

Ifremer

Ifremer
Unité de Recherche - Recherches et Développemets Technologiques
Laboratoire Comportement des Structures en Mer
Yanira Guanche Garcia, Marc Prevosto, Christophe Maisondieu

BRGM
Direction DRP, Unité R3C
Thomas Bulteau, François Paris

Juillet 2015



Géosciences pour une Terre durable

brgm



MASTENV

Méthodes d'Analyse des Séries Temporelles Environnementales:

Valeurs extrêmes - Extreme Values



MASTENV

***Méthodes d'Analyse des Séries
Temporelles Environnementales***

Valeurs extrêmes - Extreme Values

**Y. Garcia, M. Prevosto, C. Maisondieu
T. Bulteau, F. Paris**

**Ifremer
BRGM**

Abstract

The knowledge of extreme events of environmental variables is an issue of increasing concern to the scientific community. Within the branch of coastal and ocean engineering there are many fields of application where an accurate estimation of long term return period events is needed, i.e. coastal defenses design, coastal flooding management, estimation of changes of the littoral morphology, offshore and onshore renewable energy devices design, etc. But not only the engineers are concerned by extreme events; biological communities in the open sea and estuarine systems are also exposed to extreme events that may affect its natural development (i.e. extreme sea levels in estuarine environments could raze fields of plants not able to deal with salt in just a few hours).

But the analysis of environmental variables and their extreme behaviors is not an easy task. In most cases the problem to be solved presents a multivariate nature, which makes it of a special complexity. For instance, in the case of flooding analysis, an estimation of joint probability density of astronomical tide, storm surge and waves is needed; or in the case of floating offshore devices the parameters of interest would be waves, winds and currents.

Complex mathematical methods and techniques are needed to be able to estimate the joint probability of occurrence of more than one variable at the same time. In addition, some considerations must be taken regarding the quality of the data in order to be able to guarantee the achievement of reliable estimations.

In this document, an analysis of the problem of extreme value analyses, the existing methods, their limitations, and some hints to proceed are exposed. The document is divided as follows: a first part will introduce the main problems or aspects that need to be considered concerning data and exposes the main methods to solve univariate problems and the way to estimate statistical parameters; a second part is focused on the main multivariate methods; then, a fourth part will focus on the principal techniques to estimate the uncertainties; finally in the last part the available softwares useful for EV analysis are exposed.

In the appendices is presented the application of these methods in two different case studies, one proposed by IFREMER and based on the design of an offshore structure mooring and the second one proposed by BRGM with a focus on a practical application to flood risk in coastal areas.

MASTENV

Méthodes d'analyse des Séries temporelles Environnementales

Index

1.	INTRODUCTION	7
2.	UNIVARIATE ANALYSIS	8
2.1.	Block Maxima vs. Peaks Over Threshold	8
2.2.	Parameters estimation methods	10
2.2.1.	Comparison between methods	11
2.2.2.	Comparison applied to a Generalized Extreme Value distribution fit	12
2.2.3.	Comparison applied to a Generalized Pareto distribution fit	14
2.2.4.	General comments on the GEV and GPD comparisons	17
2.3.	Statistical tests	17
2.3.1.	Akaike criterion	18
2.3.2.	Anderson-Darling test	18
2.3.3.	Cramér-von Mises criterion	18
2.3.4.	D'Agostino's K^2 test	18
2.3.5.	Jarque-Bera test	19
2.3.6.	Kolmogorov-Smirnov test	19
2.3.7.	Likelihood ratio test	20
2.3.8.	Pearson's χ^2 test	20
2.3.9.	PP plot	20
2.3.10.	QQ plot	21
2.3.11.	Spearman's rank correlation coefficient ρ	21
2.4.	Identified problems/limitations	21
2.4.1.	Sample Size	21
2.4.2.	Sample Variability	28
2.4.3.	Independence between events	31
2.4.4.	Covariates influence	34
2.4.5.	Plotting position	37
2.4.6.	Measurement scale influence	39
2.4.7.	Location influence	40
2.4.8.	Database influence	42
3.	MULTIVARIATE ANALYSIS	49
3.1.	Parametric models	49

3.2.	Conditional extreme models	50
3.3.	Copula models	50
3.4.	Extreme dependence models.....	51
3.5.	Max-stable models	51
3.6.	Bayesian inference models.....	52
3.7.	Methods of Reliability Estimation	52
3.7.1.	Level I.....	52
3.7.2.	Level II.....	53
3.7.3.	Level III.....	54
4.	CHARACTERISING UNCERTAINTIES	56
4.1.	Standard Likelihood Theory or Theoretical Delta Method.....	56
4.2.	Resampling Confidence Intervals or Bootstrapping Method	57
4.3.	Bayesian Confidence Intervals or Credible Intervals.....	57
5.	SOFTWARES.....	58
5.1.	R.....	58
5.2.	Matlab.....	63
5.3.	S/S-Plus	66
5.4.	Others	67
6.	REFERENCES	69
ANNEX 1. CASE STUDY 1.....		73
A.	Tension Model Determination	73
B.	T^{100} Estimation.....	77
a.	Direct Estimation of T^{100}	77
b.	Estimation of T^{100} by estimating H_s^{100} and W^{100} independently.....	78
c.	Estimation of T^{100} by estimating W^{100} conditioned to H_s^{100}	79
d.	Estimation of T^{100} by estimating H_s^{100} conditioned to W^{100}	80
e.	Heffernan and Tawn 2004.....	81
1.	<i>Extraction of declustered maxima.</i>	81
2.	<i>Pre-treatment of the marginal distributions</i>	82
3.	<i>Estimation of the dependence coefficients (a and b)</i>	83
4.	<i>Simulation process</i>	83
5.	<i>Estimation of T^{100} from the simulated data.</i>	85
f.	Mínguez et al 2013	85
1.	<i>Fitting the marginal distributions of W</i>	86
2.	<i>Fitting the conditional distributions of H_s</i>	86

3. Fitting the parameters to a 3 rd order polynomial.....	86
4. Application to the use of IFORMS.....	87
5. Estimation of T^{100} from the IFORMS.....	88
g. Estimation of T^{100} assuming that H_s and W are completely independent.....	88
C. Comparison	90
ANNEX 2. CASE STUDY 2.....	92
A. Case Study	92
B. Independent events selection.....	93
5-days block maxima	93
<i>FL-POT</i>	94
C. Multivariate Extreme Value Analysis	96
<i>Heffernan and Tawn 2004</i>	96
<i>Join Sea</i>	97
D. Cases	97
<i>Rescaling process</i>	97
E. Extreme simulations results	100
Case 1. 5 days block maxima + Join Sea	100
Case 2. 5 days block maxima + Heffernan and Tawn + Rescaling.....	100
Case 3. Flooding level POT + Join Sea.....	100
Case 4. Flooding level POT + Heffernan and Tawn	101
F. 100 years return period triplets selection	101
Case 1. 5 days block maxima + Join Sea	101
Case 2. 5 days block maxima + Heffernan and Tawn + Rescaling.....	102
Case 3. Flooding level POT + Join Sea.....	103
Case 4. Flooding level POT + Heffernan and Tawn	104
<i>Triplets selection</i>	105
G. Storms from the historical record (16.5 years)	107
H. Flooding simulations (MARS).....	107
I. Results	109
J. Comments	111

Figures Index

Figure 1. Comparison between Weibull, Gumbel and Fréchet distributions.....	9
Figure 2. Significant wave height time series (blue line) and annual maxima (red circles).	12
Figure 3. Return Period Comparison – GEV.....	14
Figure 4. Significant wave height time series (blue line) and peaks over a threshold of 8.75m (black stars).....	15
Figure 5. Return Period Comparison – GPD (threshold=8.75m).	17
Figure 6 GPD fit for different sample sizes.....	22
Figure 7. Empirical distributions of POT samples for the different considered periods of time with its GPD fit.	22
Figure 8. Evolution of scale and shape parameters of the GPD.....	23
Figure 9. GEV fit for different sample sizes.....	24
Figure 10. Annual maxima empirical distributions for the different lengths samples with its GEV fits.	24
Figure 11. Evolution of scale, shape and location parameters of the GEV distribution.	25
Figure 12. Gumbel fit for different sample sizes.....	25
Figure 13 Annual maxima empirical distributions for the different lengths samples with its Gumbel fits.....	26
Figure 14. Evolution of scale and location parameters of the Gumbel distribution.....	26
Figure 15. Exponential fit for different sample sizes.....	27
Figure 16. Empirical distributions of POT samples for the different considered periods of time with its GPD fit.	27
Figure 17 Evolution of the scale parameter of the Exponential distribution.....	28
Figure 18. Sample variability considering 6 different samples of 10 years length.....	29
Figure 19. Sample variability considering 3 different samples of 20 years length.....	30
Figure 20. Sample variability considering 2 different samples of 30 years length.....	31
Figure 21. Gumbel Probability Plot for different block sizes in a GEV analysis.....	Erreur ! Signet non défini.
Figure 22. Exponential Probability Plot for different independence lapses in a POT analysis.....	33
Figure 23. Seasonality.....	34
Figure 24. Comparison between winter (Oct-Mar) and summer (Apr-Sep) monthly maxima with Annual Maxima.	35
Figure 25 Seasonal [spring (MAM), summer (JJA), autumn (SON) and winter (DJF)] monthly maxima comparison.....	35
Figure 26 Monthly maxima comparison.	36
Figure 27. Seasonality of the GEV parameters.....	37
Figure 28. Rose directions (standard-left rose and quantiles-right rose).	37
Figure 29. Plotting position comparison.	39
Figure 30. Seasonal distribution of the 20y return period H_s (C. M. Izaguirre 2011).....	41
Figure 31. Spatial variability of the shape parameter of extreme wave height distribution around the world. Only statistically significant values at the 90% confidence interval are dotted (from Izaguirre et al 2011).....	41
Figure 32. Spatial variability of the shape parameter of extreme wave height distribution using the NOAA database.	42
Figure 33. Databases time series (IHData-Grey, ERA40-Purple, HOMERE-green, NOAA-red and BoBWA-blue).....	44
Figure 34. GEV distribution at 2.5W 45N for the different databases.....	45
Figure 35. Gumbel distribution at 2.5W 45N for the different databases.....	46
Figure 36. GP distribution 2.5W 45N location for the different databases.	47
Figure 37. Exponential distribution at 2.5W 45N for the different databases.....	48
Figure 38. Limit state surface graphical and its linear approximation for a 2-D case.....	54

Figure 39. GEV and Gumbel fittings for H_s annual maxima.....	75
Figure 40. GPD and Exponential fittings for H_s peaks over threshold.....	75
Figure 41. GEV and Gumbel fittings for W annual maxima.....	76
Figure 42. GPD and Exponential fittings for W peaks over threshold.....	76
Figure 43. GEV and Gumbel fittings for T annual maxima.....	78
Figure 44. GPD and Exponential fittings for W peaks over threshold.....	78
Figure 45. Heteroscedastic model fit to W conditioned to H_s	79
Figure 46. Heteroscedastic model fit to H_s conditioned to W	80
Figure 47. Peaks Over Threshold of T	81
Figure 48. Declustered maxima.....	82
Figure 49 H_s and W samples.....	84
Figure 50. H_s - W joint distribution comparison between empirical and sampled data.....	84
Figure 51. Third order polynomial fitted to the GEV parameters.....	87
Figure 52. Third order polynomial fitted to the GPD parameters.....	87
Figure 53. H_s - W contour plot.....	88
Figure 54. GEV and Gumbel fittings for T annual maxima.....	89
Figure 55. GPD and Exponential fittings for W peaks over threshold.....	90
Figure 56. Study location and data points of the different variables.....	92
Figure 57. Time series of wave height at P81, river level and sea water level.....	93
Figure 58. Graphical scheme of the 5-days block maxima procedure followed.....	94
Figure 59. River mouth level scheme.....	95
Figure 60. α estimation.....	95
Figure 61. Flooding level peaks over threshold.....	96
Figure 62. Scheme of the rescaling process.....	98
Figure 63. Marginals rescaled.....	99
Figure 64. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 1.....	100
Figure 65. Joint distributions of the 3 variables sampled and rescaled(H_s , SWL and RL).Case 2.....	100
Figure 66. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 3.....	100
Figure 67. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 4.....	101
Figure 68. 100-years return period surface of H_s , SWL and RL.....	101
Figure 69. 100-years return period contours of H_s and SWL for different values of RL.....	102
Figure 70. 100-years return period surface of H_s , SWL and RL.....	102
Figure 71. 100-years return period contours of H_s and SWL for different values of RL.....	103
Figure 72. 100-years return period surface of H_s , SWL and RL.....	103
Figure 73. 100-years return period contours of H_s and SWL for different values of RL.....	104
Figure 74. 100-years return period surface of H_s , SWL and RL.....	104
Figure 75. 100-years return period contours of H_s and SWL for different values of RL.....	105
Figure 76. Example of the triplets selection procedure.....	106
Figure 77. Left: H_s - St scatter of the 16.5 y time series and the 100 y return period triplets. Right: H_s - Tp scatter of the 16.5 y time series and the 100 y return period triplets.....	106
Figure 78. Storm profile normalized.....	107
Figure 79. Storm profiles from the historical records.....	108
Figure 80. Delay (hours) between variables peaks for the historical storms.....	109
Figure 81. Statistics of the delays estimated.....	109
Figure 82. Results in the river mouth level.....	110
Figure 83. Results in the max flooded surface.....	110

Tables Index

Table 1. GEV parameters estimate using different methods and/or softwares.	13
Table 2. GPD parameters estimate using different methods and/or softwares.	16
Table 3. Threshold that provides a sample of peaks over threshold with $\lambda=1$ for the different time series.	22
Table 4. GEV parameters for the different 10-years length samples.	29
Table 5. GEV parameters for the different 20-years length samples.	30
Table 6. GEV parameters for the different 30-years length samples.	30
Table 7. Plotting position formulas.	38
Table 8. Return period (years) for the 12.61m event according the different plotting position formulae.	38
Table 9. Comparison of the measurement scale influence with the variable H_s	39
Table 10. Comparison of the measurement scale influence with the variable W	40
Table 11. Summary of the characteristics of the different databases used.	44
Table 12. GEV distribution parameters estimated for the different databases.	45
Table 13. Gumbel distribution parameters estimated for the different databases.	46
Table 14. GP distribution parameters estimated for the different databases.	46
Table 15. Exponential distribution parameters estimated for the different databases.	47
Table 16. Parameters fitted for H_s	74
Table 17. 100 years return period estimations for H_s	74
Table 18. Parameters fitted for W	74
Table 19. 100 years return period estimations for W	74
Table 20. Parameters fitted for T	77
Table 21. 100 years return period estimations for T	77
Table 22. Heteroscedastic parameters estimated.	79
Table 23. Heteroscedastic parameters estimated.	80
Table 24. Distribution fits.	86
Table 25. Distribution fits.	89
Table 26. 100 years return period for T	89
Table 27. Comparison of the results obtained for T^{100}	90
Table 28. Summary of the triplets selected.	106

1. INTRODUCTION

The knowledge of extreme events of environmental variables is an issue of increasing concern to the scientific community. Within the branch of coastal and ocean engineering there are many fields of application where an accurate estimation of long term return period events is needed, i.e. coastal defenses design, coastal flooding management, estimation of changes of the littoral morphology, offshore and onshore renewable energy devices design, etc. But not only the engineers are concerned by extreme events; biological communities in the open sea and estuarine systems are also exposed to extreme events that may affect its natural development (i.e. extreme sea levels in estuarine environments could raze fields of plants not able to deal with salt in just a few hours).

But the analysis of environmental variables and their extreme behaviors is not an easy task. In most cases the problem to be solved presents a multivariate nature, which makes it of a special complexity. For instance, in the case of flooding analysis, an estimation of joint probability density of astronomical tide, storm surge and waves is needed; or in the case of floating offshore devices the parameters of interest would be waves, winds and currents.

Complex mathematical methods and techniques are needed to be able to estimate the joint probability of occurrence of more than one variable at the same time. In addition, some considerations must be taken regarding the quality of the data in order to be able to guarantee the achievement of reliable estimations.

In this document, an analysis of the problem of extreme value analyses, the existing methods, their limitations, and some hints to proceed are exposed. The document is divided as follows: a first part will introduce the main problems or aspects that need to be considered concerning data and exposes the main methods to solve univariate problems and the way to estimate statistical parameters; a second part is focused on the main multivariate methods; then, a fourth part will focus on the principal techniques to estimate the uncertainties; finally in the last part the available softwares useful for EV analysis are exposed.

In the appendices is presented the application of these methods in two different case studies, one proposed by IFREMER and based on the design of an offshore structure mooring and the second one proposed by BRGM with a focus on a practical application to flood risk in coastal areas.

2. UNIVARIATE ANALYSIS

2.1. Block Maxima vs. Peaks Over Threshold

There are two main methods to solve univariate extreme value analysis: block maxima or peaks over threshold methods.

In block maxima methods, the time series is divided into consecutive blocks of data with a constant size. From each block the maxima is extracted, constituting these way the extremes sample.

A sample of block maxima can be assumed to follow a Generalized Extreme Values distribution (GEV) (Leadbetter G. 1983). This distribution takes the form:

$$GEV(x; \mu, \sigma, \xi) = \exp \left\{ - \left(1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right)^{-1/\xi} \right\}$$

Where μ , σ and ξ are the location, scale and shape parameters, respectively.

Depending on the value of the shape parameter, the GEV distribution is also known as Weibull distribution for $\xi < 0$; Gumbel distribution in the case of $\xi = 0$ and Fréchet when $\xi > 0$.

On the other hand, peaks over threshold methods are based on creating the sample from exceedances above a certain threshold.

A sample of exceedances above a high threshold can be approached by a Generalized Pareto Distribution (GPD) (Pickands 1975), which can be written as:

$$GPD(x; u, \sigma, \xi) = 1 - \left(1 + \xi \frac{x - u}{\sigma} \right)^{-1/\xi}$$

Where σ and ξ are the scale and shape parameters, respectively; and u is the threshold.

Considering the value of the shape parameter, the GP Distribution presents different behavior: for $\xi < 0$ (Weibull domain), it presents an upper bound at $-\sigma/\xi$, in the case of $\xi = 0$ it follows an exponential distribution and for $\xi > 0$ (Fréchet domain) there is no upper bound.

In Figure 1, it is graphically represented the different behavior of the GP distribution depending on the value of the shape parameter. As seen, the Weibull distribution (pink line) presents an asymptotic horizontal behavior, while the Fréchet distribution (purple line) shows a high probability of occurrence of great values.

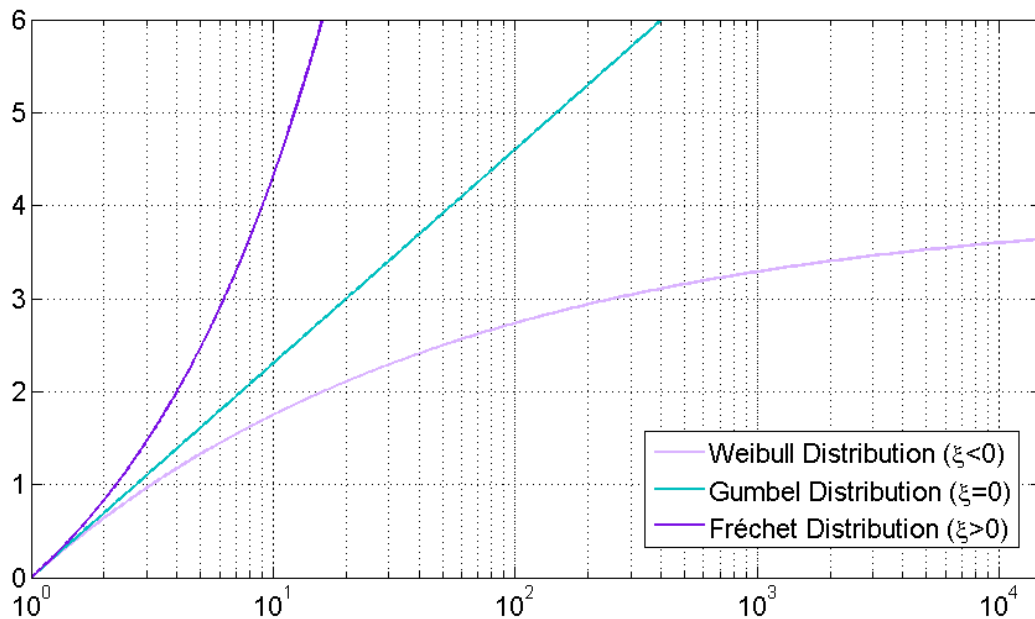


Figure 1. Comparison between Weibull, Gumbel and Fréchet distributions.

The decision of considering a block maxima or a POT method is not trivial and must be taken considering the problems that each one presents.

Peaks over threshold methods have two main disadvantages: threshold selection and independence between events:

1. Threshold selection: the threshold will influence the tail of the distribution and it should be selected by an agreement between bias and variance. The threshold should be high enough to be in the domain of asymptotic validity of the GPD in order to limit bias in the parameters estimation, but it should not be too high to limit the variance of the estimated parameters. There are two main tools to help in the threshold selection (Méndez 2006), (Jonathan 2013), (Davison 1990), (R. L. Smith 2004):

- Mean residual life plot
- Stability of shape and modified scale parameters.

2. Independence between events: only by setting a threshold the independence between events is not guaranteed. So, a minimum time-lapse between peaks should be established. This time-lapse should consider the typical storm duration, thus is dependent of the study site location (Méndez 2006).

Block maxima methods present the disadvantage of establishing the block size. This can be easily avoided by assuming yearly maxima. Then the secondary problem that may arise is the limited length of the resulting sample (only 1 event per year) and thus higher uncertainties in estimated parameters. The independence between extremes is almost guaranteed when considering yearly maxima.

2.2. Parameters estimation methods

Once the model has been selected, there are several methods to estimate the associated parameters.

Here, we provide a brief description of the main methods used and a comparison between them, according to (Mackay 2011).

Methods:

- **Moment based methods:** based on obtaining the estimators by equating the sample mean and variance with the population mean and variance and rearranging:
 - Method of moments (MOM)
 - Probability weighted moments (PWM)
 - Generalized probability weighted moment (GPWM)MOM, PWM and GPWM always exist and are easy to compute. For samples where shape is negative, they can produce estimates low biased but with large RMSE.
 - Hybrid-MOM:** always provide a solution, either the one from MOM or the one from PWM
 - L-moments:** used mainly in hydrological problems. It is a particular case of PWM, it can be expressed as a linear combination of PWMs
- **Maximum Likelihood (ML):** consists of maximizing the likelihood function of independent observations. There are some functions for which no ML solution exists, especially when the shape parameter is negative (Hosking 1987). A derived form of the ML method is the penalized maximum likelihood (PML), where the likelihood function is multiplied by a penalty function. With this penalty function the possible values of the shape parameters can be constrained.
- **Maximum entropy (ME):** based on maximizing the entropy equation. The estimators obtained for the GPD with the maximum entropy method are equivalent to those obtained with maximum likelihood if the threshold is chosen before the parameters are estimated.
- **Likelihood moment (LM):** is a hybrid solution between moment-based methods and likelihood methods. This method always provides a solution, is always feasible and easy to compute.
- **Least-squares minimization (LS):** there are two types of least-squares methods:
 - Minimizing the sum of squared differences between empirical and model quantiles (Moharran et A. K. and Kapoor 1993).
 - Minimizing the squared differences between empirical and model distribution fits (Luceño 2006).While the second method could provide results strongly biased for small, the first one could have convergence problems.

- **Empirical percentile method (EPM):** based on equating percentiles of the empirical and theoretical distribution functions.
It always provides a solution and it is always feasible but it can be very computationally demanding for large samples.
- **Bayesian methods (BM):** They can produce very good results but they are difficult to implement and are often computationally demanding. To achieve a better efficiency, (J. a. Zhang 2009) proposed a new Bayesian method based on maximum likelihood. They can require an expert knowledge to provide initial values, and thus, could result in a subjective method.
- **Robust methods (RM):** These methods are less sensitive to outliers (outliers are those events that are distant from the rest of observations, (Grubbs 1696)), thus, their use in extreme values analysis is under controversy because these highest values (the outliers) introduce important information about the tail behavior into the model. Moreover it has been proved that these methods do not perform as well as ML method (Mackay et al in 2011 mentioned (Juarez 2004)).

2.2.1. Comparison between methods

For positive shape parameter:

- Moment methods present lower RMSE (root-mean-square-error) but a strong bias.
- ML, LS, EPM and BM present strong bias and high RMSE for small sample sizes ($n < 50$).
- For large sample sizes ($n > 200$) BM has lower bias and its RMSE is similar to the ML but BM is faster to be computed (when using the method proposed by (J. a. Zhang 2009)).
- LM presents the lowest RMSE for small sample sizes.

For negative shape parameter:

- Moment methods present lower bias but a larger RMSE.
- LM consistently has lower RMSE and small bias.
- ML and PML present small bias and low variance but sometimes do not reach a solution for small samples.
- LS has low bias but not so low RMSE.
- EPM presents strong bias and high RMSE for small sample sizes, decreasing at large sample sizes (but it is very slow to compute).
- BM presents high bias and RMSE for small sample sizes but low bias and RMSE for large sample sizes and it is very fast to compute (when using the method proposed by (J. a. Zhang 2009)).

To show the differences between some of the methods proposed to estimate the parameters two examples were carried on.

2.2.2. Comparison applied to a Generalized Extreme Value distribution fit

Considering a significant wave height time series from the IHData database (2.4.8) located at the coordinates 2.5W 45N, with a temporal coverage of 67 years (hourly data from January 1948 until February 2014), a block maxima was used to obtain the sample of extremes. The size of the block was one year, so the 67 annual maxima of significant wave height were extracted from the time series. Figure 2 shows the hourly time series of H_s at the selected location where the annual maxima are marked with red circles.

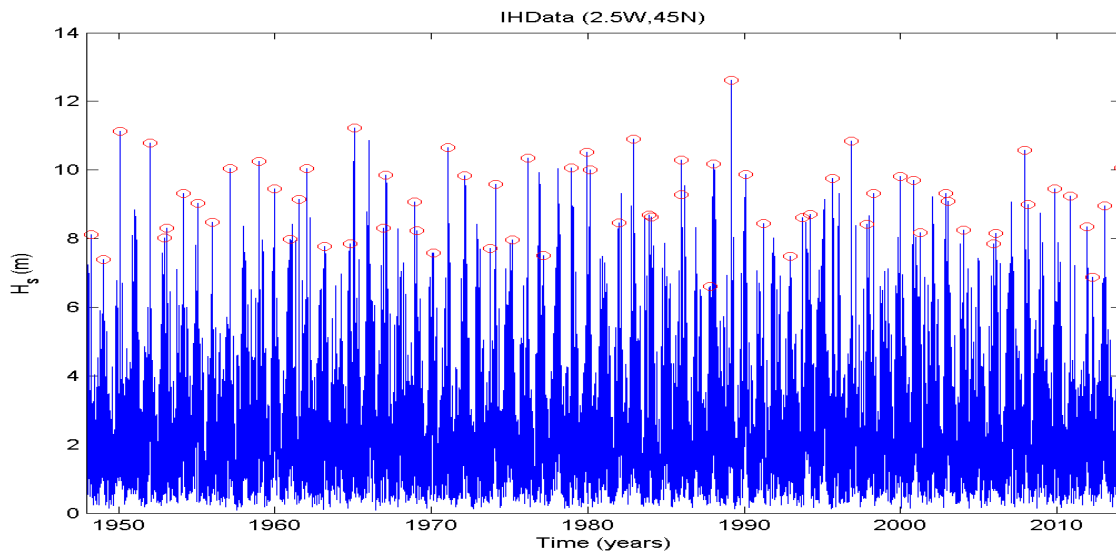


Figure 2. Significant wave height time series (blue line) and annual maxima (red circles).

Parameters estimation:

For the parameters estimation different softwares or packages were used (5):

Matlab: the already implemented functions available in Matlab.

WAFO with its available options:

- PWM-Probability Weighted Moments Method.
- ML-Maximum Likelihood Method.
- MPS-Maximum Product of Spacings Method.

EVIM.

extRemes with its optimization options:

- ML-NM-Maximum Likelihood method with Nelder-Mead optimization method (Nelder 1965).
- ML-CG-Maximum Likelihood method with Conjugated Gradients optimization method (Fletcher 1965).
- ML-BFGSqN-Maximum Likelihood method with BFGS quasi-Newton optimization method.
- ML-NM-Maximum Likelihood method with L-BFGS-B optimization method.
- ML-NM-Maximum Likelihood method with simulated annealing Belisle optimization method (Belisle 1992).

In the table below, all the obtained results are summarized. Note that for the results from the extRemes package instead of Confidence Intervals the value refers to the root mean square error of the parameter estimation.

		Mu	Mu-CI or RMSE	Sigma	Sigma-CI	Xi	Xi-CI or RMSE
MATLAB	ML	8.663	(8.374, 8.952)	1.095	(0.910, 1.318)	-0.183	(-0.326, -0.039)
WAFO	PWM	8.669	(8.368, 8.972)	1.130	(0.915, 1.345)	-0.207	(0.037, 0.377)
	ML	8.663	(8.393, 8.933)	1.095	(0.895, 1.296)	-0.183	(0.049, 0.317)
	MPS	8.634	(8.347, 8.922)	1.153	(0.922, 1.384)	-0.161	(-0.021, 0.344)
EVIM	ML	8.663	--	1.095	--	-0.183	--
extRemes	ML-NM	8.664	(0.14742)	1.095	(0.10351)	-0.183	(0.07322)
	ML-CG	8.663	(0.14741)	1.095	(0.10349)	-0.183	(0.07324)
	ML-BFGSqN	8.663	(0.14741)	1.095	(0.10349)	-0.183	(0.07324)
	ML-LBFGS-B	8.663	(0.14741)	1.095	(0.10349)	-0.183	(0.07324)
	ML-SAB	8.662	(0.14756)	1.096	(0.10375)	-0.180	(0.07468)

Table 1. GEV parameters estimate using different methods and/or softwares.

Comparison:

- The results proposed by almost every method (except WAFO-MPS) give similar estimations.
- Within the options available in the extRemes software, for the data analyzed, the optimization method 'Conjugate Gradients' did not get an optimum and the maximum number of iterations was reached.
- With the software extRemes, the standard error, confidence intervals and the covariate matrix of the parameters is directly obtained when fitting the data. With WAFO and Matlab, the confidence intervals of the parameters are given.
- Instead of providing different methods to fit the model, the extRemes software provides different optimization methods to achieve a solution with maximum likelihood estimators: this may offer a solution for cases where computational time becomes a problem.

Next figure shows the empirical data (annual maxima) against all the fittings estimated. The differences between Matlab, EVIM and extRemes are negligible.

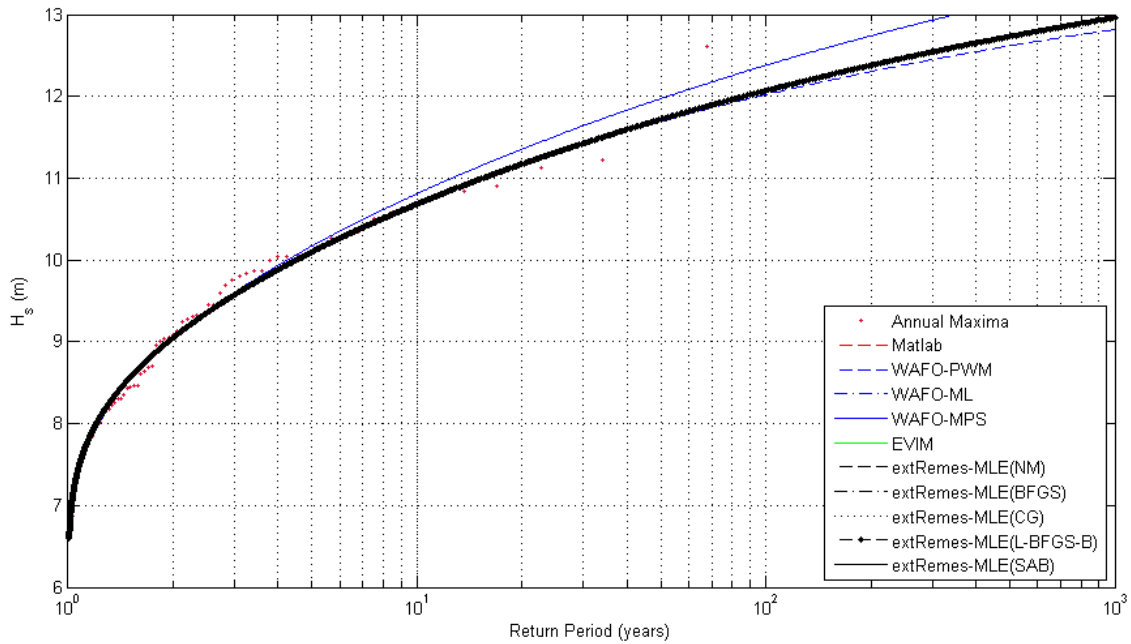


Figure 3. Return Period Comparison – GEV.

2.2.3. Comparison applied to a Generalized Pareto distribution fit

Taking as initial data the same significant wave height time series as the previous comparison; a Generalized Pareto Distribution was fitted to independent peaks over threshold of H_s . In order to obtain a similar number of events as in the annual maxima tests shown before (2.2.2), the threshold was set at 8.75 m and the independence temporal criterion between peaks was set at 3 days (72 hours). This threshold was chosen in order to get a λ (mean number of events per year) close to 1; this way the obtained results could be compared to a GEV fitting to annual maxima. The independence temporal criterion between events has been set at 3 days due to the location of the data; in the North Atlantic area it can be assumed a storm development time of 3 days (Méndez 2006). Figure 4 shows the hourly time series of H_s at the selected location where the independent peaks over the threshold are marked with black stars.

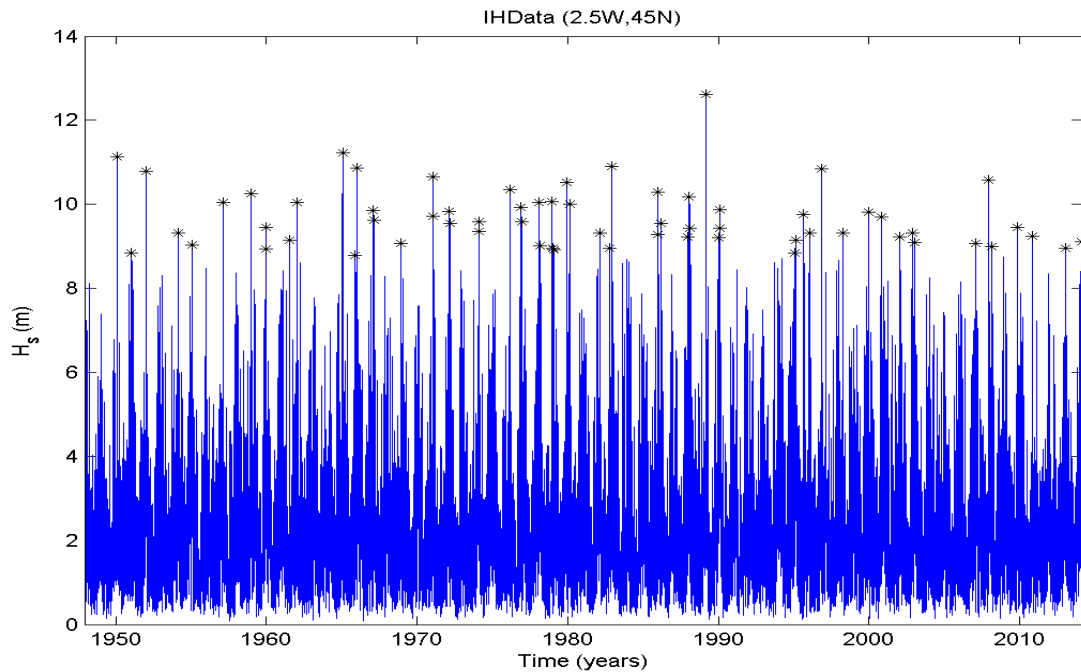


Figure 4. Significant wave height time series (blue line) and peaks over a threshold of 8.75m (black stars).

Parameters estimation:

For the parameters estimation different softwares or packages were used:

Matlab: the already implemented functions available in Matlab.

WAFO with its available options:

- PWM-Probability Weighted Moments Method.
- ML-Maximum Likelihood Method.
- MPS-Maximum Product of Spacings Method.
- MOM-Method of Moments.
- LS-Least Squares Method.
- Pickand's-Pickand's Estimator Method.

EVIM.

extRemes with its optimization options:

- ML-NM-Maximum Likelihood method with Nelder-Mead optimization method (Nelder 1965).
- ML-CG-Maximum Likelihood method with Conjugated Gradients optimization method (Fletcher 1965).
- ML-BFGSqN-Maximum Likelihood method with BFGS quasi-Newton optimization method.
- ML-NM-Maximum Likelihood method with L-BFGS-B optimization method.
- ML-NM-Maximum Likelihood method with simulated annealing Belisle optimization method (Belisle 1992).

It has been also implemented the Likelihood Moment Estimation proposed by (J. Zhang 2007). The obtained results for shape and scale parameters are similar to the solution

provided by the Maximum Likelihood Method but this method avoids convergence issues, computational difficulties and the absence of solution for shape parameters higher than one.

In the table below, all the obtained results are summarized. Note that for the results from the extRemes package instead of Confidence Intervals the value refers to the root mean square error of the parameter estimation.

		Sigma	Sigma-CI or RMSE	Xi	Xi-CI or RMSE
MATLAB	ML	1.144	(-0.057, 1.530)	-0.222	(-0.387, 0.856)
	LM (J. Zhang 2007)	1.144	--	-0.222	--
WAFO	PWM	1.354	(0.834,1.873)	-0.434	(0.069,0.798)
	ML	1.144	(0.805,1.483)	-0.222	(0.053,0.391)
	MPS	1.108	(0.747,1.468)	-0.157	(-0.108,0.422)
	MOM	1.264	(0.820, 1.709)	-0.339	(0.056, 0.622)
	LS	1.354	(0.834, 1.873)	-0.434	(0.069, 0.798)
	Pickand's	1.330	(NaN,NaN)	-0.576	(NaN,NaN)
EVIM	ML	1.144	--	-0.222	--
extRemes	ML-NM	1.144	(0.169)	-0.222	(0.084)
	ML-CG	1.143	(0.169)	-0.222	(0.084)
	ML-BFGSqN	1.143	(0.169)	-0.222	(0.084)
	ML-LBFGS-B	1.143	(0.169)	-0.222	(0.084)
	ML-SAB	1.145	(0.169)	-0.221	(0.085)

Table 2. GPD parameters estimate using different methods and/or softwares.

Comparison:

- Compared with GEV results, there are greater variations within results obtained from different methods. For this case, WAFO offers a greater amount of methods available to solve the fitting process.
- From the options available in the WAFO software, and for the data analyzed, solution achieved with a Maximum Product of Spacings method gives highest return levels of wave heights while the Pickand's estimator gives the lowest results.
- Comparing different softwares but the same method (maximum likelihood estimator), the solution achieved is the same (black line overlapping red, pink and green lines in Figure 6).
- With the software extRemes, the standard error, confidence intervals and the covariate matrix of the parameters are directly obtained when fitting the data. With WAFO and Matlab, the confidence intervals of the parameters are given.

Next figure shows the empirical data (peaks over a threshold of 8.75 meters of wave height) against all the fittings estimated.

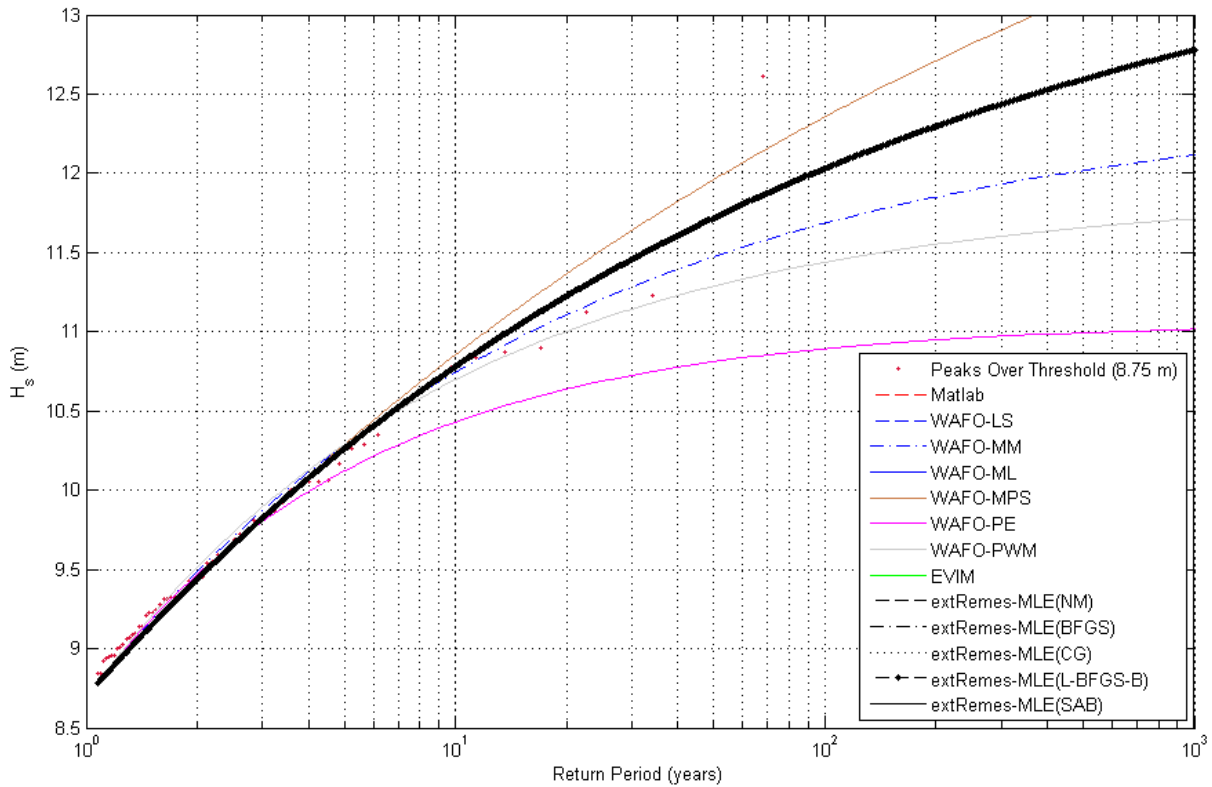


Figure 5. Return Period Comparison – GPD (threshold=8.75m).

