

Regional frequency analysis of extreme storm surges along the French coast

L. Bardet¹, C.-M. Duluc¹, V. Rebour¹, and J. L'Her²

¹Institut de Radioprotection et de Sûreté Nucléaire (IRSN), DEI, SARG, BEHRIG, BP17, 92 262 Fontenay aux Roses Cedex, France

²Centre d'Etudes Techniques Maritimes et Fluviales, Technopôle Brest-Iroise, BP5, 29 280 Plouzané, France

Received: 7 May 2010 – Revised: 14 February 2011 – Accepted: 19 March 2011 – Published: 14 June 2011

Abstract. A good knowledge of extreme storm surges is necessary to ensure protection against flood. In this paper we introduce a methodology to determine time series of skew surges in France as well as a statistical approach for estimating extreme storm surges. With the aim to cope with the outlier issue in surge series, a regional frequency analysis has been carried out for the surges along the Atlantic coast and the Channel coast. This methodology is not the current approach used to estimate extreme surges in France. First results showed that the extreme events identified as outliers in at-site analyses do not appear to be outliers any more in the regional empirical distribution. Indeed the regional distribution presents a curve to the top with these extreme events that a mixed exponential distribution seems to recreate. Thus, the regional approach appears to be more reliable for some sites than at-site analyses. A fast comparison at a given site showed surge estimates with the regional approach and a mixed exponential distribution are higher than surge estimates with an at-site fitting. In the case of Brest, the 1000-yr return surge is 167 cm in height with the regional approach instead of 126 cm with an at-site analysis.

1 Introduction

The impact of astronomical tide and radiational tide on the sea is theoretically known. A sea level can be thus predicted owing to the harmonic components of the tide signal (Whitcombe, 1996; Simon, 2007). However, the tide level which is observed can be different from the tidal prediction, mainly because of meteorological phenomena (Bode and Hardy, 1997; Olbert and Hartnett, 2010). For instance, a

depression caused by a storm could produce a rise in water level. The difference between the observed and predicted sea levels is called a “surge”.

Combined with other extreme events (high-tide, high waves, intense precipitation), a surge can contribute to flooding or destruction of coastal facilities (Gerritsen, 2005; De Zolt et al., 2006). The most serious case in modern history took place in the Netherlands during the storm of February 1953. A surge of almost three meters in height in places was combined with a high spring tide, causing the flood of a great part of the country (Gerritsen, 2005). Thus estimating extreme storm surges is essential to partly ensure protection against potential flood.

High return period surges can be estimated on the basis of observed surges by using statistical models. The annual maxima method (Thompson et al., 2009; Haigh et al., 2010) or the Peaks Over Threshold (POT) method (Pandey et al., 2004; Van Den Brink et al., 2005; Fawcett and Walshaw, 2007; Haigh et al., 2010) are widely used to estimate extreme events for many environmental processes: precipitation, wave, sea level, sea surge, river discharge, earthquake, wind... In order to study quite substantial samples of surges, with data which are representative of extreme events of surge, the POT methodology has been preferred in this study.

But the current fitting ways on the basis of local analyses cannot recreate some extreme events called “outliers” (Mason, 1992). An outlier is an exceptional event in a sample: it is an observation whose value is significantly distant from the values of the other observations of the same sample. In particular, a statistical fitting of the sample is not representative of this exceptional observation and the confidence intervals are often inadequate. As an example, Fig. 1 shows the surge data set at Brest, in which the event corresponding to the storm of 1987 is an outlier.



Correspondence to: L. Bardet
(lise.bardet@irsn.fr)

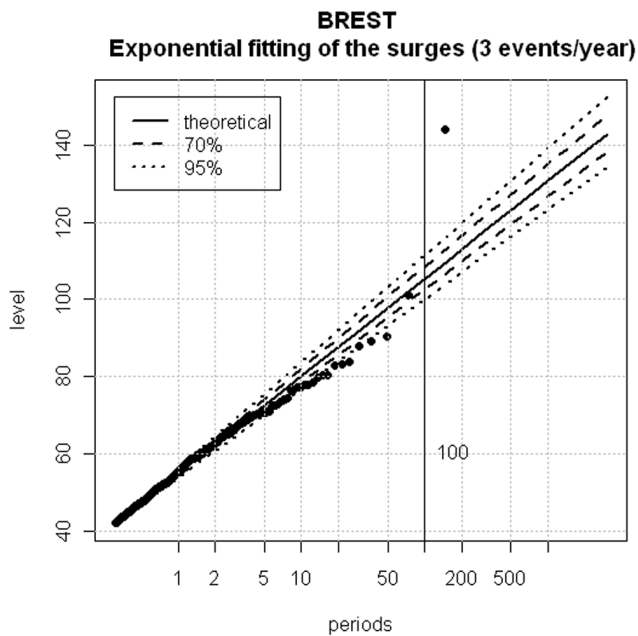


Fig. 1. Exponential fitting of the data set of surges at Brest (1846–2009). The 70% and 95% lines correspond to confidence intervals, the surge levels are in centimeters, and the return periods in years. The surge of 144 cm in height corresponding to the storm of 1987 is significantly different from the other surges of the sample; the theoretical distribution is not satisfying for this event.

If the assumption of random sampling at a given site cannot be excluded (the data set could present a surge with a return period greater than the duration of the observation period), this situation should be rare. But outliers are observed at many French harbours along the Atlantic coast and the Channel coast (storms of February 1953, December 1979, October 1987, December 1999...). This fact questions the reliability of the results obtained with current extrapolations on the basis of local observations.

To cope with this issue, the development of a statistical methodology appears necessary to take into account the outlier phenomenon in a better way. The regional frequency analysis described hereafter makes it possible to supplement local information, in particular for the extreme events. This methodology is widely used in hydrology to estimate extreme river flows (Ouarda et al., 1999; Kjeldsen et al., 2002; Ribatet, 2007; Saf, 2009). However it has seldom been applied to sea surges (Van Gelder and Neykov, 1998; Mai Van et al., 2007) and in particular in France. Thus this methodology represents a different approach than the current ways to estimate extreme storm surges in France.

2 From sea level data to a sample of surges

Observed and predicted tide levels have been collected from the *Système d'Observation du Niveau des Eaux Littorales* (SONEL) and the *Service Hydrographique et*



Fig. 2. Location of the different sites of the study, focused on the surges along the Atlantic coast and the Channel coast.

Océanographique de la Marine (SHOM) in France, with the aim of obtaining the most complete possible data on the Atlantic Coast and the Channel Coast. Sea level values have also been collected for a harbour on the Mediterranean coast (Toulon) in order to compare the statistical behaviours of surges in the Atlantic and the Mediterranean Sea in view of further studies.

This present study has been carried out with sea level data of 21 French harbours, approximately distributed in a uniform way along the Atlantic coast and the Channel coast (Fig. 2). Figure 3 illustrates the periods of continuous observation by the tide gauges for all the harbours. The tide gauges supply hourly tide levels. It is interesting to note that, since the beginning of the '90s, numerical tide gauges have been installed in various harbours and measure more precise values of the observed water level, owing to a time step of 10 min. But these data have not been used in this study in order to only capture homogeneous data throughout the periods of observation.

Table 1 reports the global period and the effective duration of the observations at each site. The effective duration can be different from the global period of observation because of tide gauge stops and malfunctions, or losses of data which lead to periods without observation of the sea level. When one or several hourly levels are consequently missing, the corresponding days are completely withdrawn from the global observation period to determine the effective duration of the sample. Therefore, the effective duration of observation is the sum of all the complete days of observation by the tide gauges.

