvil S, Guez L, Lott F, Ghattas J, Braconnot  
398 P, Marti O, Meurdesoif Y, Bopp L (2013a) Impact of the LMDZ atmospheric grid configuration on the climate and  
399 sensitivity of the IPSL-CM5a coupled model. *Climate Dynamics* 40(9-10):2167–2192
- 400 Hourdin F, Grandpeix JY, Rio C, Bony S, Jam A, Cheruy F, Rochetin N, Fairhead L, Idelkadi A, Musat I, Dufresne  
401 JL, Lahellec A, Lefebvre MP, Roehrig R (2013b) LMDZ5b: the atmospheric component of the IPSL climate model with  
402 revisited parameterizations for clouds and convection. *Climate Dynamics* 40(9-10):2193–2222
- 403 Katz R (2010) Statistics of extremes in climate change. *Climatic Change* 100(1):71–76
- 404 Katz R, Brown B (1992) Extreme events in a changing climate: Variability is more important than averages. *Climatic  
405 Change* 21(3):289–302
- 406 Leadbetter MR, Lindgren G, Rootzén H (1983) *Extremes and related properties of random sequences and processes*, vol 21.  
407 Springer-Verlag New York
- 408 Lilliefors HW (1969) On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the  
409 American Statistical Association* 64(325):387–389
- 410 Lorenz EN (1963) Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences* 20(2):130–141
- 411 Lucarini V, Faranda D, Turchetti G, Vaienti S (2012) Extreme value theory for singular measures. *Chaos: An Interdisci-  
412 plinary Journal of Nonlinear Science* 22(2):023,135
- 413 Lucarini V, Faranda D, Wouters J, Kuna T (2014) Towards a general theory of extremes for observables of chaotic dynamical  
414 systems. *Journal of Statistical Physics* 154(3):723–750
- 415 Nikulin G, Kjellström E, Hansson U, Strandberg G, Ullerstig A (2011) Evaluation and future projections of temperature,  
416 precipitation and wind extremes over Europe in an ensemble of regional climate simulations. *Tellus A* 63(1):41–55
- 417 Pickands III J (1975) Statistical inference using extreme order statistics. *The Annals of Statistics* pp 119–131
- 418 Poli P, Hersbach H, Tan D, Dee D, Thepaut JN, Simmons A, Peubey C, Laloyaux P, Komori T, Berrisford P, et al (2013)  
419 The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century  
420 assimilating surface observations only (ERA-20C)
- 421 Rivas-Martínez S, Asensi A, Díez-Garretas B, Molero J, Valle F (1997) Biogeographical synthesis of Andalusia (southern  
422 Spain). *Journal of Biogeography* 24(6):915–928
- 423 Terzago S, Cassardo C, Cremonini R, Fratianni S (2010) Snow precipitation and snow cover climatic variability for the  
424 period 1971–2009 in the southwestern Italian Alps: The 2008–2009 snow season case study. *Water* 2(4):773–787
- 425 Yiou P, Nogaj M (2004) Extreme climatic events and weather regimes over the North Atlantic: When and where? *Geophysical  
426 Research Letters* 31(7)
- 427 Yiou P, Dacunha-Castelle D, Parey S, Huong Hoang T (2009) Statistical representation of temperature mean and variability  
428 in Europe. *Geophysical Research Letters* 36(4)