Hydrol. Earth Syst. Sci. Discuss., 8, 8609–8638, 2011
www.hydrol-earth-syst-sci-discuss.net/8/8609/2011/
doi:10.5194/hessd-8-8609-2011
© Author(s) 2011. CC Attribution 3.0 License.

Hydrology and Earth System Sciences Discussions

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Nonstationarities in the occurrence rates of flood events in Portuguese watersheds

A. T. Silva[1], M. M. Portela[1], and M. Naghettini[2]

[1]Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal
[2]Federal University of Minas Gerais, Belo Horizonte, Brazil

Received: 13 September 2011 – Accepted: 14 September 2011
– Published: 23 September 2011

Correspondence to: A. T. Silva (artur.tiago.silva@ist.utl.pt)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

An exploratory analysis on the variability of flood occurrence rates in Portuguese watersheds is made, to ascertain if that variability is concurrent with the principle of stationarity. A peaks-over-threshold (POT) sampling technique is applied to 10 long series of mean daily streamflows and to 4 long series of daily rainfall in order to sample the times of occurrence (POT time data) of the peak values of those series. The kernel occurrence rate estimator, coupled with a bootstrap approach, was applied to the POT time data to obtain the time dependent estimated occurrence rate curves, $\hat{\lambda}(t)$, of floods and extreme rainfall events. The results of the analysis show that the occurrence of those events constitutes an inhomogeneous Poisson process, hence the occurrence rates are nonstationary. An attempt was made to assess whether the North Atlantic Oscillation (NAO) casted any influence on the occurrence rate of floods in the study area. Although further research is warranted, it was found that years with a less-than-average occurrence of floods tend to occur when the winter NAO is in the positive phase, and years with a higher occurrence of floods (more than twice the average) tend to occur when the winter NAO is in the negative phase. The authors conclude that the mathematical formulation of the flood frequency models relying on stationarity commonly employed in Portugal, should be revised in order to account for nonstationarities in the occurrence rates of such events.

1 Introduction

Nowadays, there seems to be a consensus among the scientific community that, due to climate change, there is an intensification of the hydrological cycle (Bates et al., 2008), assumably leading to more frequent and more intensive extreme hydrological phenomena, like floods and droughts. Milly et al. (2008) argue that such climate change undermines stationarity – a basic assumption that historically has assisted practice and research in the fields of hydrology and water resources management.

HESSD
8, 8609–8638, 2011

CC BY

In Portugal, however, the generality of research on hydrological modeling relies on the principle of the stationarity of hydrological time series (Quintela and Portela, 2002), namely the academic research related to flood hydrology (Lança, 2000; Simões, 2002; Dias, 2008; Delgado, 2007), although general circulation models (GCMs) based on some scenarios of greenhouse gas emissions show a trend for an aggravation of extreme precipitation events in northern Portugal (Santos et al., 2002, p.167). In such framework it is important to carry out research that could ascertain whether or not the hydrological time series that are used in the design of water resources systems exhibit definite signs of nonstationarity. If they do, the mathematical formalism applied to the planning and management of such structures must be revised, namely the statistical analyses that are usually employed in studies on extreme hydrological phenomena, such as extreme rainfall events and floods.

The aim of the present study was to carry out an exploratory analysis on the variability of the streamflow regime in Mainland Portugal, in what concerns the occurrence rate of floods, and ascertain if that variability is concurrent with the hypothesis of nonstationarity.

Following this Introduction, in Sect. 2, the data set used in this study is presented. This paper focuses primarily on mean daily streamflow data at 10 gauging stations geographically spread over Mainland Portugal. Long series of daily rainfall records at rain gauges located in 4 of the studied watersheds were also utilized. The peaks-over-threshold (POT) sampling technique is applied to extract the dates (POT time data) and peak values (POT value data) of flood events and extreme rainfall events, respectively, from the streamflow and rainfall samples. A brief preliminary analysis is then carried out to characterize the multi-year variability of the sampled POT data. In Sect. 3, a nonparametric occurrence rate estimator, namely, the kernel technique (Diggle, 1985; Mudelsee et al., 2003) was applied to the POT time data sampled in Sect. 2 in order to estimate the flood occurrence rate, $\lambda(t)$, in the studied watersheds. This method, coupled with a bootstrap confidence band, provides a powerful and reliable characterization of the multi-year variability of the obtained flood occurrence rates (Mudelsee

HESSD
8, 8609–8638, 2011

et al., 2004).

The North Atlantic Oscillation (NAO) is a pattern in climate variability associated with a redistribution of atmospheric masses. Several studies have established links between the phases of the NAO and both rainfall and river flow in the western Iberian Peninsula in winter months (Corte-Real et al., 1998; Rodríguez-Puebla et al., 2001; Trigo et al., 2002b, 2004; Morán-Tejeda et al., 2011; Lorenzo-Lacruz et al., 2011). In a study more congruent with hydrological extremes, Trigo et al. (2005) has established links between the NAO and rainfall triggering of landslides in a region North of Lisbon. Given the aforementioned evidence of the role of the NAO in modulating rainfall and river flow in western Iberian watersheds, an exploratory analysis on the relationship between flood occurrence in the studied watersheds and the NAO is carried out in Sect. 4.

Finally, in Sect. 5, the most relevant conclusions of the article are drawn and opportunities of future research are presented.

2 Data

2.1 Streamflow and rainfall data

This paper focuses primarily on flood occurrence rates in 10 unregulated Portuguese catchments. The data set consists of two types of hydrological time series – identified in Table 1: (i) *mean daily streamflow* series at 10 stream gauging stations that define the analyzed catchments, geographically spread over Mainland Portugal; (ii) *daily rainfall* at rain gauges located in 4 of the studied watersheds. It should be mentioned that the period of records and number of years shown in those tables refer to hydrological years, which in Portugal begin on 1 October.

Figure 1 shows the location of the streamflow and rain gauging stations as well as the outline of the catchments under study, whose areas are shown in Table 1.

The utilized data set is somewhat representative of the data that is usually used by researchers and practitioners in hydrology and water resources systems planning and management in Portugal

It should be stressed that the research underlying this paper is primarily concerned with the mean daily streamflow data. The daily rainfall data was utilized for the purpose of checking whether there are significant discrepancies between peaks in rainfall and streamflow, which would eventually suggest that the behaviour exhibited by the streamflows was under a significant anthropogenic influence.

2.2 Peaks-over-threshold (POT) data

The *peaks-over-threshold* (POT) approach to hydrological frequency analysis, consists of applying a sampling technique to a time series that retains the peak values that exceed a given truncation level usually called *base level* or *threshold* (Lang et al., 1999). Todorovic and Zelenhasic (1970) and Todorovic (1978) introduced the marked point model for flood analysis that relies on the POT sampling technique.

Conventional analysis by the POT sampling technique consists of the following: from a continuous time series x_t, with t running continuously from t_0 to t_n, the i-th peak value x'_i corresponds to the largest value among those exceeding the threshold u, in the time interval $[t_0 + t_i, t_0 + t_i + \Delta t]$, comprised in $[t_0, t_n]$. Formally,

$$\{x'_i\} = \max \{x_t | x_t > u\}_{t=t_0+t_i}^{t=t_0+t_i+\Delta t_i} \tag{1}$$

If the time of occurrence of x'_i is denoted by $T_i \in [t_0, t_n]$ and if there are m flood episodes of this kind, $\{T_i, x'_i\}_{i=1}^{m}$ represents a marked point process for the variable X.

Usually, the peak values, x'_i, are supposed identically independently distributed (i.i.d.) variables. The time of occurrence of x'_i, that is T_i, is associated with a Poisson process, with the number of occurrences in a given interval (e.g. a year) being a Poisson variate with parameter λ (Cunnane, 1979). In the context of POT sampling of hydrological variables, the time series discretized in daily intervals, such as those

HESSD
8, 8609–8638, 2011

CC BY

employed here, though not rigorously continuous, conform to the general requirements of a Poisson point process, as previously defined.

