Atmospheric Research 293 (2023) 106934

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Statistical analysis of extreme and record-breaking daily maximum temperatures in peninsular Spain during 1960–2021

Jorge Castillo-Mateo *, Ana C. Cebrián, Jesús Asín

Department of Statistical Methods, University of Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain

ARTICLE INFO

Keywords:
Climate change
Extreme temperature
Heatwave
Record test

ABSTRACT

This work analyses the effects of global warming in the upper extremes of daily temperature series over Spain. This objective implies specific analysis, since time evolution of mean temperature is not always parallel to evolution of the extremes. We propose the use of several record tests to study the behavior of the extreme and record-breaking events in different temperature signals, at different time and spatial scales. The underlying idea of the tests is to compare the occurrence of the extreme events in the observed series and the occurrence in a stationary climate. Given that under global warming, an increasing trend, or an increasing variability, can be expected, the alternative is that the probability of the extremes is higher than in a stationary climate. Some of the tests, based on a permutation approach, can be applied to sets of correlated series and this allows the analysis of short periods of time and regional analysis, where series are measured in close days and/or locations. Using these tests, we evaluate and compare the effects of climate change in temperature extreme and record-breaking events using 36 series of daily maximum temperature from 1960 to 2021, all over peninsular Spain. We also compare the behavior in different Spanish regions, in different periods of the year, and in different signals such as the annual maximum temperature. Significant evidences of the effect of an increasing trend in the occurrence of upper extremes are found in most of Spain. The effects are heterogeneous within the year, being autumn the season where the effects are weaker and summer where they are stronger. Concerning the spatial variability, the Mediterranean and the North Atlantic region are the areas where the effects are more and less clear, respectively.

1. Introduction

In the framework of climate change, there are many works that analyze the evolution of mean temperature over time, and the existence of an increasing trend is generally accepted; see Sánchez-Lugo et al. (2019) for a review and Peña-Angulo et al. (2021) for a study on the Spanish mainland. However, changes in variability and extremes of temperature are also relevant (Schär et al., 2004). The interest of analyzing whether the occurrence of extreme and record-breaking temperatures is affected by climate change is clear (Kysely, 2010; Coumou et al., 2013; Saddique et al., 2020; Om et al., 2022). The reason is that some of the most serious consequences of global warming on human health and other fields, such as agriculture or energy consumption, are often related to the occurrence of increasingly intense extremes (Tan et al., 2007; Coumou and Rahmstorf, 2012). The Mediterranean region has been referenced as a hot spot of climate change, this is, a region whose climate is especially responsive to global warming (Diffenbaugh and Giorgi, 2012; Lionello and Scarascia, 2018; Tuel and Eltahir, 2020). In fact, this region has suffered warming 20% faster than the rest of the globe (MedECC, 2020). This makes the study of temperature extremes in the Spanish mainland of special interest (Cos et al., 2022).

The numerous studies to analyze the evolution of mean temperature have been favored by the availability of simple distribution-free statistical tests, such as Mann–Kendall (MK) test (Mann, 1945; Kendall and Gibbons, 1990) and easy-to-use software to compute them. Extreme- and record-event type statistics exist since the 70s (Feller, 1991; Arnold et al., 1998; Bunge and Goldie, 2001) but they have not been widely applied within the climate community. The existing analysis of record-breaking temperatures usually aim to describe observed records and to identify the role of different factors in their occurrence (Xu et al., 2021; Zhang et al., 2021). Other works analyze the projected behavior of extreme records under different emission scenarios (Xu and Wu, 2019; Fischer et al., 2021; Yu et al., 2023).

Specific tools are required to analyze the tails of temperature distribution since their evolution may not be parallel to the evolution of the mean. In addition, if the magnitude of a trend in the mean is small in

* Corresponding author.
E-mail addresses: jorgecm@unizar.es (J. Castillo-Mateo), acebrian@unizar.es (A.C. Cebrián), jasin@unizar.es (J. Asín).

https://doi.org/10.1016/j.atmosres.2023.106934
Received 31 March 2023; Received in revised form 1 June 2023; Accepted 19 July 2023
Available online 25 July 2023

terms of the variability of the series, the effect on the extremes of that trend may be difficult to detect. This is the case of global warming trends in daily temperatures. On the other hand, in aggregated data like global mean temperatures, the natural variability is smaller and the effect of the climate change is easier to detect. For this reason, the studies to quantify the effects of global warming in the extremes have focused on the analysis of annual or monthly summaries of temperatures (Zorita et al., 2008; Coumou et al., 2013; van der Wiel and Bintanja, 2021; Salameh et al., 2019), although in this context the study of the effect in daily temperatures is more important because averaged or summarized data can under-represent warm or cool periods with a persistence of only a few days (Yosef et al., 2021).

Some works to analyze the effects of climate change on record-breaking temperatures use probabilistic properties of the occurrence of records in independent and identically distributed (i.i.d.) series to quantify their evolution (Redner and Petersen, 2006; Coumou et al., 2013; Wergen et al., 2014). Gouet et al. (2020) characterize the probabilities of record in sequences of variables with a linear trend in location, but they require restrictive assumptions about the distribution of temperature. In this context, it is of great interest to use the probabilistic properties of records to develop formal statistical tools, however, only a few works try to provide hypothesis tests to objectively establish the effects of climate change in the extreme and record-breaking temperatures. In this line, Benestad (2003), Benestad (2004) proposed record-statistic tests to detect non-stationarities based on records in forward and backward series and Monte Carlo integration, and applied them to analyze spatially aggregated monthly mean temperatures. Benestad (2008) underscored the utility of these tests in terms of evaluating trends in extremes. Meehl et al. (2009) and Anderson and Kostinski (2011) studied the ratio of daily and monthly record high maximum temperatures to record low minimum temperatures averaged across the US and compared it with the expected ratio under stationary conditions. Cebrián et al. (2022) and Castillo-Mateo (2022) provide a wide family of distribution-free record tests which can be applied without any assumption about the temperature distribution. Another important advantage of these tests is their high power even when the underling trend is small compared to the variability of data, what makes them useful to the analysis of any type of temperature signals, including daily signals. They can be applied both locally and regionally. Given that the effects of climate change show important differences depending on the climate and region, these tests are useful tools to assess and evaluate those differences. The analysis using these tools is highly facilitated by the `R` package (R Core Team, 2022) `RecordTest` (Castillo-Mateo, 2023; Castillo-Mateo et al., 2023). All the tests and graphical tools used in this work, as well as other tools based on records (Foster and Stuart, 1954; Diersen and Trenkler, 1996; Benestad, 2003; Benestad, 2004), are implemented in this package.

The contribution of this work is the use of the statistical tools developed by Cebrián et al. (2022) and Castillo-Mateo (2022) to assess and analyze the effect of global warming in the extreme and record-breaking events in different signals of 36 daily maximum temperature series in peninsular Spain over 1960–2021. In addition, a modification of the tests based on a permutation approach is proposed. This new approach can be used with correlated series, which allows the application of the tests in smaller regions and shorter periods of time. Using these tests, we evaluate and compare the effects of climate change in the extremes over different Spanish regions, globally and in different periods of the year.

The outline of the paper is as follows. Section 2 presents the time series of observed temperature data in 36 Spanish stations and an exploratory data analysis. Subsequently, the section introduces the methodology to detect non-stationarity in the occurrence of extreme and record-breaking events. Section 3 shows the analysis of the evolution over time of extreme and record-breaking temperatures in different daily temperature signals using Spanish series. It also shows relationships between records in the observed series and in temperature series at geopotential levels from ERA5 reanalysis. Section 4 concludes the paper with a discussion and conclusions.

2. Data and methods

2.1. Data

The database in this study has been extracted from the European Climate Assessment & Dataset (ECA&D; Klein Tank et al., 2002). It includes surface observations of daily maximum temperature (T_x) in °C in 36 stations located in the Spanish territory within the Iberian Peninsula, in the period 1960–2021, see Fig. 1. The Iberian Peninsula includes areas with very different climates (Chazarra-Bernabé et al., 2022). To a great extent, climate in inland Spain is temperate with a dry summer (Cs, Köppen classification), but some areas in the Central Plateau, Southeast of Andalusia and Ebro valley have a semi-arid climate (BS). North coast and Atlantic North coast show a temperate climate with no dry season (Cf). Mediterranean coast has a Cs climate in the North and BS in the Southeast. Our database includes series representing the different climate zones and also different elevation, 13 stations are over 500 m a.s.l., 5 stations over 800 m a.s.l. and one in the Central Mountains reaches 1894 m a.s.l. Climate summary measures of the locations are shown in Table S1 from Supplementary material.

All the series in the database have less than 100 missing observations in the considered period, this is, less than 0.5% of data are missing. Missing data are not removed from the series since the statistical tools applied in the analysis are weakly affected by a small percentage of missing data. Observations corresponding to February 29 are removed from the dataset for convenience. Furthermore, observed temperature data was rounded to the nearest tenth of a °C, resulting in a small probability of ties. Our analysis only considers strong records, this is, values strictly higher than the existing record. This means that the results are conservative in the sense that the number of records without the rounding effect could be around a 3% higher.

2.1.1. Exploratory data analysis

This sections shows some exploratory analysis to describe how the distribution of T_x has changed in time over the observed period 1960–2021. In particular, we compare the evolution over time of the median and the 5th and 95th percentiles of daily temperature. The aim of this analysis is to study if the effects of climate change are similar in the center and in the tails (extremes) of temperature distribution. If there exist relevant differences, the conclusions about the effects of global warming in the mean would not be valid for the tails, revealing the need of a specific analysis for the evolution of extreme and record-breaking temperatures.

