Stochastic Environmental Research and Risk Assessment manuscript No.
(will be inserted by the editor)

Modelling and projecting the occurrence of bivariate extreme heat events using a nonhomogeneous common Poisson shock process

Abaurrea, J., Asín, J. and Cebrián A.C.

Received: date / Accepted: date

Abstract This work shows the modelling process of the occurrence of extreme heat events in daily maximum and minimum temperatures of Barcelona (Spain). The suggested model, a nonhomogeneous bivariate Poisson process, allows us to account for the mutual dependence between the two types of events, and to study their time evolution and relationship with temperature predictors. A wide kit of tools for validating the model, including an influence analysis, is provided.

This model is a useful downscaling tool to obtain medium-term local projections of the occurrence of extreme heat events in a climate change scenario, using the temperature trajectories from general circulation models. Due to the generalised global warming of climate change scenarios, it is necessary to check that the values used to project do not extrapolate the conditions of the data used to fit the model. Different solutions for dealing with this extrapolation problem are suggested.

The results of the fitted model for the temperature series of Barcelona in 1951-2005, and the future projections of the extreme heat events for the period 2031-2060, using three global circulation model trajectories under the SRES A1B scenario, are discussed.

Mathematics Subject Classification (2000) 60G99 · 62J12 · 62M20 · 62N99 · 62P12

University of Zaragoza
Dpto. Métodos Estadísticos. Calle Pedro Cerbuna, 12. Zaragoza 50009 (Spain)
Tel.: +34 976762885
E-mail: acebrian@unizar.es

University of Zaragoza
Dpto. Métodos Estadísticos. Calle Pedro Cerbuna, 12. Zaragoza 50009 (Spain)

1 Introduction

Heat waves are known to produce serious impacts on ecosystems, human health and economy; some examples are the heat waves of 1980 and 1995 in the United States and 2003 in Europe. Palecki et al (2001), Grintzevitch (2006) and other studies about the effect of extreme heat on mortality rates, found that both maximum and minimum temperatures have an impact on human health. Hence, an increasing number of heat wave definitions include information on both temperatures, see for example Plummer et al (1999), and Tryhorn and Risbey (2006). The U.S. National Weather Service issues an excessive heat watch, when apparent temperature in excess of 105^oF (41^oC) during the day, and 80^oF (27^oC) during the night are forecast for two consecutive days.

Climate extremes are a growing problem according to the results by Plummer et al (1999), or Mannshardt-Shamseldin et al (2010) who reported considerable increases in the frequency and intensity of precipitation and temperature extreme events during the 20^{th} century. In addition, Cubasch et al (2001), Meehl and Tebaldi (2004), Kharin and Zwiers (2005), Meehl et al (2005), Tryhorn and Risbey (2006), Abaurrea et al (2007) and other authors suggest that heat waves will become more frequent and more severe during the 21^{st} century. Given that more information about future climate behaviour at local and finer time scales is essential for preventing climate impacts, see Sivakumar and Christakos (2011), the interest of characterizing and projecting maximum and minimum temperature extreme events is clear.

This work suggests a bivariate model for describing the occurrence of extreme heat events in maximum and minimum daily temperatures. Instead of using two different univariate models, our approach takes advantage from the information provided by the dependence between both temperatures. The suggested model is a common Poisson shock process, which can be described by three independent Poisson processes. It allows us to characterize the occurrence over time of both types of extreme events, their mutual dependence and relationship with different predictors. The model is also a downscaling tool which can be applied to obtain projections of the occurrence of extreme heat events, at a daily scale, under a climate change scenario.

Section 2 describes the common Poisson shock process and the methodology for modelling the occurrence process of extreme heat events. Section 3 summarizes the results and the modelling process of the series of Barcelona (Spain). The use of the model to obtain projections for the period 2031-60, in the climate change scenario SRES A1B, is shown in Section 4. Section 5 summarizes the most relevant conclusions.

2 Methodology

2.1 Bivariate point processes for the occurrence of extremes

The excess over threshold approach is often used to analyse extreme events in climatology and other environmental sciences. In this approach, an observation whose value exceeds a reference threshold is considered extreme. A well-known result of extreme value theory states that, asymptotically with increasing thresholds, the extreme events occur according to a Poisson process (PP), see Coles (2001). When a constant occurrence rate is not a reasonable assumption, a nonhomogeneous Poisson process (NHPP) should be considered, see for example Ogata (1988) or Evin and Favre (2013). This is the case for the temperature extremes, where a seasonal behavior and a trend must be expected, see Abaurrea et al (2007). In this context, a model which generalizes the univariate NHPP to the bivariate case seems a reasonable option for the joint modelling of the occurrence of extreme events in maximum and minimum temperature. This model should be flexible enough to allow the inclusion of predictors, to model a nonhomogeneous behaviour and the dependence between the two components.

There are different types of bivariate point processes. Lindskog and McNeil (2003) suggested the common Poisson shock process, Pfeifer and Neslehová (2004) used a copula based model, and Bauerle and Grubel (2005) introduced the TaS (Thinning and Shifts) model, trying to avoid some limitations of the two previous approaches. Since this work focuses on the event occurrence times, we opt for the common Poisson shock process (CPSP).

The CPSP is a continuous time model, but in most applications the available data are discrete. However, when the time unit is short compared with the length of the recorded period and the occurrence rate low, the use of a continuous time model can be justified.

2.1.1 Common Poisson shock process

A CPSP, denoted by N, is a multivariate point process which assumes that there is an underlying PP of shocks N_0, that can yield d different types of events. The counting process of the j^{th} type of events, for $j = 1, \ldots, d$, is the component process N_j.

Let us define the vectors $\mathbf{I}_r = (I_{1,r}, \ldots, I_{d,r})'$ for $r = 1, \ldots, N_0(t)$, formed by the binary variables $I_{j,r}$, which indicate if an event of type j occurs at the r^{th} point of N_0. The CPSP assumes that $\mathbf{I}_r$ are i.i.d multivariate Bernoulli vectors. Hence, each shock corresponds to a new independent event, but the trigger variables for the different type of events may be dependent. Each component process can be expressed as a sum of independent binary variables,

$$N_j(t) = \sum_{r=1}^{N_0(t)} I_{j,r}.$$

Since each $N_j(t)$ results from an independent thinning of the PP N_0, they are PPs themselves. However, the total process, $N(t) = \sum_{j=1}^{d} N_j(t)$, is not a Poisson but a compound Poisson process, where the compounding distribution is the distribution of $\sum_{j=1}^{d} I_{j,r}$.

Decomposition of a CPSP. Herein, the model is restricted to the bivariate case $d = 2$, where events can be divided into 3 types, $(1,0), (0,1), (1,1)$, depending on the component where they have been observed. Given two component processes N_1 and N_2, see Figure 1, the total process can be decomposed into three marginal processes which are independent: $N_{(1)}$, $N_{(2)}$ and $N_{(12)}$ which include the events occurring only in the N_1 process, only in N_2, and the events occurring simultaneously, respectively. Their intensities are denoted $\lambda_{(1)}, \lambda_{(2)}$ and $\lambda_{(12)}$. The distribution of the CPSP is totally specified by the three marginal processes and $N_1 = N_{(1)} + N_{(12)}$ and $N_2 = N_{(2)} + N_{(12)}$ with intensities $\lambda_1 = \lambda_{(1)} + \lambda_{(12)}$ and $\lambda_2 = \lambda_{(2)} + \lambda_{(12)}$. A basic assumption of this model is that the dependence between N_1 and N_2 occurs only via the simultaneity of the events, so that independent component processes are obtained when $\lambda_{(12)} = 0$. An equivalent way to specify the model is by means of a single PP N_0, with intensity $\lambda = \lambda_{(1)} + \lambda_{(2)} + \lambda_{(12)}$, and an independent marking process of its points with probabilities $p_{(1)} = \lambda_{(1)}/\lambda$, $p_{(2)} = \lambda_{(2)}/\lambda$ and $p_{(12)} = \lambda_{(12)}/\lambda$.

