Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006

Short title: Trends in extreme precipitation by non-stationary POT analysis

Santiago Beguería[†]

Aula Dei Experimental Station – CSIC, Zaragoza, Spain

Marta Angulo-Martínez

Aula Dei Experimental Station – CSIC, Zaragoza, Spain

Sergio M. Vicente-Serrano

Pyrenean Institute of Ecology – CSIC, Zaragoza, Spain

J. Ignacio López-Moreno

Pyrenean Institute of Ecology – CSIC, Zaragoza, Spain

Ahmed El-Kenawy

Pyrenean Institute of Ecology – CSIC, Zaragoza, Spain

† Corresponding author email: sbegueria@eead.csic.es; telephone: +34.976.716.158; fax: +34.976.716.019

Abstract: Most applications of the extreme value theory have assumed stationarity, i.e. that the statistical properties of the process do not change over time. However, there is evidence suggesting that the occurrence of extreme events is not stationary but changes naturally, as it has been found for many other climate variables. Of paramount importance for hazard analysis is whether the observed precipitation time series exhibit long-term trends or cycles; such information is also relevant in climate change studies. In this study the theory of non-stationary extreme value analysis was applied to data series of daily precipitation using the peaks-over-threshold approach. A Poisson/Generalized Pareto model, in which the model parameters were allowed to vary linearly with time, was fitted to the resulting series of precipitation event's intensity and magnitude. A log-likelihood ratio test was applied to determine the existence of trends in the model parameters. The method was applied to a case study in northeast Spain, comprising a set of 64 daily rainfall series from 1930 to 2006. Statistical significance was achieved in less than 5% of the stations using a linear non-stationary model at the annual scale, indicating that there is no evidence of a generalised trend in extreme precipitation in the study area. At the seasonal scale, however, a significant number of stations along the Mediterranean (Catalonia region) showed a significant decrease of extreme rainfall intensity in winter, while experiencing an increase in spring.

Keywords: extreme events, precipitation, time series analysis, climate variability, regional climate change, non-stationary extreme value analysis, Iberian Peninsula, Spain

1. Introduction

Analysis of the characteristics of extreme precipitation over large areas has received considerable attention, mainly due to its implications for hazard assessment and risk management (Easterling et al., 2000). Hazardous situations related to extreme precipitation events can be due to very intense rainfall, or to the persistence of rainfall over a long period of time. Reliable estimates of the probability of extreme events are required for land planning and management, the design of hydraulic structures, the development of civil protection plans, and in other applications.

The extreme value (EV) theory provides a complete set of tools for analyzing the statistical distribution of extreme precipitation, allowing for the construction of magnitude–frequency curves (Hersfield, 1973; Reiss and Thomas, 1997; Coles, 2001; Katz et al., 2002). Derived statistics such as quantile estimates (the average expected event for a given return period) have been widely used to express the degree of hazard related to extreme precipitation at a given location. In most cases the analysis of extreme events has been reduced to the daily intensity, although other important aspects, including event duration and magnitude, have recently been incorporated into the analysis of extreme precipitation (Beguería et al., 2009).

An important assumption of the classical EV theory refers to the stationarity of the model, which implies that the model parameters do not change over time. However, climatic series are known to be non-stationary, and hence many studies have been devoted to analyzing the occurrence of temporal trends and cycles, including the frequency and severity of extreme events (Smith, 1989; Karl et al., 1995; Karl and Knight, 1998; Groisman et al., 1999). In addition, several models have highlighted a likely increase in the frequency of extreme events under modified greenhouse gas emissions scenarios (Meehl et al., 2000; Groisman et al., 2005; Kyselý and Beranová, 2009), so advances in methodologies for assessing such changes are needed.

The importance of these issues has stimulated the development of techniques to identify trends in extreme events. A large number of studies are based on analyzing the time variation of statistics defined by fixed magnitudes or quantiles (Karl et al., 1995; Groisman et al., 1999; Brunetti et al., 2004; López-Moreno et al., 2009). These approaches can provide information about changes in high precipitation values, but not necessarily about the most extreme precipitation events, which by definition occur intermittently. For this reason other approaches have been developed to analyze changes and trends in hydrological extremes based on parametric approaches, which are based on an extension of the EV theory, namely the non-stationary extreme value (NSEV) theory (i.e., Coles, 2001). NSEV methods allow analysis of the time dependence of extreme precipitation within the extreme value theory context, enabling estimation of the time evolution of the most extreme precipitation events. The NSEV theory is the most reliable framework for analyzing the time variation of extreme events, providing the means to address important issues including the occurrence of trends and cycles in data series, as well as co-variation with other climatologic and meterological factors.

Most applications of the NSEV theory, however, have been based on block maxima data, in which only the *n* highest observations are retained at regular time intervals. The annual maxima method is a common example of this approach. The resulting data series are fitted to extreme value distributions such as the Gumbel distribution or the Generalized Extreme Values distribution (Hershfield, 1973), and there are examples of non-stationary analysis of extreme rainfall using the GEV distribution (Nadarajah, 2005; Pujol et al., 2007). There has been criticism of the waste of useful information when the block maxima approach is applied to datasets containing data in addition to the maxima (Coles, 2001; Beguería, 2005). As an alternative, the peaks-over-threshold (POT) approach is based on sampling all observations exceeding a given threshold value, and fitting the resulting data series to the Exponential or the Generalized Pareto (GP) distributions (Cunnane, 1973; Madsen and Rosbjerg, 1997).

There are some studies demonstrating non-stationarity on POT data. Li et al. (2005) used a split-sample

approach to demonstrate temporal changes on extreme precipitation in Western Australia. Hall and Tajvidi (2000) used a moving kernel sampling approach to analyze non-stationarity on the GP parameters of extreme temperature and wind speed data. Non parametric approaches (split sampling and moving kernel sampling) are interesting as a preliminary tool for exploring the presence of non-stationarity on POT data, but do not provide a functional relationship between the GP distribution parameters and time. Parametric methods to account for this relationship have been developed based on the maximum likelihood estimation (MLE) method (Smith, 1999; Coles, 2001), but the examples of their application to hydroclimatic data are relatively scarce. One of the first applications of the MLE method to model time dependence on temperature and precipitation POT data is due to Smith (1999). Chavez-Demoulin and Davison (2005) used a penalized likelihood approach and generalized additive modeling (GAM) to fit the co-variation between GP distribution parameters and other covariates. Nogaj et al. (2006) used MLE for fitting linear and quadratic trends to the parameters of temperature extremes over the North Atlantic region. A similar approach was used by Laurent and Parey (2007) and Parey (2007) for temperature extremes in France. Méndez et al. (2006) analyzed long time trends and seasonality of POT wave height data. Yiou et al. (2006) analyzed trends of POT discharge data in the Czech Republic. Abaurrea et al. (2007) used MLE to model covariation between the scale parameter of a GP distribution and mean temperature data to check for trends of POT temperature series. Applications of the NSEV theory to POT precipitation data have been very scarce, due mostly to difficulties inherent to the irregular and clustered character of rainfall time series. The study by Smith (1999) is one of the very few references, and it does not incorporate recent advances in POT analysis of precipitation data such as declustering (i.e., using series of rainfall events instead of the original daily series), nor compares the MLE method with other alternative methods such as the non-parametric ones.

Whether there are trends in extreme precipitation records is currently of major interest in climate change studies, and the evaluation of observational datasets at regional scales is still a key research focus for

better understanding of the implications of climate change for precipitation extremes. In this study we explain the principles of application of non-stationary peaks-over-threshold analysis to time series of extreme precipitation events, considering both their intensity and magnitude. One non-parametric (moving kernel sampling) and one parametric (MLE) methods are presented and practical aspects on their use are discussed. We illustrate the use of these techniques with a case study based in the northeast sector of the Iberian Peninsula, which has a strong Atlantic–Mediterranean climate gradient. The purpose of the study is twofold: i) to expand current knowledge of the temporal evolution of extreme precipitation in the Iberian Peninsula, and ii) to demonstrate the application of non-stationary peaks-over-threshold analysis to detection of changes in the most extreme precipitation events.

2. Methods

2.1. Stationary peaks-over-threshold analysis

A classical approach to EV analysis involves fitting a given probability distribution function to the highest values of a data series, in order to obtain reliable estimates of quantiles. The peaks-over-threshold (POT) approach is based on sampling only the observations above a given threshold value x_0 (Cunnane, 1973; Madsen and Rosbjerg, 1997). The choice on the value of x_0 allows control over the amount of data in the analysis.

Once it is established that the value of x_0 is sufficiently high, the inter-event arrival times $A = t_n - t_{n-1}$ can be considered random, and the process is defined as a Poisson process with average occurrence frequency λ (number of events per year). The probability distribution of a POT variable with random occurrence times belongs to the Generalized Pareto (GP) family (Leadbetter et al., 1983). The Poisson/GP

model has frequently been used for EV analysis (van Montfort and Witter, 1986; Hosking and Wallis, 1987; Wang, 1991; Madsen and Rosbjerg, 1997; Martins and Stedinger, 2001; Beguería et al., 2009). The GP distribution is a flexible, long-tailed distribution described by a shape parameter κ and a scale parameter α. It has the cumulative distribution function (Rao and Hamed, 2000):

$$P(X > x \mid X > x_0) = \left(1 + \kappa \frac{x - x_0}{\alpha}\right)^{-1/\kappa}, \qquad \text{(eq. 1)}$$

where α can take any arbitrary value, and κ controls the shape of the function, and hence the more or less pronounced character of its right tail. For $\kappa > 0$ the distribution is long-tailed (longest for smaller values of κ), and for $\kappa < 0$ it becomes upper-bounded, with the endpoint at $-\alpha/\kappa$. In the special case where $\kappa = 0$ the GP distribution yields an exponential distribution. It is possible to calculate the expected value of the *n*-year return period quantile, x_n, as well as its standard deviation, σ_{xn} (for more details see Rao and Hamed, 2000).