2.2.4. General comments on the GEV and GPD comparisons

- In the WAFO software they recommend using PWM method since it works for a wider range of parameters and have the same asymptotic properties as the Maximum Likelihood (ML) method.
- The Pickand's estimator method and Least Squares (LS) estimator also work for any value of the shape parameter ξ
- The ML method is only useful when $\xi \leq 1$, the PWM when $\xi > -0.5$ and the MOM when $\xi > -0.25$.
- The variances of the ML estimates are usually smaller than those of the other estimators. However, for small sample sizes (less than 50 events) it is recommended to use the PWM, MOM or MPS if they are valid.

2.3. Statistical tests

Goodness-of-fit tests and measures encompass all the methods and techniques available to determine how well fits a model to the actual observations. The majority of these tests try to quantify the discrepancy between the model and the observations but there are also graphical tools that might be helpful.

Some of those statistical tests that may be helpful in the model fitting process to validate or not a given statistical model are summarized below. More detailed information and examples of the different methods and tests to assess in any statistical analysis could be found in (Corder 2014), (Greenwood 1996)

2.3.1. Akaike criterion

Akaike criterion is a measure of the relative quality of a statistical model, for a given dataset:

$$AIC = 2k - 2 \ln(L)$$

Where k is the number of parameters in the statistical model and L is the maximized value of the likelihood function for the estimated model.

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value.

2.3.2. Anderson-Darling test

The Anderson-Darling test is a statistical test to decide whether a given data sample of data follows a given probability distribution or not.

$$n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 w(x) dF(x)$$

Where the hypothesized distribution is $F(x)$, the empirical sample is $F_n(x)$ and $w(x)$ is a weighting function.

If $w(x) = 1$, this test is the same as the Cramér-von Mises test. (T. W. Anderson, Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes 1952) (T. W. Anderson, A test of goodness-of-fit 1954).

2.3.3. Cramér-von Mises criterion

It is a criterion used for judging the goodness of fit of a cumulative distribution function compared to a given empirical distribution function, or for comparing two empirical distributions. Applied to one sample takes the form:

$$T = n\omega^2 = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{2i-1}{2n} - F(x_i) \right]^2$$

Where n is the size of the sample and $F(x)$ is the hypothesized distribution.

If the value of T is larger than the tabulated value (this value depends on the size of the sample and on the significance level), then the hypothesis that the empirical data follow the distribution $F(x)$ should be rejected. (T. W. Anderson 1962)

2.3.4. D'Agostino's K^2 test

This is a test to establish whether or not a sample comes from a normally distributed population. The test is based on transformations of the sample kurtosis and skewness

coefficients, and combining them to produce a statistic that is χ^2 distributed when the sample follows a normal distribution.

From a sample, its skewness and kurtosis can be defined as:

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}} \quad g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3$$

Where n is the sample size, and \bar{x} is the sample mean.

Transforming g_1 and g_2 into Z_1 and Z_2 as proposed by (R. D'Agostino 1970), the statistics K^2 can be estimated:

$$K^2 = Z_1(g_1)^2 + Z_2(g_2)^2$$

The hypothesis of the sample being normally distributed will be accepted if the statistics K^2 is χ^2 -distributed (R. B. D'Agostino 1990)

2.3.5. Jarque-Bera test

The Jarque-Bera test is a goodness of fit test to determine if a sample data have the skewness and kurtosis that corresponds to a normal distribution.

If the data comes from a normal distribution, the JB statistic asymptotically has a χ^2 distribution with two degrees of freedom.

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4}(K-3)^2 \right)$$

Where n is the sample size, S is the sample skewness and K the sample kurtosis. (Jarque, Efficient tests for normality, homocedasticity and serial independence of regression residuals 1980) (Jarque, Efficient tests for normality, homocedasticity and serial independence of regression residuals: Monte Carlo evidence 1981) (Jarque, A test for normality observations and regression residuals 1987).

2.3.6. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov is a nonparametric test for one dimensional distributions that can be used to compare a sample with a reference probability distribution, or to compare two samples. The Kolmogorov-Smirnov statistic for a given empirical distribution $F_n(x)$ compared to a known distribution $F(x)$ is:

$$D_n = \sup_x |F_n(x) - F(x)|$$

Where \sup denotes the supremum of the set of distances.

If the sample $F_n(x)$ comes from the distribution $F(x)$, then D_n converges to 0 in the limit when n goes to infinity (Massey 1951). Thus, in practice, this statistic is better for large samples.

2.3.7. Likelihood ratio test

The likelihood ratio test is a statistical test used to compare the fit of two models. It requires nested models; it means that if we want to compare a null model which considers a certain number of parameters x with an alternative model that considers y parameters, all the x parameters considered in the null model must be also considered in the alternative model.

$$D = -2 \ln \left(\frac{\text{likelihood for null model}}{\text{likelihood for alternative model}} \right)$$
$$= -2 \ln(\text{likelihood for null model}) + 2 \ln(\text{likelihood for alternative model})$$

The alternative model fits significantly better and should thus be preferred to the null model if the difference D is higher than χ^2 with n degrees of freedom. Being n the increment of parameters between the two models compared ($y-x$).

2.3.8. Pearson's χ^2 test

This test can be used to assess two types of comparison: tests of goodness of fit and tests of independence. A test of goodness of fit establishes whether or not an observed frequency distribution differs from a theoretical distribution. A test of independence assesses whether paired observations on two variables, expressed in a contingency table, are independent of each other.

The test statistic for a goodness of fit comparison takes the form:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Where O are the observed values, E the expected ones (expected according to the hypothetical distribution assumed) and n the size of the sample. A step-by-step guide of these tests can be found in (Greenwood 1996).

2.3.9. PP plot

The PP plot is a probability plot that helps in assessing how close two datasets are. The two cumulative distribution functions (cdfs) of the datasets are plotted against each other. The pattern of points in the plot is used to compare the two distributions. If the two distributions being compared are identical, the P-P plot follows the 45° line $y = x$.

2.3.10. QQ plot

The QQ plot is a graphical method for comparing two probability distributions by plotting their quantiles against each other. The pattern of points in the plot is used to compare the two distributions. If the two distributions being compared are identical, the Q–Q plot follows the 45° line $y = x$.

2.3.11. Spearman's rank correlation coefficient ρ

The Spearman's rank correlation coefficient ρ is a nonparametric measure of statistical dependence between two variables. If there are no repeated data values, a perfect Spearman correlation of $+1$ or -1 occurs when each of the variables is a perfect monotone function of the other.

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

Where x and y are the two variables and \bar{x} and \bar{y} are their means. More information regarding this coefficient and other nonparametric measures can be found in (Corder 2014).

2.4. Identified problems/limitations

The analysis of environmental variables encompasses multiple particularities. Often the available data to work with are not good enough in terms of recorded data size or quality of the recorded data. Moreover, there are also aspects that have to be considered when applying extreme value analysis to environmental data such as dependences existent between events or the presence of covariates in the variable analyzed that may have an influence in our study. Here some of these limitations and concerns are going to be depicted and graphically illustrated through examples.

2.4.1. Sample Size

The size of the sample we are working with may become a problem when dealing with extreme value analysis because the uncertainty in parameters and return value estimates will increase (Jonathan 2013).

To graphically see the influence of the sample size in extreme value analysis some tests were carried out. All these tests have been made considering a time series of significant wave height at the coordinates 2.5W 45N. These data comprise 67 years of hourly significant wave heights (1948-2014).

GPD fitting comparison with different time series lengths

From the time series, different sample sizes were taken and a Peaks Over Threshold method was applied considering a mean number of events per year equal to 1 and independence lapse of 3 days (Méndez 2006). With these conditions, the resulting thresholds are those shown in the table below. Fitting a GPD for all these cases and comparing them graphically we obtain the results shown in Figure 6; the black dots represent the peaks for the sample that covers 1948-2014 and the lines the different fittings.

To help the understanding of Figure 6, in Figure 7 are shown the empirical distributions of the peaks over threshold of the sub-samples considered (1948-1958, 1948-1968, 1948-1978, 1948-1988, 1948-1998 and 1948-2008) and its correspondent GPD fit.

Years	1948-1958	1948-1968	1948-1978	1948-1988	1948-1998	1948-2008	1948-2014
Threshold	8.45	8.5	8.75	8.85	8.79	8.72	8.67

Table 3. Threshold that provides a sample of peaks over threshold with $\lambda=1$ for the different time series.

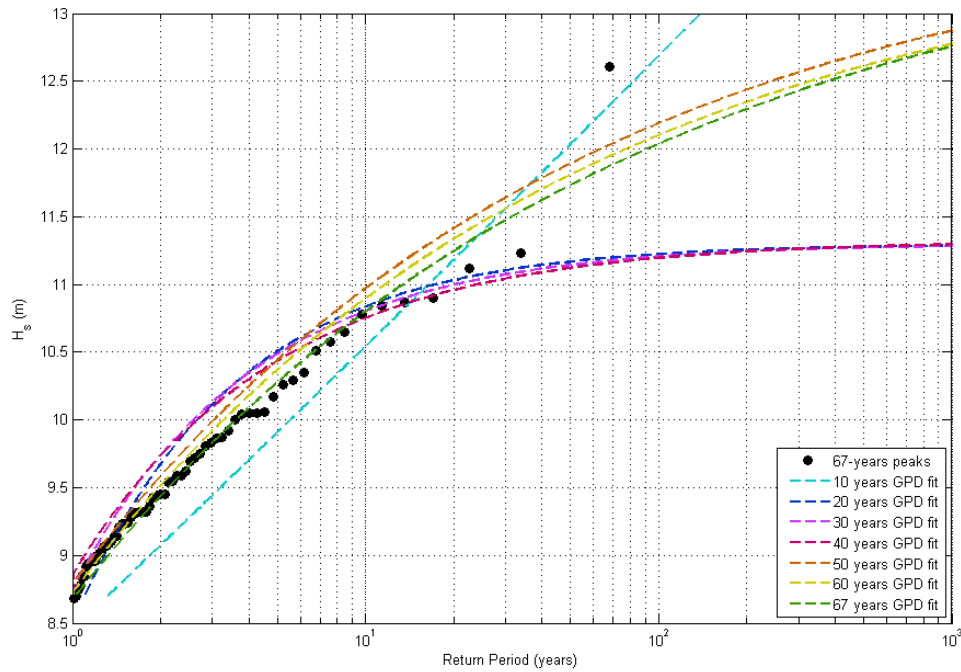


Figure 6 GPD fit for different sample sizes.

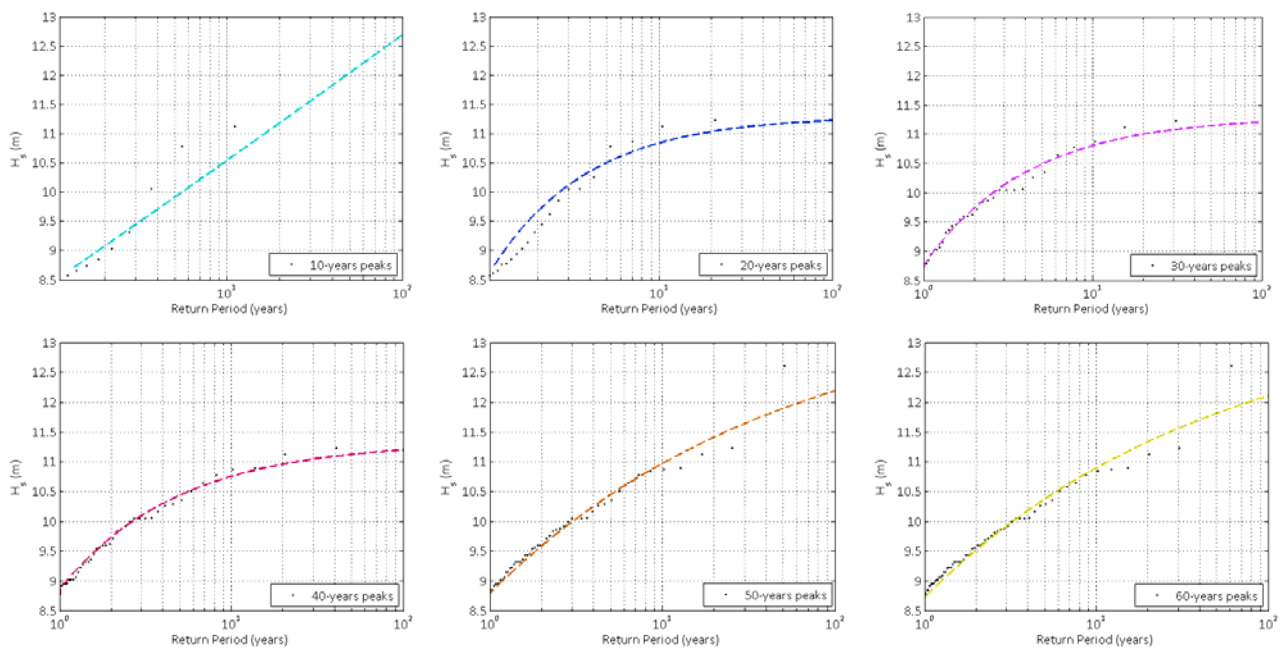


Figure 7. Empirical distributions of POT samples for the different considered periods of time with its GPD fit.

As seen, there is one event above 12.5 meters (exactly it is 12.61 meters and it was recorded on the 26th of February of 1989) which seems to control the discrepancy between the fittings. For the records that include this event (orange line for the data covering 1948-1998; yellow line, 1948-2008 and green line, 1948-2014), the expected return wave heights are much higher than for smaller records (up to 1 meter of difference, from 11 to 12 meters, for the 100-year return value of H_3). It can also be seen that the smallest sample, the one covering 1948-1958, presents a fitting clearly different from those with a higher number of events, indicating there are probably too few data points to fit a GPD

Plotting the fitted GPD parameters for all the cases we can see how they show a monotonous trend as the number of years (and thus, the sample size) increases above 10 years (Figure 8).

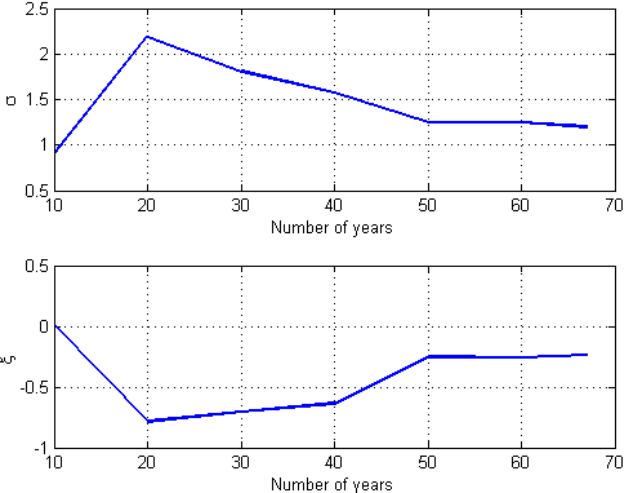


Figure 8. Evolution of scale and shape parameters of the GPD.
GEV fitting comparison with different time series lengths

The same analysis but considering a GEV was also made. In this case, the events considered were annual maxima. Figure 9 shows the GEV fittings for different sample sizes and in black dots the annual maxima for the sample that covers 1948-2014. To help the comprehension of Figure 9, in Figure 10 are shown the empirical distributions of the annual maxima of the sub-samples considered (1948-1958, 1948-1968, 1948-1978, 1948-1988, 1948-1998 and 1948-2008).

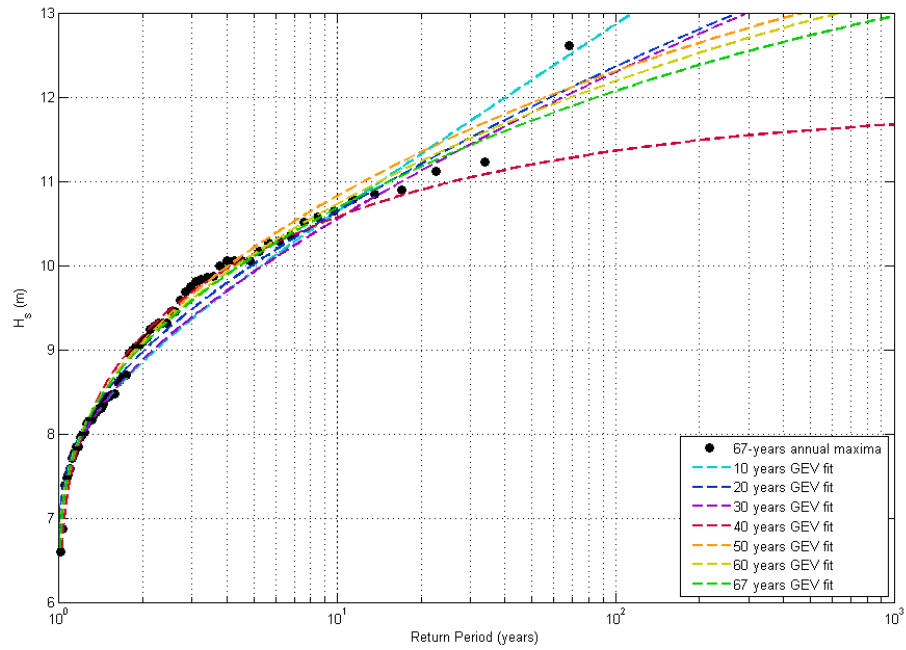


Figure 9. GEV fit for different sample sizes.

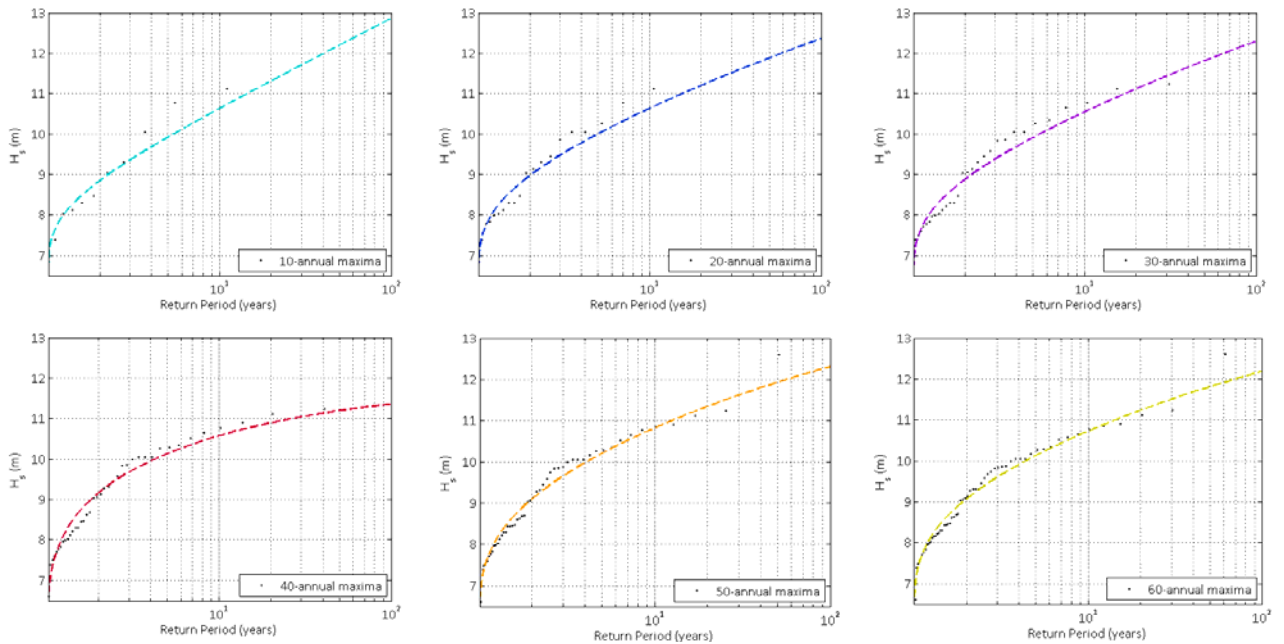


Figure 10. Annual maxima empirical distributions for the different lengths samples with its GEV fits.

The evolution of the estimated GEV parameters is shown in (Figure 11). As seen, the location and scale parameters increase as sample size increases. But the shape parameter presents a decreasing trend with the sample size. In the three parameters estimation appears a breaking point in the tendencies for the sample of 40 years; this fact can be related to the extreme event occurred in 1989, because the 40-years sample is the first one that records this event and more sensitive to these event than the longer ones (50, 60 and 67-years samples).

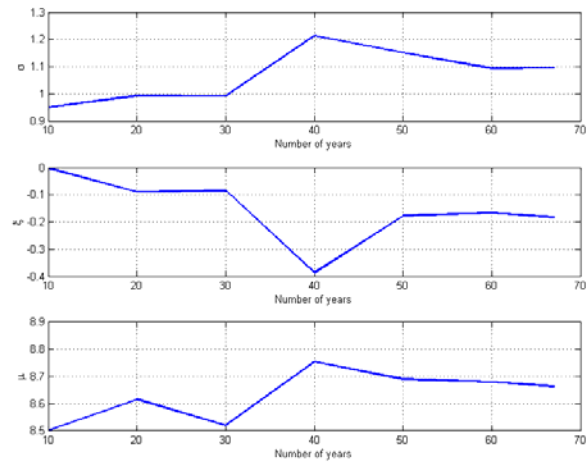


Figure 11. Evolution of scale, shape and location parameters of the GEV distribution.

Gumbel fitting comparison with different time series lengths

The Gumbel distribution is a particular case of the GEV distribution, with a shape parameter equals to zero. In Figure 12 the obtained results with this distribution are shown. This comparison was also made considering annual maxima; in Figure 13 the annual maxima of the different samples are depicted as well as their Gumbel fits.

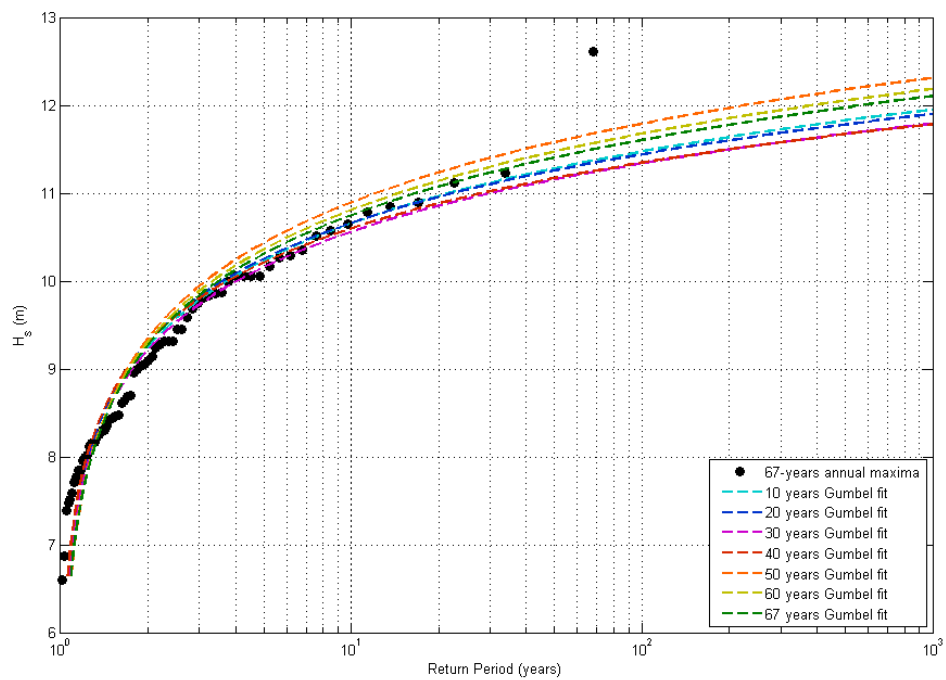


Figure 12. Gumbel fit for different sample sizes.

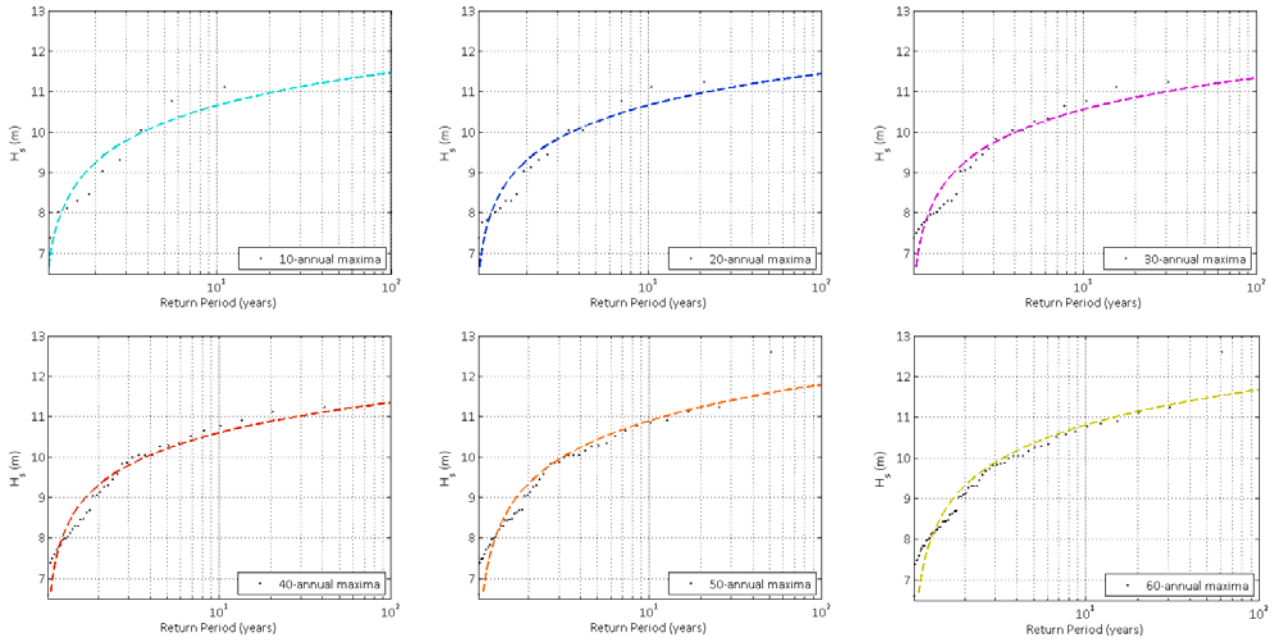


Figure 13 Annual maxima empirical distributions for the different lengths samples with its Gumbel fits.

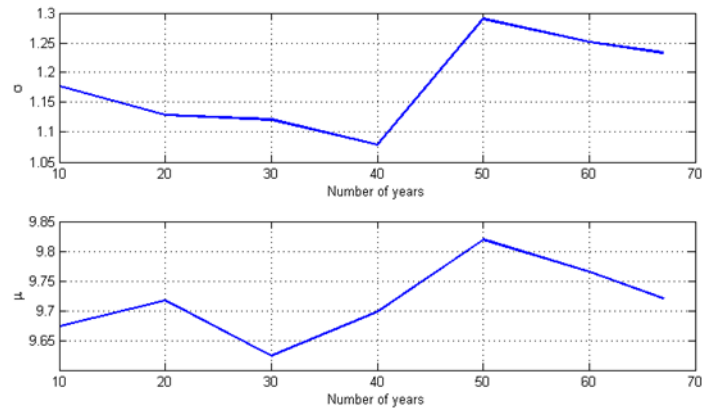


Figure 14. Evolution of scale and location parameters of the Gumbel distribution.

In Figure 14 is plotted the evolution of the Gumbel parameters as the sample size varies. The scale parameter presents a decreasing trend until the 40-years sample where it changes from a value lower than 1.1 to 1.3. The location parameter does not present a clear trend related to the sample size and its value does not vary that much: from 9.65 until 9.8.

Exponential fitting comparison with different time series lengths

Parallel to the relation existing within the Gumbel and the GEV distributions, there is a particular case of the GP distribution where the shape parameter is equal to zero: the exponential distribution. To make this comparison the peaks over threshold (with a suitable threshold that assures a mean number of 1 event per year) were used. The fitted distributions to different sample sizes are shown in Figure 15 and the peaks for the different sample sized are depicted in Figure 16 as well as their exponential fits. The evolution of the

parameters' values is shown in Figure 17. As seen, after some initial fluctuations, the value of the scale parameter tends to establish around 1 as the sample size increases.

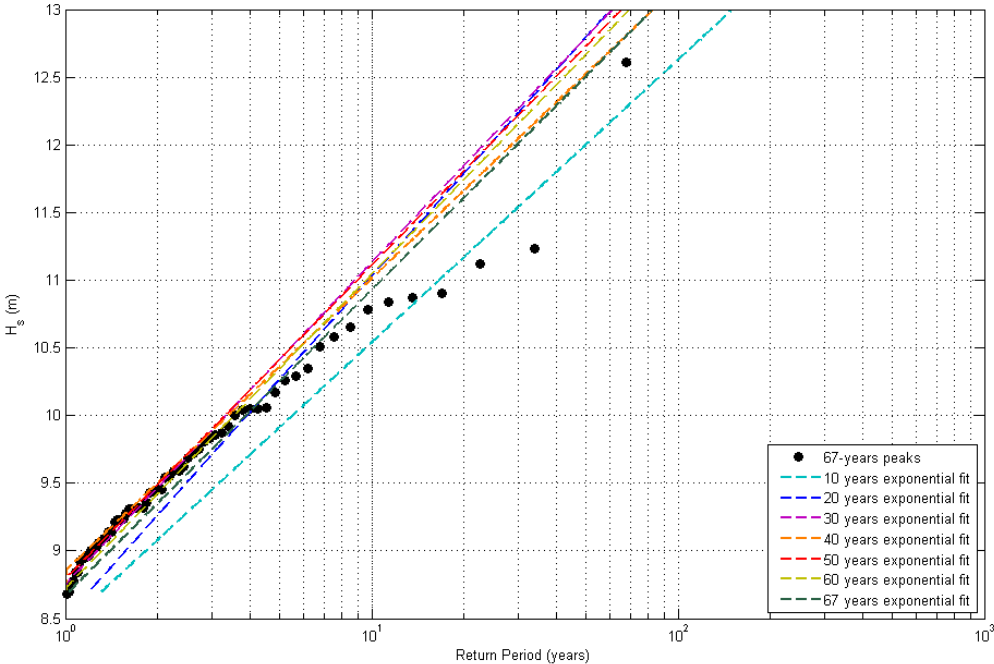


Figure 15. Exponential fit for different sample sizes.

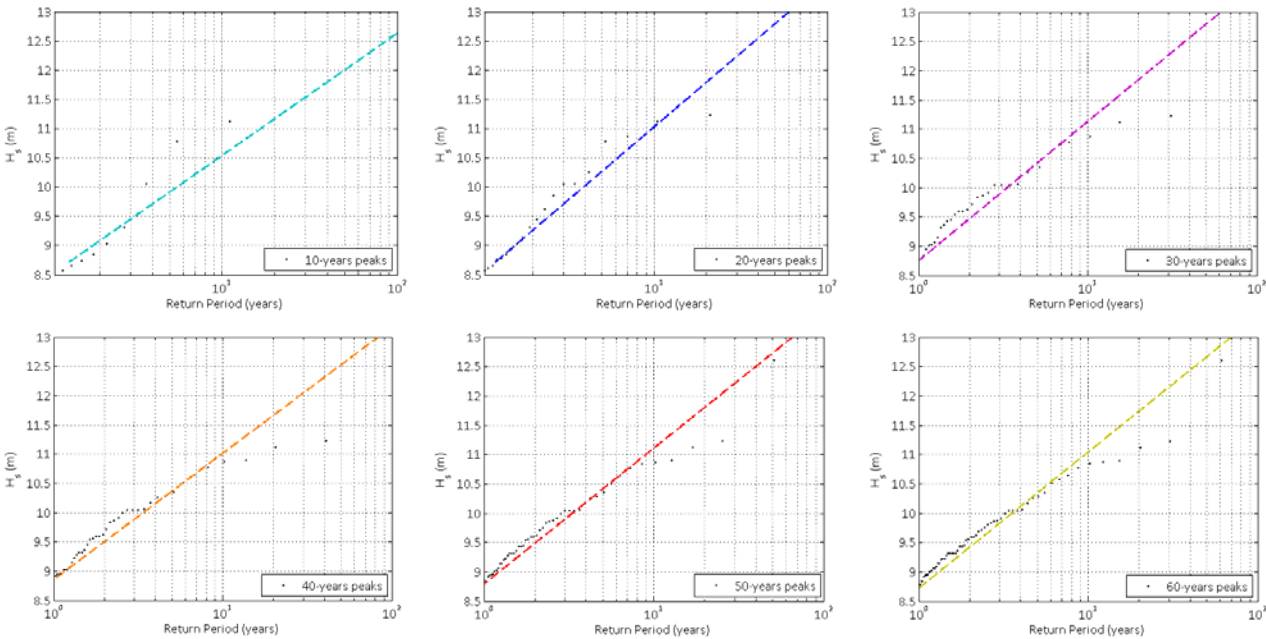


Figure 16. Empirical distributions of POT samples for the different considered periods of time with its GPD fit.

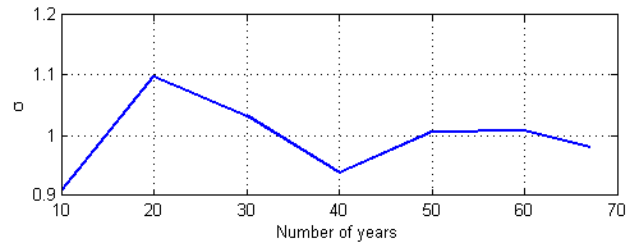


Figure 17 Evolution of the scale parameter of the Exponential distribution.

Comments regarding the sample size

As seen, the sample size is an important issue when dealing with extreme value analysis. When the available data are not long enough might be a practical solution to complement the existent data with data with similar extremal characteristics; for instance using data from more than one location or combining more than one source of data. For example, in (Minguez 2013) they combine the use of reanalysis and buoy data in order to enlarge the available data and thus could make a better estimation of the extreme behavior.

Another option could be to synthetically sample data that keep the same characteristics than the original ones. For example, in the case of sampling waves (C. a. Guedes Soares 1996) or (Longuett Higgins 1952) propose different methods to treat this issue.

From a different approach, instead of trying to enlarge the sample a practical option would be to constrain one or more parameters before fitting the model. The big drawback of this is that it requires expert knowledge and can be subjective.

Finally, regarding the decision between selecting a GEV or a GPD model; in terms of sample size, the POT (Peaks Over Threshold) typically allows taking into account more observations than block maxima (yearly maxima for instance) for the same period of time when λ (mean number of events per year) is larger than 1 (more observations of high values).

Complementing the sample with other data with similar extremal characteristics (for example using data from more than one location; or combining more than one source: (R. T. Minguez 2013) uses reanalysis and buoy data combined).

2.4.2. Sample Variability

Another aspect to be considered apart from the sample size is its variability. This concept it is intimately related to the previous one because as the sample size becomes smaller, its variability increases. To show this issue, the same 67 years' time series of wave height used in the previous section (2.4.1). From this time series, different samples of shorter lengths were compared (10, 20 and 30 years). Initially, with the six different 10-year samples available within the entire time series (1948-1958, 1958-1968, 1968-1978, 1978-1988, 1988-1998, 1998-2008) a GEV was fitted to each of them considering annual maxima. Results can be seen in Figure 18.

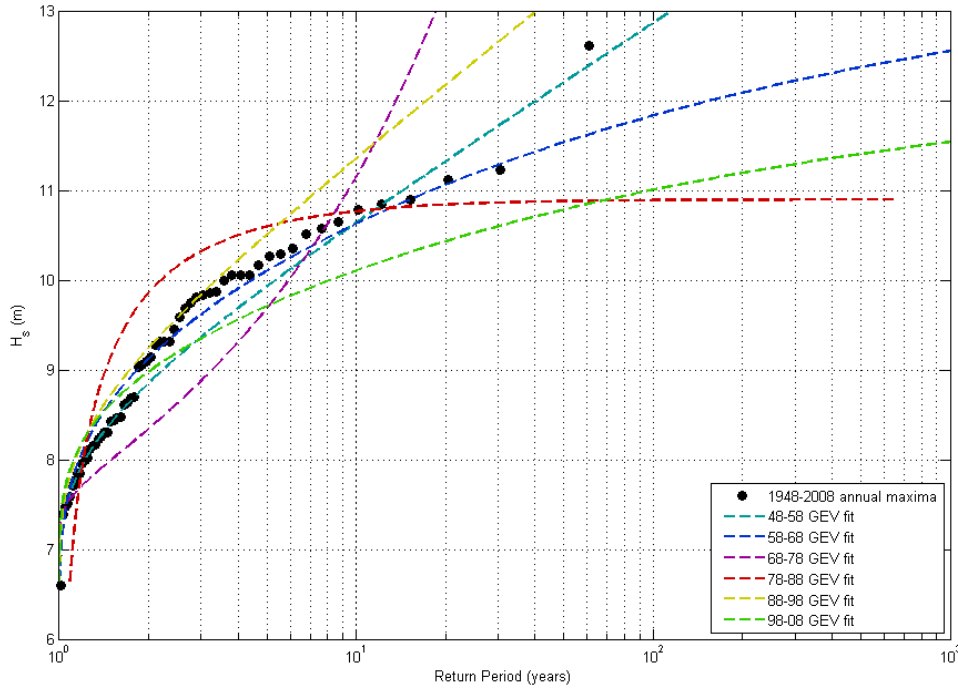


Figure 18. Sample variability considering 6 different samples of 10 years length.

As seen, the estimated fit to each 10-year time series presents a high variation. In the table below, the estimated parameters for each sample are summarized as well as the 100-year return value of H_s . The location parameter seems to keep stable around 8.5m except for two of the samples: 1968-1978 and 1978-1988 where in the first one it is only of 8m and in the second one goes up to 9.3m. This, together with the great variations in the shape parameter (with values that range from -1.09 in the 1978-1988 sample to 0.551 in the 1968-1978 sample) give explanation to the differences depicted in Figure 18: the red line presents a saturated extreme distribution, where there are no values above 11m while the purple line represents a heavy tailed distribution and therefore the 100-years return value for this fitting it is above 18m. Although the rest of the fittings seem not to be so different, the 100-years return value associated to them present variations up to 3m (from 11m in the sample of 1998-2008 to 14.1m in the one that covers 1988-1998)

Years	1948-1958	1958-1968	1968-1978	1978-1988	1988-1998	1998-2008
Location (μ)	8.503	8.753	8.060	9.341	8.834	8.697
Scale (σ)	0.950	1.047	0.695	1.701	1.104	0.790
Shape (ξ)	-0.001	-0.212	0.551	-1.091	0.012	-0.216
H_s R=100y (m)	12.87	11.83	>18	10.89	14.1	11

Table 4. GEV parameters for the different 10-years length samples.

Then, the same analysis was made but considering samples of 20 years length. Thus, up to 3 samples can be extracted from the 67-years' time series: 1948-1968, 1968-1988 and 1988-2008. The GEV fits to the annual maxima of these three samples are depicted in Figure 19, while the estimated parameters are summarized in Table 5. In this case, the variations presented by the three distributions fitted are much lower; however, the sample that covers 1968-1988 presents a different behavior than the other two having a high negative shape

parameter and thus a lower 100-years return value; because its fit seems to have a horizontal asymptote around 11m; which means that extreme events above this value are rarely expected.

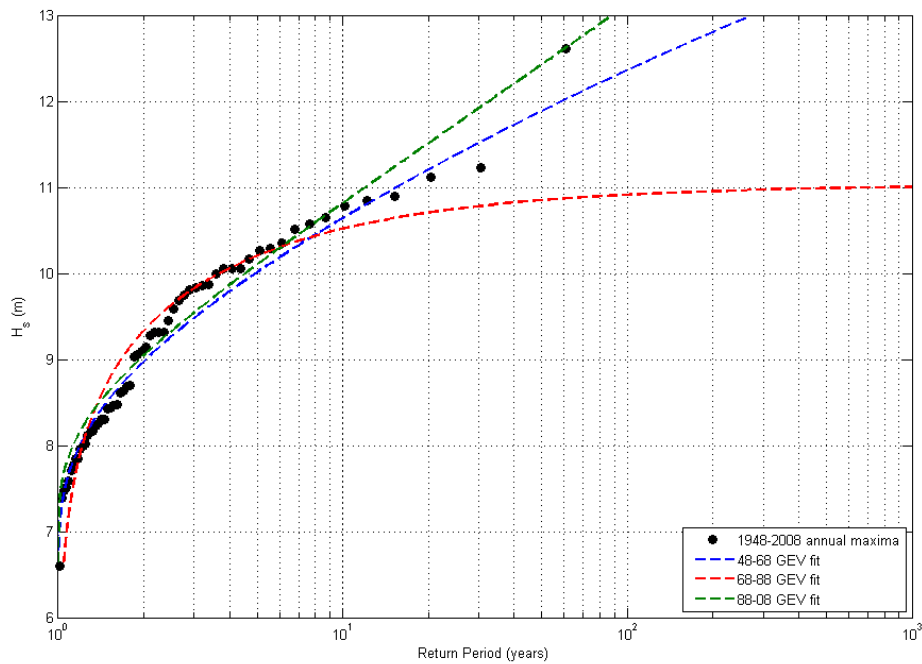


Figure 19. Sample variability considering 3 different samples of 20 years length.