The selected peaks must meet the independence condition. Several criteria has been presented in the literature in order to verify this hypothesis (Lang et al., 1999). The criterion used in the present study, for the mean daily flow series, (NERC, 1975; Cunnane, 1979) determines that peaks should be separated in time by three times the time to peak and, furthermore, that the flow between two consecutive peaks should decrease below as much as two thirds of the first peak. Due to the relatively small areas of the watersheds being analyzed (Table 1), as most of the Portuguese watersheds, times of concentration lower than 1 day are expected, except for Quinta das Laranjeiras – S10. For that reason and given the daily time-step of the data, we considered that for the purpose of applying the forementioned independence criterion, the time to peak equals approximately one day at the analyzed gauging stations.

The independence criterion applied to the daily rainfall series is similar to the one applied to the mean daily flow series, the only difference being that the daily rainfall must decrease to zero between two peaks.

The selection of the threshold, u, is a procedure that involves a great level of subjectivity (Beguería, 2005). Lang et al. (1999) reviewed a number of systematic approaches to carry out this selection. Lang et al. (1999) remark, however, that there is no universal and unequivocal method for selecting the threshold. In the present study, in what concerns the mean daily flows, the threshold adopted equals 7 times the long term mean daily flow, or modulus, which according to Quintela (1984) provides a lower limit to identify the flood occurrences. Figure 2a shows the mean daily flows sample at Fragas da Torre (S5) and the corresponding threshold. The criterion adopted to determine the threshold, as applied to the daily rainfall records, considered that the mean number of over-threshold values should be similar to the one obtained for the mean daily flows. Such criterion resulted in a threshold equal to 12 times the mean daily rainfall.

After the threshold u was defined, the correlation structure of the peaks or exceedances was analyzed, namely the lag-one and lag-two autocorrelation coefficients, as exemplified in Fig. 2b. No significant serial correlation was found that could invalidate the application of the selected thresholds.

Furthermore, tests were carried out to determine whether or not the thresholds were adequate: the mean number of over-threshold events in a year and the mean exceedance above threshold, tests no. 1 and 2, respectively, reviewed by Lang et al. (1999). The application of such tests did not invalidate the adopted threshold selection criterion. As an example, the results of the application of the tests to sample S5 – Fragas da Torre are presented in Fig. 2b and c, which show that the selected u is located in the domain where the mean number of exceedances is decreasing and the mean exceedance above threshold is approximately a linear function of u.

Furthermore it should be mentioned that, though it is not presented in this paper, the procedure, suggested by Davison (2003, p.286), of imposing slight variations of threshold levels, around the adopted ones, did not cause any significant changes to the overall results of Sects. 3 and 4, and to the final conclusions as well.

2.3 Preliminary data analysis

With the application of a POT sampling technique (Eq. 1), two types of data were obtained: *time data*, $\{T_i\}_{i=1}^{m}$, which are the instants of occurrence of extreme events, and *peak value data*, $\{x'_i\}_{i=1}^{m}$, which relates the magnitude of the events themselves.

Although this study concerns primarily with the forementioned time data, an exploratory analysis was made concerning the multi-year variability of the peak values, $\{x'(i)\}_{i=1}^{m}$, for both the mean daily flow and the daily rainfall data. This analysis was performed by obtaining the series of running average exceedances over successive periods of 15 years. Figure 3 shows those series, made dimensionless by their respective means, for both streamflow and rainfall data. In the horizontal axis of that figure, the year k refers to the hydrologic year $k/(k+1)$, which starts on 1 October of year k as previously mentioned. A visual analysis of Fig. 3 suggests that overall the samples do not exhibit trends or a significant multi-year persistences above or below

HESSD
8, 8609–8638, 2011

mean values. It should be stressed that the beginning and the ending dates of the running averages vary according to the sampling period of each case study.

A preliminary analysis was also carried out concerning the parameter λ (occurrence rate) which describes the Poisson processes that underly the POT approach. Any counting process with stationary independent increments, i.e. the distribution of the number of events occurring in any interval depending only on the length of the interval, is a homogeneous Poisson process with parameter λ. With the introduction of a time dependence in the process, the Poisson parameter becomes a time function, $\lambda(t)$. This time dependence is useful for studying nonstationarities in $\lambda(t)$. In the framework of this paper, $\lambda(t)$ denotes the occurrence rate of flood events (Mudelsee, 2010, p.249) when referring to mean daily flow, and the occurrence rate of extreme rainfall events when referring to daily rainfall.

When the process is homogeneous, λ is a constant and the variable $w_i = (T_i - t_0)/(t_n - t_o)$ (normalized times of events above threshold) is distributed as order statistics of random sample from the uniform distribution in the interval [0, 1]. A graphical analysis of that hypothesis can be performed by plotting the empirical distribution function of w_i, $\hat{F}(w)$, and checking if there are any significant departures from the uniform distribution $F(w) = w$, $0 \leq w \leq 1$ (Davison, 2003, p.277–278). Figure 2e exemplifies the former graphical analysis applied to sample S5 – Fragas da Torre where the solid grey lines indicate the significance for a Kolmogorov-Smirnov statistic at levels 5 % and 1 %. Such figure suggests that the rate of the process may be non-uniform. It was verified that the generality of the samples exhibited significant departures from the uniform distribution.

An additional evaluation of whether there was any significant increase or decrease in the intensity of those processes focused on the number of extreme events per hydrologic year, a discrete variable denoted by λ_k for the hydrologic year $k/(k+1)$. For each sample of both streamflow and rainfall data sets, λ_k was smoothed by means of a LOWESS (locally weighted regression) curve (Cleveland, 1979), using a smoothing parameter $f = 0.2$, which was the smallest f that enabled an adequate smoothing of

HESSD
8, 8609–8638, 2011

CC BY

the data. The former analysis is represented in Fig. 4, where, analogously to Fig. 3, the year k in the horizontal axis refers to the hydrologic year $k/(k+1)$.

Figure 4 shows that there is some variability in the occurrence rate of extremes in the observed samples. Most notably, there seems to be an increase in the intensity of the Poisson process in the period ranging from the late 1950s to the late 1960s in all of the samples that cover this period, though in some samples it is more pronounced than in others.

3 Nonparametric occurrence rate estimation

3.1 General remarks

In this section a nonparametric method for analyzing nonstationarities in the occurrence rate $\lambda(t)$ of floods and extreme rainfall events is presented - the *kernel occurrence rate estimation* method, or *kernel technique* (Diggle, 1985; Mudelsee et al., 2003). According to Mudelsee et al. (2004), who reviewed a number of different approaches to analyze trends in the occurrence of extreme events, the kernel technique is one of the most powerful approaches because it does not impose parametric restrictions, allows for nonlinear and nonmonotonic trends, and, coupled with a bootstrap approach, provides reliable confidence bands, which are essential for interpreting the significance of apparent trends in $\lambda(t)$.

The kernel technique is applied to the POT time data sampled from the original mean daily flow and daily rainfall time series. After the estimated flood and extreme rainfall occurrence rate curves, $\hat{\lambda}(t)$, are obtained, a pointwise bootstrap confidence band is constructed around them, allowing for a more rigorous interpretation of the results.

