Modelling intersite dependence for regional frequency analysis of extreme marine events

Jérôme Weiss, Pietro Bernardara, Michel Benoit

Modelling intersite dependence for regional frequency analysis of extreme marine events

marine events

Jérôme Weiss [1,2], Pietro Bernardara [1,3], Michel Benoit [1,2]

---

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

<u>Abstract</u>

40

41   The duration of observation at a site of interest is generally too low to reliably estimate

42   marine extremes. Regional frequency analysis (RFA), by exploiting the similarity between

43   sites, can help to reduce uncertainties inherent to local analyses. Extreme observations in a

44   homogeneous region are especially assumed to follow a common regional distribution, up to a

45   local index. The regional pooling method, by gathering observations from different sites into

46   a regional sample, can be employed to estimate the regional distribution. However, such a

47   procedure may be highly affected by intersite dependence in the regional sample. This paper

48   derives a theoretical model of intersite dependence, dedicated to the regional pooling method

49   in a "*peaks over threshold*" framework. This model expresses the tendency of sites to display

50   a similar behavior during a storm generating extreme observations, by describing both the

51   storm propagation in the region and the storm intensity. The proposed model allows the

52   assessment of i) the regional effective duration of the regional sample and ii) different

53   regional hazards, e.g., return periods of storms. An application to the estimation of extreme

54   significant wave heights from the numerical sea-state database ANEMOC-2 is provided,

55   where different patterns of regional dependence are highlighted.

56   *Keywords:* regional frequency analysis, pooling, intersite dependence, extremes, significant

57   wave heights

58

59

60

61

62

1   Introduction

The design of off-shore structures, or coastal protections preventing shoreline areas from marine flooding, particularly requires an accurate estimation of the probability of occurrence of extreme marine events (e.g., extreme storm surges or wave heights). High return levels can be inferred through a local statistical analysis of extremes, from a time series observed at a given site. However, a potential issue is the local duration of observation, generally too low to accurately estimate return levels of interest. For example, wave records from buoys are usually shorter than 20-30 years.

Regional frequency analysis (RFA) can help to reduce these uncertainties, by exploiting the information shared by similar sites in a region. When based on the index-flood method [*Dalrymple*, 1960], RFA assumes that extreme observations in a homogeneous region follow a common regional probability distribution, up to a local index representing the local specificities of each site.

A possible approach to estimate the parameters of the regional distribution is the regional pooling method [*Bernardara et al.,* 2011]. The principle is to pool the data normalized by the local index in a single regional sample, the latter being used to fit the regional distribution. This method is also referred to as the station-year method [*Buishand*, 1991] and illustrates the principle of "*trading space for time*". However, it assumes intersite independence [*Cunnane*, 1988; *Madsen and Rosbjerg*, 1997; *Stewart et al.*, 1999], which cannot be deemed realistic: indeed, for example, a storm is likely to generate dependent extremes at different sites. Thus, *Dales and Reed* [1989] and *Stewart et al.* [1999] questioned the relevance of regional pooling when intersite dependence is ignored, and showed its approximate nature. Intersite dependence in regional pooling is actually closely related to the concept of regional effective duration [*Bernardara et al.*, 2011].

87   The regional effective duration, denoted by D_{eff}, can be defined as the effective

88   duration of observation of the regional sample filtered of any intersite dependence. For

89   example, if the times series recorded in different sites from a given region are considered

90   independent, pooling data from 10 sites, each having 30 years of observation, is equivalent to

91   sample 300 years of "effective duration". This is not the case in the presence of intersite

92   dependence. At the same time, the highest independent normalized observation in the region

93   is viewed as the largest in D_{eff} years of record. It can be used to both reflect the relevance of

94   RFA to a local analysis and to estimate empirical regional return periods. However,

95   *Kergadallan* [2013] pointed out that one limitation of RFA is the difficulty to evaluate D_{eff}.

96   As an illustration, most of regional pooling studies are based on a simplifying hypothesis. For

97   example, *Hjalimarsom and Thomas* [1992], *Bernardara et al.* [2011] and *Bardet et al.* [2011]

98   assumed D_{eff} as the sum of all the local durations, hence assuming intersite independency.

99   *Dalrymple* [1958] expressed that records cannot be expanded to yield an effective duration

100   equal to the sum of local durations. In this work, it is conversely assumed that D_{eff} can be

101   formulated as the typical local duration, implicitly considering perfect intersite dependence.

102   The actual value is likely to lie between these two extreme cases. A realistic estimation of D_{eff}

103   requires a proper characterization of intersite dependence.

104   A consequence of intersite dependence is a loss of information [*Reed*, 1994]. For

105   example, when a storm impacts several sites, there is redundancy of information because

106   observed extremes stem from the same meteorological event. Several studies assessed the

107   effects of intersite dependence in the framework of RFA. For example, the effective size of

108   samples is reduced [*Bayazit and Önöz*, 2004; *Buishand*, 1991; *Kjeldsen and Rosbjerg*, 2002;

109   *Madsen and Rosbjerg*, 1997; *Rosbjerg and Madsen*, 1996]. *Castellarin et al.* [2008] also

110   observed a decrease of the power of the homogeneity test proposed by *Hosking and Wallis*

111   [1993]. *Stedinger* [1983], *Hosking and Wallis* [1988] and *Rosbjerg and Madsen* [1996]

112 showed that ignoring intersite dependence in RFA leads to an underestimation of the variance

113 of return levels estimates. When at-site distributions are the main interest, *Smith* [1990]

114 suggests to initially ignore intersite dependence and then correcting a posteriori the regional

115 variance.

116   A simple way to take into account intersite dependence is to remove it. Some authors

117 proposed its filtering through a spatial declustering procedure, where events impacting several

118 sites are counted only once. To estimate extreme surges with RFA, *Bernardara et al.* [2011]

119 and *Bardet et al.* [2011] formed the regional sample with the highest observations among

120 extremes occurring within 72 hours in the study area. However, the major disadvantage of

121 such an approach is a significant loss of information on the spatial dynamics of extremes

122 generated by a single storm. Moreover, this approach does not introduce any technique to

123 estimate D_{eff}.

124   Intersite dependence can also be modeled. *Cooley et al.* [2012] and *Bernard et al.*

125 [2013] deplored the lack of an explicit modeling of intersite dependence for RFA.

126 Nevertheless, *Renard and Lang* [2007] and *Renard* [2011] represented the dependence of

127 extreme rainfalls at different sites with elliptical copulas. Extremes at two different sites were

128 also regionally modeled by *Buishand* [1984], through bivariate extreme value theory. An

129 alternative approach, dedicated to annual maxima, was proposed by *Dales and Reed* [1989]; it

130 links distributions of the regional maximum and the typical regional data through an effective

131 number of independent sites.

132   Most of the papers cited above analyzed series of annual maxima. Yet, an alternative

133 way is to consider exceedances over a high threshold with the "*peaks over threshold*" (POT)

134 method [*Davison and Smith*, 1990]. Its superiority over methods based on annual maxima, for

135 both local and regional estimation of extremes, was demonstrated by *Madsen et al.* [1997a],

136 *Madsen et al.* [1997b] and *Arns et al.* [2013]. Besides, the POT framework is more physically

137   appealing to handle intersite dependence. For annual maxima, this one is characterized on a

138   yearly basis, and may thus be difficult to interpret: for example, series of annual maxima

139   observed at two distinct sites can be highly statistically correlated, without necessarily being

140   caused by the same meteorological phenomena. Conversely, the POT framework allows

141   reasoning at the scale of the physical event, provided that the concurrence of observations at

142   different sites can be carefully defined [*Mikkelsen et al.*, 1996; *Stewart et al.*, 1999]. In

143   particular, storms generating extreme observations offer an intuitive framework to deal with

144   intersite dependence in a POT approach. *Weiss et al.* [2014] characterized storms through the

145   gathering of extremes neighbors in space and time, and described a procedure to detect them

146   in the context of marine extremes. These storms allow to naturally define the concurrence of

147   observations at the scale of the physical event.