This decomposition into three independent PPs allows the model to be readily applied for data modelling. It can be easily generalized to the nonhomogeneous case, by permitting the marginal intensities to be a function of a vector of time-varying predictors $\mathbf{x}(t)$ and using a logarithmic link to guarantee that the intensities are positive, $\lambda(t|\mathbf{x}(t)) = \exp(\beta'\mathbf{x}(t))$. The predictors also help to model the dependence induced by the systematic part of the three intensities and consequently, in the nonhomogeneous case, the assumption of independence between the marginal processes becomes conditional independence given the predictor values.

2.2 Estimating the model

The estimation of a nonhomogeneous CPSP (NHCPSP) reduces to the estimation of the three independent marginal NHPPs, which can be carried out by maximum likelihood. The loglikelihood of a NHPP, see Coles (2001), is given by

$$LL(t_1, \ldots, t_n; \beta) = -\int_0^T \lambda(t; \beta)dt + \sum_{i=1}^{n} \log \lambda(t_1, \ldots, t_n; \beta),$$

where t_i are the observed occurrence times, T the length of the observed period and $\lambda(t; \beta) = \exp(\beta'\mathbf{x}(t))$ the intensity given the predictors. The integral $\int_0^T \lambda(t; \beta)dt$ is easily calculated since, given the usual information available, the predictors are considered constant between times t and $t + 1$. Covariate selection is based on likelihood ratio tests with a significance level $\alpha = 0.05$.

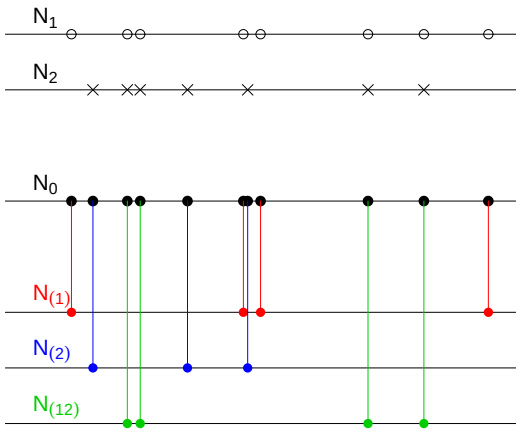

Fig. 1 Decomposition of the CPSP with component processes N_1 and N_2 into the three independent marginal PPs, $N_{(1)}$, $N_{(2)}$ and $N_{(12)}$.

The AIC criterion, as the likelihood function, can be used for the predictor selection but none of them is useful for comparing models with a different response, such as the three marginal processes of a CPSP, or models for different locations. As a performance measure, we suggest the square correlation coefficient between the empirical and the cumulative fitted intensities $\int_{l_1}^{l_2} \hat{\lambda}(t;\beta)dt$, calculated in periods of adequate length. This measure is analogous to the R^2 coefficient in multiple regression.

2.3 Validation analysis

The assumptions to be checked are the NHPP behavior of the three marginal processes and their mutual independence. The validation of the marginal processes, including an influence analysis to study the robustness of the model, is described in the next section. The independence assumption is checked using the test by Abaurrea et al (2013).

2.3.1 Checking the Poisson character of the marginal processes

The usual approach to check a univariate NHPP with intensity $\lambda(t)$ is to test if the process resulting from the time transformation $t^* = \int_0^t \lambda(u)du$ is a unit rate homogeneous Poisson process (HPP). Then, the validation analysis is standard and it is based on the study of the uncorrelation (using for example the Pearson's test) and the exponentiality of the inter-event distances $d_i^* = t_i^* - t_{i-1}^*$ or, equivalently, the uniform character of $\exp(-d_i^*)$, using the

Kolmogorov-Smirnov (KS) test. An example of the application of these validation techniques can be found in Abaurrea et al (2007).

A complementary approach to validate a NHPP, N, is the analysis of raw residuals, see Baddeley et al (2005), defined as,

$$r(t_1, t_2) = R(t_2) - R(t_1) = N(t_2) - N(t_1) - \int_{t_1}^{t_2} \hat{\lambda}(u)du$$

where $R(t) = N(t) - \int_0^t \hat{\lambda}(u)du$. These residuals are useful to check the intensity function since, under the model assumptions, $E[r(t_1, t_2)] \approx 0$. There are different types of raw residuals: scaled (such as the Pearson residuals), based on disjoint or overlapping intervals, etc. Residuals plots versus time are useful, but lurking plots, see Baddeley et al (2005), are more adequate to analyse the relationship with predictors which are non monotonous functions of time. The reason is that raw residuals are not linked to a point but to a time interval where the predictor may not be constant or monotonous. Approximate confidence envelopes can be calculated in both types of plots.

2.3.2 Identifying influential observations

In order to identify if one observation has an excessive influence on the fitted model, an influence plot based on the ratio likelihood test is suggested. An indicator variable equal to 1 for the observation analysed is defined. Then, a model including the interactions between this variable and the predictors of interest is fitted. A ratio likelihood test, which compares this model with the original one, is used to check if the effect on the response of the considered predictor values of this observation and the rest of the sample are significantly different. The p-values obtained for each observation are plotted, together with the 0.1 and the 0.05 values as a reference. It must be taken into account that multiple tests are being performed, and even if the null hypothesis is true, it must be expected that the test will reject H_0 in 5% of the cases. This tool allows us to analyse the influence of an observation on a particular predictor, on all the predictors in the model, or on a group of them. In the application in Section 3, for example, the effect on all the temperature predictors in the model is analysed. This approach can also be applied to analyse not one, but a set of observations; in the application, the influence of all the observations in each year is analysed jointly.

As a complementary tool, a β influence plot based on a leave-one-out cross-validation procedure is suggested. The model is fitted omitting each time one observation from the original sample, and the coefficients of the k^{th}-predictor resulting from each fit are plotted. If observation i is influential, the coefficient $\hat{\beta}_k^{(i)}$, fitted without observation i, will be very different from the other coefficients $\hat{\beta}_k^{(j)}$ with $j = 1, \ldots i-1, i+1, \ldots, n$. It is noteworthy that the influence graph and the β influence plots are not equivalent: the former allows the effect of one, or more predictors, to be different for the observation under study, while in the latter the effect of the observation is eliminated from the whole