HESSD
8, 8609–8638, 2011

3.2 Kernel occurrence rate estimation

The kernel technique is a nonparametric method developed by Diggle (1985) for smoothing point process data. For estimating the intensity of a point process such as the time-dependent occurrence rate, $\lambda(t)$, this technique may be formulated as:

$$\hat{\lambda}(t) = h^{-1} \sum_{i=1}^{m} K\left(\frac{t - T_i}{h}\right) \quad (2)$$

where K is the kernel function and h is the bandwidth. A Gaussian kernel, $K(y) = (2\pi)^{(-1/2)} \exp(-y^2/2)$, was used as it can be efficiently calculated in Fourier space and yields a smooth estimated occurrence rate, $\hat{\lambda}(t)$ (Mudelsee et al., 2004, 2006). The units of $\hat{\lambda}(t)$ are d^{-1}, i.e. the number of occurrences above threshold per day at a given point in time, t. However, to facilitate the interpretation of the results, $\hat{\lambda}(t)$ was multiplied by 365.25, such that, for a given instant, t, the $\hat{\lambda}(t)$ indicates the estimated number of occurrences above threshold per year (yr^{-1})

The application of Eq. (2) may lead to a boundary bias near t_0 and t_n consisting of an underestimation of $\lambda(t)$ due to the nonexistence of data outside the interval $[t_0, t_n]$. This boundary effect can be reduced by generating *pseudodata*, i.e. pseudo extreme events, pT, outside of the observation interval, before estimating $\lambda(t)$. The straightforward method of *reflection* was used to generate pseudodata (on the left side, for $t < t_0$: $pT(i) = t_0 - [T_i - t_0]$, covering an amplitude of 3 times h before t_0; analogously on the right side, for $t > t_n$). Pseudodata generation is equivalent to the extrapolation of the empirical distribution of events near the boundaries, hence the estimation of $\lambda(t)$ near the boundaries of the observation period should be analyzed with caution (Mudelsee et al., 2004; Mudelsee, 2010, p.251; Mudelsee, 2011). Considering $T^\dagger$ as the original point data, T, augmented by the pseudodata, pT, and $m^\dagger$ as the total number of points in $T^\dagger$, Eq. (2) can be rewritten as:

HESSD
8, 8609–8638, 2011

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

$$\hat{\lambda}(t) = h^{-1} \sum_{i=1}^{m^\dagger} K\left(\frac{t - T_i^\dagger}{h}\right). \quad (3)$$

Figure 5 shows the estimated occurrence rates $\lambda(t)$ for the S8 – Ponte Juncais sample, with and without pseudodata generation, exemplifying the correction of the boundary bias via pseudodata generation.

The selection of the bandwidth, h, determines the bias and variance properties of the occurrence rate estimator $\hat{\lambda}(t)$: a small h results in fewer data points that effectively contribute to the kernel estimation, which leads to a reduced bias and a high variance; on the other hand, a large h leads to an *oversmoothing* of the estimator, resulting in a small variance and increased bias. The selection of the bandwidth can be seen as a compromise between those two cases. Furthermore, since there is a high seasonal variability of the hydrologic regime in Portugal and, as the objective is to describe multi-year variability of $\lambda(t)$, the bandwidth should be considerably higher than the year to avoid the effect of the seasonal variability.

In this study the selection of the bandwidth used a straightforward method: Silverman's *rule of thumb* (Silverman, 1986, p.48). This rule defines h as:

$$h = 0.9\, A^{-1/5} \quad (4)$$

with:

$$A = \min\{\mathrm{STD}, \mathrm{IQR}\} \quad (5)$$

where STD and IQR are, respectively, the standard deviation and the interquartile range of the POT time data. Silverman (1986, p.48) comments that this method produces an adequate choice of bandwidth for many purposes. The bandwidths obtained range from 1141 days (sample S6) to 2157 days (sample S9), which is adequate for describing multi-year variability, because it smooths all the within-the-year seasonal variability.

HESSD
8, 8609–8638, 2011

CC BY

3.3 Bootstrap confidence band

As described in Sect. 3.2, a nonparametric method for estimating a varying point process intensity, the kernel method (Diggle, 1985), was used to characterize the occurrence rate of floods in the watersheds analyzed in this study. Point estimates may be difficult to interpret without some measurement of the uncertainties associated with those estimates.

For the purpose of quantifying the uncertainties associated with the results of Eq. (2), a pointwise confidence band was constructed around $\hat{\lambda}(t)$, by means of bootstrap simulations (Cowling et al., 1996; Mudelsee, 2011). This procedure consisted of drawing a set of n flood dates from the original POT data, augmented by pseudodata, with replacement, and calculating $\hat{\lambda}^{*}(t)$ after Eq. (2), using the resampled data and the same bandwidth, h. The resampling-estimation procedure was repeated until 2000 estimated curves $\hat{\lambda}^{*}(t)$ were obtained. Finally, an algorithm developed by Cowling et al. (1996) and used by Mudelsee et al. (2003, 2004) and Mudelsee (2011), namely the percentile-t type confidence band was applied pointwise to the 2000 $\hat{\lambda}^{*}(t)$ to construct a 90 % bootstrap confidence band around $\hat{\lambda}(t)$.

3.4 Results

The techniques described in Sects. 3.2 and 3.3 were applied to the POT time data sampled from the mean daily streamflow series and the daily rainfall series to obtain, respectively, estimated flood occurrence rates at the streamflow gauging stations, and, correspondingly, estimated extreme rainfall events occurrence rates at the rain gauging stations.

The results are presented in Fig. 6, which shows that the occurrence rates of extreme events in both mean daily flow and daily rainfall time series exhibit significant multi-year variability. For example: in graph S2 of Fig. 6, the peak of $\hat{\lambda}(t)$ in the 1960s is significantly higher than the upper limit of the confidence band at 1990.

HESSD
8, 8609–8638, 2011

The results of Fig. 6 also show that there are some trends in the intensity of the inhomogeneous Poisson process that are exhibited consonantly among the analyzed mean daily streamflow samples, such as: (a) a peak in $\hat{\lambda}(t)$ in the early 1960s is visible in all the graphs with available data in those years, which indicates a higher frequency of extreme events; (b) in graphs S8, S9 there is a peak in the 1930s followed by a decrease in intensity until a minimum is reached in the late 1940s (also visible in graph S7); (c) the graphs with data until the late 1990s and 2000s exhibit lower occurrence rates in the more recent years.

The fact that such trends are visible in mean daily flow series from unregulated rivers that are geographically distributed around the Portuguese territory (Fig. 1), suggests that those trends are not due to possible anthropogenic influences in the watersheds. The $\hat{\lambda}(t)$ estimates obtained for the rainfall data corroborate this hypothesis since they exhibit some of the same trends that are visible in the flow data of the catchments under the influence of those particular rainfall gauging stations (the rainfall-flow influence relationships are visible in Fig. 1: the rainfall in station R1 has influence in streamflows at S6, R2 at S3, R3 at S4, and R4 at S2 and S5).

4 Relationship between flood occurrence and the North Atlantic Oscillation

The North Atlantic Oscillation (NAO) is a prominent and recurrent pattern in climate variability of the Northern Hemisphere, which refers to a redistribution of atmospheric masses between the Arctic and the subtropical Atlantic (Hurrell et al., 2003). Studies carried out by Hurrell (1995) and Trigo et al. (2002a) have established links between the NAO phase and precipitation in western Europe. Although the NAO is evident throughout the year, its activity and impact on European surface climate is greater during the winter season (Osborn et al., 1999; Morán-Tejeda et al., 2011). There is also a number of studies on the influence of the NAO on precipitation and river flow in the western Iberian Peninsula in winter months (Rodríguez-Puebla et al., 2001; Corte-Real et al., 1998; Trigo et al., 2002b, 2004; Morán-Tejeda et al., 2011; Lorenzo-Lacruz

HESSD
8, 8609–8638, 2011

CC BY

et al., 2011), which have shown that when the NAO is in its negative phase precipitation and river flows tend to be above normal and vice-versa. Trigo et al. (2005) studied the influence of the NAO on monthly and seasonal precipitation on an area prone to landslides located north of the capital of Portugal, Lisbon, and concluded that the large inter-annual variability of winter precipitation in that region is largely modulated by the NAO mode.