148   Thus, very few studies addressed the issue of intersite dependence for RFA in a POT

149   framework. *Roth et al.* [2012] used the model of *Dales and Reed* [1989] by grouping POT

150   data into seasonal blocks. It can be argued, that defining the concurrence of observations

151   through a wide temporal block (the season) may result in a loss of information on both the

152   spatial coverage and the intensity of the physical events generating extremes. *Mikkelsen et al.*

153   [1996], *Rosbjerg and Madsen* [1996], *Madsen and Rosbjerg* [1998] and *Madsen et al.* [2002]

154   proposed regional regression models, which are not based on the index-flood procedure used

155   in this paper. Similar to geostatistics, their models explicitly account for intersite correlation,

156   where the concurrence of observations is defined through the overlap of POT data in a short

157   time window. *Madsen and Rosbjerg* [1997] corrected the variance of the regional distribution

158   parameters with an effective number of independent sites, based on a regional average

159   correlation coefficient. However, in the latter references, although the concurrence of

160   observations is defined in a physically appealing way, only the pairwise dependence is

161   modeled. Moving towards a more global model of intersite dependence indicating, for

162   example, the tendency of sites in a region to behave similarly during a storm, would help to

163   characterize different regional hazards.

164   The estimation of extreme events by RFA allows to tackle the open question of the

165   difference between regional and local return period. In particular, note that to estimate the

166   return period of a storm affecting a given area, synoptic variables are usually defined first.

167   *Della-Marta and Pinto* [2009] characterized a storm by the minimum central pressure and the

168   maximum vorticity reached during its track; *Pinto et al.* [2012] used the wind speed

169   maximum; a more general spatial index, reflecting both the magnitude and the spatial extent,

170   was defined by *Della-Marta et al.* [2009], who then estimated the return period from these

171   synoptic variables. By construction, such an estimate corresponds to a "regional" return

172   period, namely the return period of a storm which can occur anywhere in the study area.

173   However, for practical applications (e.g., protection design), a local return period must be

174   estimated. For example, it is clear that a storm whose regional return period is 50 years will

175   not generate everywhere in the area wave heights (or storm surges) corresponding to a 50-

176   year return period. In particular, the link between the regional return period of a storm and the

177   return period of a given observed variable generated by the storm at a particular location

178   remains unknown. Note that *Della-Marta et al.* [2009] showed that regional return periods

179   share up to about half of the variability of the local return periods. In this study, we will show

180   how a proper treatment of intersite dependence can help to describe the relation between the

181   regional and the local return period of a storm.

182   The objective of this paper is to develop a global model of intersite dependence for

183   RFA, specifically dedicated to the regional pooling method and POT data, by reasoning at the

184   storm scale. Distributions of the regional storm maximum and the typical regional storm data

185   are linked through a function of *regional dependence*, describing both the propagation of

186   storms and their regional intensity. The proposed model allows the derivation of different

187   regional hazards and the regional effective duration.

188   The model of regional dependence is developed in section 2, including its implications

189   on the regional pooling method (section 2.4). An application to the estimation of extreme

190   significant wave heights from the numerical database ANEMOC-2 is shown in section 3.

191   **2   <u>Methodology</u>**

192   2.1   <u>Extraction of storms</u>

193   To characterize the intersite dependence, it is first necessary to define the simultaneity

194   of observations in space. If data are sampled every hour, the reference for simultaneity can be,

195   for example, the hourly scale. However, as extreme oceano-meteorological conditions can last

196   from several hours to several days, the temporal dimension should be added to describe the

197   spatial dependence. In this paper, the scale of the physical events generating marine extremes

198   (storms) is taken as the reference to define the simultaneity of observations in space.

199   A storm is thus directly characterized through the variable of interest (e.g., wave

200   height or storm surge), being defined as a physical event generating marine extremes in at

201   least one site in the study area. In the literature, the tracking of storms often relies on a nearest-

202   neighbor search in space and time [e.g., *Leckebusch et al.*, 2008; *Renggli et al.*, 2010]. A spatio-

203   temporal declustering procedure is thus employed to detect storms and to reflect their

204   propagation in space and time. In particular, extremes neighbors in space and time are

205   supposed to stem from the same storm. The storm extraction algorithm in the context of

206   marine extremes is described in *Weiss et al.* [2014], trying to reproduce at best the physical

207   dynamics of the storms, while taking into account the spatio-temporal resolution of

208   observations. Moreover, a "double-threshold" approach is employed to separate physical

209   considerations from statistical ones [*Bernardara et al.,* 2014].

210   At a given site, the impact of a storm is characterized by observations exceeding the

211   "physical threshold" q_p, defined as the p-quantile of the initial time series, with p close to 1. In

212   order to get independent data at site scale, only the peak value W_s^i is retained to summarize

213   the storm s at site i (which implies that all other extremes occurring during that storm are

214   discarded).

215   Only the most intense storm events are now considered for statistical aspects. New

216   thresholds, denoted u and higher than the quantiles q_p, are selected corresponding to the

217   occurrence of λ storms per year on average at each site. In particular, if d_i years of data are

218   available at site i, the $n_i = \lambda d_i$ highest W_s^i are retained to form the n_i-sample X_s^i. The

219   "statistical threshold" u_i, exceeded on average λ times per year, is then defined as the smallest

220   observation from X_s^i (minus an infinitesimal quantity). Storms are then statistically redefined:

221   if site i was impacted by storm s, it is from now on impacted by s if and only if u_i is exceeded.

222   2.2   Regional frequency analysis

223   Extreme events are estimated in this paper from exceedances over a high threshold.

224   According to *Pickands* [1975], the Generalized Pareto Distribution (GPD) represents the

225   natural distribution for such exceedances. For ease of notation, the index s denoting the storm

226   is omitted in this section. For site i, let u_i be the storm threshold which is exceeded on average

227   λ times per year. The n_i-sample X^i, denoting the exceedances of u_i, is assumed to be drawn

228   from a GPD: $X^i \sim GPD(u_i, \alpha_i, k_i)$, where $\alpha_i > 0$ and k_i are, respectively, a scale and a shape

229   parameter. In particular, the p-quantile of X^i is:

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

230   The right tail of the GPD is bounded when $k_i < 0$, and unbounded when $k_i \geq 0$. The T-year

231   return level, i.e., the value exceeded on average once every T years, is given by $x^i_{1-1/\lambda T}$

232   [*Rosbjerg*, 1985].

233   A homogeneity hypothesis is required for RFA based on the index-flood method.

234   Observations from sites coming from a homogeneous region are supposed to follow the same

235   regional probability distribution, up to a local index representing the local specificities of a

236   site. In this paper, homogeneous regions are formed following *Weiss et al.* [2014], where

237   typical storm footprints are identified with a clustering algorithm based on a criterion of storm

238   propagation. In particular, sites from a given region are likely to be impacted by the same

239   storms, and any storm impacting a region is likely to remain enclosed in this region.

240   For a homogenous region of N sites, let μ_i be the local index of the site $i = 1,\dots,N$. The

241   normalized variable $Y = X^i / \mu_i$ is supposed to be independent of i, with cumulative distribution

242   function (c.d.f.) F_r. *Roth et al.* [2012] showed that dealing with exceedances over a high

243   threshold necessarily implies that the local index has to be a multiple of this threshold. Here,

244   as in *Roth et al.* [2012] and *Weiss et al.* [2014], μ_i is therefore chosen as the threshold u_i. This

245   implies that $Y \sim GPD(1, \gamma, k)$, where: i) the regional scale parameter satisfies $\gamma = \alpha_i / u_i$ and ii)

246   the shape parameter $k_i = k$ is constant over the region. From these relationships, $X^i \sim GPD(u_i,$

247    $\gamma u_i, k)$. For site i, the T-year return level is obtained by multiplying the regional T-year return

248   level by the local index: $x^i_{1-1/\lambda T} = u_i y_{1-1/\lambda T}$.

249   The two regional parameters (γ, k) can be estimated with the regional pooling method.

250   However, as sites in a region are likely to be impacted by the same storms, a strong intersite

251   dependence is expected. If ignored, this may affect the estimation process. Thus, this

252   dependence is firstly modeled as outlined in section 2.3, before the regional pooling method is

253   described in section 2.4.

254   2.3   <u>Modeling of regional dependence</u>

255   *2.3.1   Notations*

256   Let λ_r be the mean annual number of storms in the region and Z_s^i the Bernoulli variable

257   which is 1 if storm s impacts site i and 0 otherwise. When storm s impacts site i, the observed

258   normalized extreme with c.d.f. F_r is denoted by $Y_s^i = X_s^i / u_i$. Note that $Y_s^i \geq 1$. The storm s can

259   be summarized in the region by the multivariate random variable $\eta_s = (\eta_s^1,...,\eta_s^N)$, where

260   $\eta_s^i = Y_s^i Z_s^i$. The storm regional maximum is then defined as $M_s = \max\limits_{i=1,..,N} \eta_s^i$. As at least one site

261   is impacted by the storm s, $M_s \geq 1$.

