HAL open science

# How historical information can improve estimation and prediction of extreme coastal water levels: application to the Xynthia event at La Rochelle (France)

Thomas Bulteau, Déborah Idier, Jérôme Lambert, Manuel Garcin

▶ **To cite this version:**

Thomas Bulteau, Déborah Idier, Jérôme Lambert, Manuel Garcin. How historical information can improve estimation and prediction of extreme coastal water levels: application to the Xynthia event at La Rochelle (France). Natural Hazards and Earth System Sciences, 2015, 15, pp.1135-1147. 10.5194/nhess-15-1135-2015 . hal-01160309

**HAL Id: hal-01160309**

**https://brgm.hal.science/hal-01160309v1**

Submitted on 5 Jun 2015

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

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

Distributed under a Creative Commons Attribution 4.0 International License

Nat. Hazards Earth Syst. Sci., 15, 1135–1147, 2015
www.nat-hazards-earth-syst-sci.net/15/1135/2015/
doi:10.5194/nhess-15-1135-2015
© Author(s) 2015. CC Attribution 3.0 License.

Natural Hazards and Earth System Sciences
Open Access

# How historical information can improve estimation and prediction of extreme coastal water levels: application to the Xynthia event at La Rochelle (France)

**T. Bulteau, D. Idier, J. Lambert, and M. Garcin**

BRGM, 3 avenue C. Guillemin, 45060 Orléans Cedex 2, France

*Correspondence to:* T. Bulteau (t.bulteau@brgm.fr)

Received: 14 October 2014 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 20 November 2014
Accepted: 10 May 2015 – Published: 5 June 2015

**Abstract.** The knowledge of extreme coastal water levels is useful for coastal flooding studies or the design of coastal defences. While deriving such extremes with standard analyses using tide-gauge measurements, one often needs to deal with limited effective duration of observation which can result in large statistical uncertainties. This is even truer when one faces the issue of outliers, those particularly extreme values distant from the others which increase the uncertainty on the results. In this study, we investigate how historical information, even partial, of past events reported in archives can reduce statistical uncertainties and relativise such outlying observations. A Bayesian Markov chain Monte Carlo method is developed to tackle this issue. We apply this method to the site of La Rochelle (France), where the storm Xynthia in 2010 generated a water level considered so far as an outlier. Based on 30 years of tide-gauge measurements and 8 historical events, the analysis shows that (1) integrating historical information in the analysis greatly reduces statistical uncertainties on return levels (2) Xynthia's water level no longer appears as an outlier, (3) we could have reasonably predicted the annual exceedance probability of that level beforehand (predictive probability for 2010 based on data until the end of 2009 of the same order of magnitude as the standard estimative probability using data until the end of 2010). Such results illustrate the usefulness of historical information in extreme value analyses of coastal water levels, as well as the relevance of the proposed method to integrate heterogeneous data in such analyses.

## 1 Introduction

Extreme value theory has been widely used to estimate the highest values of coastal water levels (WL). Within risk analyses, the knowledge of extreme WL and their associated annual probabilities of exceedance or return periods are required for dimensioning coastal defences or for designing WL scenarios useful in flooding hazard estimations.

A first approach consists of performing a classical extreme value analysis (EVA) directly on tide-gauge observations (this approach is called direct) (Arns et al., 2013). However, such a method provides limited extrapolation time. Indeed, it is generally considered that one should not estimate levels whose return periods exceed 4 times the data-span to keep uncertainties manageable (Pugh, 2004), whereas the analysis is fully constrained by the duration of observations (a few decades at most). In addition, direct methods are sensitive to outliers (Tawn and Vassie, 1989), those particularly extreme values much higher than other observations, thus making results even more uncertain. An outlier might be an extreme manifestation of the random variable we want to analyse or it can be a realisation of a different random process or an error in recording or reporting the measurement (Grubbs, 1969). In the first case, the outlying observation should be kept in the sample as it provides valuable information on the random variability inherent in the data (Mazas and Hamm, 2011).

An alternative to the direct approach consists of performing an EVA to the random atmospheric surge signal and then combining it with the deterministic tidal probability distribution (Tawn and Vassie, 1989; Batstone et al., 2013), thus allowing extrapolation to larger return periods while being less sensitive to outliers (Haigh et al., 2010). Such an indi-

rect method assumes surges and tides are independent. This assumption being wrong in some places (Idier et al., 2012), methods have been developed to take into account this partial dependency (Mazas et al., 2014). However, the results are not yet fully satisfactory, with for instance a notable offset between direct and indirect methods within the interpolation domain (i.e. where return periods are less than the duration of observation). Moreover, even if this approach allows estimating WL of longer return periods, it is still constrained by the information measured by the tide gauge. Consequently, outliers might not be better described by the final distribution (typically if the associated atmospheric surges are outliers in their own distribution), making the estimation of their return periods problematic. For instance, the maximum hourly WL recorded at La Rochelle (8.01 m above Z.H. (Zéro Hydrographique)) during the storm Xynthia that hit the French Atlantic coast on 28 February 2010 causing 47 deaths (Bertin et al., 2012), still appears as an outlier using an indirect approach and the estimation of its return period is not relevant (Duluc et al., 2014).

Another possibility is to use regional frequency analysis (RFA) to artificially increase the duration of observation , thereby reducing uncertainties (Duluc et al., 2014; Weiss et al., 2014a, b). Outliers may thus be better described by the distribution as their representativity might increase. RFA consists of pooling together observations from several sites inside a homogeneous region, assuming the highest observations in that region follow a common regional probability distribution, up to a local scale factor representing specific characteristics of each site. However, this approach raises the issues of the definition of homogeneous regions and the intersite dependency. Using an RFA of skew surges, Duluc et al. (2014) estimated a return period of Xynthia's WL greater than 1000 years, although they acknowledged uncertainties were large.

The above-described techniques are all initially based on WL measurements. In the past, before the era of systematic gauging, extreme events also happened. For those generating marine submersion, testimonies exist which report the inundated places. This information is often partial, in the sense that most of the time it indirectly indicates that the sea-level was at least higher than a given mark, but not which water level was actually reached. Recently, Hamdi et al. (2014) proposed a method to integrate historical information in extreme surge frequency estimation, using the maximum likelihood estimators for the distribution parameters. However, this method requires the knowledge of historical surges, a piece of information rarely found in archives (see e.g. Baart et al., 2011). The added value of using historical information in EVA has been widely recognised for the last 30 years in the domain of hydrology (see e.g. Benito et al. (2004) for a review). Among the statistical techniques developed to combine both sources of data (recent observations and historical information), Bayesian methods provide the most flexible and adequate framework because of their natural ability for handling uncertainties in extreme value models (Reis and Stedinger, 2005; Coles and Tawn, 2005). Surprisingly, we found only one reference (Van Gelder, 1996) developing such a method for sea water levels. Van Gelder (1996) set up a Bayesian framework to account for known historical sea floods in the estimation of sea dikes design level in the Netherlands. The method consists of using historical data as prior information to estimate an a priori distribution for the parameters of the probability distribution. However, the method cannot deal with partial information (an estimation of the historical water level is needed), implying that a lot of historical information cannot be integrated in such a framework.