Traditionally, the NAO index has been defined as the difference in normalized surface pressure between Iceland (Stykkisholmur) and the Azores archipelago (Ponta Delgada). In recent decades, however, researchers have found that, during the winter season, stations located in the Iberian Peninsula could be used with some advantages over Ponta Delgada: Lisbon was used by Hurrell (1995) as the southern station, and Jones et al. (1997) used Gibraltar. In this work it was decided to use the Iceland-Gibraltar index developed by the Climate Research Unit (Jones et al., 1997). Also, the winter season has been defined as November to March (NDJFM) in accordance with Jones et al. (1997) and Trigo et al. (2005), although other definitions of the winter season are also used (Osborn et al., 1999). Figure 7 shows the winter NAO indices used in this study along with a LOWESS (locally weighted regression) curve (Cleveland, 1979) fitted to the data using a smoothing parameter $f = 0.1$ in order to obtain an adequate multi-year smoothing.

The objective of this section is to ascertain whether or not the phase of the NAO has an influence on the occurrence rate of floods in the studied watersheds. By visually comparing Figs. 7 and 6 it is apparent that the peak in the estimated occurrences rate in the 1960s corresponds to a prolonged NAO negative phase in the same years. However, in the 1930s, where there are peaks in $\hat{\lambda}(t)$ in graphs S8 and S9 of Fig. 6, there is no similar negative phase in the winter NAO indices.

A quantitative analysis of the relationship between the winter NAO indices and the occurrence of floods was made on a discrete annual time basis. Figure 8 shows, for each streamflow sample, the number of floods per hydrologic year, λ_k (discrete POT time data, Fig. 4), plotted against the winter NAO indices of the corresponding years.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

8, 8609–8638, 2011

Nonstationarities in the occurrence rates of flood events

A. T. Silva et al.

cc BY

Such results suggest that years with positive NAO indices have a lower number of floods than years with negative NAO indices. Although that correlation does not seem to be particularly strong, Fig. 8 clearly shows that for every analyzed sample: (i) the majority of years without floods have positive NAO indices, and (ii) the years with the highest flood occurrence do not occur in positive NAO phases.

Figure 9 shows the same results as the previous analysis, but here λ_k data from all the samples were made dimensionless by their respective mean values $\overline{\lambda}$, and plotted together against the winter NAO indices. In that figure, box plots were drawn to represent the dispersion of the NAO indices for the values of λ_k that are lower than the mean ($\lambda_k/\overline{\lambda} < 1$); between one and two times the mean ($1 \leq \lambda_k/\overline{\lambda} < 2$); and higher than two times the mean ($\lambda_k/\overline{\lambda} \geq 2$). The results clearly show that (i) years with a less-than-average number of floods tend to occur when the NAO is in a positive phase, and (ii) years with a higher number of floods (more that twice the average) tend to occur when the NAO is in a negative phase.

Although the relationship between the NAO and the occurrence of floods in Portuguese watersheds requires further investigation, the results of Figs. 8 and 9, and the comparison of Figs. 6 and 7, indicate that an increase in the rate of flood occurrence might be related to a decrease of NAO indices and vice-versa. That accordance is not strong enough to establish a cause-and-effect type of relationship. However it does merit to be scoped in future research.

5 Conclusions and future research

The current consensus on the effects of climate change on the hydrological cycle compromises the stationarity of hydrological time series. The generality of the research on hydrological modelling in Portugal, and particularly the research involving flood hydrology, relies on the stationary principle. The objective of the research underlying

HESSD

8, 8609–8638, 2011

CC BY

this paper was to make an exploratory analysis on the variability of flood occurrence rates in Portuguese watersheds, and to ascertain if that variability is concurrent with the principle of stationarity.

For that purpose, a peaks-over-threshold (POT) sampling technique was applied to 10 long series of mean daily streamflows and to 4 long series of daily rainfall in order to sample the times of occurrence (POT time data) of the peak values of those series. A preliminary analysis of the POT time data suggested that the occurrence rates of those events were non-uniform. The kernel occurrence rate estimator (Diggle, 1985; Mudelsee et al., 2003), coupled with a bootstrap approach, was applied to the POT time data to obtain the time dependent estimated occurrence rate curves, $\hat{\lambda}(t)$, of floods and extreme rainfall events.

The achieved results clearly show that, in the studied watersheds, the occurrence of floods constitutes an inhomogeneous Poisson process, hence the flood occurrence rates are nonstationary. There is a number of similarities in the behaviour of the estimated flood occurrence rates among the various samples, such as a peak in flood occurrence rates in the 1960s and the 1930s. Such peaks are also visible in the results of the application of the kernel technique to the daily rainfall data, from rain gauges located within 4 of the analyzed watersheds. The similarities in the behaviour of $\hat{\lambda}(t)$ among different watersheds that are geographically spread over the Portuguese territory, and between rainfall and streamflow suggests that the observed trends are inherent to the natural multi-year variation of the hydrological cycle, as opposed to potential anthropogenic influence on the catchments themselves. Furthermore, as mentioned in Sect. 2.2, the observed trends in the estimates of $\lambda(t)$ are robust against threshold selection in the POT sampling of extreme value data.

There has been a number of research papers that link the phases of the North Atlantic Oscillation to the climate variability in western Iberia in general and in Portugal in particular (Corte-Real et al., 1998; Trigo et al., 2002b, 2004, 2005; Morán-Tejeda et al., 2011). An attempt was made to assess whether or not the NAO index casted any influence on the flood occurrence rates in Portuguese watersheds. This was done

HESSD
8, 8609–8638, 2011

by comparing the winter (NDJFM) NAO index with the number of floods within the hydrologic year, λ_k. The results of that analysis show that, although the correlation is not particularly strong, the years with less floods tend to happen when the winter NAO is in its negative phase, and vice-versa. This subject does, however, merit further research, namely on the possibility of establishing projections of the flood occurrence rates if reliable long-range forecasts of the NAO are made available (see Sutton and Allen, 1997; Rodwell et al., 1999).

In conclusion, the results presented in this paper suggest that the mathematical formulation of the flood frequency models relying on stationarity, as those more commonly applied in Portugal, should be revised in view of accounting for nonstationarities in the occurrence rate of floods and extreme rainfall events.

Overall the obtained results lay the foundation for future research on nonstationary hydrological modelling of floods in Portuguese watersheds, such as a hybrid Poisson – extreme value distribution model (Mudelsee, 2011, p.264) which encompasses a nonparametric description of the time dependence via the inhomogeneous Poisson process and a parametric extreme value distribution. Such modeling of extreme hydrological phenomena, coupled with a potential dependence on climate patterns such as the NAO, could be of use to researchers and practitioners that deal with uncertainty in water resources systems planning and management, associated with hydrological extremes.

Acknowledgements. The study was supported by the project CapWEM (Capacity development in Water Engineering and Environmental Management DCI-ALA/19.09.01/10/21526/254-922/ALFA III (2010)55), financed by the European Commission (ALFA III Programme); and the Brazilian National Council of Scientific and Technological Development (CNPq), through grant numbered 201526/2010-7 (PDE). The authors also wish to thank João Côrte-Real of the University of Évora, whose insights are gratefully acknowledged.

HESSD

8, 8609–8638, 2011

CC BY

References

Bates, B., Kundzewicz, Z., Wu, S., and Palutikof, J.: Climate change and water, Tech. rep., Intergovernmental Panel on Climate Change, 2008. 8610

Beguería, S.: Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value, J. Hydrol., 303, 215–230, 2005. 8614

Cleveland, W.: Robust locally weighted regression and smoothing scatterplots, J. Am. Stat. Assoc., 74, 829–836, 1979. 8616, 8622

Corte-Real, J., Qian, B., and Xu, H.: Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation, Int. J. Climatol., 18, 619–635, 1998. 8612, 8621, 8624

Cowling, A., Hall, P., and Phillips, M.: Bootstrap confidence regions for the intensity of a Poisson point process, J. Am. Stat. Assoc., 91, 1516–1524, 1996. 8620

Cunnane, C.: A note on the Poisson assumption in partial duration series models, Water Resour. Res., 15, 489–494, 1979. 8613, 8614

Davison, A. C.: Statistical models, Cambridge Univ Pr, 2003. 8615, 8616

Delgado, J. M. M.: Flood discharge evaluation under scarcity of data, Master's thesis, Instituto Superior Técnico, 2007. 8611