Table 1. Characteristics of the different sites of the study: global period of observation, effective duration of observation in years, and linear trend of the sea level evolution (centimeters per year).

site	period	effective duration	observed sea level trend (cm y ⁻¹)
Dunkerque	1956–2008	39	0.1826
Calais	1941–2008	31	0.0110
Boulogne	1941–2007	21	0.2192
Dieppe	1954–1973	18	0.4383
Le Havre	1938–2008	38	0.1826
Cherbourg	1943–2008	34	0.1461
St Malo	1992–2004	11	0.4383
Roscoff	1973–2008	33	0.2192
Brest	1846–2008	148	0.1134
Le Conquet	1970–2008	35	0.2922
Concarneau	1971–2008	9	0.2557
Port Tudy	1966–2003	29	0.1461
Le Crouesty	1996–2008	6	−0.2557
St Nazaire	1957–1988	21	0.0731
St Gildas	1962–2000	24	0.0183
Sables d’Olonne	1965–2008	18	0.2922
La Rochelle	1941–2008	24	0.1826
Port-Bloc	1959–2009	21	0.1826
Arcachon	1967–2008	9	−0.0731
Bayonne	1967–2008	32	0.0731
St-Jean-de-Luz	1942–2008	36	0.1461
Toulon	1961–2007	23	0.0219

The average annual predicted levels have a constant mean over all the observation period, whereas the observed levels increase during the same time, mainly because of climate change (Plag, 2006; Pirazzoli, 2007). It is interesting to note that a possible subsidence due to tectonic or sedimentary processes could also play a role in the sea level evolution which is observed at the tide gauges (Turner, 1991; Testut et al., 2006; Haigh et al., 2010) by modifying the tidal datum (the “zero” value for tides), for instance. To compare the observed and predicted sea level data, observations at each site have been corrected so that sea level rise did not affect the average annual observed levels. Some studies consider precise trends of sea level rise owing to polynomial methods for instance (Laborde et al., 2009; Vilibić and Šepić, 2010), or dividing the global period of observation into several time periods. However, climate change is still somewhat unknown; it is very variable and we do not have enough hindsight. Thus, this study considers a linear trend over all the observation period to qualify the sea level rise. Figure 4 presents the case of Brest, where the average rise of sea level was evaluated to 0.11 cm yr⁻¹ over the period 1846–2009. Table 1 summarizes the trends of the sea level rise which have been calculated at each harbour.

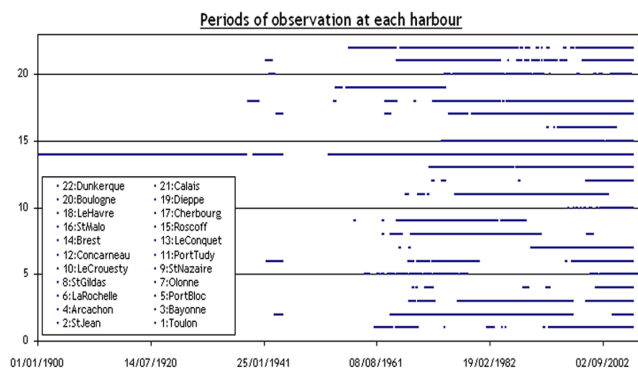


Fig. 3. Detail of the continuous periods of observation by the tide gauges at each harbour for the period 1900–2009. Before the year 1900, only the tide gauge at Brest recorded sea level data (until 1846).

Starting from hourly tidal level data, hourly surges and surges called “skew surges” (Van Den Brink et al., 2003) can be calculated (Fig. 5). A skew surge is the difference between the highest observed level and the highest predicted level, for a same high tide (Simon, 1994). These maximum levels can occur at slightly different times.

When the observed and predicted tide maxima do not coincide, the meteorological impact on the water level is not strictly separated from the astronomical impact with a hourly surge (Sterl et al., 2009). Moreover, hourly surges can be meaningless, in particular at half-tide, when the signal of the observed tide is shifted in time compared to the signal of the predicted tide (as on Fig. 5), as often actually occurs in the time series (Simon, 1994; Vilibić and Šepić, 2010). On the contrary, skew surges allow time shifts between the two signals, given that skew surges are the difference between two maxima without necessarily occurring at the same time. Thus, only skew surges have been considered in this study to limitate the “wrong” surges. But we can point out that the highest surges do not necessarily occur at a high tide and so the surge estimates can be underrated.

For this first approach in the extreme surges study, skew surges are determined thereafter for each site at an hour either side of the predicted high tide (“skew surge at more or less an hour”, see Fig. 5). A skew surge at more or less an hour is actually the difference between the maximum of the hourly observed levels and the relative maximum of the hourly predicted levels. These surges are not calculated by interpolation of the tide signals, which would make it possible to determine the real highest levels of tide as illustrated on Fig. 5. So they are subject to hourly data sampling and can underestimate or overestimate the “real” skew surges. The potential bias in the surge estimates could be evaluate in further studies on the basis of skew surges determined with a tide interpolation by a cubic spline method (Fritsch and Carlson, 1980), for instance.

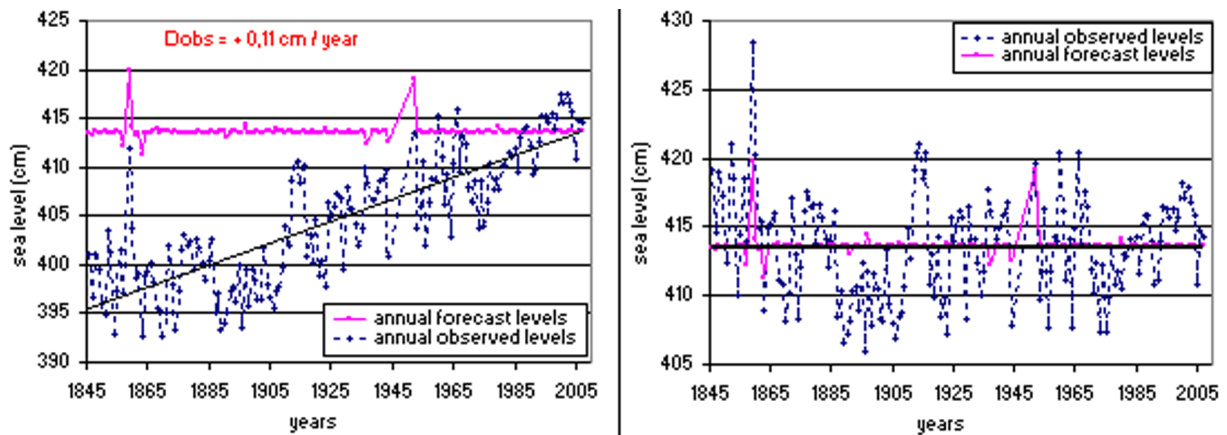


Fig. 4. Evolution of observed and predicted sea levels at Brest (1846–2009), before (left) and after (right) correction of the relative sea level rise by considering a linear trend for the observations. The “peaks” for annual predicted levels are explained by the limited amounts of data considered for the corresponding years (for instance, only 10% of the observed data are available for the year 1859. These 10% of data were measured during the month of December for which the tidal predictions were relatively important, which explains the suddenly high value of the annual predictions average for this year).