In the hydrology field, Reis and Stedinger (2005) developed a Bayesian Markov chain Monte Carlo (MCMC) approach to tackle the issue of integrating partial historical information within EVA. The essence of the approach is to incorporate partial historical data into the model likelihood as censored observations. In the present study, we build on this approach to develop a Bayesian MCMC method adapted for EVA of coastal water levels (called HIBEVA, for Historical Information in Bayesian Extreme Value Analysis, hereafter). We notably take into account the influence of mean sea-level rise on tide-gauge data and historical information. We also take advantage of the Bayesian framework to derive predictive return levels (Coles and Tawn, 2005). In particular, we investigate whether it is possible to better predict the probability of future extreme coastal WL by considering partial historical information. As a case study, we apply the HIBEVA method to the site of La Rochelle and investigate whether (Q1) integrating historical information significantly reduces statistical uncertainties; (Q2) the WL reached during Xynthia in 2010 is really an outlier; (Q3) it would have been possible to predict the annual exceedance probability of that level before it happened.

Section 2 describes the HIBEVA method. The case study at La Rochelle is then presented in Sect. 3. In Sect. 4, results are discussed and some conclusions and perspectives that such a method opens for extreme statistics are drawn in Sect. 5.

## 2 The HIBEVA method

### 2.1 Theoretical model

The model chosen to represent and extrapolate extreme values of WL is the generalised Pareto distribution (GPD), applied to a peaks-over-threshold (POT) sample. This extreme value model has been widely used and is most commonly recommended as it makes use of all the high values for the period under study to adjust the parametric distribution (Coles, 2001; Hawkes et al., 2008). Bernardara et al. (2014) recommend a double-threshold ($u_p$, $u_s$) approach to deal with auto-correlated environmental variables in a POT frame-

work. First, physical de-clustering is performed by selecting a proper physical threshold $u_p$ above which only the maximum WL value is selected for each event that exceeds this threshold. The independence of the maximum WL selected is ensured by setting a minimum interval between peak water levels. This interval is typically chosen to be representative of storm duration on the site under study. The value for $u_p$ is set so that a sample of several hundred peak values can be selected to include both moderate and strong storm events. In practice, this corresponds to a number $n$ of events per year between 5 and 10 in average. The second step of the double-threshold approach is a statistical optimization consisting in selecting a relevant value of the statistical threshold $u_s$ ($u_s > u_p$), which is used in the formulation of the GPD, limiting both bias and variance (Bernardara et al., 2014). The choice of $u_s$ is driven by classical visual tools such as mean residual life and parameters stability plots (see Coles, 2001).

The GPD is a distribution with two parameters ($\sigma$ – scale parameter, and $\xi$ – shape parameter). For a given threshold $u_s$, the cumulative distribution function (CDF) of the GPD is equal to the probability $P(X \le x \mid X > u_s)$, where the random variable $X$ describes observed peak water levels, and it can be written as follows:

$$G_{(\xi,\sigma)}(x) = \begin{cases} 1 - \left(1 + \dfrac{\xi(x-u_s)}{\sigma}\right)_+^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\dfrac{(x-u_s)}{\sigma}\right) & \text{if } \xi = 0 \end{cases} \quad \text{for } x > u_s, \tag{1}$$

where $\sigma > 0$ and the notation $y_+$ for $y \in R$ is defined as $y_+ = \max(y, 0)$. The support of the distribution is $u_s < x \le u_s - (\sigma / \xi)$ if $\xi < 0$ and $x > u_s$ if $\xi \ge 0$. Whereas $\sigma$ represents the scale of the distribution (in units of $x$), $\xi$ controls the behaviour of the distribution's tail. If $\xi < 0$, the distribution is bounded, we are in the Weibull domain. If $\xi > 0$ (resp. $= 0$), the distribution is unbounded, we are in the Fréchet (resp. Gumbel) domain. Contrary to the Weibull domain, a small change of $\xi$ in the Fréchet domain involves significant changes of the distribution.

### 2.2 Bayesian framework

In contrast with classical statistical methods used to compute the parameters of the distribution and to derive extreme values (e.g., maximum likelihood, method of moments, probability weighted moments...), Bayesian techniques provide a natural framework to deal with uncertainties. They are designed to obtain the full posterior distribution of variables of interest and not only point estimates (Coles and Tawn, 2005).

Let us denote by $\boldsymbol{\theta}$ the vector of parameters $(\xi, \sigma)$. Its posterior distribution is related to the likelihood of data through Bayes' theorem:

$$f(\boldsymbol{\theta}|D) = \frac{f(D|\boldsymbol{\theta})\, f(\boldsymbol{\theta})}{f(D)}, \tag{2}$$

where $f(D|\boldsymbol{\theta})$ is the likelihood function of a set of observations $D$ given the parameters vector, $f(\boldsymbol{\theta})$ is the prior distribution of the parameters and $f(D)$ is a normalising constant depending only on the observations. $f(\boldsymbol{\theta})$ translates the prior knowledge one may have about the parameters. In our study, we have no prior information about GPD parameters for our data set. Consequently, we use a non-informative flat prior ($f(\boldsymbol{\theta}) \propto 1$) (Payrastre et al., 2011). In that case, $f(\boldsymbol{\theta}|D)$ is proportional to the likelihood function.

To sample effectively the posterior distribution of interest, we use a Markov chain Monte Carlo (MCMC) algorithm. MCMC algorithms allow sampling values of the parameters from the posterior distribution, without computing the normalising constant. In this study, the Metropolis–Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) is used to generate a set of 50 000 vectors $\boldsymbol{\theta}$ with density $f(\boldsymbol{\theta}|D)$. The convergence of the chain is checked numerically with the Geweke test (Geweke, 1992) and visually with trace plots. We can then compute the corresponding quantiles of WL according to the GPD. In particular, the mode of the set of vectors $\boldsymbol{\theta}$ can be retrieved as the vector maximising the likelihood function (because of the proportionality between $f(\boldsymbol{\theta}|D)$ and $f(D|\boldsymbol{\theta})$). The associated quantiles $x_T$ correspond therefore to the maximum likelihood estimates for WL. Credibility intervals on WL can also be estimated based on the large set of quantile values. Results can be displayed on a return-level plot once the correspondence between quantiles $x_T$ ($x_T > u_s$) and return periods $T$ has been set up as follows:

$$P(X > x_T) = \frac{1}{nT}, \tag{3}$$

where $n$ is the mean number of exceedances of threshold $u_p$ per year. The quantile $x_T$ is said to be the standard estimative $T$-year return level and it is exceeded once on average by a peak event in $T$ years. Conversely, $T$ is said to be the standard estimative return period of level $x_T$. Since $P(X > x_T) = P(X > u_s)\, P(X > x_T | X > u_s)$, Eq. (3) can be rewritten in a more suitable form to construct a return-level plot:

$$1 - G_{\boldsymbol{\theta}}(x_T) = \frac{1}{\lambda T}, \tag{4}$$

where $\lambda = nP(X > u_s)$ is the mean number of exceedances of threshold $u_s$ per year.

One main advantage of the Bayesian analysis is the possibility to integrate all the available information in a unique predictive distribution for extreme WL values (Coles and Tawn, 2005), which is defined as follows:

$$P(X \le x | X > u_s, D) = \int_{\boldsymbol{\theta}} G_{\boldsymbol{\theta}}(x)\, f(\boldsymbol{\theta}|D)\, d\boldsymbol{\theta}. \tag{5}$$

Thus, the predictive distribution of a new observation $x$ (given it is greater than $u_s$) can be easily estimated as

the mean of GPD values calculated at $x$ for the entire set of sampled parameters and can be represented on a return-level plot after solving the equation $\widetilde{T} = 1 \big/ (\lambda P(X > \widetilde{x}_T | X > u_s, D))$, where $\widetilde{x}_T$ is the predictive return level associated with the predictive return period $\widetilde{T}$. Although the terminology of predictive return period is loose, it is useful in order to maintain comparison with the standard analogue $T$ (Coles and Tawn, 2005). Since all the uncertainty information has been integrated in the final result, credibility intervals are no longer defined. Instead, the value $p = 1 \big/ \left(n\widetilde{T}\right)$ can be interpreted as the probability that, given all the available information, a future peak WL will exceed $\widetilde{x}_T$.