The choice of an appropriate threshold value, x_0, is an important step in POT modelling, as this parameter controls the sample size, and also affects the assumption of independence of arrival times and the adequacy of the GP distribution (Valadares-Tavares and Evaristo-da Silva, 1983). The suitability of the P/GP model for a POT series defined by a threshold value x_0 can be tested by the mean excess plot. Given a rainfall amount at time *t*, x_t, a new variate $y_t = x_t - x_0$ can be constructed giving the exceedance over the threshold if $x_t > x_0$. The mean excess plot represents the average threshold exceedances $E(y_t \mid x_t > x_0)$ against the value of x_0. The mean excess plot allows definition of an appropriate threshold value as the lowest value of x_0 for which $E(y_t)$ is a linear function of x_0 (Coles, 2001). Besides, goodness-of-fit tests have been developed for assessing if the time arrival of an event greater than the threshold fit a Poisson process, and if the magnitude of the exceedances over the threshold fit a GP distribution. For the events to be considered time-independent it is necessary that the dispersion index *DI* (the ratio between the variance and the average

occurrences per year) of the event series must be approximately 1. Confidence limits for *DI* can be obtained from a chi^2 distribution at *n*-1 degrees of freedom, *n* being the number of years of the series (Cunnane, 1979). Similarly, the one sample Kolmogorov-Smirnov test can be used to check the validity of the GP distribution for the POT series.

Several numerical methods exist for obtaining sample estimates of the P/GP model parameters (Rao and Hamed, 2000). In this study we adopted the maximum likelihood method, which has the advantage over other procedures (such as the probability-weighted method) of being flexible enough to accommodate both stationary and non-stationary models.

2.2. The non-stationary POT model

In the classical POT theory, the model parameters are assumed to remain stationary over time. In contrast, in non-stationary POT (NSPOT) analysis the model parameters are allowed to vary. Non-stationarity can arise in time series of extreme values due to seasonal and interannual effects including varying atmospheric circulation patterns and synoptic types predominating in different months, or arising from long-term cycles or trends, for example as a consequence of climate change. In fact, it is appropriate to suggest that non-stationarity is more common than stationarity in series of climate extremes, and to recommend the use of the NSPOT methodology unless the assumption of stationarity can be proven for a given data series.

Non-stationarity of the POT model parameters can be modelled in several ways, and we will refer here to the non-parametric and the parametric NSPOT models. It should be noted that in this study the terms non-parametric and parametric refer only to the way in which time dependence of the model parameters was modelled, but in both cases the P/GP model was used. Letting $P/GP(x_0,\alpha,\kappa,\lambda)$ be a stationary Poisson/Generalized Pareto model with threshold x_0 and parameters α, κ and λ, as defined in the previous section, the non-parametric NSPOT model is defined by:

$$\left(X_s > x - x_{0,s} \mid X_s > x_{0,s}\right) \sim P/GP\left(\alpha_s, \kappa_s, \lambda_s\right), \quad \text{(eq. 2)}$$

where $x_{0,s}$, α_s, κ_s and λ_s are the model parameters for a given time period s. These are assumed to be independent of the parameters for other time periods, so they can be estimated directly from a sample of data for that period. The simplest approach is to split the complete time series into *n* periods of equal length. A split-sample approach was followed, for example, by Li et. al (2005) for identifying non-stationarity of POT rainfall data in Australia. Another approach is to use a moving time kernel of fixed time span *m*, to obtain subsamples for each year of the time series. Such an approach was used, for example, by Hall and Tajvidi (2000) to model non-stationarity of annual maxima and POT temperature and wind data. Independent models can also be fitted to POT data for different months, or for each season (Beguería et al., 2009). A similar approach was used by Brath et al., (2001) for annual maxima precipitation data.

Time dependence is modelled in a more sophisticated way in the parametric NSPOT model:

$$\left(X(t) > x - x_0(t) \mid X(t) > x_0(t)\right) \sim GP\left(\alpha(t), \kappa(t), \lambda(t)\right), \quad \text{(eq. 3)}$$

where a functional relationship exists between the model parameters and time. This could be linear, for example:

$$\pi(t) = \beta_0 + \sum_{i=1}^{p} \beta_i t^i,$$

where $\pi(t)$ represents one of the NSPOT model parameters, β_i $\left(0 \le i \le p\right)$ are the linear coefficients, and *p* is the polynomial order. A simple linear model (n = 1) is normally used (Smith, 1999; Katz et al., 2002), but higher order linear and even nonlinear relationships may also be used. Although high order polynomials allow fitting of almost any time series they can easily lead to unreliable models due to overfitting, so linear or log-linear models are usually preferred when searching for trends in the occurrence of extreme events. It is possible to define a model with time dependence in all parameters, or with a combination of stationary and

non-stationary parameters. As no data sub-sampling is required, all the data in the POT series are used in fitting the model, which constitutes one of the main advantages of the parametric NSPOT model over the non-parametric model. Maximum likelihood estimations of the parameters β_i can be easily obtained (Coles, 2001).

As a consequence of the large number of alternative model formulations, it is important to have an objective method for model selection, enabling choice of the simplest model that explains the largest amount of variance in the data. The process usually involves: i) the definition of a set of candidate NSPOT models; ii) fitting the model parameters; and iii) using a log-likelihood ratio test based on the distance statistic D (e.g., Coles, 2001):

$$D = -2\{\ell_1(M_1) - \ell_0(M_0)\}, \qquad \text{(eq. 4)}$$

where ℓ_1 and ℓ_0 are the maximized log-likelihoods of two nested models, M_1 and M_0, respectively, the former being more complex (i.e. having a larger number of parameters) than the latter. Large values of D indicate that model M_1 explains substantially more of the variance in the sample than M_0, and low values of D suggest that the gain in explanatory capacity of M_1 over M_0 does not compensate for its greater complexity. The model M_0 can be rejected for $D > c_\alpha$, where c_α is the $(1-\alpha)$ quantile of a χ^2_k distribution, k being the difference in the number of parameters between M_1 and M_0, and α the significance level of the test. In order to check the power of the *D* statistic test a numerical experiment was carried out involving the generation of a high number (10,000 replications) of synthetic time series following i) stationary and ii) linearly time-dependent P/GP models and having the same length than the series used in this study was carried out. The results were satisfactory, since the test yielded approximately the expected number of false positives at the significance level $\alpha = 0.05$ (5%).

3. Study case

3.1. Study area and data set

The study area is in the north-eastern quadrant of Spain (Fig. 1), and has boundaries corresponding to administrative units that include 18 provinces with a total area of 160,000 km^2. The study area has contrasting relief, with the Ebro Depression located in the centre (~200 m a.s.l.) and surrounded by high mountain ranges. The Cantabrian Range and the Pyrenees delimit the area to the north, with maximum elevations above 3,000 m a.s.l. The Iberian Range (maximum elevations 2,000–2,300 m a.s.l) encloses the Ebro Depression to the south. To the west of the study area the main unit is the Meseta, a plain structure with elevations of 700–1,000 m a.s.l. To the east, parallel to the Mediterranean coast, the Ebro Depression is enclosed by the Catalan Coastal Range (maximum elevation 1,000–1,200 m a.s.l.). The climate of the area is dominated by westerly fluxes from the west and northwest, although rainfall also responds to the influence of the Mediterranean Sea.

Data were collected from a network of 64 climate observatories that have recorded daily precipitation data continuously for the period 1930–2006 (Figure 1). The data series underwent a quality control process that included reconstruction, gap filling and homogeneity testing (Vicente-Serrano et al., 2010). A declustering algorithm was applied to the original daily series to obtain series of precipitation events. A precipitation event was defined as a series of consecutive days with precipitation, so a period of one or more days with zero precipitation was the criterion used to separate events. Declustering has important consequences for the analysis, as it allows elimination of undesirable serial dependence in the point process (Coles, 2001; Beguería, 2005). It also allows for a more complete analysis compared with a daily series, as the magnitude of the events can be also analyzed. Hence, two variables were obtained for each precipitation event: its magnitude (the total amount of precipitation cumulated over the event, i.e. the sum of all daily precipitations, in mm) and its

intensity (the highest daily precipitation during the event, in mm d^{-1}). Each event was assigned to the last day of the cluster, which enabled construction of a time series of precipitation events. Five data sets were constructed: one including all the events, and four seasonal datasets including only the events occurring in winter, spring, summer and autumn, respectively.

3.2. Analysis design

POT series were created from the original event magnitude and intensity series by selecting an appropriate threshold value for each series using the mean excess plot method. POT series were constructed for the complete series and for the four seasonal event series.

Non-parametric NSPOT models were obtained by fitting, for each year t_n (1939>n>1997), an independent P/GP model by fitting a subsample of all the events in the POT series within the 20-year period (t_{n-10}, t_{n+10}). This allowed time series of the model parameters (scale and shape) and the 10-year return period quantiles (q_{10}) to be obtained for the period 1940-1996. These time series were used as an exploratory tool due to the reduced time span of the series.

Parametric NSPOT P/GP models were fitted to the POT series using a linear relationship for parameters α and κ and time. The models were stationary with respect to the threshold value, x_0, which was kept constant and not time-dependent. The maximum likelihood method was used for fitting the models. The *D* statistic test was used to determine the significance of the linear model in each case. Time series were constructed of the model parameters, and of q_{10} and their 95% confidence intervals. Plots of these series, together with their non-parametric counterparts, were produced, and the two approaches were visually compared for inconsistencies. Finally, maps of the difference between q_{10} estimates by a linear NSPOT model at the end and the beginning of the study period 1930-2006 were produced for visualizing the results.

All calculations (mean excess plots, parameter estimation, model selection and factor analysis) were performed in R (R Development Core Team, 2008) using the ismev and the evd packages (Coles, 2001; Stephenson, 2002). Mapping was performed using ArcGIS 9.3 (ESRI, 2009).

3.3. Results

A threshold value corresponding to the 90th percentile of the distribution of the event's intensity and magnitude for the 76 years of data was established with the help of the mean excess plot (Fig. 2). The use of a percentile instead of an absolute value (in mm) allowed use of a single value to define the threshold for all the data series. The absolute value corresponding to the 90th percentile was readily determined for each data series. Several percentiles were initially considered as possible threshold values, and their validity was checked for each series by means of the linearity of the mean excess plot, plus the use of the *DI* test for the event's arrival times and the K-S test for adequacy of the GP distribution. The 90th percentile represented a conservative choice, as a lower value could have been selected for several stations. However, according to the P/GP model theory it represented a good threshold for all the event series. It must be noted that, when working with daily precipitation data, it is common to use values higher than the 95th percentile as the threshold value, and often the 99^{th} percentile is used. This results in an event arrival frequency of (0.05*365) 18.25 events per year for the 95th percentile, (0.01*365) 3.65 events per year for the 99th percentile, etc. The case is different when analyzing precipitation event data instead of daily data, since the number of events per year is substantially reduced. As a consequence, the same percentiles applied to event data result in significantly lower arrival frequencies. In the case reported in this study the average arrival frequency was 4.14±0.59 events per year. Time plots of the POT series did not allow distinguishing differences in the density of the variable, as expected from a series of extreme (and thus rare) events (Fig. 3).