262   *2.3.2   Distribution of the storm regional maximum*

263   First, note that due to the statistical redefinition of storms at the end of section 2.1, Z_s^i

264   takes the value 1 with probability λ / λ_r, independently of i. Moreover, by the regional

265   homogeneity hypothesis from section 2.2, the distribution of η_s^i does not depend on i:

$$\forall x \geq 1, \ P(\eta_s^i > x) = P(Z_s^i = 1)P(\eta_s^i > x | Z_s^i = 1) = \frac{\lambda}{\lambda_r}(1 - F_r(x)) \tag{2}$$

266   For $x \geq 1$, the distribution of M_s can be obtained through the following decomposition:

$$P(M_s > x) = P(\max\limits_{i=1,..,N} \eta_s^i > x) = \sum_{i=1}^{N-1} P(\eta_s^i > x, \max\limits_{j=i+1,..,N} \eta_s^j \leq x) + P(\eta_s^N > x) \tag{3}$$

267   Now, as the distribution of η_s^i is independent of i:

$$P(M_s > x) = P(\eta_s^1 > x)[1 + \sum_{i=1}^{N-1} P(\max\limits_{j=i+1,..,N} \eta_s^j \leq x | \eta_s^i > x)] \tag{4}$$

268   From (2), this leads to:

$$P(M_s > x) = P(\eta_s^1 > x)\varphi(x) = (1 - F_r(x))\frac{\lambda}{\lambda_r}\varphi(x) \tag{5}$$

269   where

$$\varphi(x) = 1 + \sum_{i=1}^{N-1} P(\max_{j=i+1,..,N} \eta_s^j \leq x \mid \eta_s^i > x), \ x \geq 1 \tag{6}$$

270   The distribution of M_s can be thus written in terms of the regional distribution F_r and φ.

271   *2.3.3   Characterization of the regional dependence*

272   The function φ reflects the regional dependence. Situations of independence and

273   perfect dependence, illustrating extreme cases of dependence, can be reinterpreted through

274   equation (6) with $x = 1$. In particular, the region is *regional-independent* ($r - \perp\!\!\!\perp$) if and only if

275     $\varphi \equiv N$; in that case, a storm impacts only one site in the region, whatever its intensity.

276   Conversely, the region is *perfectly regional-dependent* ($p-rd$) if and only if $\varphi \equiv 1$; a storm

277   impacts every site in the region and, whatever its intensity, the generated (normalized)

278   extremes vary the same way. Between these two extremal situations, φ takes values between 1

279   and N.

280   By construction, φ relates both the storm propagation in the region and the storm

281   intensity. It expresses the tendency of sites to display a similar behavior during a storm. The

282   regional dependence is stronger when φ is small, hence indicating that most of the sites are

283   impacted by a storm, and are likely to react the same way in terms of normalized extremes.

284   φ is influenced by the number N of sites in the region. In order to compare φ between

285   different regions, the effect of N can be removed through the following adimensional

286   function:

$$\Phi(x) = \frac{N - \varphi(x)}{N - 1}, \ x \geq 1 \tag{7}$$

287   where Φ, lying between 0 and 1, is near to 1 when regional dependence is strong.

288   *2.3.4   Assessment of regional hazards*

289   A regional hazard is an event occurring at the regional scale, whose probabilistic

290   description is related to collective risk assessment. The following examples of regional

291   hazards are expressed in terms of the function of regional dependence φ.

292   A first example is the mean number $\beta_s(x)$ of impacted sites with normalized intensity

293   larger than $x \geq 1$ when the storm regional maximum is larger than x:

$$\beta_s(x) = E[\sum_{i=1}^{N} 1_{\eta_s^i > x \mid M_s > x}] \tag{8}$$

294   From equation (5):

$$\beta_s(x) = NP(\eta_s^1 > x \mid M_s > x) = N\frac{P(\eta_s^1 > x)}{P(M_s > x)} = \frac{N}{\varphi(x)} \tag{9}$$

295   In particular, the mean number of impacted sites during any storm is given by $\beta_s(1) = N / \varphi(1)$.

296   Note that this is coherent with the definitions of *regional-independence* and *perfect regional-*

297   *dependence*.

298   Another example is the evaluation of the regional return period of a particular storm,

299   and how it is related to its local return period. Let s be a given storm, and denote by $x \geq 1$ its

300   corresponding normalized intensity. The *regional return period* of s, T_r, is defined as the

301   average time between storms impacting *at least one site* in the region with a normalized

302   intensity greater than x, i.e.:

$$T_r = \frac{1}{\lambda_r P(M_s > x)} \tag{10}$$

303   The *local return period* of s, T, is defined as the average time between storms

304   impacting *a given site* in the region with a normalized intensity greater than x:

$$T = \frac{1}{\lambda(1 - F_r(x))} \tag{11}$$

305   From (5), T_r and T are related through:

$$T_r = \frac{T}{\varphi(x)} \qquad (12)$$

306   2.4   <u>Regional pooling method</u>

307   *2.4.1   Construction of the regional sample*

308   The regional pooling method is used to estimate the regional distribution F_r. However,

309   due to the presence of intersite dependence, events impacting several sites must be counted

310   only once. Storms presented in section 2.1 are a convenient way to filter intersite dependence,

311   as each storm describes the regional footprint of a particular event generating extremes.

312   In particular, the distribution of the storm regional maximum M_s is now assumed to be

313   the same as the regional distribution F_r. This assumption was implicitly made in *Bernardara*

314   *et al.* [2011] and *Bardet et al.* [2011], where the regional distribution was estimated from the

315   highest normalized surges occurred within 72 hours in the region. In other words, the

316   distribution of the maximum of a regional cluster is identical to the distribution of a generic

317   element of this cluster. The same assumption is often made in a POT time series framework,

318   as explained by Anderson in the discussion of the paper by *Davison and Smith* [1990]: "*this*

319   *apparent paradox is a consequence of length-biased sampling: a randomly chosen*

320   *exceedance has a disproportionate chance of coming from a large cluster, and in large*

321   *clusters there tend to be large excesses.*" However, in practice, the validity of this assumption

322   must be verified. For example, the two-sample Anderson-Darling test [*Scholz and Stephens*,

323   1987] can be performed at each site i to evaluate the null hypothesis that Y_s^i and M_s have the

324   same distribution.

325   If n_r independent storms are observed in the region, the regional sample is thus formed

326   by the n_r-sample of storm regional maxima M_s, and corresponds to D_{eff} years of regional

327   effective duration.

328   The assumption that the storm regional maximum M_s is the same as the regional

329   distribution F_r depends on the data at hand. When this hypothesis is not verified, the

330   following alternative strategies nevertheless allow to perform a RFA:

331   i)   Remove sites of which Anderson-Darling p-values are too low (for example, lower

332   than 0.01) to accept this hypothesis. The application of the model of regional

333   dependence and the estimation of F_r can then be performed on the remaining sites.

334   ii)   Form the regional sample with random (normalized) observations from each storm,

335   instead of using the storm regional maxima M_s. F_r can still be estimated by pooling,

336   directly from this new regional sample. However, the simplified model of

337   dependence (section 2.4.3) is not valid anymore, as φ is not a constant function. It

338   would be possible to update equation (14) by letting the regional effective duration

339   depend on regional quantiles.

340   iii) Use another method to perform the RFA, e.g., the regional L-moments method of

341   *Hosking and Wallis* [1997]. The model of regional dependence developed in this

342   paper, dedicated to the pooling method, does not apply anymore in this case.

343   *2.4.2   Estimation of the regional distribution F_r.*

344   The two regional parameters (γ, k), see section 2.2, are estimated from the regional

345   sample. Penalized maximum likelihood estimation (PMLE) [*Coles and Dixon*, 1999] is used

346   in this study. The principle is to combine the efficiency of maximum likelihood estimators for

347   large sample sizes and the reliability of the probability weighted moment estimators for small

348   sample sizes. In particular, high estimates of the shape parameter k are penalized. PMLE is

349   implemented in the function fitgpd of the POT package [*Ribatet*, 2007], in the statistical

350   computing environment R (R Development Core Team, 2013). Uncertainties on estimates of

351   (γ, k) are here assessed with a bootstrap procedure: 10,000 replications of the (γ, k) values are

352   obtained with PMLE from resamples of the regional sample.