To compare the evolution of the distribution of T_x, we plot a kernel density estimation of T_x for July 15, in two reference periods 1961–1990 and 1991–2020, and the corresponding 5th, 50th and 95th empirical percentiles. The densities are estimated using data in a centered window of 21 days. Fig. 2 shows the plots for Sevilla and Barcelona-Fabra, as an illustration of two different climates. Same plots for San Sebastián and Zaragoza, and plots for January 15 are shown in Figs. S1 and S2 from Supplementary material, respectively. In Sevilla, a clear positive shift of the entire distribution is observed in the last period. However, Barcelona-Fabra shows a different evolution in the center of the distribution and in the extremes, since the increase of the median is higher than in the 5th and 95th percentiles. Figs. S3 and S4 from Supplementary material show the difference between both 30-year periods in the mean value and the standard deviation across days within year for Sevilla, Barcelona-Fabra, Zaragoza, and San Sebastián.

Fig. 3 represents the difference between the period 1991–2020 minus the period 1961–1990 of the median and the 95th percentile of T_x. The percentiles are computed daily using a centered window of 31 days. The different evolution of the differences between the two periods reveals that the long-term trend in the center of the distribution is

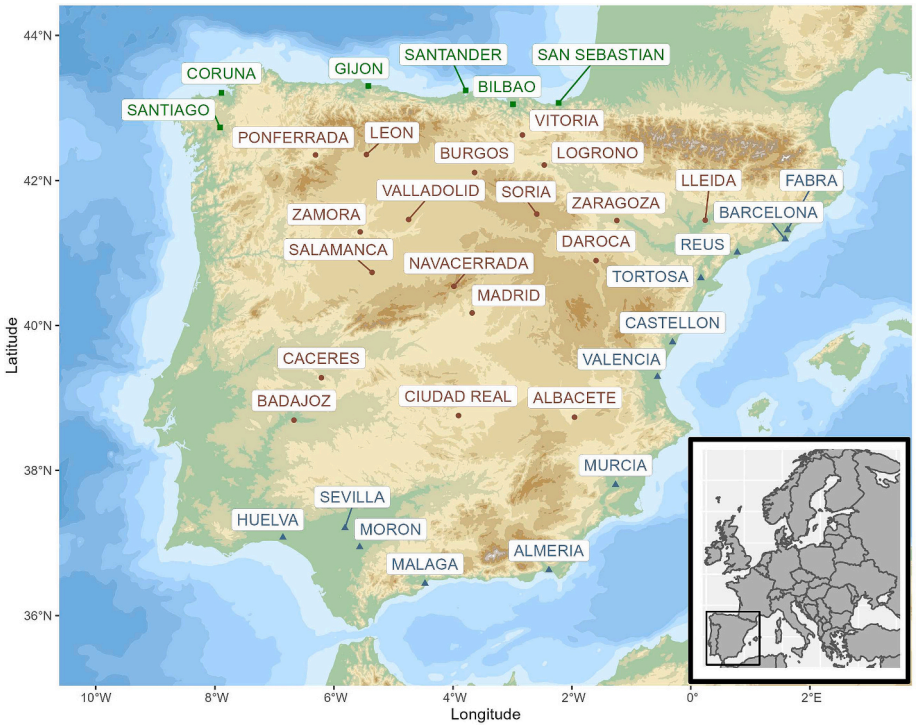

Fig. 1. Map of the 36 Spanish stations, with different climate regions: North Atlantic (green square), Continental (red dot), and Mediterranean (blue triangle).

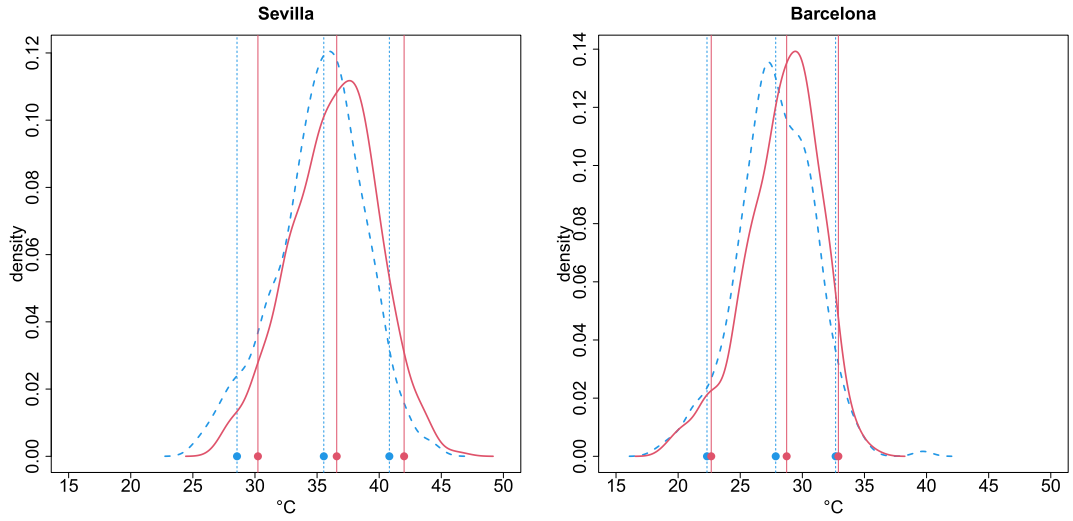

Fig. 2. Kernel density estimation of T_x in Sevilla and Barcelona-Fabra on July 15 in the periods 1961–1990 (blue dashed) and 1991–2020 (red solid). Vertical lines show the 5th, 50th and 95th empirical percentiles.

different to the trend in the extremes. In Barcelona-Fabra, the difference between the two periods is always positive in both the median and the 95th percentile. In Sevilla, the difference is also positive except for the 95th percentile in the early autumn. In Sevilla, the difference in the mean is primarily higher or equal than the difference in the 95th percentile throughout the year. This effect is observed in Barcelona-Fabra only during the summer.

To analyze spatially the different evolution of the median and the 95th percentile, we estimate a linear trend over time (°C/decade) of the median and the 95th percentile of T_x in each location and month using quantile regression. We calculate the difference between those trends and we plot, for each location, the mean of those differences in the 12 months versus elevation (m a.s.l.) at the location, see left plot in Fig. 4. Positive values indicating trends of the median higher than trends in the 95th percentile are observed in 29 locations, in particular in all locations over 500 m a.s.l. Negative values are observed only in 7 locations near the coast: San Sebastián, Bilbao and Santiago (Cantabrian coast), and Tortosa, Barcelona-Fabra, Barcelona-Airport and Almería (Mediterranean coast). This suggests that, in some areas near the coast, the increase of the extremes tends to be stronger than the increase in the mean. More details for Sevilla, Barcelona-Fabra, Zaragoza, and San Sebastián are shown in Fig. S5 from Supplementary material.

To study if the behavior of the trends is different within the year, the right plot in Fig. 4 summarizes, for each month, the differences between trends in the 36 locations using a boxplot. It shows that positive differences are observed in most of the locations (positive boxplot-median) in all the months except, January, April and November. That means in particular that, in most locations, trend in the median is higher than in

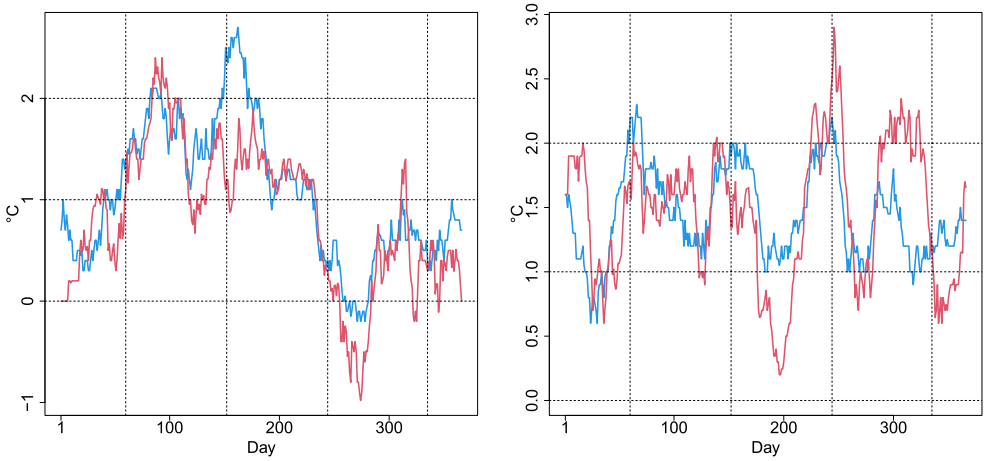

Fig. 3. Differences of the median (blue) and the 95th percentile (red) between the periods 1961–1990 and 1991–2020 of T_x in Sevilla (left) and Barcelona-Fabra (right).

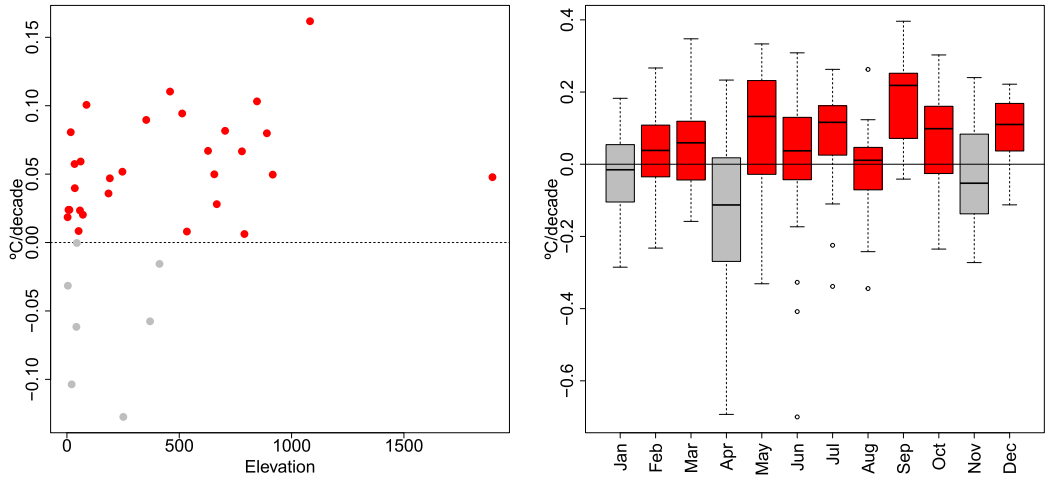

Fig. 4. Mean over the 12 months of the differences between the time trend of the median and the 95th percentile of T_x for each location versus elevation; positive values in red (left). Boxplots by month of the previous differences between trends in the 36 locations; months with positive medians of the differences in red (right).

the 95th percentile in the extended summer period from May to October.

