ISBN 978-82-326-0994-9 (printed version)
ISBN 978-82-326-0995-6 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2015:169

Oleh Karpa

Development of Bivariate Extreme Value Distributions for Applications in Marine Technology

NTNU
Norwegian University of
Science and Technology
Faculty of Engineering
Science and Technology
Department of Marine Technology

NTNU – Trondheim
Norwegian University of
Science and Technology

Oleh Karpa

Doctoral Thesis

Doctoral theses at NTNU, 2015:169

NTNU

NTNU – Trondheim
Norwegian University of
Science and Technology

Oleh Karpa

Development of Bivariate Extreme Value Distributions for Applications in Marine Technology

Thesis for the degree of Philosophiae Doctor

Trondheim, June 2015

Norwegian University of Science and Technology

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

ISBN 978-82-326-0994-9 (printed version)
ISBN 978-82-326-0995-6 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2015:169

Printed by Skipnes Kommunikasjon as

I dedicate this thesis to all those heroes who have suffered and laid down their lives for a free and independent Ukraine since the Revolution of Dignity and the outbreak of the Russian "hybrid" military intervention in Ukraine

Abstract

The extreme value theory for applications in such a responsible branch of industry as offshore and maritime engineering requires a robust, straightforward and reliable method for estimating the statistics of extremes. A method must be able to extract as much statistical information as possible from a recorded time series of data. In addition, a method must be capable to utilize the information regarding the temporal dependence structure of the process, as well as spatial dependence characteristics of the given time series in the bivariate case.

In this thesis, a newly developed method for the purpose of predicting extremes associated with the observed process is studied thoroughly and improved. The method is referred to as the average conditional exceedance rate (ACER) method. It avoids the problem of having to decluster the data to ensure independence, which is a requisite component in the application of, for example, the standard peaks-over-threshold (POT) method. Moreover, the ACER method is specifically designed to account for statistical dependence between the sampled data points in a precise manner. The proposed method also targets the use of sub-asymptotic data to improve prediction accuracy. The research shows that the ACER method, if properly implemented, is able to provide a statistical representation with error bounds of the exact extreme value distribution given by the data. In the first part of the thesis, the method is demonstrated in detail by application to both synthetic and real environmental data. From a practical point of view, it appears to perform better than the POT and block maxima methods, and, with an appropriate modification, it is directly applicable to non-stationary time series.

In the second part of the thesis, the ACER method for estimation of extreme value statistics is extended in a natural way to also cover the case of bivariate time series. This is achieved by introducing a cascade of conditioning approximations to the exact bivariate extreme value distribution. The results show that when the cascade converges, an accurate empirical estimate of the extreme value distribution can be obtained. It is also revealed that the possible functional representation of the empirically estimated bivariate ACER surface can be derived from the properties of the extreme-value copula.

In this thesis, application of the bivariate ACER method is substantially studied for bivariate synthetic data. Finally, performance of the method is demonstrated for measured coupled wind speed and wave height data as well as simultaneous wind speed measurements from two separate locations.

Contents

List of figures

List of tables

Preface

This dissertation is submitted to the Norwegian University of Science and Technology (NTNU) for conclusive fulfillment of the requirements for the degree of Philosophiae Doctor.

The results presented in this doctoral thesis were obtained during my employment at the Centre for Ships and Ocean Structures (CeSOS) of the Department of Marine Technology, NTNU, Trondheim, in the period September 2009–February 2013. This research work has been performed under the wise and effective guidance of my supervisor, Professor of the Department of Mathematical Sciences and Centre for Ships and Ocean Structures, Doctor of Technology, Arvid Næss.

The financial support of the present research project, with the assigned number 40069400, was provided by the Research Council of Norway (NFR) through the Centre for Ships and Ocean Structures (CeSOS) at the Norwegian University of Science and Technology.

Acknowledgements

No man is an island,
entire of itself;
Every man is a piece
of the continent,
a part of the main.

– John Donne,
Meditation XVII

Indeed, no man is an island, and one cannot do any scientific work solely by oneself. In the same way, it would not have been possible to complete the present research without kind support and invaluable guidance from the people I would especially like to thank.

First and foremost, I offer my utmost gratitude to my supervisor, Professor Arvid Næss of the Department of Mathematical Sciences and Centre for Ships and Ocean Structures. In particular, I thank him for his goodwill to hire me in the first place. I greatly appreciate his guidance, many ideas and comments on our joint research as well as his moral support, which encouraged me to move forward. There is no doubt that one simply could not wish for a better or friendlier supervisor. Definitely, there would not be any Dr. Karpa without Professor Næss, and I am indebted to him more than he probably knows.

I thank Dr. Anatoly Lavrenyuk for recommending me to Dr. Oleg Gaidai. I also appreciate Dr. Gaidai for providing a recommendation for me to Professor Arvid Næss. In addition, I am grateful to Dr. Gaidai, Dr. Olexandr Batsevych, and Dr. Mahdi T. Sichani for their valuable support in the early stages of my project.

I would also like to thank the Centre for Ships and Ocean Structures (CeSOS), part of NTNU, who have been my employer while I was working on my thesis. I am particularly grateful to Professor Torgeir Moan, to Sigrid Bakken Wold, senior executive officer, and to Karelle Gilbert, former secretary, for the support and assistance they have given me anytime I needed it.

My sincere thanks go to my many friends and colleagues, who, in some way or another, have supported and strengthened me. Particularly, I would like to thank the following wonderful people: Professor Luis Volnei Sudati Sagrilo of the Federal University of Rio de Janeiro, Professor Paul C. Kettler, Daniel de Almeida Fernandes and his wife Renata Kurimoto Marques, Terje Nilsen, Yuriy Zamchalkin, Taras Pauk, Yuriy Tabin and Rostik Slipetskyy. Rostik once said to me “If every

scientist would have written his thesis as 'promptly' as you have, humans would still have lived in caves!" – A good joke and a good incentive for me.

Last, but certainly not least, I would like to thank my family for their unending support, encouragement and prayers. My parents, Ihor and Nadiya; my sister, Dr. Maryana Mykolaychuk and my little niece Myroslava; my grandmother Hanna; and more distant relatives. It definitely would have been extremely arduous to complete this research without support from my dearest ones.

Oleh Karpa

L'viv – Trondheim, May 2015

Chapter 1

Introduction

In this chapter, the main statements of the problem in extreme value statistics are presented. The essences of some background cornerstones necessary for better understanding of the work and its continuity is included. The structure of the thesis is also addressed in this chapter. The main focuses of the thesis are: the detailed investigation of the performance of recently developed average conditional exceedance rate (ACER) method on the ground of a significant amount of both simulations and observational studies; development of the ACER method for the bivariate case and investigation of its properties with implication of a certain copula models.

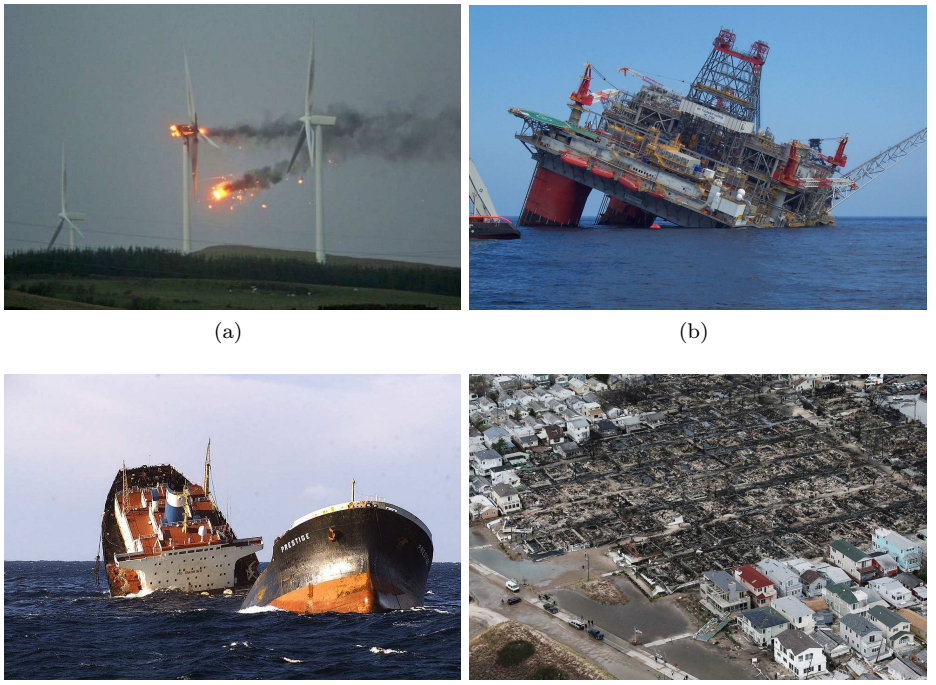
1.1 Motivation

In nature, extreme events occur almost worldwide and cause incalculable human losses, in addition to billions of dollars in damages, each year. For example, according to the Intergovernmental Panel on Climate Change (IPCC), in 2012 overall damages caused by natural disasters came to US $170 billion and insured losses US $70 billion. Of the 905 documented natural loss events, 93% of which were weather-related disasters, 45 percent were meteorological events (storms), 36 percent were hydrological events (floods), 7 percent were geophysical events (earthquakes and volcanic eruptions). Deaths during natural catastrophes in 2012 stood at 9,600 – substantially below the 10-year annual average of 106,000 (Löw, 2015). Thus, although 2012 was a moderate year, this statistics is still high.

In 2011 IPCC estimated that annual losses have ranged since 1980 from a few billion to above US$ 200 billion (in 2010 dollars), with the highest value for 2005 (the year of Hurricane Katrina)(NOAA, 2015). The global weather-related disaster losses reported over the last few decades reflect mainly monetized direct damages to assets, and are unequally distributed. Loss estimates are lower bound estimates because many impacts, such as loss of human lives, cultural heritage, and ecosystem services, are dimensionless, and thus they are poorly reflected in estimates of losses (Smith and Katz, 2013).

The effects of weather-related hazards, such as windstorms, hurricanes, typhoons and landslides, are also known to have been severe. Earthquakes, nature's

most destructive force, caused the highest economic losses. Figure (1.1) illustrates some of the possible damages caused by weather-related disasters.



(a) (b)

(c) (d)

Figure 1.1: Impact of extreme events: a) A wind turbine caught fire in high winds. Scotland, December 2011; b) Thunder Horse oil platform damaged during Hurricane Dennis. Gulf of Mexico, July 2005; c) The "Prestige" oil tanker burst during a storm off Galicia, then split in half and sank. Spanish coast, 13-19 of November 2002; d) Horrific impact of Hurricane Sandy. Northeastern United States, October 2012.

"Prepared is protected", therefore, studying the statistics of extreme events and predicting their occurrence is a first and important step in the mitigation of these disasters. Thus, extreme value theory is clearly appropriate and useful for predicting extreme events related to large-scale environmental processes. In addition, this theory and its practical implementation are also highly important in areas of application such as the reliability of structures or, more generally, failure analysis of engineering systems, resistance or capacity estimation of engineering materials, and the effect on marine structures of environmental processes such as wind, ocean waves and currents. The design of the structures subjected to environmental loads, e.g., offshore platforms, ships, aircraft, and tall buildings, include analyses of load effects, particularly of extreme loads and load effects. Thus, the development of efficient and accurate methods for estimating extreme value statistics of load effect processes is crucial for the development of rational design provisions for offshore

structures, as an example.

The major problem associated with extreme value prediction in practice has been the analysis and application of certain probability models and probability distributions with unknown parameters to be estimated based on the observed data.

1.2 Background

The classical univariate theory of extreme values was developed by Maurice R. Fréchet, Sir Ronald A. Fisher and Leonard H.C. Tippett in the twenties of the twentieth century. Boris Gnedenko and Emil J. Gumbel showed that the largest or smallest value from a set of independently distributed random variables tends to an asymptotic distribution that only depends on that of the basic variable. After standardization using suitable norming and centering constants, the limit distribution is shown to belong to one of three types, as noted by Gnedenko and Laurens de Haan. In Chapter 2 these results are presented in more detail.

In practice, extreme value statistics have often been based on these asymptotic results. Thus, parameter estimation problems for the asymptotic distributions are directly related to problems of optimization and parametric curve fitting to the observed data.

Although univariate techniques are now standard, these methods for estimating extremes (wind speeds, wave heights, structural responses, etc.) from observed data series are commonly based on assuming either that epochal extreme values are Gumbel distributed or by adopting a peaks-over-threshold (POT) approach, assuming that the exceedances above high thresholds follow a generalized Pareto distribution. An important weakness of these approaches is that they depend on adopting asymptotic distributions, which cannot be verified in practice due to the difficulty of ascertaining that the data at hand are of a truly asymptotic character.

Naess and Gaidai (2008) developed an accurate and efficient method for estimating extreme values of stochastic processes under certain conditions on the dependence structure of the time series under study. The method is based on Monte Carlo simulations. It is therefore highly suitable for use in the estimation of extreme values of combined stochastic load effect models because Monte Carlo simulations are very often possible in such cases. However, in some cases, the adopted assumption about the dependence structure, which amounts to the so-called Poisson assumption of independent up-crossing events of high response levels, may be inaccurate.

This has resulted in a method that appears to be more appropriate for the purpose of predicting extremes than the traditional methods. It is based on a cascade of conditioning approximations that makes it easy to account for dependence effects in the data series. This method is referred to as the ACER method, and its development for extreme value prediction based on sampled data has been described by Naess and Gaidai (2009), Naess, Gaidai, and Batsevych (2009).

However, it is clear that a timescale must be included in conjunction with time series phenomena, for example, exceedance times and exceedances are jointly visualized in a scatterplot. This was a first step towards using multivariate data.

The mathematical theory of multivariate extremes is a novel and rapidly growing field. Several areas are well developed, and there are the Block and Threshold models that are analogous to the univariate case. For the univariate case, these models have only asymptotic justifications, and their suitability for any practical application must be verified with care. The first problem that arises when working in a multidimensional context is the lack of a "natural" definition of extreme values; essentially, this is due to the fact that different concepts of ordering are possible. In addition, dimensionality creates difficulties for both model validation and computation, and models are less fully prescribed by the general theory.

In two and more dimensions, there does not exist a simple distinction in three basic domains of attraction, and there is no reason for the univariate marginals of a multivariate distribution to share the same type of limiting Extreme Value probability law. This should be modelled using copulas, in which the marginals will no longer represent a problem.

A potential method for investigating bivariate data consists of studying the dependence function and the marginals separately. Because copulas describe and model the dependence structure between random variables, independently of the marginal laws involved, it is intuitive to use and develop a theory of copulas. This approach not only simplifies the analysis of the phenomenon under investigation but also provides the possibility of introducing new parameters for the characterization of the extreme behaviour of the system.

Furthermore, in a multidimensional environment, the issue of dependence between different variables plays a fundamental role. Indeed, quantifying dependence is a central theme in probabilistic and statistical methods for multivariate extreme values. An important activity within the present Ph.D. project is to investigate the possibility of extending the univariate ACER method to the bivariate case. A preliminary study conducted by Naess (2011) indicates that this is possible, which would represent a major achievement in the analysis of bivariate extreme value data. The bivariate ACER function would make it possible to represent, in a statistical sense, the exact bivariate extreme value distribution that is inherent in the data. In the general case, this has not previously been achieved.

1.3 Overview of the Thesis

The thesis is written in the following order:

Chapter 1 The overview of the problem of the univariate extreme value analysis, as well as bivariate extreme value analysis – a special case of the multivariate data analysis, is given in this Chapter. The outline of the thesis and summary of main contributions of the research are emphasized in this Chapter.

Chapter 2 The cornerstones of the classical univariate theory of extreme values are presented in Chapter 2. The most essential conclusions and derivations of the annual or block maximum method and the Peaks-Over-Threshold or POT method are mentioned in this Chapter. Particularly, the precise algorithms for practical application of both methods are presented. In addition, the accurate expressions

for a T-year return period level η^{Tyr} for each of the two approaches are derived in Chapter 2. This material lays the background for understanding of the univariate ACER method, which is discussed in Chapter 3, as well as for its straightforward application for comparison purposes.

Chapter 3 As previously mentioned, the major problem of extremes is an analysis of certain probability distributions with unknown parameters to be estimated based on the observed data. Although extreme value analysis has its peculiarities, it cannot be examined in an isolated manner. Therefore, the analysis of extreme values must be embedded in other approaches of main stream statistics, such as data analysis, nonparametric curve estimation, survival analysis, time series analysis, regression analysis, robust statistics and parametric inference.

The research activities presented in Chapter 3 of this Ph.D. project involve several tasks. The ACER method for univariate extreme value prediction is studied in more details and developed further. Particularly, the main emphasis is placed on the parameter estimation problems for the asymptotic distributions and on the related problems of optimization methods for non-linear least squares parametric curve fitting to the observed data. The improved optimization algorithm based on the constrained minimization for the specific problem is proposed and discussed in Section 3.4.1.

Chapter 3 presents a detailed comparison between T-year return levels predicted by three estimation approaches (ACER, block maxima and POT). The study is based on the Monte Carlo-simulated data, as well as on real environmental data sets. Methods for bootstrap 95% confidence intervals of the predicted extreme value are also discussed.

A notable achievement of the present research is the development and introduction of the robust and straightforward standalone application that implements the ACER method for practical applications (Karpa, 2012).

Finally, note that the considerable part of the material presented in Chapter 3 has been published in the journal papers, notably in Karpa and Naess (2013) and Naess, Gaidai, and Karpa (2013).

Chapter 4 This Chapter is focused on representations and modeling techniques for extremes of a bivariate process. Particularly, by analogy with the Fisher-Tippett-Gnedenko Theorem given in Chapter 2 for the univariate case, property of the limiting joint extreme value distribution for the bivariate case is described. In addition, Chapter 4 addresses the main concepts of the approach to modelling of bivariate extreme value distributions using copulas. These ideas appear to be vital for the bivariate ACER method presented in the next Chapter.

Chapter 5 A lack of data means that the precision of extreme value estimates is often poor. The way to overcome this problem is to incorporate additional information, suggesting the use of multivariate, particularly bivariate, models. Additionally, in many cases, it is of interest to be able to make predictions about extensional effects, but full random field modelling is impossible. In such cases,

bivariate extreme value distributions can be very helpful in making indicative predictions of extensional effects. Thus, Chapter 5 of the thesis is the conversion to the bivariate case and development of the model for the exact bivariate extreme value distribution that is inherent in the sampled data. In Chapter 5, the conditioning approach of the ACER method is developed for implementation in the bivariate framework. The possible functional representation for prediction purposes of the bivariate ACER surface is developed by virtue of the extreme value copulas. The optimization methods for non-linear least square problems for surface matching are also explored in Chapter 5.

A significant part of Chapter 5 addresses analyses of the overall performance of the bivariate ACER method by application to synthetic bivariate data, as well as to bivariate field measurements of oceanographic data.

The material presented in Chapter 5 has been partially published in the journal papers, that is in Naess and Karpa (2015a,b).

Chapter 6 This Chapter provides abstract conclusions and presents achievements of the research. It also lists possible directions for future research.

References Lists the references used in the thesis.

Chapter 2

Classical Univariate Extreme Value Theory

The classical univariate theory of extreme values was developed by Fréchet (1927), and Fisher and Tippett (1928). Gnedenko (1943) and Gumbel (1958b) showed that the largest value from a set of independently distributed random variables tends to an asymptotic distribution that only depends upon that of the basic variable. After standardization using suitable norming and centering constants, the limit distribution is shown to belong to one of three types, as pointed out by Gnedenko (1943) and de Haan (1976).

In this chapter, the distributions of the largest values of a given sample is considered and two different approaches are presented. We commence by considering the idea of max-stable distributions. Then, the limit distributions of maxima are calculated by using the annual block maximum method. The three types of extreme value distribution, notably, the Gumbel, the Fréchet, and the Weibull distributions are presented. Followed by this, the Generalized Extreme Value distribution, which covers all the three types of limit distributions, is derived. A useful tools and criteria for characterizing in a synthetic way the probabilistic structure of the extremes are also provided. In Section 2.2, the Peaks-Over-Threshold method is presented. In compliance with it, the extremes are studied by considering their exceedances over a given threshold. The Generalized Pareto distribution is derived. Estimating the T-year return period level is of particular interest for both approaches.

2.1 Annual Maximum Models

In this section, we focus on the statistical behaviour of the maximum value of a sequence of independent identically distributed random variables. The asymptotic behaviour of the distribution function of extreme values and its limiting form for large N are discussed.

2.1.1 Max-stability and Three Types of Extreme Value Distributions

Let us consider a sequence $X_1, \ldots, X_N$ of independent identically distributed random variables with the common distribution function $F_x(\eta)$. In practice, the X_i are typically realizations of a stochastic process $X(t)$ measured at equidistant discrete times, e.g., hours, days, etc. The key concept of the extreme value theory is the largest value of this sequence, which can be defined as follows:

$$\hat{X}_N = \max\{X_1, \ldots, X_N\}. \tag{2.1}$$

The objective is to analyse the statistical properties of the random variable $\hat{X}_N$. Theoretically, under i.i.d. assumption, the exact distribution function $F(\eta)$ of $\hat{X}_N$ is defined as:

$$\begin{aligned} F(\eta) = \text{Prob}\big(\hat{X}_N \leq \eta\big) &= \text{Prob}\big(X_N \leq \eta, \ldots, X_1 \leq \eta\big) \\ &\overset{i.i.d}{=} \prod_{i=1}^{N} \text{Prob}\big(X_i \leq \eta\big) = \big[F_x(\eta)\big]^N. \end{aligned} \tag{2.2}$$

Thus, the problem concerning extreme values in principle can be solved if the distribution $F_x(\eta)$ is known and the sample size N is given. Nevertheless, these conditions are not available in most practical cases. First, the sample size N might be not known if, e.g., the data are censored and some amount of information is lost. In addition, the underlaying distribution function $F_x(\eta)$ is very rarely known in practice. Clearly, one can use standard statistical methods for estimating $F_x(\eta)$ from the given data and then replace the theoretical one in Eq.(2.2) with this estimate. However, the uncertainty of relatively low significance in the estimate of $F_x(\eta)$ can lead to substantial inconsistency for $F_x^N(\eta)$ as N increases.

This implies that a reasonable alternative approach is to accept that $F_x(\eta)$ is unknown. Thereafter, it is essential to find distribution families $G(\eta)$ that approximate $F_x^N(\eta)$ asymptotically, which can only be estimated from the largest values $\hat{X}_N$ (Coles, 2001). A preliminary study of the behaviour of $F_x^N(\eta)$ is required for this purpose. It is necessary to take into account that for any level η smaller than the right endpoint $\omega(F_x) := \sup\{\eta : F_x(\eta) < 1\}$ of the support of F_x, the following holds:

$$\lim_{N\to\infty} F_x^N(\eta) = \begin{cases} 1, & \text{if } F_x(\eta) \equiv 1, \\ 0, & \text{otherwise,} \end{cases} \tag{2.3}$$

which is the limit distribution of maximum degenerates to a point mass on $\omega(F_x)$. The standard way to operate is to apply a linear transformation of the form:

$$\hat{X}_N^* = \frac{\hat{X}_N - a_N}{b_N}, \tag{2.4}$$

where $\{a_N : \forall N\ a_N \in \mathbb{R}\}$ and $\{b_N : \forall N\ b_N \in \mathbb{R}^+\}$ are the sequences of location and scaling constants, respectively, that are required to stabilize $\hat{X}_N$ as N increases and to ensure that the limit $\lim_{N\to\infty} F_x^N\big(a_N + b_N\,\eta\big)$ exists and is a non-degenerate distribution. This procedure (similar to the one used in deriving the

Central Limit Theorem) aims to identify a non-degenerate limit distribution after a suitable renormalization of the variables involved. In connection with this, an important definition can be given, cf., Salvadori et al. (2007).

Definition 2.1. ***(Maximum domain of attraction).*** The distribution F_x is said to belong to the maximum domain of attraction of the non-degenerate distribution G if sequences of constants $\{a_N\}$ and $\{b_N > 0\}$ exist such that:

$$\lim_{N\to\infty} F_x^N\left(a_N + b_N\,\eta\right) \;=\; G(\eta)\,. \tag{2.5}$$

It means that convergence occurs at the continuity points of the limiting function G, cf. Leadbetter et al. (1983).

The entire range of possible limit distributions $G(\eta)$ is given by The Extremal Types Theorem. This key result was discovered first by Fisher and Tippett (1928) and later proved in complete generality by Gnedenko (1943). Note that de Haan (1976) used a more recent, simple approach to prove this result.

Theorem 2.1. (Fisher-Tippett-Gnedenko). *Let $X_1, \ldots, X_N$ constitute a sample of independent and identically distributed random variables, and let $\hat{X}_N = \max\{X_1, \ldots, X_N\}$. If normalizing sequences $\{a_N\}$ and $\{b_N > 0\}$ exist such that for $\hat{X}_N^*$ given by Eq.* (2.4) *holds*

$$\lim_{N\to\infty} \mathit{Prob}\left\{\hat{X}_N^* \le \eta\right\} \;=\; G(\eta),\; \eta \in \mathbb{R}, \tag{2.6}$$

where $G(\eta)$ is a non-degenerate distribution, then G belongs to one of the following three types of limit (or asymptotic) distributions of maxima:

$$\textit{Gumbel (Type I)}: \quad G_0(\eta) \;=\; \exp\left[-\exp\left(-\frac{\eta-\mu}{\sigma}\right)\right],\; -\infty < \eta < \infty; \tag{2.7}$$

$$\textit{Fréchet (Type II)}: \quad G_{1,\alpha}(\eta) = \begin{cases} 0, & \eta \le \mu, \\ \exp\left[-\left(\frac{\eta-\mu}{\sigma}\right)^{-\alpha}\right], & \eta > \mu; \end{cases} \tag{2.8}$$

$$\textit{Weibull (Type III)}: \quad G_{2,\alpha}(\eta) = \begin{cases} \exp\left[-\left(-\frac{\eta-\mu}{\sigma}\right)^{\alpha}\right], & \eta < \mu, \\ 1, & \eta \ge \mu. \end{cases} \tag{2.9}$$

Here, $\mu \in \mathbb{R}$ is a location parameter, $\sigma \in \mathbb{R}^+$ is a scale parameter, and $\alpha \in \mathbb{R}^+$ is a shape parameter.

In this way, Theorem (2.1) states that the rescaled sample $\left\{\hat{X}_N\right\}$ converges in distribution to a variable having a distribution within one of the families of the extreme value distributions named the Gumbel, Fréchet and Weibull. The implication of the Theorem (2.1) means that if $\hat{X}_N$ can be stabilized with suitable sequences $\{a_N\}$ and $\{b_N\}$, the corresponding normalized variable $\hat{X}_N^*$ has a limiting distribution that must be one of the three types of extreme value distribution. The significance of this result is that irrespective of the underlying distribution

$F_x(\eta)$ of the $X(t)$, the extreme value $\hat{X}_N^*$ is governed in the limit exclusively by one of the three types of extreme value distributions. Thus, in this sense the theorem provides an extreme value analog of the Central Limit Theorem. In Table (2.1) some underlying distributions $F_x(\eta)$ frequently used in applications are listed are listed together with the corresponding limiting distributions of maxima (cf., e.g. Salvadori et al. (2007)).

Underlying distribution, $F_x(\eta)$	Distribution of maxima, $G(\eta)$
Gaussian	Gumbel
Exponential	Gumbel
Rayleigh	Gumbel
Lognormal	Gumbel
Gamma	Gumbel
Gumbel	Gumbel
Pareto	Fréchet
Cauchy	Fréchet
Fréchet	Fréchet
Uniform	Weibull
Weibull	Weibull

Table 2.1: Domain of attraction of maxima of some underlying distributions.

It is observed that the three types of asymptotic distributions in Theorem (2.1) have different forms of behaviour. For the Weibull distribution its right endpoint $\omega(G_{2,\alpha})$ is finite, whereas for both Gumbel and Fréchet distributions it is infinite. Clearly, the density of the Gumbel distribution decays exponentially, while the PDF of the Fréchet has a polynomial type of decrease. Different behaviour of the three types EVD is caused by the different forms of tail behaviour for the distribution function $F_x(\eta)$ of the X. In particular, if it is known or can be assumed with high certainty that the underlying probability distribution has a cumulative distribution function of the exponential-type, then the distribution of maxima converges to the Gumbel (Type I) distribution (Ochi, 1990). A more precise definition of the exponential-type distribution is given by von Mises as the distribution that satisfies the following condition:

$$\lim_{\eta\to\infty} \frac{d}{d\eta}\left[\frac{1-F_x(\eta)}{f_x(\eta)}\right] = 0\,. \tag{2.10}$$

The exponential type distribution is unlimited towards the extreme value and all moments exist.

The Fréchet extreme value distributions emerge from the initial distribution, which has an infinite right endpoint and has only a finite number of moments. Distributions that have this property are called Cauchy-type distributions. The Cauchy-type distribution must satisfy the condition

$$\lim_{\eta\to\infty} \{1-F_x(\eta)\}\cdot\eta^k = a, \tag{2.11}$$

where k, $a \in \mathbb{R}^+$ and moments of order less than k only exist. (Ochi, 1990)

The Type III extreme value distribution is related to underlying initial distributions that have either an upper or a lower limit boundary, that is, limited-type distributions.

Apparently, in applications of extreme value theory there might be certain difficulties associated with choosing the asymptotic distribution. First, the available dataset does not allow to make an accurate and robust decision on the appropriate extreme value distribution family. Even when such a decision is made, further propositions presume the choice to be correct, and do not allow to account for the uncertainty involved, even if the uncertainty appears to be significant. Therefore, the models in Theorem (2.1) can be unified into a single compact family of models termed the Generalized Extreme Value (GEV) distribution.

Theorem 2.2. (GEV distribution of maxima). *Under the conditions of Theorem 2.1, $G_\gamma(\eta)$ is a member of the GEV family of maxima given by, for $\left\{\eta : 1+\gamma \, \frac{\eta-\mu'}{\sigma'} > 0\right\}$,*

$$G_\gamma(\eta) = \exp\left\{-\left[1+\gamma\left(\frac{\eta-\mu'}{\sigma'}\right)\right]^{-1/\gamma}\right\}, \tag{2.12}$$

where $\mu' \in \mathbb{R}$ is a location parameter, $\sigma' \in \mathbb{R}^+$ is a scale parameter, and $\gamma \in \mathbb{R}$ is a shape parameter. The limit case $\gamma = 0$ yields the Gumbel (G_0) distribution, given by Eq. (2.7)

The Fréchet distribution is obtained for $\gamma > 0$, whereas the Weibull distribution corresponds to the case $\gamma < 0$ in this parameterization. The subset of the GEV family with $\gamma = 0$ is interpreted as the limit of (2.12) as $\gamma \to 0$, leading to the Gumbel family with distribution function (2.7). Coles (2001) asserts that the unification of the original three families of limiting distributions of maxima into a single family simplifies statistical implementation. The data themselves determine the most appropriate type of tail behavior via the introduced parameter γ. Moreover, uncertainty in the parameter γ can be considered to be a measure of the uncertainty involved in choosing one of the original asymptotic distribution as most appropriate for a given dataset. At the same time, fitting the GEV distribution is simply a curve fitting procedure, regardless of the available information about the nature of the given data.

Extreme value distributions are characterized by their max-stability.

Definition 2.2. ***(Max-stable distribution).*** The distribution G is called max-stable if there exist sequences of constants $\{a_N\}$ and $\{b_N > 0\}$ such that:

$$G^N\left(a_N + b_N\,\eta\right) \;=\; G(\eta)\,. \tag{2.13}$$

This definition implies that the relocated and rescaled maximum under the distribution function G is distributed according to G. For the standard Gumbel distribution the constants are $b_N = \log N$ and $a_N = 1$, for the standard Fréchet distribution $b_N = 0$ and $a_N = N^{1/\alpha}$. Finally, for the standard Weibull distribution

$b_N = 0$ and $a_N = N^{-1/\alpha}$. In general, a distribution is max-stable if, and only if, it belongs to the GEV family (Coles, 2001; Salvadori et al., 2007).

The following theorem provides a description of the maximum domain of attraction. According to von Mises (1936); Jenkinson (1955); Falk and Marohn (1993), these sufficient conditions are proven to be widely and easily applicable.

Theorem 2.3. *Suppose that $F_x(\eta)$ has a positive derivative $f_x(\eta)$ on $[\eta_0, \omega(F_x))$. Then, F_x belongs to the domain of attraction of each of the three types of extreme value distributions if the following are satisfied, respectively:*

$$\textit{Type I}: \qquad \lim_{\eta\to\infty} \frac{f_x(\eta)}{1-F_x(\eta)} = c, \ \textit{for some } c \in (0, \infty); \tag{2.14}$$

$$\textit{Type II}: \qquad \lim_{\eta\to\infty} \frac{\eta\, f_x(\eta)}{1-F_x(\eta)} = \alpha > 0; \tag{2.15}$$

$$\textit{Type III}: \qquad \lim_{\eta\uparrow\omega(F_X)} \frac{\big(\omega(F_x)-\eta\big) f_x(\eta)}{1-F_x(\eta)} = \alpha > 0. \tag{2.16}$$

The following theorem provides necessary and sufficient conditions for the distribution $F_x(\eta)$ to belong to a given domain of attraction for maxima, cf. Leadbetter et al. (1983). This criterion was introduced by Gnedenko (1943).

Theorem 2.4. (Max-domain of attraction for maxima). *The necessary and sufficient conditions for the distribution function $F_x(\eta)$ of the random variables of the independent and identically distributed sequence $\{X_j\}$ to belong to each of three types are as follows:*

Type I: *There exists some strictly positive function $g(\eta)$ such that*

$$\lim_{\eta\to\infty} \frac{1-F_x\big(\eta + t\, g(\eta)\big)}{1-F_x(\eta)} = \exp(-t), \ \textit{for all } t \in \mathbb{R}; \tag{2.17}$$

It may be shown that $\int_0^\infty \big(1-F_x(u)\big)\mathrm{d}u < \infty$ when the Type I limit holds, and one appropriate choice of g is given by $g(\eta) = \int_\eta^{\omega(F_X)} \big(1-F_x(u)\big)\mathrm{d}u / \big(1-F_x(\eta)\big)$ for $\eta < \omega(F_x)$;

$$\textit{Type II:} \quad \omega(F_x) = \infty \ \textit{and} \ \lim_{\eta\to\infty} \frac{1-F_x(t\,\eta)}{1-F_x(\eta)} = t^{-\alpha}, \ \forall t \in \mathbb{R}^+; \tag{2.18}$$

$$\textit{Type III:} \quad \omega(F_x) < \infty \ \textit{and} \ \lim_{h\downarrow 0} \frac{1-F_x\big(\omega(F_x)-t\,h\big)}{1-F_x\big(\omega(F_x)-h\big)} = t^{\alpha}, \ \forall t \in \mathbb{R}^+, \tag{2.19}$$

where $\alpha \in \mathbb{R}^+$ indicate the shape parameter of the Type II and Type III distributions, respectively.

As the corollary of Theorem 2.4, the norming constants a_N and b_N can be calculated as follows (Salvadori et al., 2007):

$$\text{Type I (Gumbel):} \quad a_N = F_X^{-1}\left(1 - \frac{1}{N}\right), \quad b_N = F_X^{-1}\left(1 - \frac{1}{N\,e}\right) - a_N\,; \tag{2.20}$$

$$\text{Type II (Fréchet):} \quad a_N = 0\,, \qquad b_N = F_X^{-1}\left(1 - \frac{1}{N}\right); \tag{2.21}$$

$$\text{Type III (Weibull):} \quad a_N = \omega(F)\,, \qquad b_N = a_N - F_X^{-1}\left(1 - \frac{1}{N}\right), \tag{2.22}$$

where F_X^{-1} is the quantile function associated with the distribution F_X.

2.1.2 Estimation of the T-year return level

Oceanographic data, as any environmental data, are usually collected using a daily or even more frequent (e.g., ten minutes, one hour, three hours, etc.) temporal resolution. For instance, consider a sample of maximum one-hour wind gust, or a sample of three-hour significant wave height. In practical applications, the interest is often focussed on the block (year or any other period, such as a month or season) maxima. Consequently, the independent observations $X_1, \ldots, X_N$ can be partitioned into k consecutive independent blocks of length m for some large value of m. k is the number of observed years (blocks), and m is generally chosen to correspond to the number of observations in a year or any other period. Then, the maximum observation is calculated for each block. This generates a sample of block maxima of size k, $\hat{X}_{m_1}, \ldots, \hat{X}_{m_k}$. This sample is then used to estimate the parameters of the extreme value distribution (the distribution of maxima). If blocks contain data observed during a time period of one year, then the block maxima are termed the annual maxima.

One of the major objectives of statistical analysis is estimating the high quantiles of the distribution function. Particularly, in common terminology of the extreme value theory, the concept of T-year return level is used. We introduce the T-year return level η^{Tyr} as the threshold such that the mean number of exceedances over η^{Tyr} within the time span of length T years (or any other period such as days, months or seasons) is equal to 1. In this context, note that there is one observation per each year (or other considered period), cf., Reiss and Thomas (2007)

Now, if we consider random variables $\hat{X}_1, \ldots, \hat{X}_T$ with the common distribution function G, then the T-year return level η^{Tyr} is the solution of the following equation:

$$\mathbb{E}\left(\sum_{i=1}^{T} \mathbf{1}\left(\hat{X}_i \geq \eta\right)\right) = 1\,. \tag{2.23}$$

Apparently,

$$\mathbb{E}\left(\sum_{i=1}^{T}\mathbf{1}\left(\hat{X}_i \geq \eta\right)\right) = \sum_{i=1}^{T}\text{Prob}\left(\hat{X}_i \geq \eta\right) = T\left(1 - G(\eta)\right). \tag{2.24}$$

Therefore,

$$\eta^{\text{Tyr}} = G^{-1}(1 - 1/T). \tag{2.25}$$

Thus, η^{Tyr} is the $(1-1/T)$ quantile of the distribution G. In other words, level η^{Tyr} is exceeded by the annual (block) maximum in any particular year (block) with probability $1/T$. Therefore, in case of the Generalized extreme value distribution, estimates of extreme quantiles of GEV are obtained by inverting Eq. (2.12):

$$\eta^{\text{Tyr}} = \begin{cases} \mu - \frac{\sigma}{\gamma}\left\{1 - [-\log(1 - 1/T)]^{-\gamma}\right\}, & \text{for } \gamma \neq 0, \\ \mu - \sigma\log\left\{-\log(1 - 1/T)\right\}, & \text{for } \gamma = 0; \end{cases} \tag{2.26}$$

Note that plot of η^{Tyr} versus $-\log(1 - 1/T)$ on a logarithmic scale is a return level plot. It exhibits linear behaviour in the case $\gamma = 0$, the plot is convex with asymptotic limit as $T \to \infty$ at $\mu - \sigma/\gamma$ for $\gamma < 0$, whereas for the case $\gamma > 0$ the plot is concave and has no finite bound (cf. Coles (2001)). Because of the simplicity of interpretation and construction of the graph, the return level plot is particularly convenient as a diagnostic tool.

2.2 Threshold Models

The approach described in previous Section 2.1 represents the standard procedure to analyze the extreme values. Theorems (2.1) and (2.2) provide a model for the distribution of block maxima. In application, the data are sectioned into blocks of equal length and the extreme value distribution is fitted to the set of block maxima. However, the approach has some drawbacks. First of all, the choice of block size generates uncertainty caused by significant bias or large variance. Blocks of too small size mean that approximation by the asymptotic distribution is likely to be poor, leading to bias in estimation and extrapolation. Large blocks, on the other hand, generate few block maxima, leading to large estimation variance. But above all, the block maxima method discards important sample information. In fact, the extremal behavior of a phenomenon usually lasts for some period in a given block. In this way, the most extreme event, as a rule, occurs in a cluster of neighboring rough events. Hoverer, only the most extreme observation will generate the annual maximum. As a consequence, all the remaining information concerning the extremal dynamics developed during the preceding and succeeding periods will be discarded. In this section, an alternative approach to the analysis of this type of extremal behavior is given. The maxima is analyzed via the Peaks-Over-Threshold (POT) method, which avoid the procedure of blocking. According to Balkema and de Haan (1974); Pickands (1975), the POT method is based on the so-called Generalized Pareto (GP) distribution.

2.2.1 The Generalized Pareto distribution of maxima

As before, let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with the common distribution function $F_X(\eta)$. In practice, the X_i are realizations of a stochastic process $X(t)$. It is natural to accept that the extremal behaviour of the process can be modeled by those events that exceed some high threshold u, that is $\{X_i > u\}$. In other words, a description of the stochastic behavior of extreme events is given by the excess values above a given threshold. It can be described by the conditional probability $F_u(y) := \mathrm{Prob}\big(X - u \le y \,|\, X > u\big)$:

$$1 - F_u(y) \;=\; \mathrm{Prob}\big(X > u + y \,|\, X > u\big) \;=\; \frac{1 - F_X(u+y)}{1 - F_X(u)}\,,\; y > 0\,. \tag{2.27}$$

Because in most practical applications, the underlying distribution $F_X(\eta)$ is unknown, so is the probability in (2.27). Therefore, a natural way to proceed is to approximate the conditional law (2.27) by a distribution that would be independent of the distribution F_X. The following Pickands-Balkema-de Haan theorem (Balkema and de Haan, 1974; Pickands, 1975) is fundamental in the analysis of maxima using the POT method, and implicitly defines the Generalized Pareto distribution.

Theorem 2.5. (Pickands-Balkema-de Haan). *Let $X_1, \ldots, X_N$ be a sample of independent and identically distributed random variables governed by distribution F_X. If F_X satisfies the conditions of Theorem (2.2) (and, hence, Theorem (2.1)), then, for $u \gg 1$ the conditional distribution $F_u(y)$ of the exceedances $(X - u)$ can be approximated as*

$$\lim_{u\to\infty} F_u(y) \;=\; W_\gamma(y) \;:=\; 1 - \Big(1 + \gamma\,\frac{y}{\tilde\sigma}\Big)^{-1/\gamma}\,,\; y > 0\,, \tag{2.28}$$

for $(1 + \gamma y/\tilde\sigma) > 0$. Here $\tilde\sigma > 0$,

$$\tilde\sigma = \sigma' + \gamma(u - \mu') \tag{2.29}$$

and γ $(-\infty < \gamma < \infty)$ are, respectively, scale and shape parameters; μ' and σ' are location and scale parameters of the GEV distribution G_γ given by Eq. (2.12) in Theorem (2.2). The shape parameter γ is equal to that of the corresponding GEV distribution G_γ in (2.12). In the limit case $\gamma = 0$, $W_\gamma(y)$ reduces to the Exponential distribution with parameter $1/\tilde\sigma$

$$W_0(y) \;:=\; 1 - \exp\Big\{-\frac{y}{\tilde\sigma}\Big\}\,,\; y > 0\,, \tag{2.30}$$

The distribution defined in Eq. (2.28) (and Eq. (2.30) in case $\gamma = 0$) is called the Generalized Pareto (GP) distribution (in the canonical form).

The asymptotic result (2.28) implies that the Generalized Pareto distribution can be used to represent the conditional cumulative distribution function of the excess $y = X - u$ of the observed variate X over the threshold u, given that $X > u$ for sufficiently large u. The assumption of a Poisson process model for the exceedance times combined with GP distributed excesses can be shown to lead

to the generalized extreme value (GEV) distribution for corresponding extremes (Leadbetter et al., 1983; Leadbetter, 1995). In this way, Balkema and de Haan (1974); Pickands (1975) have shown that, asymptotically, the excess values above a high level will follow a GP distribution if, and only if, the parent distribution belongs to the domain of attraction of one of the extreme value distributions. Moreover, the parameters of the GP distribution of threshold excesses $W_\gamma(y)$ are uniquely determined by those of the associated GEV distribution of block maxima (2.12). Note that values of the GEV parameters are generally speaking dependent on the block size N, whereas the calculation of the scale parameter $\tilde{\sigma} = \sigma' + \gamma(u - \mu')$ is unperturbed by the changes in μ' and σ' which are self-compensating, and the shape parameter of the GP distribution γ is invariant to block size. Thus, the shape parameter γ is dominant in determining the qualitative behaviour of the generalized Pareto distribution (Coles, 2001). The cases $\gamma > 0$, $\gamma = 0$ and $\gamma < 0$ correspond to Fréchet (Type II) (2.8), Gumbel (Type I) (2.7), and Weibull (Type III) (2.9) domains of attraction, respectively, cf., Section 2.1.1.

Interestingly enough, there is the simple analytical relationship between Generalized Pareto distribution W and extreme value distribution G (cf., e.g. Reiss and Thomas (2007)):

$$W(y) \;=\; 1 + \log\, G(y)\,, \quad \text{if } \log\, G(y) > -1\,. \tag{2.31}$$

Thus, the duality between the extreme value distributions (see page 9) and the GP families can be described schematically as follows (Reiss and Thomas, 2007; Salvadori et al., 2007):

$$\text{Gumbel, } G_0 \quad \Leftrightarrow \quad \text{Exponential, } W_0(y) \;=\; 1 - \exp\left\{-\frac{y-\mu}{\sigma}\right\},\; y \geq \mu\,; \tag{2.32}$$

$$\text{Fréchet, } G_{1,\alpha} \quad \Leftrightarrow \quad \text{Pareto, } W_{1,\alpha}(y) \;=\; 1 - \left(\frac{y-\mu}{\sigma}\right)^{-\alpha},\; y \geq \mu + \sigma\,; \tag{2.33}$$

$$\text{Weibull, } G_{2,\alpha} \quad \Leftrightarrow \quad \text{Beta, } W_{2,\alpha}(y) \;=\; 1 - \left(-\frac{y-\mu}{\sigma}\right)^{\alpha},\; \mu - \sigma \leq y \leq \mu, \tag{2.34}$$

where, $\mu \in \mathbb{R}$ is a location parameter, $\sigma \in \mathbb{R}^+$ is a scale parameter, and $\alpha \in \mathbb{R}^+$ is a shape parameter. Note that the Exponential distribution function $W_0(y)$ is equal to zero for $y < \mu$; the Pareto distribution is equal to zero for $y < \mu + \sigma$; the Beta distribution (properly reparametrized) is equal to zero for $y < \mu - \sigma$ and equal to 1 for $y > \mu$.

2.2.2 Selection of the threshold

From Theorem (2.5) the following algorithm for extreme value modeling emerges. It is assumed that the initial unprocessed observed data $X_1, \ldots, X_N$ are independent and identically distributed measurements of a process $X(t)$ governed by $F_X(\eta)$. In addition, the POT method is build on the property of the data, according to which peaks over the selected threshold should occur randomly in time according to an approximate Poisson process (Leadbetter, 1983; Leadbetter et al., 1989). In

practice, however, extreme events of $X(t)$ often are dependent and have a tendency to cluster, which violates the Poisson assumption. Therefore, some declustering algorithm should be performed in advance to select the largest value in each of the clusters, and then use a Poisson distribution for the number of clusters. The selected observations should be separated by amount of time sufficient for the exceedances to be independent, see, for example, Reiss and Thomas (2007).

After the sample has been processed, its extremes $X_{(1)}, \ldots, X_{(k(u))}$ are defined as those observations that exceed the selected high threshold u: $\forall\, i\ X_{(i)} > u$, where $k(u)$ is the number of observations that exceed u. Then, the threshold excesses are calculated as $y_i = X_{(i)} - u$, for $i = 1, \ldots, k(u)$. According to Theorem (2.5), the y_i can be regarded as independent realizations of a random variable whose distribution can be approximated by a generalized Pareto distribution $W_\gamma(y)$. This implies fitting a GP distribution to the obtained excesses y_i, followed by model verification and extrapolation.

The POT method allows to analyse considerably more data and, hence, more information about the extremal behaviour of the process than the block maxima approach. Unlike the latter, the first method interprets to be extreme those observations which exceed a high threshold. The challenge of choosing the threshold is analogous to the choice of block size in the block maxima approach, implying a balance between bias and variance. In this case, too low threshold level yields bigger amount of exceedances and reduces the variance. At the same time, it is more likely to violate the asymptotic basis of the model ($u \to \infty$), leading to bias. Too high threshold level will generate only a few excesses, so that the variance of the fitted GP model will be high (Coles, 2001). There are two standard methods to define as low as possible threshold that can provide a reasonable approximation. One consist in assessment of the stability of parameter estimates based on the fitting of model across a range of different thresholds. The other method consists in exploration of the data prior to model estimation.

The second method is based on the properties of the generalized Pareto distribution. If Y is governed by a GP distribution $W_\gamma(y)$ (2.28) with parameters σ and γ, then the mean value $\mathbb{E}(Y)$ is

$$\mathbb{E}(Y) = \begin{cases} \frac{\sigma}{1-\gamma}\,, & \gamma < 1\,, \\ \infty\,, & \gamma \geq 1; \end{cases} \tag{2.35}$$

Under the assumption that the generalized Pareto distribution is a valid model of the excesses of a threshold u_0 generated by the observations of $X(t)$, Eq. (2.35) implies:

$$\mathbb{E}(X - u_0 \,|\, X > u_0) = \frac{\sigma_{u_0}}{1-\gamma}\,,\ \gamma < 1\,, \tag{2.36}$$

where σ_{u_0} denote the scale parameter of the model that depends on the threshold u_0, see Eq. (2.29). If the generalized Pareto distribution is a valid model for the excesses of the threshold u_0, then the model should also be true for all thresholds $u > u_0$. In this case, due to (2.29), it can easily be verified that $\sigma_u = \sigma_{u_0} + \gamma\,(u - u_0)$ (note that the shape parameter γ remains independent of u the GPD is valid).

Hence,

$$\mathbb{E}(X - u \,|\, X > u) \;=\; \frac{\sigma_{u_0} + \gamma\,(u - u_0)}{1 - \gamma}\,,\; \gamma < 1\,. \tag{2.37}$$

Note that for the given sample $X_1, \ldots, X_N$, $\mathbb{E}(X - u \,|\, X > u)$ denotes the mean of the excesses of the threshold u. Therefore, Eq. (2.37) means that if the generalized Pareto distribution is true model, then the mean of the excesses of the threshold u is a linear function of u. The sample mean of the excesses of u provides the empirical estimates of $\mathbb{E}(X - u \,|\, X > u)$ for each value of u. The estimates are expected to change linearly with u, at those levels of u for which the GP distribution is valid. This leads to the following practical procedure. First, the R-element grid of u-levels as linearly spaced values between $X_{\min}$ and $X_{\max}$ is generated. The exceedances $X_{(1)}, \ldots, X_{(k(u_r))}$ are defined for each element u_r of the grid for $r = 1, \ldots, R$. The sample mean is estimated:

$$\hat{m}(u_r) \;=\; \frac{1}{k(u_r)} \sum_{i=1}^{k(u_r)} \big(X_{(i)} - u_r\big)\,,\; r = 1, \ldots, R\,. \tag{2.38}$$

Plot $\big\{\big(u_r,\, \hat{m}(u_r)\big),\, r = 1, \ldots, R\big\}$ is called the mean residual life plot. Above a threshold u_0 at which the generalized Pareto distribution becomes a valid model for the distribution of excesses, the mean residual life plot should exhibit approximately linear behaviour in u. Confidence intervals can be added to the plot based on the approximate normality of sample means (Coles, 2001).

2.2.3 T-year return period levels

The T-year return period level of a given process, in years, is defined as the inverse of the probability that the specified value η^{Tyr} will be exceeded once every T years.

Suppose that a generalized Pareto distribution $W_\gamma(y)$ (2.28) with parameters σ and γ is a reliable model for exceedances of a threshold u by the observations $X_1, \ldots, X_N$ of a process $X(t)$. Then, from (2.27) and (2.27) it follows that for some $x = (y + u) > u$

$$\mathrm{Prob}\big(X > x \,|\, X > u\big) \;=\; \Big(1 + \gamma\,\frac{x - u}{\sigma}\Big)^{-1/\gamma}\,. \tag{2.39}$$

If λ_u denotes the probability of X exceeding the threshold u, that is, $\lambda_u = \mathrm{Prob}\,(X > u)$, then

$$\mathrm{Prob}\,(X > x) \;=\; \lambda_u \Big(1 + \gamma\,\frac{x - u}{\sigma}\Big)^{-1/\gamma}\,. \tag{2.40}$$

Now, it is assumed that N measurements $X_1, \ldots, X_N$ of a process $X(t)$ were taken during n_y observation years. In this way, it implies the presumption that there are NT/n_y observations during T years. Hence, the level η^{Tyr} that is exceeded on average once in T years, i.e., once every NT/n_y observations is the solution of

$$\lambda_u \Big(1 + \gamma\,\frac{\eta^{\mathrm{Tyr}} - u}{\sigma}\Big)^{-1/\gamma} \;=\; \frac{n_y}{NT}\,. \tag{2.41}$$

Rearranging,

$$\eta^{\text{Tyr}} = u + \frac{\sigma}{\gamma}\left[\left(\frac{\lambda_u NT}{n_y}\right)^{\gamma} - 1\right], \tag{2.42}$$

unless $\gamma = 0$, in which case

$$\eta^{\text{Tyr}} = u + \sigma \log\left(\frac{\lambda_u NT}{n_y}\right). \tag{2.43}$$

Estimation of return levels requires the substitution of parameter values by their estimates. Parameters σ and γ are substituted by the corresponding maximum likelihood estimates. An estimate of the probability of an individual observation exceeding the threshold u, λ_u, can be done in a natural way by introducing the mean exceedance rate of the threshold u

$$\hat{\lambda}_u = \frac{k(u)}{N},$$

where N is the total number of observations, whereas $k(u)$ is the number of observations that exceed u. Note that because the number of exceedances of u follows the binomial distribution $\mathcal{B}(N, \lambda_u)$, $\hat{\lambda}_u$ is also the maximum likelihood estimate of λ_u. Therefore, Eqs.(2.42) and (2.43) can be modified as follows:

$$\eta^{\text{Tyr}} = \begin{cases} u + \frac{\sigma}{\gamma}\left[\left(\frac{k(u)T}{n_y}\right)^{\gamma} - 1\right], & \text{for } \gamma \neq 0\,; \\ u + \sigma \log\left(\frac{k(u)T}{n_y}\right), & \text{for } \gamma = 0\,. \end{cases} \tag{2.44}$$

Chapter 3

The Average Conditional Exceedance Rates (ACER) Method. Univariate Case

3.1 Problem Statement

Let us consider a stochastic process $X(t)$ that has been observed over a time interval such as $\big(0,\ \mathcal{T}\big)$. A finite number of measured values $X_1, \ldots, X_N$, which have been derived from the observed process, are at our disposal. Assume that the data series X_j, $j = 1, \ldots, N$ is allocated to the discrete times $t_1, \ldots, t_N$ in $\big(0,\ \mathcal{T}\big)$. This could simply be the observed values of $X(t)$ at each t_j, $j = 1, \ldots, N$, or it could be average values or peak values over smaller time intervals centred at the t_js. The object of our interest is the extreme value $\hat{X}_N = \max\{X_j; j = 1, \ldots, N\}$. Our goal is to accurately determine the distribution function $F(\eta) = \text{Prob}\big(\hat{X}_N \leq \eta\big)$ $= \text{Prob}\big(X_N \leq \eta, \ldots, X_1 \leq \eta\big)$ of the $\hat{X}_N$. Specifically, we want to accurately estimate $F(\eta)$ for large values of η.

An underlying premise for the development here is that a rational approach to study the extreme values of the sampled time series is to consider exceedances of the individual random variables X_j above given thresholds, as in the POT method. Another appropriate concept is to analyse the highest observations in each independent group of exceedances above a threshold, as required by the Method of Independent Storms (Cook, 1982), or simply to consider a block maxima as in classical extreme value theory. The alternative approach of considering the exceedances by upcrossing of given thresholds by a continuous time stochastic process was developed by Naess et al. (2007); Naess and Gaidai (2008).

The latter approach originally derived by Naess and Gaidai (2009) would be the appropriate way to address the recorded data time series of environmental loads, such as wind speeds, wave heights observed at a given location, etc.

3.2 Cascade of Conditioning Approximations

In the following, we outline the principle behind a cascade of approximations based on conditioning, where the first approximation is a one-step memory approximation and thus carries some resemblance to a Markov chain approximation. However, it is emphasized that this first approximation is not equivalent to such an approximation (Naess and Moan, 2012; Naess and Gaidai, 2008, 2009).

As discussed above, the joint distribution function $F(\eta)$ cannot generally be estimated directly from the data. However, by introducing a cascade computable sequence of conditioning approximations $F_k(\eta)$ of $F(\eta)$, such that $F_k(\eta) \to F(\eta)$ as k increases, this problem can be solved in practice in a very efficient and elegant manner. The definition of $F_k(\eta)$ is as follows.

From the definitions of $F(\eta)$ and $\hat{X}_N$, we obtain

$$\begin{aligned} F(\eta) &= \text{Prob}\big(\hat{X}_N \le \eta\big) = \text{Prob}\big(X_N \le \eta, \ldots, X_1 \le \eta\big) \\ &= \text{Prob}\big(X_N \le \eta | X_{N-1} \le \eta, \ldots, X_1 \le \eta\big) \\ &\qquad \cdot \text{Prob}\big(X_{N-1} \le \eta, \ldots, X_1 \le \eta\big) \\ &= \prod_{j=2}^{N} \text{Prob}\big(X_j \le \eta \,| X_{j-1} \le \eta, \ldots, X_1 \le \eta\big) \cdot \text{Prob}\big(X_1 \le \eta\big) \end{aligned} \tag{3.1}$$

Initially, we assume that all variables X_j are statistically independent. This implies that the first approximation of the cascade is obtained as follows:

$$\begin{aligned} F(\eta) &= \prod_{j=1}^{N} \text{Prob}\big(X_j \le \eta\big) = \prod_{j=1}^{N} \big(1 - \alpha_{1j}(\eta)\big) \\ &\approx F_1(\eta) := \exp\left(-\sum_{j=1}^{N} \alpha_{1j}(\eta)\right), \ \eta \to \infty, \end{aligned} \tag{3.2}$$

where $\alpha_{1j}(\eta) := \text{Prob}\big(X_j > \eta\big)$, $j = 1, \ldots, N$, and $F_1(\eta)$ is defined by the last equality of Eq. (3.2); the assignment symbol $:=$ means "by definition". Note that the approximation $\exp(-x) \approx 1 - x$ is accurate to within .5% for values of $|x|$ as high as 0.1, and the accuracy rapidly increases for decreasing values of $|x|$. This in turn justifies its application here because the probability to exceed a high level η tends to zero.