353   *2.4.3   Simplification of the model of regional dependence*

354   The regional pooling method presented in this paper assumes that the distribution of

355   the storm regional maximum M_s is the same as the regional distribution F_r. The model of

356   regional dependence in section 2.3 can thus be simplified. Indeed, from (5), this assumption

357   implies that φ becomes a constant function:

$$\forall x \geq 1, \ P(M_s > x) = 1 - F_r(x) \leftrightarrow \varphi(x) = \frac{\lambda_r}{\lambda} \tag{13}$$

358   As φ is constant, the way sites react during a storm does not depend on the intensity of the

359   storm. Similarly, *Dales and Reed* [1989] applied their model to rainfall annual maxima and

360   observed that the effective number of sites, summarizing the spatial dependence, did not seem

361   to depend on a particular regional intensity.

362   *2.4.4   The regional effective duration D_{eff}*

363   The pooling procedure yields D_{eff} years of regional effective duration. D_{eff} is closely

364   related to the degree of regional dependence; in particular, D_{eff} is expected to be low when

365   regional dependence is strong.

366   First, the two simplistic situations of regional dependence (section 2.3.3) are

367   considered. Let $\bar{d} = \sum d_i / N$ be the mean local duration, where d_i is the local duration of

368   observation at site i and N is the number of sites in the region. If the region is $r - \perp\!\!\!\perp$, a storm

369   impacts only one site in the region. In that case, each observation from any site brings new

370   information, and D_{eff} can be written as the sum of all the local durations: $D_{eff} = N\bar{d}$.

371   Conversely, if the region is $p - rd$, a storm impacts every site in the region. Here, the typical

372   local duration of one site constitutes D_{eff}, as the information from other sites is purely

373   redundant. This can be reflected by taking, for example, D_{eff} as the mean local duration: $D_{eff} = $

374     $\bar{d}$. It is now assumed that, between these two extremal cases, D_{eff} can be more realistically

375   expressed by:

$$D_{eff} = \varphi \, \bar{d} \qquad (14)$$

376   where φ, lying between 1 and N, is the degree of regional dependence. Note that the situations

377   of $p-rd$ and $r-\perp\!\!\!\perp$ are respectively obtained for $\varphi = 1$ and $\varphi = N$. From equation (13), stating

378   that $\varphi = \lambda_r / \lambda$, its theoretical value is $D_{eff} = \lambda_r \bar{d} \, / \lambda$.

379   The mean annual number of storms in the region λ_r can be naturally estimated by n_r /

380     $\bar{d}$, where n_r is the number of observed storms. An estimate of D_{eff} is then:

$$\hat{D}_{eff} = \frac{n_r}{\lambda} \qquad (15)$$

381   Let $n_{r,t}$ be the number of observed storms during year $t = t_1,\ldots,t_\tau$ in the region, where t_1 and t_τ

382   indicate the first and the last year of observation in the region, respectively. The overall

383   number of observed storms n_r is obtained by summing the $n_{r,t}$ for $t = t_1,\ldots,t_\tau$. By assuming that

384   the $n_{r,t}$ are independent and identically distributed with common mean λ_r and standard

385   deviation σ_r, the central limit theorem followed by the Slutsky's lemma allow to derive new

386   confidence intervals for D_{eff} :

$$[\hat{D}_{eff} - z_{1-\alpha/2} \frac{\hat{\sigma}_r \sqrt{\tau}}{\lambda}, \hat{D}_{eff} + z_{1-\alpha/2} \frac{\hat{\sigma}_r \sqrt{\tau}}{\lambda}] \qquad (16)$$

387   where $z_{1-\alpha/2}$ is the quantile of order $1-\alpha/2$ of the standard normal distribution, $\hat{\sigma}_r$ is the

388   empirical standard deviation of the $n_{r,t,}$ and τ is the number of years of observation in the

389   region.

390   Note that (15) can be used even if periods of observations are different, and in the

391   presence of missing data. This formula also guarantees that $\hat{D}_{eff} \geq \max\limits_{i=1,..,N} d_i$, coherently with

392   what might be expected from the regional effective duration. Besides, it reflects the

393   importance to extract storms such that their mean annual occurrence λ at the local scale is

394   common to all sites.

395   As F_r is estimated from D_{eff} years of pooled data, the underlying principle is that any

396   site in the region can be indifferently impacted by a given storm. Parenthetically, with no

397   preferential storm track in the region, the regional pooling method is coherent with the

398   identification of storms footprints to form homogeneous regions. In particular, the regional

399   sample illustrates that, for a generic site, λ storms per year, on average, were observed during

400   D_{eff} years. D_{eff} thus helps to reflect the relevance of RFA to a local analysis. Indeed, pooling

401   enables to estimate extreme events at site i from D_{eff} years of data, compared to d_i years for a

402   local analysis.

403   *2.4.5   Evaluation of storm return periods*

404   The regional pooling method allows to distinguish between local and regional return

405   periods of normalized storm events (see section 2.3.4 for the corresponding definitions), both

406   at the empirical and theoretical levels. Let s be a given storm from the regional sample, and

407   denote by x its corresponding normalized intensity.

408   Using the Weibull plotting position, its empirical local return period $\widetilde{T}_{s,loc}$ is:

$$\widetilde{T}_{s,loc} = \frac{D_{eff}+1}{n_r+1-\text{rank}(s)} \tag{17}$$

409   where rank(s) denotes the rank of s in the regional sample. For example, if s is the most

410   intense storm observed in the regional sample, then $\widetilde{T}_{s,loc}$ is about D_{eff} years. Besides, the

411   theoretical local return period $\overline{T}_{s,loc}$ of s is given by equation (11). We recall that $\overline{T}_{s,loc}$

412   corresponds to the theoretical return period of storm s at site scale (i.e., at any site of the

413   region). Using (12) and (13), the theoretical regional return period $\overline{T}_{s,reg}$ is given by:

$$\overline{T}_{s,reg} = \frac{\lambda}{\lambda_r}\overline{T}_{s,loc} \tag{18}$$

414   The empirical regional return period $\widetilde{T}_{s,reg}$ is linked with $\widetilde{T}_{s,loc}$ through a similar relation.

415 **3** **Application**

416 3.1 Data used

417   ANEMOC-2 (Atlas Numérique d'États de Mer Océaniques et Côtiers - Numerical

418 Atlas of Oceanic and Coastal Sea states) is a numerical sea-state hindcast database covering

419 the Atlantic Ocean over the period 1979-2009 (31 years). It has been developed at Saint-

420 Venant Laboratory for Hydraulics and EDF R&D LNHE [*Laugel*, 2013]. The simulations of

421 wave conditions have been carried out with the third-generation spectral wave model

422 TOMAWAC [*Benoit et al.*, 1996] forced by wind fields from the CFSR reanalysis database

423 [*Saha et al.*, 2010].

424   The spatial resolution of the so-called "oceanic mesh" of ANEMOC-2 ranges from

425 about 120 km over the Northern part of the Atlantic Ocean down to about 20 km along the

426 European coast and 10 km along the French coast. Note this grid is supplemented by a

427 "coastal mesh" whose resolution is finer on the continental shelf, in the Channel and along the

428 French coast. For the present study, however, only data from the oceanic mesh is used, and

429 only a subset of 1847 nodes amongst the 13426 nodes of the full oceanic mesh is selected, at

430 locations plotted in Figure 1.

431   Among the wave parameters available with an hourly resolution in ANEMOC-2, we

432 consider here the significant wave height, denoted H_s, which is usually the preferred

433 parameter to summarize sea state intensity. TOMAWAC computes this wave height from the

434 zero-order moment of the wave spectrum. Hourly series of significant wave heights H_s over

435 the period 1979-2009 are thus extracted for the 1847 selected sites. The objective here is to

436 apply the proposed methodology i) to characterize the regional dependence over this area and

437 ii) to estimate extreme H_s by the regional pooling method.

438   3.2   Preparation of data for RFA

439   More details for this section may be found in *Weiss et al.* [2014], where the same

440   dataset was used.

441   The physical extraction of storms generating extreme H_s, described in the beginning of

442   section 2.1 with $p = 0.995$, leads to 5939 storms. *Weiss et al.* [2014] performed a sensitivity

443   analysis and found that storms are properly detected when $p = 0.995$. A quick analysis reveals

444   that, on average: i) there are 192 storms per year in the study area, ii) a storm impacts 38 sites

445   and iii) a storm lasts 12.5 hours at site scale. These storms serve to form physically

446   homogeneous regions, by detecting the most typical storms footprints in the study area. The

447   footprints of the storms of 15-18 February 1986, 11-13 December 1990 and 23-24 January

448   2009 (Klaus) are shown in Figure 2.