To sum up, the exploratory analysis shows that the time evolution of the central part of the distribution of T_x is not always parallel to the evolution of the tails. Consequently, the estimated effect of global warming in the mean should not be used to describe the effects in records and extreme temperatures. Due to the characteristics of extremes, which are rare by definition, specific statistical tools are needed to analyze their evolution. Another important feature of the series to be considered in their analysis is the spatial and temporal dependence between them.

2.2. Methods

Given a series (X_t), an observation X_i is called an upper record if it has a higher value than the previous observations, this is, if $X_i > \max_{t<i}\{X_t\}$. Analogously, X_i is called a lower record if $X_i < \min_{t<i}\{X_t\}$. All the properties for upper records are valid for lower records because $\min_{t<i}\{X_t\} = -\max_{t<i}\{-X_t\}$. The sequence of record indicator binary variables (I_t), with I_t taking value 1 if a record is observed at time t and 0 otherwise, characterizes the occurrence of records in a series. From this sequence, the number of records up to time t is defined by $N_t = \sum_{i=1}^{t} I_i$.

There are some probabilistic results that characterize the occurrence of records in a series (X_t) of i.i.d. continuous random variables. The first result states that the variables (I_t) are mutually independent with I_t following a *Bernoulli*(p_t) distribution, where the probability of record at time t is

$$p_t = P(I_t = 1) = 1/t, \qquad t = 1, 2, \ldots.$$

This means that in a stationary climate, the probability of observing a record decreases over time but there is always a positive probability of occurrence. Concerning the behavior of the number of records in i.i.d. series, the variables N_t have an asymptotic normal distribution where the expected number of records is $E[N_t] = \sum_{i=1}^{t} p_i = \sum_{i=1}^{t} 1/i$ and the variance $Var[N_t] = \sum_{i=2}^{t} p_i(1-p_i)$.

2.2.1. Record tests for non-stationarity detection

The null hypothesis H_0 of the tests used in this work is that the probability of record at each time t in a series of length T is the probability of record under the stationary condition characterized by i.i.d. series, this is, $1/t$. Under climate change, it is expected that temperatures show an increasing trend and/or an increasing variability. Under those conditions, the probabilities of upper record are higher than in i.i.d. series. Consequently, the following one-sided alternative hypothesis is

considered,

$$\begin{aligned} &H_0 : p_t = 1/t, \quad t = 2, \ldots, T. \\ &H_1 : p_t > 1/t, \quad \text{for at least one} \quad t = 2, \ldots, T. \end{aligned} \tag{1}$$

The details of the tests can be found in Cebrián et al. (2022). They proposed a family of distribution-free tests to detect deviations from i.i.d. series in the tails of the distribution using the probabilistic properties of the occurrence of records. The underlying idea of these tests is to compare the expected behavior of the occurrence of records in i.i.d. series with the behavior over time of the observed records. Although the tests detect any deviation from i.i.d. series, if the analyzed series is formed by independent observations with no seasonal behavior, deviations from the i.i.d. hypothesis suggest the existence of trends or changing variability, which are the usual features expected under climate change.

The tests assume the availability of $M \geqslant 1$ mutually independent series of length T. These series can be series measured at different spatial points or series obtained from splitting the original data. It is noteworthy that the tests do not require that the M series have the same distribution. From these M series, the series of binary variables $(I_{t1}), (I_{t2}), \ldots, (I_{tM})$, and the series of number of records $(N_{t1}), (N_{t2}), \ldots, (N_{tM})$ are obtained.

Record tests based on N_T. The most basic statistic is the total number of records in the observed period of length T in the M series:

$$N = \sum_{m=1}^{M} N_{Tm} = \sum_{m=1}^{M} \sum_{t=1}^{T} I_{tm}.$$

Under the null hypothesis H_0, N_T is asymptotically normal when T and/or M tend to ∞. The power of N is improved by weighting the record indicators according to their position in the series, this is,

$$\mathcal{N} = \sum_{m=1}^{M} \sum_{t=1}^{T} w_t I_{tm}.$$

The idea is to use weights which are increasing functions of t since, given that records become less likely for increasing time, the occurrence of a record at a high t gives more evidence against the null hypothesis H_0. Although different weights can be used, Cebrián et al. (2022) shows that the weights $w_t = t^2/(t-1)$ ($w_1 = 0$) give the locally most powerful unbiased score test. Under the null hypothesis H_0, $\mathcal{N}$ is still asymptotically normal in M. Using this asymptotic distribution, the p-value of the test to study the previous hypothesis is $P(Z > (\mathcal{N}_0 - 0.5 - \mu)/\sigma)$ where μ and σ^2 are the mean and variance of the statistic under the null hypothesis H_0, $\mathcal{N}_0$ is the observed statistic, Z is a standard normal variable and 0.5 is a continuity correction. These tests are implemented with the function `N.test` in `Recordtest`.

Further, the asymptotic normal distribution of the statistics N and $\mathcal{N}$ can be used to compute reference intervals (RI's) of the number of records and of the weighted number of records up to time t. This is useful to implement plots to analyze the evolution of the number of records over time. These graphical tools are implemented with the function `N.plot` (Castillo-Mateo et al., 2023).

Joining information from different types of records. It is noteworthy that four different types of records can be obtained from one series. The upper and lower records in the forward series, $X_1, X_2, \ldots, X_T$, and in the backward series obtained when the order of the variables is reversed, $X_T, \ldots, X_2, X_1$. Fig. 5 illustrates the times of occurrence of the four types of records in T_x series for June 30 in Barcelona-Fabra. The advantage of considering four types of records is that many different hypotheses can be studied combining them adequately. Further, the power of the tests based on different types of records is higher than those using only one type (Foster and Stuart, 1954; Diersen and Trenkler, 1996; Cebrián et al., 2022). Herein, the superscripts L and B indicate lower records and records in backward series. For example, I_t^L denotes the binary variables for lower records, and I_t^{BL} for lower records in the backward series.

Under the null hypothesis of i.i.d. series, all variables $I_t, I_t^L, I_t^B, I_t^{BL}$ follow a *Bernoulli*$(1/t)$ distribution. Using this property, it is easy to combine the information provided by them in different statistics that allow the study of different alternative hypotheses. It is noteworthy that by considering different types of records in different subseries and in the forward and the backward series, we obtain information from the tails of temperature distribution, not only from the observed record events. Similarly to other approaches to study the extremes of a distribution, such as annual maxima or peak over threshold methods, this information can be used to characterize the behavior of the upper tail, or even both tails, of temperature distribution.

In particular, we can define statistics to analyze:

- The effect of an increasing trend in the upper tail, by considering the upper records in the forward and also in the backward subseries, $\mathcal{N}_{upp} = \mathcal{N} - \mathcal{N}^B$.
- The effect of an increasing trend in the lower tail, by considering now the lower records, $\mathcal{N}_{low} = \mathcal{N}^{BL} - \mathcal{N}^L$.
- The effect of an increasing trend in both tails, by considering both the lower and the upper records in both the forward and the backward subseries, $\mathcal{N}_{both} = \mathcal{N} - \mathcal{N}^L - \mathcal{N}^B + \mathcal{N}^{BL}$.

Note that the sign of the statistic for each type of record is positive or negative according to whether a higher or lower number of records is expected under the considered alternative hypothesis H_1. The effect of an increasing trend in the lower tail implies that the probability of lower record is smaller than $1/t$. In the backward series an increasing trend becomes a decreasing trend so the probabilities of upper and lower re-

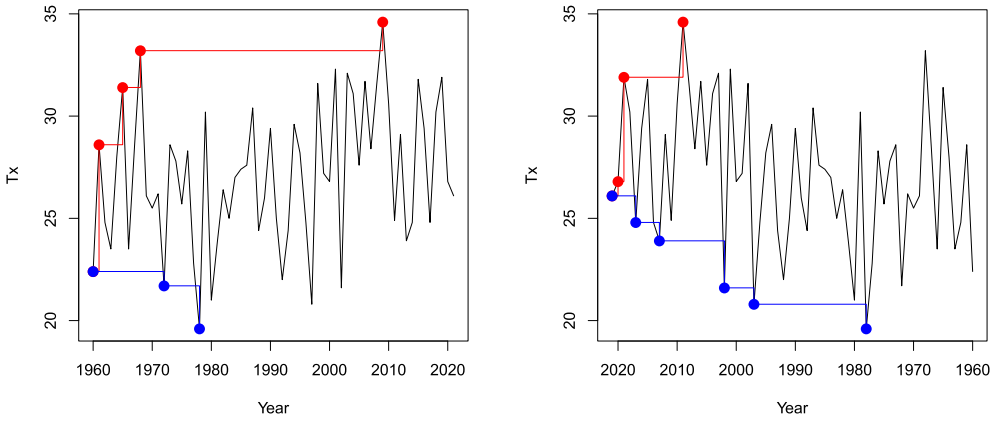

Fig. 5. T_x series for June 30 in Barcelona-Fabra and occurrence times of upper records (red) and lower records (blue) in the forward (left) and backward (right) series.

cords under H_1 are respectively lower and higher than $1/t$; see Fig. 5 as illustration.