Within the Bayesian framework, it is therefore possible to calculate and compare both standard estimative return levels $x_T$ (equal to what would have been obtained using a classical maximum likelihood estimator) and predictive return levels $\widetilde{x}_T$. While the predictive return levels incorporate all the uncertainty information, standard estimative return levels can be associated with credibility intervals which provide an overview of the uncertainty related to the quantiles $x_T$ when visualised on a return-level plot.

Finally, it is worth noting that for large return periods, the annual exceedance probability of a given level is directly available reading a return-level plot constructed with peak event return periods, contrary to the peak event exceedance probability of that level. Indeed, the former is equal to $1:T$ (or $1:\widetilde{T}$) whereas the latter is given by Eq. (3) $1:nT$ (or $1:n\widetilde{T}$) (cf. Appendix A).

### 2.3 Likelihood formulation

The formulation of the likelihood function in Eq. (2) depends on the characteristics of observations $D$ (Payrastre et al., 2011). We can split the likelihood function into two parts, thus separating the systematic period (with systematic tide-gauge records) and the historical period:

$$f(D|\boldsymbol{\theta}) = \underbrace{F(D_{\text{sys}}|\boldsymbol{\theta})}_{\text{systematic likelihood}} \cdot \underbrace{f(D_{\text{his}}|\boldsymbol{\theta})}_{\text{historical likelihood}} . \tag{6}$$

Let us assume we have a number $s$ of systematic tide-gauge observations above $u_s$ $(x_1, \ldots, x_s)$ and a historical period of $n_y$ years with $H = h$ events above a perception threshold $X_0$ $(X_0 > u_s)$. The $h$ events above $X_0$ during the historical period are supposed to be exhaustive. This is a necessary condition. Historical information can be of different types. The number $h$ can thus be broken down into $h_1$ historical events whose water levels are known $(y_1, \ldots, y_{h_1})$, a number $h_2$ of historical events that exceeded the perception threshold $X_0$ but whose exact water levels are not known and $h_3$ historical events whose water levels are known to be within a given range of values (lower bounds $y_1^{\text{lb}}, \ldots, y_{h_3}^{\text{lb}}$ larger than $X_0$; upper bounds $y_1^{\text{ub}}, \ldots, y_{h_3}^{\text{ub}}$). The general expression of the likelihood of systematic data is the following:

$$f\left(D_{\text{sys}}|\boldsymbol{\theta}\right) = \prod_{i=1}^{s} g_{\boldsymbol{\theta}}(x_i), \tag{7}$$

where $g_{\boldsymbol{\theta}}$ is the probability density function of the GPD for parameters $\boldsymbol{\theta}$.

The general expression of the likelihood of historical data is the following:

$$f(D_{\text{his}}|\boldsymbol{\theta}) = P(H|\boldsymbol{\theta}) \prod_{j=1}^{h_1} g_{\boldsymbol{\theta},\, X|X>X_0}\left(y_j\right) \tag{8}$$
$$\prod_{l=1}^{h_3} \left[ G_{\boldsymbol{\theta},\, X|X>X_0}\left(y_l^{\text{ub}}\right) - G_{\boldsymbol{\theta},\, X|X>X_0}\left(y_l^{\text{lb}}\right) \right].$$

The first term of the right-hand side is the probability of observing $h = h_1 + h_2 + h_3$ events above $X_0$ during $n_y$ years whereas the two product terms specify the historical information for $h_1$ and $h_3$ historical events. Considering that the peaks exceeding $u_s$ occur as a Poisson process (Coles, 2001), then the number of exceedances of $u_s$ in $n_y$ years follows a Poisson distribution of parameter $\lambda n_y$. Consequently, the number of exceedances of $X_0$ in $n_y$ years follows a Poisson distribution of parameter $\lambda n_y P(X > X_0 | X > u_s) = \lambda n_y [1 - G_{\boldsymbol{\theta}}(X_0)]$. Thus,

$$P(H|\boldsymbol{\theta}) = \frac{\left(\lambda n_y \left[1 - G_{\boldsymbol{\theta}}(X_0)\right]\right)^h}{h!} \exp\left(-\lambda n_y \left[1 - G_{\boldsymbol{\theta}}(X_0)\right]\right). \tag{9}$$

Replacing Eq. (9) into Eq. (8) and since $g_{\boldsymbol{\theta},\, X|X>X_0}(x) = g_{\boldsymbol{\theta}}(x) \big/ (1 - G_{\boldsymbol{\theta}}(X_0))$ and $G_{\boldsymbol{\theta},\, X|X>X_0}(x) = G_{\boldsymbol{\theta}}(x) \big/ (1 - G_{\boldsymbol{\theta}}(X_0))$, Eq. (8) becomes the following:

$$f(D_{\text{his}}|\boldsymbol{\theta}) = \frac{\left(\lambda n_y\right)^h}{h!} \exp\left(-\lambda n_y \left[1 - G_{\boldsymbol{\theta}}(X_0)\right]\right) \tag{10}$$
$$(1 - G_{\boldsymbol{\theta}}(X_0))^{h_2} \prod_{j=1}^{h_1} g_{\boldsymbol{\theta}}\left(y_j\right) \prod_{l=1}^{h_3} \left[ G_{\boldsymbol{\theta}}\left(y_l^{\text{ub}}\right) - G_{\boldsymbol{\theta}}\left(y_l^{\text{lb}}\right) \right].$$

So far, we have implicitly considered that the POT sample represents a stationary process. This assumption is systematically made in the hydrology field (Gaume et al., 2010). However, extreme WL exhibit long-term trends that cannot be ignored. Over the 20th century, these trends have been shown to be similar to those of mean sea level (MSL) at most locations worldwide (Woodworth et al., 2011). To account for this behaviour in the systematic data set, the linear trend is calculated for the entire tide-gauge record and removed from the data. Then data are adjusted to have a mean sea-level equal to that of the reference year of interest. The historical perception threshold must also be corrected for the MSL rise (and called hereafter the adjusted perception threshold).

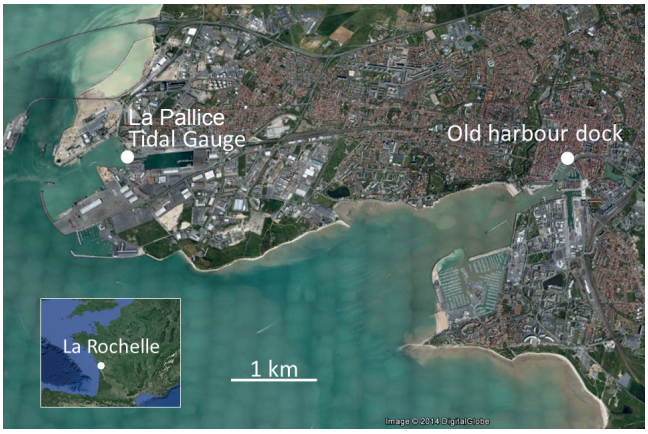


**Figure 1.** Study site and water level data localisation.

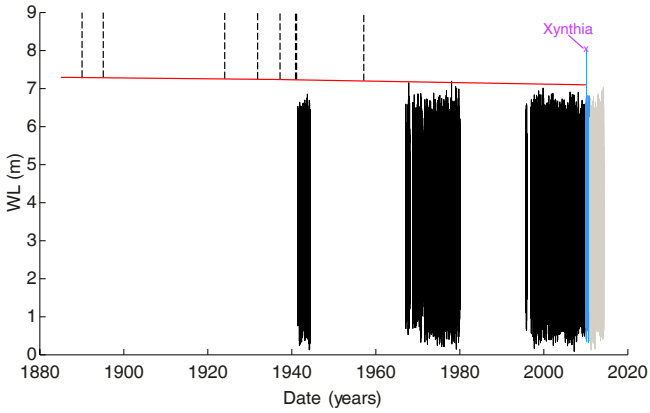


**Figure 2.** Input data: hourly tide-gauge measurements after removing the linear trend (black: until the end of 2009; blue: 2010; grey: 2011–2013) and historical information (black dotted lines). The red line represents the position of the perception threshold (7.1 m Z.H. in 2010). It varies with time as a consequence of the mean sea-level rise.