Years	1948-1968	1968-1988	1988-2008
Location (μ)	8.6165	8.8848	8.7065
Scale (σ)	0.9920	1.3662	0.9148
Shape (ζ)	-0.0890	-0.6376	0.0207
H_s R=100y (m)	12.37	10.92	13.1

Table 5. GEV parameters for the different 20-years length samples

And finally, using 30-year time series, the same test produces the results shown in Figure 20 and summarized in Table 6. In this third case, as expected, the variations between the samples are much more reduced. Although the parameters estimated differ a bit from one sample to the other, the 100-years return value is practically identical and both distributions present a similar behavior not being heavy tailed nor horizontally asymptotic.

Years	1948-1978	1978-2008
Location (μ)	8.520	8.814
Scale (σ)	0.991	1.142
Shape (ζ)	-0.085	-0.187
H_s R=100y (m)	12.30	12.34

Table 6. GEV parameters for the different 30-years length samples.

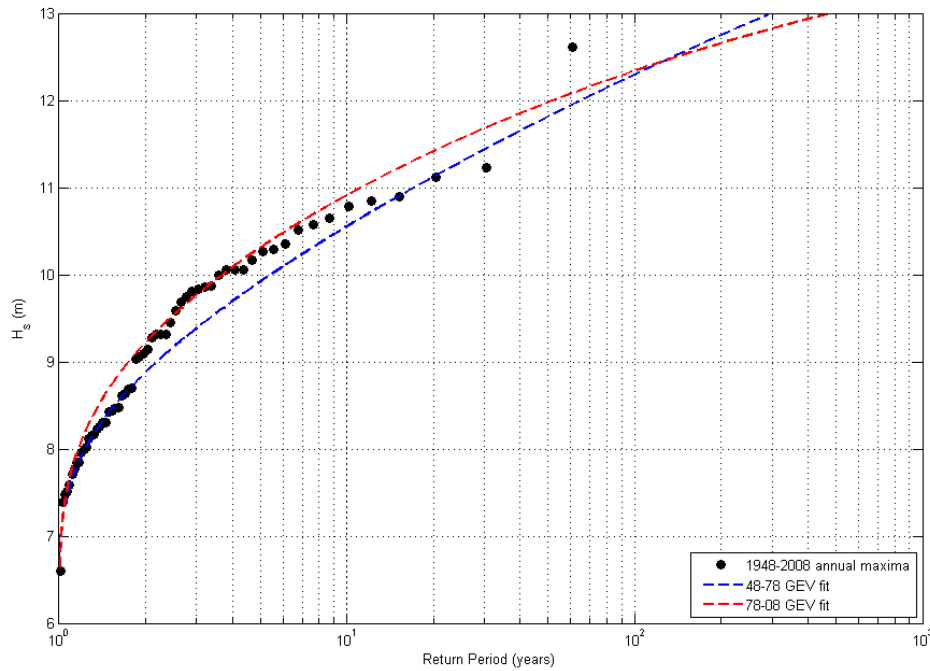


Figure 20. Sample variability considering 2 different samples of 30 years length. From these three figures (Figure 18 to Figure 20) it can be stated that as the sample size increases, its variability in terms of extreme value analysis decreases, this can be confirmed with the estimated parameters values expressed in the tables.

These last three graphs are also a good illustration of the minimum sample size requirement to perform an extreme value analysis: with 10 years only, the results obtained with the GEV are not meaningful as the results vary too much between samples, even for return periods smaller than 10 years. They also show the limitation of extrapolation time: extremes should not be extrapolated for too high return periods compared to the period of time covered by the sample.

2.4.3. Independence between events

The temporal independence between events in extreme value analysis has to be guaranteed in order to achieve accurate results. In the cases of block maxima, this independence is almost guaranteed depending on the size of the blocks, but in the case of peaks over threshold it has to be considered.

In case of developing a Peaks Over Threshold analysis, as stated in (Méndez 2006) *"the time span should be taken as the optimal compromise between the minimum time interval over which the Poisson process may be assumed to be valid [Luceño et al. 2006], the physical time interval to guarantee the independency between consecutive storms and the length of the time series"*. In their study, located in the Atlantic coast of the United States, they tested different values of time span up to 10 days and although the Poisson assumption was slightly better satisfied with 6 days than with 3, they chose a time span of 3 days to reach a good compromise between Poisson assumptions and longer sample size which rebound in smaller confidence intervals.

Numerous authors have developed studies to analyze the time span necessary to be considered in different parts around the globe. Among them, (Gulev 2001) provided a study of the extra-tropical cyclone variability in the Northern hemisphere winter. They stated that the modal life time of cyclones in the Northern hemisphere is 2-3 days, with 50% of all cyclones existing from 1 to 4 days. Considering the different basins, the eastern Pacific lifetime is shorter (2 to 4 days), while the 60% of cyclones in the Atlantic last from 2 to 6 days with a mean length of 4 days. The Arctic region is characterized by the shortest cyclone lifetimes with a mean of less than 3 days.

To illustrate the influence of this issue, the same time series of significant wave eight used in the precious sections (2.4.1 and 2.4.2) has been used. From this time series, a block maxima analysis was made, varying the block size: annual, monthly, weekly, daily and 6-hourly maxima. In **Erreur ! Source du renvoi introuvable.**, a graphical representation in a Gumbel Probability scale for the different block size maxima data compared to the Generalized Extreme Value distribution fitted with the annual maxima is represented.

The probability of occurrence of H_s varying the block size (or the independence time lapse in case of GPD) is related in the following way:

$$P(H_s^{life} \leq h) = P(H_s^{sea-state} \leq h)^{N^\circ sea-states/life}$$

The probability of H_s below a certain h within the entire sample is equal to the probability of H_s above the same h within the block size at an n power. Being n the number of these block sizes within the entire sample ($n=N^\circ sea-states/life$; being *life* the entire sample and $N^\circ sea-states$ the number of clusters in which it has been divided). For example, in case of comparing the probability of occurrence of annual, monthly and weekly maxima the relation between them would be:

$$P(H_s^{annualMAX} \leq h) = P(H_s^{monthlyMAX} \leq h)^{12} = P(H_s^{weeklyMAX} \leq h)^{52} = \dots$$

From the figure it can also be seen that as the block size used decreases, the maxima depart more and more from the GEV fitting. This is because in a small block size (i.e. less than one week) the independence between maxima events cannot be directly assumed. This kind of graphs allows the designer to select the appropriate block size previous to the model fitting. By selecting a smaller block size, the number of events will be higher and thus the variance of the model parameters will be reduced.

In case of needing more data, the block size might become smaller than that required to guarantee the independence between events; an extremal index could then be introduced. This extremal index takes into account the dependence between the events and allows using the data as independent events. In this case, the general formula above written would become:

$$P(H_s^{life} \leq h) = P(H_s^{sea-state} \leq h)^{N^\circ sea-states/life \cdot \alpha(h)}$$

With α as the extremal index function of h . The extremal index is interpreted as being the inverse of the average duration of the extreme events or in other words, the parameter that enables us to measure the degree of clustering of extremes. However, its estimation is not a trivial issue. Some approaches could be found in (R. a. Smith 1994), (Hamidieh 2010) or (Leadbetter 1983).

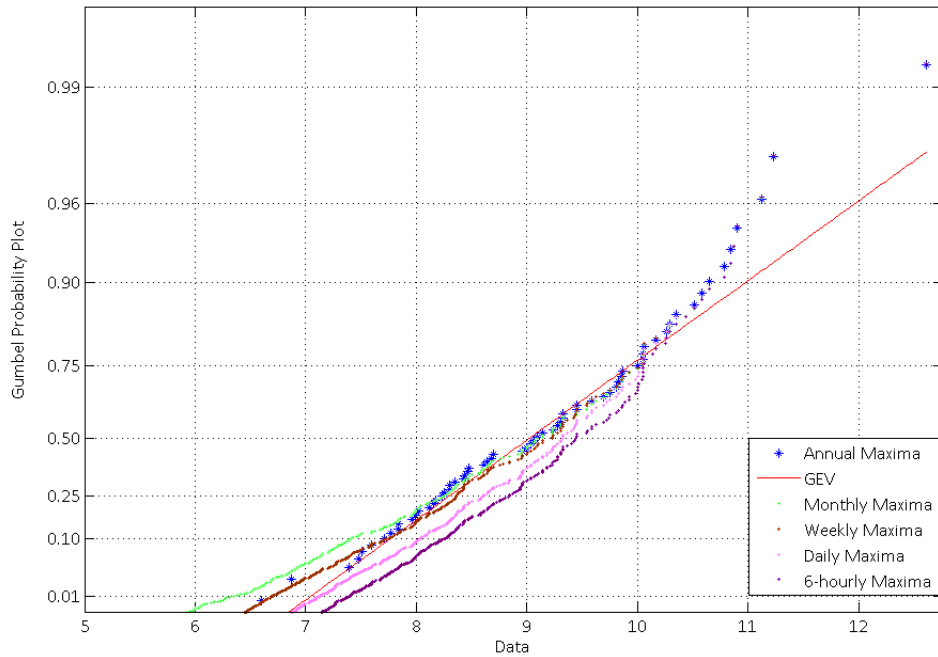


Figure 21. Gumbel Probability Plot for different block sizes in a GEV analysis.

The same analysis has been undertaken considering a Peak Over Threshold sample fitted to a Generalized Pareto distribution with the same initial wave height time series at the same location. The time span between events in the POT selection varied from 12 hours to 7 days (12 hours, 1 day, 2 days, 3 days, 5 days and 1week), while the threshold was varying in such a way that guaranteed a fixed number of 1 event per year ($\lambda=1$). The results have been plotted in an exponential probability plot in Figure 22. As seen, while for the highest values the independence is maintained regardless the time span, for smaller values and probabilities this cannot be guaranteed.

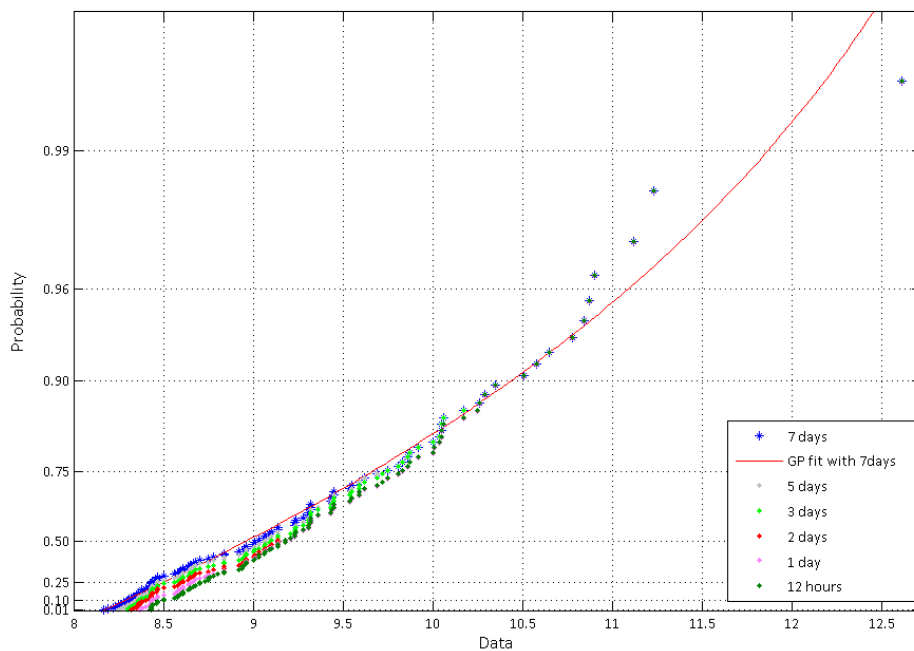


Figure 22. Exponential Probability Plot for different independence lapses in a POT analysis.

2.4.4. Covariates influence

Previous to the performance of an Extreme Value Analysis the possible existence of covariates which make that our data vary systematically should be checked. In the particular case of wave height, the most important covariates are seasonality and direction. In Figure 23 is depicted the seasonality present in the significant wave height record at the location 2.5W 45N. As seen, the magnitude (for example, waves above 8 meters) of the wave height is higher in the winter months (from October to April), while during the summer months only few values are higher than 6 meters height. This is clearly a presence of seasonality in the data record.

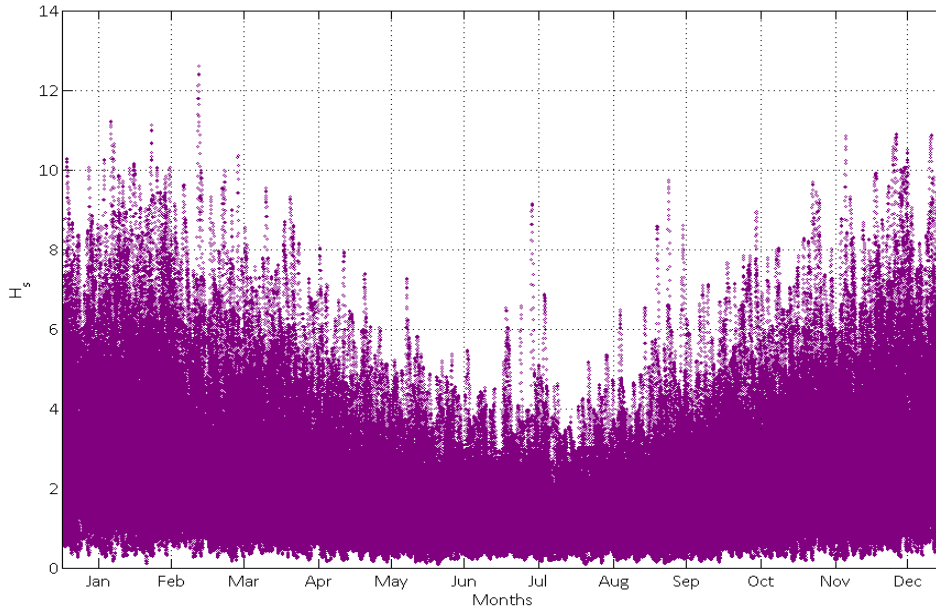


Figure 23. Seasonality.

This temporal behavior should be considered in an extreme value analysis (Menéndez 2009). In order to test if this behavior will influence the results some comparisons can be made. With the data depicted in Figure 23, the monthly maxima were extracted and divided between those belonging to winter months (October, November, December, January, February and March) and those to summer months (April, May, June, July, August and September). With those two databases, a GEV was fitted. As seen in Figure 24, there is a notable difference between both extreme models for summer (purple data) and winter (blue data). In black are represented the annual maxima and a GEV fitted to them. The resulting fit of annual maxima is quite similar to the winter maxima curve for return period values above 2 years; this is an indicator that in this particular case seasonality is not so important in terms of extremes. This is because all the annual maxima occurred in winter.

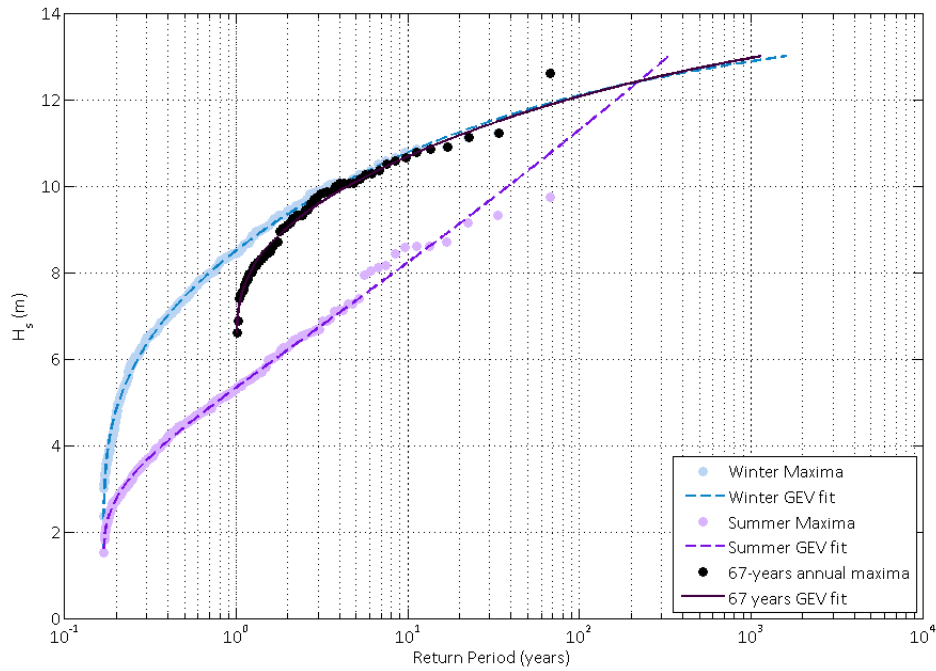


Figure 24. Comparison between winter (Oct-Mar) and summer (Apr-Sep) monthly maxima with Annual Maxima.

The same analysis was done considering the seasonality by splitting the monthly maxima into 4 seasons: winter (December, January and February), spring (March, April and May), summer (June, July and August) and autumn (September, October and November). The comparison of the GEV fits to all of the seasons and the annual maxima are presented in Figure 25. Similarly to the previous analysis, the annual maxima fit is quite similar to the winter maxima for return periods greater than 2 years.

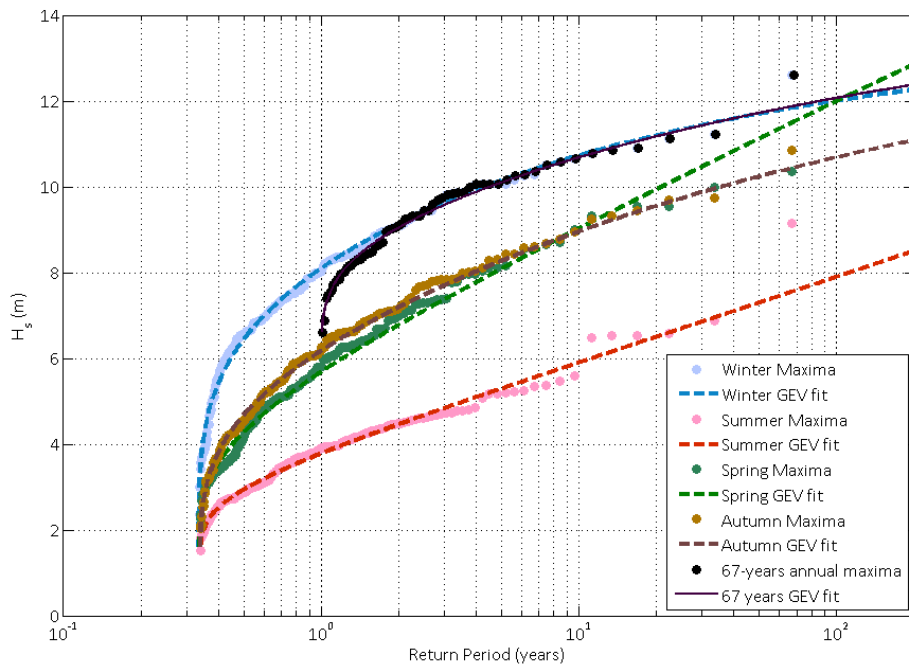


Figure 25 Seasonal [spring (MAM), summer (JJA), autumn (SON) and winter (DJF)] monthly maxima comparison.

But this cannot be extended up to a monthly analysis of the data, because not all the annual maxima occur in the same month. Because of that, the annual maxima distribution and fit shown in Figure 26 present higher values than the monthly data.

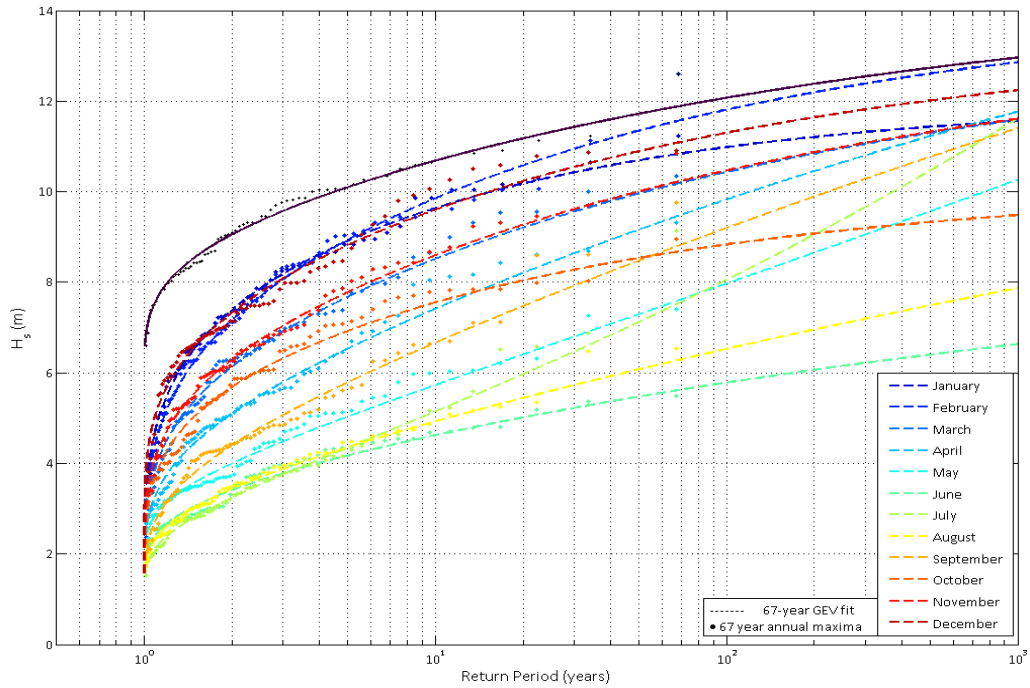


Figure 26 Monthly maxima comparison.

Dealing with covariate influence in terms of parameters estimation is not a trivial problem. The estimation of the parameters becomes more difficult and sometimes unfeasible when they are function of a certain covariate(s). In the case of introducing seasonality to study the generalized extreme value distribution parameters, the IH-Ameva software was used. This software, as a package for Matlab has already implemented the functions to fit the GEV parameters under different assumptions. In order to use it, monthly maxima were used (because, as seen before, seasonality has no influences on yearly maxima or seasonal maxima). The expressions fitted by the model are summarized below.

GEV Model:

$$F(x; \mu, \psi, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\psi} \right) \right]^{-1/\xi} \right\}$$

$$\mu(t) = \beta_0 + \sum_{i=1} [\beta_{2i-1} \cos(i\omega t) + \beta_{2i} \sin(i\omega t)] + \beta_{T1} t + \sum_k \phi_k n_{k,t}$$

$$\psi(t) = \exp \left(\alpha_0 + \sum_{i=1} [\alpha_{2i-1} \cos(i\omega t) + \alpha_{2i} \sin(i\omega t)] + \beta_{T2} t + \sum_k \phi_k n_{k,t} \right)$$

$$\xi(t) = \gamma_0 + \sum_{i=1} [\gamma_{2i-1} \cos(i\omega t) + \gamma_{2i} \sin(i\omega t)]$$

Model to Fit:

$$\mu(t) = \beta_0 + \sum_{i=1} [\beta_{2i-1} \cos(i\omega t) + \beta_{2i} \sin(i\omega t)]$$

$$\psi(t) = \exp \left(\alpha_0 + \sum_{i=1} [\alpha_{2i-1} \cos(i\omega t) + \alpha_{2i} \sin(i\omega t)] \right)$$

$$\xi(t) = \gamma_0 + \sum_{i=1} [\gamma_{2i-1} \cos(i\omega t) + \gamma_{2i} \sin(i\omega t)]$$

The results obtained by the model and the appearance of the seasonal behavior of the fitted parameters are shown below.

$$\beta_o = 4.748, \beta_1 = 1.871, \beta_2 = 0.370$$

$$\alpha_o = 0.262, \alpha_1 = 0.421, \alpha_2 = 0.089, \alpha_3 = -0.110, \alpha_4 = 0.042$$

$$\gamma_o = -0.12839, \gamma_1 = -0.163, \gamma_2 = -0.032$$

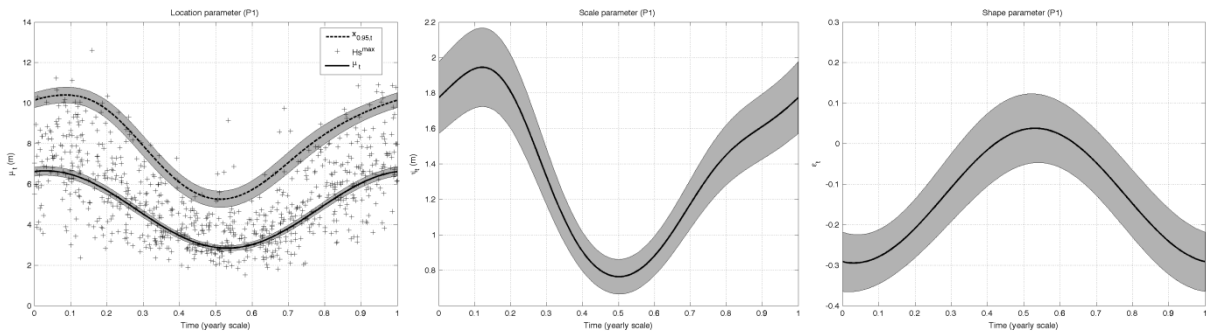


Figure 27. Seasonality of the GEV parameters.

As another example, Figure 28 shows the distribution of wave heights according to its direction at the same location and for the same record. It can be seen that the majority of waves comes from the 4° quadrant, from Westerly directions. In addition, the highest waves only arrive from directions between 270°-300°, thus, the direction is a covariate that should be taken into account in any study of extremes made in that area.

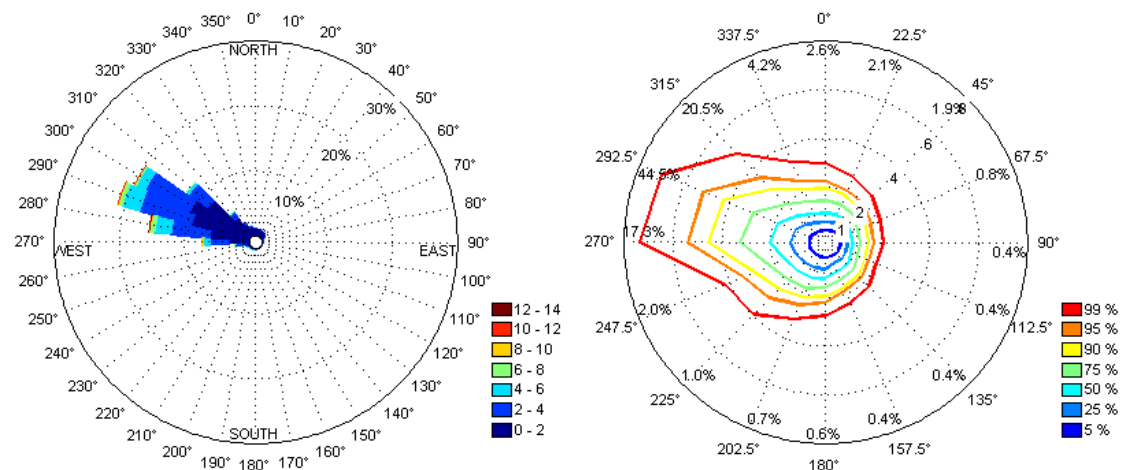


Figure 28. Rose directions (standard-left rose and quantiles-right rose).

2.4.5. Plotting position

Within the scientific community there is still no agreement related to the plotting position of the empirical distribution of data. This problem influences in the graphical

representation and it can lead into different visual appreciations of the goodness of fit of a model. A general formula to estimate the probability of empirical data is given by:

$$p_i = \frac{i-a}{N+1+b}$$

Where a and b are constants, i is the position of each observation within the sample once it has been increasingly sorted and N is the sample size.

A summary of the most common plotting position formulas is given in the table below.

<i>Author</i>	<i>a</i>	<i>b</i>	<i>p_i</i>	<i>Year</i>
Weibull	0	0	$\frac{i}{N+1}$	1939
Beard	0.31	-0.62	$\frac{i-0.31}{N+0.38}$	1943
Gringorten	0.44	-0.88	$\frac{i-0.44}{N+0.12}$	1963
Hazen	0.5	-1	$\frac{i-0.5}{N}$	1914
Cunnane	0.4	-0.8	$\frac{i-0.4}{N+0.2}$	1978

Table 7. Plotting position formulas.

In order to give an example of the difference between these different methods a graphical comparison has been made. From the same wave height time series as used before, the annual maxima has been extracted and to this sample the different formulae shown in Table 7 has been applied. Then, the results in terms of return period, $R=1/(1-p)$, are plotted in Figure 29. In this figure it has also been depicted a GEV distribution fitted to the sample of annual maxima. As seen, considering one plotting method or another would give a different visual impression of the goodness of fit of the distribution. These differences are only appreciable in the upper values, giving notably different empirical return periods for the highest events. For example, the maximum event of the record (12.61 meters of significant wave height) presents an empirical return period varying from 68 to 134 years depending on the plotting position method (Table 8).

	Weibull	Beard	Cunnane	Gringorten	Hazen
R in years for the 12.61m event	68	98	112	120	134

Table 8. Return period (years) for the 12.61m event according the different plotting position formulae.

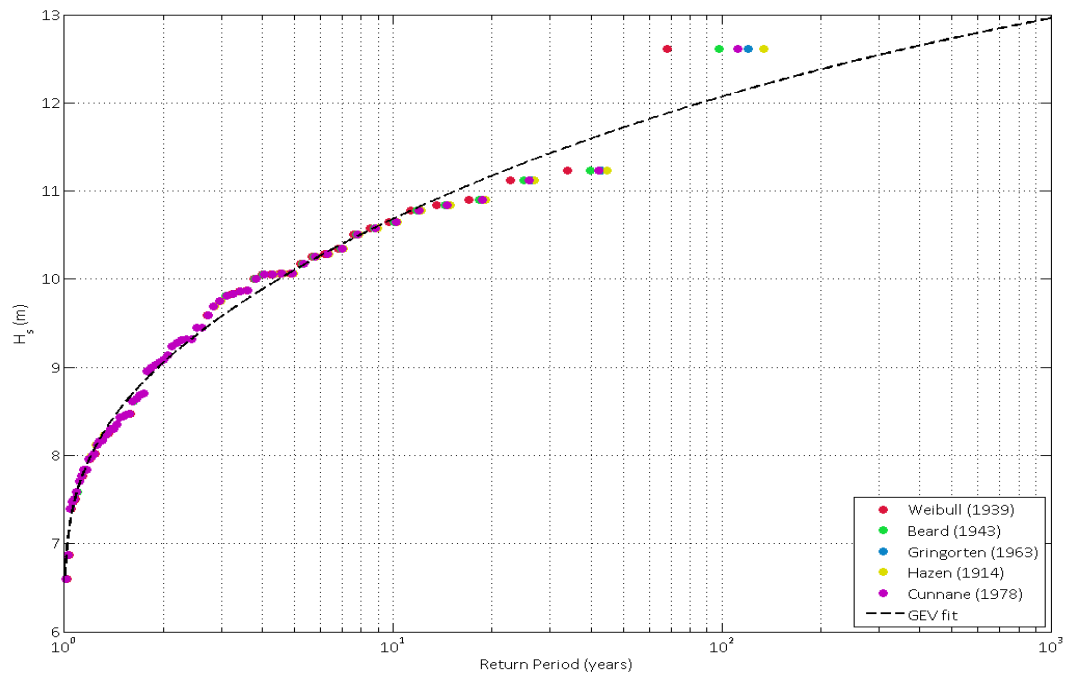


Figure 29. Plotting position comparison.

Note that in all of the graphical representations included within this report the Weibull plotting method has been used.

2.4.6. Measurement scale influence

The results obtained with an extreme value analysis generally are not invariant to the measurement scale. It may happen that a simple pre-treatment becomes really helpful or even indispensable to achieve a good EV model fit.

For example, let assume that the parameter of relevance in our study is H_s^2 with a return period of 100 years. To estimate its value two approaches can be followed:

1. Fit H_s , estimate $(H_s)_{100y}$ and then make the square of it: $((H_s)_{100y})^2$.
2. Make the square of the variable, H_s^2 and then fit this new variable and obtain $(H_s^2)_{100y}$.

With the same significant wave height time series as in the previous tests, this comparison has been made (considering 4 different fittings: GEV, Gumbel, GPD and Exponential).

	GEV	Gumbel	GPD	Exponential
$((H_s)_{100y})^2$	145.44	134.56	144.96	173.98
$(H_s^2)_{100y}$	146.5	134.5	144.1	160.2

Table 9. Comparison of the measurement scale influence with the variable H_s .

The same analysis was also made to the variable wind speed (W), the obtained results are shown in the table below.

	GEV	Gumbel	GPD	Exponential
$((W)_{100y})^2$	711.8	641.1	666.1	791.3
$(W^2)_{100y}$	720.5	638.4	669.1	752.7

Table 10. Comparison of the measurement scale influence with the variable W

As seen in the tables before, there are discrepancies within the two ways of estimating the value. It can be also seen that these discrepancies vary between the variable analyzed and also within the distribution fitted to estimate the 100-years return period. In the case of H_s , the discrepancies are lower than for the variable W , and regarding to the distribution fitted: the exponential one presents the highest discrepancies while the Gumbel and the GEV the GPD the lowest ones. Thus, for those cases where our study is clearly influenced not by the variable itself but by a transformation of it (squared, cubic, normalized ...) it may be advisable to test different variations applied to the initial data.

2.4.7. Location influence

Several studies have already shown the differences on marine climate regarding the location (Gulev 2001) (C. M. Izaguirre 2011), (C. M. Izaguirre 2012). Independently to the database or the method used there is a spatial variability along the oceans that is relevant, thus, higher waves are expected in the northern hemisphere in the North Atlantic or the Bering Sea while smaller waves are more present in the Atlantic and Pacific oceans at equatorial latitudes.

In Figure 30, extracted from (C. M. Izaguirre 2011), it is represented the seasonal 20-year return period significant wave height in a global scale. In this study the authors used satellite data to analyze the spatial variability of extreme waves. As seen, there are clearly differentiated behaviors according to the location.

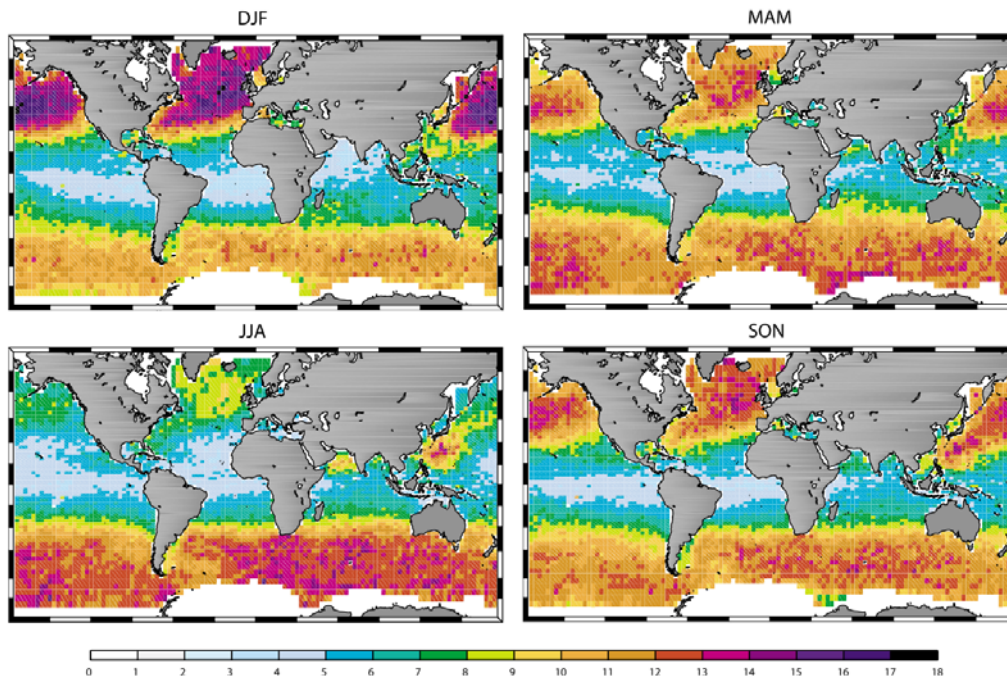


Figure 30. Seasonal distribution of the 20y return period H_s (C. M. Izaguirre 2011).

In this study they have also performed an adjustment of the annual maxima to a GEV distribution at each location. Figure 31 shows the variability of the shape parameter around the world. They state that '*Tropical areas where high cyclone and typhoon activity occur present a Fréchet like behavior: the Caribbean, the Gulf of Mexico, the western tropical Pacific (Philippine Sea) and French Polynesia. Areas of swell generation, such as the North Atlantic Ocean or the Austral Ocean (Roaring Forties) present Weibull distribution. This aspect of the behavior of the shape parameter is especially relevant when it comes to the estimation of high return-period quantiles, such as the 100 or 50 year return level values.*'

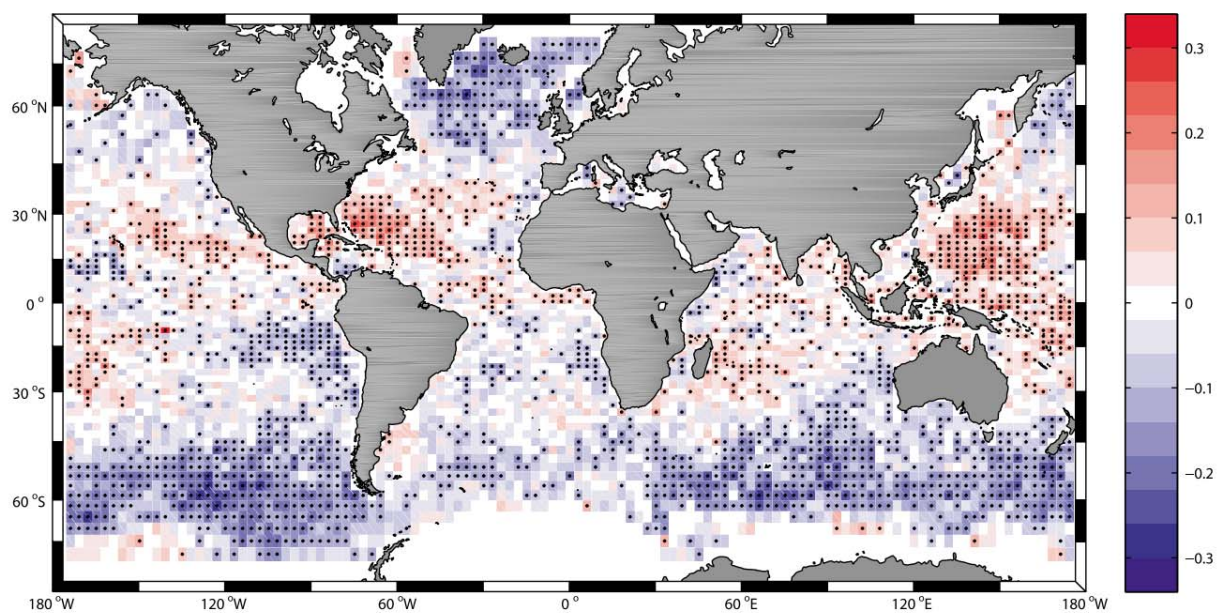


Figure 31. Spatial variability of the shape parameter of extreme wave height distribution around the world. Only statistically significant values at the 90% confidence interval are dotted (from Izaguirre et al 2011).

In Figure 32, the same analysis as that proposed by (C. M. Izaguirre 2011) is shown, but considering the NOAA database. As seen, there are areas where the shape parameter is clearly positive (around the tropical areas of the Caribbean Sea or in front of the coasts of Australia); these values of the shape parameter are related to Fréchet distributions, and it means heavy tailed extremes distribution. These last figures clearly represent the spatial variability in terms of marine climate extremes, and thus, the importance of the location where our study takes place.

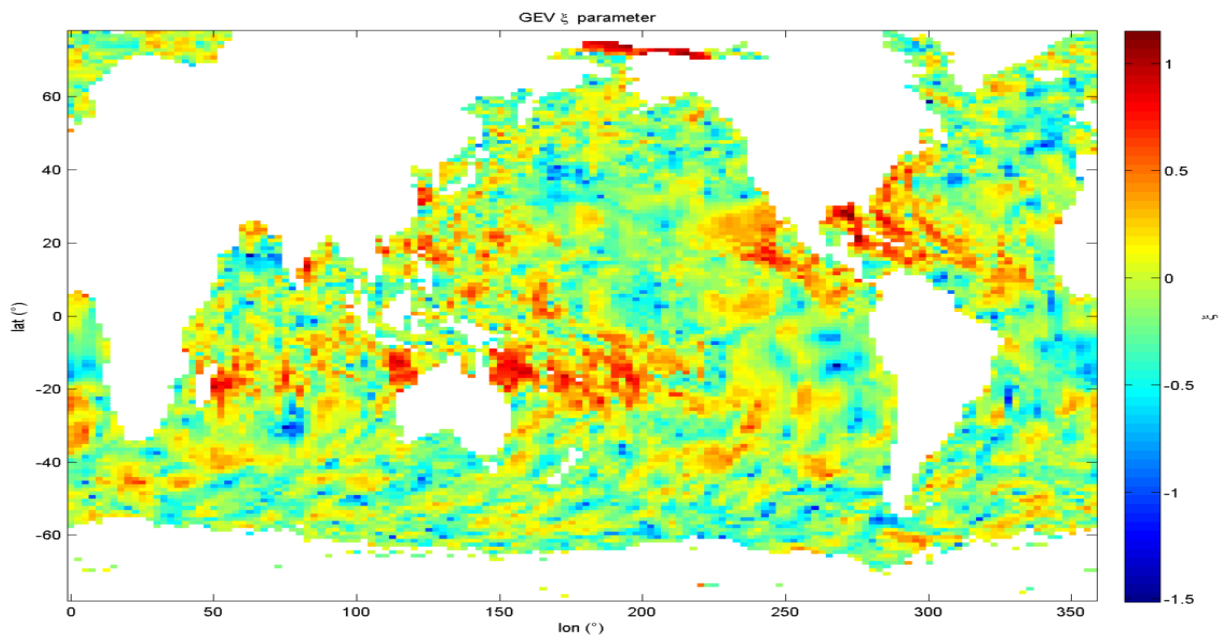


Figure 32. Spatial variability of the shape parameter of extreme wave height distribution using the NOAA database.

