Utilisation des données historiques dans l'analyse régionale des aléas maritimes extrêmes : la méthode FAB

Roberto Frau

UNIVERSITÉ PARIS-EST

École Doctorale SIE

Laboratoire d'Hydraulique Saint-Venant

Thèse

Présentée pour l'obtention du grade de DOCTEUR

DE L'UNIVERSITE PARIS-EST

par

Roberto Frau

Utilisation des données historiques dans l'analyse régionale des aléas maritimes extrêmes : la Méthode FAB

Spécialité : Génie Côtier

Soutenue le 13 novembre 2018 devant un jury composé de :

Présidente du jury	**Pr. Anne-Catherine Favre**	(Grenoble INP, LTHE, ENSE^3)
Rapporteur	**Dr. Luc Hamm**	(Artelia Eau & Environnement)
Rapporteur	**Pr. Fernando Mendez**	(Universidad de Cantabria)
Examinateur	**Pr. Thierry Sauzeau**	(Université de Poitiers)
Directeur de thèse	**Dr. Pietro Bernardara**	(CEREA, EDF R&D)
Co-directeur de thèse	**Dr. Damien Violeau**	(LHSV, EDF R&D)
Co-encadrant de thèse	**M. Marc Andreevsky**	(EDF R&D)

Thèse effectuée au sein du **Laboratoire d'Hydraulique Saint-Venant**

de l'Université Paris-Est

6, quai Watier

BP 49

78401 Chatou cedex

France

Financements : ANR (bourse CIFRE # 2015/0925) et EDF R&D

Résumé

La protection des zones littorales contre les agressions naturelles provenant de la mer, et notamment contre le risque de submersion marine, est essentielle pour sécuriser les installations côtières. La prévention de ce risque est assurée par des protections côtières qui sont conçues et régulièrement vérifiées grâce généralement à la définition du concept de niveau de retour d'un événement extrême particulier. Le niveau de retour lié à une période de retour assez grande (de 1000 ans ou plus) est estimé par des méthodes statistiques basées sur la Théorie des Valeurs Extrêmes (TVE). Ces approches statistiques sont appliquées à des séries temporelles d'une variable extrême observée et permettent de connaître la probabilité d'occurrence de telle variable.

Dans le passé, les niveaux de retour des aléas maritimes extrêmes étaient estimés le plus souvent à partir de méthodes statistiques appliquées à des séries d'observation locales. En général, les séries locales des niveaux marins sont observées sur une période limitée (pour les niveaux marins environ 50 ans) et on cherche à trouver des bonnes estimations des extrêmes associées à des périodes de retour très grandes. Pour cette raison, de nombreuses méthodologies sont utilisées pour augmenter la taille des échantillons des extrêmes et réduire les incertitudes sur les estimations. En génie côtier, une des approches actuellement assez utilisées est l'analyse régionale.

L'analyse régionale est indiquée par Weiss (2014) comme une manière très performante pour réduire les incertitudes sur les estimations des événements extrêmes. Le principe de cette méthodologie est de profiter de la grande disponibilité spatiale des données observées sur différents sites pour créer des régions homogènes. Cela permet d'estimer des lois statistiques sur des échantillons régionaux plus étendus regroupant tous les événements extrêmes qui ont frappé un ou plusieurs sites de la région.

De récents travaux sur la collecte et l'utilisation des événements du passé dans les analyses statistiques des extrêmes ont montré le rôle essentiel de ceux-ci sur les estimations. Lorsqu'on estime des événements extrêmes, les événements historiques, s'ils sont disponibles, doivent être considérés pour calculer des niveaux de retour fiables. Etant collectés grâce à différentes sources, les événements du passé sont des observations ponctuelles qui ne proviennent d'aucune série temporelle. Pour cette raison, dans la plupart des cas, aucune information est disponible sur les éventuels événements extrêmes qui sont survenues avant et après cette observation du passé. Cela ainsi que le caractère particulier de chaque événement historique ne permet pas son utilisation dans une analyse régionale classique.

Une méthodologie statistique appelée FAB qui permet de réaliser une analyse régionale tenant en compte les données historiques est développée dans ce manuscrit. Elaborée pour des données POT (Peaks Over Threshold), cette méthode est basée sur une nouvelle définition d'une durée d'observation, appelée durée crédible, locale et régionale et elle est capable de tenir en compte dans l'analyse statistique les trois types les plus classiques de données historiques (données ponctuelles, données définies par un intervalle, données au dessus d'une borne inferieure). En plus, une approche pour déterminer un seuil d'échantillonnage optimal est définie dans cette étude.

La méthode FAB est assez polyvalente et permet d'estimer des niveaux de retour soit dans un cadre fréquentiste soit dans un cadre bayésien. Une application de cette méthodologie est réalisée pour une base de données enregistrées des surcotes de pleine mer (données systématiques) et 14 surcotes de pleine mer historiques collectées pour différents sites positionnés le long des côtes françaises, anglaises, belges et espagnoles de l'Atlantique, de la Manche et de la mer du Nord.

Enfin, ce manuscrit examine la problématique de la découverte et de la validation des données qui représentent les évènements du passé.

Using historical data in the Regional Analysis of extreme coastal events: the FAB method

Abstract

The protection of coastal areas against the risk of flooding is necessary to safeguard all types of waterside structures and, in particular, nuclear power plants. The prevention of flooding is guaranteed by coastal protection commonly built and verified thanks to the definition of the return level's concept of a particular extreme event. Return levels linked to very high return periods (up to 1000 years) are estimated through statistical methods based on the Extreme Value Theory (EVT). These statistical approaches are applied to time series of a particular extreme variable observed and enables the computation of its occurrence probability.

In the past, return levels of extreme coastal events were frequently estimated by applying statistical methods to time series of local observations. Local series of sea levels are typically observed in too short a period (for sea levels about 50 years) in order to compute reliable estimations linked to high return periods. For this reason, several approaches are used to enlarge the size of the extreme data samples and to reduce uncertainties of their estimations. Currently, one of the most widely used methods in coastal engineering is the Regional Analysis.

Regional Analysis is denoted by Weiss (2014) as a valid means to reduce uncertainties in the estimations of extreme events. The main idea of this method is to take advantage of the wide spatial availability of observed data in different locations in order to form homogeneous regions. This enables the estimation of statistical distributions of enlarged regional data samples by clustering all extreme events occurred in one or more sites of the region.

Recent investigations have highlighted the importance of using past events when estimating extreme events. When historical data are available, they cannot be neglected in order to compute reliable estimations of extreme events. Historical data are collected from different sources and they are identified as data that do not come from time series. In fact, in most cases, no information about other extreme events occurring before and after a historical observation is available. This, and the particular nature of each historical data, do not permit their use in a Regional Analysis.

A statistical methodology that enables the use of historical data in a regional context is needed in order to estimate reliable return levels and to reduce their associated uncertainties.

In this manuscript, a statistical method called FAB is developed enabling the performance of a Regional Analysis using historical data. This method is formulated for POT (Peaks Over

Threshold) data. It is based on the new definition of duration of local and regional observation period (denominated credible duration) and it is able to take into account all the three typical kinds of historical data (exact point, range and lower limit value). In addition, an approach to identify an optimal sampling threshold is defined in this study. This allows to get better estimations through using the optimal extreme data sample in the FAB method.

FAB method is a flexible approach that enables the estimation of return levels both in frequentist and Bayesian contexts. An application of this method is carried out for a database of recorded skew surges (systematic data) and for 14 historical skew surges recovered from different sites located on French, British, Belgian and Spanish coasts of the Atlantic Ocean, the English Channel and the North Sea. Frequentist and Bayesian estimations of skew surges are computed for each homogeneous region and for every site.

Finally, this manuscript explores the issues surrounding the finding and validation of historical data.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

La connaissance de la probabilité d'occurrence des événements maritimes extrêmes est indispensable pour concevoir et préserver les installations côtières et, notamment, les parcs nucléaires en bord de mer contre le risque inondation. Ces probabilités peuvent être estimées à partir des éléments de base de la Théorie des Valeurs Extrêmes et, en particulier, du concept de niveau de retour : un événement associé à une période de retour de T ans correspond à un événement qui a une probabilité d'être dépassé une fois tous les T ans en moyenne.

Dans le passé, les niveaux de retour des aléas maritimes extrêmes étaient évalués à travers des analyses statistiques locales. Le principe de ces analyses est d'utiliser la TVE sur des événements extrêmes observés sur un unique site. Dans la plupart de cas, les résultats des applications locales montrent d'importantes incertitudes liées aux estimations des événements extrêmes.

Pour ces raisons, différentes approches alternatives comme l'analyse régionale et l'utilisation des données historiques sont actuellement utilisées en génie côtier pour améliorer les estimations des événements extrêmes.

L'analyse régionale est une méthodologie statistique qui permet d'identifier tous les événements extrêmes qui se sont produits dans une région et de les utiliser dans l'analyse statistique. Cette approche conduit souvent à des estimations plus fiables à travers l'exploitation d'un plus grand nombre des données extrêmes.

Différents types d'approches régionales ont été proposés depuis le 1960 pour les applications environnementales. En ingénierie maritime, l'approche la plus récente et complète est l'Analyse Fréquentielle Régionale proposée par Weiss (2014).

L'apport des données historiques dans les analyses statistiques locales est une autre approche utilisée qui permet d'obtenir des estimations liées à des incertitudes réduites. La plupart de ces analyses sont permises grâce à l'utilisation du concept de seuil de perception (Gaume et al., 2010 ; Payrastre et al., 2011 ; Bulteau et al., 2015 ; Hamdi et al., 2015). Cette notion permet d'évaluer la période d'observation historique en supposant l'exhaustivité des données historiques au-dessus de ce seuil de perception.

La combinaison de ces deux approches pourrait contribuer à améliorer encore plus les estimations des événements extrêmes. Cependant, l'utilisation des données historiques dans un contexte régional est un sujet compliqué à traiter en génie maritime. En effet, la connaissance de la période d'observation est indispensable pour réaliser une analyse statistique. Les aléas maritimes historiques manquent souvent d'informations supplémentaires concernant leurs périodes d'observations. Pour cette raison, aucune hypothèse d'exhaustivité ne peut être vérifiée et donc le concept de seuil de perception ne peut pas être utilisé. La définition d'un nouvel élément qui permet d'estimer de manière crédible la durée d'observation des données historiques est requise.

Une méthodologie régionale (appelé FAB) basée sur le nouveau concept de durée crédible est développé dans ce manuscrit. Elle permet de considérer tous les types des données historiques les plus communes dans l'analyse régionale des événements maritimes extrêmes. Enfin, des estimations fréquentistes et bayésiennes peuvent être évaluées avec l'application de cette méthode.

1.1 Thesis context

1.1.1 Industrial framework

The protection of coastal infrastructures and, in particular, nuclear power plants from flooding is an overriding priority for EDF. The characterisation of extreme sea weather conditions allows the design of suitable protections in order to guarantee the safety of nuclear fleets against flood risk.

Fig. 1 - The Blayais Nuclear Power Plant located in the Gironde estuary near Blaye (France)

EDF manages the flood risk of its nuclear fleets originated by external hazards in accordance with the French regulatory framework (ASN, 2013). In particular, different situations of flood risk are defined, all of which must be used to design protections in order to preserve the safety of nuclear fleets. Flood of a platform on which a nuclear power plant is installed can provoke water infiltration in internal rooms containing important elements for the nuclear safety. Floods can also indirectly induce negative effects such as the accumulation of different types of debris in water intake of pump stations that can cause problems on cooling circuits.

The flood risk linked to high sea levels requires the computation of a high sea level of reference. This high sea level of reference enables the design and the verification of protections for nuclear power plants located on the coastline.

The French regulatory framework (ASN, 2013) provides guidelines for the evaluation of the high sea level of reference of each particular nuclear power station. The high sea level of reference must be calculated as the sum of the maximum astronomical tide, the upper bound of 70% confidence interval of the skew surge associated to a return period of 1000 years and the mean sea level trend. If the maximum astronomical tide is evaluated by results obtained from local or regional measurement stations, extreme skew surges are evaluated by statistical methods.

The estimation of skew surges associated to a return period of 1000 years can be performed both by local analysis and by regional analysis applied to time series of reliable observations. Moreover, the ASN (2013) clarifies that historical skew surges have to be considered in statistical estimations. Local analysis has several limitations as, for instance, a proper consideration of outliers (very rare events) in the extreme skew surge estimations. For this reason, the use of a regional analysis is allowed by the ASN (2013) only if this statistical methodology is able to show the suitability of the observed outliers.

The processing of a proper regional analysis methodology that allows the exploitation of all spatial data and historical skew surges available is needed in order to be able to estimate extreme skew surges that properly take into account the outliers.

1.1.2 Scientific framework

For the design of defence's coastal infrastructures, the estimation of the occurrence probability of extreme sea water levels is required. Probabilities of extreme events can be computed following the basic elements of Extreme Value Theory (EVT).

EVT is a specific branch of statistics that was introduced at the beginning of the 20^{th} century by Frechet (1927) and Fisher and Tippett (1928). This is later developed by the studies of Gnedenko (1943), Weibull (1951), Gumbel (1958), Picklands (1975), Coles (2001) and Beirlant et al (2004). EVT provides a set of probability distributions for extremes that enables the estimation of the occurrence probability of extreme events through the use of the concept of return period. An event associated to a return period of T years corresponds to the event that has a probability 1/T to be exceeded every year over a period of T years or, in other words, it is the event that is exceeded on average once in T years.

Statistical approaches associated to the basic elements of EVT are used in order to estimate sea levels linked to high return periods. In particular, the two most common statistical approaches for practical extreme value analysis rely respectively on Block Maxima (or Annual Maxima) data and on Peaks Over Threshold data (Coles, 2001). BM data typically converges in a Generalised Extreme Value distribution while POT data classically converges in a Generalised Pareto Distribution.

Moreover, direct or indirect estimation of extreme sea levels can be computed depending on which variable is analysed. Direct estimations are calculated by statistical analyses applied directly on sea level without considering its deterministic and stochastic parts. In fact, sea level can be schematically represented as the overlap of the predicted astronomical tide and the surge (respectively its deterministic and its stochastic part, Fig. 8). Indirect approaches consider separately these two components of sea level assuming extreme value analyses of surges as more convenient. Astronomical tides and surges are successively recombined in order to estimate extreme sea levels. In particular, convolution methods might be used to define the probability distribution of sea levels in indirect approaches (Kergadallan, 2013). Indirect approaches are preferred for sites in which tidal range between high and low tides are significant (Weiss, 2014).

Regardless of the maritime variable considered, efficient statistical methods and a sufficient number of data are required in order to get reliable estimations of extremes. In the past, data originated from a single tide gauge was used to fit statistical distributions for a particular site and extreme estimations produced were usually full of uncertainties. This may be explained by the fact that the duration of recordings could not be too long to estimate reliable extreme variables. This is the case of sea levels where usually recordings last from 30 to 50 years. For this reason, estimations of extreme sea levels (associated to a return period of up to 1000 years) achieved by local analysis are in the most part not suitable to be used for engineering applications.

Nowadays, innovative statistical methods are used to enlarge the sample of extreme data. Regional analysis and the use of historical data are two relevant approaches that can applied to improve extreme estimations.

Regional analysis is able to exploit the wide availability of data in different locations. Creating larger samples of regional data typically results in reduced uncertainties of the extreme estimations.

Historical data are generally used to extend local samples considering very strong events occurred in the past. In this way, statistical analyses are applied to a bigger sample of extreme data.

The combination of these two methodologies and therefore a regional analysis with historical data allows the extension of extreme data samples, and in effect exploiting all the spatial and historical information available. The creation of a regional sample that includes historical data further improves the estimations of extreme events.

1.2 State of art

In this section, methods developed and currently used for environmental applications by different authors on the use of historical data and on the regional analysis are presented.

1.2.1 Regional Analysis

Regional Analysis is a statistical methodology that is able to exploit all the data available in a region to reduce uncertainties on the extreme estimations. The main idea of Regional Analysis is to pool similar sites into a homogeneous region, to create an extended regional data sample and, in this way, to estimate a regional distribution.

Regional methods are widely used to improve statistical estimations of extreme events linked to environmental applications since 1960. Dalrymple (1960), Stedinger (1983), Cunnane (1988), Madsen and Rosbjerg (1997), Ouarda et al. (1999), De Michele and Rosso (2001), Javelle et al. (2002), Kjeldsen et al. (2002), Merz and Bloschl (2003), Viglione et al. (2007) and Saf (2009) use Regional Analysis to compute extreme river discharges, Schaefer (1990), Hosking and Wallis (1997), Alila (1999) and Borga et al. (2005) to estimate extreme rainfalls and Goel et al. (2004), Sotillo et al. (2006) and Escalante-Sandoval (2008) to calculate extreme winds.

Regional Analysis for several maritime variables is more recent. In order to improve estimations of extreme significant wave heights, Goda et al. (2010) and Goda (2011) use the regional method for 11 sites located in the Japanese east coasts and, in the same way, Val Gelder et al., 2000 applied a regional methodology to 9 locations of the North Sea. In addition, Regional Analysis is applied also to improve extreme estimations of the tsunamis run-up for 114 sites of the Pacific Ocean by Hosking (2012) and of the sea levels for 13 sites located in North Sea by Val Gelder and Neykov, 1998.

Moreover, several applications for skew surges has been achieved mainly in last years. Bernardara et al. (2011) and Bardet et al. (2011) estimate extreme skew surges by a regional analysis applied to respectively 18 and 21 sites located in French coasts of the Atlantic Ocean and the English Channel.

Last recent regional methodology is proposed and used for skew surges by Weiss et al. (2013) and Weiss (2014). This Regional Frequency Analysis approach is based on the index-flood method proposed by Dalrymple (1960) permitting to estimate a regional probability distribution that is common to all extreme observations up to a local index representing the local specificities of a site. Extreme events collected from different sites of a same homogeneous region follow the same regional probability (Hosking and Wallis, 1997; Weiss, 2014). Regional probability distribution must be computed for independent regional observations to estimate correctly extreme events. For this reason, Weiss (2014) proposes a model to cluster different storms and to treat statistically independent regional data in order to avoid a likely redondance of regional events.

Results of the RFA application shown as the regional analysis is able to reduce the uncertainties on the estimations of return levels. In particular, results obtained by Weiss (2014) shows that return levels computed by a classical local analysis contains 55% more of uncertainties on average in each site for return levels linked to 100 years of return period compared to the same local return levels computed by regional analysis. In addition, local outliers have typically a less extraordinary and unique nature in the regional analysis.

Nevertheless, the RFA approach can be applied only for continuous time series of gauged coastal events. Further details of this methodology are set out in Chapter 3.

1.2.2 Historical data in extreme value analyses

The use of historical data in the estimation of extreme events represents another method to improve the reliability of return levels. Historical data enables typically the extension of extreme data samples in which statistical analyses are applied. Moreover, they permit the consideration of all the extraordinary events that are happened in the past.

The extension of an extreme data sample using historical data is possible only after their collection and their validation. The collection and the validation of historical data is required before any statistical analysis that allows the use of them. In addition, the knowledge of the period in which historical data are the biggest ones observed is important for the statistical analysis of the extreme events (Leese, 1973; Prosdocimi, 2017). For these reasons, the collection and reconstruction of exceptional historical coastal events (Parent et al., 2007; Pouvreau, 2008; Garnier and Surville, 2010; Brehil, 2014; Brehil et al.,2014; Peret and Sauzeau, 2014; Giloy et al., 2017) and the reconstruction of local time series during the past period have been performed for sites located in the French Atlantic coasts (Ferret, 2016).

When historical data are collected and validated, they can be used to improve the estimations of extreme events. In the past, historical data have been used to get better estimations of extreme events by Benson (1950), Condie and Lee (1982), Cohn (1984), Hosking and Wallis (1986a), Hosking and Wallis (1986b), Stedinger and Cohn (1986), Stedinger and Baker (1987), Ouarda et al. (1998) and Benito et al. (2004).

Miquel (1981) and Lang et al. (1997) modelled statistically systematic POT data above a threshold and historical POT data above a perception threshold (Renewal Method). This formulation enables the use of different types of historical data in a local statistical analysis and it can be used only after the verification of the hypothesis of exhaustiveness of the historical data above the perception threshold. For coastal events, no local or regional information on the historical period is frequently available.

Some statistical studies enabling the use of historical data in local statistical analysis have been developed in hydrology. These studies (Gaume et al., 2010; Payrastre al., 2011; Payrastre et al., 2013) are based on the concept of the perception threshold and they are performed to estimate Bayesian return levels linked to high return period.

In coastal engineering, the use of historical data was being less developed in the past due to the weak availability of historical data. Currently, scientific studies that use coastal historical data are increasingly performed to improve estimations of extreme coastal events.

In particular, Baart et al. (2011) estimate return levels of storm surges using three reconstructed historical surges occurred in Dutch coasts in the 18th century. Hamdi et al. (2015) propose two different approaches to consider different types of historical information in statistical analysis. The first approach is block maxima method (BMH) and it can be performed for the three typical types of historical data. The second approach (the peaks-over-threshold method POTH) allow the use of two different types of historical data: historical maxima (HMax) data and over-a-threshold supplementary (OTS) data. These methodologies based on the perception threshold concept are applied for historical skew surges recovered in the site of La Rochelle (France). Bulteau et al. (2015) proposes the HIBEVA method to treat historical data in the Bayesian analysis of extreme sea levels. This method is elaborate for POT data. Using 8 historical sea levels, they show by the use of the historical threshold the value of the historical events in a statistical analysis of extreme events.

All of these studies applied to coastal variables are developed for local analysis.

1.2.3 Historical data in regional analyses

More complex and difficult is to find scientific studies that use historical data to get better regional return levels. The combination of these two methodologies has been recently used in hydrology by Nguyen et al. (2014) and Sabourin and Renard (2015). In particular, a multivariate peaks-over-threshold model is proposed by this last study. This model is applied to four French catchments in which historical data are available. Nguyen et al. (2014) propose a regional method that allows the use of historical data. This method is based on the regional method proposed by Hosking and Wallis (1997) and on the Bayesian approach formulated by Gaume et al. (2010) to introduce historical data in a local analysis. The application of this method provides regional Bayesian estimations adjusted with its credibility intervals. This method is applied for BM data obtained to two French catchments.

In coastal engineering, the combination of regional analysis and historical data is actually unexplored.

1.3 Main objectives

A scientific challenge currently exists to define a new methodology that combines two different approaches in the statistical estimations of extreme coastal events: regional analysis and the use of historical data. This thesis focusses on the development of this new methodology as well as its application on a skew surge database.

Many scientific studies have highlighted the importance of using historical data in the estimations of extreme coastal events. Exceptional events of the past have to be considered in statistical analyses of extreme coastal events in order to get reliable estimations. The addition of these data in local statistical analysis allows the estimations of extreme events associated to reduced uncertainties. In particular, the majority of these studies introduce historical data by the use of the concept of “perception threshold”. This concept is based on the exhaustiveness hypothesis of the observation period of historical events. Proving this hypothesis is currently challenging for the majority of coastal events of the past.

On the other hand, regional approaches permit the definition of larger extreme data samples considering all the extreme events impacting sites within a region. This allows the estimation of regional and local return levels associated to reduced uncertainties compared with return levels computed through local analyses. Many regional methodologies had been proposed in several scientific studies, but the RFA approach (Weiss, 2014) can be considered as one of the most complete and flexible regional methodologies for the estimations of extreme coastal events.

However, this approach can only be carried out with a database composed of systematic events. In particular, local continuous time series enables the computation of the period of observation of the regional events in the RFA approach. This effective regional duration is an essential element to estimate regional and local extreme events. On the contrary, historical data represents typically isolated events of past periods or periods without systematic measurements and no further observations are locally available for these periods. In addition, a historical event is not always described by an exact data value. Different types of historical data are often available and the RFA approach enables the use of exact data values.

The RFA method is developed for POT data. For this reason, the definition of the sampling threshold is important to achieve the best performance of the statistical method. No approaches for the identification of the local sampling thresholds are proposed in the RFA method. Another limit of the RFA approach is the impossibility to estimate return levels in a statistical Bayesian framework. In fact, the RFA approach was only developed to obtain frequentist estimations of extreme events. For these reasons, a new methodology, called FAB method, that preserves all the basic elements of the RFA approach is proposed in this manuscript.

The FAB method overcomes all of these issues caused by the addition of historical data in a regional analysis. This method is based on the new concept of “credible duration” that allows the estimation of the period of observation of the regional extreme data sample.

The FAB method permits the use of all most common types of historical data through the formulation of a new regional likelihood. Moreover, a weighting approach based on the definition of primary and secondary parameters is proposed to identify optimal local sampling thresholds for regional extreme data samples composed by POT data. Finally, frequentist or Bayesian estimations can be indifferently computed through the FAB method. The option to get frequentist or Bayesian estimations of regional and local return levels is user defined. This method is detailed in Chapter 3.

An application of the FAB method to a skew surge database composed by systematic and historical data is performed in Chapter 4. Frequentist and Bayesian return levels are estimated and appropriately compared.

Nevertheless, the discovery of valid historical data is a complicated process. Chapter 2 of this manuscript discusses in detail the difficulties faced during the collection and the validation of historical data, and an example procedure is proposed.

Chapter 2

HISTORICAL DATA

Ce chapitre vise à définir la donnée historique, sa nature et les problématiques associées à sa collection et validation.

Dans ce manuscrit, les données historiques sont des données ponctuelles qui représentent des évènements remarquables survenus pendant des périodes dans lesquelles les instruments de mesure n'ont pas enregistré la variable considérée. Les trois types les plus communs de données historiques seront considérés dans cette étude : la donnée exacte, un intervalle de valeurs et une borne inférieure dépassée par la valeur historique lors de l'événement.

La connaissance d'une ou plusieurs événements historiques est un sujet difficile à traiter. En effet, les données historiques sont liées à différentes types d'informations et de sources qui ne sont pas le plus souvent facilement interprétables et accessibles. Par conséquent, la collecte des données historiques est une procédure compliquée. Toute l'information historique repérée doit être exploitée et, dans la mesure du possible, convertie en données historiques que pourraient être ensuite utilisée pour des applications statistiques d'estimation des événements extrêmes. Pour ces raisons, une approche de collecte et de validation des données historiques suivie d'un exemple d'application est présentée dans ce chapitre. La collecte doit être démarrée à partir des bases des données historiques déjà existantes et étendue à des sources accessibles. Ensuite, une validation de toutes les données collectées doit être réalisée grâce à l'apport essentiel d'un historien compétant sur la période examinée.

2.1 Historical data

Historical data are typically quantitative or qualitative isolated data values representing an exceptional event of the past. In this manuscript, historical data are defined as all quantitative isolated data values representing an extraordinary event which has occurred in a period in which the considered variable was not recorded by a measuring device. For this reason, historical data represent not only events of the past but also exceptional events in which, for whatever reason, the measuring device was not working.

The discovery of historical coastal events that can potentially provoke coastal flooding is a major challenge for an engineer dealing with estimations of extreme coastal events. Finding quantitative data values of historical events is a complex procedure. Information of a particular past event can be used to detect or reconstruct a quantitative data associated to that event. After a wide investigation of different sources, this is also possible by means of a good knowledge of the period of the event that only a qualified historian may have. An approach to collect and validate historical data is proposed in the following chapter.

Over the last twenty years, there have been a number of extraordinary storms such as Xynthia (February 2010) and Lothar and Martin (December 1999) storms, which impacted the European Atlantic coasts. These storms caused several disasters and loss of human life. These have been sometimes defined in the media as exceptional events that have never previously occurred on French coasts until now, however this is incorrect. A wide investigation of all the historical storms which impacted a particular site is required to define the rarity of an extreme event. In particular, Garnier and Surville (2010) identify several catastrophic flood events on the French Atlantic coasts during the last 500 years with a similar vulnerability to the Xynthia storm. For this reason, also if a recent exceptional storm can be perceived by people as the strongest event ever happened in a particular location, it may have previously occurred because history often repeats itself.

Furthermore, Bulteau et al. (2015) conclude that, considering historical traces in La Rochelle, the Xynthia storm is linked to a shorter return period. For this reason, considering all the traces of the past in the statistical analysis of extreme events could help to reduce the impact of natural calamities along the coasts.

Several studies on past coastal floods impacting French coasts (Peret, 2004; Lambert and Garcin, 2013; Lang and Coeur, 2014; Peret and Sauzeau, 2014) show the difficulties faced researching useful information to retrace the history.

Any type of information is useful to retrace the past and to collect, reconstruct or validate a historical data. For example, a painting, a newspaper of the time, a preserved manuscript, a document from municipal archives, an ecclesiastical text, some private letters sent to ask for help from friends and family, a water mark left in an ancient church or building are only a few sources containing useful information of the past which can be used to retrace a historical event. In addition, the vulnerability evaluation of the society of the period is required in order to understand the relevance of a particular past event. This can be used to reply to the ever more frequently asked question regarding the occurrence of a remarkable flood.

Although finding the occurrence of an exceptional past event is challenging, the quantification of a historical event can be considered more complex. In fact, only a small portion of the identified historical events may be expressed with some quantitative data, for instance, the case of water marks or overflows of sea levels in a coastal town. In these and other cases, many attempts have to be performed to retrace and rebuild the numerical value of the considered variable that represents the historical event.

Numerical values of historical events, more simply denominated historical data, can be used in the statistical analysis for extreme events. The use of historical data in the statistical analysis improves significantly the estimations of extreme coastal events (Bulteau et al., 2015; Hamdi et al., 2015). In particular, Hamdi et al. (2015) show as the use of historical data are considered at La Rochelle decreases the uncertainties linked to the estimations of extreme skew surges. For this reason, the use of past events in the extremes' statistical analysis is required to get estimations of extreme events as reliable as possible.

Historical data can be regrouped in three major data types depending on the type of historical information found:

- Type I or exact data;
- Type II or data range;
- Type III or lower limit value of data.

Type I is a data point, which can be considered as the most precise type of historical data.

Type II is a numerical data range that contains the extreme data representing the historical event. In this case, two numerical values define the past event: the lower bound and the upper bound. This type of data may be considered, for instance, when different historical sources describe the same historical event with different quantitative values (Fig. 2).

Type III represents the minimum value that the extreme data has attained during the historical event. In this case, no more information is provided by the sources. This type of data is used when, for instance, a flooding of a port or a road is widely documented. In this case, the event has at least attained the port/road level. This kind of data, even though it is not very precise, can be used in extreme analysis.

The three different types of historical data can be seen as the result of the wide investigation on a multiplicity of sources. By the definition of this three types of most common historical data, many documentations of the past that may seem worthless are useful to reconstruct historical data exploitable in statistical analysis.

In any case, the analysis of different sources is complex due to their variety and to their degree of reliability. In particular, the reliability of each source depends on many factors. For instance, a historical data found in a newspaper tends to be overestimated in order to permit to the journal editor of the period to sell as many copies as possible.

In addition, all types of historical data are often associated to significant uncertainties related to the data value. These uncertainties can depend on the accuracy of the source and on the period of the event. In fact, when an extreme event of the past is documented, the period of the event must be analysed in order to understand, for instance, which measuring instruments were used, or the vulnerability of the society of that period to that particular event.

For these reasons, only a few quantitative information is currently available for coastal variables (in particular for skew surges).

2.2 Collection and validation approach

In the past, several attempts to collect information on historical floods have been made in France (Roche et al., 2014; Daubord et al., 2015; Lang et al., 2016) and in UK (Haigh et al., 2015; Haigh et al., 2017).

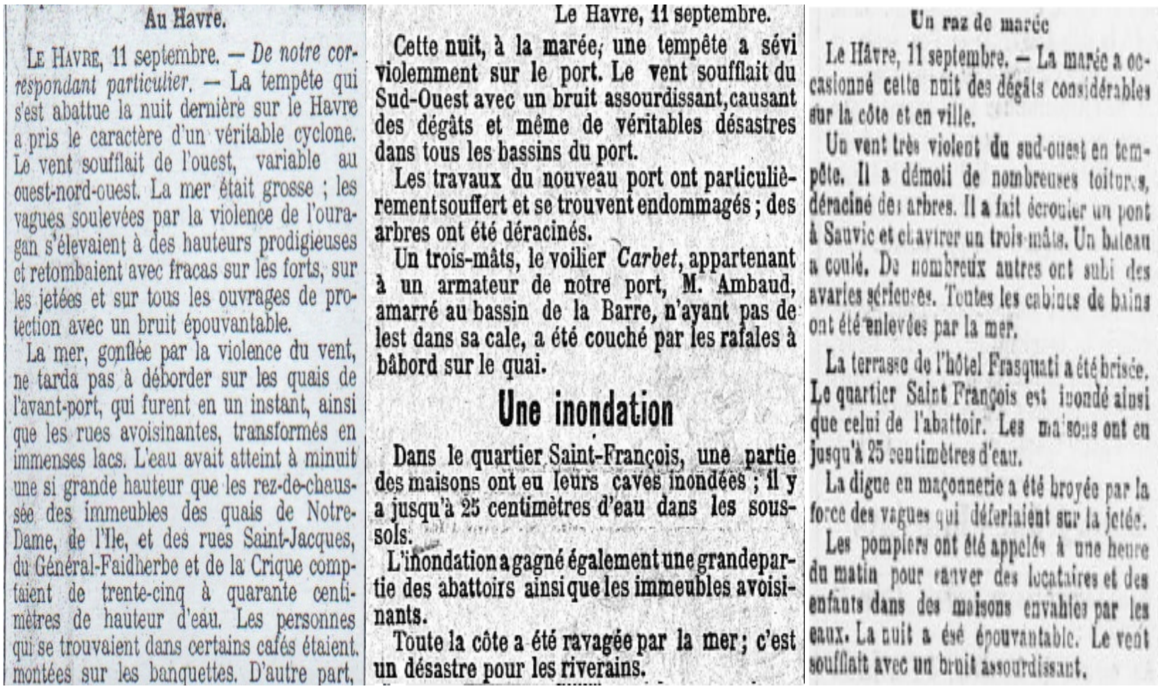
Nevertheless, the creation of a historical database of coastal events is needed in order to use historical data in statistical analyses of extreme events. An approach that enables the collection, the reconstruction and the validation of historical events is proposed.

2.2.1 Historical data collection

The collection of historical data is a complex task. Historical data can be available, for instance, in private, ecclesiastical or municipal archives or in many digitalised newspapers of the time and not ever you are aware of the availability of every single source containing important information on a past event. Besides, many sources refer to the original documents that are not readily accessible. If accessible, a supplementary work of reading and interpretation is necessary to extrapolate useful and quantitative data concerning a particular historical event.

Several phases are required to collect historical data. First, a deep analysis of previous studies that have already reviewed some past floods is suggested. This step allows the definition of a first basis of most documented extreme events occurred in the past. Then, all available sources as newspapers of the time or meteorological bulletins must be analysed in order to find past events or other useful quantitative information on events already known. The credibility of sources has to be checked. In fact, not all sources can be considered as accurate and valid.

Fig. 2 shows the difficulties faced on the interpretation and the credibility of sources describing historical sea levels. These three different newspapers (from the left Météo Paris, Le Petit Journal and Ouest-Éclair) provide different quantitative information for the same historical coastal flood occurred the 11th of September 1903 at Le Havre. The first one states that the sea level flooded roads and buildings by 35-40 centimetres, the second one said that the underground was flooded by 25 centimetres and the last one that the houses was submerged to 25 centimetres of water.



Au Havre.

LE HAVRE, 11 septembre. — *De notre correspondant particulier.* — La tempête qui s'est abattue la nuit dernière sur le Havre a pris le caractère d'un véritable cyclone. Le vent soufflait de l'ouest, variable au ouest-nord-ouest. La mer était grosse ; les vagues soulevées par la violence de l'ouragan s'élevaient à des hauteurs prodigieuses et retombaient avec fracas sur les forts, sur les jetées et sur tous les ouvrages de protection avec un bruit épouvantable.

La mer, gonflée par la violence du vent, ne tarda pas à déborder sur les quais de l'avant-port, qui furent en un instant, ainsi que les rues avoisinantes, transformés en immenses lacs. L'eau avait atteint à minuit une si grande hauteur que les rez-de-chaussée des immeubles des quais de Notre-Dame, de l'Ile, et des rues Saint-Jacques, du Général-Faidherbe et de la Crique comptaient de trente-cinq à quarante centimètres de hauteur d'eau. Les personnes qui se trouvaient dans certains cafés étaient montées sur les banquettes. D'autre part,

Le Havre, 11 septembre.

Cette nuit, à la marée, une tempête a sévi violemment sur le port. Le vent soufflait du Sud-Ouest avec un bruit assourdissant, causant des dégâts et même de véritables désastres dans tous les bassins du port.

Les travaux du nouveau port ont particulièrement souffert et se trouvent endommagés ; des arbres ont été déracinés.

Un trois-mâts, le voilier *Carbet*, appartenant à un armateur de notre port, M. Ambaud, amarré au bassin de la Barre, n'ayant pas de lest dans sa cale, a été couché par les rafales à bâbord sur le quai.

Une inondation

Dans le quartier Saint-François, une partie des maisons ont eu leurs caves inondées ; il y a jusqu'à 25 centimètres d'eau dans les sous-sols.

L'inondation a gagné également une grande partie des abattoirs ainsi que les immeubles avoisinants.

Toute la côte a été ravagée par la mer ; c'est un désastre pour les riverains.

Un raz de marée

Le Hâvre, 11 septembre. — La marée a occasionné cette nuit des dégâts considérables sur la côte et en ville.

Un vent très violent du sud-ouest en tempête. Il a démoli de nombreuses toitures, déraciné des arbres. Il a fait écrouler un pont à Sauvic et chavirer un trois-mâts. Un bateau a coulé. De nombreux autres ont subi des avaries sérieuses. Toutes les cabines de bains ont été enlevées par la mer.

La terrasse de l'hôtel Frasquati a été brisée. Le quartier Saint François est inondé ainsi que celui de l'abattoir. Les maisons ont eu jusqu'à 25 centimètres d'eau.

La digue en maçonnerie a été broyée par la force des vagues qui déferlaient sur la jetée.

Les pompiers ont été appelés à une heure du matin pour sauver des locataires et des enfants dans des maisons envahies par les eaux. La nuit a été épouvantable. Le vent soufflait avec un bruit assourdissant.

Fig. 2 – The flood occurred during the day of 11th September 1903 at Le Havre described by three different newspapers: Météo Paris, Le Petit Journal and Ouest-Éclair

Three different measurements for a same event are available. For this reason, a critical analysis on the source has to be properly performed. This analysis can be achieved by a historian that gives a degree of credibility to all types of sources discovered.

Sources provide not ever a direct quantitative information on the particular event of the past. Sometimes other quantitative data are available. This additional information can enable the reconstruction of a particular variable investigated.

For this reason, an approach to reconstruct historical skew surges is defined thanks to the stage performed by Florian Regnier at EDF R&D LNHE (2017). In particular, sea levels, astronomical tides, street or pier levels of the time are some quantitative data that can be used to reconstruct historical skew surges.

Fig. 3 shows the scheme that enables the retracing of a historical skew surge value neglecting the subsidence. In particular, knowing that the sea level exceeded the dock by the measure of *h* centimetres and we are able to know the dock level of the time *H*, we can compute the skew surge value as the difference between *H+h* and the astronomical tide level in the day of the event (for French ports they can be provided by the SHOM).

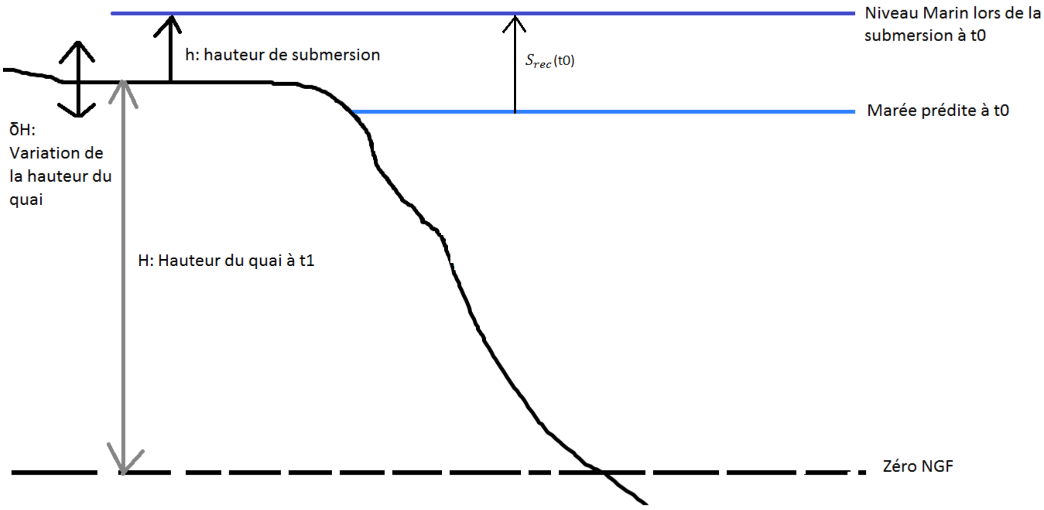


Fig. 3 - Scheme used to reconstruct 11 historical skew surges

Obviously, other factors have to be considered in this reconstruction. For instance, also if it is not difficult to know an actual dock level, it is less easy to know that level in a particular day of the past. Battles, world wars, reconstructions and restorations were made in every location and it is very hard to discover the exact level of a street or a building in the past. In addition, municipal archives are not always easily accessible to find this type of past information. This is a promising way to retrace historical skew surge but it is very challenging to realise. A similar approach can be formulated for other types of coastal variables.

In any case, the collection and reconstruction processes lead to identify some historical events that have to be validated. The use of invalid historical events could provide erroneous statistical estimations of extreme event. For this reason, historical events need a validation.

2.2.2 Historical data validation

The positive feedback of the historical expert on past event concerned allows the validation of historical data. In fact, knowing perfectly the period in which historical event is occurred, the historian can evaluate if that event is effectively happened. In addition, the historical expert can provide the degree of a credibility of a source and, in particular, if exaggerations have been written.

For instance, many deaths recorded in some archives after a storm can be also caused by other factors as plague or other diseases that impacted the society in the same time of the extreme event. These factors linked to the society have to be clearly considered when we validate an extreme event of the past.

The validation of one historical event by the expert is a long process. For this reason, some alternative ways to pre-validate a historical data can be considered.

In particular, for coastal historical events, many meteorological reanalyses and numerical models are currently available. Meteorological analyses help to know if it might be possible that a past event is occurred or not and if an exceptional coastal variable was feasible that day. Meteorological reanalyses contribute then to pre-validate historical events and they cannot replace the historical expert in any way. In particular, wind and pressure data are useful to validate physically the occurrence of a likely historical skew surge. In particular, a surge is principally generated to two factors: a low pressure and a high wind speed (and wind direction) at the sea surface (Heaps, 1983; Cariolet et al., 2011; Regnier et al., 2017). For this reason, if wind and pressure data or reanalyses are available, we can realise if an extreme skew surge can be happened during the day of the past event.

In fact, a validation of historical events based on reanalysis is not enough. Historical elements come from sources and period have to be taken in account in the validation process. Only the historical expert can definitively validate a historical event.

The approach proposed to validate historical events is composed by two phases. The first one is the pre-validation process that must be carried out through the support of physical factors extracted from proper reanalysis. The second and last phase is the final validation of a historical event. This must be performed by an historian after a wide investigation on sources of the period. The historian can focus on some elements (as the society, the epoch, the source etc.) that reanalyses cannot consider but that could be important to validate a historical event.

2.2.3 Numerical models and their role for historical events

Numerical models are another very promising approach that might be used in the future to pre-validate some particular extreme events of the past. In any case, this approach could be used only after a robust validation of the numerical model for well-known recorded extreme events, namely systematic events. The main idea is to compare historical data with the extreme values obtained by a numerical model.

Differently from meteorological models, this approach might also allow the pre-validation of a historical quantitative value in some special cases. In particular, an event of the past impacted the site A and the site B. The site A is provided by systematic data and the site B by historical data. Assuming that the site A is well modelled and, for another more recent event in which systematic data are available for the sites A and B, the numerical model provides good estimations, the historical data founded in the site B can be pre-validate by the numerical model.

This possible type of pre-validation of historical events can be allowed only after a robust validation of the numerical model for systematic extreme events. When the model is considered valid for extremes, the pre-validation of historical skew surge values by the numerical model can be achieved. An application for skew surges has been performed by Cécile Lalanne during her stage at EDF R&D LNHE (2018).

In particular, a numerical model of skew surges based on TELEMAC-2D is used. The validation of this model for extremes and some implementations to make it utilisable for extreme comparison is carried out. This work will be object to a talk that it will held during the XXVth TELEMAC-MASCARET User Conference (TUC) 2018 from 9th to 11th October in Norwich (UK). For more details, the conference paper is shown in Annexe D.

2.2.4 An example of historical skew surge collection

A preliminary collection of historical skew surges was performed through an event-by-event approach, gathering together skew surges of a same past event recorded in different locations.

Fourteen extreme events associated to 14 historical skew surges (further details on these historical data are available in Chapter 4) were collected during the first part of this PhD study for 3 French sites (La Rochelle, Dieppe and Dunkirk).

This collection was successively completed by Florian Regnier during his internship at EDF R&D LNHE in 2017. His work has contributed to review a total of 74 additional historical events occurred between 1705 and 1953 located in French side of English Channel and in French Atlantic coasts. Depending on types of numerical elements available, we focused especially on 5 of these 74 reviewed historical events that allowed the reconstruction of 11 skew surges: the event of the 1st January 1877 that impacted especially the Brittany and the Pays de la Loire regions, the event happened between the 4th and the 6th of December 1896 in Brittany and in the south of England, the storms occurred during the 11th September 1903 and during the 3rd February 1904 in the English Channel and the event of the 13th-14th March 1937 that impacted the French Atlantic coasts. In addition, other 6 historical skew surges were identified during this internship, for a total of 17 historical skew surges. All of these data collected must be after validated and, for this reason, an example of the pre-validation of historical events by meteorological reanalysis is shown in the following.

The collection and reconstruction performed have been the subject of an oral talk at the Conference EVAN 2017 (Advances in Extreme Value Analysis and application to Natural Hazard) held in Southampton (UK) between 5th and 7th September 2017 (Annexe D).

However, the collection of historical data, and in particular historical skew surges, is a common need for the scientific community dealing with the statistical estimations of extreme coastal events. For this reason, a Working Group is formed in France. The main aim of this group is to create a database of all historical events that are susceptible to cause a coastal flood (Giloy et al., 2018). For the moment, this French Working Group leaded by the IRSN is composed of EDF R&D, SHOM, Artelia, BRGM and the University of Poitiers. Each partner provides its knowledge on past events. Every historical event and their numerical values associated (as skew surges or sea levels) must be checked and validated by all members of the group including the historian.

The pre-validation approach is detailed in the following of this section with a practical example on the historical skew surges.

2.2.4.1 Pre-validation of historical skew surges by reanalysis

An application of pre-validation is performed for skew surges. Based on wind and pressure reanalysis of the 20th century, this analysis checks if an important skew surge is physically possible the day of the past event.

Many daily weather reports of the 20th century are available in France (Météo France) and in UK (Met Office). Fig. 4 shows the pressure chart and the daily weather report of Météo France for the 1st January 1877. French daily weather reports are available since 1857 and they contain pressure and wind data of the time.

Historical reports are sometimes inaccurate and wind intensities, that not always refer to the right phenomenon, are computed by the Beaufort scale. Beaufort wind force scale is an empirical measure that relates wind speed to observed conditions inshore and offshore. Accurate wind speeds and detailed wind directions is hardly interpretable from this scale.