449   Storms are then statistically redefined, following the methodology presented in section

450   2.1. In particular, $\lambda = 1$ storm per year are now observed, on average, at each site; 1340 storms

451   are thus retained among the 5939 initial ones. Site i is therefore characterized by the sample

452   of H_s over the threshold u_i exceeded on average once per year; the sample size is 31, as 31

453   years of data are available. These thresholds, represented in Figure 3, are also the local indices

454   used for RFA (section 2.2). These storms serve to i) check the statistical homogeneity of the

455   physically homogeneous regions, ii) deal with regional dependence and iii) estimate extreme

456   events with the regional pooling method.

457   RFA can thus be performed on each of the six homogeneous regions delineated in

458   *Weiss et al.* [2014], see Figure 4.

459   3.3   Regional pooling method

460   For each of the six regions, the regional sample is constructed by pooling the observed

461   normalized storm regional maxima, following section 2.4.1. To check whether the storm

462   regional maxima are sampled from the regional distribution F_r, the two-sample Anderson-

463   Darling test is performed. The p-values for the null hypothesis that Y_s^i and M_s have the same

464   distribution, for each site i from a given region, are higher than 0.01 for 95% of all sites.

465   Therefore, it may be reasonably assumed that i) the model of regional dependence can be

466   simplified and ii) F_r can be estimated from the regional sample.

467   *3.3.1   Measures of regional dependence*

468   For each region, Table 1 provides some measures of regional dependence defined in

469   sections 2.3.3, 2.3.4 and 2.4.4. Note that there are no missing values in ANEMOC-2 data, and

470   periods of observations are the same for all sites (with a common local duration of $d = 31$

471   years).

472   Storms are, respectively, most and least most frequent in regions 2 and 6, with 25.2

473   and 2.7 storms per year on average. This is explained here by their size: regions 2 and 6 are,

474   respectively, the largest and the smallest in terms of the number of sites. To compare the

475   degree of regional dependence between regions, this size effect can be removed through the

476   adimensional function Φ defined in equation (7). The regional dependence is thus the

477   strongest in region 5, meaning that sites in this region tend to behave highly similarly during a

478   storm: a large proportion of them are impacted, and the normalized extremes are likely to vary

479   the same way. Conversely, the regional dependence is the weakest in region 2. This can be

480   precised by considering the mean number $\beta_s(1)$ of impacted sites during a storm, see equation

481   (9). Indeed, on average, a storm in region 2 only impacts 18.9 sites (4% of the region),

482   whereas a storm in region 5 impacts 56.6 sites (24% of the region).

483   The regional effective duration D_{eff} and its corresponding 95% confidence interval are

484   estimated following equations (15) and (16). For example, pooling data from the 234 sites of

485   region 5 enables to get a regional sample with $D_{eff} = 128$ years of independent observations.

486   Note that taking into account the regional dependence considerably reduces what would be

487   obtained under the assumption of intersite independence (in that case, $D_{eff} = Nd = 7254$ years).

488   Figure 5 shows the evolution of the regional return period T_r against the local return

489   period T for each region, see equation (12). Note that curves for regions 3 and 4 are very close

490   to each other, giving the impression of being superimposed. The simplified model of regional

491   dependence implies that T_r and T are linearly related. For fixed T, T_r is, respectively, the

492   lowest and the highest in regions 2 and 6. For example, Table 1 gives 100_r, i.e., T_r

493   corresponding to $T = 100$ years. $100_r = 3.964$ years in region 2: about every four years on

494   average, a storm in this region causes at least one local 100-year event. Besides, although

495   region 1 is much larger that region 3, note that their 100_r estimates are similar (about 10

496   years). If intersite dependence was assumed, then this quantity would have been, in

497   proportion, much higher in region 3 than in region 1. However, the compensation is due to a

498   stronger regional dependence in region 1. Thus, modeling the regional dependence allows a

499   more realistic assessment of regional hazards.

500   *3.3.2   Estimation of extreme H_s*

501   For each of the six homogeneous regions, the regional GPD parameters (γ, k) are

502   estimated following the procedure outlined in section 2.4.2. These quantities are given in

503   Table 2, as well as the 100-year regional return level $y_{0.99}$. The shape parameter k is positive

504   (corresponding to an unbounded GPD) in regions 1, 4 and 6, suggesting a higher intensity of

505   extreme H_s. The return level plots for each of the six regional distributions, together with the

506   95% confidence intervals obtained by bootstrap, are given in Figure 6. Note that the plotting

507   position depending on the regional effective duration (equation (17)) is used to represent

508   observations from the regional sample, hence allowing the estimation of empirical return

509   periods (see section 3.3.3 for an application to the most intense storms observed).

510   At-site return levels are obtained by multiplying regional return levels by the local

511   indices. Figure 7 shows the map of the estimated at-site 100-year H_s. Estimates for coastal

512   areas are not shown because, as mentioned in section 3.1, the present analysis uses data from

513   the oceanic model of ANEMOC-2, whose resolution is not sufficient for these coastal areas

514   and which includes only parts of the shallow-water effects. In a follow-up of this study, data

515   from the coastal model may improve the simulated sea states in coastal areas. One can note on

516   Figure 7 that 100-year H_s estimates display a coherent spatial pattern, with lower values near

517   the West European coasts. The highest return levels are obtained for sites located in the north-

518   central part of the study area (up to 29.65 m). Note that these estimates are comparable to

519   those from *Caires and Sterl* [2005] and *Weiss et al.* [2014].

520   Compared to a local statistical analysis, the regional pooling method can help to

521   reduce uncertainties in the estimation of extreme events, at a given site. Indeed, extrapolations

522   from a local analysis would be based here on $d = 31$ years of observations; the regional

523   pooling makes available $D_{eff} > d$ years of data for any site.

524   *3.3.3   Return periods of the most intense storms observed*

525   From the numeric database ANEMOC-2, storms with the highest normalized intensity

526   were observed on February 1986, February 1979, December 1990, February 1988, December

527   1989 and January 2009, respectively in region 1, 2, 3, 4, 5 and 6. Figure 2 displays these

528   storms which occurred in regions 1, 3 and 6.

529   As an application of section 2.4.5, return periods of these storms are provided in Table

530   3, both at the local and regional scales (empirical and theoretical). For example, in region 3,

531   the empirical local return period $\tilde{T}_{s,loc}$ of the storm of 11-13 December 1990 is estimated at

532   280 years. Its theoretical counterpart is quite close ($\overline{T}_{s,loc} = 367$ years), as indicated by the

533   return level plot for this region (Figure 6). As such a storm was observed only once in 31

534   years of observations in this region, its empirical regional return period $\tilde{T}_{s,reg}$ is logically

535   estimated at 31 years. From equation (18), the model of regional dependence predicts a

536   theoretical regional return period $\overline{T}_{s,reg}$ of 41 years.

4   Conclusions

By exploiting the information shared by statistically similar sites, regional frequency analysis (RFA) can reduce uncertainties in the estimations of high return levels, when at-site durations of observations are short. It is assumed that, in a homogeneous region, extreme observations follow a common regional probability distribution, up to a local index representing the local specificities of a site.

The method of regional pooling is employed in this paper, where normalized observations from different sites are gathered into a regional sample to estimate the regional distribution. In particular, this pooling procedure allows to define D_{eff} years of regional effective duration. D_{eff} is actually closely related to the degree of intersite dependence: for example, D_{eff} is expected to be low when the dependence is strong. Intersite independence is a usual assumption in the literature and practice, although unrealistic: a storm is likely to generate dependent extremes at different sites. We have therefore proposed a theoretical frame to model intersite dependence for the regional pooling method.

Storms are here identified by detecting physical events generating extremes in at least one site in the study area; their spatio-temporal propagation is taken into account through the gathering of extremes neighbors in space and time. Storms allow to naturally define the concurrence of observations at the scale of the physical event, hence enabling to perform a RFA within a "*peaks over threshold*" framework. These storms represent a convenient way to describe regional dependence. In particular, they are the basis to i) construct the regional sample, by filtering the redundancy of information and ii) model the regional dependence.