Once this is done, Eq. (10) becomes

$$f\left(D_{\text{his}}\,|\,\boldsymbol{\theta}\right)=\prod_{m=1}^{n_y}\left[\frac{\lambda^{h_m}}{h_m!}\exp\left(-\lambda\left[1-G_{\boldsymbol{\theta}}\left(X_{0,m}\right)\right]\right)\right. \tag{11}$$
$$\left.\left(1-G_{\boldsymbol{\theta}}\left(X_{0,m}\right)\right)^{h_{2,m}}\prod_{j=1}^{h_{1,m}}g_{\boldsymbol{\theta}}\left(y_j\right)\prod_{l=1}^{h_{3,m}}\left[G_{\boldsymbol{\theta}}\left(y_l^{\text{ub}}\right)-G_{\boldsymbol{\theta}}\left(y_l^{\text{lb}}\right)\right]\right],$$

where $X_{0,m}$ is the adjusted perception threshold for historical year $m$ and $h_{1,m}, h_{2,m}, h_{3,m}$ are, respectively, the numbers of historical events with known WL, with unknown WL and with WL within a range of values, that exceeded $X_{0,m}$ during year $m$. $h_m$ is the total number of historical events that exceeded $X_{0,m}$ during year $m$ ($h_m=\sum_{i=1}^{3}h_{i,m}$).

## 3 Application to the Xynthia event at La Rochelle

### 3.1 Study site and data

The study site is La Rochelle (west Atlantic coast of France, Fig. 1), focusing on the tide gauge located at La Pallice harbour (about 30 years of data until 2013). The highest recorded sea-level is 8.01 m Z.H. and it occurred during Xynthia at high water on 28 February 2010 (see Fig. 2). As a comparison, the highest tidal level estimated from tidal components analysis is 6.86 m Z.H. (SHOM, 2013).

As highlighted in the introduction, to illustrate the usefulness of the developed HIBEVA method, we investigate whether: (Q1) integrating historical information significantly reduces statistical uncertainties, (Q2) the WL of 8.01 m Z.H. reached during Xynthia is really an outlier, (Q3) it would have been possible to predict the annual exceedance probability of that level beforehand.

Four cases are considered, applying the HIBEVA method, respectively, to the following: (case 1) the systematic data until year 2009, (case 2) the systematic data including Xynthia's year (2010), (case 3) the systematic data until year 2009 with historical information, (case 4) the systematic data including Xynthia's year (2010) with historical information.

Whereas all cases are useful to answer our first point (Q1), cases 2 and 4 aim more specifically at investigating the outlier nature of Xynthia's WL (Q2), and cases 1, 3 and 4 aim at studying the capability of the HIBEVA method to predict the exceedance probability of Xynthia's WL beforehand (Q3).

Regarding the systematic data until 2010, the tide gauge provides about 27 years of data. Figure 2 shows the data after removing the linear trend ($1.9\pm0.1\,\text{mm}\,\text{yr}^{-1}$, in agreement with the study of Gouriou et al. (2013) in the same area) and adjusting it to the mean sea-level of 2010 (calculated from the same data set and equal to 3.93 m Z.H.).

Concerning historical events, the data set is based on two analyses of archives: Garnier and Surville (2010) and Lambert (2014). A convenient perception threshold is the altitude of the old harbour dock of La Rochelle, which has remained unchanged over the studied period (first identified event: 1890). When the dock is mentioned as flooded, the water level is considered to have reached at least the dock altitude. Following the notations of Sect. 2, we are in a case where $h_1=h_3=0$. Based on a digital terrain model (DTM) (Litto3D ®, horizontal resolution 1 m, vertical accuracy 0.15 m), the mean altitude of the dock, calculated from 457 points surrounding it, is $X_0=7.1\pm0.1$ m Z.H. A total of eight flooding events of the old harbour dock are identified since 1890 (Table 1 and Fig. 2). Original archives can be found in the above-mentioned references. The entire historical period covers 94.4 years (including gaps in the systematic period). As explained in Sect. 2.3, this historical data set must be corrected for the mean sea-level rise. Since the systematic data trend is close to the global sea-level rise trend (see e.g. Church and White, 2011) and the vertical land motion at the study site (monitored by GPS station since 2001) is negli-

**Table 1.** Summary of the eight historical flooding events that submerged the old harbour dock since 1890. The altitude of the old harbour dock is 7.1 m Z.H. Each event reported here has therefore generated a WL higher than 7.1 m Z.H. back in the year of the event. Notations for the sources: GS – Garnier and Surville (2010); L – Lambert (2014).

| Date (dd/mm/yyyy) | Sources | WL reached, corrected for year 2010 (m Z.H.) |
|---|---|---|
| 22–23/01/1890 | GS; L | > 7.29 |
| 10–11/02/1895 | GS; L | > 7.29 |
| 08–09/01/1924 | GS; L | > 7.25 |
| 10/11/1931 | L | > 7.25 |
| 13–14/03/1937 | L | > 7.24 |
| 16/11/1940 | GS; L | > 7.23 |
| 16/02/1941 | GS; L | > 7.23 |
| 15/02/1957 | GS | > 7.20 |

gible (Santamaría-Gómez et al., 2012), we can assume that the relative sea-level rise in the La Rochelle area is equal to the absolute global sea-level rise. In other locations where this cannot be assumed, regional estimations of sea-level rise should be used instead. Making the hypothesis that this result is valid over the long term, we use the global sea-level rise rate of Church and White (2011) over the period 1880–1935 (global linear trend of $1.1 \pm 0.7\,\mathrm{mm\,yr^{-1}}$) to adjust the perception threshold $X_0$ for each year since 1890 and until 1935. For the period 1936–2010, we use the one calculated previously from local tide-gauge measurements. For example, $X_0$ adjusted for year 1890 becomes $X_{0,1} = X_0 + 0.0019(2010\text{-}1936) + 0.0011\ (1936\text{–}1890) \simeq 7.29\,\mathrm{m}$.