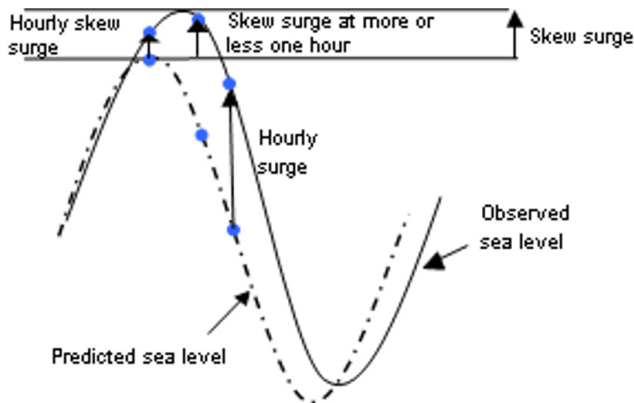


Fig. 5. Definitions of the different surges, with an hourly sampling. In this scenario, the observed tide signal is shifted in time in reference to the predicted tide signal.

Non-significant data (tide gauge malfunction, bad transcription) have been eliminated for the highest surges (higher than 30 cm) owing to a visual analysis of the time series of surges and sea levels. No replacement of these data has been done, such as with a linear interpolation from neighbouring data (Vilibić and Šepić, 2010), in order to only capture the “real” data. The eventual drifts in records, due to measuring defaults of the bottom and atmospheric pressure sensors at the gauging station (Testut et al., 2006), are not treated in this study and so might affect the future estimation of extreme storm surges. This choice was taken given the real complexity to identify these drifts and to correct them (Testut et al., 2006).

A criterion of at least 3 days has been applied between two successive skew surges so that the highest surges are retained. This period of time is supposed to be sufficient to ensure in-

dependence between the surges. This assumption of independence is necessary (Coles, 2001) for statistical analyses according to the POT method, which is carried out thereafter. However, it is interesting to point out that some studies call into question the independence of data for the reliability of the results with the POT method (Fawcett and Walshaw, 2007). This point is discussed in Sect. 5.3.

Lastly, a common threshold of 20 cm, named threshold of preselection, has been applied to the surges at each site. It is not linked with a statistical analysis but is a way to target the events which are representative for the study, by eliminating negative surges and the lowest positive surges. This threshold is arbitrary, as it is not dependent on the surge series. However, it was checked at each site that the selected value of 20 cm was low enough for not impacting the later fitting of the highest surges, the threshold of adjustment being significantly higher than the threshold of preselection.

It is important to point out that the program described in the previous parts can not extract reliable skew surges for the site of Toulon. Indeed, the observed tide signal in the Mediterranean Sea is much more susceptible to be impacted by the surges because of the low amplitude of the theoretical tide than for the sites along the Atlantic coast. This special feature leads to a great variability of the tide signal at Toulon and the observed levels at high tide can not be spotted as well as for the other sites, preventing the calculation of the skew surges. In addition, the observed tide signal can be shifted several hours compared to the predicted tide signal for the series of Mediterranean levels. However, the program calculates skew surges at more or less an hour either side of the predicted high tide and so can not process such a difference in time between the tide signals. As a consequence, the site of Toulon is not integrated into the study in the following parts.

3 Statistical software “Renext”

The statistical analyses of data sets of skew surges, as built previously, have been carried out with version 0.4–1 (work version) of the statistical software “Renext”. This software is used under the mathematical environment of R.

The Renext program, developed by the *Institut de Radio-protection et de Sûreté Nucléaire* (IRSN), implements the Peaks Over Threshold (POT) method according to the theory of J. Miquel (Miquel, 1981; Lang et al., 1997). In the POT method, the exceedances of surges over a selected threshold are fitted and the distribution of the events is supposed to follow a Poisson process (Lang et al., 1997; Coles, 2001). With this methodology, it is possible to target the significant events which are representative of the extreme events to study, which is not possible with the maxima method.

Among its functions, Renext offers a varied choice of statistical distributions: exponential, Weibull, Generalized Pareto, mixed exponential... It can test the stationarity of a sample and can process potential historical data. The parameters of the distributions are estimated by maximising the likelihood (Coles, 2001).

This statistical software has been placed at public disposal since June 2010 under the CRAN web site. A qualification file validates the main functions and the results obtained with the program, in particular by comparing with the results of another statistical software (ASTEX) in collaboration with the *Centre d’Etudes Techniques Maritimes et Fluviales* (CETMEF).

4 Regional frequency analysis of storm surges in France

4.1 Principle and interest of the method

A limitation in studies of extreme sea surges is the limited amount of data. Indeed very little information can be available at a given site or the data can be incomplete. This lack of data can lead to significant uncertainties in statistical estimating of extreme surges (Van Den Brink et al., 2005). The regional frequency analysis is a way to enlarge samples, by using available data at sites different from the site of interest. However, all the sites composing the region statistically have to present a similar hydrological behaviour.

The regional approach used here is inspired by the theory of Hosking and Wallis (Hosking and Wallis, 1997) and the studies carried out essentially for river flows (Ouarda et al., 1999; Kjeldsen et al., 2002; Ouarda et al., 2008; Saf, 2009) or for rainfall (St-Hilaire et al., 2003; Abolverd and Khalili, 2010) with the index flood model.

To take into account some potential local effects of amplification or extenuation of the surges due to characteristics of the site (bathymetry (Weaver and Slinn, 2010), configuration), the surges are standardized by being divided by the annual empirical surge at each harbour. The annual empirical

surge is determined as the observed surge which is exceeded once a year on average, and therefore it is a characteristic of the considered harbour. In practice, for D years of observation, the annual empirical surge is the $(D+1)$ highest observed surge.

Then the standardized surges are gathered to create a regional data set which is submitted to the statistical analysis. It is interesting to point out there can be a possible inter-site dependence between the data (a same strong storm may produce two important surges at two different harbours, especially for nearby sites). This issue is discussed in Sect. 5.3.

Results are finally transposed at each site by multiplying the regional value by the annual empirical surge of each site.

Standardization by the annual empirical surge is an arbitrary choice. However, an additional study has shown that standardization by the empirical surge, which is exceeded three times a year on average, gives similar theoretical surges at each site.

4.2 Choice of the region

We reiterate that the site of Toulon is not integrated into the study because of an impossibility to retrieve skew surges (see Sect. 2).

Inside a homogeneous region, the at-site frequency distributions of surges have to be the same except for a site-specific scale factor (few variations of the shape factor must be observed between the different sites). In order to analyse the variability of shape and scale factors of the at-site distributions, the observed surges have been fitted with the Weibull distribution and for a threshold which is exceeded three times a year on average. For this first approach for estimating extreme surges, the threshold is arbitrary because the optimum choice of the threshold is too difficult to evaluate with the available durations of the series (Van Den Brink et al., 2005). It has been chosen to have a quite substantial sample of data which are representative of the highest events of surge. Without being the best possible, these fittings are acceptable as a whole for the sites of the region. The numerical tests of Kolmogorov-Smirnov (Van Zyl, 2011; Liao and Shimokawa, 1999) show, for instance, a relatively good adequation between the theoretical curve and the empirical distribution at each site (the p -values of the tests presented in Table 2 are all higher than 5%, which is in general the criterion of acceptance).

The values of the different factors of the distributions for the various samples are reported in Table 2.

We can note that shape factors of the at-site distributions are roughly similar. Thus, the various sites seem to have hydrological behaviours which statistically correspond enough to constitute a homogeneous region.