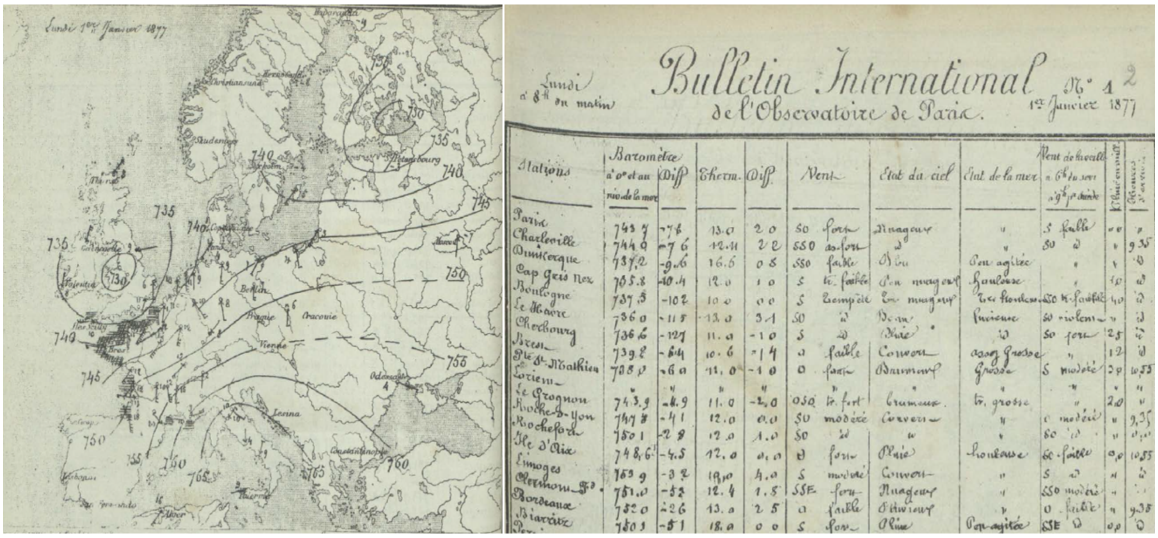


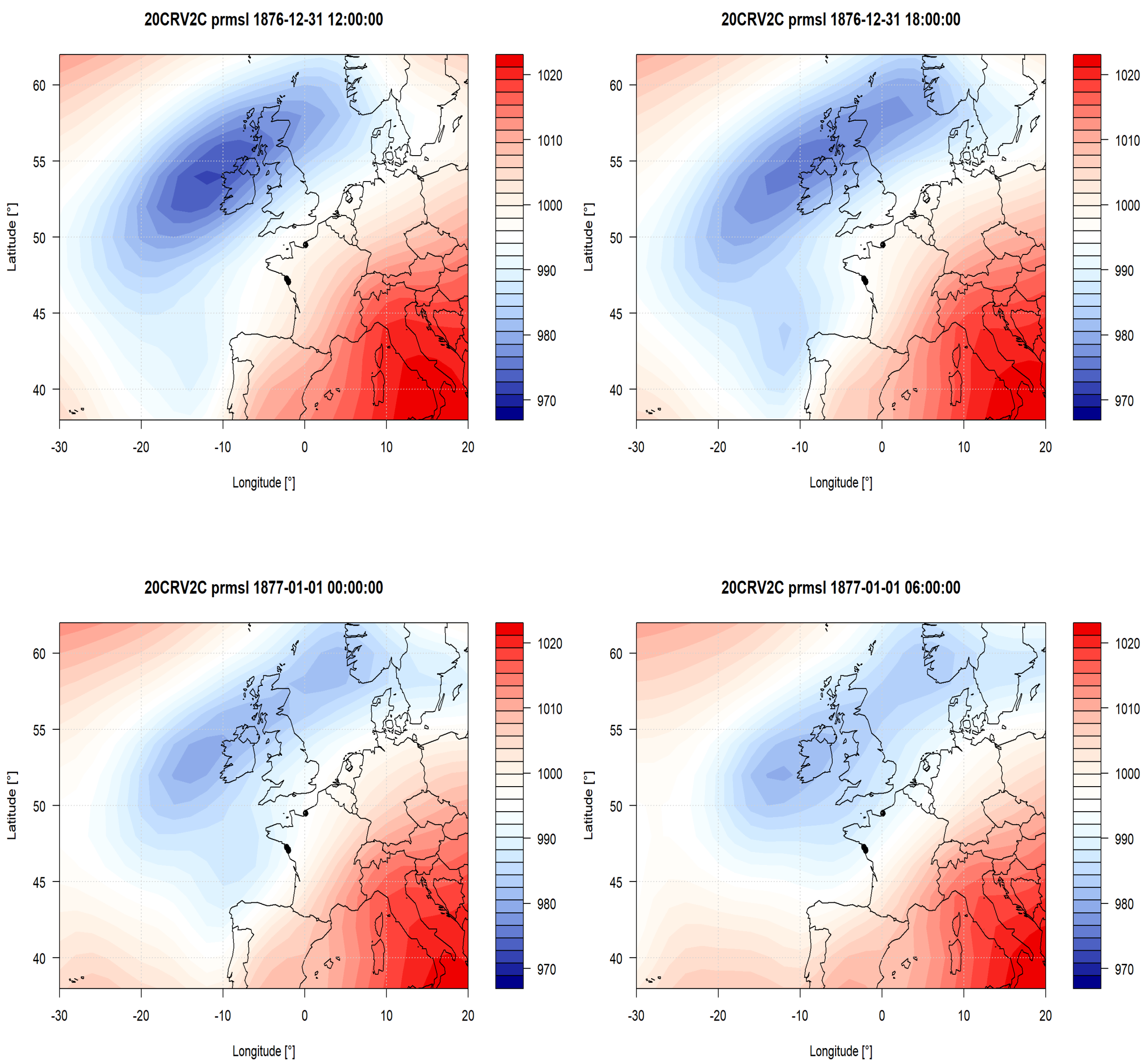
Fig. 4 – Daily weather report on the left and Pressure chart on the right for the 1st January 1877 (source Météo France)

Wind and pressure reanalysis of 20th century are available and more easily exploitable. The reanalysis 20CRV2C of the pressure at the sea surface (hPa) and of wind direction and wind speed (m/s) have been used to pre-validate some historical skew surges recovered. An example of pre-validation by reanalysis of the historical event of 1st January 1877 impacted Le Havre, Saint Nazaire and Les Moutiers-en-Retz is showed above.

Meteorological reanalysis 20CRV2C are available every 6 hours and the pressure sea surface, direction and speed of the wind for the areas affected are analysed before, during and after the storm.

In this way, synoptic charts of sea surface pressure and wind speed/direction are generated (Fig. 5 and Fig. 6). As you can see in Fig. 5, a strong depression impacted the Europe during the 1st January 1877. In the same moment, very fast winds impacted French coasts (Fig. 6) and, in particular, the black points (Le Havre, Saint Nazaire and Les Moutiers-en-Retz). Focusing on Fig. 5 and Fig. 6, an extreme skew surge was physically possible that day.

Finally, meteorological reanalysis can be a useful tool to better assess and know if a remarkable skew surge could be happened. In any case, in order to fully validate a past skew surge, the opinion of the expert on that period is needed.



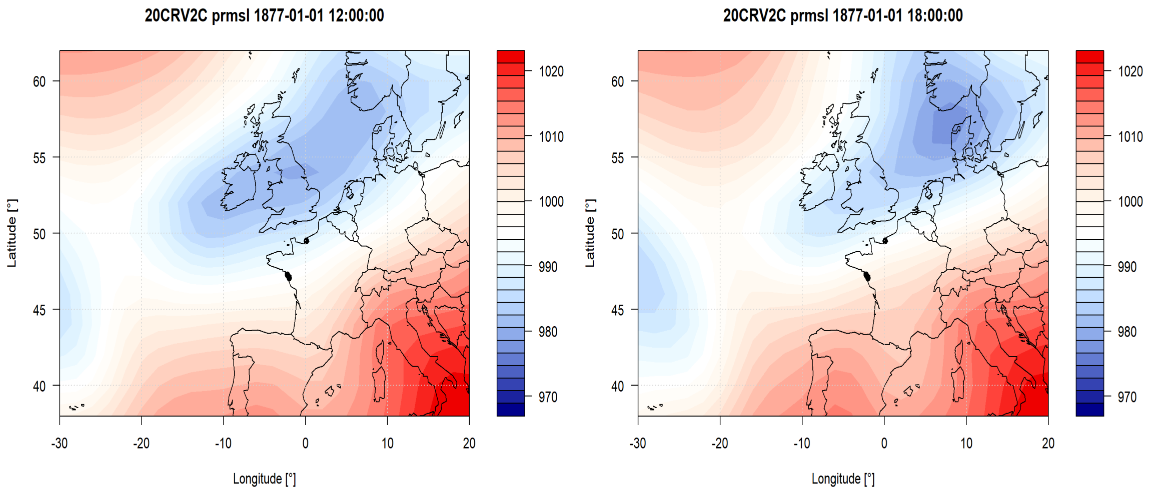
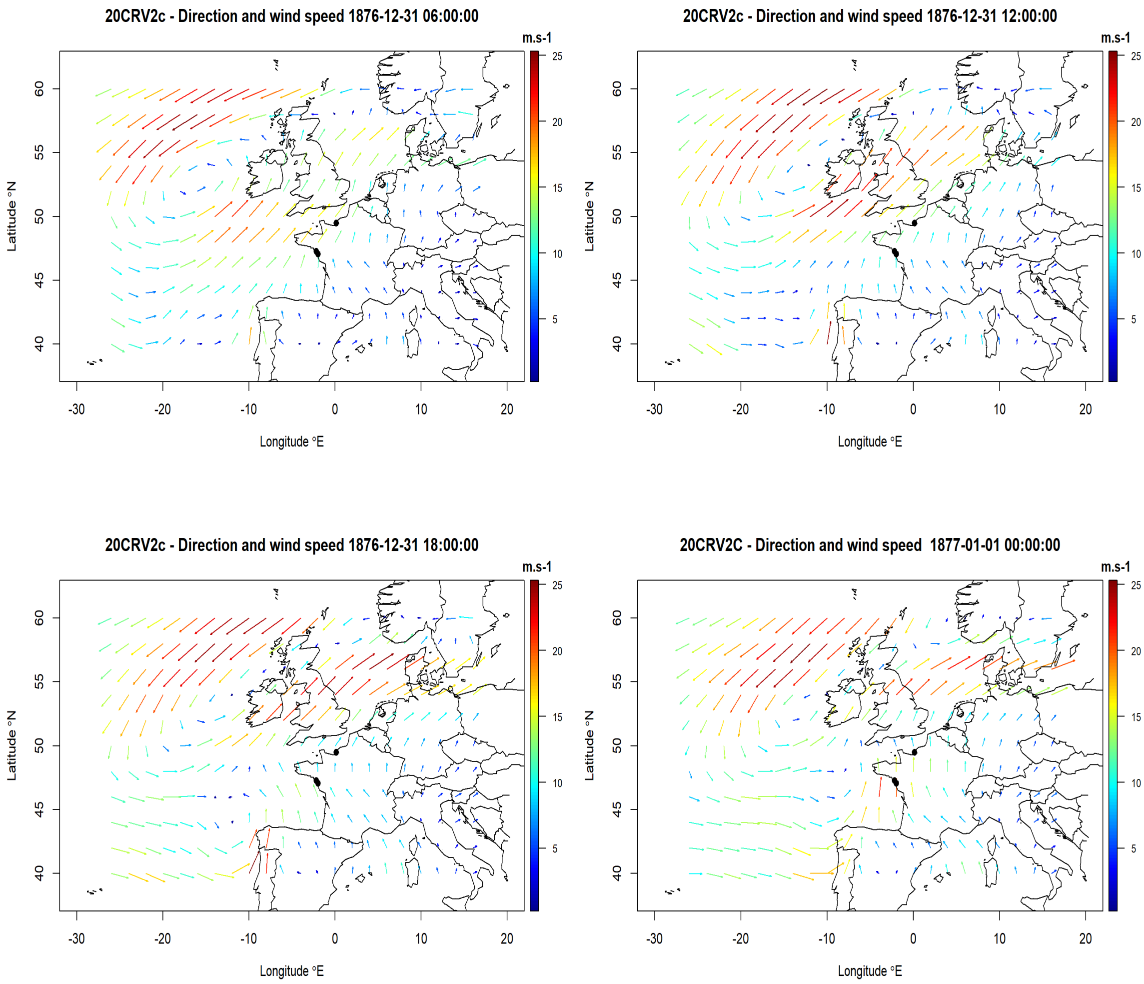


Fig. 5 – Synoptic charts generated by reanalysis 20CRV2C of the pressure on sea surface (hPa) every 6 hours between the 31st December 1876 and the 1st January 1877



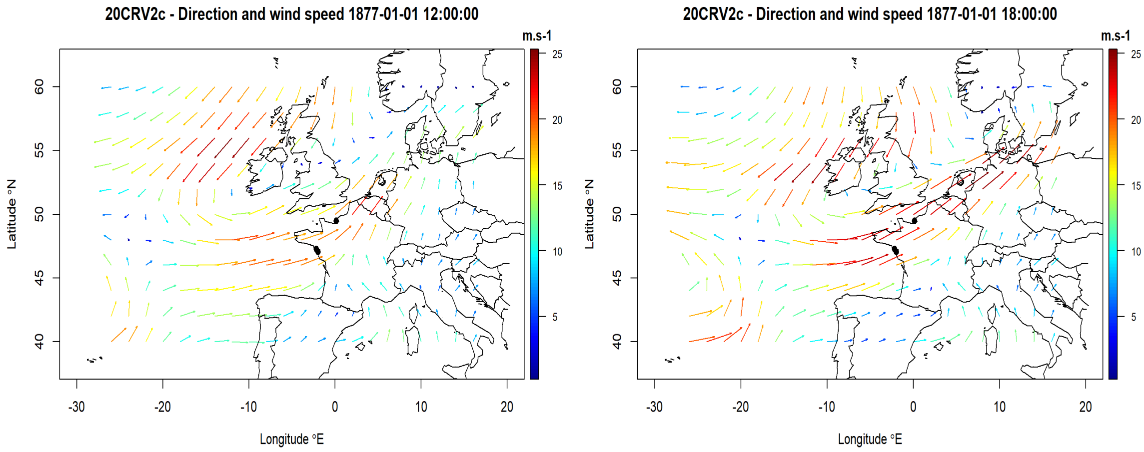


Fig. 6 - Synoptic charts generated by reanalysis 20CRV2C of wind direction and wind speed (m/s) every 6 hours between the 31st December 1876 and the 1st January 1877

2.2.4.2 Final validation by the historical expert

Even though the pre-validation based on physical considerations could provide good results, the historical data need a final validation of the historical expert. These validation is a long process and actually it has not been yet performed. For this reason, these collected and pre-validated historical skew surges cannot be still used for statistical applications of extreme events' estimations.

Chapter 3

FAB METHOD

Une méthodologie statistique appelée FAB est proposée dans ce manuscrit pour utiliser tous les trois types les plus communs de données historiques dans une analyse régionale des aléas maritimes extrêmes.

Cette méthode est fondée sur les principaux éléments de l'approche AFR proposée par Weiss (2014) pour les données systématiques et sur la notion de durée crédible. Ce nouveau concept est basé sur l'hypothèse crédible d'absence de tendance sur l'occurrence des tempêtes sur la période de collecte des données en Europe. La période d'observation locale et régionale des données historiques (appelée respectivement durée crédible locale et durée crédible régionale) peut être ainsi évaluée.

La méthode FAB permet l'utilisation des trois différents types des données historiques grâce à la définition d'une nouvelle fonction de vraisemblance pénalisée. Cinq différents types de sous-fonctions sont considérés dans cette fonction de vraisemblance : la vraisemblance des données systématiques, les trois vraisemblances pour chaque type de donnée historique et la fonction de pénalisation. Cette nouvelle fonction permet d'estimer les deux paramètres de la loi statistique régionale.

Par ailleurs, une approche de pondération réalisée sur des paramètres liés à l'analyse statistique est présentée par la méthode FAB. Cette approche permet d'identifier les seuils optimaux d'échantillonnage.

Enfin, des estimations fréquentistes ou bayésiennes peuvent être obtenues par l'application de la méthode FAB.

3.1 Introduction

Statistical estimations of extreme coastal variables and, in particular of extreme sea levels and extreme skew surges, are necessary to protect coastal nuclear fleets from the risk of flooding. Several statistical methodologies applied to a sample of extreme observations allows the computation of return levels linked to high return periods.

For sea levels and skew surges, local analysis does not permit generally to get reliable estimations of extreme values due to the limited recording period of observations (in our case of study, 40-50 years gauged on average). More extreme data are needed to extend extreme data samples and to reduce uncertainties on the extremes' estimations. For this reason, Regional Analysis is currently one of the most used statistical approach that take advantage of the wide spatial availability of tide gauges' records to extend extreme data samples.

Weiss (2014) proposes a detailed Regional Frequency Analysis approach (hereinafter its approach is mentioned as RFA) for coastal hazards. The RFA method enables the pooling of different sites considered as physically and statistically similar in a homogeneous region. In this way, extended regional extreme data samples in which a frequentist statistical analysis is applied are created. Being regional extreme data sample biggest than local extreme data sample, regional extreme estimations are generally linked to lower uncertainties than local extreme estimations.

More recently, several studies on the use of historical data in a local statistical analysis of extreme sea levels (Bulteau et al., 2015) and extreme skew surges (Hamdi et al., 2015) show that these past observations can improve extremes' estimations. In fact, representing typically very strong events not recorded by a gauge, historical data cannot be neglect in a statistical analysis of extreme events.

Unfortunately, the RFA approach can be applied only to time series of gauged data (hereinafter called as systematic data) and not to historical data. In order to be correctly applied, RFA method requires the knowledge of the observation period linked to the extreme data sample. Past observations are frequently linked with any period of observation because they do not come from continuous time series of data. No information in what happened before and after a historical event is usually available. In addition, they might be found in different types of data (for more details see Chapter 2) that cannot be used in a RFA application.

Facing these issues, a method called hereinafter FAB (from the name of the authors Frau, Andreevsky and Bernardara) is developed in this manuscript. Preserving the main concepts of the RFA approach, it allows the use historical data in the statistical analysis by the new definition of the local and the regional credible duration. This concept enables the estimation of the credible period of the past observations. Moreover, all of the three types of historical data can be used in the FAB application through the definition of two different likelihoods for systematic and historical data. Frequentist or Bayesian estimations can be computed through the application of this method.

After a short recall on the concepts of the RFA approach, the FAB method is afterwards illustrated in this chapter.

3.1.1 The RFA approach (Weiss, 2014)

The RFA approach is a statistical method proposed by Weiss (2014) for coastal hazards that enables the extension of extreme data samples. This is possible through the pooling in a regional extreme data sample of extreme data observed in sites considered physically and statistically similar. RFA method is based on the laws of regionalisation illustrated by Hosking & Wallis (1997). The formation of regions, the creation of regional samples with a defined duration and the computation of regional and local return levels are the key points of the RFA approach. All of these elements are summarised in next paragraphs.

3.1.1.1 Formation of regions

Before the definition of regional extreme data samples on which the statistical analysis is performed, the formation of homogeneous regions is required. In the RFA method, physical considerations are firstly used to form physical regions. Nevertheless, the statistical homogeneity of these physical regions must be successively verified (Hosking and Wallis, 1993).

The formation of physical homogeneous regions is based on the identification of storm clusters. A storm is defined by Weiss et al. (2014a) as an extreme physical event that impacts at least one site in a particular area. A cluster for each storm is created considering the spatiotemporal propagation of an extreme event. In fact, extreme values that are spatiotemporal neighbours are supposed to belong to the same storm.

Observations are considered as extremes when they are higher than a local physical threshold defined by a *p*-value. In addition to the *p*-value, a storm cluster is defined by other two parameters depending on spatiotemporal considerations. When an extreme data is detected in a particular site, we have to check if during Δ hours extreme data are detected in the η-nearest sites. In this way, all of these extreme data founded in the η-nearest sites during Δ hours are merged together and they belong to a same storm. Then, three parameters (p, Δ, η) representing the spatiotemporal propagation of a physical extreme event are used to define a storm cluster in the RFA method.

This definition of storm clusters enables the formation of physical regions. In fact, they are identified as the most typical storm footprints. Computing the probability $p_{i,j}$ that a site i is impacted by the same storm of a site j, the dissimilarity index $d_{i,j}=1- p_{i,j}$ can be evaluated. The dissimilarity index $d_{i,j}$ can be employed in the hierarchical clustering method of Ward (1963) that enables the pooling of sites in a defined number of regions. The most typical configuration of storms footprints or, the best number of regions, is evaluated through the significant jump of dendrogram heights (Mojena, 1977) in the RFA method. Physically homogeneous regions represent therefore the most typical impact area of storms.

The statistical homogeneity of each physical homogeneous region founded must be verified in order to apply statistical methods to regional extreme data sample. Hosking and Wallis (1993) propose a test to check this homogeneity by the computation of a measure H representing the degree of statistical homogeneity. This procedure is widely detailed in the study of Weiss et al. (2014a).

Nevertheless, this statistical homogeneity test proposed by Hosking and Wallis (1993) and used in the RFA method cannot be directly applied to storm clusters got by previous physical thresholds (corresponding to a particular p-value). In particular, when dealing with physical or statistical parameters, Bernardara et al. (2014) recommend to use different sampling thresholds (double threshold approach). If a physical threshold is used to detect storms and to reproduce their spatiotemporal dynamics, another threshold, called statistical threshold, has to be defined to consider all the different statistical aspects. Performing a statistical analysis of extreme events, statistical threshold is considered as higher than physical threshold. For this reason, the test that checks the statistical homogeneity is performed on a reduced physical storm clusters occurred in every region. In fact, the number of extreme observations considered in each physical storm cluster is decreased due to the new higher statistical thresholds selected. Finally, statistical thresholds allow the check of the statistical homogeneity of physical regions. Regions are then considered as physically and statistically homogeneous. Obviously, if statistical homogeneity hypothesis is satisfied, data coming from different sites of a homogeneous physical region fit the same regional probability distribution. This enables the estimation of regional return levels.

In the RFA approach, local statistical thresholds correspond to a common value of λ that represents the number of local exceedances per year. Specifically, local extreme data exceed statistical threshold on average λ times per year

3.1.1.2 Regional data samples and return levels

A regional pooling method is used in RFA approach to create regional extreme data sample. A regional statistical distribution can be defined only for a regional extreme data sample filtered from intersite dependence.

The pooling of extreme data observed in different locations of the region is allowed only after a normalisation of these data by a local index (Dalrymple, 1960). Local index preserves local features of the observations allowing the use of extreme events in a regional sample. In particular, local extreme data sample X_i can be normalized through a local index μ_i. Normalised data $Y_i = X_i/\mu_i$ are then supposed independent from the site i. Roth et al. (2012) recommend the use of a local index proportional to the statistical threshold for a sample of POT data. For this reason, RFA method consider the local index equal to the local statistical threshold $\mu_i = u_i$.

Nevertheless, a regional extreme data sample must be constituted by independent events. For this reason, only the maximum normalized observation of every regional storm M_s is considered for the creation of the regional sample.

The regional distribution can be defined only after verifying that normalised observations Y_i^s of each site i follow the same distribution of M_s. For this reason, the two-sample Anderson-Darling test (Scholz and Stephens, 1987) is proposed by the RFA method. Anyway, alternative tests can be carried out to verify this assumption. If this assumption is verified, the regional statistical distribution can be evaluated for a regional sample composed by independent extreme normalised data M_s.

Picklands (1975) suggests to fit the Generalised Pareto Distribution (GPD) to extreme data when dealing with POT exceedances. The GPD is the regional distribution used in the RFA method.

In particular, for a site i, the local extreme data sample X_i formed by data over the statistical threshold u_i can be fitted by a GPD as follows: $X_i \sim GPD\ (u_i, \alpha_i, k_i)$. The scale parameter is $\alpha_i > 0$, the shape parameter is k_i (positive values of k_i to an unbounded GPD) and the generic p-quantile of the local extreme data sample X_i is defined in Eq. 3.1 as:

$$x_p^i = \begin{cases} u_i - \dfrac{\alpha_i}{k_i}\left(1 - (1-p)^{-k_i}\right), & k_i \neq 0 \\ u_i - \alpha_i \log(1-p), & k_i = 0 \end{cases} \tag{3.1}$$

Furthermore, RFA approach requires local statistical thresholds corresponding to a fixed value of λ of exceedances per year. For this reason, the T-year return level of the local extreme data sample X_i can be defined as $x^i_{1-1/\lambda T}$ (Rosbjerg, 1985).

These concepts can be extended to a regional context through the use of the local indexes. The regional GPD distribution can be fit to the regional extreme data sample Y^r as formulated in Eq.3.2:

$$Y^r \sim GPD\ \left(\frac{u_i}{\mu_i}, \frac{\alpha_i}{\mu_i}, k_i\right) = GPD\ (1, \gamma, k) \tag{3.2}$$

The regional scale parameter γ and the shape parameter k are estimated through the application of Penalized maximum likelihood estimation (PMLE) to the regional extreme data sample. Coles and Dixon (1999) recommend the use of this method that allows to combine the efficiency of maximum likelihood estimators for large sample sizes and the reliability of the probability weighted moment estimators for small sample sizes penalizing high values of shape parameter.

Moreover, RFA approach permit the consideration of seasonal effects during the estimation process of regional distribution. In fact, Jonathan et al. (2008) and Jonathan and Ewans (2013) pointed out the relevance catching covariate effects on statistical estimations of extreme ocean events. These seasonal effects are taken into account through a mix of regional GPD $(1, \gamma_{r,c}, k_r)$ in which the scale parameter varies according with the season considered. The number of seasons (4 seasons proposed for the skew surge application) is considered equal to the number of co-variables C. In this way, 8 sub models are defined in accordance to parameter values of distribution $(\gamma_r^0, \gamma_r^1, \gamma_r^2, k_r)$. The easier model is the exponential distribution $(\gamma_r^1 = \gamma_r^2 = k_r = 0)$ and the most complicated is the mixed GPD computed with cosine and sine terms $(\gamma_r^0, \gamma_r^1, \gamma_r^2, k_r \in$

R^4). The best sub model is chosen by the AIC criterion (Di Baldassarre et al., 2009; Laio et al., 2009; Mendez et al., 2008).

Defining the regional *T*-year return level as $y^r_{1-1/\lambda T}$, the local T-year return level is computed in Eq.3.3 as:

$$u_i \cdot y^r_{1-1/\lambda T} = x^i_{1-1/\lambda T} \tag{3.3}$$

RFA method allows the computation of local return levels calculated through the estimation of regional return levels and the computation of the local index u_i. These estimations are also affected by the λ value considered as the same in every site of the region.

Finally, some particular regional elements are detailed in this last part of the paragraph in order to understand the issues faced by introducing historical data in the RFA. The application of regional pooling method provides an effective duration of the regional extreme data sample D_{eff} in years. The quantification of the effective duration enables the knowledge of the gain that the application of a regional analysis brings rather a local one (Weiss et al., 2014b). This variable measures the period in years in which the regional extreme data sample is observed. Regional effective duration D_{eff} is calculated as the product between the mean of site durations $\bar{d}$ belonging to the region and the degree of regional dependence φ. This degree of regional dependence φ corresponds to a value that ranges between 1 and N sites of regions. It represents the propensity of all regional sites to have the same behavior during a storm. More φ is close to 1 and more the dependency of regional sites is strong. On the contrary, more φ is close to N and more the regional sites behave in an independent way during a storm. The degree of regional dependence can be computed as $\varphi = \lambda_r/\lambda$ where λ_r is the mean annual number of storms in the region ($\lambda_r = n_r/\bar{d}$) considering that each regional storm impact every site of the region.

3.2 FAB Method

FAB method is the regional approach proposed in this manuscript that enables the estimations of extreme coastal events using historical data. This methodology is developed starting from the principal elements of the RFA approach. RFA method is exclusively developed for time series of gauged data with a known period of observation. The use of different types of historical data linked to an unknown period of observation in the RFA approach is not allowed. For these reasons, although the FAB method preserves and employs a similar procedure to that used for a RFA application (formation of regions, regional pooling method and estimations of local and return levels), the revision of existing regional elements and the creation of new regional concepts are required in order to use historical data in a regional analysis.

The main challenge in the extreme events' statistical analysis using historical events concerns the knowledge of the observation period of the whole extreme data sample and, in particular, of the historical data. The observation period of the extreme data sample, hereinafter called duration, represents the time period in which all extreme data are observed. This duration is known for extreme systematic data obtained from time series of gauged observations but it is unknown for the historical data. In fact, historical observations are frequently data points that are not linked with any time series. Only specific investigations at a local scale on past events or credible hypotheses can provide the time period of historical data.

For the most of French sites in which time series of sea levels and skew surges are available, additional information about the duration of the historical data is not available. Then, in order to use historical data in the statistical analysis of extreme events, credible hypotheses on the observation period of historical events have to be realised.

Some studies (Gaume et al., 2010; Payrastre et al., 2011; Bulteau et al., 2015, Hamdi et al., 2015) use a perception threshold to consider historical information in the statistical analysis. The perception threshold is based on the hypothesis of exhaustiveness of the historical period. For this reason, the perception threshold can be used only a wide investigation in a particular location in which historical data are available. In particular, the use of this element in a statistical analysis means that historical data above a high threshold (the perception threshold) are considered as the only extreme events occurred during the day of the first historical data and the day of the first systematic data. At the moment, this hypothesis is really strong for the most of sites in which gauged observations of sea levels and skew surges are available. In particular, no information about historical periods of extreme skew surges of the past is currently available. This does not enable to suppose a likely exhaustiveness of historical skew surges.

When exhaustiveness hypotheses of the perception threshold can not be satisfied, other credible hypotheses have to be formulated in order to use historical data in the statistical analysis of extreme events.

For these reasons, the FAB method introduces the new concept of credible duration firstly at local scale. This concept enables the estimation of a credible period for the past observations.

Historical data are so observed during a credible historical period. This period is based on credible hypothesis that extreme data with a known period of observations (systematic data) has the same frequency of occurrence than the extreme historical observations. By the formulation of this hypothesis, the credible historical duration can be estimated and the whole credible duration of an extreme data sample composed by systematic and historical data can be known. Furthermore, the FAB method extends the concept of credible duration at regional scale through its use in the regional pooling method already employed in the RFA approach. In particular, the new local credible duration is considered when pooling data of different sites of the region. For this reason, the duration of the regional extreme data sample is not longer effective but it becomes as well credible.

Another challenge for the regional analysis of extreme events is the use of most traditional types of historical events available (Chapter 2). For a statistical analysis of POT data, the three types of historical events can be considered during the process of the parameters' estimation of the considered statistical distribution. FAB method uses a Penalised Maximum Likelihood Estimation (PMLE) approach to estimate the parameters of the statistical distribution of the regional extreme data sample. The likelihood is formulated separately for systematic and historical data. This likelihood composed by the systematic and historical formulation and by the penalisation of the shape parameter of the statistical distribution is successively maximised in order to estimate the parameters of the Generalised Pareto Distribution (GPD).

In addition, when dealing with POT data, a main point in the statistical analysis is the choice of a good sampling threshold. For local or regional analysis, a bad sampling threshold can generally impact the estimations of the extreme events. In addition, in the FAB context, the sampling threshold (or statistical threshold) depends directly on credible duration. For these reasons, a procedure to select the optimal sampling threshold is proposed and suggested for FAB applications.

Finally, FAB method permits to compute Frequentist or Bayesian regional and local estimations of extreme coastal events.

All of the challenges explained in this introduction to the FAB method and the approaches in which FAB method deals with them are exposed one-by-one in this Chapter.

3.2.1 Addition of historical data

The application of the FAB methodology to local extreme data samples is achieved following the main steps of the RFA approach. In fact, the addition of historical data to the FAB method does not modify the global regional framework used in the RFA approach. The formation of regions, the definition of regional samples of independent extreme data and the estimation of regional and local return levels are the major phases used to perform the FAB method. Anyway, a wide review of each of these methodological steps and the definition of new conceptual elements are achieved to use the historical data in the regional approach. The FAB method based on the new concept of credible duration can be then applied to local extreme data samples containing historical data.

In particular, originated from different sources, historical data is an extreme data value representing a storm that impacted a particular site in the past or during a dysfunction period of the tidal gauge located in the same particular site. The fact that historical events are isolated does not allow the knowledge of what it happened before and after a particular past event in terms of storm occurrence.

The knowledge of the storm occurrence on average per year λ is essential for the application of a generic statistical analysis on POT extreme data. This occurrence per year λ above a sampling threshold might be computed through the observation period in years of historical events.

The FAB method provides credible hypothesis based on the frequency of gauged data to remedy to the absence of information on past period. The addition of historical data in a local sample modifies the number of extreme events per year λ exceeding a particular statistical threshold. When only systematic data are available, the generic occurrence λ of gauged data can be computed as:

$$\lambda = \frac{n}{d} \tag{3.4}$$

where n is number of data above the sampling threshold and d is the duration in years of the time series of the gauged observations. Below, Fig. 7 shows that for systematic records of skew surges the duration d (in blue) is well-known. Fixing a sampling threshold (in red in the image below), only few skew surges are considered as extreme (the green crosses) and the number of extreme events per year λ exceeding this threshold is calculated by the formulation given in Eq.3.4.

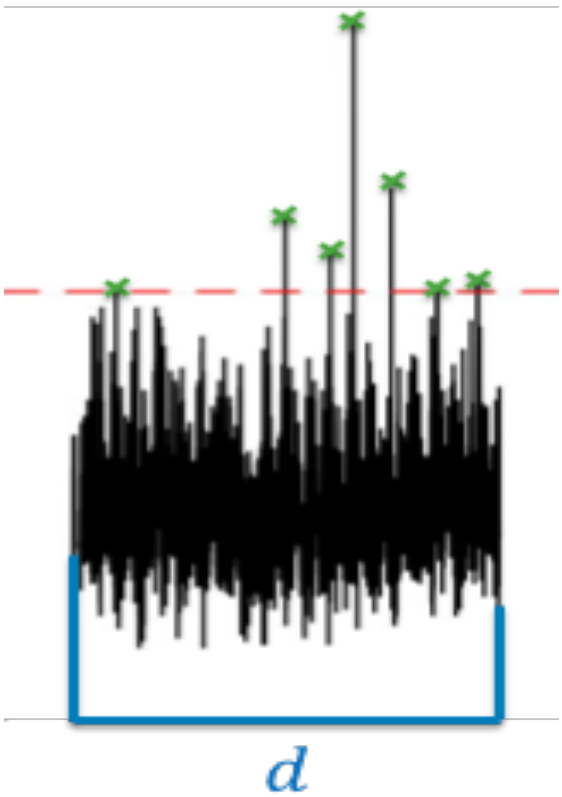


Fig. 7 – Scheme for λ computation on systematic time series

Therefore, fixing a sampling threshold for an extreme data sample composed by historical and systematic events, the number of n values exceeding the fixed threshold is known. This is not enough to calculate the duration of the whole extreme data sample. In fact, the computation of the observation period *d* of this sample can be provided by the knowledge of the storm occurrence λ.

Furthermore, the duration *d* can be defined as the sum of the duration d_{syst} of systematic data and the duration d_{hist} of historical data. The observation period of systematic data d_{syst} for the systematic part of the extreme data sample is is equal to the period of the gauged observations. On the other hand, the observation period of systematic data d_{hsst} for the historical part of the extreme data sample is unknown and the duration *d* cannot be calculated.

In the FAB framework, the local duration *d* permits the computation of the local storm occurrence λ, the degree of regional dependence φ and the regional credible duration D_{cr}. For this reasons, the historical duration d_{hist} is needed to apply the FAB method to a regional extreme data sample of systematic and historical events. In addition, the relevance of the regional credible duration D_{cr} in the FAB method is also underlined in the estimation of local and return levels process. Regional T-year return level $y^{r}_{1-1/\lambda T}$ and the local T-year return level $x^{i}_{1-1/\lambda T}$ depend on the λ value through which D_{cr} is estimated.

3.2.2 Local and Regional Credible duration

The main new concepts in which the FAB method is based are the local and regional credible durations. These elements allow the use of historical data in a regional analysis of extreme events.

If historical data are available in a particular site, its local duration contain an additional period given by historical events. Splitting in two parts the duration of a whole extreme data sample at this particular site, the historical period in which historical data exceed a threshold is unknown. The computation of this observation period is enabled by the formulation of credible hypothesis.

After the examination of storm frequency on several scientific studies (Barriendos et al., 2003; Hanna et al., 2008; Matulla et al., 2008; Allan et al., 2009; Barring and Fortuniak, 2009; Ferreira et al., 2009; Wang et al., 2009; Wang et al., 2011; Hartmann et al., 2013), the lack of a trend on storm frequency during the 20th century is assumed as the credible hypothesis. This hypothesis allows the estimation of the credible duration.

In particular, for skew surges (the specific coastal variable analysed by this method in Chapter 4) a supplementary test on the longest skew surge series is performed (Annexe F). Anyway, this hypothesis must be verified before the application of the FAB method for any coastal variables. For non-stationary datasets, this methodology can be adapted through the adjustment of hypothesis formulated.

With this credible hypothesis, the number of systematic data per year λ_{syst} above a statistical threshold u_i is equal to the number of historical data per year λ_{hist} above the same statistical threshold u_i for a particular site i. Recalling that $d_{hist}=n_{hist}/\lambda_{hist}$ and stating that $\lambda=\lambda_{syst}=\lambda_{hist}$, the formulation of the local credible duration $d_{cr,i}$ for a site i composed by systematic and historical data as follows:

$$d_{cr,i} = d_{syst,i} + d_{hist,i} = \frac{n_{syst}}{\lambda_{syst}} + \frac{n_{hist}}{\lambda_{hist}} = \frac{n_{syst}}{\lambda} + \frac{n_{hist}}{\lambda} \tag{3.5}$$

The local credible duration based on credible hypothesis is now defined. Recalling that the correct value of λ is known, the additional part of local duration provided by the availability of historical data is then calculated as $d_{hist}=n_{hist}/\lambda$. The knowledge of this observation period allows the use of historical data in regional analysis.

Moreover, the use of historical data also provides an additional of the observation period of the regional extreme data sample. Recalling that the regional duration is defined as the mean of local durations filtered by any intersite dependence, the availability of historical events in one or more sites of the region provides an additional local duration and consequently an additional regional duration.

Regional duration was in RFA approach computed by effective local durations. With historical data, a regional credible duration $D_{cr,r}$ can be computed for a particular region r. The regional credible duration $D_{cr,r}$ depending on the mean of local credible durations of the N sites belonging to the region r and it is formulated as follows:

$$D_{cr,r} = \varphi \times \overline{d_{cr}} = \frac{\lambda_r}{\lambda} \times \sum_{i=1}^{N} \frac{d_{cr,i}}{N} \tag{3.6}$$

Obviously, if no historical data are available in any sites, local duration is effective in every site of the region and regional duration is computed as for the RFA approach.

The regional credible duration represents the duration of the regional extreme data sample after filtering from any intersite dependence. This filtration is achieved through the use of the degree of regional dependence φ and it enables to compute a correct duration of the regional sample of extreme data. The degree of regional dependence φ depends not only on the λ but also on the mean annual number of storms λ_r in the region r.

The mean annual number of storms λ_r is defined as the ratio between the number of regional data N_r and the mean of local credible durations. In this way, the factor λ_r is computed without taking into account the different behaviour that sites can have when a storm impacts the region. Forming regional extreme data sample, it is important to underline that not all storms impacting the region are observed in each site of the region. For this reason, the computation of a degree of regional dependence φ is necessary to know the real duration of a regional sample of extreme data. After this further formulations, regional credible duration $D_{cr,r}$ can be expressed as follows:

$$D_{cr,r} = \frac{\lambda_r}{\lambda} \times \overline{d_{cr}} = \frac{N_r}{\overline{d_{cr}}} \times \frac{\overline{d_{cr}}}{\lambda} = \frac{N_r}{\lambda} \quad (3.7)$$

More the regional duration is high and more extreme estimations can be properly calculated. The introduction of historical data increases the regional duration and then they improve the reliability on the extreme estimations. In addition, the regional credible duration can be separated in two parts depending on systematic and historical data as follows:

$$D_{cr,r} = D_{cr,r,syst} + D_{cr,r,hist} = \frac{N_{r,syst} + N_{r,hist}}{\lambda} \quad (3.8)$$

Dividing historical data in the three most common types of historical data, Eq.3.8 can be reformulated as follows:

$$D_{cr,r,hist} = D_{cr,r,hist,I} + D_{cr,r,hist,II} + D_{cr,r,hist,III} = \frac{N_{r,hist,I} + N_{r,hist,II} + N_{r,hist,III}}{\lambda} \quad (3.9)$$

Eq.3.9 is useful for next formulations of likelihood functions for systematic and historical data.

3.2.3 The use of different types of historical data

The use of the three different types of historical data impacts all methodological phases used in FAB method. In particular, the formation of regions and the creation of regional extreme data sample processes are formulated in the RFA approach to use extreme exact data. Then, only type I of historical data (exact historical data) can be considered in these phases of the regional statistical analysis.

FAB method permits to consider all the three types of historical data in these two phases using the mean value of the range for the historical data type II and the lower limit value for the historical data of type III.

In particular, considering the case in which a historical data of type III $x_{hist\,III,i}$ is available in the site i and it is higher than the physical threshold p_i and (or) the statistical threshold u_i, $x_{hist\,III,i}$ is the data value used to form the regions and (or) to evaluate the maximum normalized value M_s of the storm s during the definition process of the regional extreme data sample.

Considering another case in which a historical data of type II [$x_{higher\ II,i}$, $x_{lower\ II,i}$] is available in the site i and $x_{lower\ II,i}$ is higher than the physical threshold p_i and (or) the statistical threshold u_i, the mean between is the data value used to form the regions and (or) to evaluate the maximum normalized value M_s of the storm s during the definition process of the regional extreme data sample.

In addition, the particular case in which a historical data of type II [$x_{higher\ II,i}$, $x_{lower\ II,i}$] is available in the site i and $x_{lower\ II,i}$ is lower than the physical threshold p_i and (or) the statistical threshold u_i but the $x_{higher\ II,i}$ is higher than the physical threshold p_i and (or) the statistical threshold u_i has to be examined. The means between $x_{higher\ II,i}$ and the threshold p_i or u_i are used respectively for the two considered regional phases.

Type II or type III of historical data can belong to a regional extreme data sample through the approach exposed above. In this case, they recover their original nature before the estimation of the regional distribution. Specifically, normalized data ranges are considered when historical data of type II is the maximum normalized data M_s of the storm s impacted the region (Y^r_s=[$x_{higher\ II,i}/\mu_i$, $x_{lower\ II,i}/\mu_i$]). In the particular case in which historical data of type II is a M_s that belongs to a regional extreme data sample with a $x_{lower\ II,i} < u_i$, these normalized data ranges are considered (Y^r_s=[$x_{higher\ II,i}/\mu_i$, u_i/μ_i]). On the other hand, all possible normalized data above its normalized lower limit are considered when historical data of type III is a maximum normalized data M_s belonging to a regional extreme data sample ($Y^r_s = x_{hist\ III,i}/\mu_i$).

Anyway, a more structured approach to consider the three different types of historical data is used during the estimation of the regional distribution. In fact, all types of historical data can be considered in the statistical analysis by the formulation of separately likelihood for systematic and historical data. This enables the estimation of the regional distribution's parameters considering correctly all types of historical information available.

3.2.4 Penalised maximum likelihood

The regional extreme data sample $Y^r = \{Y^r_1, \ldots, Y^r_h\}$ composed by a number h of independent storms is fitted by a GPD distribution (Eq.3.2). The estimation of the regional GPD parameters is performed by the Penalized Maximum Likelihood Estimator (PMLE). In general, the penalization of the likelihood enables the use of additional information into the inference compared with that supplied by the extreme data sample (Coles and Dixon, 1999). As suggested by Coles and Dixon (1999) and by Weiss (2014), a likelihood penalization is used to penalise positive values (< 1) of the shape parameter k as follows:

$$P(k) = \exp\left(-m\left(\frac{1}{1-k} - 1\right)^{d} \quad if\ 0 < k < 1 \right. \tag{3.10}$$

where the parameters m and d are recommended to be set to unity by Coles and Dixon (1999) after some experimentations on their performances. In addition, if $k>1$ the penalty is null $P(k)=0$ and, in the contrary, if $k<0$ the penalty is equal to 1.

This penalty function permits the exploitation of the efficiency of maximum likelihood for big extreme sample and, at the same time, of the accuracy of the probability weighted moment for small extreme sample. The penalized likelihood function and the penalized log-likelihood function used in FAB method is defined by the Eq.3.11 and Eq.3.12 as follows:

$$\mathcal{L}_{pen}(Y^r|\lambda,\theta) = \mathcal{L}(Y^r|\lambda,\theta) * P(k) \tag{3.11}$$

$$\ell_{pen}(Y^r|\lambda,\theta) = \ell(Y^r|\lambda,\theta) + \ln(P(k)) \tag{3.12}$$

where the vector of parameters θ contains the scale γ and shape k parameters of the regional GPD. The likelihood function $\mathcal{L}$ and the log-likelihood function ℓ are defined for different types of data in next paragraphs.

3.2.5 Likelihood formulation with historical data

The likelihood function $\mathcal{L}$ for the regional extreme data sample $Y^r = \{Y^r_{syst}, Y^r_{hist}\}$ composed by systematic and historical data can be calculated considering two different likelihood functions for systematic and for historical data (Miquel, 1981; Cohn and Stedinger, 1987; Lang et al., 1997). The likelihood function $\mathcal{L}$ and the log-likelihood function ℓ for the regional sample Y^r is then formulated as follows:

$$\mathcal{L}(Y^r|\lambda,\theta) = \mathcal{L}\left(Y^r{}_{syst}|\lambda,\theta\right) * \mathcal{L}(Y^r{}_{hist}|\lambda,\theta) \tag{3.13}$$

$$\ell(Y^r|\lambda,\theta) = \ell\left(Y^r{}_{syst}|\lambda,\theta\right) + \ell(Y^r{}_{hist}|\lambda,\theta) \tag{3.14}$$

These likelihood and log-likelihood expressions define two different likelihood functions for systematic and historical data. The formulation of these two likelihood ($\ell\left(Y^r{}_{syst}|\lambda,\theta\right)$ and $\ell(Y^r{}_{hist}|\lambda,\theta)$) is necessary to identify the penalised likelihood used to estimate the parameters of the regional distribution of the whole regional extreme data sample Y^r.

3.2.5.1 Likelihood for a sample of systematic and historical data

The likelihood for a generic POT data sample *Y* composed of *h* systematic extreme data observed in *d* years is formulated as follows (Miquel, 1981):

$$\mathcal{L}(Y|\lambda,\theta) = P(n_1) * \ldots * P(n_d) * \prod_{i=1}^{h} f(Y_i|\theta) \tag{3.15}$$

The term of the left part of Eq.3.15 represents the Poisson process of occurrence of the observed data assuming the independence of observations. In particular, $P(n_d)$ is the probability to observe n_d peaks in the year *d*. It can be formulated as:

$$P(n_d) = e^{-\lambda}\frac{\lambda^{n_d}}{n_d!} \tag{3.16}$$

The second term of the Eq.3.15 corresponds to the product of the density functions of the statistical distribution for each observed data. Eq.3.15 can be reformulated as follows:

$$\mathcal{L}(Y|\lambda,\theta) = e^{-\lambda}\frac{\lambda^{n_1}}{n_1!} * \ldots * e^{-\lambda}\frac{\lambda^{n_d}}{n_d!} * \prod_{i=1}^{h} f(Y_i|\lambda,\theta) = \prod_{i=1}^{d} P(n_i) * \prod_{i=1}^{h} f(Y_i|\theta) \quad (3.17)$$

From Eq.3.17, the log-likelihood can be defined as:

$$\ell(Y|\lambda,\theta) = \sum_{i=1}^{d} ln(P(n_i)) + \sum_{i=1}^{h} ln\big(f(Y_i|\theta)\big) = \sum_{i=1}^{d}(-\lambda + n_i \ln(\lambda) - \ln(n_i!)) + \sum_{i=1}^{h} ln\big(f(Y_i|\theta)\big) \quad (3.18)$$

The maximisation of likelihood must be performed for the λ and for the parameters θ of the statistical distribution. The maximisation of the likelihood for the λ concerns only the first term of the Eq. 3.17 and corresponds to the λ value computed until now in Eq.3.4. For this reason, this first term of the likelihood that not depends on the parameters θ can be neglected in the maximisation process.

For a stationary process in which the λ value is considered constant for a whole extreme data sample (the case of credible duration) observed in D_{cr} years as formulated in Eq.3.8, the formulations of likelihood and log-likelihood for a regional extreme data sample composed of h_{syst} systematic data observed in $D_{cr,syst}$ and h_{hist} historical data observed in $D_{cr,hist}$ data are exposed are described as follows:

$$\mathcal{L}(Y^r|\lambda,\theta) = \prod_{i=1}^{D_{cr}} P(n_i) * \prod_{i=1}^{h_{syst}} f\big({Y^r}_{i,syst}|\theta\big) * \prod_{i=1}^{h_{hist}} f\big({Y^r}_{i,hist}|\theta\big) \quad (3.19)$$

$$\ell(Y^r|\lambda,\theta) = \sum_{i=1}^{D_{cr}} ln(P(n_i)) + \sum_{i=1}^{h_{syst}} ln\Big(f\big({Y^r}_{i,syst}|\theta\big)\Big) + \sum_{i=1}^{h_{hist}} ln\Big(f\big({Y^r}_{i,hist}|\theta\big)\Big) \quad (3.20)$$

As for Eq.3.17 and Eq.3.18, only the first terms of the Eq. 3.19 and Eq.3.20 are concerned to the maximisation of the likelihood for the λ. The likelihood maximisation corresponds to the λ value computed in Eq.3.7. For this reason, this first terms of the likelihood and log-likelihood that not depends on the parameters θ are neglected hereinafter in the definitions of the particular likelihood functions for systematic and historical data.