### 3.2 Results

The first step of the double-threshold approach detailed in Sect. 2.1 is the physical de-clustering of systematic data. With a minimal duration of 72 h (typical storm duration on the French Atlantic coast) between two peaks to ensure their independence, the physical threshold $u_\mathrm{p}$ is chosen so that $n$, the mean number of peak events per year, is about 10. Then, the statistical threshold $u_\mathrm{s}$ is selected using the classical tools described in Sect. 2. This provides a threshold $u_\mathrm{s} = 6.68\,\mathrm{m}$ for the case with the smallest data set, i.e. case 1. For this case, the mean number of peak WL that exceed that threshold per year is $\lambda = 2.9$. It is estimated as the number of peak WL exceeding $u_\mathrm{s}$ divided by the effective duration of the systematic period (about 26 years for case 1). For sake of intercomparison, the threshold $u_\mathrm{s}$ is kept constant for every case (1 to 4). It should be noted that the rate $\lambda$ could be treated as uncertain under the Bayesian framework, thus making the problem tri-dimensional. In that case, the likelihood of systematic data (cf Eq. 7) should be modified to account for the probability of observing $s$ peak WL during the systematic period. However, to simplify the presentation, we chose to fix $\lambda$ at the proportion observed in the systematic data set.

Results are presented in Fig. 3 and Table 2. As a general comment, whatever the case, predictive return levels are uniformly above standard estimates (Fig. 3). This is a consequence of the parameter uncertainty they account for (Coles and Tawn, 2005). At low levels, there is little difference between predictive and standard return levels. At higher levels, the difference becomes larger as a consequence of the increasing parameter uncertainty.

First, we focus on the impact of historical information on the standard estimative return levels $\mathrm{WL}_T$ ($T = 50$, 100 or 500 years) as well as on their associated credibility intervals (Q1) (Table 2). When historical data are taken into account, the values of $\mathrm{WL}_T$ increase whatever the considered return period (cases 3 and 4 vs. cases 1 and 2, respectively). When systematic data until the end of 2010 are used (cases 2 and 4), the precision related to estimated return levels also changes. In particular, the integration of historical data divides by a factor of 2 the relative widths of the credibility intervals whatever the return period. When systematic data until the end of 2009 are used (cases 1 and 3), we notice the relative widths of the credibility intervals are almost the same, with a slight increase for case 3 where historical data are integrated. This can be explained by a shift of the distribution of the GPD parameters towards the Fréchet domain (i.e. positive values of $\xi$) (Fig. 3 $\mathrm{a}_1$ and $\mathrm{a}_3$). Indeed, as mentioned in Sect. 2.1, a small change of $\xi$ in the Fréchet domain involves significant changes of the distribution. Consequently, credibility intervals are wider if the distribution of the GPD parameters lies in the Fréchet domain than if it lies in the Weibull domain. If we consider that case 4 is the reference, then integrating historical data in case 3 leads to more accurate values of $\mathrm{WL}_T$ compared to case 1, where no historical information is used, while keeping the relative width of the credibility interval similar to case 1. Thus, integrating historical information in the EVA of WL reduces uncertainties with more accurate and/or more precise estimative return levels.

Now, we investigate the outlier nature of Xynthia's WL (Q2), comparing standard estimative return periods for cases 2 and 4 (Fig. 3). In case 2, the bivariate posterior probability density contours of $(\xi, \sigma)$ shows that the shape parameter of the distribution's mode is positive, indicating a heavy-tailed distribution. There is also a large variability of $\xi$, resulting in extremely large credibility intervals for the highest return periods. This is due to the value of the highest point (Xynthia) which is about 0.8 m above the second highest and could be reasonably considered as an outlier. The standard estimative return period of Xynthia's WL is 320 years which is much larger than 4 times the observation period ($4 \times 27 = 108$ years) and therefore highly uncertain. In case 4, the bivariate density plot shows that $\xi$ is better constrained: the historical information has greatly reduced the uncertainties on $\xi$ as it can be seen on the credibility intervals (Fig. 3 $\mathrm{b}_4$). In this case, the water level reached during

**Table 2.** Standard estimative return values of WL and widths of the associated 95 % central credibility intervals (absolute – ΔCI and relative – ΔCI / $WL_T$) for several return periods $T$ according to each case: (1) systematic data until the end of 2009, (2) systematic data until the end of 2010, (3) systematic data until the end of 2009 and historical information, (4) systematic data until the end of 2010 and historical information.

| Case | $T$ (years) | $WL_T$ (m Z.H.) | ΔCI (m Z.H.) | ΔCI / $WL_T$ (%) |
|---|---|---|---|---|
| 1 | 50 | 7.19 | 0.62 | 9 |
| | 100 | 7.24 | 0.85 | 12 |
| | 500 | 7.35 | 1.64 | 22 |
| 2 | 50 | 7.51 | 1.46 | 19 |
| | 100 | 7.68 | 3.39 | 44 |
| | 500 | 8.15 | 5.67 | 70 |
| 3 | 50 | 7.46 | 0.61 | 8 |
| | 100 | 7.61 | 0.91 | 12 |
| | 500 | 8.00 | 2.10 | 26 |
| 4 | 50 | 7.56 | 0.70 | 9 |
| | 100 | 7.76 | 1.08 | 14 |
| | 500 | 8.33 | 2.63 | 32 |

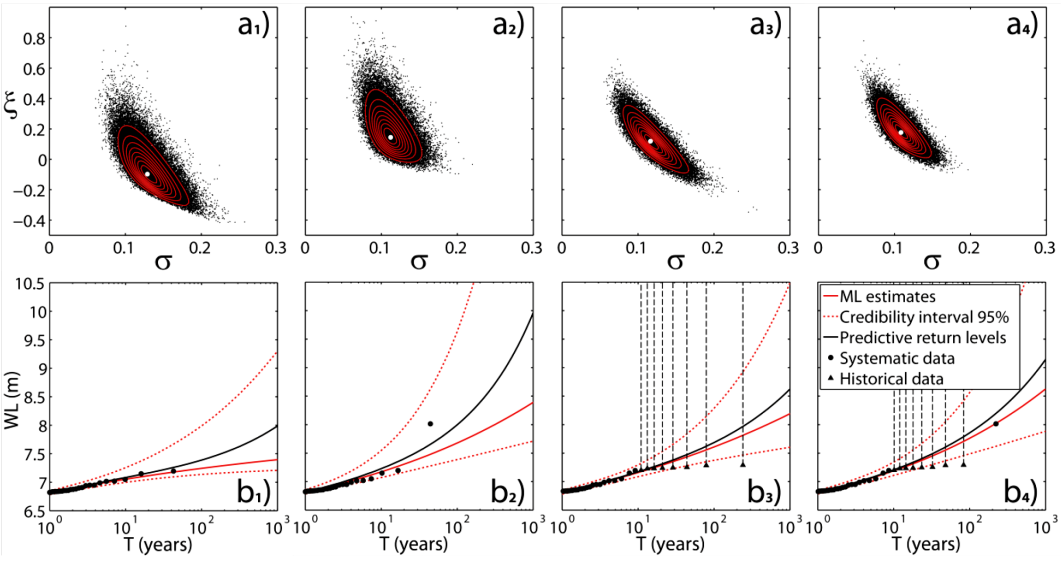


**Figure 3.** Results: **(a)** bivariate density contours of parameters $\xi$ and $\sigma$. Black dots represent the Markov chain of 50 000 simulations. The white dot is the mode of the bivariate density. **(b)** return-level plots for the following cases: (1) systematic data until the end of 2009, (2) systematic data until the end of 2010, (3) systematic data until the end of 2009 and historical information, (4) systematic data until the end of 2010 and historical information. The plotting positions of systematic and historical data are defined using the method developed by Naulet (2002) (see Appendix B).

Xynthia no longer appears as an outlier. The standard estimation of its return period is 220 years.