2.4.8. Database influence

Within the different met-ocean databases available there are some discrepancies. Different sources of data (buoy measurements, results from reanalysis models, satellite observations ...) or different initial conditions and hypothesis within the models will lead into differences between different sources of information. These differences may conduct to differences in the extreme value analysis and on the expected return values. In order to quantify these discrepancies, various databases at the same location have been compared.

Databases

The data bases used for this were the followings:

- NOAA

The NOAA (National Oceanic and Atmospheric Administration) from the EEUU develop databases for different ocean and atmospheric variables, as well as forecast and real-time observations of several meteo-ocean phenomena (<http://www.nodc.noaa.gov/>). Their wave database is an open source database available for any utilization but in the models they have used to develop their data they have not considered currents or tides.

They provide different spatial grids of resolution ($1^\circ \times 1.25^\circ$, $0.5^\circ \times 0.5^\circ$, $0.5^\circ \times 2.5^\circ$) and different temporal coverage but all of them with a 3-hourly temporal resolution.

The time series from NOAA used in this study covers 17 years, from 1/2/1997 until 31/10/2013, with data 3-hourly and with data from the parameters: H_s (m), T_p (s), Dir ($^\circ$), Wind Speed (m/s) and Wind Direction ($^\circ$).

- ERA40

The ERA40 (http://data-portal.ecmwf.int/data/d/era40_daily/) is a reanalysis product of the ECMWF (European Center for Medium-Range Weather Forecasts) of atmospheric and surface conditions. It is also an open source database but it is only available for research purposes. In their models they did not take into account currents or tides.

With a 6-hourly temporal resolution and a spatial resolution of $2.5^\circ \times 2.5^\circ$, they provide data of H_s (m), T_m (s), Dir ($^\circ$), Wind Speed (m/s) and Wind Dir ($^\circ$) for a period of 46 years, covering from 1/9/1957 until 31/8/2002.

- HOMERE

The HOMERE database is a reanalysis database developed by IFREMER, which is available for all purposes and that, takes into account currents and tides in their models (http://www.previmer.org/en/produits/hindcast_sea_states_homere).

They provide an hourly time series that spans 19 years covering from 1/1/1994 until 31/12/2012. It has been developed with an unstructured mesh of 100.000 nodes that cover the French Coast. The parameters provided are: H_s (m), T_p (s), Dir ($^\circ$) Wind Speed (m/s), Wind Dir ($^\circ$), Currents, sea level and swells and seas.

- BoBWA

The BoBWA (Bay of Biscay Wave Atlas) is a reanalysis database developed by BRGM available for all purposes but that does not take into account in their models currents or tides (<http://bobwa.brgm.fr/>).

The grid used has a spatial resolution of $0.5^\circ \times 0.5^\circ$ in the North Atlantic Coast and $0.1^\circ \times 0.1^\circ$ around the French Coast. This grid has a temporal resolution of 6 hours, but in the nodes of the grid they have hourly temporal resolution. It has a temporal coverage of 45 years from 1/1/1958 to 31/08/2002 and with data from the parameters: H_s (m), T_m (s), T_p (s) and Dir ($^\circ$).

- IH-DATA

IH-DATA is a reanalysis meteo-ocean database developed by IH-Cantabria, freely available only for research purposes (<http://www.ihcantabria.com/es/software/item/719-ihdata>). They have produced different databases of wind fields, wave's data and sea level. The wave database has been developed for different grids (with variable spatial resolution: $1.5^\circ \times 1^\circ$ -

Global; 0.25° x 0.25°- Europe; 0.125° x 0.125°-Mediterranean Sea) covering different spatial areas. The model used to generate the wave database does not take into account the currents or tides.

Within the products they provide, in this study it has been used a time series of 67 years of hourly data covering the period from 1/1/1948 to 28/2/2014 and with data for the parameters: H_s (m), T_p (s), T_m (s), Dir (°).

Databases Influence

To show the existent differences between databases, time series of significant wave height of the different sources mentioned above (NOAA, ERA40, IH, HOMERE and BoBWA) at the same location (2.5W 45N) have been compared. Their characteristics are resumed in the table below.

<i>Database</i>	<i>Initial date</i>	<i>Final date</i>	<i>N° Years</i>	<i>Maxima</i>
IH-GOW	1/1/1948	28/2/2014	67	12.61
ERA40	1/9/1957	31/8/2002	46	10.42
NOAA	1/2/1997	31/10/2013	17	10.92
HOMERE	1/1/1994	31/12/2012	19	12.36
BoBWA	1/10/1958	31/08/2002	44	13.51

Table 11. Summary of the characteristics of the different databases used.

Figure 33 shows the time series of the five different databases used along a temporal axis. As seen, the IH GOW database is the one covering the largest time span (dots in grey), while the shortest one is the NOAA (red dots).

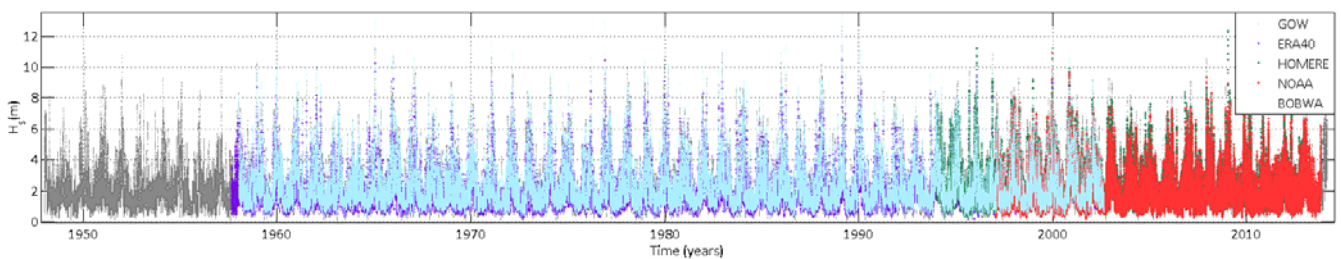


Figure 33. Databases time series (IHData-Grey, ERA40-Purple, HOMERE-green, NOAA-red and BoBWA-blue).

With these data some distributions were fitted. Extracting the annual maxima of each database a GEV distribution and a Gumbel distribution (particular case of the GEV distribution with the shape parameter fixed to zero) were fitted. On the other hand, a Peaks Over Threshold approach was performed to all of the databases guaranteeing a mean number of events per year equal to 1 ($\lambda=1$). With these POT samples, a GPD distribution and an Exponential distribution (being the Exponential distribution the same as fixing the shape parameter to zero within a GPD) were fitted.

GEV fitting comparison

In Figure 34 the GEV comparison carried out is shown. The dots represent the annual maxima for each database and the solid line the GEV distribution fitted to these maxima. It can be seen that there are some discrepancies between the different databases. While fittings from GOW, ERA40 and BoBWA present a Weibull behavior with shape parameters lower than 0, the results from NOAA and HOMERE present a heavy tailed distribution, with Fréchet behavior and positive shape parameters. This could be related to the fact that those two databases (NOAA and HOMERE) are the ones with shortest time coverage (17 and 19 years respectively) and an annual block size to the GEV might not be enough.

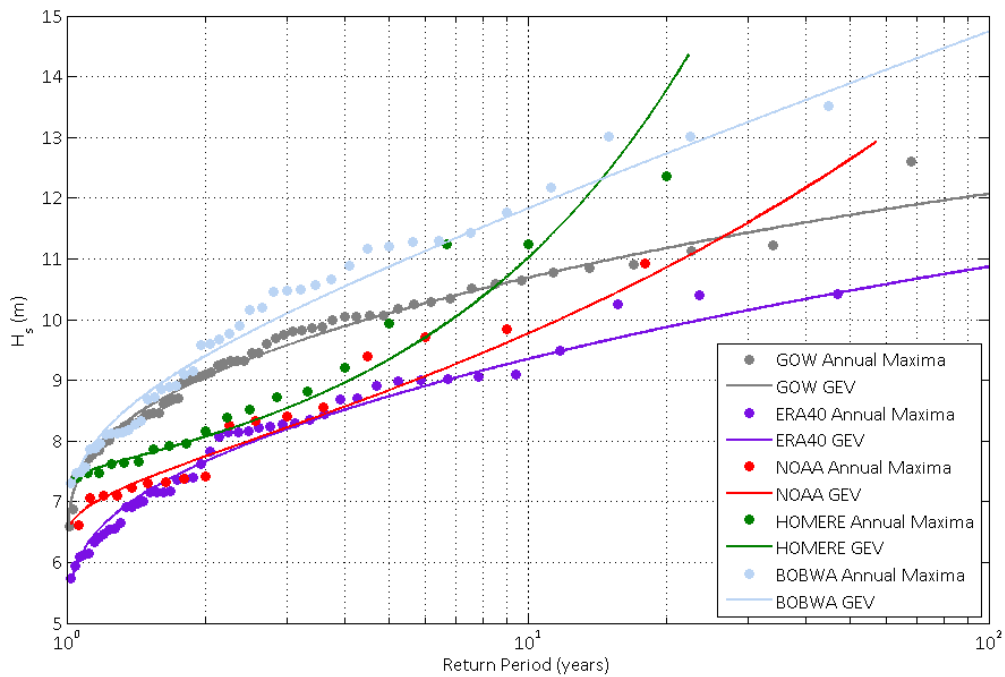


Figure 34. GEV distribution at 2.5W 45N for the different databases.

Database	Location (μ)	Scale (σ)	Shape (ξ)	H_s R=100y (m)
IH-GOW	8.663	1.095	-0.183	12.08
ERA40	7.276	1.087	-0.153	10.87
NOAA	7.463	0.741	0.274	14.1
HOMERE	7.842	0.544	0.744	>20
BoBWA	8.906	1.298	-0.010	14.75

Table 12. GEV distribution parameters estimated for the different databases.

Gumbel fitting comparison

Figure 35 represents the Gumbel fitting comparison. The Gumbel distribution is the particular case of a Generalized Extreme Value distribution with the shape parameter equal to 0. By fixing this parameter the behavior obtained does not vary that much for all the databases but still there are discrepancies of almost 3 meters within the different databases in terms of the 100-year return value estimates.

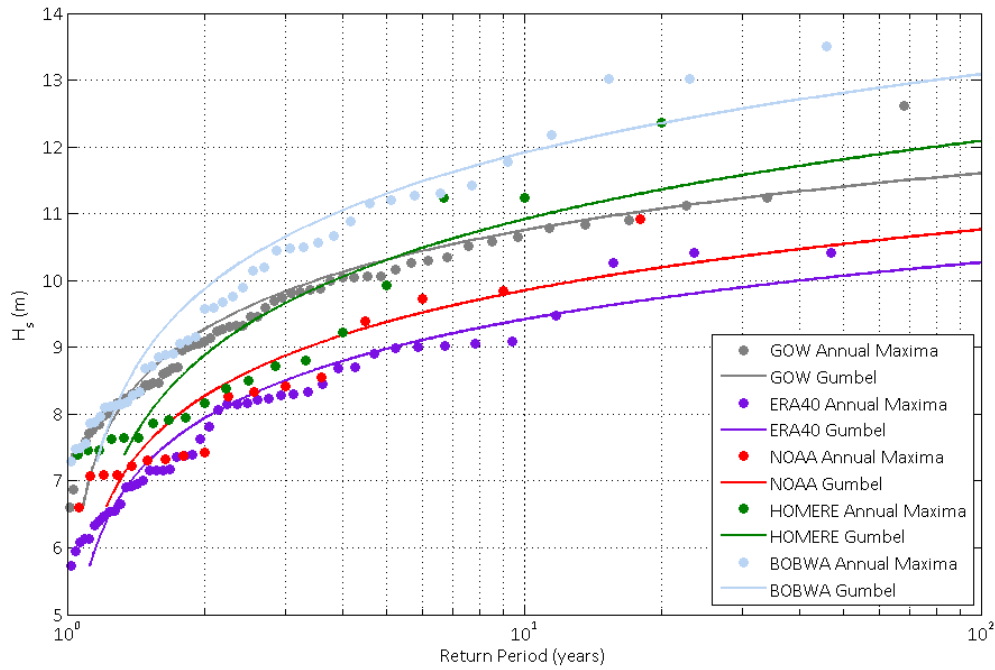


Figure 35. Gumbel distribution at 2.5W 45N for the different databases.

Database	Location (μ)	Scale (σ)	H_s R=100y (m)
IH-GOW	9.721	1.233	11.61
ERA40	8.389	1.231	10.26
NOAA	8.751	1.315	10.73
HOMERE	9.500	1.694	12.09
BoBWA	10.493	1.696	13.09

Table 13. Gumbel distribution parameters estimated for the different databases.

GPD fitting comparison

With the samples obtained by performing a Peaks Over Threshold analysis to all of the databases a General Pareto Distribution was fitted. The comparison is shown in Figure 36. As seen, the behavior of the different databases is as same as the one seen with the GEV: shorter databases (NOAA and HOMERE) present heavy tailed distributions while longer database present bounded distributions.

Database	Scale (σ)	Shape (ξ)	Threshold	H_s R=100y (m)
IH-GOW	1.199	-0.236	8.670	12.04
ERA40	1.263	-0.295	7.350	10.53
NOAA	0.545	0.383	7.570	14.2
HOMERE	0.561	0.586	7.830	>19
BoBWA	1.831	-0.305	8.935	13.46

Table 14. GP distribution parameters estimated for the different databases.

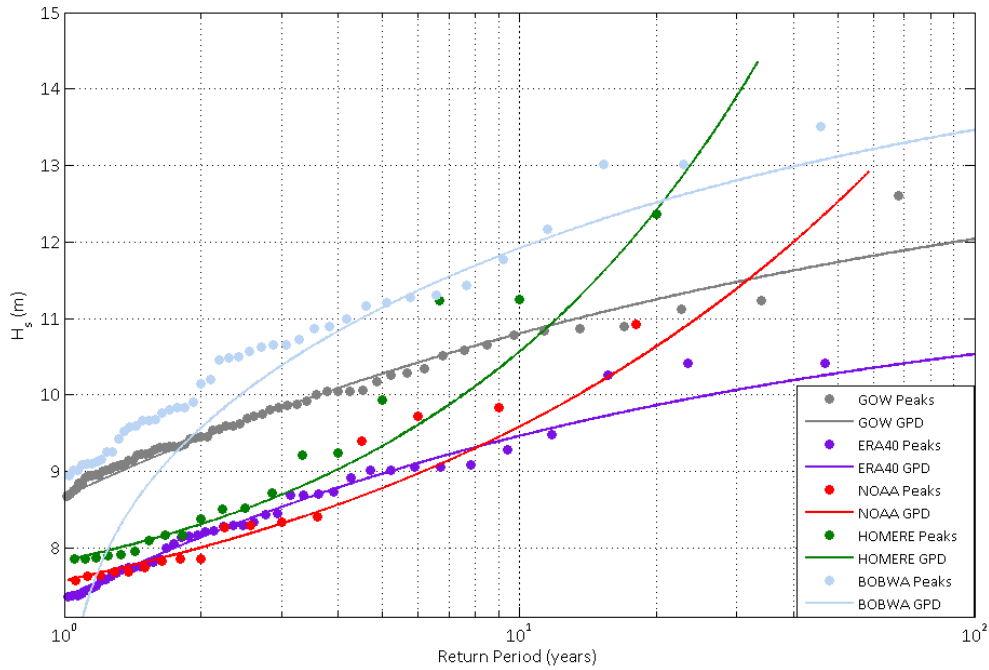


Figure 36. GP distribution 2.5W 45N location for the different databases.

Exponential fitting comparison

As it was done with the GEV, the shape parameter has been fixed to analyze its influence. In this case, fixing the shape parameter to 0 within a Generalized Pareto Distribution is the same as fitting an Exponential Distribution. Results to this test are shown in Figure 37. As it can be seen in the graphical representation, by fixing the shape parameter, the heavy tailed behavior of the shorter databases (NOAA and HOMERE) is avoided; but there are discrepancies of up to 4 meters between the highest estimator of the 100year return value (for BoBWA database) and the lowest one (for NOAA database).

Database	Scale (σ)	Threshold	H_s R=100y (m)
IH-GOW	0.980	8.670	13.18
ERA40	0.970	7.350	11.81
NOAA	0.823	7.570	11.36
HOMERE	1.070	7.830	12.76
BoBWA	1.397	8.935	15.37

Table 15. Exponential distribution parameters estimated for the different databases.

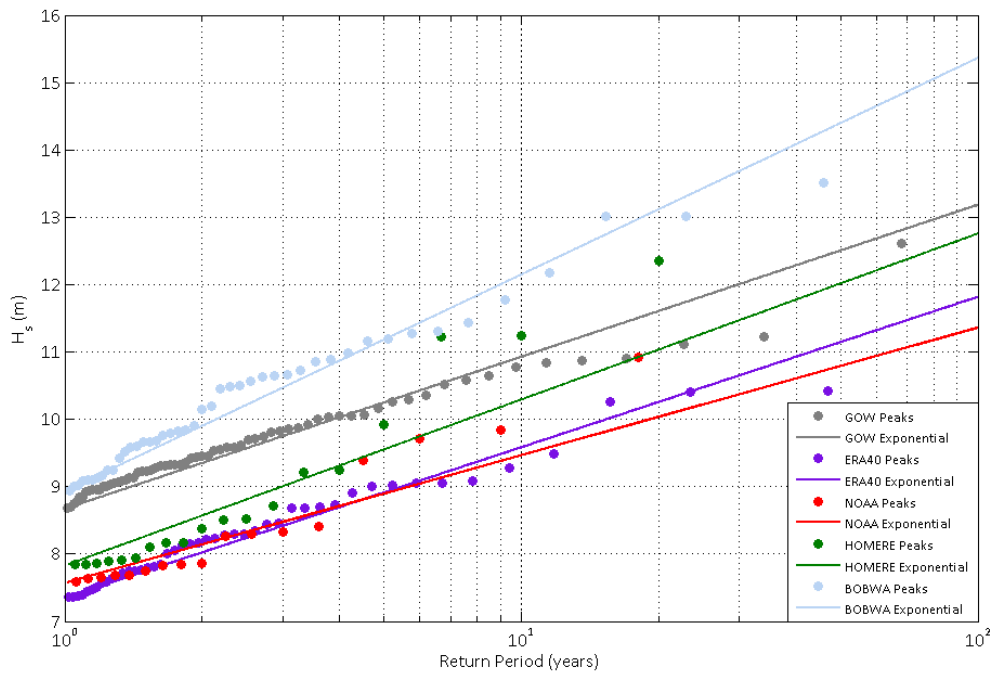


Figure 37. Exponential distribution at 2.5W 45N for the different databases.

The results shown for the different distributions (from Figure 34 to Figure 37), are an example of the important discrepancies existent between databases. In the case of the GEV and the GP distributions (Figure 34 and Figure 36, respectively); the time series from NOAA and HOMERE present a Fréchet behavior, with heavy tailed distributions and positive shape parameters. While the other three databases present shape parameters close to 0 in the GEV distribution case (Table 12) and negative in the GPD case (Table 14). In case of fitting a Gumbel or an Exponential distributions (Figure 35 and Figure 37, respectively), because of having fixed the shape parameter to be equal to 0, the behavior of the 5 databases is similar. In terms of 100-years return period value, there are great discrepancies between databases and between distributions fitted, having values between 10.26m and 15.37m for those distributions without shape parameter (Gumbel and Exponential) and higher variations in the two distributions where the shape parameter was involved (GEV and GPD).

3. MULTIVARIATE ANALYSIS

The characterization of marine climate in terms of extreme events, and therefore the design or coastal defenses or coastal management requires a careful treatment of the variables involved. Wave climate is usually described in terms of sea-states variables like significant wave height, H_s , and spectral peak period, T_p . But, other variables might be of interest in the process like wave mean direction, θ_m , sea level, η , or even wind parameters like wind speed and wind direction.

The complexity of a multivariate extreme values analysis relies in the lack of knowledge of the dependences between variables in the tail of the distribution. Then, we need to make an estimation of the tail behavior of the multivariate distribution.

Multivariate extreme value distribution is a wide research topic, with many contributions in the last years. As an example, (Jonathan 2013) and (Monbet 2007) give a survey of the existent models to deal with multivariate analyses.

Generally speaking, existent multivariate approaches for solving extremes can be divided into 5 main groups. In the following sections (3.1 to 3.5) those models will be succinctly explained. Then, in section 3.6 Bayesian inference techniques will be briefly presented and in section 3.7 the characteristics of reliability methods will be explicated.

3.1. Parametric models

The aim of the parametric models is try to describe the tail dependences between variables using a finite number of parameters. The initial approach to achieve this would be to impose a given distribution (bivariate Normal, multivariate kernel ...) to the variables of our study and then estimate its parameters.

In the last decades, a large number of parametric models have been proposed trying to establish the relationship between different met-ocean variables. For example (Ferreira 2002) propose a parametric model of significant wave height and mean period based on a bivariate density function or (C. C. Guedes Soares 2000) provides a bivariate model based on ARMA (autoregressive moving average) techniques to establish the joint distribution of wave height and wave period. In (Cai 2011), they propose a multivariate model able to deal with more than two variables also based on ARMA models, although in the paper they only apply it to the simulation of significant wave height wave period. A review in this kind of models could be find in (Kotz 2000), (Beirlant 2004) or (Monbet 2007).

The main drawback of these methods is that the estimation of parameters might become a tedious or even unaffordable if the number of variables is too high, therefore they are limited to 2 or even 3 variables. Moreover, they require the imposition of a particular structure to the extreme tail behavior of both the marginal and the joint distribution. This may not be accurate in all the cases and also may lead to common misunderstandings. For example, in the case of two variables that both follow a normal distribution it does not imply that the joint distribution follows a binormal distribution.

3.2. Conditional extreme models

Parametric models showed some limitations in terms of number of variables involved and difficulties to assign a given distribution to an empirical sample. To overcome this were developed the conditional extreme models. These models are based on taking each variable as the conditioning variable by turns and model the remaining variables conditioned to the large values of the one being the conditioning one. For example, in a bivariate case the model process would encompass modeling $X_2|X_1=x_1$ for large x_1 and $X_1|X_2=x_2$ for large x_2 .

The two main advantages of these methods are that they can be extended to higher dimensions than the other methods easily and also to samples in which not all the variables are simultaneously extreme.

(Haver 1985) applied a conditional extreme model for significant wave height and wave period based on a Weibull distribution for the significant wave height and a log-normal for the wave period. The selection of those distributions may be questionable, so (Heffernan 2004) went a step forward avoiding these problems. Their model can be succinctly described as follows:

- i) Transform each variable into a Gumbel distribution.
- ii) Estimate the dependence between variables for extremes.
- iii) Sample values that follow these dependences.
- iv) Transform them back into the original scale.

The major advantages of their model are that there is no need of estimating the coefficient of tail dependence, it can be applied to higher dimension problems and there is no need of assuming given distributions to the marginals. Moreover, compared to max-stable models this kind of models admits all forms of extremal dependence.

Recently, the method proposed by (Heffernan 2004) has been introduced into the modeling of flood risk analyses by (Wyncoll 2013) sampling extreme hydraulic loading events (rainfall) in several locations within a spatial grid. Moreover it has been also introduced into the analysis of coastal flooding and structural design by sampling extreme met-ocean conditions, waves and winds, (Gouldby 2014). This method have been compared with two other methods for a bivariate case of extreme rainfall and storm surge and its application to flooding risk in (Zheng 2013)

3.3. Copula models

Copula models are based on Sklar's theorem which says that any multivariate distribution can be written in terms of univariate marginal distribution functions of the variables involved and a copula function that describes the dependence between the variables (Sklar 1996).

In the last years, the use of copula functions to approach a multivariate analysis has become widely popular. In the field of coastal and ocean sciences applications can be found in (De Waal 2005), (De Michele 2007), (Salvadori 2011), (T. M. Wahl 2012) or (T. B. Wahl 2013).

The Copula function to calculate the joint probabilities can belong to different Copula families: i.e.; the Elliptical, the Normal, or the Archimedean family among others. For the case of hydrological sciences, the most often used Copulas belong to the Archimedean family due to their flexibility and because they are easy to construct and able to model a wide range of dependence structures (Favre 2004).

The utility of these models to solve several problems have been demonstrated. But, although they represent a promising alternative to address multivariate problems, they present two main drawbacks: the first one is that for a practitioner not familiar with Copulas the selection of the best Copula function to be used could be difficult; and the second one is that these models become not practical for more than 2-3 variables. More details regarding these methods could be found in (Nelsen 1999).

3.4. Extreme dependence models

Extreme dependence model consist of estimating directly from the sample, a tail dependence coefficient η (a constant in $[0, 1]$) which quantifies the extremal dependence for the joint distribution. When η is equal to 1 it means the distribution is asymptotically dependent, otherwise it is asymptotically independent.

In (Ledford 1997) they introduce the estimation of the dependence coefficient within joint tail regions for those cases where the variables involved have Gumbel marginal. In addition, (Schlather, Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution 2001) provides examples for a wide range of multivariate extreme value distributions. There is also a summarized η estimation explanation in (Jonathan 2013).

The main disadvantage of these methods is that they require a careful treatment of the obtained results because any misunderstood in the asymptotic dependence-independence assumption may change drastically the occurrence rate of joint events. But on the contrary, they are useful because they can be applied to different forms of extremal dependence.

3.5. Max-stable models

The theory of multivariate max-stable processes is the generalization of the univariate max-stable theory used to motivate the choice of the Generalized Extreme Value Distribution in the case of independent and identically distributed (i.i.d.) events. Models based in this theory and the their basic principles can be found in (de Haan 1984) (Schlather, Models for Stationary Max-Stable Random Fields 2002), (Padoan 2010)

Max-stable models are a very practical option for spatial problems. They consist of defining the distribution of maxima (also called the componentwise maximum, which is the maximum value across all the observations for that variable) and make the extreme value analysis to sample extracted from the data.

These models are useful to model certain types of dependence structure (asymptotic dependence and perfectly independence), but not asymptotically independence.

Because they only apply to componentwise maxima, they may therefore exaggerate dependence effects (Jonathan 2013).

3.6. Bayesian inference models

Generally speaking, the goal of Bayesian inference is to update prior information about estimated parameters of a model with information gained from observation to generate improved posterior information. It enables the combination of data from observation with past experience and expertise in a consistent manner.

Regarding the existent literature related Bayesian inference to extremes the reader can be addressed to the reviews developed by (S. Coles 2001) or (Beirlant 2004). And more specifically, related to metocean applications: (S. a. Coles 2005) use Bayesian inference methods to estimate extreme surges along the coast of UK or in (Scotto 2007) they develop a method that combines Bayesian theory and extreme value techniques to obtain long-term estimations of significant wave height in the North Sea based in a Markov Chain Monte Carlo Scheme.

However, Bayesian inference presents some difficulties. Its main drawback is its subjective nature; different practitioners would set different prior information and this would lead into different results for the same initial data. Moreover, parameter estimation using Bayesian inference in high dimensions can become a very computational demanding problem.

3.7. Methods of Reliability Estimation

In engineering, a reliable design of any structure must be done so as to ensure that during its service life, some limit states are not exceeded. By service life is meant the period of time during which the structure must maintain acceptable safety, functionality and aspect conditions without requiring any rehabilitation operation. The limit state refers to the threshold above which the structure is not able to accomplish any of the functions for which it was designed (Melchers 1999).

From a practical point of view, in the maritime structure design, reliability methods are divided into Level I, Level II and Level III methods. By the application of a reliability method, a design that ensures safety, service and operation over the structure's lifetime can be formulated.

3.7.1. Level I

Level I methods are those based on safety factors estimation. There are two main branches within these methods: methods based on global safety factors, and those based on partial safety factors for all the involved variables. With the first set, a minimum safety factor

that the infrastructure has to ensure over its lifetime is estimated. This coefficient is obtained from the main failure modes and lifetime of the structure, and from the limit states, ultimate and service. On the other hand, methods based on the estimation of partial safety factors disaggregate the safety factor into several partial factors. Each of them is obtained through the quantiles of the different involved variables. Global safety factors are easier to be estimated than partial factors, but they might be more imprecise.

Level I methods represent an easy and traditional approach to structural design. They are very useful and easily applied methods, but they are not able to inform about the possibility of not satisfying the requirements. Examples of these methods applied to coastal structures can be found in (Burcharth 1999) or (Kim 2013).

3.7.2. Level II

Level II and III methods are based on estimating the failure probability, p_f . An analytical solution of p_f , where all the involved variables are taken into account, can be expressed as:

$$p_f = \int_{g(x_1, x_2, \dots, x_n) \leq 0} f_{x_1, x_2, \dots, x_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

Where $g(x_1, x_2, \dots, x_n) \leq 0$ is the equation that defines the limit state of the structure design and $f_{x_1, x_2, \dots, x_n}(x_1, x_2, \dots, x_n)$ is the density function of all the variables involved in terms of loads and resistances. Usually, the achievement of an analytical solution of this integration is not possible. There are different approaches to estimate the failure probability. In the case where the involved variables are not correlated it is possible to estimate the two first moments of the joint distribution f and proceed with them to the design. But, in the case of correlated variables, they must first be transformed into independent variables. That way, the joint distribution can be expressed as a product of marginal distributions.

Level II methods are also known as FOSM (First Order Second Moment) methods. The theoretical basis of FOSM lies in the statement that satisfactory estimates of the parameters of a distribution may be given by first order approximations of Taylor series expansions of second moment parameters (e.g. mean and variance) of a random variable calculated from samples (Melchers 1999). Practically, FOSM methods simplify the ultimate limit state equation to a straight line (in the case of being in dimension 2), and instead of using the actual joint distribution of the variables, uses its two first moments: its mean and variance. With that, it is possible to define the point on the linear approximation for which the distance is the minimum to the central value of the joint distribution (its mean). This point of the linear approximation will be the design point (Figure 38) and distance β represents the reliability index as a constant measurement, to determine whether the system is safe or not.

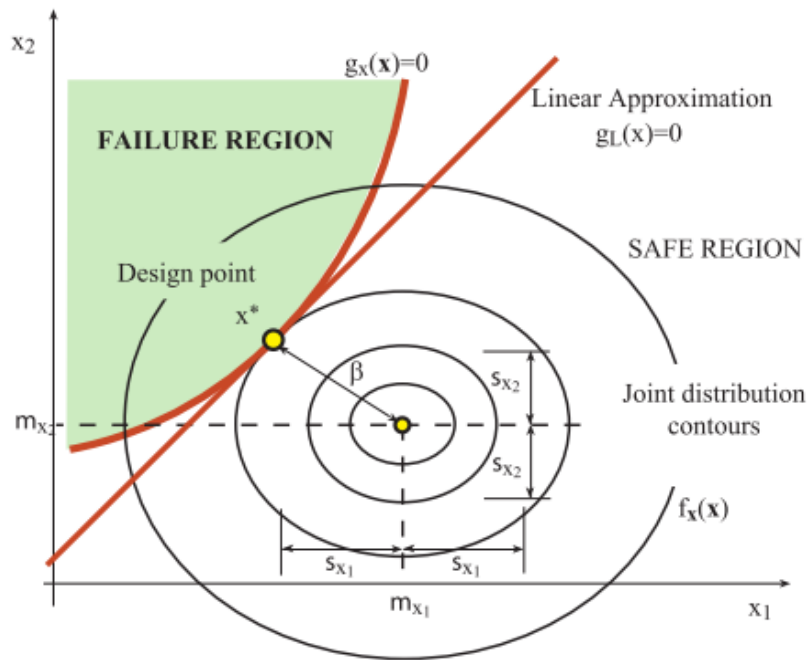


Figure 38. Limit state surface graphical and its linear approximation for a 2-D case.

3.7.3. Level III

Finally, Level III methods are those that try to solve equation of p_f by integration techniques (analytically or numerically), or by using simulation techniques. Only in a few cases does an analytical solution exist, so normally simulation techniques are used. Level III methods based on the estimation of the failure region can be split into FORM (First Order Reliability Methods), and SORM (Second Order Reliability Methods). On the other hand, Level III methods that consist of the synthetic generation of involved variables are generally based on simulation techniques such as Monte Carlo (Metropoli 1949).

FORM. First Order Reliability Methods

While FOSM estimates the failure probability considering the two first moments of the random variables (mean and variance), FORM methods take into account the actual distributions of the variables but keep on using a linear approximation of the limit state problem. The process undertaken by FORM methods is similar to FOSM ones, the main difference is related to the transformation applied to the random variables to transfer them to the multivariate normal space. More detailed information regarding these methods could be found in: (Freudenthal 1956), (Hasofer 1974), (Rackwitz 1978) and (Ditlevsen 1973) among others.

SORM. Second Order Reliability Methods

There are cases where it is not feasible to approximate the state limit equation to a straight line (in the case of 2D). This occurs when the state limit equation is a sharp curve, or

when this curve is not so sharp in the original space but is sharp in the transformed space. When this occurs, the problem can be solved by using SORM (Second Order Reliability Methods).

These second order methods are based on an approximation of the state limit equations by parabolic or spherical approximations in the surroundings of the maximum probability design point. The results obtained by these methods are very accurate and can be more efficient than Monte Carlo simulation techniques at estimating extreme percentiles.

Among these methods can be listed the approaches proposed by: (Breitung 1984), (Castillo, Estimating extreme probabilities using tail simulated data. 1996), (Castillo, High probability one-sided confidence intervals in reliability models 1997), (Haskin 1996), (Koyluoglu 1994).

i-FORM. Inverse First Order Reliability Methods

(Winterstein 1993) introduced a novelty way to deal with the problem of a multivariate extreme values analysis. In their approach, instead of estimating β as the minimum distance between the joint distribution barycenter and the limit state equation, they initially specify a probability of failure and thus a value for β . Then they search for all the possible FORM with a certain return period (hyper-sphere with radius β). This way they obtain all the maximum responses the structure must withstand. This new approach allows the determination of multivariate contours associated to a previously specified return period.

4. CHARACTERISING UNCERTAINTIES

In any statistical analysis it is needed to quantify the accuracy of the obtained estimations. In term of economic losses or casualties, the characterization of these uncertainties acquires a vital importance when the study's objectives are the extreme values analysis of environmental variables for any coastal protection design or management.

Quantify the uncertainty of an estimation made based on probabilities is a complex issue because there are several sources that may introduce uncertainties in our estimation: error induced by experimental measurements (buoys), error induced by the models and the hypothesis they assume (reanalysis models), error induced by the assumption of a statistical distribution or method and the estimation of its parameters (block maxima or POT, GEV or GPD), etc.

In this study there is no intention to solve this issue, but here are summarized some of the most used methods to give an estimation of the uncertainties in terms of probabilities.

A practical form to characterize the uncertainty of an estimated value is giving a confidence interval around it. A confidence interval is a range of values (interval) around the estimated parameter value with a confidence level. This confidence level indicates the probability that this interval captures the parameter estimation given several samples. For example, if we have a sample of a variable which its mean value is estimated to be 5, with a confidence interval of (4,6) with a confidence level of 99% it means than in 1% of samples of this variable we will obtain a mean out of the interval (4,6). In practice, confidence intervals are usually estimated with a 95% confidence level, but in graphical representations it is also usual to represent 50% and 99%.

4.1. Standard Likelihood Theory or Theoretical Delta Method.

For a large sample size n and assuming that a proposed model is valid, the distribution of the ML estimators $\hat{\theta}$ is approximately multivariate normal with mean θ (the true parameter values) and covariance matrix given by the inverse of the observed information matrix $I_o(\theta)$,

$$I_o(\theta) = -\frac{\partial^2 l(\theta)}{\partial \theta_i \partial \theta_j}$$

Evaluated at $\theta = \hat{\theta}$. If an arbitrary term in the inverse of $I_o(\theta)$ is s_{ij} , the square root of the diagonal entry s_{ii} is approximately the standard error, $se(\hat{\theta}_i)$, of the i^{th} component of the ML estimators vector $\hat{\theta}$. Therefore confidence intervals for θ_i can be obtained in the form $[\hat{\theta}_i - z_{\alpha} se(\hat{\theta}_i), \hat{\theta}_i + z_{\alpha} se(\hat{\theta}_i)]$, where $z_{0.95} = 1.96$ gives a 95% confidence interval (Méndez 2006).

This method has the advantage of being an analytical approach and being fast to compute. But in the other hand it can be only applied to Maximum Likelihood estimators and might not be very accurate for small samples.

4.2. Resampling Confidence Intervals or Bootstrapping Method

This method is based on simulating numerous random samples with the same size of the original one and estimate same extreme value model parameters with each of them. Then, from the range of estimators obtained, confidence intervals can be estimated. If quantiles are estimated for each set of estimators, it is also possible to derive confidence intervals on quantiles.

This method has the advantage of being generally accurate and applicable to any estimation method. But it can become computationally demanding for large samples.

4.3. Bayesian Confidence Intervals or Credible Intervals

Credible intervals or also called Bayesian Confidence Intervals in Bayesian statistics are analogous but not equal to confidence intervals in frequentist statistics.

To express uncertainty, the frequentist approach uses a confidence interval: a range of values designed to include the true value of the parameter with some minimum probability, say 95%. So that out of every 100 experiments run start to finish, at least 95 of the resulting confidence intervals will be expected to include the true value of the parameter. In the other 5 experiments, the true value of the parameter might be outside of the confidence intervals.

Bayesian approaches formulate the problem differently. Instead of saying the parameter simply has one true value, the Bayesian paradigm says the parameter's value is fixed but has been chosen from some probability distribution - known as the prior probability distribution. This *prior* might be known or it might be an estimation. If no knowledge is available *a priori*, the prior probability distribution might be the uniform distribution. Within the Bayesian inference some data are collected, and then the probability distribution of the parameter given the data is calculated. This new probability distribution is called the *a posteriori probability* or simply the *posterior*. Bayesian approaches summarize the uncertainty by giving a range of values on the posterior probability distribution that includes for instance 95% of the probability; this is called a *95% credibility interval*. Contrary to confidence intervals in the frequentist approach, the actual value of the parameter has a particular probability (e.g. 95%) of being in the credibility interval given the data actually obtained.

Within this Bayesian framework, there is the possibility of avoiding the credibility intervals working directly with the predictive probability distribution. This distribution aggregates all the uncertainties combining and integrating them into a one single probability distribution. According to (S. a. Coles 2005) and applied to a lock maxima case, it can be defined by:

$$P(Z < z_p | H) = \int_{\Theta} G(z_p; \theta) f(\theta | H) d\theta,$$

Where Z are the annual maxima, $G(z_p; \theta)$ is the conditional annual maximum GEV distribution with model parameters $\theta = (\mu, \sigma, \xi)$; and $f(\theta | H)$ is the posterior term being H a subset of the observed data that contains all of the relevant information on extreme events. The reader can refer to (S. a. Coles 2005) to have a more detailed explanation of this approach as well as a practical application.

5. SOFTWARES

Here, a summary of available softwares that solve extreme values analysis or at least help with time series analysis, most of them are explained in the review of softwares and its features done by (Stephenson 2006). They are divided following a writing language criteria

5.1. R

BSquare Conduct Bayesian quantile regression for continuous, discrete and censored data.

`dqreg` Plots posterior density for parametric basis functions.

`qreg` MCMC code for the quantile regression model of Reich and Smith, 2013.

`qreg_spline` Quantile regression using spline basis functions for the quantile process.

`qr_plot` Plot of posterior distribution of effects

<http://www4.stat.ncsu.edu/~reich/QR/>

copula Modeling multivariate dependences with copulas.

The copula package provides:

- Classes of commonly used copulas including elliptical (normal and t; `ellipCopula`), Archimedean (Clayton, Gumbel, Frank, Joe, and Ali-Mikhail-Haq; `archmCopula`), extreme value (Gumbel, Husler-Reiss, Galambos, Tawn, and t-EV; `evCopula`), and other families (Plackett and Farlie-Gumbel-Morgenstern).
- Methods for density, distribution, random number generation (`dCopula`, `pCopula` and `rCopula`); bivariate dependence measures (ρ , τ , etc), perspective and contour plots.
- Functions for fitting copula models including variance estimates (`fitCopula`).
- Independence tests among random variables and vectors.
- Serial independence tests for univariate and multivariate continuous time series.
- Goodness-of-fit tests for copulas based on multipliers, and the parametric bootstrap, with several transformation options.
- Bivariate and multivariate tests of extreme-value dependence.
- Bivariate tests of exchangeability.

evd univariate and multivariate parametric EV distribution .

Extends simulation, distribution, quantile and density functions to univariate and multivariate parametric extreme value distributions, and provides fitting functions which calculate maximum likelihood estimates for univariate and bivariate maxima models, and for univariate and bivariate threshold models.

evdbayes functions for the bayesian analysis of extreme value models.

`Accept Rate` Compute suited proposal standard deviations for the MCMC algorithm.

`dinfo` Show means, variances and modes for beta and gamma distributions.

Jumps Definition These are function useful to define reversible jumps for the algorithm. There is currently three functions: `jacFun` which evaluates the jacobian, `movType1` which returns a special proposal of type 1 and `movTyp2` which returns a special proposal of type2.

`lh` Calculate log-likelihoods for the `gev`, orderstatistics or point process models.

`mc.quant` Compute `gev` quantiles from samples stored within a Markov chain, corresponding to specified probabilities in the upper tail.

`mposterior` Maximizing prior and posterior distributions for the location (with optional trend), scale and shape parameters under the `gev`, order statistics or point process models.

`posterior` Constructing MCMC samples of prior and posterior distributions for the location (with optional trend), scale and shape parameters under the GEV, order statistics or point process models.

`posterior.mix` Constructing MCMC samples of prior and posterior distributions for the location (with optional trend), scale and shape parameters under the GEV, order statistics or point process models. Note this function is only effective for a mixture denned in the prior distribution.

`prior` Constructing prior distributions for the location, scale and shape parameters using normal, beta or gamma distributions. A linear trend for the location can also be specie, using a prior normal distribution centered at zero for the trend parameter.

`rainfall` A numeric vector of length 20820 containing daily aggregate rainfall observations, in millimeters, recorded at a rain gauge in England over a 57 year period, beginning on a leap year. Three of these years contain only missing values.