Dias, A. T. G. d. S.: Caudais instantâneos máximos anuais em Portugal Continental, Proposta de regionalização, Master's thesis, Instituto Superior Técnico, 2008. 8611

Diggle, P.: A kernel method for smoothing point process data, J. Roy. Stat. Soc. C, 34, 138–147, 1985. 8611, 8617, 8618, 8620, 8624

Hurrell, J.: Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation, Science, 269, 676, 1995. 8621, 8622

Hurrell, J., Kushnir, Y., Ottersen, G., and Visbeck, M.: An overview of the North Atlantic oscillation, in: The North Atlantic Oscillation: Climatic Significance and Environmental Impact, Geophysical Monograph, edited by: Hurrell, J., Kushnir, Y., Ottersen, G., and Visbeck, M., AGU American Geophysical Union, 134, 1–36, 2003. 8621

Jones, P., Jonsson, T., and Wheeler, D.: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland, Int. J. Climatol., 17, 1433–1450, 1997. 8622

Lança, R. M. M.: Contribuição para o estudo de cheias recorrendo a um modelo distribuído, Master's thesis, Universidade de Évora, 2000. 8611

CC BY

Lang, M., Ouarda, T., and Bobée, B.: Towards operational guidelines for over-threshold modeling, J. Hydrol., 225, 103–117, 1999. 8613, 8614, 8615
Lorenzo-Lacruz, J., Vicente-Serrano, S. M., López-Moreno, J. I., Gonz'alez-Hidalgo, J. C., and Morán-Tejeda, E.: The response of Iberian rivers to the North Atlantic Oscillation, Hydrol. Earth Syst. Sci., 15, 2581–2597, doi:10.5194/hess-15-2581-2011, 2011. 8612, 8621
Milly, P., Betancourt, J., Falkenmark, M., Hirsch, R., Kundzewicz, Z., Lettenmaier, D., and Stouffer, R.: Stationarity Is Dead: Whither Water Management?, Science, 319, 573, 2008. 8610
Morán-Tejeda, E., Ignacio, L., Antonio, C., and Sergio, M.: Evaluating Duero's basin (Spain) response to the NAO phases: spatial and seasonal variability, Hydrol. Process., 25, 1313–1326, 2011. 8612, 8621, 8624
Mudelsee, M.: Climate Time Series Analysis: Classical Statistical and Bootstrap Methods, Springer, Dordrecht, 2010. 8616, 8618
Mudelsee, M.: The Bootstrap in Climate Risk Analysis, in: In Extremis: Disruptive Events and Trends in Climate and Hydrology, edited by: Kropp, J. P. and Schellnhuber, H. J., Springer, 45–58, 2011. 8618, 8620, 8625
Mudelsee, M., Börngen, M., Tetzlaff, G., and Grünewald, U.: No upward trends in the occurrence of extreme floods in central Europe, Nature, 425, 166–169, 2003. 8611, 8617, 8620, 8624
Mudelsee, M., Börngen, M., Tetzlaff, G., and Grünewald, U.: Extreme floods in central Europe over the past 500 years: Role of cyclone pathway "Zugstrasse Vb", J. Geophys. Res., 109, D23101, doi:10.1029/2004JD005034, 2004. 8611, 8617, 8618, 8620
Mudelsee, M., Deutsch, M., Börngen, M., and Tetzlaff, G.: Trends in flood risk of the River Werra (Germany) over the past 500 years, Hydrolog. Sci. J., 51, 818–833, 2006. 8618
NERC: Flood Studies Report, vol. I, Natural Environment Research Council, 1975. 8614
Osborn, T., Briffa, K., Tett, S., Jones, P., and Trigo, R.: Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model, Clim. Dynam., 15, 685–702, 1999. 8621, 8622
Quintela, A. C.: Curso Internacional de Hidrologia Operativa, vol. 2, Direcção-Geral dos Recursos e Aproveitamentos Hidráulicos, 1984. 8614
Quintela, A. C. and Portela, M. M.: A modelação hidrológica em Portugal nos últimos 25 anos do século XX nas perspectivas determinística e estocástica, Recursos Hídricos, 23, 7–22, 2002. 8611

HESSD
8, 8609–8638, 2011

Nonstationarities in the occurrence rates of flood events

A. T. Silva et al.

CC BY

Rodríguez-Puebla, C., Encinas, A. H., and Sáenz, J.: Winter precipitation over the Iberian peninsula and its relationship to circulation indices, Hydrol. Earth Syst. Sci., 5, 233–244, doi:10.5194/hess-5-233-2001, 2001. 8612, 8621

Rodwell, M., Rowell, D., and Folland, C.: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate, Nature, 398, 320–323, 1999. 8625

Santos, F., Forbes, K., and Moita, R.: Climate change in Portugal: scenarios, impacts and adaptation measures: SIAM Project, Gradiva, 2002. 8611

Silverman, B.: Density estimation for statistics and data analysis, vol. 26, Chapman & Hall/CRC, 1986. 8619

Simões, J. M, D. A.: Modelação distribuída de cheias utilizando suporte raster, Master's thesis, Instituto Superior Técnico, 2002. 8611

Sutton, R. and Allen, M.: Decadal predictability of North Atlantic sea surface temperature and climate, Nature, 388, 563–567, 1997. 8625

Todorovic, P.: Stochastic models of floods, Water Resour. Res., 14, 345–356, 1978. 8613

Todorovic, P. and Zelenhasic, E.: A stochastic model for flood analysis, Water Resour. Res., 6, 1641–1648, 1970. 8613

Trigo, R., Osborn, T., and Corte-Real, J.: The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms, Climate Res., 20, 9–17, 2002a. 8621

Trigo, R., Osborn, T. R., and Corte-Real, J.: Influência da Oscilação do Atlântico Norte no Clima do Continente Europeu e no Caudal dos Rios Ibéricos Atlânticos, Finisterra, 37, 5–31, 2002b. 8612, 8621, 8624

Trigo, R., Pozo-Vázquez, D., Osborn, T., Castro-Díez, Y., Gámiz-Fortis, S., and Esteban-Parra, M.: North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula, Int. J. Climatol., 24, 925–944, 2004. 8612, 8621, 8624

Trigo, R., Zezere, J., Rodrigues, M., and Trigo, I.: The influence of the North Atlantic Oscillation on rainfall triggering of landslides near Lisbon, Nat. Hazards, 36, 331–354, 2005. 8612, 8622, 8624

HESSD

8, 8609–8638, 2011

Nonstationarities in the occurrence rates of flood events

A. T. Silva et al.

Table 1. Mean daily streamflow and daily rainfall data. Sample code, name, period of records and catchment area of the 10 streamflow gauging stations.

Mean daily streamflow samples			
Code	Name	Period of records (no. of years)	Catchment area (km^2)
S1	Castelo Bom	1957/58–2003/04 (47)	897
S2	Castro Daire	1945/46–2003/04 (59)	291
S3	Cunhas	1949/50–2005/06 (57)	338
S4	Ermida Corgo	1956/57–2005/06 (50)	291
S5	Fragas da Torre	1946/47–2005/06 (60)	660
S6	Monte da Ponte	1959/60–1991/92 (33)	701
S7	Odivelas	1934/35–1966/67 (33)	431
S8	Ponte Juncais	1918/19–1974/75 (57)	604
S9	Ponte Sta. Clara-Dão	1921/22–1972/73 (52)	177
S10	Quinta das Laranjeiras	1942/43–2005/06 (64)	3464
Daily rainfall samples			
Code	Name	Period of records (no. of years)	
R1	Almodôvar	1959/60–1999/00 (61)	
R2	Alturas do Barroso	1946/47–1995/96 (50)	
R3	Campeã	1959/60–1994/95 (36)	
R4	Castro Daire	1916/17–2000/01 (85)	