Finally, we evaluate if we could have predicted the exceedance probability of Xynthia's WL before it happened (Q3), by comparing results of cases 1, 3 and 4 in terms of standard estimation and prediction (Fig. 3). Because the calculated return periods of Xynthia's WL are large (typically greater than 100 years) and it makes more sense to speak about predictive exceedance probabilities rather than predictive return periods (see Sect. 2.2), we will compare results in terms of annual probabilities of exceedance (see Sect. 2.2 and Appendix A). Then we shall recall that the prediction for a Xynthia-like WL can be interpreted as the probability that next year's maximum WL (e.g. in 2010 if we are in 2009) will exceed Xynthia's WL. In case 1, the shape parameter of the distribution's mode is slightly negative, which indicates a bounded distribution with a maximum of $u_s - \sigma / \xi =$ 7.98 m. This is lower than Xynthia's WL, which implies the standard estimation of the return period for a Xynthia-like WL is not defined. The obtained prediction of the annual

probability of exceedance of a Xynthia-like WL for 2010 is $p \simeq 1:1060$ years. In case 3, the bivariate density plot for the GPD parameters shows that the value of $\xi$ for the distribution's mode is now positive compared to case 1, indicating a heavy-tailed distribution. The dispersion of $\xi$ is slightly lower than in case 1, but its distribution is shifted towards the Fréchet domain, resulting in larger credibility intervals as mentioned previously (Fig. 3 $b_3$). The standard estimation of the return period of a Xynthia-like WL is about 520 years. Considering the predictive return levels, the annual probability of exceedance of a Xynthia-like WL is $p \simeq 1:270$ years. Thus, by considering historical data, the predictive probability of having an annual maximum WL in 2010 of at least 8.01 m is about 5-fold the one calculated in case 1 where no historical information is used. Finally, results of case 4 can be used to estimate the predictive quality of the method for this event on the study site: interestingly, there is not much difference between cases 3 and 4. The bivariate density plot of case 4 shows that $\xi$ is slightly greater with smaller dispersion and that $\sigma$ is a bit more constrained. Consequently, the return-level plots are very similar in both cases. The standard estimation of the annual exceedance probability of Xynthia's WL is $1:220$ years, a value close to $1:270$ years predicted back in 2009 (case 3). Thus, at the end of 2009, applying the HIBEVA method to the available systematic data at that time, together with historical information, we could have predicted the right order of magnitude of the annual exceedance probability of a Xynthia-like WL.

## 4 Discussion

By integrating historical information in the extreme value analysis of WL, the proposed method allows a better assessment of standard estimative return levels while reducing statistical uncertainties. This has been verified on the site of La Rochelle. Furthermore, the HIBEVA method allows placing extreme events which can be considered as outliers in classical EVA, in a broader context, thus relativising their uniqueness. The standard estimation of the return period of the WL reached during Xynthia in the complete analysis at La Pallice (case 4, $T = 220$ years) is significantly lower than the previous estimate of Duluc et al. (2014) using the same systematic data set ($T > 1000$ years, see Sect. 1). Going one step further, the method, applied on the full data set (systematic data until 2013 and historical information, all data adjusted to the mean water level of year 2013), provides a standard return period of a Xynthia-like WL of about 270 years. It is the smallest return period we found in the literature regarding Xynthia's WL, tending to show it is probably closer to 100 years than to 1000 years. In terms of prediction, the method provides a probability of about $1:180$ years that the maximum WL in 2014 exceeds that of Xynthia.

However, like other EVA approaches, the HIBEVA method relies on some approximations and assumptions (both on the data and the statistical model).

First, the use of historical data leads to uncertainties at two levels. Within this study, we assume WL values at the tide gauge of La Pallice and inside the harbour of La Rochelle (about 5 km apart) are comparable. Due to local effects, this might not be exactly the case. This is a primary difficulty when using historical data: most of the time, historical observations are not made at the tide-gauge location. One solution to deal with this issue, although beyond the scope of this paper, would be hydrodynamic modelling of last decades' events to statistically quantify the WL offset, called $DZ$ hereafter. A second source of uncertainty is the estimation of the perception threshold. Most of archives' information deals with water flooding a given area without more detailed information. In the present study, archives indicate that the old harbour dock was flooded without specifying the water entrance location on the dock. We assume the threshold to be the mean dock altitude, but this is an approximation. It should be noted that since the distribution of $\xi$ lies mostly in the Fréchet domain (especially in cases 2, 3 and 4), the standard and predictive estimation of large WL should be highly sensitive to the $DZ$ and $X_0$ parameters. Nevertheless, the selected values (no $DZ$ and $X_0 = 7.1$ m) lead to a standard estimation of the return periods of the 8 historical events ranging from about 10 to 15 years (Fig. 3, $b_4$). Such return periods appear consistent with the observed probability of flooding events (8 in 94.4 years) and as a first approximation, our choice seems reasonable. For other applications, where $DZ$ and $X_0$ could be more difficult to estimate, the Bayesian framework should make the integration of $DZ$ and $X_0$ uncertainties within the HIBEVA method possible (Reis and Stedinger, 2005).

The statistical model also contains uncertainties. In the POT/GPD model, a main source of uncertainties is the choice of the systematic statistical threshold $u_s$. Estimated quantiles are indeed highly dependent on the threshold, the selection of which is sometimes difficult and often subjective (Li et al., 2012). In our case study, even if there are still some uncertainties in the statistical threshold selection (done for case 1, see Sect. 3.2), the resulting estimative distribution passed two statistical adjustment tests with a 0.05 level of risk ($\chi^2$ with 10 classes, Greenwood and Nikulin, 1996; and Kolmogorov-Smirnov, see e.g. Shorack and Wellner, 2009). A second source of uncertainty comes from the seasonal and interannual variability of WL which has not been considered in our model. Regarding the seasonal variability, if $u_s$ is chosen high enough, which is the case here, the selected events occur mainly in the winter period (October to March for the French Atlantic coast) and we can reasonably consider that seasonal variability is negligible in the POT sample. Interannual variability, on the other hand, can lead to significant variations of extreme values in time as highlighted by the work of Menéndez et al. (2009). Finally, we have fitted the GPD directly on WL measurements, so even with additional

historical information, our approach could be classified as direct (see Sect. 1). As such, additional uncertainties may be involved for high return values compared to an indirect approach (Haigh et al., 2010). However, integrating historical information in an indirect approach is challenging. It would require the characterisation of historical events in terms of surges rather than WL, a piece of information rarely found in archives as mentioned in Sect. 1. It would thus also require the knowledge of historical tides which might be difficult to estimate as the tide is not a stationary process, as highlighted for instance by studies of sea-level rise influence on tidal harmonics (Pickering et al., 2012).

As described in Sect. 3 and Fig. 3, the bivariate distribution of the GPD parameters for our case study at La Rochelle lies mostly in the Fréchet domain. A consequence is that small changes of $\xi$ can generate significant changes in the return-level plots, especially in the tail of the distribution. Therefore, regarding the above-described approximations and assumptions done in the present study, the estimated values of the return period of Xynthia's WL should be considered with caution, and interpreted as orders of magnitude rather than exact values.

## 5 Conclusion

To reduce statistical uncertainties and to address the issue of outliers in extreme value analyses of coastal water levels, we developed a Bayesian method to integrate historical information (even partial) of past events that occurred before the era of systematic gauging. The proposed method, inspired from previous works in the hydrology field, is adapted to POT sample of coastal sea levels, taking into account the influence of mean sea-level rise. It provides standard estimative as well as predictive return levels, the latter being particularly useful for decision makers.