In general, the variables X_j are statistically dependent. In such a case, a one-step memory conditioning will measurably account, to a certain extent, for dependence between the X_js. Therefore, the next approximation is obtained by assuming that

$$\text{Prob}\big(X_j \le \eta | X_{j-1} \le \eta, \ldots, X_1 \le \eta\big) \approx \text{Prob}\big(X_j \le \eta | X_{j-1} \le \eta\big), \tag{3.3}$$

for $j = 2, \ldots, N$. Therefore, by this approximation,

$$\begin{aligned} F(\eta) &\approx \prod_{j=2}^{N} \mathrm{Prob}\big(X_j \le \eta \,|\, X_{j-1} \le \eta\big) \cdot \mathrm{Prob}\big(X_1 \le \eta\big) \\ &= \prod_{j=2}^{N} \big(1 - \alpha_{2j}(\eta)\big) \cdot \big(1 - \alpha_{11}(\eta)\big) \\ &\approx F_2(\eta) \coloneqq \exp\left(-\sum_{j=2}^{N} \alpha_{2j}(\eta) - \alpha_{11}(\eta)\right), \\ &\qquad\qquad \eta \to \infty, \end{aligned} \tag{3.4}$$

where we have introduced the notation $\alpha_{2j}(\eta) \coloneqq \mathrm{Prob}\big(X_j > \eta \,|\, X_{j-1} \le \eta\big)$ for $2 \le j \le N$.

In the next step, through conditioning on one additional data point, the third level of approximation is achieved. It is assumed that

$$\begin{aligned} &\mathrm{Prob}\big(X_j \le \eta | X_{j-1} \le \eta, \ldots, X_1 \le \eta\big) \\ &\approx \mathrm{Prob}\big(X_j \le \eta | X_{j-1} \le \eta, X_{j-2} \le \eta\big) \end{aligned} \tag{3.5}$$

for $j = 3, \ldots, N$. By adopting this approximation in Eq (3.1), it is obtained that

$$\begin{aligned} F(\eta) &\approx \prod_{j=3}^{N} \mathrm{Prob}\big(X_j \le \eta | X_{j-1} \le \eta, X_{j-2} \le \eta\big) \\ &\qquad \cdot \mathrm{Prob}\big(X_2 \le \eta \,|\, X_1 \le \eta\big) \cdot \mathrm{Prob}\big(X_1 \le \eta\big) \\ &= \prod_{j=3}^{N} \big(1 - \alpha_{3j}(\eta)\big) \cdot \big(1 - \alpha_{22}(\eta)\big) \cdot \big(1 - \alpha_{11}(\eta)\big) \\ &\approx F_3(\eta) \coloneqq \exp\left(-\sum_{j=3}^{N} \alpha_{3j}(\eta) - \alpha_{22}(\eta) - \alpha_{11}(\eta)\right), \\ &\qquad\qquad \eta \to \infty\,, \end{aligned} \tag{3.6}$$

where we have introduced the notation

$$\alpha_{3j}(\eta) \coloneqq \mathrm{Prob}\big(X_j > \eta \,|\, X_{j-1} \le \eta, X_{j-2} \le \eta\big),\ 3 \le j \le N\,.$$

It is realized that by continuing this conditioning process, a general kth approximation will be obtained. By introducing the notation

$$\alpha_{kj}(\eta) \coloneqq \mathrm{Prob}\big(X_j > \eta \,|\, X_{j-1} \le \eta, \ldots, X_{j-k+1} \le \eta\big),\ 1 \le k \le j \le N.$$

Evidently, $\alpha_{kj}(\eta)$ denotes the probability of exceedance of X_j conditioned on $k-1$ immediately preceding non-exceedances. Therefore, the general approximation will assume the form ($k \ge 2$),

$$\begin{aligned} F(\eta) \approx F_k(\eta) \coloneqq \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta) - \sum_{j=1}^{k-1} \alpha_{jj}(\eta)\right), \\ \eta \to \infty, \end{aligned} \tag{3.7}$$

It should be emphasized that the one-step memory approximation adopted in Eqs. (3.3) and (3.4) is not a Markov chain approximation, cf. Smith (1992); Coles (1994); Smith et al. (1997); furthermore, the k-step memory approximations in Eq. (3.7) do not lead to kth-order Markov chains either (Yun, 1998, 2000). An effort to relinquish the Markov chain assumption to obtain an approximate distribution of clusters of extremes was reported by Segers (2005).

Summarizing the cascade, the following has been obtained for $\eta \to \infty$, cf. Karpa and Naess (2013):

$$
\begin{aligned}
F_1(\eta) &= \exp\left(-\sum_{j=1}^{N} \alpha_{1j}(\eta)\right);\\
F_2(\eta) &= \exp\left(-\sum_{j=2}^{N} \alpha_{2j}(\eta) - \alpha_{11}(\eta)\right);\\
F_3(\eta) &= \exp\left(-\sum_{j=3}^{N} \alpha_{3j}(\eta) - \alpha_{22}(\eta) - \alpha_{11}(\eta)\right);\\
&\vdots\\
F_k(\eta) &= \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta) - \sum_{j=1}^{k-1} \alpha_{jj}(\eta)\right);\\
&\downarrow k \text{ increases}\\
F_N(\eta) &\equiv F(\eta) = \mathrm{Prob}(X_N \le \eta, \ldots, X_1 \le \eta)\,.
\end{aligned}
\tag{3.8}
$$

Thus, we have constructed a sequence or cascade $\{F_k(\eta)\}_{k=1}^{N}$ of conditional probability distributions that has the target distribution function $F(\eta)$ of the extreme value $\hat{X}_N$ as the limit. The process of constructing the sequence and its properties follow from the definition of the extreme value distribution $F(\eta)$ and the properties of conditional probability.

It is natural to assume that the degree of dependence between observed values $\{X_j\}$ of the time series $X(t)$ separated by k lags decreases as $k \uparrow N$ and is equal to zero for $k = N$. Therefore, for this cascade of approximations to have practical significance, it is implicitly assumed that there is a k_e satisfying $k_e \ll N$ such that effectively, $F(\eta) = F_{k_e}(\eta)$. Then, $F_1(\eta) \le F_2(\eta) \le \ldots \le F_{k_e}(\eta) = F(\eta)$. Note that for k-dependent stationary data sequences, that is, for data where X_i and X_j are independent whenever $|j - i| > k$, then $F(\eta) = F_{k+1}(\eta)$ exactly and $\lim\limits_{N\to\infty} F_1(\eta) = \lim\limits_{N\to\infty} F(\eta)$ (Watson, 1954). In fact, it can be shown that $\lim\limits_{N\to\infty} F_1(\eta) = \lim\limits_{N\to\infty} F(\eta)$ is true for conditions weaker than k-dependence (Leadbetter et al., 1983). However, for finite values of N, the picture is considerably more complex, and purely asymptotic results should be used with some caution. Cartwright (1958) used the notion of k-dependence to investigate the effect on extremes of correlation in sea wave data time series.

It will be verified that the property $k_e \ll N$ is indeed satisfied for the type of data analysed in the present research. Furthermore, under this assumption, for

all k values of interest, $k \ll N$, implying that $\sum_{j=1}^{k-1} \alpha_{jj}(\eta)$ is generally negligible in comparison with $\sum_{j=k}^{N} \alpha_{kj}(\eta)$. This leads to the approximation, which is applicable to both stationary and non-stationary data,

$$F(\eta) \approx F_k(\eta) \approx \exp\left(-\sum_{j=k}^{N} \alpha_{kj}(\eta)\right), \; k = 1, 2, \ldots, \tag{3.9}$$

Therefore, estimation of the extreme value distribution using the described conditioning approach reduces to estimation of the set of $\alpha_{kj}(\eta)$ functions (Naess and Gaidai, 2009).

It is important to note that under the assumption that all of the variables X_j are statistically independent, it follows from the definition of $\alpha_{1j}(\eta)$ that $\sum_{j=1}^{N} \alpha_{1j}(\eta)$ is equal to the expected number of exceedances of the threshold η during the time interval $(0, \mathcal{T})$. Thus, Eq. (3.2) expresses the assumption that the sequence of exceedance events follow a Poisson distribution. Consequently, this clarifies the essence of Eq. (3.7) by interpreting the expressions $\sum_{j=k}^{N} \alpha_{kj}(\eta) + \sum_{j=1}^{k-1} \alpha_{jj}(\eta) \approx \sum_{j=k}^{N} \alpha_{kj}(\eta)$ as the expected effective number of exceedances subjected to $k-1$ immediately preceding non-exceedances.

3.3 Estimation of the ACER

The concept of the average conditional exceedance rate (ACER) of order k is now introduced as follows:

$$\varepsilon_k(\eta) \; = \; \frac{1}{N-k+1} \sum_{j=k}^{N} \alpha_{kj}(\eta) \,, \; k = 1, 2, \ldots \tag{3.10}$$

By combining Eqs. (3.10) and (3.9), it is instantly obtained that

$$F(\eta) \approx F_k(\eta) \approx \exp\Big(- (N-k+1)\, \varepsilon_k(\eta) \Big) \,. \tag{3.11}$$

In practice, there are typically two scenarios for the underlying process $X(t)$. We may consider it to be either a stationary process or even an ergodic process. The other alternative is to view $X(t)$ as a process that depends on certain parameters whose variation in time may be modelled as an ergodic process in its own right. For each set of parameter values, the premise is that $X(t)$ can be modelled as an ergodic process. This is the scenario that can be used to model the long-term statistics of wind speed (Naess, 1984).

For both of these scenarios, the empirical estimation of the ACER function $\varepsilon_k(\eta)$ proceeds in a completely analogous manner by counting the total number of

favourable incidents, that is, exceedances conditional on the requisite number of preceding non-exceedances, for the total data time series and then finally dividing by $N - k + 1 \approx N$. Note that the expressions given thus far also apply to the case of non-stationary time series.

The numerical estimation of the ACER functions is based on counting the requisite events. First, it starts by defining random functions:

$$\begin{aligned} A_{jk}(\eta) &= \mathbf{1}\{X_j > \eta,\, X_{j-1} \leq \eta,\, ...,\, X_{j-k+1} \leq \eta\} \\ B_{jk}(\eta) &= \mathbf{1}\{X_{j-1} \leq \eta,\, ...,\, X_{j-k+1} \leq \eta\}\,, \end{aligned} \tag{3.12}$$

for $k \geq 2$, $j = k, \ldots, N$, where $\mathbf{1}\{\mathcal{A}\}$ denotes the indicator function of some event $\mathcal{A}$. Then,

$$\alpha_{jk}(\eta) \;=\; \frac{\mathbb{E}[A_{kj}(\eta)]}{\mathbb{E}[B_{kj}(\eta)]},\; k \geq 2,\; j = k, \ldots, N\,, \tag{3.13}$$

where $\mathbb{E}[\cdot]$ is the expectation operator. Under the assumption that the process $X(t)$ is ergodic, we have $\varepsilon_k(\eta) = \alpha_{kk}(\eta) = \ldots = \alpha_{kN}(\eta)$. Then, by replacing ensemble means with corresponding time averages, it may be assumed that for the time series at hand,

$$\varepsilon_k(\eta) \;=\; \lim_{N\to\infty} \frac{\sum\limits_{j=k}^{N} \mathrm{a}_{kj}(\eta)}{\sum\limits_{j=k}^{N} \mathrm{b}_{kj}(\eta)}\,. \tag{3.14}$$

We use $\mathrm{a}_{kj}(\eta)$ and $\mathrm{b}_{kj}(\eta)$ to designate the realizations of the random functions $A_{kj}(\eta)$ and $B_{kj}(\eta)$, respectively, for the observed time series. Then, the estimate $\hat{\varepsilon}_k(\eta)$ of the ACER function $\varepsilon_k(\eta)$ can be expressed as:

$$\hat{\varepsilon}_k(\eta) \;=\; \frac{\sum\limits_{j=k}^{N} \mathrm{a}_{kj}(\eta)}{\sum\limits_{j=k}^{N} \mathrm{b}_{kj}(\eta)}\,. \tag{3.15}$$

Figure (3.1) illustrates the process of counting the exceedances. In this figure, the black points indicate non-exceedances of the level $\eta = 1.5$. These points are taken into account for $\mathrm{b}_{kj}(\eta)$. All of the observations marked by circles ($\circ$) denote exceedances considered for $\mathrm{a}_{1j}(\eta)$. The exceedances that have an additional mark in the form of a star ($*$) within a rectangle are considered for $\mathrm{a}_{2j}(\eta)$. To count the exceedances with two immediately preceding non-exceedances, that is $\mathrm{a}_{3j}(\eta)$, only observations with ($\square$) marks are used. Finally, three dots ($\cdots$) indicate data points that should be used to count $\mathrm{a}_{kj}(\eta)$ with $k > 3$. From this consideration, it is evident that for $\varepsilon_k(\eta)$ with a lower degree of conditioning k, significantly more data are available, which reduces statistical uncertainty in the estimation.

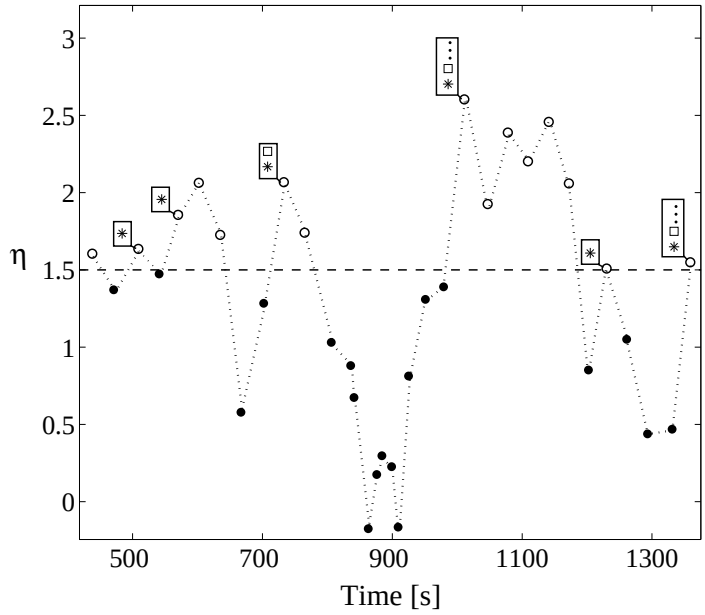


Figure 3.1: Counting of the exceedances of level $\eta = 1.5$ for sampled time series: • – non-exceedances; ∘ – unconditional exceedances considered for $\mathrm{a}_{1j}(\eta)$; ∗ – marks for conditional exceedances considered for $\mathrm{a}_{2j}(\eta)$; □ – marks for conditional exceedances considered for $\mathrm{a}_{3j}(\eta)$. Three dots $(\cdots)$ within rectangles indicate exceedances that could be considered for $\mathrm{a}_{kj}(\eta)$ with $k > 3$.

Evidently, $\lim_{\eta\to\infty} \mathbb{E}[B_{kj}(\eta)] = 1$. Therefore, it is convenient to introduce the modified ACER function as follows:

$$\tilde{\varepsilon}_k(\eta) = \frac{\sum_{j=k}^{N} \mathbb{E}[A_{kj}(\eta)]}{N-k+1}\,, \tag{3.16}$$

for which the following holds: $\lim_{\eta\to\infty} \tilde{\varepsilon}_k(\eta)/\varepsilon_k(\eta) = 1$. It turns out that the modified ACER function $\tilde{\varepsilon}_k(\eta)$ for $k \geq 2$ is easier to use for non-stationary or long-term statistics than $\varepsilon_k(\eta)$. Because the focus is on the values of an ACER function at the extreme levels, any function that provides a correct estimate at these levels may be used. This is true for $\tilde{\varepsilon}_k(\eta)$. However, some care should be exercised when selecting the data to be used in the statistical analysis because this modified version of the ACER function exhibits a strong spurious decrease for lower η levels, a tendency that increases with increasing k (Naess et al., 2010).

To show why $\tilde{\varepsilon}_k(\eta)$ defined by Eq. (3.16) may be applicable for non-stationary

time series, from Eqs. (3.9) and (3.13), it is recognized that

$$\begin{aligned} F(\eta) &\approx \exp\left(-\sum_{j=k}^{N}\alpha_{kj}(\eta)\right) = \exp\left(-\sum_{j=k}^{N}\frac{\mathbb{E}[A_{kj}(\eta)]}{\mathbb{E}[B_{kj}(\eta)]}\right) \\ &\underset{\eta\to\infty}{\simeq} \exp\left(-\sum_{j=k}^{N}\mathbb{E}[A_{kj}(\eta)]\right). \end{aligned} \tag{3.17}$$

If the time series can be partitioned into L blocks such that $\mathbb{E}[A_{kj}(\eta)]$ remains approximately constant within each block and such that $\sum_{j\in C_\ell}\mathbb{E}[A_{kj}(\eta)] \approx \sum_{j\in C_\ell}\mathrm{a}_{kj}(\eta)$ for a sufficient range of η values, where C_ℓ denotes the set of indices for the data in block no. ℓ, $\ell = 1,\ldots,L$, then

$$\begin{aligned} \sum_{j=k}^{N}\mathbb{E}[A_{kj}(\eta)] &= \sum_{\ell=1}^{L}\sum_{j\in C_\ell}\mathbb{E}[A_{kj}(\eta)] \\ &\approx \sum_{\ell=1}^{L}\sum_{j\in C_\ell}\mathrm{a}_{kj}(\eta) = \sum_{j=k}^{N}\mathrm{a}_{kj}(\eta)\,. \end{aligned} \tag{3.18}$$

Hence, we may write

$$F(\eta) \approx \exp\Big(-(N-k+1)\hat{\bar{\varepsilon}}_k(\eta)\Big)\,,\quad \eta\to\infty\,, \tag{3.19}$$

where

$$\hat{\bar{\varepsilon}}_k(\eta) = \frac{1}{N-k+1}\sum_{j=k}^{N}\mathrm{a}_{kj}(\eta)\,. \tag{3.20}$$

It is of interest to note which events are actually counted for the estimation of the various $\varepsilon_k(\eta)$, $k \geq 2$. Let us start with $\varepsilon_2(\eta)$. It follows from the definition of $\varepsilon_2(\eta)$ that $\varepsilon_2(\eta)\cdot(N-1)$ can be interpreted as the expected number of exceedances above the level η, satisfying the condition that an exceedance is counted only if it is immediately preceded by a non-exceedance. A reinterpretation of this is that $\varepsilon_2(\eta)\cdot(N-1)$ is equal to the average number of groups of exceedances above η, for the realizations considered within the time interval $\big(0,\ \mathcal{T}\big)$, where a group of exceedances is defined as a maximum number of consecutive exceedances above η. In general, $\varepsilon_k(\eta)\cdot(N-k+1)$ then equals the average number of groups of exceedances above η during the observation period $\big(0,\ \mathcal{T}\big)$, separated by at least $k-1$ non-exceedances (Naess and Gaidai, 2009). If the analysed time series is obtained by extracting local peak values from a narrow band response process, it is interesting to note the similarity between the ACER approximations and the envelope approximations for extreme value prediction (Naess and Gaidai, 2008; Vanmarcke, 1975).

Alternative statistical approaches that account for the effect of clustering on the extreme value distribution were studied in papers by Leadbetter (1983); Hsing (1987, 1991); Leadbetter (1995); Ferro and Segers (2003); Robert (2009). In these works, the emphasis was on the notion of an extremal index, which characterizes

the clumping or clustering tendency of the data and its effect on the extreme value distribution. However, note that these effects are automatically accounted for by the ACER functions (Naess, Gaidai, and Karpa, 2013).

Now, we can consider the problem of estimating a confidence interval for $\varepsilon_k(\eta)$. In the case in which several realizations of the time series $X(t)$ are provided, or the time series can be appropriately sectioned into R blocks, say, that is several annual realizations, or realizations of other duration of time, the sample estimate of $\varepsilon_k(\eta)$ would be:

$$\hat{\bar{\varepsilon}}_k(\eta) = \frac{1}{R}\sum_{r=1}^{R}\hat{\varepsilon}_k^{(r)}(\eta)\,, \tag{3.21}$$

where R is the number of realizations, and the $\hat{\varepsilon}_k^{(r)}(\eta)$ can be estimated using either the result from Eq. (3.15), that is

$$\hat{\varepsilon}_k^{(r)}(\eta) \;=\; \frac{\sum\limits_{j=k}^{N_r} \mathrm{a}_{kj}^{(r)}(\eta)}{\sum\limits_{j=k}^{N_r} \mathrm{b}_{kj}^{(r)}(\eta)}\,, \tag{3.22}$$

in case the considered stochastic process is assumed to be stationary, or the result from Eq. (3.20):

$$\hat{\varepsilon}_k^{(r)}(\eta) \;=\; \frac{1}{N_r-k+1}\sum_{j=k}^{N_r}\mathrm{a}_{kj}^{(r)}(\eta)\,, \tag{3.23}$$

for the non-stationary time series. The index (r) here refers to sample number r and N_r denotes the size of the sample number r, such that for N_r values holds $\sum\limits_{r=1}^{R} N_r \;=\; N$.

After the sample average conditional exceedance rate has been estimated, the 95% confidence interval for $\varepsilon_k(\eta)$ can then be determined. First, the sample standard deviation $\hat{s}_k(\eta)$ can be estimated by the basic formula

$$\hat{s}_k^2(\eta) \;=\; \frac{1}{R-1}\sum_{r=1}^{R}\Big(\hat{\varepsilon}_k^{(r)}(\eta)-\hat{\varepsilon}_k(\eta)\Big)^2\,. \tag{3.24}$$

Assuming that realizations are independent, a good approximation of the 95% confidence interval for the value $\varepsilon_k(\eta)$ is $CI = \Big(CI^-(\eta),\; CI^+(\eta)\Big)$, where

$$CI^{\pm}(\eta) \;=\; \hat{\varepsilon}_k(\eta) \;\pm\tau\cdot\frac{\hat{s}_k(\eta)}{\sqrt{R}}, \tag{3.25}$$

can be compared with Karpa and Naess (2013). Here, $\tau = t^{-1}\Big((1-0.95)/2,\; R-1\Big)$ is the corresponding quantile of the Student's t-distribution with $R-1$ degrees of freedom, cf. Rees (2001).

Alternatively, and as it also applies to the non-stationary case, as mentioned previously in Section 3.2, it is consistent with the adopted approach to assume that the stream of conditional exceedances over a threshold η constitute a Poisson process, possibly non-homogeneous. This leads to the following result:

$$\mathrm{Var}\left[\sum_{j=k}^{N} A_{kj}(\eta)\right] = \mathbb{E}\left[\sum_{j=k}^{N} A_{kj}(\eta)\right] = (N+k-1)\,\tilde{\varepsilon}_k(\eta)\,. \tag{3.26}$$

Therefore, using the normal distribution with mean $(N+k-1)\,\tilde{\varepsilon}_k(\eta)$ and variance $(N+k-1)\,\tilde{\varepsilon}_k(\eta)$ to approximate the Poisson distribution, for high levels η, limits of a 95 % confidence interval of $\tilde{\varepsilon}_k(\eta)$, and also $\varepsilon_k(\eta)$, can be estimated as,

$$CI^{\pm}(\eta) = \hat{\varepsilon}_k(\eta)\big(1 \pm \frac{1.96}{\sqrt{(N-k+1)\hat{\varepsilon}_k(\eta)}}\big)\,. \tag{3.27}$$

This, primarily, should be applied to the case when only one realization of the time series is available.

3.4 Functional Representation of the ACER Function

Using only the empirically estimated average conditional exceedance rates, the extreme value estimation cannot be achieved. The parametric representation of the ACER curve as a continuous function of η is required. It shall be derived in this section.

Unfortunately, the available extreme values observed from the sampled time series do not necessarily constitute the asymptotic distribution, or at least proving that they are truly asymptotic is in fact a nontrivial task. This implies the relevance of expanding the domain of our interest to the sub-asymptotic levels. More reason for the justification of this effort is the fact that the ACER functions allow us to use sub-asymptotic data with low statistical uncertainty, which is clearly an advantage.

Two scenarios for a time series and their distribution laws have been considered in this research. First, we focused on the case in which the underlying distribution belongs to the domain of attraction of the asymptotic extreme value distribution of the Gumbel type (2.7). This approach was originally derived and studied by Naess and Gaidai (2009); Naess et al. (2010). The extension of the asymptotic distribution to a parametric class of extreme value distribution tails that to some extent have the capacity to capture sub-asymptotic behaviour is more transparent, and perhaps more obvious, for the Gumbel case, cf. Naess et al. (2013). The second scenario was obtained assuming that the asymptotic behaviour of the underlying process takes form of the so-called generalized extreme value (GEV) distribution (2.12). In this approach, Naess et al. (2013) proposed that the class of parametric functions needed for the prediction of extreme values for the general case can be modelled on the relation between the Gumbel distribution and the general extreme value distribution.

3.4.1 The Asymptotic Gumbel case

The implication of the asymptotic distribution being of the Gumbel type on the possible sub-asymptotic parametric forms of the $\varepsilon_k(\eta)$ function cannot easily be determined in any detail. It cannot be done using only the sampled time series and estimated $\hat{\varepsilon}_k(\eta)$ as a basis for the estimation of extremes either. However, using the asymptotic extreme value distribution of the Gumbel type as a guide, it is argued by Naess and Gaidai (2009) that the behaviour of the ACER function in the tail is typically dominated by a function of the form $\exp\{-a(\eta-b)^c\}$. Here, a, b and c are suitable constants, whereas the values of η start from an appropriate tail level η_1 such that $b \leq \eta_1 \leq \eta$. Hence, it is assumed that

$$\varepsilon_k(\eta) = q_k(\eta)\, \exp\left\{ - a_k(\eta - b_k)^{c_k} \right\}, \;\; \eta \geq \eta_1 \geq b_k \,, \tag{3.28}$$

where the function $q_k(\eta)$ is slowly varying compared to the exponential function $\exp\{-a_k(\eta-b_k)^{c_k}\}$ and a_k, b_k, and c_k are suitable constants that will generally be dependent on k. Note that the values $c_k = q_k = 1$ correspond to the Gumbel asymptotic form, which implies that the Gumbel distribution would be obtained in the special case where the extreme value data are truly asymptotic.

Note that from Eq. (3.28), it follows that

$$-\log\left|\log\left(\frac{\varepsilon_k(\eta)}{q_k(\eta)}\right)\right| = -c_k \log(\eta - b_k) - \log(a_k)\,, \tag{3.29}$$

Therefore, under the employed assumptions, a plot of $-\log\left|\log\left(\varepsilon_k(\eta)/q_k(\eta)\right)\right|$ versus $\log(\eta - b_k)$ will exhibit an almost perfectly linear behaviour in the tail, which suggests a linear extrapolation strategy.

Naess and Gaidai (2009) argued that although the function $q_k(\eta)$ is generally not a constant function, its variation at the tail levels is often significantly slower compared to the function $\exp\{-a_k(\eta-b_k)^{c_k}\}$. This allows us to replace the function $q_k(\eta)$ by a constant value, q_k, potentially by adjusting the tail marker η_1. The validity of this approach has been successfully demonstrated by Naess and Gaidai (2008); Naess, Gaidai, and Haver (2007); Naess, Gaidai, and Teigen (2007) for mean up-crossing rate estimation for extreme value analysis of the response processes related to different dynamic systems.

Thus, for the estimation problem, it is assumed that the ACER functions can be represented in the tail as follows,

$$\varepsilon_k(\eta) = q_k\, \exp\left\{ - a_k(\eta - b_k)^{c_k} \right\}, \;\; \eta \geq \eta_1 \geq b_k \,, \tag{3.30}$$

where a_k, b_k, c_k, and q_k are appropriate constants. In a certain sense, the expression in Eq. (3.30) introduces a minimal class of parametric functions that can be used for our purpose. This representation of the ACER function makes it possible to achieve a few important tasks (Naess et al., 2013). First, the class is sufficiently flexible to capture, to a certain extent, the sub-asymptotic behaviour of any extreme value distribution, that is, asymptotically Gumbel. Particularly, the parametric class contains the asymptotic form given by $c_k = q_k = 1$ as a special case. Finally, the parametric functions agree with a wide range of known special cases, of which a

very important example is the extreme value distribution for a regular stationary Gaussian process, which has $c_k = 2$.

This immediately puts us in a position to apply a linear extrapolation strategy for deep tail prediction problems.

The tail marker η_1

Initially, it is important to discuss in some detail how the tail marker η_1 is defined. First, $\log(\hat{\varepsilon}_k(\eta))$ should be plotted versus levels of η. Visual inspection of this plot helps to identify the tail marker η_1. The value chosen for η_1 corresponds to the beginning of regular tail behaviour of the ACER function on a logarithmic scale in a sense to be discussed below.

Better understanding of the behaviour of the ACER function requires analysis of its components $\mathrm{a}_{kj}(\eta)$ and $\mathrm{b}_{kj}(\eta)$. We discovered in (3.22) and (3.23) that

$$\hat{\varepsilon}_k^{(r)}(\eta) \;=\; \frac{\mathrm{a}_k^{(r)}(\eta)}{\mathrm{b}_k^{(r)}(\eta)}\,, \;\; \text{or} \;\; \hat{\varepsilon}_k^{(r)}(\eta) \;=\; \frac{\mathrm{a}_k^{(r)}(\eta)}{N_r - k + 1}\,, \tag{3.31}$$

depending on whether the concerned process is stationary or non-stationary, respectively. Here, we use $\mathrm{a}_k^{(r)}(\eta) := \sum_{j=k}^{N} \mathrm{a}_{kj}^{(r)}(\eta)$ and $\mathrm{b}_k^{(r)}(\eta) := \sum_{j=k}^{N} \mathrm{b}_{kj}^{(r)}(\eta)$ to ease the notation. As before, the index (r) refers to the sample number r, $1 \le r \le R$.

First, we analyse the behaviour of function $\mathrm{b}_k(\eta)$, that is, the sum of realizations of the random function $B_{kj}(\eta)$, see Eq. (3.12). As previously discussed in Section 3.3, $\lim_{\eta\to\infty} \mathbb{E}[B_{kj}(\eta)] = 1$. This implies that $\lim_{\eta\to\infty} \mathrm{b}_k(\eta) = N - k + 1$ holds for $\mathrm{b}_k(\eta)$. In addition, from the definition of $B_{kj}(\eta)$, it follows that $\mathrm{b}_k(\eta)$ increases monotonically from zero to $N - k + 1$ on its support – the interval $\big[X_{(1)},\, X_{(N)}\big]$, where $X_{(k)}$ is the k-th order statistic of the given sample. The typical behaviour of $\mathrm{b}_k(\eta)$ is illustrated in Figure (3.2). Functions $\mathrm{b}_k^{(r)}(\eta)$ for four out of twenty realizations of a generated narrow band process with $5{\cdot}10^4$ data points per one realization are plotted versus η/σ levels for $k = 16$.

Thus, because functions $\mathrm{b}_k^{(r)}(\eta)$ are monotonically increasing, their reciprocals decrease monotonically on the sub-asymptotic levels of η, on which we are focused. $\mathrm{b}_k^{(r)}(\eta)$ degenerates to the value $1/(N - k + 1)$ in the far tail.

The aforementioned implies that variation in the $\varepsilon_k(\eta)$ substantially depends on the variation in the function $\mathrm{a}_k(\eta)$. From the definition of $\mathrm{a}_{kj}(\eta)$ as the realization of the random function $A_{kj}(\eta)$, cf. Eq. (3.12), it follows that the total number of observations that exceed level η and succeed $k - 1$ non-exceedances is low for lower magnitudes of η because most of the data exceed these levels. As η reaches values that are typical for the given time series, the function $\mathrm{a}_k(\eta)$ increases due to the increased availability of the appropriate observations. After missing the extremum at a certain point, the function $\mathrm{a}_k(\eta)$ begins to decrease on the interval that corresponds to sub-asymptotic and asymptotic levels of η because the number of exceedances of these levels decreases. Therefore, it can be concluded that the function $\mathrm{a}_k(\eta)$ reaches its local maximum point close to the vicinity of the mode of

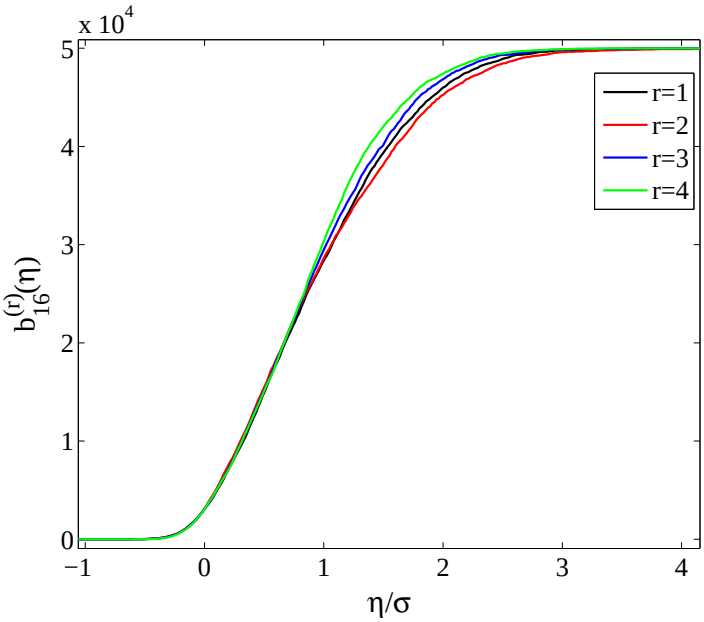


Figure 3.2: Behavior of the estimated functions $\mathrm{b}_k^{(r)}(\eta)$ for a stationary process for $k = 16$.

the observed time series and subsequently decreases as η tends to the right endpoint of the process.

Therefore, when choosing the level of the beginning of regular behaviour of $\mathrm{a}_k(\eta)$, the mode of the analysed time series should be taken into consideration. The attention should be drawn to the fact that the time series might be multimodal, or in other words, the probability density function of the time series might have several local maxima. In this case, the rightmost local maximum point should be considered. After passing this point function, $\mathrm{a}_k(\eta)$ decreases regularly, so does the $\varepsilon_k(\eta)$ function. If R realizations of the time series are available, the tail marker η_1 can, for instance, be defined as $\eta_1 = \max\left\{\eta_1^{(r)};\ r = 1, \ldots, R\right\}$, where $\eta_1^{(r)} = \arg\max_{\eta \in V}\left\{\mathrm{a}_k^{(r)}(\eta)\right\}$ for each $r = 1, \ldots, R$, and V here denotes the neighborhood of the argument of rightmost local maximum point of the PDF. Since the histogram of the given data set is a simple and reliable diagnostic tool for defining modes, hence it would always be useful to plot it prior to beginning the ACER analysis.

Figures (3.3) - (3.5) illustrate the above-described process for estimating the tail marker η_1 for three types of processes: a generated Gaussian narrow band process, a generated bimodal process and observed real data, respectively. In Figures (3.3a), (3.4a) and (3.5a), the histogram is plotted for each time series. As clearly shown in Figure (3.4a), the probability density function of the generated bimodal process has two local maxima, of which we are interested only in the rightmost. For the generated Gaussian narrow band process and the observed hourly maximum of the

3 s wind gust recorded during 12 years, both are unimodal processes that can also be affirmed by their histograms (3.3a) and (3.5a), respectively. The behaviour of the $\mathrm{a}_k(\eta)$ function for each example is shown in Figures (3.3b), (3.4b) and (3.5b). Four realizations of $\mathrm{a}_k(\eta)$, that is, $\mathrm{a}_k^{(r)}(\eta)$, $r = 1, \ldots, 4$, are plotted against normalized levels η/σ. It is clear that for all three cases functions, $\mathrm{a}_k^{(r)}(\eta)$ reach their local maxima on levels of local maxima of the corresponding PDFs, or close to the neighborhood of these levels. Thereafter, curves of $\mathrm{a}_k^{(r)}(\eta)$ subside regularly on the sub-asymptotic levels. Finally, plots of the estimated ACER function with the chosen levels of conditioning k are presented in Figures (3.3c), (3.4c) and (3.5c). The regular behaviour of the $\varepsilon_k(\eta)$ can be recognized within the context of the behaviour of the corresponding $\mathrm{a}_k^{(r)}(\eta)$ functions. The tail marker η_1 is automatically estimated using the Matlab-based standalone downloadable application created to implement the ACER method (Karpa, 2012). The tail marker η_1 is shown by the red diamond on the $\varepsilon_k(\eta)$ line. As a final remark, note that within the aforementioned software, the estimated tail marker η_1 can be considered rather as a hint, and expert users can change its value during use.

Estimation of parameters

We now proceed to the question of finding the optimal values of the parameters a, b, c, q of the ACER fit:

$$\varepsilon(\eta) = q \, \exp\big\{ - a(\eta - b)^c \big\}, \quad \eta \geq \eta_1 \geq b \,. \tag{3.32}$$

The index k is suppressed here and wherever applicable to simplify the notation.

Tentatively, we remove from consideration the very tail of the data, where uncertainty is high. As a practical procedure, Naess and Gaidai (2009) proposed neglecting data points where the relative confidence band width is greater than some constant δ. In this manner, we discard all points for which the following holds:

$$\frac{\tau \, \hat{s}(\eta)/\sqrt{R}}{\hat{\varepsilon}(\eta)} > \delta \,, \tag{3.33}$$

where the value chosen for δ depends on the actual accuracy of the considered data tail, and its value would typically be in the interval (0.5, 1]. Setting this parameter equal to 1 leaves sufficient data points for the weighted optimization problem and also ensures that no complex numbers will occur during calculations (Karpa, 2012).

Now, it is expedient to adopt the mathematical programming approach by optimizing the fit on the log level. Naess and Gaidai (2009) presented the following mean square error function to be minimized:

$$F(a, b, c, q) = \sum_{i=1}^{I} w_i' \big(\log \hat{\varepsilon}(\eta_i) - \log q + a(\eta_i - b)^c \big)^2 \,, \tag{3.34}$$

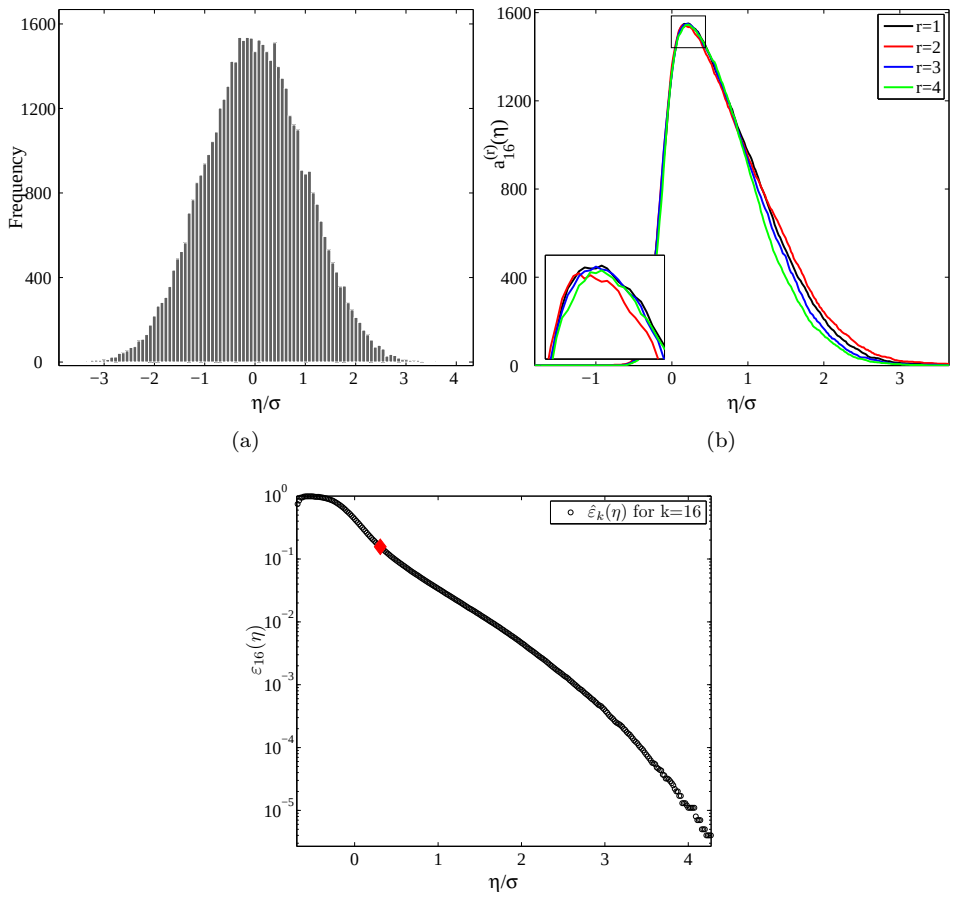


Figure 3.3: Estimation of the tail marker η_1 of the $\varepsilon_k(\eta)$ for $k = 16$ for the synthetic Gaussian narrow band process with twenty realizations. (a) – histogram of the process; (b) – plot of the four out of twenty estimated functions $\mathrm{a}_{16}^{(r)}(\eta)$; (c) – plot of the $\hat{\varepsilon}_{16}(\eta)$ against η/σ on a logarithmic scale with marked tail marker $\eta_1 = 0.31\sigma$; $\sigma = 1.8$.

where $\eta_1 < \ldots < \eta_I$ are the levels at which the ACER function has been empirically estimated, and $w_i' = w_i / \sum_{j=1}^{I} w_j$ with

$$w_i = \frac{1}{\Big[\log CI^{+}(\eta_i) - \log CI^{-}(\eta_i)\Big]^{\theta}}\,,\ \theta = 1 \text{ or } 2\,, \tag{3.35}$$

denoting the normalized weight factors that place more emphasis on the more reliable estimates of the $\hat{\varepsilon}_k(\eta_i)$ (Karpa and Naess, 2013). The choice of weight factor is, of course, to some extent arbitrary, and if it is considered more appropriate to

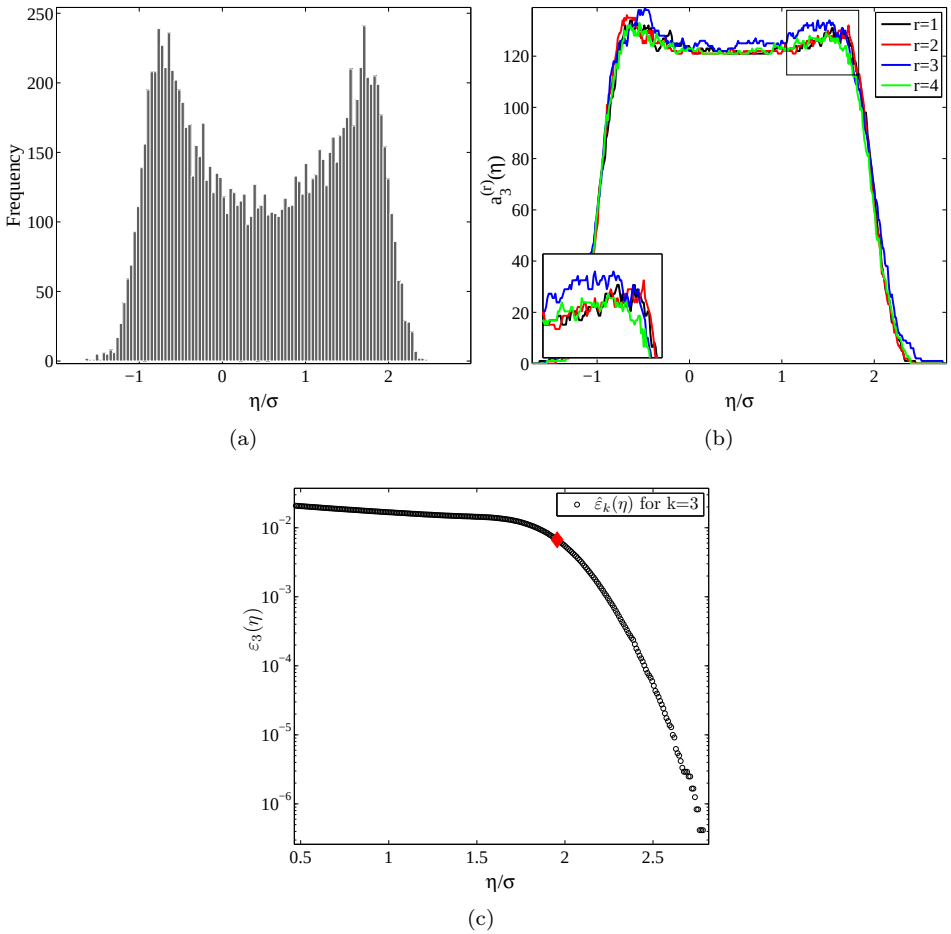


Figure 3.4: Estimation of the tail marker η_1 of the $\varepsilon_k(\eta)$ for $k = 3$ for the synthetic bimodal narrow band process with 200 realizations. (a) – histogram of the process; (b) – plot of the four out of 200 estimated functions $\mathrm{a}_3^{(r)}(\eta)$; (c) – plot of the $\hat{\varepsilon}_3(\eta)$ against η/σ on a logarithmic scale with marked tail marker $\eta_1 = 2\sigma$; $\sigma = 2.7 \cdot 10^3$.

place greater emphasis on the larger data, this can be simply achieved by assigning a value of 1 to the variable θ in the definition of w_i in Eq. (3.35). The optimization task can also be viewed as a weighted linear regression problem, as it will be discussed hereinafter. From this perspective, it emerges that the best linear unbiased estimators (BLUE) are obtained for another type of the weight factors (Draper and Smith, 1998; Montgomery et al., 2001). In this case, the 'best' weight factors are presented as follows,

$$w_i = \frac{1}{\mathrm{Var}\big(\log \hat{\varepsilon}_k(\eta_i)\big)} \,. \tag{3.36}$$

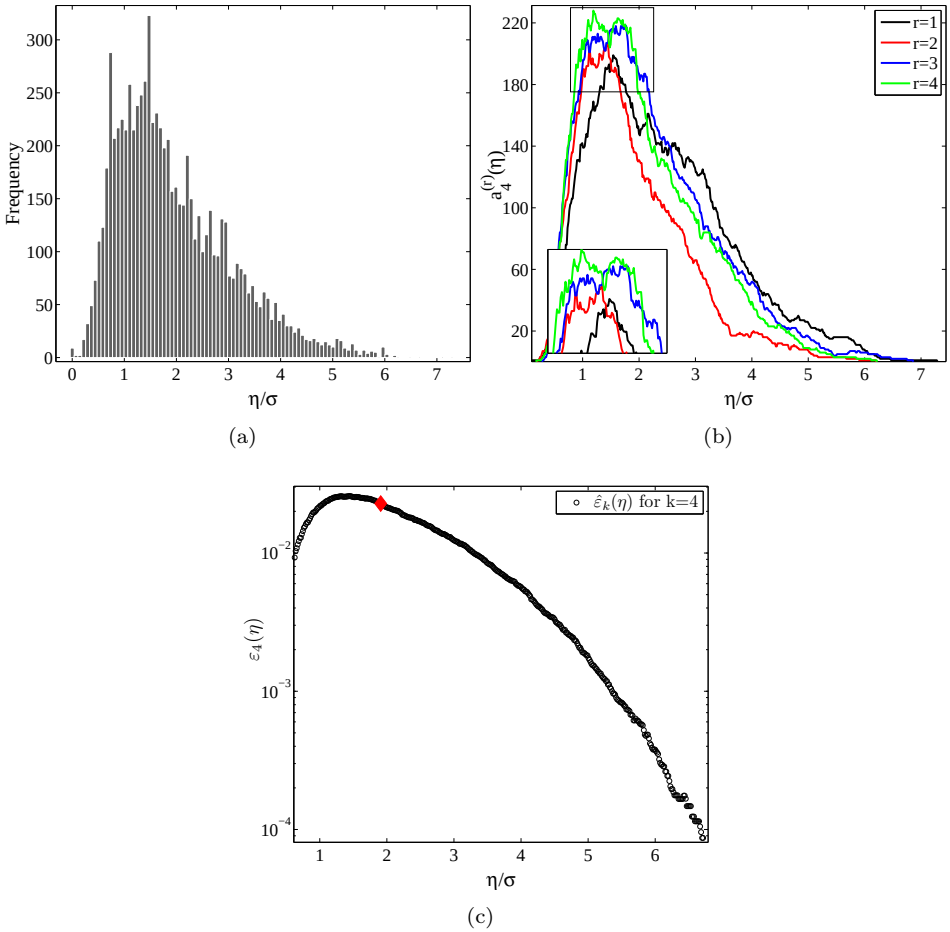


Figure 3.5: Estimation of the tail marker η_1 of the $\varepsilon_k(\eta)$ for $k = 4$ for the wind speed data observed during twelve years. (a) – histogram of the process; (b) – plot of the four out of twelve estimated functions $\mathrm{a}_4^{(r)}(\eta)$; (c) – plot of the $\hat{\varepsilon}_4(\eta)$ against η/σ on a logarithmic scale with marked tail marker $\eta_1 = 1.9\sigma$; $\sigma = 5.5$.

However, in practice, there are no benefits of the choice presented by Eq. (3.36) over the one defined by Eq. (3.35). In fact, the weight factors have disadvantages in both cases. Thus, in the alternative case given by Eq. (3.36), the summation in Eq. (3.34) would have to stop at those values of η_i for which at least one R realizations $\hat{\varepsilon}^{(r)}(\eta_i)$ becomes zero. This usually occurs at undesirably small values of η_1. In case of using Eq. (3.35), the summation in Eq. (3.34) would have to stop at the high values of η_i for which $CI^{-}(\eta_i)$ becomes negative. For this reason, the condition given by Eq. (3.33) has been applied in advance. Finally, it appears that using the weight factors defined by Eq. (3.35) retains considerably more significant data points than the other one.

It is discussed at some length by Naess and Gaidai (2009) that for the optimization task at hand, the Levenberg-Marquardt least-squares optimization method (Gill et al., 1981; Forst and Hoffmann, 2010) combined with the error function in Eq. (3.34) can be used. After the initial values q_0, b_0, a_0 and c_0 for the parameters are chosen, this method has usually worked quite well. We use $c_0 = 2$ as the starting value of the corresponding parameter. Then, it appears to be natural to use the mean value of the concerned time series ($\equiv \hat{\mu}$) as the initial value b_0. If this value is higher than the tail marker η_1, then it should obviously be replaced by the latter, as it follows from the assumption in Eq. (3.30). The availability of these two values makes it possible to find the remainder of the initial parameters. By taking the logarithm of both sides of Eq. (3.30), the following simple linear regression problem arises:

$$\log \hat{\varepsilon}(\eta_i) = -a(\eta_i - b_0)^{c_0} + \log q\,, \tag{3.37}$$

where $\log \hat{\varepsilon}(\eta_i)$ is the regressand and $(\eta_i - b_0)^{c_0}$ is the regressor of the model. The solution of this problem provides us with the initial values of two other parameters, q and a: the computed slope of the regression fit gives $-a_0$, and the intercept of the regression fit gives $\log q_0$.

As discussed by Naess et al. (2013), if the parameter c is equal to 1.0 or close to 1.0, the optimization problem becomes ill-posed or close to ill-posed because there is an infinite number of (b, q) values that provide exactly the same value of $F(a, b, c, g)$. Therefore, there is no well-defined optimal solution in parameter space. A possible method to facilitate this problem is to fix the q value, and one option is $q = 1$. Another choice for fixing the q is the ACER function estimated at the mean value of the concerned time series, that is, $q = \hat{\varepsilon}(\hat{\mu})$ (Valberg, 2010).

Note that the optimization problem specified by Eq. (3.34), was implemented in MATLAB, version 7.9.0.529 (R2009b) and earlier. Unfortunately, the embedded Levenberg-Marquardt algorithm does not handle constraints of any type (Matlab, 2009), although it is advisable to consider the constrained optimization as the one that can provide more reasonable values for the parameters. It is known, however, that the Levenberg-Marquardt method can be applied for solving constraint optimization problems (Kanzow et al., 2004). This algorithm has already been implemented by (Lourakis, 2004) and can also be implemented in MATLAB. Hence, a natural direction for future research involves employing the Levenberg-Marquardt algorithm for constrained optimization.

In general, to improve the robustness of the results, it is recommended that a nonlinearly constrained optimization be applied (Forst and Hoffmann, 2010). From our experience, the trust-region approach for constrained minimization (Sun and Yuan, 2006; Moré and Sorensen, 1983; Wu and Wu, 1991) is well suited for the task. The constrained optimization problem with the objective function defined in Eq. (3.34) and nonlinear inequality constraints is written as:

$$\begin{cases} F(a, b, c, q) \to \min\,; \\ \log q - a(\eta_i - b)^c \le 0\,,\ i = 1, \ldots, I\,; \\ \{a, b, c, q\} \ \in\ \mathcal{S}^{(4)}\,, \end{cases} \tag{3.38}$$

where the constraints domain $\mathcal{S}^{(4)}$ is defined as:

$$\mathcal{S}^{(4)} = \big\{\{a,\, b,\, c,\, q\} \in \mathbb{R}^4 \;\big|\; a,\, c,\, q \in (0,\; +\infty);\; b \in (b_{\min},\, \eta_1]\big\}. \tag{3.39}$$

The system of inequality constraints is evident from the definition of the estimate $\hat{\varepsilon}(\eta)$ of the ACER function, given by Eqs. (3.15) or (3.20), as it follows that for all η holds $\hat{\varepsilon}(\eta) \leq 1$. This implies that under the assumption we made on the functional representation of the ACER in Eq. (3.30), $\forall\, i = 1, \ldots, I$ $\log \hat{\varepsilon}(\eta_i) = \log q - a(\eta_i - b)^c \leq 0$.

The bounds, defined by the constraints domain $\mathcal{S}$, for the values of a, c and q are evident. These parameters should remain positive to ensure that the properties of the cumulative distribution function defined by Eq. (3.11) are fullfilled. In addition, it was reported by Karpa and Naess (2013) that the c value can be further bounded. Based on the underlying statistics of environmental data, particularly of wind speed data, it appears to be reasonable to restrict the range of the c values to the interval $1 \leq c \leq 3$. Finally, regarding the margins of the b value, the right bound is defined by the tail marker η_1, which implies that in this case $\forall\, \eta \geq \eta_1$, the difference $(\eta - b)$ remains non-negative. The left end point of the location parameter b is, generally speaking, unlimited; therefore, the $b_{\min}$ value can be set to $-\infty$. However, in our research, we used the value that also reflects the underlying statistics of the observed data, namely, the minimal possible value of the considered process, which is $b_{\min} = 0$ in the case of wind speeds.

Scrutiny of the tail representation of the ACER function and the objective function in Eq. (3.34) ascertains that if b and c are fixed constants, the optimization problem reduces to a standard weighted linear regression problem of the form:

$$y_i = -a\, x_i + \log q, \tag{3.40}$$

with the corresponding mean square error function:

$$F(a, q; b, c) = \sum_{i=1}^{I} w_i' \big(y_i - \log q + a\, x_i\big)^2, \tag{3.41}$$

where $y_i = \log \varepsilon(\eta_i)$ and $x_i = (\eta_i - b)^c$, $i = 1, \ldots, I$. Thus, with both b and c fixed, the optimal values of a and q are found using closed-form weighted linear regression formulas in terms of normalized weights w_i' derived from Eq. (3.35), y_i and x_i, cf., e.g., Ryan (2008); Carroll and Ruppert (1988):

$$\begin{aligned} a^*(b, c) \;&=\; -\,\frac{\sum\limits_{i=1}^{I} w_i'(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{I} w_i'(x_i - \overline{x})^2} \\ &=\; -\,\frac{\mathrm{Cov}(x,\, y)}{\mathrm{Var}(x)}, \end{aligned} \tag{3.42}$$

and

$$\log q^*(b, c) \;=\; \overline{y} + a^*(b, c)\, \overline{x}, \tag{3.43}$$

where $\overline{x} = \sum_{i=1}^{I} w_i' x_i$ with a similar definition of $\overline{y}$, and as it is observed after the second equal sign in Eq. (3.42), $\mathrm{Cov}(\cdot, \cdot)$ and $\mathrm{Var}(\cdot)$ are the weighted sample covariance and variance, respectively.

Now, the nonlinear optimization procedures are to be applied for finding the optimal values of the parameters b and c. We substitute the expressions in Eqs. (3.42) and (3.43) into the objective function in (3.41) such that the modified mean square error function to be minimized takes the following form:

$$\widetilde{F}(b,c) = F\big(a^*(b,c),\, b,\, c,\, q^*(b,c)\big) = \mathrm{Var}(y) - \frac{\mathrm{Cov}^2(x,\, y)}{\mathrm{Var}(x)}\,, \tag{3.44}$$

The Levenberg-Marquardt method may now be applied to the function $\widetilde{F}(b,c)$ to determine the optimal b^* and c^*. In the present research, the interior-point approach to constrained minimization (Byrd et al., 2000; Waltz et al., 2006) was also used for this purpose. The optimization problem with bound constraints only is stated as:

$$\begin{cases} \widetilde{F}(b,c) \to \min\,, \\ b_{\min} < b \le \eta_1\,,\ 0 < c < +\infty\,. \end{cases} \tag{3.45}$$

The same considerations for the lower and upper bounds of the b and c values, which were previously discussed, are relevant here. It is also possible to subject the minimization to the nonlinear inequalities of the form

$$\overline{y} - \mathrm{Cov}(x,\, y)(\overline{x} - x_i)/\mathrm{Var}(x) \le 0\,,\ i = 1, \ldots, I\,, \tag{3.46}$$

which emerges from the corresponding set of constraints presented in (3.38) by replacing $\log q$ and a with the corresponding expressions from Eqs. (3.42) and (3.43). The use of this set of nonlinear inequality constraints induces a direction for further study. However, the minimization problem (3.45) with bound constraints only works well and thus far appears to be robust. The final step is now to apply Eqs. (3.42) and (3.43) to find the optimal a^* and q^*.