3.2.5.2 Likelihood formulation for systematic data

The definition of the likelihood functions for the systematic part of the regional extreme data sample composed of h_{syst} number of data observed in $D_{cr,syst}$ as follows:

$$\mathcal{L}\left(Y^r{}_{syst}|\lambda,\theta\right) = \prod_{i=1}^{h_{syst}} f\left(Y^r{}_{i,syst}|\theta\right) \tag{3.21}$$

$$\ell\left(Y^r{}_{syst}|\lambda,\theta\right) = \sum_{i=1}^{h_{syst}} ln\left(f\left(Y^r{}_{i,syst}|\theta\right)\right) \tag{3.22}$$

In particular, replacing the density function of the GPD for the parameters expressed in Eq.3.2, the log-likelihood of Eq.3.22 can be reformulated as follows (Grimshaw, 1993):

$$\ell\left(Y^r{}_{syst}|\lambda,\theta\right) = -h_{syst} * \ln(\gamma) + \left(\frac{1}{k} - 1\right) \sum_{i=1}^{h_{syst}} \ln\left(1 - \frac{kZ_{i,syst}}{\gamma}\right) \tag{3.23}$$

where the $z_{i,syst} = (Y_{i,syst} - 1)/\gamma$ for the regional analysis case in which the regional location parameter is equal to 1.

3.2.5.3 Likelihood formulation for historical data

The likelihood function for the historical part of the regional extreme data sample composed of $h_{\mathrm{hist}}=h_{\mathrm{hist,I}}+h_{\mathrm{hist,II}}+h_{\mathrm{hist,III}}$ number of data observed (the three different types of historical data defined in Chapter 2) in $D_{cr,hist}=D_{cr,hist,I}+D_{cr,hist,II}+D_{cr,hist,III}$ years can be formulated as follows:

$$\mathcal{L}(Y^r{}_{hist}|\lambda,\theta) = \mathcal{L}\left(Y^r{}_{hist,I}|\lambda,\theta\right) * \mathcal{L}\left(Y^r{}_{hist,II}|\lambda,\theta\right) * \mathcal{L}\left(Y^r{}_{hist,III}|\lambda,\theta\right) \tag{3.24}$$

$$\ell(Y^r{}_{hist}|\lambda,\theta) = \ell\left(Y^r{}_{hist,I}|\lambda,\theta\right) + \ell\left(Y^r{}_{hist,II}|\lambda,\theta\right) + \ell\left(Y^r{}_{hist,III}|\lambda,\theta\right) \quad (3.25)$$

The likelihood and log-likelihood for the type I of historical data (exact data) corresponds to the likelihood function of systematic data (Eq.3.21 and Eq.3.22):

$$\mathcal{L}\left(Y^r{}_{hist,I}|\lambda,\theta\right) = \prod_{i=1}^{h_{hist,I}} f\left(Y^r{}_{i,hist,I}|\theta\right) \quad (3.26)$$

$$\ell\left(Y^r{}_{hist,I}|\lambda,\theta\right) = \sum_{i=1}^{h_{hist,I}} ln\left(f\left(Y^r{}_{i,hist,I}|\theta\right)\right) \quad (3.27)$$

Type II and type III of historical data represent respectively a data range ($Y^r{}_{higher,hist,II,}$ $Y^r{}_{lower,hist,II}$) and a lower limit value of the historical data ($Y^r{}_{lower,hist,III}$). For these reasons, the likelihoods and log-likelihoods of these two particular types of historical data can be defined as follows:

$$\mathcal{L}\left(Y^r{}_{hist,II}|\lambda,\theta\right) = \prod_{i=1}^{h_{hist,II}} \left(F\left(Y^r{}_{i,higher,hist,II}|\theta\right) - F\left(Y^r{}_{i,lower,hist,II}|\theta\right)\right) \quad (3.28)$$

$$\ell\left(Y^r{}_{hist,II}|\lambda,\theta\right) = \sum_{i=1}^{h_{hist,II}} ln\left(F\left(Y^r{}_{i,higher,hist,II}|\theta\right) - F\left(Y^r{}_{i,lower,hist,II}|\theta\right)\right) \quad (3.29)$$

$$\mathcal{L}\left(Y^r{}_{hist,III}|\lambda,\theta\right) = \prod_{i=1}^{h_{hist,III}} \left(1 - F\left(Y^r{}_{i,lower,hist,III}|\theta\right)\right) \quad (3.30)$$

$$\ell\left(Y^{r}{}_{hist,III}|\lambda,\theta\right)=\sum_{i=1}^{h_{hist,III}} ln\left(1-F\left(Y^{r}{}_{i,lower,hist,III}|\theta\right)\right) \tag{3.31}$$

where F is the cumulative function of the statistical distribution used. For a GPD of parameters equal to those expressed in Eq.3, the log-likelihoods functions for each type of historical data can be reformulated as follows:

$$\ell\left(Y^{r}{}_{hist,I}|\lambda,\theta\right)=-h_{hist,I}*\ln(\gamma)+\left(\frac{1}{k}-1\right)\sum_{i=1}^{h_{hist,I}}\ln\left(1-\frac{kZ_{i,hist,I}}{\gamma}\right) \tag{3.32}$$

$$\ell\left(Y^{r}{}_{hist,II}|\lambda,\theta\right)=\sum_{i=1}^{h_{hist,II}} ln\left(\left(\left(kZ_{i,higher,hist,II}-1\right)^{\frac{1}{k}}\right)-\left(\left(kZ_{i,lower,hist,II}-1\right)^{\frac{1}{k}}\right)\right) \tag{3.33}$$

$$\ell\left(Y^{r}{}_{hist,III}|\lambda,\theta\right)=\frac{1}{k}\sum_{i=1}^{h_{hist,III}} ln\left(kZ_{i,lower,hist,III}-1\right) \tag{3.34}$$

where the $Z_{i,hist,I}=(Y_{i,hist,I}-1)/\gamma$, $Z_{i,higher,hist,II}=(Y_{i,higher,hist,II}-1)/\gamma$, $Z_{i,lower,hist,II}=(Y_{i,lower,hist,II}-1)/\gamma$ and $Z_{i,lower,hist,III}=(Y_{i,lower,hist,III}-1)/\gamma$ for a regional location parameter is equal to 1.

The formulations of Eq.3.23, Eq.3.32, Eq.3.33 and Eq.3.34 enable the definition of each element of the penalized log-likelihood of Eq.3.12:

$$\ell_{pen}(Y^{r}|\lambda,\theta)=\ell\left(Y^{r}{}_{syst}|\lambda,\theta\right)+\ell\left(Y^{r}{}_{hist,I}|\lambda,\theta\right)+\ell\left(Y^{r}{}_{hist,II}|\lambda,\theta\right)+\ell\left(Y^{r}{}_{hist,III}|\lambda,\theta\right)+\ln(P(k)) \tag{3.35}$$

Eq.3.9 can be now maximised (PMLE) and the regional parameters are estimated.

3.2.6 Choice of statistical threshold

The regional analysis based on a POT approach requires the definition of local sampling thresholds for each site. In this way, extreme events can be detected in each site and regional extreme data samples can be defined.

A double threshold approach (Bernardara et al., 2014) is used in the RFA method. For this reason, the "physical" thresholds that is equal to the 0.995 p-value computed in each site do not correspond to the higher "statistical" or "sampling" threshold. Sampling threshold represents the extreme bound above which all data can be considered as independent extreme events.

Every sampling threshold can be represented by the value of number of occurred storms λ on average per year above this threshold. In the RFA, the value of λ storms per year is considered identic for each site of the same region.

Varying the λ value, the value of the sampling threshold changes in each site and consequently every local extreme data sample is different. In particular, a too low λ value leads the performance of a statistical analysis with the biggest extreme events available in each site. In this case, the more usual inconvenient is that the number of data is not enough to compute proper estimations. On the contrary, a too high λ value leads the computation of a statistical distribution with many events that obviously they cannot considered as extremes. For these reasons, FAB method proposes an appropriate approach to find an optimal λ value, and consequently optimal sampling thresholds. This approach enables the performance of an efficient statistical analysis.

The definition of optimal sampling thresholds is a complex subject and no univocal methods exist currently in literature. FAB method proposes the selection of the optimal sampling threshold through the use of several indicators or parameters depending on this threshold. Both downstream parameters and upstream parameters of the statistical process can be considered. The sensitivity analysis on these parameters is conducted by a weighting calculation in order to compute the best value of λ that corresponds consequently to optimal values of statistical thresholds.

FAB method performs a sensitivity analysis for a total of twelve parameters depending on λ value. Although twelve parameters are chosen for the FAB method, the scientific expert can consider other parameters in according to the statistical analysis performed in the application of this approach.

In any case, five primary parameters and seven secondary parameters are considered in order to simplify the individuation of the optimal λ value.

Primary parameters represent all tests necessary to perform the statistical analysis. Without the verification of these tests, the statistical analysis cannot perform. Primary parameters enable the removing of all possible λ values in which tests are not verified. The five primary tests are applied to the regional sample of extreme values for a range of λ values between 0,25 and 2 by a step of 0.01. This range is defined between these two values in order to consider values of λ neither too low nor too high. In particular, under the value of 0.25 storms per year only few extreme data are considered to get a valid statistical distribution and over 2 storms per year the statistical analysis starts to be performed not only considering extreme events.

Secondary parameters are not essential variables for the statistical process but they can be considered by the expert as equally important for the performance of the statistical analysis. Both types of factors are used in the sensitivity analysis. The seven secondary indicators of the sensitivity analysis are evaluated as the primary parameters for each possible λ value.

The best regional sample in which primary and secondary parameters are globally the best ones corresponds to the selection of the optimal λ. In order to find an univocal optimal λ, a weighting computation is proposed. The weighting analysis is applied to all secondary parameters corresponding to λ-cases in which primary parameters are verified. In particular, if extreme samples are not suitable for a regional analysis, it is useless to calculate a weighting for secondary parameters. This enables the simplification of the process of identification of the optimal λ value. The optimal λ corresponds to the possible λ-cases in which the sum of each secondary parameter's weighting is the highest. In fact, if all parameters are analysed only visually, many combinations appear to be appropriate and it is not easy to find the best one.

Finally, this procedure based on a weighting analysis allows the definition of the optimal λ for each physical and statistical region. In particular, the statistical verification of physical regions is one of the five primary tests in the proposed approach and, for this reason, the weighting computation is evaluated only -λ-cases corresponding to physical and statistical homogeneous regions. In particular cases in which only a little percentage of λ values between the all possible values of λ are selected due to the computation of this particular primary test for a considered region, a further division in two physical homogeneous regions is suggested. In this case, a new analysis on primary and secondary parameters has to be performed for the new two regions defined.

The other four primary tests proposed by the FAB method concern the stationarity test, the chi-squared test χ^2, the Kolmogorov-Smirnov test and the test to detect outliers in regional extreme data samples.

The seven secondary parameters considered concern the data number of regional sample, the regional duration, the value of χ^2, the scale parameter of the regional GPD, the shape parameter of the regional GPD, the adimensional degree of regional dependence $\boldsymbol{\Phi}$ and the estimated regional return level associated to a return period T of 1000 years.

All details of primary and secondary tests mentioned here are exposed in the following. In any case, it is recommended by this study to give a quick look on regional return level plots that correspond to the optimal λ value selected by the approach here proposed.

The sensitivity analysis proposed by the FAB method is flexible and enables the addition or removing of other tests that can be considered as significant for the statistical analysis performed. In addition, if parameters are considered more significant than others, a double or a multiple weighting value can be computed for the relevant parameter or a manual check can be performed by the expert in order to verify if that variable assumes or not an acceptable value for the considered statistical analysis.

3.2.6.1 Primary parameters

Primary parameters represent basic statistical tests that must be satisfied for the performance of a statistical analysis. These tests are applied to a regional extreme data sample corresponding to a particular λ value. Five primary tests are considered in the FAB method.

Stationarity test

The stationarity on the intensity of regional extremes is required to perform the considered statistical analysis.

This stationarity can be checked by stationarity test applied to the storm intensities of regional sample (Hosking and Wallis, 1993). This test verifies if the difference between the mean of two determined regional sub-samples is significant for a level of risk. In particular, the regional sample is stationary if the difference of the previous two means is lower than the admissible difference linked to the risk level. In this study, the risk level considered is 10%.

Homogeneity test

A basic hypothesis on a regional analysis of extreme values is the statistical homogeneity of the region. Without this homogeneity the regional analysis cannot be performed (Hosking and Wallis, 1997; Weiss, 2014).

The statistical homogeneity of a physical region is tested through the homogeneity test proposed by Hosking and Wallis (1997). This test evaluates the heterogeneity value H. If the value of H is higher than 2, the region is considered as heterogeneous and the regional analysis cannot be performed. However, as Weiss (2014) states in his study, the heterogeneity of a regional extreme data sample can be generated by discordant sites. In this case, the value of the heterogeneity H has to be computed without considering discordant sites.

Pearson's chi-squared test χ^2

The chi-squared test is used to know the goodness of a fit (Cochran, 1952; Chernoff and Lehmann, 1954). Applied to the regional GPD distribution, this test provides the dispersion value $S\text{-}\chi^2$ between the observed frequency distribution of the regional sample and the theoretical distribution. If this value is lower than a limit value S_{lim} linked to a significance level and to the degree of freedom, no indications are provided the statistical distribution is good for the regional extreme data sample. For this type of test, the degree of freedom has to be equal to a considered number of classes minus one reduced by the number of the estimated parameters of the distribution and by one. In our case, a significance level (p-value) of 5% and a number of 10 classes corresponding to seven degrees of freedom are used. With the considered parameters, the $S\text{-}\chi^2$ may not exceed a $S_{lim\text{-}0.05}$ of 14.067.

Kolmogorov-Smirnov test

The regional analysis assumes that the distribution of the maximum of regional storms M_s is the same as the normalised values Y_i observed in every site of the region. This assumption is checked by the two-sample Kolmogorov-Smirnov test (Smirnov, 1939) applied for each site of the region. This test is based on the formulation of null hypothesis that M_s and Y_s have the same distribution. Without the verification of this assumption, the regional distribution cannot be performed. The test computes p-values between the M_s sample and each sample of Y_i. This p-value has to be higher than a limit p-value corresponding in this study to 0.01.

Test to detect outliers

The absence of outliers in the extreme data sample is required for engineering applications linked to the coastal protections' design of Nuclear Power Plants (ASN, 2013). In according to the definition of outlier (Barnett and Lewis, 1994), a test to detect outliers in a sample of extremes is used (Hubert and Van der Veeken, 2008; Weiss, 2014). This test verifies that the considered regional extreme data sample does not contain outliers.

3.2.6.2 Secondary parameters

Secondary parameters are factors considered as significant for the statistical analysis performed. In the process to choose the optimal statistical threshold, only secondary parameters calculated for λ-cases in which all of the primary tests are satisfied, are used to perform the weighting analysis.

Seven secondary parameters are chosen for FAB application. These parameters are described in the following.

Number of regional extreme data

A sufficient number of extreme data is required to get reliable estimations of extreme events. In fact, extreme events' estimations of a sample of few extreme data are typically linked to high uncertainties. For this reason, a sufficient number of extreme data is required to define accurate statistical distributions. In the FAB method, the number of extreme events contained in the regional sample depends on the λ value and, consequently, on local sampling thresholds (Eq.3.4). In particular, more the value of λ is high and more events are contained in the regional extreme data sample.

Regional credible duration

Extreme data samples observed for a long period provide frequently estimations of extreme events linked to low uncertainties. In the FAB method, the period of observation of the regional extreme data sample (the regional credible duration) depends on the value of λ. High values of regional duration around the optimal λ value are preferred. In addition, it is preferred to define a statistical distribution stable around the optimal λ value. Being the regional credible duration a parameter of the statistical distribution, stable value of this parameters are preferred.

Adimensional degree of regional dependence

The adimensional degree $\boldsymbol{\Phi}$ of regional dependence is a statistical factor depending on the degree φ of regional dependence (recalling that it is a value between 1 and N sites of the region) as follows:

$$\Phi = \frac{N - \varphi}{N - 1} \tag{3.36}$$

This parameter assumes values between 0 and 1 representing a weak or a strong regional dependence. Being φ function of the value of λ, the value of $\boldsymbol{\Phi}$ depends consequently on the same λ value. For the same reasons exposed for the regional credible duration, a stability on the parameter $\boldsymbol{\Phi}$ is needed for each λ-case.

Value of S-χ^2

Results of the primary Pearson's chi-squared test can be used as a secondary parameter. In particular, more the dispersion value S-χ^2 is low and more the considered statistical distribution is suitable for the extreme data sample. A low value of S-χ^2 is preferred to a highest one.

Scale parameter of regional GPD

The stability of parameters of the statistical distribution around the optimal value of λ is suggested to perform a statistical analysis. In the FAB method, the regional GPD' scale parameters for the four seasons are estimated for the regional extreme data sample by the Penalised Maximum Likelihood Estimation (PMLE). For this reason, the stability of all four scale parameters is required for each value of λ.

Shape parameter of regional GPD

As for the scale parameters, the stability of regional GPD' shape parameter is looked around the optimal λ.

Regional return level for a return period T of 1000 year

The stability of return levels linked to high return period is preferred around the optimal value of λ. In the FAB method, regional return levels are computed for each value of λ. A stability coefficient is computed for every corresponding to the regional return level linked to a return period of 1000 years.

In addition, return level linked to other return periods or also a particular confidence interval can be considered as a secondary parameter instead of the parameter proposed here.

3.2.6.3 Weighting analysis on secondary parameters

A weighting analysis is only performed for all the chosen secondary parameters corresponding to λ-cases in which primary tests are satisfied. This analysis is based on the computation of a weighting for each secondary parameter. The measure of weighting has to be based on criteria provided by the expert for each of these parameters in order to get better estimation of extreme events. In particular, the stability of the regional duration, the adimensional degree of regional dependence, the scale and shape parameters of regional GPD and the regional return level, the small value of the S-χ^2 and high value of the number of regional data and the regional duration are required for the application of the FAB method. A different measure of weighting is estimated for the criteria defined for each of seven secondary parameters. The stability, the minimisation and the maximisation are the three different types of criteria considered in this analysis. Eight values of single weighting (both stability and maximisation criteria are assumed for the regional duration) are summed for each possible λ-case. More details of this weighting calculation for the three types of criteria are provided in the following.

Weighting computation for the stability of parameters

A measure of stability *M* has to be first computed for the parameters *t* that require the stability. For every parameter *t*, this measure *M* is calculated only for the *N* possible λ-cases in which primary parameters are verified:

$$M_{i_0,t} = \sum_{i=1,i\neq i_0}^{R} \left| \frac{t(\lambda_i) - t(\lambda_{i_0})}{\lambda_i - \lambda_{i_0}} \right| \tag{3.37}$$

the parameter *t* is evaluated for the particular λ_{i0} and also for the λ_i-cases around the λ_{i0}. An interval of R=40 values (R/2 before and R/2 after the considered value of λ_{i0}) is considered in this study. The measure of stability $M_{i0,t}$ is computed for each of the *N* λ-cases. More the measure of stability $M_{i0,t}$ of the parameter *t* for the considered λ_{i0} is higher and more the values of $t(\lambda_i)$ vary around the considered value of $t(\lambda_{i0})$. In addition, this measure $M_{i0,t}$ provides more relevance to the values of $t(\lambda_i)$ nearest to the considered λ_{i0}.

Now, after the computation of $M_{N,t}$ measures of stability $M_{i0,t}$, the weighting measure $W_{i0,t}$ can be formulated for each of the *N* possible λ-case as follows:

$$W_{i_0,t} = 1 + (N-1)\frac{\max(M_{N,t}) - M_{i_0,t}}{\max(M_{N,t}) - \min(M_{N,t})} \tag{3.38}$$

where the value of the weighting measure $W_{i0,t}$ varies between 1 and *N*. In for two different N cases in which the measure $M_{i0,t}$ is similar then, their weightings $W_{i0,t}$ have likewise similar values. More the value of the weighting $W_{i0,t}$ is high and more the parameter *t* computed for λ_{i0} is stable.

This type of weighting measure is calculated to evaluate the stability of the five parameters mentioned above. For the particular case of the four scale parameters of the regional GPD, a weighting measure is estimated for every season and a simple average of these four weighting values is considered in the analysis.

Weighting computation for the minimisation of parameters

The weighting measure $W_{i0,t}$ that allows the minimisation of a parameter t is estimated as follows:

$$W_{i_0,t} = 1 + (N-1)\frac{\max(t_N) - t(\lambda_{i_0})}{\max(t_N) - \min(t_N)} \tag{3.39}$$

the value of this weighting measure varies between 1 and N. More the value of the weighting $W_{i0,t}$ for the parameter t is high and more the parameter t computed for λ_{i0} is small. In the FAB method, this type of weighting is calculated only for the value of S-χ^2.

Weighting computation for the maximisation of parameters

This measure of weighting enables the maximisation of the parameters t as follows:

$$W_{i_0,t} = 1 + (N-1)\frac{t(\lambda_{i_0}) - \min(t_N)}{\max(t_N) - \min(t_N)} \tag{3.40}$$

where the value of the weighting measure $W_{i0,t}$ varies between 1 and N. This weighting formulation is similar to the last one but not identical. In fact, more the value of the weighting $W_{i0,t}$ of the parameter t is high and more the parameter t computed for λ_{i0} is big. This last type of weighting criteria is used for the number of regional data and for the regional duration.

The sum of the eight weighting measures is performed for each possible value of λ_{i0}. The value λ_{i0} with the highest total weighting is defined as the optimal λ.

Moreover, this approach that permits the computation of the optimal value of λ is enough flexible. In fact, further primary or secondary parameters can be added or removed from the analysis depending on the aim of the study. For instance, higher return levels can be preferred when dealing with nuclear safety. In this case, an additional weighting measure that maximise the return level parameter has to be considered.

3.2.7 Frequentist return levels

Regional extreme data samples are fitted to a GPD (1, *γ, k*) created by the unitary parameter of location and the estimated scale *γ* and shape *k* parameters. Seasonal effects are considered in the estimation of the regional GPD parameters (performed by the PMLE) through the use of four seasons. Differently to the RFA approach, FAB method considers only mixed GPD with $k \neq 0$. In fact, particular exponential cases ($k=0$) provide frequently return levels linked to high return periods lower compared to that computed by a GPD with $k \neq 0$. The model that fits better regional observations can be chosen only between the remaining four possible models proposed in the RFA approach (GPD, GPD_{cos}, GPD_{sin} and $GPD_{cos\ sin}$) with a no-zero value of shape parameter *k*. The best seasonal GPD model used to fit our regional sample is defined by the Akaike Information Criterion (Di Baldassarre et al., 2009; Laio et al., 2009; Mendez et al., 2008).

Now, the p-quantile y^r_p of the regional cumulative distribution function F_r corresponding to $y^r_{1-1/\lambda T}$ (Rosbjerg, 1985) can be computed as follows:

$$y_p^r = y^r_{1-\frac{1}{\lambda T}} = 1 - \frac{\gamma}{k}(1-(1-p)^k) \tag{3.41}$$

The regional quantile associated to a return level T is equal to:

$$\mathrm{y_T^r} = \mathrm{F_r^{-1}}\left(1 - \frac{1}{\lambda \mathrm{T}}\right) \tag{3.42}$$

Similarly, the return level linked to a return level of T for a site i can be calculated by the use of the local index:

$$x_T^i = F_i^{-1}\left(1 - \frac{1}{\lambda T}\right) = u_i\, F_r^{-1}\left(1 - \frac{1}{\lambda T}\right) = u_i y_T^r \tag{3.43}$$

Eq.3.41, Eq.3.42 and Eq.3.43 are used to compute regional and local return levels that can be illustrated in the regional return level plots.

Uncertainties on regional and local return levels are computed by the parametric bootstrap method proposed in the RFA approach. The bootstrap allows the simulation of new regional samples M'_r starting from the estimated distribution of the original regional observations (Efron, 1979; Weiss, 2014). New regional GPD parameters are then estimated as before by a PMLE. These new parameters enable the estimation of the statistical distribution F'_r of the new regional extreme data sample. This bootstrap process is replicated for U times. The particular quantile of the new regional distributions generated corresponds to the investigated confidence intervals. Regional confidence intervals linked to particular return levels are estimated.

In addition, a similar bootstrap procedure is applied to local estimations in order to compute their confidence intervals. Being local return levels related to regional estimations by the corresponding local index, the bootstrap has to consider in this case the variability of the regional return level and the local index. Local confidence intervals are computed as for regional confidence intervals through a bootstrap replicated for U times. Further details of this procedure are provided by Weiss (2014).

In particular cases, the bootstrap method can provide upper confidence intervals that for small return period they result bounded and for high return period they become unbounded. This variation on the curve behaviour may occur when high confidence levels (typically over 90%) are computed for regional extreme data samples generated by very high sampling thresholds. In particular, this is frequently caused by the variability of the regional GPD shape parameter k estimated by the PMLE as a value close to 0. In fact, resampling only few very extreme data, the new shape parameter k' of the new regional extreme data sample F_r' generated by bootstrap could be estimated for a value of opposite sign.

3.2.8 Bayesian return levels

The FAB method can be also used to estimate Bayesian return levels. This is useful for experts that would like to introduce in the statistical analysis priori information on the observations or that would like simply estimate return levels in a Bayesian framework. In fact, most of the authors that deal with historical data prefer to use a Bayesian inference to estimate the extreme events.

Bayesian estimations can be computed for a regional extreme data sample. This analysis leads the computation of predictive return levels y^*_R, standard estimative return levels y_R and their associated Credibility Intervals. Further details about the Bayesian inference are provided in the Annexe B.

Regional GPD parameters' vector θ composed by the regional scale γ and regional shape k parameters is defined by its posterior distribution (Eq.B.2.1). It is computed considering a non-informative prior probability distribution $f(\theta) \propto 1$ (Payrastre et al., 2011; Bulteau et al., 2015) and the penalised likelihood formulation for systematic and historical data (Eq.3.35).

In particular, the non informative-prior distribution is enough common in literature when historical data are used. It supposes that no knowledge a priori on the parameters of the statistical distribution is available. The possible formulation of a priori distribution is a difficult topic. As stated by Coles (1999), improper priors cause some problems in Bayesian estimations mainly during the resolution of the Markov Chain Monte Carlo.

The posterior distribution of the regional GPD parameters' vector θ can be computed through a MCMC algorithm. Several chains containing a number U of vectors θ of the two regional GPD parameters are calculated by the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). Convergence between chains and in chains are verified by the Gelman and Rubin test (Gelman and Rubin, 1992) and the most recent Brooks and Gelman correction of the previous test (Brooks and Gelman, 1998).

Contrary to the estimations of regional GPD parameters for frequentist return levels, the posterior distribution does not take into account the seasonality on the estimation of the regional GPD scale parameter γ. For this reason, a number U of vectors θ with a unique value of regional GPD scale parameter γ is sampled. The location parameter is equal to 1 for each of the U iterations.

The knowledge of the posterior distribution of the regional GPD parameters' vector θ allows the computation of the predictive distribution, standard estimative return levels and credibility intervals.

The predictive distribution is estimated calculating the mean of the different regional GPD distributions created for each value of θ. In addition, a burn-in of U' iterations of posterior distribution is suggested to estimate reliable predictive return levels.

Return levels obtained computing the regional GPD distribution with the mode of the vectors θ_U of parameters correspond to return levels computed in the frequentist framework. These quantiles can be defined as standard estimative return levels. Uncertainties associated to this return levels are defined by the credibility intervals. They are identified as the regional GPD distribution computed by the corresponding quantile of the vectors θ_U of regional parameters.

Return level plot of each of the three variables can be figured out recalling the relationship between quantiles x_T and return periods T equal to: $P(X>x_T)=1/\lambda T$.

Finally, the three different regional quantiles $y_{R,T}$ associated to a return period T enable to know local quantiles through the local index. Local quantiles $x_{i,T}$ of the predictive distribution, credibility intervals and standard estimative return levels are computed for a particular return period T as: $x_{i,T}=y_{R,T}/u_i$.

Further considerations are needed for the local computation of the credibility intervals. In fact, the variability of local return levels $x_{i,T}$ considers only the variability of $x_{R,T}$ (the variability of u_i is not considered in this approach).

Chapter 4

FAB APPLICATION

Une application de la méthode FAB à une base des données des surcotes de pleine mer, c'est à dire la partie considérée comme aléatoire du niveau marin, est développée dans ce chapitre.

La base de données est constituée par des sériés temporelles de surcotes de pleine mer systématiques provenant de 74 ports situés dans l'Atlantique, dans la Manche, dans la Mer d'Irlande et dans la Mer du Nord, et par 14 surcotes de pleine mer historiques.

La première étape de la FAB méthode consiste à former des régions physiques par la méthode de clustering des tempêtes définie dans l'approche AFR. Cette méthode de clustering est fondée sur la définition de trois paramètres physiques (p, Δ, η). Pour des séries temporelles avec une différente durée, une analyse de sensibilité de ces paramètres est conseillée. La formation des régions physiques est recommandée pour une fenêtre temporelle dans laquelle la plupart des marégraphes sont en fonction.

Les autres étapes de la méthode FAB sont réalisées exclusivement pour les régions comprenant des sites avec des données historiques disponibles (Région 1 et Région 2). Les seuils optimaux sont évalués pour ces deux régions qui sont ensuite statistiquement vérifiées. Des niveaux de retour fréquentistes et bayésiens sont estimés et une comparaison préliminaire des résultats est montrée. Enfin, la méthode FAB est appliquée sur la même base de données sans surcotes historiques. Cette analyse supplémentaire permet d'indiquer le rôle des données historiques dans l'analyse régionale des événements extrêmes.

4.1 Introduction

FAB method is applied to a database of systematic and historical skew surges illustrated in the following. This application permits the practical steps of the FAB methodology. Starting from a database of systematic and historical variables, this method enables the estimations of regional and local return levels.

After the collection of systematic and historical data from different locations, FAB method requires to pool sites in homogeneous regions in order to form regional samples and to treat them with a statistical analysis. In particular, physical homogeneous regions are formed through the definition of three parameters p, Δ and η. These parameters leading the formation of storm clusters have to be calibrated for time series with different periods of observation. For this reason, a sensitivity analysis of these three parameters is performed in this application. This sensitivity analysis is recommended when the FAB method is applied to a database of time series with different recording periods. Moreover, it is important to recall that in the FAB application case of the paper illustrated in Annexe A, the three parameters to form physical homogeneous regions were not calibrated. In that case, the parameters used were the same proposed in the RFA approach.

Now, physical homogeneous regions can be statistically verified. A double threshold approach is used and so a statistical threshold has to be defined in every site of the region. The statistical threshold chosen corresponds to a number of storms λ per year above this threshold. This λ value is common for every site in the region. The optimal value of $\lambda_{opt,r}$ has to be found for each physical region r in order to get the best performances of this methodology on the extreme estimations. The identification of the $\lambda_{opt,r}$ based on a weighting analysis of sensitivity indicators is performed for the skew surge database.

The regional extreme data sample is then formed and the degree of dependence as the local and regional credible durations are computed. The GPD distribution is estimated for the sample of regional storms considering the seasonality of skew surges in the frequentist framework. No seasonality on skew surges is considered in the estimations of Bayesian return levels. Frequentist and Bayesian regional return levels and frequentist and Bayesian local return levels of each site by the use of the local indexes are estimated. This can allow the comparison between frequentist and Bayesian estimations always recalling the different concept of probabilities that the two statistical inferences have (Annexe B).

In such way, FAB method is applied to a database of 74 skew surge time series composed by systematic and historical data. In particular, the process of formation of physical and statistical regions states that sites in which historical skew surges are available belong to the Region 1 and the Region 2. For this reason, the FAB method is only focused on the frequentist and Bayesian estimations of return levels of these two regions.

4.2 Skew surge database

In this application, the maritime variable analysed is the skew surge. A skew surge database composed of systematic and historical skew surges is used to perform the FAB method.

Systematic skew surges are generated from time series recorded by tide gauges in different sites located along the coasts of Atlantic Ocean, the North Sea and the English Channel. Conversely, historical skew surges are collected in some French sites of Atlantic Ocean and the English Channel.

For this reason, the database of skew surges used in this application is formed by a systematic skew surge database in which historical skew surges are successively added.

4.2.1 Definition of skew surge

The sea level can be schematically shown as the overlap of two contributions: the predicted astronomical tide (deterministic part of sea level) and the instantaneous surge or residual (stochastic part of water level provoked by the meteorological variables).

As Fig. 8 shown, skew surge is defined as the difference between the maximum water level measured by a tide gauge and the maximum predicted astronomical tide computed during the same tidal cycle (Simon, 2007; Kerdagallan, 2013; Weiss, 2014).

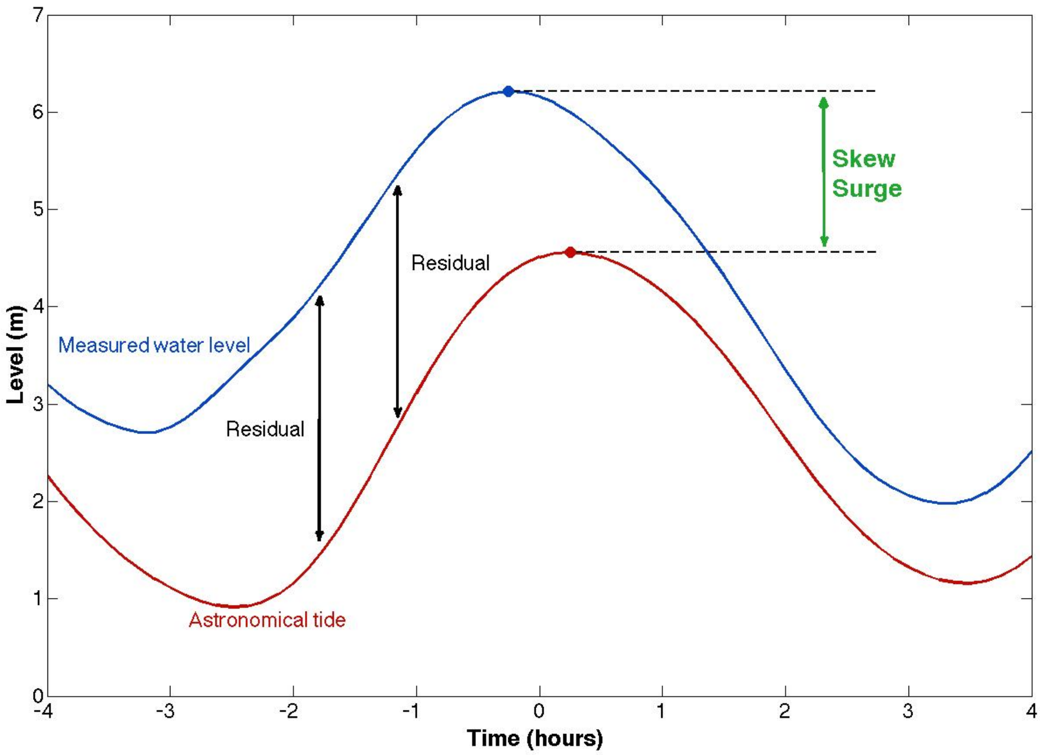


Fig. 8 – Definition of skew surge. Source: SurgeWatch Glossary (University Southampton)

Skew surges are preferred to instantaneous surges when dealing with extreme sea level. In fact, notwithstanding it is complicate get accurate and continuous time series of instantaneous surges, they are not influenced by a possible time lag between astronomical tides and sea levels. In addition, astronomical tides and residuals can have a possible dependence in the instant of maximum sea level and this has to be considerate in the computation of extreme sea levels (Weiss, 2014).

Finally, the knowledge of predicted astronomical tides and sea levels enables the calculation of time series of systematic skew surges and historical skew surge values.

4.2.2 Systematic skew surge database

The database of systematic skew surges is generated through sea levels data recorded by tide gauges. The quality of these water level measurements is checked by different national ocean data centers that provided these data. A big effort has been done during the acquisition of tidal gauge measurements in order to create a database of systematic skew surges as biggest as possible.

In addition, astronomical tides or residuals must be available to compute time series of systematic skew surges.

Data recorded by 74 tidal gauges located in French, Spanish, British and Belgian coasts of Atlantic Ocean, English Channel, Irish Sea and North Sea are considered in this study. The location of the 74 sites in which measurements of water levels are available is shown in Fig. 9. Time series of skew surges are successively computed for each of the 74 sites.

Sea level measurements of the 23 tidal gauges placed in French coasts are provided by the SHOM (Service Hydrographique et Océanographique de la Marine) and REFMAR (Réseaux de référence des observations marégraphiques) every hour. Time series of skew surges for French sites are created by the astronomical tides generated through the official software PREDIT provided by the SHOM.

Sea level measurements of 46 tidal gauges placed in British coasts are provided by the BODC (British Oceanographic Data Center) every hour until 1992 and every 15 minutes after 1993. In addition, the BODC supplies also residuals. This allows the computation of astronomical tides and consequently of the skew surges for each tidal cycle.

Only 2 Spanish sites in which water levels are provided by IEO (Instituto Español de Oceanografía) every hour are considered. Skew surges are calculated for the 2 Spanish sites by astronomical tides computed through the software SHOMAR (provided by the SHOM).

MVB (Meetnet Vlaamse Banken) has kindly provided the water levels for 3 Belgian sites (Nieuwpoort, Oostende et Zeebrugge) and the astronomical tides associated every 5 minutes. Skew surges are then calculated for these 3 Belgian ports.

Fig. 9 - Location of 74 tide gauges considered in this study. Source of the map: Google Maps.

This database can be considered as an update of the skew surge database used by Weiss (2014) for the application of the RFA approach. In particular, French and British data are available until 2017, data from sites like Dielette (France), Port Ellen (UK), Portbury (UK) and Bournemouth (UK) are exploitable, Belgian tide gauge measurements are considered and an important restoration of the time series at Saint-Nazaire had been done by the SHOM (Ferret, 2016) and these data are now available since 1821 (this time series' reconstruction has been used in this study even though the SHOM had recently retired these data to definitively validate them).

Before the computation of systematic skew surges for all sites, the sea levels measured in each site must be corrected by a likely significant eustatism. The eustatism is the modification of mean sea level caused mainly by ocean floor motion and ice sheet melting. Not ever the eustatism gives a remarkable contribution in sea levels and so not ever it has to be taken into account in sea level corrections. Eustatism is considered significant when the regression curve slop of annual mean of sea levels can be practically considered void (p-value of T-Student test lower than 5% (Watson, 2016)). The eustatism for 67 of 74 sites was already calculated by Weiss (2014) until the year 2010. In cases in which eustatism is considered as significant, sea levels (recorded before 2010) are corrected by the same eustatism computed by Weiss (2014) apart from the site of Saint-Nazaire in which the eustatism has been entirely recalculated. For sea levels between 2011 and 2017, eustatism is considered irrelevant (only 6-7 years more of water level measurements).

Sea levels recorded by tide gauge at Saint-Nazaire has been corrected by an eustatism of 1,76 mm/year until the year 1893. Before this period, the regression curve slop is broken and not more tendency on sea levels are detected. For the 7 additional sites used in this study (3 British sites, 3 Belgian sites and 1 French site), no eustatism is detected and sea levels have no need to be corrected.

Every site has its own period of tidal recordings that depends on many factors. For instance, this could depend on the time in which tidal gauge is put into operation or damage on measurements during a big storm. The longest period of sea levels measurements is available for the tide gauge located at Brest: 156.57 years of recordings since 1846 (Fig. 10).

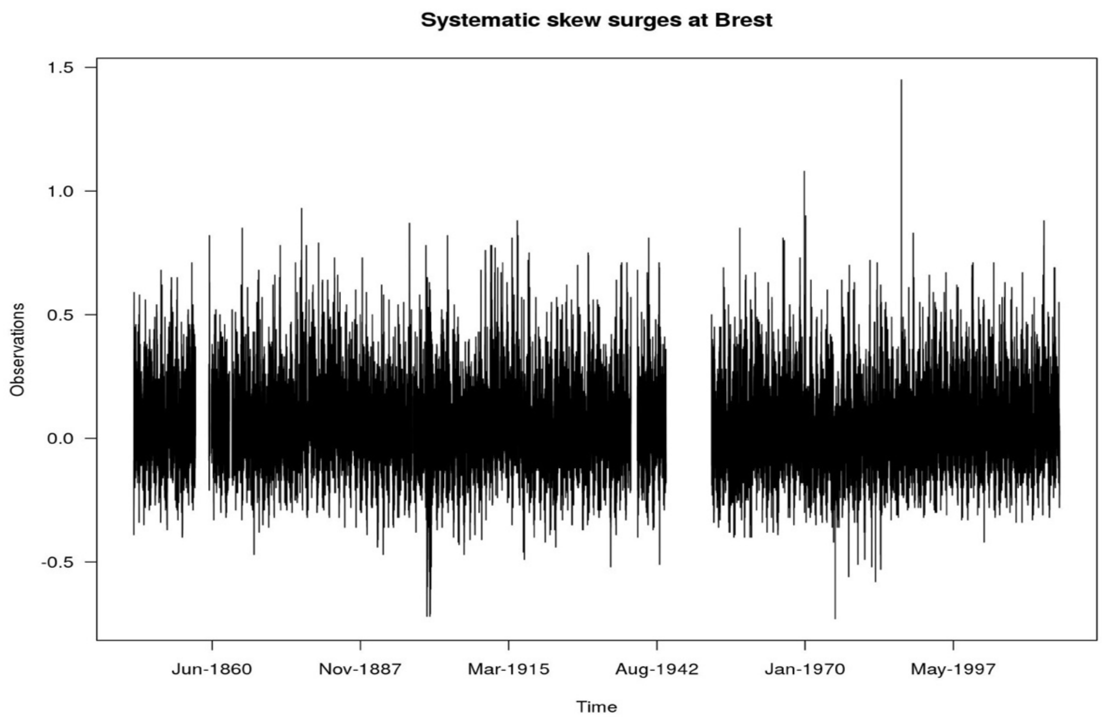


Fig. 10 - The longest time series of skew surge database (measurements recorded by Brest tidal gauge)

4.2.3 Historical skew surges

Historical skew surges are collected through a deep investigation on several sources. Historical skew surges are punctual numerical values of skew surges that are not associated to any time series of water levels recorded by the tidal gauge.

In this application, only the 14 historical data recovered in 3 different sites (La Rochelle, Dieppe and Dunkirk) during the first part of this PhD are considered. The additional 17 historical skew surges founded by Florian Regnier during his internship at EDF R&D LNHE in 2017 are not taken into account. This allows a correct comparison between results of this application and results of previous studies (e.g. the study presented in the Annexe A). In addition, the most of these 17 historical skew surges are recovered in locations in which any time series of systematic skew surges is available and, for this reason, not all of them could have been exploitable in this study.

However, although the 14 historical data considered may not seem a large number of additional data, their contribution in statistical analysis of extreme events can be considerable. When historical data are available, it is unusual to get directly the skew surge value from documentations. For this reason, maximum sea level and the associated maximum astronomical tide have to be known in order to compute the skew surge corresponding to the event of the past period.

In particular, the studies of Garnier and Surville (2010), Gouriou (2012) and Brehil (2014), performed on the collection or on the modelling of historical events at the site of La Rochelle, allow the recovering of 9 historical skew surges (Tab. 1) that are added to the time series of systematic skew surges (Fig. 11).

Gouriou (2012) recovered in archives of the tidal gauge at La Rochelle the 4 oldest historical skew surges (1866, 1867, 1869 and 1872). The historical skew surges of 1924 and 1940 are instead computed through the study of Garnier and Surville (2010). In fact, this study provides the maximum water level and the maximum astronomical tide for the two events of the past.

The three most recent historical events recovered at La Rochelle (1941, 1957, 1999) are calculated by numerical simulations (Brehil, 2014). Fig. 11 and Tab. 1 illustrate the historical skew surge happened on 27th December 1999 (Martin storm) that is computed by the difference between maximum water level and predicted max tidal level provided by Brehil, 2014. This is the biggest skew surge never seen at La Rochelle. Unfortunately, the tidal gauge located in La

Rochelle has not operated in that day. Despite several questions are asked on the truthfulness of this large value, the historical skew surge of 2.17 meters is considered in this application.

Fig. 11 shows these 9 historical skew surges (the green crosses) in the time series of systematic skew surges at La Rochelle. Historical skew surges are isolated data points and, for this reason, the historical period of observations, or better said, what it is happened before or after these historical events is unknown.

Historical skew surges – La Rochelle								
11 Jan. 1866	*27 Jul. 1867*	*20 Jan. 1869*	*10 Dec. 1872*	*9 Jan. 1924*	*16 Nov. 1940*	*16 Feb. 1941*	*15 Feb. 1957*	*27 Dec. 1999*
1.11 m	**0.8 m**	**0.87 m**	**0.96 m**	**1.27 m**	**1.49 m**	**1.37 m**	**1.11 m**	**2.17 m**

Tab. 1 – Historical skew surges recovered at the site of La Rochelle

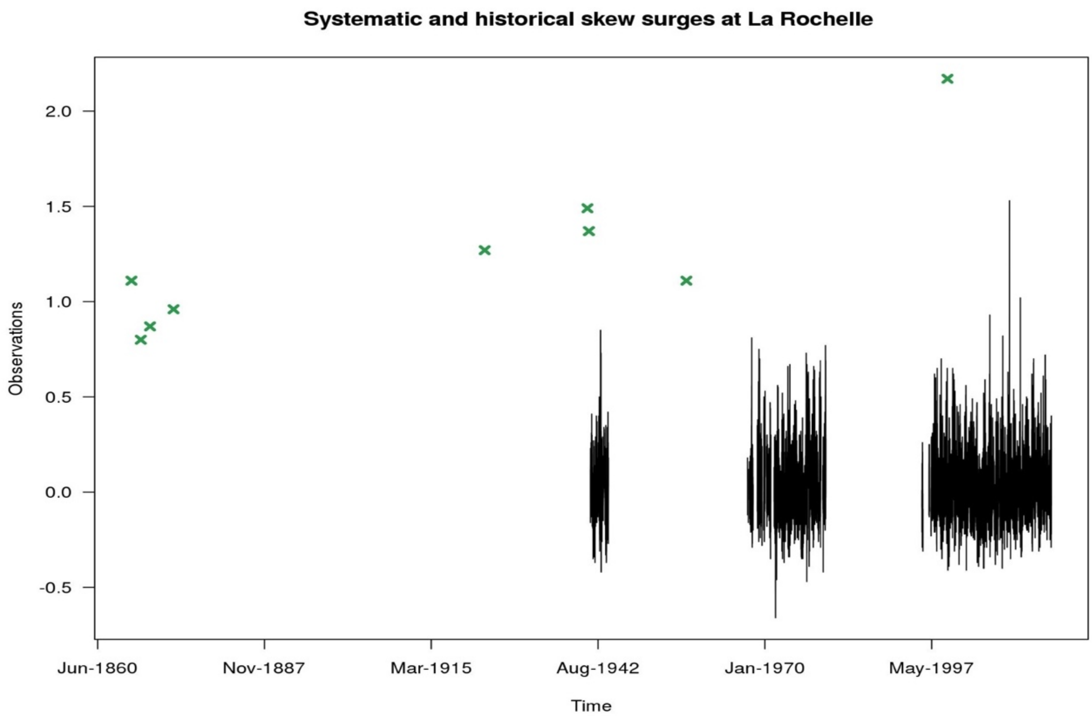


Fig. 11 – Systematic skew surge recorded by the tide gauge at La Rochelle (in black)and historical skew surges (in green) recovered for La Rochelle

The technical report of the project NIVEXT (Daubord et al., 2015) provides the only one historical value of skew surge recovered at the site of Dieppe (Tab. 2). As done before for La

Rochelle, this historical skew surge is added to the time series of systematic skew surges recorded by Dieppe's tidal gauge (Fig. 12).