`rl.pred` Produce return level plots depicting prior and posterior predictive GEV distributions.

`rl.pst` Produce return level plots depicting prior and posterior distributions of GEV quantiles.

evir Extreme Values in R.

Functions for extreme value theory, which may be divided into the following groups; exploratory data analysis, block maxima, peaks over thresholds (univariate and bivariate), point processes, GEV/GPD distributions.

extRemes

A suite of functions for carrying out analyses on the extreme values of a process of interest; be they block maxima over long blocks or excesses over a high threshold.

It provides a windows GUI for the '`ismev`' package, allowing easy use of the tools that the `ismev` package provides, in addition to a few extra useful functions.

<http://www.ral.ucar.edu/staff/ericg/extRemes/>

extremevalues Detect extreme values in one dimensional data.

`evGui` Opens a Graphical User Interface and plots results. Options of the extreme value package functions can be set and results are updated instantly. Includes a code generator button.

`extremevalues` This package offers outlier detection and plot functions for univariate data. The package is the implementation of the outlier detection methods introduced in the reference below. Briefly, the methods work as follows. Using a subset of the data, the parameters for a model distribution are estimated using regression of the sorted data on their QQ-plot positions. A value in the data is an outlier when it is unlikely to be drawn from the estimated distribution. There are two methods to determine the "unlikelyness". The first, called "Method I", determines the value above which less than ρ observations are expected, given the total number of observations in the data. Here ρ is a parameter which should have a value of 1 or less. The second notion of unlikelyness uses the fit residuals. Extremely large or small values are outliers when their residuals are above or below a confidence limit α , to be determined by the user.

`getOutliers` is a wrapper function for `getOutliersI` and `getOutliersII`. Both methods use the subset of y -values between the F_{min} and F_{max} quantiles to fit a model cumulative density distribution. Method I detects outliers by checking which are below (above) the limit where according to the model distribution less than $\rho[1]$ ($\rho[2]$) observations are expected (given $\text{length}(y)$ observations). Method II detects outliers by finding the observations (not used in the fit) whose fit residuals are below (above) the estimated confidence limit $\alpha[1]$ ($\alpha[2]$) while all lower (higher) observations are outliers too.

`invErf` Inverse error function

`outlierPlot` This is a wrapper for two plot functions which can be used to analyze the results of outlier detection with the extreme values package.

`pareto` Pareto density distribution, quantile function and random generator.

`ismev` univariate extreme value modeling.

Functions to support the computations carried out in 'An Introduction to Statistical Modeling of Extreme Values' by Stuart Coles. The functions may be divided into the following groups; maxima/minima, order statistics, peaks over thresholds and point processes.

`lmom` functions to compute L-moment estimates for extreme value distribution parameters.

Functions related to L-moments: computation of L-moments and trimmed L-moments of distributions and data samples; parameter estimation; L-moment ratio diagram; plot vs. quantiles of an extreme-value distribution.

L-moments are measures of the location, scale, and shape of probability distributions or data samples. They are based on linear combinations of order statistics. Hosking (1990) and Hosking and Wallis (1997, chap. 2) give expositions of the theory of L-moments and L-moment ratios. Hosking and Wallis (1997, Appendix) give, for many distributions in common

use, expressions for the L-moments of the distributions and algorithms for estimating the parameters of the distributions by equating sample and population L-moments (the “method of L-moments”). This package contains R functions that should facilitate the use of L-moment-based methods. For each of 13 probability distributions, the package contains functions to evaluate the cumulative distribution function and quantile function of the distribution, to calculate the L-moments given the parameters and to calculate the parameters given the low-order L-moments.

lmomRFA functions for regional frequency analysis (RFAA).

This package implements methods described in the book “Regional frequency analysis: an approach based on L-moments” by J. R. M. Hosking and J. R. Wallis. It is a supplement to the **lmom** package, which implements L-moment methods for more general statistical applications.

lmomco compute L-moments, trimmed L-moments, L-comoments, and probability-weighted moment estimation for extreme value distributions.

The package is a comparatively comprehensive implementation of the theory of L-moments.

L-moment, probability-weighted moment (PWM), and parameter estimation for numerous familiar and not-so-familiar distributions are provided. L-moment estimate for the same distributions are provided. L-moments are analogous to product moments; however, L-moments have many advantages including unbiasedness, robustness, and consistency. L-moments can outperform maximum likelihood for small to moderate samples.

POT POT analysis for both univariate and bivariate cases .

Some functions useful to perform a Peak Over Threshold analysis in univariate and bivariate cases. Some features:

- Analysis-of-deviance tables for univariate and bivariate cases.
- GPD for univariate and bivariate cases.
- Dependence measures for EVA
- Extremes clusterization
- Extremal Index: estimation, plot, density.
- Estimate confidence intervals and convergence assessment
- Threshold Selection: dispersion index, mean residual life and threshold choice plots
- Markov Chain: fitting MC to POT, simulation.
- Compute SampleL-moments
- Graphical Diagnostic: the univariate/bivariate GPD Model, Markov Chains for all exceedances, the bivariate EV distribution model, pp plot, qq plot, return level plot, spectral density

QRM

Routines developed by Alexander McNeil, then included in the packaged **QRMLib** (see **EVIS** in section **S/S-Plus**), based on his book: **Quantitative Risk Management: Concepts, Techniques, and Tools** by Alexander J. McNeil, Rüdiger Frey, & Paul Embrechts.

Includes tools related to Copula models, GEV and GPD fitting and graphical checking plots like PP or QQ plots among other features. It is focused on econometrics and finances.

Rmetrics

Open source solution for teaching financial market analysis and valuation of financial instruments. With hundreds of functions build on modern methods, Rmetrics combines explorative data analysis and statistical modeling and rapid model prototyping. Rmetrics covers Time Series Econometrics, Hypothesis Testing, GARCH Modeling and Volatility Forecasting, Extreme Value Theory and Copulae, Pricing of Derivatives, Portfolio Analysis, Design and Optimization.

<https://www.rmetrics.org/>

SpatialExtremes

SpatialExtremes is a package that provides tools for the statistical modeling of spatial extremes. The main types of statistical models thus far implemented are:

- Max-stable processes
- Conditional max-stable simulations
- Bayesian Hierarchical models
- Copula approaches
- GEV and GP distributions

<http://spatialextremes.r-forge.r-project.org/index.php>

texmex conditional EVA approach of H&T 2004.

Statistical extreme value modeling of threshold excesses, maxima and multivariate extremes. Univariate models for threshold excesses and maxima are the GP, and GEV model respectively. These models may be fitted by using maximum (optionally penalized) likelihood, or Bayesian estimation, and both classes of models may be fitted with covariates in any/all model parameters. Model diagnostics support the fitting process. Graphical output for visualizing fitted models and return level estimates is provided. For serially dependent sequences, the intervals declustering algorithm of Ferro and Segers is provided, with diagnostic support to aid selection of threshold and declustering horizon. Multivariate modeling is performed via the conditional approach of Heffernan and Tawn, with graphical tools for threshold selection and to diagnose estimation convergence.

VGAM allows for modeling parameters as linear or smooth functions of covariates.

Vector generalized linear and additive models, and associated models (Reduced-Rank VGLMs, QuadraticRR-VGLMs, ReducedRank VGAMs). This package fits many models and distribution by maximum likelihood estimation (MLE) or penalized MLE. Also fits constrained ordination models in ecology.

5.2. Matlab

EVIM

Open source toolbox developed in Matlab to analyze and work with extreme value analysis. Its functions can be divided into: preliminary data analyses, distribution and tail estimation and predictions at the tail.

i) preliminary data analyses:

`emplot` This function plots the empirical distribution of a sample.

`meplot` This function plots the sample mean excesses against thresholds.

`qplot` This function provides a QQ plot of data against the GPD in general.

`hillplot` This function plots the Hill estimate of the tail index against the upper-order statistics (number of exceedances) or against different thresholds.

`qgev` This function calculates the inverse of a GEV distribution.

`qgpd` This function calculates the inverse of a GPD.

`pgev` This function calculates the value of a GEV distribution.

`pgpd` This function calculates the value of a GPD.

`rgev` This function generates a particular number of random samples from a GEV distribution.

`rgpd` This function generates a particular number of random samples from a GPD.

`records` This function creates a MATLAB data structure showing the development of records in a data set and expected record behavior of an i.i.d. data set with the same sample size.

`block` This function divides the input data vector into blocks and calculates block maxima or minima.

`findthresh` This function finds a threshold value such that a certain number (or percentage) of values lie above it.

ii) distribution and tail estimation:

`exindex` This function estimates the extremal index using the blocks method. The extremal index is utilized if the elements of the sample are not independent.

`pot` This function fits a Poisson point process to the data. This technique is also known as the peaks over threshold (POT) method.

`plot_pot` This function is a menu-driven plotting facility for a POT analysis.

`decluster` This function declusters its input so that a Poisson assumption is more appropriate over a high threshold. The output is the declustered version of the input. Suppose that there are 100 observations, among which there are four exceedances over a high threshold. Further suppose that these exceedances are the 10th, 11th, 53th, and 57th observations. If the `gap` option is entered in the above function as 0.05, the 11th and 57th observations are skipped, since there are fewer than five observations (5 percent of the data) between these exceedances and the nearest previous exceedances. This function is useful when exceedances are coming from a Poisson process, since Poisson processes assume low density (no clustering) and independence.

`gev` This function fits a GEV distribution to the block maxima of its input data.

`plot_gev` This function is a menu-driven plotting facility for a GEV analysis.

`gpd` This function fits a GPD to excesses over a high threshold.

`plot_gpd` This function is a menu-driven plotting facility for a GPD analysis.

iii) predictions at the tail:

`gpdq` This function calculates quantile estimates and confidence intervals for high quantiles above the threshold in a GPD analysis.

`shape` This function is used to create a plot showing how the estimate of the shape parameter varies with threshold or number of extremes.

`quant` This function creates a plot showing how the estimate of a high quantile in the tail of data based on a GPD estimation varies with varying threshold or number of exceedances.

<http://www.sfu.ca/~rgencay/index.html>

EXTREMES

EXTREMES is an interface developed in Matlab to work with tools written in C++. This software is dedicated to extreme values study, focused on extreme quantiles estimation and model selection for distribution tails. Its functions can be divided into:

i) Statistical functions:

- Sample simulation
- Plotting distribution related functions
- Parameter estimation
- Non parametric estimation of density
- Parametric estimation of quantiles
- Anderson-Darling or Cramer-von Mises test

ii) EVA functions:

- Estimation of extreme value analysis for both GEV and GP.
- Graphical and numerical tests (Hill, Dekkers ...) to check model fittings.
- Extreme quantiles estimation

iii) Other procedures:

- Bayesian regularization procedure
- Central fits tests (-Darling or Cramer-von Mises) to check data distribution in both central and extremal ranges.

<http://mistis.inrialpes.fr/software/EXTREMES/>

IH-AMEVA

AMEVA free software is a compendium of Matlab based functions to study and characterize environmental variables. It is divided into different modules all of them independent. This modules are:

i) Calibration:

Set of functions that allow a directional calibration by using quantiles and assuming a smooth directional variation by means of splines.

ii) Descriptive statistic:

Package developed to perform statistical analysis of environmental variables: clusterization techniques, several diagnosis plots, persistence analysis, temporary maximum and occurrence charts among others.

iii) Distributions fit:

Fitting functions to different known distributions (normal, exponential, rayleigh, gamma, gumbel, extreme value, logistic, t-student, lognormal, weibull, weibul min, weibull max, beta, pareto) for both kind of data: continuous and discrete. Also the possibility of several diagnosis plots is available: time series, PDF, CDF, PP plot, QQ plot and probability plot. Confidence intervals can be also estimated.

iv) Extremal statistic: GEV and POT-Pareto Poisson:

GEV tool is used to fit monthly or yearly maxima. This maxima can be obtained with the function specially determined for it. Within the GEV some settings are displayed: introduction of seasonality, long term trend or influence of covariates in the parameters estimation; criteria to stop the parameters optimization (akaike or chi-square criteria).

POT tool is used to fit independent events above a certain threshold and the annual frequency is fitted with a Poisson model. The sample of extreme events can be provided to the tool or extracted from a raw data sample. Some settings refereeing to the POT model are the definition of the threshold, the significance of the parameters or the timelaps between two consecutive events to be considered independent.

For both models some diagnosis plots can be generated: QQ plot, PP plot, autocorrelation, partial autocorrelation, aggregated quantiles, and plots referred to each parameter estimated.

Apart from these main modules it also has some secondary models of applicability to the main modules:

a) Clusterization tools: It includes three kind of clusterization techniques such as Self-Organizing-Maps (SOM); Maximi-Dissimilarity Algorithm (MDA) and K-Means Algorithm (KMA). The number of clusters can be changed as well and diagnosis plots are also generated.

b) Temporal maxima: Data selection can be performed at 5 different levels:

Weekly, monthly or yearly maxima

Weekly, monthly or yearly mean

Weekly, monthly or yearly quantile

Data above a certain threshold

Independent peaks above a certain threshold

c) Persistence analysis: This tool allows obtaining the array of durations for one, two or three variables that follow some restricting criteria.

d) Occurrence plots: It allows obtaining the occurrence frequency or the number of events between two variables. It is also possible to obtain monthly or sectorial occurrence charts. The number of classes in which the variables will be divided can be given.

e) Homocedastic/Heterocedastic models: With this tool homocedastic (constant variance and constant/linearly variable or cuadratically variable mean) and heterocedastic

(constant/linearly variable or quadratically variable mean and variance) regression models can be performed.

<http://ihcantabria.com/IHAMEVA/>

WAFO

WAFO is an open source toolbox of Matlab routines for statistical analysis and simulation of random waves and random loads.

i) Fatigue Analysis:

- Fatigue life prediction for random loads
- Theoretical density of rain flow cycles
- Sea modeling
- Simulation of linear and non-linear Gaussian waves
- Estimation of frequency spectra and directional wave spectra
- Model spectra (JONSWAP, Torsethaugen, Bretschneider, Pierson-Moskowitz,...)
- Joint wave height, wave steepness, wave period distributions

ii) Statistics:

- Extreme value analysis
- Multivariate Gaussian probabilities and expectations
- Profile log-likelihood analysis
- Regression analysis
- Hypothesis tests (one-sample and two sample T-test, Wilcoxon signed rank and rank-sum test,....)
- Kernel density estimation (Fast data-gridders with various bandwidth selectors)
- Hidden markov models
- Probability distributions and random number generators (Beta, Chi-2, Exponential, Gamma, GPD, GEV, Gumbel, Inverse Gaussian, Lognormal, Normal, Rayleigh, Students-T, Weibull,...)
- Bootstrap
- Design of Experiments

iii) Numerics:

- Derivatives (gradient, hessian, jacobian) (Available in next release)
- Integration (Gaussian quadratures (1D and 2D), simpson, Clenshaw-Curtis)
- Smoothing spline

<http://www.maths.lth.se/matstat/wafo/index.html>

5.3. S/S-Plus

EVIS - Alexander McNeil's Routines

Routines developed by Alexander McNeil, then included in the packaged QRMLib, based on his book: Quantitative Risk Management: Concepts, Techniques, and Tools by Alexander J. McNeil, Rüdiger Frey, & Paul Embrechts. (see QRM in section R)

Includes tools related to Copula models, GEV and GPD fitting and graphical checking plots like PP or QQ plots among other features. It is focused on econometrics and finances.

Initially is an open source software to be acquired after registration on the author's web page.

<http://www.macs.hw.ac.uk/~mcneil/book/QRMLib.html>

S+FinMetrics – EVANESCE : René Carmona's Routines

The Extreme Value Analysis Employing Statistical Copula Estimation (EVANESCE) library for S-PLUS FinMetrics module provides a set of functions for bivariate extreme value analyses using parametric and non-parametric copula estimation methods. It is contributed by Rene Carmona and described in Carmona (2001) and Carmona and Morrison (2001). Some of the original functions are renamed and model objects restructured when the library is incorporated into FinMetrics to be consistent with the other extreme value theory library in the package (EVIS by Alexander McNeil). This document gives an overview of the copula concept and their implementation. The detailed function documentation is available both in S-PLUS FinMetrics Reference Manual and the product online help.

<http://www.princeton.edu/~rcarmona/SVbook/svbook.html>

Stuart Coles's Routines

Prof. Coles has developed some functions in S. I haven't been able to easily find them in Internet and its description, but some packages in R (ismev, eXtremes) are based in his functions so may not be so different from them.

5.4. Others

Dataplot

DATAPLOT is an open source multi-platform software to perform scientific, engineering, statistical, mathematical, and graphical analysis with a multipatform. It is basically focused on graphical analysis and its capabilities include:

- Summary Graphics: charts, histograms, bar plots, X-Y plots, ...
- Diagrammatic Graphics: graphs with symbols (geometric, electrical, logical,...)
- Graphical Data Analysis: Histograms, scatter plots, probability plots, qq plots...
- Exploratory Data Analysis: Box plots, stem-and-leaf diagrams, bootstrap plots...
- Time Series Analysis: Lag plots, autocorrelation plots, spectral plots...
- Smoothing: Moving average, general least squares polynomial smoothing, robust...
- Fitting: Linear, polynomial, multi-linear, general spline, graphical residual analysis...
- General Data Analysis: t-test, Chi-squared tests, F tests, ANOVA...
- Statistics/Probability Calculations: cdf's, pdf's, random number generation, simulation, moment calculations...

- Multivariate Analysis: Star plots, profile plots, principal components, Andrews curves, scatter plot matrix, condition plots, multiplotting

<http://www.itl.nist.gov/div898/software/dataplot/index.htm>

Hyfran

Private software (300 \$ the basic version and 450 \$ the version plus) with tools for statistical analysis of extreme events.

Some of its features can be summarized as:

- Verification of independence or homogeneity.
- Fitting to known distributions: exp, GP, GEV, Gumbel, Weibull, Normal, LogNormal, Gamma, Poisson, ...
- Different fitting methods: method of moments, probability weighted moments, ML.
- Estimation of quantiles and confidence intervals.
- Test and graphics to decision support.

<http://www.wrpllc.com/books/HyfranPlus/hyfranplusdescrip.html>

STABLE

Calculates stable densities, cumulative distribution functions and quantiles. Also includes stable random number generation and maximum likelihood estimation of stable parameters using a fast 3-dimensional cubic spline interpolation of stable densities.

Initially is an open source software, it can be downloaded from the author's page. A registration is demanded for knowledge purposes.

<http://academic2.american.edu/~jpnolan/stable/stable.html>

6. REFERENCES

- Anderson, T. W. and Darling, D. A. «A test of goodness-of-fit.» *Journal of American Statistical Association*, 1954: 49, 765-769.
- Anderson, T. W. and Darling, D. A. «Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes.» *Annals of Mathematical Statistics*, 1952: 23, 193-212.
- Anderson, T. W. «On the distribution of the two-sample Cramer-von Mises criterion.» (*Annals of Mathematical Statistics*) 33, n° 3, 1148-1159 (1962).
- Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. *Statistics on Extremes: Theory and Applications*. Wiley, 2004.
- Belisle, C. J. P. «Convergence theorems for a class of simulated annealing algorithms R^d .» *Journal of Applied Probability*, 1992: 29, 885-895.
- Breitung, K. «Asymptotic approximations for multinormal integrals.» *Journal of Engineering Mechanichs*, 1984: 100, 3, 357-366.
- Burcharth, H.F. and Sorensen, J.D. «Design of vertical wall caisson breakwaters using partial safety factors.» *Proceedings of the 26th Conference on Coastal Engineering*, 1999: 2138-2151.
- Cai, Y. «Multivariate time series simulation.» *Journal of Time Series Analysis*, 2011: 32, 566-579.
- Castillo, E., Solares, C. and Gomez, P. «Estimating extreme probabilities using tail simulated data.» *International Journal of Approximate Reasoning: official publication of the North American Fuzzy Information Processing Society*, 1996: 17,2-3,163-190.
- Castillo, E., Solares, C. and Gomez, P. «High probability one-sided confidence intervals in reliability models.» *Nuclear science and engineering*, 1997: 126, 158-167.
- Coles, S. *An introduction to the statistical modelling of extreme values*. London: Springer-Verlag, 2001.
- Coles, S. and Tawn, J. «Bayesian modelling of extreme surges on the UK east coast.» *Philosophical Transactions of the Royal Society A*, 2005: 363, 1387-1406.
- Corder, G. W. and Foreman, D. I. «Nonparametric statistics: A step-by-step approach.» (Wiley) 2014.
- D'Agostino, R. «Transformation to normality of the null disstribution of g_1 .» (*Biometrika*) 57, n° 3, 679-681 (1970).
- D'Agostino, R., Belanger, A. and D'Agostino, R. Jr. «A suggestion for using powerful and informative tests for normality.» (*The American Statician*) 44, n° 4, 316-321 (1990).
- Davison, A. C. and Smith, R. L. «Models for exceedances over high thresholds (with discussion).» *Journal of the Royal Statistical Society*, 1990: 52, 393-442.
- de Haan, L. «A spectral representation for max-stable processes.» *The Annals of Probability*, 1984: 12, 4, 1194-1204.
- De Michele, C., Salvatore, G., Passoni, G. and Vezzoli, R. «A multivariate model of sea storms using copulas.» *Coastal Engineering*, 2007: 54, 734-751.
- De Waal, D. J. and van Gelder, P. H. A. J. M. «Modelling of extreme wave heights and periods through copulas.» *Extremes*, 2005: 8, 345-356.
- Ditlevsen, O. «Structural reliability and the invariance problem.» *Solid Mechanichs Report. University of Waterloo.*, 1973.
- Favre, A. C., El Adlouni, S., Perreault, L., Thiémonge, N. and Bobée, B. «Multivariate hydrological frequency analysis using copulas.» *Water Resources Research*, 2004: 40, DOI: 10.1029/2003WR002456.
- Ferreira, J. A. and Guedes Soares, C. «Modelling bivariate distributions of significand wave height and mean wave period.» *Applied Ocean Research*, 2002: 24, 31-45.

- Fletcher, R. and Reeves, C. M. «Function minimization by conjugate gradients.» *Computer Journal*, 1965: 7, 148-154.
- Freudenthal, A. M. «Safety and the probability of structural failure.» *Transactions, ASCE*, 1956: 121, 1337-1397.
- Gouldby, B., Méndez, F.J., Guanche, Y., Rueda, A. and Minguez, R. «A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis.» *Coastal Engineering*, 2014: 88, 15-26.
- Greenwood, P. E. and Nikulin, M. S. «A guide to chi-squared testing.» (Wiley) 1996.
- Grubbs, F. «Procedures for detecting outlying observations in samples.» *Technometrics*, 1966: 11 (1), 1-21.
- Guedes Soares, C. and Ferreira, A. M. «Representation of nonstationary time series of significant wave height with autoregressive models.» *Probabilistic Engineering Mechanics*, 1996: 139-148.
- Guedes Soares, C., Cunha, C. «Bivariate autoregressive models for the time series of significant wave height and mean period.» *Coastal Engineering*, 2000: 40, 297-311.
- Gulev, S. K., Zolilna, O. and Grigoriev, S. «Extratropical cyclone variability in the Northern Hemisphere winter from the NCEP/NCAR reanalysis data.» *Climate Dynamics*, 2001: 17,795-809.
- Hamidieh, K. «On the estimation of the extremal index based on scaling and resampling.» *Journal of Computational and Graphical Statistics*, 2010: DOI: 10.1198/jcgs.2009.08065.
- Haskin, F.E., Staple, B.D. and Ding, C. «Efficient uncertainty analyses using fast probability integration.» *Nuclear engineering and Design*, 1996: 166, 2, 225-248.
- Hasofer, A. M. and Lind, N., C. «Exact and invariant second moment code format.» *Journal of Engineering Mechanics*, 1974: 111-121.
- Haver, S. «Wave climate of northern Norway.» *Applied Ocean Research*, 1985: 7, 85-92.
- Heffernan, J. E. and Tawn, J. A. «A conditional approach for multivariate extreme values.» *Journal of the Royal Statistical Society*, 2004: 66, 3, 497-546.
- Hosking, J. R. M., and Wallis, J. R. «Parameter and quantile estimation for the generalized Pareto distribution.» *Technometrics*, 1987: 29,339-349.
- Izaguirre, C., Méndez, F. J., Menéndez, M. and Losada, I. «Global extreme wave height variability based on satellite data.» *Geophysical Research Letters*, 2011: 38, L10607.
- Izaguirre, C., Menéndez, M., Camus, P., Méndez, F. J., Minguez, R. and Losada, I. J. «Exploring the interannual variability of extreme wave climate in the Northeast Atlantic Ocean.» *Ocean Moelling*, 2012: 59-60, 31-40.
- Jarque, C. M. and Bera, A. K. «A test for normality observations and regression residuals.» (International Statistical Review) 55, n° 2, 890-892 (1987).
- Jarque, C. M. and Bera, A. K. «Efficient tests for normality, homocedasticity and serial independence of regression residuals.» (Economic Letters) 6, n° 3, 255-259 (1980).
- Jarque, C. M. and Bera, A. K. «Efficient tests for normality, homocedasticity and serial independence of regression residuals: Monte Carlo evidence.» (Economics Letters) 7, n° 4, 313-318 (1981).
- Jonathan, P. and Ewans, K. «Statistical modelling of extreme oceans environments for marine design: a review.» *Ocean Engineering*, 2013: 62, 91-109.
- Juarez, S. F. and Schucany, W. R. «Robust and efficient estimation for generalized Pareto distributions.» *Extremes*, 2004: 7 (3),237-251.
- Kim, S.W. and Suh, K.D. «Performance analysis of vertical breakwaters designed by partial safety factor method.» *Journal of Coastal Research*, 2013: 65, 295-301.
- Kotz, S. and Nadarajah, S. *Extreme Value Distribution: Theory and Applications*. London: Imperial College Press, 2000.

- Koyluoglu, H.U. and Nielsen, R. K. «New approximations for SORM integrals.» *Structural Safety*, 1994: 13, 235-246.
- Leadbetter G., Lindgren, G. and Rootzen, H. *Extremes and related properties of random sequences and processes*. New York: Springer, 1983.
- Leadbetter, M., Lindgren, G. and Rotzen, H. «Extremes and related properties of random sequences and processes.» *Springer-Verlag. New York*, 1983.
- Ledford, A. and Tawn, J. «Modelling dependence within joint tail regions.» *Journal of the Royal Statistical Society*, 1997: 59, 2, 475-499.
- Longuet Higgins, M. S. «On the statistical distribution of the heights of sea waves.» *Journal of Marine Research*, 1952: 11, 245-266.
- Luceño, A. «Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators.» *Computational Statistical Data Analysis*, 2006: 51, 904-917.
- Mackay, E. B. L., Challenor, P. G. and Bahaj, A., S. «A comparison of estimators for the generalised Pareto distribution.» *Ocean Engineering*, 2011: 38, 1338-1346.
- Massey, F. J. «The Kolmogorov-Smirnov test for goodness of fit.» (*Journal of the American Statistical Association*) 46, n° 68-78 (1951).
- Melchers, R. E. *Structural reliability analysis and prediction*. New York: Wiley, 1999.
- Méndez, F., Menéndez, M., Luceño, A. and Losada, I.J. «Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model.» *Journal of Geophysical Research*, 2006.
- Menéndez, M., Méndez, F. J., Izaguirre, C., Luceño, A. and Losada, I. J. «The influence of seasonality on estimating return values of significant wave height.» *Coastal Engineering*, 2009: 56, 211-219.
- Metropoli, N. and Ulam, S. «The monte carlo method.» *Journal of the American Statistical Association*, 1949: 44, 335-341.
- Mínguez, R., Tomas, A., Méndez, F. J. and Medina, R. «Mixed extreme wave climate model for reanalysis databases.» *Stochastic Environmental Research and Risk Assessment*, 2013: 27, 4, 757-768.
- Moharran, S. H., Gosain, et P. N. A. K. and Kapoor. «A comparative study for the estimators of the generalized Pareto distribution.» *Journal of Hydrology*, 1993: 150, 169-185.
- Monbet, V., Ailliot, P. and Prevosto, M. «Survey of stochastic models for wind and sea state time series.» *Probabilistic Engineering Mechanics*, 2007: 22, 113-126.
- Nelder, J. A. «A simple algorithm for function minimization.» *Computer Journal*, 1965: 7, 308-313.
- Nelsen, R. B. *An introduction to copulas*. New York: Lecture Notes in Statistics, 139, Springer, 1999.
- Padoan, S. A., Ribatet, M. and Sisson, A. «Likelihood-based inference for Max-Stable processes.» *Journal of the American Statistical Association*, 2010: DOI: 10.1198/jasa.2009.tm08577.
- Pickands, J. «Statistical inference using extreme order statistics.» *Annals of Statistics*, 1975: 3, 119-131.
- Rackwitz, R. and Fiessler, B. «Structural reliability under combined load sequences.» *Computational Structure*, 1978: 9, 489-494.
- Salvadori, G., De Michele, C. and Durante, F. «On the return period and design in a multivariate framework.» *Hydrological Earth System Science*, 2011.
- Schlather, M. «Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution.» *Statistics and Probability Letters*, 2001: 53, 325-329.
- Schlather, M. «Models for Stationary Max-Stable Random Fields.» *Extremes*, 2002: 5, 33-44.

- Scotto, M.G. and Guedes Soares, C. «Bayesian inference for long-term prediction of significant wave height.» *Coastal Engineering*, 2007: 5, 393-400.
- Sklar, A. «Random variables, distribution functions and Copulas - A personal look backward and forward.» *Distributions with Fixed Marginals and Related Topics. IMS Lectures Notes*, 1996: 28.
- Smith, R. and Weissman, I. «Estimating the Extremal Index.» *Journal of the Royal Statistical Society*, 1994: 56 (3),515-528.
- Smith, R. L. *Statistics of extremes, with applications in environment, insurance and finance*. London: Chapman & Hall, 2004.
- Stephenson, A. and Gilleland, E. «Software for the analysis of extreme events: The current state and future directions.» *Extremes*, 2006: 8,87-109.
- Wahl, T., Bender, J. and Jensen, J. «Copula functions as a useful tool for coastal engineers.» *Proceedings of the EVAN 2013*, 2013.
- Wahl, T., Mudersbach, C. and Jensen, J. «Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas: a multivariate statistical approach based on Copulas.» *Natural Hazards and Earth System Sciences*, 2012: 12, 495-510.
- Winterstein, S. R., Ude, T., C., Cornell, C. A. Bjerager, P. and Haver, S. «Environmental parameters for extreme response inverse FORM with omission factors.» *Proceedings of the ICOSSAR*, 1993.
- Wyncoll, D. and Gouldby, B. «Integrating a multivariate extreme value method within a flood risk analysis model.» *Journal of Flood Risk Management*, 2013: DOI: 10.1111/jfr3.12069 .
- Zhang, J. and Stephens, M. A. «A new and efficient estimation method for the generalized Pareto distribuion.» *Technometrics*, 2009: 51 (3), 316-325.
- Zhang, J. «Likelihood Moment Estimation for the Generalized Pareto Distribution.» *Australian and New Zealand Journal of Statistics*, 2007: 49, 69-77.
- Zheng, F., Westra, S., Leonard, M. and Sisson, S. «Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk.» *Water Resources Research*, 2013.

ANNEX 1. CASE STUDY 1

This annex refers to the first of the two proposed cases within the project.

In this first case of study an analysis of a variable related to the design and estimation of maritime engineering is studied. In order to do so, a theoretical case of tension is proposed. This load variable can be assumed to be a simple approach of the mooring lines tension in a floating offshore structure. Both waves and winds have been considered in this example.

A. Tension Model Determination

This first case study proposed is related to the design and calculation of any coastal or offshore structure where both waves and winds are met-ocean variables to be considered.

The data used to this case study are those provided by the reanalysis database developed by IH Cantabria, which cover a time span of 67 years of data: from 1948 until 2014 with an hourly temporal resolution. The location selected for this example is a point located in front of the French Atlantic coast in the Biscay Bay near Bordeaux: 2.5 W, 45 N.

Considering a theoretical case of a floating offshore structure, the tension of its mooring lines can be estimated as a function of the waves and the wind to which the structure will be affected during its life time.

$$T = f(H_s, W)$$

For the sake of simplicity a simple relation of square significant wave height (H_s) and winds magnitude (W) is going to be considered.

$$T = f(H_s, W) = \alpha_1 H_s^2 + \alpha_2 W^2$$

Where α_1 and α_2 are two coefficients that may quantify the influence of each variable (H_s and W) into the tension estimation.

In order to estimate high return values of T (let say 100 years return value) as a function of H_s and W , there are different approaches depending on the dependence between the variables and also the complexity of its joint distribution.

Estimation of the coefficients α_1 and α_2

The coefficients α_1 and α_2 are going to be set in such a way that guarantee the same influence of both variables in the equation for a certain return period (let say 100 year return period).

$$\alpha_1 (H_s^{100})^2 = \alpha_2 (W^{100})^2$$

There is needed another equation to be able to set the values of the coefficients. Because, the theoretical nature of this example a simple equation that allows this estimation has been considered. In case of a real design process, there may exist structural or geometrical impositions that should be considered.

$$1 = \alpha_1^2 + \alpha_2^2$$

The values of H_s^{100} and W^{100} are estimated by fitting an extremes distributions to the empirical time series. For both variables two approaches were conducted: on one hand annual maxima from the time series were extracted and they were fitted to a Generalized Extreme Value distribution and to a Gumbel distribution; and on the other hand a Peaks Over Threshold analysis was performed to the time series and the sample of peaks was fitted to a Generalized Pareto distribution and to an Exponential distribution. The POT method was undertaken considering an independence lapse of 3 days between events and the threshold was set to guarantee a mean of 1 event per year. The method used to obtain the parameters of the distributions was Maximum Likelihood for all the distributions. Next figures (Figure 39 to Figure 42) show the different fittings for both variables and in the table below there are summarized the values of 100 years return period.

Distribution	Location (μ)	Scale (σ)	Shape (ζ)
GEV	8.66	1.095	-0.183
Gumbel	9.721	1.233	-----
Distribution	Threshold (u)	Scale (σ)	Shape (ζ)
GPD	8.67	1.199	-0.236
Exponential	8.67	0.980	-----

Table 16. Parameters fitted for H_s

	GEV	Gumbel	GPD	Exponential
H_s^{100} (m)	12.07	11.6	12.04	13.19

Table 17. 100 years return period estimations for H_s .

Note that in all the graphical representations included within this annex the Weibull plotting method for the empirical distribution, within those methods explained in 2.4.5, has been used.

Distribution	Location (μ)	Scale (σ)	Shape (ξ)
GEV	20.542	1.546	-0.065
Gumbel	22.318	1.966	-----
Distribution	Threshold (u)	Scale (σ)	Shape (ξ)
GPD	20.38	2.238	-0.316
Exponential	20.38	1.684	-----

Table 18. Parameters fitted for W

	GEV	Gumbel	GPD	Exponential
W^{100} (m/s)	26.68	25.32	25.81	28.13

Table 19. 100 years return period estimations for W

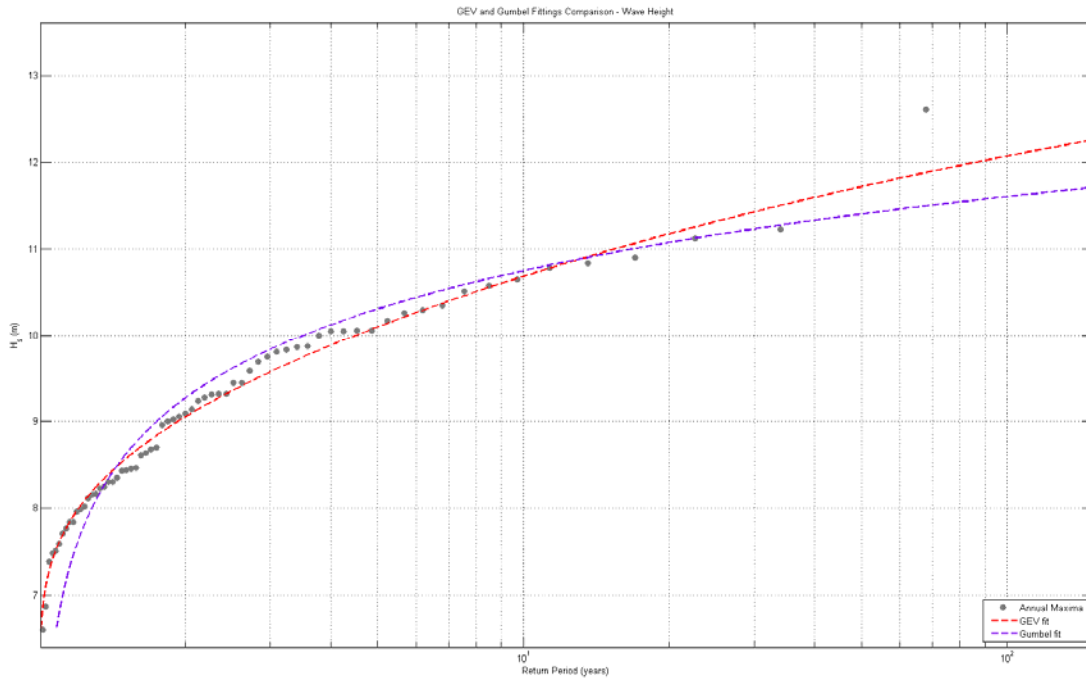


Figure 39. GEV and Gumbel fittings for H_s annual maxima.

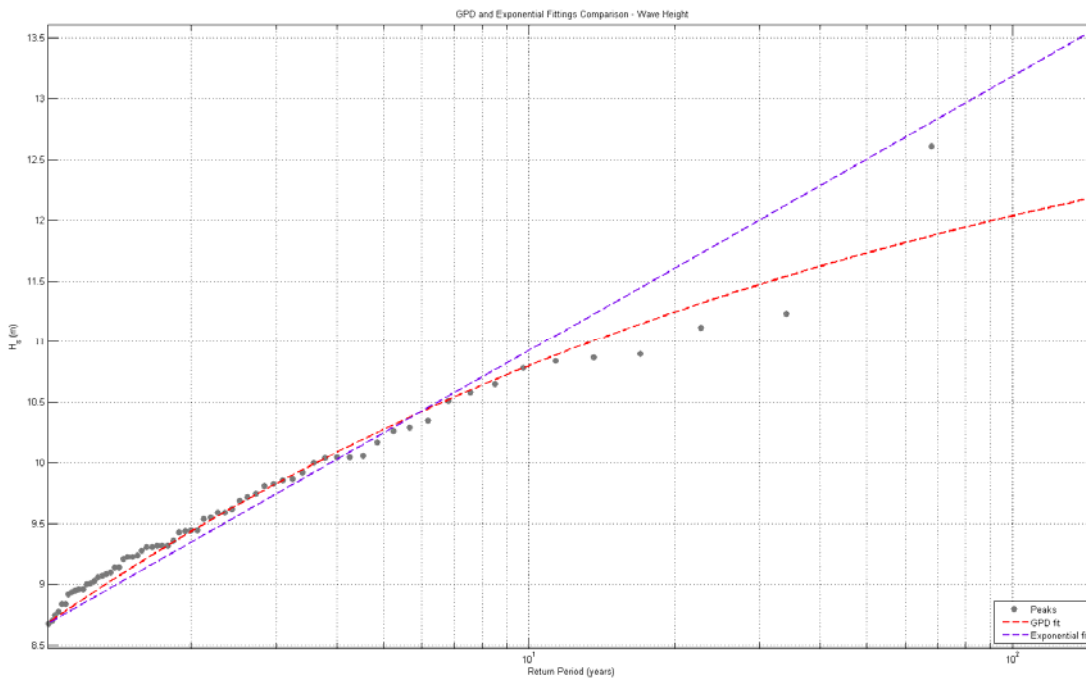


Figure 40. GPD and Exponential fittings for H_s peaks over threshold.

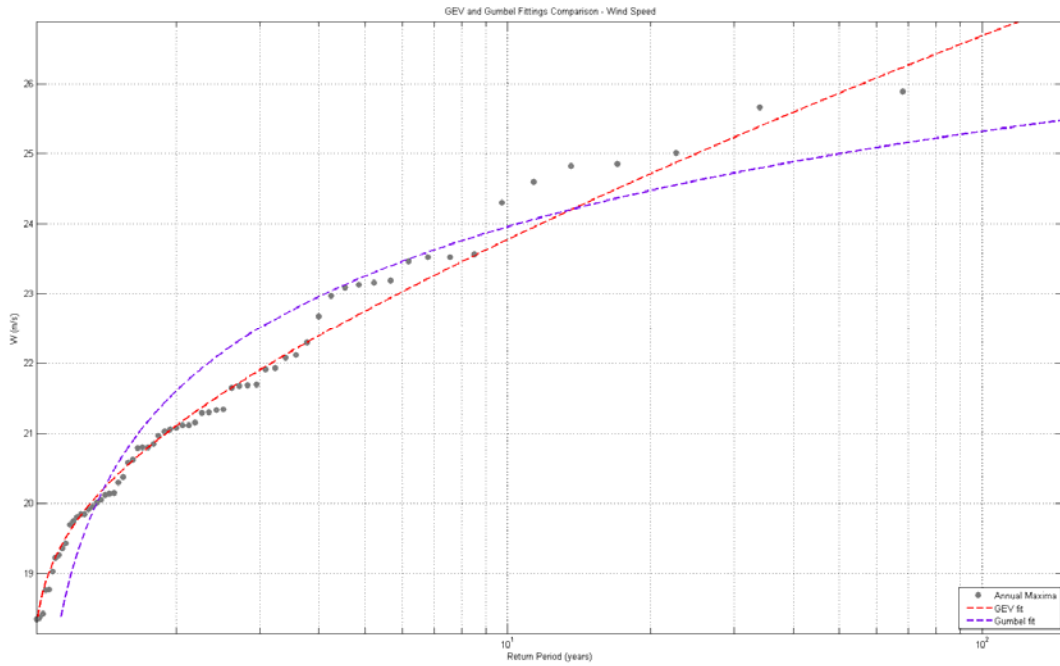


Figure 41. GEV and Gumbel fittings for W annual maxima.