Table 2. Shape factors, scale factors, and Kolmogorov-Smirnov tests of the at-site distributions (Weibull) for thresholds corresponding to three events a year in average.

site	threshold (cm)	KS test ^a (<i>p</i> -value)	shape	scale
Arcachon	43.8	0.89	1.009	11.649
Bayonne	33.8	0.76	0.929	10.391
Boulogne	51.3	0.89	1.016	13.016
Brest	43	0.93	1.028	10.984
Calais	53.1	0.97	1.109	12.272
Cherbourg	38.4	0.97	1.018	8.555
Concarneau	41.7	0.88	1.022	11.291
Dieppe	49.1	0.99	1.154	9.873
Dunkerque	64	0.87	0.940	14.043
La Rochelle	44.5	0.27	1.065	12.910
Le Havre	52.3	0.81	0.928	13.083
Le Conquet	40.2	0.94	0.962	9.724
Le Croüesty	42.7	0.93	1.443	14.736
Olonne	45.8	0.90	1.184	11.172
Port Bloc	44.3	0.50	1.034	10.874
Port Tudy	41.1	0.40	0.976	10.625
Roscoff	38.9	0.88	1.114	9.533
Saint Malo	45.5	0.88	1.067	14.426
St Jean	28.7	0.84	1.201	7.335
St Gildas	43.6	0.42	1.082	13.415
St Nazaire	53.9	0.96	1.236	13.338
mean			1.072	11.583
standard deviation			0.122	1.979

^a Test of goodness-of-fit of Kolmogorov-Smirnov.

4.3 Validation of the homogeneity of the region

Several ways exist to check the homogeneity of a region but this study is restricted to the method described by Hosking and Wallis (1997). Indeed, the tests of Hosking and Wallis (1997) are perhaps currently some of the most frequently applied tests of regional homogeneity (Kjeldsen et al., 2002; Castellarin et al., 2008; Saf, 2009; Abolverdi and Khalili, 2010) and in particular are used for some studies of extreme water levels or surges (Van Gelder and Neykov, 1998; Mai Van et al., 2007). The tests have been partly carried out with the package named “RFA” which is available under the environment of R. This package was worked out by M. Ribatet during his studies on regional frequency analysis (Ribatet, 2007) and on the basis of the Hosking and Wallis theory (1997).

Hosking and Wallis (1997) based their regional approach on L-moment ratios (combination of weighted moments) determined with the statistics of variation and skewness factors. These L-moment ratios, named L-CV and L-skewness, are calculated at each site and must be similar in a homogeneous region.

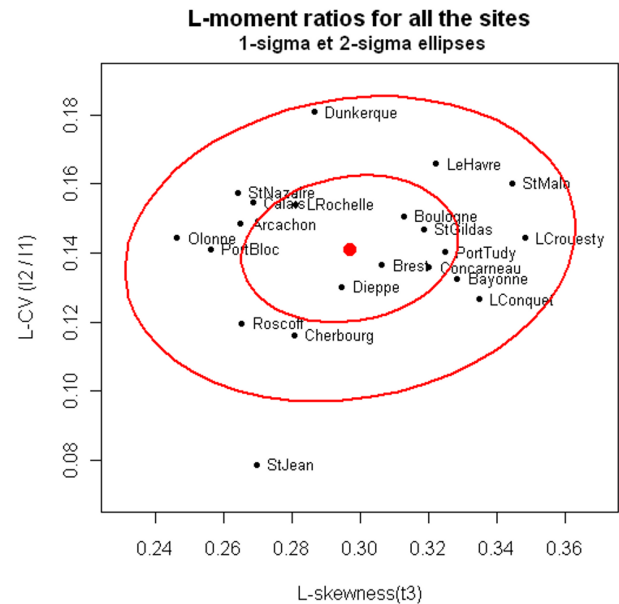


Fig. 6. Homogeneity test for the region with all the available harbours: comparison of the L-moment ratios with the graph of L-skewness versus L-CV. The point corresponding to the sample of Saint-Jean-de-Luz is outer the 2-sigma ellipse and it seems to be an unusual site for the region.

A visual assessment of the dispersion of the at-site L-moment ratios can be obtained by plotting them on graphs of L-skewness versus L-CV, as shown in Fig. 6. Sites whose L-moments are notably different from those of the other sites are excluded from the data set owing to ellipses adapted to the data (Mai Van et al., 2007; Van Gelder and Neykov, 1998). The 1-sigma ellipse (39.4% confidence) and the 2-sigma ellipse (86.5% confidence) are commonly used to adapt ellipses to a data set. According to Van Gelder and Neykov (1998), the unusual sites are outside the inner 1-sigma ellipse or outer the 2-sigma ellipse.

The surge sample of Saint-Jean-de-Luz does not appear to be homogeneous with the other samples composing the region. A reason could be the particular configuration of the site: a bay almost enclosed by a set of dikes (the recording tide gauge is located at the extremity of the dike of Socoa). To respect the homogeneity test, we chose to exclude the data set of Saint-Jean from the regional sample. However, some studies (Ribatet et al., 2006) seem to show that working with relatively large and homogeneous regions may lead to more accurate results than working with smaller and highly homogeneous regions. Thus, the influence of this data set in the regional sample is considered in parallel thereafter in each step of the study.

Hosking and Wallis (1997) also define a heterogeneity measure H for the whole region, from L-moment ratios. The sampling variability of a theoretically homogeneous region is quantified in order to compare it with the sampling variability of the region to study. This test is based on numerical

simulations according to the Monte Carlo method (Saporta, 2006; Castellarin et al., 2008) which consists of simulating many artificial samples of random variables whose distributions are known and carrying out the calculation for each sample before synthesizing all the results.

The heterogeneity measure H is then given by the following formula:

$$H = \frac{V_{\text{obs}} - \mu_v}{\sigma_v} \quad (1)$$

With V_{obs} : observed value of V ,

V : weighted standard deviation for L-CV ratios,

μ_v : mean of the values of V obtained by Monte Carlo simulations,

σ_v : standard deviation of the values of V obtained by Monte Carlo simulations.

Hosking and Wallis (1997) consider that a region can be regarded as “acceptably homogeneous” if $H < 1$, “possibly heterogeneous” if $1 \leq H < 2$, and “definitely heterogeneous” if $H \geq 2$.

The region constituted by the surges along the Atlantic coast and the Channel coast (except Saint-Jean-de-Luz) presents a heterogeneity measure of $H = -0.85$, so the region can be considered homogeneous. Moreover, it should be noted that the regional data set including the surges at Saint-Jean-de-Luz is also homogeneous ($H \sim -0.88$) according to this test. It is interesting to point out that cross-correlation between the sites of the region can affect the result of the heterogeneity measure (Castellarin et al., 2008). Hosking and Wallis (1997, p. 71) state that positive correlation among sites is the most likely cause for negative values. This point is discussed in Sect. 5.1.

4.4 Empirical distribution of the standardized surges

The regional observations consist of data from 1846 until 2009. In this study we make the hypothesis of an effective duration of 601 yr. This duration is determined as the sum of the effective durations of the at-site samples which constitute the region, that is to say without the samples of Toulon and Saint-Jean-de-Luz. Actually it is probably a too high duration given the dependence that can exist between the regional observations (for instance, two important surges occurring at two different harbours but resulting from the same strong storm). A trail to deepen this issue is described in the discussion.

The strongest surge ever measured (Dunkerque, 213 cm) has been included in the regional data set. It was observed during the storm of February 1953 and corresponds to an observed high tide level of 7.90 m according to the Port Authority Of Dunkerque for a predicted level of 5.77 m (estimate with the harmonic method nowadays used by the SHOM).