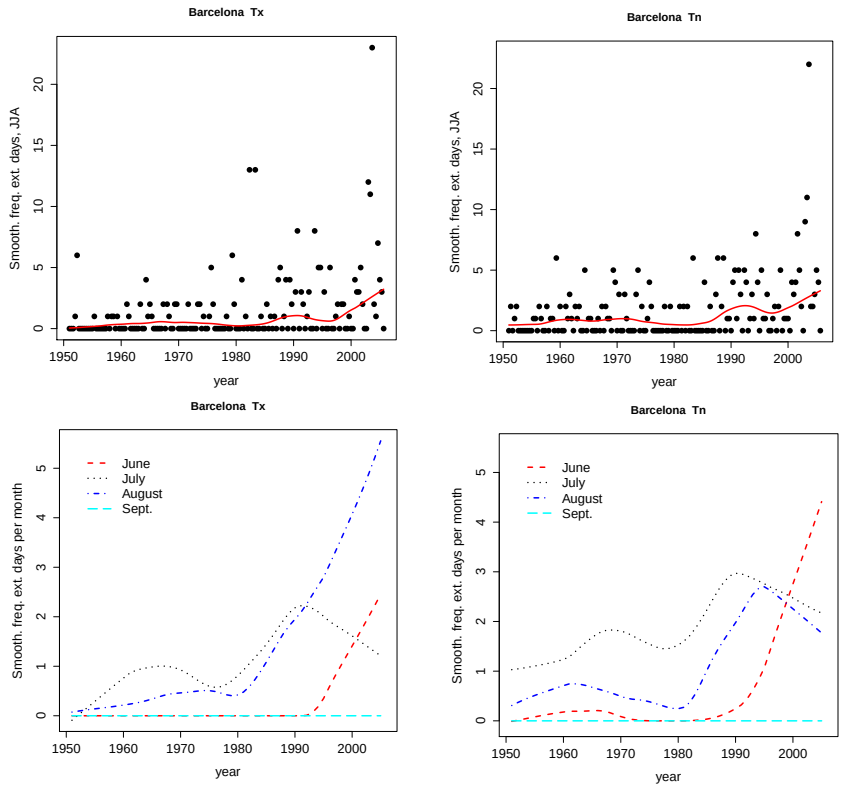

Fig. 2 Time and seasonal evolution of the frequency of summer extreme heat days

model. In addition, the likelihood test analyses the global effect on the model, while in the β influence plot the effect on each $\hat{\beta}_k$ is analysed separately. As the previous tool, this plot can be applied omitting a set of observations each time, instead of one.

3 Modeling Barcelona extreme heat events

3.1 Describing the data

The daily maximum and minimum temperature series of Barcelona for the period [1951, 2005], Tx and Tn herein, are used for illustration. Since in this region extreme heat events occur just in summer, only the observations from May to September (MJJAS) are considered. The exploratory analysis of temperature extreme behavior is applied to the values exceeding the 95^{th} percentile of the daily data from June to August, in the reference period 1971-2000. These percentiles are 31.8^oC for Tx and 22^oC for Tn.

non-isolated %	Tx				Tn			
Period	MJJAS	Jun.	Jul.	Aug.	MJJAS	Jun.	Jul.	Aug.
[1951, 1975]	75%	86%	65%	86%	46%	50%	49%	42%
[1976, 1994]	56%	86%	51%	56%	59%	50%	55%	65%
[1995, 2005]	86%	89%	85%	86%	80%	76%	71%	89%

Table 1 Percentage of non-isolated extremes in Tx and Tn for summer and for each month.

Time and seasonal evolution of the number of extreme heat days in Tx and Tn. The monthly frequency of extreme days for June, July and August, and the corresponding lowess with a 20% window are plotted in Figure 2 (top). The lowess (with 0 iterations and a 40% window) separated for each month (June, July, August and September) are displayed at the bottom plots; May is not shown since no extreme event was observed. The time evolution in the different months is not parallel. The most remarkable feature of the Tx evolution is the strong increase observed in August from 1980 onwards. June shows an increase from about 1990 in both Tx and Tn.

Tail dependence in Tx and Tn. A coefficient for summarizing the extremal dependence between two variables is defined in terms of their transformed uniform marginals U and V

$$\chi = \lim_{u \to 1} \chi(u)$$

with $\chi(u) = P(U > u | V > u)$, see Coles et al (1999). Barcelona shows a strong tail dependence between Tx and Tn since $\hat{\chi}(0.95) = 0.56$ and $\hat{\chi} = 0.4$, while under independence they should be 0.05 and 0, respectively.

In order to explore the simultaneity of Tx and Tn hot observations, the concept of non-isolated extremes is defined: a non-isolated extreme is an extreme of Tx (Tn) such that at least one of the observations of Tn (Tx) around it, is also extreme. Table 1 shows the percentage of non-isolated extremes with respect to the total number of extreme observations, for the whole summer and for each month, during three consecutive time intervals. The simultaneity rate is high but the complementary percentage is not negligible. There is an increase of the occurrence of non-isolated extremes in the period [1995, 2005], which is greater in Tn.

To sum up, the exploratory analysis shows a time evolution seasonally heterogeneous, and a time-varying dependence between the occurrence of both type of extremes. This result confirms that a bivariate model, such as the NHCPSP, is required for describing the occurrence and the dependence between the Tx and Tn extremes.

3.2 Definition of the extreme events

An extreme heat event (EHE) only in Tx is a run of consecutive days where Tx exceeds U_x but Tn does not exceed U_n, with U_x and U_n their respective extreme thresholds. An EHE only in Tn is defined analogously and a simultaneous EHE is a run where both Tx and Tn exceed their extreme thresholds. U_x and U_n are the 95^{th} percentiles used in the exploratory analysis. Tools for selecting an adequate threshold can be found in Coles (2001) and in Abaurrea et al (2007).

As mentioned before, the occurrence of the extreme events can be modelled by a point process, but it is necessary to assign an occurrence point to each event. In only Tx and Tn events, the occurrence point is located at the maximum temperature time, and in the simultaneous events at the time where the sum of both temperatures is maximum. In the Barcelona temperature series, there are 97 EHEs only in Tx, 114 only in Tn and 82 simultaneous EHEs.

3.3 Predictors

The fitted model will be used to obtain projections of the occurrence of EHEs in climate change scenarios, see Section 4.3, so we need predictors whose future 21^{st} century values can be obtained. The temperature projections from general circulation models, after being rescaled to fit the local characteristics, are reliable at an aggregated time scale but not at a daily one, see Wang et al (2012). For this reason, only aggregated temperature signals are considered as potential predictors.

- **Seasonal terms**. They are defined as the part of the annual harmonic signals corresponding to the five summer months, that is, $\cos(2k\pi i_t)$ and $\sin(2k\pi i_t)$ with i_t =121/365,..., 273/365, the time index from the 1^{st} of May to the 30^{th} of September and $k = 1, 2, ...$
- **Temperature terms**. A clear influential factor on EHE occurrence is the current temperature level. The moving average of Tx and Tn in 15 or 31 day intervals around t, Tx_{m15}, Tn_{m15}, Tx_{m31} and Tn_{m31}, are used to define short term temperature predictors. The decadal temperature trend is another variable to be considered, but this type of predictors was discarded since the model could not be used to project due to extrapolation problems, see Section 4.2.
- **Interaction terms**. Interaction terms between the harmonics and the temperature predictors are considered to allow different temperature effects over the summer.

3.4 Fitting process

The result of the initial predictor selection process for $N_{(1)}$ is,

$$\log(\hat{\lambda}_{(1)}(t)) = -21.4 - 2.0\cos(2\pi i_t) - 0.34\sin(2\pi i_t) + 0.07Tx_{m15} - 0.02Tn_{m31}.$$