Historical skew surges				
Dieppe	**Dunkirk**			
17 Dec. 2004	*29 Nov. 1897*	*1 Mar. 1949*	*1 Feb. 1953*	*2 Jan. 1995*
1.3 m	**1.75 m**	**1.56 m**	**2.22 m**	**1.18 m**

Tab. 2 - Historical skew surges recovered at the site of Dieppe and Dunkirk

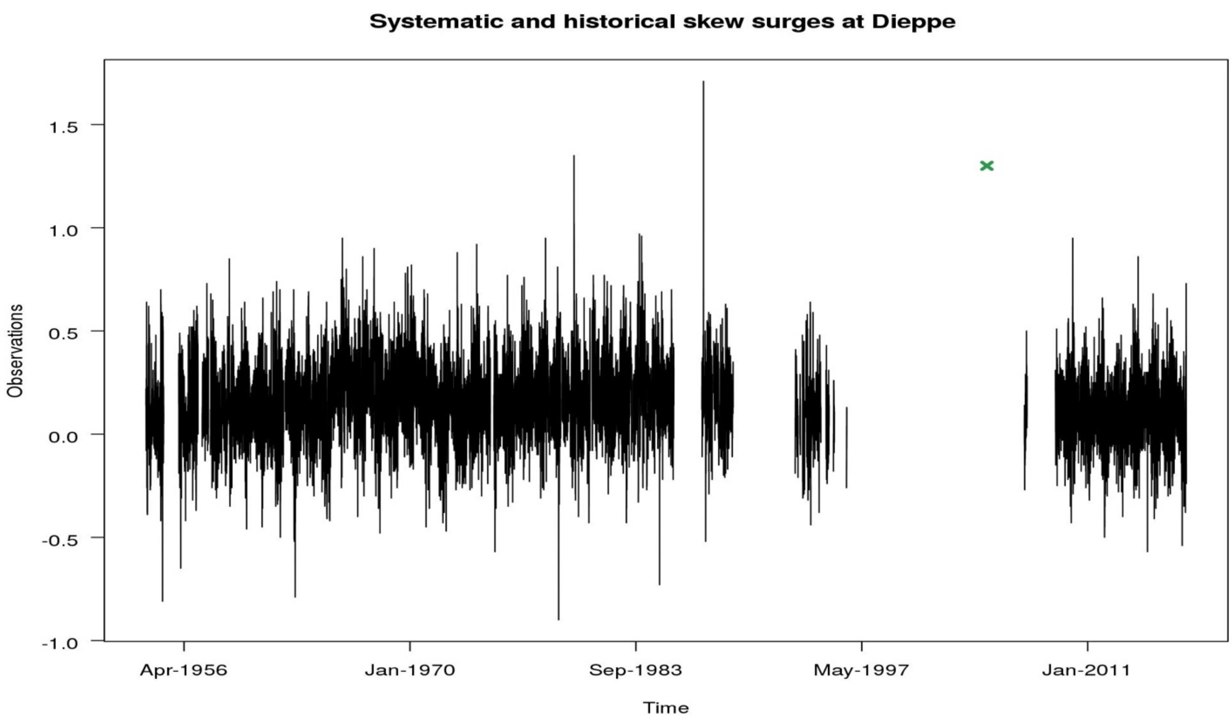


Fig. 12 - Systematic skew surge recorded by the tide gauge at Dieppe (in black) and historical skew surges (in green) recovered for Dieppe

On the contrary, several documents are available for the site of Dunkirk. Four additional historical events, are taken into account in this application (Tab. 2). In particular, Dunkirk was impacted by storms in 1897, 1949, 1953 and 1995 that provoked significant skew surges. All of these skews surges are collected by different sources.

Le Cornec and Peeters (2009) state that the maximum water level reached 7.36 meters at Dunkirk on the 29th November 1897 (the municipal archives of Dunkirk is the primary source). The SHOM provides astronomical tide values for many locations by the website "*maree.shom.fr*". Astronomical tides calculated at Dunkirk by the SHOM for that day are equal to 5.71 and 5.78 meters. Making the assumption that the eustatism is the same computed by Weiss (2014) for the period 1956-2010 (1.5 mm/year), the historical skew surge of 1.75 meters is considered.

Le Gorgeu and Guitonneau (1954) state that the maximum sea level of 7.3 meters was reached on the 1st March 1949 at Dunkirk. The associated astronomical tides of 5.77 and 5.83 meters is provided by the SHOM website. Taking into account the eustatism of 1,5 mm/year at Dunkirk (Weiss, 2014), the historical skew surge of 1.56 meters has been computed.

The historical skew surge of 2.22 meters (including the eustatism) has been founded at Dunkirk on the 1st February 1953 by an internal study performed in the past at EDF R&D. This value was already used as systematic skew surge by Weiss (2014). However, a different value of skew surge (2.13 meters) has been used by Bardet et al. (2011) for this event.

Maspataud (2011) evaluates a skew surge of 1.15 meters at Dunkirk on 2nd January 1995 (primary source is Service Maritime de Nordn-S.I.L.E.-Les Dunes de Flandres). Considering the same eustatism of Weiss (2014), the historical skew surge of 1.18 meters is considered.

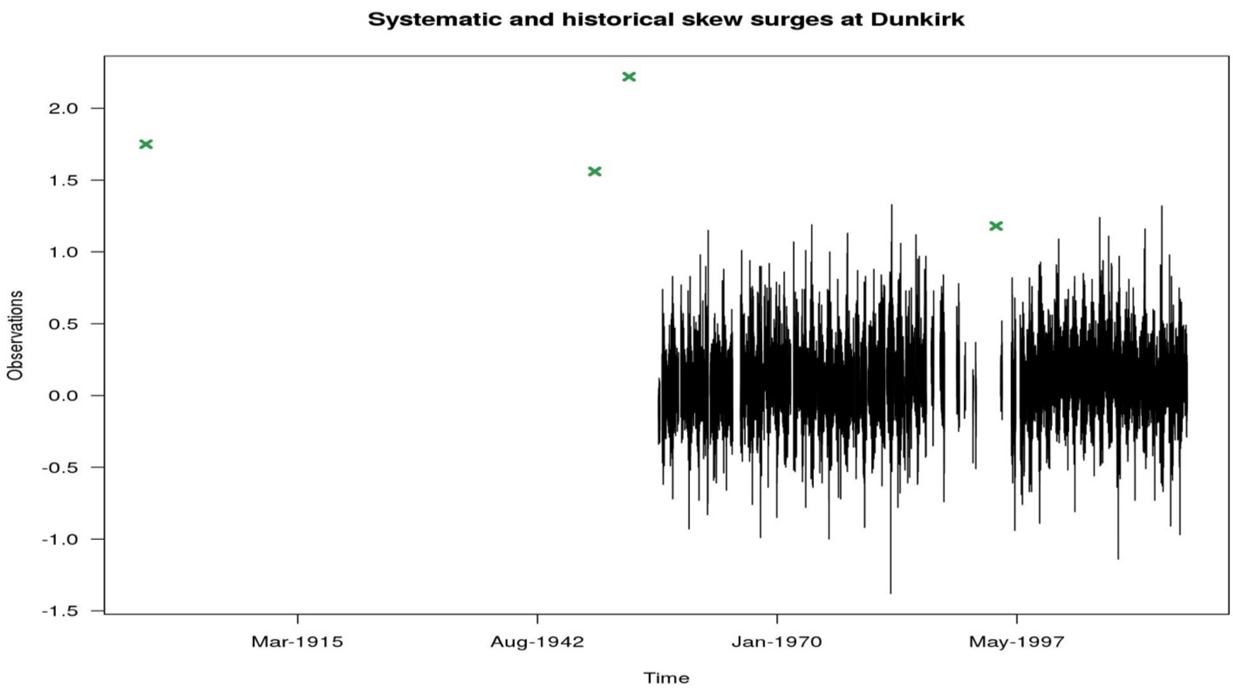


Fig. 13 - Systematic skew surge recorded by the tide gauge at Dunkirk (in black) and historical skew surges (in green) recovered for Dunkirk

The historical skew surges recovered at Dunkirk are merged with time series of systematic skew surges recorded by tidal gauge located at Dunkirk (Fig. 13).

Even if for historical skew surges computed at Dunkirk the eustatism calculated for the period 1956-2010 is taken into account, historical skew surges calculated at La Rochelle are not corrected by a likely eustatism. In fact, no information is known about a possible correction already done (mainly for the 3 historical skew surges originated by numerical models). The consideration of an eustatism calculated between 1941-2010 (Weiss, 2014) for many events of the past (exactly 4 of 9 events founded) happened in 19th century could not be proper. For these reasons, historical skew surges recovered at La Rochelle have not been corrected by any eustatism.

The eustatism is not taken into account for the historical skew surge at Dieppe because, being a skew surge quite recent, the eustatism is considered as negligible.

Notwithstanding a complete documentation of past events is provided by authors of scientific studies, historical numerical values must be constantly criticized because the value recovered not always is the correct one. In particular, historical skew surges of the 19th century precise to the centimeter could called into question. In addition, some historical skew surges recovered at Immingham and Portsmouth by other sources for some events happened after the year 2010 had different numerical values compared to the new systematic skew surges computed at these sites (the difference is of few centimeters).

For these reasons, these 14 historical skew surges recovered have been validated as proposed in Chapter 2. For the moment, any validation has been performed for these events.

4.3 Physical homogeneous regions

Physical homogeneous regions are formed in FAB method by a process based on the computation of probabilities p_{ij} that storms occurred in a site *i* might impact also the site *j* and vice versa.

This method to form homogeneous regions is the same proposed in the RFA approach. In this approach, the process to form regions was principally developed for continuous time series of modelled significant wave heights in which data are observed for a common period in every site. When data come from different tide gauges, the recording period can be not the same in each location.

For this reason, the probabilities $p_{i,j}$ that a storm impacts the site *i* and the site *j* for a same storm have to be computed for a period in which the most part of tide gauges are operating. A particular period has to be defined to form homogeneous regions. This allows the optimisation of formation of physical homogeneous regions. In addition, a sensitivity analysis on the three parameters that lead the formation of physical region is recommended. Further details are illustrated in the following.

All the probabilities p_{ij} should be ideally computed from data approximately recorded for a same period of time at each site in order to form accurate typical storms footprints and consequently precise physical regions. More time periods of tide gauge recordings from different sites are common, more the probabilities p_{ij} are less impacted by mutual time periods of data and they could be properly hierarchically compared and clustered. In fact, hierarchical clustering method provides better typical storm footprints if applied to probabilities p_{ij} that are computed for approximately the same period of time. Computing probabilities p_{ij} for a same time period at each site is possible only when dealing with modelled time series that are generated for a same time period.

In this application, skew surges are available in 74 tide gauges that start and finish to operate in different periods. In addition, they have different periods of failure on gauging. For this reason, a time window in which most of tide gauges operate has to be chosen in order to estimate proper values of p_{ij} and to get accurate physical homogeneous regions.

Fig. 14 illustrates the number of working tide gauges during every year. Tide gauges never work at the same time.

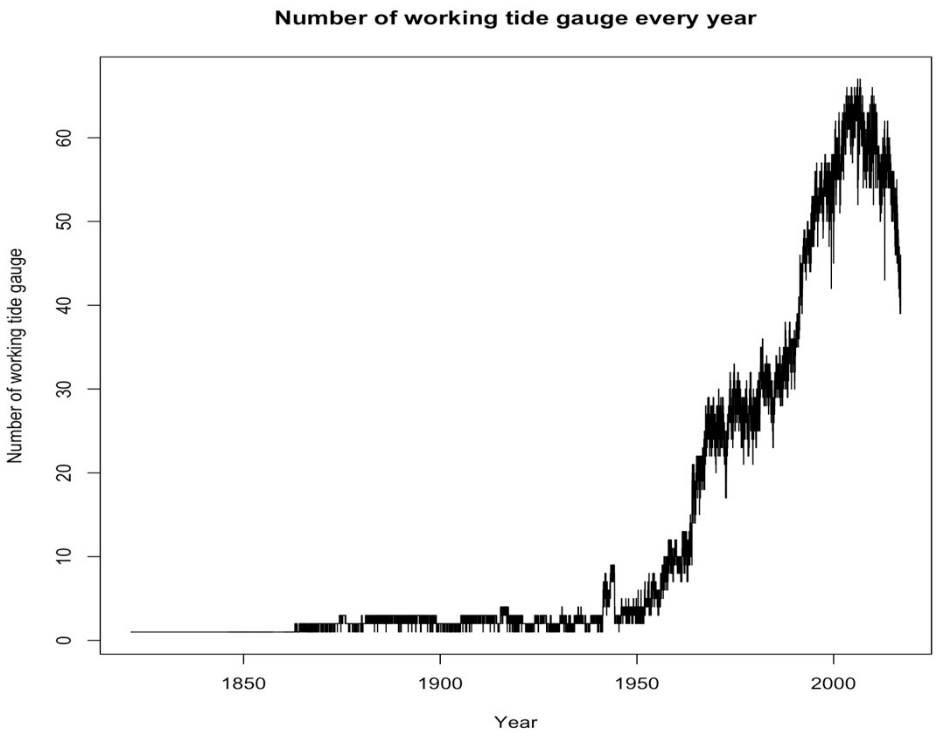


Fig. 14 – Number of operating tide gauges per year

A clear breaking point on the number of working tide gauges is marked between the beginning and the end of the year 1950. In fact, the most part of the 74 tide gauges available in this database starts to operate and to measure water levels from the beginning of the year 1950. For this reason, the time window in which the probabilities p_{ij} are considered to form physical homogeneous regions lasts 67 years (from 1951 to 2017). In this time window approximately 35 of 74 tide gauge are operating on average.

Another shortest time window with more tide gauges operating on average could be chosen but it is important to specify that more the time window is reduced and more the probabilities p_{ij} might be inaccurate. On the contrary, the extension of the time window before 1951 would mean that the probabilities p_{ij} are computed also for time periods in which only 2 tide gauges work on average.

In this application, typical storm footprints are computed only for storms occurred after the 31st December 1950.

A generic site *i* must typically have a computing occurrence probability between all other interested sites in order to use a hierarchical method. In fact, if a site *i* has not common period with a site *j*, the probability p_{ij} cannot be computed. Unfortunately, 5 of our 74 sites face with this problem. These 5 sites are removed when hierarchical clustering method is applied.

After the formation of physical regions, they can be reintroduced in the region in which they evidently belong. In this study, the sites removed are Saint-Servan, Dielette, Le Crouesty, Port Bury and Moray Firth. A particular case arises for the site of Le Crouesty. Removing the other 4 sites, the site of Le Crouesty have any common period with the site of Le Verdon. A choice of which site had to be removed is carried out depending on length of time series. Tide gauge at Le Verdon operates for the most part of the years between 1951-2000 and until to November 2000 while the tide gauge at Le Crouesty starts the recordings from the end of the year 2000.

A test removing the site of Le Crouesty instead of the site of Verdon is performed at the end of this analysis. Using the parameters for storm clusters chosen in the following, the physical homogeneous regions are precisely the same of that founded in this application (Fig. 21).

4.3.1.1 Parameters to detect storms

The formation of physical homogeneous regions is applied only to storms occurred after the year 1950. Notwithstanding, storm clusters have to be created for all the available events. In fact, the determination of clusters for all the events is necessary to define the regional extreme data samples composed by all extreme events occurred in the considered region.

Physical storms are detected by a spatiotemporal declustering and they are defined by three parameters (p, Δ, η). Extremes happened in neighbour sites η during a temporal interval Δ are generated by the same storm. In particular, two or more extreme values that exceed the local physical threshold (represented by the p-value) belong to the same storm if they occur in η-nearest sites during an interval of Δ hours.

The value of 0.995 is widely used in literature of extreme events (Mendez et al., 2008; Di Baldassarre et al., 2009; Laio et al., 2009; Weiss, 2014) for the parameter p and, for this reason, a sensitivity analysis on the other two parameters Δ and η is performed.

The variation of these two parameters could generate significant differences in the formation of storm clusters. This would impact the formation of physical and statistical homogeneous regions and successively regional and local estimations of extreme events. In fact, storm clusters are physical element that represent an effective storm. The goodness of representation of its physical process depends on the parameters used to form clusters. For this reason, a sensitivity analysis on these parameters that generate storm clusters have to be performed.

The same sensitivity analysis is recommended for each application of the FAB method to databases of variables recorded in different locations.

Sensitivity analysis for the temporal parameter *Δ*

Before to perform the analysis performed for the temporal parameter Δ, it is important to recall that the identification of a valid temporal parameter Δ is necessary to get a satisfactory representation of storms. In particular, a single storm can be identified like two or more different storms if the value of Δ hours is too little and, on the contrary, if the value of Δ hours is excessively high, two or more different storms can be identified as a single one.

For this reason, three different values of temporal parameter Δ hours are considered to detect storms in this application. This analysis is performed separately for storms happened before the year 1951 that are not used on the process of formations of physical regions and for storms occurred after the year 1950. In fact, as a result of the previous study (Fig.3), the duration of a storm propagation Δ has to be higher when only few tide gauges operate. More time is needed to a storm moving to another operating location that it should ideally be more far. For this reason, storms happened before the year 1951 and after the year 1950 are independently analysed.

In particular, the duration of a storm, or better the difference in time between the first and the last extreme event of a storm cluster, is examined for the three different values of Δ hours: 25 hours, 48 hours and 72 hours. A value of 25 hours instead of 24 hours is considered to allow the realisation of a time of at least two tidal cycles in a same location. Histograms of number of storms for 12 hours' progressive intervals of storm durations are produced to interpret easily the results of this sensitivity analysis.

The results of this analysis are illustrated below for a spatial parameter of $\eta=16$. In any case, analyses with other spatial parameters have been performed. Behaviours of histograms with other spatial parameters are similar to these of Fig. 15 and Fig. 16.

A Δ of 25 hours is selected as the most satisfactory value by the analysis performed for storms happened after the year 1950. In fact, as the histograms of Fig. 15 and Fig. 16 shown, more the value of Δ hours is higher and more the storms have a longer duration. The histogram for a Δ of 72 hours (on the right side of Fig. 16) illustrates that more than 50 storms of the 1036 storms detected in total last more than 6 days and two storms last approximately 25 days. The detection of a storm that lasts 25 days has not any physical sense. Similar considerations can be pointed out for the case of a Δ of 48 hours in which 1189 storms are detected. This case is illustrated in the histogram on the left side of Fig. 16.

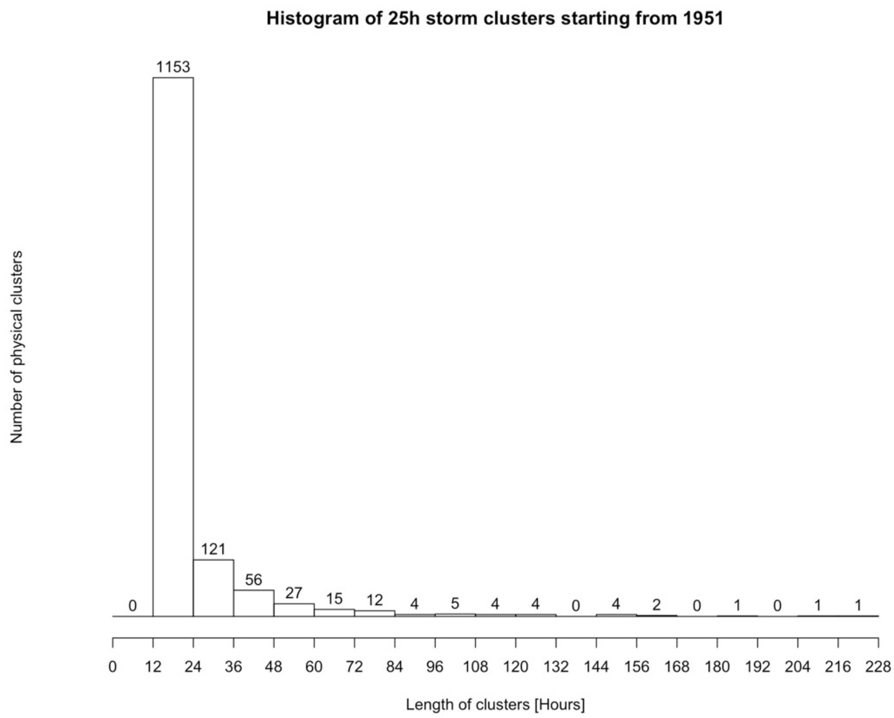


Fig. 15 – Histogram of the number of storm clusters created by a Δ=25 hours and a η=16 in function of their storm duration (in hours) for storms happened after the year 1950

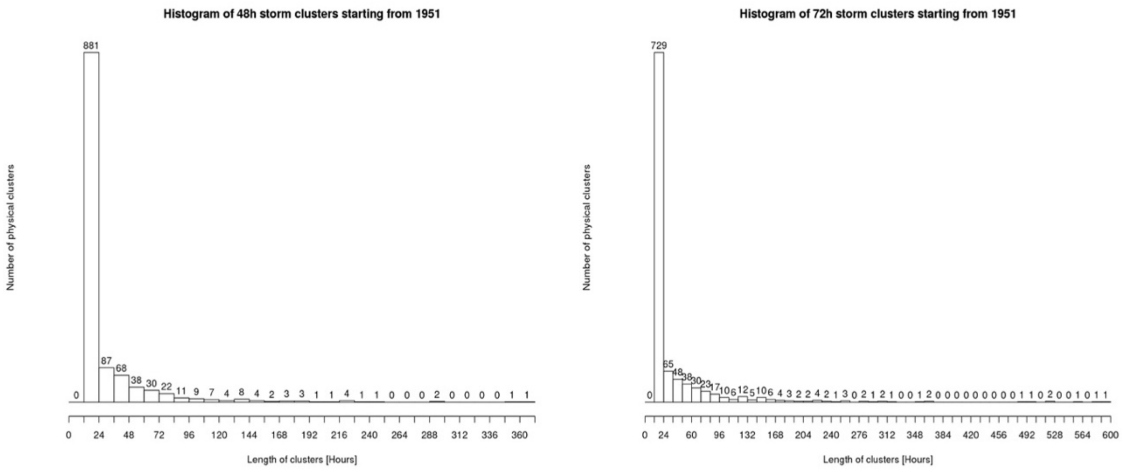


Fig. 16 - Histograms of the number of storm clusters created by a Δ=48 hours (on the left) or a Δ=72 hours (on the right) and a η=16 in function of their storm duration (in hours) for storms happened after the year 1950

Conversely, for a *Δ* of 25 hours, only 9 storms of the 1410 detected last more than 6 days and, in addition, the biggest one lasts 9 days. A temporal parameter *Δ* of 25 hours can be considered satisfactory for this application and it is used to detect storms happened after the year 1950.

The same analysis is performed for storms occurred before the year 1951 (Fig. 17 and Fig. 18).

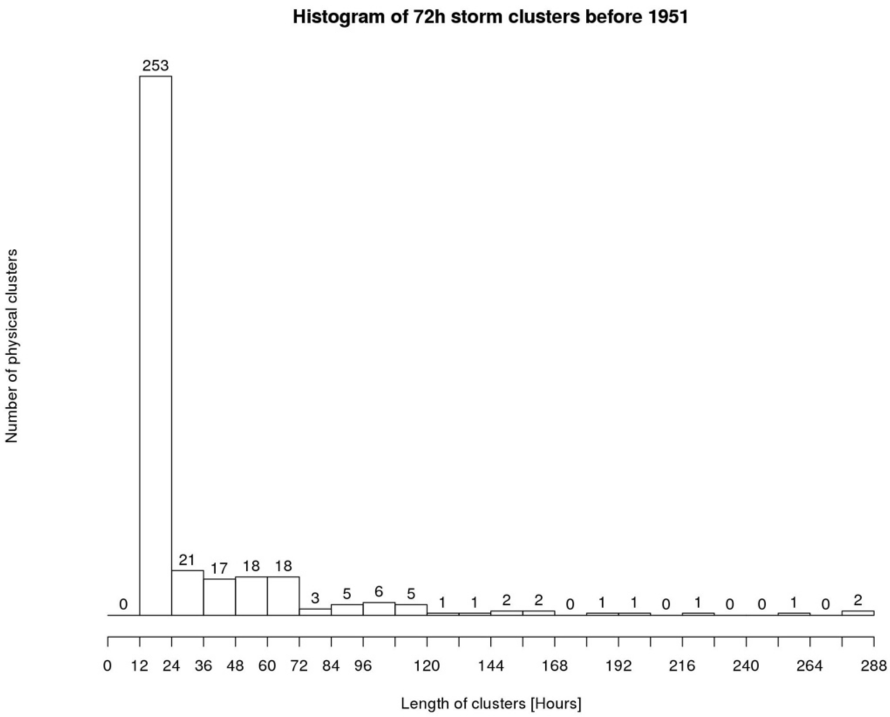


Fig. 17 - Histogram of the number of storm clusters created by a Δ=72 hours and a η=16 in function of their storm duration (in hours) for storms happened before the year 1951

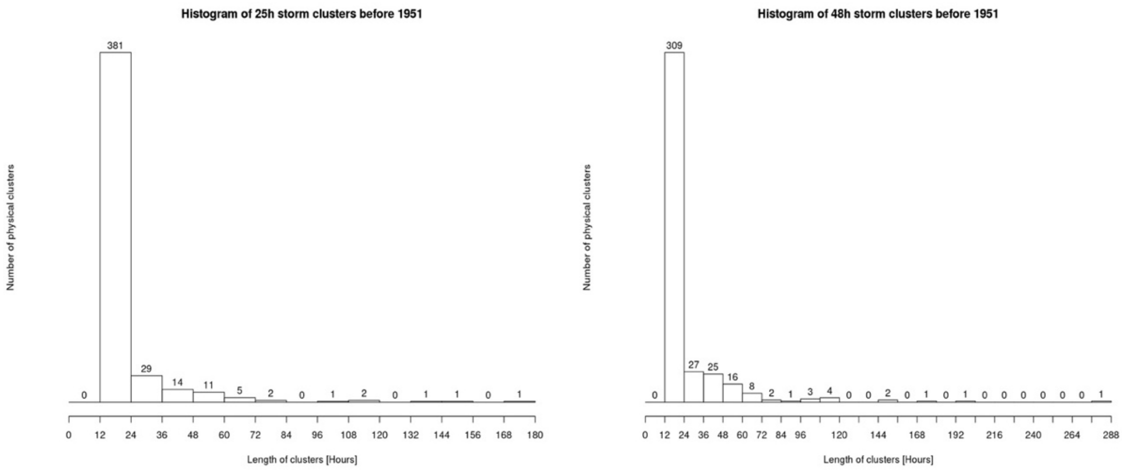


Fig. 18 - Histograms of the number of storm clusters created by a Δ=25 hours (on the left) or a Δ=48 hours (on the right) and a η=16 in function of their storm duration (in hours) for storms happened before the year 1951

During the years before the 1951, only few tide gauges operate. The sensitivity analysis of the temporal parameter Δ is performed for values of 25 hours, 48 hours and 72 hours. The histogram for a Δ of 25 hours (on the left side of Fig. 18) shows that 381 storms of 448 detected last less than 1 day. In addition, only a few number of storms has a duration higher than 48 hours. On the contrary, the histogram created for a Δ of 48 hours (on the right side of Fig. 18) and 72 hours (Fig. 17) shown as the storm duration is more uniform. A shorter percentage of storms for these two cases last more than 1 day and only few storms last more than 6 days (the biggest one lasts approximately 10 days). Notwithstanding both cases give satisfactory results, a Δ of 72 hours is considered for storms occurred before the 1951 to give a sufficient time to a storm to be detected to a farther operating location. Finally, storms happened before 1951 are created for a Δ of 72 hours.

These results obtained by this sensitivity analysis confirm that the division performed for extreme storms before the year 1951 and after the year 1950 had to be performed. In periods of time with less working tide gauges, the Δ value is higher than in time periods in which several tide gauges work.

Sensitivity analysis for the spatial parameter η

The choice of a valid temporal parameter Δ of 25 hours for the storms occurred after the year 1950 allows the performance of the sensitivity analysis for several values of the spatial parameter η. In particular, a stability on the composition of physical regions is investigated. A valid η value is important to detect correctly a storm.

A same storm can be detected as two or more different storms if the spatial parameter η is a little value. On the contrary, two or more storms can be detected as a single storm if the spatial parameter η is assumed as a high value.

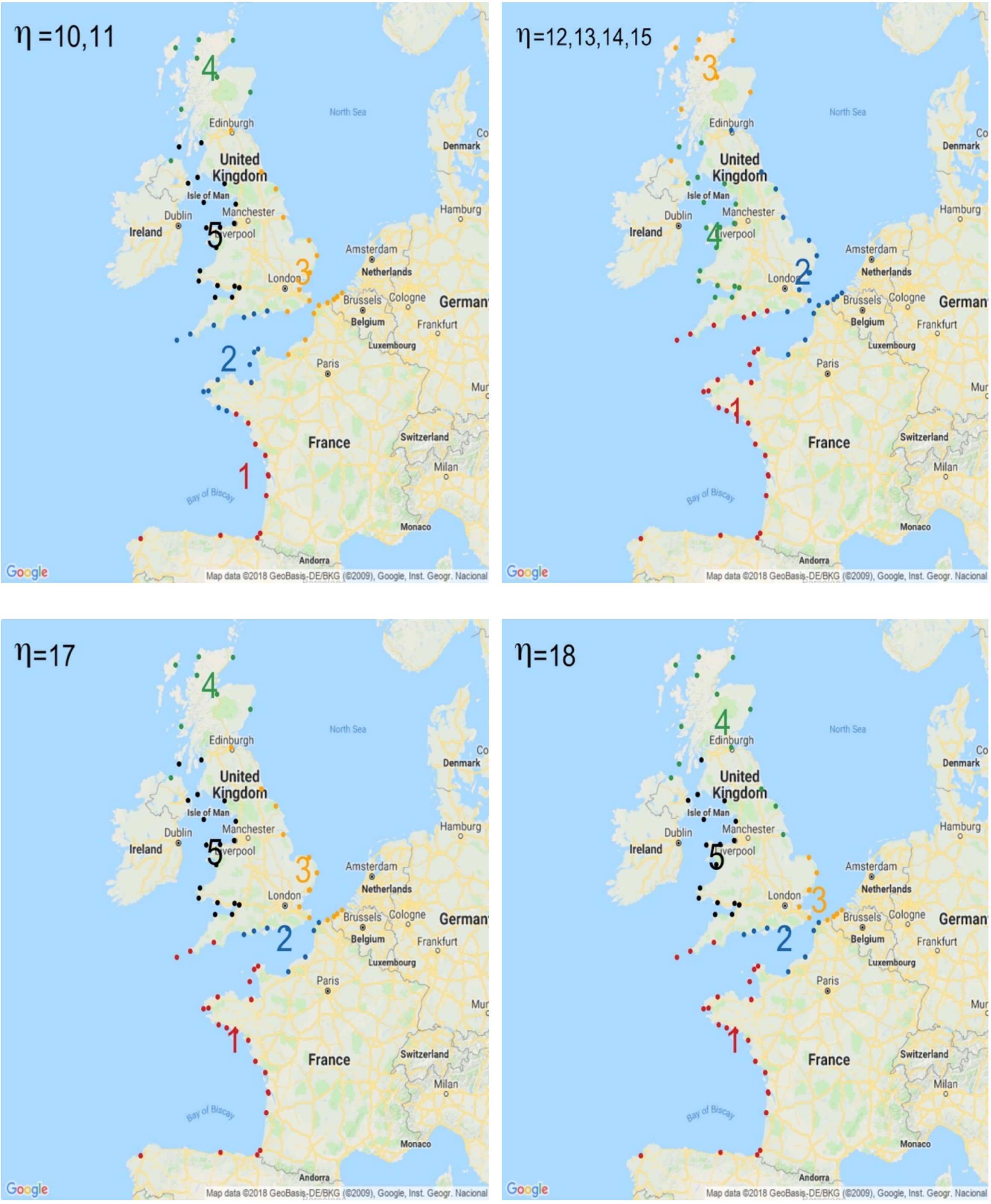


Fig. 19 – Physical regions formed for different values of the spatial parameter η (10,11,12,13,14,15,17,18) around the value parameter of η=16 selected and using the selected temporal parameter Δ of 25 hours. Source of the map: Google Maps

Several compositions of physical regions are obtained varying the spatial parameter η between 10 and 18. These 9 cases are considered in this analysis. The η value has to be a value neither too big and nor too small compared with the total number of available locations.

Fig. 19 figures out the compositions of the physical regions for these different values of η. For a spatial value between 12 and 16, the composition of physical regions is the same. The higher parameter η between 12 and 16 is considered for this study as the valid spatial parameter that provides more extended clusters.

The spatial parameter η of 16 nearest sites is used for the formation of storm clusters happened before and after the year 1950. Storm clusters are formed in this study for the following parameters: $p=0.995$, $\eta=16$ and $\Delta=72$ *hours* or $\Delta=25$ *hours* depending respectively on the period of the event: the first value of the parameter Δ is used for storms before the year 1951 and the second value of the parameter Δ for storms after the year 1950.

4.3.1.1 Physical homogeneous regions

The formation of storm clusters of skew surges enables the formation of physical homogeneous regions. The method to form physical regions is the same used in the RFA approach.

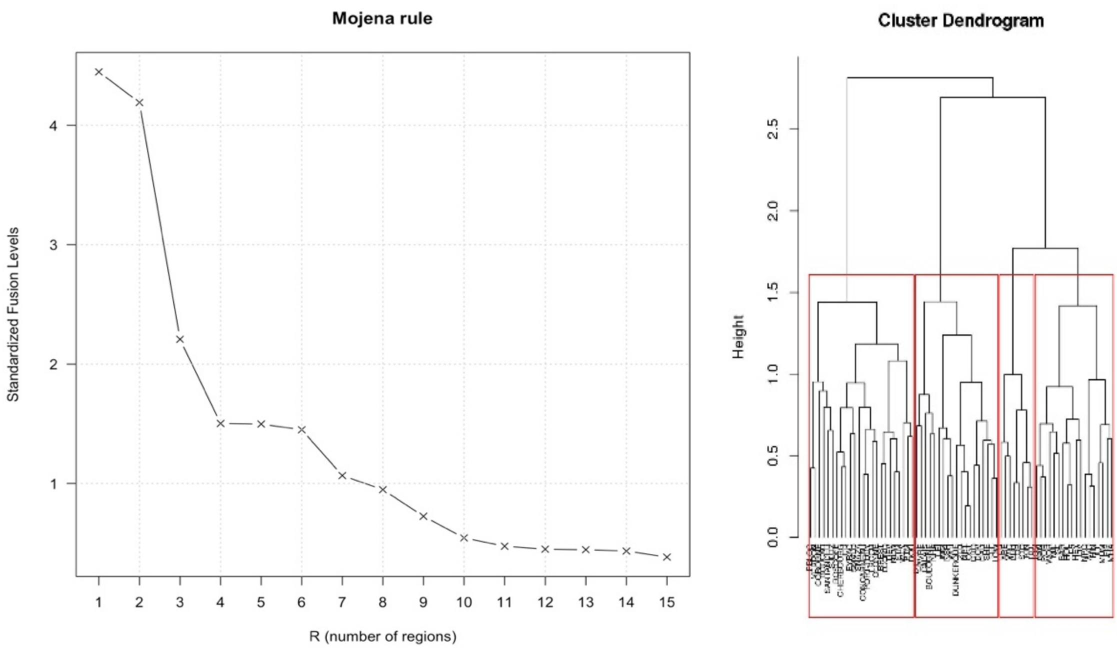


Fig. 20 – The graph of the Mojena's stopping rule (on the left) and the cluster dendrogram (on the right)

The probabilities p_{ij} are computed for storms occurred after the year 1950. The resulting dissimilarity indexes d_{ij} can be now estimated. The hierarchical clustering method proposed by Ward (1963) is performed in order to select the most typical configuration of storms' footprints through the most significant jump of the dendrogram heights (Mojena, 1977). For this application, the most significant jump of the Mojena's stopping rule is obtained for 4 regions (Fig. 20).

Fig. 21 – The 4 physical homogeneous regions founded by the typical storm footprints for events happened after the year 1950. Source of the map: Google Maps

Four physical homogeneous regions are thus formed for the 74 sites of the skew surges' database (Fig. 21). The composition of the physical regions is similar to that calculated by Weiss (2014) of Fig. 22 and to the typical storms' footprints obtained by Haigh et al. (2016) for the extreme events of the SurgeWatch database.

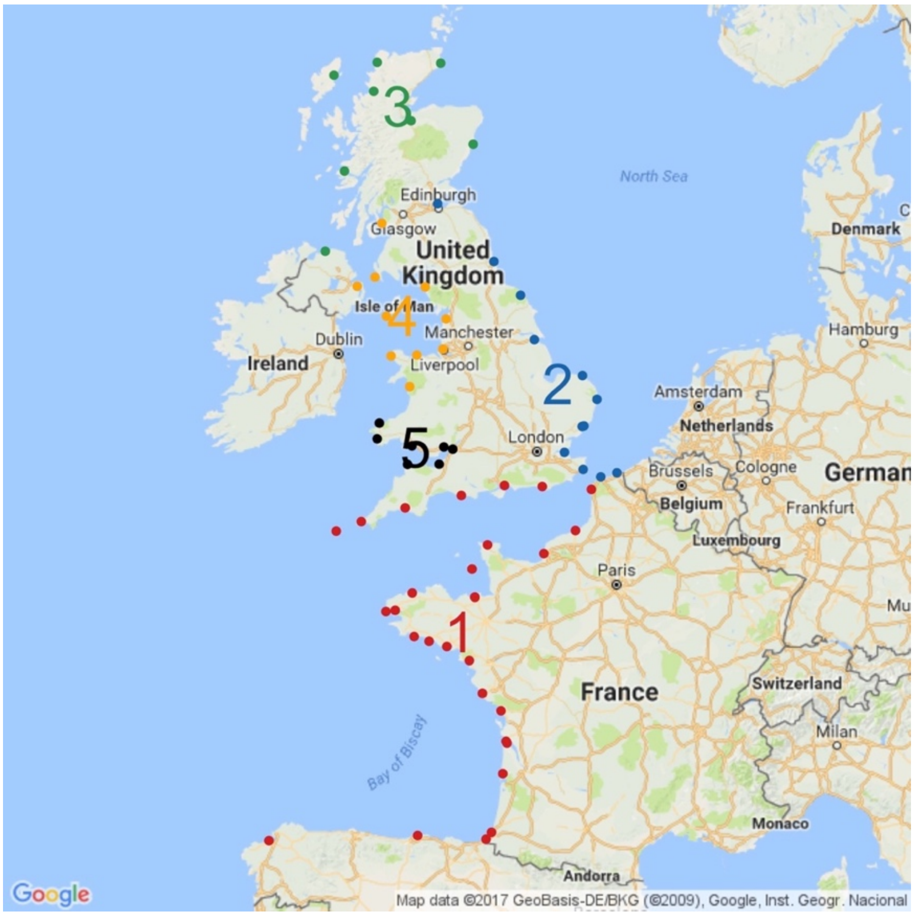


Fig. 22 – The five physical and statistical regions founded in the PhD thesis of Weiss 2014

In particular, Weiss (2014) obtained in his study 5 physically and statistically homogeneous regions (Fig. 22) selecting $p=0.995$, $\Delta=24h$ and $\eta=14$ as valid parameters. These values are founded by Weiss (2014) after an analysis on the biggest storms well-known as, for instance, the Lothar and Martin storms that impact principally coasts of the Bay of Biscay between 26th and 28th December 1999.

However, the statistical threshold of a particular site is higher compared with its physical threshold. Statistical thresholds have to be chosen in such a way to get a λ extreme values on average each year in every site. For this reason, local statistical thresholds are computed fixing a common λ value for all sites of the region. A value of $\lambda=1$ has been used in the RFA application for skew surges by Weiss (2014) in all the five regions.

The slight differences in the composition of physical regions obtained in the FAB application and the RFA application performed by Weiss (2014) impact especially the sites located in the English Channel. Different borders between the Region 1 and the Region 2 are caused by the introduction of the three Belgian sites that extend the Region 2 and by the correction of the R function "*hclust*" used previously by Weiss (2014) to apply the hierarchical clustering method of Ward (1963). In particular, Murtagh and Legendre (2014) affirm that the R function '*hclust*' must be used with squared Euclidean distances d_{ij} to implement correctly the Ward method. In this application, squared distances d_{ij} are properly used.

4.4 Computation of optimal sampling threshold

In the FAB method's application, the optimal λ is the results of the analysis of all the more important parameters of the regional analysis. In particular, it is selected through a weighting analysis as the best λ case between 176 different λ-cases. The optimal λ case corresponds to a regional extreme data sample that satisfies all the primary tests and has obtained the best global weighting value on secondary parameters.

In this application, physical homogeneous regions containing sites with available historical data are considered. Historical skew surges are available for three sites: La Rochelle located in the Region 1 and Dieppe and Dunkirk in the Region 2.

For this reason, the primary tests are performed in the regional extreme data sample of Region 1 and Region 2 for every λ value between 0.25 to 2 by steps of 0.01. The 176 λ-cases for each region are analysed. Only λ-cases verified are successively used for the weighting analysis of secondary parameters.

Depending on the λ value, these parameters are computed in the frequentist analysis' case for the Region 1 (Fig. 23) and for the Region 2 (Fig. 24).

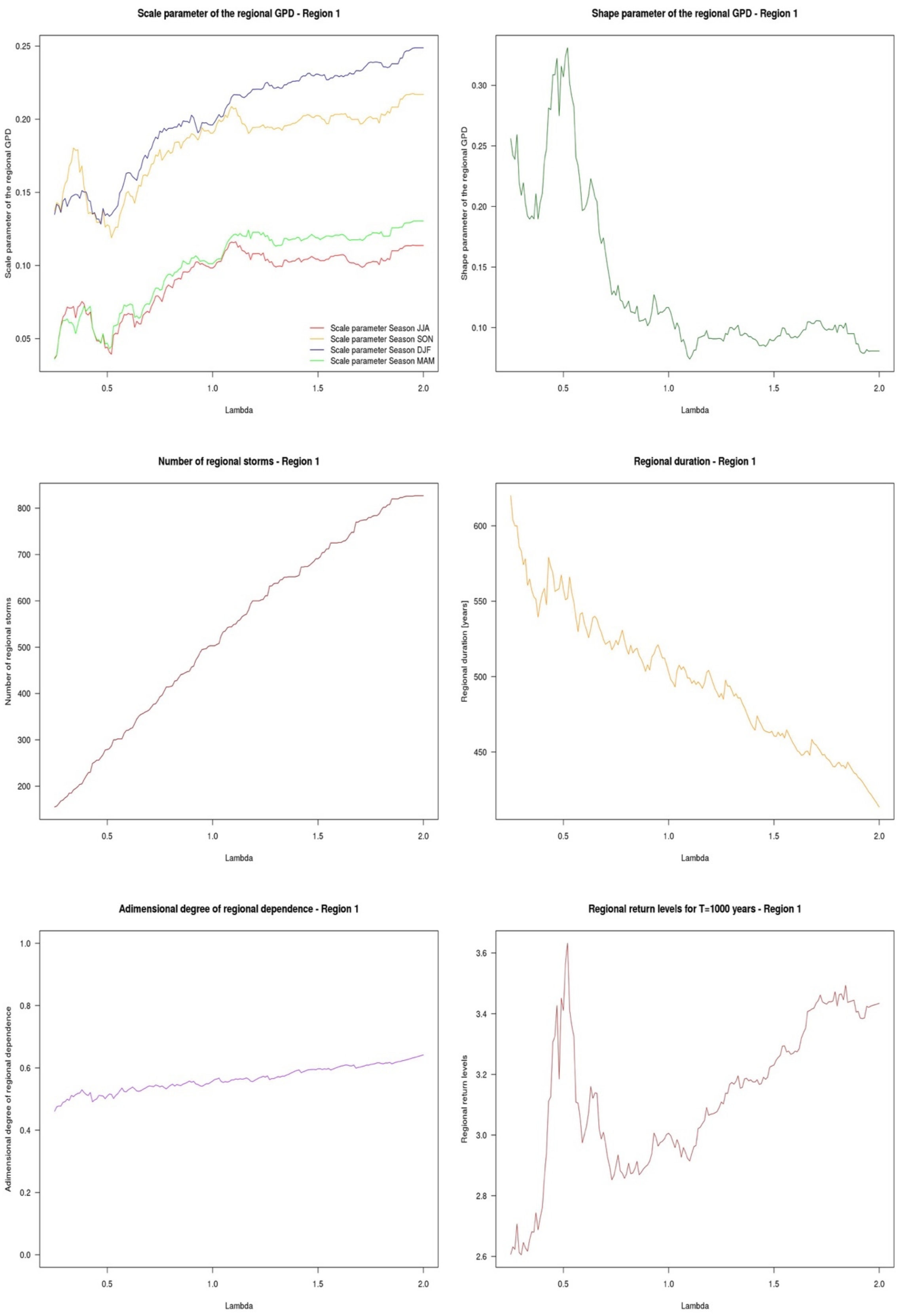


Fig. 23 – Sensitivity analysis of secondary parameters used for the Region 1

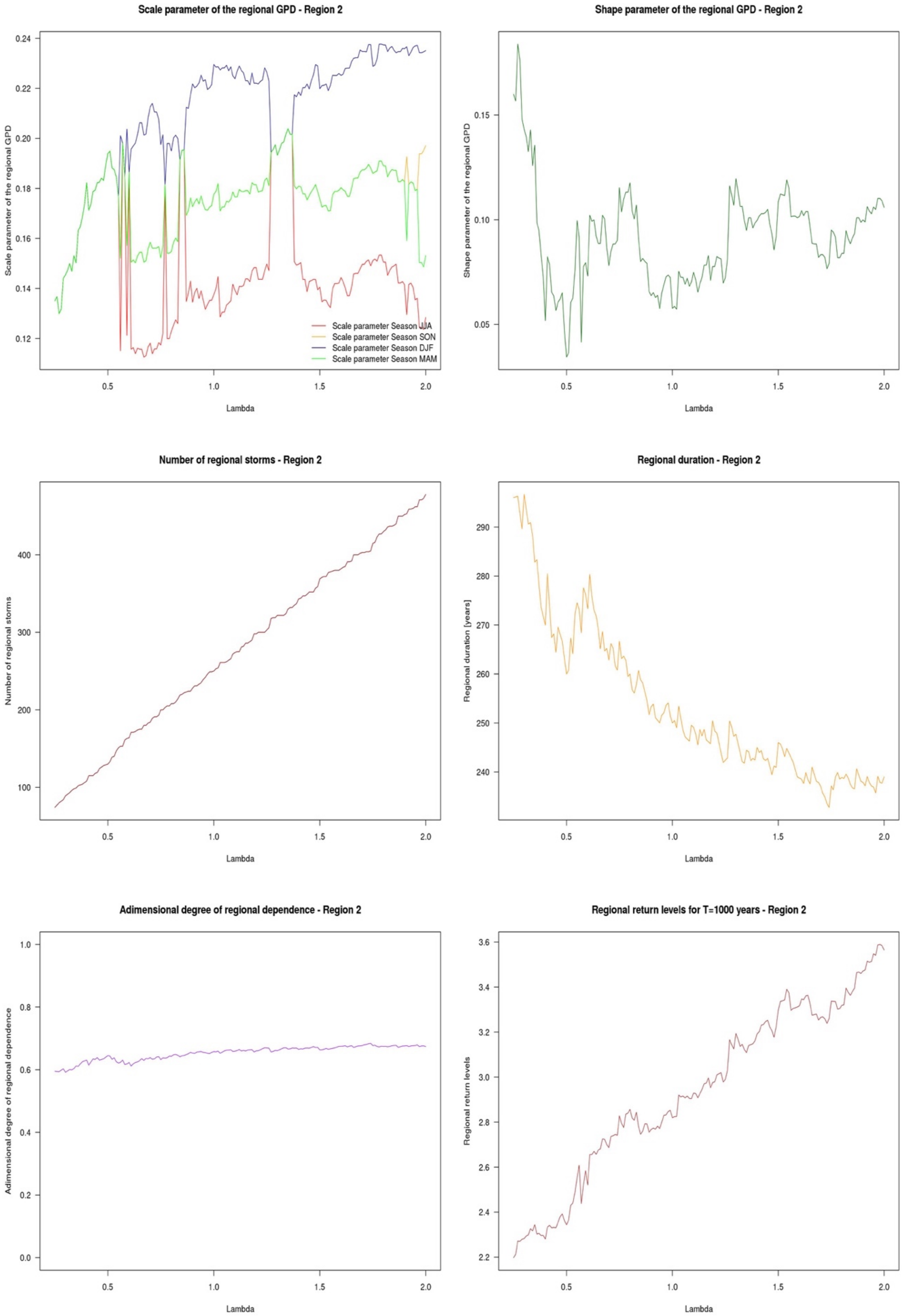


Fig. 24 – Sensitivity analysis of secondary parameters used for the Region 2

Furthermore, secondary parameters are computed for the Bayesian analysis' case for the Region 1 and for the Region 2 through the use of the "*nsRFA*" R package. This package enables the solving of the MCMC algorithm and the computation of the posterior distribution of the regional GPD parameters.