It is also known that an important surge was observed in Port-Bloc during the storm of 1999, but the tide gauge could

not measure it because of a malfunction. However it is interesting to note that the surge corresponding to the storm of the 24 January 2009 was measured at Port-Bloc (100 cm). This surge is not an outlier but it is the most important surge ever measured by the tide gauge at this site.

Moreover, storm Xynthia of the 28 February 2010 led to significant surges along a part of the Atlantic coast: the tide gauge of La Rochelle measured a surge of about 1.50 m during high tide. This event is an outlier for the sample of surges at La Rochelle but it has not been used in this study which has been carried out before the storm. However it is interesting to note that this surge, standardized by the annual empirical surge at La Rochelle, would belong to the most important events of the regional data set (standardized surge near to 2.5).

Figure 7 represents the empirical distribution of the standardized surges. The first part of the distribution is linear, then a curve is observed around the value 1.5 (normalized surge). However, no outlier is observed as in the case of Brest, for instance. The events associated to significant storms such as that of 1987 are present in the data set but do not constitute outliers for the regional sample.

4.5 Fitting with the mixed exponential distribution

Thereafter the regional data set of surges is supposed to be stationary, in other words the average rate of events in a year is approximately constant in time. However, we can note that only the surge data of Brest are available for the period before 1938.

The regional data set has been fitted with various statistical distributions (exponential, Generalized Pareto, mixed exponential), for a threshold which is exceeded once a year on average. So the analyzed sample consists of about 600 values, what is largely sufficient to fit the data set with confidence. The graphic results for the exponential and Pareto distributions are shown on Fig. 8.

Relying on the graphs, the adjustments with the exponential and Pareto distributions are not representative enough of the regional data set, considering a threshold which is exceeded once a year on average. Indeed, these distributions cannot recreate the curve which is observed in the empirical distribution.

To fit with the mixed exponential distribution here appears more appropriate for the sample. The mixed exponential distribution, defined for this study as a combination of two exponential distributions, should recreate the curve in a better way. The graph in Fig. 9 presents the curve fitting corresponding to the theoretical distribution.

By comparing the empirical and theoretical distributions, for a threshold which is exceeded once a year on average, the adjustment of the regional data set seems to be more acceptable with the mixed exponential distribution than with the exponential or Pareto distributions.

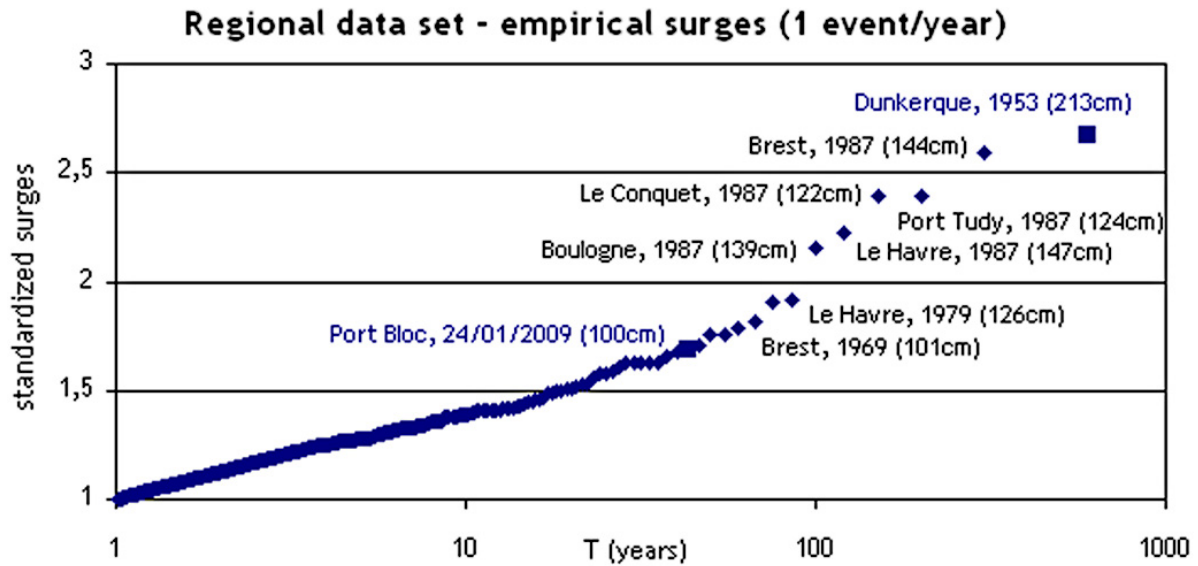


Fig. 7. Empirical distribution of the surges of the regional data set.

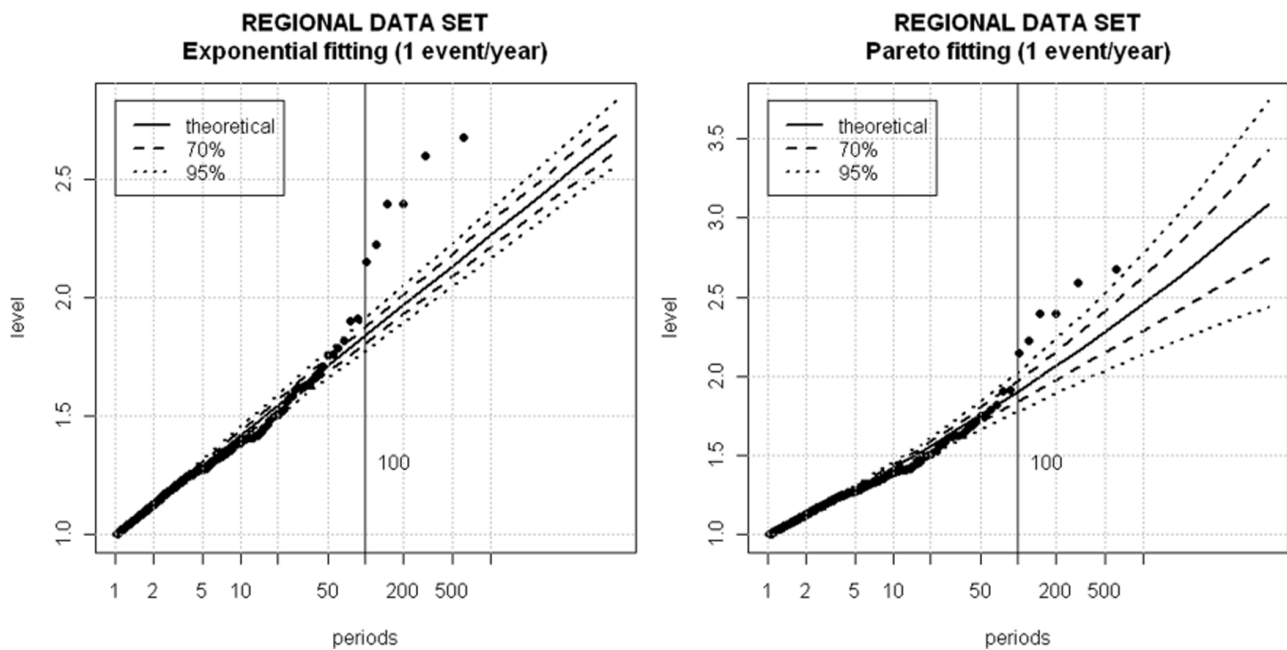


Fig. 8. Adjustments of the regional data set with the exponential and Pareto distributions. The 70% and 95% lines correspond to confidence intervals, the surge levels do not have units (normalized surges), and the return periods are in years.