The results of the *DI* and K-S tests allowed accepting the hypothesis of the convenience of the P/GP

model for the POT event series above the 90th percentile for a majority of stations, both at the annual and the seasonal time scales (Table 1). The K-S test was positive in all cases, demonstrating the goodness of fit of the POT series generated this way to the GP distribution. The *DI* test for the independence of the point process was not positive for all stations, although the percentage of positive cases seemed appropriate for accepting the P/GP model for the whole dataset. Other studies have shown a better result of the *DI* test for POT event series above the 90th percentile (Beguería, 2005), although in the context of stationary POT analysis. It is possible that non-stationarity affects the Poisson assumption of the P/GP model, most especially in the presence of a temporal trend. This hypothesis, however, could not be validated in this study, since no significant differences were found between the P/GP model parameter trends of positive and negative results of the *DI* test.

The POT series that resulted from taking the exceedances over the threshold value were fitted to linear NSPOT models, and their significance was checked using a log-likelihood ratio test against a stationary model. At the annual scale only 11.0% (seven cases) and 4.7% (three cases) of the linear NSPOT models were significantly better than a stationary one for the event intensity and magnitude, respectively (Table 1). At the seasonal scale, winter and spring showed a higher fraction of significant non-stationary models, raising up to 20 (13 cases) and 17% (11 cases) for the event intensity, respectively, while for the event magnitude and for both variables in the other seasons the fraction of positive NSPOT models was lower than 10%. Time series of the model parameters and the q_{10} show the importance that non-stationarity had in certain series (Figs 4 and 5). Although for some series there were no major differences in q_{10} between the beginning and the end of the study period (e.g. the series X287 for the event's intensity), in many cases the estimated q_{10} at both extremes differed by more than 50% (e.g. series X164). The time evolution of q_{10} depended on the evolution of both α (scale) and κ (shape), although the effect of the latter parameter was much stronger. The amplitude of the confidence band was also well correlated to the value of κ, with a negative relationship. It is

noteworthy that, although a linear model was used for both parameters, their combined effect can lead to a curvilinear relationship between q_{10} and time, although the highest / lowest values of q_{10} always occurred at the beginning / end of the time period.

Comparison of the non-parametric and the parametric approaches to NSPOT modelling (dots and lines in Figures 4 and 5) shows the main advantages of the latter approach: i) to provide estimates of extreme events at the beginning and the end of the time series, which are lost in the non-parametric model due to the moving window necessary to obtain subsamples; and ii) to offer robustness against the occurrence of very extreme events, which have a major impact on the non-parametric model due to the shorter time span of the subsamples. This is clearly shown by the magnitude of the standard error of q_{10}, which was much higher for the non-parametric model (Fig. 6). The standard error of the parametric NSPOT model was, in average, equivalent to that of the (parametric) stationary POT model, illustrating one of the principal advantages of the parametric NSPOT approach.

High order polynomial NSPOT models were also fitted to the data (Fig. 7). Models of increasing complexity were tested for significance against the stationary model, and the most parsimonious (i.e. with the lowest polynomial order) was chosen. Following this approach, statistical significance was found in 45 (70.3%) models for intensity and 42 (65.5%) for magnitude. The models had orders between 5 and 9 for both α and κ, which reflects quite a high complexity. It is evident that high order polynomials are more flexible for fitting complex data, but that comes at the expense of reduced reliability. In fact, the models have the same drawbacks as the non-parametric approach: i) excessive impact of outlier observations resulting in over-fitting, and ii) non-reliability at the beginning and the end of the study period, as shown by the large increase in the uncertainty bands in many cases (e.g. series X9198 in Fig. 7, and short-dashed lines in Fig. 6).

With respect to the time series of q_{10}, differences were occasionally found between adjacent

observatories, reflecting the influence of single, very localized extreme events. Maps of q_{10} where produced to visualize the spatial distribution of the q_{10} difference at the beginning and the end of the period of study, and to check the regional coherence of the stations for which a linear NSPOT model was significant (Figs. 8 and 9). Differences in q_{10} were not very high for both the event intensity and magnitude at the annual level (ranging between -10 and +30% in most cases), and the spatial distribution of the few significant series was random, suggesting that no trends exist in the study area.

At the seasonal level, however, the differences between the beginning and the end of the period of study were larger, especially for the event's intensity in winter (for which a decrease of both intensity a magnitude was recorded in most of the region) and spring (with a generalised increase). For these seasons a cluster of stations with significant NSPOT models appeared close to the Mediterranean coast line (Catalonia) for the event's intensity, suggesting that the results found with respect to the time variation of q_{10} are consistent in this area. For the event's intensity in summer and autumn, and for the event's magnitude in all four seasons, the spatial patterns were weaker and the distribution of the significant series seemed random.

4. Discussion and conclusions

The stationarity hypothesis is an important assumption of conventional extreme events analysis, although it has been shown to do not hold for many environmental variables. Methodologies such as the one used in this study are able to accommodate non-stationarity on POT time series, and even allow testing for stationarity. Non-stationary POT models allow finding temporal trends in extreme event series, and are of special interest in the context of climate change studies. At this respect, regional assessments comprising many data series within a region could aid identification of relevant patterns in the time evolution of extreme events.

The non-stationary peaks-over-threshold (NSPOT) methodology facilitates assessment of the temporal occurrence of extreme climatic events or other environmental variables, allowing time series of associated

statistics of applied interest (such as low-frequency quantiles) to be obtained. Once a threshold value in the form of a quantile value is set (here we used the 90th quantile of the complete time series of precipitation events), a Poisson/Generalized Pareto model can be defined in which the values of the model parameters (the scale and shape of the distribution) are time dependent. Such a model can be fitted using the maximum likelihood method. One of the difficulties in using the NSPOT model lies in the definition of the time dependency model, as many alternative relationships can be proposed. A test based on the *D* statistic, representing the gain in log-likelihood between two alternative models, can be used as a statistically sound model selection criterion. The proposed test penalizes the increased complexity involved in higher order models, thus constituting a trade-off between gain in explanatory power and complexity. Here we propose to use this test to establish the statistical significance of the non-stationary model against the stationary one. In addition to the parametric NSPOT model we used a non-parametric approach involving the fitting of a series of stationary models to subsamples of the original time series, following a moving window approach. This method was used as an exploratory tool, as it has many drawbacks affecting its reliability as a tool to model non-stationarity in extreme event analysis.

We applied this methodology to a dataset of peaks-over-threshold rainfall events magnitude and intensity series covering the northeast of Spain. At the annual scale, our results showed that at most stations a linear NSPOT model was not significant according to the *D* statistic test, and the spatial distribution of the significant stations was apparently random. Hence no strong evidence was found in support of a change in the frequency and magnitude of extreme events in the region. The results from the non-parametric approach, which are not tied to any parametric trend model, showed a complex time pattern in most cases, with a strong influence of extremely high events. High order polynomial NSPOT models were able to fit such complex temporal patterns, but were constrained by the same drawbacks found in the non-parametric approach. At the seasonal scale, a higher number of stations for which a linear NSPOT model was significant

were found for the event's intensity in winter (decreasing) and spring (increasing). A large fraction of this stations were clustered spatially in the Mediterranean part of the study region (Catalonia), suggesting that these results are consistent in that area. Difficulties for fitting linear NSPOT models to the observed data may be due to the fact the model parameters were set as time dependent only. This situation could be improved if covariates such as the NAO index or other tele-connexions would be included in the analysis. Previous studies (Lana et al., 1995; Martín-Vide et al., 2008; Vicente-Serrano et al., 2009) shown that extreme precipitation in north-east Spain may be linked to westerly flows associated with Atlantic depressions and other influences from the east (Mediterranean), most especially during winter, spring, and autumn.

Acknowledgements

This research has been supported by the following projects: CGL2008-00831/BTE, CGL2006-11619/HID, CGL2008-01189/BTE, and CGL2008-1083/CLI funded by the Spanish Commission of Science and Technology and FEDER; EUROGEOSS (FP7-ENV-2008-1-226487) and ACQWA (FP7-ENV-2007-1-212250) funded by the VII Framework Programme of the European Commission; "Las sequías climáticas en la cuenca del Ebro y su respuesta hidrológica" and "La nieve en el Pirineo Aragonés: distribución espacial y su respuesta a las condiciones climáticas" funded by "Obra Social La Caixa" and the Aragón Government; and "Programa de grupos de investigación consolidados" funded by the Aragón Government.

Reference list

Abaurrea J, Asín J, Cebrián AC, Centelles A. 2007. Modeling and forecasting extreme hot events in the

central Ebro valley, a continental-Mediterranean area. *Global and Planetary Change* 57: 43–58.

Beguería S. 2005. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology* 303: 215–230.

Beguería S, Vicente-Serrano SM, López-Moreno JI, García-Ruiz JM. 2009. Annual and seasonal mapping of peak intensity, magnitude and duration of extreme precipitation events across a climatic gradient, North-East Spain. *International Journal of Climatology* 29(12): 1759—1779.

Brath A, Castellarin A, Montanari A. 2001. At-site and regional assessment of the possible presence of non-stationarity in extreme rainfall in northern Italy. *Physics and Chemistry of the Earth (B)* 26(9): 705–710.

Brunetti M, Buffoni L, Mangianti F, Maugeri M, Nanni T. 2004. Temperature, precipitation and extreme events during the last century in Italy. *Global and Planetary Change* 40(1-2): 141–149.

Chavez-Demoulin V, Davison AC. 2005. Generalized additive modelling of sample extremes. *Applied Statistics* 54: 207–222.

Coles S. 2001. *An introduction to statistical modeling of extreme values*. Springer-Verlag: London.

Cunnane, C. 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. *Journal of Hydrology* 18 (3/4): 257–271.

Cunnane, C. 1979. A note on the Poisson assumption in partial duration series models. *Water Resources Research* 15: 489 – 494.

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO. 2000. Climate extremes: observations, modeling, and impacts. *Science* 289(5487): 2068–2074.

ESRI (Environmental Systems Resource Institute). 2009. ArcGIS 9.3. ESRI, Redlands, California.

Groisman PYa, Karl TR, Easterling DR, Knight RW, Jamason PB, Hennessy JK, Suppiah R, Page ChM,

Wibig J, Fortuniak K, Razuvaev VN, Douglas A, Forland E, Zhai PM. 1999. Changes in the probability of heavy precipitation: important indicators of climatic changes. *Climatic Change* 42(1): 243–283.