In the Bayesian case, same types of primary and secondary parameters are used. All the five primary parameters proposed do not depend on the framework of the statistical analysis. On the contrary, three of the seven secondary parameters are differently estimated in the two statistical framework of analysis: the two regional GPD parameters and the regional return levels linked to a return period T of 1000 years. In particular, the mode of the $U=100000$ iterations of regional GPD parameters' vectors θ is computed for each value of λ between 0.25 and 2 (Fig. 25 and Fig. 26). Standard estimative regional return levels linked to a return period T of 1000 years can be now computed for each value of λ.

These return levels could correspond to that computed in frequentist context for a PMLE without considering the seasonality of skew surges. In fact, in the particular case in which the seasonality of skew surges is not considered in frequentist estimations, results of primary and secondary parameters are identical for the frequentist and for the Bayesian analysis' cases.

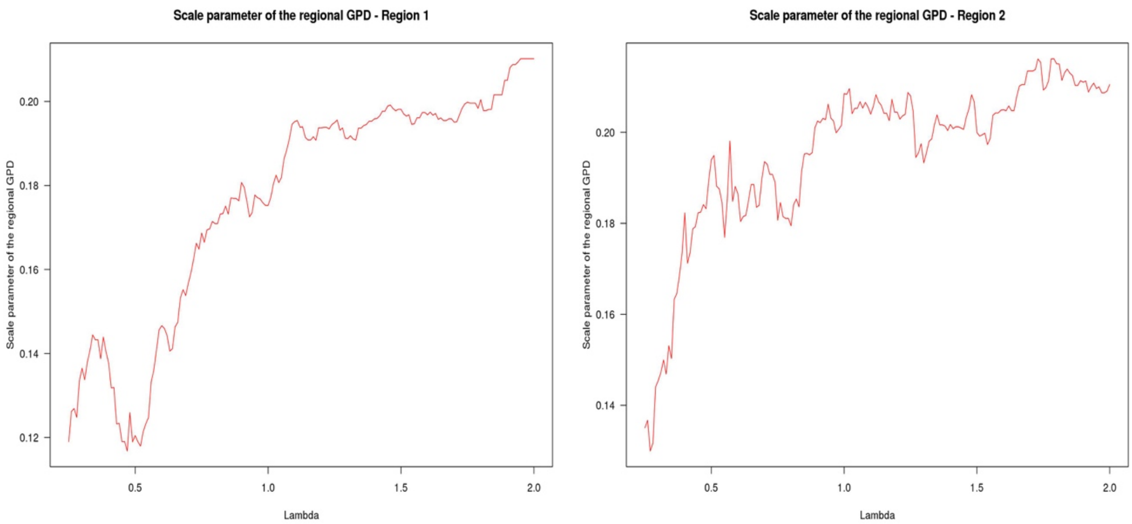


Fig. 25 – Sensitivity analysis of regional Bayesian scale parameter for Region 1 and Region 2

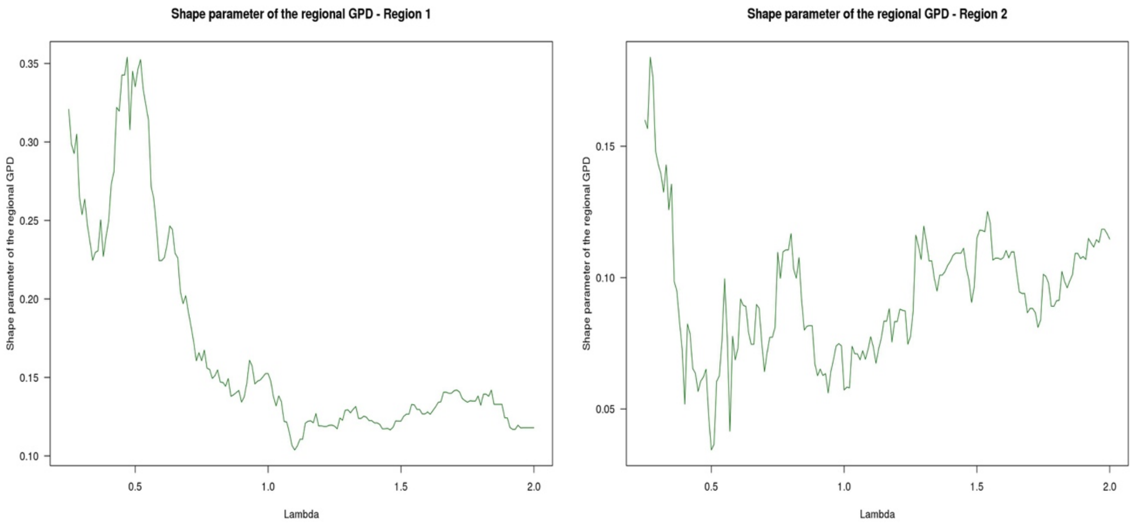


Fig. 26 – Sensitivity analysis of regional Bayesian shape parameter for Region 1 and Region 2

The weighting analysis of the secondary parameters provides the same optimal λ value for the frequentist and Bayesian cases. In particular, the optimal λ value for the Region 1 corresponds to a value of 0.84 storms on average per year and to a value of 1.79 for the Region 2.

A further test is performed for the frequentist case. Using the weighting of the upper value of the regional confidence interval for a return period T of 1000 years instead of the regional return level linked to a return period T of 1000 years. Same optimal λ values are identified through this alternative test in both regions.

In any case, a particular clarification has to be provided for the application performed in the scientific publication for NHESS (Annexe A). In that case, the sampling threshold of Region 1 was identified for a λ value of 0.36. This λ value had been defined after a visual look on a total of twelve parameters for each possible value of λ. In that case, sensitivity analysis' results had been translated in an optimal value of threshold without the application of an unequivocal approach of computation. In particular, several other values of λ can be considered as optimal in that application. That choice was considered the better one to illustrate the benefit of the use of historical data on the regional analysis. Moreover, in the study of Annexe A, the regional analysis was applied to a database of skew surges not yet updated with the three Belgian sites. Physical homogeneous regions were calculated without performing a sensitivity analysis on parameters used to identify the storm clusters.

4.5 Physical and statistical regions

The statistical homogeneity of physical regions is verified only for the particular region considered. A physical homogeneous region linked to the optimal λ value has already been statistically verified. In fact, the statistical homogeneity test is a primary test and it is already verified for optimal λ values. Furthermore, , the statistical homogeneity of all physical homogeneous regions for the optimal values of λ obtained by in the previous analysis has to be tested in order to show the definitive map of the physical and statistical homogeneous regions.

For the two optimal λ, regions except the Region 4 of Fig. 21 are statistically homogeneous. For this reason, a further division into two different sub-regions has to be performed for Region 4. The new sub-regions are retested by the homogeneity test and now they appear as statistically homogeneous. Finally, five physical and statistical homogeneous regions are founded in this study for the both optimal λ identified for the Region 1 and the Region 2. The five physical and statistical homogeneous regions are shown by the map of Fig. 27.

Fig. 27 – Physical and statistical regions for lambda=0.84 and lambda=1.79. Source of map: Google Maps

4.6 Frequentist estimations of return levels

The statistical analysis is performed for the two regional extreme data samples defined for the Region 1 and the Region 2. Estimations of the frequentist return levels are computed for several return periods *T*.

Tab. 3 and Tab. 5 illustrate all the main elements of the regional analysis obtained by the application of the FAB method in a framework of frequentist estimation.

Tab. 4 and Tab. 6 show the results of the regional statistical analysis for the Region 1 and the Region 2. In particular, regional return levels associated to a return period of 1000 years, their upper bounds of the 90% confidence intervals and the percentage of the relative width of these intervals are figure out.

These results are displayed in the return level plots of Fig. 28 and Fig. 29. Confidence intervals are calculated by a bootstrap of $U=10^5$ new regional extreme data samples.

In addition, empirical regional extreme data are positioned in return levels by the use of a Weibull position (Weibull, 1951) in Fig. 28 and Fig. 29. Historical data are displayed by red dots in order to recognise the significance of historical data compared with systematic data. Grey dotted lines are used to illustrate the 90% confidence intervals.

Regional parameters of the Region 1				
Optimal λ	*Scale param. of regional GPD (season DJF)*	*Shape parameter of regional GPD*	*Regional credible duration (years)*	*Adimensional degree of regional dependence* $\boldsymbol{\Phi}$
0.84	**0.197**	**0.112**	**517.86**	**0.544**

Tab. 3 – Regional parameters computed by the FAB method applied to the Region 1

Results for the Region 1		
Regional quantiles for T=1000ys	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
2.89	**3.39**	**34**

Tab. 4 - Regional return levels computed by the FAB method applied to the Region 1

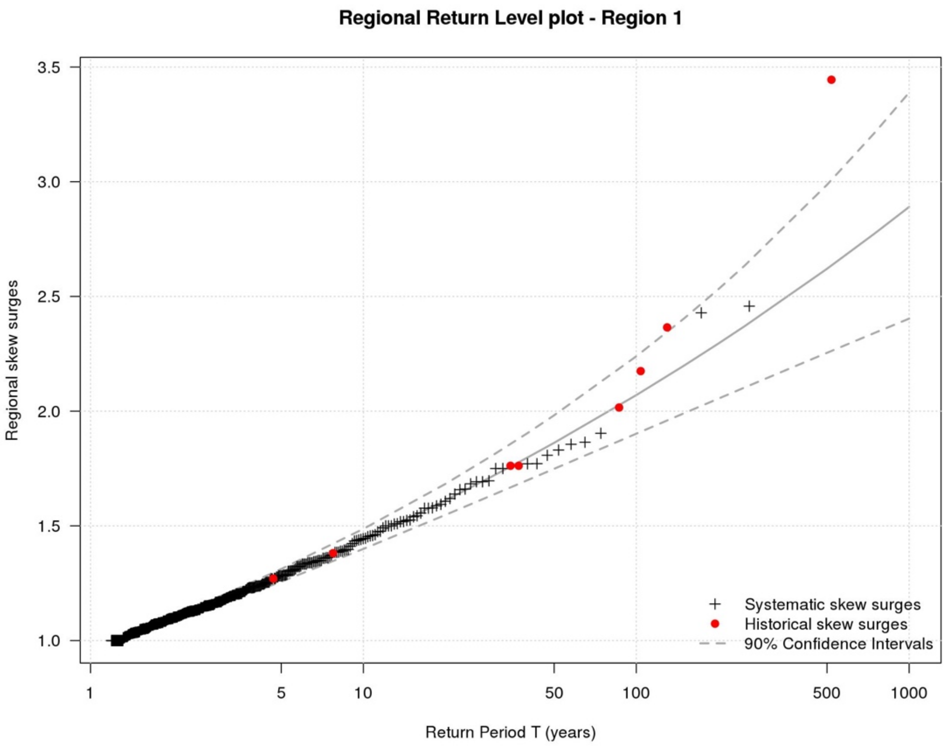


Fig. 28 – Regional return level plot for the Region 1

Fig. 28 show the return level plot of the Region 1. The biggest event is represented by an historical data. In particular, it corresponds to the storm that impacts La Rochelle in the year 1999. Notwithstanding this event seems very far from the others, no outliers are detected in this regional extreme data sample.

Regional parameters of the Region 2				
Optimal λ	*Scale param. of regional GPD (season DJF)*	*Shape parameter of regional GPD*	*Regional credible duration (years)*	*Adimensional degree of regional dependence*
1.79	**0.238**	**0.082**	**238.55**	**0.675**

Tab. 5 - Regional parameters computed by the FAB method applied to the Region 2

Results for the Region 2		
Regional quantiles for T=1000ys	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
3.30	**3.99**	**40**

Tab. 6 - Regional return levels computed by the FAB method applied to the Region 2

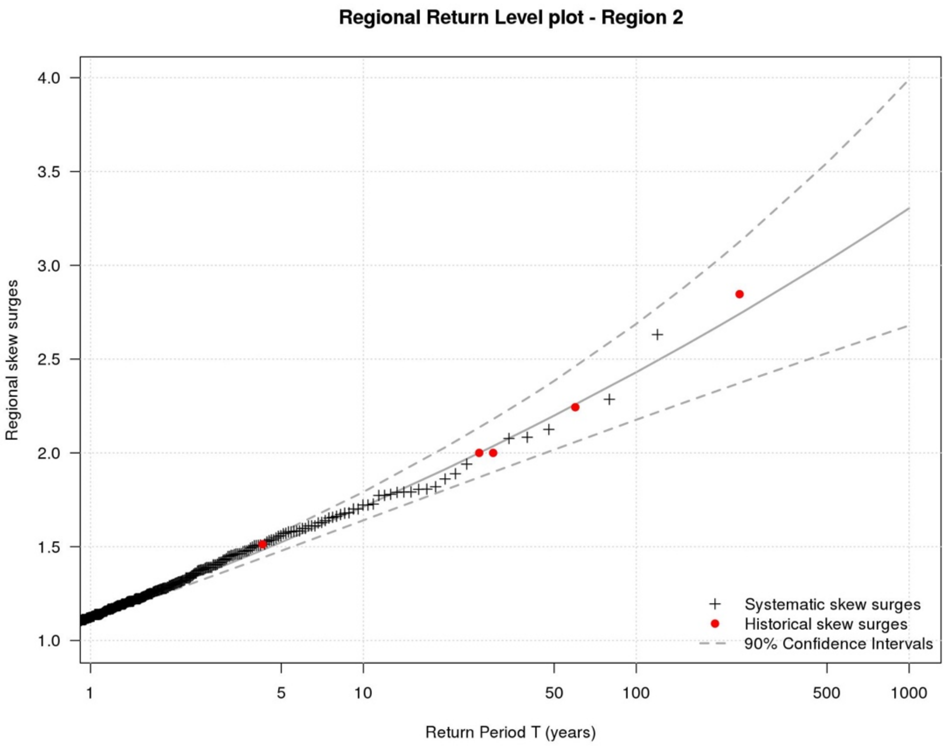


Fig. 29 – Regional return level plot for the Region 2

Fig. 29 displays the return level plot of the Region 2 in which the biggest event corresponds to an historical data. In this case, this historical event is located within the 90% confidence intervals.

The use of historical data in the regional analysis could impact differently the local estimations. In fact, sites of a region preserve their local features by local indexes that assume different value in each site. For this reason, two sites (randomly chosen) belonging to the Region 1 (La Rochelle) and to the Region 2 (Calais) are examined in the same way as the regional estimations. Tab. 7 shows their local return levels associated to a return period of 1000 years.

The relative width of the 90% confidence intervals for the site of La Rochelle is little higher than the Region 1. This means that the uncertainties linked to the return levels associated to a return period of 1000 years are bigger at La Rochelle than in this region.

On the contrary, the site of Calais has the same relative width of the 90% confidence intervals of the Region 2.

Local return level plots for these two sites are illustrated in Fig. 30. Empirical regional extreme data are locally calculated by the local index and illustrated in these return levels.

La Rochelle			**Calais**		
Quantiles for T=1000ys (meters)	*Upper CI 90% for T=1000ys (meters)*	*ΔCI90 / xT=1000ys (%)*	*Quantiles for T=1000ys (meters)*	*Upper CI 90% for T=1000ys (meters)*	*ΔCI90 / xT=1000ys (%)*
1.82	**2.15**	**36**	**1.78**	**2.14**	**40**

Tab. 7 - Results for the sites of La Rochelle (Region 1) and Calais (Region 2)

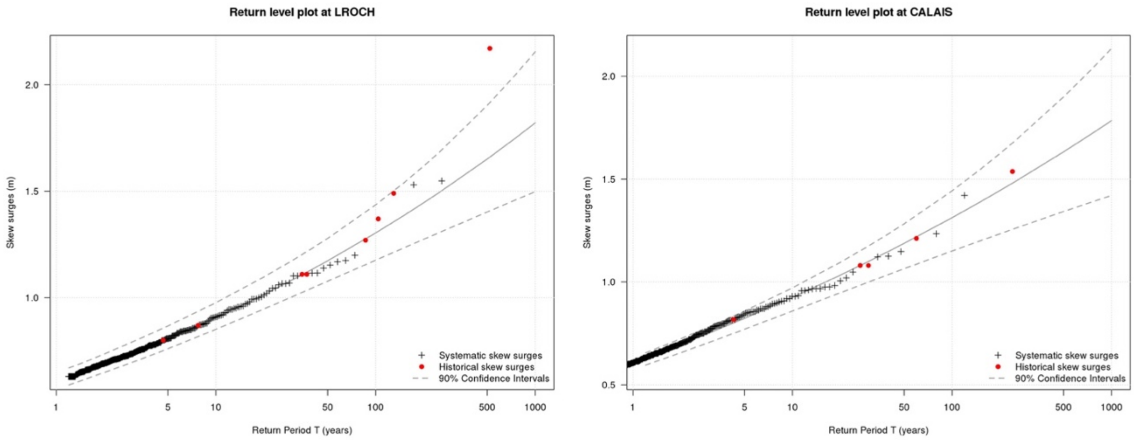


Fig. 30 – Local return levels plots for the sites of La Rochelle (on the left) and Calais (on the right)

4.7 Estimations without historical data

The same frequentist analysis has been performed using the same database without considering historical data in order to realise the influence of the historical data in the regional analysis. A comparison between the results obtained through the FAB method performed for the skew surge database with and without historical data is displayed.

In the case in which a database is composed by only systematic data, the regional duration is considered effective like for the RFA approach.

Using only systematic skew surges, the performance of weighting analysis to compute the optimal value of λ has to be repeated. In fact, the optimal λ can be identified as an another λ value between all the 176 λ-cases when a different database is used to generate the regional extreme data sample.

This frequentist analysis is performed for the Region 1 and the Region 2 as in the previous case. The optimal value of λ is 0.84 for the Region 1 (the same λ value identified through the weighting analysis for a database with historical data) and 1.7 for Region 2.

Frequentist regional estimations are computed for Region 1 and the Region 2 and they are shown in Tab. 8 and Tab. 9.

Regional return levels associated to a return period of 1000 years and their upper bounds of 90% confidence intervals are shown in Tab. 10 and Tab. 11 for the Region 1 and the Region 2. In addition, the relative widths of 90% confidence intervals are computed.

Regional return level plots for the Region 1 and the Region 2 are illustrated in Fig. 31 and Fig. 32.

Regional parameters of the Region 1 (without the use of historical data)				
Optimal λ	*Scale param. of regional GPD (season DJF)*	*Shape parameter of regional GPD*	*Regional effective duration (years)*	*Adimensional degree of regional dependence*
0.84	**0.199**	**0.052**	**514.28**	**0.543**

Tab. 8 - Regional parameters computed by the FAB method applied to the Region 1 without historical data

Regional parameters of the Region 2 (without the use of historical data)				
Optimal λ	*Scale param. of regional GPD (season DJF)*	*Shape parameter of regional GPD*	*Regional effective duration (years)*	*Adimensional degree of regional dependence*
1.7	**0.243**	**0.036**	**235.88**	**0.677**

Tab. 9 - Regional parameters computed by the FAB method applied to the Region 2 without historical data

Results of Tab. 8 and Tab. 9 are compared with these obtained by the use of historical data in the regional analysis in Tab. 3 and Tab. 5.

The application of FAB method with historical skew surges enables the estimation of a higher regional GPD shape parameter and of a bit higher regional duration for the Region 1. Other values figured out in Tab. 8 and in Tab. 3 are very close.

Also for the Region 2, the regional GPD shape parameter and the regional duration decrease without the use of historical data.

Results for these regions point out a common rise of the GPD scale parameter using historical data. In addition, a little growth of the period of observed regional extreme events has to be identified when historical data are used.

Results for the Region 1 (without the use of historical data)		
Regional quantiles for T=1000ys	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
2.55	**2.90**	**28**

Tab. 10 - Regional return levels computed by the FAB method applied to the Region 1 without historical data

Results for the Region 2 (without the use of historical data)		
Regional quantiles for T=1000ys	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
2.94	**3.45**	**34**

Tab. 11 - Regional return levels computed by the FAB method applied to the Region 2 without historical data

Results of Tab. 10 and Tab. 11 are compared with these obtained in Tab. 4 and Tab. 6.

For both the regions, the introduction of the extraordinary historical events tends to increase the value of return levels. In addition, the relative width of the confidence intervals is a bit lower (5-6 %) when historical data are not used. This trend is due to the exceptionality of the extreme past events. In fact, the higher skew surges of the both regional extreme data samples are historical data. In any case, this value is acceptable for a statistical analysis.

Moreover, the use of historical data has increased the regional durations of the both regional extreme data samples. In particular, historical data allows the detection of three additional storms in the Region 1 and of five additional storms in the Region 2. In addition, other historical data are the biggest normalised data of five storms detected in the Region 1.

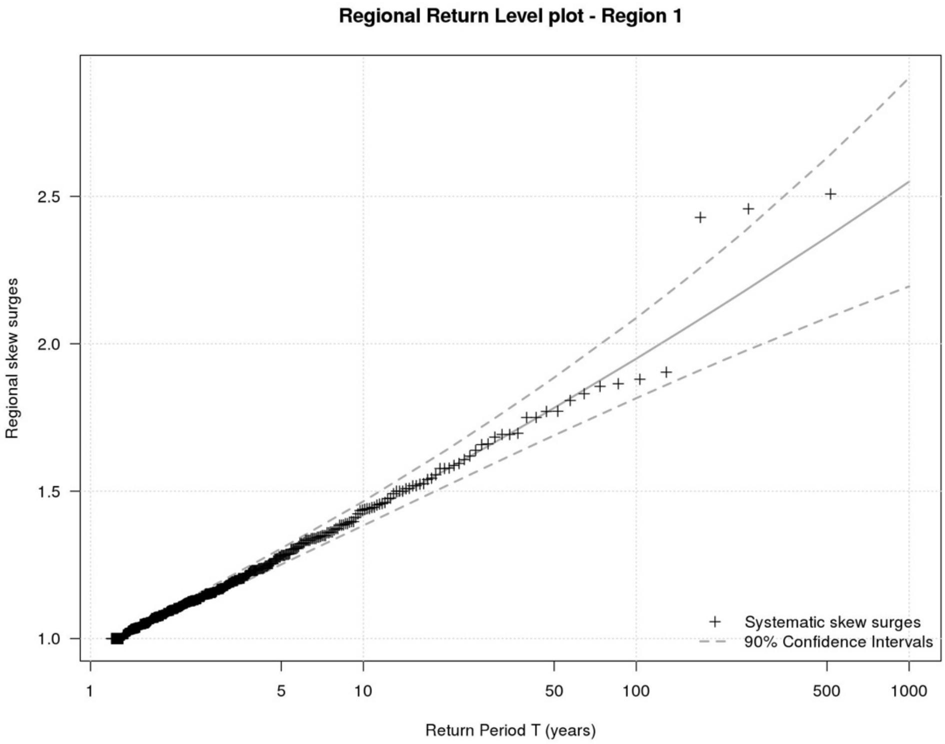


Fig. 31 – Regional return level plot for the Region 1 without the use of historical data

Fig. 31 shows graphically the decrease of high return levels when historical data are not considered in the statistical analysis.

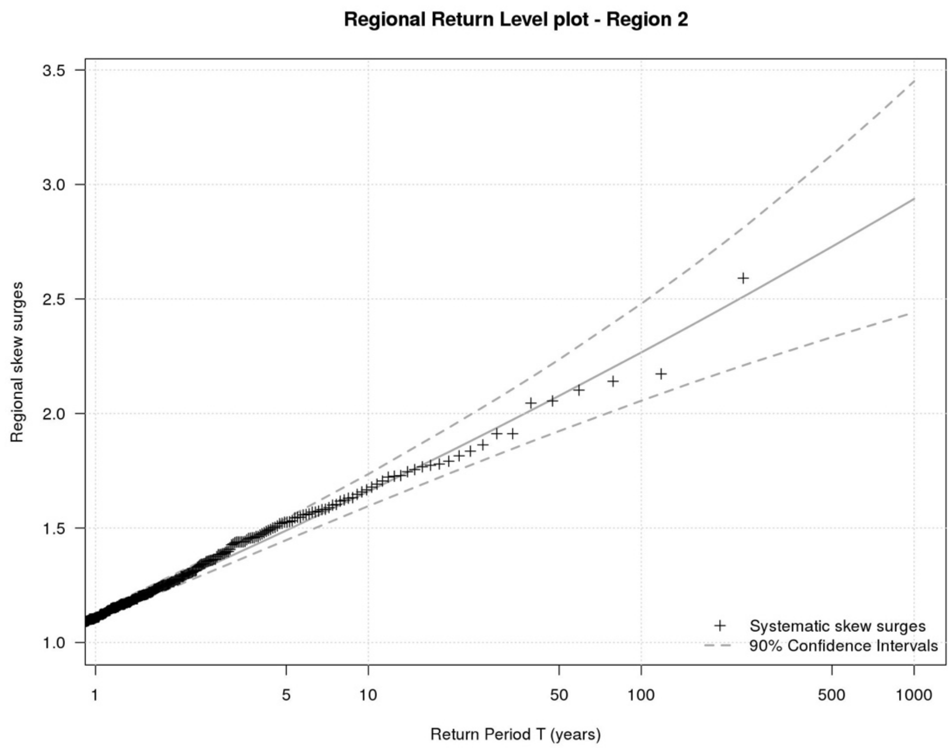


Fig. 32 – Regional return level plot for the Region 2 without the use of historical data

Fig. 32 illustrates the return level plot of the Region 2 without the five historical data used in the previous analysis.

However, regional return levels are normalised values and, for this reason, it is appropriate to compare also local return levels in order to provide global conclusions on this analysis of comparison.

Results at La Rochelle and at Calais are shown in Tab. 12 and Fig. 33 and they are compared respectively with Tab. 7 and Fig. 30. Same considerations on return levels can be pointed out for the local cases. In fact, return levels linked to a T of 1000 years for the two sites of La Rochelle and Calais are lower than these obtained with historical data. The difference for these estimations is of 21 centimetres for the site of La Rochelle and of 16 centimetres for the site of Calais.

La Rochelle			Calais		
Quantiles for T=1000ys (meters)	*Upper CI 90% for T=1000ys (meters)*	*ΔCI90 / xT=1000ys (%)*	*Quantiles for T=1000ys (meters)*	*Upper CI 90% for T=1000ys (meters)*	*ΔCI90 / xT=1000ys (%)*
1.61	**1.84**	**29**	**1.62**	**1.90**	**35**

Tab. 12 - Results for the sites of La Rochelle (Region 1) and Calais (Region 2) without historical data

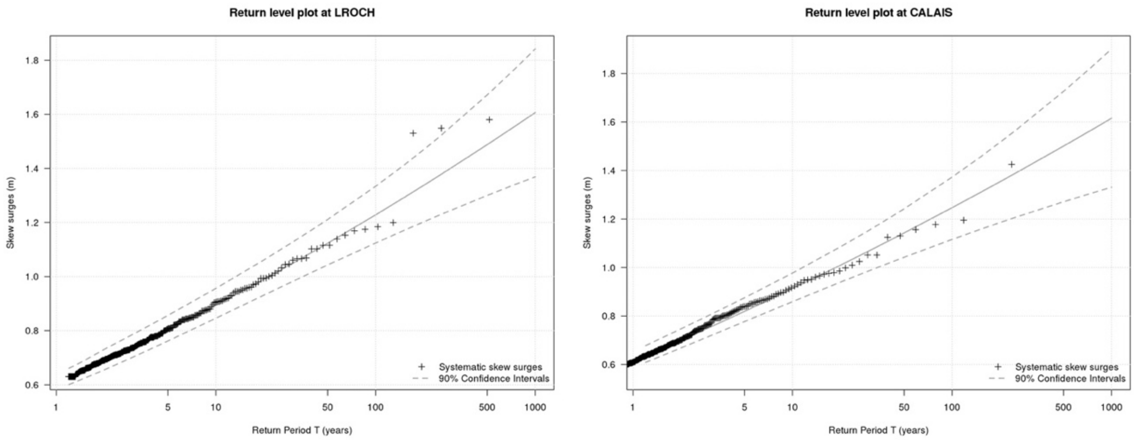


Fig. 33 – Local return level plots for the sites of La Rochelle (on the left) and Calais (on the right) without the use of historical data

In conclusion, for this particular application, historical data does not reduce the uncertainties on estimations of regional and local extreme events. This reduction could be obtained with the use of a big number of historical events.

On the contrary, the consideration of the extraordinary historical events has permitted the extension of the period of observation of the regional extreme data sample and the computation of more reliable extreme events. In particular, additional exceptional storms are detected and they have been used to perform this statistical analysis.

4.8 Bayesian estimations of return levels

Bayesian estimations of regional and local return levels associated to several return periods T are computed through the application of the FAB method. In particular, Region 1 and Region 2 are analysed in the following. Predictive return levels y^{*}_{R}, standard estimative return levels y_R and their Credibility Intervals are estimated and figured out in a same return level plot.

In any case, the Bayesian statistical framework requires the definition of a posterior distribution of the vector of parameters of the statistical distribution considered in order to estimate return levels.

In this analysis, the posterior distribution of the regional GPD parameters' vector θ is computed through the generation of three chains of one million of iterations U. The convergence of the MCMC process for these vectors θ must be achieved in every chain and between all the chains. The Gelman and Rubin test (Gelman and Rubin, 1992) is performed to test this convergence. This test allows the computation of a degree of convergence called Potential Scale Reduction Factor. A value close to 1 for the PSRF has to be computed in order to verify the convergence (it is suggested to get at least a PSFR smaller than 1.05).

The PSRF is estimate for a total of one million of iterations. Fig. 34 shows how this factor assumes a value very close to 1 for each regional GPD parameter of the Region 1 and Region 2.

In addition, the convergence is also verified by the test of Brooks (Brooks and Gelman, 1998). This test is an implementation of Gelman and Rubin test and allows the estimation of another degree of convergence called Multivariate PSFR. This factor considers the likely variability that the sample could have.

The MPSFR must assume lower values than 1.2 to verify the convergence of MCMC iterations. MPSFR is considerably smaller than 1.2 for the Region 1 and the Region 2.

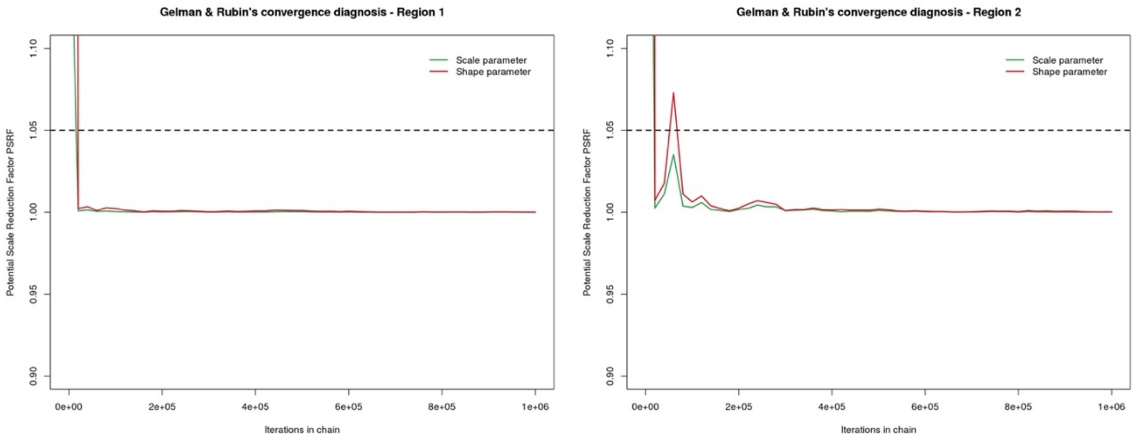


Fig. 34 – Progressive PSFR values of tests of Gelman and Rubin for Region 1 and Region 2

Values of the scale and the shape parameters obtained for every MCMC iteration and for each chain is shown in Fig. 35 and Fig. 36.

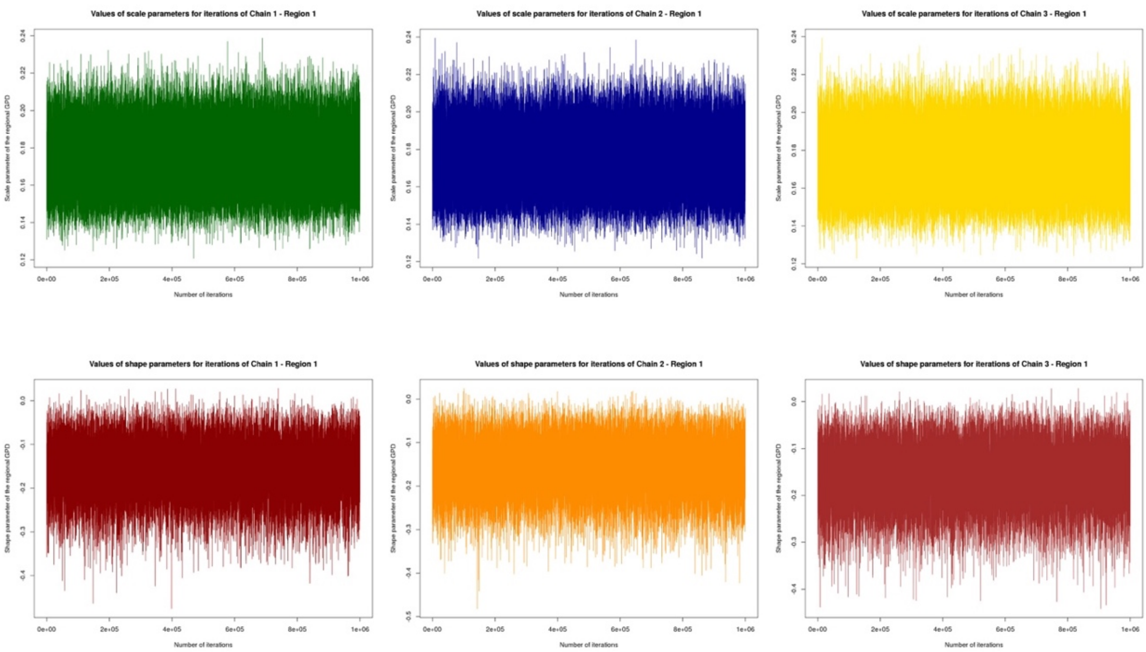


Fig. 35 – Values of regional GPD scale and shape parameters of each iteration for the 3 chains of Region 1

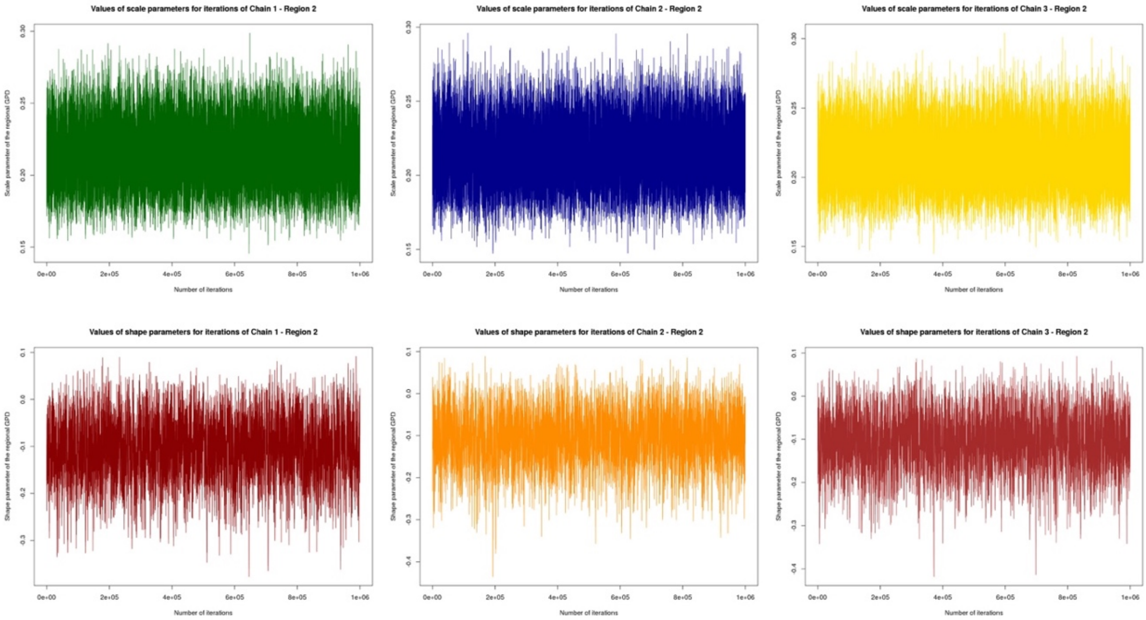


Fig. 36 - Values of regional GPD scale and shape parameters of each iteration for the 3 chains of Region 2

The verification of the convergence of the MCMC process enables the use of the posterior distribution of the regional GPD parameters' vector θ in the Bayesian estimation process of return levels.

The Bayesian estimation of regional return levels can be now performed for the Region 1 and the Region 2. The results of this statistical analysis are are shown in Tab. 13 and Tab. 14. In particular, the predictive return levels are computed after a burn-in of $U'=10^5$ iterations.

Parameters and results for the Region 1						
Optimal λ	*Scale param. of estimative regional GPD*	*Shape param. of estimative regional GPD*	*Regional estimative quantiles for T=1000ys*	*Regional predictive for T=1000ys*	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
0.84	**0.175**	**0.144**	**2.99**	**3.14**	**3.91**	**44**

Tab. 13 - Useful parameters and results computed by the Bayesian analysis on the Region 1

Parameters and results for the Region 2						
Optimal λ	*Scale param. of estimative regional GPD*	*Shape param. of estimative regional GPD*	*Regional estimative quantiles for T=1000ys*	*Regional predictive for T=1000ys*	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
1.79	**0.216**	**0.089**	**3.30**	**3.51**	**4.56**	**53**

Tab. 14 - Useful parameters and results computed by the Bayesian analysis on the Region 2

Tab. 13 and Tab. 14 display the regional results of the Bayesian analysis performed through the FAB method. Regional factors as the regional credible duration or the adimensional degree of regional dependence ***Φ*** are not figured because they are identic to that shown in Tab. 3 and Tab. 5. In fact, the Bayesian framework of estimation does not modify the computation of the main elements of the regional analysis. On the contrary, estimations of the statistical distribution's parameters and return levels can assume different values.

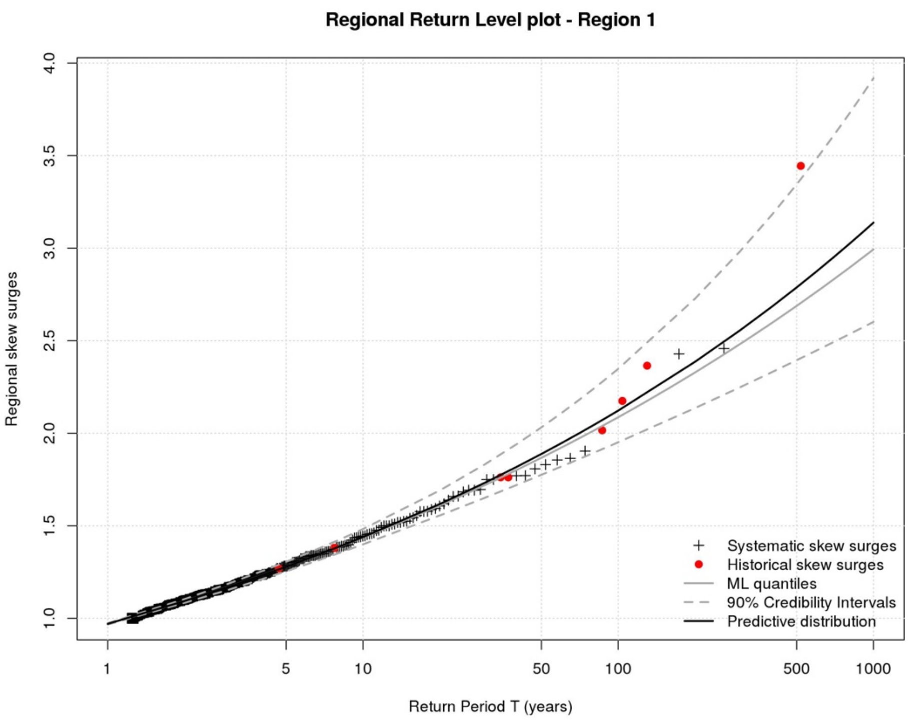


Fig. 37 – Regional return level plot of Bayesian estimations for Region 1

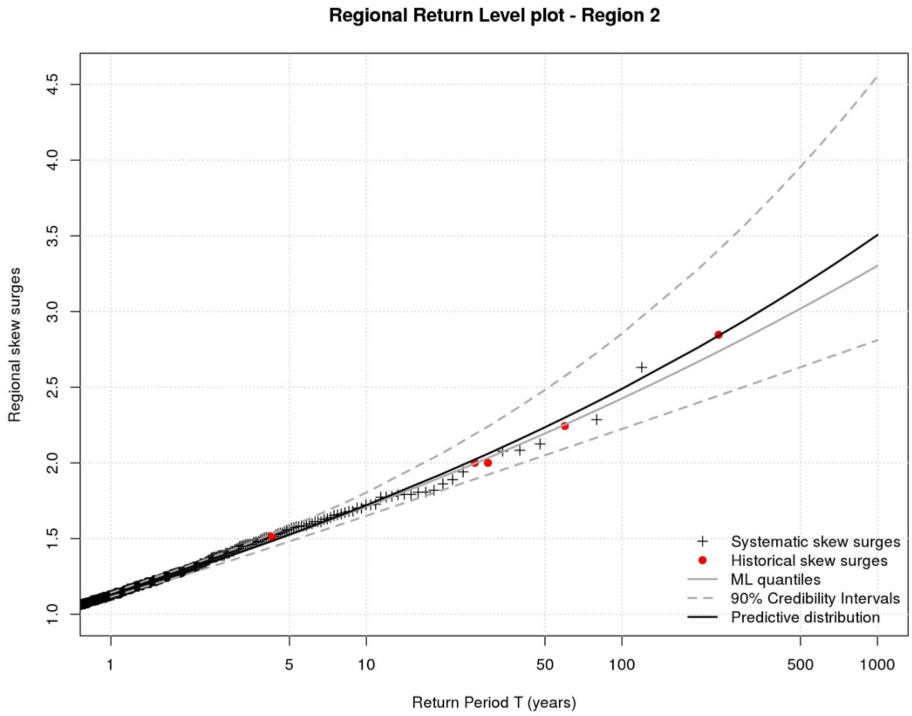


Fig. 38 - Regional return level plot of Bayesian estimations for Region 2

Fig. 37 and Fig. 38 display return level plots of Region 1 and Region 2. They highlight how the predictive return levels (in black) linked to a generic return period *T* are bigger than the estimative return levels (in grey) associated to the same generic return period *T*. In addition, Credibility Intervals are illustrated by the grey dotted lines. In these figures, historical data are represented as red dots in order to point out their relevance compared with the systematic events. All the empirical events are located in the return level plot by a Weibull position (Weibull, 1951).

However, these results can be locally evaluated in order to know the effective impact of these estimations. Local return levels are then computed from regional return levels by the use of local indexes. Same sites of the frequentist analysis are considered: La Rochelle for the Region 1 and Calais for the Region 2.

La Rochelle			Calais		
Estimative quantiles for T=1000ys (meters)	*Predictive quantiles for T=1000ys (meters)*	*Upper CI 90% for T=1000ys (meters)*	*Estimative quantiles for T=1000ys (meters)*	*Predictive quantiles for T=1000ys (meters)*	*Upper CI 90% for T=1000ys (meters)*
1.88	**1.97**	**2.47**	**1.78**	**1.89**	**2.46**

Tab. 15 – Bayesian estimations for the sites of La Rochelle (Region 1) and Calais (Region 2)

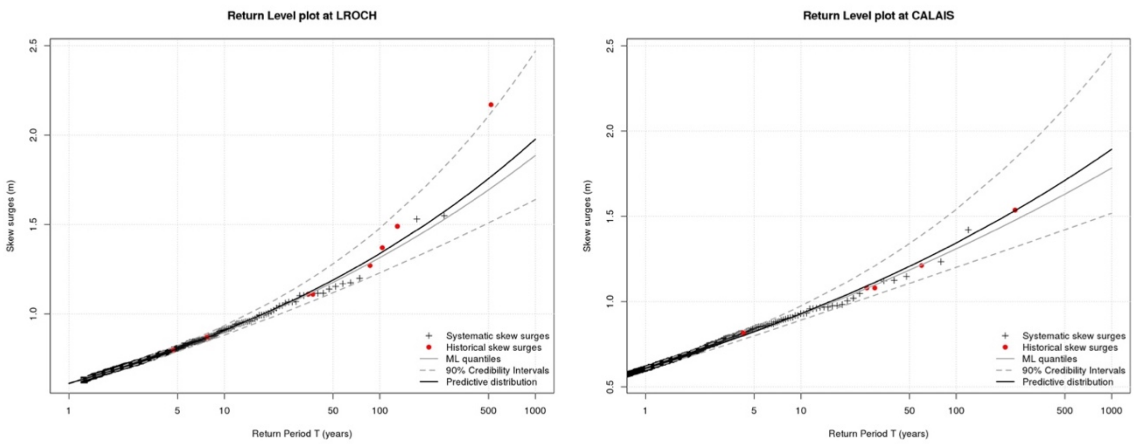


Fig. 39 - Return level plots of Bayesian estimations for La Rochelle (site belonging to Region 1) and Calais (site belonging to Region 2)

Same conclusions supplied for the regional Bayesian estimations can be provided for the local Bayesian estimations. Predictive return levels linked to a particular return period T are higher than standard estimative return levels estimated for the same particular return period.

Further clarifications have to be provided for the computation of the relative width of the Credibility Intervals. Regional and local CI have the same relative width in this analysis. In fact, the computation of regional and local CI considers in a Bayesian application of the FAB method only the variation of the return levels and not of the local indexes (as performed in frequentist by the bootstrap methodology).

4.9 Frequentist and Bayesian estimations

The comparison of the results of the frequentist and Bayesian statistical frameworks is not totally correct due to two different concepts of probability. Nevertheless, a preliminary comparison is performed in order to realise the differences obtained computing frequentist and Bayesian estimations with the FAB method.

In the Bayesian context, FAB method is performed considering the seasonal effects of the skew surges in the estimation process of the regional GPD parameters. For this reason, a particular case of frequentist estimations that not consider the seasonality of the skew surges has to be achieved. This case enables the preliminary comparison of these two statistical frameworks of estimation.

The value of optimal λ is the same for Region 1 and for Region 2 (respectively 0.84 and 1.79) for this frequentist case. Results and regional return levels are displayed in Tab. 16 and Fig. 40.

The comparison between the 90% Confidence Intervals of Tab. 16 with the previous 90% Credibility Intervals computed in the previous Bayesian analysis has to be carefully performed. In fact, in the frequentist framework the Confidence Intervals are calculated through a bootstrap method while in the Bayesian framework the Credibility Intervals are computed by the definition of the posterior distribution.