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

8, 8609–8638, 2011

Nonstationarities in the occurrence rates of flood events

A. T. Silva et al.

cc BY

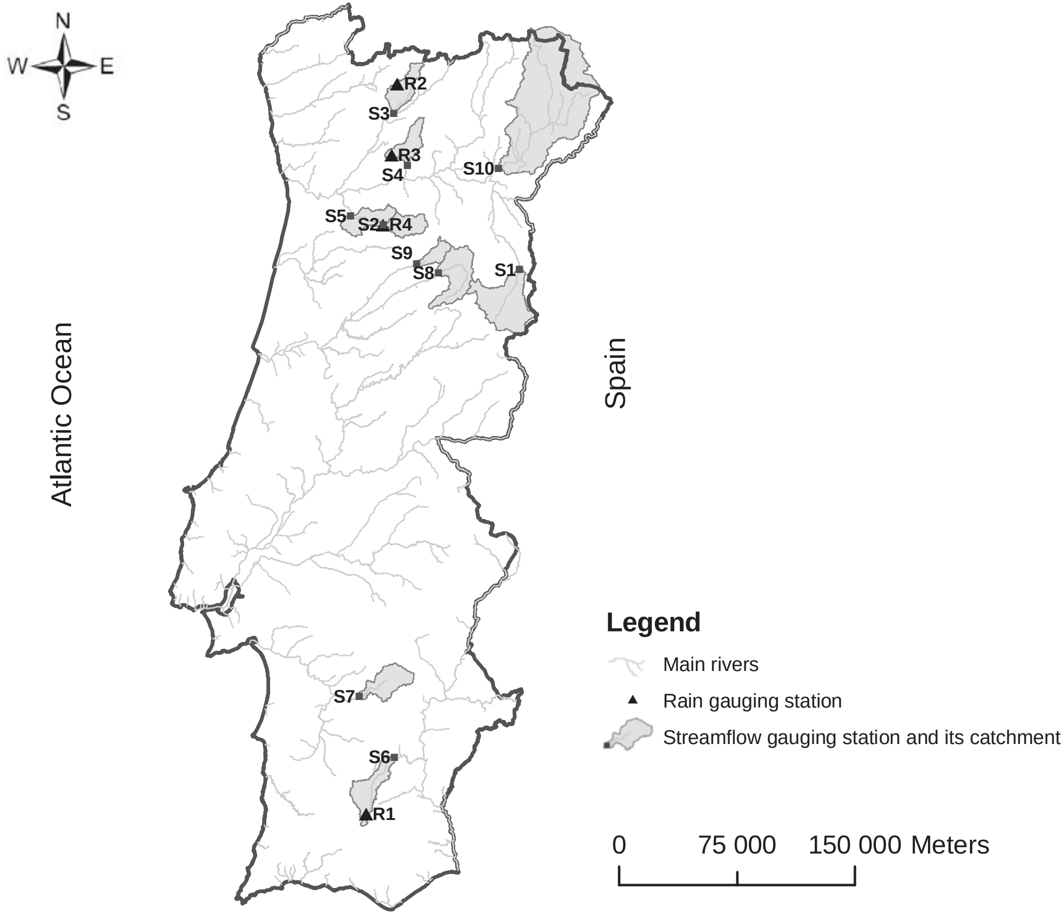

Fig. 1. Map of Portugal. Location of the rainfall and streamflow gauging stations, as identified by the codes presented in Table 1. The shaded areas correspond to the catchments under study.

HESSD

8, 8609–8638, 2011

cc BY

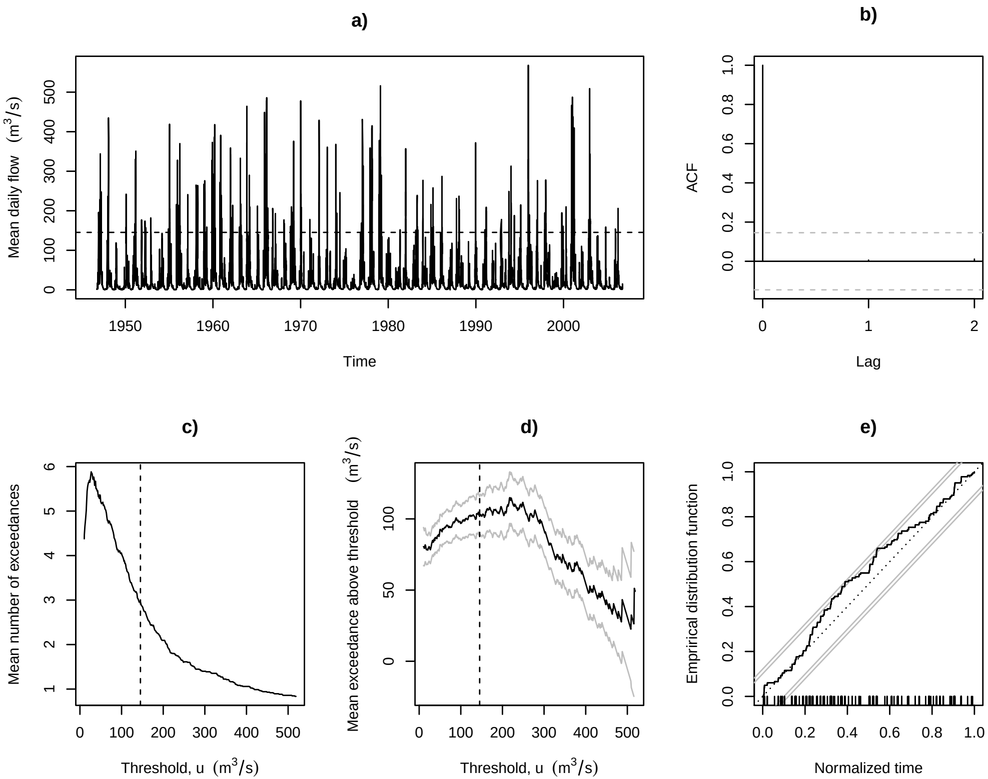

Fig. 2. Peaks-over-threshold sampling technique applied to sample S5 – Fragas da Torre. **(a)** Mean daily flow sample, **(b)** autocorrelation function (ACF) of the exceedances with a 95 % confidence band, **(c)** mean number of exceedances in a year, **(d)** mean exceedance over the threshold with a 95 % confidence band, and **(e)** empirical distribution function of the times of the events. The selected threshold is identified by the dashed lines in panels **(a)**, **(c)** and **(d)**. In panel **(e)**, the vertical ticks show the points in time associated with the flood occurrences, the dotted diagonal line shows the uniform distribution and the solid diagonal lines show significance for a Kolmogorov-Smirnov statistic at 5 % and 1 %.

HESSD
8, 8609–8638, 2011

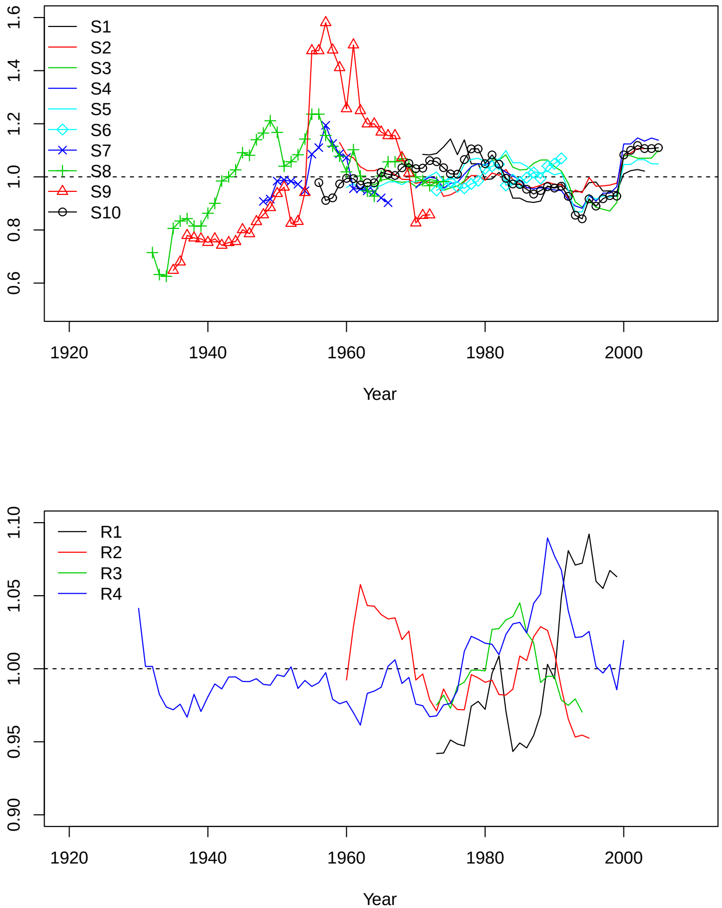

Fig. 3. Streamflow (top panel) and rainfall (bottom panel) data: running average exceedances above threshold over successive periods of (the previous) 15 years, and made dimensionless by the long term mean.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD
8, 8609–8638, 2011