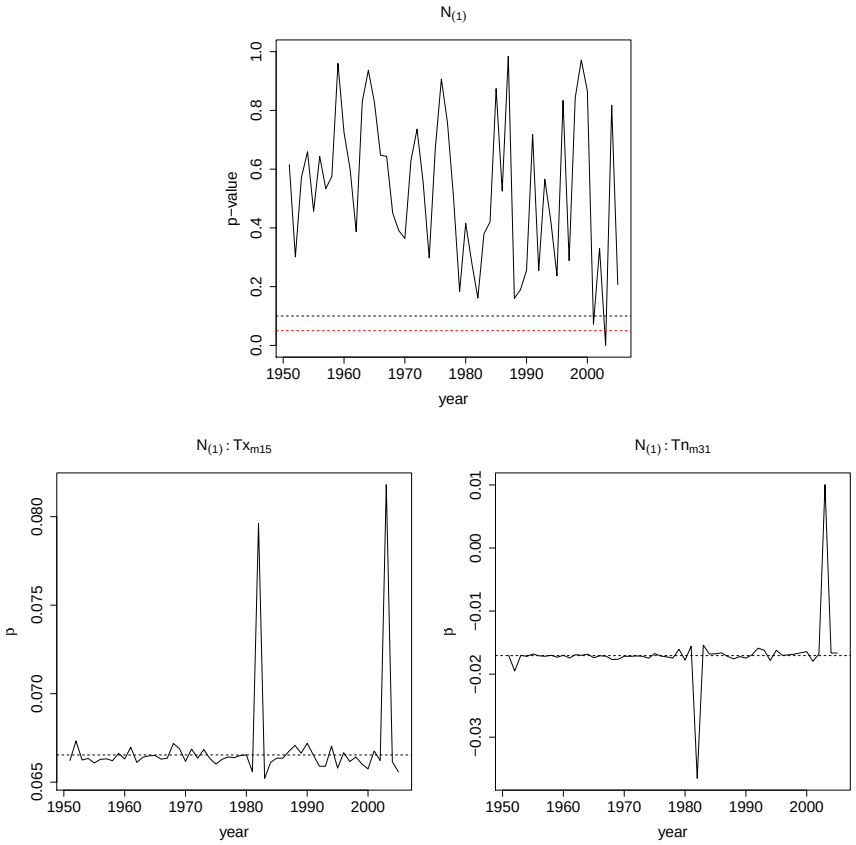

Fig. 3 Plots of the influence analysis of the first fitted model for $N_{(1)}$, Barcelona

The residuals of this model pass the validation analysis, with a Pearson autocorrelation p-value of 0.83 and a KS p-value of 0.17. An influence analysis is performed for the observations in each year. The influence graph, the first plot in Figure 3, shows that the year 2003 has a significantly different temperature effect with a p-value lower than 0.001. This yields a high residual and a high influence on the fitted coefficients, as it shows the β influence plots of the temperature predictors, Figure 3 (bottom). Also the year 1981 has an important influence on the fitted coefficients.

To get a more robust model, a new predictor selection process, adding the quadratic terms of the temperature variables as possible predictors, was carried out. It was observed empirically that these terms help to smooth the influence of the observations. The estimated coefficients of the intercept and the significant predictors are shown in the first row of Table 2. Both Tx_{m15} terms are significant and the p-value of Tn_{m31} is 0.07; although this predictor is not significant at a 0.05 level, it is not removed since the results of the

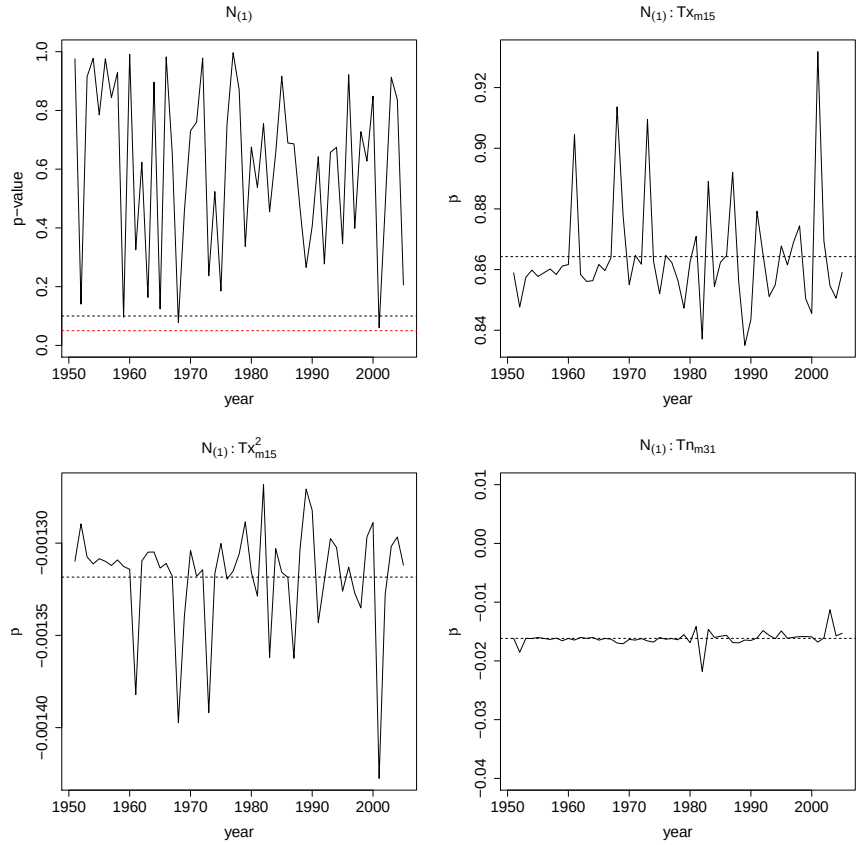

Fig. 4 Influence plots of the final model for $N_{(1)}$, Barcelona

Marginal processes	$\hat{\beta}_0$	$\cos(2\pi i_t)$	$\sin(2\pi i_t)$	Tx_{m15}	Tx_{m31}	Tn_{m15}	Tn_{m31}
$N_{(1)}$	-138.6	1.2	0.69	0.86			-0.001
(quadratic terms)				-0.02			
$N_{(2)}$	-68.0	-0.48	0.55			0.63	-0.03
(quadratic terms)						-0.001	
$N_{(12)}$	-23.0	-0.26	-0.02	0.03		0.10	-0.06

Table 2 Fitted models for the three marginal processes in Barcelona

validation analysis are better, see the first row in Table 3. This model is more robust, as the influence analysis shows, see Figure 4. The lowest p-value in the 55 influence tests is 0.065, for the year 2001. The rest of the validation analysis is satisfactory, as shows the first row in Table 3 and the plot of the Pearson residuals versus time in Figure 5.

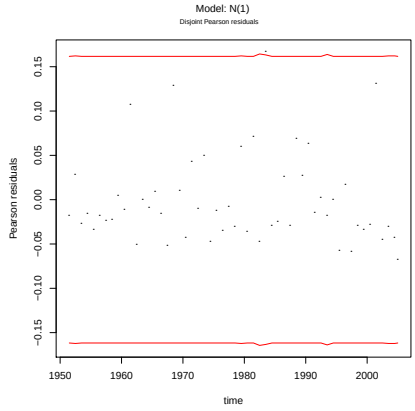

Fig. 5 Pearson residuals from the final model for $N_{(1)}$, Barcelona

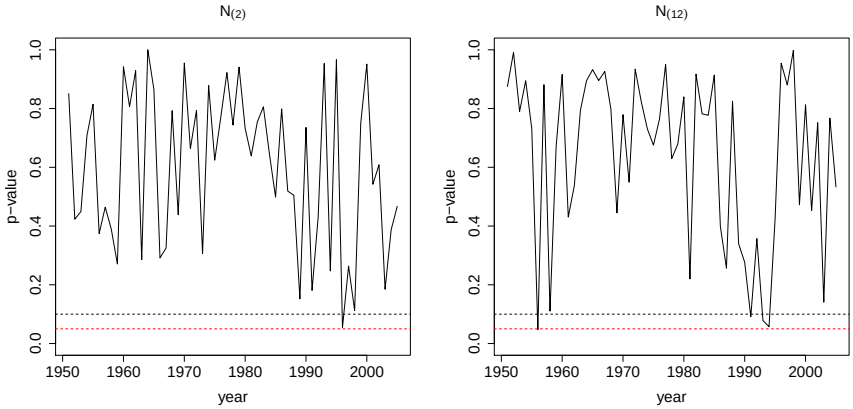

Fig. 6 Influence plots of the final models for $N_{(2)}$ and $N_{(12)}$, Barcelona

The final models for $N_{(2)}$ and $N_{(12)}$ resulting from similar modelling processes are shown in Table 2. The first harmonic is significant in all the models; Tx_{m15} is required in $N_{(1)}$ and $N_{(12)}$ (the two processes defined from Tx), while both Tn_{m15} and Tn_{m31} are included in the processes defined from Tn, $N_{(2)}$ and $N_{(12)}$. These final models satisfy the validation analysis, see Figure 6 and Table 3 where R^2 and the p-values of the Pearson correlation test and the KS test for checking the uniform distribution are summarized. The lowest p-values of the influence test are 0.06 (year 1996) in $N_{(2)}$ and 0.05 and 0.06 (years 1956 and 1994) in $N_{(12)}$.