The application of the method on the site of La Rochelle in France illustrates the usefulness of historical information in reducing statistical uncertainties in EVA and relativising apparent outliers such as Xynthia's WL. In particular, it shows that, back in 2009 before the storm, we could have predicted the right order of magnitude of the annual exceedance probability of a Xynthia-like WL. These results are particularly important for raising awareness among decision makers and eventually enhancing preparedness for future flooding events. However, some uncertainties remain in the data and the statistical model, and because of the high variability of the GPD tail in the Fréchet domain, numerical values presented in this paper should be considered as indicative only.

The method opens a large field of possibilities for engineers wishing to put into perspectives classical extreme value analyses of water levels with the richness of historical information on coastal floods. Furthermore, beyond the integration of historical information in the EVA of WL, the proposed method should allow combining data of different natures together with associated uncertainties. For instance, future research may focus on combining tide-gauge data not only with historical data but also with model outputs, thus filling the holes during tide-gauge failures for example.

## Appendix A: Relation between annual exceedance probability and peak event return period

Let us denote with "maxy", the annual maximum. Using Eq. (3), the probability that the annual maximum of WL is greater than $x_T$ is

$$P\,(\text{maxy}\,(\text{WL}) > x_T) = 1 - P(X \le x_T)^n = 1 - \left(1 - \frac{1}{nT}\right)^n. \quad \text{(A1)}$$

For large return periods $T$, more precisely when $nT \gg 1$, Eq. (A1) becomes

$$P\,(\text{maxy}\,(\text{WL}) > x_T) \simeq \frac{1}{T}, \quad \text{(A2)}$$

which can be interpreted as follows: the standard estimation of the annual exceedance probability of level $x_T$ is $1:T$ years. Thus, in that case, the annual exceedance probability is directly available reading a return-level plot constructed with peak event return periods contrary to the peak event exceedance probability (cf Eq. 3).

Similarly, in the case of the predictive distribution, we obtain

$$P\,(\text{maxy}\,(\text{WL}) > \widetilde{x}_T) \simeq \frac{1}{\widetilde{T}}, \quad \text{(A3)}$$

which can be interpreted as follows: the probability that, given all the available information, next year's maximum WL will exceed $\widetilde{x}_T$, is $1:\widetilde{T}$ years.

## Appendix B: Plotting positions

The method of plotting positions used in this paper is based on the one developed by Naulet (2002) which is itself based on the formulation of Hirsch and Stedinger (1987). The plotting positions are used only for plotting return-level estimates in Fig. 3, they are not involved in the model fitting process.

Let us consider a number $m$ of perception thresholds $X_{0,k} (1 \le k \le m)$ with $X_{0,1} = u_s < X_{0,2} < \ldots < X_{0,m+1} = \infty$ defined over the entire period. The objective is to calculate the empirical exceedance probability $\hat{P}_i$ of each observed water level $X_i$ (systematic or historical). In the case of historical censored observations, $X_i$ is taken either as the corresponding lower bound for historical events that exceeded a value but whose exact water levels are not known, or as the middle value of the corresponding range for historical events whose water levels are known to be within a given range of values. Let $P_{X_{0,k}}$ be the exceedance probability of perception threshold $X_{0,k}$. If $X_{0,k} < X_i < X_{0,k+1}$, then the probability $\hat{P}_i$ must verify $P_{X_{0,k+1}} < \hat{P}_i < P_{X_{0,k}}$.

The exceedance probabilities of the perception thresholds are determined as follows:

$$P_{X_{0,k}} = P\left(X \ge X_{0,k}\right) = P\left(X \ge X_{0,k+1}\right) + P\left(X_{0,k} \le X < X_{0,k+1} \middle| X < X_{0,k+1}\right) P\left(X < X_{0,k+1}\right). \quad \text{(B1)}$$

Therefore:

$$P_{X_{0,k}} = P_{X_{0,k+1}} + P_{c_k}\left(1 - P_{X_{0,k+1}}\right), \quad \text{(B2)}$$

where $P_{c_k}$ is the conditional probability of threshold $X_{0,k}$. The probabilities $P_{X_{0,k}}$ can then be calculated step by step (since $P_{X_{0,m+1}} = 0$) as soon as the probabilities $P_{c_k}$ are estimated:

$$\hat{P}_{c_k} = \frac{A_k}{A_k + B_k + C_k}, \quad \text{(B3)}$$

where $A_k$ is the number of events (systematic and historical periods) with a WL $x$ such that $X_{0,k} \le x < X_{0,k+1}$, $B_k$ is the number of events (systematic and historical periods) with a WL $x$ such that $x < X_{0,k}$, and $C_k$ is the number of events that did not reach the perception thresholds $X_{0,1}, \ldots, X_{0,k}$ during the periods of definition of the thresholds. For example, if $X_{0,2}$ is defined for 5 years with 2 events above it during these 5 years and $X_{0,3}$ is defined for 10 years with 1 event above it during these 10 years, then $C_3$ is estimated as follows: $C_3 = 15\lambda - 3$, with $\lambda$ the mean number of exceedances of threshold $u_s$ per year (see Sect. 2.2).

The empirical exceedance probabilities $\hat{P}_i^k$ $(1 \le i \le A_k)$, or plotting positions, of the $A_k$ events with WL between $X_{0,k}$ and $X_{0,k+1}$ ranked in descending order, are finally calculated

with the following formula:

$$\hat{P}_i^k = \hat{P}_{X_{0,k+1}} + \left(\hat{P}_{X_{0,k}} - \hat{P}_{X_{0,k+1}}\right) \frac{i-a}{A_k + 1 - 2a} \qquad \text{(B4)}$$

$$1 \leq i \leq A_k,$$

where $a$ is a constant between 0 and 0.5 characterising the spacing between plotting positions. For the present work, we used the value 0.4 (Cunnane, 1978).

*Acknowledgements.* This work was supported by BRGM (Histo-Stat project). Observations at La Pallice tide gauge are the property of SHOM and Grand Port Maritime de La Rochelle, and they are available on the website http://refmar.shom.fr. All calculations and figures have been realised with a home-made code in Matlab©. The MCMC part of the code is inspired by packages *nsRFA* and *evdbayes* of the *R* statistical software. Also, the authors would like to thank G. Le Cozannet for his contribution on sea-level rise and J. Rohmer for his help on *R* packages. Last, we gratefully acknowledge helpful comments from E. Bradshaw and an anonymous reviewer that have significantly improved this paper.

Edited by: M. Parise

## References

Arns, A., Wahl, T., Haigh, I. D., Jensen, J., and Pattiaratchi, C.: Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practice, Coast. Eng., 81, 51–66, doi:10.1016/j.coastaleng.2013.07.003, 2013.

Baart, F., Bakker, M. A. J., van Dongeren, A., den Heijer, C., van Heteren, S., Smit, M. W. J., van Koningsveld, M., and Pool, A.: Using the 18th century storm-surge data from the Dutch Coast to improve the confidence in flood-risk estimates, Nat. Hazards Earth Syst. Sci., 11, 2791–2801, doi:10.5194/nhess-11-2791-2011, 2011.

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., and Hunt, T.: A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines, Ocean Eng., 71, 28–39, doi:10.1016/j.oceaneng.2013.02.003, 2013.

Benito, G., Lang, M., Barriendos, M., Llasat, M., Frances, F., Ouarda, T., Thorndycraft, V., Enzel, Y., Bardossy, A., Couer, D., and Bobbe, B.: Use of systematic, paleoflood and historical data for the improvement of flood risk estimation: review of scientific methods, Nat. Hazards, 31, 623–643, 2004.