As a final comment, it is worth pointing out that optimal values of the parameters a, b, c and q may also be determined using a sequential quadratic programming (SQP) method incorporated in the NAG Numerical Library (Numerical Algorithms Group, 2010). Research on the efficiency of this toolbox for the described problem is to be performed in the future.

3.4.2 The General case

Thus far, we have analysed the class of parametric functions presented in Eq. (3.30) as the possible sub-asymptotic parametric forms of $\varepsilon_k(\eta)$. This approach was based on the assumption that the asymptotic distribution is of the Gumbel type. In this subsection, the general case will be discussed.

Although the extension of the asymptotic Gumbel case to the proposed class of sub-asymptotic distributions was fairly evident, this is not equally so for the general case. Naess et al. (2013) assumed that the class of parametric functions

required for predicting the extreme values for the general case can be modelled on the relation between the Gumbel distribution and the so-called generalized extreme value distribution (GEV), cf. Leadbetter et al. (1983); Coles (2001); Reiss and Thomas (2007). This approach leads to the assumption that for independent data in the general case, the ACER function $\varepsilon_1(\eta)$ can be expressed in the tail asymptotically as

$$\varepsilon_1(\eta) \underset{\eta\to\infty}{\simeq} \Big[1+\xi\big(a(\eta-b)\big)\Big]^{-1/\xi} \tag{3.47}$$

where $a>0$, b, are ξ are constants such that $\big[1+\xi\big(a(\eta-b)\big)\big]>0$. This representation follows from the explicit form of the generalized extreme value distribution, expressed by Eq. (2.12) in Section 2.1. For the dependent data, it is reasonable to use a type of approximation similar to that for the asymptotic Gumbel case, cf. Section 3.4.1. In this case, the sub-asymptotic part of the tail of the ACER function is assumed to follow a curve largely of the form $\big[1+\xi\big(a(\eta-b)^c\big)\big]^{-1/\xi}$ for the considered range of $\eta:\ \eta\geq\eta_1\geq b$, where $a>0$, $c>0$, $\xi>0$ and b are proper constants (Naess et al., 2013). The tail marker η_1 is chosen to correspond to the beginning of regular tail behaviour of the ACER function on a logarithmic scale as before, cf., Section 3.4.1. Hence, it is assumed that

$$\varepsilon_k(\eta) = q_k(\eta)\Big[1+\xi_k\big(a_k(\eta-b_k)^{c_k}\big)\Big]^{-1/\xi_k},\ \ \eta\geq\eta_1\geq b_k\,, \tag{3.48}$$

where the function $q_k(\eta)$ is slowly varying compared with the function $\Big[1+\xi_k\big(a_k(\eta-b_k)^{c_k}\big)\Big]^{-1/\xi_k}$ and $a_k>0$, $c_k>0$, $\xi_k>0$ and b_k are suitable constants that, generally speaking, are dependent on k.

It is worth considering two special cases of the general expression in Eq. (3.48). The values $q_k(\eta)\equiv 1$ and $c=1$ correspond to the asymptotic limit given by (3.47). When $\xi_k=0$, the general expression (3.48) degenerates to the asymptotic Gumbel case $q_k(\eta)\exp\big\{-a_k(\eta-b_k)^{c_k}\big\}$ presented in Eq. (3.28).

An alternative equivalent form to Eq. (3.48) is to assume that

$$\varepsilon_k(\eta) = \Big[1+\xi_k\big(a_k(\eta-b_k)^{c_k}+d_k(\eta)\big)\Big]^{-1/\xi_k},\ \ \eta\geq\eta_1\geq b_k\,, \tag{3.49}$$

where the function $d_k(\eta)$ is weakly varying compared with the function $a_k(\eta-b_k)^{c_k}$. The equivalence between Eqs. (3.48) and (3.49) can easily be derived by setting the value of ξ_k to zero in both equations, causing degeneration to the asymptotic Gumbel case (3.28). Therefore, it may be simply concluded that $d_k(\eta)=\log q_k(\eta)$. However, for estimation purposes, the form given by Eq. (3.48) is preferable because it leads to simpler estimation procedures, which will be discussed later.

For practical recognition of the ACER functions given by (3.48), it is convenient to assume that the unknown function $q_k(\eta)$ varies sufficiently slowly to be replaced by a constant. In general, $q_k(\eta)$ is not constant, but its variation in the tail region is assumed to be sufficiently slow to allow for its replacement by a constant. Hence, as in the Gumbel case, it is essentially assumed that $q_k(\eta)$ can be replaced by a

constant q_k for the considered range of $\eta : \eta \geq \eta_1$ with an appropriately chosen tail marker η_1. Again, for simplicity of notation, the index k will be suppressed. Consequently, the ACER function can now be written as

$$\varepsilon(\eta) = q\big[1 + \tilde{a}(\eta - b)^c\big]^{-\gamma}, \quad \eta \geq \eta_1 \geq b\,, \tag{3.50}$$

where $\tilde{a} = \xi a$, $\gamma = 1/\xi$.

Prior to the data analysis, the tail marker η_1 is estimated. The value chosen for η_1 is identified from visual inspection of the plot $\big(\eta_i, \log\hat{\varepsilon}(\eta_i)\big)$, $i = 1, \ldots, I$, and it corresponds to the beginning of regular tail behaviour as discussed in Section 3.4.1. However, note that the problem of defining the tail marker η_1 for time series with heavy tails requires further scrutiny in future research.

The parameters q, $\tilde{a}$, b, c and γ are determined as the solution of the log-level optimization problem, as for the Gumbel case. The mean square error function for the estimated $\hat{\varepsilon}(\eta)$ to be minimized in the general case is written as

$$F(\tilde{a}, b, c, q, \gamma) = \sum_{i=1}^{I} w_i' \Big[\log\hat{\varepsilon}(\eta_i) - \log q + \gamma \log\big(1 + \tilde{a}(\eta_i - b)^c\big)\Big]^2, \tag{3.51}$$

where the normalized weight factors $w_i' = w_i / \sum_{j=1}^{I} w_j$ are defined as previously through Eq. (3.35). Because the summation in Eq. (3.51) has to stop before $CI^-(\eta)$ becomes negative for the very tail of the data, the condition (3.33) has to be previously applied, as it was for the Gumbel case.

It is now possible to proceed to the problem of finding the optimal parameters of (3.51). An option for estimating the five parameters $\tilde{a}$, b, c, q, γ is again to use the Levenberg-Marquardt least squares optimization method Naess et al. (2013). As previously discussed for the Gumbel case (see Section 3.4.1), it is also reasonable to apply the constrained minimization based on the trust-region approach, cf. Sun and Yuan (2006); Moré and Sorensen (1983), and on the interior-point approach for constrained minimization (Byrd et al., 2000; Waltz et al., 2006). In this case, the optimization problem with only the bound constraints is written as:

$$\begin{cases} F(\tilde{a}, b, c, q, \gamma) \to \min\,, \\ \{\tilde{a}, b, c, q, \gamma\} \;\in\; \mathcal{S}^{(5)}\,, \end{cases} \tag{3.52}$$

where the constraint domain $\mathcal{S}^{(5)}$ is similar to the one given by Eq. (3.39) and is now defined as:

$$\mathcal{S}^{(5)} = \big\{\{\tilde{a},\, b,\, c,\, q,\, \gamma\} \in \mathbb{R}^5 \;\big|\; \tilde{a},\, c,\, q,\, \gamma \in (0,\; +\infty);\; b \in (b_{\min},\, \eta_1]\big\}. \tag{3.53}$$

Note that for future research on the general case, it would be relevant, inter alia, to include the nonlinear inequality constrain of the form $\log q - \gamma \log\big(1 + \tilde{a}(\eta_i - b)^c\big) \leq 0$.

In addition, the estimation problem can be simplified by reducing it to a standard weighted linear regression problem. It is easy to see that for the fixed values $\tilde{a}$, b and c in Eq. (3.51), the optimal values of γ and $\log q$ are found using closed-form weighted linear regression formulas in terms of w_j, $y_i = \log\hat{\varepsilon}(\eta_i)$

and $x_i = 1 + \tilde{a}(\eta_i - b)^c$ (Naess et al., 2013). Here, the optimal values of γ and $\log q$ are given by relations analogous to Eqs. (3.42) and (3.43), respectively. To calculate the final optimal set of parameters, the Levenberg-Marquardt method, or other methods that tolerate constraints, may then be used on the function $\widetilde{F}(\tilde{a}, b, c) = F\big(\tilde{a},\, b,\, c,\, q^*(\tilde{a},\, b,\, c),\, \gamma^*(\tilde{a},\, b,\, c)\big)$ to find the optimal values $\tilde{a}^*$, b^* and c^*. Then, the corresponding γ^* and q^* can be calculated.

The NAG Numerical Library (Numerical Algorithms Group, 2010) may also be used to apply a sequential quadratic programming method to find the optimal parameters $\tilde{a}^*$, b^*, c^*, q^* and γ^*.

3.4.3 The T-year return level by the ACER

As discussed in Section 2.1.2, the T-year return level η^{Tyr} is a threshold that is exceeded by the observation in the given year (or any other period) once in T years (or any other periods), that is, with probability $1/T$:

$$\mathrm{Prob}\{X > \eta^{\mathrm{Tyr}}\} = \frac{1}{T} \tag{3.54}$$

Thus, evidently,

$$F^{1\mathrm{yr}}(\eta^{\mathrm{Tyr}}) = 1 - \frac{1}{T}\,, \tag{3.55}$$

where $F^{1\mathrm{yr}}(\eta)$ denotes the distribution function of the yearly extreme value.

Returning to the ACER function, it is observed in Section 3.2 and from Eq. (3.11) that the cumulative distribution can be effectively expressed as:

$$F(\eta) = F_k(\eta) = \exp\Big(- (N - k + 1)\, \varepsilon_k(\eta)\Big). \tag{3.56}$$

The definition of the average conditional exceedance rate (3.10) implies that the $F(\eta)$ here is the distribution of the extreme value of $X(t)$ during the observation period $\big(0,\ \mathcal{T}\big)$ because the expression $\varepsilon_k(\eta) \cdot (N - k + 1)$ is the expected effective number of exceedances subjected to $k - 1$ immediately preceding non-exceedances and recorded for the duration $\big(0,\ \mathcal{T}\big)$. Now, let us assume that the time interval $\big(0,\ \mathcal{T}\big)$ takes n_y observation years (or corresponding periods). In this case, over the time of one year, the expected effective number of conditional exceedances is $\varepsilon_k(\eta) \cdot (N - k + 1)/n_y$. Therefore,

$$F^{1\mathrm{yr}}(\eta) = \exp\left(- \frac{(N - k + 1)}{n_y}\, \varepsilon_k(\eta)\right). \tag{3.57}$$

After matching Eqs. (3.55) and (3.57), it clearly emerges that

$$\varepsilon_k(\eta^{\mathrm{Tyr}}) = - \frac{n_y}{(N - k + 1)} \log\left(1 - \frac{1}{T}\right). \tag{3.58}$$

Here, the precise explicit form of the T-year return level η^{Tyr} expressed in terms of the fitted ACER function for the asymptotic Gumbel case is:

$$\eta^{\mathrm{Tyr}} \;=\; \left[\frac{1}{a} \log\left(- \frac{q\,(N - k + 1)}{n_y\, \log(1 - 1/T)}\right)\right]^{1/c} + b\,, \tag{3.59}$$

with the estimated optimal parameters a, b, c and q in Eq. (3.32). In the general case, the T-year return level η^{Tyr} is expressed as:

$$\eta^{\mathrm{Tyr}} = \left\{ \frac{1}{\tilde{a}} \left[\left(- \frac{n_y \log(1 - 1/T)}{q\,(N - k + 1)} \right)^{-1/\gamma} - 1 \right] \right\}^{1/c} + b \,, \tag{3.60}$$

where the parameters $\tilde{a}$, b, c, q and γ are optimally estimated for Eq. (3.50).

3.4.4 Estimation of the 95% confidence interval

After we have determined the T-year return level of the extreme value distribution given by a particular ACER function as provided by the fitted parametric curve, the 95% confidence interval of the deep tail quantile has to be estimated.

To obtain a reasonable and sufficiently precise estimate of the confidence interval, a bootstrapping method is recommended (Efron and Tibshirani, 1993). However, this procedure cannot be applied to observed real data because, e.g., sampling with replacement would demolish the dependence structure of the stochastic process. Therefore, this method will be presented in the section on extreme value prediction for synthetic data.

A simple first estimation of a confidence interval for the extreme value predicted using the parametric ACER function was proposed by Naess and Gaidai (2009); Naess et al. (2009). It starts from re-anchoring the empirical confidence band to the optimally fitted curve. For this purpose, the individual confidence intervals $\mathrm{CI}^{\pm}(\eta_i)$ for the point estimates $\hat{\varepsilon}(\eta_i)$ of the ACER function are centred on the fitted one, that is:

$$CI^{\pm}(\eta_i) \;=\; \varepsilon(\eta_i) \;\pm \tau \cdot \frac{\hat{s}(\eta_i)}{\sqrt{R}} \,, \;\; i = 1, \ldots, I \,, \tag{3.61}$$

where the ACER curve $\varepsilon(\eta_i)$ is defined either by Eq. (3.32) or by Eq. (3.50).

Subsequently, in the work of Naess, Gaidai, and Batsevych (2009), the authors perturbed the optimal values of each parameter linearly around their 5%-neighborhood to construct a range of ACER curves. Only those curves that remained within the re-anchored empirical 95% confidence interval were taken into consideration. Evidently, each of the selected curves provides a prediction for the extreme return level of interest. Thus, the ultimate values determine an optimized confidence interval of the desired return value; see Naess and Gaidai (2009); Naess et al. (2009, 2010).

Comparison of the described procedure with the results obtained using a non-parametric bootstrapping method for the synthetic data revealed that the size of the 95% confidence interval is underestimated in the first case. Therefore, the authors attempted to eliminate this inconsistency in the length of the 95% confidence band in their later research. Karpa and Naess (2013); Naess et al. (2013) hypothesized that the class of parametric curves specified by Eq. (3.32) or by Eq. (3.50) fully describes the behaviour of the ACER functions in the tail. Under this assumption, corresponding parametric curves are fitted to the given set of points $\big(\eta_i,\, CI^{+}(\eta_i)\big)$ and $\big(\eta_i,\, CI^{-}(\eta_i)\big)$ for $i = 1, \ldots, I$ of the re-anchored confidence band (3.61). The

same fitting procedure as described above in Section 3.4.1 is adopted to find the optimal curve. Apparently, the points (3.61) are considered for levels η_i bounded by the tail marker η_1 and the condition given by Eq. (3.33) as before. Regarding the selection the weight factors of the corresponding mean square error function, the same weights as for the ACER optimization are used here. This appears to be reasonable because as presented by Eq. (3.35), the weights give less significance to more uncertain points in the tail as required. Furthermore, the fitted curves are extrapolated to the level of interest and thereby provide the first estimate of a 95% confidence interval of the predicted return level Naess et al. (2013). This procedure appears to provide confidence intervals that are consistent in length but slightly shifted compared with the results obtained using a non-parametric bootstrapping method. A comparison of both these methods of estimating the confidence intervals is presented in the section Section 3.5.1 on extreme value prediction for synthetic data.

As a final point, it has been observed, in particular by Karpa and Naess (2013), that the predicted return value is not very sensitive to the choice of tail marker η_1, provided it is chosen with some care. This property is easily recognized by looking at the way the optimized fitting is performed. If the tail marker is in the appropriate domain of the ACER function, the optimal fitted curve does not appreciably change by moving the tail marker (Naess et al., 2013).

3.5 Numerical Illustrations

3.5.1 Monte Carlo analysis of synthetic independent wind data

In this section, we illustrate the performance of the ACER method and the estimation of the 95% confidence interval. The extreme value statistics will be analysed by application to synthetic data for which the exact extreme values can be calculated (Naess and Clausen, 2001).

Let us consider 20 years of synthetic wind speed data with 100 observations per year, amounting to 2000 data points. This is not a considerable amount of data for detailed statistics; however, this case may represent a real situation in which only a limited data sample is available. Therefore, it is crucial to utilize all available data when providing extreme value estimates. As it will be shown, the tail extrapolation technique provided by the ACER method performs better than asymptotic methods such as POT or Gumbel.

As previously performed by Naess and Clausen (2001); Naess and Gaidai (2009), we also assumed that the underlying normalized stochastic process $X(t)$ is a stationary Gaussian process with a zero mean value and a standard deviation equal to one. It was also assumed that the mean zero up-crossing rate $\nu^+(0)$ is such that the product $\nu^+(0)T = 10^3$, where $T = 1$ year. According to Naess and Clausen (2001), this appears to be typical for the wind speed process. Using the Poisson assumption and the Rice formula, see, for example, Naess and Moan (2012), the

distribution of the yearly extreme value of $X(t)$ is then calculated using the formula

$$\begin{aligned} F^{1\mathrm{yr}}(\eta) &= \exp\left\{-\nu^{+}(\eta)T\right\} = \exp\left\{-\nu^{+}(0)T\exp\left(-\frac{\eta^2}{2}\right)\right\} \\ &= \exp\left\{-10^3\exp\left(-\frac{\eta^2}{2}\right)\right\}, \end{aligned} \tag{3.62}$$

where $T = 1$ year, $\nu^{+}(\eta)$ is the mean up-crossing rate per year, and η is the scaled wind speed. Using expression (3.62), the 100-year return period value $\eta^{100\mathrm{yr}}$ is then calculated as the solution to equation $F^{1\mathrm{yr}}(\eta^{100\mathrm{yr}}) = 1 - 1/100$. Thus, the 100-year value is $\eta^{100\mathrm{yr}} = 4.80$.

Extreme value analysis of wind data is generally performed based on the fact that the peak events extracted from measurements of the wind speed process should be separated by 3-4 days. This is done to obtain approximately independent data, as required by the POT method. (Naess and Haug, 2010).

Accordingly, the Monte Carlo simulated peak event data to be used for the synthetic example were generated from the following extreme value distribution, which was also used by Naess and Clausen (2001):

$$F^{3\mathrm{d}}(\eta) = \exp\left\{-q\exp\left(-\frac{\eta^2}{2}\right)\right\}, \tag{3.63}$$

where $q = \nu^{+}(0)T = 10$, which corresponds to $T = 3.65$ days, such that $F^{1\mathrm{yr}}(\eta) = \left(F^{3\mathrm{d}}(\eta)\right)^{100}$. This implies the necessity to generate 100 data points from one year; thus, in total (i.e., in 20 years), the data amounted to 2000 data points.

To obtain an idea about the performance of the ACER, POT and Gumbel methods, 200 independent 20-year Monte Carlo simulations were conducted by analogy with the work by Naess, Gaidai, and Karpa (2013) as follows. For the ACER method, it was naturally decided to analyse data with no conditioning on previous observations. Because the generated data points (i.e., $T = 3.65$ days maxima) are independent, the ACER function $\varepsilon_k(\eta)$ is independent of k. Therefore, we set $k = 1$. Because the 200 independent simulations were performed in a loop and because there was no capability to analyse each sample for the purposes of defining the tail marker η_1, it was decided to fix this variable to the level 2.3. This decision was based on the analysis of several independent 20-year simulations. To estimate a 95% confidence interval for each estimated value of the ACER function $\varepsilon_1(\eta_i)$ for the chosen range η_i, $i = 1, \ldots, I$ of η values, the required standard deviation $\hat{s}_k(\eta)$ in Eq. (3.25) was based on 20 estimates of the ACER function using the yearly data. This provided a 95% confidence band on the optimally fitted curve based on 2000 data points. From these data, the predicted 100 year return level was obtained using Eq. (3.59) from the relation $\hat{\varepsilon}_1(\eta^{100\mathrm{yr}}) = -\log(0.99)/100$. A non-parametric bootstrapping method was also used to estimate a 95% confidence interval based on 1000 resamples of size 2000. For this purpose, we used the Matlab (2009) Statistics Toolbox routine `bootstrp` to sample with replacement from 2000 Monte Carlo simulated $T = 3.65$ days peak events. The 100-year return level $\eta^{100\mathrm{yr}}$ was estimated using the ACER method for each of the 1000 replicate samples.

The POT prediction of the 100-year return level was based on using the maximum likelihood estimates (MLE) of two parameters of the generalized Pareto

distribution for a specific choice of threshold. One of the unfortunate features of the POT method is that the predicted 100-year value may vary significantly with the choice of the threshold u_0 (Karpa and Naess, 2013; Naess et al., 2013), Therefore, for the synthetic data, we also followed the standard recommended procedures for identifying a suitable threshold (Coles, 2001). Again, because 200 independent Monte Carlo simulations were performed in a loop, it was decided to fix the threshold at the level $u_0 = 3$. On average, this choice retains 200 exceedances over the threshold out of 2000 data points, which is sufficient to provide the MLE (Braunstein, 1992). The decision was based on the analysis of a mean residual life plot, as shown in Figure (3.6).

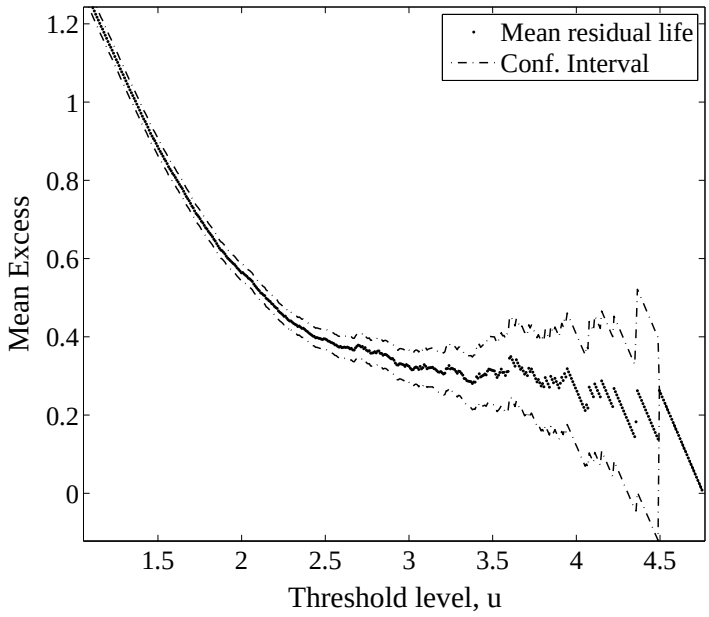


Figure 3.6: Mean residual life plot for 20 years Monte Carlo simulated 3.65 days wind peaks.

The 95% confidence interval was obtained using the non-parametric bootstrap method as it was done for the ACER method. The Matlab (2009) gpfit routine for the generalized Pareto distribution was applied to 1000 samples that had been obtained by sampling with replacement from 2000 Monte Carlo simulated data. Here, for the given threshold $u_0 = 3$, we produced a sample of 1000 POT predictions, from which the 95% CI was estimated. In addition, for comparison purposes, the 95% confidence interval was also estimated from the parametrically bootstrapped PDF of the generalized Pareto estimate for the given threshold $u_0 = 3$. A sample of 1000 data sets was used.

Note that under the assumptions we made for this model and the fact that the normal distribution belongs to the Type I domain of attraction (Leadbetter et al., 1983), the true asymptotic distribution of maxima is the Gumbel extreme value

distribution (2.7). This means that the shape parameter of the GEV distribution in (2.12) is $\gamma = 0$, which immediately implies that the true asymptotic distribution of threshold excesses is the exponential distribution of the form (2.30). The POT method used here, however, is based on adopting the generalized Pareto distribution $W_\gamma(y)$, Eq. (2.28). The reason is simply that this is the recommended procedure (Coles, 2001), which is somewhat unfortunate but understandable because the GP distribution provides greater flexibility in terms of curve fitting (Naess et al., 2013). The correct asymptotic distribution of exceedances was also used on this example. However, poor results for the estimated return period values were obtained. The price to pay for using the GP distribution is that the estimated parameters may easily lead to an asymptotically inconsistent extreme value distribution.

The 100 year return level predicted by the Gumbel method was based on using the method of moments for parameter estimation on the sample of 20 yearly extremes. This choice of estimation method is due to the small sample of extreme values. The 95% confidence interval was obtained from the parametrically bootstrapped PDF of the Gumbel prediction. This was based on a sample of 10,000 data sets of 20 yearly extremes. The results obtained using the method of moments were compared with the corresponding results obtained using the maximum likelihood method. Although there were individual differences, the overall picture was one of very good agreement.

Table (3.1) compares the predicted values and confidence intervals for a selection of 20 over 200 simulated cases together with average values. It is observed that the average of the 200 predicted 100 -year return levels is slightly better for the ACER method than for both the POT and the Gumbel methods. More significantly, however, the 100 year return levels predicted by the ACER method vary from 4.44 to 5.22, whereas the same for the POT method vary from 4.12 - 5.62; for the Gumbel method based on the method of moments, the range is 4.27 - 5.77, and the range is 4.26 - 5.41 for the maximum likelihood estimation approach. Hence, in this case, the ACER method performs consistently better than both of these methods. It is also observed from the estimated 95% confidence intervals that the ACER method, as implemented in our research, provides higher accuracy than the other two methods. Finally, note that the confidence intervals of the 100 year return levels estimates by the ACER method obtained using either the simplified extrapolated confidence band approach or by non-parametric bootstrapping are very similar except for a slight mean shift. As a final comparison, the 200 bootstrapped confidence intervals obtained for the ACER method missed the target value $\eta^{100\text{yr}} = 4.80$ only three times; for the Gumbel method, six times; and for the POT method, the number of misses was 37 times for non-parametric bootstrapping and 35 times for parametric bootstrapping.

An example of the ACER plot and the results obtained for one set of data are presented in Figure (3.7). The predicted 100-year value is 4.85 with a predicted 95% confidence interval of (4.45, 5.09). Figure (3.8) presents POT predictions based on MLE for different thresholds in terms of the number n of data points above the threshold. The predicted value is 4.7 at $n = 204$, whereas the 95% confidence interval is (4.25, 5.28). The same data set as in Figure (3.7) was used. This was also used for the Gumbel plot shown in Figure (3.9). In this case, the predicted value based on the method of moments (MM) is $\eta_{\text{MM}}^{100\text{yr}} = 4.75$, with a parametric bootstrapped

No.	A $\hat{\eta}^{100}$	A CI	A BCI	G $\hat{\eta}^{100}_{\text{MM}}$	G BCI_{MM}	G $\hat{\eta}^{100}_{\text{MLE}}$	G BCI_{MLE}	GP $\hat{\eta}^{100}$	GP n.p.BCI	GP p.BCI	Exp $\hat{\eta}^{100}$
1	4.62	(4.41, 4.85)	(4.40, 5.02)	5.26	(4.60, 6.04)	5.08	(4.53, 5.62)	4.99	(4.08, 5.76)	(4.43, 5.84)	5.18
2	5.07	(4.64, 5.37)	(4.63, 5.54)	4.71	(4.31, 5.21)	4.81	(4.40, 5.20)	4.67	(4.10, 5.17)	(4.28, 5.08)	5.26
3	4.63	(4.27, 4.81)	(4.29, 4.93)	4.79	(4.32, 5.36)	4.92	(4.43, 5.40)	4.56	(4.17, 4.97)	(4.14, 5.00)	5.01
4	4.94	(4.44, 5.30)	(4.52, 5.34)	5.15	(4.54, 5.89)	4.98	(4.50, 5.44)	5.03	(4.19, 5.89)	(4.43, 5.90)	5.17
5	5.10	(4.61, 5.42)	(4.69, 5.43)	4.88	(4.41, 5.44)	5.09	(4.55, 5.60)	4.85	(4.37, 5.50)	(4.35, 5.46)	5.22
10	4.84	(4.34, 5.10)	(4.48, 5.18)	4.81	(4.36, 5.36)	4.81	(4.40, 5.23)	4.90	(4.40, 5.52)	(4.32, 5.67)	5.07
20	4.78	(4.46, 5.10)	(4.49, 5.29)	4.90	(4.40, 5.49)	5.06	(4.52, 5.57)	4.69	(4.24, 5.14)	(4.27, 5.20)	5.19
30	5.10	(4.50, 5.35)	(4.57, 5.47)	5.31	(4.69, 6.08)	5.41	(4.79, 6.00)	5.37	(4.65, 6.44)	(4.59, 6.49)	5.23
40	4.55	(4.27, 4.71)	(4.33, 4.86)	4.73	(4.28, 5.25)	4.73	(4.33, 5.16)	4.51	(4.14, 4.91)	(4.15, 4.92)	5.06
50	5.01	(4.37, 5.30)	(4.47, 5.36)	5.01	(4.46, 5.69)	5.06	(4.50, 5.57)	4.84	(4.42, 5.36)	(4.28, 5.47)	5.17
60	4.99	(4.66, 5.47)	(4.70, 5.47)	4.70	(4.35, 5.15)	4.88	(4.48, 5.29)	4.44	(4.25, 4.69)	(4.17, 4.63)	5.54
70	4.95	(4.53, 5.19)	(4.54, 5.33)	5.22	(4.63, 5.92)	5.24	(4.69, 5.78)	5.07	(4.45, 5.85)	(4.48, 5.79)	5.27
80	4.85	(4.41, 5.13)	(4.45, 5.25)	4.94	(4.42, 5.56)	4.82	(4.38, 5.23)	4.81	(4.13, 5.39)	(4.29, 5.39)	5.17
90	4.80	(4.35, 5.13)	(4.37, 5.04)	5.05	(4.49, 5.76)	5.10	(4.56, 5.64)	4.79	(4.31, 5.32)	(4.27, 5.43)	5.16
100	4.76	(4.42, 4.95)	(4.45, 5.19)	4.63	(4.26, 5.09)	4.84	(4.40, 5.26)	4.38	(4.19, 4.60)	(4.12, 4.61)	5.27
120	5.02	(4.52, 5.29)	(4.58, 5.41)	5.21	(4.62, 5.93)	5.21	(4.65, 5.75)	5.03	(4.31, 5.75)	(4.45, 5.70)	5.34
140	4.88	(4.41, 5.10)	(4.47, 5.18)	4.89	(4.43, 5.44)	5.01	(4.53, 5.47)	4.84	(4.37, 5.38)	(4.32, 5.49)	5.15
160	4.95	(4.48, 5.20)	(4.56, 5.28)	4.58	(4.23, 5.00)	4.77	(4.38, 5.16)	4.56	(4.23, 5.01)	(4.20, 4.92)	5.12
180	5.06	(4.55, 5.38)	(4.64, 5.48)	5.25	(4.65, 5.99)	5.05	(4.56, 5.51)	5.06	(4.35, 5.74)	(4.46, 5.72)	5.45
200	4.61	(4.26, 4.80)	(4.30, 4.98)	4.85	(4.36, 5.46)	4.86	(4.39, 5.35)	4.61	(4.16, 5.10)	(4.20, 5.08)	5.11
Av.200	**4.82**	(4.40, 5.09)	(4.47, 5.21)	**4.86**	(4.39, 5.44)	**4.88**	(4.43, 5.33)	**4.74**	(4.26, 5.27)	(4.27, 5.28)	**5.17**
Min 200	4.44			4.27		4.26		4.12			4.81
Max 200	5.22			5.77		5.41		5.62			5.54

Table 3.1: Return level estimates and 95% CI (BCI = CI by bootstrap) for three methods. A $\hat{\eta}^{100}$ - 100-years return level estimated by the ACER method; A CI - ACER confidence interval estimated by the simplified extrapolation approach; A BCI - ACER confidence interval estimated by the bootstrap; G $\hat{\eta}^{100}_{\text{MM}}$ - 100-yr return level estimated by the Gumbel method using the method of moments; G BCI_{MM} - 95% CI by bootstrap for the Gumbel method using the method of moments; G $\hat{\eta}^{100}_{\text{MLE}}$ - 100-yr return level estimate by the Gumbel method with maximum likelihood approach applied; G BCI_{MLE} - 95% CI by bootstrap for the Gumbel method using MLE; GP $\hat{\eta}^{100}$ - 100-yr return level estimate for the POT method based on the Generalized Pareto distribution; GP n.p. BCI - 95% CI by non-parametric bootstrap for the GP distribution; GP p. BCI - 95% CI by parametric bootstrap for the POT method; Exp $\hat{\eta}^{100}$ - 100-yr return level estimate for the POT method using Exponential distribution.

95% confidence interval of (4.34, 5.27). Prediction based on the Gumbel-Lieblein BLUE method (GL), cf., e.g., Cook (1985), is $\eta_{\mathrm{GL}}^{100\mathrm{yr}} = 4.73$, with a parametric bootstrapped 95% confidence interval equal to (4.35, 5.14).

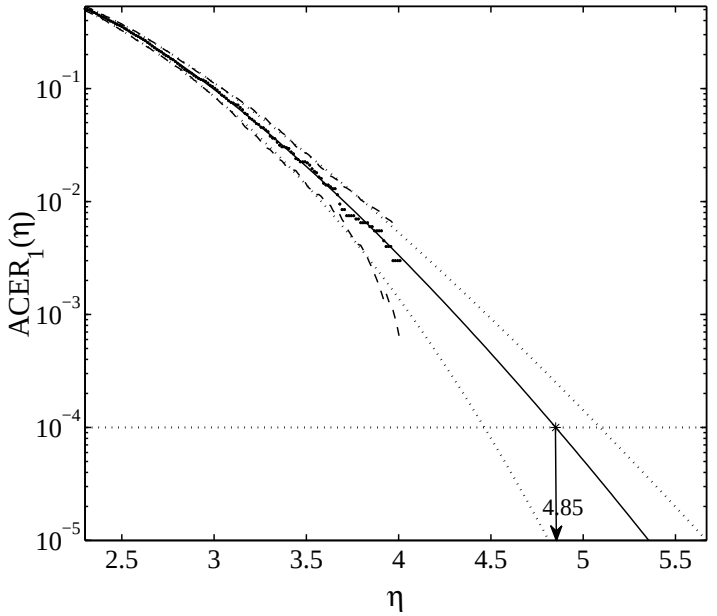


Figure 3.7: Synthetic data ACER $\hat{\varepsilon}_1$, Monte Carlo simulation ($*$); optimized curve fit (—); empirical 95% confidence band (- -); optimized confidence band ($\cdots$). Tail marker $\eta_1 = 2.3$

3.5.2 Wind speeds from 5 stations

Extreme wind speed prediction is an important issue for designing structures exposed to weather variations. Significant efforts have been devoted to the problem of predicting extreme wind speeds on the basis of data measured by various authors over several decades, see, for example, Perrin et al. (2006); Cook and Harris (2004); Harris (2001); Palutikof et al. (1999); Naess (1998a); Cook (1982) for extensive references to previous works.

In this section, we present the analysis of real wind speed data using the ACER method to obtain numerical estimates of extreme wind speeds. Measurements were provided by the Norwegian Meteorological Institute (2012) and downloaded from the Climate Data Web Services of the Institute. The hourly maximum of the 3 s wind gust (10 meters above the ground) was analysed for five weather stations off the coast of Norway: at A – Torsvåg Fyr weather station (station number 90800), at B – Hekkingen Fyr station (no.88690), at C – Nordøyan Fyr (station number 75410), at D – Sula station (no. 65940) and at E – Obrestad Fyr weather station (no. 44080). Figure (3.10) shows the geographical position of each station.

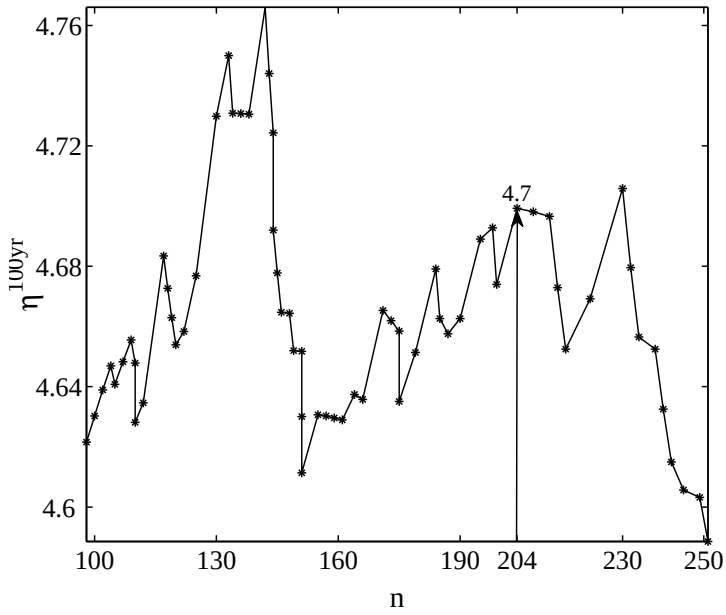


Figure 3.8: The point estimate $\tilde{\eta}^{100\text{yr}}$ of the 100-year return period value based on 20 years synthetic data as a function of the number n of data points above threshold. The return level estimate = 4.7 at $n = 204$.

Data were recorded for 13 years (1997-2010) at Torsvåg, for 14 years (1998-2012) at Hekkingen (station B), for 13 years (1999-2012) at station C, for 12 years (1998-2010) at Sula (station D), and for 16 years (1994-2010) at station E. The objective of the analysis is to estimate a 100-year return level of wind speed. Variations in the wind speed caused by seasonal variations in the wind climate during the year makes the wind speed a non-stationary process on the scale of months. Moreover, due to global climate change, yearly statistics may vary on the scale of years. The latter is, however, a slow process, and for the purpose of long-term prediction, we assume here that a quasi-stationary model of the wind speeds applies within a time span of 100 years. This may not be entirely true, however (Naess et al., 2013). For analysis of the data with the ACER method using the Matlab-based standalone downloadable application for general use that has been created for this purpose (Karpa, 2012), the data series were divided into one-year records. In this way, the standard deviation of the ACER function estimates can be calculated fairly accurately.

Figures (3.5.2) - (3.13) present plots of the time series observed from the Torsvåg, Sula and Obrestad stations, respectively.

Note that the samples from the Torsvåg Fyr and Obrestad Fyr stations contain outlying observations, such as 45.3 m/s in June 06, 1997; 43.7 m/s in May 10, 2001; and 60.8 m/s in September 09, 2008 for the Obrestad Fyr station and 45.3 m/s in July 12, 1998 and July 31, 1999 for the Torsvåg Fyr station. Such wind

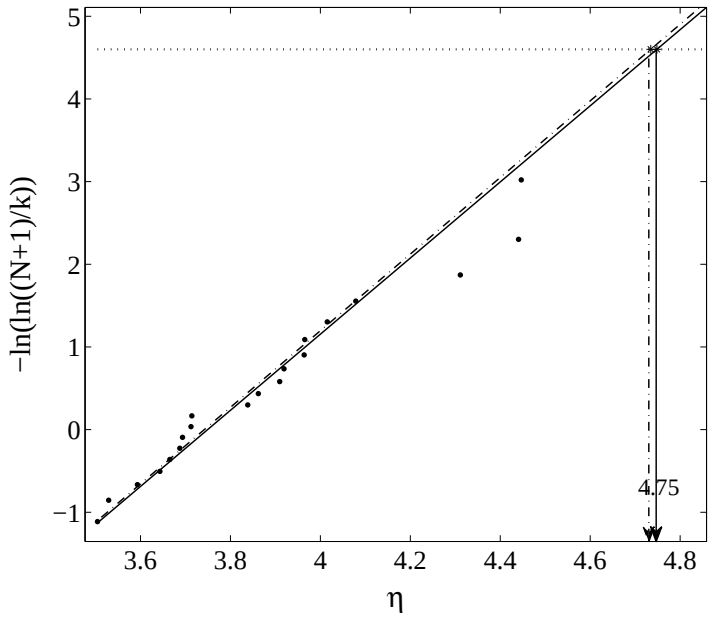


Figure 3.9: The point estimate $\tilde{\eta}^{100\text{yr}}$ of the 100-year return period value based on 20 years synthetic data. Lines are fitted by the method of moments – solid line (—) and the Gumbel-Lieblein BLUE method – dash-dotted lite (– · –). The return level estimate by the method of moments is 4.75, by the Gumbel-Lieblein BLUE method is 4.73

speeds are clearly spurious for the corresponding time periods and latitudes. Moreover, observations from the weather stations in the close vicinity of Obrestad Fyr confirm that no heavy storm occurred during the period in question, whereas no information from the stations in the vicinity of Torsvåg Fyr is available. Therefore, the outliers from the Obrestad Fyr station have to be rejected, whereas the outliers from Torsvåg Fyr are retained, primarily to show the invariance of the ACER method to observations of this type.

In Figures (3.5.2) - (3.18), $\hat{\varepsilon}_k(\eta)$ is plotted versus wind speeds η/σ for different values of k for five stations. The figures reveal that there is a significant dependence between consecutive data, which is clearly reflected in the effect of conditioning on previous data values.

Notably, the dependence effect is to some extent already captured by $\hat{\varepsilon}_2$, that is, by conditioning only on the value of the previous data point. Subsequent conditioning on more than one previous data point does not lead to substantial changes in ACER values, particularly for tail values. On the other hand, to fully elucidate the dependence structure of these data, it was necessary to carry the conditioning process to higher orders of k. It is clearly observed that the dependence between the data largely accounted for by $k = 48$ because there is a marked degree of convergence in the tail of $\hat{\varepsilon}_k(\eta)$ for $k \geq 48$. Note that $k = 48$ clearly corresponds

to exceedances separated by two days for hourly observations. Here, $\hat{\varepsilon}_{96}$, that is, four-days declustered data, is considered to represent the final converged results. This means that $\hat{\varepsilon}_{96} \approx \hat{\varepsilon}_k$ in the tail for $\forall k > 96$; thus, for any practical application, there is no need to consider conditioning of an even higher order than 96.

Clearly, the most important information for the prediction of a 100-year value is provided by the far tail the ACER function. Therefore, from a practical perspective, for extreme value estimation, $\hat{\varepsilon}_1(\eta)$ can be used. The reason for this is that Figures (3.5.2) - (3.18) reveal that all of the ACER functions converge in the tail.

Figure 3.10: Map of Norway with marked weather stations in decreasing order from the top: A – Torsvåg Fyr weather station (station number 90800); B – Hekkingen Fyr (88690); C – Nordøyan Fyr (75410); D – Sula station (65940); E – Obrestad Fyr (44080).

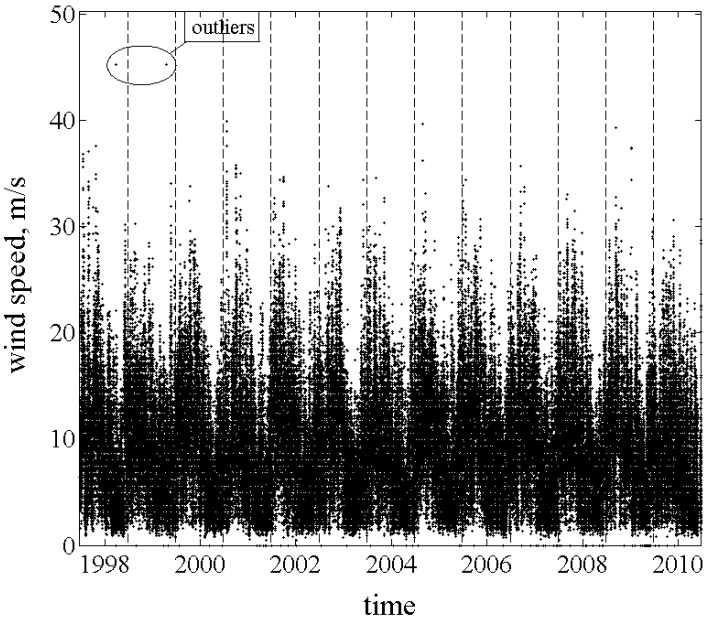


Figure 3.11: Observations from Torsvåg Fyr station (A)

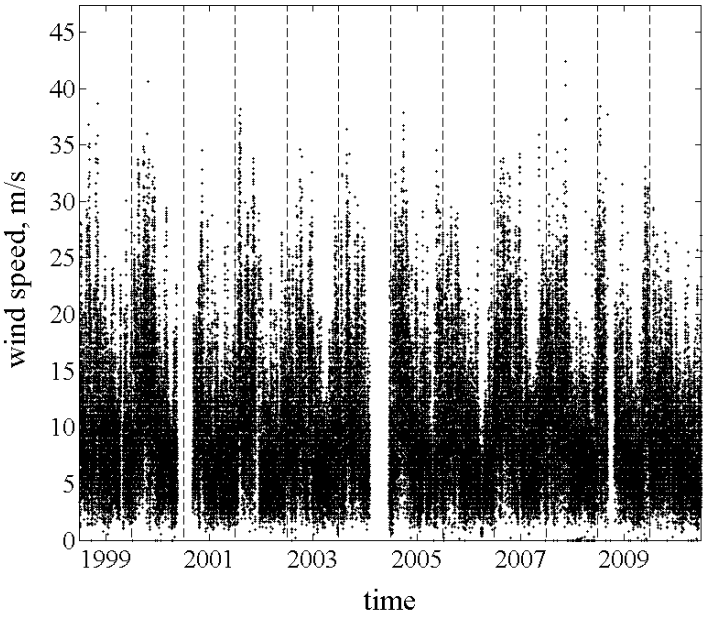


Figure 3.12: Observations from Sula station (D)

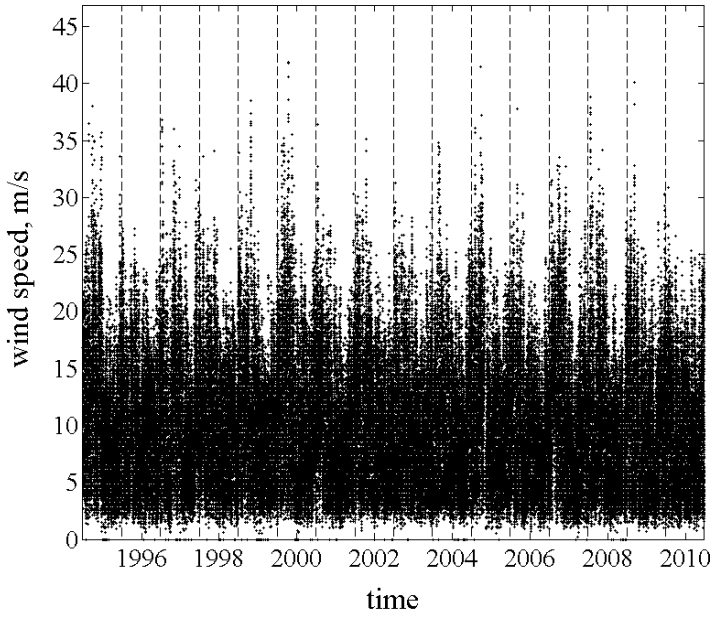


Figure 3.13: Observations from Obrestad Fyr station (E)

This clearly demonstrates the power of an ACER function plot as a diagnostic tool to determine the value of k required for extreme value estimation in a particular case. Although there are significant dependence effects for the lower wind speeds, for the extreme wind speeds, these effects are largely absent. This makes it possible to choose $k = 1$, which makes considerably more data available for estimation, with a possible reduction of uncertainty in estimation as a result.

In Table (3.2), the results for the parametric estimation of the 100-year return value using the ACER program (Karpa, 2012) and its 95% CI are listed together with the optimal parameters q, b, a and c of the $\varepsilon_1(\eta)$ curve.

Note that the estimated values of the shape parameter c are typical for the underlying statistics of wind speeds for Norway considered in the present work.

Figures (3.5.2) - (3.23) present the results from the parametric estimation of the return value and its 95% CI for $\hat{\varepsilon}_1(\eta)$ for each station.

The annual maxima method is applied to the wind gust data to compare the estimated 100-year return level values. The Gumbel estimate $\hat{\eta}^{100yr}$ is based on the method of moments (MM) and the Gumbel-Lieblein BLUE method (GL), cf., e.g., Cook (1985). A computer program was written in the Matlab language to implement both methods. Figures (3.5.2) - (3.28) present the observed yearly extremes extracted from the hourly data together with fitted straight lines on the Gumbel probability plot. Here, the 100-year return level values for the first station with outliers included are $\hat{\eta}^{100yr}_{MM} = 51.33$ m/s and $\hat{\eta}^{100yr}_{GL} = 51.57$ m/s, whereas in the case of rejected outlying observations, $\hat{\eta}^{100yr}_{MM} = 44.31$ m/s and $\hat{\eta}^{100yr}_{GL} = 45.84$ m/s. For the Hekkingen Fyr and Nordøyan Fyr stations, $\hat{\eta}^{100yr}_{MM} = 58.1$ m/s with

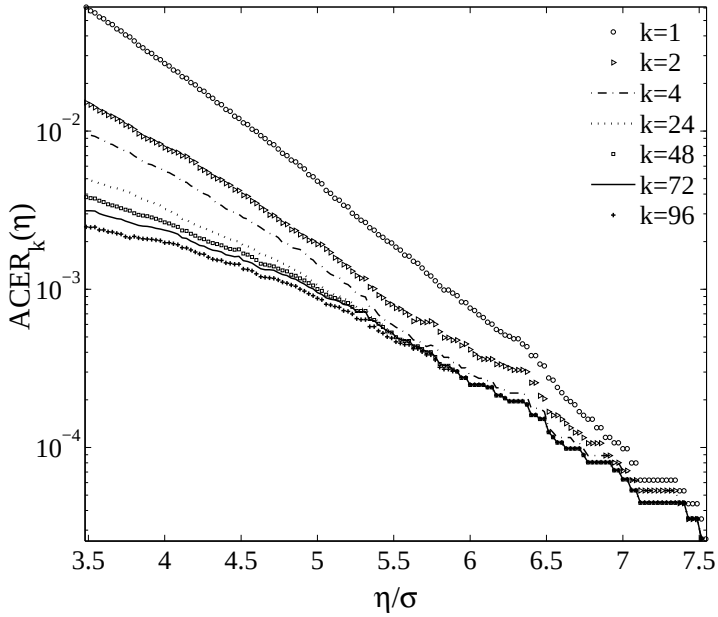


Figure 3.14: A - Torsvåg Fyr wind speed statistics, 13 years of hourly data; outliers are included. Comparison between ACER estimates for different degrees of conditioning; $\sigma = 5.30$ m/s.

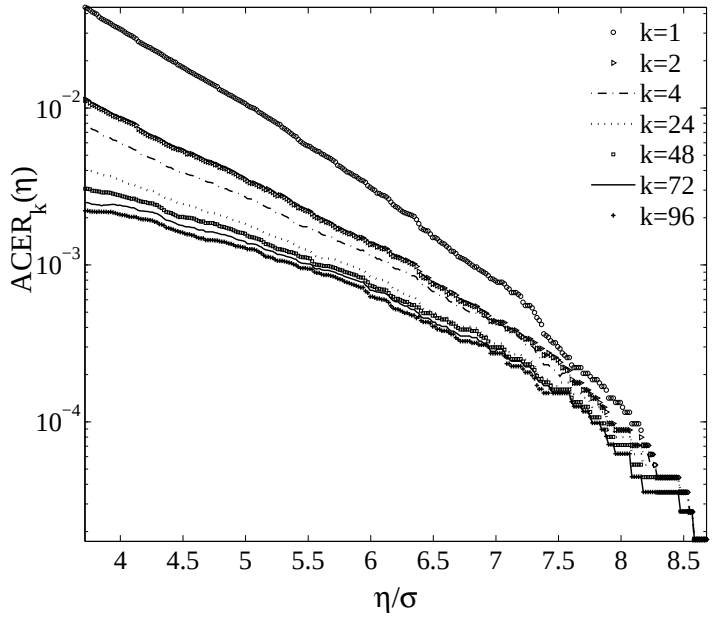


Figure 3.15: B - Hekkingen wind speed statistics, 14 years of hourly data. Comparison between ACER estimates for different degrees of conditioning; $\sigma = 5.72$ m/s.

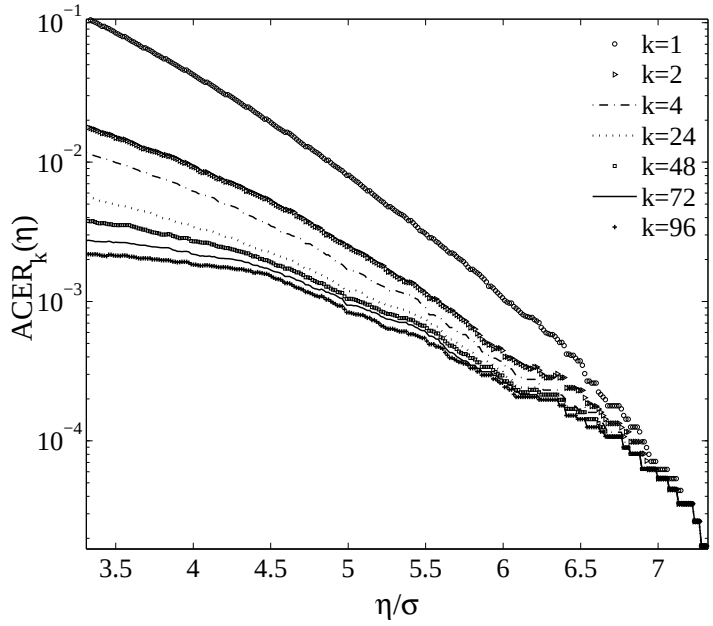


Figure 3.16: C - Nordøyan wind speed statistics, 13 years of hourly data. Comparison between ACER estimates for different degrees of conditioning; $\sigma = 6.01$ m/s.

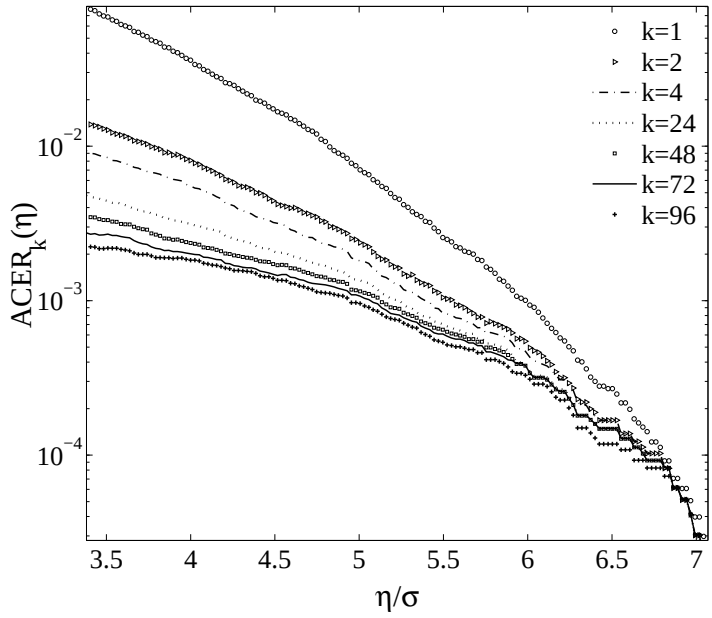


Figure 3.17: D - Sula wind speed statistics, 12 years of hourly data. Comparison between ACER estimates for different degrees of conditioning; $\sigma = 5.49$ m/s.

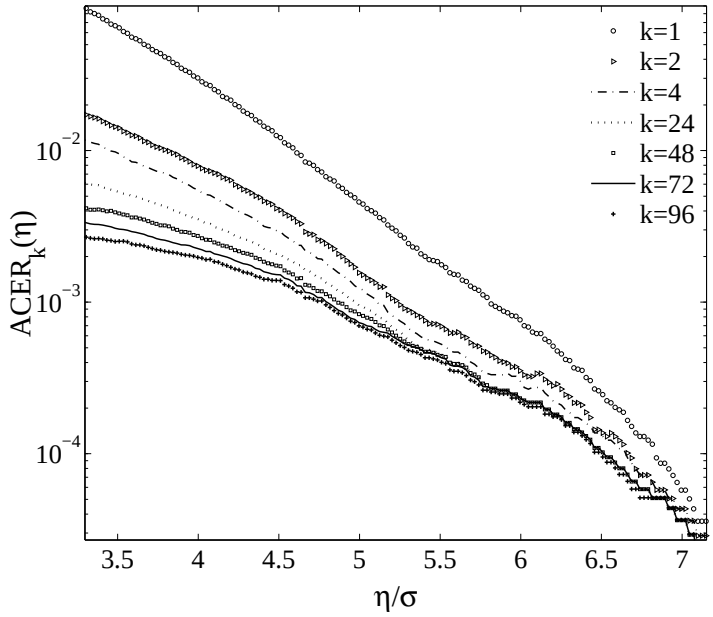


Figure 3.18: E - Obrestad Fyr wind speed statistics based on 16 years of hourly data. Comparison between ACER estimates for different degrees of conditioning; $\sigma = 5.47$ m/s.

Station	$\hat{\eta}^{100yr}$	95% CI ($\hat{\eta}^{100yr}$)	optimal parameters			
			q^*	b^*	a^*	c^*
A - Torsvåg; incl. outliers	47.5	(42.1, 50.7)	0.44	9.02	0.1	1.33
A - Torsvåg; without outliers	47.2	(39.9, 50.6)	0.47	8.49	0.09	1.36
B - Hekkingen	60.5	(53.1, 64.9)	0.27	0	0.008	1.78
C - Nordøyan	51.9	(48.4, 53.1)	1.02	0	0.008	1.9
D - Sula	46.3	(43.4, 47.8)	0.58	0	0.005	2.07
E - Obrestad	48.4	(43.2, 50.7)	0.29	12.34	0.13	1.27

Table 3.2: Results of the optimization procedure used to estimate the 100-year return value by the ACER method for $\hat{\varepsilon}_k(\eta)$ with $k = 1$ for all weather stations.