The empirical and the cumulative fitted rates $\int_{t_1}^{t_2} \hat{\lambda}(u)du/(t_2 - t_1)$, calculated in moving 5-month long intervals, are plotted in Figure 7 for the three marginal processes. The observed evolution is satisfactorily reproduced by the

Process	R^2	KS pv	Pearson pv
$N_{(1)}$	0.75	0.39	0.28
$N_{(2)}$	0.46	0.40	0.97
$N_{(12)}$	0.73	0.60	0.62

Table 3 Summary of the goodness of fit and validation analysis for the marginal processes, Barcelona

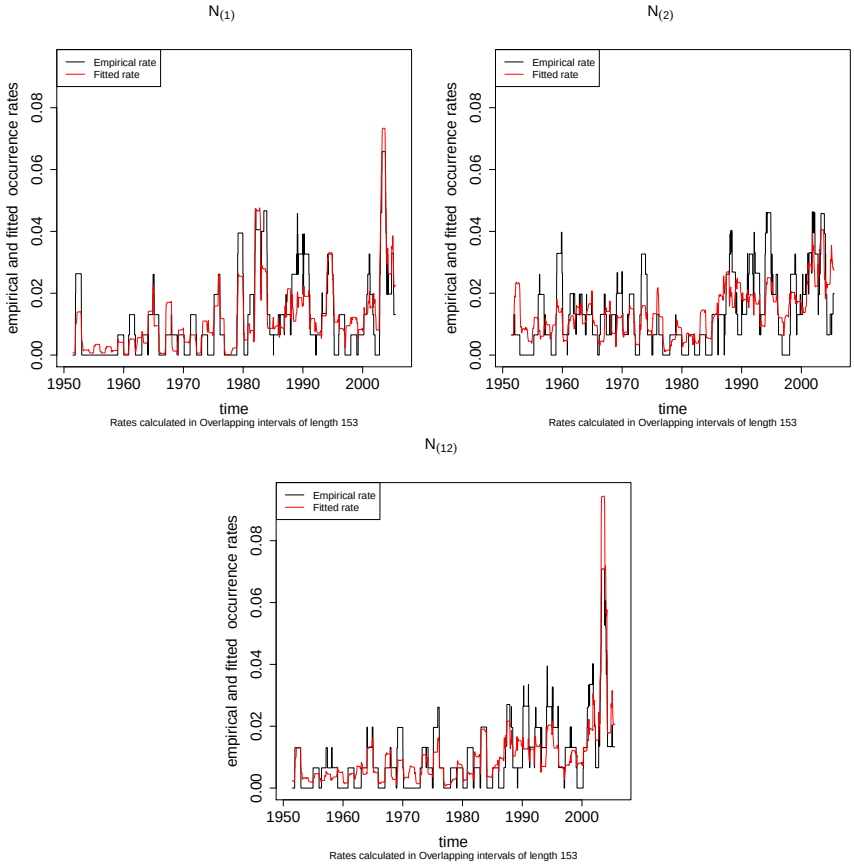

Fig. 7 Empirical and fitted intensities of the NHCPSP marginal processes, Barcelona.

fitted intensity, and a clear increase of the occurrence rate from about the middle 80's is observed in $N_{(2)}$ and $N_{(12)}$, and a bit earlier in $N_{(1)}$. This increase becomes steeper from 2000 onwards. The R^2 measures are around 0.75 for $N_{(1)}$ and $N_{(12)}$ and 0.46 for $N_{(2)}$.

The assumption of independence of the marginal processes is checked using the test by Abaurrea et al (2013) with $n = 1000$ simulations. The p-value is 0.936 and independence cannot be rejected.

	$\bar{\lambda}_{(1)}$	$\bar{\lambda}_{(2)}$	$\bar{\lambda}_{(12)}$	$\bar{\lambda}_1$	$\bar{\lambda}_2$	$\bar{\lambda}_{(12)}/\bar{\lambda}$
[1951, 1955]	0.004	0.011	0.005	0.009	0.016	0.249
[1981, 1985]	0.022	0.009	0.009	0.031	0.018	0.225
[2001, 2005]	0.029	0.029	0.081	0.110	0.110	0.582

Table 4 Mean fitted intensities in three 5-year periods.

Interpretation of the fitted models. The mean fitted intensities of the three marginal and the two component processes are calculated for three 5-year long periods located at the beginning, in the middle and at the end of the record, see Table 4. The occurrence rate is increasing in all the marginal processes, and $N_{(12)}$ shows, in relative terms, the highest increase. Its proportion with respect to the total intensity $\bar{\lambda} = \bar{\lambda}_{(1)} + \bar{\lambda}_{(2)} + \bar{\lambda}_{(12)}$ is summarized in the last column, and it has doubled in the last period. This is a relevant result, since the combination of extreme hot days and nights is more harmful for human health and for energy demand.

4 Projection of the occurrence of EHEs in a climate change scenario

General Circulation Models (GCMs) are the best tool to reproduce atmospheric variables at a monthly or seasonal scale over broad areas, see IPCC (2007) and Jeong et al (2012). However, GCMs cannot be directly applied to obtain variables at a daily or local scale, as it is often necessary for climate change impact studies. The interest of projections of extreme events at these scales is underlined by Wang et al (2012) and Casanueva et al (2013).

The model developed in the previous sections can be used as a downscaling tool for obtaining daily scale projections from wider time scale predictors. More precisely, in this section the fitted model is used to obtain medium-term projections for the EHE evolution in a climate change scenario, using predictors based on the temperature GCM trajectories, properly downscaled to fit the local characteristics.

4.1 GCM data

Initially, seven GCM daily trajectories were available from the WCRP CMIP3 Multi-Model Data webpage[1] and the ENSEMBLES distribution portal at the Max Planck Institute for Meteorology[2]. Four of them (CGCM3.1, MIROC3.2, ECHAM5 and HadGEM) are among the best five in reproducing the sea level pressure patterns in the North Atlantic-Europe region, from 23 GCMs of IPCC

[1] https://esg.llnl.gov:8443/home/publicHomePage.do
[2] http://www.mad.zmaw.de/projects-at-md/ensembles/

AR4, see van Ulden and van Oldenborgh (2006). According to Errasti et al (2011), three of them (MIROC3.2, ECHAM5 and HadGEM) are among the best five, from 24, in reproducing the seasonal cycle and the variability of the sea level pressure and the surface air temperature over the Iberian Peninsula. We use for illustration the period 2031-60 and the scenario A1B, which describes a future world of rapid economic growth, a global population that peaks in the mid-century and more efficient technologies.