Frequentist estimations for the Region 1 and Region 2					
Region 1			**Region 2**		
Optimal λ	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*	*Optimal* λ	*Regional upper CI 90% for T=1000ys*	*ΔCI90 / yR,T=1000ys (%)*
0.84	**3.59**	**37**	**1.79**	**4.01**	**40**

Tab. 16 - Useful parameters and results computed by the Bayesian analysis on the Region 1 without seasonality

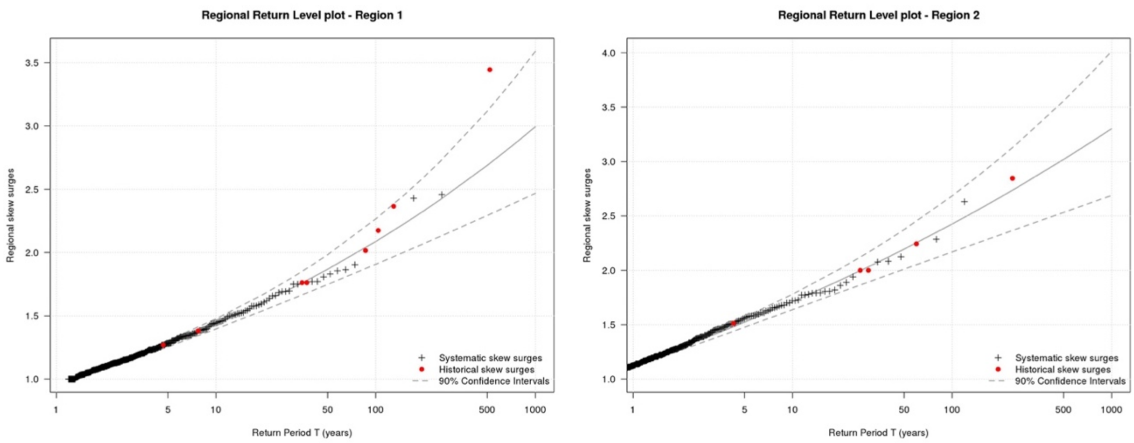


Fig. 40 – Regional return level plot of frequentist estimations for Region1 and Region 2 without seasonality

Comparing regional results obtained in Tab. 16 and those illustrated in Tab. 13 and Tab. 14, Credibility Intervals tend to be more unbounded than Confidence Intervals in this application. This is pointed out by the relative widths of the Credibility Intervals that assume higher values of percentage for a return period *T* of 1000 years compared with Confidence Intervals linked to the same return period *T*.

Same considerations can be provided for the local estimations of Tab. 17 and Fig. 41 compared with Bayesian local estimations of Tab. 15 and Fig. 39.

La Rochelle		**Calais**	
Upper CI 90% for T=1000ys (meters)	*ΔCI90 / yR,T=1000ys (%)*	*Upper CI 90% for T=1000ys (meters)*	*ΔCI90 / yR,T=1000ys (%)*
2.28	**39**	**2.15**	**40**

Tab. 17 - Frequentist estimations for the sites of La Rochelle (Region 1) and Calais (Region 2) without seasonality

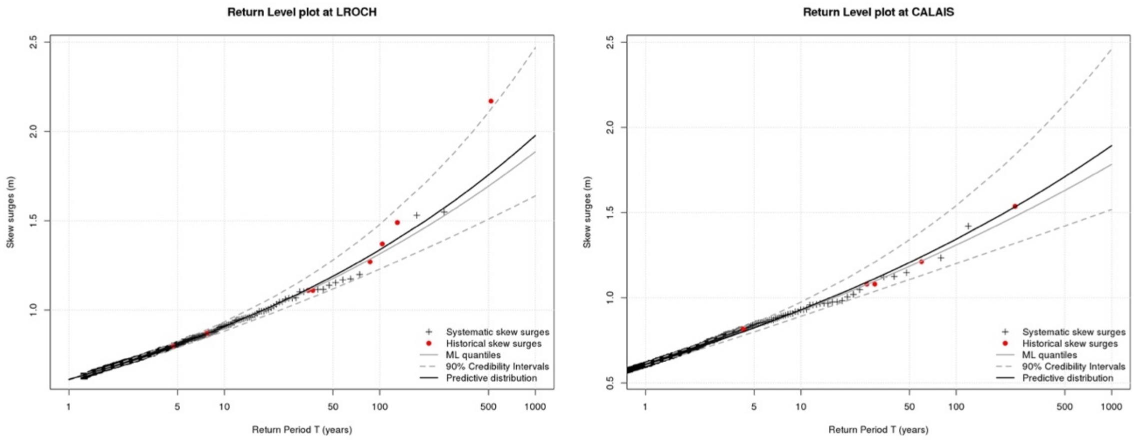


Fig. 41 - Return level plots of frequentist estimations for La Rochelle (site belonging to Region 1) and Calais (site belonging to Region 2) without seasonality

As for Bayesian estimations, results and local return levels are figured out for the site of La Rochelle (Region 1) and for the site of Calais (Region 2).

Significant differences are identified between the upper 90% Confidence Interval for a return period T of 1000 years computed by this particular frequentist analysis (Tab. 17) and the upper 90% Credibility Interval for the same return period *T* (Tab. 15). For the site of La Rochelle, this difference is of 18 centimetres and for Calais is of 30 centimetres.

Finally, for this case of study, Bayesian estimations obtained by the application of the FAB method provide higher uncertainties linked to very high return period *T* than the frequentist estimations linked to the same return period *T*. On the other hand, the use of the Bayesian framework allows the utilisation of a prior information that the expert could know and the computation of the predictive distribution of the extreme data sample.

All frequentist and Bayesian estimations of this study obtained with and without the use of historical data are synthetized in the summary table displayed in the Annexe E.

Chapter 5

CONCLUSIONS

La méthode FAB est une approche statistique qui permet l'utilisation des données historiques dans un contexte régional d'estimation des événements maritimes extrêmes. Cette méthode vise à répondre à plusieurs questions sur l'utilisation des données historiques en régionale. En particulier, les questions liées à la période d'observation de ces événements ainsi que les différents types de données disponibles sont traitées à travers le nouveau concept de durée crédible et une nouvelle définition de la fonction de vraisemblance régionale.

Une application de la méthode FAB à une base des données des surcotes de pleine mer avec et sans données historiques est réalisé pour comprendre l'importance des événements du passé. En fait, pour les deux régions analysées, les événements historiques sont les données extrêmes les plus grandes dans l'échantillon régional. Le petit nombre de données historiques disponibles ne permet pas d'avoir des améliorations en termes d'incertitudes liées aux niveaux de retour.

Plusieurs améliorations peuvent être apportées à cette méthode comme, par exemple, l'introduction des seuils de perception pour les sites dans lequel l'exhaustivité pourrait être vérifiée dans le futur. Par ailleurs, la méthode de formation des régions pourra être améliorée avec la considération d'autres variables ou ré-analyses. L'analyse régionale centrée est considérée aussi comme une autre piste d'amélioration.

5.1 Conclusions

Adequate coastal protections are needed to protect all infrastructures located on coastlines from floods. The design of these protections requires the statistical estimations of extreme coastal events and, in particular, of extreme sea levels and skew surges. Nevertheless, huge uncertainties are frequently associated to the estimation of extreme events when local analyses are performed. This is typically due to the short time series of sea levels measured by tide gauges. In order to improve extreme estimations, two approaches are often employed in literature: the regional analysis and the use of historical data.

The use of historical data enables the extensions of local extreme data samples considering exceptional events of the past. The regional analysis allows the extension of extreme data samples considering extreme events occurred in a defined region. A regional analysis approach, called RFA, has previously been developed for coastal events. The combination of these two different approaches can further improve the estimations of extreme coastal events.

The FAB method aims to improve the reliability of statistical estimations of extreme coastal events using regional analysis and historical events together. In particular, the combination of these two approaches enables the performance of a statistical analysis on a larger regional extreme data sample also composed of the exceptional past events available.

Several issues are linked to the discovery and the use of historical data. In particular, historical data are not easily available and therefore their collection is complex. For this reason, Chapter 2 exposes all the issues linked to the discovery of a historical event. Historical information is available in several sources that are sometimes difficult to access. When information on a historical event is available, this has to be translated to numerical data in order to be used in a statistical process. In fact, historical information does not always contain the exact historical data that represents the event considered. For this reason, three most common types of historical data that depend on the historical information available are defined: exact data, data range and lower limit value of data.

Issues linked to the discovery of historical information lead to the definition of a simplified collection and validation approach. This approach recommends to firstly analyse all the already existing historical databases in order to detect the most known historical events that happened in the past. The next step is to research all the sources accessible in order to collect the highest number of historical data. All of these historical data collected must be successively validated.

A preliminary validation can be performed by meteorological reanalyses but the final validation is achieved through specific knowledge from a historian on the period of the event. For this reason, a historical event can be validated by a historian. A functional example of this approach is performed for skew surges in the end of Chapter 2.

When historical data are validated, they have to be used in the statistical analysis in order to get reliable estimations of extreme events. Statistical analyses for POT data require the knowledge of the period of observation of the considered extreme events' sample. Historical data are typically discovered as isolated exceptional events and, for this reason, no period of observation can be associated to these types of data. Regional analyses and, in particular, the RFA approach need the knowledge of the period of observation of the whole regional extreme data sample in order to estimate regional and local return levels. When systematic and historical events are available, the observation period of the regional extreme data sample has to be estimated. Based on the main elements of the RFA approach, the FAB method proposes the use of the credible hypothesis in which no trends on storm frequency exist during the 20th century in the period of data observation. This hypothesis enables the definition of the new concepts of local and regional credible durations. These durations represent the period of observation of a specific extreme data sample.

In any case, different types of historical data can be available. For this reason, a particular penalised likelihood has to be defined in order to estimate appropriate parameters of the regional statistical distribution. Different likelihood functions are formulated for systematic and each of the three types of historical data. The FAB method enables the estimation of the regional GPD parameters through the maximisation of the whole penalised likelihood composed of systematic and historical likelihood functions.

The composition of a regional sample of independent extreme events is performed by the definition of local sampling thresholds in each site of a particular region. In Chapter 3, an approach to define an optimal sampling threshold in each site of a particular region is proposed. This approach is used in the FAB method and it is based on a weighting process. In particular, primary and secondary parameters have to be chosen in order to perform this analysis. Primary parameters are statistical tests necessary to perform the considered statistical analysis. Secondary parameters are elements considered by the expert as important for the statistical process. The weighting analysis is achieved for several local thresholds and is performed only on secondary parameters for threshold cases in which primary parameters are verified. This approach allows the definition of optimal local sampling thresholds through the detection of the optimal

λ value (number of occurred local storms per year). Finally, the selection of the optimal threshold is important to reach good performance of the FAB method in computing estimations of extreme events.

In addition, the FAB method allows the calculation of return levels in the frequentist and Bayesian statistical frameworks. Several experts consider the Bayesian context as the most adequate to statistically treat historical data. For this reason, the FAB method permits the choice of the statistical framework of estimation. Frequentist or Bayesian return levels can both be estimated by the application of the FAB method.

An application of the FAB method to a skew surge database is performed in Chapter 4. The skew surge database is composed of time series of systematic skew surges gauged in 74 sites of the Atlantic Ocean, English Channel, Irish Sea and North Sea and of 14 historical skew surges available in three of the 74 sites. Physical homogeneous regions are created by the use of the same approach used in the RFA method. These regions are based on the definition of storm clusters. Storm clusters are created by the use of three parameters (p, Δ, η). Data exceeding a physical threshold defined by the local p-value that occur in neighbour sites η during a temporal interval Δ are generated by the same storm. These parameters are calibrated by a sensitivity analysis recommended for time series linked to a different gauging period. The next steps of the FAB methodology are performed only for the regions in which historical data are available (Region 1 and Region 2). For these regions, an optimal λ value is computed and physical regions are statistically verified. Frequentist and Bayesian return levels are then estimated for Region 1 and Region 2.

In addition, FAB method is performed for the same database composed only of systematic data. Results indicate that the use of exceptional historical data can enrich the regional extreme data sample. In fact, the use of additional extreme events that occurred in the past increases the length of the extreme data sample. However, relative widths of 90% Confidence Intervals for return levels associated to $T=1000$ years do not show improvements on the uncertainties linked to the estimations for these two regions. This is probably due to the reduced number of available historical data. For this reason, the formation of a database of historical events can further improve estimations of extreme events.

Finally, a preliminary comparison between frequentist and Bayesian estimations of return levels is performed. In this analysis, frequentist return levels are associated to lower uncertainties than return levels computed by the Bayesian statistical analysis. This could be due to the two different methods of computation of the Confidence and Credibility Intervals. In any case, the Bayesian framework allows the computation of predictive return levels and the introduction of priori information in the estimation process.

Lastly, the FAB method is a regional statistical approach that enables the estimations of extreme coastal events using both systematic and historical data.

5.2 Future works and perspectives

The FAB method developed in this study allows the estimation of the period of observation of the historical data in a particular location and, then, in the region. In fact, information on the period of observation is frequently unavailable for coastal events of the past.

In any case, wide investigation of past events occurring in a particular site could provide some supplementary information that allows the knowledge of their period of observation above a perception threshold. For this reason, future studies on the occurrence of the most exceptional storms of the past are required.

Currently, the FAB method cannot take into account this particular information that could improve estimations of extreme events extending local and regional durations. For this reason, the development of the FAB method v2.0 is planned in the near future.

The main idea of this upgrade is to compute an effective duration for sites in which historical data with perception thresholds are available and a credible duration for sites in which no additional information of the historical period is available. In particular, a new definition of the regional credible duration and a new formulation of the penalised maximum likelihood are required when the concept of perception threshold is introduced in a regional analysis. Upgrading the FAB method will allow the use of any type of information on the period of observation of a particular event of the past which will further improve the estimation of extreme events.

This upgrade can obviously be used in cases in which the exhaustiveness of a historical event above a perception threshold is widely guaranteed by an expert of the period. This exhaustiveness can be assured only in some areas in which satisfactory documentations are available. For this reason, the concept of the credible duration still has an important role in the FAB method.

In fact, it enables the estimation of the observation's period of the historical events in sites in which the collection of supplementary information is challenging.

The collection and the validation of a larger number of historical data is another way to improve estimations of extreme coastal events using the FAB methodology. In fact, the extension of the historical part of the database would allow the use of more exceptional events in the regional analysis. In particular, more storms are detected in the region and, then, a larger regional extreme data sample observed in a higher regional period is created. The use of a regional data sample composed by a higher number of extraordinary events could improve the statistical estimations of return levels.

Furthermore, in the FAB method, the regions are based on the definition of storm clusters formed by the use of three parameters. These parameters would assume different values for each storm. In fact, they depend on physical spatio-temporal processes that are singular for every exceptional event. For this reason, the clustering method used in this study could, in particular cases, produce two storm clusters for a single storm or a single storm cluster for two different storms. Different solutions can be performed to create more precise storm clusters and, then, more accurate physical regions:

- the use of additional physical variables such as wind speeds or sea level-pressures;
- the use of meteorological reanalyses of wind speeds and sea level-pressures;
- the definition of different parameters to detect storm clusters that depend on a particular season considered.

In addition, regional extreme data samples could also be formed through the use of the spatial extremogram. In this case, the region is formed around a defined target-site (Hamdi et al., 2016). This specific statistical tool enables the computation of the extremal dependence value between a target-site and its neighbour sites. Consequently, the region can be defined as the influence area of the target-site.

Finally, the FAB method can only be applied for a single coastal variable. Nevertheless, more than one variable can provoke extreme meteorological conditions. For this reason, the simultaneous consideration of two or more variables could provide a better characterisation of the extreme events. In particular, a bivariate or multivariate statistical analysis can be performed by the definition of a model of dependence between the different variables considered.

Annexe A

HD IN REGIONAL ANALYSIS

1.1 Historical data in regional analysis (2018)

Nat. Hazards Earth Syst. Sci., 18, 949–962, 2018
https://doi.org/10.5194/nhess-18-949-2018

Natural Hazards and Earth System Sciences
Open Access
EGU

The use of historical information for regional frequency analysis of extreme skew surge

Roberto Frau[1,2], **Marc Andreewsky**[1], **and Pietro Bernardara**[3]

[1]EDF R&D Laboratoire National d'Hydraulique et Environnement (LNHE), Chatou, 78401, France
[2]Université Paris-Est, Saint-Venant Hydraulics Laboratory (ENPC, EDF R&D, CEREMA), Chatou, 78401, France
[3]CEREA, Joint Laboratory École des Ponts ParisTech – EDF R&D, Université Paris-Est, Marne-la-Vallée, 77455, France

Correspondence: Roberto Frau (roberto.frau@edf.fr)

Received: 25 October 2017 – Discussion started: 2 November 2017
Revised: 8 February 2018 – Accepted: 27 February 2018 – Published: 26 March 2018

Abstract. The design of effective coastal protections requires an adequate estimation of the annual occurrence probability of rare events associated with a return period up to 10^3 years. Regional frequency analysis (RFA) has been proven to be an applicable way to estimate extreme events by sorting regional data into large and spatially distributed datasets. Nowadays, historical data are available to provide new insight on past event estimation. The utilisation of historical information would increase the precision and the reliability of regional extreme's quantile estimation. However, historical data are from significant extreme events that are not recorded by tide gauge. They usually look like isolated data and they are different from continuous data from systematic measurements of tide gauges. This makes the definition of the duration of our observations period complicated. However, the duration of the observation period is crucial for the frequency estimation of extreme occurrences. For this reason, we introduced here the concept of "credible duration". The proposed RFA method (hereinafter referenced as FAB, from the name of the authors) allows the use of historical data together with systematic data, which is a result of the use of the credible duration concept.

1 Introduction

Flood risk is a noteworthy threat for anthropic and industrial settlement along the coast. Coastal defence constructions have to withstand extreme events. Therefore, in this context, a statistical estimation of extreme marine variables (such as sea levels or skew surges, the difference between the maximum observed sea level and the maximum predicted astronomical tide level during a tidal cycle; Simon, 2007; Weiss, 2014) occurrence probability is necessary. These variables associated with high return periods (up to 1000 years; ASN, 2013) are estimated by statistical methods linked to the extreme value theory (EVT) (Fréchet, 1928; Gnedenko, 1943; Gumbel, 1958; Picklands, 1975; Coles, 2001).

In the past, time series recorded at a single site had been used to estimate return levels for marine hazards. However, these extreme quantiles are linked to huge estimation uncertainties due to local recordings being too short (typically 30 to 50 years long). A greater amount of data is therefore necessary to obtain reliable estimations of extreme marine hazards.

For this reason and thanks to the availability of several local datasets, the regional frequency analysis (RFA) is employed to reduce statistical uncertainties, allowing the use of bigger extreme samples. The primary purpose of RFA is to cluster different locations in a region and to use all the data together (Cunnane, 1988; Hosking and Wallis, 1997; Van Gelder and Neykov, 1998; Bernardara et al., 2011; Weiss, 2014). RFA is based on the index flood principle (Dalrymple, 1960) that introduces a local index for each site in order to preserve their peculiarities in a region. A regional distribution can be defined only after checking its statistical homogeneity (Hosking and Wallis, 1997). The probability distribution of extreme values must be the same everywhere in the region to allow the fitting of a probability distribution with a big data sample. Several applications exist for marine variables, although these are mostly for skew surges (Bernardara et al., 2011; Bardet et al., 2011; Weiss et al., 2013).

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

Specifically, Weiss (2014) proposed a RFA approach to treat extreme waves and skew surges. According to this approach, homogeneous regions are first built relying on typical storm footprints and their statistical homogeneity is subsequently tested. The addition of physical elements allows for the inclusion of to regions that are both physically and statistically homogeneous. Furthermore, this methodology, and the pooling method in particular, yields an equivalent regional duration in years that enables the evaluation of the return periods of extreme events.

As proved in hydrology, extending datasets with historical data or paleofloods is an alternative way to reduce uncertainties in the estimation of extremes (Benson, 1950; Condie and Lee, 1982; Cohn, 1984; Stedinger and Cohn, 1986; Hosking and Wallis, 1986a, 1986b; Cohn and Stedinger, 1987; Stedinger and Baker, 1987; Ouarda et al., 1998; Benito et al., 2004; Reis and Stedinger, 2005; Payrastre et al., 2011; Sabourin and Renard, 2015) as well as representing better potential outliers. Previous works on local sea levels (Val Gelder, 1996; Bulteau et al., 2015) and on local skew surges (Hamdi et al., 2015) pointed out that the incorporation of historical data leads to a positive effect on extreme frequency analysis.

The combination of these two approaches increases the amount of available data allowing to get better extreme return levels. The use of extraordinary floods events in RFA has already been explored by Nguyen et al. (2014) for peak discharges and Hamdi et al. (2016) have explored it in the frame of oceanic and meteorological variables.

In this study, we investigate the use of historical data in RFA of systematic measurements of skew surge levels. The proposed FAB method (a declination of RFA using credible duration) allows the estimation of reliable extreme skew surge events by taking into account all the available non-systematic records of skew surge levels. Specifically, historical data are defined as all of the skew surge events that were not recorded by a tide gauge, and consequently, systematic data are all the measurements of skew surge recorded by tide gauge. One obvious and very promising way of collecting information not recorded by a tide gauge is to investigate and reconstruct historical events that occurred before the starting date of gauge observations. This activity is also known as data archaeology and has been demonstrated very promising and useful for several study in the literature. However, this definition also includes isolated data points reconstructed during gaps in the systematic measurements. This is because statistically speaking they could be treated similarly to the isolated data reconstructed from before the gauged period. To find historical events it is thus necessary to go back in time and look for information before the gauging period or during each data gap in a time series. During extreme storms, tide gauge may be affected by a partial or total failure in the sea level measurements and so this kind of event cannot always be detected. Thanks to non-systematic records, we can hold a numerical value to these ungauged extreme events. This allows to use a larger number of storms in the statistical assessment of the extreme skew surges.

Baart et al. (2011) pointed out the hitch in recovering accurate historical information. A historical event is the result of a wide investigation of several qualitative and quantitative sources. Historical data are collected from different sources as newspapers of the period, archives, water marks, palaeographic studies etc. For this reason, each historical data has its own origin and its quality has to be checked (Barriendos et al., 2003). Therefore, a critical analysis of historical data must be performed before applying a statistical approach.

Using historical data with time series data recorded by a tide gauge is not a simple process. Extreme statistical analysis requires a knowledge of the time period for which all the events above an extreme threshold are known (Leese, 1973; Payrastre et al., 2011). For the gauged, or preferably, systematic data, this time is represented by the recording period (i.e. systematic duration). As historical data are isolated data points, the duration of the observation is not defined. What happened during the time period between two or more past extreme events remains unknown. However, the coverage period estimation of historical data is a key step in extreme value statistical models (Prosdocimi, 2018). Reasonable computation of historical duration is the base of reliable statistical estimation of rare events.

In the past, to overcome these issues the perception threshold concept has been introduced. The perception threshold represents the minimum value over which all the historical data are reported and documented (Payrastre et al., 2011; Payrastre et al., 2013). The hypothesis is thus set that all the events over the perception threshold must have been recorded and the time period without records corresponds to time period without extreme events. This allows the estimation of the coverage period and the indirect assumption that the collected historical data are exhaustive throughout the coverage period assessed. So far, the perception threshold has allowed the use of historical information in extreme frequency analysis (Payrastre et al., 2011; Payrastre et al., 2013; Bulteau et al., 2015) but the hypothesis for using this approach seems sometimes difficult to prove.

Indeed, the perception threshold is usually defined equal to a historical event value, based on the principles that, if any other historical event would have occurred it would have been recorded (Payrastre et al., 2011; Bulteau et al., 2015). This is quite a sweeping hypothesis, which is difficult to verify and validate. Moreover, the duration associated to the perception threshold is usually defined as the duration between the oldest historical data and the start of the systematic series (Payrastre et al., 2011; Hamdi et al., 2015).

The main point of difference in this paper is the proposal of a new approach in order to deal with these questions in a more satisfactory way.

Note that the new approach is required for the application of the RFA, at least in the POT approaches (Van Gelder and Neykov, 1998; Bernardara et al., 2011; Bardet et al., 2011;

Weiss et al., 2014b). Indeed, the definition of the perception threshold as equal to a historical value precludes the possibility of using the RFA methodology proposed by Weiss (2014). Specifically, in Weiss regional methodology, the number of extreme events per year λ must be equal in each site of a region in order to estimate a regional distribution. This is impossible to achieve if the threshold is linked to the characteristics of the local observations.

In addition, when historical information (i.e. non-systematic data) is collected during a total failure of tide gauge, this influences the frequency of the event. One could only re-estimate λ if the credible duration is re-estimated (for details, see the Sect. 2).

In order to deal with these questions in this paper, a different approach is proposed to estimate the equivalent duration of historical data. This evaluation is based on the hypothesis that the frequency of the extreme events is not changing significantly in time. In other words, the average number of extreme events should be the same (obviously this has to be carefully verified and demonstrated case by case). In this way, we can assess the coverage period of historical data. Based on the frequency value of systematic storms occurrence, our hypothesis introduces a physical and tangible element in the estimation of historical duration. For this reason, we state that this advantage allows us to supplant the concept of a perception threshold providing by a credible estimation of the unknown historical duration used to estimate the extreme quantiles.

Thanks to this new concept of duration in this study, the FAB methodology is developed in order to use historical data and systematic data together in a RFA of extreme events. This methodology, described in detail in Sect. 2, is applied to the extreme skew surges given by 71 tide gauges located on the British, French and Spanish coasts and 14 historical collected skew surges is shown in Sect. 4. In addition, a functional example on the computation of credible duration is illustrated in Sect. 3.

2 FAB Methodology

2.1 The concept of credible duration

In the extreme value analysis (EVA) of natural hazards, the estimations of extreme quantiles are performed by applying statistical tools to a dataset of extreme variables. Usually, these datasets of extremes can be generated from continuous time series data or by taking extremes as the maximum value in several predefined time blocks (block maxima approach) or fixing a sampling threshold (peak-over-threshold (POT) approach).

Based on a POT approach, our study examines datasets of skew surges. From now on, we use skew surge as our chosen extreme variable but all concepts that we will describe can be extended also to other natural hazards.

In the POT framework, an extreme sampling threshold has to be considered in order to extract the dataset of extreme skew surges from continuous time series data of systematic skew surges recorded by tide gauge. In this way, the extreme sample is created and the extreme quantiles computation of skew surge can be performed according to the statistical distribution used. The relationship between return periods T and quantiles x_T is expressed as follows:

$$P(X > x_T) = \frac{1}{\lambda T}, \tag{1}$$

where $P(X > x_T)$ is the exceedance probability of x_T or the probability that in every λT years an event X, stronger than x_T, will occur. By using Eq. (1), the results can be illustrated through a return level plot.

The variable λ represents the number of skew surge events per year that exceed the sampling threshold. In other words, λ is the frequency of storm occurrence and it can be expressed as the ratio of the number of skew surges n over the sampling threshold to the recording time of systematic skew surge measurements in years d (hereafter called systematic duration).

$$\lambda = \frac{n}{d} \tag{2}$$

Knowledge of n and d is necessary to compute the frequency of extreme skew surges occurrence λ and to estimate extreme quantiles of skew surges. In other words, a duration estimation is needed to transform the probability [no units] in frequency [time^{-1} dimension].

When historical skew surges are available, the computation of λ is a complex topic. In fact, although the number of historical data n_{hist} over the sampling threshold is known, we do not have any information on what happens during non-gauged periods. It is thus impossible to define a duration over which the n_{hist} data are observed. In other words, regardless of the selection of sampling threshold value, we need this information to define the frequency of occurrence of events over the threshold (see the Eq. 2). We called the missing information "historical duration" d_{hist}, or in other words, the time period for which historical skew surges above the sampling threshold are estimated to be exhaustively recorded to the best of our knowledge. The inclusion of historical data in a systematic dataset of extreme skew surge is shown as an example in Fig. 1. It needs the formulation of a credible hypothesis in order to estimate the duration of the whole dataset composed by systematic and historical records.

Taking advantage of the perfect knowledge of systematic skew surges, the basic hypothesis proposed here for the computation of historical duration is that the frequency of storms λ remains unchanged over time. More precisely, λ_{syst}, namely the number of systematic skew surges per year over the sampling threshold, is supposed to be equal to λ_{hist}, namely the number of historical skew surges per year over

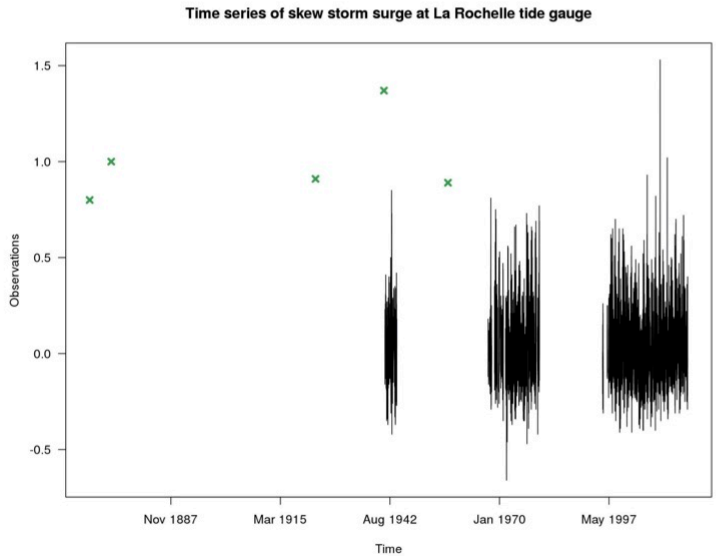


Figure 1. Illustrative example of the skew surge dataset at the La Rochelle tide gauge. In black the systematic skew surge and in green the historical isolated data points (non-systematic data) are shown.

the same sampling threshold.

$$\lambda = \lambda_{\text{syst}} = \lambda_{\text{hist}} \tag{3}$$

Equation (3) is in accordance with the lack of a significant trend on the extreme skew surges frequency during the second part of the 19th, 20th and the first part of 21st century in the regions considered in this study (Fig. 5).

According to current scientific literature, based on past observations, frequency of storms has no defined trend (Hanna et al., 2008; Barring and Fortuniak, 2009; MICORE Project, 2009). Therefore, the hypothesis that storm frequency is constant has been assumed. No significant trends are detected in terms of magnitude or frequency of storms between 1780 and 2005 in the North Atlantic and European regions since the Dalton minimum (Barring and Fortuniak, 2009). Over the past 100 years in the North Atlantic basin, no robust trends in annual numbers of tropical storms, hurricanes and major hurricanes have been identified by the last IPCC report of 2013. Hanna et al. (2008), Matulla et al. (2008) and Allan et al. (2009) state that observations have no clear trends over the past century or longer with substantial decadal and longer fluctuations with the exceptions of some regional and seasonal trends (Wang et al., 2009, 2011). Moreover, MICORE (2009) in the "Review of climate change impacts on storm occurrence" affirms that results from coastal areas do not detect any significant trend in the frequency of storms for the existing and available datasets.

Therefore, the hypothesis that storm frequency is constant may be a reasonable working hypothesis for demonstrating this approach. Obviously, it is crucial to test this hypothesis before the application of the approach in the present state to any new dataset. Note that the methodology could be adapted quickly in the future for non-stationary datasets, where the introduced concept can be adjusted. For skew surges, the validity this hypothesis is ensured.

The assumption that the annual frequency of skew surges has been constant in Europe during the last two centuries is warranted by the several references mentioned above. A solid check must be done for other marine variables to know if they can fit this hypothesis as well.

A new duration of observation d_{cr} is defined as "credible" (i.e. based on credible hypothesis) considering a steady λ. This novelty is expressed as follows:

$$d_{\text{cr}} = \frac{n_{\text{tot}}}{\lambda} = \frac{n_{\text{syst}} + n_{\text{hist}}}{\lambda} = \frac{n_{\text{syst}}}{\lambda} + \frac{n_{\text{hist}}}{\lambda}. \tag{4}$$

Equation (4) shows how the number of extreme data is divided in two sides according to the nature of data: n_{syst} is the number of extreme systematic data above the sampling threshold u and n_{syst} is the number of extreme historical data over the same threshold u. Therefore, the credible duration also depends on the sampling threshold value.

Moreover, this splitting indirectly provides the coverage period of historical skew surges in addition to the duration of recorded data. The historical side of the duration $d_{\text{cr}_{\text{hist}}}$ can be formulated as follows:

$$d_{\mathrm{cr}}^{\mathrm{hist}} = \frac{n_{\mathrm{hist}}}{\lambda}. \tag{5}$$

The coverage period on average that every historical skew surge holds is called associated duration $d_{\mathrm{cr,ass}}$.

$$d_{\mathrm{cr,ass}} = \frac{d_{\mathrm{cr}}^{\mathrm{hist}}}{n_{\mathrm{hist}}} = \frac{1}{\lambda} \tag{6}$$

Equation (6) displays how the associated duration can be evaluated just by knowing λ. Therefore, every extreme data not originated from a recorded time series has an associated duration. This is also the case for historical events that occurred during a time gap in gauged skew surges series.

A functional example is illustrated in Sect. 3 in order to explain all of the notions put forth up to this point and their relevance to the sampling threshold choice.

2.2 The use of credible duration in RFA framework

Weiss (2014) defined a methodology to fit a regional probability distribution for sea extreme variables. This approach enables the putting together of data from different locations and the forming of regions. The probability distribution of extreme skew surges must be the same inside a region. Thus, the extreme data of the same region are gathered together through the use of a normalisation parameter, which preserves the peculiarities of each site and the regional probability distribution defined as such. The formation of regions based on typical storm footprints (Weiss et al., 2013) is used to create physically and statistically homogeneous regions and it leads to knowledge of a master component of this regional process: the regional effective duration D_{eff}.

The regional effective duration D_{eff} represents the effective duration (in years), filtered of any intersite dependence, in which the number of regional data N_{r} are observed. D_{eff} is defined as the product of the degree of regional dependence $\varphi[1; N]$ and the mean of local durations:

$$D_{\mathrm{eff}} = \varphi \times \overline{d} = \frac{\lambda_{\mathrm{r}}}{\lambda}\overline{d} = \frac{N_{\mathrm{r}}}{\lambda}, \tag{7}$$

where λ_{r} is the number of regional storms per year $\lambda_{\mathrm{r}} = N_{\mathrm{r}}/\overline{d}$ and $\overline{d}$ is the mean of local observations' duration of the N regional sites: $\overline{d} = \sum_{i=1}^{N} d_i/N$.

The local frequency of storms λ must be the same across the region. More details about the D_{eff} are clarified in Weiss et al. (2014b) for the classic RFA approach, which is set for systematic data.

The knowledge of the regional sample duration, as exposed by a generic frequency analysis in the previous paragraph of this section, is a required factor. In fact, the evaluation of the regional return period T_{r} of a particular regional quantile $x_{T_{\mathrm{r}}}$ is expressed as in Eq. (8) as follows:

$$T_{\mathrm{r}} = \frac{1}{\lambda_{\mathrm{r}} P\left(X_s > x_{T_{\mathrm{r}}}\right)}, \tag{8}$$

where $P\left(X_s > x_{T_{\mathrm{r}}}\right)$ is the probability that a storm s impacts at least one site in the region with a normalised intensity greater than $x_{T_{\mathrm{r}}}$. Therefore, thanks to the regional distribution F_{r}, the local return period of the storm s is calculated as follows:

$$T = \frac{1}{\lambda\left(1 - F_{\mathrm{r}}(x_{T_{\mathrm{r}}})\right)}. \tag{9}$$

In addition, Weiss et al. (2014b) use for the computation of the empirical local return periods a Weibull plotting position. This element is correlated with the regional duration D_{r}:

$$\widetilde{T}_{s,\mathrm{loc}} = \frac{D_{\mathrm{r}} + 1}{n_{\mathrm{r}} + 1 - \mathrm{rank}(s)}, \tag{10}$$

here the biggest storm $s_{\max}$ of the regional sample is linked to a local return period thanks to D_{r}.

However, the need to employ historical data in the extreme analysis has been considered as highly relevant by scientific community to get reliable quantiles and to not miss significant information. In the RFA, the use of historical records is allowed by the new notion of credible duration. A new way to process historical skew surges at regional scale is introduced. The addition of one or more historical skew surges in one or more sites of a region requires the knowledge of the duration of each site. Thanks to the credible duration concept applied previously to a single site, a coverage period for both types of data can be evaluated. It allows switching from the notion of the regional effective duration $D_{\mathrm{r,eff}}$ to the regional credible duration $D_{\mathrm{r,cr}}$:

$$D_{\mathrm{r,cr}} = \varphi \times \overline{d} = \varphi \sum_{i=1}^{N} \frac{d_{i,\mathrm{cr}}}{N}. \tag{11}$$

Regional credible duration depends also on the degree of regional dependence φ that increases when the data are independent. Moreover, Eq. (11) shows that a local duration surplus consequently brings an excess of the coverage period of regional sample. With some simplifications, the Eq. (11) can thus be reformulated as follows:

$$D_{\mathrm{r,cr}} = \frac{\lambda_r}{\lambda}\overline{d} = \frac{N_r}{\lambda}. \tag{12}$$

The lack of a trend on storms' frequency is a necessary condition to assume a steady λ. If this hypothesis is satisfied, the regional credible duration can be computed.

The regional frequency analysis of skew surges including historical data, or FAB method, can be performed and these two combined approaches ensure better estimations of extreme values.

3 Illustration of credible duration

In order to better illustrate the concept of credible duration and how it varies according to different thresholds, a specific illustration on connection between these variables is described in this section by the example shown in Fig. 1.

Table 1. Summary table of several variables involved in the computation of credible duration for 4 different cases of study.

	Case 1	Case 2	Case 3	Case 4
Lambda λ	0.5	0.23	0.2	0.07
Sampling threshold (m)	0.7	0.77	0.81	1.02
Systematic duration (years)	32.88	32.88	32.88	32.88
Historical credible duration (years)	10	21.74	20	14.29
Number of historical data OT	5	5	4	1
Associated duration (years)	2	4.35	5	14.29
Credible duration (years)	42.88	54.62	52.88	47.17

Figure 1 shows the functional example of the whole skew surge series recorded at the La Rochelle tide gauge (France). Systematic skew surges are represented by the continuous series of data plotted in black from about 1940 to 2017. They are observed during a fixed period equal to the recording time period, in this case, 32.88 years.

The historical skew surges are artificial data generated by ourselves for this descriptive illustration and they are represented as green cross points. As they are not supposed to be obtained by continuous systematic records of skew surge measurements, the duration of their support is unknown.

As expressed above, d_{cr} and in particular $d_{\mathrm{cr}}^{\mathrm{hist}}$ depends on the sampling threshold value u. Equations (4) and (5) show this dependence through the λ value.

The application of the new proposed concept allows the computation of the historical duration of skew surges that represents a basic way to take into account the historical data on the estimation of extreme values.

Although the period of systematic observation is known and fixed by tide gauge's records, the coverage period of historical observations is associated with the value of sampling threshold u. The λ value is computed from the systematic measurements (Eq. 2). As a result of the hypothesis that the frequency of storms λ has no trend (Eq. 3), this λ value is then used to compute the duration of the historical records. Moreover, the hypothesis of Eq. (3) leads to different values of historical duration for each λ selected value and consequently, for each likely sampling threshold. Four cases of different values of high sampling threshold are examined (Fig. 2 and Table 1); these particular cases are chosen in order to let us know the effects of different sampling thresholds on the computation of the credible duration.

Case 1 uses a sampling threshold u_{c1} corresponding to 0.5 extremes per year ($\lambda = 0.5$) and lead to a credible duration of 42.88 years (Table 1). A higher threshold u_{c2} is used in Fig. 2 for the Case 2 ($\lambda = 0.23$) and the credible duration grows until 54.62 years as the number of historical data over the threshold does not change. In Case 3, the threshold (u_{c3}) increases ($\lambda = 0.2$) but the credible duration decreases (52.88 years) because the smallest value of historical skew surge does not exceed the threshold. On average, the coverage period of each historical event still rises. Case 4 displays the use of a very high threshold u_{c4} ($\lambda = 0.07$) to define the sample of extremes. In this case, only one historical event is bigger than threshold and its coverage period, or associated duration, corresponds to the historical duration. When the historical data value is larger, the probability that the historical event covers a longer period is higher.

As stated above, the four cases (resume in Fig. 2 and Table 1) show the behaviour of credible duration for different values of threshold. As threshold value increases, associated duration increases as well. Equation (4) expresses the credible duration as the sum of systematic duration and historical duration. Systematic duration is steady and it is independent from threshold value. For this reason, the variation of credible duration depends only on the variation of historical duration. In particular, when the threshold is extremely high, it is almost sure that no extreme skew surges occurred during the coverage period of historical data. If the threshold is lower, then it is less certain that all extreme skew surges have been observed in the historical period.

However, it is not always true that the credible duration is biggest when threshold grows (Fig. 3). In fact, the credible duration and historical duration depends not only on the threshold value but also on the number of historical events exceeding it. If the threshold is extremely high, not all historical data are taken into account in the extremes' sample. On the contrary, the associated duration of each set of historical data increases when the threshold grows for the reasons explained in the previous paragraph (shown by the blue line in Fig. 3).

In conclusion, the sampling threshold choice is an important step in the extreme quantiles estimation allowing the computation of the credible duration. Finally, with all of these elements we can deal with historical data in a regional framework.

4 Application of the FAB method to a real dataset

4.1 Systematic skew surge database and historical data

The systematic skew surge database is originated from a time series of sea levels recorded by 71 tide gauges located along French, British and Spanish coasts of the Atlantic Ocean, the English Channel, the North Sea and the Irish Sea (Fig. 4).

French data are provided by SHOM (Service Hydrographique et Océanographique de la Marine, France), English data by BODC (British Oceanographic Data Center, UK) and Spanish data by IEO (Instituto Espanol de Oceanografia, Spain). All data except the Spanish data, are updated to January 2017. Every site has different systematic durations depending on the tide gauge records. Starting from 1846, the longest time series is recorded at Brest for a systematic period of 156.57 years. Sea levels are recorded hourly except for British ports. BODC supplies sea levels every 15 min from 1993 onwards.

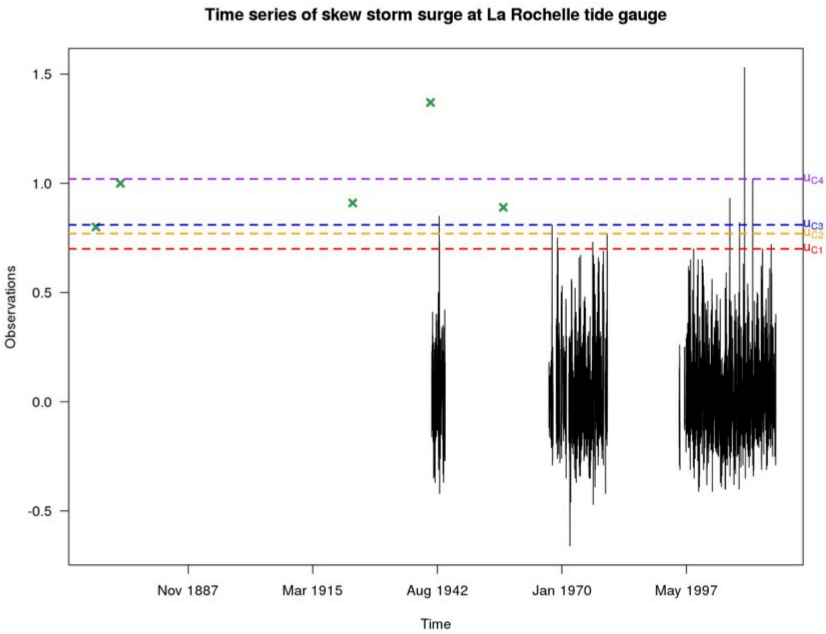


Figure 2. Skew surge dataset of Sect. 3 (composed of systematic and historical data) used in the computation of credible duration with 4 different thresholds u_{c1}, u_{c2}, u_{c3}, u_{c4}. The four sampling thresholds u_c are used linked to a rising value o λ ($\lambda_{c1} = 0.5$, $\lambda_{c2} = 0.23$, $\lambda_{c3} = 0.2$, $\lambda_{c4} = 0.07$).

Skew surge data are generated at each site from the difference between the maximum sea levels recorded around the time of the predicted astronomical high tide and the selfsame predicted astronomical high tide. Before computing skew surge, sea levels must be corrected by a potentially long-term alteration of mean sea levels, or eustatism. Only sea level data with significant linear trends are edited. More details on the computation of skew surges are described by Simon (2007), Bernardara et al. (2011) and Weiss (2014). In this way, the systematic database of 71 skew surge series is created.

Finding accurate historical data is complex and difficult and at the time of writing only 14 historical skew surges are available. These historical records are located in 3 of the 71 ports: 9 at La Rochelle (Gouriou, 2012; Garnier and Surville, 2010; Breilh, 2014; Breilh et al., 2014), 4 at Dunkirk (DREAL, 2009; Parent et al., 2007; Pouvreau, 2008; Maspataud, 2011) and 1 at Dieppe (provided by the Service hydrographique et océanographique de la Marine SHOM, France). The historical data are introduced in corresponding skew surge series of the concerned sites.

This skew surge database of systematic and historical data is exploited in this study for the estimation of regional return periods.

4.2 Formation of regions

The method to form homogeneous regions proposed by Weiss et al. (2014a) is used to get physically and statistically homogeneous regions from the database shown above. This methodology is based on a storm propagation criterion identifying the most typical storms footprints. The configuration of parameters to detect storms is set using the same criteria as Weiss (2014). The most typical storm footprints are revealed by Ward's hierarchical classification (Ward, 1963) and correspond to four physical regions. The homogeneity test (Hosking and Wallis, 1997) is applied to check if all regions are also statistically homogeneous. Only one physical region is not statistically homogeneous, and after an inner partition and a further check, five physically and statistically homogeneous regions are obtained.

Moreover, a double threshold approach (Bernardara et al., 2014) is employed to separate the physical part from the statistical part of the used extreme variables and two different thresholds are used: the first to define physical storms (physical threshold) and the second to define regional samples (statistical sampling threshold, or more simply, sampling threshold).

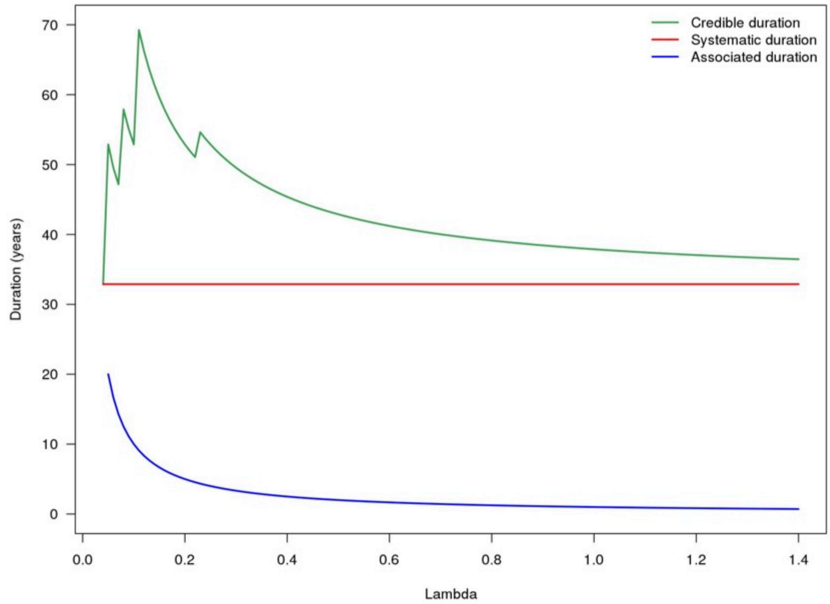


Figure 3. Fictive example of skew surge dataset of systematic and historical data – Credible duration (green line), systematic duration (red line) and associated duration (blue line) for several threshold values.

Figure 5 displays the five regions founded out through the application of Weiss' methodology. Although a different input database is used, the number and the position of the regions are similar to Weiss' study.

The regionalisation process shows that only Region 1 and Region 2 hold historical data. The three sites with historical data belong to these two regions: La Rochelle to Region 1 and Dieppe and Dunkirk to Region 2.

The region with more historical information is the Region 1 (nine historical skew surges) and the results of FAB method proposed are focused on this region.

4.3 Focus on Region 1

4.3.1 Statistical threshold and credible duration

The sample of regional extremes is analysed with statistical tools in order to estimate regional return levels associated with very extreme return periods. A regional pooling method is used to select independent maximum data. A regional sample is formed through normalised local data. Maximum local data are divided by a local index, equal to the local sampling threshold value (Weiss, 2014). Moreover, this methodology enables us to estimate the regional distribution F_r considered as the distribution of the normalised maximum data above the statistical sampling threshold.

This threshold corresponds to a storm's frequency λ, which is assumed equal at all sites of a region. The sampling threshold value selection, or more accurately the λ value selection, is a complicated topic (Bernardara et al., 2014). There is no an unequivocal approach for defining the exact sampling threshold. Twelve sensitivity indicators depending on all possible λ values are proposed to assign the best λ value for the regional analysis of extreme skew surges using historical data: the statistical homogeneity test (Hosking and Wallis, 1997), the stationarity test applied to storm intensities of regional sample (Hosking and Wallis, 1993), the chi-squared test for the regional distribution (Cochran, 1952), the statistical test to detect outliers in regional sample (Barnett and Lewis, 1994; Hubert and Van der Veeken, 2008; Weiss, 2014), the value of regional credible duration, the number of regional sample data, the value of the degree of regional dependence φ, the visual look of the regional return plot and the stability of regional credible duration, local return levels, shape parameter and scale parameter of the GPD distribution (Coles, 2001). The λ value selection can not be carried out through one of these methods and only the combination of all indicators can enable us to appoint the optimal λ. All sensitivity indicators have the same relevance in the proposed sensitivity analysis.