Then, with the mixed exponential distribution and an average rate of one event a year, the value of the 1000-yr return surge is estimated to 3.017 (normalized value) with the value 3.636 for the higher 70% confidence limit.

To estimate the value of the 1000-yr return surge at each site, the regional surge is multiplied by the at-site annual empirical surge.

The surges estimated with regional frequency analysis are then compared with the surges obtained with a Weibull at-site distribution and for a threshold corresponding to three events a year on average. The estimates of the 1000-yr return surges are reported in Table 3.

Globally, a regional analysis gives higher theoretical surges compared to at-site fittings. The reason is the presence of the scarcest regional events which generate a curve to the

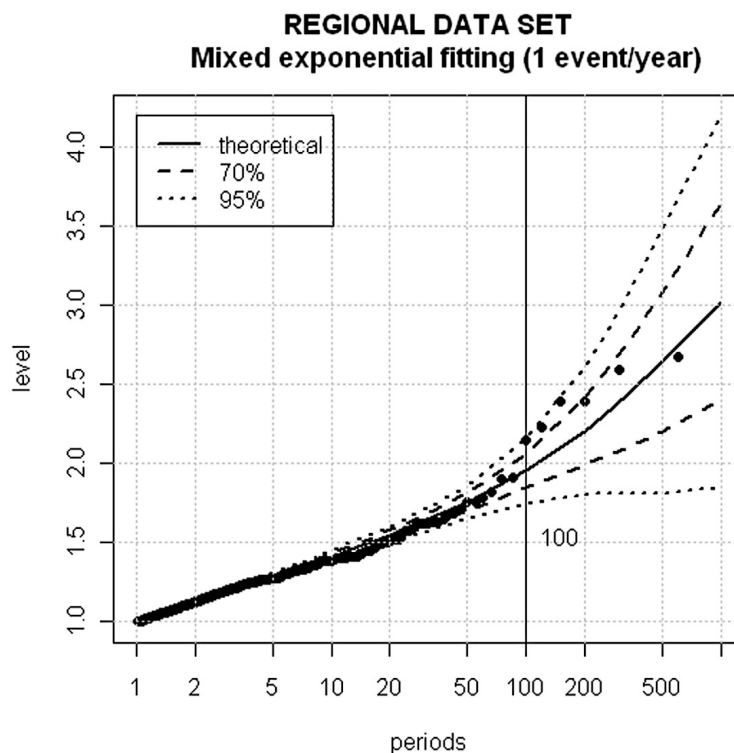


Fig. 9. Fitting of the regional data set with the mixed exponential distribution. The 70% and 95% lines correspond to confidence intervals, the surge levels do not have units (normalized surges), and the return periods are in years. The parameters of the distribution are $\text{rate1} \sim 1.7784$ (rate for the first exponential density), $\text{rate2} \sim 5.9283$ (rate for the second exponential density), and $\text{prob1} \sim 0.0361$ (probability weight for the first exponential density).

top of the distribution, whereas at-site distributions are approximately linear for most samples.

In addition, the arbitrary choice of the distribution and the threshold at each site do not affect this observation. Indeed, a fitting optimization leads only to a few centimeters of difference for the surge estimates (less than 5 cm of variation at Brest for the higher 70% confidence limit of the 1000-yr return surge).

The statistical fitting of the regional sample including the data at Saint-Jean-de-Luz gives very similar results (less than 2 cm of variation at each site), compared to the fitting of the data set without Saint-Jean-de-Luz.

5 Discussion

5.1 Homogeneity of the selected region

This research has shown the homogeneity of the French surges of the region composed of the harbours of the Atlantic coast and the Channel coast. These sites have hydrological behaviours of surges which are statistically similar. In addition, the integration or not of the surges at Saint-Jean-de-Luz to the regional data set does not compromise the homogeneity of the region, the impact being negligible on the estimates of the extreme surges. The homogeneity of the region has

also been checked when excluding the surge sample at Brest (which has the largest observation period – 148 yr – and an outlier). Thus, the region composed by the sites along the Atlantic coast and the Channel coast seems to be robust, with surge estimates that do not significantly depend on the sites chosen to be integrated to the regional data set.

However, the negative value of the heterogeneity measure according to the test of Hosking and Wallis (1997) seems to indicate a possible correlation among the different sites of the region. Some studies have shown the presence of cross-correlation may significantly reduce the power of the heterogeneity measure and lead, for instance, to the error which considers a region as acceptably homogeneous whereas it is a possibly heterogeneous cross-correlated region (Castellarin et al., 2008). It could thus be interesting to refine the tests of homogeneity by analysing the possible correlations between the different sites.

In addition, a site-by-site analysis with a discordancy measure, as the one proposed by Hosking and Wallis (1997) and widely applied in hydrological studies (Van Gelder and Neykov, 1998; Abolverd and Khalili, 2010; Saf, 2009, 2010), could be used to study the homogeneity of the region with a more refined scale.

The case of Toulon could also be interesting to analyse, in particular to compare the statistical behaviours of the surges along the Atlantic coast or the Mediterranean Sea, the hydrographical basins and low-pressure weather systems being very different. However, a previous suitable retrieval of the skew surges for the Mediterranean sites is necessary.

Some local components of surges can also exist. A prevailing wind, the bathymetry (Weaver and Slinn, 2010), or the particular configuration of a site (such as the bay of Saint-Jean-de-Luz?) could increase or decrease the surge phenomenon. Then these components question the use of the regional frequency analysis because of a potential failure of homogeneity. The local specificities could be taken into account in a better way, owing to a more refined standardization of the surges. A new definition could be $(surge-a)/b$, with a and b which are dependent of the site (for instance, a could represent the sea level rise which is observed at the site and b could represent the annual empirical surge at the site).

Lastly, we can imagine that the batch of outliers – identified with at-site analyses – could be the consequence of phenomena different from the ones at the origin of all the other observed surges. In this case, and although the mixed exponential distribution can be used for a heterogeneous population, the approach to determine extreme sea surges could be completely different, with careful attention to the instigator phenomena of the surges. To go further in this direction, a detailed analysis is necessary to know for each site the weather conditions in favour of extreme sea surges. It is interesting to point out that some studies based on wind models over the North Sea and associated with climate change (considering estimated CO₂ concentrations according to a period in the future) seem to suggest the existence of a second population in the extreme wind distribution, some extratropical “superstorms” with more extreme winds than expected from extrapolation, originating from a different kind of meteorological system (Van Den Brink et al., 2005).

5.2 Comparison of the results obtained by the usual fittings or the regional frequency analysis

Generally, the surges which are estimated at each site by a regional frequency analysis are greater than the estimates obtained by at-site fittings. These differences are more or less significant according to the site and can be important for the harbours where no outlier has been measured.

Indeed, with the regional frequency analysis, an extreme event which is observed at one site of the region is considered to have the same probability to be observed in another site of the same hydrological region. Therefore, the surge estimates integrate the potential occurrence of outliers, even for a harbour where no outlier has been observed.

Table 3. Annual empirical surges and 1000-yr return surges at each site for at-site fitting or regional frequency analysis.

site	annual empirical surge (cm)	1000-yr return surges (cm)			
		at-site fitting		regional fitting	
		quantile	CL +70% ^a	quantile	CL +70% ^a
Dunkerque	79.6	192	212	240	289
Calais	66.9	140	155	202	243
Boulogne	64.8	157	180	196	236
Dieppe	60.4	114	128	182	220
Le Havre	66	179	201	199	240
Cherbourg	48.5	100	110	146	176
St Malo	61	147	175	184	222
Roscoff	49.7	97	106	150	181
Brest	55.5	126	133	167	202
Le Conquet	50.9	125	139	153	185
Concarneau	54.6	128	155	165	198
Port Tudy	51.9	131	147	157	189
Le Crouesty	56.3	105	122	170	205
St Nazaire	69.1	126	138	208	251
St Gildas	59.8	135	152	180	217
Olonne	57.7	111	123	174	210
La Rochelle	60.1	136	153	181	218
Port Bloc	59.1	126	142	178	215
Arcachon	57.7	135	165	174	210
Bayonne	46.2	131	149	139	168

^a CL+70% : value of the surge at the highest 70% confidence limit.