We can also analyze the effect of an increasing variability in the tails, taking into account that the increase will lead to a higher number of lower and upper records in the forward series and lower in the backward,

$$\mathscr{N}_{var} = \mathscr{N} + \mathscr{N}^{L} - \mathscr{N}^{B} - \mathscr{N}^{BL}.$$

Under H_0, all these statistics have an asymptotic normal distribution with zero mean, and the corresponding p-values can be obtained as usual. The function `foster.test` in `RecordTest` implements these tests. The power of these tests with $M > 12$ and $T > 50$ is high even with trends in location around 1% of the standard deviation but it decreases with lower M (Cebrián et al., 2022).

2.2.2. Record tests for change-point detection

Castillo-Mateo (2022) proposed three distribution-free statistics to detect a change-point if the record occurrence stops being stationary. The statistics test the null hypothesis H_0 in (1) against the two-sided alternative hypothesis

$$H_1 : p_t = 1/t, \quad t = 1, \ldots, t_0, \qquad \text{and} \qquad p_t \neq 1/t, \quad t = t_0 + 1, \ldots, T,$$

where t_0 is the change-point. The test statistic for change-point detection in the record occurrence is $\mathscr{K} = \max_{1 \leqslant t \leqslant T} |K_T(t)|$, where

$$K_T(t) = \frac{N_t - E(N_t)}{\sqrt{Var(N_T)}} - \frac{Var(N_t)}{Var(N_T)} \frac{N_T - E(N_T)}{\sqrt{Var(N_T)}}.$$

The change-point $\widehat{t}_0$ is defined as $\widehat{t}_0 = \arg\max_{1 \leqslant t \leqslant T} |K_T(t)|$. Under the null hypothesis H_0, the distribution of $\mathscr{K}$ is Kolmogorov in the limit as $T \to \infty$. Thus in a two-sided test for a change-point in the record occurrence, the null hypothesis H_0 is rejected if $\mathscr{K} > k_{\alpha/2}$, where $k_{\alpha/2}$ is the upper $\alpha/2$th quantile of the Kolmogorov distribution and α is the significance level for the test; and a significant change-point occurs at time $\widehat{t}_0$ if the null hypothesis H_0 is rejected.

Henceforth, according to Castillo-Mateo (2022), the change-point statistic also considers M series and weights for the record indicators with $w_t = \sqrt{t^2/(t-1)}$ ($w_1 = 0$). The statistic does not follow the Kolmogorov distribution, but the p-value can be estimated using Monte Carlo simulations. Castillo-Mateo (2022) also includes a detailed Monte Carlo analysis of the power and ability of this statistic to detect the actual change-point. The estimator is right-sided biased, which can be usefully interpreted as the time when the underlying process that drives the temperature distribution truly affects the records in the observed data. In other words, the change-point is determined by the occurrence of a record; if no records are observed, there is no change-point. This bias decreases significantly when the number of series M is increased or when the effect of the change becomes greater under the alternative hypothesis H_1. This test is implemented with the function `change.point` in `RecordTest`.

2.2.3. Applying the tests to temperature series

Temperatures, as most environmental time series, measured in an intra-annual temporal scale show seasonal behavior and often serial correlation. Since the previous record tests detect any deviation from i.i.d. series, when the aim is to detect deviations provoked by climate change (increasing trend or changing variability), we need to remove first seasonality and serial correlation from the series. Another practical limitation of the previous tests is that they require M independent series, and independence is a quite restrictive assumption in climate series. This section shows how to prepare the data and a modification of the tests to solve these limitations.

Seasonal and serially correlated series. A common approach in environmental studies to deal with seasonality and serial correlation is to split the series of daily observations into 365 series, so that each series contains the observations from one calendar day across T years:

$$\begin{array}{c} \\ \text{Year } 1 \\ \text{Year } 2 \\ \vdots \\ \text{Year } T \end{array} \begin{array}{c} \begin{array}{cccc} \text{Day } 1 & \text{Day } 2 & \cdots & \text{Day } 365 \end{array} \\ \begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,365} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,365} \\ \vdots & \vdots & & \vdots \\ X_{T,1} & X_{T,2} & \cdots & X_{T,365} \end{pmatrix} \end{array}. \tag{2}$$

The series in each column consists of independent observations with no seasonal behavior and no serial correlation. The resulting 365 series do not necessarily have the same distribution because the distribution in December is probably different from the distribution in August. However, this is not an assumption of the tests, since the probability of record under the null hypothesis H_0 does not depend on the distribution of the series.

The transformation of one series into $M = 365$ subseries is also useful to obtain a high number M of series to apply the tests, and consequently to increase their power. The problem is that, given that the series are measured in consecutive times (days in this case), they will probably show a strong dependence between them, and the tests require independent series. If the number of dependent series is high enough, the simplest option is to extract among them a subset of M independent series. This can be done, selecting series separated by a fixed distance (for example, 10 days) or applying an approach based on Pearson correlation tests (Castillo-Mateo et al., 2023). A second option is to use tests that allow dependent series, as described below.

Tests for a set of M dependent series. The asymptotic distribution of the record statistics relies on the assumption that the M studied series are independent. To avoid this restriction, we propose an alternative approach where the p-values of the statistics defined in Section 2.2.1 are computed using permutation techniques.

Permutation tests only rely on the assumption of exchangeability under the null hypothesis (Welch, 1990). A sample is exchangeable if any permutation of it has the same joint probability distribution. In the record tests, there is a sample of $t = 1, \ldots, T$ observations of a vector of M variables $(X_{t1}, X_{t2}, \ldots, X_{tM})$. Under the null hypothesis, the T observations of the vector $(X_{t1}, X_{t2}, \ldots, X_{tM})$ are independent with the same multivariate distribution, so that permutations of rows, see the data structure in (2), are exchangeable. Note that the M variables in the vector may be correlated between them or may have different marginal distributions, the only required assumption is that the multivariate distribution of the vector is the same over the T years.

If observations are exchangeable under H_0, and all the possible permutations are considered to compute the p-value, the resulting test yields the exact significance level. In general, a random sample of possible permutations must be used because the total number of different permutations is too high. The p-value is computed as the proportion of samples whose statistic value is greater or equal than the observed statistic. A number of 10,000 permutations gives a good approximation to the p-value.

3. Results

3.1. Analysis of daily temperature series in one location

This section aims to show how the record tests previously defined work, and the different hypotheses that can be analyzed with them. To that end, a detailed analysis is presented using the longest series available, the daily maximum temperature series in Barcelona-Fabra from 1914 to 2021.

The observed number of upper and lower records in the period 2001–2021 in the 365 daily series is 224 and 23, respectively. The number of upper records is 10 times the number of lower records while, in a stationary climate, the expected number of records in both cases is

the same and equal to 78.5. In this case, where the observed period is $T = 108$ years long, even this simple exploratory analysis suggests the effects of an increasing trend, but in many series, the observed period is shorter and more formal inference tools are needed to evaluate the effect of global warming.

Before applying the record tests, given the seasonal behavior and serial correlation of daily temperature in Barcelona-Fabra, the processing tools described in Section 2.2 have to be used. The resulting final data is a subset of $M = 46$ uncorrelated series of length $T = 108$ with no seasonal behavior and no serial correlation.

We start by analyzing graphically the evolution of the number of records over time, to detect possible deviations from the stationary behavior and to identify when they appear. To that end, the mean, in M series, of the number of records up to time t, $\overline{N}_t = \sum_{m=1}^{M}\sum_{i=1}^{t} I_{im}/M$, for $t = 1, \ldots, T$, is plotted versus t together with RI's for the number of records in i.i.d. series. An analogous plot is obtained for the weighted number of records $\overline{\mathscr{N}}_t = \sum_{m=1}^{M}\sum_{i=1}^{t} w_i I_{im}/M$. Fig. 6 summarizes the evolution of the four types of records, in both cases. In the four types of records there is a significant evidence of a non-stationary behavior in the upper tail (forward and backward upper records), and even earlier in the lower tail (forward and backward lower records). The evidence is clearer, and consequently it is detected earlier, in the backward series and using weights.

To obtain more formal conclusions about the effect of global warming in extreme temperature, we can apply the record tests and study different hypotheses. To analyze the behavior of the upper records, we apply the test $\mathscr{N}$ with the alternative that the probabilities of the upper records are higher than in i.i.d. series. The p-value is $1.8e-05$, so that the null hypothesis is rejected at any usual significance level. To analyze the behavior of the upper extreme temperatures, not only the records, we apply the test $\mathscr{N}_{upp}$; the null hypothesis is also rejected with a p-value $8.4e-7$. To study the lower tail and both tails simultaneously, we have to use the statistics $\mathscr{N}_{low}$, and $\mathscr{N}_{both}$, respectively, they yield p-values equal to $7.5e-7$, and $8.8e-12$. We conclude that there is evidence of the effect of an increasing trend in both the upper and the lower tail and this evidence is stronger when information of both tails is joined. Since effects of an increasing trend are detected in both tails, it can be expected that the variability is not increasing; this is confirmed with the p-value of $\mathscr{N}_{var}$ test, 0.5066, that leads to not reject the null hypothesis. Finally, the change-point test detects a significant change-point at time $\widehat{t}_0 = 64$ (1977) with a p-value 0.0029.

3.2. Spatial analysis of temperatures

Climate change is a spatial phenomenon, so the interest lies in the analysis of the behavior of the temperature records and extremes over the peninsular Spain. To make fair comparisons of the behavior across different locations, the same period of time should be analyzed in all of them, so that the set of 36 series from 1960 to 2021 introduced in Section 2.1 is considered.

It is noteworthy that since 36 series are analyzed, corrections for multiple comparisons should be applied. Given that the test statistics from nearby locations will be possibly dependent, and although this dependence may become negligible between the farthest locations, we opted for applying the conservative Benjamini-Yekutieli method (Benjamini and Yekutieli, 2005). This method controls the expected proportion of false rejections and can be applied to p-values from dependent statistics.