Bertin, X., Bruneau, N., Breilh, J.-F., Fortunato, A.B., and Karpytchev, M.: Importance of wave age and resonance in storm surges: The case Xynthia, Bay of Biscay, Ocean Model., 42, 16–30, doi:10.1016/j.ocemod.2011.11.001, 2012.

Church, J. A. and White, N. J.: Sea-level rise from the late 19th to the early 21st century, Surv. Geophys., 32, 585–602, 2011.

Coles, S.: An Introduction of Statistical Modeling of Extreme Values, Springer, London, 2001.

Coles, S. and Tawn, J.: Bayesian modelling of extreme surges on the UK east coast, Phil. Trans. R. Soc. A., 363, 1387–1406, 2005.

Cunnane, C.: Unbiased plotting positions – A review. J. Hydrol., 37, 205–222, 1978.

Duluc, C.-M., Deville, Y., and Bardet, L.: Extreme sea level assessment: application of the joint probability method at Brest and La Rochelle and uncertainties analysis, La Houille Blanche, 1, 11–17, doi:10.1051/lhb/2014002, 2014.

Garnier, E. and Surville, F.: La tempête Xynthia face à l'histoire: submersions et tsunamis sur les littoraux français du Moyen Âge à nos jours, edited by: Le Croît Vif, Saintes, 2010.

Gaume, E., Gaàl, L., Viglione, A., Szolgay, J., Kohnovà, S., and Blöschl, G.: Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites, J. Hydrol., 394, 101–117, 2010.

Geweke, J.: Evaluating the accuracy of sampling-based approaches to calculating posterior moments, edied by: Bernardo, J. M., Berger, J. O., Dawid, A. P., Smith, A. F. M., in: Bayesian Statistics 4, Oxford University Press, Oxford, 169–193, 1992.

Gouriou, T., Martín Míguez, B., and Wöppelmann, G.: Reconstruction of a two-century long sea level record for the Pertuis d'Antioche (France), Cont. Shelf Res., 61–62, 31–40, 2013.

Greenwood, P. E. and Nikulin, M. S.: A guide to chi-squared testing, Wiley, New York, 1996.

Grubbs, F. E.: Procedures for detecting outlying observations in samples, Technometrics, 11, 1–21, doi:10.1080/00401706.1969.10490657, 1969.

Haigh, I. D., Nicholls, R., and Wells, N.: A comparison of the main methods for estimating probabilities of extreme still water levels, Coast. Eng., 57, 838–849, 2010.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V.: Use of historical information in extreme surge frequency estimation: case of the marine flooding on the La Rochelle site in France, Nat. Hazards Earth Syst. Sci. Discuss., 2, 5647–5688, doi:10.5194/nhessd-2-5647-2014, 2014.

Hastings, W. K.: Monte Carlo sampling methods using markov chains and their applications, Biometrika, 57, 97–109, 1970.

Hawkes, P. J., Gonzales-Marco, D., Sánchez-Arcilla, A., and Panayotis, P.: Best practice for the estimation of extremes: a review, J. Hydraul. Res., 46, 324–332, 2008.

Hirsch, R. M. and Stedinger, J. R.: Plotting positions for historical floods and their precision, Water Resour. Res., 22, 715–727, 1987.

Idier, D., Dumas, F., and Muller, H.: Tide-surge interaction in the English Channel, Nat. Hazards Earth Syst. Sci., 12, 3709–3718, doi:10.5194/nhess-12-3709-2012, 2012.

Lambert, J.: Contribution au recensement de submersions marines historiques liées aux tempêtes dans le secteur de La Rochelle (Charente-Maritime), BRGM Report BRGM/RP-63723-FR, 2014.

Leadbetter, M.: Extremes and local dependence in stationary sequences, Probab. Theory Relat. Fields, 65, 291–306, 1983.

Li, F., Bicknell, C., Lowry, R., and Li, Y.: A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset, Coast. Eng., 69, 1–11, doi:10.1016/j.coastaleng.2012.05.006, 2012.

Mazas, F. and Hamm, L.: A multi-distribution approach to POT methods for determining extreme wave heights, Coast. Eng., 58, 385–394, 2011.

Mazas, F., Kergadallan, X., Garat, P., and Hamm, L.: Applying POT methods to the Revised Joint Probability Method for determining extreme sea levels, Coast. Eng., 91, 140–150, doi:10.1016/j.coastaleng.2014.05.006, 2014.

Menéndez, M., Méndez, F. J., and Losada, I. J.: Forecasting seasonal to interannual variability in extreme sea levels, ICES J. Mar. Sci., 66, 1490–1496, 2009.

Metropolis, N., Rosenbluth, A. W., Teller, A. H., and Teller, E.: Equations of state calculations by fast computing machines, J. Chem. Phys., 21, 1087–1092, 1953.

Naulet, R.: Utilisation de l'information des crues historiques pour une meilleure prédétermination du risque d'inondation. Application au bassin de l'Ardèche à Vallon Pont-d'Arc et Saint Martin

d'Ardèche, PhD thesis, Université Joseph Fourrier – Grenoble 1 (France) and INRS eau terre environnement (Québec), 2002.

Payrastre, O., Gaume, E., and Andrieu, H.: Usefulness of historical information for flood frequency analyses: developments based on a case study, Water Resour. Res., 47, W08511, doi:10.1029/2010WR009812, 2011.

Pickering, M. D., Wells, N. C., Horsburgh, K. J., and Green, J. A. M.: The impact of future sea-level rise on the European Shelf tides, Cont. Shelf Res., 35, 1–15, 2012.

Pugh, C.: Changing Sea Levels – Effects of Tides, Weather and Climate, Cambridge University Press, Cambridge, 2004.

Reis, D. S. and Stedinger, J. R.: Bayesian MCMC flood frequency analysis with historical information, J. Hydrol., 313, 97–116, 2005.

Santamaría-Gómez, A., Gravelle, M., Collilieux, X., Guichard, M., Martín Míguez, B., Tiphaneau, P., and Wöppelmann, G.: Mitigating the effects of vertical land motion in tide gauge records using a state-of-the-art GPS velocity field, Global Planet. Change, 88–99, 6–17, doi:10.1016/j.gloplacha.2012.07.007, 2012.

SHOM: Références Altimétriques Maritimes – édition 2013, ISBN:978-2-11-097286-6, 2013.

Shorack G. R. and Wellner J. A.: Empirical Processes with Applications to Statistics, Society for Industrial and Applied Mathematics, Philadelphia, 2009.

Tawn, J. A. and Vassie, J. M.: Extreme sea-levels: the joint probabilities method revisited and revised, in: Proceedings of the Institute of Civil Engineering Part 2, 249–442, 1989.

Van Gelder, P. H. A. J. M.: A new statistical model for extreme water levels along the Dutch coast, edited by: Tickle, K. S., Goulter, I. C., Xu, C. C., Wasimi, S. A., Bouchart, F., Stochastic Hydraulics '96, Proceedings of the 7th IAHR International Symposium, Balkema, Rotterdam, 243–249, 1996.

Weiss, J., Bernardara, P., and Benoit, M.: Formation of homogeneous regions for regional frequency analysis of extreme significant wave heights, J. Geophys. Res. Oceans, 119, 2906–2922, doi:10.1002/2013JC009668, 2014a.

Weiss, J., Bernardara, P., and Benoit, M.: Modeling intersite dependence for regional frequency analysis of extreme marine events, Water Resour. Res., 50, doi:10.1002/2014WR015391, 2014b.

Woodworth, P. L., Menéndez, M., and Gehrels, W. R.: Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels, Surv. Geophys., 32, 603–618, doi:10.1007/s10712-011-9112-8, 2011.