A skew surge's frequency of 0.36 is chosen for Region 1 thanks the sensitivity analysis of the statistical factors men-

Figure 4. Location of the 71 tide gauges used in this study.

Figure 5. The five physically and statistically homogeneous regions of the 71 tide gauges.

tioned above. Taking $\lambda = 0.36$ results in having, on average, one extreme skew surge over the sampling threshold every 2.78 years at each site. This threshold selection permits the computation of the credible duration for every site belonging to the Region 1. In Region 1, all ports except for La Rochelle have only systematic skew surge and their credible duration, equal to systematic duration, is determined and independent from the sampling threshold value.

On the contrary, the site of La Rochelle has nine historical skew surges and its credible duration depends both on systematic duration and historical duration. The systematic duration does not change and it is equal to the recording period of La Rochelle tide gauge. As shown in the fictive illustration of Sect. 3, the historical duration differs as threshold value changes.

The credible duration at La Rochelle is 57.88 years and the contribution of historical data is almost the 43 % of the total duration. Obviously, without these nine historical data, the coverage period of extreme data at La Rochelle would have been less than 25 years, equivalent to the historical duration. All the historical data are over the chosen threshold and, consequently, the associated duration of each historical data is equal to 2.77 years.

4.3.2 Regional credible duration

Equation (11) enables us to compute the regional credible duration for Region 1. The introduction of nine historical data in the local time series and only eight in the regional sample after the use of the regional pooling method, grants an earning of 11.11 years of regional duration. As Table 2 shows, historical data increase the credible durations for sites where there are historical records and consequently the regional credible duration likewise increases. In Region 1, only the site of La Rochelle increases its duration. This contribution allows for the increase of the regional credible duration.

Regional credible duration depends not only on the mean of local duration but also on the degree of regional dependence ϕ (Eq. 11). Obviously, this degree does not change with only nine more added data. In conclusion, the rise of the regional credible duration of Region 1 is produced mostly by the increase of credible duration at La Rochelle.

Table 2 shows the difference with the FAB method that allows to take into account historical data and a traditional method of RFA (Bernardara et al., 2011; Weiss, 2014) that does not enable the use of historical data. Figure 6 displays how the more λ decreases in value, the more the difference between regional credible duration with the nine historical data (blue line) and the regional effective duration without historical data (green line) for Region 1 increases (red line). This is due to the greater certainty that historical data above the sampling threshold are the biggest over a longer period.

4.3.3 Regional return levels

Regional return levels for Region 1 are obtained through the fit of a generalised Pareto distribution (Fig. 7). Penalised maximum likelihood estimation (PMLE) (Coles and Dixon, 1999) is used to estimate the regional GPD parameters (γ, k). As shows Fig. 7, the regional distribution is unbounded due to a positive shape parameter k. The return level plot

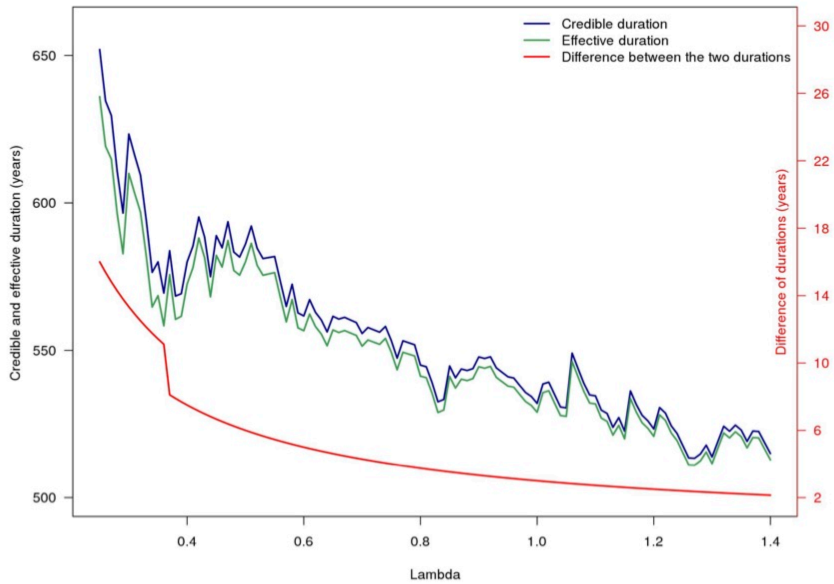


Figure 6. Variation in regional credible duration (taking into account the nine historical data) and regional effective durations (only with systematic data) and their difference (in red) for Region 1.

Table 2. Summary table of the results obtained with the application of FAB method for Region 1.

	RFA without historical data	RFA with nine historical data (FAB, 2017)
Lambda λ	0.36	0.36
Local duration on average (years)	41.4	42.36
Regional credible duration (years)	558.33	569.44
Intensity of biggest normalised surge	2.13 (SD)	3.01 (HD)

of Region 1 is produced with the 95 % confidence intervals generated by a bootstrap of observed storms. A Weibull plotting position is employed for the regional empirical return period. Return levels for each site are achieved through multiplication between regional return levels and local indices, equivalent to the local sampling threshold.

Red dotted points represent the eight historical data taken into account in Region 1. As it appears, the benefit of using historical data is not only an increase of the coverage period of extreme data but also to be able to consider in a frequency analysis when non-systematic storms occurred. The historical skew surges used in this study represent very big storms and for this reason it is newsworthy to take into account their large magnitude. Indeed, the biggest normalised skew surge considered in Region 1 is 3.01 compared with 2.13 of the regional sample generated without historical data (Table 2).

5 Conclusions

The application of the FAB method, based on the new concept of credible duration, allows more reliable extreme estimations of return levels to be obtained due to the inclusion of historical data to a systematic skew surge database.

The introduction of historical data on a frequency analysis of extremes is an innovative way to increase the reliability of estimations. In the past, the estimation of the period covered by the historical data collection has usually limited the use of historical data. The concept of a perception threshold was employed to use historical data even though its hypothesis is quite broad and difficult to verify. For this reason, the new concept of credible duration has been formulated that is based on physical considerations supported by several scientific studies. This leads to use historical isolated data points on the EVA.

The hypothesis, that the frequency of storms λ remains unchanged in time and on which credible duration concept is based, is satisfactory for skew surge. For other variables, a hypothesis check has to be done. Following the hypothesis check, all the considerations that we raised in this study can also be applied for other natural hazards.

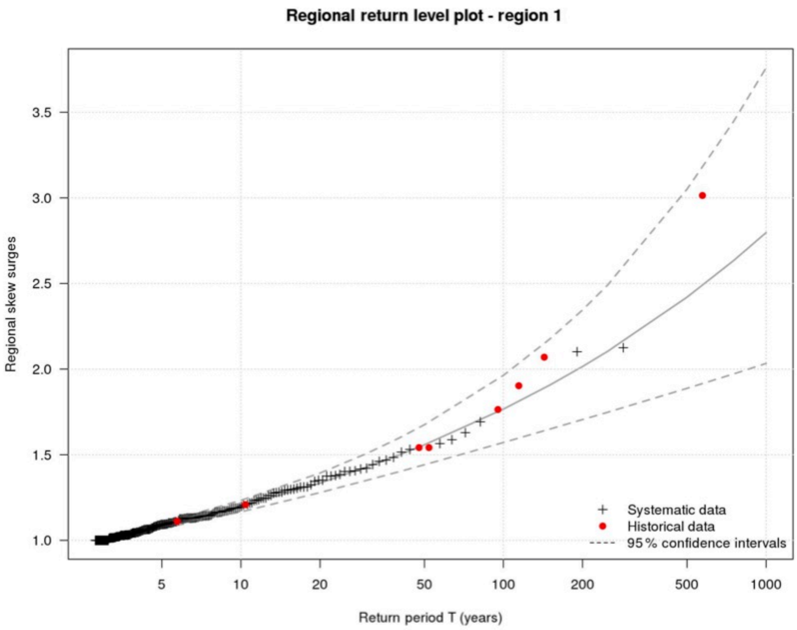


Figure 7. Return level plot for the regional distribution of Region 1 with the confidence intervals of 95 % estimated by bootstrap. Red points represent the regional historical data and black crosses the regional systematic data.

When historical data are available, the credible duration depends on sampling threshold. This novelty allows the expansion of the credible duration concept to a regional scale. The RFA is another efficient way to reduce uncertainties on the extreme estimations. These two ways can be used together by the proposed FAB method.

The results describe the relevance of using historical records on a regional analysis both in terms of regional duration and of storm magnitude. Sometimes, these data can represent very large events and so we have to consider them during the estimation of extreme variables. Moreover, the existence of historical records enables to raise the coverage period of the extreme sample.

No extreme data can be missed in an extreme frequency analysis. It is very important to use all of the available data in order to get increasingly more reliable extreme estimations and regional analysis and historical data are two tools that can be merged to get satisfactory results.

New perspectives are opened to improve extreme quantiles estimations by using historical data that is becoming increasingly more available through different investigations on several qualitative and quantitative sources. The aggregation of all of these data in a historical database is a need to exploit the maximum amount of extremes.

Furthermore, every historical data value is associated with different uncertainties and it is challenging to be able to take into account these uncertainties in the regional analysis.

Data availability. The systematic data can be obtained by request addressed to the corresponding author: SHOM (Service Hydrographique et Océanographique de la Marine) for French data, the BODC (British Oceanographic Data Centre) for British data and the IEO (Instituto Español de Oceanografía) for Spanish data.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The permission to publish the results of this ongoing research study was granted by the EDF (Electricité De France). The results in this paper should, of course, be considered as R&D exercises without any significance or embedded commitments upon the real behaviour of the EDF power facilities or its regulatory control and licensing. The authors would like to acknowledge the SHOM (Service Hydrographique et Océanographique de la Marine, France), the BODC (British Oceanographic Data Centre, UK) and the IEO (Instituto Español de Oceanografía, Spain) for providing the data used in this study.

Edited by: Ira Didenkulova
Reviewed by: two anonymous referees

References

Allan, R., Tett, S., and Alexander, L.: Fluctuations in autumn-winter severe storms over the British Isles: 1920 to present, Int. J. Climatol., 29, 357–371, https://doi.org/10.1002/joc.1765, 2009.

ASN: Guide de l'ASN n° 13 : Protection des installations nucléaires de base contre les inondations externes, 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 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, https://doi.org/10.5194/nhess-11-2791-2011, 2011.

Bardet, L., Duluc, C.-M., Rebour, V., and L'Her, J.: Regional frequency analysis of extreme storm surges along the French coast, Nat. Hazards Earth Syst. Sci., 11, 1627–1639, https://doi.org/10.5194/nhess-11-1627-2011, 2011.

Barnett, V. and Lewis, T.: Outliers in Statistical Data, 3rd Edn., John Wiley, Chichester, 1994.

Barriendos, M., Coeur, D., Lang, M., Llasat, M. C., Naulet, R., Lemaitre, F., and Barrera, A.: Stationarity analysis of historical flood series in France and Spain (14th–20th centuries), Nat. Hazards Earth Syst. Sci., 3, 583–592, https://doi.org/10.5194/nhess-3-583-2003, 2003.

Barring, L. and Fortuniak, K.: Multi-indices analysis of southern Scandinavian storminess 1780-2005 and links to interdecadal variations in the NW Europe-North sea region, Int. J. Climatol., 29, 373–384, https://doi.org/10.1002/joc.1842, 2009.

Benito, G., Lang, M., Barriendos, M., Llasat, M., Frances, F., Ouarda, T., Thorndycraft, V., Enzel, Y., Bardossy, A., Cœur, D., and Bobée, 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.

Benson, M. A.: Use of historical data in flood-frequency analysis, Eos, Transactions American Geophysical Union, 31, 419–424, https://doi.org/10.1029/TR031i003p00419, 1950.

Bernardara, P., Andreewsky, M., and Benoit, M.: Application of regional frequency analysis to the estimation of extreme storm surges, J. Geophys. Res., 116, C02008, https://doi.org/10.1029/2010JC006229, 2011.

Bernardara, P., Mazas, F., Kergadallan, X., and Hamm, L.: A two-step framework for over-threshold modelling of environmental extremes, Nat. Hazards Earth Syst. Sci., 14, 635–647, https://doi.org/10.5194/nhess-14-635-2014, 2014.

Breilh, J. F.: Les surcotes et les submersions marines dans la partie centrale du Golfe de Gascogne : les enseignements de la tempete Xynthia, PhD thesis, Sciences de la Terre, Université de La Rochelle, France, available at: https://tel.archives-ouvertes.fr/tel-01174997/document (last access: 20 March 2018), 2014.

Breilh, J. F., Bertin, X., Chaumillon, E., Giloy, N., and Sauzeau, T,: How frequent is storm-induced flooding in the central part of the Bay of Biscay?, Global Planet. Change, 122, 161–175, https://doi.org/10.1016/j.gloplacha.2014.08.013, 2014.

Bulteau, T., Idier, D., Lambert, J., and Garcin, M.: How historical information can improve estimation and prediction of extreme coastal water levels: application to the Xynthia event at La Rochelle (France), Nat. Hazards Earth Syst. Sci., 15, 1135–1147, https://doi.org/10.5194/nhess-15-1135-2015, 2015.

Cochran, W. G.: The $\chi 2$ Test of Goodness of Fit, Ann. Math. Stat., 23, 315–345, 1952.

Cohn, T. A.: The incorporation of historical information in flood frequency analysis, MS thesis, Dep. of Environ. Eng., Cornell Univ., Ithaca, NY, USA, 1984.

Cohn, T. A. and Stedinger, J. R.: Use of historical information in a maximum-likelihood framework, J. Hydrol., 96, 215–223, https://doi.org/10.1016/0022-1694(87)90154-5, 1987.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer, London, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Coles, S. and Dixon, M.: Likelihood-based inference for extreme value models, Extremes, 2, 5–23, https://doi.org/10.1023/A:1009905222644, 1999.

Condie, R. and Lee, K. A.: Flood frequency analysis with historic information, J. Hydrol., 58, 47–61, https://doi.org/10.1016/0022-1694(82)90068-3, 1982.

Cunnane, C.: Methods and merits of regional floods frequency analysis, J. Hydrol., 100, 269–290, https://doi.org/10.1016/0022-1694(88)90188-6, 1988.

Dalrymple, T.: Flood Frequency Analyses, Manual of Hydrology: Part 3, "Flood Flow Techniques" U.S. Geological Survey Water-Supply Paper, 1543-A, US GPO, 1960.

DREAL (Ministère de l'Ecologie, de l'Energie, du Développement durable, et de l'Aménagement du territoire), 59/62 SREI/DRNHM: Détermination de l'aléa de submersion marine intégrant les conséquences du changement climatique en région Nord – Pas-de-Calais – Etape 1: compréhension du fonctionnement du littoral, available at: http://www.hauts-de-france.developpement-durable.gouv.fr/IMG/pdf/Determination-alea-submersion-marine-integrant-consequences-changement-climatique-dreal-npdc.pdf (last access: 20 March 2018), 2009.

Fréchet, M.: Sur la lois de probabilité de l'écart maximum, Annales de la Société Polonaise de Mathématique, 6, 93–122, 1928.

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, Le croît Vif., Saintes, France, 2010.

Gnedenko, B.: Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math., 44, 423–453, 1943.

Gouriou, T.: Evolution des composantes du niveau marin à partir d'observations de marégraphe effectuées depuis la fin du 18e siècle en Charente-Maritime, Ph.D. thesis, Océanographie physique, Université de La Rochelle, France, 2012.

Gumbel, E. J.: Statistics of Extremes, Columbia University Press, New York, USA, 1958.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V.: Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France, Nat. Hazards Earth Syst. Sci., 15, 1515–1531, https://doi.org/10.5194/nhess-15-1515-2015, 2015.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V.: Development of a target-site-based regional frequency model using historical information, Geophys. Res. Abstr., EGU2016-8765, EGU General Assembly 2016, Vienna, Austria, 2016.

Hanna, E., Cappelen, J., Allan, R., Jonsson, T., Le Blancq, F., Lillington, T., and Hickey, K.: New Insights into North European and North Atlantic Surface Pressure Variability, Storminess, and Related Climatic Change since 1830, J. Climate, 21, 6739–6766, https://doi.org/10.1175/2008JCLI2296.1, 2008.

Hosking, J. R. M. and Wallis, J. R.: Paleoflood hydrology and flood frequency analysis, Water Resour. Res., 22, 543–550, https://doi.org/10.1029/WR022i004p00543, 1986a.

Hosking, J. R. M. and Wallis, J. R.: The value of historical data in flood frequency analysis, Water Resour. Res., 22, 1606–1612, https://doi.org/10.1029/WR022i011p01606, 1986b.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resour. Res., 29, 271–281, https://doi.org/10.1029/92WR01980, 1993.

Hosking, J. R. M. and Wallis, J. R.: Regional Frequency Analysis, An approach based on L-moments, Cambridge University Press, 1997.

Hubert, M. and Van der Veeken, S.: Outlier detection for skewed data, J. Chemometr., 22, 235–246, https://doi.org/10.1002/cem.1123, 2008.

IPCC: Climate change 2013: the physical science basis, Contribution to the IPCC fifth assessment report (AR5), Chapter 2: Observations: Atmosphere and Surface, 2013.

Leese, M. N.: Use of censored data in the estimation of Gumbel distribution parameters for annual maximum flood series, Water Resour. Res., 9, 1534–1542, https://doi.org/10.1029/WR009i006p01534, 1973.

Maspataud, A.: Impacts des tempêtes sur la morphodynamique du profil côtier en milieu macrotidal, Ph.D. thesis, Océanographie, Université du Littoral Côte d'Opale, France, available at: https://tel.archives-ouvertes.fr/tel-00658671/document (last access: 20 March 2018), 2011.

Matulla, C., Schoner, W., Alexandersson, H., von Storch, H., and Wang, X. L.: European storminess: late nineteenth century to present, Clim. Dynam., 31, 125–130, https://doi.org/10.1007/s00382-007-0333-y, 2008.

MICORE Project 202798: Review of climate change impacts on storm occurrence, edited by: Ferreira, O. (UALG), Vousdoukas, M. (UALG), and Ciavola, P. (UniFe), available at: http://www.micore.eu/file.php?id=4 (last access: 20 March 2018), 2009.

Nguyen, C. C., Gaume, E., and Payrastre, O.: Regional flood frequency analyses involving extraordinary flood events at ungauged sites: further developments and validations, J. Hydrol., 508, 385–396, https://doi.org/10.1016/j.jhydrol.2013.09.058, 2014.

Ouarda, T. B. M. J., Rasmussen, P. F., Bobée, B., and Bernier, J.: Utilisation de l'information historique en analyse hydrologique fréquentielle, Revue des sciences de l'eau, 11, 41–49, https://doi.org/10.7202/705328ar, 1998.

Parent, P., Butin, T., Vanhee, S., and Busz, N.: Un territoire soumis au risque de submersion marine. Les tempetes de 1949 et 1953 à Dunkerque, Institution départementale des Wateringues, available at: http://www.ingeo.fr/fichier/file/Wateringues/Affiche4-Dunkerque.pdf (last access: 20 March 2018), 2007.

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, https://doi.org/10.1029/2010WR009812, 2011.

Payrastre, O., Gaume, E., and Andrieu, H.: Historical information and flood frequency analyses: which optimal features for historical floods inventories?, La Houille Blanche – Revue internationale de l'eau, 3, 5–11, https://doi.org/10.1051/lhb/2013019, 2013.

Picklands, J.: Statistical inference using extreme order statistics, The Ann. Stat., 3, 119–131, 1975.

Pouvreau, N.: Trois cents ans de mesures marégraphiques en France : outils, méthodes et tendances des composantes du niveau de la mer au port de Brest, PhD thesis, Climatologie, Université de La Rochelle, France, available at: https://tel.archives-ouvertes.fr/tel-00353660/document (last access: 20 March 2018), 2008.

Prosdocimi, I.: German tanks and historical records: the estimation of the time coverage of ungauged extreme events, Stoch. Environ. Res.-Risk. Assess., 32, 607–622, https://doi.org/10.1007/s00477-017-1418-8, 2018.

Reis, D. S. and Stedinger, J. R.: Bayesian MCMC flood frequency analysis with historical information, J. Hydrol., 313, 97–116, https://doi.org/10.1016/j.jhydrol.2005.02.028, 2005.

Sabourin, A. and Renard, B.: Combining regional estimation and historical floods: A multivariate semiparametric peaks-over-threshold model with censored data, Water Resour. Res., 51, 9646–9664, https://doi.org/10.1002/2015WR017320, 2015.

SHOM (Service Hydrographique et Océanographique de la Marine): Références Altimétriques Maritimes, available at: http://diffusion.shom.fr/pro/risques/references-verticales/references-altimetriques-maritimes-ram.html (last access: 20 March 2018), 2017.

Simon, B.: La marée océanique côtière, Editions de l'Institut Océanographique, available at: http://diffusion.shom.fr/pro/la-maree-oceanique-cotiere.html# (last access: 20 March 2018), 2007.

Stedinger, J. R. and Baker, V. R.: Surface water hydrology: historical and paleoflood information, Rev. Geophys., 25, 119–124, https://doi.org/10.1029/RG025i002p00119, 1987.

Stedinger, J. R. and Cohn, T. A.: Flood frequency analysis with historical and paleoflood information, Water Resour. Res., 22, 785–793, https://doi.org/10.1029/WR022i005p00785, 1986.

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., and Bouchart, F., Stochastic Hydraulics '96, Proceedings of the 7th IAHR International Symposium, Mackay, Queensland, Australia, 243–249, 1996.

Van Gelder, P. H. A. J. M. and Neykov, N. M.: Regional frequency analysis of extreme water level along the Dutch coast using L moments: A preliminary study, Stochastic models of hydrological processes and their applications to problems of environmental preservation, NATO Advanced Research Workshop, 23–27 November 1998, Moscow, Russia, 14–20, 1998.

Wang, X. L., Zwiers F. W., Swail V. R., and Feng Y.: Trends and variability of storminess in the Northeast Atlantic region, 1874–2007, Clim. Dynam., 33, 1179–1195, https://doi.org/10.1007/s00382-008-0504-5, 2009.

Wang, X. L., Wan, H., Zwiers, F. W., Swail, V. R., Compo, G. P., Allan, R. J., Vose, R. S., Jourdain, S., and Yin, X.: Trends and low-frequency variability of storminess over western Europe, 1878–2007, Clim. Dynam., 37, 2355–2371, https://doi.org/10.1007/s00382-011-1107-0, 2011.

Ward, J. H.: Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc., 58, 236–244, 1963.

Weiss, J.: Analyse régionale des aléas maritimes extrêmes, PhD thesis, Mécanique des fluides, Université Paris-Est, France, available at: https://pastel.archives-ouvertes.fr/tel-01127291/document (last access: 20 March 2018), 2014.

Weiss, J., Bernardara, P., and Benoit, M.: A method to identify and form homogeneous regions for regional frequency analysis of extreme skew storm surges, Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards, 18–20 September 2013, Siegen, Germany, 2013.

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, https://doi.org/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, 5926–5940, https://doi.org/10.1002/2014WR015391, 2014b.

1.2 Regional analysis and historical data

The new concept of credible duration and the application of the Regional Analysis including historical data was presented in advance by a talk at the "Natural Hazards and climate change impacts in coastal areas" session at the EGU 2017 (European Geosciences Union General Assembly) in Wien (Austria).

Geophysical Research Abstracts
Vol. 19, EGU2017-5101-1, 2017
EGU General Assembly 2017

Including historical data on the Regional Analysis of extreme storm surges

Roberto Frau (1,2), Marc Andreewsky (1), Pietro Bernardara (2,3)

(1) Laboratoire National d'Hydraulique et Environnement (LNHE), EDF R&D, Chatou, France, (2) Université Paris Est, Saint-Venant Laboratory for Hydraulics (ENPC, EDF R&D, CEREMA), Chatou, France, (3) EDF Energy R&D UK Centre, London, UK

The occurrence of rare (up to 10^{-4} annual probability of occurrence) and extreme oceano-meteorological conditions is required to design effective coastal protections and to protect coastal areas from flooding. Statistical methods based on the extreme value theory are widely used for the characterisation of such events. Unfortunately, this was often performed from very short observations series (usually 30 years to 50 years long) and the uncertainties associated to the extrapolation are too wide to be used in a design approach. Nowadays, larger dataset are available, in particular when looking at the regional scale. Moreover, in several disciplines the historical evidence of extreme events observed before the systematic observation periods, has been provided by specific studies. In the framework, the combination between these two sources of information is shown to dramatically increase the reliability of the statistical extrapolation. Merge the two approaches is however challenging because of the need to define homogenous regions and to consider the fact that historical data are not continuous and exhaustive. Here, an overall methodology is presented which allows a robust estimation, including a selection an optimal sampling threshold and a definition a "regional credible duration", which describe the amount of information available, taking into account the unknown period of observation associated with historical events. The Regional Historical Analysis is applied on a database of extreme skew storm surges and on a several extreme historical storm surges collected from different sources.

Annexe B

THE BAYESIAN INFERENCE

2.1 The Bayesian inference

The Bayesian inference is an approach to identify the statistical inference of a sample. The Bayesian context provides a posterior probability linked to the Bayes' theorem on which the expert could able to give a state of knowledge on the particular phenomenon analysed.

Dealing with frequentist probabilities, the parameters of a statistical distribution are considered as fixed notwithstanding they are effectively unknown. In fact, they have to be estimated using a defined number of observations of a particular phenomenon and there is no way to consider them as a random variable. Conversely, the Bayesian probability allows the definition of distributions of these unknown parameters.

As stated above, Bayesian probability is associated to the formulation of Bayes' theorem (Bayes, 1763) for the posterior distribution of parameters' vector θ:

$$P(\theta|D) = \frac{P(\theta)\,P(D|\theta)}{P(D)} \qquad \text{(B2.1)}$$

the posterior distribution $P(\theta|D)$ of parameters' vector θ shows as the Bayesian inference considers the parameters of the statistical distribution not fixed but variable. The Eq.B2.1 shows the posterior distribution of parameters. It depends on the likelihood formulation $P(D|\theta)$, on the marginal likelihood $P(D)$ and on the prior distribution $P(\theta)$. This last contribution represents the belief that the expert has on the events without considering observations of the same event.

The posterior distribution enables the computation of the predictive distribution of a new event X based on the past estimations. In addition, credible intervals can be estimated through the range of vector θ of parameters. These variables can be associated to return periods and a return level plot can be produced.

Bayesian framework is widely used in scientific literature especially when historical data are considered in the statistical analysis (Gaume et al., 2010; Payrastre et al., 2011; Nguyen et al., 2014; Bulteau et al., 2015). Some authors assume that this statistical context is the most suitable to treat the various uncertainties linked to historical data (Reis and Stedinger, 2005; Coles and Tawn, 2005; Bulteau et al., 2015). In addition, if a priori information on historical events is known by the expert, this can be used in the analysis to improve the statistical estimations.

Bayesian probabilities are compared to frequentist estimations always paying particular attention to the different meanings of both probabilities. A proper comparison between the two statistical frameworks is a difficult topic because the concept of probability is differently interpreted. In particular, frequentist probability is considered as an aleatory probability that provides the occurrence's relative frequency of an event by the knowledge of repetitions of the same process performed under similar conditions. In this study, frequentist probabilities refer to the occurrence's frequencies of extreme events. On the contrary, Bayesian probability is linked to a degree of belief that you can have on the occurrence of a particular event. For this reason, Bayesian probabilities refer to the knowledge that the expert has on the event.

Finally, frequentist and Bayesian context reply to different questions about probabilities and, for this reason, they estimate extreme values that are associated to different concepts of probability.

The choice to estimate extreme values by frequentist or Bayesian approaches has to be performed by the expert depending on the type of probabilities required.

2.1.1 Basic elements for Bayesian estimations

The meaning and the computation of basic Bayesian elements are detailed in the following.

Prior distribution

The prior distribution is a key element of a Bayesian statistical analysis. It represents the unconditional probability that an event might occur without considering any evidence originated from its observation. It is the prior knowledge that the expert has on the analysed phenomenon.

Prior distributions can be generated by different methods (Carlin and Louis, 2008) for any type of prior knowledge as, for instance, regional knowledge or knowledge on past events. The definition of a prior distribution is not tricky and, as Coles (1999) stated, inapt priors lead to several problems on estimations.

In addition, not ever a prior information is available. In this study, a type of prior called as non-informative has to be used in order to get Bayesian estimations also when no prior information is available.

Posterior distribution

The distribution of the vector θ of parameters of the statistical distribution is the posterior distribution. The posterior distribution is calculated by the formulation of Bayes' theorem (Eq. B2.1). This distribution is another important element of the Bayesian approach. In fact, the Bayesian probability considers the unknown parameters of the statistical distribution as a random variable. For this reason, their distribution has to be computed in the analysis.

The posterior distribution $P(\theta|D)$ represents the probability of the vector θ of parameters given the observed events D. This distribution is computed through the formulation of a prior distribution $P(\theta)$. The prior only depends on the prior knowledge of the analysed phenomenon and not depends on the observed data D. Lastly, the posterior distribution depends on the likelihood formulation $P(D|\theta)$ and on the marginal factor $P(D)$. The posterior distribution part formed by $P(D|\theta)/P(D)$ represent the impact that the observed data D has on the probability distribution of the vector θ of parameters.

Likelihood formulation $P(D|\theta)$ is the probability to observe the data D knowing the vector θ of parameters. When historical data are used in the analysis, the likelihood formulation can be defined as that of Eq.3.35.

Marginal likelihood $P(D)$ is a normalisation factor that is identic for each vector θ of parameters.

Finally, posterior distribution can be reformulated as follows:

$$P(\theta|D) \propto P(D|\theta)\, P(\theta) \tag{B2.2}$$

The computation of a posterior distribution can be performed by the use of Markov Chain Monte Carlo methods. The use of a MCMC algorithm allows the sampling of the posterior probability distribution without computing the marginal likelihood. A sufficient number of vectors θ can be estimated for some a particular number of chains. The convergence of the MCMC results has to be verified inside every chain and for different chains.

Finally, the posterior probability is affected by the variability of observed events. The knowledge of its uncertainties can be computed by the concepts of the Credibility Intervals.

Credibility Intervals

The Credibility Intervals are ranges of values that delineate a credible region in which exists a probability (corresponding to a p-value) to find the true value of the vector θ. Being posterior distribution derived from the prior distribution, also the Credibility Intervals consider within them the knowledge of the expert. They can be computed by the posterior distribution. In fact, a particular Credibility Interval is equal to the corresponding p-quantile of the vector θ of parameters.

The Credibility intervals are different to confidence intervals. In particular, Confidence Intervals are computed considering fixed the estimated parameters but varying the extreme data sample.

In this study, the Credibility Intervals are displayed in return level plots through the calculation of the corresponding regional quantiles x_{Tr} for each return period T.

Predictive distribution

The predictive distribution is generated by the vector θ of parameters and provides a single statistical distribution in which all the possible uncertainties are integrated. This distribution can be considered as the distribution of a new future event X^* that it is not yet observed. The probability to observe a new event X^* is computed as follows:

$$P(X^*|D) = \int P(X^*, \theta|D)\, d\theta = \int P(X^*|\theta, D)\, P(\theta|D)\, d\theta \qquad \text{(B2.3)}$$

This distribution is estimated through the mean of the considered distribution created for each vector θ of parameters.

In this study, the predictive distribution is figured out in return level plots through the computation of the probability $P(X^*>x_T^*)$ that a future event X^* will be higher than the predictive return level x_T^* linked to a predictive return period T^* as follows: $P(X^*>x_T^*)=1/\lambda T^*$

Annexe C

REGIONAL BAYESIAN ANALYSIS

3.1 Regional Bayesian analysis

A preliminary version of the FAB method for Bayesian estimations was presented by a talk in the Extreme Value Analysis and application to Natural Hazard (EVAN) conference held in Southampton (UK) between the 5th and the 7th September 2017.

Third international conference on Advances in Extreme Value Analysis and Application to Natural Hazard (EVAN), 5 to 7 September 2017, Southampton, UK

The use of archaeological data in the Regional Bayesian Analysis of extreme skew storm surges

Roberto Frau[1,2], Marc Andreewsky[1], Pietro Bernardara[3]

[1]EDF R&D LNHE, 6 Quai Watier, 78401 Chatou, France. E-mail: roberto.frau@edf.fr

[2]Laboratoire d'Hydraulique Saint-Venant, 6 Quai Watier, 78401 Chatou, France.

[3]EDF R&D UK Centre, 81-85 Station Rd, Croydon CR0 2AJ, UK.

The characterization of extreme sea levels and skew storm surges is a basic need to protect coastal Nuclear Fleets from flooding. Statistical methods linked to extreme value theory allow to estimate extreme events. The extreme events are associated to very high return periods (up to 10^3 years) and the estimation methodologies in the past were applied to a short time series (usually 30 years to 50 years long) at a given site. The uncertainties associated to these extreme event estimations are obviously huge and the scientific community has started to explore alternative ways to get robust quantiles.

A way to get reliable estimations is to look at the regional scale. In the era of big data, a lot of datasets, reanalysis and new earth observations are available and it's very important to exploit them. The idea of regional analysis is to put together data from different locations and to fit a regional distribution of a biggest data sample of extremes. This allows to increase the reliability estimation of the extreme variables probability occurrence. A Regional Frequency Analysis methodology had been already performed and applied for skew storm surges.

The use of archaeological data is another way to increase the amount of extremes available. A methodology to introduce archaeological skew storm surges in a Regional Frequency Analysis of digitalized data has been developed in a frequentist framework. Usually, the nature of the archaeological data is different and sometimes they are linked to measurement uncertainties that we can't ignore in a statistical analysis of extremes. The Bayesian framework is a suitable way to treat archaeological data.

The regional bayesian analysis is applied on a database of skew storm surges and on a several archaeological skew storm surges collected from different sources.

Annexe D

HISTORICAL DATA

4.1 Collection of historical skew surges

Third international conference on Advances in Extreme Value Analysis and Application to Natural Hazard (EVAN), 5 to 7 September 2017, Southampton, UK

Increasing skew surge database with the collection of archaeological data

Florian Regnier[1], Roberto Frau[1,2], Marc Andreewsky[1], Nathalie Giloy[3]

[1]EDF R&D LNHE, 6 Quai Watier, 78401 Chatou, France. E-mail: florian-f.regnier@edf.fr

[2]Laboratoire d'Hydraulique Saint-Venant, 6 Quai Watier, 78401 Chatou, France.

[3]IRSN, Institut de Radioprotection et de Sureté Nucléaire, 31 Avenue de la Division Leclerc, 92260 Fontenay-aux-Roses, France.

The knowledge of archaeological extreme events is important to estimate reliable extreme sea levels and skew storm surges. These estimations are associated to high return periods (up to 10^3 years) and usually they are computed only using tide gauge data. This leads to marked uncertainties in statistical estimation of extreme values. A good way to reduce these uncertainties is the use of archaeological data. In order to introduce archaeological data in an extreme analysis of digitalized marine variables, several statistical methods have already been developed by the scientific community. The application of these methodologies has to be performed using the maximum amount of data available. For this reason, the collection of archaeological extreme data is an important need.

However, gathering archaeological marine information is not an easy task. Historical sea levels or skew surges can be found in several sources such as historical archives and old newspapers and their quality has to be checked. Nevertheless, for some events, the archaeological skew storm surge values can be quantified at a single site, in particular, when the maximum sea level is available.

A collection of archaeological events connected to stronger storms of the past has been done for the French and British coasts and a preliminary database has been created. Mainly, these events have generated widespread coastal flooding due to the combination of high astronomical tide and extreme skew storm surge. The biggest skew storm surges are linked to specific meteorological conditions that are available through historical meteorological observations and modern reanalysis datasets. This allows us also to carry out a meteorological analysis for every archaeological event detected. These analyses will be applied to an outstanding event impacted Brittany and Bay of Biscay roughly one century ago.

4.2 Simulations of extreme skew surges

Improving simulations of extreme skew surges through waves' contributions

Cécile LALANNE, Roberto FRAU, Marc ANDREEVSKY, Chi-Tuan PHAM
Laboratoire National d'Hydraulique et Environnement
EDF R&D
Chatou, France
cecilemarion.lalanne@gmail.com
roberto.frau@edf.fr

Vanessya LABORIE
Laboratoire d'Hydraulique Saint-Venant
Cerema EMF
Chatou, France

Cécile RAOULT
Laboratoire d'Hydraulique Saint-Venant
Chatou, France

Nathalie GILOY
IRSN
Fontenay-aux-roses, France

***Abstract*—The coastal flood risk assessment is an overriding priority for EDF to ensure the nuclear safety. For this reason, statistical methods linked to Extreme Value Theory (EVT) are carried out to evaluate extreme events associated to high return periods (up to 10^3 years). Usually, these evaluations are applied to time series from 30 to 50 years and extreme estimations are not very accurate. A potential way to improve statistical estimations of extreme events is the use of historical data ([6], [7], [4]). Before to properly use them in a statistical analysis, the validation of historical records is needed.**

Numerical models may be complementary to historical values and they may even validate historical values recovered and reconstructed from several sources. Firstly, it is necessary to achieve a deep examination of the numerical models during several well-known extreme events in order to be able to validate historical events. In this study, extreme sea levels and, in particular, extreme skew surges simulated by a TELEMAC-2D model are considered.

TELEMAC-2D allows to simulate free-surface flows in two dimensions and to compute sea levels taking into account meteorological conditions during a storm. Unfortunately, not considering waves' contributions in simulations ([15], [14]) leads to non-accurate results. Waves' contributions can represent a significant part of skew surge [21]. In the present work, waves' contributions are taken into account in the computation of the surface drag coefficient C_D, using the Charnock relation, and the consideration of wave stresses. A sensitivity analysis of the Charnock coefficient is studied to find an optimal value.
Extreme skew surges are computed from simulations and these values are compared to measurements. Better results are obtained considering waves' contributions.
The model is tested for three of the well-known storms that impacted French coasts in 1987, 1999 and 2010, respectively The Great Storm of 1987, Lothar-Martin storms and Xynthia storm.

I. INTRODUCTION

The safety of nuclear power plants located along the coasts is one of the main priorities for EDF. Indeed, due to their proximity to the sea, coastal nuclear stations are subjected to the aggressions of extreme meteo-oceanic conditions such as sea levels, surges and waves. It is crucial to provide an accurate coastal risk assessment in order to be able to design effective protections. As part of the prevention of risks, numerical models allow to simulate storm events to study the different physical variables and processes involved. In this context, a lot of effort has been spent to improve simulations of extreme sea levels. The model has to be suitable for extreme events and effective at representing skew surges and in particular the maximum skew surge, our variable of interest in this study. The skew surge is the difference between the maximum observed sea level and the maximum predicted astronomical tide level during a tidal cycle ([22], [23], [6]). The risk of coastal flooding is bigger at high water conditions and justifies working with the maximum skew surge. Skew surge time series at several locations along French and British coasts can be obtained with the model.

At the Saint-Venant Hydraulics Laboratory (LHSV), a surge numerical model based on TELEMAC-2D software was built a few years ago [15] and then globally validated with additional tests [14]. The model showed relatively bad performances for the estimation of maximum skew surges along some regions such as Pays de la Loire or Nouvelle-Aquitaine. Waves' contributions had not been taken into account yet in [14] and at least for this reason, skew surges may have been underestimated for most of the study sites along the French coastline. Storm surges are generated by the meteorological forcing, in particular wind and pressure [8], and also by the waves. Waves' contributions can be divided into three components [17]: sea surface drag coefficient modification with the nature of waves, bottom friction and wave set-up. The positive relevance to use wave set-up and atmospheric effects in simulations, for instance, through a better surface drag parameterization, has been shown as part of the Previmer-Surcotes project [13].

The aim of this study is to improve the performances of the TELEMAC-2D model in South of the North Sea, English sea, and Biscay Bay and to provide the best simulated skew surge during an extreme event. As a first step, satisfactory results for maximum skew surges for some recent and well-known storms are expected. For this reason, a comparison between observed skew surges recorded by tide gauges and simulated skew surges has to be done in order to verify the numerical model. Finally, the model may be used to validate historical skew surges. Since historical data can be associated with considerable uncertainties, simulations generated by a reliable model can help us to determine if these skew surges likely happened in the past and so if they should be taken into account in the statistical of extreme events or not.

This paper presents the implementation of the waves' contributions in the TELEMAC-2D surge model through Charnock formulation and wave stresses. In addition, a validation part with three well-known storms The Great storm of 1987, Lothar-Martin (1999) and Xynthia (2010) is carried out. All the physical processes involved and their modelling are fully described in Sect.2. Sect.3 presents the results for the estimation of the maximum skew surges for each storm in different sites along the French coasts.

II. NUMERICAL MODEL AND SIMULATIONS

In this study, TELEMAC-2D (T2D) solves the Shallow Water Equations and some user FORTRAN sub-routines (for instance, prosou.f) are adapted to simulate skew storm surges. The numerical model is based on the one of [14] but a sub-routine has been changed and some input data have been added in order to consider waves' contributions. The TELEMAC-2D model extends from 10°W to 14°E and from 42°N to 64°N and includes French and British coasts (Fig. 2). The mesh (called mesh 2 in Fig. 3) is unstructured: it is particularly refined near the coastline, with one node per kilometer. Off the French coasts, the greatest distance between two nodes is around 40 km. The bathymetry "North East Atlantic Europe" (NEA) provided by the LEGOS is used. The data base for the harmonic constants is provided by the LEGOS [11] atlas to be consistent with the bathymetry. Initial water level and tidal currents are computed from the Atlantic Ocean solution of TPXO [12] database by OSU. The bottom friction is parametrized by the Chézy formulation with a constant coefficient of 70 $m^{1/2}/s$.

The meteorological forcing is provided by The National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) [20]. In our study, mean level atmospheric pressure at the sea level and horizontal components of wind (at 10 m) are used. Selected hourly time-series variables are available from January 1979 to December 2010. Besides a great temporal resolution, the fine spatial resolution (0.301°×0.301°) is necessary to represent precisely the atmospheric phenomena. Using a Python program, CFSR data are interpolated and a single SELAFIN file containing pressures and wind velocities data is obtained. To compute simulated skew surges, two simulations are achieved (Fig. 3): the first with meteorological forcing, the second one without (only tide propagation is used). Tidal simulations have been validated previously for several French harbours [14]. However, for some sites, an error up to 30 cm has been found during high tide. In our study, skew surges are considered and particularly the maximum skew surge as extreme values are sought. Substracting maximum predicted astronomical tide level to maximum observed water levels, potentially occurring with a time lag, leads to skew surge levels. The results are compared to those observed by the French Navy Hydrographic and Oceanographic (SHOM). For each storm event, a simulation, beginning seven days before the date of the storm and ending four days after, is run. The simulation time step is 30 s, according to [15].

III. IMPROVEMENT ON EXTREME EVENTS SIMULATIONS

The quality of a storm surge model depends on the accuracy of the input data, being the meteorological forcing, the spatial and temporal resolution and also the physical processes modelled. Storm surges were not properly modelled so far because at least waves' contributions were not taken into account: only the tide and the surge induced by the atmospheric forcing were integrated in the model. In order to improve skew surges estimations using waves' contributions in our model, the parametrization of the sea surface drag coefficient has to be firstly modified. This allows to describe more precisely the air-sea interaction. Secondly, wave stresses have to be considered during the simulations.

A. Sea surface drag coefficient

The wind influence is represented by a dimensionless sea surface drag coefficient C_D. This coefficient can be calculated with several formulations and most of them depend on the wind magnitude velocity at 10 m, U_N. C_D models complex phenomena. In fact, the wind influence depends on U_N but also on the roughness of the sea surface, which is itself dependent on the wind and the distance over which it is applied (fetch) [10].

In TELEMAC-2D, the wind influence is represented by the following formulation of Flather (Fig. 1):

$$C_D = 0.565 \times 10^{-3} if U_N \leqslant 5\, m/s$$

$$C_D = (-0.12 + 0.137 U_N) \times 10^{-3} if 5\, m/s \leqslant U_N \leqslant 19.22\, m/s$$

$$C_D = 2.513 \times 10^{-3} if U_N \geqslant 19.22\, m/s$$

With this formulation the coefficient only depends on U_N, whereas the wind influence may also depends on the roughness of the sea surface induced by the waves (characterized by the sea state). Charnock formulation suggests that the roughness length z_0 of the wind profile depends on the kinematic viscosity ν in the case of weak wind or on the Charnock relation (1) in the case of strong wind (above 20 m/s), for instance during a storm [9]:

$$z_0 = (\alpha_{CH} U_{STAR}^2)/g \quad (1)$$

where α_{CH} is the dimensionless Charnock coefficient; U_{STAR}, defined by $U_N/25$ [9], is the friction velocity (m/s) and g is the gravitational acceleration (m/s²). z_0 is linked to the sea surface drag coefficient C_D according to the following relation (2):

$$C_D = \kappa^2 log(z/z_0)^{-2} \quad (2)$$

κ=0.4 is the Von Karman constant and z is the altitude (m).

The Charnock coefficient models the surface roughness of the ocean and varies in time and space. α_{CH} (usually between 0.01 and 0.04, 0.018 is a typical value) depends on the sea state and on the wage age [24]. A wave model should be used to obtain a Charnock coefficient which takes into account the sea state. For example, WaveWatchIII gives α_{CH} from 1990 to 2018, based on CFSR or ECMWF reanalysis, and those data can be read in TELEMAC-2D. The consideration of waves' contributions through this database allows to improve the estimation of surges [18]. For the purpose of studying historical storms, a database for the Charnock coefficient that goes back further in the past is needed. The spectral wave model used at the LNHE, TOMAWAC, does not allow the computation of α_{CH} for the moment. It would require some developments that is why, as a first step, the formulation of Charnock has been implemented in TELEMAC-2D with a α_{CH} as a parameter fixed by the user and thus constant in time and space. The Charnock formulation gives more flexibility for the range of value of the drag coefficient. Higher values can be reached for the higher wind speed (increasing α_{CH}) in comparison with the formulation of Flather (Fig. 1). Thus, the Charnock coefficient can be used to strengthen, or not, the wind influence, depending on the value of α_{CH}. However, recent studies ([19], [5]) have shown that for winds greater than 33 m/s, the drag coefficient starts to decrease (Fig. 1). Hence, the Charnock formulation is not correct anymore and other formulations like Makin [16] should be used instead. In this paper, the maximum wind measured during the three considered storms is below 33 m/s so Charnock formulation has been kept.

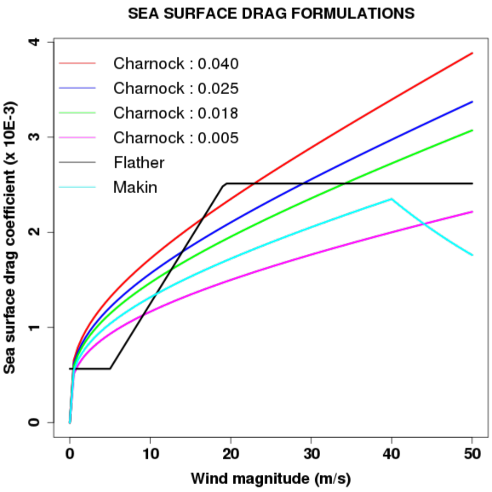


Figure 1. *Comparison between various formulations for the sea drag coefficient C_D and analysis of the influence of the Charnock parameter on this coefficient.*

Given that the performances of TELEMAC-2D were not homogeneous along the French coastline [14], a regional division (only determined by the latitude) based on French geographical areas is carried out. It is a first approach which has to be improved. Thus, four regions have been defined (Fig. 2): Hauts-de-France/Normandy, Brittany, Pays de la Loire and Nouvelle-Aquitaine. For each area, a different α_{CH} is applied, more appropriate locally, waiting to be able to calculate α_{CH} for each point of the mesh considering the sea state. The values for the Charnock coefficient have been chosen after several tests, depending on the results of our TELEMAC-2D model with the Flather formulation (if the maximum skew surge simulated by [14] was under the SHOM maximum skew surge, a high coefficient is fixed and conversely). In Sect. 3, details will be given about the α_{CH} used for each storm.