To project the temperature in a small area, the GCM Tx and Tn trajectories must be downscaled to fit the local characteristics of the climate. We opted for the following bias correction approach,

$$y_{sca}(t) = \frac{y_{GCM}(t) - \hat{y}_{GCM,20c3M}(t)}{\hat{s}_{GCM,20c3M}(t)} \hat{s}_{obs}(t) + \hat{y}_{obs}(t)$$

where y_{GCM}(t) is the daily GCM trajectory to be scaled; $\hat{y}_{.}(t)$ is the fitted value in day t from a regression model with the daily temperature as response and the annual harmonic terms as predictors; $\hat{s}_{.}(t)$ is the squared root of the fitted value for day t from a regression model with response the square of the residuals obtained in the previous regression and the annual harmonic terms as predictors. The subscripts $_{obs}$ and $_{GCM,20c3M}$ indicate if the regression models have been fitted to the observed data in 1971-2000, or to the GCM trajectories for the 20c3M scenario. More details can be found in Abaurrea et al (2014).

A future GCM trajectory downscaled with the previous procedure is considered reliable, if the downscaled GCM trajectory for the 20c3M scenario reproduces satisfactorily the global distribution of the observed temperatures, according to a KS test. From the seven trajectories initially considered, only E5r1, E5r3 and E5r4, the three trajectories from ECHAM5, and the one from CGCM3 passed that checking analysis.

4.2 Extrapolation analysis

Statistical downscaling models require that the relationship between predictor and predictand is unchanged between the present and the future climate, see for example Fowler et al (2007) and Schmith (2008). Given that climate change scenarios show a general global warming trend and that the range of current and future values of some temperature-related variables is quite different, that assumption cannot be guaranteed even for medium-term projections. Sharif et al (2013) and Wang and Chen (2013) note that regression-type models can account for variability to an acceptable level but frequent extrapolation beyond the historical conditions may be unreliable. Hence, in order to get reliable results, the fitted models should not be used under severe extrapolation conditions.

We suggest different approaches to avoid this extrapolation problem. The first one is to eliminate from the set of possible predictors those with future values far outside the range used to fit the model. The second is to check each day of the trajectory in order to eliminate from the projection the days under

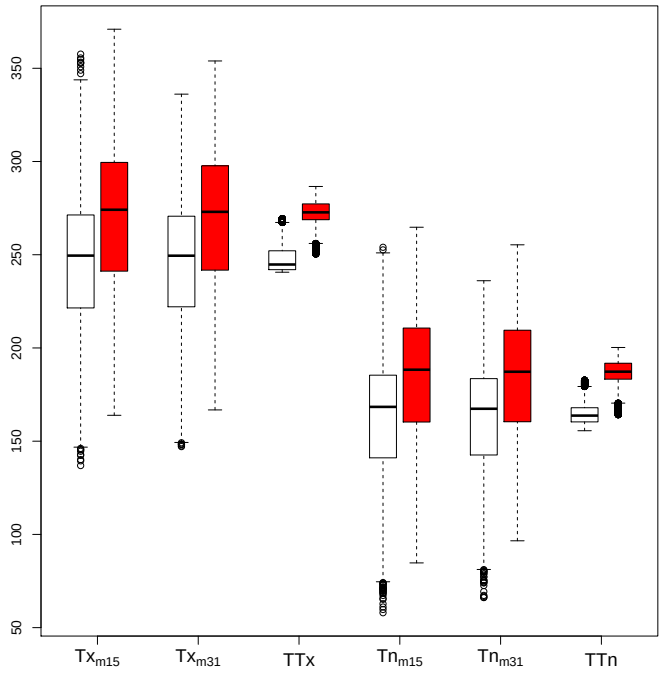

Fig. 8 Boxplots of short and long term temperature predictors (in Celsius degree tenths) using observed values (white boxplots) and scaled values from E5r1-A1B (coloured boxplots)

severe extrapolation. Finally, to reduce the effect of the observations which are very far from the center of the distribution, robust measures such as the trimmed mean should be used to summarize and to obtain conclusions, see Section 4.3.

Predictors with future values far outside its range in the fitting period. For illustrating this problem, the boxplots of some temperature predictors calculated with the downscaled trajectory E5r1 for the period 2031-60 and the observed series are shown in Figure 8. The high increase in the long term temperature predictors, TTx and TTn (defined as the lowess with a 30% window of Tx and Tn) would lead to extrapolation problems in most of the values and, consequently, they were eliminated from the set of potential predictors. As a general rule, the predictors whose future range have a low or no intersection with the observed one should not be included in the model. In this case, the variables with longer aggregation time periods show less variability, and the percentage of values leading to extrapolation is greater.

Trajectory	E5r1	E5r3	E5r4	CGCM3
	2.96	2.24	1.39	30.22

Table 5 Percentage of days under extrapolation in the downscaled trajectories, scenario A1B 2031-60.

Checking extrapolation conditions of the trajectories. A day of a trajectory should be projected only if its predictor values are not under extrapolation, so a preliminary analysis to check it must be carried out.

We apply the following procedure to check both marginal and multivariate extrapolation:

Marginal checking. The day t is not under extrapolation and is projected if its values in all the predictors are lower than their corresponding maxima in the fitting period.

Multivariate checking. The day t is projected if its Mahalanobis distance (with respect to the observed mean vector and covariance matrix) is lower than the maximum of the Mahalanobis distances of the observed values or, alternatively, if all the predictor values are lower than their 90^{th} percentiles in the observed period.

The left top plot in Figure 9 shows the bivariate distribution of the two predictors in the model for $N_{(1)}$, Tx_{m15} and Tn_{m31}, estimated from the observed data with a Gaussian density kernel based on 50 points. The points of the E5r1-A1B trajectory under extrapolation, according to the previous criteria, are plotted in the same graph. The other plots in Figure 9 show the marginal densities estimated by a kernel, the marginal values of the points under extrapolation, and the observed maximum and 90^{th} percentile (the reference limits of the extrapolation analysis). The election of these limits is heuristic and the plots help us to calibrate them, checking the location of the points eliminated due to extrapolation with respect to the observed distribution of the predictor.

The percentage of days under extrapolation in 2031-60 for the four downscaled trajectories is summarized in Table 5. If the percentage of days not projected due to extrapolation in a trajectory is high, the whole trajectory is eliminated from the study. This is the case of the CGCM3 trajectory, with more than 50% of the observations under extrapolation in August.

4.3 Projections for the period 2031-60

To obtain the projected intensity of the EHE occurrence in the period 2031-60, the downscaled GCM values of the predictors in that period are used as input of the model. Figure 10 shows the lowess of the three marginal intensities obtained from the models fitted to the observed data and to the three downscaled ECHAM5 trajectories. The characterization of the future distribution of the EHE occurrence is usually more interesting than the daily projections. To describe the future seasonal pattern, the quantiles of the intensities for

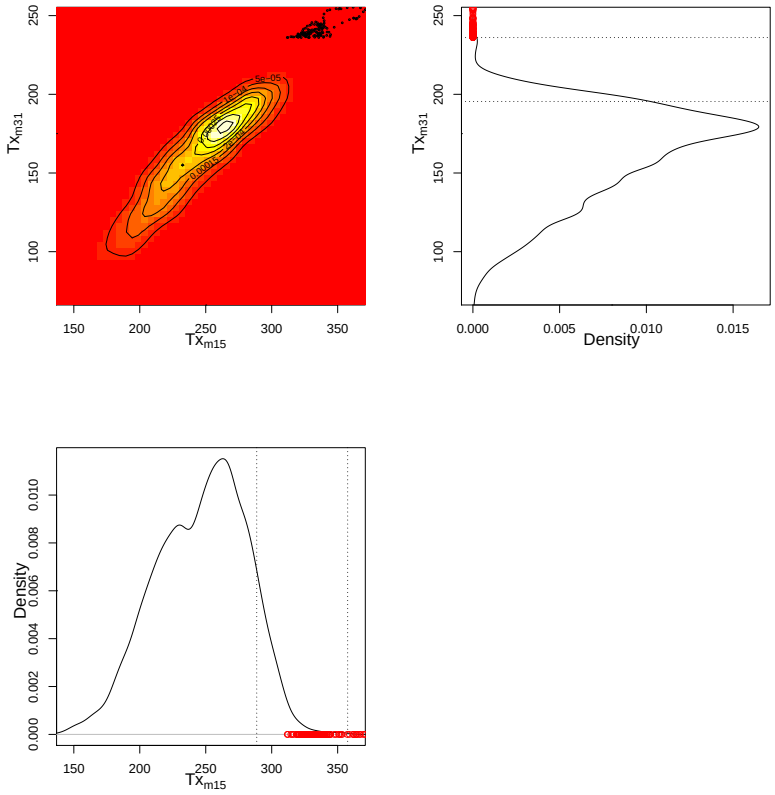

Fig. 9 Bivariate distribution of the two predictors in the model for $N_{(1)}$ estimated from the observed data. The points are the E5r1-A1B observations under extrapolation. Right top and bottom plots display the marginal distributions.