Groisman PYa., Knight RW, Easterling DR, Karl TR, Hegerl GC, Razuvaev VN. 2005. Trends in intense precipitation in the climate record. *Journal of Climate* 18(9): 1326–1350.

Hall P., Tajvidi N. 2000. Non-parametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science* 15: 153–167.

Hershfield DM. 1973. On the probability of extreme rainfall events. *Bulletin of the American Meteorological Society* 54: 1013–1018.

Hosking JRM, Wallis JR. 1987. Parameter and quantile estimation for the Generalized Pareto distribution. *Technometrics* 29(3): 339–349.

Karl T, Knight RW. 1998. Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States. *Bulletin of the American Meteorological Society* 79(2): 231–241.

Karl TR, Knight RW, Plummer N. 1995. Trends in high-frequency climate variability in the Twentieth Century. *Nature* 377(6546): 217–220.

Katz RW. 1999. Extreme value theory for precipitation: sensitivity analysis for climate change. *Advances in Water Resources* 23: 133–139.

Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* 25(8-12): 1287–1304.

Kyselý J, Beranová R. 2009. Climate change effects on extreme precipitation in central Europe: uncertainties of scenarios based on regional climate models. *Theoretical and Applied Climatology*, doi: 10.1007/s00704-008-0014-8.

Lana X, Mills GF, Burgueño A. 1995. Daily precipitation maxima in Catalonia (North-East Spain): expected values and their spatial distribution. *Int. J. Climatol.* 15: 341-354.

Lana X, Burgueño A, Martínez MD, Serra C. 2005. Statistical distributions and sampling strategies for the analysis of extreme dry spells in Catalonia (NE Spain). *Journal of Hydrology* 324(1-4): 94–114.

Laurent C, Parey S. 2007. Estimation of 100-year return-period temperatures in France in a non-stationary climate: results from observations and IPCC scenarios. *Global and Planetary Change* 57: 177–188.

Leadbetter MR, Lindgren G, Rootzén H. 1983. *Extremes and related properties of random sequences and series.* Springer Verlag: New York.

Li Y, Cai W, Campbell EP. 2005. Statistical modeling of extreme rainfall in southwest Western Australia. *Journal of Climate* 18: 852–863.

López-Moreno JI, Beniston M. 2009. Daily precipitation intensity projected for the 21st century: seasonal changes over the Pyrenees. *Theoretical and Applied Climatology* 95: 375–384.

Madsen H., Rosbjerg D. 1997. The partial duration series method in regional index-flood modeling. *Water Resources Research* 33: 737–746.

Martins ES, Stedinger JR. 2001. Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research* 37(10): 2551–2557.

Martín-Vide, J.; Sánchez-Lorenzo, A.; López-Bustins, J.A.; Cordobilla, M.J.; García-Manuel, A. and Raso, J.M. (2008): Torrential Rainfall in Northeast of the Iberian Peninsula: Synoptic patterns and WeMO influence. *Advances in Science and Research*, 2, 99-105.

Meehl GA, Zwiers F, Evans J, Knutson Th, Mearns L, Whetton P. 2000. Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. *Bulletin of the*

American Meteorological Society 81(3): 427–436.

Méndez FJ, Menéndez M, Luceño A, Losada IJ. 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model. *Journal of Geophysical Research C: Oceans* 111(7): C07024, doi: 10.1029/2005JC003344.

Nadarajah S. 2005. Extremes of daily rainfall in West Central Florida. *Climatic Change* 69: 325-342.

Nogaj M, Yioui P, Parey S, Malek F, Naveau P. 2006. Amplitude and frequency of temperature extremes over the North Atlantic region. *Geophysical Research Letters* 33: L10801, doi:10.1029/2005GL024251.

Parey S, Malek F, Laurent C, Dacunha-Castelle D. 2007. Trends and climate evolution: statistical approach for very high temperatures in France. *Climatic Change* 81: 331–352.

Pujol N, Neppel L, Sabatier R. 2007. A multivariate regional test for trend detection in extreme rainfall: the case of extreme daily rainfall in the French Mediterranean area. *Water Resources Research* 44(8): W08419, doi:10.1029/2007WR006268.

R Development Core Team. 2008. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna. Available online at http://www.R-project.org.

Rao AR, Hamed KH. 2000. *Flood Frequency Analysis*. CRC Press: Boca Ratón, Florida, USA.

Reiss R-D, Thomas M. 1997. *Statistical analysis of extreme values*. Birkhäuser: Basel-Boston-Berlin.

Smith RL. 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science* 4: 367–93.

Smith RL. 1999. Trends in rainfall extremes. Unpublished paper, available at http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html.

van Montfort MAJ, Witter JV. 1986. The generalized Pareto distribution applied to rainfall depths. *Hydrological*

Sciences Journal 31(22): 151-162.

Stephenson AG. 2002. evd: Extreme Value Distributions. *R News* 2(2): 31–32. Accessible online at http://CRAN.R-project.org/doc/Rnews/.

Vicente-Serrano S.M., Santiago Beguería, Juan I. López-Moreno, Ahmed M. El Kenawy y Marta Angulo. (2009): Daily atmospheric circulation events and extreme precipitation risk in Northeast Spain: the role of the North Atlantic Oscillation, Western Mediterranean Oscillation, and Mediterranean Oscillation. Journal of Geophysical Research-Atmosphere. 114, D08106, doi:10.1029/2008JD011492.

Vicente-Serrano SM, Beguería S, López-Moreno JI, García-Vera MA, Stepanek P. 2010. A complete daily precipitation database for North-East Spain: reconstruction, quality control and homogeneity. *International Journal of Climatology,* doi: 10.1002/joc.1850.

Valadares-Tavares L, Evaristo da Silva J. 1983. Partial duration series method revisited. *Journal of Hydrology* 64: 1–14.

Wang QJ. 1991. The POT model described by the Generalized Pareto distribution with Poisson arrival rate. *Journal of Hydrology* 129: 263–280.

Yiou P, Ribereau P, Naveau P, Nogaj M, Brázdil R. 2006. Statistical analysis of floods in Bohemia (Czech Republic) since 1825. *Hydrological Sciences Journal* 51: 930–945.

Tables

Table 1. Fraction of stations yielding a positive result (significance level $\alpha = 0.05$) of the DI test for the Poisson model of arrival times, K-S test for the Generalized Pareto model of exceedances and D test for a linear NSPOT model.

	DI test		K-S test		*D* test	
	Intensity	Magnitude	Intensity	Magnitude	Intensity	Magnitude
Year	0.89	0.91	1.00	1.00	0.11	0.047
Winter	0.88	0.97	1.00	1.00	0.20	0.063
Spring	0.97	0.92	1.00	1.00	0.17	0.078
Summer	0.92	0.94	1.00	1.00	0.078	0.094
Autumn	0.89	0.92	1.00	1.00	0.094	0.063

Figure captions

Figure 1. Map of the study area and spatial distribution of the observatories.

Figure 2. Mean residual life plots for four observatories: x axis, event intensity (mm d^{-1}); y axis (left), average exceedance over threshold (mm d^{-1}) and confidence band at $\alpha = 0.05$ (dashed line); y axis (right), cumulative distribution of the variable. The straight lines indicate the absolute value (mm d^{-1}) corresponding to the 90th percentile.

Figure 3. Peaks-over-threshold (POT) time series over the annual 90th percentile at four stations. Location of the stations is shown in Fig 1.

Figure 4. Time variation of the GP model scale and shape parameters and of the 10-year expected return period q_{10} for the event intensity (mm d^{-1}) at four stations, as estimated by non-parametric (circles) and parametric linear (solid lines) NSPOT approaches. The dashed lines in the q_{10} plots are uncertainty bands for $\alpha = 0.05$.

Figure 5. Time variation of the GP model scale and shape parameters and of the 10-year expected return period q_{10} for the event magnitude (mm) at four stations, as estimated by non-parametric (circles) and parametric linear (solid line) NSPOT approaches. The dashed lines in the q_{10} plots are uncertainty bands for $\alpha = 0.05$. Location of the stations is shown in Fig 1.

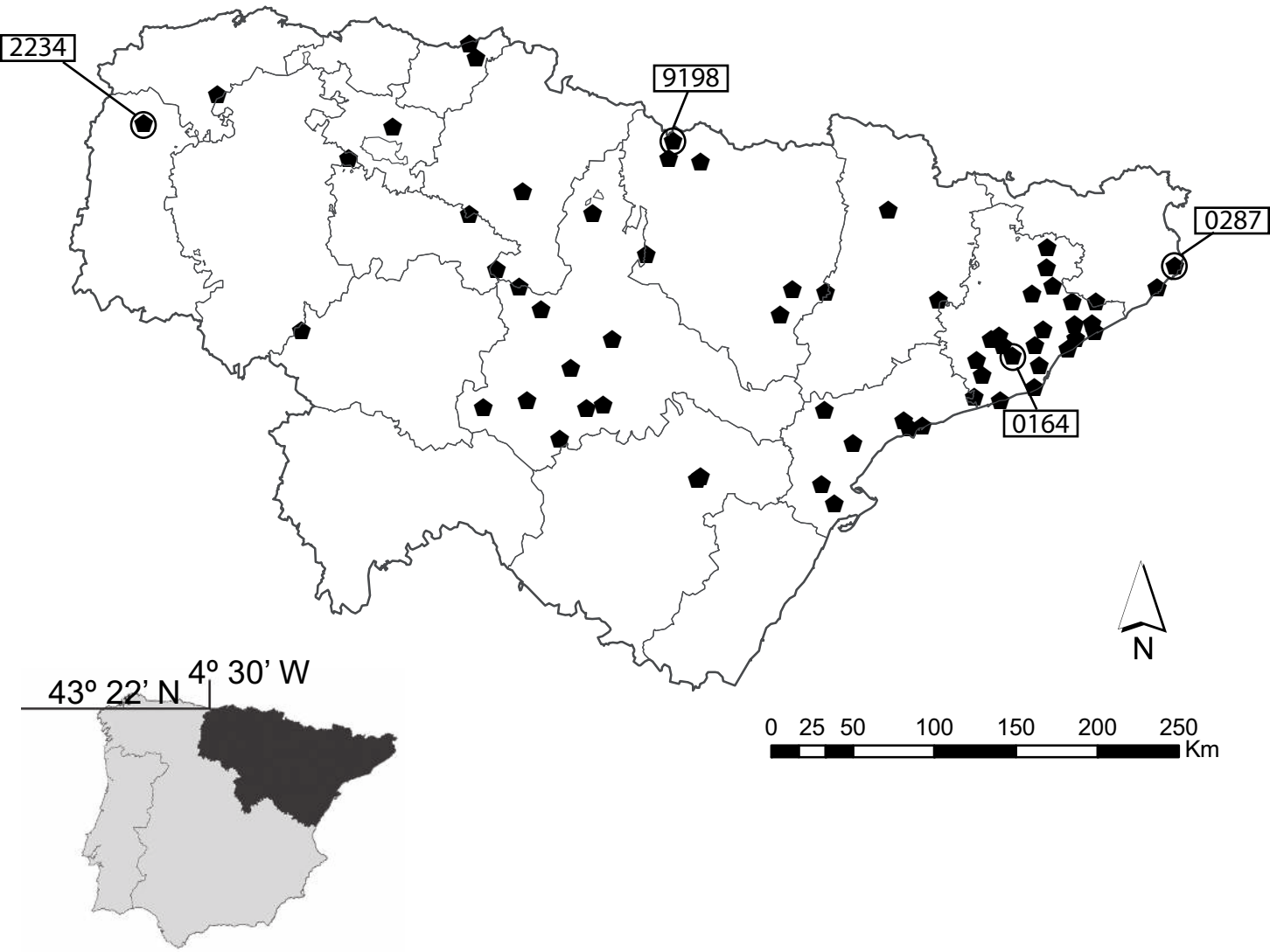
Figure 6. Model uncertainty: time variation of the 10-year return period event intensity standard error (mm d^{-1}) by a non-parametric NSPOT model (dots), a linear NSPOT model (solid line), a high order polynomial NSPOT model (short-dashed line) and a stationary POT model (long-dashed line) in four stations. Location of the stations is shown in Fig 1.