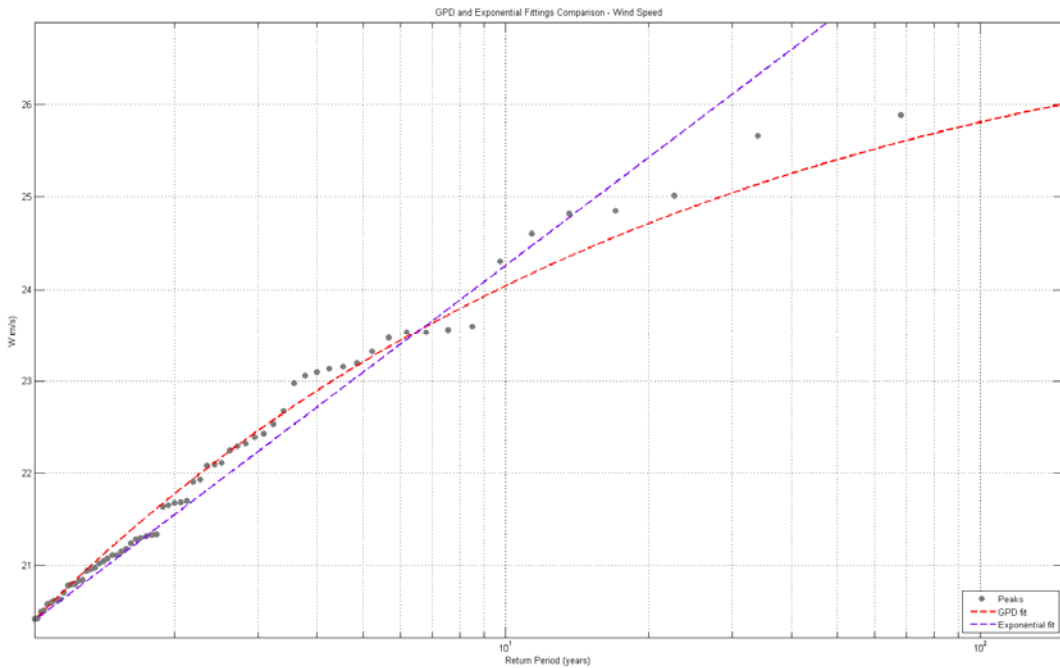


Figure 42. GPD and Exponential fittings for W peaks over threshold.

It can be stated from the figures and the return period values that the variable H_s presents behavior more upper bounded than the wind speed. Furthermore, for both variables, the fittings made based on the annual maxima provide more conservative estimations than those made based on the peaks over threshold. According to the values obtained with the different adjustments, a design value of 12 meters is assumed for H_s^{100} and a wind velocity of 26.25 m/s for W^{100} . Having this, the values of α_1 and α_2 can be estimated.

$$\alpha_1 (H_s^{100})^2 = \alpha_2 (W^{100})^2 \rightarrow 12^2 \alpha_1 = 26.25^2 \alpha_2 \rightarrow \alpha_1 = 4.785 \alpha_2$$

$$1 = \alpha_1^2 + \alpha_2^2 = \alpha_2^2 (1 + 4.785^2)$$

$$\alpha_2 = 0.205 \quad \alpha_1 = 0.979$$

B. T¹⁰⁰ Estimation

Once the model of T has been established, the 100 years return period value of T can be estimated. In order to do so, different approaches are going to be considered

a. Direct Estimation of T¹⁰⁰

The estimation of the coefficients α_1 and α_2 allows us to have an empirical estimation of the mooring lines tension distribution as a function of H_s and W :

$$T = 0.979H_s^2 + 0.205W^2$$

Then the estimation of the 100 return period estimate of T can be directly be done in the same way it was done with the wave and wind variables by fitting them a distribution.

Below, in Figure 43 is shown the adjustment of the T annual maxima to a GEV and a Gumbel distributions. While in Figure 44 are depicted the fittings to a GPD and an Exponential distributions made based on the Peaks Over Threshold. The samples of Peaks is obtained with a threshold that assures a mean of 1 event per year ($\lambda=1$) and with a minimum lap between peaks of 72 hours. The method used to obtain the parameters of the distributions was Maximum Likelihood for all the distributions. As seen in the table summarizing the 100 return period estimations of T, the distributions fitted considering the shape parameter (GEV and GPD) provide estimations around 265, while when fixing the shape parameter equal to zero the Gumbel distribution predicts a lower value: 243 and the exponential a much higher one: 275.

Distribution	Location (μ)	Scale (σ)	Shape (ζ)
GEV	148.3	28.34	-0.045
Gumbel	181.74	40.108	-----
Distribution	Threshold (u)	Scale (σ)	Shape (ζ)
GPD	149	29.389	-0.0751
Exponential	149	27.325	-----

Table 20. Parameters fitted for T

	GEV	Gumbel	GPD	Exponential
T ¹⁰⁰	266.2	243	264.6	275

Table 21. 100 years return period estimations for T

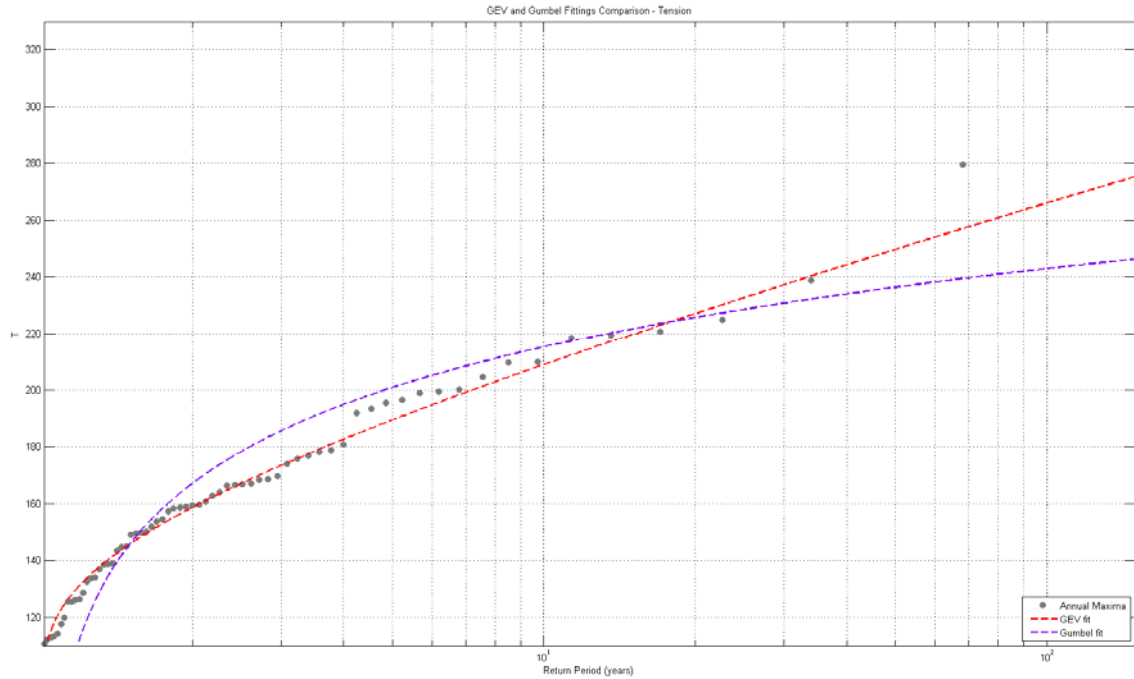


Figure 43. GEV and Gumbel fittings for T annual maxima

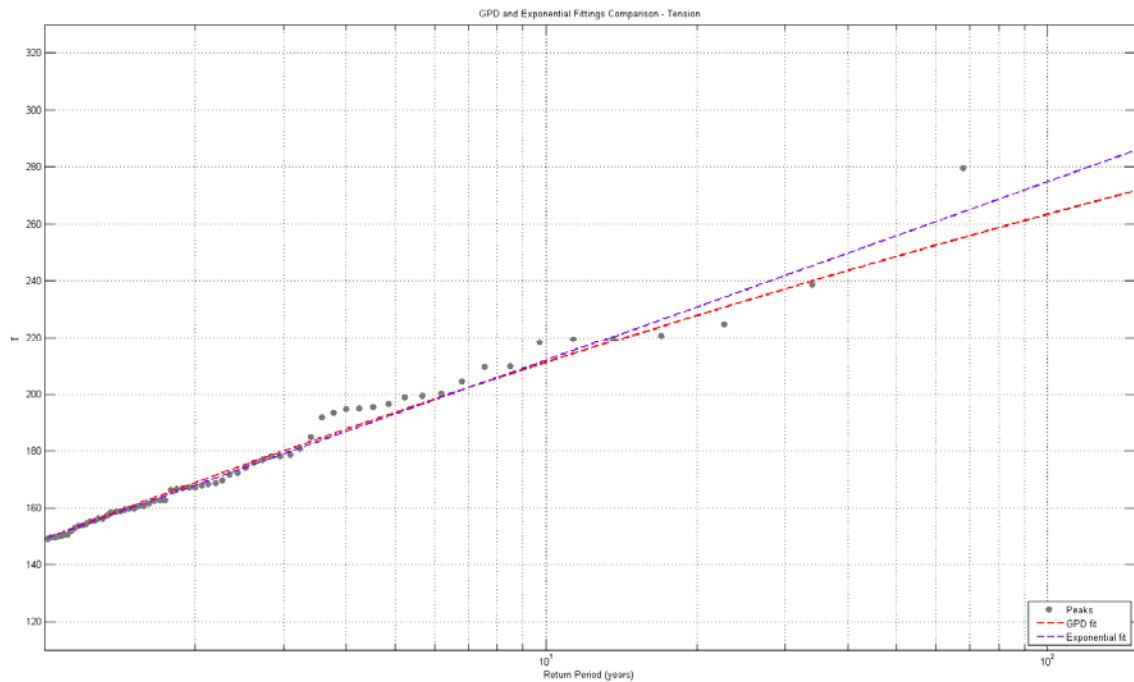


Figure 44. GPD and Exponential fittings for W peaks over threshold

b. Estimation of T^{100} by estimating H_s^{100} and W^{100} independently

An initial estimation of the T^{100} value can be done by directly estimating it by using the values of H_s^{100} and W^{100} . This is an estimation that assumes complete dependence between the variables and also, assumes both 100-year return values (for H_s and W) occur at the same time. Thus, generally it is a very conservative method.

$$T^{100} = 0.979(H_s^{100})^2 + 0.205(W^{100})^2$$

$$T^{100} = 0.979(12)^2 + 0.205(26.25)^2 = 282.23$$

$$T^{100} = 282.23$$

c. Estimation of T^{100} by estimating W^{100} conditioned to H_s^{100}

The model of dependence of this method is complete dependence as in the previous method shown. The difference relays in not considering a simultaneous occurrence of H_s^{100} and W^{100} . In order to do so, a heteroscedastic model was fitted to the joint distribution of H_s and W where the mean expected value was fitted to a quadratic polynomial and the standard deviation to a linear function. A heteroscedastic model is that in which the standard deviation around the mean is not considered constant, it can be assumed to follow a linear or a quadratic function.

$$\mu_w(H_s) = p_1 + p_2 H_s + p_3 H_s^2$$

$$\sigma_w(H_s) = p_4 + p_5 H_s$$

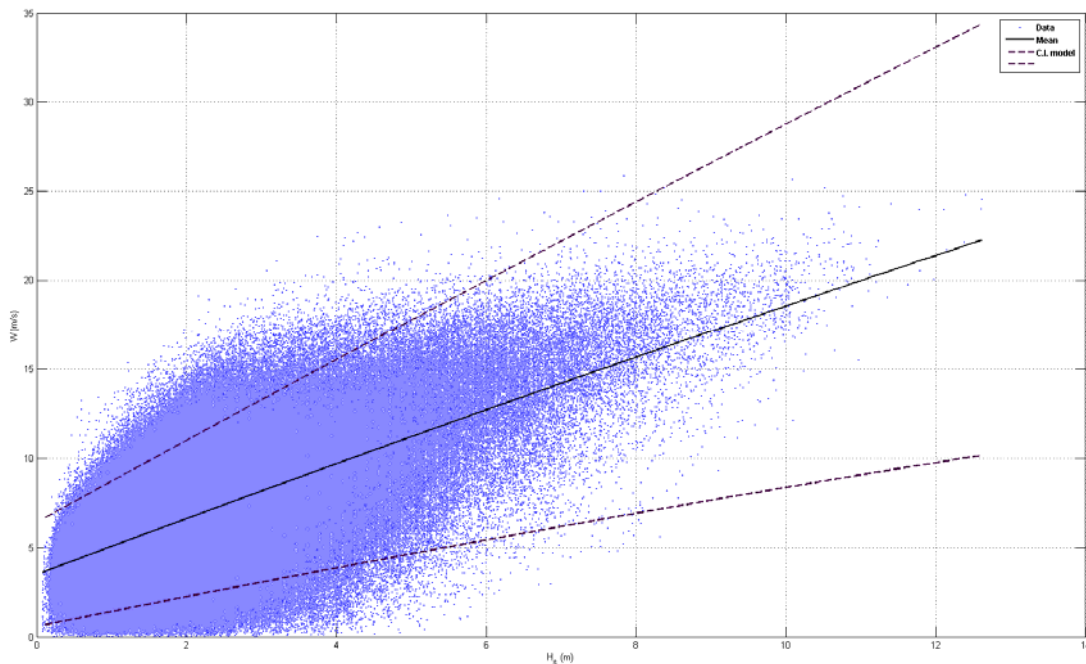


Figure 45. Heteroscedastic model fit to W conditioned to H_s .

p_1	p_2	p_3	p_4	p_5
3.506	1.576	-0.007	1.773	0.443

Table 22. Heteroscedastic parameters estimated

$$\mu_w(H_s) = 3.506 + 1.576H_s - 0.007H_s^2$$

$$\sigma_w(H_s) = 1.773 + 0.443H_s$$

By introducing the value of H_s^{100} of 12 meters obtained before, the mean value of W associated with this value of H_s^{100} could be estimated and then the expected value of T^{100} since complete dependence is assumed.

$$\mu_{W^{100}} = 3.506 + 1.576 \cdot 12 - 0.007 \cdot 12^2 = 21.71$$

$$\sigma_{W^{100}} = 1.773 + 0.443 \cdot 12 = 7.089$$

$$T^{100} = 0.979(12)^2 + 0.205(21.41)^2 = 237.59$$

$$\mathbf{T^{100} = 237.59}$$

d. Estimation of T^{100} by estimating H_s^{100} conditioned to W^{100}

In a parallel way that done before, the value of H_s^{100} can be estimated conditioned to the estimated value of W^{100} . To do so, the following model was fitted.

$$\mu_{H_s}(W) = p_1 + p_2W + p_3W^2$$

$$\sigma_{H_s}(W) = p_4 + p_5W$$

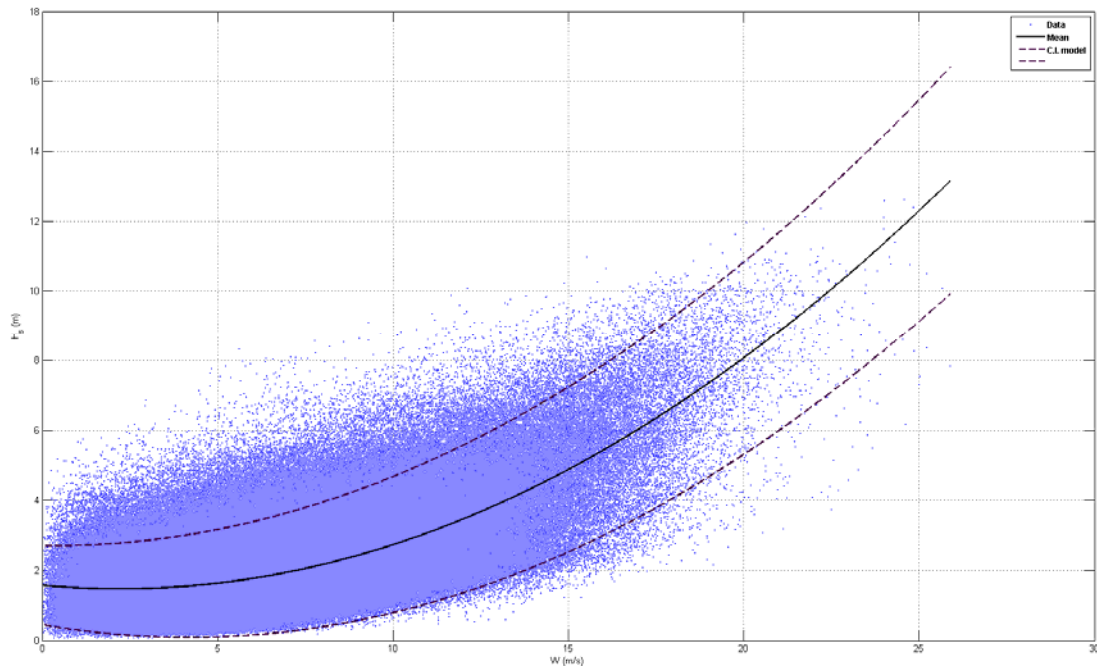


Figure 46. Heterocedastic model fit to H_s conditioned to W .

p_1	p_2	p_3	p_4	p_5
1.574	-0.092	0.021	0.685	0.05

Table 23. Heterocedastic parameters estimated

$$\mu_{H_s}(W) = 1.574 - 0.092W + 0.021W^2$$

$$\sigma_{H_s}(W) = 0.685 + 0.05W$$

By introducing the value of W^{100} of 26.25 m/s obtained before, the mean value of H_s^{100} could be estimated and then the expected value of T^{100} .

$$\mu_{H_s^{100}} = 1.574 - 0.092 \cdot 26.25 + 0.021 \cdot 26.25^2 = 13.62$$

$$\sigma_{H_s^{100}} = 0.685 + 0.05 \cdot 26.25 = 1.998$$

$$T^{100} = 0.979(13.62)^2 + 0.205(26.25)^2 = 322.87$$

$$\mathbf{T^{100}=322.87}$$

e. Heffernan and Tawn 2004

The method proposed by (Heffernan 2004) to sample variables interrelated can be summarized in the following steps. This method is useful when the extreme tail behavior of the dependence of the variables is unknown, and allows an artificial enlargement of dataset. In order to illustrate the method, an example of simulating significant wave height and wind speed will be carried out:

1. Extraction of declustered maxima.

From the time series independent maxima should be extracted to serve as initial data for the method. The way of selecting the maxima may vary from one application to another, in this case a Peak Over Threshold analysis to the variable T has been done. Being T a function of H_s and W ($T = 0.979H_s^2 + 0.205W^2$) and setting the threshold at the 80% percentile and a lag between events of 3 days (Figure 47).

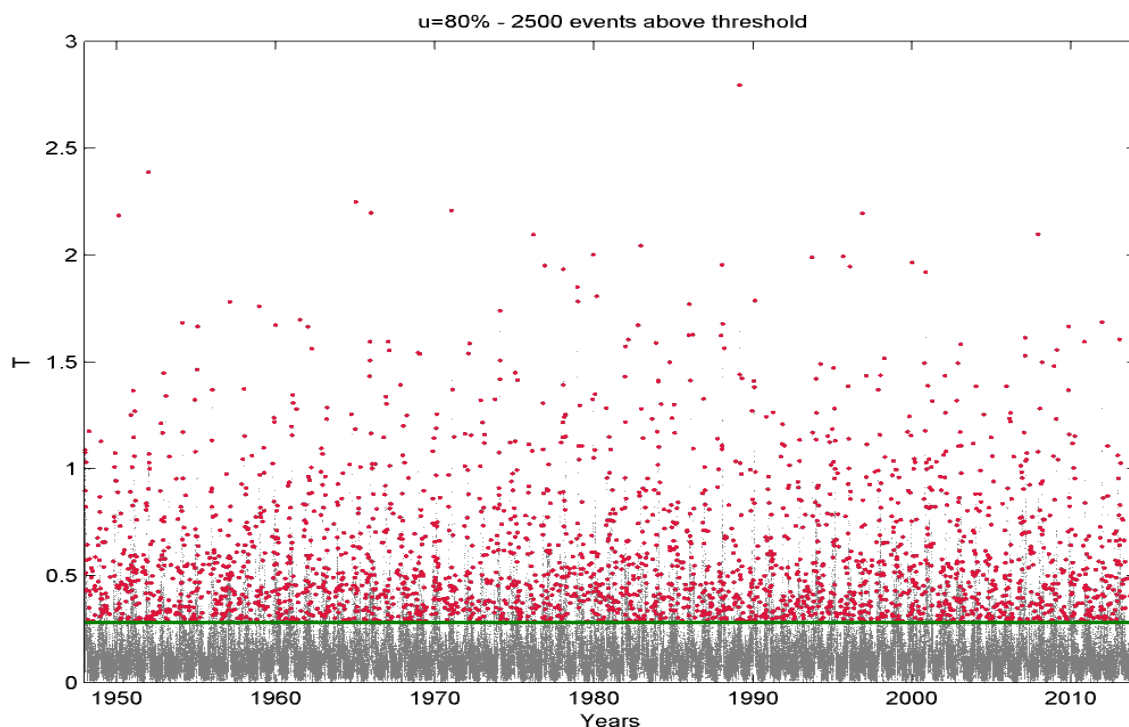


Figure 47. Peaks Over Threshold of T

With this, a set of 2500 maxima were obtained. These maxima are depicted in red in Figure 48.

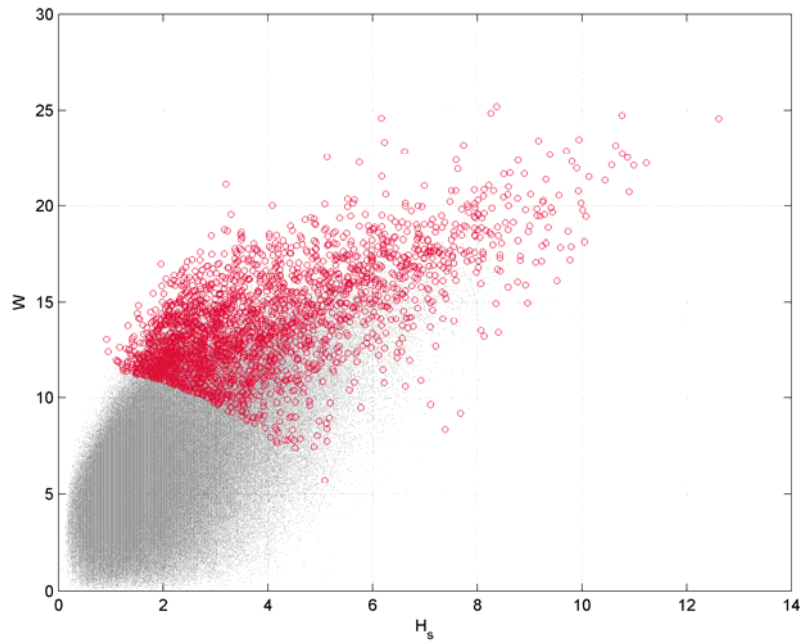


Figure 48. Declustered maxima.

2. Pre-treatment of the marginal distributions

2.1. Setting thresholds: Two thresholds must be set to perform the method. The first one is the threshold above which the variables dependence is going to be estimated: dependence threshold. The second one is the threshold above which a GPD is fitted to each variable and is assumed within the sampling process (GPD threshold). In accordance with (Gouldby, A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis 2014):

- Dependences threshold: 85% percentile.
- GPD threshold: 95% percentile.

To set the dependences and GPD thresholds, a sensibility analysis should be done. In this case, since the variables and databases used are the same as those used by (Gouldby, A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis 2014), this selection criterion was not necessary and their same thresholds were used.

2.2. GPD fit to the upper 5% tail to each variable.

2.3. Re-escalation of the variables: Up to the 95% percentile threshold the empirical distribution is taken; above, the empirical distribution is re-escalated using the values of the GPD fit as proposed by (S. a. Coles 1991)

$$\hat{F}_{x_i}(x) = \begin{cases} 1 - \{1 - \tilde{F}_{x_i}(u_{x_i})\} \{1 + \xi_i(x - u_{x_i}) / \beta_i\}_+^{-1/\xi_i} & \text{for } x > u_{x_i} \\ \tilde{F}_{x_i}(u_{x_i}) & \text{for } x \leq u_{x_i} \end{cases}$$

Where ξ_i and β_i are the GPD parameters and u_i is the threshold.

2.4. Transformation of the marginal distributions to a standard Gumbel scale.

$$Y_i = -\log\left[-\log\left\{\widehat{F}_{X_i}(X_i)\right\}\right] = t_i(X_i; \varphi_i; \tilde{F}_{X_i}) = t_i(X_i) \quad \text{for } i=1, \dots, d$$

Where $\varphi_i = (\beta_i, \xi_i)$ are the marginal parameters.

3. Estimation of the dependence coefficients (a and b)

A multivariate non-linear regression is fitted and the dependence coefficients a and b are estimated.

$$Y_{-i} | Y_i = aY_i + Y_i^b W \quad \text{for } Y_i > v$$

Where a and b are vectors of parameter to each pair-wise of variables, v is the specified threshold and W is a vector of the residuals. The model is fitted using maximum likelihood assuming the residuals follow a normal distribution.

4. Simulation process

4.1. Once the marginals are all in a Gumbel scale they can be directly compared to obtain the proportion of events where each variable Y_i is a maximum. (*) In our case of study the proportion of events with the variable H_s as the maximum is 51.4% while the percentage of events where the wind speed is maximum is 48.6%. These proportions will be used within the Monte Carlo simulation.

4.2. Sampling by Monte Carlo simulation considering the a and b coefficients related to each variable, and introducing a residual from the empirical data residuals (assuming to be uniformly distributed).

4.3. During the Monte Carlo simulation, rejection sampling is used to ensure the initial proportions (*) of each variable being maxima. The Monte Carlo simulation is done in two steps: in a first step datasets (pairs of wave and wind data) considering the wave height as the conditioning variable are sampled and then in a second step are sampled the datasets considering the wind speed as the conditioning variable. All the datasets sampled in each step have to ensure that the variable considered as conditioning variable is maxima, thus the proportions estimated at 4.1 are guaranteed; so any sampled dataset that does not satisfy this condition is rejected and sampled again.

4.4. Transformation back from the Gumbel to the original scale of the sampled variables, taking into account the empirical distribution below the GPD threshold and the GPD fits for the upper tail (above the GPD threshold).

In Figure 49 the marginal distributions of H_s and W obtained through a 1.000 years simulation are shown; while in Figure 50 the joint distribution of both variables is represented. In grey dots there are represented the original 60-years time series of H_s and W , in orange are marked the maxima declustered from the historical record and in purple are depicted the 1.000 years simulated time series.

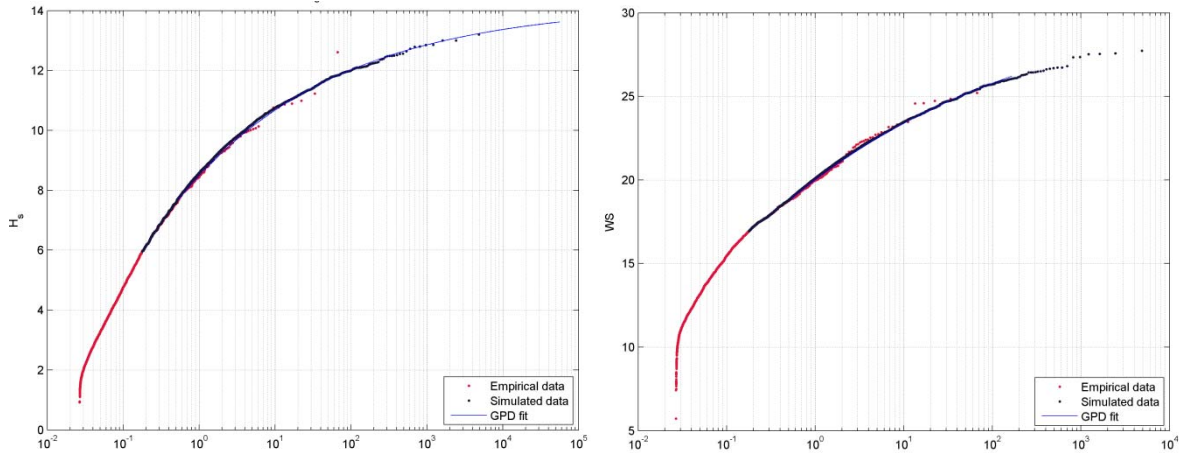


Figure 49 H_s and W samples

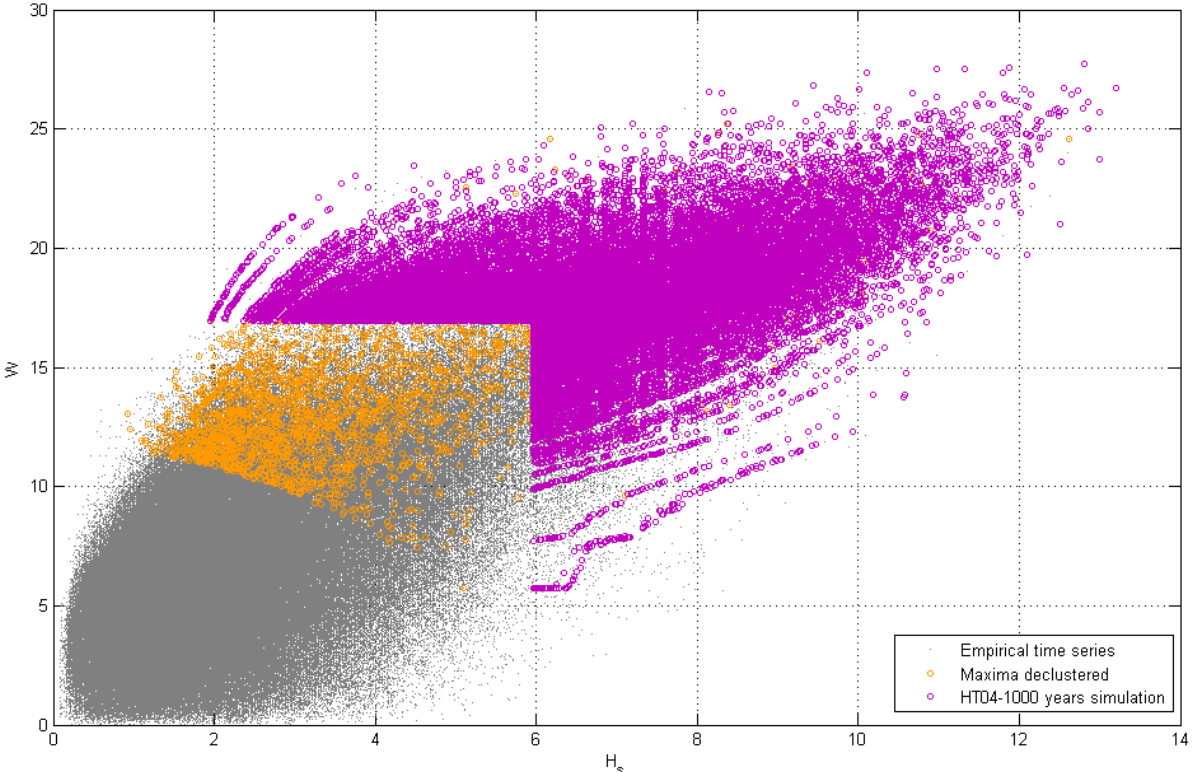


Figure 50. H_s - W joint distribution comparison between empirical and sampled data.

5. Estimation of T^{100} from the simulated data.

Once H_s and W have been sampled, the tension T can be calculated. Then, the value of T^{100} can be assumed to be the tenth biggest value of the sample because 1000 years have been simulated.

Considering the sampled events of the variables related to T (H_s and W), a certain return period value of T could be estimated.

A value of $T^{100} = 266.57$ has been obtained

Based on the theory of conditioned probabilities T^{100} can be estimated in the following way:

$$Pr(T > T^{100}) = Pr(T > T^{100} | T > T^u) Pr(T > T^u)$$

Assuming T^u as the tension associated to the dependence thresholds ($H_s^u = 5.96$ m and $W^u = 16.94$ m/s), it is $T^u = 93.67$.

$Pr(T > T^t | T > T^u)$ is the probability of T being above T^t conditioned to T being above T^u and $Pr(T > T^u)$ is the probability of T being above T^u .

$Pr(T > T^u)$ could be empirically estimated by calculating the ratio between the number of sampled events with T above T^u and the entire synthetic time series of T (purple dots). Then, $Pr(T > T^u) = 0.6555$.

Independently, $Pr(T > T^t)$ could be theoretically estimated for a given return period, t . $Pr(T > T^t) = 1/(t \cdot n)$; with n being the number of events per year: in this case it has been sampled 7.67 events/year (number of events above the dependence threshold from the 2500 declustered events from the historical data divided by the number of years) and t the return period, $t = 100$. With these two values ($t = 100$ and $n = 7.67$), we obtain a value of $Pr(T > T^{100}) = 0.0013$

With that the value of T^{100} that assures $Pr(T > T^t | T > T^u) Pr(T > T^u) = Pr(T > T^{100}) / Pr(T > T^u)$ can be estimated. In this case, the value of T^{100} has been estimated to be **266.57**.

f. Mínguez et al 2013

In 2013 (R. G. Mínguez 2013) performed a methodology to determinate the joint probability of occurrence of interrelated variables in such a way that allow to avoid the decision of working with the entire time series (point-in-time distribution) or just with the extremes (extreme value distribution).

This method consists basically on assuming one variable to be conditioned to the other one. In our case we have assumed the significant wave height conditioned to the wind speed. The method can be summarized in the following steps:

1. Fitting the marginal distributions of W

The entire time series of wind speed is fitted to a GEV distribution, the election of this distribution corresponds to the nature itself of wind data but for other environmental variable another distribution may be chosen. To appropriately reproduce the extremes, a GPD is fitted to the events above a certain threshold; in this case the threshold has been set at 17.1 m/s. The values obtained for the two distributions fitted are showed below:

GEV	μ	σ	ζ
	5.37	2.93	-0.063
GPD	u	σ	ζ
	17.1	1.54	-0.084

Table 24. Distribution fits

2. Fitting the conditional distributions of H_s

The conditional distribution of significant wave height for given values of wind speed is fitted to a GEV. In order to do so, the values of H_s are divided into groups of similar value of wind speed, i.e.:

$$\begin{aligned}
 &H_s|0 < W < 1 \\
 &H_s|1 < W < 2 \\
 &\dots \\
 &H_s|19 < W < 20 \\
 &H_s|20 < W < 22 \\
 &H_s|W < 22
 \end{aligned}$$

As done with the wind speed, a GPD is fitted to all the groups in order to correctly characterize the upper tail of each distribution. The threshold has been set to be the 95% percentile.

3. Fitting the parameters to a 3rd order polynomial.

The parameters of the distributions fitted before are fitted to a third order polynomial. Then for the GEV parameters would be:

$$\begin{aligned}
 \mu(W) &= p_1^\mu W^3 + p_2^\mu W^2 + p_3^\mu W^1 + p_4^\mu \\
 \sigma(W) &= p_1^\sigma W^3 + p_2^\sigma W^2 + p_3^\sigma W^1 + p_4^\sigma \\
 \zeta(W) &= p_1^\zeta W^3 + p_2^\zeta W^2 + p_3^\zeta W^1 + p_4^\zeta
 \end{aligned}$$

In Figure 51, the polynomial fitted for the three parameters (red line) as well as the points based to the fitting (black dots) are shown.

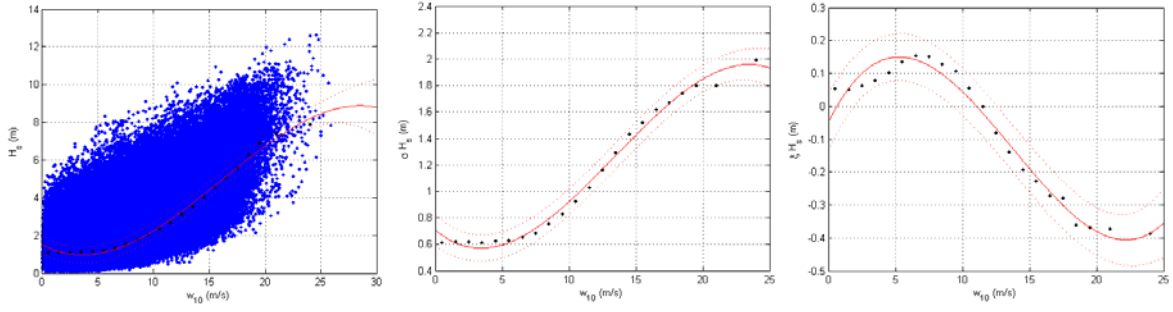


Figure 51. Third order polynomial fitted to the GEV parameters

The same process has been done to the GPD parameters:

$$\begin{aligned}
 u(W) &= p_1^u W^3 + p_2^u W^2 + p_3^u W^1 + p_4^u \\
 \sigma(W) &= p_1^\sigma W^3 + p_2^\sigma W^2 + p_3^\sigma W^1 + p_4^\sigma \\
 \xi(W) &= p_1^\xi W^3 + p_2^\xi W^2 + p_3^\xi W^1 + p_4^\xi
 \end{aligned}$$

Below, Figure 52 shows the fitted polynomial (red line) and the values used (black dots).

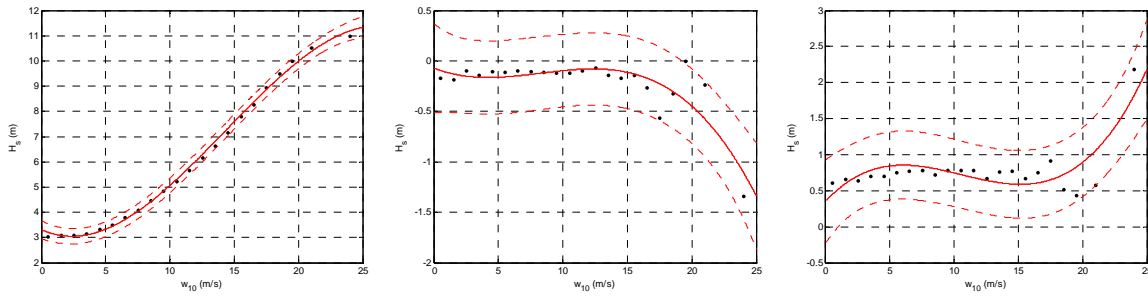


Figure 52. Third order polynomial fitted to the GPD parameters

4. Application to the use of IFORMS

Once the conditional distributions have been fitted the joint n-year environmental contours can be represented, being n the return period. This can be done by application of the Rosenblatt transformation:

$$\begin{aligned}
 \Phi(z_1) &= F_w(w) \\
 \Phi(z_2) &= F_{H_s|W}(h_s)
 \end{aligned}$$

Where $F_w(w)$ is the marginal distribution of wind speed and $F_{H_s|W}(h_s)$ is the conditional distribution of significant wave height for given values of wind speed. The contours are defined by those points that satisfy the equation

$$z_1^2 + z_2^2 = \beta^2$$

Where β is obtained from the return period: $\Phi(\beta) = 1 - \frac{1}{N}$, being N the number of independent sea states in n years. A detailed description of the IFORMS method could be found in (Hasofer 1974).

This way, Figure 53 represents the scatter plot of the 60-years of hourly data of wind speed and significant wave height with the contours for n -years, with $n=0.01, 1, 5, 50, 100$ and 500 years.

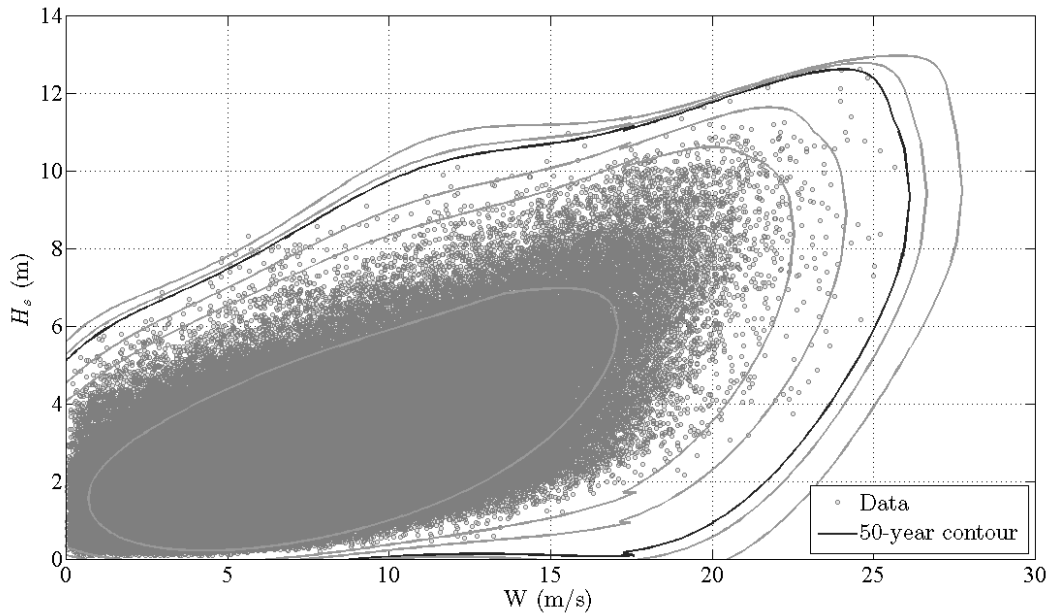


Figure 53. H_s - W contour plot.

5. Estimation of T^{100} from the IFORMS

From the contours depicted in Figure 53, the T^{100} can be estimated. Considering the points that define the 100-years return period contour T is calculated. From these values of T , selecting the maxima is obtained the T^{100} .

In our particular case, a value of $T^{100} = 289.37$ has been obtained.

g. Estimation of T^{100} assuming that H_s and W are completely independent

An estimation of the tension considering that the variables involved are completely independent has been carried out. In order to do so, to avoid the existent dependences in the original time series the variable W has been shifted half a year, creating this way a new data set where the independence between variables is assured.

