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.

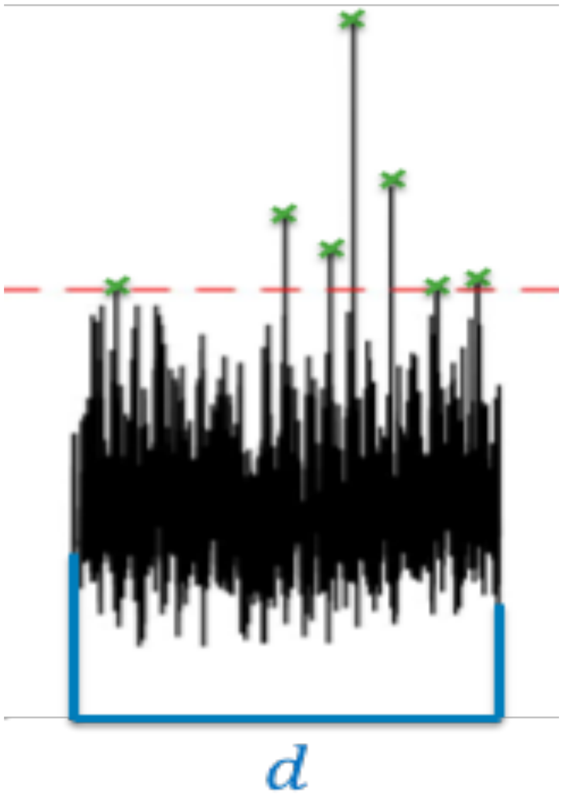

Fig. 7 – Scheme for λ computation on systematic time series

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:

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

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:

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

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:

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

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

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

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:

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

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):

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

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:

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

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

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:

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

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

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):

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

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

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:

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

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

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

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

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 $S\text{-}\chi^2$.

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≠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≠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:

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

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

$$\mathrm{y}^{\mathrm{r}}_{\mathrm{T}} = \mathrm{F}_{\mathrm{r}}^{-1}\left(1 - \frac{1}{\lambda\mathrm{T}}\right) \tag{3.42}$$

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:

$$x^i_T = F_i^{-1}\left(1 - \frac{1}{\lambda T}\right) = u_i\, F_r^{-1}\left(1 - \frac{1}{\lambda T}\right) = u_i y^r_T \tag{3.43}$$

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).