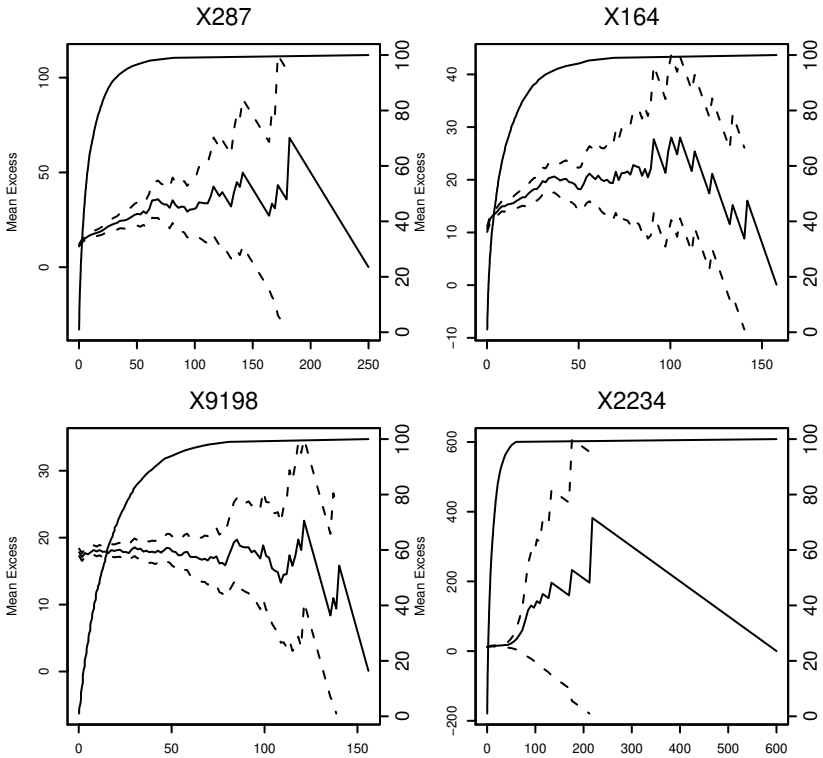
Figure 7. Time variation of the GP model scale and shape parameters and of the 10-year expected return

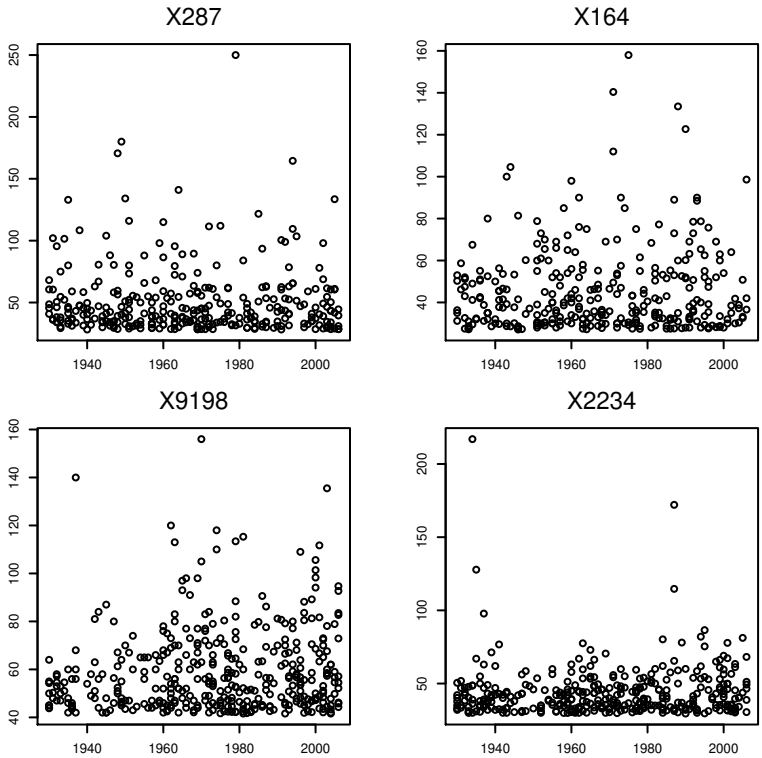
period q_{10} for the event intensity (mm d^{-1}) at six stations, as estimated by non-parametric (circles) and parametric high order polynomial (solid lines) NSPOT approaches.. The dashed lines in the q_{10} plots are uncertainty bands for $\alpha = 0.05$. Location of the stations is shown in Fig 1.

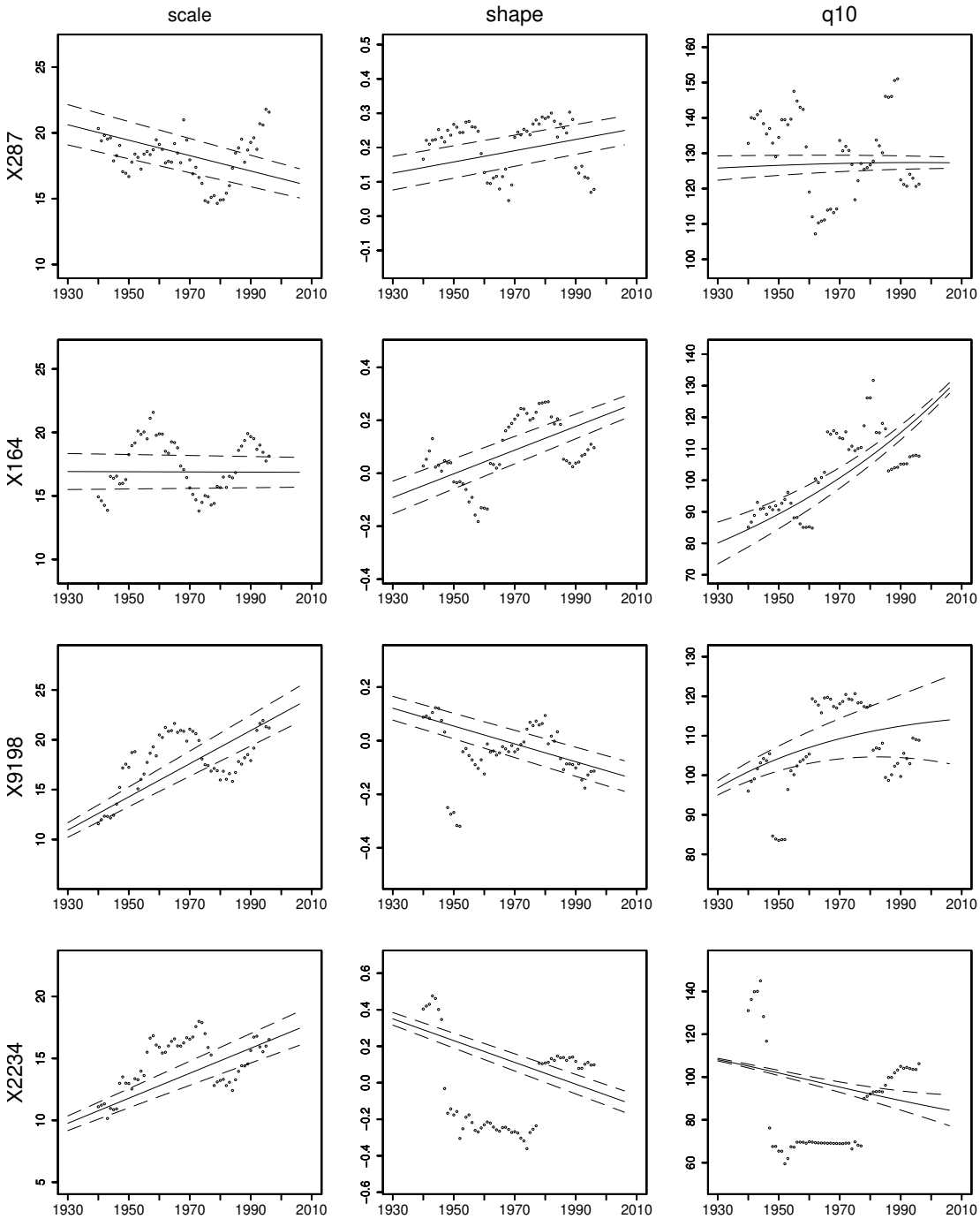
Figure 8. Percent variation of the 10-years return period quantile (q_{10}) over the period 1930-2006, $(q_{10}(2006) - q_{10}(1930))/q_{10}(1930) \times 100$, as estimated by a linear NSPOT model: event's intensity (a) and magnitude (b). The black dots indicate stations for which the linear NSPOT model was significantly better than the stationary one.