With this modified dataset the 100 years return period estimate of T can be estimated in the same way as it was done in 3.1. To do so, the time series of T is obtained through the same equation as before,

$$T = 0.979H_s^2 + 0.205W^2$$

To the annual maxima extracted from T time series a GEV and a Gumbel distributions has been fitted. These distributions are shown in Figure 54. Moreover, a GPD and an Exponential distribution have been fitted to a sample of peaks extracted from the T time series. These peaks were extracted assuming a threshold that assures a mean of 1 event per year ($\lambda=1$) and with a minimum lap between peaks of 72 hours. The fitted distributions together with the peaks are shown in figure Figure 55. In the fitting of all the distributions, the method used to obtain the parameters was Maximum Likelihood. Below the figures there is a table summarizing the 100 return period estimations of T and the parameters of the distributions fitted. All the values obtained for T^{100} with the four distributions fitted using a dataset where the independence between H_s and W is assured are in the range 154-172. These values are much lower than the ones obtained with the original dataset (between 243 and 275 depending on the distribution fitted) and this is because of the existent dependence within variables of environmental nature such as waves and winds (with high winds there might be high waves). Thus, it may not be accurate to design assuming independence.

Distribution	Location (μ)	Scale (σ)	Shape (ξ)
GEV	100.27	16.15	-0.032
Gumbel	119.77	22.79	-----
Distribution	Threshold (u)	Scale (σ)	Shape (ξ)
GPD	100.7	16.13	-0.044
Exponential	100.7	15.45	-----

Table 25. Distribution fits

	GEV	Gumbel	GPD	Exponential
T^{100}	169.3	154.6	171.9	167.9

Table 26. 100 years return period for T

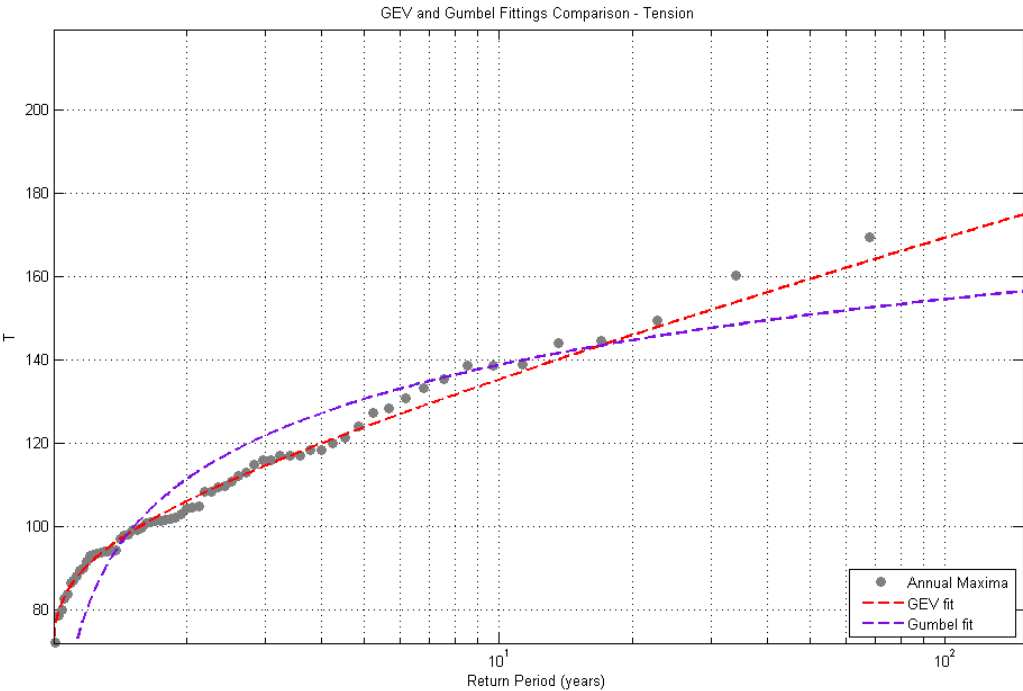


Figure 54. GEV and Gumbel fittings for T annual maxima.

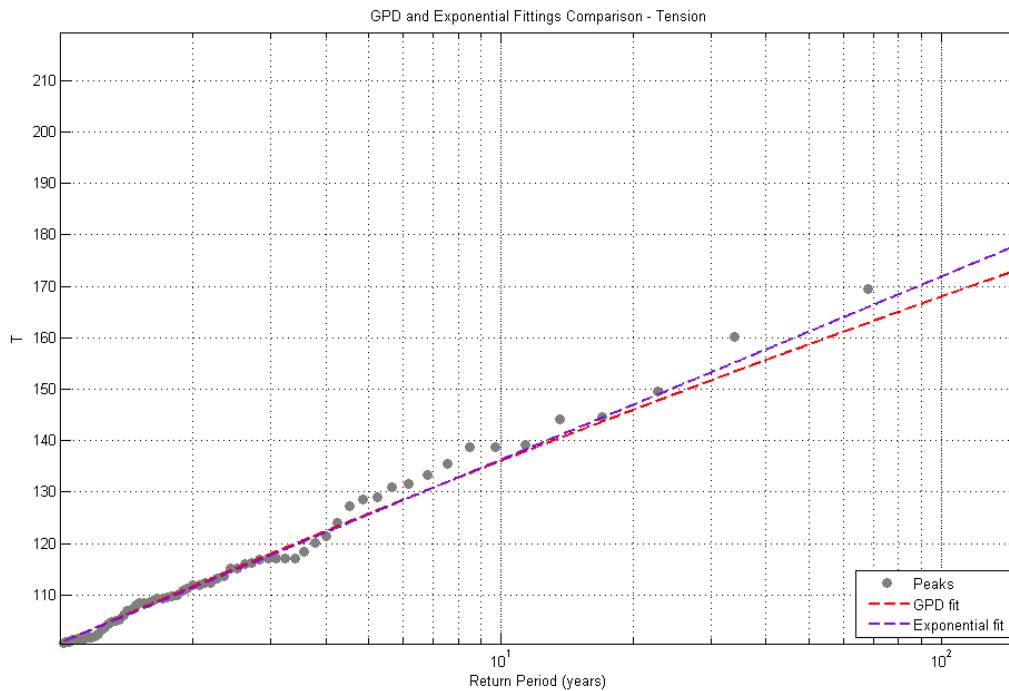


Figure 55. GPD and Exponential fittings for W peaks over threshold

C. Comparison

In the table below is summarized a brief comparison of the different estimations obtained for T^{100} .

	T^{100}
Empirical distribution of T (GEV fit)	266.2
Empirical distribution of T (Gumbel fit)	243.0
Empirical distribution of T (GPD fit)	263.6
Empirical distribution of T (Exponential fit)	275.0
With W^{100} and H_s^{100} completely dependent	282.23
With W conditioned to H_s^{100}	237.59
With H_s conditioned to W^{100}	322.87
Heffernan and Tawn 2004	266.57
Mínguez et al 2013 + IFORM	289.37
Empirical distribution of T, W shifted (GEV fit)	169.3
Empirical distribution of T, W shifted (Gumbel fit)	154.6
Empirical distribution of T, W shifted (GPD fit)	171.9
Empirical distribution of T, W shifted (Exponential fit)	167.9

Table 27. Comparison of the results obtained for T^{100}

The values estimated for T^{100} vary from 154.6 obtained with the application of a Gumbel distribution to the modified dataset that assures the independence between the variables and 322.87 as a result of considering the value of wind speed for 100 years and the value of wave height conditioned to it. Such a big range in the estimations can be related to the different nature of the methods applied.

The estimation of T^{100} based on W^{100} and $H_s^{100} | W^{100}$ gives the biggest value estimated due to the quadratic relation fitted that gives a value of $H_s^{100} | W^{100}$ of 13.6 m., while when analyzed independently $H_s^{100}=12\text{m}$.

Fitting the empirical distribution of T to a GEV with annual maxima or a GPD with a $\lambda=1$ gives similar results; around 260. But if the scale parameter is forced to be equal a zero then the Gumbel fit (done with the annual maxima) gives a lower estimation (243) and the Exponential fit (done with the same peaks extracted as for the GPD) results in a higher estimation (275).

On the other hand, considering fully dependent the variables and estimating T^{100} directly applying to the model established the values of H_s^{100} and W^{100} , gives a conservative value of T^{100} with of 282.2.

The application of more complex and elaborated methods like the one proposed by Heffernan and Tawn results in value of: 266.57. While applying the conditioned model proposed by Mínguez et al together with IFORM contours provides a value of 289.37 which is higher than those obtained by the direct analysis of T .

Finally, the values obtained with the assumption of H_s and W being completely independent (whatever the distribution fitted it is) are much lower to all the rest that consider somehow the existent dependence within waves and winds. This way, any design made under this assumption may be not be accurate or safe enough. This enhances the importance of a correct characterization of these dependences within variables.

ANNEX 2. CASE STUDY 2

This annex refers to the second case study proposed within the MASTENV project.

In this second case of study, an analysis of the flooding risk in a certain area of the French coast is studied. This application involves the selection of multivariate data sets to be introduced in a wave propagation model (SWAN) and a latter flooding model (MARS).

A. Case Study

The main objective of this second case study is the determination of the 100 years return period of the flooding level in a certain region. In order to do this, the extreme events (wave height, sea water level and river level) needed to be introduced in the flooding model as input have to be determined.

a. Location and data available

The study proposed is located in the southern French coast, in the region of Hyères. At this location the combination of the storm surge and the set up related to the incident waves together with high levels of the river may provoke flooding events. To study this problem, flooding event simulations can be conducted with the MARS model. To be able to run the model, the input variables should be accurately estimated. In Figure 56, the location of the study as well as the situation of the available data is represented.

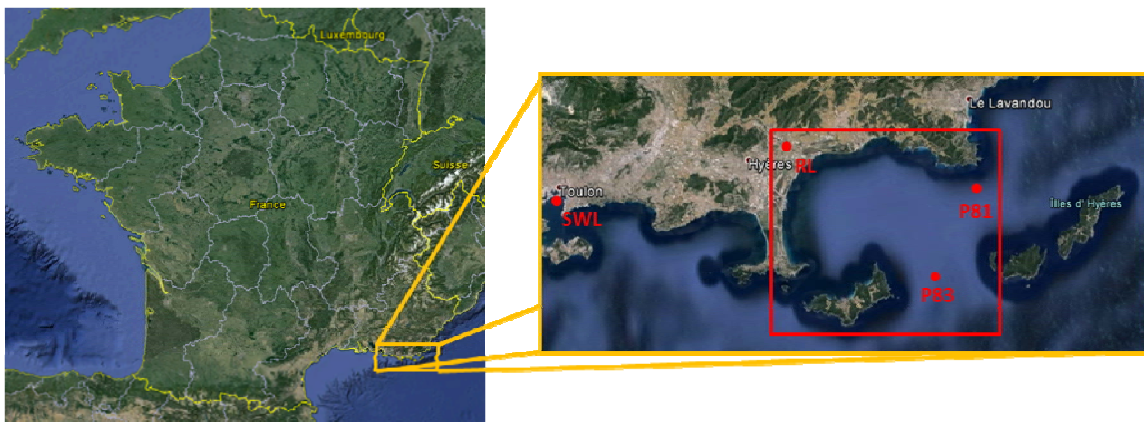


Figure 56. Study location and data points of the different variables

There are available data records from the three main parameters involved in the process: river level, sea level and waves. The record of sea water level is located in the gauge of Toulon, although this gauge is not directly in Hyères, not big differences in terms of sea level are expected. Two points of reanalysis wave data are located near the area of study (P81 and P82) and there is also a gauge which records river level measurements at Saint-Eulalie. Nearshore wave data are needed for the flooding model; these data are obtained after propagating the offshore conditions until the shore line with the SWAN model.

Not all the data have the same temporal coverage and resolution, and there are also gaps in the time series. In Figure 57, there are represented the time series of wave height, river level and sea water level without gaps in any of the three variables.

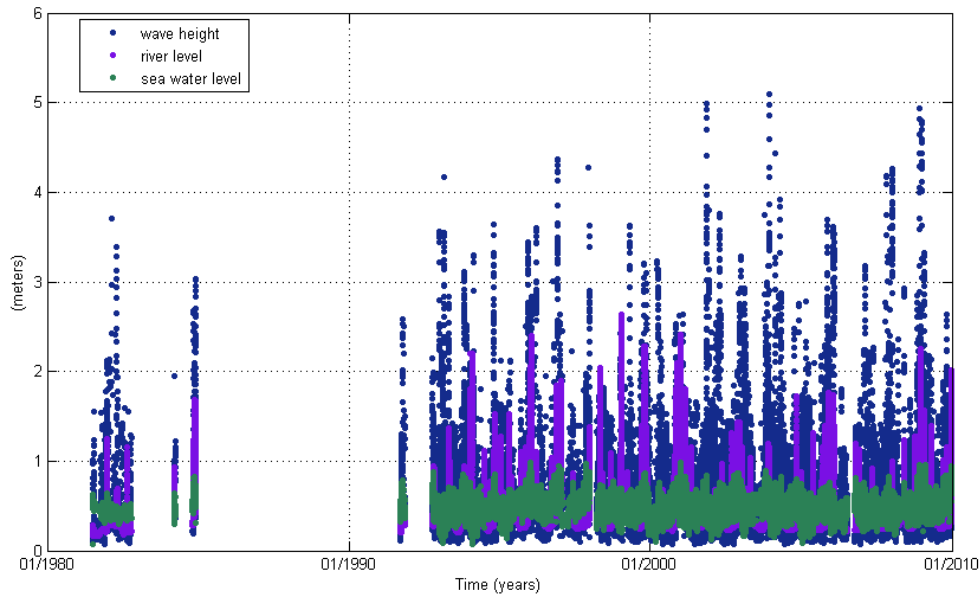


Figure 57. Time series of wave height at P81, river level and sea water level

b. Problem presentation

There are multiple possible approaches to deal with the problem presented. In this case of study we are going to focus in one and analyze its different options. The scheme of the procedure followed in this approach is as follows:

1. Independent events selection.
2. Multivariate Extreme Value Analysis method application.
3. 100 years Return Period triplets (H_s , SWL, RL) selection.
4. Propagation of the 100y. RP triplets to the nearshore.
5. Flooding model with the propagated 100y RP cases.
6. Estimation of 100y RP flooding levels.

In the following sections all this steps will be described in detail.

B. Independent events selection

The use of multivariate extreme value methods usually require a previous declusterization of the data in order to work with a subset of independent extreme events. This first step is then of great importance because it would lead into different results.

Two procedures of selection were studied, one based on a Peaks Over Threshold method and another one based on a Block Maxima method.

5-days block maxima

Considering the River Level, a subset of 5-days block maxima was extracted. Then, the other two variables were selected as the maximum within a 2.5 days window centered in the river

level maxima. In Figure 58 there is a graphical representation of the procedure followed to extract the independent events.

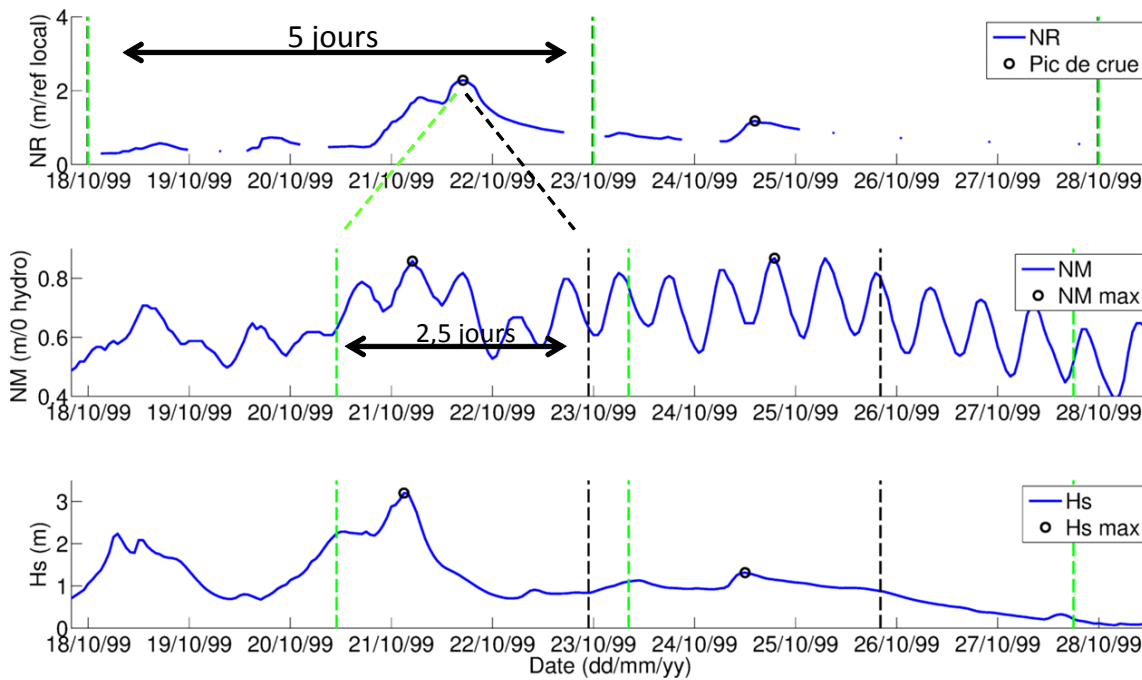


Figure 58. Graphical scheme of the 5-days block maxima procedure followed

By means of this technique a subset of 1200 independent triplets (H_s , SWL and RL) were extracted from the 16.5 years' time series.

FL-POT

The other way of extracting the independent events used was by means of a Peaks Over Threshold method. In order to do so, a theoretical flooding level (FL) parameter combining the 3 variables involved in the process was constructed. This parameter should include the relative weighted influence of the river level respect to the sea parameters. In a theoretical case, and avoiding the effect of the waves, we can assume that the level at the river mouth is the addition of the sea water level plus the river level multiplied by a coefficient α (see Figure 59).

$$RML = SWL + \alpha RL$$

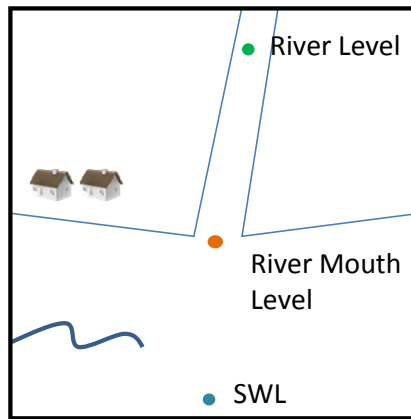


Figure 59. River mouth level scheme

To estimate this different influence of parameters (estimate α) some theoretical tests with the flooding model were carried out. These theoretical tests comprised all the combinations between 6 values of SWL (0, 0.2, 0.4, 0.6, 0.8 and 1 m) and 7 values of river discharge (0, 50, 100, 150, 200, 250 and 300 m³/s) which in terms of river level, RL, gives values of (0, 1.127, 1.578, 1.928, 2.225, 2.487 and 2.724 m). This made a total of 42 tests. The results of these tests are depicted in Figure 60. From these results a value of $\alpha=0.11$ (slope of the regression) could be estimated.

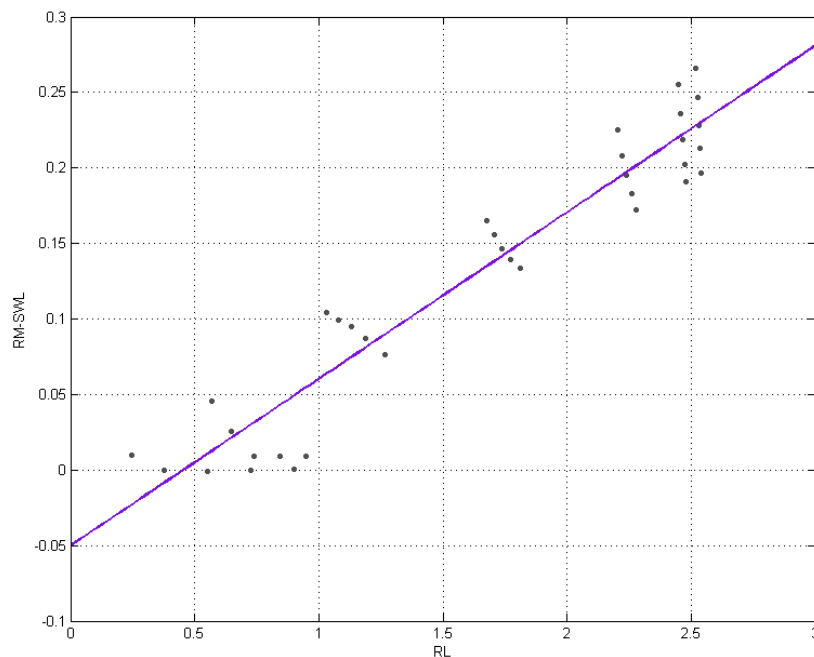


Figure 60. α estimation

Then, the flooding level could be calculated by:

$$FL = SWL + Set\ up + 0.11\ RL$$

To extract the independent events, a Peaks Over Threshold method was applied to this flooding level variable. The threshold of the POT was set in such a way that assured a subset of 1200 events (to be able to compare with the other method) and a lag between events of 1

day minimum was established. In Figure 61, the peaks extracted from the time series are depicted.

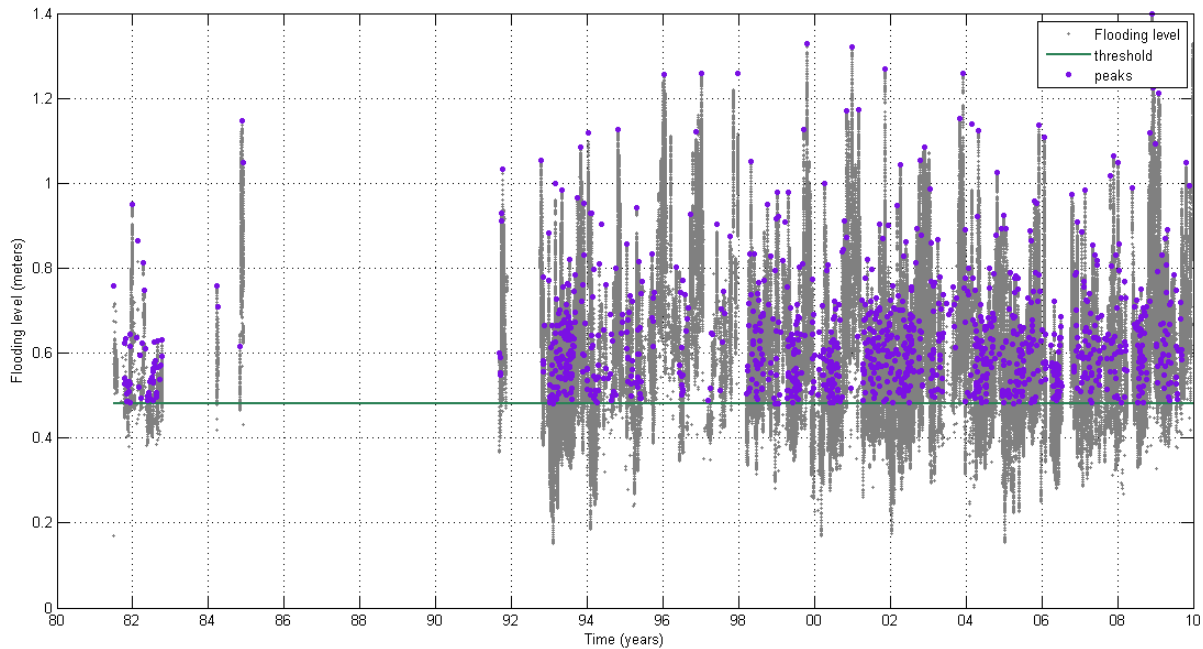


Figure 61. Flooding level peaks over threshold

The subset of triplets (H_s , SWL and RL) associated to the FL peaks extracted from the POT analysis is then the second subset of independent events used in the extreme value analysis methods.

C. Multivariate Extreme Value Analysis

Once a subset of independent extremes has been extracted from the times series, a multivariate extreme value analysis can be done. Two different methods to do it were used: the method proposed by (Heffernan 2004) and the Join Sea method, a software developed by HR Wallingford. The three variables sampled are significant wave height (H_s), sea water level (SWL) and river level (RL). Other parameters related to the waves (peak period, T_p and peak direction, θ_p) are sampled conditioned to H_s . In the case of the peak period, this conditioned sampling was done considering the wave steepness relationship.

Heffernan and Tawn 2004

The Heffernan and Tawn method (Heffernan 2004) is a method that can be classified as a conditioned extreme model. This method can be succinctly described as follows:

- i) Transform each variable into a Gumbel distribution.
- ii) Estimate the dependence between variables for extremes.
- iii) Sample values that follow these dependences.
- iv) Transform them back into the original scale.

The major advantages of their model are that there is no need of estimating the coefficient of tail dependence, it can be applied to higher dimension problems and there is no need of assuming given distributions to the marginals.

More detailed information was specified in the Annex1. Case Study 1, and bibliography regarding this model and its application to different studies are summarized in 3.2.

Join Sea

Join Sea is a software package coded in Fortran 77 developed by HR Wallingford together with the University of Lancaster in the late 90s. It is an open source code but it requires the use of NAG libraries (Numerical Algorithms Groups), which requires a license.

The Join Sea code allows a statistical analysis of input data to obtain estimates:

- Marginal distributions and extreme values of water levels, wave heights and steepness (which is directly related to T_p and is more robust to statistical calculations)
- Extremes of the joint distribution of H_s and SWL,
- Distributions and extreme values of structural variables (such as run-up, or overtopping).

It has the advantage that includes the variability of the wave period (or steepness) which plays an important role in certain structural features.

The method is divided into 5 steps:

1. Pre-treatment of the input data.
2. Marginal fitting.
3. Correlation between variables estimation.
4. Extremes simulation.
5. Extreme value analysis.

D. Cases

Combining the 2 ways of extracting the independent extreme events and the 2 methods of multivariate extreme value analysis, 5 different cases were tested:

- Case 1. 5 days block maxima + Join Sea
- Case 2. 5 days block maxima + Heffernan and Tawn with rescaling the marginals
- Case 3. Flooding level POT + Join Sea
- Case 4. Flooding level POT + Heffernan and Tawn

Rescaling process

In Case 2 the dataset of independent events extracted through the method of 5 days block maxima is used as input for the model proposed by (Heffernan 2004). In order to follow the same procedure as in the Join Sea method, after the simulation process a rescaling of the

simulated events is performed. This rescaling is an option available in the Join Sea software. With this option, the use of improved marginal is allowed to get a better estimation of the simulated data. In the case we are studying the three variables only share a common period of 16.5 years of data, although the time series of wave height and sea water level is longer. By this rescaling we can take into account the marginals of the variables as long as they are, and thus, a better estimation of the GPDs of their right tails. The rescaling proposed within Join Sea is as follows:

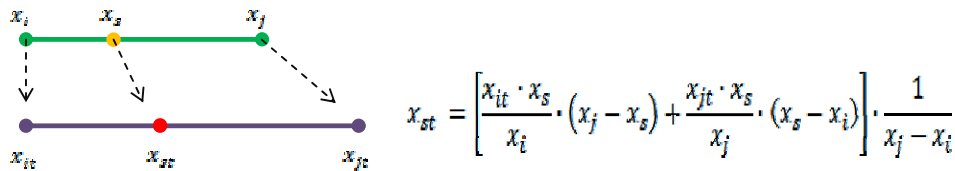


Figure 62. Scheme of the rescaling process

In Figure 62 is represented a scheme of the rescaling process which is implemented in the Join Sea software and which has been used in Case 2. Being x_i and x_j , two values of the empirical marginal distribution of independent events (in our case, that is the declustered data, the subset of 1200 triplets) and x_s , a simulated value lying between x_i and x_j ; the rescaled value of x_s will be x_{st} , which lies between x_{it} and x_{jt} , being x_{it} and x_{jt} two values of the marginal used to rescale (in our case, that is the 5 days block maxima marginal of the entire time series, not only of the 16.5 years of common data). Note that this can be done as the way of selecting the independent events is the same for the 16.5 years' time series and for the entire marginal of the three variables. For this reason, this rescaling process cannot be applied to the cases where the independent events selection was done through the POT to the FL.

In Figure 63 the results of the rescaling process applied to the three variables (H_s , SWL and RL) is presented. In red dots there are plotted the declustered marginals, it means, the 1200 independent events extracted from the 16.5 years' time series. Black dots represent the simulated data with Heffernan and Tawn considering as input the red dots. In grey dots it is represented the distribution of 5 days block maxima for the entire time series (31 years in the case of H_s , 27.1 years of sea water level data and 37.8 years of river level). Finally, the green dots represent the simulated data once they have been rescaled to consider the influence of the entire time series and not only the one of 16.5 years.

As seen; for the SWL variable there are no big differences before and after the rescaling. But, for the other two variables the differences are bigger. For example, in the H_s , there is a difference of around 0.5 meters for 100 years return period.

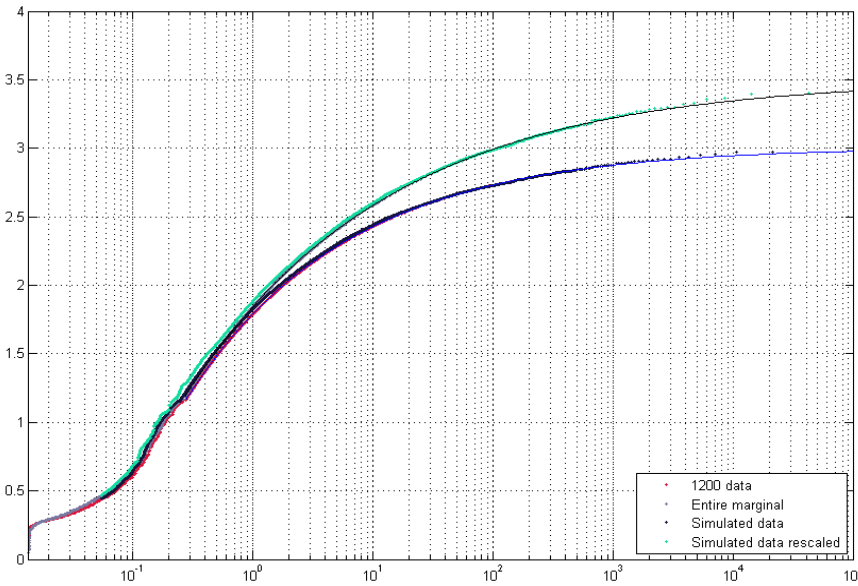
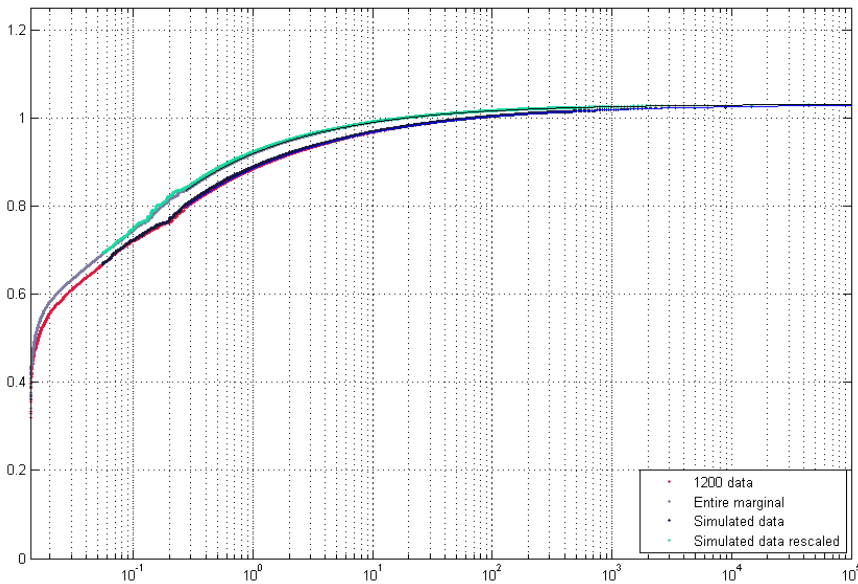
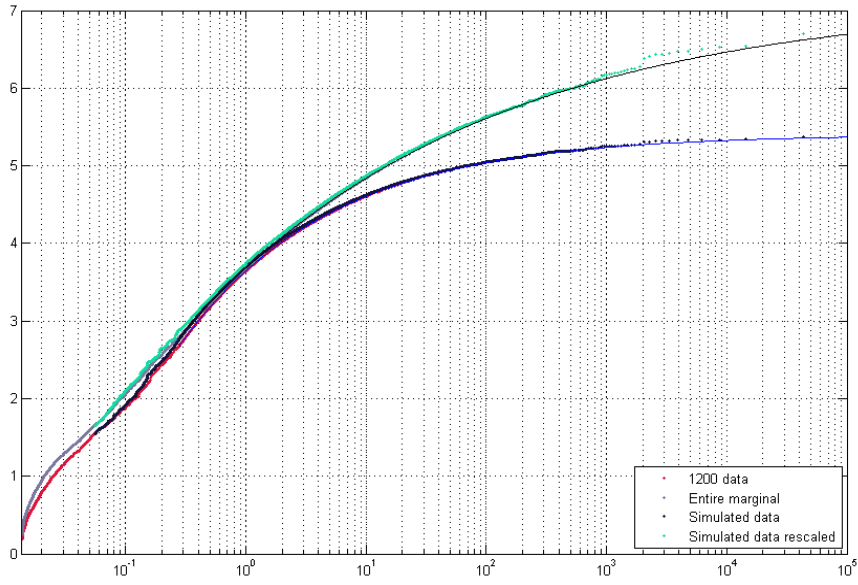


Figure 63. Marginals rescaled

E. Extreme simulations results

Case 1. 5 days block maxima + Join Sea

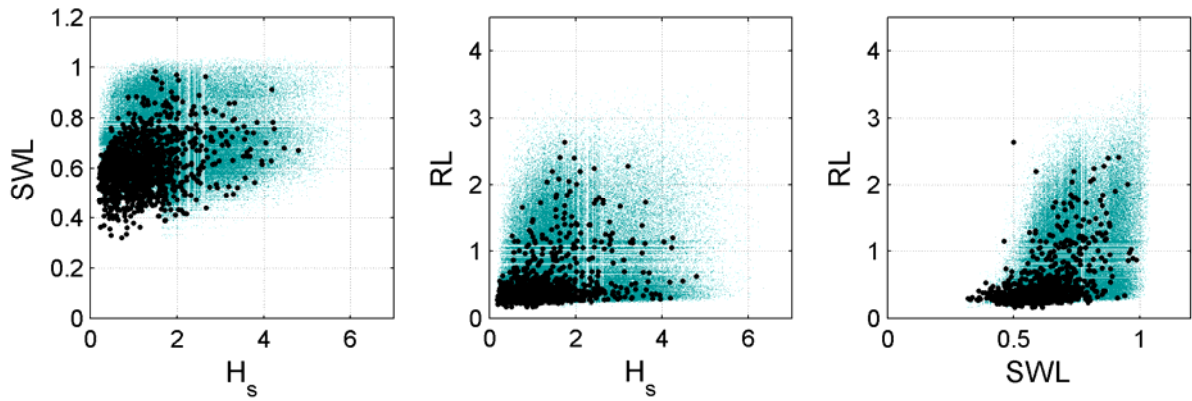


Figure 64. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 1.

Case 2. 5 days block maxima + Heffernan and Tawn + Rescaling

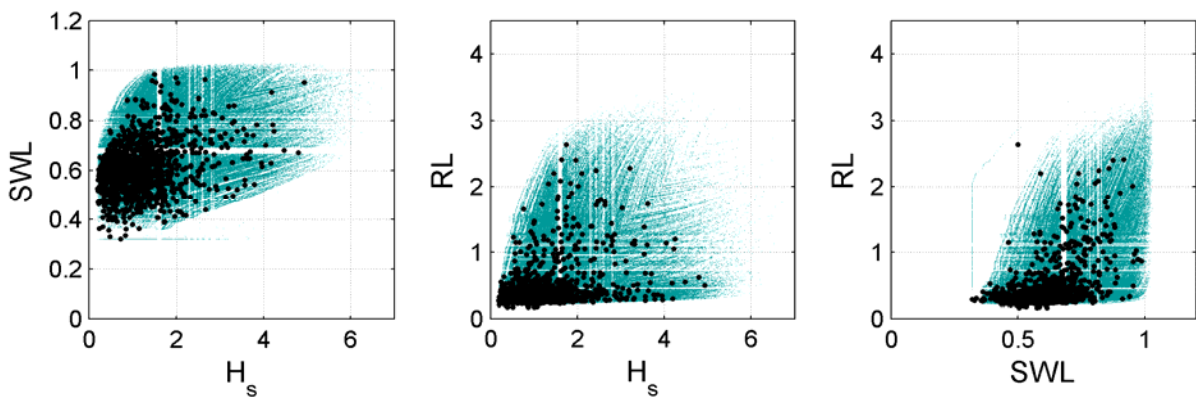


Figure 65. Joint distributions of the 3 variables sampled and rescaled (H_s , SWL and RL). Case 2.

Case 3. Flooding level POT + Join Sea

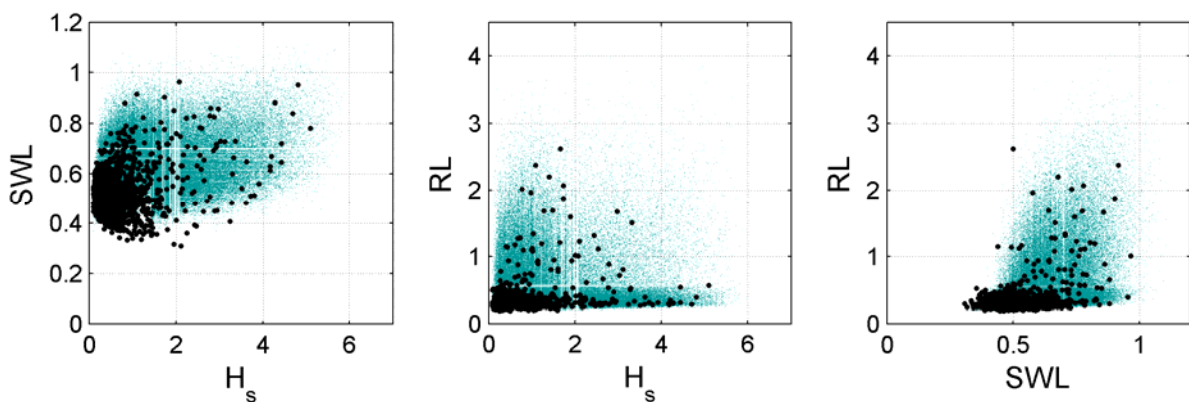


Figure 66. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 3.

Case 4. Flooding level POT + Heffernan and Tawn

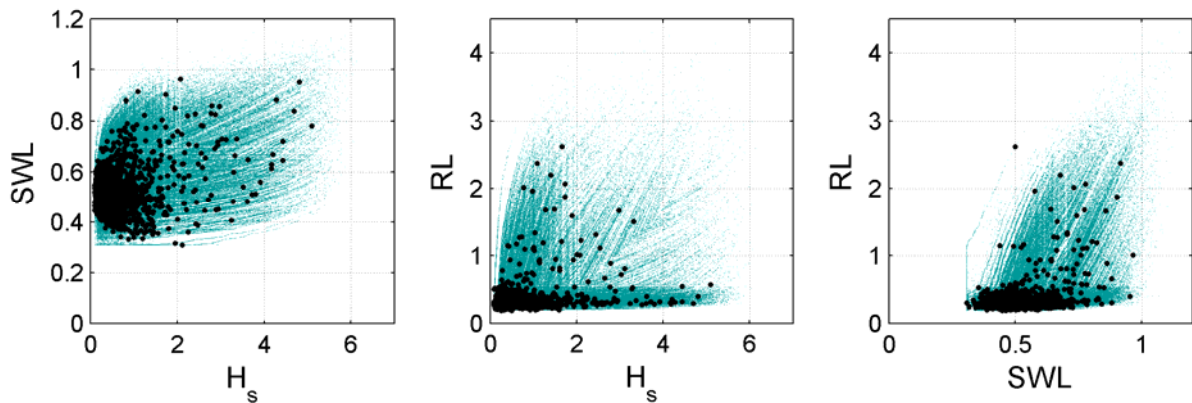


Figure 67. Joint distributions of the 3 variables sampled (H_s , SWL and RL). Case 4.

Figure 64 to Figure 67, represent the joint distributions of the three variables simulated according to the 4 cases proposed.

F. 100 years return period triplets selection

In order to select the 100 years return period triplets (H_s , SWL and RL), the surfaces and contour lines of the joint occurrence were plotted. In Figure 68 until Figure 75 the 100y RP surfaces and contours for the 4 different cases are showed. As it can be seen, there are differences between the centennial surfaces (and in consequence the contours) obtained. These differences are related to both: the two different methods to sample synthetic extremes and the two different ways of extracting the independent events.