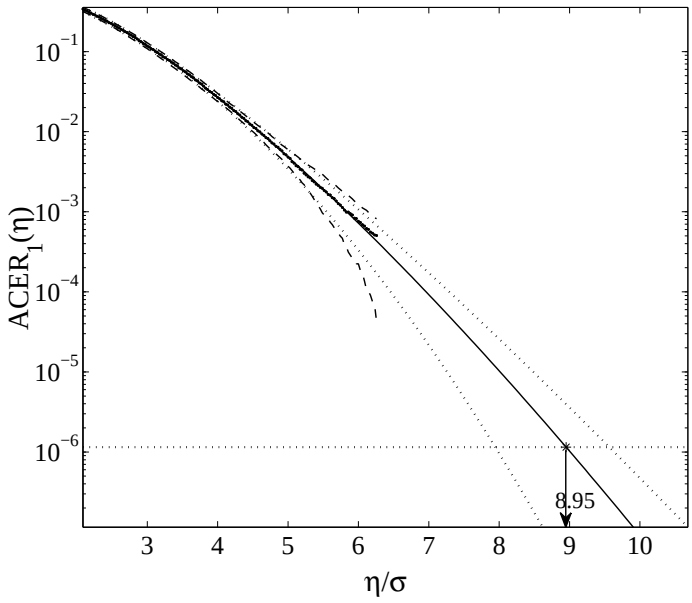


Figure 3.19: A - Torsvåg Fyr: plot of $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); optimized curve fit (—); empirical 95% CI (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 2.07\,\sigma$; $\sigma = 5.30$ m/s.

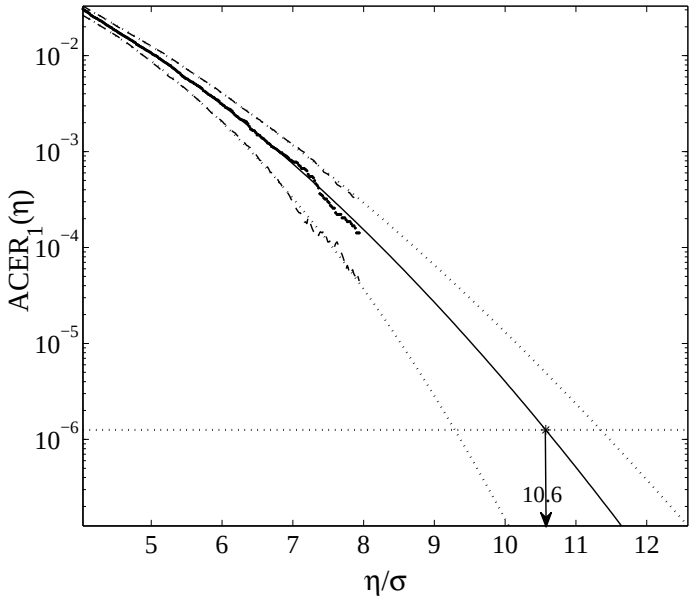


Figure 3.20: B - Hekkingen Fyr: plot of $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); optimized curve fit (—); empirical 95% CI (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 4.02\,\sigma$; $\sigma = 5.72$ m/s.

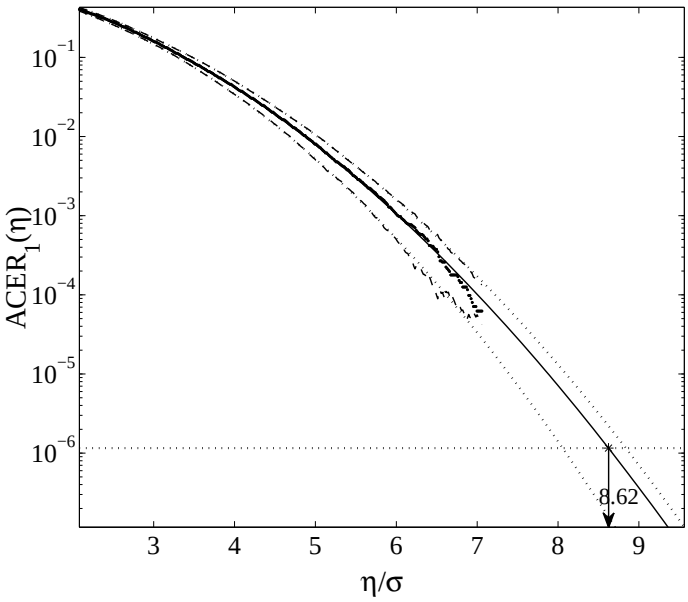


Figure 3.21: C - Nordøyan Fyr: plot of $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); otimized curve fit (—); empirical 95% CI (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 2.08\,\sigma$; $\sigma = 6.01$ m/s.

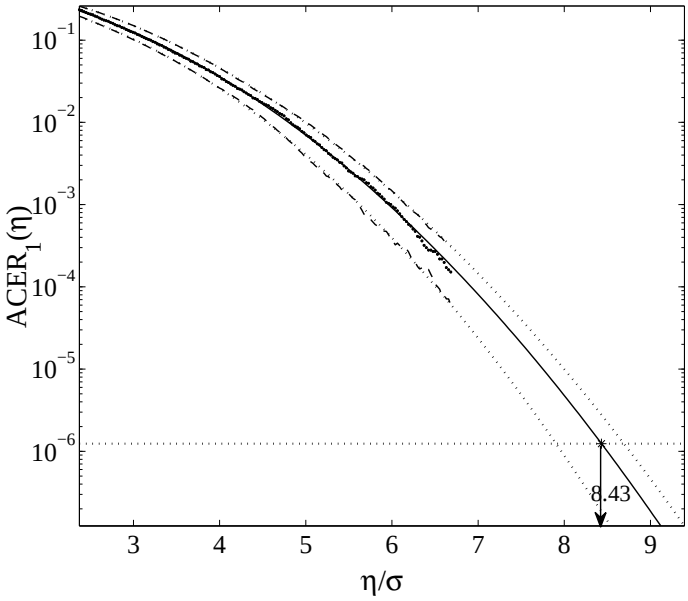


Figure 3.22: D - Sula: plot of $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); optimized curve fit (—); empirical 95% CI (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 2.36\,\sigma$; $\sigma = 5.49$ m/s.

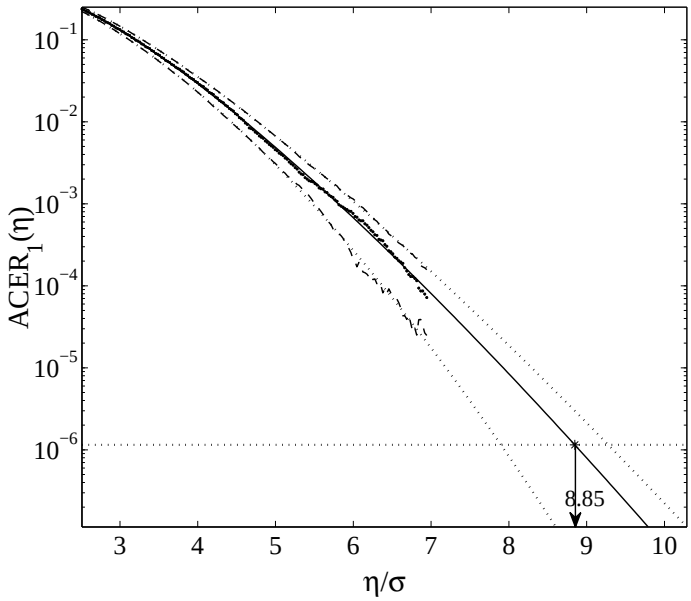


Figure 3.23: E - Obrestad Fyr: plot of $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); optimized curve fit (—); empirical 95% CI (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 2.5\,\sigma$; $\sigma = 5.47$ m/s.

$\hat{\eta}_{GL}^{100yr} = 60.63$ m/s and $\hat{\eta}_{MM}^{100yr} = 51.5$ m/s with $\hat{\eta}_{GL}^{100yr} = 55.5$ m/s, respectively. Finally, for the Sula and Obrestad Fyr stations, the 100-year return level values are $\hat{\eta}_{MM}^{100yr} = 48.66$ m/s with $\hat{\eta}_{GL}^{100yr} = 52.9$ m/s and $\hat{\eta}_{MM}^{100yr} = 48.59$ m/s with $\hat{\eta}_{GL}^{100yr} = 53.79$ m/s, respectively.

Although the Gumbel-Lieblein BLUE method is considered to be one of the best available conventional Gumbel methods, the application of the GL method requires tables of the BLUE coefficients that are not available for annual data with a sample size $N > 25$, cf., Harris (2001). The observed results reveal the sensitivity of this method to outliers, as well as for the method of moments. Additionally, note that the Gumbel-Lieblein BLUE method appears to have a tendency to overestimate the predicted return level values, whereas the method of moments appears to be reasonably stable for the studied sets of data (Karpa and Naess, 2013).

The POT method was also applied to the wind gust time series. Following WAFO-group (2000), the data were declustered beforehand. Declustering was performed in such a way that peak events separated by 3.5 days or more were extracted from the measured data and selected for the analysis to achieve approximate independence of the exceedances (Naess, 1998b). Figures (3.5.2) - (3.33) present POT estimates $\hat{\eta}^{100yr}$ for different threshold numbers based on the MLE, cf., e.g., Coles (2001). Estimations were obtained using the Matlab (2009) Statistics Toolbox routine gpfit. It is interesting to observe the unstable characteristics of the estimates over a range of threshold values, whereas they are quite stable on either side of this

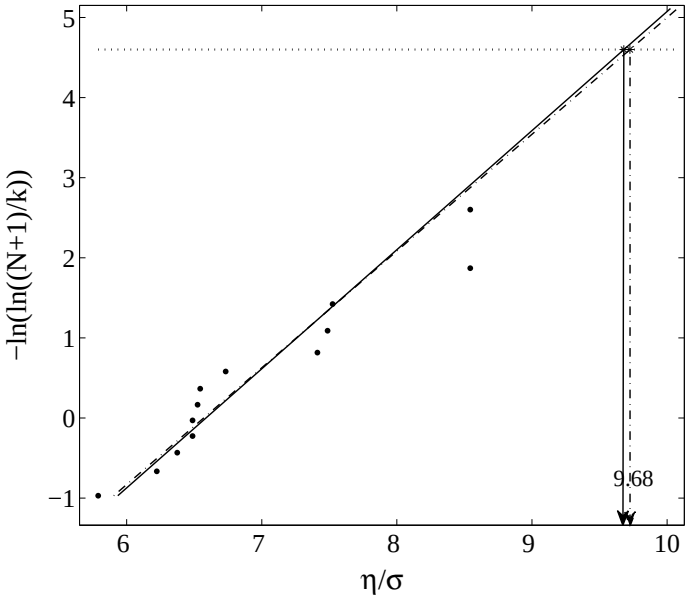


Figure 3.24: A - Torsvåg Fyr: the point estimate $\hat{\eta}^{100yr}$ of the 100-year return period value by the Gumbel method. Lines are fitted by the method of moments (—) and the Gumbel-Lieblein BLUE method (– · –); $\sigma = 5.30$ m/s.

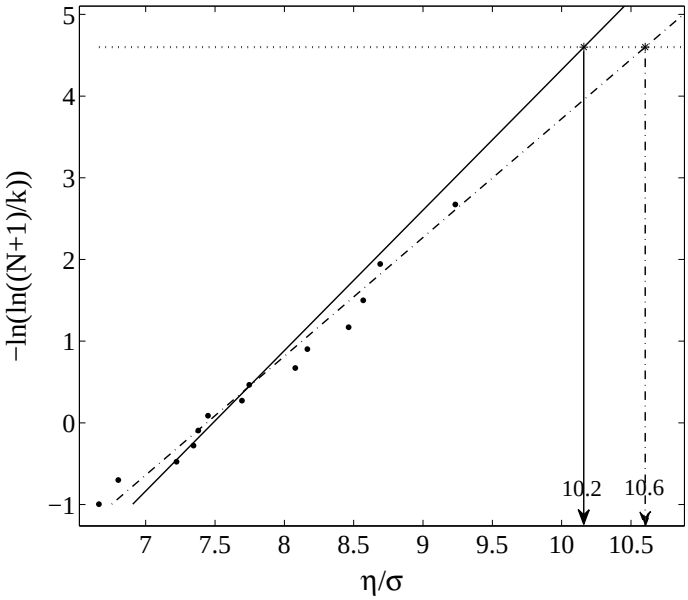


Figure 3.25: B - Hekkingen Fyr: the point estimate $\hat{\eta}^{100yr}$ of the 100-year return period value by the Gumbel method. Lines are fitted by the method of moments (—) and the Gumbel-Lieblein BLUE method (– · –); $\sigma = 5.72$ m/s.

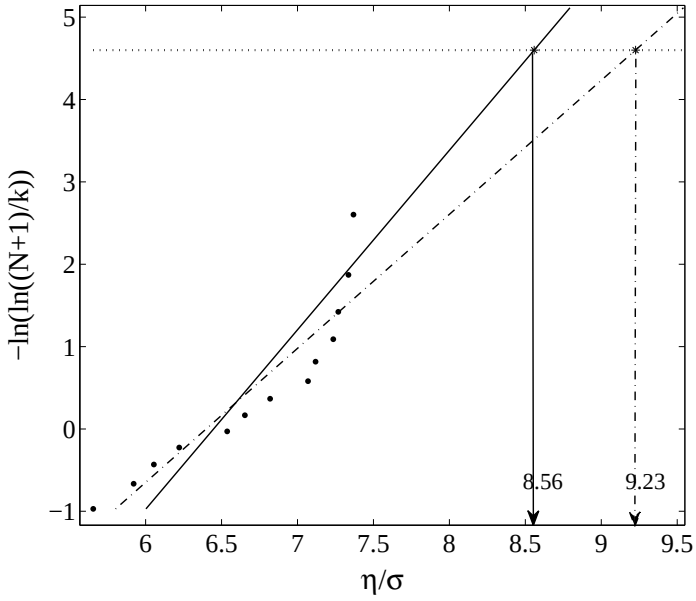


Figure 3.26: C - Nordøyan Fyr: the point estimate $\hat{\eta}^{100yr}$ of the 100-year return period value by the Gumbel method. Lines are fitted by the method of moments (—) and the Gumbel-Lieblein BLUE method (−·−); $\sigma = 6.01$ m/s.

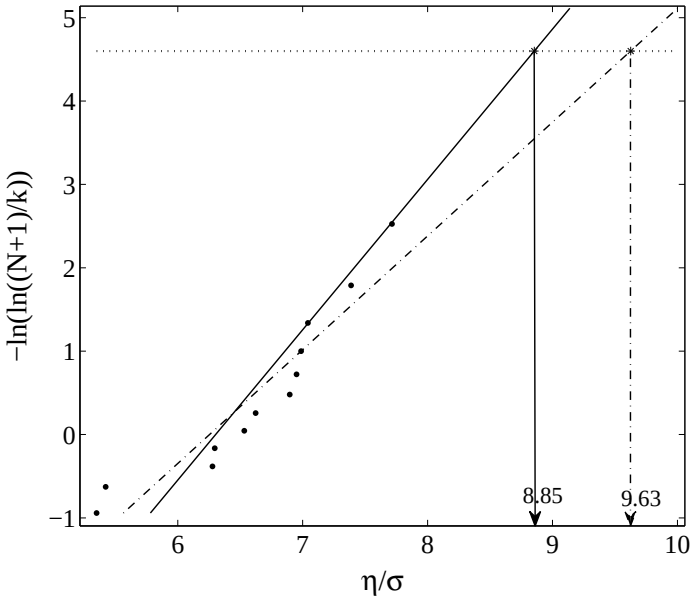


Figure 3.27: D - Sula: the point estimate $\hat{\eta}^{100yr}$ of the 100-year return period value by the Gumbel method. Lines are fitted by the method of moments (—) and the Gumbel-Lieblein BLUE method (−·−); $\sigma = 5.49$ m/s.

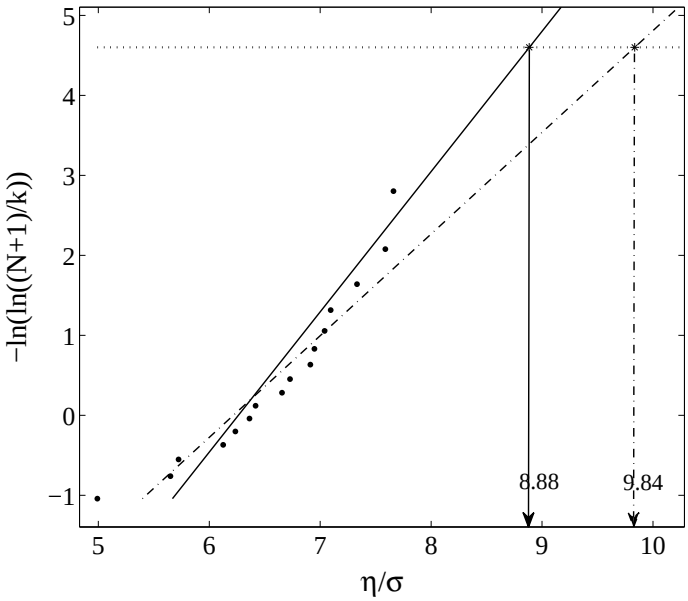


Figure 3.28: E - Obrestad Fyr: the point estimate $\hat{\eta}^{100yr}$ of the 100-year return period value by the Gumbel method. Lines are fitted by the method of moments (—) and the Gumbel-Lieblein BLUE method (– · –); $\sigma = 5.47$ m/s.

range, providing predictions that are more in line with the results from the other two methods.

In Tables (3.3) - (3.4), the 100-year return period values are listed together with the predicted 95% confidence intervals for all methods and each station. For the case of the annual maxima method, the 95% confidence intervals are estimated from a parametric bootstrapping of the Gumbel estimates based on a sample of 10,000 data sets of 13, 14, 13, 12 and 16 yearly extremes. For the POT method, the bootstrapped 95% confidence intervals are estimated using Matlab (2009) Statistics Toolbox routine bootstrp. Ten thousand samples are generated by sampling with replacement from the observed exceedances above a high threshold.

3.5.3 Bootstrapping for ACER

In this section, we discuss the performance of the ACER scheme in estimating the 95% CI on the basis of the observed real environmental data. Previously, the 100-year return wind speeds from the five weather stations and their confidence intervals were determined through ACER analysis with different degrees of conditioning. As mentioned in Section 3.4.4, estimation of the 95% CI for the predicted extreme value by the extrapolation of confidence bands is a rather simple first estimation of a confidence interval. Of course, the bootstrapping methods would provide more reasonable and accurate estimates. However, this primarily applies to synthetic data with a known underlying distribution because then a bundle of

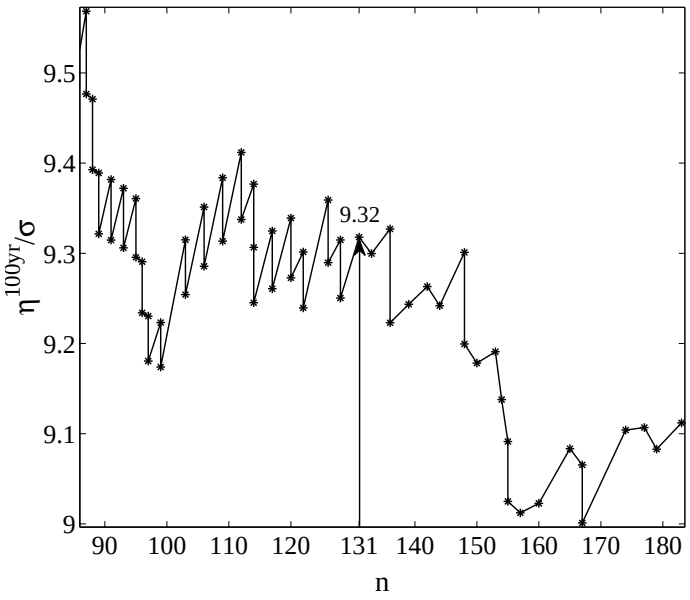


Figure 3.29: A - Torsvåg Fyr, POT approach: the point estimate $\tilde{\eta}^{100yr}$ of the 100-year return period value as a function of the number n of data points above threshold. The return level estimate 49.41 m/s is at $n = 131$; $\sigma = 5.30$ m/s.

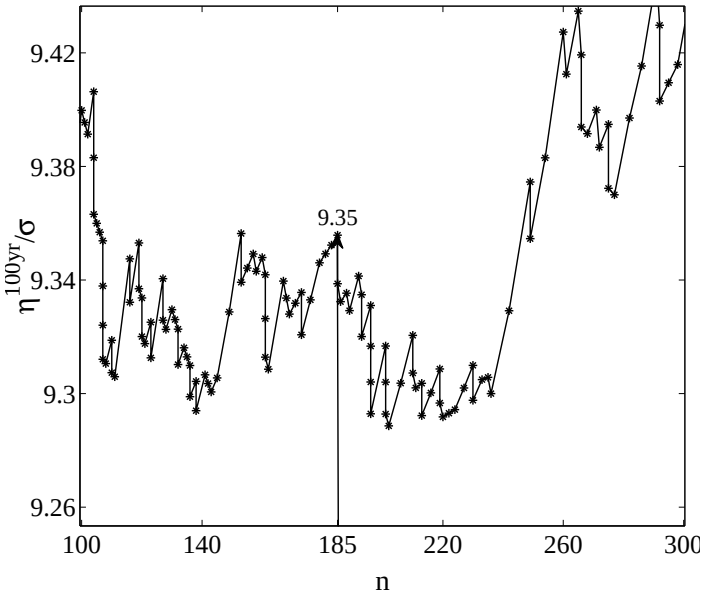


Figure 3.30: B - Hekkingen Fyr, POT approach: the point estimate $\tilde{\eta}^{100yr}$ of the 100-year return period value as a function of the number n of data points above threshold. The return level estimate 53.48 m/s is at $n = 185$; $\sigma = 5.72$ m/s.

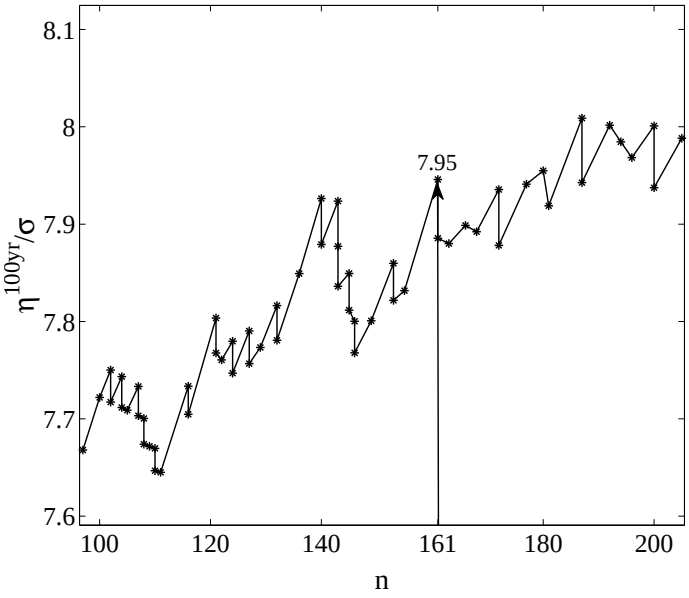


Figure 3.31: C - Nordøyan Fyr, POT approach: the point estimate $\tilde{\eta}^{100yr}$ of the 100-year return period value as a function of the number n of data points above threshold. The return level estimate 47.8 m/s is at $n = 161$; $\sigma = 6.01$ m/s.

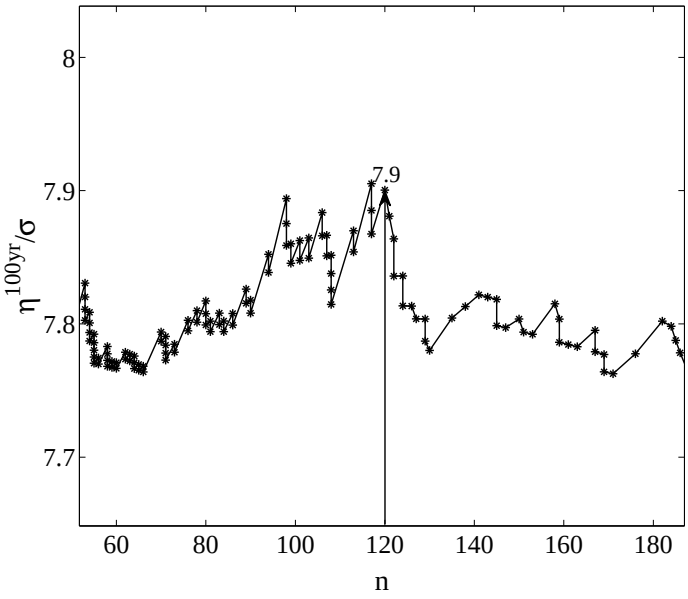


Figure 3.32: D - Sula, POT approach: the point estimate $\tilde{\eta}^{100yr}$ of the 100-year return period value as a function of the number n of data points above threshold. The return level estimate 43.42 m/s is at $n = 120$; $\sigma = 5.49$ m/s.

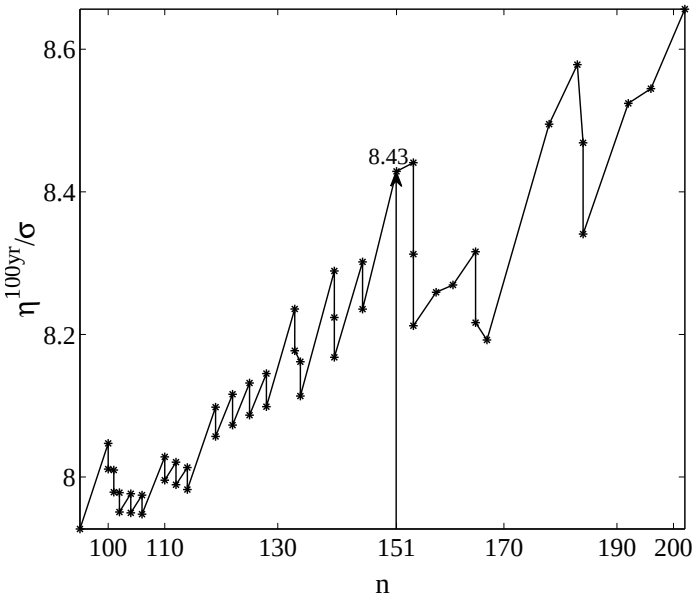


Figure 3.33: E - Obrestad Fyr, POT approach: the point estimate $\tilde{\eta}^{100yr}$ of the 100-year return period value as a function of the number n of data points above threshold. The return level estimate 46.1 m/s is at $n = 151$; $\sigma = 5.47$ m/s.

replicate time series can be generated and the corresponding ACER functions can be estimated and extrapolated to the extreme value levels. Nevertheless, it is important to examine the possible methods for constructing confidence intervals from the bootstrap for the observed data. Parametric and non-parametric models will be considered to compare the results.

Non-parametric bootstrap

To perform the non-parametric bootstrap in our case, it is natural to use a sampling with replacement from the set of observed data Efron and Tibshirani (1993). The scheme implies an equal occurrence probability for any observation, which means that the observations are independent. This inevitably involves that only the case of ACER analysis with no conditioning, e.g. $\varepsilon_k(\eta)$ with $k = 1$, and its results can be considered here.

To perform the analysis, 1000 replicated samples were generated for each of the five weather stations. Note that for clarity, the values of the tail marker η_1 and the parameter δ were kept the same as for the ACER analysis of the original time series. After generating a sample of size 10^3 of 100-year return levels estimated using the ACER method, the right and left ends of the 95% confidence interval are defined by the 2.5th and 97.5th percentiles, respectively.

In Table (3.5), the bootstrap-estimated 95% confidence intervals (95% BCI) in the second column are listed together with the average of 10^3 MC $\hat{\eta}^{100yr}$ in the

third column for all stations. For comparison purposes, the initial results of the ACER analysis are shown in the fourth and fifth columns (cf. Table (3.2)).

As shown in Table (3.5), the results were as expected: the mean value of the 1000 samples of the 100-year return period is close to the initial $\hat{\eta}^{100yr}$, and the confidence interval estimated by the ACER extrapolation is approximately 40%

Station	Method	Spec	$\hat{\eta}^{100yr}$, m/s	95% CI ($\hat{\eta}^{100yr}$), m/s
A - Torsvåg; incl. outliers	ACER, various k	1	47.46	(42.11, 50.71)
		2	48.18	(41.48, 51.31)
		4	46.96	(42.25, 49.63)
		24	48.36	(43.44, 51.63)
		48	47.54	(43.46, 49.75)
		72	47.44	(44.39, 48.79)
		96	48.78	(44.53, 51.61)
	Annual maxima	MM	51.33	(43.08, 61.57)
		GL	51.57	(44.24, 60.67)
	POT	–	49.41	(40.95, 59.42)
A - Torsvåg; without outliers	ACER, various k	1	47.21	(39.94, 50.60)
		2	47.79	(41.13, 50.93)
		4	46.32	(42.00, 49.04)
		24	47.22	(43.26, 50.04)
		48	46.38	(43.60, 48.19)
		72	46.32	(44.24, 47.37)
		96	47.80	(44.45, 49.95)
	Annual maxima	MM	44.31	(39.36, 50.39)
		GL	45.84	(40.72, 52.41)
	POT	–	42.62	(39.01, 47.31)
B - Hekkingen	ACER, various k	1	60.47	(53.1, 64.9)
		2	62.23	(53.3, 70.0)
		4	63.03	(53.0, 74.5)
		24	60.63	(51.3, 70.7)
		48	60.44	(51.3, 77.0)
		72	58.06	(51.2, 66.4)
		96	59.19	(52.0, 68.3)
	Annual maxima	MM	58.10	(50.8, 67.3)
		GL	60.63	(53.0, 70.1)
	POT	–	53.48	(48.9, 57.0)

Table 3.3: Predicted 100-year return period levels for Torsvåg Fyr and Hekkingen Fyr weather stations by the ACER-method for different degrees of conditioning, Annual maxima and POT methods, respectively.

shorter than the 95% BCI due to the higher uncertainty related to the sampling with replacement. In addition, the ACER 95% CI is slightly shifted to the left, which is caused by the behaviour of the $CI^{-}(\eta)$ curve in the tail.

The relative difference and visual comparison of the 95% CI estimated by both methods is presented in Table (3.6). The initial 95% CI estimated by the ACER

Station	Method	Spec	$\hat{\eta}^{100yr}$, m/s	95% CI ($\hat{\eta}^{100yr}$), m/s
C - Nordøyan	ACER, various k	1	51.85	(48.4, 53.1)
		2	51.48	(46.1, 54.1)
		4	52.56	(46.7, 55.7)
		24	52.90	(47.0, 56.2)
		48	54.62	(47.7, 57.6)
		72	53.81	(46.9, 58.3)
		96	54.97	(47.5, 60.5)
	Annual maxima	MM	51.5	(45.2, 59.3)
		GL	55.5	(48.0, 64.9)
	POT	–	47.8	(44.8, 52.7)
D - Sula	ACER, various k	1	46.33	(43.41, 47.77)
		2	46.81	(44.08, 49.04)
		4	47.99	(44.80, 50.57)
		24	46.65	(44.10, 48.07)
		48	46.83	(44.28, 48.03)
		72	45.80	(43.01, 46.96)
		96	45.69	(42.32, 47.01)
	Annual maxima	MM	48.66	(41.58, 57.58)
		GL	52.90	(44.29, 63.39)
	POT	–	43.42	(39.07, 47.80)
E - Obrestad	ACER, various k	1	48.38	(43.18, 50.74)
		2	48.11	(42.38, 50.69)
		4	48.81	(42.34, 51.59)
		24	47.90	(42.87, 50.53)
		48	48.90	(43.82, 50.72)
		72	49.47	(44.06, 51.52)
		96	48.55	(43.46, 49.96)
	Annual maxima	MM	48.59	(42.10, 56.84)
		GL	53.79	(46.16, 63.53)
	POT	–	46.10	(41.00, 55.00)

Table 3.4: Predicted 100-year return period levels for Nordøyan Fyr, Sula and Obrestad Fyr weather stations by the ACER-method for different degrees of conditioning, Annual maxima and POT methods, respectively.

Station	95% BCI	Av. 10^3	A 95% CI	$\hat{\eta}^{100yr}$
A - Torsvåg; incl. outliers	(42.78, 57.80)	48.29	(42.11, 50.71)	47.46
A - Torsvåg; without outliers	(42.58, 51.46)	47.18	(39.94, 50.60)	47.21
B - Hekkingen	(53.93, 75.04)	62.52	(53.12, 64.92)	60.47
C - Nordøyan	(48.69, 57.08)	52.32	(48.43, 53.14)	51.85
D - Sula	(44.64, 50.80)	47.23	(43.41, 47.77)	46.33
E - Obrestad	(42.54, 53.73)	48.12	(43.18, 50.74)	48.38

Table 3.5: Comparison table for 95%CI by the non-parametric bootstrap and the ACER extrapolation of confidence bands.

method together with the estimated $\hat{\eta}^{100yr}$ are represented by red lines, whereas the bootstrapped 95% CI is shown as a blue line. Note that in the case of the analysis of data from the Torsvåg station without outliers, the initial A 95% CI is 20% wider (-20% in the table) than the bootstrapped CI.

Thus, the results of the analysis reveal that estimation of the 95% CI through the extrapolation of confidence bands using the ACER scheme can generally provide a reasonable estimate. Although this method is a first-stage estimation, it is not as time consuming as the bootstrap method and provides 95% CI estimation for any degree of conditioning k.

Parametric bootstrap

Now let us discuss possible approaches for the parametric bootstrap based on the parametrically fitted ACER curve. As derived in Section 3.2 and Section 3.3, the extreme value distribution $F(\eta)$ associated with the observed time series is represented by the average conditional exceedance rates as follows

$$\begin{aligned} F(\eta) &\approx \exp\Big[-(N-k+1)\,\varepsilon_k(\eta)\Big] \\ &\approx \exp\Big[-q_k\,(N-k+1)\,\exp\big\{ -a_k(\eta-b_k)^{c_k}\big\}\Big]\,,\ \ \eta\geq\eta_1\geq b\,. \end{aligned} \tag{3.64}$$

where the exactness of $F(\eta)$ is provided by conditioning on previous $k-1$ data points, which captures the essential structure of the given time series. Clearly, Eq.(3.64) does not provide us with any opportunities for the prompt deduction of the explicit form of the underlying distribution $F_x(\eta)$ or for revealing the dependence relation of the data that we need. Therefore, the only way to progress in this study is to assume independence of the data and to provide the analysis solely for $\varepsilon_k(\eta)$, $k=1$. The assumption yields:

$$\Big[F_x(\eta)\Big]^N = F(\eta) \approx \exp\Big[-q\,(N-k+1)\,\exp\big\{ -a(\eta-b)^c\big\}\Big]\,,\ \ \eta\geq\eta_1\geq b\,. \tag{3.65}$$

Thus, a general form of the cumulative distribution function would be :

$$F_x(\eta) \approx \exp\Big[- q \, \exp\big\{ - a(\eta - b)^c \big\}\Big] \,, \;\; \eta \geq \eta_1 \geq b \,. \tag{3.66}$$

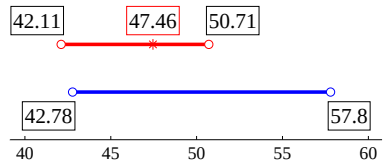
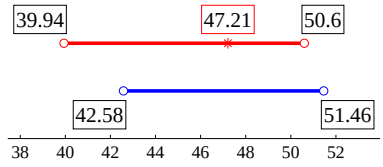
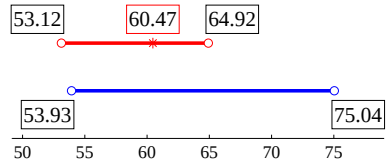
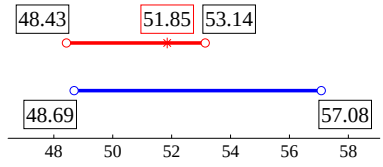
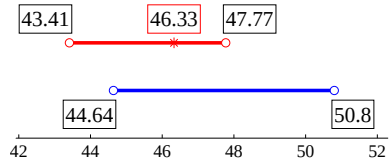
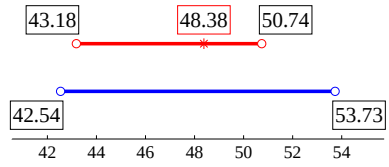
Station	Rel.diff	Intervals comparison
A - Torsvåg; incl. outliers	43%	
A - Torsvåg; without outliers	-20%	
B - Hekkingen	44%	
C - Nordøyan	44%	
D - Sula	29%	
E - Obrestad	32%	

Table 3.6: Comparison table for 95%CI by the non-parametric bootstrap and the ACER extrapolation. Relative difference is displayed in the second column. Red intervals on top visualize the initial ACER CI with marked $\hat{\eta}^{100yr}$; blue intervals from bottom – non-parametrically bootstrapped CI.

It will be referred to as 'Case a' of the parametric bootstrap. To be able to generate samples above the tail marker η_1, we modify CDF $F_X(\eta)$ to obtain a monotone distribution function $\tilde{F}_X(\eta)$, $\eta \geq \eta_1$, such that $\tilde{F}_X(\eta_1) = 0$, as follows:

$$\tilde{F}_X(\eta) = \frac{F_X(\eta) - F_X(\eta_1)}{1 - F_X(\eta_1)} \tag{3.67}$$

In addition to the approximation (3.66) and modification (3.67), it was decided to apply substitution $\exp(-x) \approx 1 - x$ in Eq. (3.66) to obtain the following representation (referred to as 'Case b'):

$$F'_X(\eta) \approx 1 - q \, \exp\{-a(\eta - b)^c\}\,, \;\; \eta \geq \eta_1 \geq b\,, \tag{3.68}$$

with the corresponding modification

$$\tilde{F'}_X(\eta) = \frac{F'_X(\eta) - F'_X(\eta_1)}{1 - F'_X(\eta_1)} = 1 - \exp\{-a\left((\eta - b)^c - (\eta_1 - b)^c\right)\}. \tag{3.69}$$

As we can see, the parameter q is eliminated in Case b. This was indeed the reason to employ Case b because we assumed that the experimental uncertainty of the bootstrap could potentially be reduced.

Finally, by applying inverse transformations to Eqs. (3.67) and (3.69), we obtain expressions to be used directly in the generation of samples. Thus, the inverse for Eq. (3.67) for 'Case a' has the form

$$\eta = \left\{ -\frac{1}{a} \log \left[-\frac{1}{q} \log\Big(U \cdot \big(1 - F_X(\eta_1)\big) + F_X(\eta_1) \Big) \right] \right\}^{1/c} + b\,, \tag{3.70}$$

where U is a random variable uniformly distributed at [0, 1], and $F_X(\eta_1)$ is value of the CDF (3.67) at the point η_1. To generate samples from the CDF (3.69) for 'Case b', the corresponding inverse is given by the equation:

$$\eta = \left[-\frac{1}{a} \log\left(1 - U\right) + (\eta_1 - b)^c \right]^{1/c} + b\,, \tag{3.71}$$

with the random variable U uniformly distributed at [0, 1].

Finally, the sample size of a synthetic time series should be defined. For this purpose, it was naturally decided to use the original sample with respect to the number of observations above the tail marker η_1. To randomize the sample size of each generated sample, we used a binomial model. The number of elements in each time series was decided to be random and binomially distributed, with the number of trials equal to the original sample size. The observed exceedance probability above the tail marker η_1, that is, the ratio of the number of the observed exceedances over η_1 to the sample size, defines the probability of success for each trial. In the following Table (3.7), these numbers, i.e., the probability of success, number of η_1 exceedances and sample size, are presented for each station:

We note that for the parametric bootstrap regarding the Torsvåg station, we considered only the case with outliers included. In addition, the data in the table reveal that for the Hekkingen station, the chosen tail marker η_1 does not allow a sufficient amount of data for the ACER analysis to be retained. Therefore, the uncertainty of the analysis is rather high, and the results might not be reliable, as it will be shown below.

Station	$p_0 = n(\eta \geq \eta_1)/N$
A - Torsvåg	$0.34 = 38\,421/112\,898$
B - Hekkingen	$0.03 = 3\,437/112\,170$
C - Nordøyan	$0.40 = 45\,575/112\,770$
D - Sula	$0.24 = 22\,839/96\,785$
E - Obrestad	$0.24 = 33\,594/139\,253$

Table 3.7: Probability of success, number of η_1 exceedances and sample size for five stations.

Case a

One-thousand time series were drawn from the CDF (3.67) using Eq. (3.70) for five weather stations. All required parameters q, b, a and c of the optimized curve $\varepsilon_k(\eta),\ k = 1$ for the simulation were taken from Table (3.2).

A sample of 10^3 of the of 100-year return levels $\hat{\eta}^{100yr}$ estimated by the ACER method defines a 95% confidence interval.

In the following Table (3.8), bootstrap-estimated 95% confidence intervals (95% BCI) are listed for five stations together with a graphic comparison between the initial 95% CI depicted by a red line and 95% BCI shown by a blue line.

The results presented in Table (3.8) show that the confidence intervals estimated by both methods have nearly identical lengths. The initial 95% confidence interval for three stations is approximately 30% shorter then the parametric 95% BCI, whereas in the other two stations it is longer, which indicates the same level of uncertainty for both approaches. At the same time, it is clear that the existent essential systematic error causes a considerable shift of the bootstrapped confidence interval to the right.

Case b

Similarly, 1000 simulations were performed using the CDF (3.69) via Eq. (3.71) for the five weather stations. A sample of 10^3 of the of 100-year return levels $\hat{\eta}^{100yr}$ estimated by the ACER method defines a 95% confidence interval.

Table (3.9) presents the bootstrap-estimated 95% confidence intervals (95% BCI) for the five stations. In addition, a graphic comparison between the initial 95% CI (depicted by a red line) and 95% BCI (blue line) is shown.

The results in Table (3.9) indicate that the 'Case b' scheme yields 95% confidence intervals that are identical to those obtained from the 'Case a' approach. As shown, the length of the initial ACER CI is approximately 20% shorter then the BCIs in three cases. In this way, it is reasonable to state that the stochastic uncertainty, which, properly speaking, is described by the confidence interval, is virtually equal for the 'Case b' parametric bootstrap and the ACER approach. However, a salient shift of the BCI to the right reflects the substantial systematic error of the considered approach.

Thus, we can ascertain that the considered parametric bootstrap approaches

provide the 95% CI estimate, the consistency of which is however questionable. In addition, the considered parametric bootstrap can be used only with the data assumed to be independent, and as for any bootstrap approach, its execution requires considerable time.

Thus, summarizing the above discussion, we can conclude that the adoption of

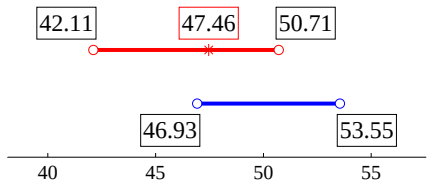
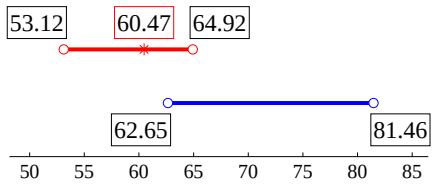
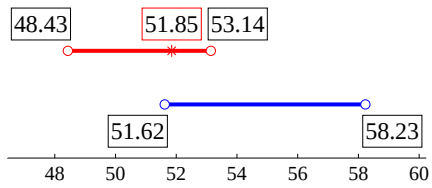
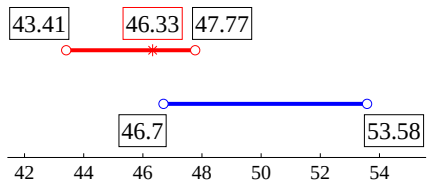
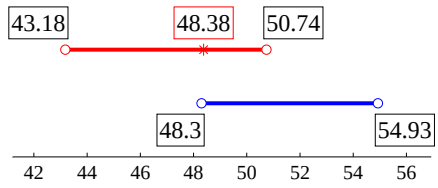
Station	Rel.diff	Intervals comparison
A - Torsvåg;	-30%	
B - Hekkingen	37%	
C - Nordøyan	29%	
D - Sula	37%	
E - Obrestad	-14%	

Table 3.8: Parametric bootstrap for estimation the 95%CI ('Case a' approach). Relative difference is displayed in the second column. Red intervals on top visualize the initial ACER CI with marked $\hat{\eta}^{100yr}$; blue intervals from bottom show parametrically bootstrapped CI.

the method of extrapolating the confidence bands using the ACER fit can remain the main approach for estimating the 95% CI of the T-year return level. The main advantages of this approach are the prompt and straightforward realization, its applicability to any degree of data conditioning k, and its estimated confidence interval appears to be credible.

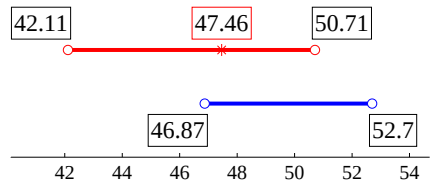
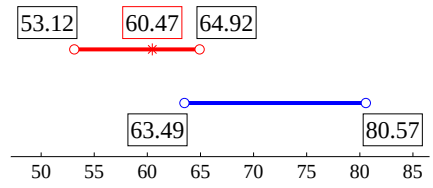
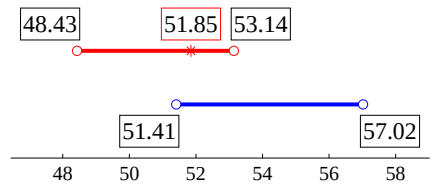
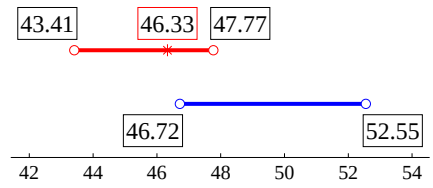
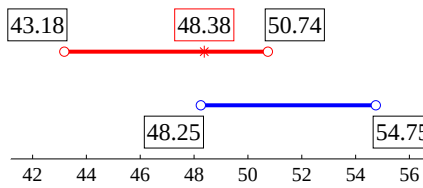
Station	Rel.diff	Intervals comparison
A - Torsvåg;	-48%	
B - Hekkingen	31%	
C - Nordøyan	16%	
D - Sula	25%	
E - Obrestad	-16%	

Table 3.9: Parametric bootstrap for estimation the 95%CI ('Case b' approach). Relative difference is displayed in the second column. Red intervals on top visualize the initial ACER CI with marked $\hat{\eta}^{100yr}$; blue intervals from bottom show parametrically bootstrapped CI.

3.5.4 Synthetic data from the heavy tail distribution for the general ACER

In this section, the performance of the general ACER approach discussed in Section 3.4.2 will be illustrated. We will analyse the extreme value statistics for the synthetic data with known exact extreme values.

As discussed in the introduction to Section 3.4, the ACER function represented by a four-parameter curve Eq. (3.28) was developed for the analysis of processes with Gumbel-type behaviour on asymptotic levels. Instead, the five-parameter general ACER case Eq. (3.51) applies to the processes that follow distributions with heavier tails. By this, we imply that it is known or can be assumed with a high certainty that the CDF does not have an exponential-type tail behaviour (Ochi, 1990). Thus, for our synthetic example for the general case, it was decided to consider the following underlying heavy tail distribution:

$$F_X(\eta) = 1 - \left(1 + \gamma \, \frac{\eta^2}{\eta + a}\right)^{-\frac{1}{\gamma}}, \; \eta \geq 0, \tag{3.72}$$

such that $\left(1 + \gamma \, \frac{\eta^2}{\eta+a}\right) > 0$.

It is easy to ascertain that the distribution Eq. (3.72) does not satisfy neither the sufficient conditions Eq (2.10) and Eq. (2.14) mentioned in Section 2.1.1 nor the criterion Eq. (2.17).

To generate a random variable X that follows the distribution $F_X(\eta)$, the inverse transformation of the Eq. (3.72) was applied. Quite simple derivations end up with the following expression for X:

$$X = \frac{1}{2}\left(K + \sqrt{K^2 + 4Ka}\right), \tag{3.73}$$

where $K = \frac{1}{\gamma}\left((1-U)^{-\gamma} - 1\right)$ and $U \sim \mathcal{U}(0,\, 1)$ is a uniformly distributed random variable.

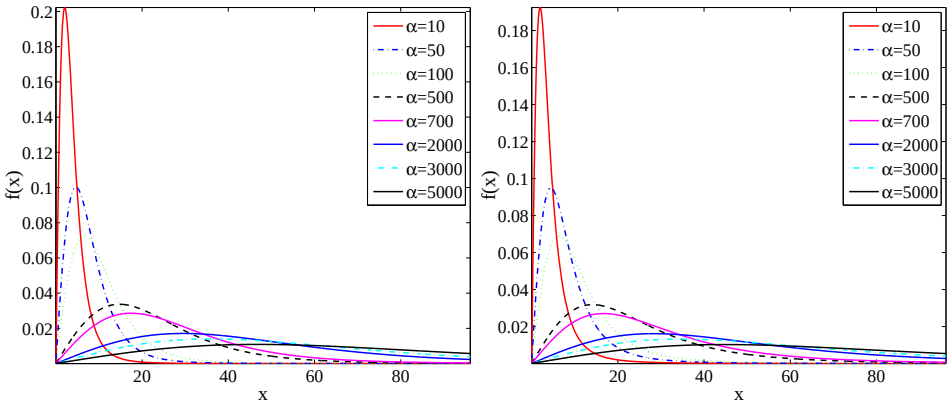
This setup allows us to generate a sample of N independent values $X_1, \ldots, X_N$ that follow the distribution (3.72). In compliance with the above, the distribution function $G(\eta)$ of the extreme value $M_X = \max\{X_1, \ldots, X_N\}$ would simply be

$$G(\eta) = \Pr\{X_1 \leq \eta, \ldots, X_N \leq \eta\} = \{F_X(\eta)\}^N. \tag{3.74}$$

Any T-year return period value η^{Tyr} can be calculated as the solution of the equation $F_X(\eta^{\mathrm{Tyr}}) = \left(1 - 1/T\right)^{1/N}$, provided N here is the amount of data per year.

As the final point, it was ascertained that to obtain a true fat tail distribution, values of the α parameter should be in range $\alpha > 10^3$, while the γ values should be $\gamma < 1$. Figure (3.34) illustrates the behaviour of the density function for a range of values of parameter α for fixed $\gamma = 0.3$ (3.34a) and fixed $\gamma = 0.5$ (3.34b).

These figures reveal that in both examples, the PDFs behave in much the same way; however, it was decided to consider two cases, namely, (a) with $\gamma = 0.3$ and $\alpha = 2 \cdot 10^3$ and (b) with $\gamma = 0.5$ and $\alpha = 3 \cdot 10^3$. Now we can proceed directly to the simulations. Let us assume we stock 30 years of synthetic observations, such that the annual record contains $N = 300$ data points such that the total amount of data



(a) (b)

Figure 3.34: Plot of the probability density function of of the distribution $F_x(\eta)$ (3.72) for different values of parameter a. a) parameter $\gamma = 0.3$; b) parameter $\gamma = 0.5$.

contains 9000 observations. This amount is not large for a detailed analysis, yet it brings the study closer to a real situation when a limited data sample is available. Thus, for the considered case (a), the 10-year and 100-year values are respectively $\eta^{10\text{yr}} = 273.54$ and $\eta^{100\text{yr}} = 410.83$, while in case (b), they are $\eta^{10\text{yr}} = 615.33$ and $\eta^{100\text{yr}} = 1201.4$.

To obtain an idea about the performance of the general ACER method as compared to the POT method and the annual maxima method based on the GEV distribution, 1000 independent 30-year Monte Carlo simulations were performed for both scenarios. For the ACER method, it was evident to analyse the ACER curve $\hat{\varepsilon}_k(\eta)$ for $k = 1$ with no conditioning on previous observations, because the generated from Eq. (3.73) data points are i.i.d. With regard to the value of the tail marker η_1, it was ascertained that the rational level to fix this parameter to is $1.7\,\sigma$ for both scenarios.

The POT prediction was based on using the maximum likelihood estimates (MLE) of two parameters of the generalized Pareto distribution for a chosen threshold level. Mean residual life plots for several realizations revealed the reasonable values of this parameter for both cases. Thus, the threshold u_0 was fixed at the level 80 for the case (a), whereas $u_0 = 130$ in case (b).

Predictions of the return levels by the annual maxima method was based on using the maximum-likelihood estimation of the three-parameter generalized extreme value (GEV) distribution. Sample of 30 yearly extremes was used for the analysis.

In the present research, it was decided to confine the analysis to the prediction of the return levels without estimating 95% confidence intervals. Substantially, this decision was based on the fact that the method of extrapolating the confidence bands using the ACER fit in general case has a tendency to overestimate

the interval. This makes the estimate inconsistent. Thus, the problem of estimating confidence interval for general ACER requires further analysis and will be considered in future research.

The results of the selected 16 out of 1000 independent 30-year Monte Carlo simulations for the case (a) are presented in Table (3.10) and for the case (b) – in Table (3.11). It is observed that the estimated average of the 1000 predicted

No.	A $\hat{\eta}^{10\text{yr}}$	GEV $\hat{\eta}^{10\text{yr}}_{\text{MM}}$	GP $\hat{\eta}^{10\text{yr}}$	A $\hat{\eta}^{100\text{yr}}$	GEV $\hat{\eta}^{100\text{yr}}_{\text{MM}}$	GP $\hat{\eta}^{100\text{yr}}$
1	255.00	262.35	256.55	356.83	387.46	348.04
2	277.04	287.41	271.20	414.88	441.28	375.22
4	290.89	280.07	280.12	449.23	402.07	398.71
6	241.94	247.41	246.59	329.62	387.94	330.11
8	270.22	312.28	279.48	408.86	1,005.45	412.71
10	274.90	285.93	263.11	410.76	476.50	354.61
20	234.60	247.26	251.25	297.99	497.86	332.70
40	289.51	276.98	273.23	466.16	410.66	391.82
60	279.20	274.45	265.66	433.34	377.20	373.51
80	274.21	262.68	265.81	405.97	352.91	366.85
100	317.99	315.65	294.61	531.29	495.21	425.23
200	272.32	249.49	251.88	416.23	280.01	338.20
400	258.68	255.50	252.84	372.61	350.30	341.71
600	303.02	316.49	302.45	492.93	605.70	465.36
800	286.66	293.54	285.37	434.46	429.82	409.83
1000	265.12	274.15	267.07	382.77	418.09	370.81
Av. 1000	**275.08**	**272.45**	**268.73**	**412.63**	**421.95**	**376.21**
Min 1000	223.76	217.71	222.62	276.66	237.07	275.66
Max 1000	346.39	369.41	322.64	626.83	1,185.20	512.28
2.5^{th} P	239.28	234.43	238.43	313.60	278.81	308.45
97.5^{th} P	321.58	322.62	306.17	553.71	731.30	470.31
True		**273.54**			**410.83**	

Table 3.10: Predicted return period levels by the General ACER method, A $\hat{\eta}^{\text{Tyr}}$, Annual maxima method, GEV $\hat{\eta}^{\text{Tyr}}_{\text{MM}}$, and POT method, GP $\hat{\eta}^{\text{Tyr}}$, respectively, for the (a) case: $\gamma = 0.3$ and $\alpha = 2 \cdot 10^3$.

10 - year return period levels $\hat{\eta}^{10\text{yr}}$ by the annual maxima method (GEV $\hat{\eta}^{10\text{yr}}_{\text{MM}}$ in the Tables) is more precise in both cases (a) and (b). However, the average of the 1000 predicted 100 - year return levels $\hat{\eta}^{100\text{yr}}$ by the general ACER method is considerably more accurate and closer to the theoretical value than by the other methods.

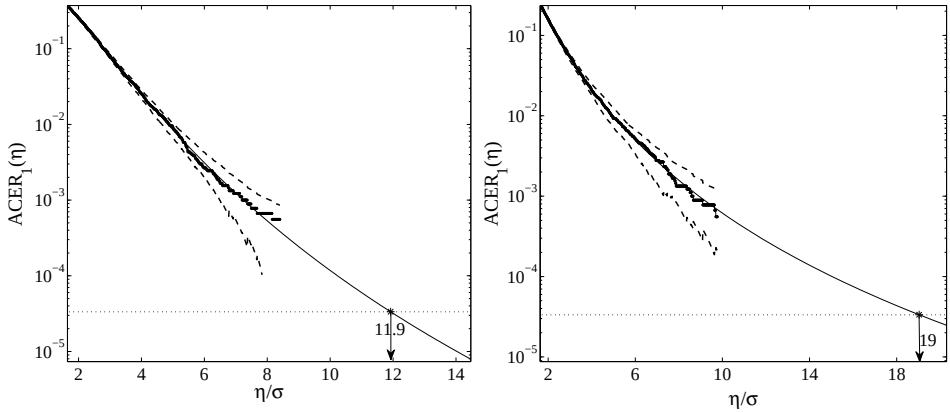
Note that based on the available sample of 1000 predicted return levels, it is possible to estimate its 2.5^{th} and 97.5^{th} percentiles. These values can be considered as left and right bounds of the 95% confidence interval. Thus, in case (a), 95% of 1000 predicted 100 - year return levels by the general ACER method lie inside the interval (313.60, 553.71), whereas the same for the block maxima method vary from (278.81, 731.30); for the POT method, the interval is (308.45, 470.31). In case (b), 95% of the predicted 100 - year return levels by the general ACER method lie within the range (738.96, 1917.79); 95% predictions by the annual maxima method stay inside the interval (639, 3173.65); for the POT method, the inter-

No.	A $\hat{\eta}^{10\text{yr}}$	GEV $\hat{\eta}^{10\text{yr}}_{\text{MM}}$	GP $\hat{\eta}^{10\text{yr}}$	A $\hat{\eta}^{100\text{yr}}$	GEV $\hat{\eta}^{100\text{yr}}_{\text{MM}}$	GP $\hat{\eta}^{100\text{yr}}$
1	560.06	511.05	548.34	937.38	752.65	880.96
2	516.25	555.51	566.93	827.67	1,454.38	965.58
4	1,018.05	800.13	921.83	3,029.60	1,373.29	2,384.55
6	500.29	526.29	492.81	806.09	949.42	745.11
8	577.33	594.20	612.07	983.42	1,240.46	1,062.23
10	635.83	621.25	614.35	1,165.83	1,334.78	1,056.05
20	612.06	649.80	608.62	1,118.95	1,200.91	1,038.75
40	575.03	555.66	568.18	983.13	791.50	934.37
60	577.29	472.66	526.79	980.00	571.50	803.73
80	755.51	717.74	722.47	1,713.92	1,274.72	1,478.62
100	561.21	539.40	569.71	975.98	1,176.86	974.35
200	540.15	521.08	507.69	921.72	952.55	776.66
400	615.20	671.99	614.36	1,184.65	1,614.68	1,106.89
600	731.42	612.51	649.23	1,648.59	855.70	1,217.49
800	707.65	598.49	625.88	1,428.97	1,094.52	1,061.48
1000	599.43	584.98	585.53	1,089.79	1,027.28	983.71
Av. 1000	**622.00**	**617.16**	**609.88**	**1,203.02**	**1,334.17**	**1,096.89**
Min 1000	447.56	426.00	453.81	618.15	484.78	624.25
Max 1000	1,018.05	1,200.37	921.83	3,029.60	15,138.90	2,384.55
2.5^{th} P	494.44	483.66	500.47	738.96	639	771.47
97.5^{th} P	796.30	839.49	751.30	1,917.79	3,173.65	1,596.32
True		**615.33**			**1201.40**	

Table 3.11: Predicted return period levels by the General ACER method, A $\hat{\eta}^{\text{Tyr}}$, Annual maxima method, GEV $\hat{\eta}^{\text{Tyr}}_{\text{MM}}$, and POT method, GP $\hat{\eta}^{\text{Tyr}}$, respectively, for the (b) case: $\gamma = 0.5$ and $\alpha = 3 \cdot 10^3$.