each day of the summer are estimated using quantile regression, see Koenker (2005). The response of these models is the fitted intensity and the predictors are the two first harmonics, which guarantee an adequate fit of the seasonal behaviour.

Some plots characterizing the future distribution of only Tx EHEs, based on the E5r1 trajectory, are shown in Figure 11. The estimated 25^{th}, 50^{th} and 75^{th} percentiles, q25, q50 and q75, of the intensities fitted to the observed data (herein denoted by simplicity observed percentiles) and to the downscaled trajectory (projected percentiles) are shown in the left plot. The percentile 90 (q90) is not calculated for the E5r1 trajectory due to the extrapolation level. The plot shows that the projected q25 is higher than the observed q50 in July and it almost reaches the observed q75 in the middle of August. All the projected q75 are higher than the observed q90 of the same Julian day. The

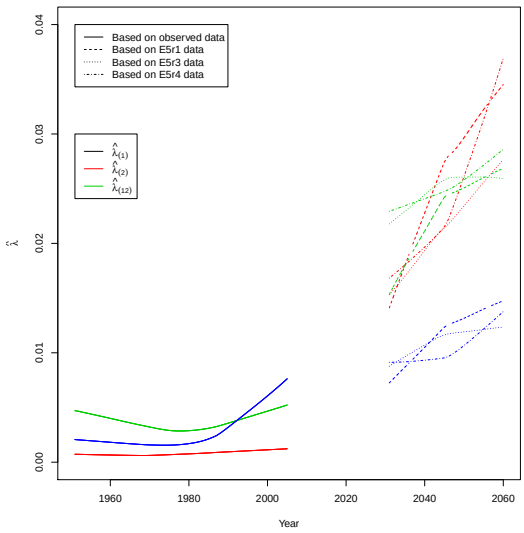

Fig. 10 Lowess of the fitted intensities for the three marginal processes using the observed data (1954-2005) and the three ECHAM5 downscaled trajectories, scenario A1B 2031-60.

right plot displays the estimated q50 of the intensities fitted to the observed data and to the projected values for the intervals, 2031-40, 2041-50 and 2051-60; a high increase is forecast from 2041 onwards.

To describe the behaviour of the projections based on the ECHAM5 trajectories, Table 6 summarizes the monthly 25% trimmed means of the intensities fitted to the observed data and to the 2031-60 trajectories under scenario A1B. The highest mean of the observed intensities is 0.023 and corresponds to $N_{(2)}$ in July. The means of the projected intensities of the three processes in July and August attain twice to four times this value. To make the comparisons easier, the projected/observed mean ratios are shown in brackets (except in May where the observed means are nearly 0). For each process and month, the ratios of the three GCM trajectories do not show relevant differences. The highest differences are found in June and September in $N_{(1)}$, due to the low observed values. Projected values in June and September are getting closer to the current means in July and August, while projected values for those months are between 2 and 9 times the observed ones. The highest ratios in July and August are found in $N_{(12)}$, and the increase in this process is so high that, although its observed means are the lowest of the three processes, its projected means are higher than those projected for $N_{(2)}$ and than most of the projected ones for $N_{(1)}$. In June and September the highest increase ratio is forecast in $N_{(1)}$ but, as in the observed case, the highest mean of the projected intensities corresponds to $N_{(2)}$.

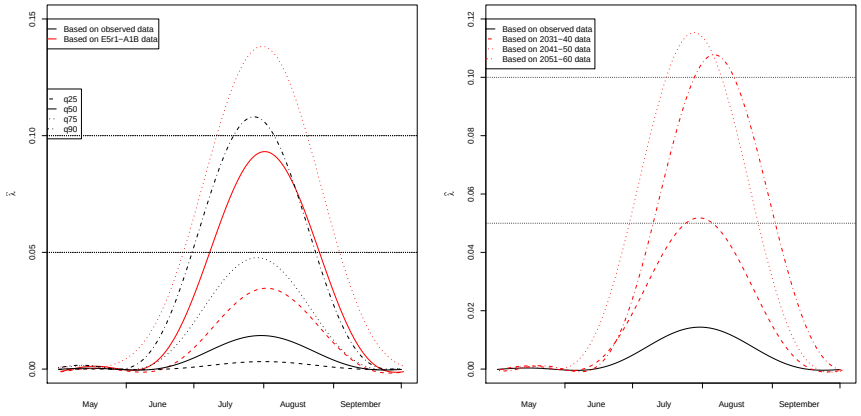

Fig. 11 Estimated percentiles of the intensities fitted to the observed data and to the E5r1 downscaled trajectory (left plot). Estimated q50 of the intensities fitted to the observed data and to three different time intervals of the projected period (right plot).

		May	June	July	August	September
$\hat{\lambda}_{(1)}$	E5r1	0.0000	0.0079 (79)	0.0703 (5)	0.0739 (6)	0.0019 (19)
	E5r3	0.0000	0.0045 (45)	0.0779 (5)	0.0669 (6)	0.0016 (16)
	E5r4	0.0000	0.0056 (56)	0.0960 (6)	0.0523 (4)	0.0013 (13)
	Observed	0.0000	0.0001	0.0155	0.0119	0.0001
$\hat{\lambda}_{(2)}$	E5r1	0.0000	0.0108 (18)	0.0525 (2)	0.0432 (2)	0.0089 (8)
	E5r3	0.0000	0.0096 (16)	0.0548 (2)	0.0448 (2)	0.0107 (10)
	E5r4	0.0000	0.0112 (19)	0.0658 (3)	0.0425 (2)	0.0060 (5)
	Observed	0.0000	0.0006	0.0234	0.0187	0.0011
$\hat{\lambda}_{(12)}$	E5r1	0.0002	0.0038 (8)	0.0704 (7)	0.0772 (8)	0.0037 (4)
	E5r3	0.0001	0.0033 (7)	0.0557 (6)	0.0872 (9)	0.0042 (5)
	E5r4	0.0001	0.0035 (7)	0.0827 (8)	0.0426 (5)	0.0029 (3)
	Observed	0.0000	0.0005	0.0099	0.0093	0.0009

Table 6 Monthly 25% trimmed means of the intensities fitted to the observed data and to the ECHAM5 downscaled trajectories under scenario A1B; in brackets, the projected/observed mean ratios.

Concerning the seasonal pattern, although the differences between months diminish, they do not disappear: the June/July ratio for the observed means is around 1/100, 1/30 and 1/20, depending on the marginal process, while for the projected values it varies between 1/18 and 1/5. A similar behaviour is shown by the September/August ratio.