3.2.1. Analysis of daily temperature series

The first hypotheses that we study are the effect of an increasing trend in the upper records and in the upper tail using statistics $\mathscr{N}$ and $\mathscr{N}_{upp}$, respectively. Fig. 7 shows the maps summarizing the results. These maps show with points in a red color scale the locations where the p-value is lower than $\alpha = 0.1$, and in a blue color scale, higher values. For the upper tail there are 29, 25, and 22 series with significant p-values at level $\alpha = 0.10, 0.05$ and 0.01 respectively, and for the upper records, 23, the same 23, and 18. There are evidences of the effect of an increasing trend in both upper records and upper tail in most of the studied locations; as expected, there are more evidences in the upper tail, where more information is available. Only in the North area, and specially in the Cantabrian coast, there is no significant evidence of the effects of an increasing trend. The effects are weaker also in other locations in the coast, for example Huelva, in the South, and Castellón and nearby locations in the Levante coast.

Analysis within the year. It is known that effects of climate change may not be homogeneous over the year, so that it is necessary to analyze the potential effects of a trend in shorter periods of time within the year, seasons or even months. In those cases, M, the number of independent series available is usually small, so that the power of the asymptotic tests is low with weak trends. Then, it is preferable to apply the tests based on permutations described in Section 2.2.1, using all the daily series available in the considered period. Observations from subsequent days are dependent, but permutation tests allow for the analysis of dependent columns in matrices like (2). Fig. 8 shows the p-values from the tests to study the effect in the upper tail of the daily maximum temperature in each season: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). It is clear that the effects are not homogeneous over the year: autumn is the season where the effects are weaker, being clearly significant only in the North part of the Mediterranean coast. On the other hand, in summer, significant evidences are found all over Spain except in Cantabrian coast and some nearby locations. Note that in this area, only in winter, there are some weak evidences of the effect of the trend in the extremes. The analysis of the upper records shows a similar pattern to the upper extremes, although the p-values tend to be slightly higher. In particular, in autumn, significant evidences at a $\alpha = 0.05$ significance

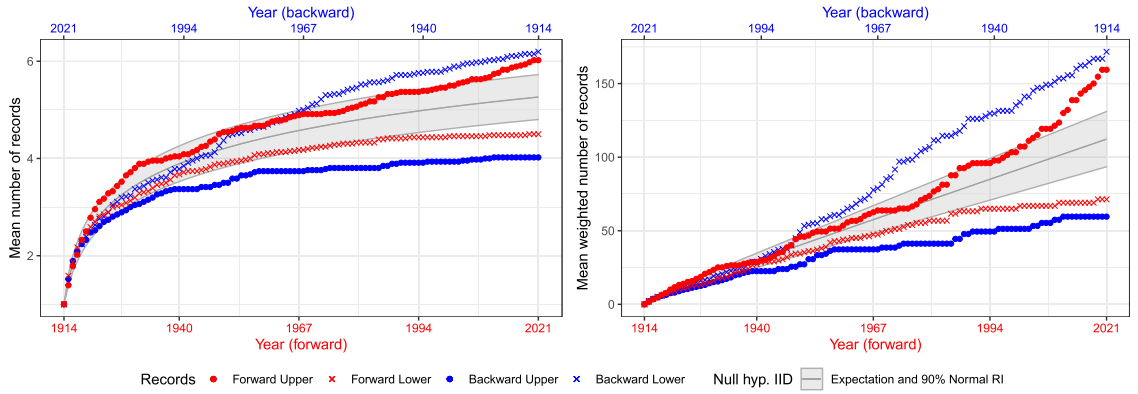

Fig. 6. Mean (left) and mean weighted (right) number of records up to time t versus t, in daily maximum temperature in Barcelona-Fabra (1914–2021).

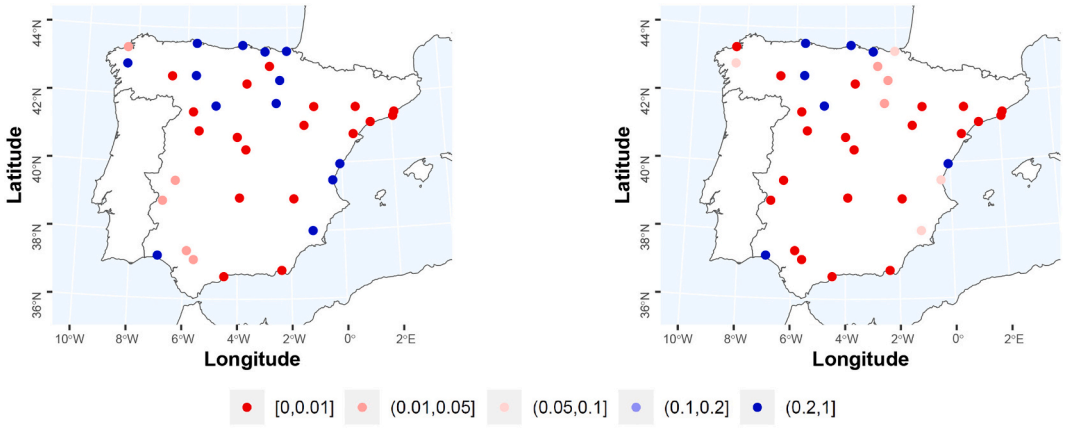

Fig. 7. P-values from the analysis of the effect of a trend in the upper records (left) and in the upper tail (right) in daily T_x series during the entire year in Spanish locations.

level are found only in some areas in the Mediterranean coast, see right plot in the last row in Fig. 8.

Even within seasons, the behavior may not be homogeneous, so that the same tests are applied by month (maps not shown). Concerning the upper extremes, no significant effect of a positive trend is found in any location, in February, March and November, and in September only in two Mediterranean locations (Barcelona-Fabra and Tortosa). On the other hand, June, July and August are the months with largest areas of Spain with evidences of the effects of a trend, both in the upper tail and the upper records. As an example, the p-values from the upper tail in June are shown in Fig. 8.

Regional analysis. Since global warming is a spatial phenomenon, it is of interest to analyze its effects at a regional scale, by joining information from all the series available in a region. The proposed tests can be used to study this type of hypothesis, using the permutation approach if it cannot be assumed that the series available in a region are independent, as it is often the case. Note that the regional analysis will be more powerful since joining series from different locations, more information is provided to the tests.

Here, we consider three regions with different climate characteristics: the North Atlantic, the Continental and the Mediterranean areas. See in Fig. 1 the locations by color in each region. The effect of a positive trend in the upper tail and in the upper records is significant at a significance level $\alpha = 0.05$ in all the regions in the four seasons. Table 1 shows the years identified by the change-point test to detect when the upper record occurrence is significantly different from the occurrence in i.i.d. series, in each season and in the whole year. The behavior is quite homogeneous all over Spain, with slight differences between the three regions, but relevant differences appear between seasons. Winter is the season where the evidences of non-stationarity in the upper records appear first, around 1975, while in autumn the change-points occur 30 year later. In spring and summer the non-stationarity is detected from around 1985.

The regional analysis is also developed by month, and the resulting p-values are summarized in Fig. 9. Again, the different p-values are shown in a red–blue scale according to their significance. In this case, in each region, the 12 p-values corresponding to one month are corrected using the Benjamini-Yekutieli approach. The results show that February, September and November are the months where the upper extremes and upper records are less affected by global warming, with no significant effect in any of the climate regions. In contrast, April and October are the months where a significant behavior is observed all over Spain. The Mediterranean region is the area where the number of months with strong evidences is higher, in all the year except the three months previously mentioned, while the North Atlantic region is the area where less evidences of non-stationarity occur; only in April and October a significant behavior is detected. Concerning the upper records, the same pattern is observed, although the effects are slightly weaker.

3.2.2. Analysis of annual maximum and annual minimum of T_x

In this section, the behavior of the records and both tails of the annual maxima and the annual minima of T_x, the daily maximum temperature, are analyzed. Unlike the case of daily series, only one series per location is available for this analysis. Given that with $M = 1$, the power of the test is low for weak trends, only the regional analysis is carried out. Table 2 summarizes the p-values of the tests to assess the effect of an increasing trend in the upper tail, the upper records, the lower tail and the lower records in the annual maximum and in the annual minimum of T_x, in the three climate regions. The effect in the tails, both in upper and lower, is significant at a significance level $\alpha = 0.05$ in all the regions. In the records, the effect is slightly weaker, specially in the lower records of the annual maxima, where the p-values are not significant at a $\alpha = 0.05$ level in any region. This implies that there are not enough evidence to state that the occurrence of lower records is not lower than $1/t$ in that signal, and could suggest an increase of variability. However, the tests to detect variability, based on statistic $\mathcal{N}_{var}$, are not significant at any usual significance level, with p-values 0.16, 0.53 and 0.74; this means that there is no significant evidence of an increase of variability in annual maximum temperature in any of the three considered regions.

3.3. Records in T_x and daily temperatures at geopotential levels

The aim of this section is to analyze the occurrence of records in temperatures at geopotential levels, 850, 700, 500 and 300 hPa, to identify whether it is related to the occurrence of records in the surface temperature.

3.3.1. Data and exploratory analysis

The considered database includes the daily temperature series at 850, 700, 500 and 300 hPa levels, measured at 12:00 from 1960 to 2021, in the 176 points from the $1^\circ \times 1^\circ$ grid 35°N–45°N and 10°W–5°E that covers the Iberian Peninsula. This database is obtained from the 5th generation ECMWF reanalysis database ERA5 available in Climate Data Store (CDS) of Copernicus Climate Change Service (C3S) (Hersbach et al., 2023), see map in Fig. S6 from Supplementary material.