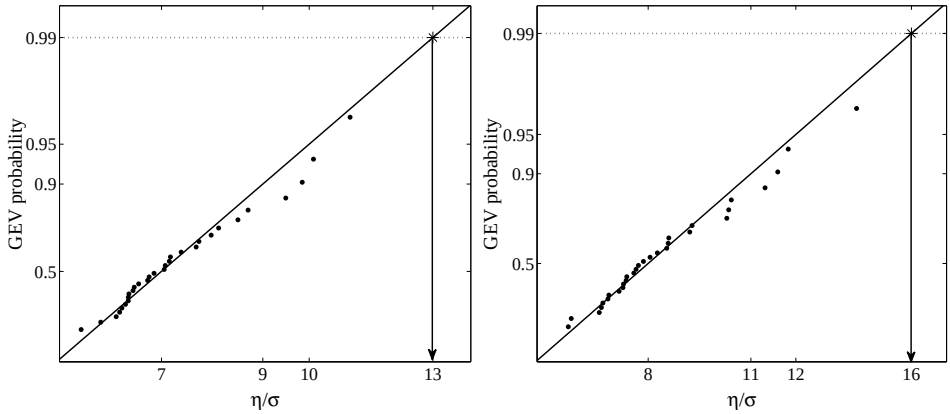
val is (771.47, 1596.32). It is easy to see that in case of the general ACER method there is a minor shift of the interval to the right in both scenarios, however for the annual maxima method the right-wing bias is significant. The POT method has the tendency to underestimate the interval in both cases. Therefore, the general ACER method performs consistently better and with higher accuracy than the other two methods. At the same time, it is important to emphasise that the general ACER case requires further research. Specifically, it concerns some aspects of the optimal curve fitting, as mentioned in Section 3.4.2, estimation of the 95% confidence intervals and analysis of the real data that follow the fat-tailed distributions (e.g., financial data) by the general ACER.

In Figure (3.35) the general ACER function and the fitted curve are plotted for the selected sets of data from the cases (a) and (b). The predicted 100-year return level $\hat{\eta}^{100\text{yr}} = 356.83$ in case (a) and $\hat{\eta}^{100\text{yr}} = 937.38$ in (b) scenario. Corresponding data sets were used in Figure (3.36) to illustrate the results of the annual maxima method by the GEV fit. The predicted values based on the maximum-likelihood estimation are $\hat{\eta}^{100\text{yr}} = 387.46$ and $\hat{\eta}^{100\text{yr}} = 752.65$ for the cases (a) and (b), respectively. Finally, Figure (3.37) presents POT predictions based on MLE for different thresholds in terms of the number n of observations above the threshold. The predicted value in case (a) is $\hat{\eta}^{100\text{yr}} = 348.04$, whereas in case (b) $\hat{\eta}^{100\text{yr}} = 880.96$.



(a) $\gamma = 0.3$, $\alpha = 2 \cdot 10^3$; $\sigma = 29.93$ (b) $\gamma = 0.5$, $\alpha = 3 \cdot 10^3$; $\sigma = 49.25$

Figure 3.35: Plot of the general ACER $\hat{\varepsilon}_1(\eta)$ on a $\log_{10}$ scale vs. η/σ (*); otimized curve fit (—); empirical 95% CI (- -). Tail marker $\eta_1 = 1.7\,\sigma$.

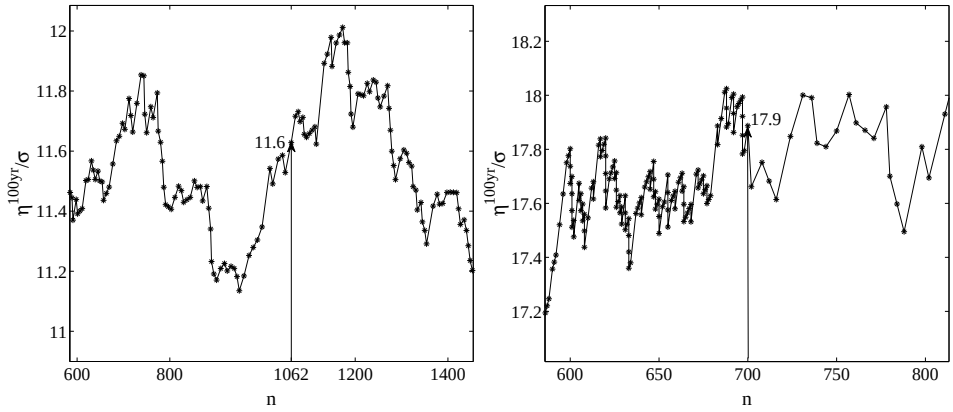


(a) $\gamma = 0.3$, $\alpha = 2 \cdot 10^3$; $\sigma = 29.93$ (b) $\gamma = 0.5$, $\alpha = 3 \cdot 10^3$; $\sigma = 49.25$

Figure 3.36: The point estimate of the 99% fractile of the Extreme value distribution by the Frechet method based on the 30 blocks of synthetic data.

3.5.5 Extreme tether tension

In this section we shall study data obtained from model tests of an offshore platform for oil production. The Heidrun tension leg platform (TLP) is a large concrete platform installed at a depth of 347 m in the Norwegian Sea. It is designed with four circular columns forming a square, with a square ring pontoon. Extensive model tests at a 1:55 scale were conducted in MARINTEK's 50 m × 80 m Ocean Basin in Trondheim in 1993. The mass of the TLP in ultimate limit state (ULS) conditions was 257, 888 tonnes, and the drift was 79.3 m. All data given here are in



(a) $\gamma = 0.3$, $\alpha = 2 \cdot 10^3$; $\sigma = 29.93$ (b) $\gamma = 0.5$, $\alpha = 3 \cdot 10^3$; $\sigma = 49.25$

Figure 3.37: The point estimate of 99% fractile of the Extreme value distribution as a function of the number n of data points above the threshold.

prototype scale. A sketch depicting the TLP from the side is shown in Figure (3.38).

The column diameter was 31 m, except in a small section in the wave zone where it was 31.6 m. The centre-to-centre distance between columns was 80.0 m. The pontoon has a rectangular cross-section, with a height 13.0 m and width 16.0 m.

The actual prototype tether group of four tethers at each column was modelled by a single equivalent tether, designed to correspond to the prototype with respect to stiffness, drag and weight properties.

The original test program included a number of different irregular wave test conditions and a large number of measuring channels, cf., (Naess et al., 2009). In this study, we concentrate on one severe ULS condition. It is specified in terms of the following sea state, which is a unidirectional (long crested) sea: significant wave height $H_s = 15.7$ m and spectral peak period $T_p = 17.8$ s.

The platform had a 45 degrees heading relative to the waves. The most heavily loaded tether is designated T10, and it is positioned towards the waves. Six different random realizations with a duration of 3 hours each were run. Thus, the resulting statistics correspond to a duration of 18 hours for the given sea state.

A particular observation from the model tests was the strongly non-Gaussian behaviour of the measured tensions, particularly in these high sea states. Thus, resonant high-frequency oscillations occurred, known as "ringing", which are excited by higher-order wave forces on columns in high and steep individual waves (Faltinsen et al., 1995; Stansberg, 1997). This comes in addition to the more commonly known "springing", excited by second-order sum-frequency forces. The extraordinary statistical behaviour was a main reason why these sea states were run with 6 realizations each. A time series sample from the measurements is shown in Figure (3.39), which clearly displays the ringing phenomenon caused by a steep wave.

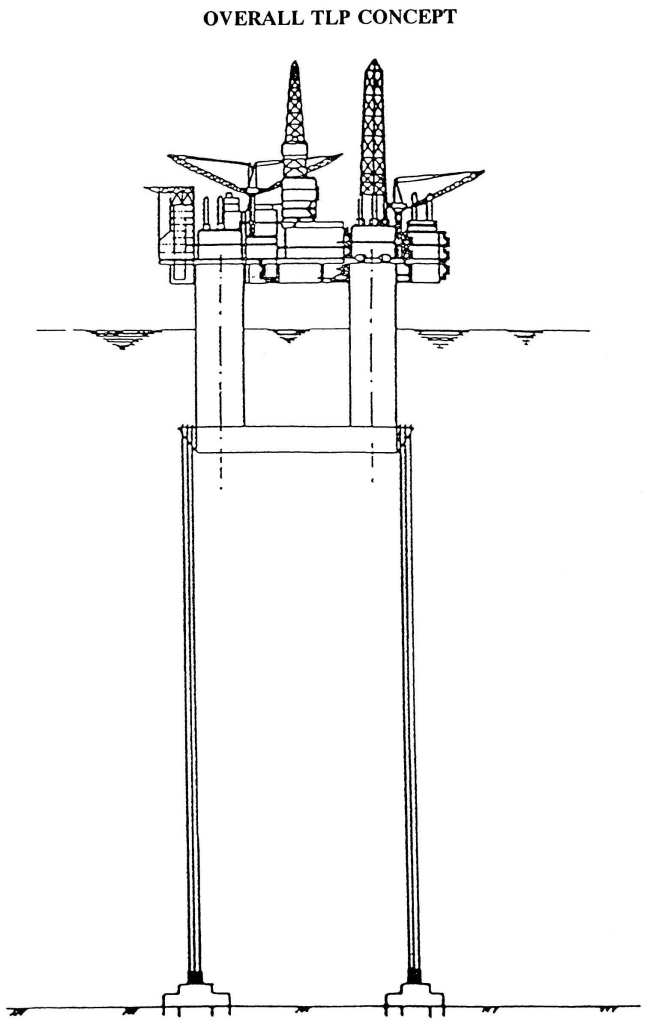


Figure 3.38: Heidrun TLP as seen from the side.

A basic statistical analysis of the 18 hours of time series for tension T10 shows that the mean tension μ = 97,537.5 kN, while the standard deviation σ = 5,698.73 kN. As shown from the plotted part of the time series in Figure (3.40), the considered process is rather densely sampled.

Evidently, heavy and extreme loads on the tether caused by a steep wave are represented by the peak values of the time series. Thus, for practical purposes, it was decided to conduct the ACER analysis for the process obtained by extracting peak events from the observed data. In Figure (3.41), the ACER functions $\varepsilon_k(\eta)$ are plotted for $k = 1, \ldots, 6$.

It is observed that there is a significant effect of dependence in the time series, which is reflected in the fact that the $\varepsilon_k(\eta)$, for $k = 2, \ldots, 6$, are noticeably smaller

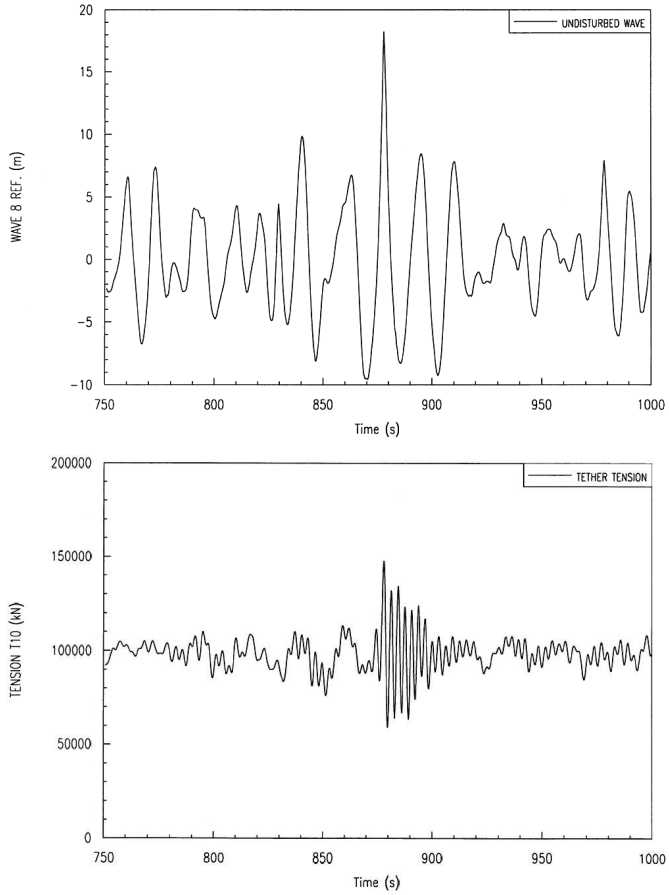


Figure 3.39: Short time series samples of wave elevation and tether tension T10, with a "ringing" event caused by a steep wave.

than $\varepsilon_1(\eta)$ over the whole range of response values. The figure reveals that there is no tendency for $\varepsilon_1(\eta)$ to merge with $\varepsilon_k(\eta)$ for $k = 2, \ldots, 6$ for the visible (available) range of tensions η. However, it is observed that a good approximation is already obtained for $k = 2$ and that convergence is certainly achieved for $k = 4$. To emphasize this point, the predicted value of the 90% fractile of the 3-hour extreme value distribution by the ACER method based on ε_1 is found to be $\eta_{0.90} = 184{,}562.64$ kN, with the 95% confidence interval (156,345.46, 210,692.98) kN. Here, parameters of the optimal curve are as follows: $q = 0.12$, $b = 109,806.47$, $a = 0.009$ and $c = 0.62$. The fractile $\eta_{0.90}$ estimated by means of ε_4 is 182,339.55 kN, with a 95% confidence interval of (158,443.62, 211,994.43) kN. The parameters of the optimal curve are as follows: $q = 0.19$, $b = 109,588.07$, $a = 0.047$ and $c = 0.48$. The tail marker is $\eta_1 =$ 112,000 kN in both cases. It is noticeable that the predicted 90% percentile value by the ACER method based on $\varepsilon_4(\eta)$ is slightly lower (1.2%) than the corresponding value based on $\varepsilon_1(\eta)$. Hence, the effect of statistical dependence in the available

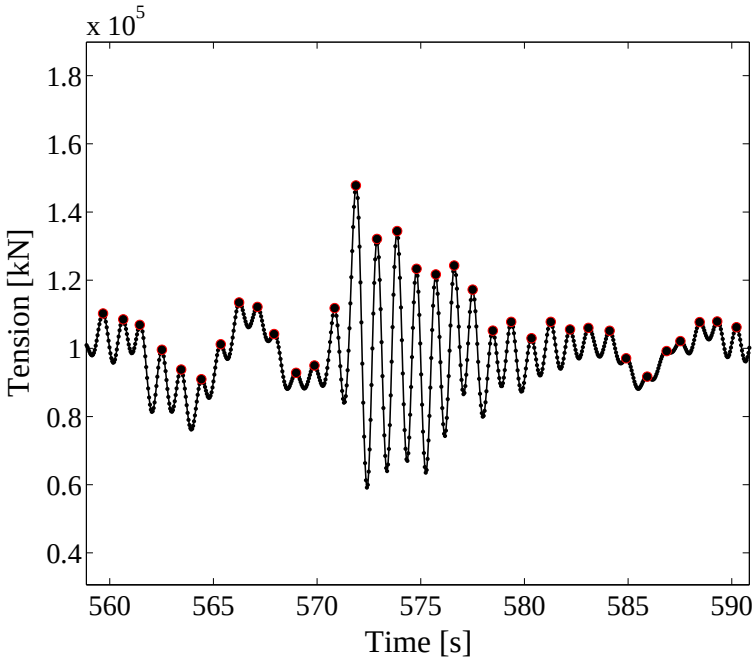


Figure 3.40: Part of the time series of tether tension T10, with marked peak events.

realization on the predicted 0.90 fractile of the 3-hour extreme value distribution is of some importance. It is also noted that the predicted statistical uncertainty is approximately the same in both cases.

To highlight the predictions based on the ACER, POT and Gumbel methods, we estimated the 99% percentile of the 3-hour extreme value distribution provided by the three methods. The obtained results are briefly summarized in Table (3.12):

Method	Spec	$\hat{\eta}_{0.99}$, kN	95% CI ($\hat{\eta}_{0.99}$), kN
ACER, various k	1	216,829.99	(170,697.58, 262,164.41)
	4	223,375.18	(177,404.33, 291,157.60)
Annual maxima	MM	218,414.13	(171,488.33, 279,250.21)
POT	–	205,754.55	(163,600.31, 273,673.86)

Table 3.12: Predicted return period levels by the ACER method for different degrees of conditioning, Annual maxima and POT methods, respectively.

Corresponding plots are presented in Figures (3.42) - (3.44). Note the large variability in the POT estimates depending on the choice of threshold.

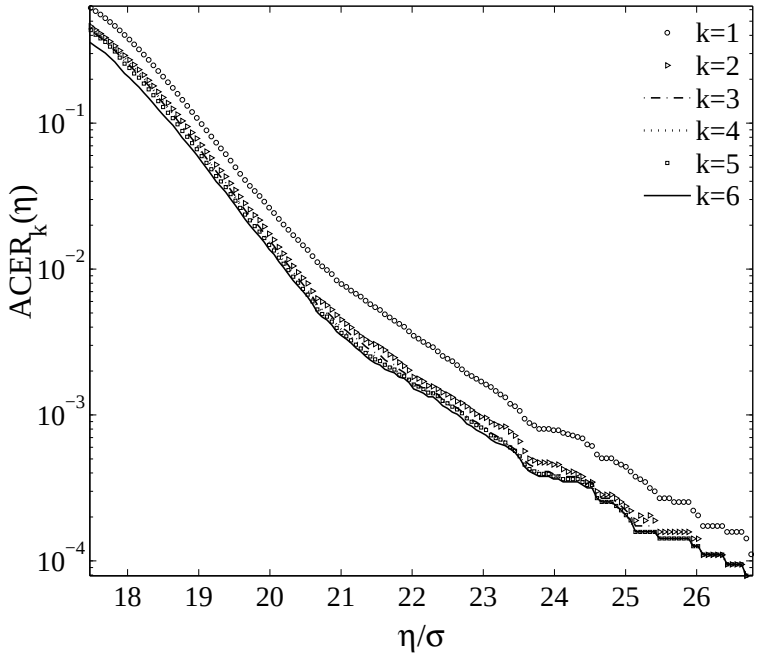


Figure 3.41: Plot of empirical ACER $\varepsilon_k(\eta)$ for different degrees of conditioning $k = 1, \ldots, 6$ based on the time series of the peak values, cf. Figure (3.40); $\sigma =$ 5,698.73 kN.

3.5.6 Analyses of the narrow-banded time series

In engineering mechanics, a classical extreme response prediction problem is the case of a lightly damped mechanical oscillator subjected to random forces. To illustrate this prediction problem, we shall investigate the response process of a linear mechanical oscillator driven by Gaussian white noise. Let $X(t)$ denote the displacement response; the dynamic model can then be expressed as $\ddot{X}(t) + 2\zeta\omega_e\dot{X}(t) + \omega_e^2 X(t) = W(t)$, where $\zeta =$ relative damping, $\omega_e =$ undamped eigenfrequency, and $W(t) =$ a stationary Gaussian white noise (of suitable intensity). By choosing a small value for ζ, the response time series will exhibit narrow band characteristics, that is, the spectral density of the response process $X(t)$ will assume significant values only over a narrow range of frequencies. This manifests itself by producing a strong beating of the response time series, which means that the size of the response peaks will change slowly in time; see Figure (3.45). A consequence of this is that neighbouring peaks are strongly correlated, and there is a conspicuous grouping of the peak values. Hence, the problem with accurate prediction because the usual assumption of independent peak values is then violated.

Many approximations have been proposed to address this correlation problem, but no completely satisfactory solution has been presented. In this section, we will show that the ACER method solves this problem efficiently and elegantly in a

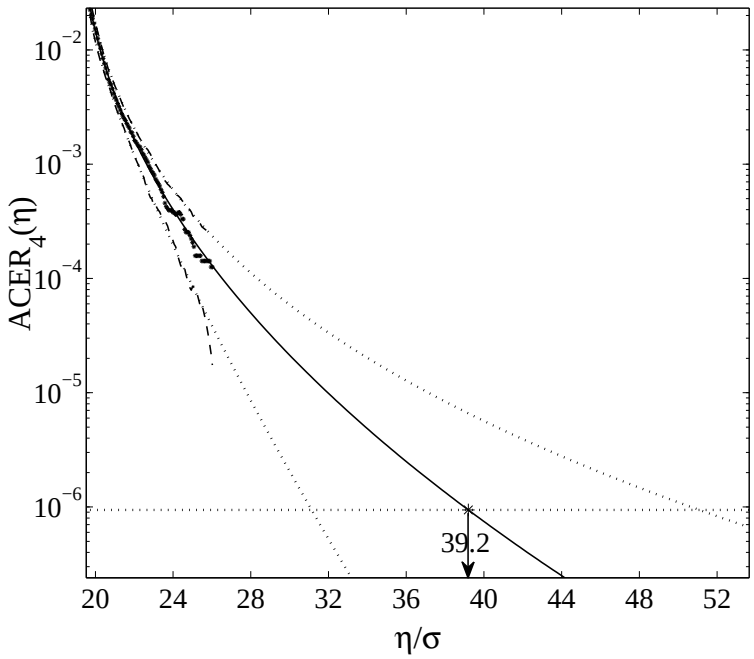


Figure 3.42: Tether tension: plot of empirical ACER $\hat{\varepsilon}_4$ on a $\log_{10}$ scale vs. η/σ (*); optimized curve fit (—); empirical 95% confidence band (- -); optimized 95% CI ($\cdots$). Tail marker $\eta_1 = 19.65\,\sigma$ kN. σ = 5,698.73 kN.

statistical sense. In Figure (3.46), some of the ACER functions for the example time series are shown. It can be verified from Figure (3.45) that there are approximately 32-35 sample points between two neighbouring peaks in the time series. To illustrate this point, we have chosen to analyse the time series consisting of all sample points. Typically, in practice, only the time series obtained by extracting the peak values would be used for the ACER analysis. In the present case, the first ACER function is then based on assuming that all of the sampled data points are independent, which is obviously completely incorrect. The second ACER function, which is based on counting each exceedance with an immediately preceding non-exceedance, is simply an upcrossing rate. Using this ACER function is largely equivalent to assuming independent peak values. It is now interesting to observe that the 25th ACER function can hardly be distinguished from the second ACER function. In fact, the ACER functions after the second do not change appreciably until one starts to approach the 32nd, which corresponds to hitting the previous peak value in the conditioning process. Therefore, the important information concerning the dependence structure in the present time series appears to reside in the peak values, which may not be very surprising. It is observed that the ACER functions exhibit a significant change in value as a result of accounting for the correlation effects in the time series. To verify the full dependence structure in the

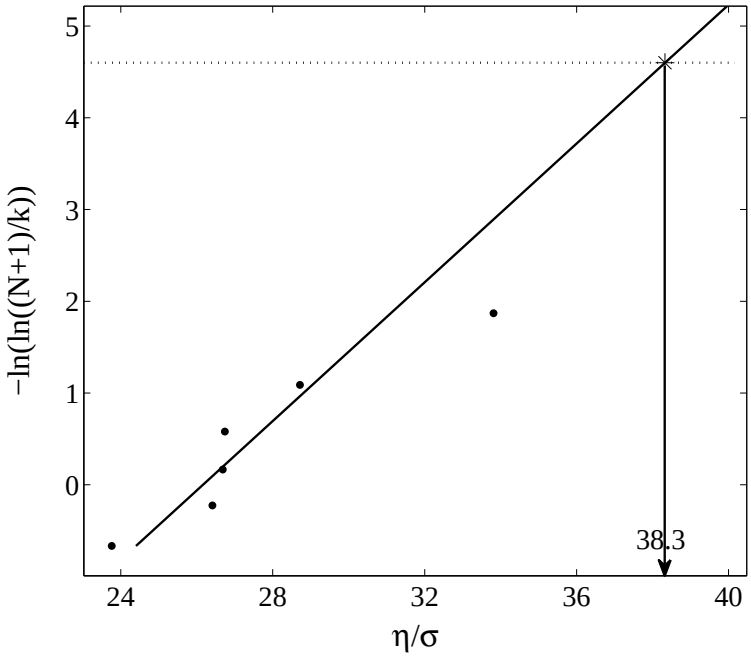


Figure 3.43: Tether tension: the point estimate $\hat{\eta}_{0.99}$ of the 0.99 fractile of the 3 hour extreme value distribution by the Gumbel method. Line is fitted by the method of moments (—). σ = 5,698.73 kN.

time series, it is necessary to continue the conditioning process down to at least the 64th ACER function. In the present case, there is virtually no difference between the 32nd and the 64th, which shows that the dependence structure in this particular time series is captured almost completely by conditioning on the previous peak value. It is interesting to contrast the method of dealing with the effect of sampling frequency discussed here with that of Robinson and Tawn (2000).

To illustrate the results obtained by extracting only the peak values from the time series, which would be the approach typically chosen in an engineering analysis, the ACER plots for this case are shown in Figure (3.47). By comparing the results from Figures (3.46) and (3.47), it can be verified that they are in very close agreement by recognizing that the second ACER function in Figure (3.46) corresponds to the first ACER function in Figure (3.47) and by noting that there is a factor of approximately 32 between the corresponding ACER functions in the two figures. This is because the time series of peak values contains approximately 32 times less data than the original time series.

In addition, it is of importance to pay particular attention to the positional relationship of the plots of the ACER functions with low degrees of conditioning for narrow-banded processes. From the discussions conducted in the previous sections, it can naturally be concluded that the ACER function with a lower degree of

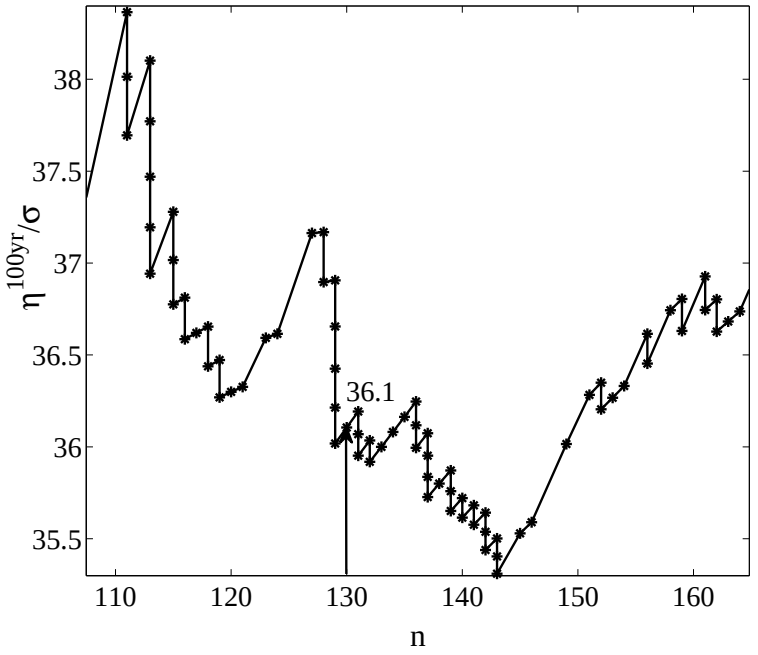


Figure 3.44: Tether tension: the point estimate $\tilde{\eta}_{0.99}$ of the 0.99 fractile of the 3 hour EVD by the POT approach as a function of the number n of data points above threshold. The 0.99 fractile estimate $36.1\,\sigma$ is at $n = 130$ for the threshold $\upsilon_0 = 21.5\,\sigma$. σ = 5,698.73 kN.

conditioning is, in general, not less than the one with a higher degree:

$$\hat{\varepsilon}_{k_i}(\eta) \geq \hat{\varepsilon}_{k_j}(\eta) \text{ for any } k_i \leq k_j. \tag{3.75}$$

Clearly, for any non-stationary time series, the law (3.75) is identically true because, for empirical ACER function estimation, Eq. (3.20) is applied. At the same time, it appears that this assertion does not necessarily apply for the stationary processes. Particularly, the rule does not hold for the densely sampled stationary narrow-band processes.

To illustrate the considered subject, let us scrutinize two stationary time series. The first process is the response of the linear oscillator to the Gaussian white noise we described in the beginning of the present section. As mentioned, two neighboring peaks of the time series are separated by approximately 32-35 sample points; see Figure (3.45). In the second case, we have an autoregressive (AR) random process, part of which is presented in Figure (3.48).

The figure reveals that the time series was sampled quite densely and that there are approximately 55-60 observations between two neighboring peak values. Therefore, the degree of conditioning $k \leq 20$ for the first case and $k \leq 30$ for the second can actually be considered as low. Now, the attentive scrutiny of the

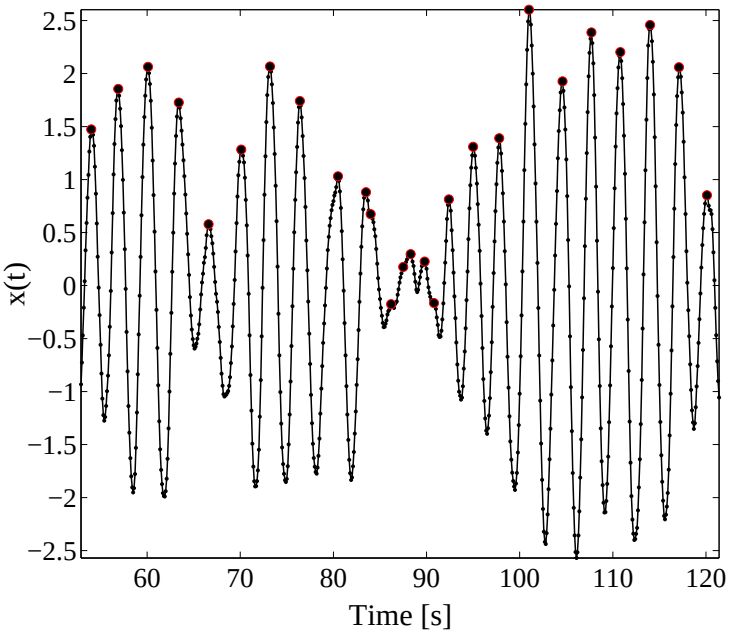


Figure 3.45: Part of the narrow-band response time series of the linear oscillator with fully sampled and peak values indicated.

logarithmic plot of the ACER functions with low k values discloses the opposite layout than the one defined by the Eq. (3.75); see Figures (3.5.6) and (3.49).

It is directly observed that conditioning on more previous observed data points yields an increase in the corresponding ACER function, e.g., in the case of the response process, the ACER function $\hat{\varepsilon}_{20}(\eta)$ with conditioning on $k-1=19$ previous observations takes on higher values than, for instance, the $\hat{\varepsilon}_k(\eta)$ for $k=2$ with conditioning on one preceding data point. At the same time, attention should be drawn to the fact that this effect vanishes after reaching a higher k degree, cf., Figure (3.46).

According to the arguments stated in Section 3.3, the ACER functions of both narrow-banded processes were estimated using Eq. (3.15), as for any stationary process, that is:

$$\hat{\varepsilon}_k(\eta) \;=\; \frac{\mathrm{a}_k(\eta)}{\mathrm{b}_k(\eta)}\,. \tag{3.76}$$

Here, to simplify the expression, we used the notations

$$\mathrm{a}_k \;\coloneqq\; \sum_{j=k}^{N} \mathrm{a}_{kj}\,; \quad \mathrm{b}_k \;\coloneqq\; \sum_{j=k}^{N} \mathrm{b}_{kj}\,,$$

where a_{kj} and b_{kj} are the realizations of the indicator functions defined by

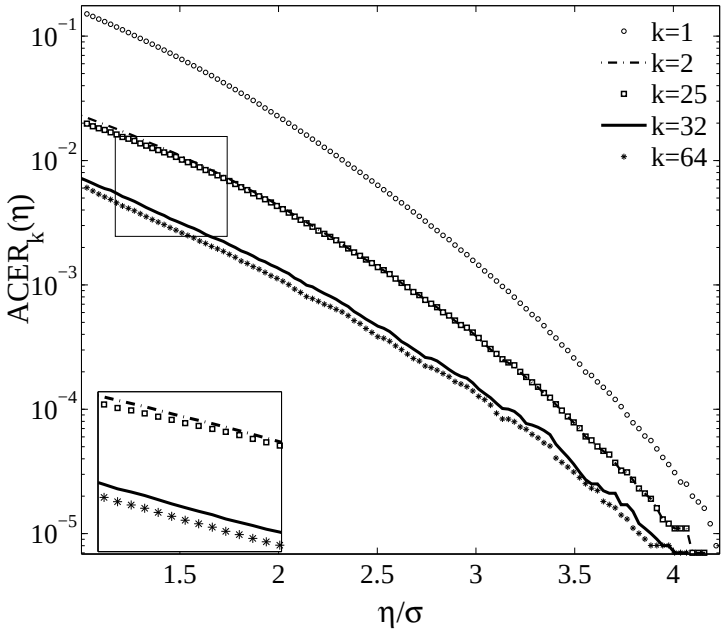


Figure 3.46: Comparison between ACER estimates for different degrees of conditioning for the narrow-band time series.

Eq. (3.12). Thus, the behaviour of the ACER function is defined by the functions $\mathrm{a}_k(\eta)$ and $\mathrm{b}_k(\eta)$.

Apparently, for every η value, the function $\mathrm{a}_k(\eta)$ represents the total number of observations that exceed the η level and follow immediately after a sequence of at least $k-1$ non-exceedances. What is crucial here is that for a densely sampled narrow-band time series, the number of η-level exceedances that come after $k-1$ consecutive non-exceedances can remain invariant. This occurs if the degree of conditioning remains low, that is, if $k-1$ is substantially less than the average number of data points between adjacent peaks. Regarding the $\mathrm{b}_k(\eta)$ function, its mathematical meaning is simply the total number of sequences of observations of length $k-1$ that do not exceed level η. Evidently, for any time series and for all η, the function $\mathrm{b}_k(\eta)$ with a lower k parameter takes on a higher value. This result is explained by the fact that the shorter sequence of non-exceedances, the greater the total number of such sequences is. To complete the discourse, let us consider plots of the corresponding $\mathrm{a}_k(\eta)$ and $\mathrm{b}_k(\eta)$ functions.

In Figures (3.5.6) and (3.51), the layout of the $\mathrm{a}_k(\eta)$ functions for both processes is demonstrated. It is observed that for all k, the functions exhibit the same behaviour and coincide almost entirely.

Corresponding parts of the $\mathrm{b}_k(\eta)$ functions are presented in Figures (3.5.6) and (3.53). These plots show that the estimated $\mathrm{b}_k(\eta)$ curve with the smallest parameter k is on top of all other curves.

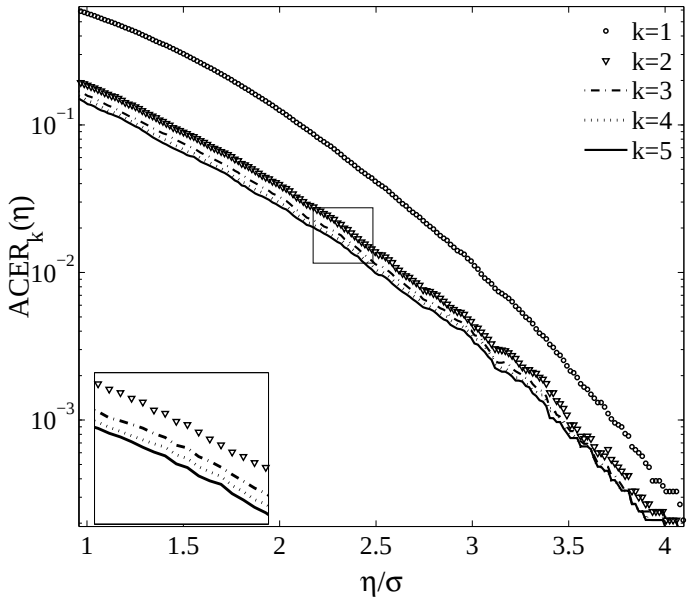


Figure 3.47: Comparison between ACER estimates for different degrees of conditioning based on the time series of the peak values, cf. Figure (3.45).

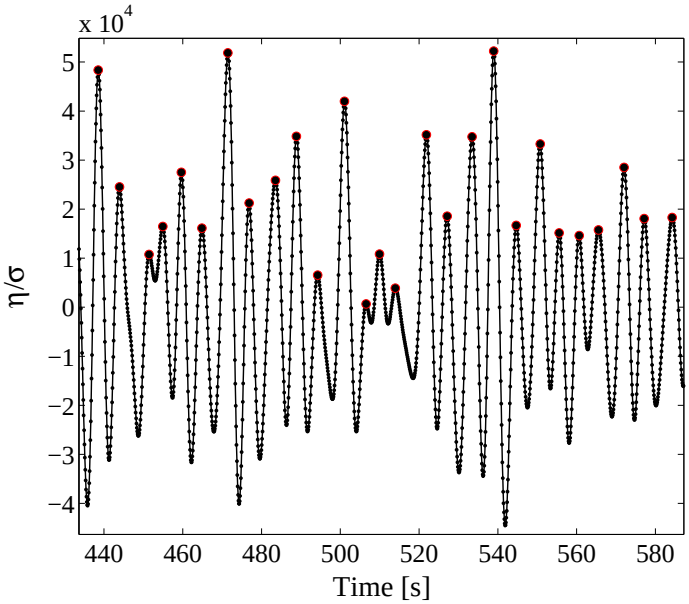


Figure 3.48: Part of the sampled AR random process with peak values indicated.

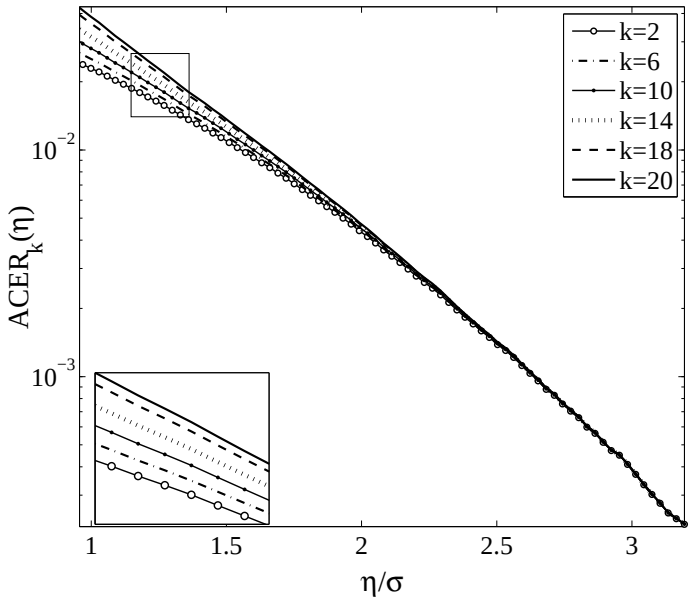


Figure 3.49: Comparison between ACER estimates for different low degrees of conditioning k for the narrow-band response process.

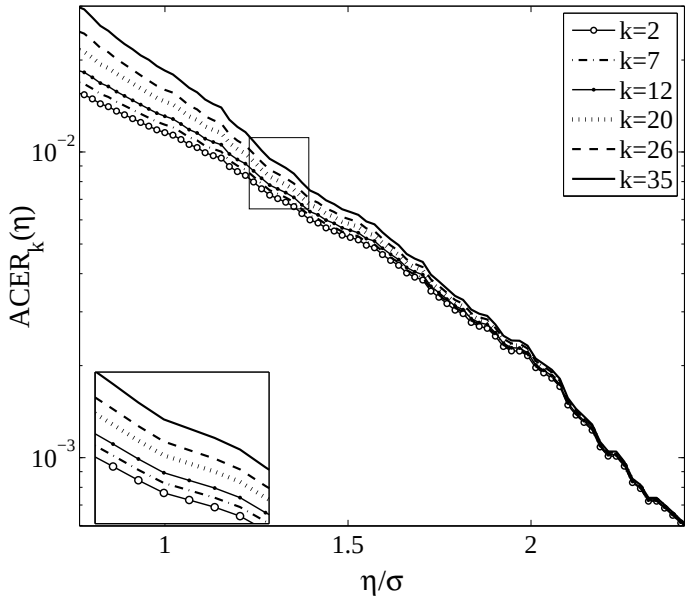


Figure 3.50: Comparison between ACER estimates for different low degrees of conditioning k for the autoregressive process.

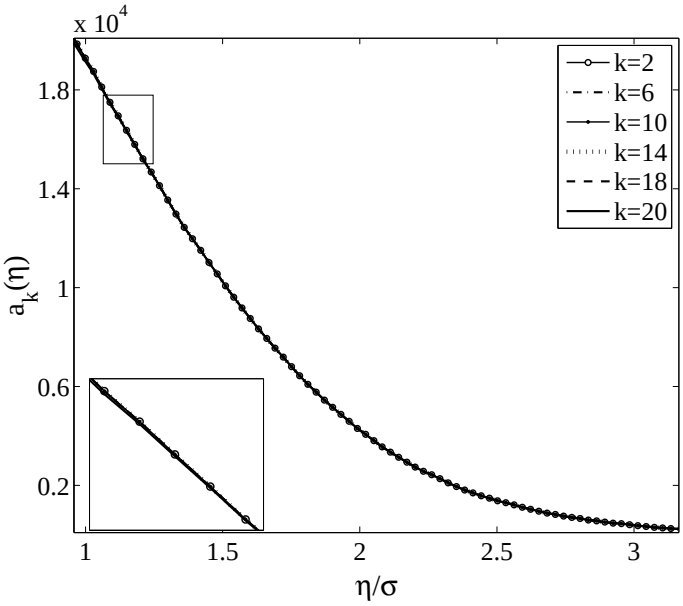


Figure 3.51: Comparison between significant parts of the estimated $\mathrm{a}_k(\eta)$ functions for low values of k parameter for the narrow-band response process.

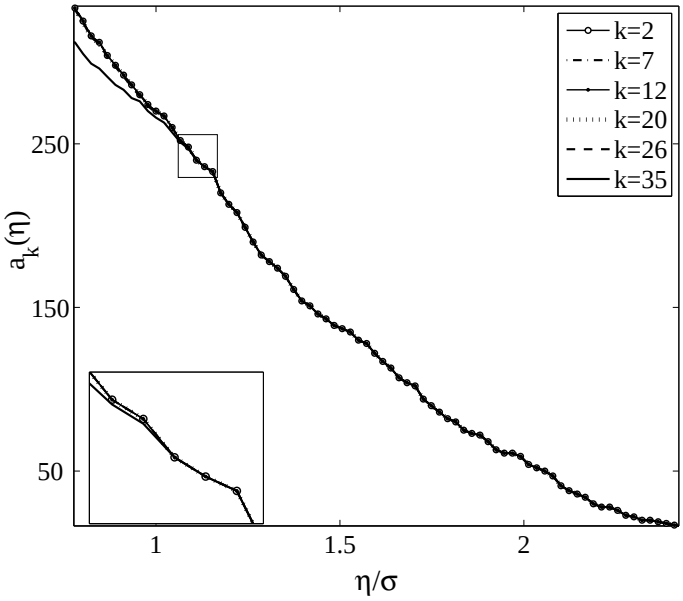


Figure 3.52: Comparison between significant parts of the estimated $\mathrm{a}_k(\eta)$ functions for low values of k parameter for the narrow-band autoregressive process.

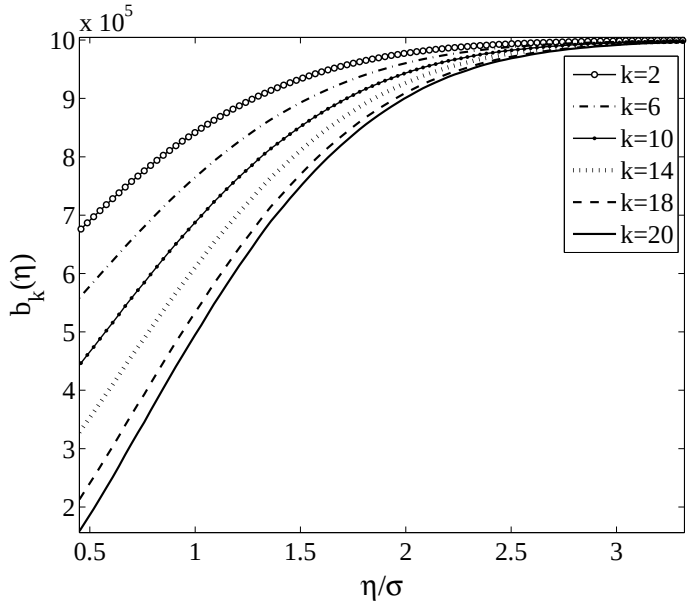


Figure 3.53: Comparison between significant parts of the estimated $b_k(\eta)$ functions for low values of k parameter for the narrow-band response process.

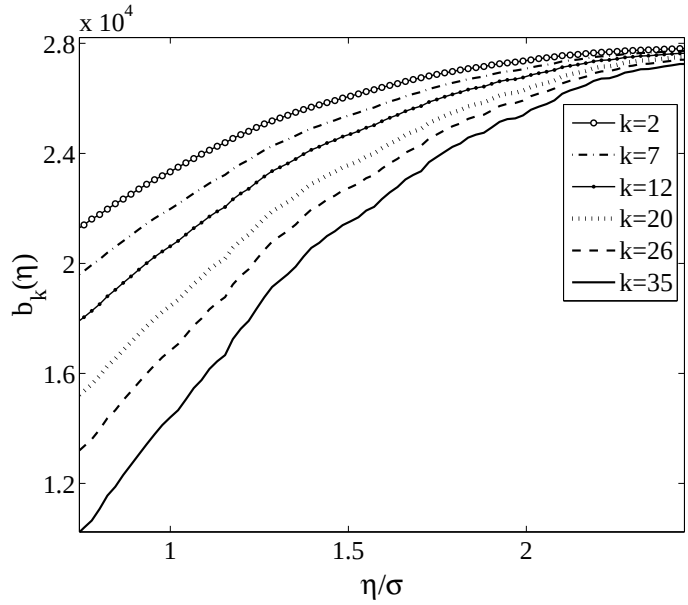


Figure 3.54: Comparison between significant parts of the estimated $b_k(\eta)$ functions for low values of k parameter for the narrow-band autoregressive process.

Thus, by comparing the results from Figures (3.5.6) - (3.5.6) and Figures (3.51) - (3.53), respectively, it becomes clear that in our case, the behaviour of the ACER functions solely depends on the layout of the corresponding $\mathrm{b}_k(\eta)$ curves. Furthermore, it is observed from Eq. (3.76) that the inverse ratio holds, which explains the discussed positional relationship between several $\hat{\varepsilon}_k(\eta)$ for the selected range of small values of the parameter k.

In conclusion, it should be noted that there is no accurate definition for the conception of a low k value that can be applied to any densely sampled narrow-band process. Individual detailed visual inspection of the data plot can hardly serve the task. Plotting the ACER curves for all $k = 1, 2, ...$ can be time consuming, and the figure itself can be overloaded and obscure. In addition, the inverse positioning of the $\hat{\varepsilon}_k(\eta)$ functions confuse the actual dependence structure of a process, which is crucial. Therefore, it appears that the only solution to this would be extracting the peak values and resumption of the ACER analysis based on the process of maxima.

Chapter 4

Bivariate Extreme Value Theory

The mathematical theory of multivariate extremes, particularly bivariate extremes, is a relatively novel field. The extension of extreme value statistics from the univariate to the bivariate case involves several challenges. First, there is no direct generalization of the univariate extreme value type theorem to the bivariate case.

Several areas, such as the Gumbel's logistic and mixed models, see Gumbel (1960b,a, 1961); Gumbel and Mustafi (1967), appear to be quite well developed. However, for the univariate case, these models have only asymptotic justifications. The later results on possible bivariate asymptotic extreme value distributions were derived on the basis of the aforementioned Gumbel's logistic and mixed models. The results are presented by Tiago de Oliveira (1982, 1984); Pickands (1981); Marshall and Olkin (1988), and others. However, these models became, in a sense, too general to be of much practical value, and their suitability for any practical application must be checked with care. It is also important to note that the recent work of Yue et al. (1999); Yue (2000, 2001b,a); Yue and Wang (2004) explored the usefulness of a bivariate extremal distribution, notably the Gumbel logistic model for describing bivariate hydrological extreme events.

Many efforts have been made to model and estimate a function that describes the dependence structure between extreme components. A considerable contribution to the development of the dependence measures is made by, e.g., de Haan and Resnick (1977); Coles and Tawn (1991, 1994); Coles et al. (1999) and Schlather and Tawn (2003); Eastoe et al. (2013). Nevertheless, there are no precise estimation tools that allow us to decide on the joint distribution of the bivariate extremes from a given set of bivariate data. Of course, the marginal data sets can be used to derive estimates of the marginal extreme value distributions, as in Zachary et al. (1998); de Haan and de Ronde (1998), but the joint distribution is still a long way off.

A popular method for addressing the problem of bivariate extremes is to adopt a copula model to represent the joint distribution structure (Nelsen, 2006). This copula is then combined with asymptotic extreme value distributions to represent the marginal distributions, typically of the GEV type (Coles, 2001). For this purpose, a range of different copula models have been proposed (see Waal and Gelder (2005); Tawn (1988)). The main problem with this approach is that it is rather

ad hoc. That is, there appears to be no theoretical justification for choosing one particular copula over the other. Even in case of the bivariate extreme value copula (Pickands, 1981; Balakrishnan and Lai, 2009), due to the properties of the dependence function, generally speaking, there are an infinite number of models.

The conditional extremes model developed by Heffernan and Tawn (2004) and subsequently used, for example, by Jonathan et al. (2010); Ewans and Jonathan (2014), provides a rather novel approach to model the marginal and dependence structure of multivariate extremes. First, the marginal distribution of each variable is transformed from the original scale to a standard Gumbel scale. Consequently, the threshold exceedences of a marginal variable are modeled independently using the generalized Pareto distribution, whereas the dependence structure is modeled for pairs of transformed variables. What counts here is that the model assumptions eventually lead to the collection of separate semiparametric conditional models being fitted through ad hoc methods. The complicated nature of the methods may restrict the applicability of the model.

Efforts have also been invested in an attempt to extend the peaks-over-threshold (POT) method to the multivariate case (Kaufmann and Reiss, 1995; Falk and Reiss, 2003). The bivariate POT framework also faces the same tasks of defining the dependence function. This has not yet resulted in a method with the same prediction capabilities as the univariate POT.

4.1 Bivariate Extreme-Value Distributions

Let us consider a bivariate stochastic process $Z(t) = \big(X(t), Y(t)\big)$, with dependent component processes $X(t)$ and $Y(t)$, which has been observed over a time interval, e.g., $(0; T)$. Suppose that the sampled values $(X_1, Y_1), \ldots, (X_N, Y_N)$ are allocated to the times $0 \leq t_1 < \ldots < t_N \leq T$. This sequence can be considered to be a series of vectors that are replicas of a random vector with a $H_z(\xi, \eta)$ distribution. In a maritime setting, the vector $\big(X(t), Y(t)\big)$ might represent pairs of wind speeds measured at two neighboring locations or simultaneously recorded wave elevation and wind speed values. By analogy with the univariate case, the classical theory of bivariate extremes is based on limiting the behaviour of the block maxima. It starts with defining the vector of component-wise maxima $\hat{Z}_N$, as follows,

$$\hat{Z}_N = \big(\hat{X}_N, \hat{Y}_N\big), \text{ with } \hat{X}_N = \max_{1 \leq i \leq N} \{X_i\} \text{ and } \hat{Y}_N = \max_{1 \leq j \leq N} \{Y_j\} \,. \tag{4.1}$$

The classical asymptotic theory of bivariate extreme values studies the $\hat{Z}_N$, as $N \to \infty$. Because the sequence of vectors $\{Z_i\}$ is assumed to be an independent version of a random vector with common bivariate distribution $H_z(\zeta)$, where we used the notation $\zeta = (\xi, \eta)$, then

$$\text{Prob}\left(\hat{Z}_N \leq \zeta\right) \;=\; \text{Prob}\left(\max_{1 \leq i \leq N} Z_i \leq \zeta\right) \;=\; \big[H_z(\zeta)\big]^N \,, \tag{4.2}$$

because $\max\limits_{1 \leq i \leq N} Z_i \leq \zeta$ holds if, and only if, $Z_1 \leq \zeta, \ldots, Z_N \leq \zeta$. Thus, $\hat{Z}_N$ in (4.2) follow $\big[H_z(\zeta)\big]^N$ if Z_i are governed by $H_z(\zeta)$.

By analogy with the univariate case, the exact distribution function $\left[H_Z(\zeta)\right]^N$ can be replaced by a limiting distribution. If

$$\lim_{N\to\infty} H_Z^N\left(a_N + b_N\,\zeta\right) \;=\; H(\zeta)\,, \tag{4.3}$$

for two-component vectors b_N and $a_N > 0$, then $H(\zeta)$ is called a bivariate extreme value distribution and $H_Z(\xi,\eta)$ belongs to the max-domain of attraction of $H(\xi,\eta)$.

Let $F_X(\xi)$ and $G_Y(\eta)$ denote margins of $H_Z(\zeta)$. From (4.3), it follows that

$$\begin{aligned} \lim_{N\to\infty} F_X^N\left(a_{N,1} + b_{N,1}\,\xi\right) \;&=\; F(\xi)\,; \\ \lim_{N\to\infty} G_Y^N\left(a_{N,2} + b_{N,2}\,\eta\right) \;&=\; G(\eta)\,, \end{aligned} \tag{4.4}$$

where $F(\xi)$ and $G(\eta)$ are marginal distributions of $H(\zeta)$. Hence, each marginal of a bivariate extreme value distribution is necessarily a univariate extreme value distribution.

Now, the task is partly resolved by recognizing that $\{X_i\}$ and $\{Y_i\}$ can be considered separately and are sequences of independent random variables. Thus, classical univariate extreme value theory applies to both components. This implies possible simplification of the presentation by assuming the X_i and Y_i variables have a known marginal distribution. Other marginal distributions, whose extremal properties are determined by the univariate characterizations, can always be transformed into this known form. The three types extreme-value distributions also can be transformed to each other. Gumbel and Goldstein (1964), Gumbel (1965), Gumbel and Mustafi (1967), Tiago de Oliveira (1962, 1984) assumed Gumbel marginals, whereas, e.g., de Haan and Resnick (1977) and Kotz and Nadarajah (2000) chose Fréchet marginals $F(a) = \exp\left(-1/a\right)$. All three types can be easily transformed to exponential variates, and, for example, Pickands (1981), Tawn (1988), Reiss and Thomas (2007) choose exponential marginals.

Without loss of generality and for simplified representations, it is assumed that the distribution functions $F_X(\xi)$ and $G_Y(\eta)$ of the components $\{X_i\}$ and $\{Y_i\}$, respectively, both standard Fréchet distributions, cf., e.g., Coles (2001):

$$F_X(\vartheta) \;=\; G_Y(\vartheta) \;=\; \exp\left(-\frac{1}{\vartheta}\right),\;\; \vartheta > 0\,. \tag{4.5}$$

The standard Fréchet distribution is the GEV distribution with parameters $\mu = 0$, $\sigma = 1$ and $\gamma = 1$. Therefore, due to the max-stability of any GEV distribution (see Definition 2.2), which in our case is fulfilled with constants $a_N = 0$ and $b_N = N$, for independent observations, holds:

$$\text{Prob}\left(\hat{X}_N/N \le \vartheta\right) \;=\; \text{Prob}\left(\hat{Y}_N/N \le \vartheta\right) \;=\; \exp\left(-\frac{1}{\vartheta}\right),\;\; \vartheta > 0\,. \tag{4.6}$$

Therefore, to be able to carry out the above results for each marginal, it is essential to introduce the re-scaled vector $\hat{Z}_N^* = \left(\hat{X}_N/N,\, \hat{Y}_N/N\right)$.

By analogy with the Fisher-Tippett-Gnedenko Theorem 2.1 for the univariate case, the following theorem describes the property of the limiting joint distribution for the bivariate case.