The distribution of the storm regional maximum is linked to the regional distribution through a function of regional dependence. This function, describing both the storm propagation in the region and the storm intensity, expresses the tendency of sites to display a similar behavior during a storm. The proposed model allows i) a proper evaluation of D_{eff} and

562 ii) the assessment of different regional hazards: for example, the mean number of impacted

563 sites during a storm, or return periods of storms both at the local and regional scales can be

564 theoretically derived.

565   An application to significant wave heights from the numerical sea-state database

566 ANEMOC-2 has been provided to demonstrate the capabilities of the model. Six

567 homogeneous regions, corresponding to the most typical storms footprints were delineated in

568 the North-East part of the Atlantic Ocean. Different patterns of regional dependence have

569 been characterized in this area, before applying the regional pooling method to estimate

570 extreme significant wave heights.

571   Although the proposed example considers significant wave heights, the method can

572 easily be applied to other marine variables. Indeed, it is variable-oriented, in the sense that

573 storms are specifically defined through to the variable of interest only. Moreover, D_{eff} can also

574 be estimated when periods of observations are not the same for all sites, and/or in the

575 presence of missing values. Future works could, for example, apply the proposed model to

576 other marine hazards (e.g., storm surges) to compare how regional dependence manifests

577 compared to significant wave heights.

578

587 has been extracted from the ANEMOC-2 database. This dataset can be obtained by request

588 addressed to the corresponding author. The use of this data is restricted to research purpose,

589 all industrial or commercial applications being excluded.

590

591 **6   <u>References</u>**
592

593 Arns, A., Wahl, T., Haigh, I. D., Jensen, J. and C. Pattiaratchi (2013), Estimating extreme
594   water level probabilities: A comparison of the direct methods and recommendations
595   for best practice, *Coastal Engineering*, 81, 51-66.
596 Bardet, L., C.-M. Duluc, V. Rebour and J. L'Her (2011), Regional frequency analysis of
597   extreme storm surges along the French coast, *Natural Hazards and Earth System
598   Sciences,* 11, 6, 1627-1639.
599 Bayazit M. and B. Önöz (2004), Sampling variances of regional flood quantiles affected by
600   intersitecorrelation, *Journal of Hydrology*, 291, 1–2, 42–51.
601 Benoit, M., F. Marcos, and F. Becq (1996), Development of a third generation shallow-water
602   wave model with unstructured spatial meshing, in *Proceedings of the 25th
603   International Conference on Coastal Engineering*, pp. 465–478, American Society of
604   Civil Engineers (ASCE), Orlando, Florida.
605 Bernard, E., Naveau, P., Vrac, M. and O. Mestre (2013), Clustering of maxima: spatial
606   dependencies among heavy rainfall in France, *Journal of Climate*, 26, 7929–7937.
607 Bernardara, P., Andreewsky M. and M. Benoit (2011), Application of the Regional Frequency
608   Analysis to the estimation of extreme storm surges, *Journal of Geophysical Research*,
609   116, C02008, 1-11.
610 Bernardara, P., F. Mazas, X. Kergadallan, and L. Hamm (2014), A two-step framework for
611   over-threshold modelling of environmental extremes, *Nat. Hazards Earth Syst. Sci.*,
612   14, 635–647.
613 Buishand, T. A. (1984), Bivariate extreme-value data and the station-year method, *Journal of
614   Hydrology*, 69, 1-4, 77–95.
615 Buishand, T. A. (1991), Extreme rainfall estimation by combining data from several sites,
616   *Hydrological Sciences*, 36, 4, 345–365.
617 Caires, S. and A. Sterl (2005), 100-year return value estimates for ocean wind speed and
618   significant wave height from the ERA-40 data, *Journal of Climate*, 18, 7, 1032-1048.
619 Castellarin, A., Burn, D. H. and A. Brath (2008), Homogeneity testing : how homogeneous do
620   heterogeneous cross-correlated regions seem ? *Journal of Hydrology*, 360, 67–76.
621 Coles, S. and M. Dixon (1999), Likelihood-Based Inference for Extreme Value Models,
622   *Extremes*, 2, 1, 5–23.
623 Cooley, D., Cisewski, J., Erhardt, R. J., Jeon, S., Mannshardt, E., Omolo, B. O. and Y. Sun
624   (2012), A survey of spatial extremes : Measuring spatial dependence and modeling
625   spatial effects, *REVSTAT*, 10, 1.
626 Cunnane, C. (1988), Methods and merits of regional flood frequency analysis, *Journal of
627   Hydrology*, 100, 1–3, 269–290.
628 Dales M. Y. and D. W. Reed (1989), *Regional flood and storm hazard assessment*, report No.
629   102, Institute of Hydrology, Wallingford, Oxon.
630 Darlymple, T. (1958), Flood frequency relations for gaged and ungaged streams, US
631   Geological Survey.

Darlymple, T. (1960), Flood Frequency Analysis, 1543-A, US Geological Survey, *Water Supply Paper*.

Davison, A. C. and R. L. Smith (1990), Models for exceedances over high thresholds, *Journal of the Royal Statistical Society Series B (Methodological)*, 52, 3, 393–442.

Della-Marta, P. M. and J. G. Pinto (2009), Statistical uncertainty of changes in winter storms over the North Atlantic and Europe in an ensemble of transient climate simulations, *Geophysical Research Letters*, 36, L14703.

Della-Marta, P. M., Mathis, H., Frei, C., Liniger, M. A., Kleinn, J. and C. Appenzeller (2009), The return period of wind storms over Europe, *International Journal of Climatology*, 29, 437– 459.

Hjalimarsom, H. and B. Thomas (1992), New Look at Regional Flood Frequency Relations for Arid Lands, *J. Hydraul. Eng.*, 118, 6, 868–886.

Hosking, J. R. M. and J. R. Wallis (1988), The effect of intersite dependence on regional flood frequency analysis, *Water Resources Research*, 24, 4, 588-600.

Hosking, J. R. M. and J. R. Wallis (1993), Some statistics useful in regional frequency analysis, *Water Resources Research*, 29, 2, 271–281.

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

Kergadallan, X. (2013), *Analyse statistique des niveaux d'eau extrêmes - Environnements maritime et estuarien*, CETMEF, Centre d'études techniques maritimes et fluviales, Compiègne, 179p (in French).

Kjeldsen, T. R. and D. Rosbjerg (2002), Comparison of regional index flood estimation procedures based on the extreme value type I distribution, *Stochastic environmental research and risk assessment*, 16, 5, 358-373.

Laugel, A. (2013), *Sea state climatology in the North-East Atlantic Ocean: analysis of the present climate and future evolutions under climate change scenarios by means of dynamical and statistical downscaling methods*, Ph.D. Thesis, Saint-Venant Laboratory for Hydraulics, Université Paris-Est, Chatou, France.

Leckebusch, G. C., Renggli, D. and U. Ulbrich (2008), Development and application of an objective storm severity measure for the Northeast Atlantic region, M*eteorologische Zeitschrift,* 17, 5, 575-587.

Madsen, H. and D. Rosbjerg (1997), The partial duration series method in regional index-flood modeling, *Water Resources Research*, 33, 4, 737–746.

Madsen, H. and D. Rosbjerg (1998), A regional Bayesian method for estimation of extreme streamflow droughts, *Statistical and Bayesian methods in hydrological sciences,* UNESCO, Paris, 327-340.

Madsen, H., Rasmussen, P. F. and D. Rosbjerg (1997a), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, *Water Resources Research*, 33, 4, 747-757.

Madsen, H., Pearson, C. P. and D. Rosbjerg (1997b), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. Regional modeling, *Water Resources Research*, 33, 4, 759-769.

Madsen, H., Mikkelsen, P. S., Rosbjerg, D. and P. Harremoës (2002), Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics, *Water Resources Research*, 38, 11, 21/1-21-11.

Mikkelsen, P. S., Madsen, H., Rosbjerg, D. and P. Harremoës (1996), Properties of extreme point rainfall III: Identification of spatial inter-site correlation structure, *Atmospheric Research*, 40, 77–98.

Pickands, J. (1975), Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics*, 3, 1, 119-131.