Case 1. 5 days block maxima + Join Sea

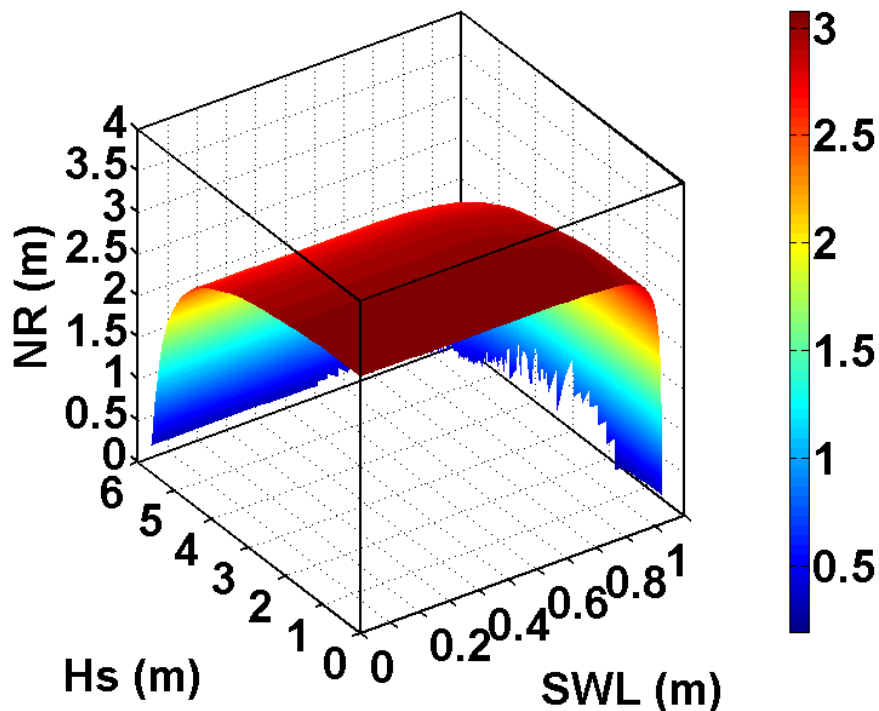


Figure 68. 100-years return period surface of H_s , SWL and RL

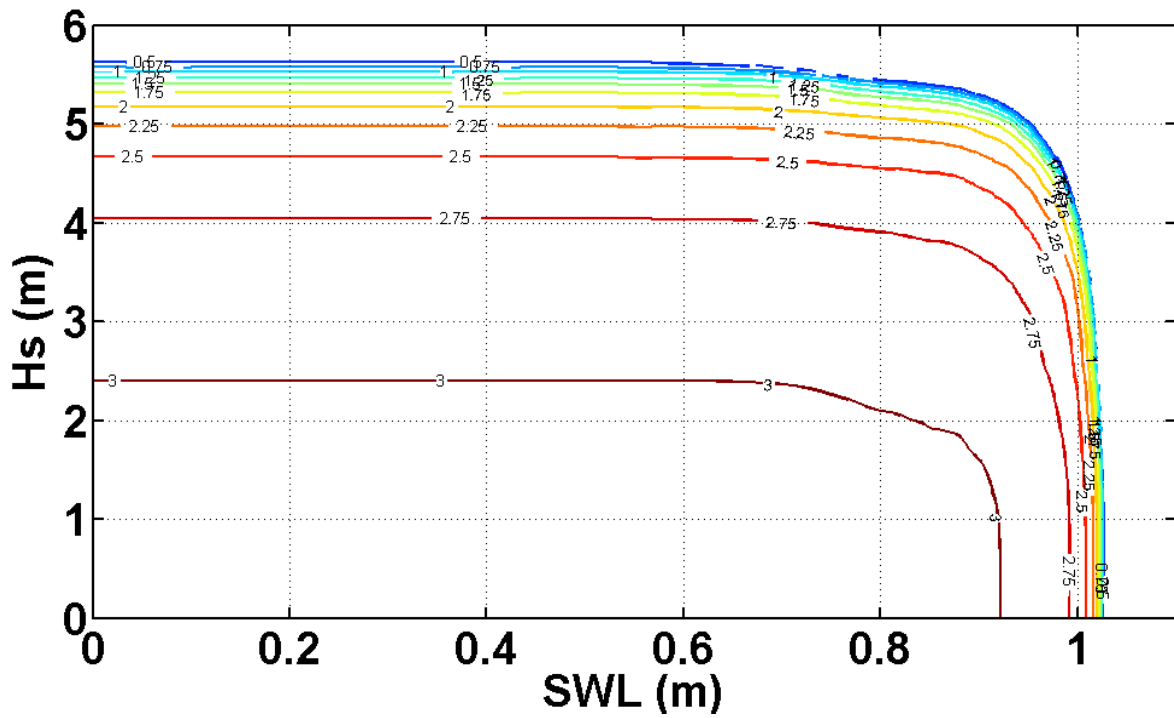


Figure 69. 100-years return period contours of H_s and SWL for different values of RL

Case 2. 5 days block maxima + Heffernan and Tawn + Rescaling

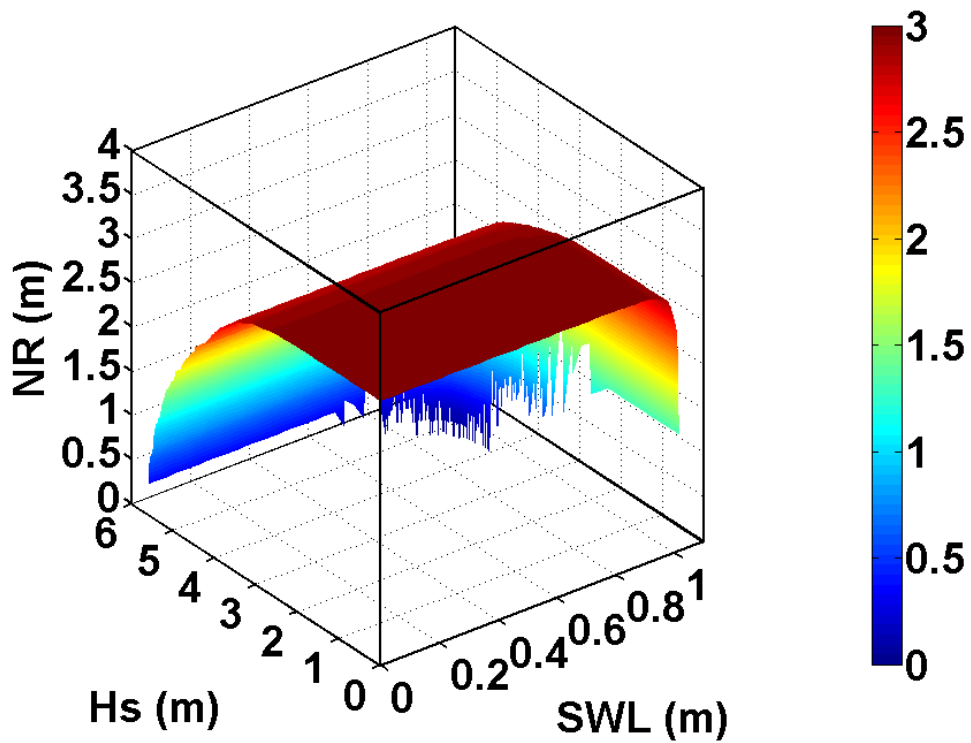


Figure 70. 100-years return period surface of H_s , SWL and RL

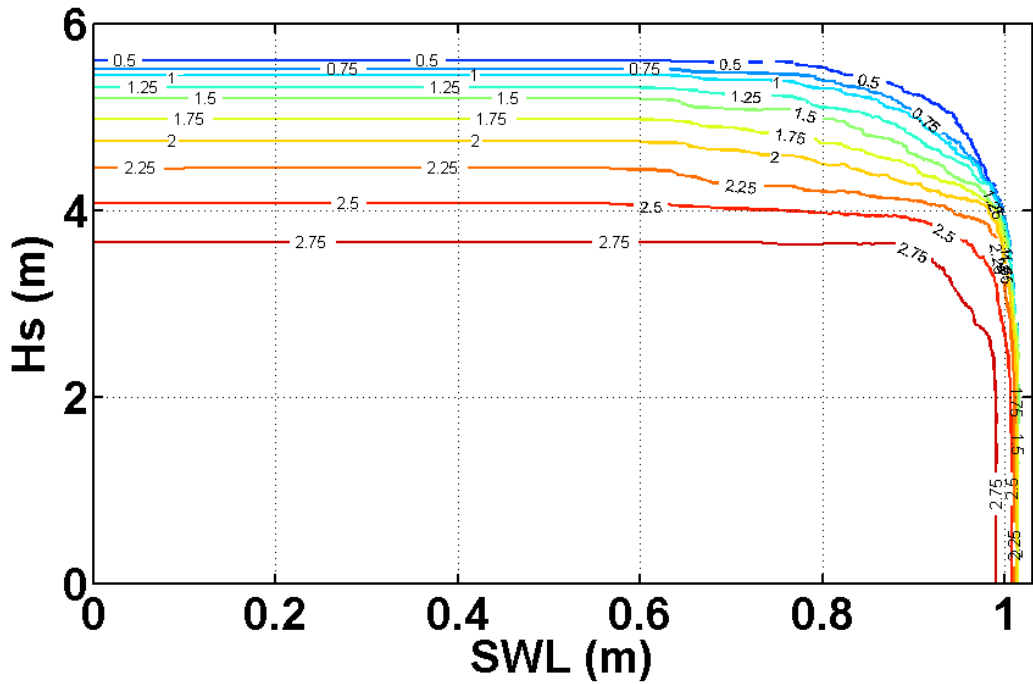


Figure 71. 100-years return period contours of H_s and SWL for different values of RL

Case 3. Flooding level POT + Join Sea

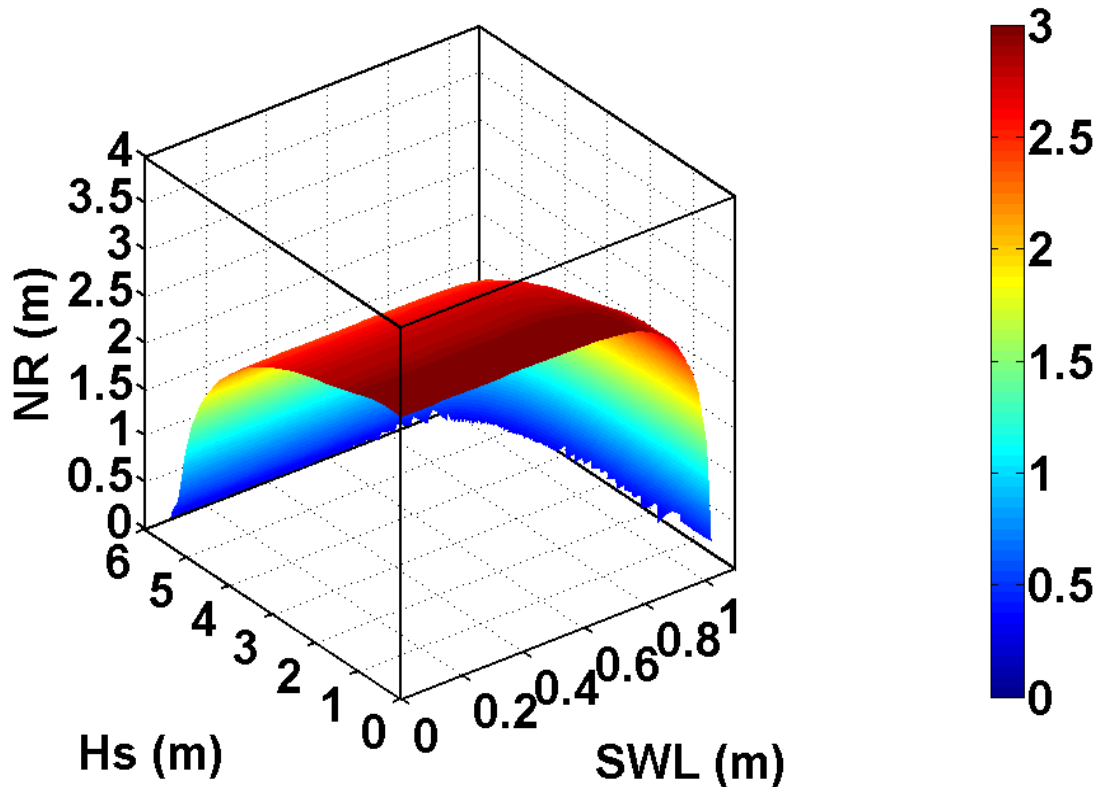


Figure 72. 100-years return period surface of H_s , SWL and RL

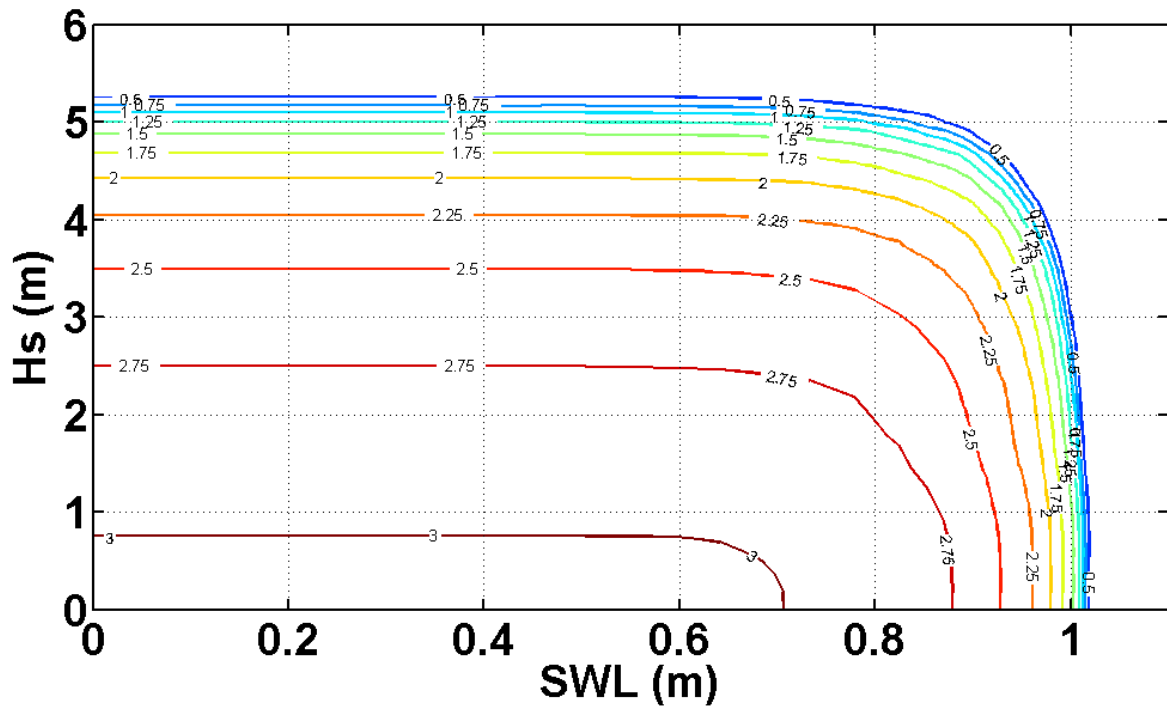


Figure 73. 100-years return period contours of H_s and SWL for different values of RL

Case 4. Flooding level POT + Heffernan and Tawn

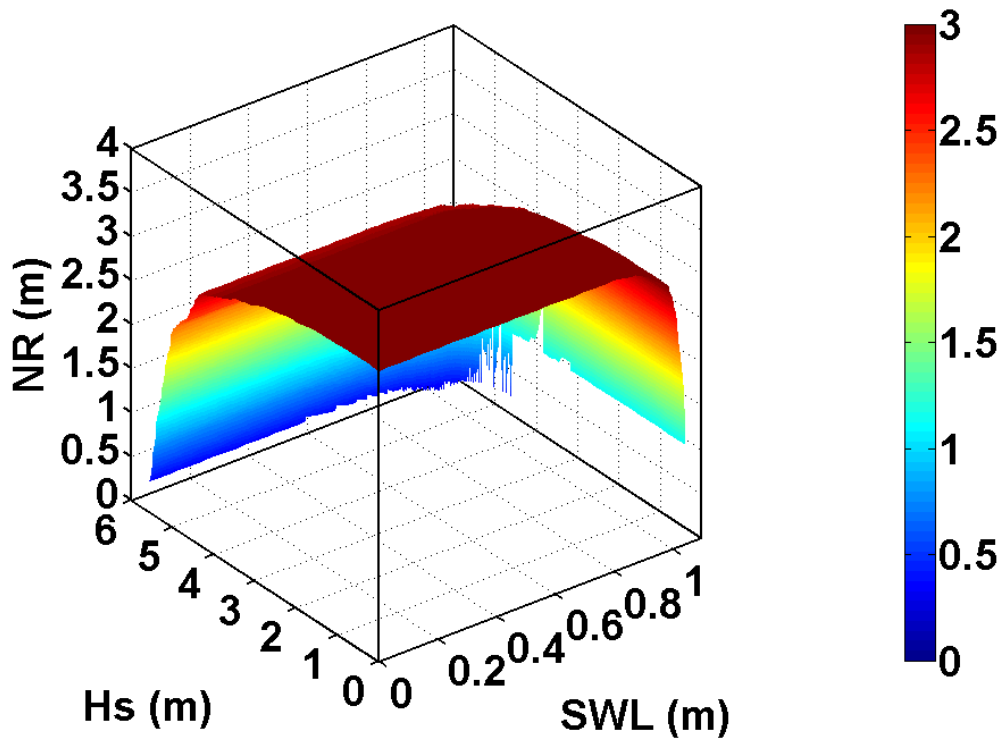


Figure 74. 100-years return period surface of H_s , SWL and RL

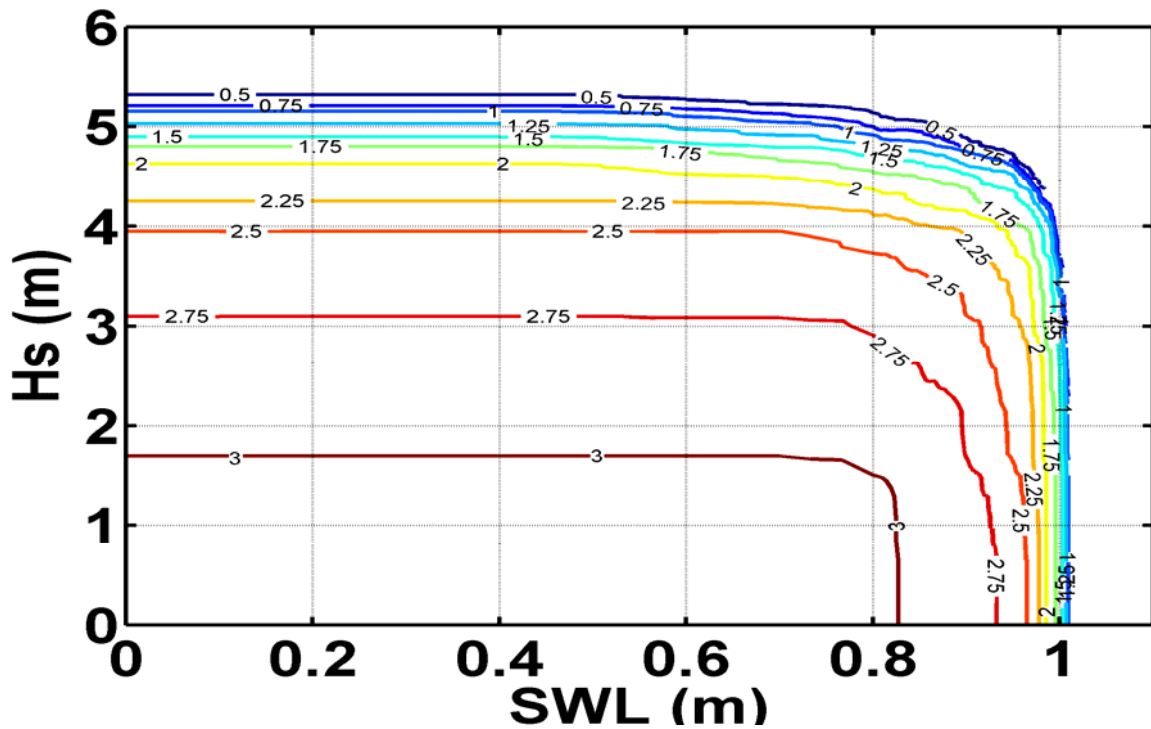


Figure 75. 100-years return period contours of H_s and SWL for different values of RL

Triplets selection

The procedure followed to obtain the 100 years return period triplets from the contour plots showed before can be summarized as follow:

- In each case a *SWL limit* was established. This SWL limit is the highest value of SWL for which the contours lines are horizontal.
- 100 years return period triplets (H_s , SWL and RL) selected will be the intersections between the contour lines at different RL values (each 0.25m) and SWL values (each 0.05 above the SWL limit) (see Figure 76).

In Figure 76 it is depicted the procedure followed. In dashed pink vertical line is defined the SWL limit in this case, and the rest of the dashed purple vertical lines are each 0.05 m of SWL. The black stars represent the points that will define the triplets, for simplicity there have been only depicted the ones corresponding to the SWL limit, the rest of triplets of this case would be the intersections between the purple dashed lines and the contour lines.

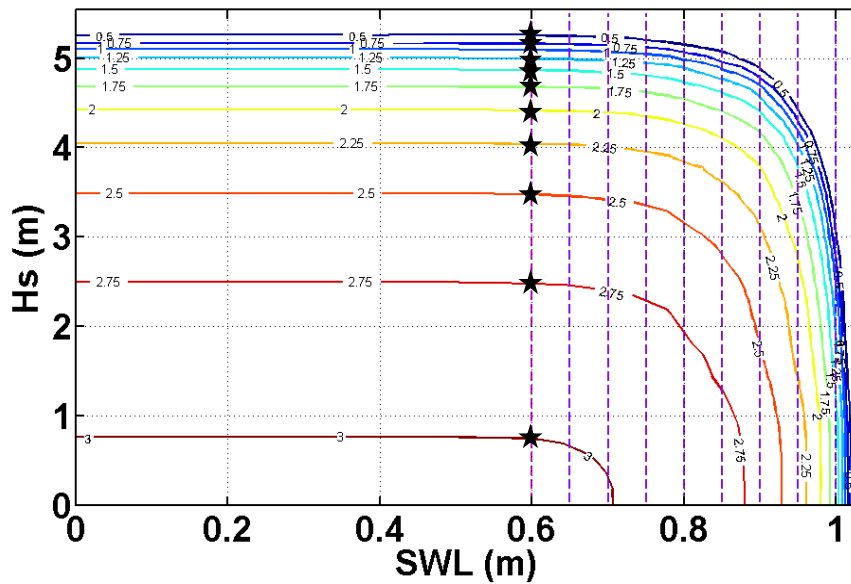


Figure 76. Example of the triplets selection procedure

This procedure has been repeated to the 4 cases of study to get the triplets to be propagated to the nearshore.

The subset of 100 years return period triplets selected is summarized in Table 28:

	SWL values	RL values	N° triplets
Case 1	0.7-1	0.5-3	74
Case 2	0.8-1	0.5-3	49
Case 3	0.6-1	0.5-3	85
Case 4	0.75-1	0.5-3	56
Total:			264

Table 28. Summary of the triplets selected.

The peak period associated has been determined through the steepness relationship. In Figure 77, the results for the 4 cases are shown. In the left plot, in grey dots there are the 16.5 years data of H_s and steepness and by colors the steepness associated to the values of H_s extracted to all the 100 years return period triplets of each case. The right side plot represents the joint distribution of H_s and T_p for all the 16.5 years' time series in grey dots and for the triplets of the 4 cases in different colors.

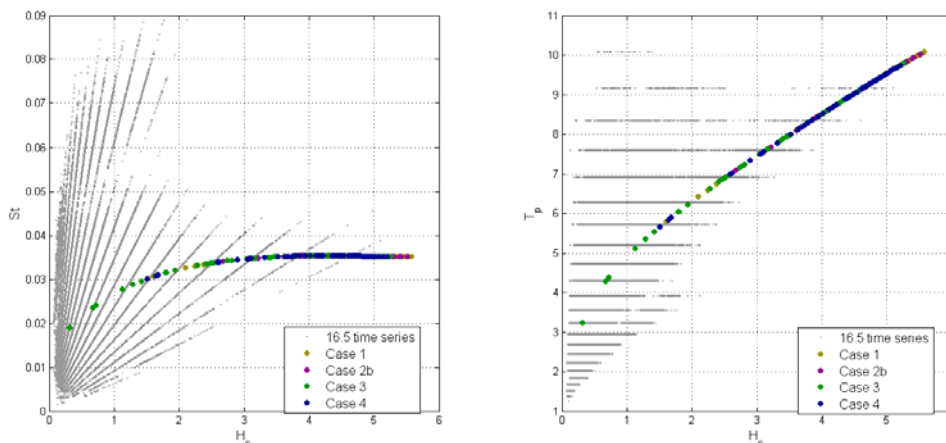


Figure 77. Left: H_s - St scatter of the 16.5 y time series and the 100 y return period triplets. Right: H_s - T_p scatter of the 16.5 y time series and the 100 y return period triplets.

The peak direction has been assumed as 180° for all the triplets. This was assumed because this is the main direction for high waves in the studied location (it is the only direction registered for waves higher than 3 meters).

G. Storms from the historical record (16.5 years)

The historical record comprised about 16.5 years of data of the three variables of interest: H_s , SWL and RL. From these data, applying a Peaks Over Threshold technique, 161 storms were extracted. These 161 storms were propagated to the nearshore by means of the SWAN model and then applied to the MARS model to get the flooding estimations.

H. Flooding simulations (MARS)

The sea states from the triplets selected (H_s , T_p and θ_w) at each case were propagated from offshore conditions to nearshore conditions by means of the SWAN model. Once the sea conditions at the nearshore were obtained they were used as input (together with the river level and sea level from the triplets) for the MARS model.

All the simulations were performed considering a storm profile (see Figure 78) extracted from the historical record. The peaks of the three variables (H_s , SWL and RL) were considered to occur simultaneously.

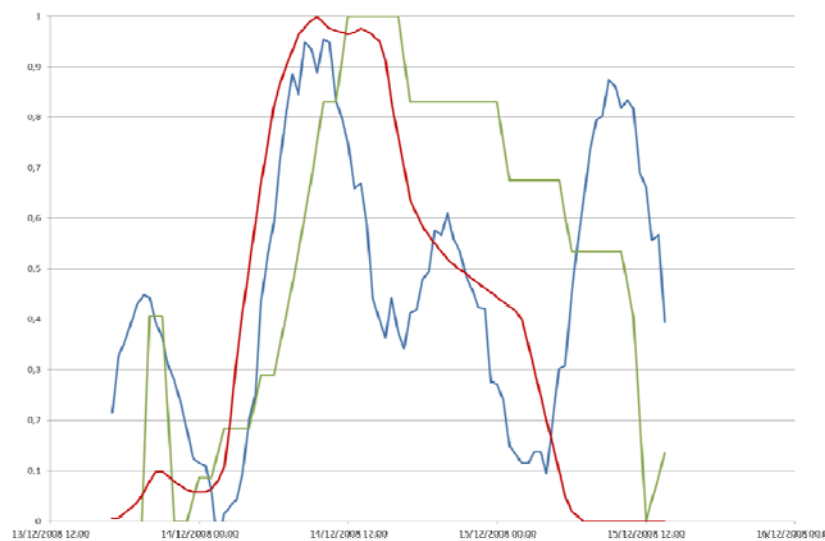


Figure 78. Storm profile normalized. Red-RL, green- H_s and blue-SWL.

To validate the assumption of the three peaks occurring at the same time, the real storm profiles from the 161 historical storms extracted from the database were analyzed. In Figure 79 the appearance of some storm profiles from the 161 historical records are shown; the delay between peaks was estimated in three ways: delay between H_s and SWL peaks (positive values for SWL occurring after H_s), delay between H_s and RL peaks (positive values for RL occurring after H_s) and delay between SWL and RL peaks (positive values for RL occurring after SWL).

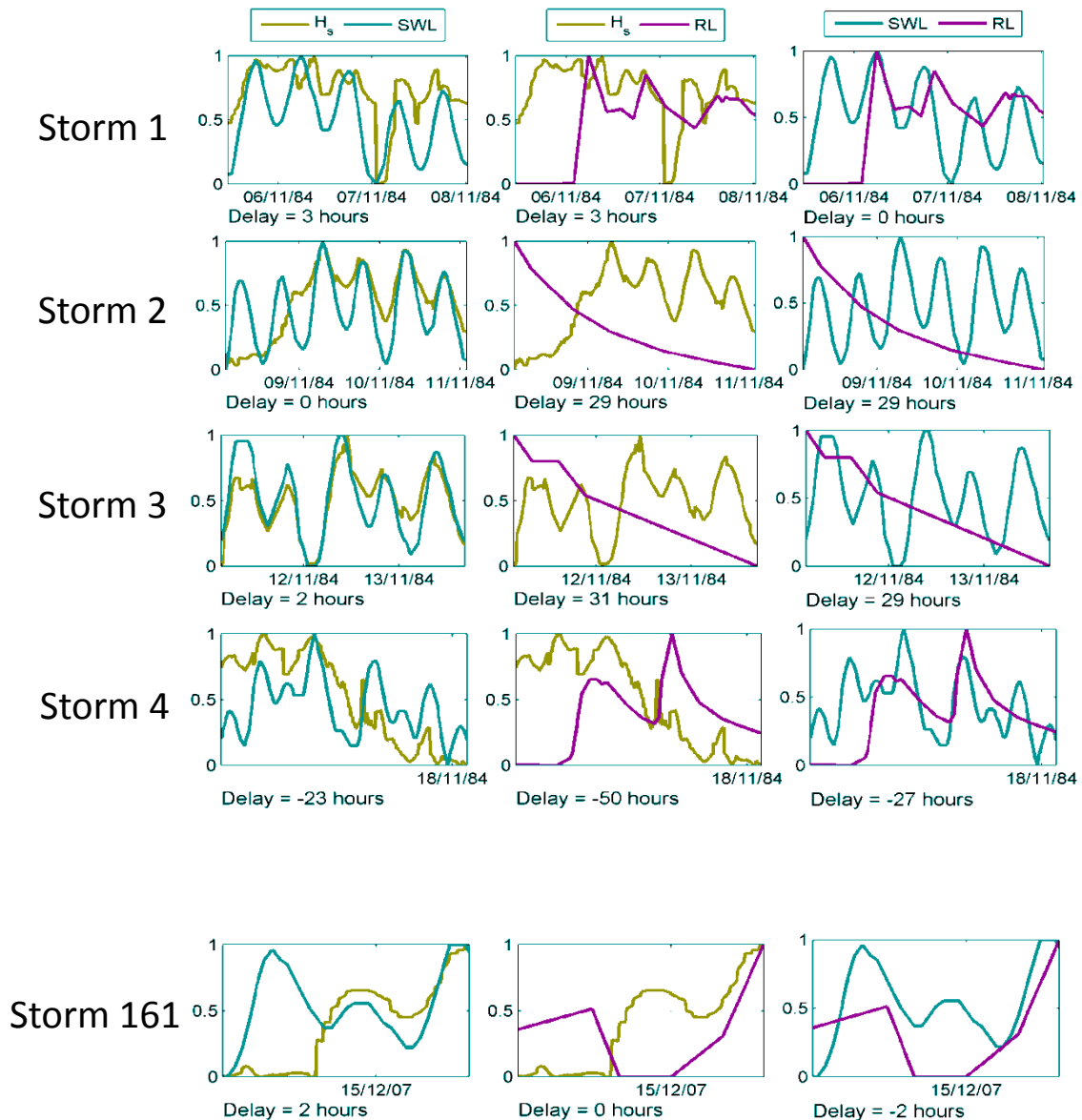


Figure 79. Storm profiles (normalized) from the historical records.

The delays estimated between variables for the 161 storms are plotted in Figure 80. And in Figure 81 are shown the boxplots of the delays with the statistics of the delay between each variable. As seen from the results depicted in Figure 81, assuming that the three peaks occur at the same time is not far away from the historical records.

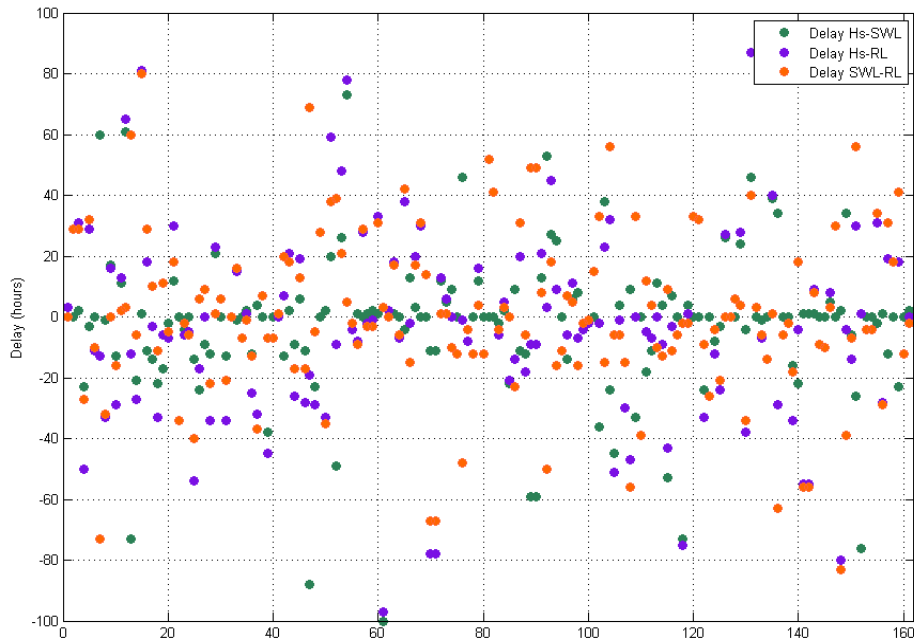


Figure 80. Delay (hours) between variables peaks for the historical storms.

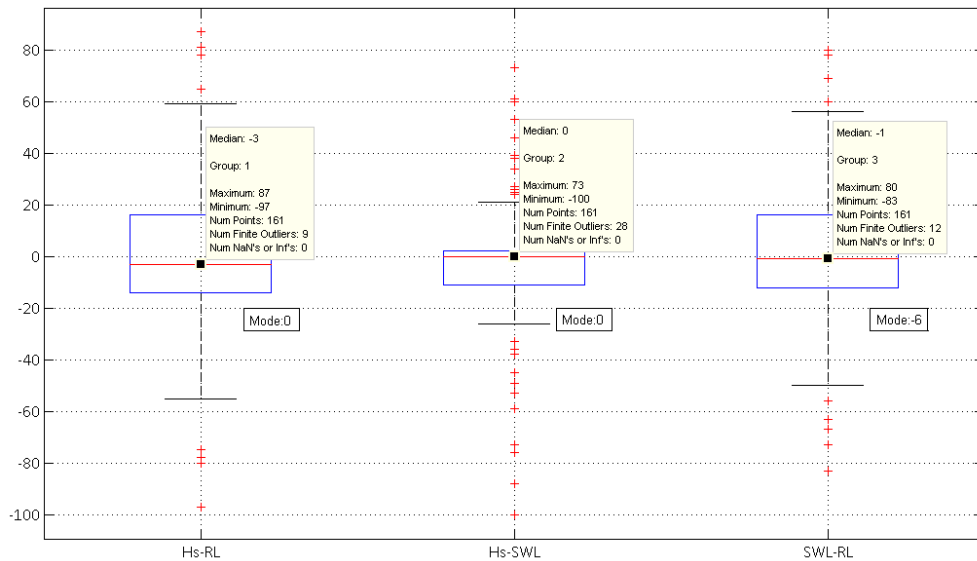


Figure 81. Statistics of the delays estimated.

I. Results

Once the simulations with the MARS model have been done, the results obtained can be compared with the estimations resulting from the historical 161 storms.

In Figure 82, the results obtained in terms of sea level at the river mouth are shown. Black dots represent the peaks from the historical storms, green lines represent a GPD fit performed to the historical storms with the 95% confidence intervals (estimated through bootstrapping). Colored dots represent the maximum value obtained from the triplets simulated according each case: in purple the case with the 5 days Block selection criteria and Join Sea simulation process, in green 5 days Block and Heffernan and Tawn simulation

model, in yellow the POT applied to a combined variable as selection criteria and the Join Sea simulation model and in blue the POT applied to a combined variable and the Heffernan and Tawn method.

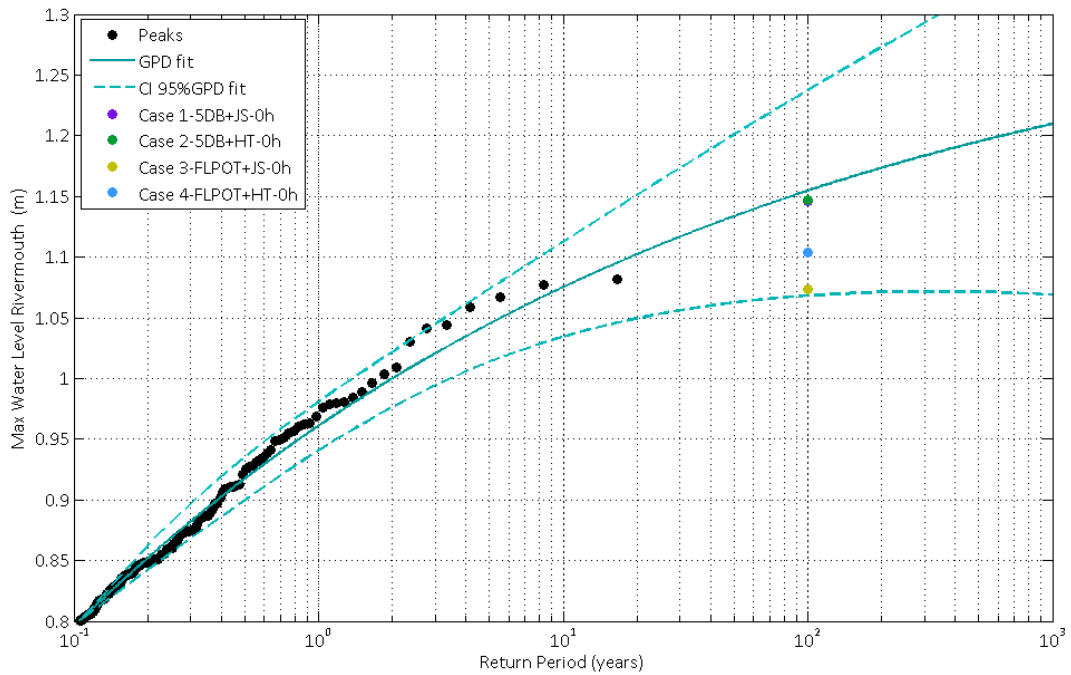


Figure 82. Results in the river mouth level

In terms of maximum flooded surface (m^2), Figure 83 represents the comparison between the simulation estimations and the historical storms. Same as before the black dots represent the historical peaks, green lines a GPD fitted to the black dots and colored dots the maximum flooded surface obtained from the triplets associated to each case proposed (the four combinations of selection procedure and simulation method already mentioned).

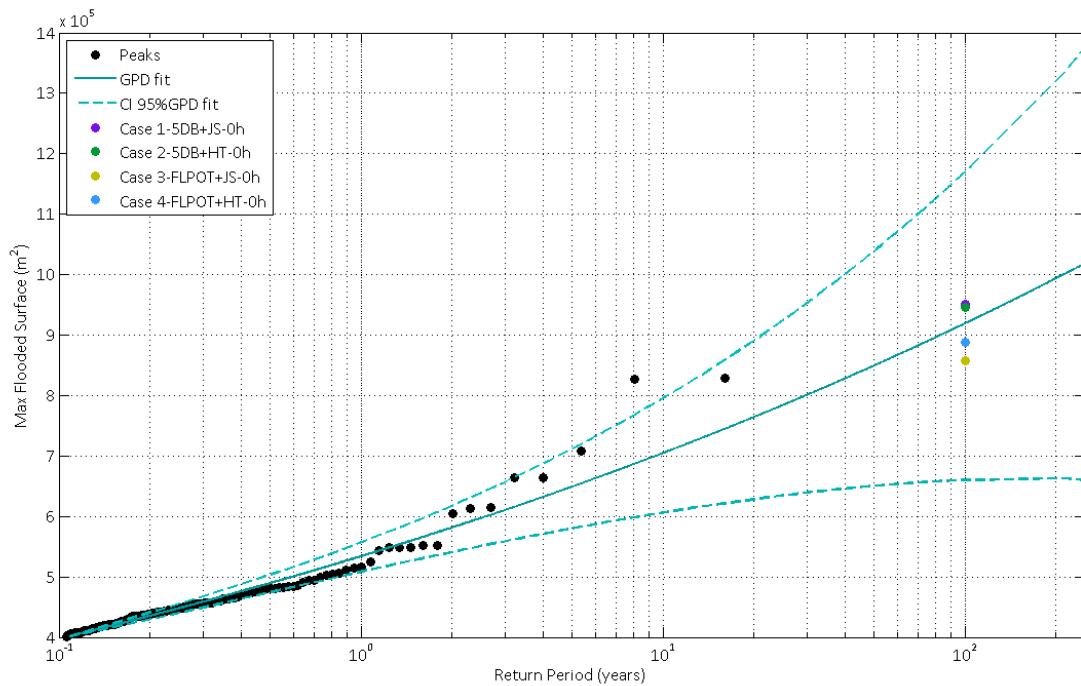


Figure 83. Results in the max flooded surface

J. Comments

From the results obtained in this Case Study 2, some comments can be extracted:

- The selection of the independent events to perform the multivariate simulation technique is really important and has direct influence in the results of the entire process. As seen in the last two plots, the results obtained with those cases where the events were selected with the 5 days block maxima are slightly higher than the other two cases: in this method, blocks of 5 days were selected and the maximum value of each variable was selected in each block. Compared to the other selection method, the triplets obtained with the 5 days blocks are higher, then the multivariate simulations gave higher results.
- The great advantage of the 5 days block selection method compared to the POT base selection method is that the first one allows a posterior rescaling of the simulated values (with both models, Join Sea and Heffernan and Tawn). This rescaling process takes into account the entire dataset of each variable. In this particular case, because of the existent gaps in the different records only 16.5 years of data could be considered comprising the three variables; but the waves and sea level record are much longer (about 30 years). Then, this rescaling 'corrects' the simulations by considering the entire records not only the common period of 16.5.
- There are no very big differences between the multivariate simulations made either with the Join Sea software or the Heffernan and Tawn method. The first method is based in Copulas functions while the second one is a conditional extreme model.
- The results presented in the previous paragraph were done in terms on maximum flooded surface over the simulated grid and sea level at the mouth of the river. Other parameters could be studied but no very big differences should be expected. In both cases, the results obtained are within the confidence interval around the GPD fitted. In the case of the sea level at the mouth river results, the four cases give results below the GPD fit while the results of the maximum flooded surface are more variable.