Theorem 4.1. *Let* $\hat{Z}_N^* = \left(\hat{X}_N/N,\, \hat{Y}_N/N\right)$ *be the re-scaled vector of componentwise maxima, where* $\hat{X}_N$ *and* $\hat{Y}_N$ *are defined by* (4.1) *and the* $(X_i,\, Y_i)$ *are independent vectors with standard Fréchet marginal distributions. If a non-degenerate distribution function* $H(\xi,\, \eta)$*, such that*

$$Prob(\hat{X}_N/N \le \xi,\, \hat{Y}_N/N \le \eta) \xrightarrow{d} H(\xi, \eta)\,, \tag{4.7}$$

exists, then $H(\xi,\, \eta)$ *has the form*

$$H(\xi, \eta) = \exp\big\{ - V(\xi, \eta)\big\};\; \xi > 0,\; \eta > 0\,, \tag{4.8}$$

where

$$V(\xi, \eta) = 2 \int_0^1 \max\left(\frac{w}{\xi}, \frac{1-w}{\eta}\right)\, dA(w)\,, \tag{4.9}$$

and A *is a distribution function on* $[0,\, 1]$ *satisfying the mean value constraint*

$$\int_0^1 w\, dA(w) = \frac{1}{2}\,. \tag{4.10}$$

If the limit in Eq. (4.7) exists, then the distribution function $H(\xi, \eta)$ defines the class of bivariate extreme value distributions. From Theorem 4.1 follows the existence of bijective correspondence between the $H(\xi, \eta)$ family and set of $A(w)$ distribution functions on $[0, 1]$ subject to the condition (4.10). If $A(w)$ is differentiable and has a density, then the entire model is considered to be a differentiable bivariate extreme value distribution. Alternatively, bivariate extreme value distributions are also generated by measures $A(w)$ that are not differentiable. It appears to be important to consider two examples that characterize two limit states of the dependence between processes $X(t)$ and $Y(t)$. When $A(w)$ is a measure that places mass 0.5 on $w = 0$ and $w = 1$, that is $A(0) = A(1) = 0.5$, condition (4.10) is trivially satisfied, and it follows that $V(\xi, \eta) = 1/\xi + 1/\eta$ by Eq (4.9); hence, for $\xi > 0$, $\eta > 0$, the bivariate extreme value distribution is

$$H(\xi, \eta) = \exp\left\{-\frac{1}{\xi} - \frac{1}{\eta}\right\} = \exp\left(-\frac{1}{\xi}\right) \exp\left(-\frac{1}{\eta}\right)\,. \tag{4.11}$$

This function corresponds to the case of independent X and Y. Similarly, if a measure $A(w)$ places unit mass on $w = 0.5$, that is, $A(0.5) = 1$, Eq. (4.10) is again satisfied trivially. Then, from Eq. (4.9), it follows that $V(\xi, \eta) = \max\big(1/\xi,\, 1/\eta\big)$, and for $\xi > 0$, $\eta > 0$, the corresponding bivariate extreme value distribution is

$$H(\xi, \eta) = \exp\left\{-\max\left(\frac{1}{\xi},\, \frac{1}{\eta}\right)\right\} = \min\left\{\exp\left(-\frac{1}{\xi}\right),\, \exp\left(-\frac{1}{\eta}\right)\right\}\,, \tag{4.12}$$

which is the distribution function of variables that are marginally standard Fréchet but which are perfectly dependent: $X = Y$ with probability 1.

An important class of differentiable bivariate extreme value distributions is the logistic family, also referred to as the Type B bivariate extreme value distribution (Balakrishnan and Lai, 2009). Developed by Gumbel (1960b) and Gumbel and Mustafi (1967), it has subsequently been widely used for modelling of environmental extremes (see, for example, works by Yue (2001b,a)). The logistic family with standard Fréchet marginals takes the following form:

$$H(\xi, \eta) = \exp\left\{ -\left(\left(\frac{1}{\xi}\right)^m + \left(\frac{1}{\eta}\right)^m \right)^{1/m} \right\}, \tag{4.13}$$

for $m \geq 1$ and $\xi > 0$, $\eta > 0$. The relation between (4.13) and general forms (4.9) and (4.8) is not trivial, yet, the distribution function (4.13) is derived for the certain density function $a(w; m)$ of $A(w)$, the expression of which can be found in the literature (see Coles (2001)). The main reason for the popularity of the logistic family is its flexibility because subject to the value of the m parameter distribution, (4.13) covers all levels of dependence from independence for $m = 1$ to perfect dependence.

Thus, overall, it can be concluded that measure $A(w)$ comprises the dependence structure between components $X(t)$ and $Y(t)$.

Because any non-degenerate univariate asymptotic extreme value distribution can be written as a GEV distribution (see Theorem 2.2):

$$F(\vartheta) = \exp\left\{ -\left[1 + \gamma \left(\frac{\vartheta - \mu}{\sigma} \right) \right]^{-1/\gamma} \right\},$$

it follows that the complete class of asymptotic bivariate extreme value distributions can be obtained simply by generalizing the marginal distributions. Specifically, using the assignments

$$\tilde{\xi} := \left[1 + \gamma_x \left(\frac{\xi - \mu_x}{\sigma_x} \right) \right]^{1/\gamma_x} \quad \text{and} \quad \tilde{\eta} := \left[1 + \gamma_y \left(\frac{\eta - \mu_y}{\sigma_y} \right) \right]^{1/\gamma_y},$$

the complete family of bivariate extreme value distributions, with arbitrary GEV margins, has a distribution function of the form

$$H(\xi, \eta) = \exp\left\{ -V(\tilde{\xi}, \tilde{\eta}) \right\},$$

given that $[1 + \gamma_x (\xi - \mu_x)/\sigma_x] > 0$, $[1 + \gamma_y (\eta - \mu_y)/\sigma_y] > 0$, and where the functions V and A satisfy (4.9) and (4.10). Note that parameters of the GEV marginal distributions are $\{\gamma_x, \mu_x, \sigma_x\}$ and $\{\gamma_y, \mu_y, \sigma_y\}$, respectively.

It is noticed that function $V(\xi, \eta)$ defined in (4.9) is homogeneous of negative degree -1, because for any constant $c > 0$, holds: $V(c^{-1}\,\xi, c^{-1}\,\eta) = c\,V(\xi, \eta)$. This property and Eq. (4.8) lead to

$$H^N(\xi, \eta) = H\left(N^{-1}\,\xi,\, N^{-1}\,\eta\right), \tag{4.14}$$

or in more compact form, $H^N(\zeta) = H(N^{-1}\,\zeta)$. In addition, it is possible to show that $H^N\big(a_N + b_N\,\zeta\big) = H(\zeta)$ for some vectors b_N and $a_N > 0$. Therefore, distribution H defined in (4.8) has the property of max-stability. Like in the univariate

case (see Section 2.1.1), it can be shown that distributions of the type (4.8) are the only max-stable distributions, accurate to within the chosen marginal distributions (Resnick, 2008). Finally, it can be proved that H is an extreme value distribution if, and only if, it is max-stable distribution.

Despite the fact that Theorem 4.1 provides a complete and straightforward description of bivariate extreme distributions, the class of possible asymptotic distributions is wide. The reason is that the only constraint on $H(\xi, \eta)$ in Eq. (4.8) is given by (4.9). Particularly, any distribution function A on [0, 1] in (4.9) that satisfies the mean value constraint $\int_0^1 w\, dA(w) = 1/2$, ensures a valid limit in (4.8). This leads to difficulties, as the asymptotic family has no finite parameterization. One reasonable way to deal with the problem is to possibly use nonparametric methods of estimation. This is also complicated by the fact that nonparametric estimators can hardly be constrained accurately to satisfy functional conditions of the type (4.10).

A possible alternative is to use parametric sub-families of distributions for A, leading to sub-families of distributions for H. In this way, only a small subset of the complete class of limit distributions for H is obtained. Yet it is possible to ensure that a wide sub-class of the entire limit family is approximated. That is, parametric families for A, and hence H, can be obtained, such that every member of the full limit class for H can be closely approximated by a member of the sub-family generated by the family of A. In principle, it is simple to require a parametric family for A on [0, 1] whose mean is equal to 0.5 for every value of the parameter. The corresponding family for H are generated then by substituting A into (4.9) and (4.8). In practice, however, it is not so easy to generate a parametric family with parameter-free mean and for which the integral in (4.9) can be controlled.

An alternative way of modeling bivariate extreme value distribution consists of studying the dependence structure and the marginals separately. Since copulas describe and model the dependence structure between random variables, independently of the marginal laws involved, it appears to be expedient to introduce the mathematical theory of copulas.

4.2 Bivariate Extreme-Value Analysis via Copulas

The copula theory is a relatively new and growing field. The construction and properties of copulas have been studied rather extensively during the last 20 years or so. Hutchinson and Lai (1990) were among the early authors who popularized the study of copulas. Nelsen (1999) presented a comprehensive treatment of bivariate copulas, while Joe (1997) devoted a chapter of his book to multivariate copulas. Further authoritative updates on copulas are given in Nelsen (2006).

In this section, the general definition of a copula will be surveyed first. All the theoretical justifications can be found in, e.g. Joe (1997); Nelsen (2006). A more detailed survey can also be found in Sempi (2003). Thereafter, the extreme value copula will be presented.

4.2.1 Basic properties of a bivariate copula

Let $\mathbb{I} = [0, 1]$ denote a unit interval.

Definition 4.1. ***(Bivariate copula).*** A bivariate copula $C(u, v)$ is a function $C : \mathbb{I} \times \mathbb{I} \to \mathbb{I}$ such that:

- For every $u, v \in \mathbb{I}$

$$C(u, 0) = 0,\ C(u, 1) = u,\ C(0, v) = 0,\ C(1, v) = v, \tag{4.15}$$

- A copula is continuous in u and v; copula satisfies the stronger Lipschitz condition:

$$\big|C(u_2, v_2) - C(u_1, v_1)\big| = |u_2 - u_1| + |v_2 - v_1|\,; \tag{4.16}$$

- For $0 \le u_1 < u_2 \le 1$ and $0 \le v_1 < v_2 \le 1$,

$$\begin{aligned} &\mathrm{Prob}\big(u_1 \le U \le u_2,\, v_1 \le V \le v_2\big) \\ &= C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) > 0\,. \end{aligned} \tag{4.17}$$

With the given definition, it is not difficult to verify that any finite convex linear combination of bivariate copulas C_i is also a bivariate copula. In other words, for $k \in \mathbb{N}$ and with $\lambda_i \ge 0$, such that $\sum_{i=1}^{k} \lambda_i = 1$, let C be given by $C(u, v) = \sum_{i=1}^{k} \lambda_i\, C_i(u, v)$. Then C is a copula.

Sklar's theorem below expounds the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals (Sklar, 1959). As before, $F_X(\xi)$ and $G_Y(\eta)$ denote the marginal distribution functions of the random variables X and Y, respectively.

Theorem 4.2. (Sklar, bivariate case). *Let $H_{XY}(\xi, \eta)$ be a joint distribution function with marginals $F_X(\xi)$ and $G_Y(\eta)$. Then, there exists a copula C such that, for all $\xi,\ \eta \in [-\infty, \infty]$,*

$$H_{XY}(\xi,\, \eta) \;=\; C\big(F_X(\xi),\ G_Y(\eta)\big)\,. \tag{4.18}$$

If F_X and G_Y are continuous, then the copula C is unique; otherwise, C is uniquely determined on (Range of F_X $\times$ Range of G_Y). Conversely, if C is a copula and F_X and G_Y are univariate distribution functions, then H_{XY} given by (4.18) *is a joint distribution function with marginals F_X and G_Y.*

This can also lead to the following essential result that plays a fundamental role in practical applications.

Corollary 4.3. *Let C, $H_{XY}(\xi, \eta)$ and $F_X(\xi)$ and $G_Y(\eta)$ be as in Theorem 4.2, and let us assume that F_X, G_Y are continuous strictly increasing marginals. Then $\forall\,(u, v) \in \mathbb{I}^2$*

$$C(u, v) \;=\; H_{XY}\big(F_X^{-1}(\xi),\ G_Y^{-1}(\eta)\big)\,. \tag{4.19}$$

In addition, it is important to make an example of the following three special copulas:

$$\begin{aligned} W_2(u,v) &= \max\{u+v-1,\,0\}\,; \\ \Pi_2(u,v) &= u\,v\,; \\ M_2(u,v) &= \min\{u,\,v\}\,. \end{aligned} \tag{4.20}$$

It is easy to verify that W_2, Π_2 and M_2 are valid copulas (Embrechts et al., 1999). Note that the copulas W_2 and M_2 provide general bounds, since for any bivariate copula C and $\forall\,(u,v) \in \mathbb{I}^2$, $W_2(u,v) \le C(u,v) \le M_2(u,v)$. In particular, for continuous random variables X and Y the following holds:

1. The variable Y is a strictly decreasing function of X with probability one if, and only if, $C_{XY} = W_2$. Random variables with copula W_2 are often called counter-monotonic.
2. The variables Y and X are independent if, and only if, their copula is Π_2: obviously, $H_{XY} = \Pi_2\big(F_X, G_Y\big) = F_X\,G_Y$.
3. The variable Y is a strictly increasing function of X with probability one if, and only if, $C_{XY} = M_2$. Random variables with copula M_2 are often called co-monotonic.

A particular subclass of copulas, called Archimedean, features many useful properties. According to Genest et al. (1998), Archimedean copulas provide a host of models that are versatile in terms of both the nature and strength of the association they induce between the variables.

In some situations, there exists a function φ such that

$$\varphi\big(C(u,v)\big) = \varphi(u) + \varphi(v). \tag{4.21}$$

Evidently, we also have an equivalent presentation: $\varphi\big(H_{XY}(\xi,\eta)\big) = \varphi\big(F_X(\xi)\big) + \varphi\big(G_Y(\eta)\big)$, that is bivariate distribution function H_{XY} can be written as a sum of functions of marginals F_X and G_Y. Because the expressions that can be used for the construction of copulas is of particular interest, it is necessary to solve the relation (4.21). First, the inverse of φ should be defined properly (Nelsen, 2006)

Definition 4.2. Let $\varphi : \mathbb{I} \to [0,\infty]$ is continuous and strictly decreasing, and $\varphi(1) = 0$. Let φ^{-1} denote the ordinary inverse function of φ. The pseudo-inverse of φ is the function $\varphi^{[-1]} : [0,\infty] \to \mathbb{I}$ given by

$$\varphi^{[-1]} = \begin{cases} \varphi^{-1} & 0 \le t \le \varphi(0) \\ 0 & \varphi(0) \le t \le \infty \end{cases}$$

Clearly, if $\varphi(0) = \infty$, then $\varphi^{[-1]} = \varphi^{-1}$. Also φ is convex if, and only if, $\varphi^{[-1]}$ is convex. Then, function $C(u,v)$ of the form

$$C(u,v) = \varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big)$$

is a copula if, and only if, the pseudo-inverse $\varphi^{[-1]}$ is a convex decreasing function. Copulas of the form above are called Archimedean copulas. The function φ is called a generator of the copula.

An important example of Archimedean copula is the Gumbel copula (also known as Gumbel-Hougard copula)

$$C(u,v) = \exp\left\{-\left[(-\log u)^{\alpha} + (-\log v)^{\alpha}\right]^{1/\alpha}\right\}\,,\ \alpha \geq 1\,. \tag{4.22}$$

In this case, $\varphi = (-\log\, t)^{\alpha}$.

4.2.2 Bivariate extreme value copula

Let random variables X and Y be joined by the copula C and also let $\hat{C}_N$ denote the copula of component-wise maxima $\hat{X}_N$ and $\hat{Y}_N$ defined in (4.1). From Theorem 3.3.1 of Nelsen (2006), it is known that

$$\hat{C}_N(u,v) = C^N\left(u^{1/N},\, v^{1/N}\right),\ (u,v) \in \mathbb{I}^2.$$

The limit of the sequence $\left\{\hat{C}_N\right\}$ leads to the following definition of an extreme value copula.

Definition 4.3. A copula $C_{_E}$ is an extreme value copula if there exists a copula C such that

$$\lim_{N\to\infty} C^N\left(u^{1/N},\, v^{1/N}\right) \,=\, C_{_E}\big(u,\, v\big),\ (u,v) \in \mathbb{I}^2\,. \tag{4.23}$$

C is said to belong to the domain of attraction of $C_{_E}$. It is also easy to verify that $C_{_E}$ satisfies the relationship

$$C_{_E}\left(u^k,\, v^k\right) \,=\, C_{_E}^k\big(u,\, v\big),\ k > 0. \tag{4.24}$$

Eq. (4.24) emphasizes a fundamental feature of EV copulas known as max-stability property, cf., e.g. Salvadori et al. (2007).

An important example of the extreme value copula is the Gumbel-Hougaard copula (4.22). Note that, in fact, there is no other Archimedean copula that is also an extreme-value copula.

Pickands (1981) obtained a general form of a bivariate extreme-value copula of a joint bivariate extreme value distribution with EV marginals $F(\xi)$ and $G(\eta)$, as follows:

$$C_{_E}\,(u,\, v) \,=\, \exp\left\{\log\,(uv)\mathcal{D}\left(\frac{\log(u)}{\log(uv)}\right)\right\},\ (u,v) \in \mathbb{I}^2\,. \tag{4.25}$$

where the dependence function $\mathcal{D}(\cdot)$ satisfies $\mathcal{D}(x) : [0,1] \to [\,\max(x,\, 1-x),\ 1]$, cf. also Gudendorf and Segers (2010). Properties of the dependence function $\mathcal{D}(\cdot)$ are as follows:

- $\mathcal{D}(0) \,=\, \mathcal{D}(1) \,=\, 1$.
- $\max(x,\, 1-x) \leq \mathcal{D}(x) \leq 1$, for $0 \leq x \leq 1$.
- $\mathcal{D}(x) \,=\, 1$ implies that $C_{_E} \,=\, u\,v$, that is, the components of a bivariate extreme value vector are independent.

- $\mathcal{D}(x) = \max(x,\, 1-x)$ implies that the components are equal with probability one.
- $\mathcal{D}$ is convex. In this way, $\mathcal{D}\big(\lambda\, x + (1-\lambda)\, y\big) \leq \lambda\, \mathcal{D}(x) + (1-\lambda)\, \mathcal{D}(y)$.
- If $\mathcal{D}_i$ are dependence functions, so is $\sum_{i=1}^{n} \alpha_i\, \mathcal{D}_i$, where $\forall i\ \alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$.

$\mathcal{D}$ may or may not be differentiable. In the former case, H has a joint density everywhere; in the latter, H has a singular component and is not differentiable in a certain region of its support.

Evidently, a set of the dependence functions $\mathcal{D}$ is infinite dimensional, which gives a large freedom in the construction of the extreme value copula. At the same time, this leads to difficulties, for no finite parametrization exists for such a family. In practical applications, only parametric subfamilies are used. However, by a careful choice, it is possible to ensure that a wide enough subclass of the entire limit family is approximated. We shall consider some key examples of $\mathcal{D}$.

Gumbel (1958a, 1960b, 1965) described two general forms for bivariate extreme value distributions in terms of the marginals (univariate extreme-value distributions). Type A bivariate extreme value distribution is also known as the mixed model. It is given by the following joint CDF:

$$H(\xi,\, \eta) \;=\; F(\xi)G(\eta)\, \exp\left\{-\theta \left[\frac{1}{\log F(\xi)} + \frac{1}{\log G(\eta)}\right]^{-1}\right\}. \tag{4.26}$$

The corresponding copula is given as

$$C_{E}(u,\, v) = uv\, \exp\left(-\theta\, \frac{\log u\, \log v}{\log uv}\right), \tag{4.27}$$

The mixed model sets

$$\mathcal{D}(x) = \theta x^2 - \theta x + 1\,,\ 0 \leq \theta \leq 1\,. \tag{4.28}$$

Type B bivariate extreme value distribution is also referred to as the logistic model (see also Eq.(4.13)):

$$H(\xi,\, \eta) \;=\; \exp\left\{-\Big[\big(-\log F(\xi)\big)^{m} + \big(-\log G(\eta)\big)^{m}\Big]^{\frac{1}{m}}\right\}, m \geq 1\,. \tag{4.29}$$

The copula that corresponds to the Type B extreme-value distribution is the Gumbel-Hougard copula described by Eq. (4.22). The logistic model sets

$$\mathcal{D}(x) = [x^m + (1-x)^m]^{1/m},\ m \geq 1\,. \tag{4.30}$$

It is clear that for $m = 1$ the components are independent, whereas for $m \to \infty$ the $M_2(u, v)$ copula is obtained, which corresponds to the complete dependence model. (see Eq. (4.20) on page 104). The Pearson product-moment correlation coefficient is $\rho = 1 - m^{-2}$ (Gumbel and Mustafi, 1967; Yue, 2001a).

Type C bivariate extreme value distribution, or the biextremal model, is an example of a non-differentiable model, cf., e.g. Tiago de Oliveira (1974, 1984). For this distribution, the joint distribution function is

$$H(\xi,\,\eta) \;=\; \exp\left[\min\left\{\log F(\xi);\, \log G(\eta) + \log F(\xi)(1-\phi)\right\}\right]\,, \qquad 0 \le \phi \le 1\,. \tag{4.31}$$

The corresponding copula is defined by the dependence function of the form

$$\mathcal{D}(x) = \max\{x,\, 1-\phi x\},\;\; 0 \le \phi \le 1\,. \tag{4.32}$$

More results for this copula can be found, in particular, in Balakrishnan and Lai (2009).

Chapter 5

The ACER Method. Bivariate Case

In reliability engineering and design of offshore structures, probabilistic approaches are frequently adopted. They require the estimation of extreme quantiles of oceanographic data based on statistical information. Due to strong correlation between such random variables as, e.g., wave heights and wind speeds, application of the multivariate, or bivariate in the simplest case, extreme value theory is sometimes necessary.

In this chapter we demonstrate that the concept of average conditional exceedance rate (ACER) can be extended in a natural way to also cover several dimensions, in particular, two. A vehicle, first obtained by Naess (2011), provides a nonparametric statistical estimate of the bivariate extreme value distribution given by a bivariate time series. It will be shown that the bivariate ACER function is able to cover both spatial and temporal dependence characteristics of the given time series. Thus, it covers all simultaneous and non-simultaneous extreme events. From a practical point of view, this makes it possible to investigate the true behavior of the bivariate extreme value distribution for a particular case, and at the same time check the validity of the proposed copula models for bivariate extremes.

As the first effort in investigating the functional representation of the empirically estimated bivariate ACER surface, the bivariate extreme value copula approach is adopted. Specifically, Naess and Karpa (2015a,b) use the asymmetric logistic and Gumbel logistic models combined with asymptotically consistent marginal extreme value distributions based on the univariate ACER functions. Since the univariate ACER functions have proved to portray accurately the marginal tail behaviour, this may offer an opportunity to verify to some extent the viability of copula models to capture the dependence structure of bivariate extreme value distributions.

5.1 Cascade of Bivariate Conditioning Approximations

Let us consider a bivariate stochastic process $Z(t) = \big(X(t), Y(t)\big)$ with dependent component processes, which has been observed over a time interval, such as $(0, \mathcal{T})$. Similarly to the univariate case, we assume that the sampled values $(X_1, Y_1), \ldots, (X_N, Y_N)$ are allocated to the (generally equidistant) discrete times

$t_1, \ldots, t_N$ in $(0, \mathcal{T})$. Our goal now is to accurately determine the joint distribution function of the extreme value vector $\hat{Z}_N = \left(\hat{X}_N, \hat{Y}_N\right)$, with components $\hat{X}_N = \max\{X_j\,; j = 1, \ldots, N\}$, and $\hat{Y}_N$ with a similar definition. Particularly, we want to accurately estimate the joint distribution $H(\xi, \eta) = \text{Prob}\left(\hat{X}_N \le \xi, \hat{Y}_N \le \eta\right) = \text{Prob}\left(X_N \le \xi, Y_N \le \eta \ldots, X_1 \le \xi, Y_1 \le \eta\right)$ for large values of ξ and η.

In the following, we outline the implementation of a cascade of approximations based on conditioning, where the first approximation is a one-step memory approximation, which may be considered a Markov-like approximation. This approximation concept was described by Naess (1985, 1990). However, it is emphasized that it is not a Markov chain approximation.

From the definition of $H(\xi, \eta)$, it follows that

$$\begin{aligned} H(\xi,\eta) &= \text{Prob}\big(X_N \le \xi, Y_N \le \eta, \ldots, X_1 \le \xi, Y_1 \le \eta\big) \\ &= \text{Prob}\big(X_N \le \xi, Y_N \le \eta \,|\, X_{N-1} \le \xi, Y_{N-1} \le \eta, \ldots, X_1 \le \xi, Y_1 \le \eta\big) \\ &\quad \cdot \text{Prob}\big(X_{N-1} \le \xi, Y_{N-1} \le \eta, \ldots, X_1 \le \xi, Y_1 \le \eta\big) \\ &= \prod_{j=2}^{N} \text{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_1 \le \xi, Y_1 \le \eta\big) \\ &\quad \cdot \text{Prob}\big(X_1 \le \xi, Y_1 \le \eta\big) \end{aligned} \tag{5.1}$$

The development of the cascade of conditioning approximations begins by considering the following basic case of independent sample points (Naess, 2011).

Independent sample points

The first approximation of the cascade is obtained by assuming that variables X_j of the component $X(t)$ are statistically independent, with the same assumption for the component $Y(t)$. In this special case, we obtain

$$\begin{aligned} H(\xi,\eta) &= \prod_{j=1}^{N} \text{Prob}\big(X_j \le \xi, Y_j \le \eta\big) \\ &= \prod_{j=1}^{N} \big\{1 - \text{Prob}(X_j > \xi) - \text{Prob}(Y_j > \eta) + \text{Prob}(X_j > \xi, Y_j > \eta)\big\}. \end{aligned} \tag{5.2}$$

Now, we introduce the assignment:

$$\begin{aligned} \alpha_{1j}(\xi;\eta) &:= \text{Prob}(X_j > \xi); \\ \beta_{1j}(\eta;\xi) &:= \text{Prob}(Y_j > \eta); \\ \gamma_{1j}(\xi,\eta) &:= \text{Prob}(X_j > \xi, Y_j > \eta), \end{aligned} \tag{5.3}$$

for $1 \le j \le N$. Note that although neither $\alpha_{1j}(\xi;\eta)$ depend on η nor $\beta_{1j}(\eta;\xi)$ depend on ξ, we keep this notation for the correctness of the further derivations.

Eq. (5.2) can then be rewritten as

$$\begin{aligned} H(\xi,\eta) &= \prod_{j=1}^{N} \big\{1 - \alpha_{1j}(\xi;\eta) - \beta_{1j}(\eta;\xi) + \gamma_{1j}(\xi,\eta)\big\} \\ &\approx H_1(\xi,\eta) = \exp\left\{ -\sum_{j=1}^{N} \big(\alpha_{1j}(\xi;\eta) + \beta_{1j}(\eta;\xi) - \gamma_{1j}(\xi,\eta)\big) \right\} ;\ \xi,\eta \to \infty\,, \end{aligned} \tag{5.4}$$

where the approximation $1 - x \approx \exp(-x)$ has been applied to the sum $\big(\alpha_{1j}(\xi;\eta) + \beta_{1j}(\eta;\xi) - \gamma_{1j}(\xi,\eta)\big)$. The relative error of this approximation is up to 0.5% for values of $|x| < 0.1$, and it rapidly decreases for decreasing values of $|x|$.

Conditioning on one and two previous sample points

In general, the variables X_j are statistically dependent, as are the variables Y_j. In this case, the first genuine conditioning approximation is obtained by neglecting all previous conditioning events except for the immediate predecessor in Eq. (5.1), that is,

$$\begin{aligned} &\mathrm{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_1 \le \xi, Y_1 \le \eta\big) \\ &\approx \mathrm{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big), \text{ for } j = 2, \ldots, N. \end{aligned} \tag{5.5}$$

Consequently, the following one-step memory approximation is adopted:

$$\begin{aligned} H(\xi,\eta) \approx \prod_{j=2}^{N} &\mathrm{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big) \\ &\cdot \mathrm{Prob}\big(X_1 \le \xi, Y_1 \le \eta\big) \end{aligned} \tag{5.6}$$

This may be rewritten as

$$\begin{aligned} H(\xi,\eta) \approx \prod_{j=2}^{N} \Big\{ &1 - \mathrm{Prob}\big(X_j > \xi \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big) \\ &- \mathrm{Prob}\big(Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big) \\ &+ \mathrm{Prob}\big(X_j > \xi, Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big) \Big\} \\ \cdot \big\{ &1 - \mathrm{Prob}(X_1 > \xi) - \mathrm{Prob}(Y_1 > \eta) + \mathrm{Prob}(X_1 > \xi, Y_1 > \eta) \big\} \end{aligned} \tag{5.7}$$

By introducing the notation

$$\begin{aligned} \alpha_{2j}(\xi;\eta) &:= \mathrm{Prob}\big(X_j > \xi \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big); \\ \beta_{2j}(\eta;\xi) &:= \mathrm{Prob}\big(Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big); \\ \gamma_{2j}(\xi,\eta) &:= \mathrm{Prob}\big(X_j > \xi, Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta\big), \end{aligned} \tag{5.8}$$

for $2 \leq j \leq N$, we obtain as in Eq. (5.4) that for high values of ξ and η, the following approximation holds:

$$\begin{aligned} H(\xi,\eta) &\approx \prod_{j=2}^{N} \Big\{1-\alpha_{2j}(\xi;\eta)-\beta_{2j}(\eta;\xi)+\gamma_{2j}(\xi,\eta)\Big\} \\ &\quad \cdot \big\{1-\alpha_{11}(\xi;\eta)-\beta_{11}(\eta;\xi)+\gamma_{11}(\xi,\eta)\big\} \\ \approx\ H_2(\xi,\eta) &= \exp\left\{-\sum_{j=2}^{N}\big(\alpha_{2j}(\xi;\eta)+\beta_{2j}(\eta;\xi)-\gamma_{2j}(\xi,\eta)\big)\right. \\ &\quad \left. -\big(\alpha_{11}(\xi;\eta)+\beta_{11}(\eta;\xi)-\gamma_{11}(\xi,\eta)\big)\right\};\ \xi,\eta\to\infty\,. \end{aligned} \tag{5.9}$$

To extend the process of approximating $H(\xi,\eta)$, the third level of approximation is achieved by assuming that

$$\begin{aligned} &\mathrm{Prob}\big(X_j\leq\xi, Y_j\leq\eta \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\ldots,X_1\leq\xi, Y_1\leq\eta\big) \\ &\approx \mathrm{Prob}\big(X_j\leq\xi, Y_j\leq\eta \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\, X_{j-2}\leq\xi, Y_{j-2}\leq\eta\big), \end{aligned} \tag{5.10}$$

for $j=3,\ldots,N$. Therefore, Eq. (5.1) can be rewritten as:

$$\begin{aligned} H(\xi,\eta) \approx \prod_{j=3}^{N} &\mathrm{Prob}\big(X_j\leq\xi, Y_j\leq\eta \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\, X_{j-2}\leq\xi, Y_{j-2}\leq\eta\big) \\ &\cdot \mathrm{Prob}\big(X_2\leq\xi, Y_2\leq\eta \,|\, X_1\leq\xi, Y_1\leq\eta\big) \\ &\cdot \mathrm{Prob}\big(X_1\leq\xi, Y_1\leq\eta\big) \end{aligned} \tag{5.11}$$

By employing the complementary event, this can be rewritten as

$$\begin{aligned} H(\xi,\eta) = \prod_{j=3}^{N} &\Big\{1-\mathrm{Prob}\big(X_j>\xi \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\, X_{j-2}\leq\xi, Y_{j-2}\leq\eta\big) \\ &\quad -\mathrm{Prob}\big(Y_j>\eta \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\, X_{j-2}\leq\xi, Y_{j-2}\leq\eta\big) \\ &\quad +\mathrm{Prob}\big(X_j>\xi, Y_j>\eta \,|\, X_{j-1}\leq\xi, Y_{j-1}\leq\eta,\, X_{j-2}\leq\xi, Y_{j-2}\leq\eta\big)\Big\} \\ &\cdot\Big\{1-\mathrm{Prob}\big(X_2>\xi \,|\, X_1\leq\xi, Y_1\leq\eta\big)-\mathrm{Prob}\big(Y_2>\eta \,|\, X_1\leq\xi, Y_1\leq\eta\big) \\ &\quad +\mathrm{Prob}\big(X_2>\xi, Y_2>\eta \,|\, X_1\leq\xi, Y_1\leq\eta\big)\Big\} \\ &\cdot\big\{1-\mathrm{Prob}(X_1>\xi)-\mathrm{Prob}(Y_1>\eta)+\mathrm{Prob}(X_1>\xi, Y_1>\eta)\big\}. \end{aligned} \tag{5.12}$$

Now, we are able to introduce the following assignment:

$$\begin{aligned}
\alpha_{3j}(\xi;\eta) &:= \mathrm{Prob}\big(X_j > \xi \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta,\, X_{j-2} \le \xi, Y_{j-2} \le \eta\big);\\
\beta_{3j}(\eta;\xi) &:= \mathrm{Prob}\big(Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta,\, X_{j-2} \le \xi, Y_{j-2} \le \eta\big);\\
\gamma_{3j}(\xi,\eta) &:= \mathrm{Prob}\big(X_j > \xi, Y_j > \eta \,|\, X_{j-1} \le \xi, Y_{j-1} \le \eta,\, X_{j-2} \le \xi, Y_{j-2} \le \eta\big),
\end{aligned} \tag{5.13}$$

for $3 \le j \le N$.

$$\begin{aligned}
H(\xi,\eta) = \prod_{j=3}^{N} &\Big\{1 - \alpha_{3j}(\xi;\eta) - \beta_{3j}(\eta;\xi) + \gamma_{3j}(\xi,\eta)\Big\}\\
&\cdot \Big\{1 - \alpha_{22}(\xi;\eta) - \beta_{22}(\eta;\xi) + \gamma_{22}(\xi,\eta)\Big\}\\
&\cdot \big\{1 - \alpha_{11}(\xi;\eta) - \beta_{11}(\eta;\xi) + \gamma_{11}(\xi,\eta)\big\}\\
\approx\; H_3(\xi,\eta) = \;\exp\Bigg\{ &-\sum_{j=3}^{N}\big(\alpha_{3j}(\xi;\eta) + \beta_{3j}(\eta;\xi) - \gamma_{3j}(\xi,\eta)\big)\\
&- \big(\alpha_{22}(\xi;\eta) + \beta_{22}(\eta;\xi) - \gamma_{22}(\xi,\eta)\big)\\
&- \big(\alpha_{11}(\xi;\eta) + \beta_{11}(\eta;\xi) - \gamma_{11}(\xi,\eta)\big)\Bigg\}\;;\; \xi,\eta \to \infty\,.
\end{aligned} \tag{5.14}$$

Conditioning on $k-1$ previous sample points

It has been observed that in the univariate case, conditioning on one previous data point is sometimes sufficient for capturing the effect of dependence in the time series to a large extent (Karpa and Naess, 2013; Naess and Gaidai, 2009). However, there are also cases in which conditioning on one previous data point is not sufficient. This can only be ascertained by a method that shows the complete picture regarding the importance of dependence on the extreme value distribution. Our proposed solution to this situation is obtained by introducing a cascade of conditioning approximations beyond the one- or two-step approximation presented above.

We start by defining the following set of events,

$$\mathcal{C}_{kj}(\xi,\eta) = \big\{X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_{j-k+1} \le \xi, Y_{j-k+1} \le \eta\big\} \tag{5.15}$$

From Eq. (5.1) and conditioning on not more than $k-1$ previous data points, where $k = 2, \ldots, N$ and $j \ge k$, it is observed that

$$H(\xi,\eta) = \prod_{j=k}^{N} \mathrm{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, \mathcal{C}_{kj}(\xi,\eta)\big) \cdot \mathrm{Prob}\big(\mathcal{C}_{kk}(\xi,\eta)\big)\,, \tag{5.16}$$

where

$$\begin{aligned}
\mathrm{Prob}\big(\mathcal{C}_{kk}(\xi,\eta)\big) = &\,\mathrm{Prob}\big(X_{k-1} \le \xi, Y_{k-1} \le \eta \,|\, \mathcal{C}_{k-1,k-1}(\xi,\eta)\big)\\
&\cdot \mathrm{Prob}\big(\mathcal{C}_{k-1,k-1}(\xi,\eta)\big)\,.
\end{aligned} \tag{5.17}$$

By introducing the notation

$$\begin{aligned}
\alpha_{kj}(\xi;\eta) &:= \mathrm{Prob}\big(X_j > \xi \,|\, \mathcal{C}_{kj}(\xi,\eta)\big);\\
\beta_{kj}(\eta;\xi) &:= \mathrm{Prob}\big(Y_j > \eta \,|\, \mathcal{C}_{kj}(\xi,\eta)\big);\\
\gamma_{kj}(\xi,\eta) &:= \mathrm{Prob}\big(X_j > \xi, Y_j > \eta \,|\, \mathcal{C}_{kj}(\xi,\eta)\big),
\end{aligned} \tag{5.18}$$

for $k \le j \le N$, it can now be shown that

$$\begin{aligned}
&\prod_{j=k}^{N} \mathrm{Prob}\big(X_j \le \xi, Y_j \le \eta \,|\, \mathcal{C}_{kj}(\xi,\eta)\big)\\
&\approx \exp\left\{ -\sum_{j=k}^{N} \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big) \right\} \;;\; \xi,\eta \to \infty\,.
\end{aligned} \tag{5.19}$$

Similarly, it is found that

$$\begin{aligned}
\mathrm{Prob}\big(\mathcal{C}_{kk}(\xi,\eta)\big) &\approx \exp\big\{ -\big(\alpha_{k-1,k-1}(\xi;\eta) + \beta_{k-1,k-1}(\eta;\xi) - \gamma_{k-1,k-1}(\xi,\eta)\big)\big\}\\
&\quad \cdot \mathrm{Prob}\big(\mathcal{C}_{k-1,k-1}(\xi,\eta)\big)\\
&\approx \exp\left\{ -\sum_{j=1}^{k-1} \big(\alpha_{jj}(\xi;\eta) + \beta_{jj}(\eta;\xi) - \gamma_{jj}(\xi,\eta)\big) \right\} \;;\; \xi,\eta \to \infty\,.
\end{aligned} \tag{5.20}$$

Hence, we finally obtain the result

$$\begin{aligned}
H(\xi,\eta) \approx\; H_k(\xi,\eta) = \exp\Bigg\{ &-\sum_{j=k}^{N} \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big)\\
&-\sum_{j=1}^{k-1} \big(\alpha_{jj}(\xi;\eta) + \beta_{jj}(\eta;\xi) - \gamma_{jj}(\xi,\eta)\big) \Bigg\} \;;\; \xi,\eta \to \infty\,.
\end{aligned} \tag{5.21}$$

Thus, based on the definition of the extreme value distribution $H(\xi,\eta)$ and the properties of conditional probability, we have constructed a set $\big\{H_k(\xi,\eta)\big\}_{k=1}^{N}$ of conditional probability distributions that converges to the target distribution $H(\xi,\eta)$ of the extreme value $\hat{Z}_N = \big(\hat{X}_N, \hat{Y}_N\big)$ in the limit as k increases.

Summarizing the cascade, we have:

$$
\begin{aligned}
H_1(\xi,\eta) &= \exp\left\{-\sum_{j=1}^{N}\big(\alpha_{1j}(\xi;\eta)+\beta_{1j}(\eta;\xi)-\gamma_{1j}(\xi,\eta)\big)\right\};\\
H_2(\xi,\eta) &= \exp\left\{-\sum_{j=2}^{N}\big(\alpha_{2j}(\xi;\eta)+\beta_{2j}(\eta;\xi)-\gamma_{2j}(\xi,\eta)\big)\right.\\
&\qquad\left.-\big(\alpha_{11}(\xi;\eta)+\beta_{11}(\eta;\xi)-\gamma_{11}(\xi,\eta)\big)\right\};\\
H_3(\xi,\eta) &= \exp\left\{-\sum_{j=3}^{N}\big(\alpha_{3j}(\xi;\eta)+\beta_{3j}(\eta;\xi)-\gamma_{3j}(\xi,\eta)\big)\right.\\
&\qquad-\big(\alpha_{22}(\xi;\eta)+\beta_{22}(\eta;\xi)-\gamma_{22}(\xi,\eta)\big)\\
&\qquad\left.-\big(\alpha_{11}(\xi;\eta)+\beta_{11}(\eta;\xi)-\gamma_{11}(\xi,\eta)\big)\right\};\\
&\vdots\\
H_k(\xi,\eta) &= \exp\left\{-\sum_{j=k}^{N}\big(\alpha_{kj}(\xi;\eta)+\beta_{kj}(\eta;\xi)-\gamma_{kj}(\xi,\eta)\big)\right.\\
&\qquad\left.-\sum_{j=1}^{k-1}\big(\alpha_{jj}(\xi;\eta)+\beta_{jj}(\eta;\xi)-\gamma_{jj}(\xi,\eta)\big)\right\};\\
&\downarrow\ (k \text{ increases})\\
H(\xi,\eta) &= \mathrm{Prob}\big(X_N\le\xi, Y_N\le\eta\ldots, X_1\le\xi, Y_1\le\eta\big)
\end{aligned}
\tag{5.22}
$$

For most applications and to have practical significance, the following assumption on this cascade of approximations is made: there is an effective k_e satisfying $k_e \ll N$ such that $H(\xi,\eta) = H_{k_e}(\xi,\eta)$. Then, $H_1(\xi,\eta) \le H_2(\xi,\eta) \le \ldots \le H_{k_e}(\xi,\eta) = H(\xi,\eta)$. Note that for a k-dependent stationary bivariate stochastic process $Z(t) = \big(X(t), Y(t)\big)$, that is, for data where Z_i and Z_j are independent componentwise whenever $|j-i| > k$, then $H(\xi,\eta) = H_{k+1}(\xi,\eta)$ exactly and $\lim_{N\to\infty} H_1(\xi,\eta) = \lim_{N\to\infty} H(\xi,\eta)$ (Watson, 1954). In fact, it can be shown that $\lim_{N\to\infty} H_1(\xi,\eta) = \lim_{N\to\infty} H(\xi,\eta)$ is true for conditions weaker than k-dependence (Leadbetter et al., 1983). However, for finite values of N, the picture is considerably more complex, and purely asymptotic results should be used with some caution.

It will be verified that the property $k_e \ll N$ is indeed satisfied for the type of data analysed in the present research. Furthermore, under this assumption, $\sum_{j=1}^{k-1}\big(\alpha_{jj}(\xi;\eta)+\beta_{jj}(\eta;\xi)-\gamma_{jj}(\xi,\eta)\big)$ is generally negligible compared to the sum

$\sum_{j=k}^{N} \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big)$. This leads to the approximation

$$H(\xi,\eta) \approx H_k(\xi,\eta) \approx \exp\left\{ -\sum_{j=k}^{N} \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big) \right\} ; \ \xi,\eta \to \infty , \tag{5.23}$$

from which it emerges that for the estimation of the bivariate extreme value distribution, it is necessary and sufficient to estimate the sequence of functions $\Big\{ \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big) \Big\}_{j=k}^{N}$ (Naess and Karpa, 2015a; Naess, 2011).

5.2 Empirical Estimation of the Bivariate ACER

To obtain a more compact representation, it is expedient to introduce the concept of the kth-order bivariate average conditional exceedance rate (ACER) function, as follows:

$$\mathcal{E}_k(\xi,\eta) \ = \ \frac{1}{N-k+1} \sum_{j=k}^{N} \big(\alpha_{kj}(\xi;\eta) + \beta_{kj}(\eta;\xi) - \gamma_{kj}(\xi,\eta)\big), \ k = 1,\, 2, \ldots \tag{5.24}$$

Hence, when $N \gg k$, we may write

$$H(\xi,\eta) \approx \exp\left\{ -(N-k+1)\, \mathcal{E}_k(\xi,\eta) \right\} , \ \xi,\eta \to \infty . \tag{5.25}$$

From this equation, the result for, e.g., $\text{Prob}(\hat{Y}_N \le \eta \,|\, \hat{X}_N \le \xi)$ follows by writing

$$\begin{aligned} \text{Prob}(\hat{Y}_N \le \eta \,|\, \hat{X}_N \le \xi) &= \frac{H(\xi,\eta)}{\text{Prob}(\hat{X}_N \le \xi)} \\ &\approx \exp\{ -(N-k+1)\big(\mathcal{E}_k(\xi,\eta) - \varepsilon_k(\xi)\big)\} ; \ \xi,\eta \to \infty , \end{aligned} \tag{5.26}$$

where $\varepsilon_k(\xi)$ is the kth-order univariate ACER function for the time series $X(t)$ (Karpa and Naess, 2013; Naess and Gaidai, 2009).

A few more details on the numerical estimation of the ACER functions are useful. We start by introducing a set of random functions. For $k = 2, \ldots, N$ and $k \le j \le N$, let

$$\begin{aligned} A_{kj}(\xi;\eta) \ &= \ \mathbf{1}\{\underline{X_j > \xi}, \\ &\qquad X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_{j-k+1} \le \xi, Y_{j-k+1} \le \eta\}; \\ B_{kj}(\eta;\xi) \ &= \ \mathbf{1}\{\underline{Y_j > \eta}, \\ &\qquad X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_{j-k+1} \le \xi, Y_{j-k+1} \le \eta\}; \\ G_{kj}(\xi,\eta) \ &= \ \mathbf{1}\{\underline{X_j > \xi, Y_j > \eta}, \\ &\qquad X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_{j-k+1} \le \xi, Y_{j-k+1} \le \eta\}; \\ C_{kj}(\xi,\eta) \ &= \ \mathbf{1}\{\mathcal{C}_{kj}(\xi,\eta)\} = \\ &\qquad \mathbf{1}\{X_{j-1} \le \xi, Y_{j-1} \le \eta, \ldots, X_{j-k+1} \le \xi, Y_{j-k+1} \le \eta\}, \end{aligned} \tag{5.27}$$

where $\mathbf{1}\{\mathcal{A}\}$ denotes the indicator function of some event $\mathcal{A}$.

From these definitions, it follows that

$$\alpha_{kj}(\xi;\eta) = \frac{\mathbb{E}[A_{kj}(\xi;\eta)]}{\mathbb{E}[C_{kj}(\xi,\eta)]}\,, \tag{5.28}$$

$$\beta_{kj}(\eta;\xi) = \frac{\mathbb{E}[B_{kj}(\eta;\xi)]}{\mathbb{E}[C_{kj}(\xi,\eta)]}\,, \tag{5.29}$$

$$\gamma_{kj}(\xi,\eta) = \frac{\mathbb{E}[G_{kj}(\xi,\eta)]}{\mathbb{E}[C_{kj}(\xi,\eta)]}\,, \tag{5.30}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator.

Assuming ergodicity of the process $Z(t) = \big(X(t), Y(t)\big)$, then clearly $\mathcal{E}_k(\xi,\eta) = \big(\alpha_{kk}(\xi;\eta) + \beta_{kk}(\eta;\xi) - \gamma_{kk}(\xi,\eta)\big) = \ldots = \big(\alpha_{kN}(\xi;\eta) + \beta_{kN}(\eta;\xi) - \gamma_{kN}(\xi,\eta)\big)$, and it may be assumed that for the bivariate time series at hand

$$\mathcal{E}_k(\xi,\eta) = \lim_{N\to\infty} \frac{\sum\limits_{j=k}^{N} \big(\mathrm{a}_{kj}(\xi;\eta) + \mathrm{b}_{kj}(\eta;\xi) - \mathrm{g}_{kj}(\xi,\eta)\big)}{\sum\limits_{j=k}^{N} \mathrm{c}_{kj}(\xi,\eta)}\,, \tag{5.31}$$

where $\mathrm{a}_{kj}(\xi;\eta)$, $\mathrm{b}_{kj}(\eta;\xi)$, $\mathrm{g}_{kj}(\xi,\eta)$ and $\mathrm{c}_{kj}(\xi,\eta)$ are the realized values of $A_{kj}(\xi;\eta)$, $B_{kj}(\eta;\xi)$, $G_{kj}(\xi,\eta)$ and $C_{kj}(\xi,\eta)$, respectively, for the observed time series.

Clearly, $\lim\limits_{\xi,\eta\to\infty} \mathbb{E}[C_{kj}(\xi,\eta)] = 1$. Hence, $\lim\limits_{\xi,\eta\to\infty} \tilde{\mathcal{E}}_k(\xi,\eta) \,/\, \mathcal{E}_k(\xi,\eta) = 1$, where

$$\tilde{\mathcal{E}}_k(\xi,\eta) = \lim_{N\to\infty} \frac{\sum\limits_{j=k}^{N} \Big(\mathbb{E}[A_{kj}(\xi;\eta)] + \mathbb{E}[B_{kj}(\eta;\xi)] - \mathbb{E}[G_{kj}(\xi,\eta)]\Big)}{N-k+1}\,. \tag{5.32}$$

The advantage of using the modified bivariate ACER function $\tilde{\mathcal{E}}_k(\xi,\eta)$ for $k \geq 2$ is that it is easier to use for non-stationary or long-term statistics than $\mathcal{E}_k(\xi,\eta)$. Because our focus is on the values of the ACER functions at the extreme levels, we may use any function that provides correct predictions of the appropriate ACER function at these extreme levels.

To demonstrate why Eq. (5.32) may be applicable for non-stationary time series, it is recognized that when $\xi,\eta \to \infty$,

$$\begin{aligned} H(\xi,\eta) &\approx \exp\{-(N-k+1)\,\mathcal{E}_k(\xi,\eta)\} \approx \exp\{-(N-k+1)\,\tilde{\mathcal{E}}_k(\xi,\eta)\} \\ &= \exp\left\{-\sum_{j=k}^{N} \Big(\mathbb{E}[A_{kj}(\xi;\eta)] + \mathbb{E}[B_{kj}(\eta;\xi)] - \mathbb{E}[G_{kj}(\xi,\eta)]\Big)\right\}. \end{aligned} \tag{5.33}$$

If the non-stationary time series can be segmented into K blocks such that $\mathbb{E}[A_{kj}(\xi;\eta)]$, $\mathbb{E}[B_{kj}(\eta;\xi)]$ and $\mathbb{E}[G_{kj}(\xi,\eta)]$ remain approximately constant within

each block and such that

$$\sum_{j\in C_i}\mathbb{E}[A_{kj}(\xi;\eta)]\approx\sum_{j\in C_i}\mathrm{a}_{kj}(\xi;\eta),$$
$$\sum_{j\in C_i}\mathbb{E}[B_{kj}(\eta;\xi)]\approx\sum_{j\in C_i}\mathrm{b}_{kj}(\eta;\xi)\text{ and}$$
$$\sum_{j\in C_i}\mathbb{E}[G_{kj}(\xi,\eta)]\approx\sum_{j\in C_i}\mathrm{g}_{kj}(\xi,\eta)$$

for a sufficient range of ξ,η values, where C_i denotes the set of indices for block no. i, $i=1,\ldots,K$, then $\sum_{j=k}^{N}\mathbb{E}[A_{kj}(\xi;\eta)]\approx\sum_{j=k}^{N}\mathrm{a}_{kj}(\xi;\eta)$ with a corresponding approximations of $\sum_{j=k}^{N}\mathbb{E}[B_{kj}(\xi;\eta)]$ and $\sum_{j=k}^{N}\mathbb{E}[G_{kj}(\xi;\eta)]$ (Naess, 2011). Hence, for a non-stationary bivariate time series, it is obtained that $(\xi,\eta\to\infty)$,

$$H(\xi,\eta)\approx\exp\{-(N-k+1)\hat{\mathcal{E}}_k(\xi,\eta)\}\,, \tag{5.34}$$

where

$$\hat{\mathcal{E}}_k(\xi,\eta)=\frac{1}{N-k+1}\sum_{j=k}^{N}\left(\mathrm{a}_{kj}(\xi;\eta)+\mathrm{b}_{kj}(\eta;\xi)-\mathrm{g}_{kj}(\xi,\eta)\right), \tag{5.35}$$

Now, we consider the problem of estimating confidence intervals for the bivariate ACER function. If several realizations of the time series $Z(t)=\big(X(t),Y(t)\big)$ are provided or if the time series can be appropriately sectioned into several records, e.g., several annual or other time span records, then the sample estimate of $\mathcal{E}_k(\xi,\eta)$ would be

$$\hat{\mathcal{E}}_k(\xi,\eta)=\frac{1}{R}\sum_{r=1}^{R}\hat{\mathcal{E}}_k^{(r)}(\xi,\eta)\,, \tag{5.36}$$

where R is the number of realizations (samples), and

$$\hat{\mathcal{E}}_k^{(r)}(\xi,\eta)=\frac{\sum_{j=k}^{N}\left(\mathrm{a}_{kj}^{(r)}(\xi;\eta)+\mathrm{b}_{kj}^{(r)}(\eta;\xi)-\mathrm{g}_{kj}^{(r)}(\xi,\eta)\right)}{\sum_{j=k}^{N}\mathrm{c}_{kj}^{(r)}(\xi,\eta)}\,, \tag{5.37}$$

for the stationary time series, or

$$\hat{\mathcal{E}}_k^{(r)}(\xi,\eta)=\frac{1}{N-k+1}\sum_{j=k}^{N}\left(\mathrm{a}_{kj}^{(r)}(\xi;\eta)+\mathrm{b}_{kj}^{(r)}(\eta;\xi)-\mathrm{g}_{kj}^{(r)}(\xi,\eta)\right), \tag{5.38}$$

for non-stationary time series, where the index (r) refers to realization no. r. The sample standard deviation $\hat{s}_k(\xi;\eta)$ can then be estimated by the standard formula,

$$\hat{s}_k^2(\xi;\eta)=\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\mathcal{E}}_k^{(r)}(\xi,\eta)-\hat{\mathcal{E}}_k(\xi,\eta)\right)^2. \tag{5.39}$$

Assuming that the realizations are independent, Eq. (5.39) leads to a good approximation of the 95 % confidence interval $\mathrm{CI} = \big(\mathrm{CI}^{-}(\eta), \mathrm{CI}^{+}(\eta)\big)$ for the value $\mathcal{E}_k(\xi, \eta)$, where

$$\mathrm{CI}^{\pm}(\xi, \eta) = \hat{\mathcal{E}}_k(\xi, \eta) \pm \tau \cdot \frac{\hat{s}_k(\xi, \eta)}{\sqrt{R}}\,, \tag{5.40}$$

and $\tau = t^{-1}\Big((1 - 0.95)/2,\ R - 1\Big)$ is the corresponding quantile of the Student's t-distribution with $R - 1$ degrees of freedom.

5.3 The T-year Return Level by the Bivariate ACER

In the previous section, we introduced the concept of bivariate average conditional exceedance rates and derived the methodology for its empirical estimation. Note that from the definition of $\mathcal{E}_k(\xi, \eta)$ follows that $\mathcal{E}_k(\xi, \eta) \cdot (N - k + 1)$ represents the expected number of the bivariate observations $Z_j = (X_j, Y_j)$ such that its components exceed corresponding levels ξ and η and follow after at least $k - 1$ previous simultaneous non-exceedances. Thus, the bivariate ACER function $\mathcal{E}_{k_e}(\xi, \eta)$, where k_e is such that

$$H(\xi, \eta) \;=\; H_{k_e}(\xi, \eta) \;=\; \exp\big\{ -(N - k + 1)\, \mathcal{E}_{k_e}(\xi, \eta)\big\}\,,$$

is able to describe the dependence structure of the considered bivariate time series, and it defines the exact joint distribution function.

This implies the capability to obtain high quantiles of the bivariate extreme value distribution. Thus, the joint T-year return period contour associated with the event that either $\hat{X}_N$ or $\hat{Y}_N$ or both are exceeded, see, for example, Salvadori et al. (2007), that is, $\Big\{\big(\hat{X}_N > \xi^{\mathrm{Tyr}}\big) \vee \big(\hat{Y}_N > \eta^{\mathrm{Tyr}}\big) \vee \big(\hat{X}_N > \xi^{\mathrm{Tyr}} \wedge \hat{Y}_N > \eta^{\mathrm{Tyr}}\big)\Big\}$, is represented by

$$1 - H^{1\mathrm{yr}}(\xi^{\mathrm{Tyr}},\, \eta^{\mathrm{Tyr}}) = \frac{1}{T}, \tag{5.41}$$

where $H^{1\mathrm{yr}}(\xi, \eta)$ is the joint distribution function of the annual maxima. Assuming that the duration of the observation period of the bivariate process $Z(t)$ is n_y years, then

$$H^{1\mathrm{yr}}(\xi, \eta) = \exp\left\{ -\frac{N - k + 1}{n_y}\, \mathcal{E}_k(\xi, \eta)\right\}. \tag{5.42}$$

From Eqs. (5.41) and (5.42), it follows that the joint T-year return levels $(\xi^{\mathrm{Tyr}},\, \eta^{\mathrm{Tyr}})$ are obtained as the solution of the implicit equation:

$$\mathcal{E}_k(\xi^{\mathrm{Tyr}},\, \eta^{\mathrm{Tyr}}) = -\,\log\left(1 - \frac{1}{T}\right) \frac{n_y}{N - k + 1}. \tag{5.43}$$

5.4 Functional Representation of the Empirically Estimated Bivariate ACER Surface

Clearly, the empirically estimated kth-order bivariate average conditional exceedance rate does not provide the necessary information for estimating high

quantiles of the joint extreme value distribution. The behaviour of the bivariate ACER as a continuous function of two variables cannot be determined using available statistical data. Therefore, the sub-asymptotic functional form of the ACER surface $\mathcal{E}_k(\xi, \eta)$ can potentially be obtained approximately through the copula representation of a bivariate extreme value distribution.

From the result by Sklar (1959), for any pair of random variables (X, Y) with marginal distribution functions $F_X(\xi)$ and $G_Y(\eta)$, the joint distribution function $H_{XY}(\xi, \eta) = \text{Prob}(X \le \xi, Y \le \eta)$ can be presented by the bivariate copula $C(u, v)$ as follows (see Theorem 4.2 on page 103): $H_{XY}(\xi, \eta) = C\big(F_X(\xi),\ G_Y(\eta)\big)$, cf., e.g., Nelsen (2006); Balakrishnan and Lai (2009). This result also applies to any bivariate extreme value distribution.

As discussed in Section 4.2.2 on page 105, Pickands (1981) discovered that a bivariate copula C_E is an extreme-value copula if and only if

$$C_E(u, v) = \exp\left\{\log(uv)\mathcal{D}\left(\frac{\log(u)}{\log(uv)}\right)\right\}\ ;\ 0 < u < 1,\ 0 < v < 1\,, \tag{5.44}$$

where the Pickands dependence function $\mathcal{D}(\cdot)$ is a convex function and satisfies $\mathcal{D}(x) : [0, 1] \to [\max(x, 1 - x),\ 1]$, cf., Gudendorf and Segers (2010).