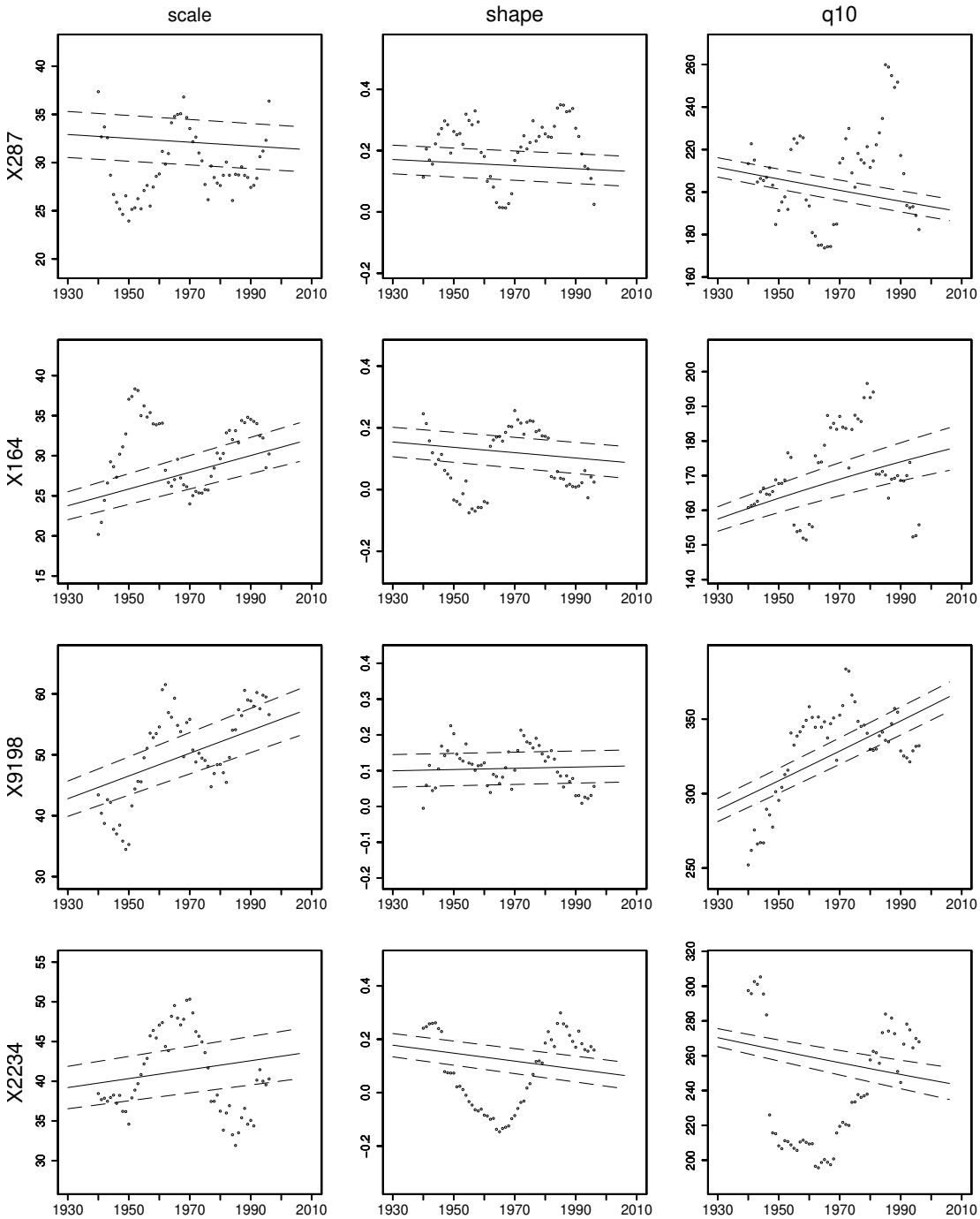
Figure 9. Percent variation of the 10-years return period quantile (q_{10}) over the period 1930-2006, $(q_{10}(2006) - q_{10}(1930))/q_{10}(1930) \times 100$, as estimated by a linear NSPOT model: seasonal event's intensity (a) and magnitude (b). The black dots indicate stations for which the linear NSPOT model was significantly better than the stationary one.