682 Pinto, J. G., Karremann, M. K., Born, K., Della-Marta, P. M. and M. Klawa (2012), Loss
683   potentials associated with European windstorms under future climate conditions,
684   *Climate Research*, 54, 1–20.
685 Reed, D. W. (1994), *Rainfall frequency analysis for flood design*, in Coping with Floods,
686   NATO ASI Series Volume 257, 59-75.
687 Renard, B. (2011), A bayesian hierarchical approach to regional frequency analysis, *Water*
688   *Resources Research*, 47, W11513, 2011.
689 Renard B. and M. Lang (2007), Use of a gaussian copula for multivariate extreme value
690   analysis : some case studies in hydrology, *Advances in Water Resources*, 30, 897–912.
691 Renggli, D., Leckebusch, G. C., Ulbrich U., Gleixner, S. N. and E. Faust (2011), The Skill of
692   Seasonal Ensemble Prediction Systems to Forecast Wintertime Windstorm Frequency
693   over the North Atlantic and Europe. *Monthly Weather Review*, 139, 3052–3068.
694 Ribatet, M. (2007), POT: Modelling Peaks Over a Threshold, *R News*, 7, 3.
695 Rosbjerg, D. (1985), Estimation in partial duration series with independent and dependent
696   peak values, *Journal of Hydrology*, 76, 183-195.
697 Rosbjerg, D. and H. Madsen (1996), The role of regional information in estimation of extreme
698   point rainfalls, *Atmospheric Research*, 42, 1-4, 113-122.
699 Roth, M., T. Buishand A., Jongbloed, G., Klein Tank, A. M. G. and J. H. van Zanten (2012),
700   A regional peaks-over-threshold model in a nonstationary climate, *Water Resources*
701   *Research*, 48, W11533.
702 Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R.,
703   Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J.,
704   Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist,
705   D., van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S.,
706   van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B.,
707   Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou,
708   C.-Z., Liu, Q., Chen, Y., Cucurull, L., Reynolds, R. W., Rutledge, G. and M. Goldberg
709   (2010), The NCEP Climate Forecast System Reanalysis, *Bulletin of the American*
710   *Meteorological Society*, 91, 1015–1057.
711 Scholz, F. W. and M. A. Stephens (1987), K-sample Anderson-Darling Tests, *Journal of the*
712   *American Statistical Association*, 82, 399, 918–924.
713 Smith, R. L. (1990), Regional estimation from spatially dependent data, *unpublished*.
714 Stedinger, J. R. (1983), Estimating a regional flood frequency distribution, *Water Resources*
715   *Research*, 19, 2, 503–510.
716 Stewart, E. J., Reed, D. W., Faulkner, D. S. and N. S. Reynard (1999), The FORGEX method
717   of rainfall growth estimation - I: Review of requirement, *Hydrology and Earth System*
718   *Sciences*, 3, 187-195.
719 Weiss, J., Bernardara, P. and M. Benoit (2014), Formation of homogeneous regions for
720   regional frequency analysis of extreme significant wave heights, submitted to *Journal*
721   *of Geophysical Research*.
722
723
724 **7   Table captions**
725

726   Table 1.   Measures of regional dependence for each region (with the number of sites N

727   indicated between parentheses): λ_r is the mean annual number of storms in the region, Φ is

728   the adimensional function of regional dependence, $\beta_s(1)$ is the mean number of impacted

729   sites during a storm, 100_r is the regional return period (in years) of the storm causing at

730   least one local 100-year event (equation (12) with $T = 100$) and D_{eff} is the regional

731   effective duration (in years, along with the 95% confidence interval).

732   Table 2.   Parameters of the regional distribution: γ (GPD scale parameter), k (GPD

733   shape parameter), $y_{0.99}$ (100-year regional return level).

734   Table 3.   Return periods (in years) of the storms with the highest normalized intensity

735   observed in each region: $\tilde{T}_{s,loc}$, $\overline{T}_{s,loc}$, $\tilde{T}_{s,reg}$ and $\overline{T}_{s,reg}$ are respectively the empirical local

736   return period, the theoretical local return period, the empirical regional return period and

737   the theoretical regional return period.

738

739 **8   <u>Figure captions</u>**
740

741   Figure 1.   Location of the 1847 sites extracted from the oceanic mesh of the ANEMOC-2

742   sea-state database.

743   Figure 2.   Footprints of the storms of a) 15-18 February 1986, b) 11-13 December 1990

744   and c) 23-24 January 2009 (Klaus), where red dots indicate the impacted sites.

745   Figure 3.   Map of threshold values of H_s exceeded on average once per year (m).

746   Figure 4.   Division into six homogeneous regions.

747   Figure 5.   Regional return period T_r against the local return period T for each region, as

748   defined in equation (12). Curves for regions 3 and 4 are superimposed.

749   Figure 6.   Return level plots of the regional distributions (crosses represent observations

750   from each regional sample), together with the 95% confidence intervals obtained by

751   bootstrap.

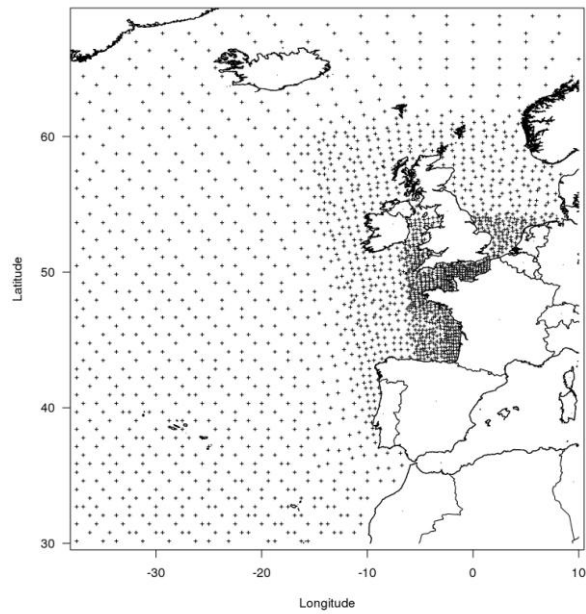752   Figure 7.   Map of estimated 100-year H_s (m).

Figure 1.   Location of the 1847 sites extracted from the oceanic mesh of the ANEMOC-2 sea-state
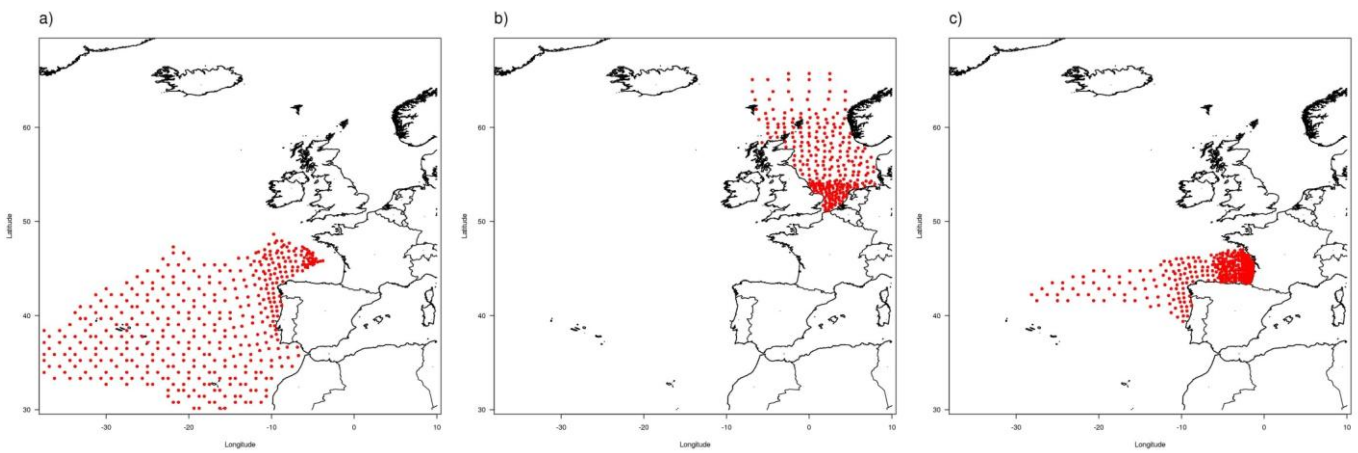
database.