Considering the above, any bivariate extreme value distribution $H(\xi, \eta)$ with marginal univariate extreme value distributions $F(\xi)$ and $G(\eta)$ is given by the formula:

$$H(\xi, \eta) = \exp\left\{\log\big(F(\xi)G(\eta)\big)\mathcal{D}\left(\frac{\log\big(F(\xi)\big)}{\log\big(F(\xi)G_Y(\eta)\big)}\right)\right\}. \tag{5.45}$$

We assume that asymptotically consistent marginal extreme value distributions $F(\xi)$ and $G(\eta)$ are represented by the corresponding univariate ACER functions, that is:

$$\begin{aligned} F(\xi) &\approx \exp\big\{-(N-k+1)\varepsilon_k^x(\xi)\big\}\,,\ \ \xi \ge \xi_1\,; \\ G(\eta) &\approx \exp\big\{-(N-k+1)\varepsilon_k^y(\eta)\big\}\,,\ \ \eta \ge \eta_1\,, \end{aligned} \tag{5.46}$$

where the sub-asymptotic functional form of the univariate ACER functions is defined previously to be

$$\begin{aligned} \varepsilon_k^x(\xi) &= q_k^x\ \exp\{-a_k^x(\xi - b_k^x)^{c_k^x}\}\,,\ \ \xi \ge \xi_1\,; \\ \varepsilon_k^y(\eta) &= q_k^y\ \exp\{-a_k^y(\eta - b_k^y)^{c_k^y}\}\,,\ \ \eta \ge \eta_1\,. \end{aligned} \tag{5.47}$$

Now, substituting Eq. (5.46) into Eq. (5.45) the following representation of the bivariate extreme value distribution applies:

$$H(\xi, \eta) = \exp\left\{-(N-k+1)\Big(\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)\Big)\mathcal{D}\left(\frac{\varepsilon_k^x(\xi)}{\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)}\right)\right\}, \quad \xi \ge \xi_1,\ \eta \ge \eta_1. \tag{5.48}$$

On the other hand, as it was discovered before, the bivariate extreme value distribution can be expressed through the bivariate ACER function, as

$H(\xi,\eta) = \exp\{-(N-k+1)\,\mathcal{E}_k(\xi,\eta)\}$ for high values of ξ and η. Thereby, according to Naess and Karpa (2015a), the functional form of the bivariate ACER surface can possibly be obtained by:

$$\mathcal{E}_k(\xi,\eta) = \Big(\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)\Big)\,\mathcal{D}\left(\frac{\varepsilon_k^x(\xi)}{\varepsilon_k^x(\xi)+\varepsilon_k^y(\eta)}\right),\ \ \xi \geq \xi_1,\ \eta \geq \eta_1. \tag{5.49}$$

Consequently, our aim now is to find the dependence function $\mathcal{D}(\cdot)$ that would provide optimal fit of the parametrical surface defined by Eq. (5.49) to the empirical bivariate ACER surface $\hat{\mathcal{E}}_k(\xi,\eta)$.

As described by Tawn (1988), subject to the form of the dependence function $\mathcal{D}(\cdot)$, different parametric bivariate extreme-value distributions can be considered. As discussed in Section 4.1, these models can be classified as differentiable, e.g., those for which a bivariate density exists, and non-differentiable models. By setting $\mathcal{D}(x) = \theta x^2 - \theta x + 1$ for $0 \leq \theta \leq 1$, the Type A bivariate extreme value distribution of two correlated Gumbel-distributed random variables, or Gumbel mixed (GM) model, is obtained (see Section 4.2.2). This differentiable model was first introduced by Gumbel (1960a,b); Gumbel and Mustafi (1967). Another differentiable model is acquired by setting $\mathcal{D}(x) = [x^m + (1-x)^m]^{1/m}$, $m \geq 1$, where parameter m describes the association between two marginals. This is the Type B distribution or Gumbel logistic (GL) model, cf., e.g. Gumbel (1960b, 1961); Gumbel and Mustafi (1967); Hougaard (1986). As discussed in Section 4.2.2, in this case, the bivariate extreme value distribution is:

$$\begin{aligned} H_G(\xi,\,\eta) \;&=\; \exp\left\{-\Big[\big(-\log F(\xi)\big)^m + \big(-\log G(\eta)\big)^m\Big]^{\frac{1}{m}}\right\} \\ &= \exp\left\{-(N-k+1)\Big[\big(\varepsilon_k^x(\xi)\big)^m + \big(\varepsilon_k^y(\eta)\big)^m\Big]^{\frac{1}{m}}\right\},\ \ \xi \geq \xi_1,\ \ \eta \geq \eta_1\,. \end{aligned} \tag{5.50}$$

Thus, the possible general functional form of the bivariate ACER surface, evidently takes the following form:

$$\mathcal{G}_k(\xi,\eta) \;=\; \Big[\big(\varepsilon_k^x(\xi)\big)^m + \big(\varepsilon_k^y(\eta)\big)^m\Big]^{\frac{1}{m}},\ \ \xi \geq \xi_1,\, \eta \geq \eta_1;\, m \geq 1. \tag{5.51}$$

The Type C distribution, also known as the biextremal model, cf. Tiago de Oliveira (1984), can be considered as an example of a non-differentiable model. In this case, the dependence function is $\mathcal{D}(x) = \max(x,\ 1-\theta x)$ for $0 \leq \theta \leq 1$, see Eq. (4.32).

In the literature, differentiable models have typically received the greatest interest, and they have been used to analyse bivariate environmental events. Yue et al. (1999); Yue (2000) applied the Gumbel mixed (GM) model to rainfall data to provide storm frequency analysis. Yue (2001b,a) also studied the Gumbel logistic (GL) model with application to flood peaks – flood volume pair of bivariate data.

In the work by Yue and Wang (2004), a comparison between the Gumbel mixed and Gumbel logistic models was performed. The authors argued that both models are appropriate and provide similar estimates of the joint distribution of two Gumbel-distributed random variables whose Pearson product moment correlation

coefficient is $0 \leq \rho \leq 2/3$. When $\rho > 2/3$, the GM model cannot be applied (Tiago de Oliveira, 1982; Yue and Wang, 2004). For these reasons, it was decided to consider the Gumbel logistic model (5.50) and (5.51) as the one that fits better the objectives of the present research.

In the dependence function $\mathcal{D}(\cdot)$ for the Gumbel logistic model, Tawn (1988) added the extra parameters ϕ and θ to obtain further flexibility. This lead to the asymmetric logistic (AL) model, which sets

$$\mathcal{D}(x) = [\phi^m x^m + \theta^m (1-x)^m]^{1/m} + (\theta - \phi)x + 1 - \theta, \tag{5.52}$$

where $0 \leq \theta \leq 1$, $0 \leq \phi \leq 1$ and $m \geq 1$. Note that when $\theta = \phi = 1$, the logistic model (4.29) is obtained. When $\theta = 1$, we get the Type C distribution (4.31). In addition, when $\theta = \phi$ we have a non-differentiable Gumbel's model, see, for example, Marshall and Olkin (1967). Thus, the asymmetric logistic is able to approximate several families of distributions.

In this case, the bivariate extreme value distribution can be expressed as:

$$\begin{aligned} H_A(\xi, \eta) \;=\; \exp\Bigg\{ -\Big[\big(-\phi \log F(\xi)\big)^m + \big(-\theta \log G(\eta)\big)^m \Big]^{\frac{1}{m}} \\ + (1-\phi)\log F(\xi) + (1-\theta)\log G(\eta) \Bigg\} . \end{aligned} \tag{5.53}$$

Using the general formula from Eq (5.49), functional form of the bivariate ACER surface in the asymmetric logistic case $\mathcal{A}_k(\xi, \eta)$ is obtained by:

$$\begin{aligned} \mathcal{A}_k(\xi, \eta) \;=& \Big[\big(\phi \varepsilon_k^x(\xi)\big)^m + \big(\theta \varepsilon_k^y(\eta)\big)^m \Big]^{\frac{1}{m}} \\ &+ (1-\phi)\varepsilon_k^x(\xi) + (1-\theta)\varepsilon_k^y(\eta)\,, \;\; \xi \geq \xi_1, \;\; \eta \geq \eta_1\,, \end{aligned} \tag{5.54}$$

where $0 \leq \theta \leq 1$, $0 \leq \phi \leq 1$ and $m \geq 1$.

As mentioned, subject to the values of parameters, the asymmetric logistic model is able to cover several other bivariate dependence structures (cf., e.g. Balakrishnan and Lai (2009)). However, it is evident that the higher number of the unknown parameters, the higher overall uncertainty of the model and its predictions are. Therefore, in the present research, it was decided to consider and analyse both logistic $\mathcal{G}_k(\xi, \eta)$ defined by Eq. (5.51) and asymmetric logistic $\mathcal{A}_k(\xi, \eta)$ defined by Eq. (5.54) bivariate ACER surfaces.

Estimation of parameters

We now proceed to the question of finding the optimal values of the parameters. First, the respective optimal parameters q, b, a and c of the univariate ACER functions given by Eq. (5.47) are estimated by the procedure described in Section 3.4.1. This procedure also automatically works up both components X and Y of the bivariate process, i.e. it generates a discrete grids $\{\xi_i\}$ and $\{\eta_j\}$ at which the ACER functions $\hat{\varepsilon}_k^x$, $\hat{\varepsilon}_k^y$ has been empirically estimated, and defines the tail markers ξ_1, η_1 with further taking into consideration only those data which are not less than

these levels. The procedure also includes rejection of the very tail of the data, where uncertainty is high (see Eq.(3.33)).

As discussed by Naess and Karpa (2015a,b), after the bivariate ACER surface $\hat{\mathcal{E}}_k(\xi,\eta)$ has been empirically estimated, it is important to introduce the following mean square error function:

$$F_{\mathcal{A}}(m,\theta,\phi) = \sum_{j=1}^{N_\eta}\sum_{i=1}^{N_\xi} w'_{ij}\left[\log\hat{\mathcal{E}}_k(\xi_i,\eta_j) - \log\mathcal{A}_k\big(\varepsilon_k^x(\xi_i),\varepsilon_k^y(\eta_j)\big)\right]^2, \tag{5.55}$$

where N_η, N_ξ are numbers of grid points $\{\xi_i\}$ and $\{\eta_j\}$, respectively, at which the bivariate ACER surface $\hat{\mathcal{E}}_k$ and corresponding univariate ACER function has been empirically estimated, and $w'_{ij} = w_{ij}/\sum\sum w_{ij}$ with

$$w_{ij} = \big(\log CI^+(\xi_i,\eta_j) - \log CI^-(\xi_i,\eta_j)\big)^{-2}, \tag{5.56}$$

denoting normalized weight factors that place more emphasis on the more reliable estimates. Here, we used the notation $\mathcal{A}_k\big(\varepsilon_k^x(\xi_i),\varepsilon_k^y(\eta_j)\big)$ to emphasis that the optimally fitted parametric curves ε_k^x and ε_k^y from Eq. (5.47) has been evaluated at points ξ_i, η_j and substituted into Eq. (5.54). Analogous mean square error function $F_{\mathcal{G}}(m)$ is introduced for the logistic model $\mathcal{G}_k(\xi,\eta)$ (5.51).

Now, we adopt the mathematical programming approach by minimizing the objective function (5.55). The optimal parameters θ^*, ϕ^* and m^* can be found as the solution of the following constrained optimization problem:

$$\begin{cases} F_{\mathcal{A}}(m,\theta,\phi) \to \min; \\ \{m,\theta,\phi\} \ \in \ \mathcal{S}, \end{cases} \tag{5.57}$$

with the constraints domain

$$\mathcal{S} = \big\{\{m,\theta,\phi\} \in \mathbb{R}^3 \ \big|\ \theta,\ \phi \ \in [0,\ 1];\ m \in [1,\ +\infty)\big\}. \tag{5.58}$$

In case of the logistic model, the problem is defined as follows:

$$\begin{cases} F_{\mathcal{G}}(m) \to \min; \\ m \geq 1. \end{cases} \tag{5.59}$$

For the optimization task, the Trust-region-reflective nonlinear least-squares optimization algorithm can be used, as well as the interior-point algorithm to find minimum of constrained nonlinear multivariable function (Matlab, 2009). These method has usually worked quite well with the chosen initial values m_0, θ_0 and ϕ_0 for the parameters. We use $\theta_0 = \phi_0 = 0.5$ as the starting value of the corresponding parameters. Then, it appears to be reasonable to use the estimate $1/\sqrt{(1-\rho)}$ as the initial value m_0. Here, ρ is the Pearson product-moment correlation coefficient between the observed data of X and Y, which has been used for the ACER analysis, that is larger than the tail markers ξ_1 and η_1, respectively.

Note on the Hausdorff distance

Scrutiny of the objective function in Eq. (5.55) reveals that it actually represents a type of distance measure between two surfaces on the log level. In this context, the problems (5.57) and (5.59) can be considered as the tasks of surface matching by minimizing the distance measure. Note that the commonly used distance measure to determine the difference between two different representations of the same 3D geometric shapes, such as point sets, curves, or surfaces, discrete or continuous, is the Hausdorff distance measure, cf., e.g. Alt and Guibas (2000); Cignoni et al. (1998).

The classical Hausdorff distance measure $d_H(A,\, B)$ (see Huttenlocher et al. (1993)) between two finite point sets $A = \{a_1, \ldots, a_p\}$ and $B = \{b_1, \ldots, b_q\}$ of sizes p and q, respectively, is defined as

$$d_H(A,\, B) \;=\; \max\Big\{ \max_{a\in A}\min_{b\in B} d(a,b),\; \max_{b\in B}\min_{a\in A} d(a,b)\Big\},$$

where $d(\cdot,\cdot)$ is some underlying norm on the points of A and B, typically, the Euclidean distance.

The term $\max_{a\in A}\min_{b\in B} d(a,b)$ finds the point $a \in A$ that is utmost from any point of B and measures the distance from this point a to its nearest neighbor in B, using the given norm $d(a,b)$. In other words, it ranges points of A in order of their distances to the nearest point of B and then uses the largest distance. So, if $\max_{a\in A}\min_{b\in B} d(a,b) = r$, then each point of A must be within distance r of some point of B, and there also is at least one point of A that is exactly at the distance r from the nearest point of B (the most mismatched point of A). Thus intuitively, the expression 'the Hausdorff distance equals r', that is $d_H(A,\, B) = r$, means that every point of A is within a distance r of some point of B and vice versa.

This definition concerns the point sets, but as long as curves, images and surfaces can be interpreted as point sets, it can be generalized to these objects (Sim et al., 1999). For example, because any 3-D surface is specified by the matrix of its values, the number of columns can be considered as the dimension of the associated Euclidean space, whereas the number of rows is treated then as the number of observations.

From the aforementioned, it naturally emerges that in our case, e.g. for the asymmetric logistic model Eq. (5.54), the optimal parameters $\{m^*,\, \theta^*,\, phi^*\} \in \mathcal{S}$ are those, that minimize the following modified objective function $\tilde{F}_{\mathcal{A}}(m,\theta,\phi)$:

$$\begin{aligned}\tilde{F}_{\mathcal{A}}(m,\theta,\phi) \;=\; d_H\Big(&\log\hat{\mathcal{E}}_k(\xi_i,\eta_j),\, \log\mathcal{A}_k\big(\varepsilon_k^x(\xi_i),\varepsilon_k^y(\eta_j)\big)\Big),\\ &i=1,\ldots,N_\xi,\; j=1,\ldots,N_\eta\end{aligned} \tag{5.60}$$

In the present research, the authors attempted to apply the target function Eg. (5.60) in the bivariate ACER analysis. The Matlab (2009) routine HausdorffDist(P,Q,lmf,dv) developed by Danziger (2009) was used for the task. The results revealed that the optimization based on the function (5.60), generally speaking, does not improves the fit. Note that the adopted routine estimates the Hausdorff distance based on the Euclidean norm. Therefore, the function $\tilde{F}_{\mathcal{A}}$

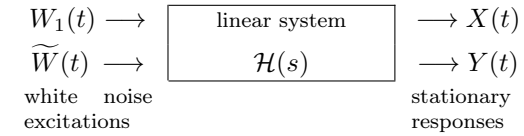
actually differs from the $F_{\mathcal{A}}$ defined in (5.55) by the lower number of summands, but above all, the main and significant disadvantage of the objective $\tilde{F}_{\mathcal{A}}$ is that it leaves out of account the weight factors $W' = \{w'_{ij}\}$, which appear to be significant in the ACER analysis. Nevertheless, the Hausdorff distance is considered an important issue in tackling the problem of surface matching. Thus, it is assumed that the possible improvements, e.g. rational including of the weight factors in the objective function $\tilde{F}_{\mathcal{A}}(m, \theta, \phi)$, will be examined in further research.

5.5 Analysis of the Simulated Data by the Bivariate ACER Method

In this section, the bivariate ACER method will be analysed by application to synthetic data. First, we will study the performance of the bivariate ACER method applied to the bivariate narrow-banded process, being the response of the linear system to the Gaussian white noise excitation. Then, the case for which the exact extreme values can be calculated will be considered.

5.5.1 Response of the linear system to the white noise excitation

Let us consider two stationary processes $X(t)$ and $Y(t)$, being the responses of a single degree of freedom linear system to the stationary Gaussian white noise excitations. The $X(t)$ component is assumed to be the response of the system to the white noise excitation $W_1(t)$, whereas the stationary process $Y(t)$ is assumed to be the response of the same linear system to the input signal $\widetilde{W}(t)$, as shown schematically in the following diagram:



The corresponding equations of motion are given by

$$\begin{aligned} \ddot{X} + 2\zeta_0\omega_0\dot{X} + \omega_0^2 X &= W_1(t), \\ \ddot{Y} + 2\zeta_0\omega_0\dot{Y} + \omega_0^2 Y &= \widetilde{W}(t), \end{aligned} \tag{5.61}$$

with specific damping constant ζ_0 and resonance frequency $\omega_0 = 2\pi/T_0$. The white noise excitation $\widetilde{W}(t)$ is given by

$$\widetilde{W}(t) = \left(\frac{\alpha}{\sqrt{\alpha^2 + (1-\alpha)^2}} W_1(t) + \frac{(1-\alpha)}{\sqrt{\alpha^2 + (1-\alpha)^2}} W_2(t) \right), \; 0 \leq \alpha \leq 1 \tag{5.62}$$

where $W_1(t)$ and $W_2(t)$ are independent stationary Gaussian white noise excitations with the autocorrelation function $\mathbb{E}\big[W_1(t)W_1(t+\tau)\big] = \mathbb{E}\big[W_2(t)W_2(t+\tau)\big] = \delta(\tau)$. The way that the excitation $\widetilde{W}(t)$ is defined implies that it is also a stationary

Gaussian white noise with the autocorrelation function $R_{\widetilde{W}}(\tau) = \mathbb{E}\big[\widetilde{W}(t)\widetilde{W}(t+\tau)\big] = \delta(\tau)$. This implies that the power spectral densities $S_{W_1}(\omega)$ and $S_{\widetilde{W}}(\omega)$ of both excitations are equal and take value $1/(2\pi)$, that is $S_{W_1}(\omega) = S_{\widetilde{W}}(\omega) = 1/(2\pi) := S(\omega)$ (cf., e.g. Newland (1993)).

In this case, the two-sided power spectral densities (PSD) of both processes $X(t)$ and $Y(t)$ are equal and given as:

$$S_X(\omega) = S_Y(\omega) = |H(\omega)|^2 S(\omega) = \frac{1}{2\pi\big((\omega_0^2-\omega^2)^2 + (2\,\zeta_0\,\omega_0\,\omega)^2\big)}\,, \tag{5.63}$$

where the frequency response function of the linear system $H(\omega)$ is:

$$H(\omega) = \frac{1}{(\omega_0^2-\omega^2) + \mathrm{i}\,(2\,\zeta_0\,\omega_0\,\omega)}\,. \tag{5.64}$$

In this way, because the excitation processes $W_1(t)$ and $\widetilde{W}(t)$ are dependent, so are the corresponding response processes $X(t)$ and $Y(t)$ (Naess and Moan, 2012). The response processes are also Gaussian with zero mean, due to the properties of a linear oscillator (see, e.g. Ochi (1990)). Then, the correlation coefficient $\rho_{XY} = \mathbb{E}[X(t)Y(t)]/\sigma_X\sigma_Y$ between the dependent processes $X(t)$ and $Y(t)$ can be obtained as:

$$\rho_{XY} = \frac{R_{XY}(0)}{\sigma_X\sigma_Y}\,, \tag{5.65}$$

where $R_{XY}(0)$ is the cross-correlation function $R_{XY}(\tau)$ between processes $X(t)$ and $Y(t)$ evaluated at point $\tau = 0$. From the spectral analysis it is known that

$$R_{XY}(\tau)|_{\tau=0} = \int_{-\infty}^{\infty} S_{XY}(\omega)e^{\mathrm{i}\omega\tau}d\omega\Big|_{\tau=0} = \int_{-\infty}^{\infty} S_{XY}(\omega)d\omega\,. \tag{5.66}$$

The cross-spectrum $S_{XY}(\omega)$ of two jointly stationary processes $X(t)$ and $Y(t)$ is found as

$$S_{XY}(\omega) = H(\omega)^*\,H(\omega)\,S_{W_1\widetilde{W}} = |H(\omega)|^2\;S_{W_1\widetilde{W}}\,, \tag{5.67}$$

where $H(\omega)^*$ is complex conjugate of the frequency response function $H(\omega)$ (Naess and Moan, 2012). The cross-spectrum $S_{W_1\widetilde{W}}$ of the excitations $W_1(t)$ and $\widetilde{W}(t)$ can easily be found as

$$S_{W_1\widetilde{W}} = \frac{\alpha}{2\pi\sqrt{\alpha^2+(1-\alpha)^2}} \tag{5.68}$$

The variances σ_X^2 and σ_Y^2 of both processes are equal and can be expressed in terms of the power spectral densities $S_X(\omega) = S_Y(\omega)$ (see Eq. (5.63)) as follows,

$$\sigma_X^2 = \sigma_Y^2 = \int_{-\infty}^{\infty} |H(\omega)|^2\,S(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\omega)|^2 d\omega\,. \tag{5.69}$$

Thus, by Eq. (5.63) and Eqs. (5.65) - (5.69), it is found that

$$\rho_{XY} = \frac{\alpha}{2\pi\sqrt{\alpha^2+(1-\alpha)^2}\,\sigma_X\sigma_Y}\int_{-\infty}^{\infty}|H(\omega)|^2\,d\omega = \frac{\alpha}{\sqrt{\alpha^2+(1-\alpha)^2}}. \quad (5.70)$$

Therefore, two dependent response processes $X(t)$ and $Y(t)$ compose a bivariate stochastic process $Z(t) = \big(X(t), Y(t)\big)$. By changing the values of α, it is possible to control the correlation coefficient ρ between the components, as illustrated in Figure (5.1). For the Monte Carlo simulations, which will be described hereinafter,

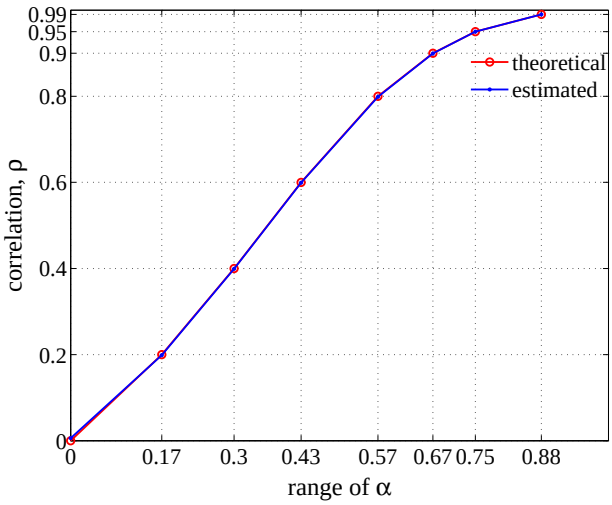
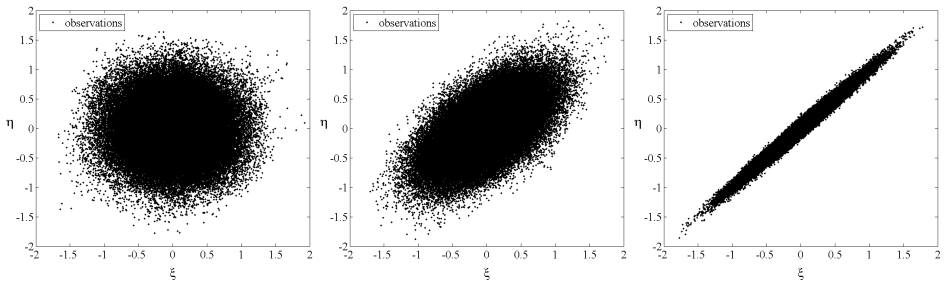


Figure 5.1: Correlation coefficient ρ between X and Y for different values of α. (red) – theoretical values; (blue) – sample estimates.

it was decided to consider values of α that result in the following values of the correlation coefficient $\rho = \{0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.99\}$. The limiting value $\alpha = 0$ yields $\rho = 0$, that is, the Gaussian response processes are uncorrelated. Another limit case $\alpha = 1$ implies that $\rho = 0$, and hence the time series $X(t)$ and $Y(t)$ are fully correlated and, in fact, are equal with probability one. In Figure (5.2), 10^5 bivariate observations of the process $Z(t)$ ($X(t)$ component on the x-axis versus $Y(t)$ on the y-axis) are plotted for three cases of values of parameter ρ to illustrate the corresponding correlation effect.

It was also decided to consider the case of fixed $\zeta_0 = 0.05$ and $\omega_0 = 3$. By this, the spectral densities of both processes $S_X(\omega) = S_Y(\omega)$ are concentrated in a narrow frequency band, see Figure (5.3). This implies that the peak events of the time series $X(t)$ and $Y(t)$ have a tendency to occur in groups, that is, the response processes are narrow banded (see Section 3.5.6). Figure (5.4) illustrates parts of the generated response processes and their peak values. First, it is observed that both time series exhibit narrow band characteristics and that the neighbouring data points of each process are strongly dependent. In addition, the correlation effect between the components $X(t)$ and $Y(t)$ is also displayed. Apparently, the



(a) $\alpha = 0$, $\rho = 0$ (b) $\alpha = 0.43$, $\rho = 0.6$ (c) $\alpha = 0.88$, $\rho = 0.99$

Figure 5.2: Scatterplot of the simulated 10^5 bivariate response events for different values of α (and ρ).

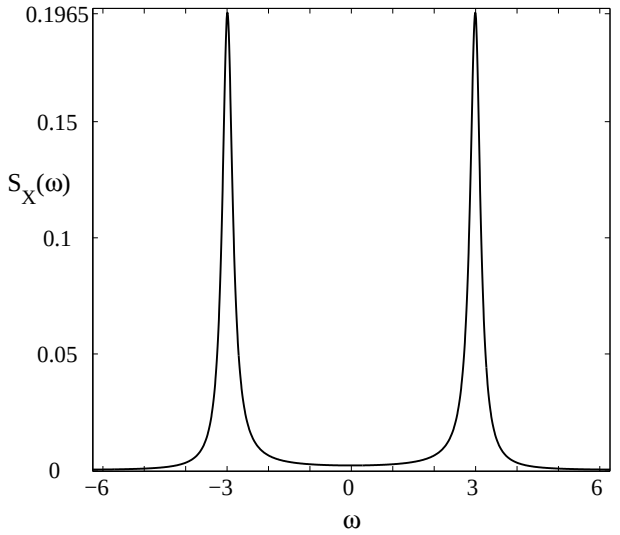


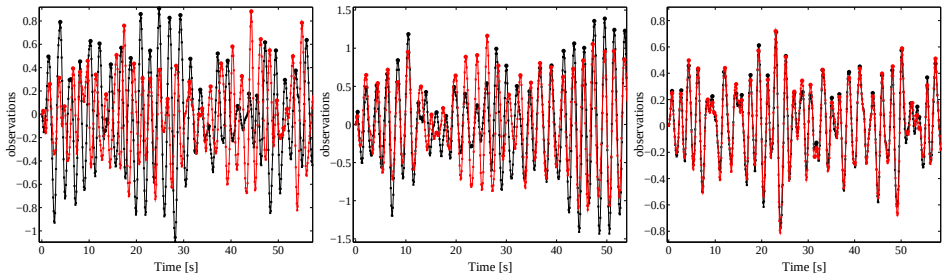
Figure 5.3: Plot of the two-sided narrow power spectral density with $\zeta_0 = 0.05$ and $\omega_0 = 3$ (see Eq. (5.63)).

peaks of both time series occur substantially independently when the ρ value is low, cf. Figure (5.4a), whereas the peak events occur on full concordance for the high values of ρ, see Figure (5.4c).

In summary, the aforementioned model allows us to generate a bivariate time series that possesses both spatial and temporal dependence characteristics, and therefore, is of particular interest for the bivariate ACER analysis.

Behaviour of the bivariate ACER surface

It appears to be important to first consider the question of typical behaviour of the bivariate ACER surface $\hat{\mathcal{E}}_k(\xi, \eta)$.



(a) $\alpha = 0$, $\rho = 0$ (b) $\alpha = 0.43$, $\rho = 0.6$ (c) $\alpha = 0.88$, $\rho = 0.99$

Figure 5.4: Part of the narrow-band response time series $X(t)$ (black) and $Y(t)$ (red) with indicated peak values, for different values of α (and ρ).

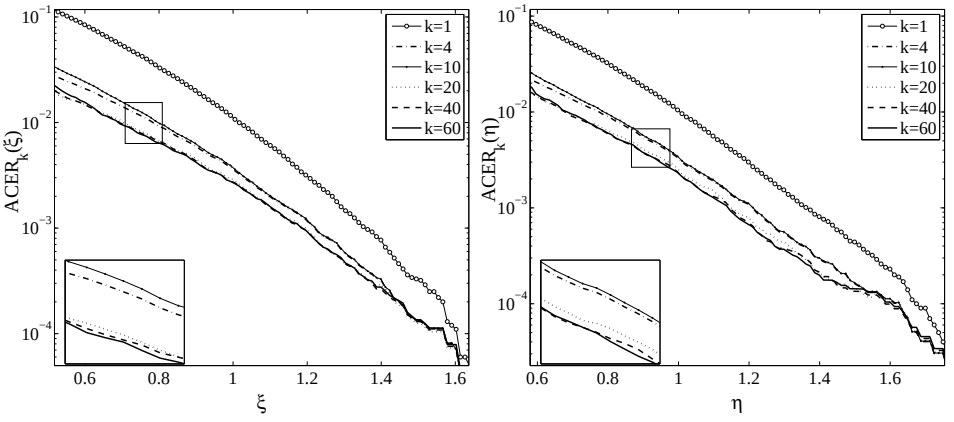
Let us consider 100 blocks of the bivariate response process $Z(t) = \big(X(t), Y(t)\big)$ with 1000 bivariate observations in one block, so that the total number of observations is 10^5 bivariate data points. This amount of data is sufficient to answer a need of a fairly accurate estimation of the empirical bivariate ACER surface $\hat{\mathcal{E}}_k(\xi, \eta)$. The Matlab (2009) Simulink model was build and used to generate two response processes $X(t)$ and $Y(t)$ for each of eight values of the correlation coefficient ρ.

First, the univariate average conditional exceedance rates $\hat{\varepsilon}_k^x(\xi)$ and $\hat{\varepsilon}_k^y(\eta)$ were estimated for different degrees of conditioning k for each value of ρ. Because the spectral densities of the response processes are equal and do not depend on the parameter α, the dependence structure, inherent in the data, remains identical for both time series for all values of ρ. Therefore, the estimated univariate ACER functions $\hat{\varepsilon}_k$ with the chosen range of k exhibit the same behavioral pattern for every response process and for every ρ value. This is confirmed by Figure (5.5). It shows the ACER functions $\hat{\varepsilon}_k^x(\xi)$ and $\hat{\varepsilon}_k^y(\eta)$ plotted for $k = 1; 4; 10; 20; 40; 60$. Without loss of generality, we chose the case $\alpha = 0.43$, $\rho = 0.6$.

Figure (5.5) reveals that there is a significant dependence effect between consecutive data. It is reflected in the fact that there is a noticeable difference between the three groups of the ACER functions. The first curve $\hat{\varepsilon}_1$ is based on the assumption that all of the sampled data are independent. It is noticeably bigger than the rest of the ACER functions and cannot be used for the further analysis, because the assumption is incorrect. The ACER function $\hat{\varepsilon}_4$ can hardly be distinguished from the $\hat{\varepsilon}_{10}$, though none of them can fully capture the dependence effect in the data. Finally, it is observed that for $k \geq 20$, where the $\hat{\varepsilon}_k$ with $k = 20$ takes into account the neighbouring peak values, the full convergence has been achieved.

This demonstrates that the dependence structure is entirely accounted for by the ACER line $\hat{\varepsilon}_k$, $k = 20$. For this reason, it was decided to focus our further attention on the ACER functions with index $k = 20$.

It should be emphasized here that for the ACER analysis of the bivariate extremes, the chosen time series has to consists of all sample points. That is, the peak values should not be extracted from the data array for the analysis, in contrast to



(a) plot for $X(t)$ (b) plot for $Y(t)$

Figure 5.5: Comparison between univariate ACER estimates for different degrees of conditioning; $\rho = 0.6$.

the univariate ACER analysis, especially of the narrow band time series. The main valid reason for this is that the cascade of bivariate conditioning approximations (see Section 5.1) is based on the bivariate sampled data allocated to the same time. In other words, observations of each component have to occur (or to be sampled) simultaneously. Evidently, this framework can be violated by extracting the peak values.

Now, the empirical bivariate ACER functions can be estimated for the same degrees of conditioning $k = 1$; 4; 10; 20; 40; 60. The cascade of $\hat{\mathcal{E}}_k(\xi, \eta)$ surfaces for the case $\rho = 0.6$ is shown in Figure (5.6). $\hat{\mathcal{E}}_k(\xi, \eta)$ with $k = 1$ is the upper most, then the following group of $\hat{\mathcal{E}}_4$ and $\hat{\mathcal{E}}_{10}$ that match on the entire $\xi - \eta$ domain. Finally, the group of surfaces with $k \geq 2$ that match entirely, is the lower most. The same arguments as in the univariate case are applied to make the decision about the bivariate ACER surface to be used in the analyses. That is, there is virtually no difference between the 20th and the 60th ACER surfaces, which shows that the dependence structure in this particular time series is captured almost completely by conditioning on previous 19 sampled values. As long as estimation of the surface $\hat{\mathcal{E}}_{20}$ is more accurate due to availability of more data, it is reasonable to choose the surface with the degree of conditioning $k = 20$.

In addition, as it is seen from the figure, the cross section of the surfaces at the high level of η gives the univariate ACER functions of the $X(t)$ observations, while the cross section at a high ξ level represents the empirical univariate ACER function of the time series $Y(t)$, respectively.

The aforementioned statements concern solely the internal coherence of the sampled data of each component of the bivariate process, that is the temporal dependence of the processes. Apparently, correlation between two components of the time series, or spatial dependence, has a substantial impact on the bivariate ACER surface $\hat{\mathcal{E}}_k(\xi, \eta)$.

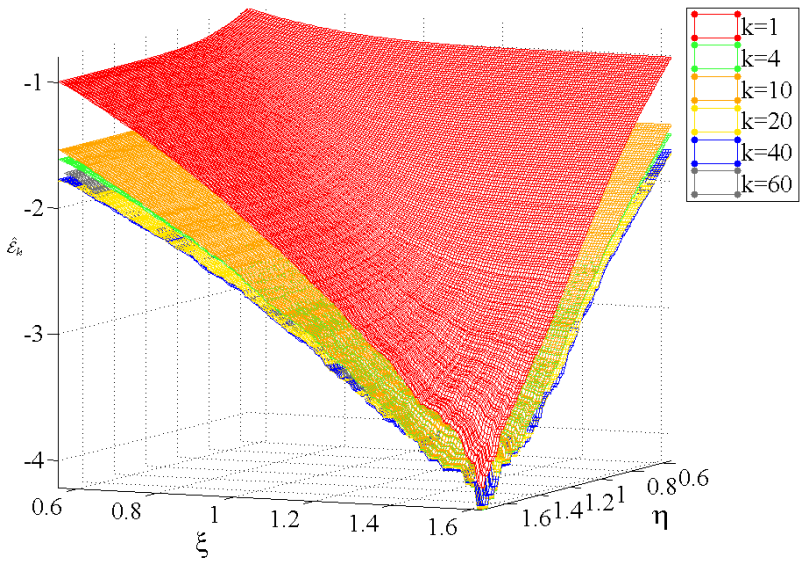


Figure 5.6: Comparison between bivariate ACER surface estimates for different degrees of conditioning. $\hat{\mathcal{E}}_k(\xi, \eta)$ surfaces are plotted on a $\log_{10}$ scale; $\rho = 0.6$.

As discussed in Section 5.4, the joint extreme value distribution $H(\xi,\, \eta)$ of the bivariate extreme value vector $\hat{Z}_N = \big(\hat{X}_N, \hat{Y}_N\big)$ with marginal extreme value distributions $F(\xi)$ and $G(\eta)$, respectively, is given by

$$H(\xi,\, \eta) = C_{\mathrm{E}}\big(F(\xi),\; G(\eta)\big)\,. \tag{5.71}$$

where the extreme value copula C_{E} is defined as

$$C_{\mathrm{E}}\,(u,\, v) = \exp\big\{\log(u\,v)\mathcal{D}\big(\log(u)/\log(u\,v)\big)\big\}$$

for $0 < u < 1$, $0 < v < 1$. It comprises information about the inter-component dependencies. In fact, the dependence structure is given by the Pickands function $\mathcal{D}(x) : [0, 1] \longmapsto [\,\max(x,\, 1 - x),\;\, 1]$ (Pickands, 1981). Two boundary values of the function $\mathcal{D}(\cdot)$ lead to two trivial cases. For statistical independent components $X(t)$ and $Y(t)$, the dependence function $\mathcal{D} = 1$, and therefore

$$H(\xi,\, \eta) = F(\xi)\, G(\eta)\,. \tag{5.72}$$

On the contrary, in case of complete dependence of $X(t)$ and $Y(t)$, the dependence function becomes $\mathcal{D}(x) = \max(x,\, 1 - x)$. This leads to the following result (cf., e.g. Nelsen (2006))

$$H(\xi,\, \eta) = \min\big\{F(\xi),\, G(\eta)\big\}\,. \tag{5.73}$$

On the other hand, it has been discovered before in Section 5.2, Eq. (5.25), that the joint extreme value distribution $H(\xi,\, \eta)$ can be presented through the bivariate ACER function $\mathcal{E}_k(\xi, \eta)$, as $H(\xi, \eta) = \exp\{-(N - k + 1)\,\mathcal{E}_k(\xi, \eta)\}$. It is also

assumed that the marginal EVD $F(\xi)$ and $G(\eta)$ are presented through the corresponding univariate ACER functions, as shown in Eqs. (5.46) and (5.47). Now, we can substitute the right-hand side of these expressions into Eqs. (5.72) and (5.73). This yields to the following results. If Eq. (5.72) holds, that is, the components $X(t)$ and $Y(t)$ are independent, then evidently,

$$\mathcal{E}_k(\xi,\eta) = \varepsilon_k^x(\xi) + \varepsilon_k^y(\eta). \tag{5.74}$$

Oppositely, for the fully dependent processes $X(t)$ and $Y(t)$, from Eq. (5.73) follows:

$$\begin{aligned} &\exp\big\{-(N-k+1)\,\mathcal{E}_k(\xi,\eta)\big\} \\ &= \min\big\{\exp\big\{-(N-k+1)\,\varepsilon_k^x(\xi)\big\},\, \exp\big\{-(N-k+1)\,\varepsilon_k^y(\eta)\big\}\big\}\,. \end{aligned} \tag{5.75}$$

And hence,

$$\begin{aligned} &-(N-k+1)\,\mathcal{E}_k(\xi,\eta) \\ &= \min\big\{-(N-k+1)\,\varepsilon_k^x(\xi),\, -(N-k+1)\,\varepsilon_k^y(\eta)\big\}\,, \end{aligned} \tag{5.76}$$

from which it finally emerges that

$$\mathcal{E}_k(\xi,\eta) = \max\big\{\varepsilon_k^x(\xi),\, \varepsilon_k^y(\eta)\big\}. \tag{5.77}$$

Therefore, behaviour of the bivariate ACER function $\mathcal{E}_k(\xi,\eta)$ varies from $\big(\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)\big)$ to $\max\big\{\varepsilon_k^x(\xi),\, \varepsilon_k^y(\eta)\big\}$ with growth of the dependence between two processes. To verify the aforementioned, we use the estimated ACER functions of the response processes $X(t)$ and $Y(t)$. The bivariate ACER surface $\hat{\mathcal{E}}_{20}(\xi,\eta)$ and the surfaces composed of the univariate ACER functions as given by Eqs. (5.74) and (5.77), are considered.

For the case $\rho = 0$, that is, for the uncorrelated components, the surfaces $\hat{\mathcal{E}}_{20}(\xi,\eta)$ and $\big(\hat{\varepsilon}_{20}^x(\xi) + \hat{\varepsilon}_{20}^y(\eta)\big)$ match entirely along the domain, within the inherent statistical uncertainty. This, in addition, indicates that the components are independent and hence, they are jointly normally distributed. Figure (5.7) shows the contour plot of the $\hat{\mathcal{E}}_{20}(\xi,\eta)$ together with $\big(\hat{\varepsilon}_{20}^x(\xi) + \hat{\varepsilon}_{20}^y(\eta)\big)$ on a $\log_{10}$ scale.

The bivariate ACER surfaces $\hat{\mathcal{E}}_{20}(\xi,\eta)$ and $\max\big\{\hat{\varepsilon}_{20}^x(\xi),\, \hat{\varepsilon}_{20}^y(\eta)\big\}$ estimated from the fully correlated observations ($\rho = 1$) also coincide entirely. In Figure (5.8), contour lines of the decimal logarithm of both surfaces are plotted together.

As a final point, the contour lines of the $\hat{\mathcal{E}}_{20}(\xi,\eta)$ surface for different values of ρ were considered. Contour lines were estimated for the same fixed range of levels of the bivariate ACER function: $\log_{10}\hat{\mathcal{E}}_{20} = \{-1.5, -1.7, -1.9, -2.1, -2.4, -2.6, -2.9, -3.2, -3.6\}$ and plotted. Figure (5.9) enables to track the change in the behaviour of the $\hat{\mathcal{E}}_{20}(\xi,\eta)$ surface as the correlation coefficient ρ increases from 0 to 1. It is observed that for low-correlated data ($\rho \leq 0.5$), the ACER surfaces behave rather identically, whereas for higher values of ρ, the contour lines tend to sharpen on the diagonal, that is, the joint distribution converges to the form $\hat{\mathcal{E}}_{20}(\xi,\eta) = \max\big\{\hat{\varepsilon}_{20}^x(\xi),\, \hat{\varepsilon}_{20}^y(\eta)\big\}$ as $\rho \to 1$.

The Monte Carlo simulation

In this part, the overall performance of the bivariate ACER method will be studied by application to repeated random samples of the generated response processes

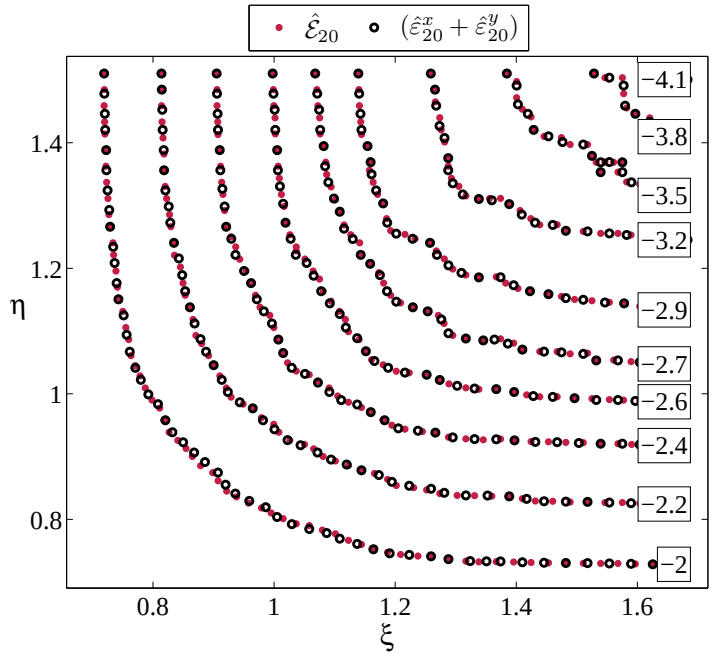


Figure 5.7: Contour plot of the ACER surface $\hat{\mathcal{E}}_{20}(\xi, \eta)$, (• •), and the surface $\big(\hat{\varepsilon}^x_{20}(\xi) + \hat{\varepsilon}^y_{20}(\eta)\big)$, (∘ ∘). Surfaces are estimated empirically for the case $\rho = 0$. Boxes indicate levels on a $\log_{10}$ scale.

of the linear system to the white noise excitations. The bivariate extreme value statistics will be based on 100 independent the Monte Carlo simulations for every value of the parameter α, that is, for eight considered cases of the correlation ρ between the components.

As before, let us consider 100 blocks of the bivariate response process $Z(t) = \big(X(t), Y(t)\big)$ with 1000 bivariate observations in one block, that is, in total, there are 10^5 bivariate data available. This amount is assumed to be satisfactory for accurate empirical estimation of the bivariate ACER function $\hat{\mathcal{E}}_k$. As discussed above, the dependence structure in this particular time series is entirely accounted for by the ACER surface estimated by conditioning on not less than 19 previous sampled values, that is, for $k = 20$. $\hat{\mathcal{E}}_{20}(\xi, \eta)$ corresponds to reaching the previous peak values of the narrow-banded processes $X(t)$ and $Y(t)$ in the conditioning process. Therefore, it appears to be efficient to perform the extreme value analysis based on this surface. In addition, the ACER surface $\hat{\mathcal{E}}_{20}$ practically coincides with the one based on the exact bivariate extreme value distribution that is inherent in the data. This makes it possible to use the accurately estimated $\hat{\mathcal{E}}_{20}$ as a benchmark for comparison with the fitted asymmetric logistic model $\mathcal{A}_k(\xi, \eta)$ from (5.54) and the fitted logistic model $\mathcal{G}_k(\xi, \eta)$ from (5.51).

Also, for comparison purposes, it was decided to consider the common approach of assuming Gumbel marginal extreme value distributions combined with a logistic

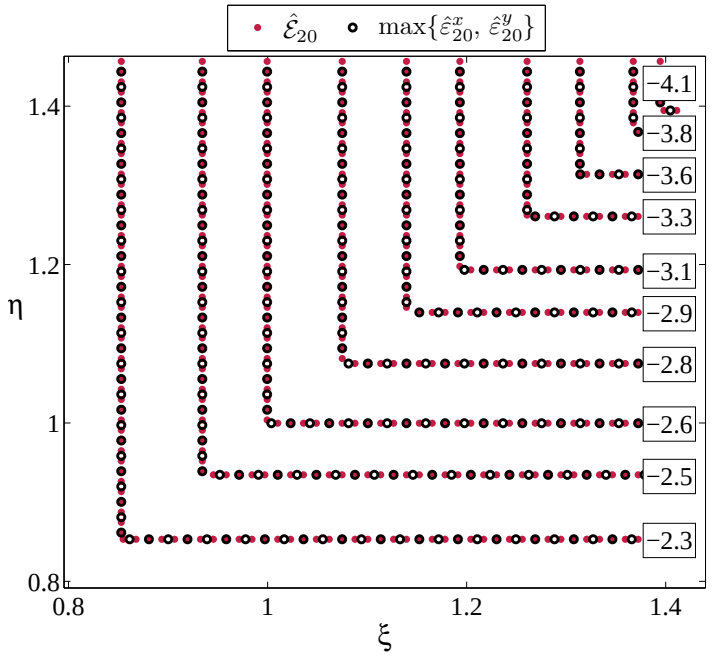


Figure 5.8: Contour plot of the ACER surface $\hat{\mathcal{E}}_{20}(\xi, \eta)$, (• •), and the surface $\max\left\{\hat{\varepsilon}^x_{20}(\xi),\, \hat{\varepsilon}^y_{20}(\eta)\right\}$, (∘ ∘). Surfaces are estimated empirically for perfectly correlated data, ($\rho = 1$). Boxes indicate levels on a $\log_{10}$ scale.

copula, that is, the Type B bivariate EVD:

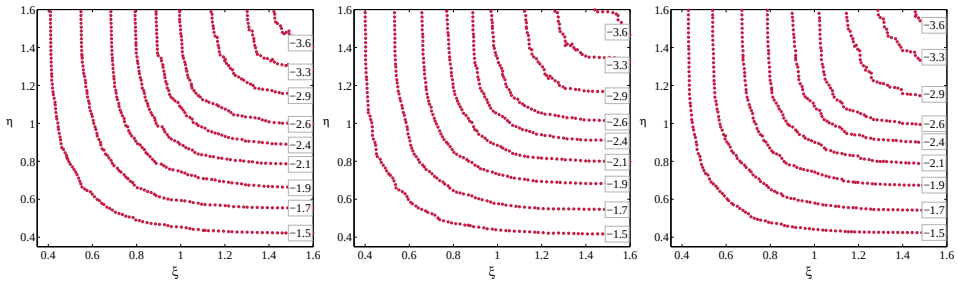
$$H_G(\xi,\, \eta) \;=\; \exp\left\{-\Big[\big(-\log F(\xi)\big)^m + \big(-\log G(\eta)\big)^m\Big]^{1/m}\right\}, \tag{5.78}$$

where

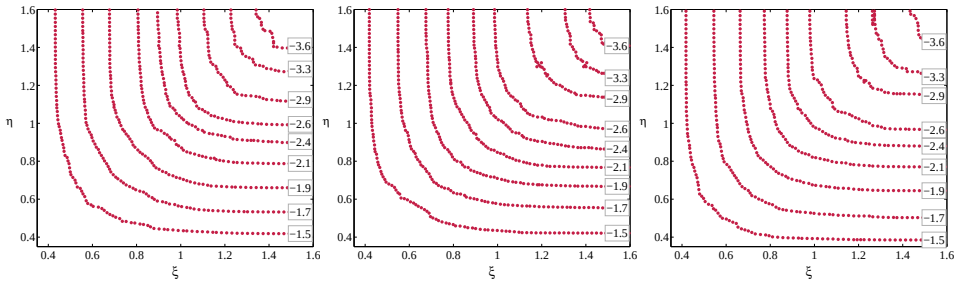
$$\begin{aligned} F(\xi) \;&=\; \exp\left\{-\exp\left(-(\xi-\mu_x)/\sigma_x\right)\right\}, \\ G(\eta) \;&=\; \exp\left\{-\exp\left(-(\eta-\mu_y)/\sigma_y\right)\right\} \end{aligned} \tag{5.79}$$

are the Gumbel marginal distributions of the extreme values $\hat{X}_N$ and $\hat{Y}_N$. As discussed by Gumbel and Mustafi (1967); Yue (2001a), the estimator of the parameter m is given by $\hat{m} \;=\; 1/\sqrt{(1-\hat{\rho})}$, where $\hat{\rho}$ is the estimated product-moment correlation coefficient between $\hat{X}_N$ and $\hat{Y}_N$. Thus, if the sample size of $\hat{Z}_N = \big(\hat{X}_N, \hat{Y}_N\big)$ is sufficiently large, the parameters μ_x, μ_y, σ_x, σ_y and m can by estimated fairly accurately.

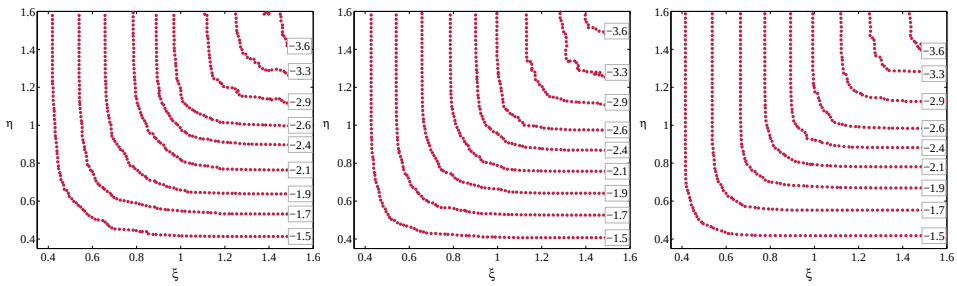
To obtain an idea about the performance of the bivariate ACER method and, in addition, to highlight the results obtained by applying the Type B bivariate extreme value distribution, 100 independent 100-block Monte Carlo simulations were conducted as follows. Based on the generated 10^5 bivariate response data, the univariate ACER functions $\hat{\varepsilon}^x_{20}(\xi)$ and $\hat{\varepsilon}^y_{20}(\eta)$, as well as the bivariate ACER surface



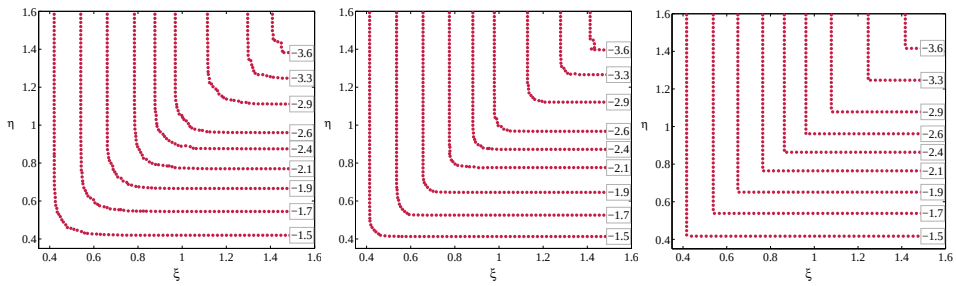
(a) $\rho = 0$ (b) $\rho = 0.2$ (c) $\rho = 0.3$



(d) $\rho = 0.4$ (e) $\rho = 0.5$ (f) $\rho = 0.6$



(g) $\rho = 0.7$ (h) $\rho = 0.8$ (i) $\rho = 0.9$



(j) $\rho = 0.95$ (k) $\rho = 0.99$ (l) $\rho = 1$

Figure 5.9: Contour plot of the empirically estimated ACER surface $\hat{\mathcal{E}}_{20}(\xi, \eta)$, (●●), for increasing values of ρ. Boxes indicate the same level lines of the $\hat{\mathcal{E}}_{20}$ on a $\log_{10}$ scale.

$\hat{\mathcal{E}}_{20}(\xi, \eta)$ were estimated. Because the response time series $X(t)$ and $Y(t)$ have the same spectral density and, therefore, identical magnitudes, it was decided to set the tail markers at the same level, that is, $\xi_1 = \eta_1 = 0.7$. Then, the corresponding optimization procedures (see Section 3.4.1) were applied to find the optimal values of the parameters of the univariate ACER functions ε^x_{20} and ε^y_{20}. Followed by that, the asymmetric logistic $\mathcal{A}_{20}(\xi, \eta)$ and the logistic $\mathcal{G}_{20}(\xi, \eta)$ models with the optimized univariate ACER marginals were fitted to the empirically estimated $\hat{\mathcal{E}}_{20}$, using the techniques defined by Eqs. (5.57) and (5.59), respectively. Finally, based on the parametric surfaces $\mathcal{A}_{20}$ and $\mathcal{G}_{20}$, the T-year return period levels were obtained as the solution of the Eq. (5.43).

The bivariate return levels predicted by the Gumbel logistic model with Gumbel marginals was based on the sample of 100 block extremes. Five estimators $\hat{\mu}_x$, $\hat{\mu}_y$, $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $\hat{m}$ of the parameters of the model were constructed using the method of moments. This method was chosen as one of the most commonly used robust methods of estimation. In addition, it was agreed that the method of moments provides rather accurate estimates based on the 100 sample points. Subsequently, the Gumbel logistic model with Gumbel marginals estimated by the method of moments will be referred to as $\mathcal{G}_{MM}(\xi, \eta)$. The T-year return period levels were obtained as the solution of the equation

$$\mathcal{G}_{MM}(\xi^{\mathrm{Tyr}}, \eta^{\mathrm{Tyr}}) = 1 - 1/T\,. \tag{5.80}$$

Obviously, 100 Monte Carlo empirically estimated bivariate ACER functions, as well as the corresponding fitted and estimated surfaces cannot provide any distinct conclusions regarding the method. Therefore, it appears to be natural to estimate the 100 average surfaces of each model. Any 3D surface is given by a matrix of Z-values calculated at each (x, y) point, together with the corresponding grid on the X-Y plane. In our case, the surfaces are designated by the matrices of values of the ACER function (empirically estimated in case of $\hat{\mathcal{E}}_{20}$, or calculated otherwise). Number of columns of the matrix is equal to N_ξ and corresponds to the number of points of division $\{\xi_i\}$ of the ξ-axis (corresponds to the X-axis), whereas number of rows is equal to N_η, that is, to the number of points $\{\eta_j\}$ on the axis η ($\equiv Y$), cf. Eq. (5.55). Note that the ξ-η grid was created prior to the Monte Carlo simulations to ensure concordance of the surfaces. Therefore, the 100 average of each surface can be estimated as the mean values for elements along the third dimension of the three-dimensional array formed by 100 matrices. In Figure (5.10), this process is illustrated graphically via the contour plot of the T-year return period levels. It is observed that the black solid lines correspond to the contours the T-year return levels of each of the fitted MC $\mathcal{A}_{20}$ surfaces, whereas the red line corresponds to the contour of the surface $\bar{\mathcal{A}}_{20}$ obtained by the averaging of the 100 fitted surfaces $\mathcal{A}_{20}$.