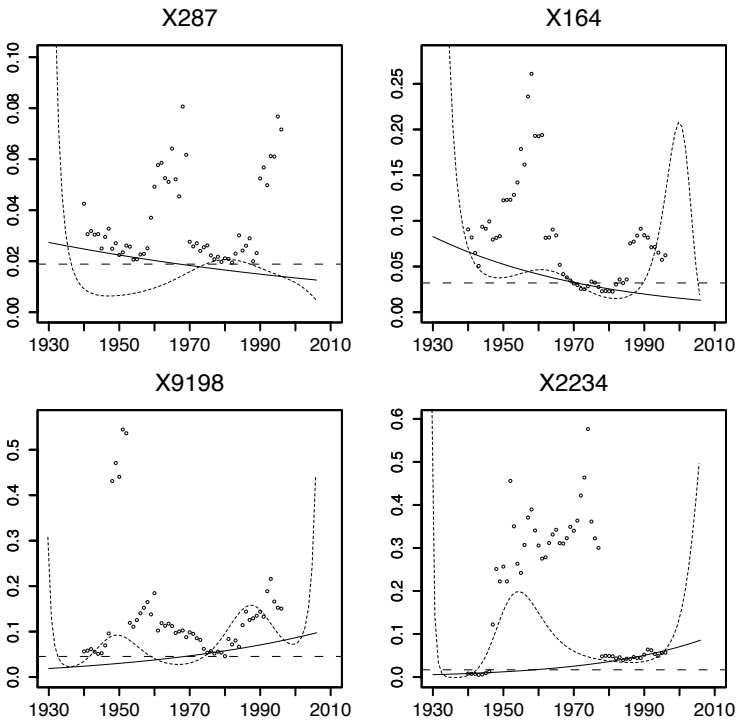

2234
9198
0287
0164
N
0 25 50 100 150 200 250
Km
43º 22' N
4º 30' W

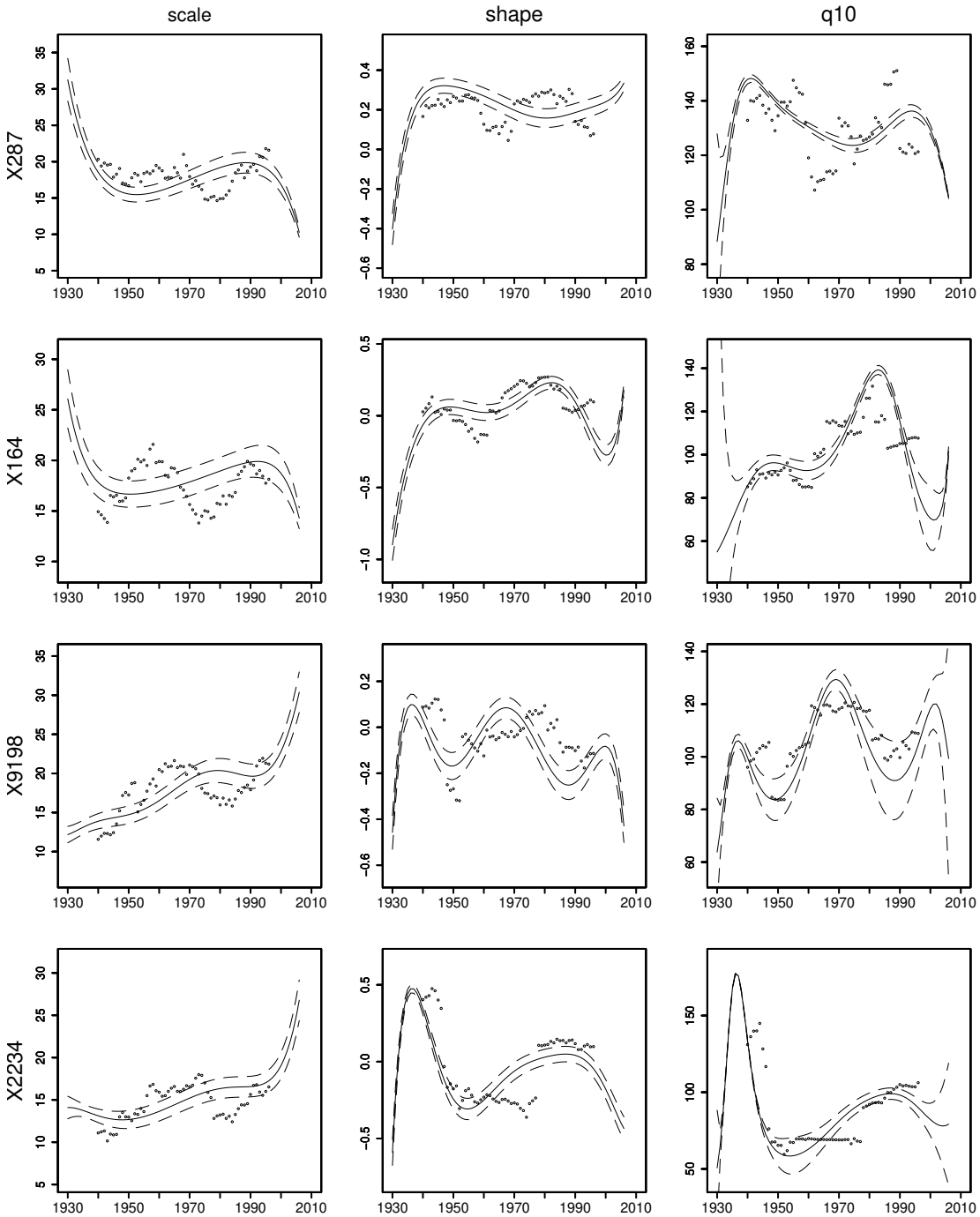

X287
X164
X9198
X2234
Mean Excess
Mean Excess

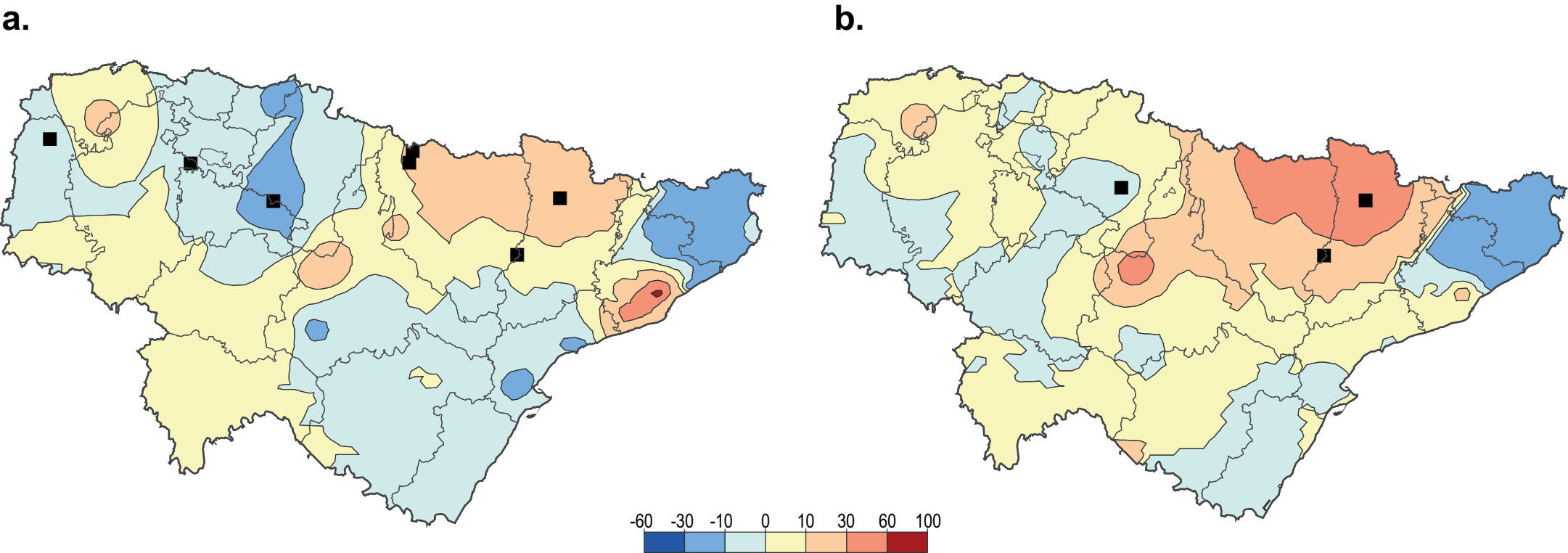

X287
X164
X9198
X2234
1940
1960
1980
2000


scale
shape
q10
X287
X164
X9198
X2234


scale
shape
q10
X287
X164
X9198
X2234
1930
1950
1970
1990
2010

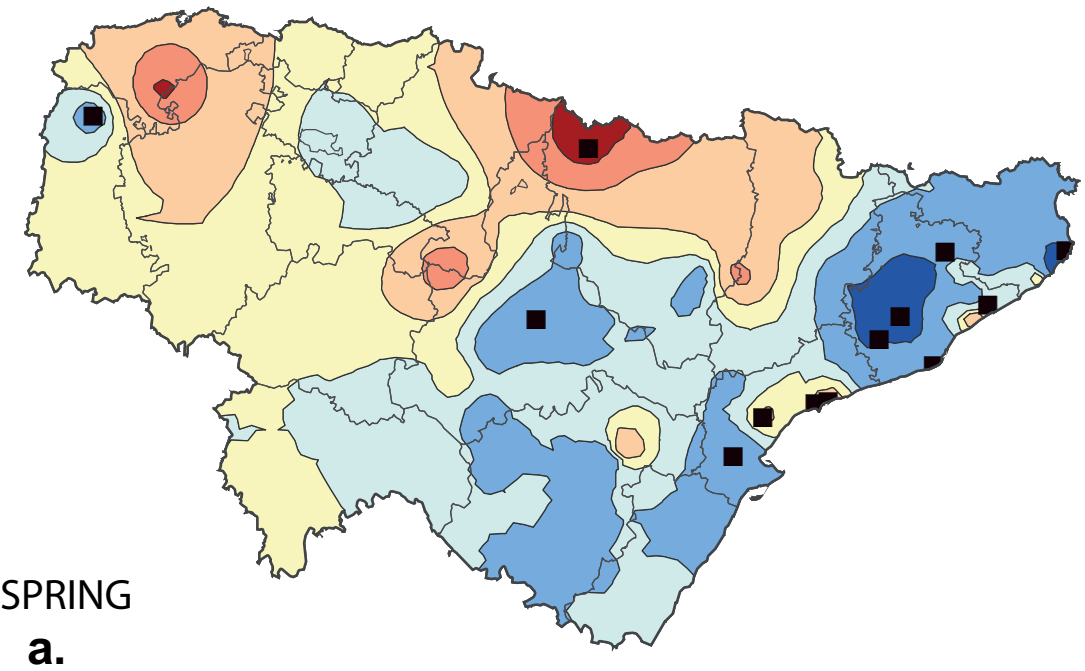

X287
X164
X9198
X2234
1930
1950
1970
1990
2010


scale
shape
q10
X287
X164
X9198
X2234

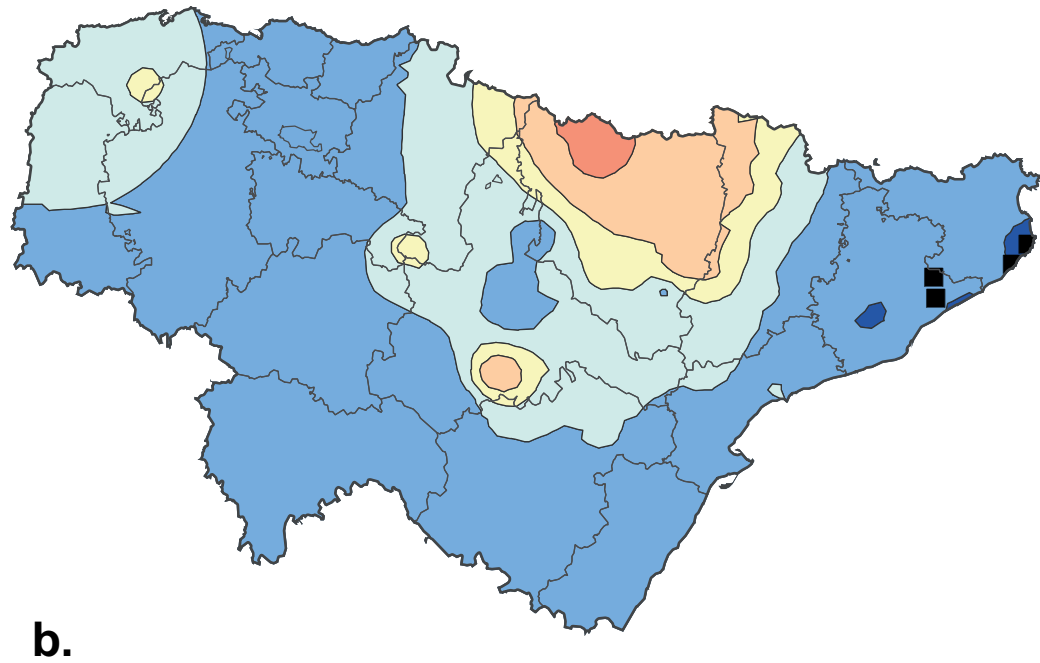

a.
b.
-60
-30
-10
0
10
30
60
100

WINTER

a.

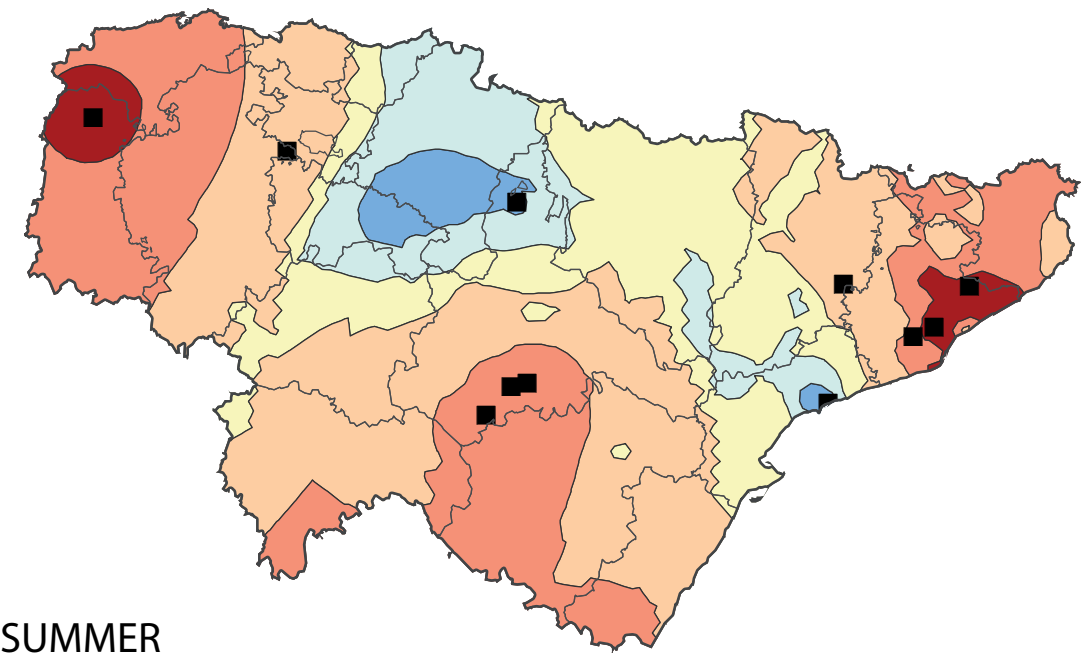


b.