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

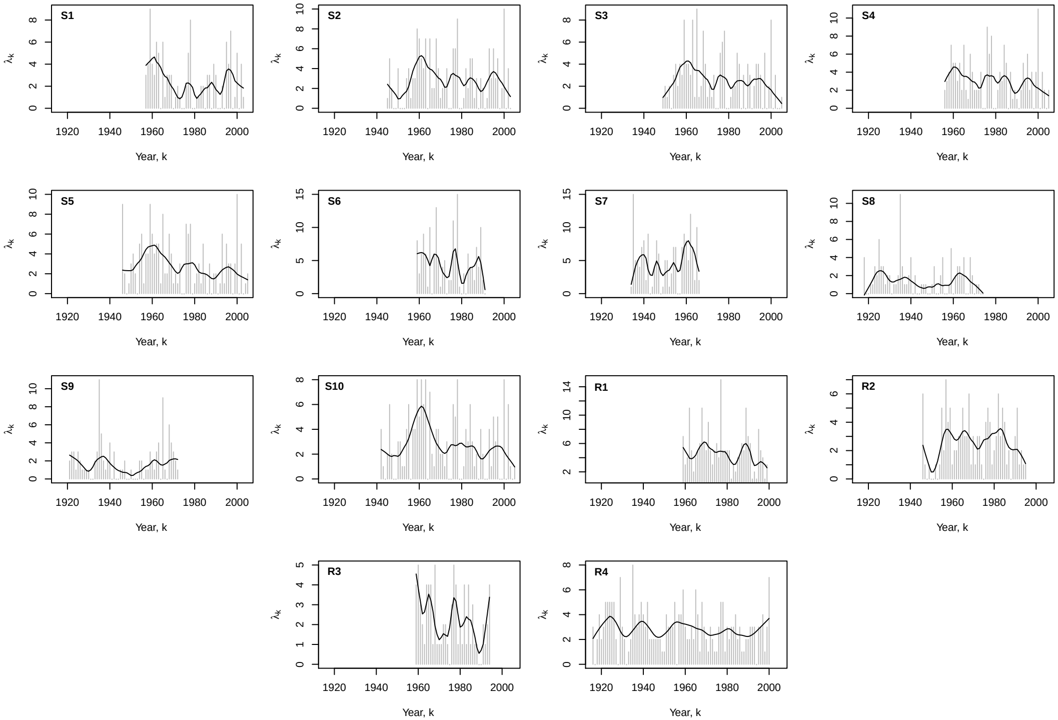

Fig. 4. Number of extremes in year k, λ_k (bars). Multi-year variability described by a LOWESS curve fitted to data, with a smoothing parameter $f = 0.2$ (solid lines).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

8, 8609–8638, 2011

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

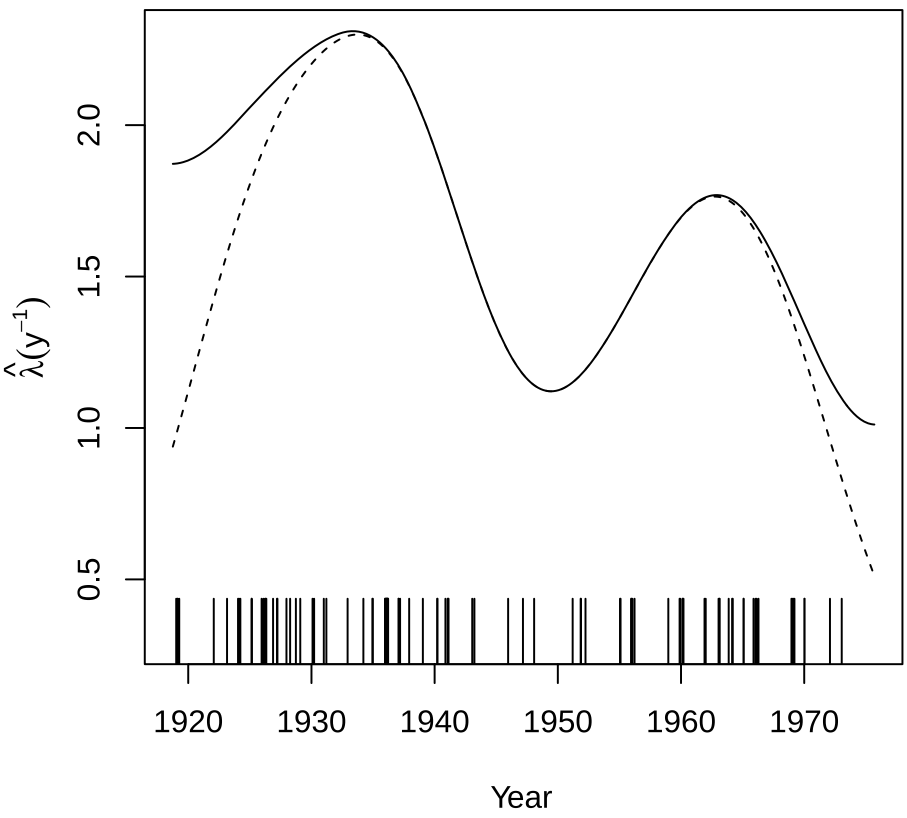

Fig. 5. Estimated flood occurrence rate at S8 – Ponte Juncais, with (solid line) and without (dashed line) pseudodata generation; flood dates (represented by vertical lines) obtained from the POT time data.

HESSD
8, 8609–8638, 2011

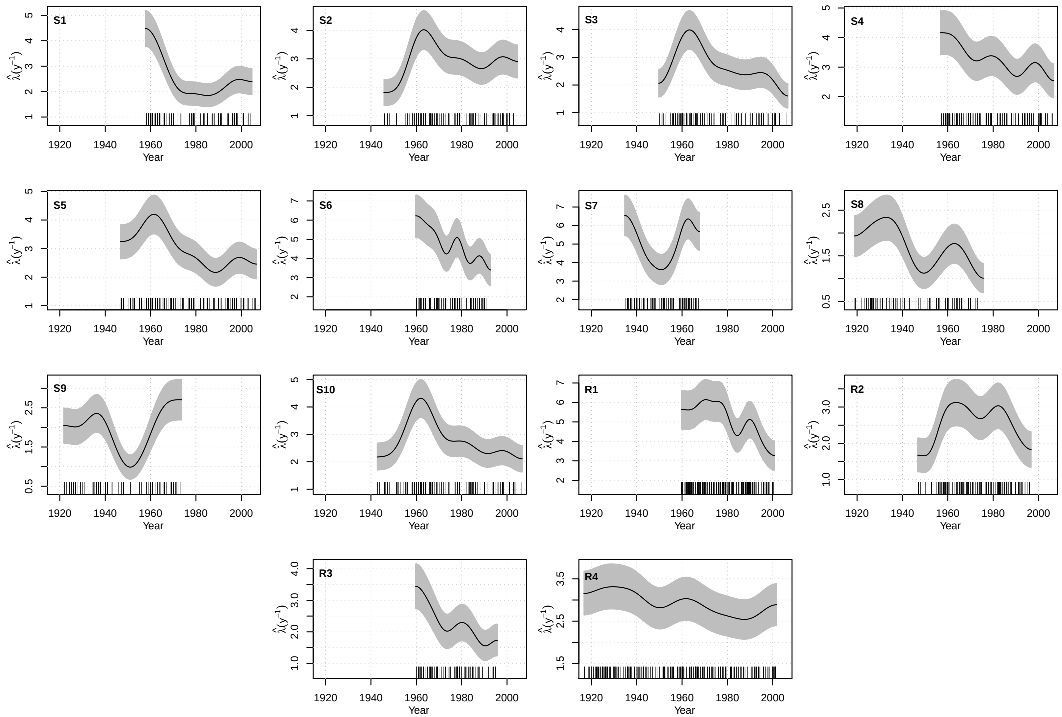

Fig. 6. Estimated flood and extreme rainfall event occurrence rate (black lines) with 95 % confidence intervals (grey area). The vertical ticks indicate the points in time when events occurred (POT time data). The code in the top left corner of each graph identifies the data series (Table 1).