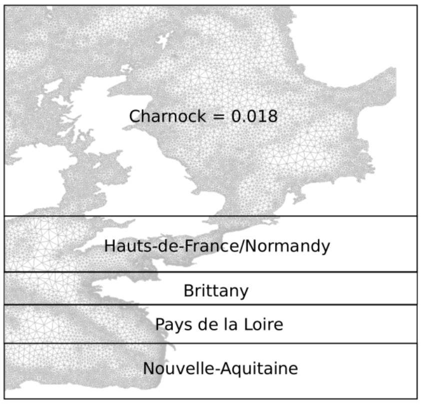


Figure 2. *Regional division for the adaptation of α_{CH}.*

B. Wave stresses

As TELEMAC-2D is used to simulate skew surges, waves are not taken into account. However, waves induce currents which may impact the surge and this effect can represent a significant part of the surge [21]. Those wave driven currents are calculated in TOMAWAC in the form of two forces F_u and F_v, called the wave stresses. The TOMAWAC software models wave propagation in coastal areas and estimates the mean characteristics of waves (water depth, direction, frequency). TELEMAC-2D is designed to be coupled with TOMAWAC but this requires to build a wave model on the same mesh as the one used for TELEMAC-2D (called mesh 2 in Fig. 3) with the determination of boundary conditions. Thus, for a first test of using wave stresses in the model, the data were taken from another project where a wave model is run with varying water level and currents due to tide (steps 1, 2 and 3 in Fig. 3). The same forcing conditions are used, but the computational domain is smaller and limited to close to the coast (called mesh 1 in Fig. 3). If the results are promising, a "real" 2-way-coupling will be implemented. With F_u and F_v as input data in

TELEMAC-2D, simulations with the contributions of wave induced currents are realized (step 4 in Fig. 3).

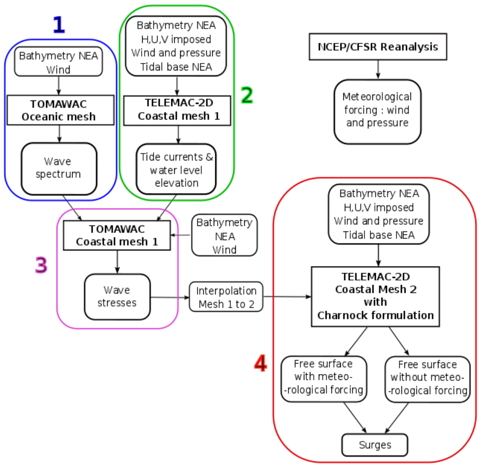


Figure 3. *Diagram of the chaining methodology to simulate surges.*

IV. Results

A. Xynthia

Xynthia is a recent well-known storm for which the SHOM collected data in plenty of ports. This case study served to calibrate our TELEMAC-2D surge model and also to estimate the contributions linked to the Charnock formulation or the wave stresses.

Xynthia was a violent storm which crossed rapidly Western Europe between the 27th of February and the 1st of March 2010. The trajectory of the storm was quite unusual, from South-West to North-East and created a particular sea state in the Bay of Biscay [1]. The waves were really short and arched. This induced the effect of increasing the sea roughness and so the drag coefficient [2]. To model this phenomenon, a Charnock coefficient of 0.04 is applied in the region of Pays de la Loire and 0.018, the typical value, everywhere else. Nine harbours are concerned: Dunkerque, Dieppe, Le Havre, Saint-Malo, Roscoff, Saint-Nazaire, La Rochelle, Port-Bloc and Boucau. The results of the TELEMAC-2D model with or without waves' contributions are compared to the SHOM observations. For all the study sites, the maximum skew surge was underestimated by the model. Nevertheless, using the Charnock formulation rather than the Flather one (Fig. 1) permitted to reduce the error between the peak of the simulated skew surge and the peak of the observed skew surge (Table 1).

The wave stresses do not have positive influences on our results, except for Le Havre. The performances of the TELEMAC-2D model are still not homogenous between all harbours: for instance, at Port-Bloc, the correct numerical value for the peak of skew surge is simulated, whereas at Saint-Nazaire, it is clearly overestimated (Fig. 4). Further tests should be conducted with a lower value of α_{CH} in Pays de la Loire to approach the maximum skew surge recorded by the SHOM. At Boucau, regardless of the modifications of the model, the same result is obtained. We will see with the other storms that the region of Nouvelle-Aquitaine shows low sensitivity to the model parameters in general. The regional division should be modified: working with smaller regions could help to describe local effects.

Table 1: Results of absolute relative error for the 9 sites for the maximum skew surges during Xynthia

Harbour	**Absolute relative error for the peak between the TELEMAC-2D model and the SHOM observations (%)**			
	Without waves' contri-butions	***With Charnock formulation only***	***With wave stresses only***	***With waves' contri-butions***
Dunkerque	16.05	1.23	17.28	1.23
Dieppe	12.63	2.11	14.74	1.05
Le Havre	35.64	25.74	7.92	1.98
Saint-Malo	16.47	4.71	17.65	4.71
Roscoff	31.67	28.33	33.33	30.0
Saint-Nazaire	19.81	26.42	33.96	24.52
La Rochelle	47.06	12.42	44.44	13.07
Port-Bloc	23.15	0.00	24.07	0.93
Boucau	43.90	39.04	46.34	41.46

In conclusion, taking into account the wind influence, through Charnock formulation, and the wave stresses helps to improve the estimation of the maximum skew surge for all sites for the Xynthia storm. To improve the results, the change of bathymetry database and the mesh refinement are prominent possibilities to take into account for future improvements. Of course, those promising results will lead to a complete coupling between TOMAWAC and TELEMAC-2D. The calibration of α_{CH} has to be refined eventually with a calculation directly in TOMAWAC.

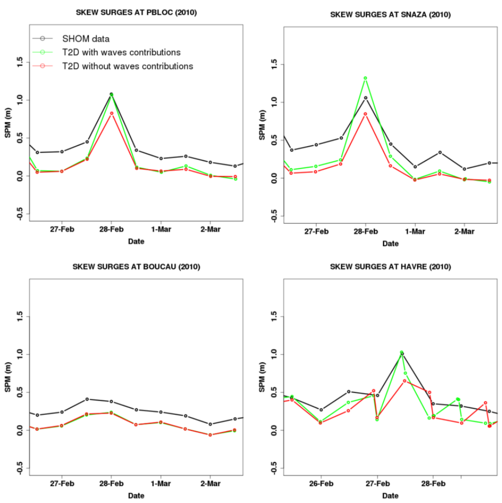


Figure 4. *Comparison between simulated skew surge (SPM) with (in green) or without (in red) the waves 'contributions with data recorded from tide gauge station (in black) during Xynthia storm.*

B. *Lothar-Martin*

Storm Lothar crossed Europe following a West-East track and peaked during the high tide of a moderate tidal range. It occurred on the December 26th, 1999. Less than 36 hours later, a second storm, called Martin, crossed France, a little further south, and affected almost of the same sites. This is quite unusual and during the tests of [15], the 1999 events were not correctly represented by the TELEMAC-2D model. Five tide gauges recorded the water level during both Lothar and Martin: Boucau, Cherbourg, Le Havre, Roscoff and Saint-Nazaire. La Rochelle tide gauge was not operating during those storms because of a general power failure. [3] simulated a skew surge value of 2.17 m for December 27th for storm Martin at La Rochelle so our results are compared with it (Fig. 5).

After some tests, the following values for the Charnock coefficient were chosen:

- 0.001 for Hauts-de-France/Normandy and Brittany,
- 0.04 for Pays de la Loire and Nouvelle-Aquitaine.

Indeed, the model used in [14] overestimated the peak of the skew surge in northern France, so a very small α_{CH} is used to reduce the wind influence and conversely for the South of France. For Cherbourg, Le Havre, Saint-Nazaire and Boucau, we manage to improve the results of the TELEMAC-2D model through waves' contributions (Fig. 5) but the numerical value of the maximum skew surge cannot be validated, except at Cherbourg. Finally, for La Rochelle, the waves' contributions lead to two skew surge peaks rather than three (Fig. 5). It could be more coherent as there is two really close storms but the simulated values are still far from measurements and the temporal occurrence is not quite exact.

To conclude, in this case, the implementation of the waves' contributions does not allow our model to describe correctly the 1999 events in all harbours. Results have been enhanced for some sites which encourages us to continue our work. As for storm Xynthia, a bathymetry and a mesh with a better resolution should have a benefit on our skew surge estimations as a precision of the geographical regions.

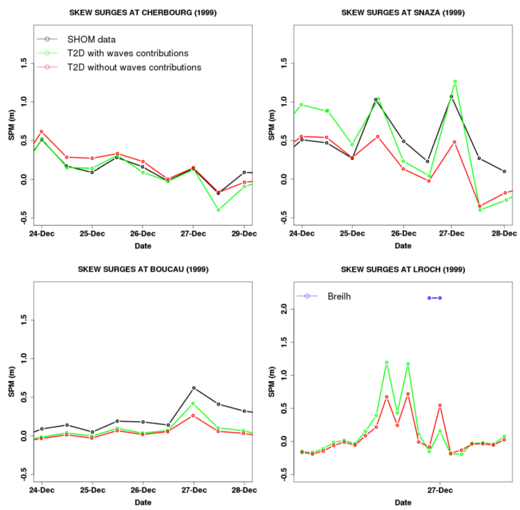


Figure 5. *Comparison between simulated skew surge (SPM) with (in green) or without (in red) the waves 'contributions with data recorded from tide gauge station (in black) during Lothar-Martin storms. Comparison with [3] (in blue) for La Rochelle.*

C. *The Great Storm of 1987*

This storm occurred in the middle of October 1987: a depression originated on the Bay of Biscay on the 15th and moved North-East. The Great Storm of 1987 impacted Brittany and then England. Eight French tide gauges recorded the sea level during this event: Dieppe, Le Havre, Cherbourg, Roscoff, Le Conquet, Port-Tudy, Verdon and Saint-Jean-de-Luz. For this study, we choose α_{CH} = 0.04 for Hauts-de-France/Normandy, α_{CH} = 0.35 for Brittany and α_{CH}= 0.018 for the other regions as the storm mainly affected the North of France.

Estimations of the maximum skew surge are improved only for six harbours. In fact, this storm does not strongly impact the sites of Le Verdon and Saint-Jean-de-Luz in which time series of skew surges are available. In addition, the Nouvelle Aquitaine region, to which these two sites belong, is poorly sensitive to the parameters of the TELEMAC-2D model. Results at Cherbourg and Roscoff (Fig. 6) allow us to get few ameliorations for the maximum skew surge. On the contrary, for Le Havre and for Port-Tudy (Fig. 6), the waves' contributions have a clear positive influence.

This case study needs a careful work especially for the regions of Hauts-de-France/Normandy and Brittany where the storm had the strongest impact. As the Great Storm of 1987 affected the English coasts too, skew surges simulations should be done for British harbours. As for the 2010 and 1999 storm events, the TELEMAC-2D model should be enhanced with more refined bathymetry and mesh. In addition, a coupling with TOMAWAC could be considered, rather than a chaining.

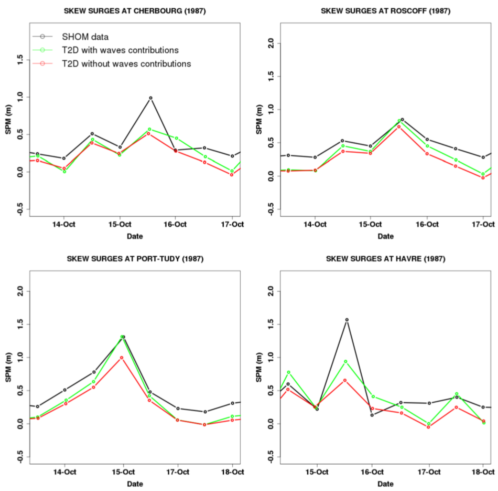


Figure 6. *Comparison between simulated skew surge (SPM) with (in green) or without (in red) the waves 'contributions with data recorded from tide gauge station (in black) during The Great Storm of 1987.*

Conclusions and Perspectives

The storm surges model based on TELEMAC-2D built and validated a few years ago ([15], [14]) has been improved through the implementation of the waves' contributions. The formulation for the sea surface drag coefficient which translates the wind influence has been modified with the Charnock formulation. A regional division has been settled to affect a particular Charnock coefficient for each area. In addition, the wave stresses are now taken into account in our simulation thanks to a chaining with TOMAWAC. For the three storms studied, an improvement, nevertheless sometime small, of our estimations of the maximum skew surge is observed in most of the sites. The examination of the TELEMAC-2D model for several well-known storms is essential to be able to study extreme historical events later and thus validate historical values.

Work is still in progress at the LNHE. A new bathymetry from the SHOM with a resolution of 100 m should be tested and a new mesh will be soon developed. Indeed, all tide gauges are located in ports so there are influenced by local effects. A coupling between TOMAWAC and TELEMAC-2D could be considered as a promising way to still improve results. The Charnock formulation is valid for winds below 33 m/s, we may change for the Makin formulation [16] for other storms. Moreover, the geographic division has to be precise and the Charnock coefficient needs to be calculated for each node of the mesh, updated at each time step. This could be possible with the calculation of the coefficient directly in TOMAWAC. One advantage will be that our model would not be dependent anymore on a database such IOWAGA from WaveWatchIII and therefore it will ensure coherence between all the data used in our storm surges simulations. In addition, British ports should be studied to complete this work, especially for The Great Storm of 1987.

References

[1] Bertin, X., Bruneau, N., Breilh, J. F., Fortunato, A. B., & Karpytchev, M. (2012). Importance of wave age and resonance in storm surges: The case Xynthia, Bay of Biscay. Ocean Modelling, 42, 16-30.

[2] Bertin, X., Li, K., Roland, A., Breilh, J. F., & Chaumillon, E. (2012). Contributions des vagues dans la surcote associée à la tempête Xynthia, février 2010. XIIèmes Journées Nationales Génie Côtier–Génie Civil.

[3] Breilh, J. F., Bertin, X., Chaumillon, É., Giloy, N., & Sauzeau, T. (2014). How frequent is storm-induced flooding in the central part of the Bay of Biscay?. Global and Planetary change, 122, 161-175.

[4] Bulteau, T., Idier, D., Lambert, J., & Garcin, M. (2015). 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, 15(6), 1135-1147.

[5] Donelan, M. A., et al. (2004). On the limiting aerodynamic roughness of the ocean in very strong winds. Geophysical Research Letters, 31(18).

[6] Frau, R., Andreewsky, M., & Bernardara, P. (2018). The use of historical information for regional frequency analysis of extreme skew surge. Natural Hazards and Earth System Sciences, 18(3), 949-962.

[7] Hamdi, Y., Bardet, L., Duluc, C. M., & Rebour, V. (2015). Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France. Natural Hazards and Earth System Sciences, 15(7), 1515-1531.

[8] Heaps, N. S. (1983). Storm surges, 1967–1982. *Geophysical Journal of the Royal Astronomical Society*, *74*(1), 331-376.

[9] Hersbach, H. (2011). Sea Surface Roughness and Drag Coefficient as Functions of Neutral Wind Speed. Journal of Physical Oceanography, 41(1), 247-251.

[10] Hervouet, J. M. (2007). Hydrodynamics of Free Surface Flows: modelling with the finite element method. John Wiley & Sons.

[11] http://sirocco.omp.obs-mip.fr/outils/Tugo/Produits/TugoProduits.htm

[12] http://volkov.oce.orst.edu/tides/

[13] Idier, D., et al. (2012). Modélisation des surcotes sur les côtes françaises, Manche et Atlantique. In Congrès SHF: Evènements extrêmes fluviaux et maritimes).

[14] Laborie, V., Sergent, P., Levy, F., Frau, R., & Weiss, J. (2015). The hydrodynamic, sea-state and infrastructures platform developed by Saint-Venant Hydraulics Laboratory and Cerema: a special focus on the TELEMAC2D surge levels numerical model of the Atlantic Ocean, the Channel and the North Sea. In 22nd TELEMAC-MASCARET User Conference 2015 (pp. 172-181).

[15] Levy, F. "Modélisation des surcotes avec Telemac 2D", internal report, 24 pages, April 2013

[16] Muller, H., Pineau-Guillou, L., Idier, D., & Ardhuin, F. (2014). Atmospheric storm surge modeling methodology along the French (Atlantic and English Channel) coast. Ocean Dynamics, 64(11), 1671-1692.

[17] Nicolle, A., Karpytchev, M., & Benoit, M. (2009). Amplification of the storm surges in shallow waters of the Pertuis Charentais (Bay of Biscay, France). Ocean Dynamics, 59(6), 921-935.

[18] Pineau-Guillou, L., Theetten, S., Dumas, F., Lecornu, F., & Idier, D. (2012, June). Prévision opérationnelle des niveaux de la mer, surcotes et décotes sur les côtes de la Manche et de l'Atlantique. In XIIèmes Journées Nationales Génie Côtier–Génie Civil, Cherbourg, 12-14 juin 2012.

[19] Powell, M. D., Vickery, P. J., & Reinhold, T. A. (2003). Reduced drag coefficient for high wind speeds in tropical cyclones. Nature, 422(6929), 279-283.

[20] Saha, S., et al. (2010). The NCEP climate forecast system reanalysis. Bulletin of the American Meteorological Society, 91(8), 1015-1058.

[21] Sheng, Y. P., Alymov, V., & Paramygin, V. A. (2010). Simulation of storm surge, wave, currents, and inundation in the Outer Banks and Chesapeake Bay during Hurricane Isabel in 2003: The importance of waves. Journal of Geophysical Research: Oceans, 115(C4).

[22] Simon, B., & Gonella, J. (2007). La marée océanique côtière. Institut océanographique.

[23] Weiss, J., Bernardara, P., Andreewsky, M., & Benoit, M. (2012). Seasonal autoregressive modeling of a skew storm surge series. Ocean Modelling, 47, 41-54.

[24] Zhou, L., Wang, A., Guo, P., & Wang, Z. (2008). Effect of surface waves on air–sea momentum flux in high wind conditions for typhoons in the South China Sea. Progress in Natural Science, 18(9), 1107-1113.

Annexe E

SUMMARY TABLE OF RESULTS

5.1 Summary table of the FAB application

Numerical results of the FAB application (Chapter 4) are here summed up as follows:

	Region 1	La Rochelle	Region 2	Calais
Quantiles for T=1000ys with Historical Data (Frequentist Estimations)	**2.89**	**1.82m**	**3.30**	**1.78m**
Quantiles for T=1000ys without HD (FE)	**2.55**	**1.61m**	**2.94**	**1.62m**
Predictive quantiles for T=1000ys with HD (Bayesian Estimations)	**3.14**	**1.97m**	**3.51**	**1.89m**
Upper CI 90% for T=1000ys with HD (FE)	**3.39**	**2.15m**	**3.99**	**2.14m**
Upper CI 90% for T=1000ys without HD (FE)	**2.90**	**1.84m**	**3.45**	**1.90m**
Upper CI 90% for T=1000ys with HD and without seasonality (FE)	**3.59**	**2.28m**	**4.01**	**2.15m**
Upper CI 90% for T=1000ys with HD (BE)	**3.91**	**2.47m**	**4.56**	**2.46m**
ΔCI/xT =1000ys with HD (FE)	**34%**	**36%**	**40%**	**40%**
ΔCI/xT =1000ys without HD (FE)	**28%**	**29%**	**34%**	**35%**
ΔCI/xT =1000ys with HD and without seasonality (FE)	**37%**	**39%**	**40%**	**40%**
ΔCI/xT =1000ys with HD (BE)	**44%**	**44%**	**53%**	**53%**

Tab. 18 – Summary table of the FAB application's results

5.1 Summary table of the FAB application

This summary table allows an easier numerical comparison of the two different statistical frameworks analysed as well as the statistical analyses performed using or not historical data.

Annexe F

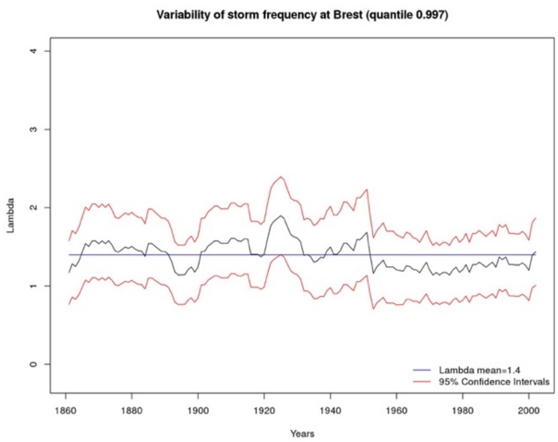
STORM FREQUENCY ANALYSIS

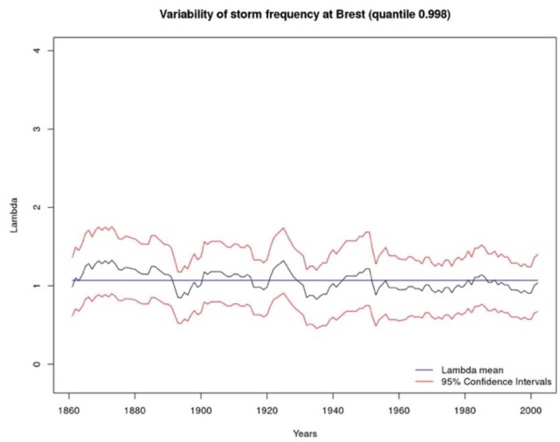
6.1 Storm frequency on the longest time series

The hypothesis of the lack of a trend on storm frequency during the 20th century, and in which the definition of the credible duration is based (Chapter 3), has been verified for skew surges through an additional test on the longest skew surge series. In fact, the availability of long skew surge time series allows the estimation of the lambda value for different periods of the 20th century. This computation can indicate if a trend on the storm frequency is present or not.

In particular, for every year y_t of recordings, a mean of lambda values is calculated considering a period of 15 years before the year y_t and 15 years after the same year y_t. The computation of lambda has been performed for different thresholds corresponding to the quantiles of 0.997, 0.998, 0.9985 and 0.999 of the considered skew surge serie. In addition, the 95% confidence intervals and a mean of lambda on the whole skew surge serie are computed.

This analysis has been performed on the longest French skew surge serie that has been recorded since the year 1846 at Brest for a total of 156.57 years and on the longest British skew surge serie that has been recorded since the year 1915 at Newlyn for a total of 87.32 years. The results of these two analyses are figured out below in Fig.42 and Fig.43.





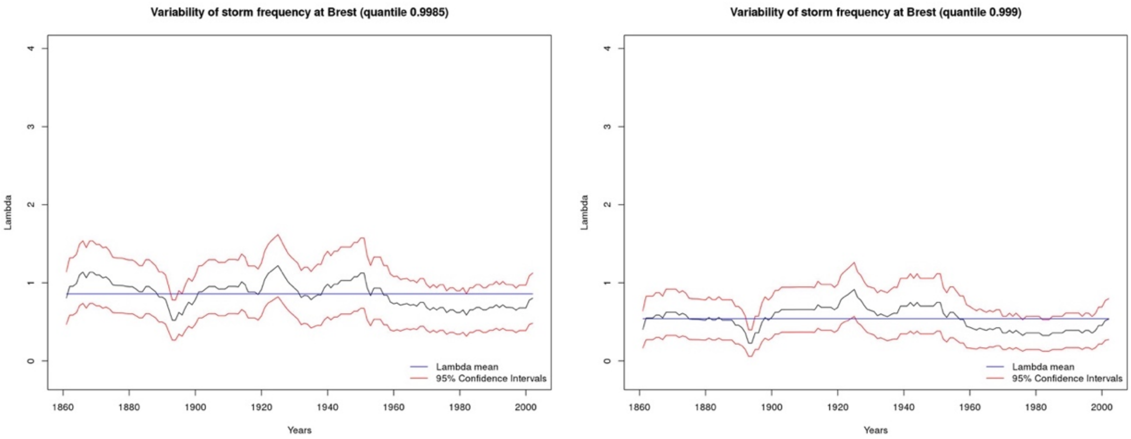


Fig. 42 – Variability of storm frequency for different thresholds corresponding to quantiles of 0.997, 0.998, 0.9985 and 0.999 at Brest tide gauge

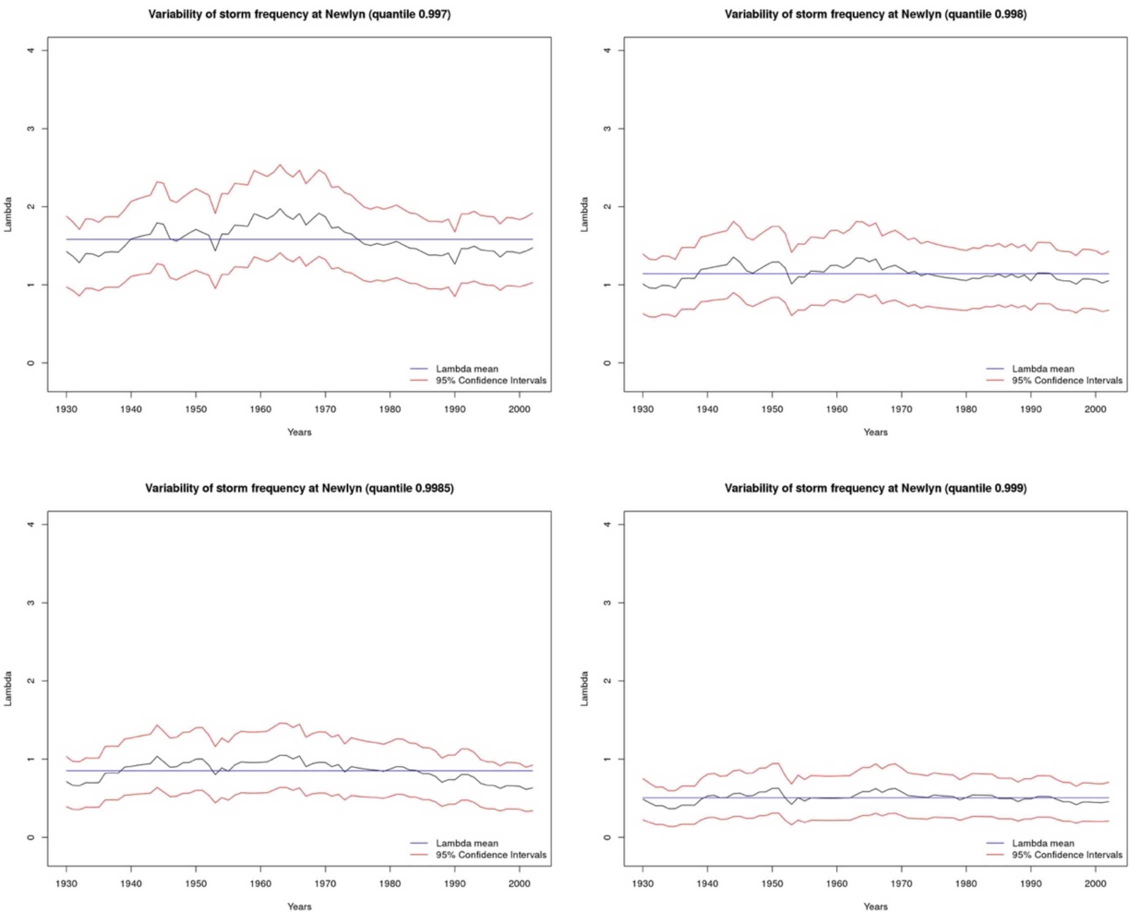


Fig. 43 - Variability of storm frequency for different thresholds corresponding to quantiles of 0.997, 0.998, 0.9985 and 0.999 at Newlyn tide gauge

The lambda values calculated for different periods of 31 years (the black line of Fig. 42 and Fig. 43) varies around the lambda mean both in Brest and in Newlyn. In addition, the lambda mean is most of the time inside the 95% confidence intervals.

These analyses in Brest and in Newlyn consolidate the credible hypothesis of the lack of a trend on storm frequency during the 20th century.

References

Alila, Y. (1999). A hierarchical approach for the regionalization of precipitation annual maxima in Canada. *Journal of Geophysical Research , 104* (D24), 31,645-31,655,.

Allan, R., Tett, S., & Alexander, L. (2009). Fluctuations in autumn-winter severe storms over the British Isles: 1920 to present. *International Journal of Climatology , 29* (3), 357-371.

ASN. (2013). *Guide pour la protection des installations nucléaires de base contre les inondations externes.* Guide n°13 de l'ASN .

Baart, F., Bakker, M. A., van Dongeren, A., den Heijer, C., van Heteren, S., Smit, M. W., et al. (2011). Using the 18th century storm-surge data from the Dutch Coast to improve the confidence in flood-risk estimates. *Natural Hazards and Hearth System Sciences , 11*, 2791-2801.

Bardet, L., Duluc, C.-M., Rebour, V., & L'Her, J. (2011). Regional frequency analysis of extreme storm surges along the French coast. *Natural Hazards and Earth System Sciences , 11* (6), 1627-1639.

Barnett, V., & Lewis, T. (1994). *Outliers in Statistical Data, 3rd Edition.* Chichester: John Wiley.

Barriendos, M., Coeur, D., Lang, M., Llasat, M., Naulet, R., Lemaitre, D., et al. (2003). Stationarity analysis of historical flood series in France and Spain (14th–20th centuries). *Natural Hazards and Earth System Sciences , 3* (6), 583-592.

Barring, L., & Fortuniak, K. (2009). Multi-indices analysis of southern Scandinavian storminess 1780-2005 and links to interdecadal variations in the NW Europe-North sea region. *International Journal of Climatology , 29* (3), 373-384.

Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances. *By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S.,* *Philosophical Transactions, Giving Some Account of the Present Undertakings, Studies and Labours of the Ingenious in Many Considerable Parts of the World , 53*, 370-418.

Beirlant, J., Goegebeur, U., Segers, J., & Teugels, J. L. (2004). *Statistics of Extremes : Theory and Applications.* Wiley.

Benito, G., Lang, M., Barriendos, M., Llasat, M. C., Frances, F., Ouarda, T., et al. (2004). Use of Systematic, Paleoflood and Historical Data for the Improvement of Flood Risk Estimation. Review of Scientific Methods. *Natural Hazards , 31* (3), 623-643.

Benson, M. A. (1950). Use of historical data in flood-frequency analysis. *Eos, Transactions American Geophysical Union , 31* (3), 419-424.

Bernardara, P., Andreewsky, M., & Benoit, M. (2011). Application of regional frequency analysis to the estimation of extreme storm surges. *Journal of Geophysical Research , 116* (C02008), 1-11.

Bernardara, P., Mazas, F., Kergadallan, X., & Hamm, L. (2014). A two-step framework for over-threshold modelling of environmental extremes. *Natural Hazards and Earth System Sciences , 14* (3), 635-647.

Borga, M., Vezzani, C., & Dalla Fontana, G. (2005). Regional rainfall depth-duration-frequency equations for an alpine region. *Natural Hazards , 36*, 221-235.

Brehil, J.-F. (2014). *Les surcotes et les submersions marines dans la partie centrale du Golfe de Gascogne : les enseignements de la tempete Xynthia.* Sciences de la Terre. Université de La Rochelle.

Breilh, J.-F., Bertin, X., Chaumillon, E., Giloy, N., & Sauzeau, T. (2014). How frequent is storm-induced flooding in the central part of the Bay of Biscay? *Global and Planetary Change , 122*, 161-175.

Brooks, S., & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics , 7* (4), 434-455.

Bulteau, T., Idier, D., Lambert, J., & Garcin, M. (2015). 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 , 15*, 1135-1147.

Carlin, B. P., & Louis, T. A. (2008). *Bayesian Methods for Data Analysis (Third ed.).* CRC Press.

Chernoff, H., & Lehmann, E. L. (1954). The Use of Maximum Likelihood Estimates in χ2 Tests for Goodness of Fit. *The Annals of Mathematical Statistics , 25* (3), 579-586.

Cochran, W. G. (1952). The χ2 Test of Goodness of Fit. *The Annals of Mathematical Statistics , 23* (3), 315-345.

Cohn, T. A. (1984). *The incorporation of historical information in flood frequency analysis.* Ithaca, N.Y.: M. S. thesis, Dep. of Environ. Eng., Cornell Univ.

Cohn, T. A., & Stedinger, J. R. (1987). Use of historical information in a maximum-likelihood framework. *Journal of Hydrology , 96* (1-4), 215-223.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* London: Springer-Verlag.

Coles, S., & Dixon, M. (1999). Likelihood-Based Inference for Extreme Value Models. *Extremes , 2* (1), 5-23.

Coles, S., & Tawn, J. (2005). Bayesian modelling of extreme surges on the UK east coast. *Philos Trans A Math Phys Eng Sci. , 363(1831):1387-406.*

Condie, R., & Lee, K. A. (1982). Flood frequency analysis with historic information. *Journal of Hydrology , 58* (1-2), 47-61.

Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. *Journal of Hydrology , 100* (1-3), 269-290.

Dalrymple, T. (1960). Flood-frequency analyses, Manual of Hydrology: Part 3. *U.S. Geological Survey Water-Supply Paper 1543-A* .

Daubord, C., André, G., Goirand, V., & Kerneis, M. (2015). Rapport technique final du projet NIVEXT. *Rapport d'étude Shom* , 448.

De Michele, C., & Rosso, R. (2001). Uncertainty assessment of regionalized flood frequency estimates. *J. Hydrol. Eng. , 6* (6), 453-459.

Di Baldassarre, G., Laio, F., & Montanari, A. (2009). Design flood estimation using model selection criteria. *Physics and Chemistry of the Earth , Parts A/B/C 34* (10-12), 606-611.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics , 7*, 1-26.

Escalante-Sandoval, C. (2008). Bivariate distribution with two component extreme value marginals to model extreme wind speed. *Atmosfera , 21* (4), 373-387.

Ferreira, O., Vousdoukas, M., & Ciavola, P. (2009). *Review of climate change impacts on storm occurrence.* MICORE Project 202798.

Ferret, Y. (2016). *Reconstruction de la série marégraphique de Saint-Nazaire.* Rapport d'étude Shom.

Fisher, R., & Tippett, L. (1928). On the estimation of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society , 24*, 180-190.

Frechét, M. (1928). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique , 6*, 93-122.

Garnier, E., & Surville, F. (2010). *La tempete Xynthia face à l'histoire.* Saintes, France: Croît vif.

Gaume, E., Gaal, L., Viglione, A., Szolgay, J., Kohnova, S., & Bloschl, G. (2010). Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites. *Journal of Hydrology* .

Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using mul- tiple sequences (with discussion). *Statistical Science , 7*, 457–511.

Giloy, N., Duluc, C. M., Frau, R., Ferret, Y., Bulteau, T., Mazas, F., et al. (2018, mai 29-31). La base de données TEMPETES : un support pour une expertise collégiale et interdisciplinaire des informations historiques de tempêtes et de submersions. *XVèmes Journées Nationales Génie Côtier* .

Giloy, N., Garnier, E., Hamdi, Y., & Duluc, C. M. (2017, September 5-7). Reconstructing historic surges in the Dunkirk Area, France. *3rd International Conference EVAN* .

Gnedenko, B. (1943). Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *Annals of Mathematics , 44* (3), 423-453.

Goda, Y. (2011). Plotting-position estimator for the l-moment method and quantile confidence interval for the gev, gpa, and weibull distributions applied for extreme wave analysis. *Coastal Engineering , 53* (2), 111–149.

Goda, Y., Kudaka, M., & Kawai, H. (2010). Incorporating of weibull distribution in l-moments method for regional frequency analysis of peak over threshold wave heights. *Proc. 32th Int. Conf. on Coastal Eng. (ICCE 2010, Shangai)* .

Goel, N. K., Burn, D. H., Pandey, M. D., & An, Y. (2004). Wind quantile estimation using a pooled frequency analysis approach. *Journal of wind engineering and industrial aerodynamics , 92* (6), 509–528.

Gouriou, T. (2012). *Evolution des composantes du niveau marin à partir d'observations de marégraphe effectuées depuis la fin du 18e siècle en Charente-Maritime.* Océanographie physique. Université de La Rochelle.

Grimshaw, S. D. (1993). Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution. *Technometrics , 35* (2), 185-191.

Gumbel, E. J. (1958). *Statistics of Extremes.* New York, U.S.: Columbia University Press.

Haigh, I. D., Ozsoy, O., Wadey, M. P., Nicholls, R. J., Gallop, S. L., Brown, T. W., et al. (2017). An improved database of coastal flooding in the United Kingdom from 1915 to 2016. *Sci. Data* , 4:170100.

Haigh, I. D., Wadey, M. P., Gallop, S. L., Loehr, H., Nicholls, R. J., Horsburgh, K., et al. (2015). A user-friendly database of coastal flooding in the United Kingdom from 1915–2014. *Sci. Data* , 2:150021.

Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., et al. (2016). Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. *Sci. Data* , 3:160107.

Hamdi, Y., Bardet, L., Duluc, C.-M., & Rebour, V. (17–22 April 2016). Development of a target-site based regional frequency model using. *European Geosciences Union General Assembly 2016 (EGU 2016)*, *18*, pp. EGU2016-8765. Vienna, Austria.

Hamdi, Y., Bardet, L., Duluc, C.-M., & Rebour, V. (2015). Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France. *Natural Hazards and Earth System Sciences , 15*, 1515-1531.

Hanna, E., Cappelen, J., Allan, R., Jonsson, T., Blancq, F. L., Lillington, T., et al. (2008). New Insights into North European and North Atlantic Surface Pressure Variability, Storminess, and Related Climatic Change since 1830. *Journal of Climate , 21*, 6739-6766.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., et al. (2013). Observations: Atmosphere and Surface. In T. D.-K. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika , 57* (1), 97-109.

Hosking, J. R., & Wallis, J. R. (1997). *Regional Frequency Analysis. An approach based on L-Moments.* Cambridge University Press.

Hosking, J. R., & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research , 29* (2), 271-281.

Hosking, J. (2012). Towards statistical modeling of tsunami occurrence with regional frequency analysis. *Journal of Math-for-Industry , 4* (2012A-6), 41-48.

Hosking, J., & Wallis, J. (1986a). Paleoflood Hydrology and Flood Frequency Analysis. *Water Resources Research , 22* (4), 543-550.

Hosking, J., & Wallis, J. (1986b). The value of historical data in flood frequency analysis. *Water Resources Research , 22* (11), 1606-1612.

Hubert, M., & Van der Veeken, S. (2008). Outlier detection for skewed data. *Journal of chemometrics , 22* (3-4), 235-246.

Javelle, P., Ouarda, T. B., Lang, M., Bobee, B., Galea, G., & Gresillon, J.-M. (2002). Development of regional flood-duration-frequency curves based on the index-flood method. *Journal of Hydrology , 258*, 249-259.

Jonathan, P., & Ewans, K. (2013). Statistical modelling of extreme ocean environ- ments for marine design: A review. . *Ocean Engineering , 62*, 91-109.

Jonathan, P., Ewans, K., & Forristall, G. (2008). Statistical estimation of ex- treme ocean environments: The requirement for modelling directionality and other covariate effects. *Ocean Engineering , 35* (11-12), 1211-1225.

Kergadallan, X. (2013). *Analyse statistique des niveaux d'eau extrêmes - Environnements maritime et estuarien.* Compiègne: CETMEF, Centre d'études techniques maritimes et fluviales.

Kjeldsen, T. R., Smithers, J. C., & Schulze, R. E. (2002). Regional frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. *J. Hydrol. , 255*, 194-211.

Laio, F., Di Baldassarre, G., & Montanari, A. (2009). Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research , 45* (7).

Lambert, J., & Garcin, M. (2013). *Submersions marines historiques.* BRGM/RP-62644-FR.

Lang, M., & Coeur, D. (2014). *Les inondations remarquables en France.* Editions Quae, Versailles.

Lang, M., Coeur, D., Audouard, A., Villanova-Oliver, M., & Pène, J. P. (2016). BDHI: a French national database on historical floods. *E3S Web of Conferences 7* .

Lang, M., Rasmussen, P., Oberlin, G., & Bobée, B. (1997). Echantillonnage par valeurs supérieures à un seuil: modélisation des occurrences par la méthode du renouvellement. *Revue des Sciences de l'Eau , 10*, 279–320.

Le Cornec, E., & Peeters, P. (2009). *Détermination de l'alea de submersion marine intégrant les conséquences du changement climatique en région Nord-Pas de Calais.*

Le Gorgeu, V., & Guitonneau, R. (1954). Reconstruction de la Digue de l'Est à Dunkerque. *Coast. Eng. , 5*, 555-586.

Leese, M. N. (1973). Use of censored data in the estimation of Gumbel distribution parameters for annual maximum flood series. *Water Resources Research , 9* (6), 1534-1542.

Madsen, H., & Rosbjerg, D. (1997). The partial duration series method in regional index-flood modeling. *Water Resources Research , 33* (4), 737-746.

Maspataud, A. (2011). *Impacts des tempetes sur la morphodynamique du profil cotier en milieu macrotidal.* France: Océanographie, Université du Littoral Cote d'Opale.

Matulla, C., Schöner, W., Alexandersson, H., von Storch, H., & Wang, X. L. (2008). European storminess: late nineteenth century to present. *Climate Dynamics , 31* (2-3), 125-130.

Méndez, F. J., Menéndez, M., Luceño, A., Medina, R., & Graham, N. E. (2008). Seasonality and duration in extreme value distributions of significant wave height. *Ocean Engineering , 35* (1), 131-138.

Merz, B., & Bloschl, G. (2003). A process typology of regional floods. *Water Resources Research , 39* (12), 1340.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys. , 21*, 1087.

Miquel, J. (1981). *Guide d'estimation des probabilités des débits de crue-Parties 1 et 2* (Vol. E43/81.45). EDF-Direction des études et recherches.

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: An evaluation. *Comput. J. , 20*, 359–363.

Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification , 31* (3), 274–295.

Nguyen, C. C., Gaume, E., & Payrastre, O. (2014). Regional flood frequency analyses involving extraordinary flood events at ungauged sites: further developments and validations. *Journal of Hydrology , 508*, 385-396.

Ouarda, T. B., Rasmussen, P., Bobée, B., & Bernier, J. (1998). Utilisation de l'information historique en analyse hydrologique fréquentielle. *Revue des sciences de l'eau , 11*, 41-49.

Ouarda, T., Lang, M., Bobée, B., Bernier, J., & Bois, P. (1999). Synthèse de modèles régionaux de crue utilisés en France et au Québec. *Revue des Sciences de l'Eau , 12*, 155-182.

Parent, P., Butin, T., Vanhee, S., & Busz, N. (2007). *Un territoire soumis au risque de submersion marine. Les tempetes de 1949 et 1953 à Dunkerque.* Retrieved from Institution départementale des Wateringues: http://www.ingeo.fr/fichier/file/Wateringues/Affiche4-Dunkerque.pdf

Payrastre, O., Gaume, E., & Andrieu, H. (2013). Historical information and flood frequency analyses: which optimal features for historical floods inventories? *La Houille Blanche - Revue internationale de l'eau , 3*, 5-11.

Payrastre, O., Gaume, E., & Andrieu, H. (2011). Usefulness of historical information for flood frequency analyses: Developments based on a case study. *Water Resources Research , 47*, W08511.

Peret, J. (2004). *Naufrages et pilleurs d'épaves sur les côtes charentaises auc XVII et XVIII siécles.* Geste Editions, La Crèche.

Peret, J., & Sauzeau, T. (2014). *Xynthia, ou la mémoire réveillée.* Geste Editions, La Crèche.

Picklands, J. (1975). Statistical inference using extreme order statistics. *Ann. Stat.* , 3(1), 119-131.

Pouvreau, N. (2008). *Trois cents ans de mesures marégraphiques en France : outils, méthodes et tendances des composantes du niveau de la mer au port de Brest.*

Prosdocimi, I. (2017). German tanks and historical records: the estimation of the time coverage of ungauged extreme events. *Stochastic Environmental Research and Risk Assessment* , 1-16.

Regnier, F., Frau, R., Andreevsky, M., & Giloy, N. (2017, September 5-7). Increasing skew surge database with the collection of archaeological data. *3rd International Conference EVAN* .

Reis, D. S., & Stendiger, J. R. (2005). Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology , 313* (1-2), 97-116.

Roche, A., Baraer, F., Le Cam, H., Madec, T., Gautier, S., Jan, G., et al. (2014, Juilliet 2-4). Projet VIMERS : une typologie des tempêtes bretonnes pour prévoir l'impact des tempêtes à venir et mieux s'y préparer. *XIIIèmes Journées Nationales Génie Côtier* .

Rosbjerg, D. (1985). Estimation in partial duration series with independent and dependent peak values. *J. Hydrol. , 76*, 183–195.

Roth, M., Buishand, T. A., Jongbloed, G., Klein Tank, A., & van Zanten, J. (2012). A regional peaks-over-threshold model in a nonstationary climate. *Water Resour. Res. , 48*, W11533.

Sabourin, A., & Renard, B. (2015). Combining regional estimation and historical floods: a multivariate semi-parametric peaks-over-threshold model with censored data. *Water Resources Reseach , 51* (12).

Saf, B. (2009). Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey. *Water Resour. Manag. , 23*, 531-551.

Schaefer, M. G. (1990). Regional analysis of precipitation annual maxima in washington state. *Water Resources Research , 26* (1), 119-131.

Scholz, F. W., & Stephens, M. A. (1987). K-Sample Anderson-Darling Tests. *Journal of the American Statistical Association , 82* (399), 918-924.

Simon, B. (2007). *La Marée océanique côtière.* Editions de l'Institut Océanographique.

Smirnov, N. V. (1939). Estimate of deviation between empirical distribution functions in two independent samples. *Bull. Moscow Univ. , 2* (2), 3-16.

Sotillo, M. G., Aznar, R., & Valero, F. (2006). Mediterranean offshore extreme wind analysis from the 44-year hipocas database: different approaches towards the estimation of return periods and levels of extreme values. *Advances in Geosciences , 7*, 275-278.

Stedinger, J. (1983). Estimating a regional flood frequency distribution. *Water Resources Research , 19* (2), 503-510.

Stedinger, J. R., & Baker, V. R. (1987). Surface water hydrology: Historical and paleoflood information. *Reviews of Geophysics , 25* (2), 119-124.

Stedinger, J. R., & Cohn, T. A. (1986). Flood Frequency Analysis With Historical and Paleoflood Information. *Water Resources Research , 22* (5), 785-793.

van Gelder, P. H., & Neykov, N. M. (23-27 November 1998). Regional frequency analysis of extreme water level along the Dutch coast using L-moments: A preliminary study. In *Stochastic models of hydrological processes and their applications to problems of environmental preservation* (pp. 14-20). Moscow, Russia: NATO Advanced Research Workshop.

van Gelder, P. H., De Ronde, J., Neykov, N. M., & Neytchev, P. (2000). Regional frequency analysis of extreme wave heights: Trading space for time. *Proc. 27th Int. Conf. on Coastal Eng. (ICCE 2000, Sydney)* .

Viglione, A., Laio, F., & Claps, P. (2007). A comparison of homogeneity tests for regional frequency analysis. *Water Resources Research , 43* (3), 1-10.

Wang, X. L., Wan, H., Zwiers, F. W., Swail, V. R., Compo, G. P., Allan, R. J., et al. (2011). Trends and low-frequency variability of storminess over western Europe. *Climate Dynamics , 37* (11-12), 2355–2371.

Wang, X. L., Zwiers, F. W., Swail, V. R., & Feng, Y. (2009). Trends and variability of storminess in the Northeast Atlantic region. *Climate Dynamics , 33* (7-8), 1179-1195.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association , 58* (301), 236-244.

Watson, P. J. (2016). A new perspective on global mean sea level (GMSL) acceleration. *Geophys. Res. Lett. , 43*, 6478–6484.

Weibull, W. (1951). A statistical distribution function of wide applicability. *J. Appl. Mech.-Trans. ASME , 18* (3), 293-297.

Weiss, J. (2014). *Analyse régionale des aléas maritimes extremes.* Paris: Sciences de l'ingénieur, Université Paris-Est.

Weiss, J., Bernardara, P., & Benoit, M. (2013). A method to identify and form homogeneous regions for regional frequency analysis of extreme skew storm surges. *Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards.* Siegen.

Weiss, J., Bernardara, P., & Benoit, M. (2014a). Formation of homogeneous regions for regional frequency analysis of extreme significant wave heights. *Journal of Geophysical Research: Oceans , 119* (5), 2906-2922.

Weiss, J., Bernardara, P., & Benoit, M. (2014b). Modeling the intersite dependance for regional frequency analysis of extreme marine events. *Water Resources Research , 50*, 5926-5940.