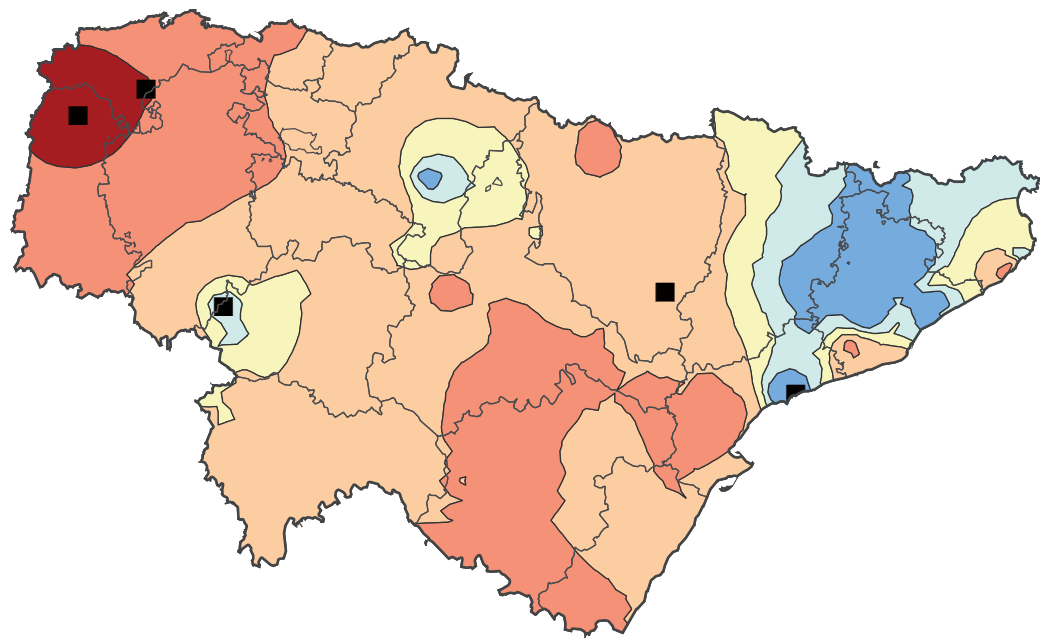


SPRING

a.

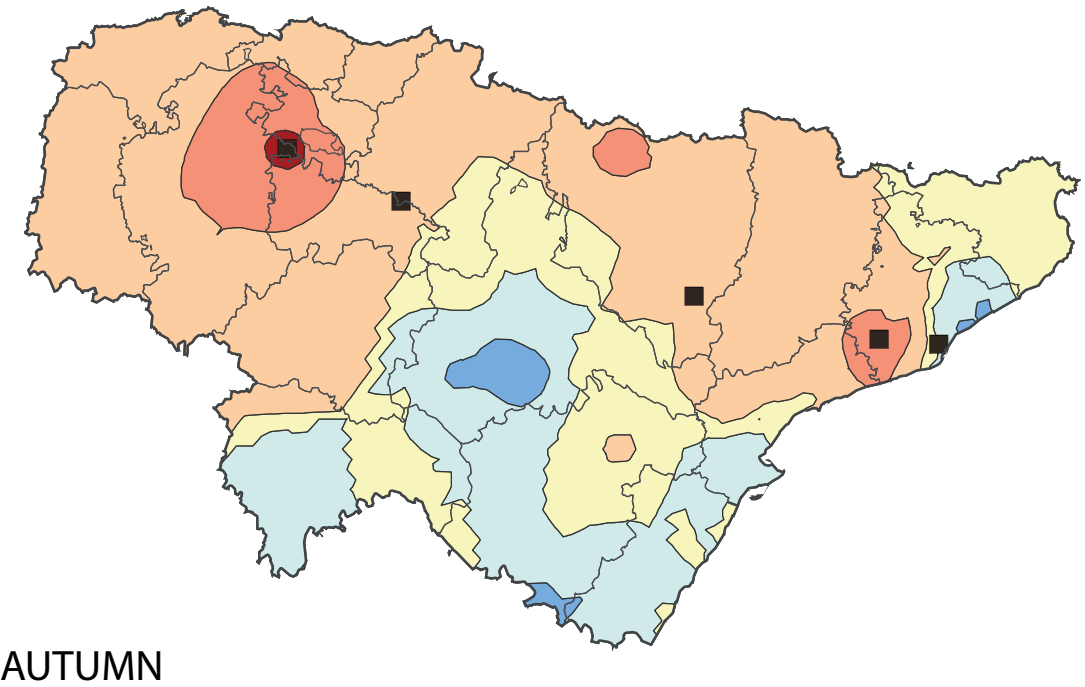


b.

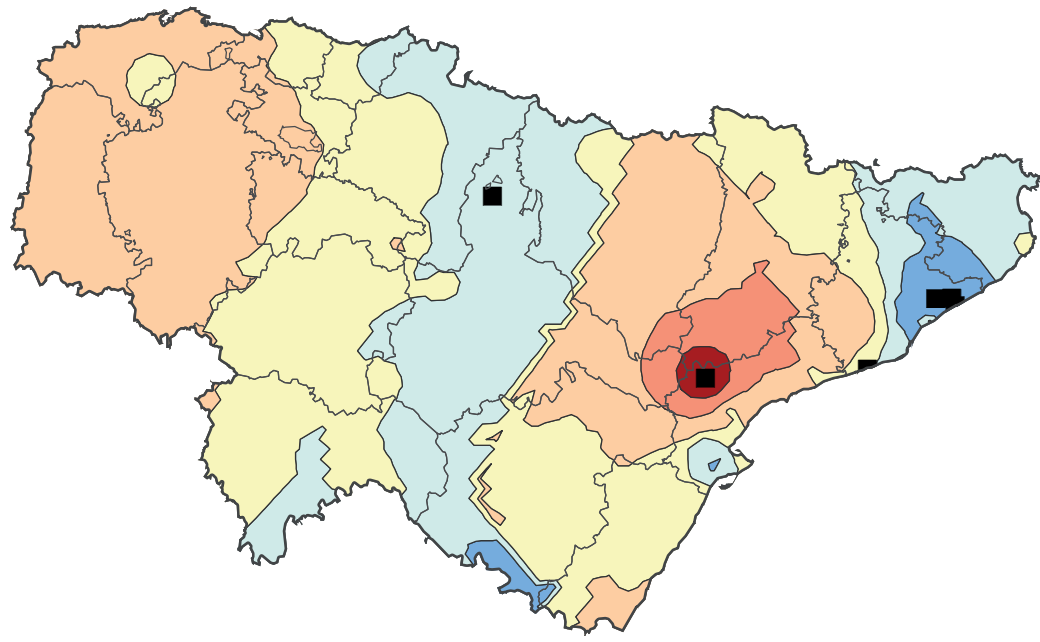


SUMMER

a.

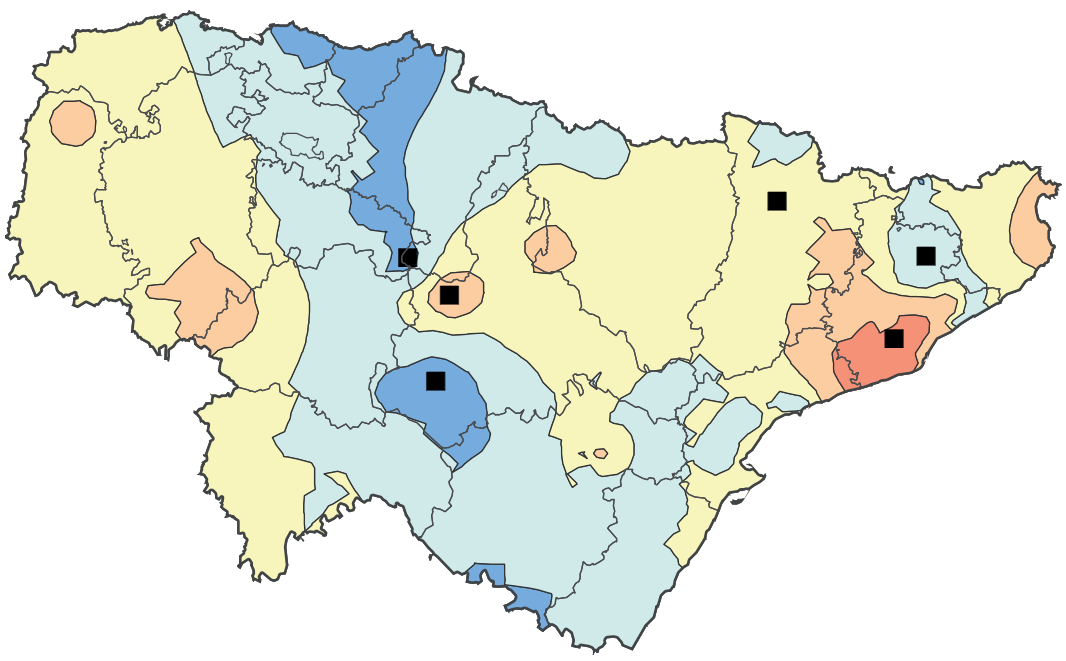
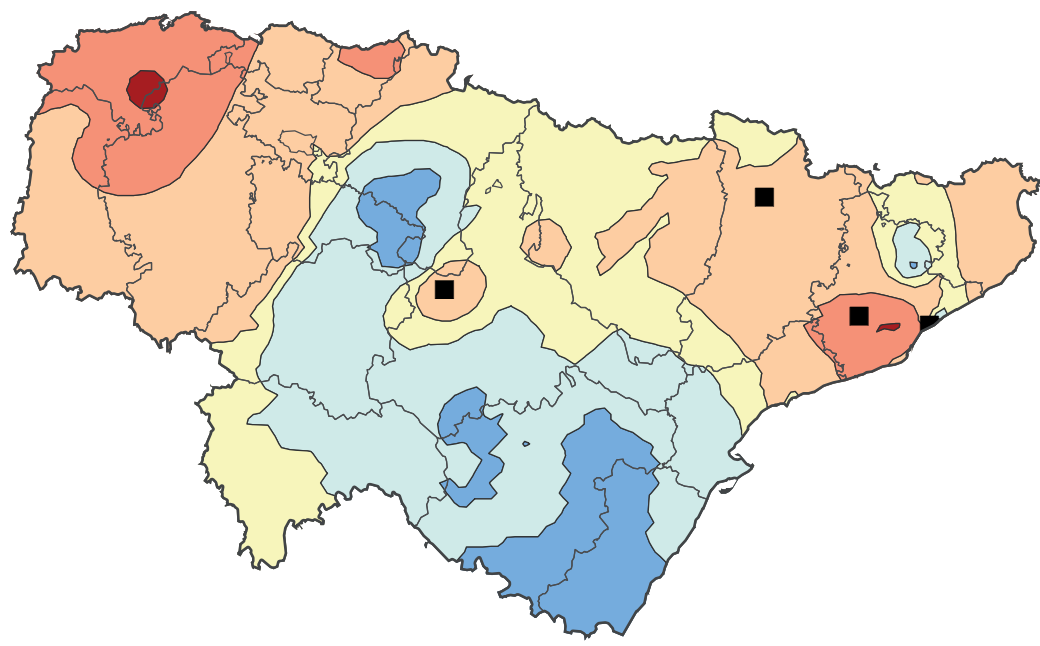
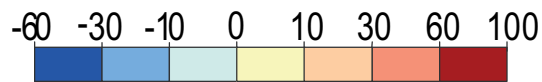


b.



AUTUMN

a.



b.