Figure 2.   Footprints of the storms of a) 15-18 February 1986, b) 11-13 December 1990 and c) 23-24

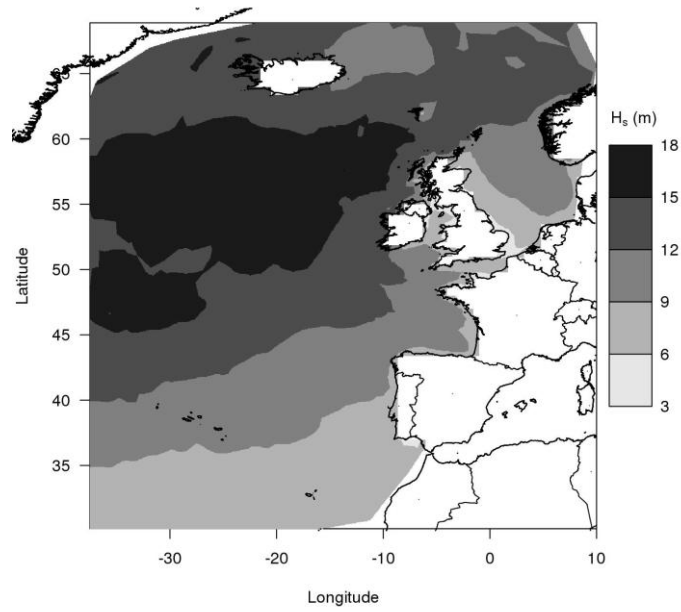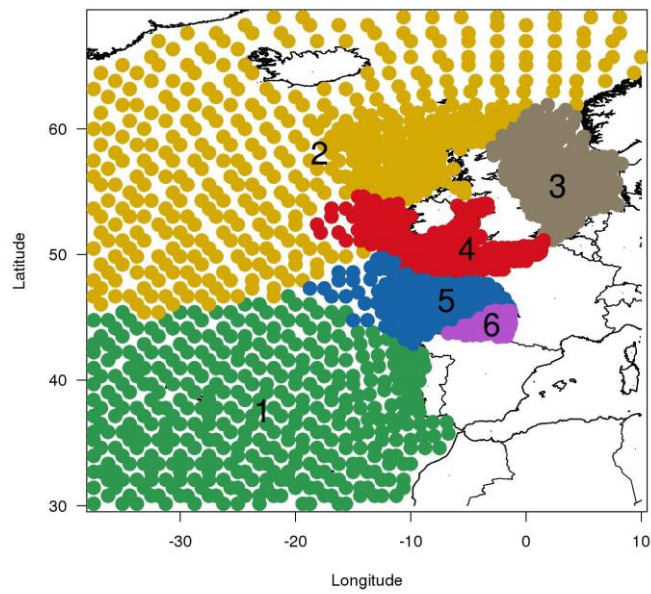January 2009 (Klaus), where red dots indicate the impacted sites.

Figure 3.   Map of threshold values of H_s exceeded on average once per year (m).

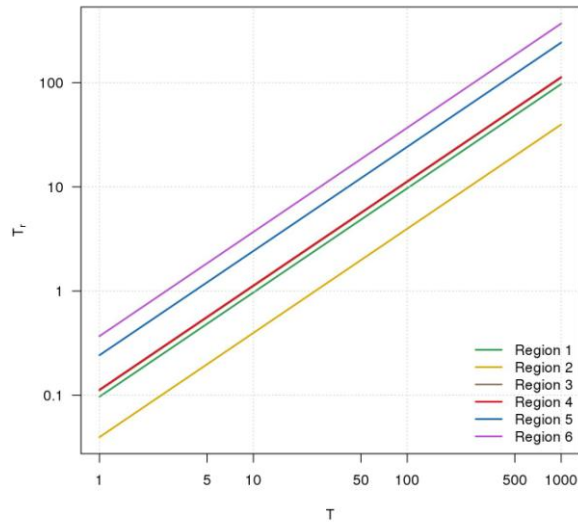Figure 4.   Division into six homogeneous regions.

Figure 5.   Regional return period T_r against the local return period T for each region, as defined in equation (12). Curves for regions 3 and 4 are superimposed.
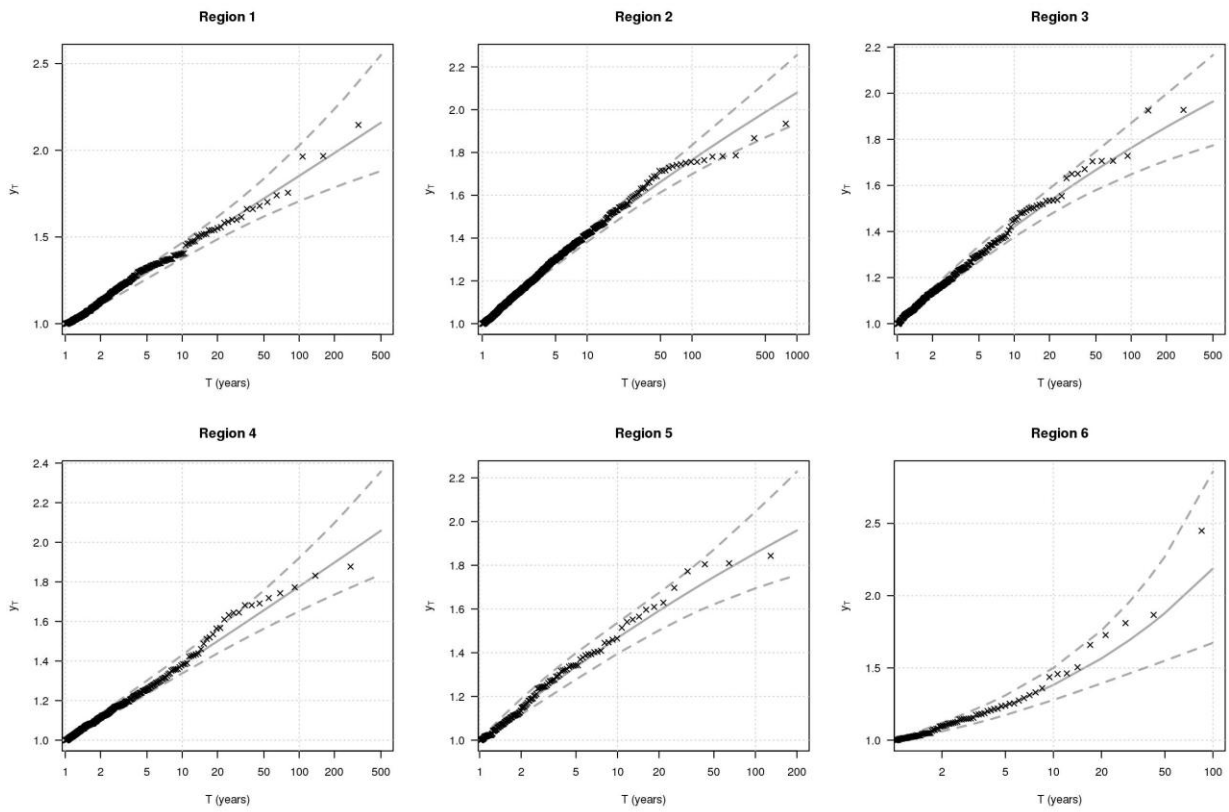

Figure 6.   Return level plots of the regional distributions (crosses represent observations from each regional sample), together with the 95% confidence intervals obtained by bootstrap.
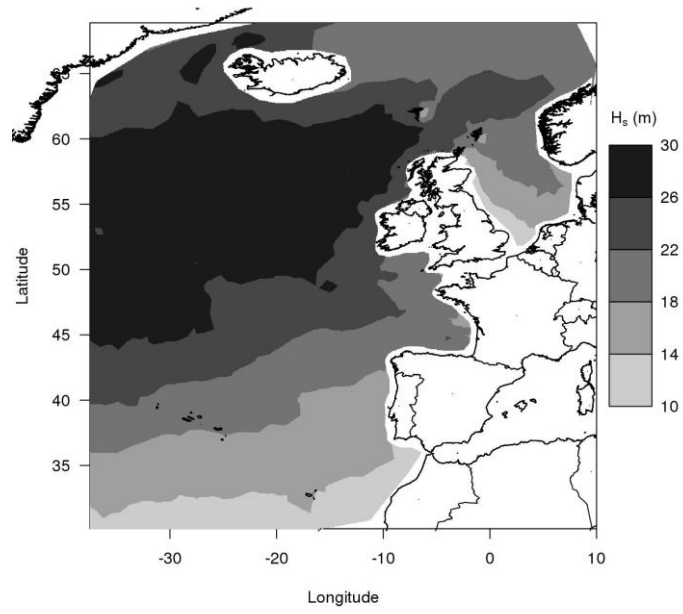
Figure 7.   Map of estimated 100-year H_s (m).