HESSD

8, 8609–8638, 2011

CC BY

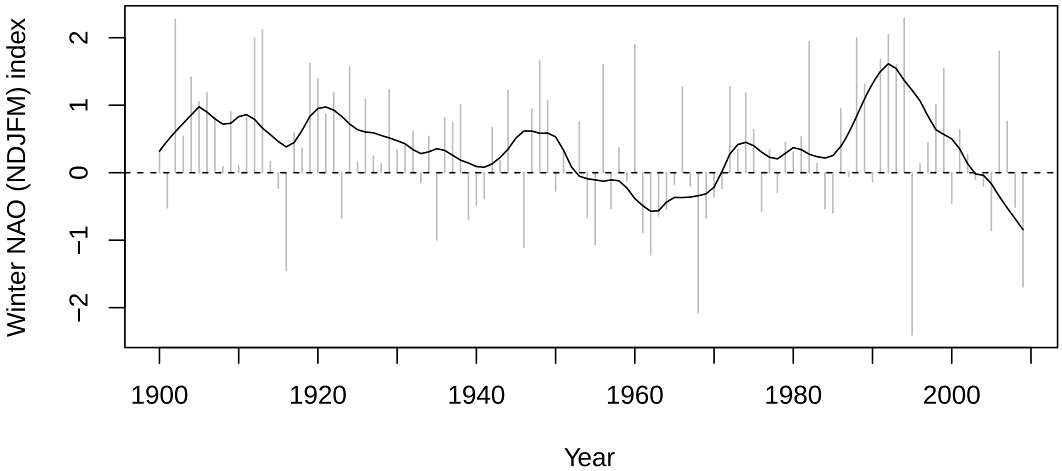

Fig. 7. Winter NAO indices based on standardized sea-level pressures differences from November to March – NAO (NDJFM) – between Iceland and Gibraltar from 1900 to 2009 (the year corresponds to November of each NDJFM average). The solid black line is a smoothing LOWESS curve.

CC BY

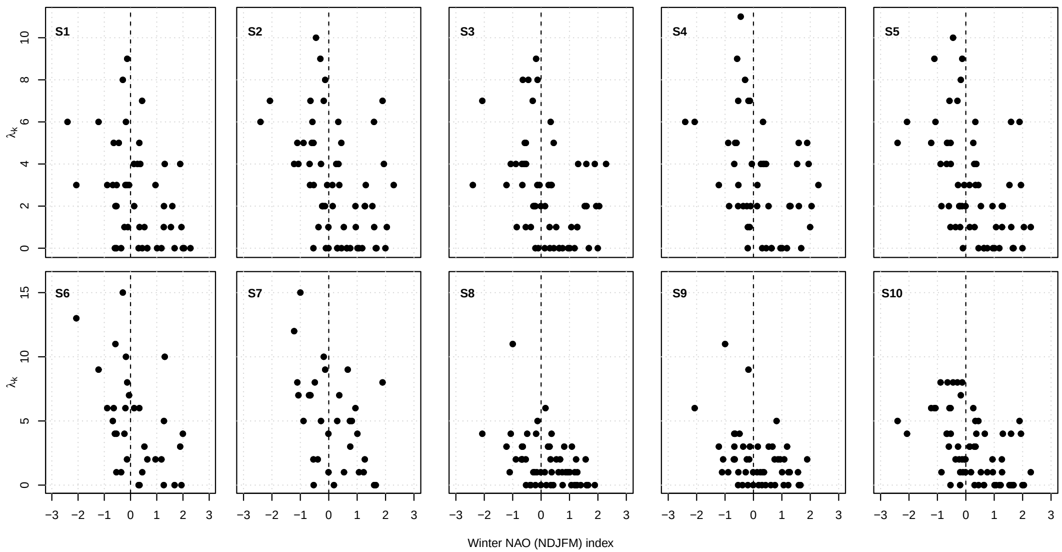

Fig. 8. Number of floods in a hydrologic year $k/(k+1)$, λ_k plotted against the winter NAO (NDJFM) index of the same year. Graphs are identified with the sample number in the top left corner.

HESSD

8, 8609–8638, 2011

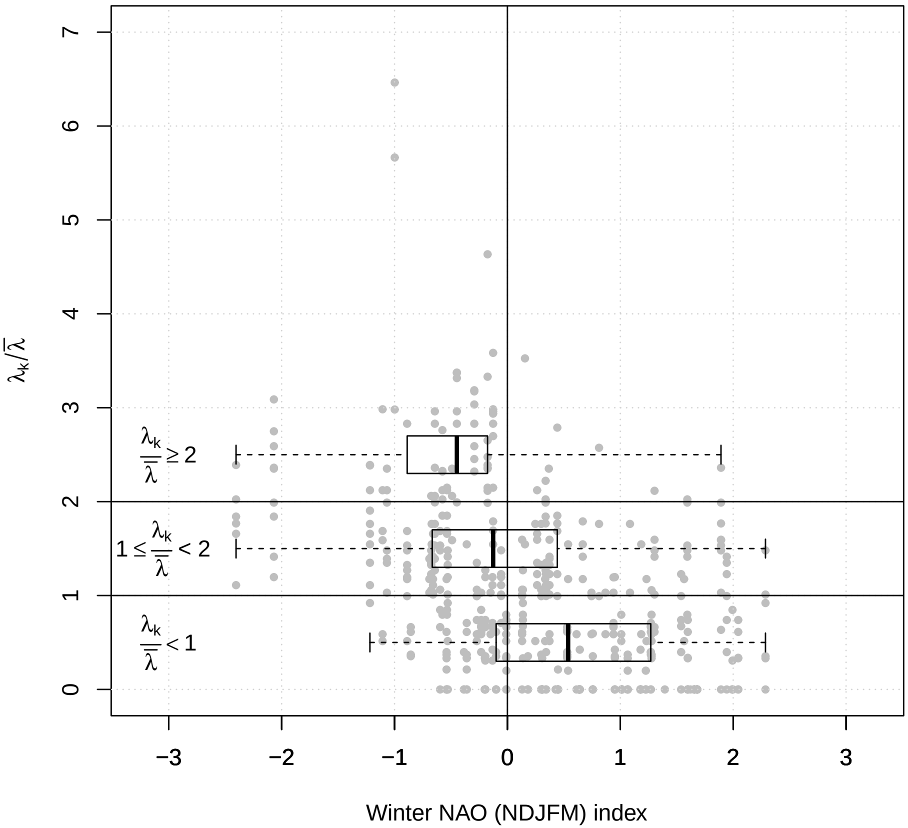

Fig. 9. Samples S1 to S10. Number of floods in a hydrologic year divided by the mean number, $\lambda_k/\bar{\lambda}$, of all the mean daily flow samples plotted together against the winter NAO (NDJFM) index of year k.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

8, 8609–8638, 2011

Nonstationarities in the occurrence rates of flood events

A. T. Silva et al.