5.3 Dependence between the surges of the regional data set

For this study, the regional effective duration is determined as the sum of all the effective durations. But the regional observations may be dependent because several events can be caused by the same storm. Taking this dependence into account should reduce the effective duration of the regional data set and thus we should expect greater results than the ones presented in this study. However, throughout the period between 1846 and 2008 (global period of observation considering all the sites), we know that we can have independent events the same day but at different sites. As an illustration, a strong storm in the area of Dieppe would produce surges at Dieppe but also at Dunkerque since the two sites are quite close. If the storm occurs during the high tide at Dieppe, the skew surges observed at Dieppe and Dunkerque this same day have high probabilities to be dependent given the low interval of time between the high tides at these sites (according to the predicted high tides given by the SHOM for the period 2008–2010, the high tide at Dunkerque occurs about one hour on average after the high tide at Dieppe). But a surge at Dunkerque and a surge at Bayonne occurring the same day are probably independent, in particular if the surges are “medium” and so most likely the consequence of a local depression.

Thus we know the real effective duration of observation for the whole region is halfway between the 162 yr of real observation (period between 1846 and 2008) and the 601 yr corresponding to the sum of all the at-site durations. A way to approximate the real effective duration could be to determine an equivalent scenario with only independent data. An equivalent duration of observation would be estimated, integrating the potential dependence between the regional observations. A similar approach already exists for the study of rainfall in the region of “Grand Lyon” in France (Chocat, 2009). However, given the complexity of the determination of independent data between the different sites, we have chosen to make the hypothesis of the sum of the at-site durations in order to point out the fact that the regional duration was higher than the period of real observation (1846–2008).

As explained before, we know an “ideal” sample is composed of events which are really independent, with a physical point of view. But this involves doing a hydrometeorological analysis to determine the physical dependencies between the events (e.g. to check if two surges result from the same storm or not). Given the difficulty of a hydrometeorological analysis, this option was rejected in this study. As for the effective duration of observation, the choice of gathering the independent data at all the sites is a hypothesis to approximate the ideal sample. Another alternative or hypothesis could be to consider only the events occurring at three days of interval, for instance, whatever the sites. But this option eliminates events which are not dependent, as described above. However, this case is worth testing to deepen the dependence issue. First results show a different empirical distribution and the presence of several outliers seems to lead to a loss of a regional sample benefit.

In addition, some studies seem to show that the direct analysis of all exceedances over a high threshold, without considering a criterion of independence of the surges but making appropriate adjustments to allow for the dependence, can reduce to negligible levels the underestimation of the parameters which is incurred with the classical POT method and the maximum likelihood approach (Fawcett and Walshaw, 2007). As a consequence, return levels would be systematically underestimated for strong correlation in the sample. Thus an additional study could be carried out in parallel, keeping all the exceedances over a given threshold in order to compare the new high-return surge estimates with the results of this present study to quantify a possible underestimation.

5.4 Stationarity of the samples of surges

This study considers the samples of surges are stationary. However, this stationarity is not necessarily verified for each sample. A reason could be the amount of gaps in the observations, more or less important according to the tide gauge and the years of observation. A more complete sampling of the observed surges would be better for a good estimating of extreme storm surges.

A seasonality effect could also impact the stationarity of the surges because the highest surges seem to occur more often in winter or autumn (Vilibić and Šepić, 2010; Tsimplis and Shaw, 2010). Dividing the global analysis by seasons could be thus imaginable to take into account this seasonality. The estimates of extreme surges would depend on the season or would be the results of a combination of the different seasonal distributions (Laborde et al., 2009).

In addition, this study only considers the impact of climate change on the mean sea level (with a correction by the average sea level rise at a given point) and not on the surges. However, the climate change can lead to evolution in the height of the surges. It is interesting to point out that several studies have examined the role of climate change in a potential evolution of storminess and the height of the surges in the future (Van Den Brink et al., 2005; Lowe and Gregory, 2005; Sterl et al., 2009). These studies are based on improved physical models forcing scenarios of a future climate system with increased atmospheric concentrations of greenhouse gases to simulate surges. These models seem to show a small increase in the extreme wind speeds over the North Sea. However, the consequences on the height of the surges differ according to the studies and the location. Lowe and Gregory (2005) have shown the surge height will mainly increase around the UK coastline, whereas Sterl and al. (2009) have found no change along the Dutch coast within the limits of natural variability. In actuality, it is really difficult to evaluate with precision the behaviour of the surges in view of changing climate, because of the many uncertainties in the emission of greenhouse gases, in the methodologies to response to these gases to estimate surges, and in the natural climate variability (Lowe and Gregory, 2005).

6 Conclusions

With the regional frequency analysis carried out in this study, extreme surges due to strong storms (1979, 1987, 1999...) that current at-site fittings cannot recreate do not appear as outliers for the regional data set, given the assumptions of this study. Therefore, the regional frequency analysis seems to be a promising way to process the outliers in at-site samples of surges, as these outliers make the usual statistical fittings really questionable.

However, the methodology presented in this study also presents some limitations linked with the dependence between the data, the physical knowledge of the sites or the phenomena, etc. Several improvement trails are worth considering, as, for instance, the case of strictly independent skew surges determined by interpolation of the tide signals. A particular issue in this study is also the amount of events corresponding to extreme storms, which does not appear sufficient. For instance, the storm of 1987 is very represented in the regional data set, whereas some extreme surges, in particular the one of 1999 which was observed in several harbours,

are not in the regional sample because of a failure of the tide gauges of the SHOM, for example. So it is necessary to enlarge the regional data set to obtain the most complete possible surge sample, with British data for instance. The validity of this method to estimate extreme sea surges could thus be more precisely evaluated.