Since the intensity of the first component process is $\lambda_1 = \lambda_{(1)} + \lambda_{(12)}$, the mean of the projected intensities for the occurrence of Tx EHEs in July and August is between 3 and 5 times the observed mean. Analogously, the projected intensity for the occurrence of Tn EHEs is between 3 and 4 times the mean of the observed λ_2.

5 Conclusions

Occurrence model of extreme heat events. A bivariate process, the non-homogeneous common Poisson shock process, is successfully applied to model the occurrence of EHEs in daily maximum and minimum temperatures, taking into account their mutual dependence. The NHCPSP can be readily applied to statistical modelling since it can be decomposed into three independent NHPPs: the process of the events only in Tx, $N_{(1)}$, the events only in Tn, $N_{(2)}$, and the simultaneous events, $N_{(12)}$. The model allows us to characterize the time evolution of the EHE occurrence and its relationship with different predictors. The dependence structure between the components is described via the simultaneous event process $N_{(12)}$, and also by conditioning the marginal intensities on adequate covariates.

The modelling process of the Barcelona temperature series is thoroughly described. In this case, only moving average temperatures and harmonic terms are used as predictors but, depending on the use of the model, other types of variables such as time trends or atmospheric variables could be considered. The application illustrates the importance of carrying out an exhaustive validation of the model, for which the necessary tools are suggested. In particular, two influence plots, one based on the likelihood ratio test and the other on a leave-one-out approach, are proposed.

Projections under climate change scenarios. The suggested model can be used to obtain medium-term regional projections of the occurrence of EHEs in climate change scenarios. It works as a downscaling tool which provides projections at a daily scale, using wider time scale temperature trajectories from the GCMs. This is important since climate projections at a daily and local scale are often essential for climate change impact studies.

As in all regression-type models, it is necessary to check that projections are not made under extrapolation conditions. Different solutions to avoid this problem are considered: the elimination of predictors which suffer an important range change under climate change scenarios, a marginal and multivariate extrapolation analysis of the GCM trajectories, and the calculation of robust measures, such as trimmed means, to summarize the results.

Barcelona results. The fitted model for Barcelona reveals an increase of the rate of the EHE occurrence from 1980 onwards, which is higher from the late 90's. There is evidence of an important increase of the intensity of the simultaneous EHEs, whose ratio with respect to the total intensity has doubled in the 2001-2005 period.

As regards the future behaviour under the A1B scenario, even though the most extreme temperature GCM trajectories are not projected due to extrapolation and conservative approaches have been applied, we obtain a clear increase of the mean intensity of the three marginal processes in 2031-60. The fastest increase is found in the intensity of the simultaneous events. Concerning the seasonal pattern, the projected values for June and September are getting closer to the current values for July and August, and the projected values for these months are between 2 and 9 times the observed ones.

Acknowledgements

The authors acknowledge the financial support from Ministerio de Ciencia e Innovación (Spanish Department of Science) and Ministerio de Medio Ambiente (Spanish Department of Environment) through the projects CGL2009-09646 and ESTCENA 2009/0017. They also thank AEMET, the Spanish meteorological agency, for the temperature data.

References

Abaurrea J, Asín J, Cebrián AC, Centelles A (2007) Modeling and forecasting extreme heat events in the central ebro valley, a continental-mediterranean area. Global and Planetary Change 57(1-2):43–58

Abaurrea J, Asín J, Cebrián AC (2013) A test of independence between non-homogeneous poisson processes and its application to heat wave modeling. submitted to Ecological and Environmental Statistics

Abaurrea J, Asín J, Cebrián AC (2014) Evaluating a simple statistical downscaling method for maximum and minimum daily temperature in a mountainous area. Preprint

Baddeley A, Turner R, Møller J, Hazelton J (2005) Residual analysis for spatial point processes. Journal of the Royal Statistical Society: Series B 67(5):617–66

Bauerle N, Grubel R (2005) Multivariate counting processes: copulas and beyond. ASTIN Bulletin 35(2):379–408

Casanueva A, Herrera S, Fernndez J, Fras MD, Gutirrez JM (2013) Evaluation and projection of daily temperature percentiles from statistical and dynamical downscaling methods. Natural Hazards and Earth System Sciences 13:2080–99

Coles S (2001) An introduction to statistical modeling of extreme values. Springer-Verlag

Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analysis. Extremes 2(4):339–65

Cubasch U, Meehl GA, Boer GJ, Stouffer M R Jand Dix, Noda A, Senior CA, Raper S, Yap KS (2001) Projections of future climate change. Climate Change 2001 The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change

Errasti I, Ezcurra A, Senz J, Ibarra-Berastegi G (2011) Validation of ipcc ar4 models over the iberian peninsula. Theor Appl Climatol 103:61–79

Evin G, Favre A (2013) Further developments of a transient poisson-cluster model for rainfall. Stoch Environ Res Risk Assess 27:831–47

Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int J Climatol 27:1547–78

Grintzevitch S (2006) Heat waves: their climatic and biometeorological nature in two north american regions. Ph D dissertation, University of Reading, School of Mathematics, Meteorology & Physics

IPCC (2007) In: Solomon (ed) Climate Change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, Cambridge University Press

Jeong D, St-Hilaire A, Ouarda T, Gachon P (2012) Comparison of transfer functions in statistical downscaling models for daily temperature and precipitation over canada. Stoch Environ Res Risk Assess 26:633

Kharin VV, Zwiers FW (2005) Estimating extremes in transient climate change simulations. J Clim 18:1156–73

Koenker R (2005) Quantile regression. Cambridge U. Press

Lindskog P, McNeil A (2003) Common poisson shock models: applications to insurance and risk modelling. ASTIN Bulletin 33:209–38

Mannshardt-Shamseldin EC, Smith RL, Sain SR, Mearns LO, Cooley D (2010) Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. Ann Appl Stat 4(1):484–502

Meehl G, Washington WM, Collins W, Arblaster J, Hu A, Buja L, Strand W, Teng H (2005) How much more global warming and sea level rise? Science 307:1769–72

Meehl GA, Tebaldi C (2004) More intense, more frequent and longer lasting heat waves in the 21st century. Science 305:994–7

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association 83(401):9–27

Palecki M, Changnon S, Kunkel K (2001) The nature and impacts of the july 1999 heat wave in the midwestern united states: learning from the lessons of 1995l. B Am Meteoroll Soc 82:1353–67

Pfeifer D, Neslehová J (2004) Modeling and generating dependent risk processes for irm and dfa. ASTIN Bulletin 34:333–360

Plummer N, Salinger M, Nicholls N, Suppiah R, Hennessy K, Leighton R, Trewin B, Page C, Lough J (1999) Changes in climate extremes over the australian region and new zealand during the twentieth century. Clim Chang 42:183–202

Schmith T (2008) Stationarity of regression relationships: Application to empirical downscaling. Journal of Climate 21(17):4529–37

Sharif M, Burn DH, Hofbauer KM (2013) Generation of daily and hourly weather variables for use in climate change vulnerability assessment. Water Resources Management 27(5):1533–50

Sivakumar B, Christakos G (2011) Climate: patterns, changes and impacts. Stoch Environ Res Risk Assess 25:443

Tryhorn L, Risbey J (2006) On the distribution of heatwaves over the australian region. Aust Meteorol Mag 55:169–82

van Ulden A, van Oldenborgh G (2006) Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in central europe. Atmos Chem Phys 6:863–81

Wang L, Chen W (2013) A cmip5 multimodel projection of future temperature, precipitation, and climatological drought in china. Int J Climatol

Wang X, Yang T, Shao Q, Acharya K, Wang W, Yu Z (2012) Statistical downscaling of extremes of precipitation and temperature and construction of their future scenarios in an elevated and cold zone. Stoch Environ Res Risk Assess 26(3):405–18