As in surface temperature signals, the effects of global warming are clear in the mean evolution of air temperature at 850, 700, 500 and 300 hPa geopotential levels. The mean temperature at the four levels shows an increase in the period 1960–2021, but the magnitude of that increase and the spatial variability are different. The effect of relief on the spatial behavior is more clear on temperature at 850 hPa level, where the

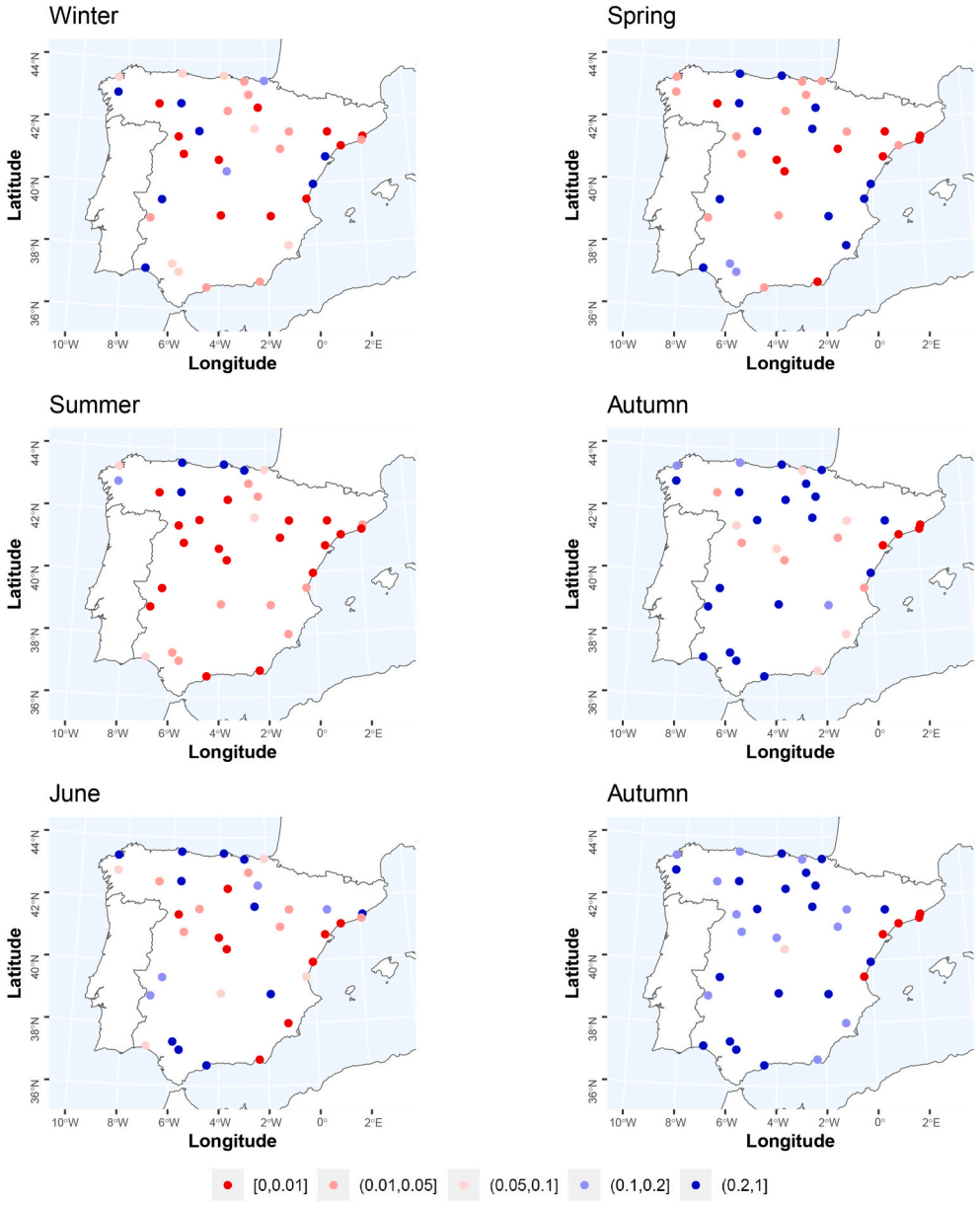

Fig. 8. P-values from the analysis of the effect of an increasing trend in the upper extremes in T_x by season in 36 Spanish locations (first and second row). Same map only for June (left plot last row) and for the upper records in autumn (right plot last row).

Table 1
Time points identified by the change-point test to detect when the upper record occurrence is significantly different from the occurrence in i.i.d. series, in each season and in the whole year.

Region	Series				
	Year	winter	spring	summer	autumn
North Atlantic	1976	1976	1980	1988	2010
Continental	1986	1973	1988	1986	2003
Mediterranean	1979	1973	1986	1986	2008

increasing trend over time of temperature varies from 0.20 in the North-West to 0.35°C/decade in the center of the Iberian Peninsula. At upper levels, the spatial variability is lower with trends from around 0.20 to 0.26°C/decade, with the highest increases in the South-West. Maps of the estimated trend at each geopotential level are shown in Fig. S7 from Supplementary material.

Working with the original time series, an absolute record is defined to be a value that exceeds all previous values in the series. Considering the 704 daily series of temperatures, one associated to each of the four geopotential levels in each of the 176 grid-points, the median number of absolute records since 1962 is 6. Given that four air temperature series are available at each grid-point we identify simultaneous records, based on the idea of compound events (Zscheischler et al., 2020), that represent the occurrence of a highly extreme situation in the geopotential temperature series. A "compound record" at a grid-point in a day is defined when simultaneous absolute records occur in at least 2 geopotential level series. Table 3 on the left summarizes dates since 1971 with compound records in at least 6 points of the grid. It shows on the right dates when the absolute record occurs simultaneously in more than two T_x series. Simultaneous compound records are temporally close to simultaneous absolute records in T_x series, for example, the episode 13–16 August 1987, or the unprecedented heatwave in 26–29 June 2019, that caused record-breaking high temperatures in about one-third areas of Europe (Sousa et al., 2019; Xu et al., 2021). An exploratory analysis of these compound records is shown in Tables S2 and S3 in the

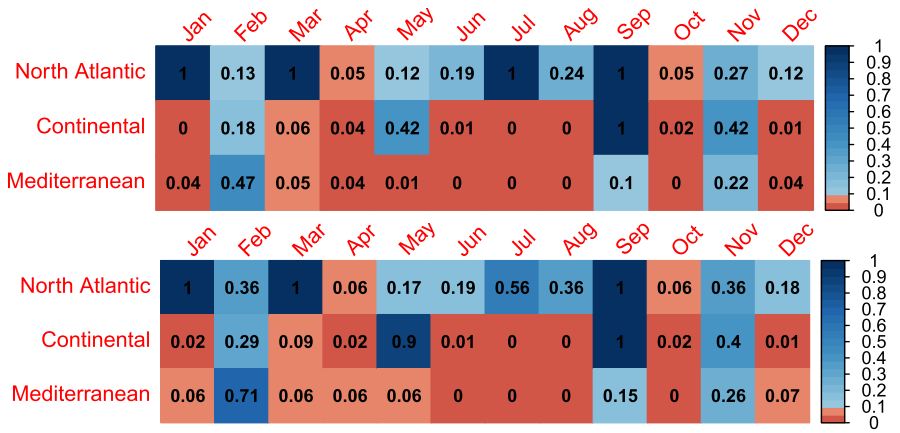

Fig. 9. P-values from the analysis of the effect of a positive trend in the upper tails (top) and upper records (bottom) in T_x by month in the three climate regions.

Table 2
P-values of the effect of an increasing trend in the upper tails (UT), the upper records (UR), the lower tail (LT) and the lower records (LR) in annual maximum and minimum of T_x.

	Annual max.				Annual min.			
Region	UT	UR	LT	LR	UT	UR	LT	LR
North Atlantic	0.000	0.002	0.023	0.104	0.007	0.009	0.013	0.013
Continental	0.000	0.000	0.021	0.121	0.041	0.067	0.005	0.002
Mediterranean	0.011	0.007	0.012	0.056	0.008	0.013	0.039	0.013

Table 3
Dates from 1971 with compound records in at least 6 points of the grid. Right column shows next dates with absolute record simultaneously in more than two T_x series.

Compound event	Absolute record in more than 2 stations
1980–8–2	
1983–7–30	
1987–8–13, 1987–8–16	1987–8–12, 1987–8–13
2019–6–26, 2019–6–27	2019–6–29
2021–7–10	

Supplementary material.

3.3.2. Time evolution of records in daily temperature at geopotential levels

Our aim in this section is to compare the time evolution of the number of records and the time where non-stationary behavior is detected in surface and geopotential level temperature series. To that end, the weighted cumulative number of records, $\overline{\mathcal{N}}_t$, in the observed temperature series T_x in a location and in temperature series at a given geopotential level in the four grid-points around that location, are plotted in the same graph. In the same plot, we show the years where the estimated change-point occurs in each of the five series. Fig. 10 shows, as an illustration, the graph for Barcelona-Fabra in July for the four geopotential levels. At 850 and 500 levels, the time evolution of the records and the estimated change-points are very similar in T_x and in the four grid-points and, at 700 hPa, in two of the grid-points. However, the estimated change-point at 300 hPa occurs later, around 1998. The same plots for Barcelona-Fabra in January and Sevilla in July are shown in Figs. S8 and S9, respectively; and a similar plots based on the 36 series is shown in Fig. S10 in the Supplementary material.

4. Discussion and conclusions

This work aims to analyze and evaluate the effects of global warming in the upper tail (extreme and record-breaking events) of daily maximum temperature series over peninsular Spain. Since, an exploratory analysis shows that the time evolution of the central part of the distribution of daily maximum temperature is not parallel to the evolution of the tails, it is necessary to perform a specific analysis of records and extreme temperatures. Here, we propose the use of statistical tools and record tests described in Cebrián et al. (2022) and Castillo-Mateo (2022) to assess the effect of global warming in the upper tail in different temperature signals in Spain, at different time and spatial scales. The analysis of short periods of time, for example months, and regional analysis imply the study of series measured in close days and/or in close locations, which are correlated. Given that the previous tests require independent series, we propose a modification based on a permutation approach to compute the p-values, which can be used with correlated series. Using these tests, we evaluate and compare the effects of climate change in the extremes and the record-breaking events all over Spain, in different Spanish regions, and in different periods of the year. All the record tests and statistical tools used in the analysis are freely available in the `R` package `RecordTest`. The proposed tests can be applied to analyze the evolution of extremes and records in observed series and also in gridded data obtained from reanalysis or Earth System Models, taking into account the spatial dependence between them.