Edited by: U. Ulbrich

Reviewed by: E. D'Onofrio and two other anonymous referees

References

- Abolverdi, J. and Khalili, D.: Development of regional rainfall annual maxima for Southwestern Iran by L-moments, *Water Resour. Manag.*, 24, 2501–2526, 2010.
- Bode, L. and Hardy, T. A.: Progress and recent developments in storm surge modeling, *J. Hydraul. Eng.*, 123, 315–331, 1997.
- Castellarin, A., Burn, D. H., and Brath, A.: Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem?, *J. Hydrol.*, 360, 67–76, 2008.
- Chocat, B.: Intensités extrêmes de pluie sur le territoire du Grand Lyon. Analyse des données du Grand Lyon. Période 1987-2007, Conseil Supérieur de la Météorologie, Rapport, 2009.
- Coles, S.: *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, 2001.
- De Zolt, S., Lionello, P., Nuhu, A., and Tomasin, A.: The disastrous storm of 4 November 1966 on Italy, *Nat. Hazards Earth Syst. Sci.*, 6, 861–879, doi:10.5194/nhess-6-861-2006, 2006.
- Fawcett, L. and Walshaw, D.: Improved estimation for temporally clustered extremes, *Environmetrics*, 18, 173–188, 2007.
- Fritsch, F. N. and Carlson, R. E.: Monotone piecewise cubic interpolation, *SIAM J. Numer. Anal.*, 17, 238–246, 1980.
- Gerritsen, H.: What happened in 1953? The big flood in the Netherlands in retrospect, *Philos. T. Roy. Soc. A*, 363, 1271–1291, 2005.
- 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.
- Hosking, J. R. M. and Wallis, J. R.: *Regional Frequency Analysis – An approach based on L-moments*, Cambridge University Press ed., 1997.
- Kjeldsen, T. R., Smithers, J. C., and Schulze, R. E.: Regional frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method, *J. Hydrol.*, 255, 194–211, 2002.
- Laborde, J. P., Gourbesville, P., and Rebour, V.: Sea Levels in Brest from 1860 to 2004: evolution of the average level and statistics of extreme sea surges, *World City Water Forum 2009, Innovation and Harmony of Waters and Cities*, Songdo Convensia, Incheon, Korea, 18–21 August, 2009.
- Lang, M., Rasmussen, P., Oberlin, G., and Bobée, B.: Echantillonnage par valeurs supérieures à un seuil: modélisation des occurrences par la méthode du renouvellement, *Revue des Sciences de l'Eau*, 10, 279–320, 1997.
- Liao, M. and Shimokawa, T.: A new goodness-of-fit test for type-I extreme-value and 2-parameter Weibull distributions with estimated parameters, *J. Stat. Comput. Sim.*, 64, 23–48, 1999.
- Lowe, J. A. and Gregory, J. M.: The effects of climate change on storm surges around the United Kingdom, *Philos. T. Roy. Soc. A*, 363, 1313–1328, 2005.
- Mai Van, C., Van Gelder, P., and Vrijling, H.: Statistical methods to estimate extreme quantile values of the sea data, *Proceedings of the Fifth International Symposium on Environmental Hydraulics*, 2007.
- Masson, J. M.: Un problème parmi d'autres dans l'analyse des distributions des variables hydrologiques: Les horsains (outliers), in: *Statistique impliquée*. Paris: ORSTOM (Colloques et Séminaires), Séminfor 5: cinquième séminaire informatique de l'ORSTOM, 1991/09/02-04, Montpellier (France), 303–311, 1992.
- Miquel, J.: Guide d'estimation des probabilités des débits de crue-Parties 1 et 2, EDF-Direction des études et recherches, E43/81.45, 1981.
- Olbert, A. I. and Hartnett, M.: Storms and surges in Irish coastal waters, *Ocean Model.*, 34, 50–62, 2010.
- Ouarda, T., Lang, M., Bobée, B., Bernier, J., and Bois, P.: Synthèse de modèles régionaux de crue utilisés en France et au Québec, *Revue des Sciences de l'Eau*, 12, 155–182, 1999.
- Ouarda, T., Bobée, B., and St-Hilaire, A.: Synthèse des développements récents en analyse régionale des extrêmes hydrologiques, *Revue des Sciences de l'Eau*, 21, 219–232, 2008.
- Pandey, M. D., Van Gelder, P. H. A. J. M., and Vrijling, J. K.: Dutch case studies of the estimation of extreme quantiles and associated uncertainty by bootstrap simulations, *Environmetrics*, 15, 687–699, 2004.
- Pirazzoli, P. A.: *Projet Discobole*, CNRS, Laboratoire de Géographie Physique (UMR no 8591), 2007.
- Plag, H.-P.: Recent relative sea-level trends: an attempt to quantify the forcing factors, *Philos. T. Roy. Soc. A*, 364, 821–844, 2006.
- Ribatet, M.: Consolidation de l'information hydrologique disponible localement et régionalement pour l'estimation probabiliste du régime des crues, PhD Thesis, Institut National Polytechnique de Grenoble, 2007.
- Ribatet, M., Sauquet, E., Grésillon, J.-M., and Ouarda, T. B. M. J.: A regional Bayesian POT model for flood frequency analysis, *Stoch. Env. Res. Risk A.*, 21, 327–339, 2006.
- Saf, B.: Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey, *Water Resour. Manag.*, 23, 531–551, 2009.
- Saf, B.: Assessment of the effects of discordant sites on regional flood frequency analysis, *J. Hydrol.*, 380, 362–375, 2010.
- Saporta, G.: *Probabilités, analyse des données et statistique*, edited by: Technip, E., Paris, 2006.
- Simon, B.: *Statistique des niveaux marins extrêmes le long des côtes de France*, SHOM Rapport no. 001/94, 1994.
- Simon, B.: *La marée océanique côtière*, Institut Océanographique-SHOM ed., 433 pp., 2007.
- St-Hilaire, A., Ouarda, T. B. M. J., Lachance, M., Bobée, B., Barbet, M., and Bruneau, P.: La régionalisation des précipitations: une revue bibliographique des développements récents, *Revue des Sciences de l'Eau*, 16, 27–54, 2003.
- Sterl, A., van den Brink, H., de Vries, H., Haarsma, R., and van Meijgaard, E.: An ensemble study of extreme storm surge related water levels in the North Sea in a changing climate, *Ocean Sci.*, 5, 369–378, doi:10.5194/os-5-369-2009, 2009.
- Testut, L., Wöppelmann, G., Simon, B., and Téchiné, P.: The sea

- level at Port-aux-Français, Kerguelen Island, from 1949 to the present, *Ocean Dynam.*, 56, 464–472, 2006.
- Thompson, K. R., Bernier, N. B., and Chan, P.: Extreme sea levels, coastal flooding and climate change with a focus on Atlantic Canada, *Nat. Hazards*, 51, 139–150, 2009.
- Tsimplis, M. N. and Shaw, A. G. P.: Seasonal sea level extremes in the Mediterranean Sea and at the Atlantic European coasts, *Nat. Hazards Earth Syst. Sci.*, 10, 1457–1475, doi:10.5194/nhess-10-1457-2010, 2010.
- Turner, R. E.: Tide gauge records, water level rise, and subsidence in the Northern Gulf of Mexico, *Estuaries*, 14, 139–147, 1991.
- Van Den Brink, H. W., Können, G. P., and Opsteegh, J. D.: The reliability of extreme surge levels, estimated from observational records of order hundred years, *J. Coastal Res.*, 19, 376–388, 2003.
- Van Den Brink, H. W., Können, G. P., and Opsteegh, J. D.: Uncertainties in extreme surge levels estimates from observational records, *Philos. T. Roy. Soc. A*, 363, 1377–1386, 2005.
- Van Gelder, P. and Neykov, N. M.: Regional Frequency Analysis of Extreme Water Levels along Dutch Coast using L-Moments, *International Scientific Conference on Stochastic models of hydrological processes and their applications to problems of environmental preservation*, Moscow, Russia, 23–27 November, 1998.
- Van Zyl, J. M.: Application of the Kolmogorov-Smirnov test to estimate the threshold when estimating the extreme value index, *Communication In Statistics Part B – Simulation & Computation*, 40, 199–219, 2011.
- Vilibić, I. and Šepić, J.: Long-term variability and trends of sea level storminess and extremes in European Seas, *Global Planet. Change*, 71, 1–12, 2010.
- Weaver, R. J. and Slinn, D. N.: Influence of bathymetric fluctuations on coastal storm surge, *Coast. Eng.*, 57, 62–70, 2010.
- Whitcombe, L. J.: A Fortran program to calculate tidal heights using the simplified harmonic method of tidal prediction, *Computers and Geosciences*, 22, 817–821, 1996.