As far as we know, this is the first study of temperature records developed in Spain that does not focus on a particular event. Sousa et al. (2019) described the intense heatwaves striking the Iberian Peninsula in early August 2018 and late June 2019. They found the Saharan air intrusions as a relevant mechanism for Iberian heatwaves. Their results are in agreement with the compound events and absolute records on late June 2019 observed in our work. Also there are some analyses about extremes, for example, Serrano-Notivoli et al. (2022) studied the heatwaves in Spain with a high-resolution gridded daily temperature dataset 1940–2014. They found a tipping point in the early 1980s from which heatwaves became more frequent, this is compatible with our results in Table 1 for summer. In other regions of the world, some authors found similar results about temperature records. McBride et al. (2022) studied the record-breaking frequencies of the highest daily maximum temperature in South Africa for the 1951–2019 period. As in this work, they found that the number of records is higher than the theoretically expected in a stationary climate. They also found that mainland stations

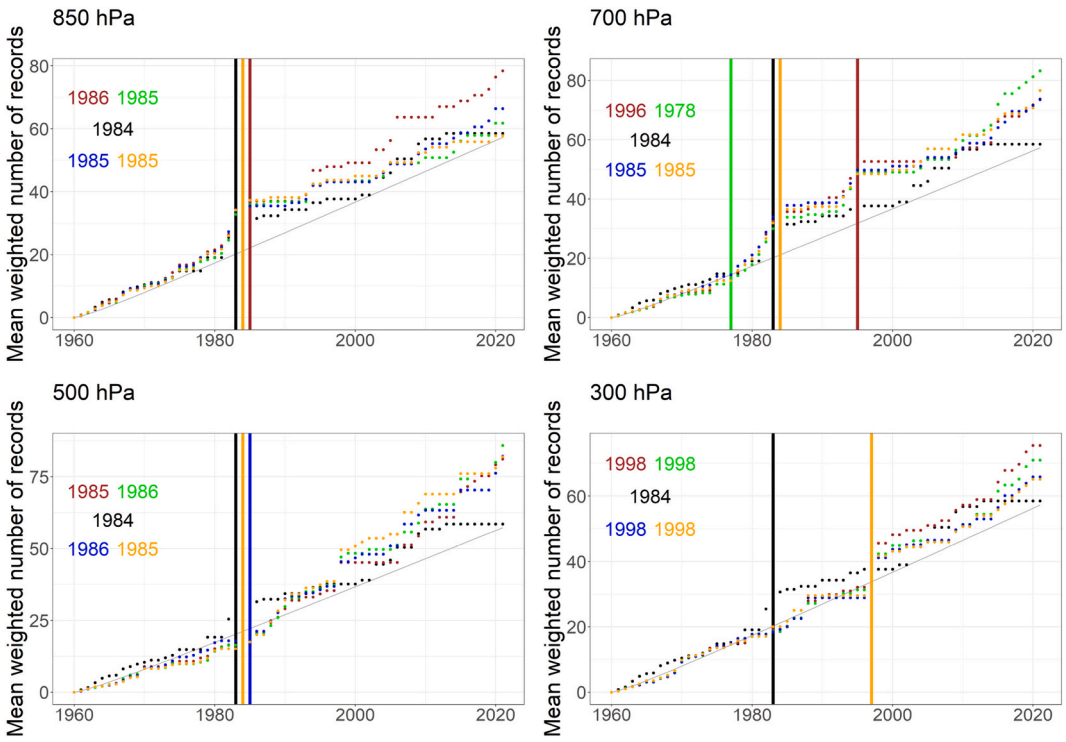

Fig. 10. Mean weighted number of records up to time t in July versus t, in T_x in Barcelona-Fabra (black) and in air temperature series in the four grid-points around Barcelona (NW: brown, NE: green, SW: blue and SE: orange) at 850, 700, 500 and 300 hPa. The years where the estimated change points occur in each of the four grid-points are shown, and plotted as vertical lines.

are more affected by the increasing number of records than stations close to the coast.

Our results yield the following conclusions about the evolution of extreme temperatures over Spain:

- Significant evidences of the effect of an increasing trend in the occurrence of upper extreme and record-breaking events in daily maximum temperature have been found in most of Spain, when the whole year is studied. Only in the North and the Cantabrian Coast areas the effects are not significant.
- The effects are heterogeneous within the year. Autumn is the season where the effects are weaker; only on the coast of Catalonia, the evidences are significant. On the other hand, summer is the season where the effects are stronger: in all the locations, except 5 locations in the North coast area, significant effects are found.
- Concerning the spatial variability, the Mediterranean region is the area where the number of months with strong evidences is higher, while the North Atlantic region is the area where less evidences of non-stationarity occur.
- No evidence of an increase of variability in daily maximum temperature has been found.

Many of these results are in agreement with those found in other areas. In Romania, Busuioc et al. (2015) detected significant increasing trends for the temperature extremes in all seasons, except for autumn; they found the highest increasing rate in summer, and that it is possibly associated with the the Atlantic multidecadal oscillation. Tošić et al. (2023) also found that, in Serbia, there are significant changes in temperature extremes consistent with warming, specially in summer, and that they have a highly positive correlation with the East Atlantic pattern. Sánchez-Benítez et al. (2020) summarized the connection between Iberian heat waves, and atmospheric circulation patterns in four weather regimes, and found that those heatwaves are associated with ridge conditions in western Europe.

The results from this work suggest that more analyses in a spatio-temporal framework about the occurrence of records and extreme temperature events in Spain are needed, and they can open new research lines. In particular, more research to study the relation between records and the atmospheric situation expressed by air temperature at different geopotential levels is of interest, and also with teleconnection patterns to find plausible physical mechanisms. It is also of interest to develop space–time models for the occurrence of records. These models should include long-term trends, space–time dependence in the tails and random effects, so that Bayesian models can provide an adequate framework. Models including atmospheric variables as covariates would be also useful as statistical downscaling models, and they could be used to obtain future projections of the behavior of records under different climate scenarios.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

To enhance reproducibility, the methodological tools used in this paper are provided in the `R` package `RecordTest` (Castillo-Mateo et al., 2023). Specific data and code scripts in `R` are available upon request to the authors.

Acknowledgement

This work has been supported in part by the Grants PID2020-116873 GB-I00 and TED2021-130702B-I00 funded by MCIN/ AEI/10.13039/ 501100011033 and Unión Europea NextGenerationEU; and the Research Group E46_20R: Modelos Estocásticos funded by Gobierno de

Aragón. Jorge Castillo-Mateo has been supported by the Doctoral Scholarship ORDEN CUS/581/2020 funded by Gobierno de Aragón. The authors thank the ECA&D project for providing the observational data, and the C3S for providing the Hersbach et al. (2023) data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.atmosres.2023.106934.

References

Anderson, A., Kostinski, A., 2011. Evolution and Distribution of Record-Breaking High and Low Monthly Mean Temperatures. J. Appl. Meteorol. Climatol. 50 (9), 1859–1871. https://doi.org/10.1175/JAMC-D-10-05025.1.

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1998. Records, Wiley Series in Probability and Statistics. John Wiley & Sons, New York. https://doi.org/10.1002/9781118150412.

Benestad, R.E., 2003. How Often Can We Expect a Record Event? Clim. Res. 25 (1), 3–13. https://doi.org/10.3354/cr025003.

Benestad, R.E., 2004. Record-Values, Nonstationarity Tests and Extreme Value Distributions. Global Planet. Change 44 (1–4), 11–26. https://doi.org/10.1016/j.gloplacha.2004.06.002.

Benestad, R.E., 2008. A simple test for changes in statistical distributions. Eos 89 (41), 389–390. https://doi.org/10.1029/2008EO410002.

Benjamini, Y., Yekutieli, D., 2005. False discovery rate–adjusted multiple confidence intervals for selected parameters. J. Am. Stat. Assoc. 100 (469), 71–81. https://doi.org/10.1198/016214504000001907.

Bunge, J., Goldie, C.M., 2001. Record sequences and their applications. Handb. Stat. 19, 277–308. https://doi.org/10.1016/S0169-7161(01)19012-7.

Busuioc, A., Dobrinescu, A., Birsan, M.-V., Dumitrescu, A., Orzan, A., 2015. Spatial and temporal variability of climate extremes in Romania and associated large-scale mechanisms. Int. J. Climatol. 35 (7), 1278–1300. https://doi.org/10.1002/joc.4054.

Castillo-Mateo, J., 2022. Distribution-free changepoint detection tests based on the breaking of records. Environ. Ecol. Stat. 29 (3), 655–676. https://doi.org/10.1007/s10651-022-00539-2.

Castillo-Mateo, J., 2023. RecordTest: Inference Tools in Time Series Based on Record Statistics, URL:https://CRAN.R-project.org/package=RecordTest, R package version 2.2.0.

Castillo-Mateo, J., Cebrián, A.C., Asín, J., 2023. RecordTest: An R package to analyse non-stationarity in the extremes based on record-breaking events. J. Stat. Softw. 106 (5), 1–28. https://doi.org/10.18637/jss.v106.i05.

Cebrián, A.C., Castillo-Mateo, J., Asín, J., 2022. Record tests to detect non-stationarity in the tails with an application to climate change. Stoch. Env. Res. Risk Assess. 36, 313–330. https://doi.org/10.1007/s00477-021-02122-w.

Chazarra-Bernabé, A., Lorenzo Mariño, B., Romero Fresneda, R., Moreno García, J.V., 2022. Evolución de los climas de Köppen en España en el periodo 1951–2020, Notas técnicas de AEMET 37, doi: 10.31978/666-22-011-4.

Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P.-A., Samsó, M., 2022. The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. Earth Syst. Dyn. 13 (1), 321–340. https://doi.org/10.5194/esd-13-321-2022.

Coumou, D., Rahmstorf, S., 2012. A Decade of Weather Extremes. Nat. Clim. Change 2, 491–496. https://doi.org/10.1038/nclimate1452.

Coumou, D., Robinson, A., Rahmstorf, S., 2013. Global Increase in Record-Breaking Monthly-Mean Temperatures. Clim. Change 118 (3–4), 771–782. https://doi.org/10.1007/s10584-012-0668-1.

Diersen, J., Trenkler, G., 1996. Records Tests for Trend in Location. Statistics 28 (1), 1–12. https://doi.org/10.1080/02331889708802543.

Diffenbaugh, N.S., Giorgi, F., 2012. Climate change hotspots in the CMIP5 global climate model ensemble. Clim. Change 114, 813–822. https://doi.org/10.1007/s10584-012-0570-x.

Feller, W., 1991. An Introduction to Probability Theory and its Applications, second ed., vol. 2. John Wiley & Sons, New York.

Fischer, E.M., Sippel, S., Knutti, R., 2021. Increasing probability of record-shattering climate extremes. Nat. Clim. Change 11 (8), 689–695. https://doi.org/10.1038/s41558-021-01092-9.

Foster, F.G., Stuart, A., 1954. Distribution-Free Tests in Time-Series Based on the Breaking of Records. J. R. Stat. Soc. B 16 (1), 1–22.

Gouet, R., Lafuente, M., López, F.J., Sanz, G., 2020. Exact and asymptotic properties of δ) records in the linear drift model. J. Stat. Mech: Theory Exp. 2020 (10), 103201 https://doi.org/10.1088/1742-5468/abb4dc.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2023. ERA5 hourly data on pressure levels from 1940 to present, Tech. Rep. Copernicus Climate Change Service (C3S), Climate Data Store (CDS). https://doi.org/10.24381/cds.bd0915c6.

Kendall, M., Gibbons, J.D., 1990. Rank Correlation Methods, A Charles Griffin Title, fifth ed. Oxford University Press, New York.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V., Petrovic, P., 2002. Daily Dataset of 20th-Century Surface Air Temperature and Precipitation Series for the European Climate Assessment. Int. J. Climatol. 22 (12), 1441–1453. https://doi.org/10.1002/joc.773.

Kysely, J., 2010. Recent severe heat waves in central Europe: how to view them in a long-term prospect? Int. J. Climatol. 30 (1), 89–109. https://doi.org/10.1002/joc.1874.

Lionello, P., Scarascia, L., 2018. The relation between climate change in the Mediterranean region and global warming. Reg. Environ. Change 18, 1481–1493. https://doi.org/10.1007/s10113-018-1290-1.

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13 (3), 245–259. https://doi.org/10.2307/1907187.

McBride, C.M., Kruger, A.C., Dyson, L., 2022. Trends in probabilities of temperature records in the non-stationary climate of South Africa. Int. J. Climatol. 42 (3), 1692–1705. https://doi.org/10.1002/joc.7329.

MedECC, 2020. Climate and Environmental Change in the Mediterranean Basin – Current Situation and Risks for the Future. First Mediterranean assessment report, Tech. Rep. Union for the Mediterranean, Plan Bleu, UNEP/MAP, Marseille, France. https://doi.org/10.5281/zenodo.4768833.

Meehl, G.A., Tebaldi, C., Walton, G., Easterling, D., McDaniel, L., 2009. Relative increase of record high maximum temperatures compared to record low minimum temperatures in the US. Geophys. Res. Lett. 36 (23) https://doi.org/10.1029/2009GL040736.

Om, K.-C., Ren, G., Kim, K.-H., Pak, Y.-I., Jong, S.-I., Kil, H.-N., 2022. Observed trends in extreme temperature events over northern part of the Korean Peninsula during 1960–2019 and a comparative overview. Atmos. Res. 270, 106061 https://doi.org/10.1016/j.atmosres.2022.106061.

Peña-Angulo, D., Gonzalez-Hidalgo, J.C., Sandonís, L., Beguería, S., Tomas-Burguera, M., López-Bustins, J.A., Lemus-Canovas, M., Martin-Vide, J., 2021. Seasonal temperature trends on the Spanish mainland: A secular study (1916–2015). Int. J. Climatol. 41 (5), 3071–3084. https://doi.org/10.1002/joc.7006.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL:https://www.R-project.org/.

Redner, S., Petersen, M.R., 2006. Role of global warming on the statistics of record-breaking temperatures. Phys. Rev. E 74 (6), 061114. https://doi.org/10.1103/PhysRevE.74.061114.

Saddique, N., Khaliq, A., Bernhofer, C., 2020. Trends in temperature and precipitation extremes in historical (1961–1990) and projected. Stoch. Env. Res. Risk Assess. 34 (10), 1441–1455. https://doi.org/10.1007/s00477-020-01829-6.

Salameh, A.A., Gámiz-Fortis, S.R., Castro-Díez, Y., Abu Hammad, A., Esteban-Parra, M.J., 2019. Spatio-temporal analysis for extreme temperature indices over the Levant region. Int. J. Climatol. 39 (15), 5556–5582. https://doi.org/10.1002/joc.6171.

Sánchez-Benítez, A., Barriopedro, D., García-Herrera, R., 2020. Tracking Iberian heatwaves from a new perspective. Weather Clim. Extrem. 28, 100238 https://doi.org/10.1016/j.wace.2019.100238.

Sánchez-Lugo, A., Berrisford, P., Morice, C., Nicolas, J.P., 2019. Global surface temperature [in "State of the climate in 2018"]. Bull. Amer. Meteor. Soc. 100 (9), 11–14. https://doi.org/10.1175/2019BAMSStateoftheClimate.1.

Schär, C., Vidale, P.L., Lüthi, D., Frei, C., Häberli, C., Liniger, M.A., Appenzeller, C., 2004. The role of increasing temperature variability in European summer heatwaves. Nature 427 (6972), 332–336. https://doi.org/10.1038/nature02300.

Serrano-Notivoli, R., Lemus-Canovas, M., Barrao, S., Sarricolea, P., Meseguer-Ruiz, O., Tejedor, E., 2022. Heat and cold waves in mainland Spain: Origins, characteristics, and trends. Weather Clim. Extrem. 37, 100471 https://doi.org/10.1016/j.wace.2022.100471.

Sousa, P.M., Barriopedro, D., Ramos, A.M., García-Herrera, R., Espírito-Santo, F., Trigo, R.M., 2019. Saharan air intrusions as a relevant mechanism for Iberian heatwaves: The record breaking events of August 2018 and June 2019. Weather Clim. Extrem. 26, 100224 https://doi.org/10.1016/j.wace.2019.100224.

Tan, J., Zheng, Y., Song, G., Kalkstein, L.S., Kalkstein, A.J., Tang, X., 2007. Heat wave impacts on mortality in Shanghai, 1998 and 2003. Int. J. Biometeorol. 51, 193–200. https://doi.org/10.1007/s00484-006-0058-3.

Tošić, I., Tošić, M., Lazić, I., Aleksandrov, N., Putniković, S., Djurdjević, V., 2023. Spatio-temporal changes in the mean and extreme temperature indices for Serbia. Int. J. Climatol. 43 (5), 2391–2410. https://doi.org/10.1002/joc.7981.

Tuel, A., Eltahir, E.A., 2020. Why is the Mediterranean a climate change hot spot? J. Clim. 33 (14), 5829–5843. https://doi.org/10.1175/JCLI-D-19-0910.1.

van der Wiel, K., Bintanja, R., 2021. Contribution of climatic changes in mean and variability to monthly temperature and precipitation extremes. Commun. Earth Environ. 2 (1), 1–11. https://doi.org/10.1038/s43247-020-00077-4.

Welch, W.J., 1990. Construction of permutation tests. J. Am. Stat. Assoc. 85 (411), 693–698. https://doi.org/10.2307/2290004.

Wergen, G., Hense, A., Krug, J., 2014. Record Occurrence and Record Values in Daily and Monthly Temperatures. Clim. Dyn. 42 (5), 1275–1289. https://doi.org/10.1007/s00382-013-1693-0.

Xu, K., Wu, C., 2019. Projected changes of temperature extremes over nine major basins in China based on the CMIP5 multimodel ensembles. Stoch. Env. Res. Risk Assess. 33 (1), 321–339. https://doi.org/10.1007/s00477-018-1569-2.

Xu, P., Wang, L., Huang, P., Chen, W., 2021. Disentangling dynamical and thermodynamical contributions to the record-breaking heatwave over Central Europe in June 2019. Atmos. Res. 252, 105446 https://doi.org/10.1016/j.atmosres.2020.105446.

Yosef, Y., Aguilar, E., Alpert, P., 2021. Is it possible to fit extreme climate change indices together seamlessly in the era of accelerated warming? Int. J. Climatol. 41, E952–E963.

Yu, Y., You, Q., Zuo, Z., Zhang, Y., Cai, Z., Li, W., Jiang, Z., Ullah, S., Tang, X., Zhang, R., et al., 2023. Compound climate extremes in China: Trends, causes, and projections. Atmos. Res. 286, 106675 https://doi.org/10.1016/j.atmosres.2023.106675.

Zhang, J., Ding, T., Gao, H., 2021. Record-breaking high temperature in Southern China in 2017 and influence from the middle-latitude trough over the East of Japan. Atmos. Res. 258, 105615 https://doi.org/10.1016/j.atmosres.2021.105615.

Zorita, E., Stocker, T., von Storch, H., 2008. How unusual is the recent series of warm years? Geophys. Res. Lett. 35 (24), L24706. https://doi.org/10.1029/2008GL036228.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R.M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M.D., et al., 2020. A typology of compound weather and climate events. Nat. Rev. Earth Environ. 1 (7), 333–347. https://doi.org/10.1038/s43017-020-0060-z.