In this manner, the average of the 100 Monte Carlo surfaces were estimated for each of eight considered values of the correlation coefficient ρ: average of the 100 empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_{20}$, the averages of the optimized asymmetric logistic and optimized logistic models, $\bar{\mathcal{A}}_{20}$ and $\bar{\mathcal{G}}_{20}$, respectively, as well as the average of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$. Figures (5.11) and (5.12) present contour plots of the 20-year, 50-year and 100-year

return period levels for the averages of the 100 Monte Carlo surfaces, for all values of ρ (and α).

The figures reveal that the estimated average of 100 MC empirical bivariate ACER surfaces $\bar{\hat{\mathcal{E}}}_{20}(\xi, \eta)$ exhibits a behaviour analogous to the one demonstrated in Figure (5.9). That is, for the low-correlated data ($\rho \leq 0.6$) the contour lines behave rather identically and have smooth curvature on the diagonal. At the same time, it is observed that the $\bar{\hat{\mathcal{E}}}_{20}$ surface captures high correlation between the data, which is indicated by the contour lines that tend to sharpen on the diagonal as ρ increases.

Figures (5.11) and (5.12) also demonstrate equally high degree of concordance between the average empirical ACER surface $\bar{\hat{\mathcal{E}}}_{20}$ and the averages of the optimized asymmetric logistic and logistic surfaces $\bar{\mathcal{A}}_{20}$ and $\bar{\mathcal{G}}_{20}$, respectively. In addition, the level of agreement between the contour lines of $\bar{\mathcal{A}}_{20}$ and $\bar{\mathcal{G}}_{20}$ allows to infer that these surfaces actually coincide entirely. Finally, regarding the average of 100 MC estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}(\xi, \eta)$, it is clearly observed that the level of agreement between the contours of $\bar{\hat{\mathcal{E}}}_{20}$ and $\bar{\mathcal{G}}_{MM}$ is high only for short return periods, that is for $T < 50$. For $T > 50$ the distinction is rather significant.

Apparently, the visual images, like any other qualitative index, have, to a certain extent, low capacity to enclose the complete information about mutual alignment of the contour lines. Therefore, it appears to be appropriate to introduce a quan-

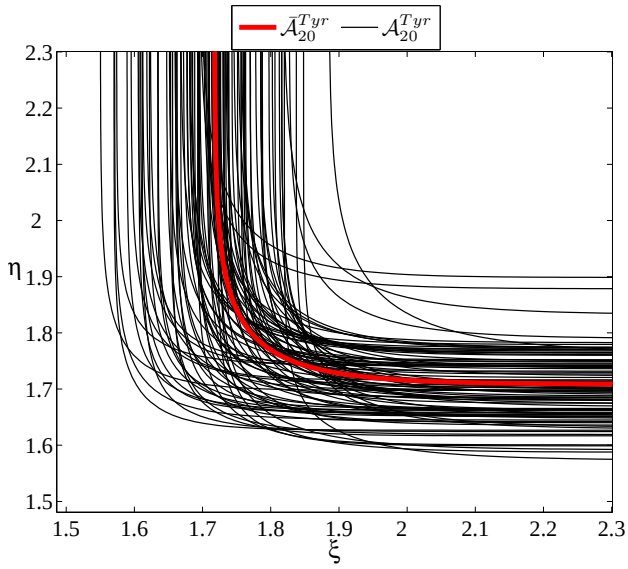
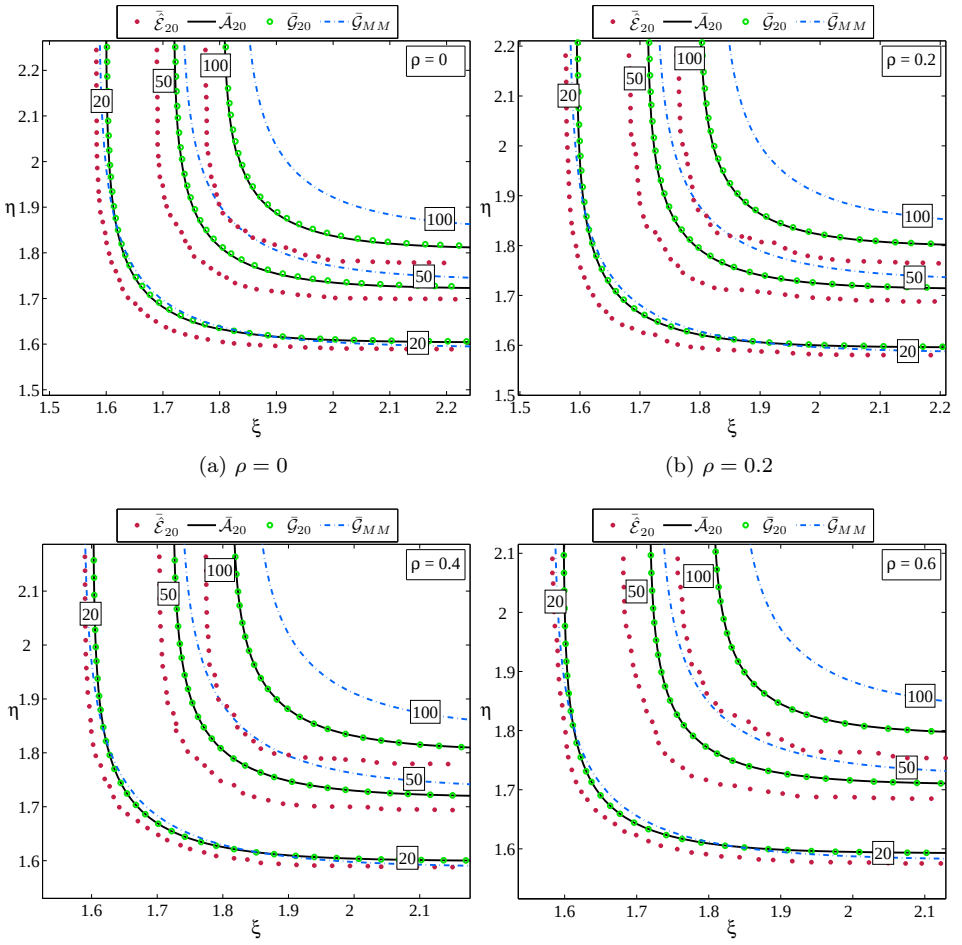


Figure 5.10: Estimation of the average of the 100 Monte Carlo simulated surfaces illustrated by the contour plot of the T-year return level. Contours of the optimized asymmetric logistic MC simulated surfaces, $\mathcal{A}_{20}$, (——); contour of the average of the 100 MC surfaces, $\bar{\mathcal{A}}_{20}$, (▬).

titative index too. One of the possible methods that can be adopted in our case to measure the reciprocal distance between two contour lines of a similar behaviour is to apply the Hausdorff distance based on the Euclidean norm $d(\cdot, \cdot)$ on a $\mathbb{R}^2$ space, cf. e.g. Alt et al. (1995); Alt and Guibas (2000). As discussed in Section 5.4, intuitively, the Hausdorff distance between 2D curves A and B given by the corresponding sets of points $A = \{a_1, \ldots, a_{N_A}\}$, $B = \{b_1, \ldots, b_{N_B}\}$, where $\forall a_i$ and

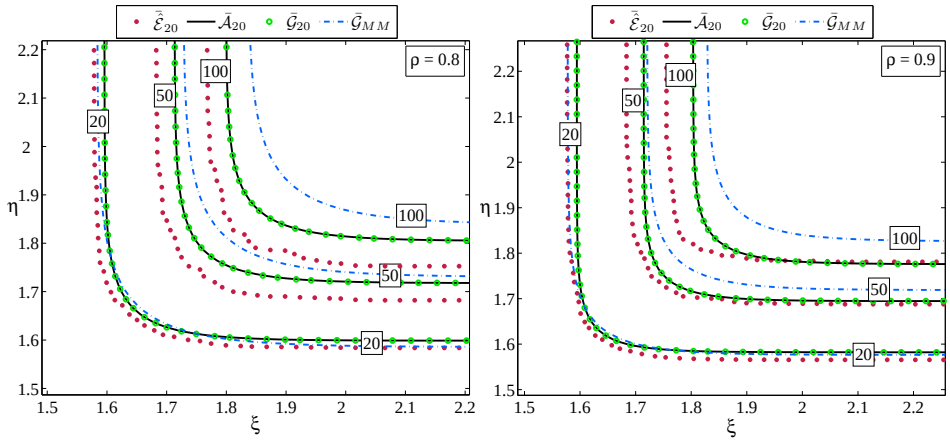


(a) $\rho = 0$ (b) $\rho = 0.2$

(c) $\rho = 0.4$ (d) $\rho = 0.6$

Figure 5.11: Response of the linear system. Contour plot of the T-year return levels (T = 20, 50 and 100 years) for the averaged surfaces: average of the 100 MC empirically estimated ACER surfaces, $\bar{\mathcal{E}}_{20}$, (●●); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_{20}$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_{20}$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (-·-). Boxes indicate return levels in years. The case of $\rho = 0$, 0.2, 0.4 and 0.6.

$\forall b_j \in \mathbb{R}^2$, is defined as follows. First, the Euclidean distances from any fixed point $a_i \in A$ to all points of B are calculated and the shortest distance is taken, that is the distance $\min_{j=1\ldots N_B} \{d(a_i, b_j)\}$. This procedure is passed across all the points $a_i \in A$, $i = 1 \ldots N_A$. Then, the maximum of the calculated minima is obtained as $\max_{i=1\ldots N_A} \min_{j=1\ldots N_B} \{d(a_i, b_j)\}$. Because this function is, generally speaking, asymmet-



(a) $\rho = 0.8$ (b) $\rho = 0.9$

(c) $\rho = 0.95$ (d) $\rho = 0.99$

Figure 5.12: Response of the linear system. Contour plot of the T-year return levels (T = 20, 50 and 100 years) for the averaged surfaces: average of the 100 MC empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_{20}$, (••); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_{20}$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_{20}$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (–·–). Boxes indicate return levels in years. The case of $\rho = 0.8$, 0.9, 0.95 and 0.99.

ric, the distance from B to A is calculated in the same manner. Finally, the largest of these two values would define the Hausdorff distance.

In our case, it was decided to consider the ACER $\hat{\mathcal{E}}_{20}$ surface as the reference and to measure the Hausdorff distance between the contours of this surface and the corresponding contours of the rest of the surfaces. Note that the magnitudes of the response time series $X(t)$ and $Y(t)$ are of the same units. It is also noted that both time series are zero-mean processes with equal standard deviation. Therefore, it was decided to calculate the Hausdorff distances directly. Otherwise, it would be necessary to initially relocate and rescale the measuring units.

In Table (5.1) the Hausdorff distances between the contours of the average ACER surface $\bar{\hat{\mathcal{E}}}_{20}$ and the rest of the surfaces are listed for all considered return periods (first column) and for all cases of the correlation between the data(second column). Distances between the contour lines of the surface $\bar{\hat{\mathcal{E}}}_{20}$ and the average of the fitted asymmetric logistic surfaces, $\bar{\mathcal{A}}_{20}$, are listed in the third column. This value, $d_H(\bar{\hat{\mathcal{E}}}_{20}, \bar{\mathcal{A}}_{20})$, is considered to be the benchmark for estimating the relative difference (listed in parentheses) with the other distances. It is clearly observed from column four of the table that the difference between the surfaces $\bar{\mathcal{A}}_{20}$ and $\bar{\mathcal{G}}_{20}$ is indeed insignificant. Moreover, in case of the highly correlated data ($\rho \geq 0.6$) the optimized logistic model $\mathcal{G}_k$ appears to provide better fit. Finally, the table data confirm that the Type B EVD given by the surface $\mathcal{G}_{MM}$ can provide satisfactory predictions for short return periods (see column 5), whereas for return period of 50 or 100 years the level of disagreement with the empirical ACER surface can be three times higher than for the optimized models.

5.5.2 Synthetic wind speeds coupled by a copula

In this section, the bivariate ACER methodology will be studied in more detail in a study on synthetic bivariate wind speed data coupled by a copula model with known extreme value distribution, and therefore a known T-year return period level $\left(\xi^{\mathrm{Tyr}}, \eta^{\mathrm{Tyr}}\right)$. The idea developed in this section is a natural extension of the research expound in Section 3.5.1. To get an idea about the performance of the bivariate ACER method and the existing Gumbel logistic model with Gumbel marginals fitted by the method of moments, Monte Carlo simulations will be carried out to produce 100 bivariate data samples. It will be observed that the predicted 100-year return period levels was consistently better for the ACER method.

First, the Gumbel-Haugaard (logistic) copula (Gumbel and Mustafi, 1967; Balakrishnan and Lai, 2009) will be considered. Then, it will be assumed that the bivariate wind speed data are combined by the Gaussian copula.

The Gumbel-Haugaard copula

Let us consider 70 years of synthetic coupled wind speed data, amounting to 7000 data points, which is not sufficient to perform a detailed bivariate statistical analysis. However, this case fully represents a real situation in which only a limited data sample is available. Under these conditions, it is crucial to utilize all available data to provide extreme value estimates.

By analogy with Section 3.5.1, it is assumed here that each component of the underlying stochastic bivariate process $Z(t) = \big(X(t),\, Y(t)\big)$ is stationary and Gaussian with a mean value of zero and standard deviation equal to one. It is also assumed that the mean zero up-crossing rates $\nu_x^+(0)$ and $\nu_y^+(0)$ are such that the products $\nu_x^+(0)T$ and $\nu_y^+(0)T$ are equal to 10^3 for $T = 1$ year. This appears to be typical for the wind speed process.

The Monte Carlo-simulated data to be used for the synthetic example are the peak events $Z^{3\mathrm{d}}(t) = \big(X^{3\mathrm{d}}(t),\ Y^{3\mathrm{d}}(t)\big)$ extracted from component-wise measurements of the wind speed process and separated by 3.65 days to obtain approximate

Return period	ρ	$d_H(\bar{\hat{\mathcal{E}}}_{20},\, \bar{\mathcal{A}}_{20})$	$d_H(\bar{\hat{\mathcal{E}}}_{20},\, \bar{\mathcal{G}}_{20})$; (rel.diff,%)	$d_H(\bar{\hat{\mathcal{E}}}_{20},\, \bar{\mathcal{G}}_{MM})$; (rel.diff,%)
20 years	0	0.035	0.037 (+7.4%)	0.045 (+29.1%)
	0.2	0.039	0.040 (+3.8%)	0.052 (+33.1%)
	0.4	0.024	0.024 (+0.4%)	0.034 (+40.9%)
	0.6	0.021	0.021 (−0.3%)	0.032 (+53.6%)
	0.8	0.020	0.020 (−0.6%)	0.025 (+25.4%)
	0.9	0.017	0.017 (−0.5%)	0.021 (+21.3%)
	0.95	0.022	0.022 (−0.2%)	0.020 (−10.0%)
	0.99	0.020	0.020 (−0.1%)	0.013 (−34.0%)
50 years	0	0.049	0.052 (+5.5%)	0.102 (+107.7%)
	0.2	0.055	0.056 (+2.8%)	0.109 (+97.7%)
	0.4	0.046	0.046 (+0.1%)	0.095 (+104.8%)
	0.6	0.052	0.052 (−0.2%)	0.103 (+99.7%)
	0.8	0.046	0.045 (−0.5%)	0.086 (+89.6%)
	0.9	0.029	0.029 (−0.3%)	0.060 (+103.0%)
	0.95	0.022	0.022 (−0.3%)	0.055 (+147.6%)
	0.99	0.021	0.021 (+0.0%)	0.049 (+139.4%)
100 years	0	0.073	0.075 (+3.7%)	0.158 (+117.6%)
	0.2	0.078	0.080 (+1.9%)	0.167 (+113.4%)
	0.4	0.083	0.083 (+0.0%)	0.166 (+99.5%)
	0.6	0.065	0.065 (−0.4%)	0.150 (+130.9%)
	0.8	0.065	0.065 (−0.2%)	0.140 (+113.4%)
	0.9	0.051	0.051 (−0.3%)	0.117 (+129.5%)
	0.95	0.028	0.028 (+0.0%)	0.091 (+226.9%)
	0.99	0.042	0.042 (+0.0%)	0.110 (+161.4%)

Table 5.1: Response of the linear system. The Hausdorff distances d_H between contour lines of the T-year return levels ($T = 20$, 50 and 100 years) of the averaged surfaces for different levels of dependence, ρ. The reference, $\bar{\hat{\mathcal{E}}}_{20}$, is the average of the 100 MC empirically estimated ACER surfaces; $\bar{\mathcal{A}}_{20}$ – av.100 of the optimized asymmetric logistic surfaces; $\bar{\mathcal{G}}_{20}$ – av.100 of the optimized logistic surfaces; $\bar{\mathcal{G}}_{MM}$ – av.100 of the estimated Gumbel logistic models with Gumbel marginals. The relative differences with respect to the benchmark $d_H(\bar{\hat{\mathcal{E}}}_{20},\, \bar{\mathcal{A}}_{20})$, are indicated in parentheses.

independence (Naess, 1998b). Thereafter, based on the Rice formula, the marginal extreme value distributions of the peak event data are

$$\begin{aligned} F^{3\mathrm{d}}(\xi) &= \exp\left\{-q_x \exp\left(-\frac{\xi^2}{2}\right)\right\} \\ G^{3\mathrm{d}}(\eta) &= \exp\left\{-q_y \exp\left(-\frac{\eta^2}{2}\right)\right\}, \end{aligned} \tag{5.81}$$

for given $q_x = \nu_x^+(0)T$ and $q_y = \nu_y^+(0)T$, where $T = 3.65$ days. It is also assumed that the dependence structure between the two marginal peak wind speed variables $X^{3\mathrm{d}}(t)$ and $Y^{3\mathrm{d}}(t)$ is described by the Gumbel-Haugaard copula (4.22) of the form

$$C(u,v) = \exp\left\{-\left[(-\log u)^m + (-\log v)^m\right]^{1/m}\right\}, \ m \geq 1, \tag{5.82}$$

where the dependence parameter m can be expressed through the Pearsonś product moment correlation coefficient ρ as follows, $m = 1/\sqrt{(1-\rho)}$, cf., e.g., Gumbel and Mustafi (1967); Balakrishnan and Lai (2009).

Accordingly, the bivariate extreme value distribution of the peak event data is:

$$\begin{aligned} H^{3\mathrm{d}}(\xi,\eta) &= \exp\left\{-\left[(-\log F^{3\mathrm{d}}(\xi))^m + (-\log G^{3\mathrm{d}}(\eta))^m\right]^{1/m}\right\} \\ &= \exp\left\{-\left[q_x^m \exp\left(-m\,\frac{\xi^2}{2}\right) + q_y^m \exp\left(-m\,\frac{\eta^2}{2}\right)\right]^{1/m}\right\}. \end{aligned} \tag{5.83}$$

Based on the approximate independence of the peak event data separated by 3-4 days, the distribution of the yearly extreme value of $Z^{3\mathrm{d}}(t)$ is then calculated by the formula

$$\begin{aligned} H^{1\mathrm{yr}}(\xi,\eta) &= \left(H^{3\mathrm{d}}(\xi,\eta)\right)^{100} \\ &= \exp\left\{-100\left[q_x^m \exp\left(-m\,\frac{\xi^2}{2}\right) + q_y^m \exp\left(-m\,\frac{\eta^2}{2}\right)\right]^{1/m}\right\}. \end{aligned} \tag{5.84}$$

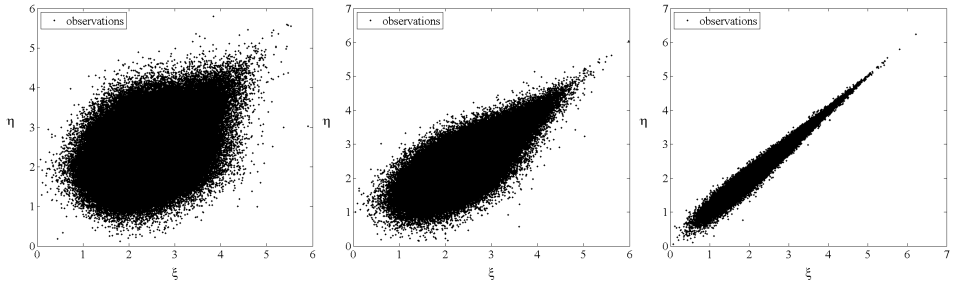
This implies that the exact T-year return period pairs $\left(\xi^{\mathrm{Tyr}}, \eta^{\mathrm{Tyr}}\right)$ are calculated as a solution of the equation

$$H^{1\mathrm{yr}}(\xi^{\mathrm{Tyr}}, \eta^{\mathrm{Tyr}}) = 1 - 1/T\,. \tag{5.85}$$

In our setup, we assumed $q_x = q_y = \nu^+(0)T = 10$. In other words, there are at the average 10 mean zero up-crossings within the period T that corresponds to $T^{3\mathrm{d}} \approx 3.65$ days. Thus, the assumption satisfies the condition of 10^3 mean zero up-crossings per year, as required (cf., e.g. Naess and Clausen (2001)). In fact, this constrains us to generate precisely 100 bivariate peak events from one year. To supply the process of empirical estimation of the bivariate ACER function with more accuracy, 70 years of observations were considered (in contradistinction to Section 3.5.1, where only 20 years were considered).

It was decided to analyse several cases of the degree of dependence between the components $X^{3\mathrm{d}}(t)$ and $Y^{3\mathrm{d}}(t)$ expressed in terms of the Pearson product-moment correlation coefficient ρ: $\rho = \{0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99\}$, so

that $m = \{1.29, 1.41, 1.58, 1.83, 2.24, 3.16, 4.47, 10\}$. Figure (5.13) illustrates the mutual alignment of the components of bivariate wind peak events for three out of eight values of ρ (and m). It is clearly observed that the spatial dependence



(a) $\rho = 0.4$, $m = 1.29$ (b) $\rho = 0.8$, $m = 2.24$ (c) $\rho = 0.99$, $m = 10$

Figure 5.13: Scatterplot of the simulated bivariate wind peak events coupled by the Gumbel-Haugaard copula for three different values of ρ (and m).

increases as $\rho \to 1$ ($\Leftrightarrow \; m \to +\infty$). Note that the mass density is concentrated mainly at the bottom left side of the data cluster.

Finally, to obtain an idea about the performance of the bivariate ACER method and the Type B extreme value distribution, that is, assuming Gumbel marginal distributions combined with a Gumbel-Hougaard copula, 100 independent 70-year Monte Carlo simulations were performed for each value of ρ. By analogy with the descriptions given on page 134, Section 5.5.1, the algorithm of bivariate extreme value analyses based on the Monte Carlo simulated data was performed according to the following steps.

1. 70 blocks of synthetic bivariate data with 100 observations per block, amounting to 7000 data points, were generated. The uniform components u and v were combined by the Gumbel-Hougaard copula in (5.82) using the Matlab (2009) Statistics Toolbox routine copularnd('Gumbel',m,N). Routine uses the Marshal-Olkin method of sampling the Archimedean copulas described by, e.g. Devroye (1986); Frees and Valdez (1998). The precise equations can be found in Melchiori (2006). Then, the peak events were sampled from the 3-day extreme value distributions $F^{3\mathrm{d}}(\xi)$ and $G^{3\mathrm{d}}(\eta)$ given by (5.81).

2. Because the peak events $X^{3\mathrm{d}}$ are assumed to be independent and separated by 3.65 days, and because the data are sampled from $F^{3\mathrm{d}}(\xi)$ independently, there is no temporal dependence characteristic of the time series. The same applies to process $Y^{3\mathrm{d}}$. This implies no need of conditioning on the previous data points when estimating the ACER function. Thus, the empirical univariate ACER functions $\hat{\varepsilon}_1^x(\xi)$ and $\hat{\varepsilon}_1^y(\eta)$ as well as the bivariate ACER surface $\hat{\mathcal{E}}_1$ were estimated based on 7000 observations.

3. The optimization procedures described in Section 3.4.1 were used to find the optimal values of the parameters q, b, a and c of the univariate ACER functions ε_1^x and ε_1^y. Note that because the generated time series follow the

identical distributions F^{3d} and G^{3d}, the tail markers ξ_1 and η_1 were fixed at the same level 2.4.

4. The asymmetric logistic $\mathcal{A}_1$ (5.54) and logistic $\mathcal{G}_1$ (5.51) models with optimized univariate ACER marginals were fitted to the empirically estimated $\hat{\mathcal{E}}_1$.
5. Five parameters of the Gumbel logistic distribution with the Gumbel marginals, given by Eqs.(5.78)-(5.79), were estimated by the method of moments. The location and scale parameters of Gumbel marginals F and G, and the dependence parameter m were estimated based on the samples of 70 annual extremes of X^{3d} and Y^{3d}, respectively.
6. Entries 1-5 were replicated 100 times in a loop to provide independent Monte Carlo simulations.
7. Average surfaces of the corresponding 100 Monte Carlo simulated surfaces were estimated, that is $\bar{\hat{\mathcal{E}}}_1$, $\bar{\mathcal{A}}_1$, $\bar{\mathcal{G}}_1$ and $\bar{\mathcal{G}}_{MM}$.

As previously mentioned, the procedure was performed eight times for each value of ρ.

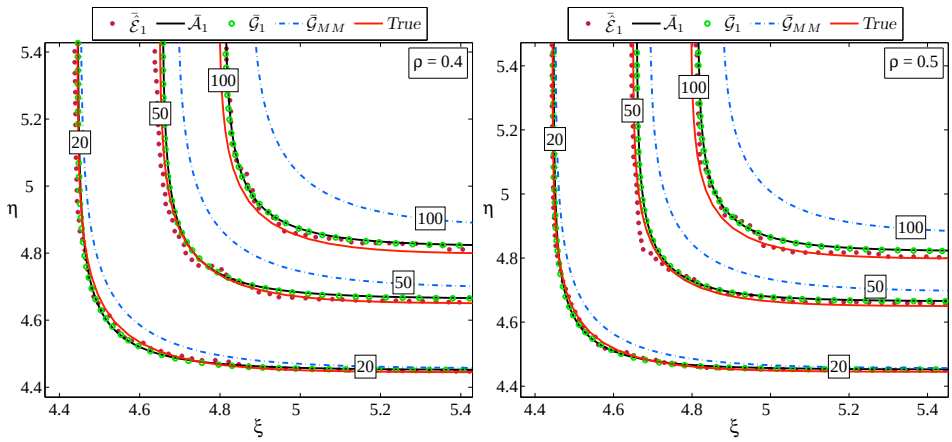
The results of the analysis are presented via contour lines of the T-year bivariate return levels for the 100 Monte Carlo average surfaces $\bar{\hat{\mathcal{E}}}_1$, $\bar{\mathcal{A}}_1$, $\bar{\mathcal{G}}_1$, and $\bar{\mathcal{G}}_{MM}$. Note that the contour lines in fact represent the solutions of the corresponding implicit equations Eq. (5.43), Eq. (5.80) and Eq. (5.85) for the exact return levels. Figures (5.14) and (5.15) present contour plots of the 20-, 50- and 100-year return period levels for the averaged surfaces.

In the first place, it is important to draw attention to the positional relationship of contours of the exact return period levels (*True*, red solid line (—)), and the average of 100 empirically estimated bivariate ACER surfaces ($\bar{\hat{\mathcal{E}}}_1$, red dots (••)). The figures reveal that with allowance for the sampling uncertainties and for a certain spacing inaccuracy of the discrete empirical bivariate ACER surface, the level of agreement between these two lines is actually remarkable for every value of return period T and for all values of parameter ρ. This allows to accept the fact that the empirically estimated bivariate ACER surface, in effect, provides an estimate of the true extreme value distribution.

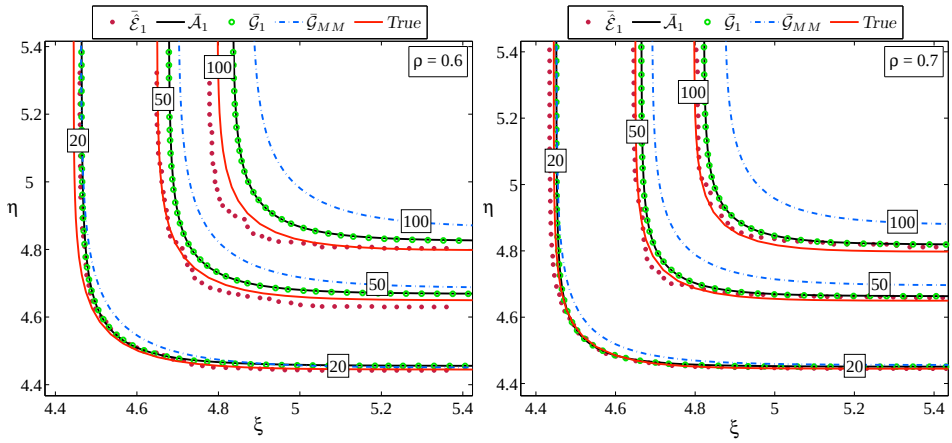
Figures (5.14) and (5.15) also reveal equally high degree of agreement between the exact bivariate yearly extreme value distribution (5.84) and the averages of the optimally fitted asymmetric logistic surfaces, $\bar{\mathcal{A}}_1$, and the optimized logistic surfaces, $\bar{\mathcal{G}}_1$. It is clearly observed that the contours of the optimally fitted surfaces $\bar{\mathcal{A}}_1$ and $\bar{\mathcal{G}}_1$ virtually match, which means that for any practical purpose there is no need to use the model with two extra parameters θ and ϕ. Finally, figures also demonstrate that the Gumbel logistic distribution with Gumbel marginals estimated by the method of moments, $\mathcal{G}_{MM}$, on average remains in acceptable agreement with the exact return levels for short periods, whereas for $T = 50$ and 100 years the divergence is significant.

In Table (5.2) the Hausdorff distances between the contours are listed for all considered return periods and for all values of the correlation coefficient ρ (see page 137 for more details). Naturally, the distances to the exact bivariate return level *True* is of interest. Because the both processes $X^{3d}(t)$ and $Y^{3d}(t)$ were sampled

from the identical distributions and are measured in the same units, the Hausdorff distances were calculated directly. It was also decided to consider the distance between the true levels *True* and the levels of $\bar{\mathcal{A}}_1$ as the reference value for the relative differences with the other distances.



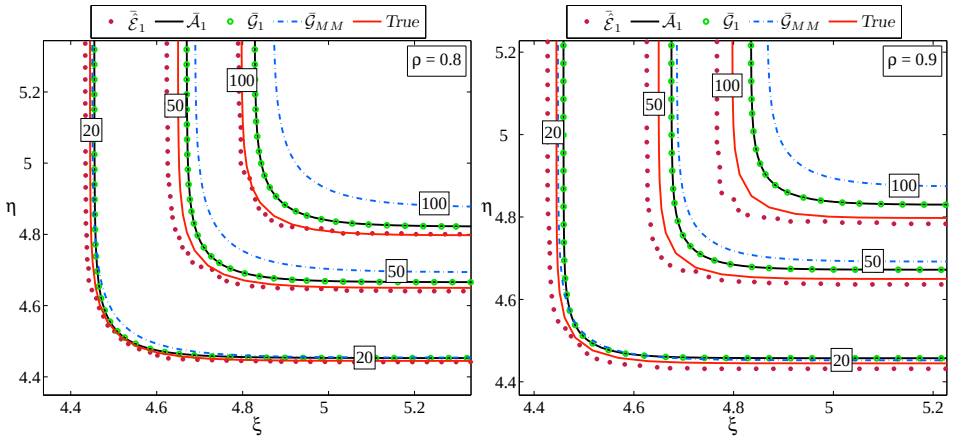
(a) $\rho = 0.4$, $m = 1.29$ (b) $\rho = 0.5$, $m = 1.41$



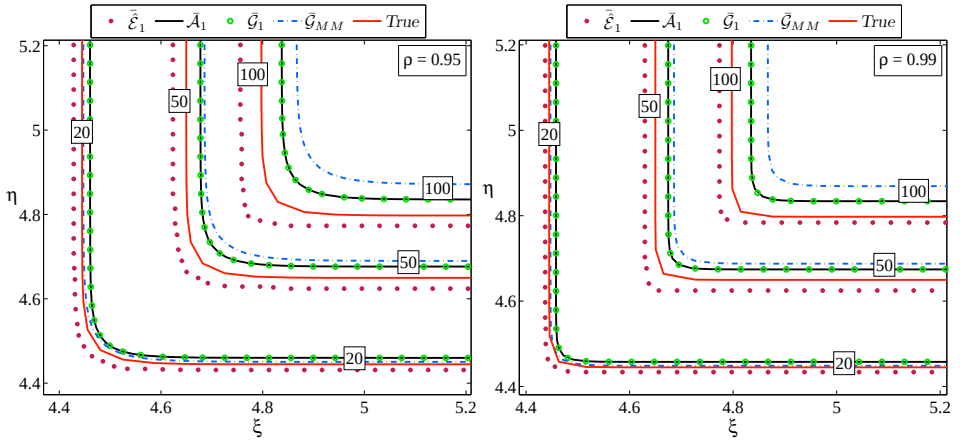
(c) $\rho = 0.6$, $m = 1.58$ (d) $\rho = 0.7$, $m = 1.83$

Figure 5.14: Synthetic winds with the Gumbel-Haugaard copula. Contour plot of the T-year return levels ($T = 20$, 50 and 100 years) for the averaged surfaces: average of the 100 Monte Carlo empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_1$, (●●); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_1$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_1$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (–·–). The exact return levels are defined by Eq.(5.85), *True*, (—). Boxes indicate return levels in years. The case of $\rho = 0.4$, 0.5, 0.6 and 0.7.

In general, the presented data confirm the conclusions we made based on the visual inspection of the Figures (5.14) and (5.15). However, it is also observed that the one-parameter logistic model $\mathcal{G}_1$, on average, provides either better estimation of the return period levels, that is, it matches with the exact levels from 5 to 10%



(a) $\rho = 0.8$, $m = 2.24$ (b) $\rho = 0.9$, $m = 3.16$



(c) $\rho = 0.95$, $m = 4.47$ (d) $\rho = 0.99$, $m = 10$

Figure 5.15: Synthetic winds with the Gumbel-Haugaard copula. Contour plot of the T-year return levels ($T = 20$, 50 and 100 years) for the averaged surfaces: average of the 100 Monte Carlo empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_1$, (●●); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_1$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_1$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (-·-). The exact return levels are defined by Eq.(5.85), *True*, (—). Boxes indicate return levels in years. The case of $\rho = 0.8$, 0.9, 0.95 and 0.99.

better than the asymmetric logistic model $\mathcal{A}_1$, or the relative difference between $d_H(\mathit{True}, \bar{\mathcal{A}}_1)$ and $d_H(\mathit{True}, \bar{\mathcal{G}}_1)$ is ignorable. In addition, the table reveals that the Hausdorff distance between the correct levels *True* and the contours of the surface $\mathcal{G}_{MM}$ can, on average, be up to seven times lager than the corresponding distances for the logistic model $\mathcal{G}_1$.

The Gaussian copula

It is known that the bivariate ACER logistic $\mathcal{G}_k(\xi, \eta)$ and asymmetric logistic $\mathcal{A}_k(\xi, \eta)$ models both emerge from the Gumbel-Haugaard copula, cf. Section 5.4.

Return period	ρ	$d_H(\mathit{True}, \bar{\mathcal{A}}_1)$	$d_H(\mathit{True}, \bar{\mathcal{G}}_1)$; (rel.diff,%)	$d_H(\mathit{True}, \bar{\mathcal{G}}_{MM})$; (rel.diff,%)
20 years	0.4	0.015	0.013 (−12.2%)	0.041 (+167.8%)
	0.5	0.011	0.010 (−10.7%)	0.038 (+248.1%)
	0.6	0.020	0.019 (−4.2%)	0.042 (+110.5%)
	0.7	0.007	0.006 (−10.5%)	0.033 (+371.3%)
	0.8	0.010	0.010 (−4.1%)	0.028 (+170.0%)
	0.9	0.015	0.016 (+2.1%)	0.020 (+34.4%)
	0.95	0.021	0.021 (+0.4%)	0.015 (−27.0%)
	0.99	0.018	0.018 (+0.1%)	0.009 (−49.1%)
50 years	0.4	0.016	0.014 (−10.6%)	0.099 (+516.0%)
	0.5	0.016	0.015 (−9.0%)	0.095 (+480.5%)
	0.6	0.030	0.029 (−4.0%)	0.095 (+218.8%)
	0.7	0.016	0.016 (−2.0%)	0.090 (+449.4%)
	0.8	0.024	0.024 (+1.1%)	0.082 (+248.8%)
	0.9	0.034	0.034 (+0.5%)	0.072 (+112.5%)
	0.95	0.041	0.041 (+0.2%)	0.066 (+62.2%)
	0.99	0.036	0.036 (+0.0%)	0.058 (+61.8%)
100 years	0.4	0.025	0.022 (−8.7%)	0.153 (+523.5%)
	0.5	0.026	0.025 (−4.4%)	0.149 (+476.6%)
	0.6	0.045	0.045 (+1.0%)	0.145 (+224.3%)
	0.7	0.032	0.032 (+1.2%)	0.143 (+344.9%)
	0.8	0.041	0.041 (+0.6%)	0.134 (+225.6%)
	0.9	0.053	0.053 (+0.3%)	0.123 (+132.2%)
	0.95	0.060	0.060 (+0.1%)	0.115 (+90.9%)
	0.99	0.054	0.054 (+0.0%)	0.106 (+96.4%)

Table 5.2: Synthetic winds with the Gumbel-Haugaard copula. The Hausdorff distances d_H between contour lines of the T-year return levels ($T = 20, 50$ and 100 years) of the averaged surfaces and the exact return level *True* (defined by Eq.(5.85)) for different levels of dependence ρ. $\bar{\mathcal{A}}_1$ – av.100 of the optimized asymmetric logistic surfaces; $\bar{\mathcal{G}}_1$ – av.100 of the optimized logistic surfaces; $\bar{\mathcal{G}}_{MM}$ – av.100 of the estimated Gumbel logistic distributions with Gumbel marginals. The relative differences with respect to the reference value $d_H(\mathit{True}, \bar{\mathcal{A}}_1)$, are indicated in parentheses.

Therefore, intuitively, it was rather expected to obtain a good performance of these models in application to the bivariate process sampled from the Gumbel-Haugaard copula. Thus, it was also decided to study the case of non-Archimedean, elliptical copula – the Gaussian copula (Joe, 1997; Balakrishnan and Lai, 2009).

As in the previous example defined on page 140), we consider 70 years of synthetic coupled wind speed data, amounting to 7000 data points. It is also assumed that the peak events $X^{3d}(t)$ and $Y^{3d}(t)$ are separated by 3.65 days and follow the same marginal distributions, as defined by Eq. (5.81) ($q_x = q_y = \nu^+(0)T = 10$). Now, it is assumed that the dependence structure between the two marginal peak wind speed variables $X^{3d}(t)$ and $Y^{3d}(t)$ is specified by the Gaussian copula for a given correlation coefficient ρ. The bivariate Gaussian (Normal) copula is given by (see corollary 4.3)

$$\begin{aligned} C(u,v) &= \Psi\big(\Phi^{-1}(u),\, \Phi^{-1}(v)\big) \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right] dsdt\,,\ 0 < \rho < 1\,, \end{aligned} \tag{5.86}$$

where Ψ denote the standard bivariate normal distribution function and Φ denote the standard univariate normal distribution function. Thus, the bivariate extreme value distribution of the peak event data in this case can be presented in a simplified form as follows,

$$\tilde{H}^{3d}(\xi,\eta) = \Psi\Big(\Phi^{-1}\left(F^{3d}(\xi)\right),\, \Phi^{-1}\left(G^{3d}(\eta)\right)\Big)\,. \tag{5.87}$$

Subsequently, the distribution of the bivariate annual extreme value is given by the formula

$$\tilde{H}^{1yr}(\xi,\eta) = \left[\Psi\Big(\Phi^{-1}\left(F^{3d}(\xi)\right),\, \Phi^{-1}\left(G^{3d}(\eta)\right)\Big)\right]^{100}. \tag{5.88}$$

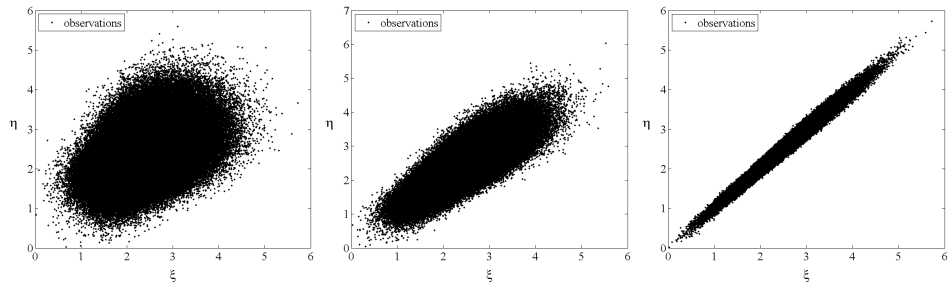
This leads to the implicit equation for the exact T-year return period pairs $\big(\xi^{Tyr},\, \eta^{Tyr}\big)$:

$$\tilde{H}^{1yr}(\xi^{Tyr},\, \eta^{Tyr}) = 1 - 1/T\,. \tag{5.89}$$

Thus, the copula type and everything that it yields are the only differences from the previous case. The other terms of the Monte Carlo framework were retained without changes.

In Figure (5.16) the bivariate process of the 3.65-day peak wind events is plotted for three out of eight values of the correlation coefficient ρ. Note that the mass density is concentrated more at the top right side of the data cluster.

It is expedient now to analyse the bivariate ACER method based on the data obtained from 100 independent Monte Carlo simulations. After the Gumbel-Haugaard copula was replaced by the Gaussian copula, all seven steps of the algorithm described on page 143 were reproduced in a similar manner eight times for each value of ρ. The results of the analysis are presented in Figures (5.17) and (5.18). Contour lines of the 20-year, 50-year and 100-year return period levels for the averaged empirical ACER surface $\bar{\tilde{\mathcal{E}}}_1$, the averaged asymmetric logistic and logistic surfaces $\bar{\mathcal{A}}_1$



(a) $\rho = 0.4$ (b) $\rho = 0.8$ (c) $\rho = 0.99$

Figure 5.16: Scatterplot of the simulated bivariate wind peak events coupled by the Gaussian copula for three different values of ρ.

and $\bar{\mathcal{G}}_1$, respectively, and the averaged surface of the Type B bivariate EVD, $\bar{\mathcal{G}}_{MM}$, are plotted together with the true return period levels calculated using Eq. (5.89).

Several key conclusions can be made based on the detailed visual inspection of the figures:

1. As in case of the Gumbel-Haugaard copula, contours of the empirically estimated discrete bivariate ACER surface $\hat{\mathcal{E}}_1$, on average, coincide with the exact return period levels depicted by red solid line. Several discrepancies (cf., e.g. Figures (5.17c), (5.18a)) can be explained by the uncertainties inherent in the sampled data and, especially, by a certain spacing inaccuracy relevant for the estimated contour lines of the discrete surface. Thus, the empirical ACER surface can potentially be a robust estimate of the exact extreme value distribution;
2. It is observed that the average of the optimized asymmetric logistic surfaces, that is $\bar{\mathcal{A}}_1$, appears to be in full agreement with the logistic model $\bar{\mathcal{G}}_1$. However, it is also evident that the degree of agreement between the exact return period levels (5.89) derived from the Gaussian copula, and the contours of both fitted AL and GL ACER models is lower as compared with the case of the Gumbel-Haugaard copula. The mismatch is specifically observable along the diagonal sections, which are of particular interest in the study of the dependence properties (cf., e.g. Nelsen (2006)). Thus, further in-depth study of the possible sub-asymptotic functional form of the empirical ACER surface is required.

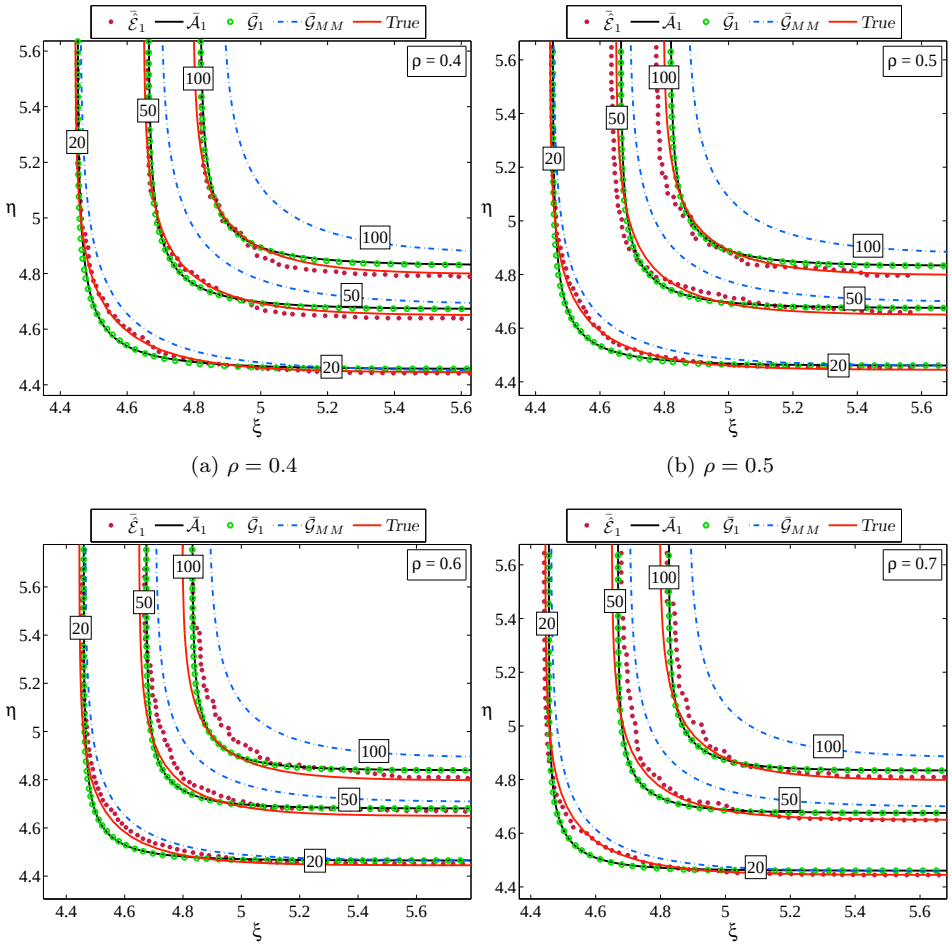
   And finally,

3. figures demonstrate that contour lines of the Type B bivariate EVD with Gumbel marginals estimated by the method of moments, that is $\mathcal{G}_{MM}$, on average, agree with the exact return levels for the return period $T = 20$ only. It is also noticed that the level lines of $\bar{\mathcal{G}}_{MM}$, in general, hold the same shape as the exact levels, though the shift from the exact levels is rather significant.

In Table (5.3) the Hausdorff distances to the exact bivariate return level *True* are listed for all considered return periods and for all values of the correlation

coefficient (see page 137 for more details).

The data in the table reveal that the one-parameter GL model $\mathcal{G}_1$ in some cases can provide, on average, better estimation of the return period levels than the AL model $\mathcal{A}_1$. In addition, an interesting comparison with the data given in Table (5.2) on page 147 can be made. Note that the comparison is correct because

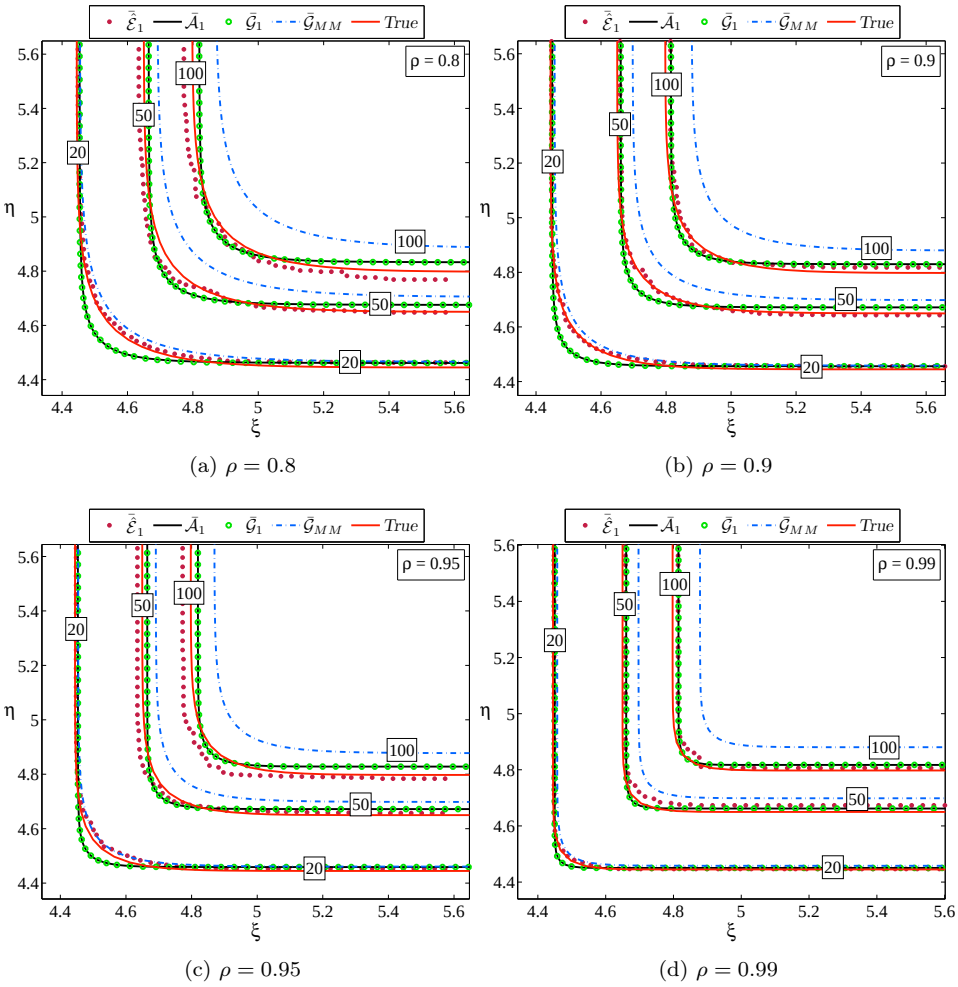


(a) $\rho = 0.4$ (b) $\rho = 0.5$

(c) $\rho = 0.6$ (d) $\rho = 0.7$

Figure 5.17: Synthetic winds with the Gaussian copula. Contour plot of the T-year return levels (T = 20, 50 and 100) for the averaged surfaces: average of the 100 Monte Carlo empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_1$, (●●); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_1$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_1$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (-·-). The exact return levels are defined by Eq.(5.89), *True*, (—). Boxes indicate return levels in years. The case of $\rho = 0.4$, 0.5, 0.6 and 0.7.

the peak wind events follow the same distribution in both cases. It appears that the distances between the exact levels and contours of the fitted surfaces $\bar{\mathcal{A}}_1$ and $\bar{\mathcal{G}}_1$ in case of the Gaussian copula are roughly two times bigger than the corresponding distances in Table (5.2) for the Gumbel-Haugaard copula. This clearly reflects the fact that the performance of the optimally fitted AL and GL models defined in



(a) $\rho = 0.8$ (b) $\rho = 0.9$

(c) $\rho = 0.95$ (d) $\rho = 0.99$

Figure 5.18: Synthetic winds with the Gaussian copula. Contour plot of the T-year return levels (T = 20, 50 and 100) for the averaged surfaces: average of the 100 Monte Carlo empirically estimated ACER surfaces, $\bar{\hat{\mathcal{E}}}_1$, (••); av.100 of the optimized asymmetric logistic models, $\bar{\mathcal{A}}_1$, (—); av.100 of the optimized logistic models, $\bar{\mathcal{G}}_1$, (○○); av.100 of the estimated Gumbel logistic models with Gumbel marginals, $\bar{\mathcal{G}}_{MM}$, (–·–). The exact return levels are defined by Eq.(5.89), *True*, (—). Boxes indicate return levels in years. The case of $\rho = 0.8$, 0.9, 0.95 and 0.99.

Eqs. (5.51) and (5.54) deteriorated for the case of the Gaussian copula.

5.6 Measured Wind Speed–Wave Height Data

In the reliability engineering and design of offshore structures probabilistic approaches are frequently adopted. They require the estimation of extreme quantiles of oceanographic data based on the statistical information. Due to strong correlation between such random variables as e.g. wave heights and wind speeds, application of the multivariate, or bivariate in the simplest case, extreme value theory is sometimes necessary.

Return period	ρ	$d_H(True, \bar{\mathcal{A}}_1)$	$d_H(True, \bar{\mathcal{G}}_1)$; (rel.diff,%)	$d_H(True, \bar{\mathcal{G}}_{MM})$; (rel.diff,%)
20 years	0.4	0.044	0.040 (−7.3%)	0.042 (−4.6%)
	0.5	0.047	0.045 (−4.9%)	0.040 (−15.6%)
	0.6	0.045	0.043 (−4.5%)	0.048 (+5.7%)
	0.7	0.056	0.055 (−2.6%)	0.032 (−42.7%)
	0.8	0.060	0.059 (−1.8%)	0.031 (−48.8%)
	0.9	0.058	0.058 (−0.8%)	0.018 (−68.5%)
	0.95	0.041	0.041 (−0.4%)	0.019 (−54.4%)
	0.99	0.022	0.022 (+0.0%)	0.015 (−31.9%)
50 years	0.4	0.023	0.022 (−3.9%)	0.101 (+331.0%)
	0.5	0.027	0.027 (−2.3%)	0.095 (+249.4%)
	0.6	0.031	0.031 (−0.5%)	0.109 (+251.9%)
	0.7	0.036	0.035 (−1.8%)	0.092 (+154.0%)
	0.8	0.043	0.043 (−1.0%)	0.082 (+91.1%)
	0.9	0.043	0.043 (−0.6%)	0.071 (+65.6%)
	0.95	0.029	0.029 (−0.3%)	0.064 (+124.4%)
	0.99	0.012	0.012 (−0.1%)	0.066 (+430.7%)
100 years	0.4	0.033	0.033 (−1.3%)	0.155 (+370.5%)
	0.5	0.035	0.035 (−0.8%)	0.146 (+316.6%)
	0.6	0.041	0.041 (−0.6%)	0.165 (+298.4%)
	0.7	0.036	0.036 (−0.4%)	0.148 (+311.8%)
	0.8	0.035	0.035 (−0.2%)	0.131 (+274.6%)
	0.9	0.032	0.032 (−0.1%)	0.121 (+277.8%)
	0.95	0.031	0.030 (−0.1%)	0.110 (+260.5%)
	0.99	0.020	0.020 (+0.0%)	0.115 (+480.7%)

Table 5.3: Synthetic winds with the Gaussian copula. The Hausdorff distances d_H between contour lines of the T-year return levels (T = 20, 50 and 100 years) of the averaged surfaces and the exact return level *True* (the reference, defined by Eq.(5.85)) for different levels of dependence ρ. $\bar{\mathcal{A}}_1$ – av.100 of the optimized asymmetric logistic surfaces; $\bar{\mathcal{G}}_1$ – av.100 of the optimized logistic surfaces; $\bar{\mathcal{G}}_{MM}$ – av.100 of the estimated Gumbel logistic models with Gumbel marginals. The relative differences with respect to the reference value $d_H(True, \bar{\mathcal{A}}_1)$, are indicated in parentheses.

Wind speed (WS - 3 hours mean) and significant wave height (Hs - total sea) data measured in the Norwegian sea at location N 65.29, E 7.32 were analysed by Naess and Karpa (2015b) to obtain numerical estimates of bivariate extremes. Figure (5.19) shows the geographical position of the measurement site. The data

Figure 5.19: Map of the part of Scandinavia with marked location.

were recorded during 54 years (1957 - 2011), eight times per day (every three hours).

Figure (5.20) presents the scatterplot of the observed data. As shown in this figure, there is a strong dependence between two time series.

The Pearson product-moment correlation coefficient is found to be $\rho = 0.79$. The Kendall's rank correlation coefficient (Kendall, 1938) is $\tau = 0.56$, and Spearman's ρ (Spearman, 1904) is equal to 0.7, which also indicates nonlinear agreement between WS and Hs.

Note that the available bivariate observations have low accuracy. This particularly applies to the significant wave height data, where the graduating mark is 0.1 meters and there are on average 98 unique numerical values of the Hs data per year. Clearly, for a fairly accurate estimation of the bivariate ACER functions, more data are required. Therefore, to increase the sample size of each record (realization), it was decided to divide the data series into 18 three-year records for the analysis. Furthermore, in this way, the standard deviation of the ACER function estimates can be calculated fairly accurately.

The univariate ACER functions $\hat{\varepsilon}_k$ were estimated first using the Matlab-based standalone downloadable application (Karpa, 2012). In Figures (5.21a) - (5.21b), $\hat{\varepsilon}_k$ is plotted versus different levels of wind speeds and wave heights, respectively, for different values of k. From both figures, it is clearly observed that there is a significant time dependence between WS observations and between Hs data. It is

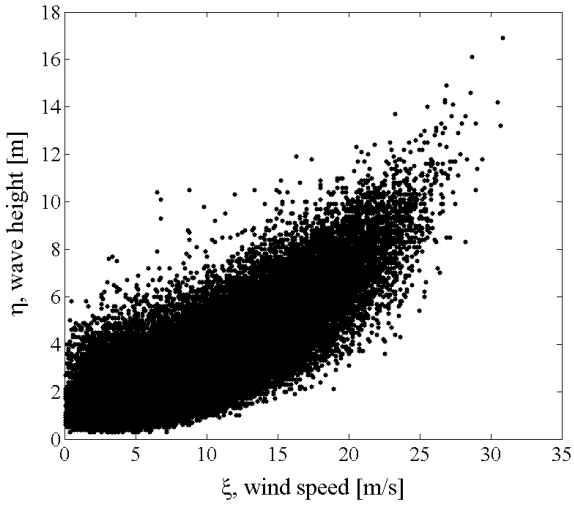
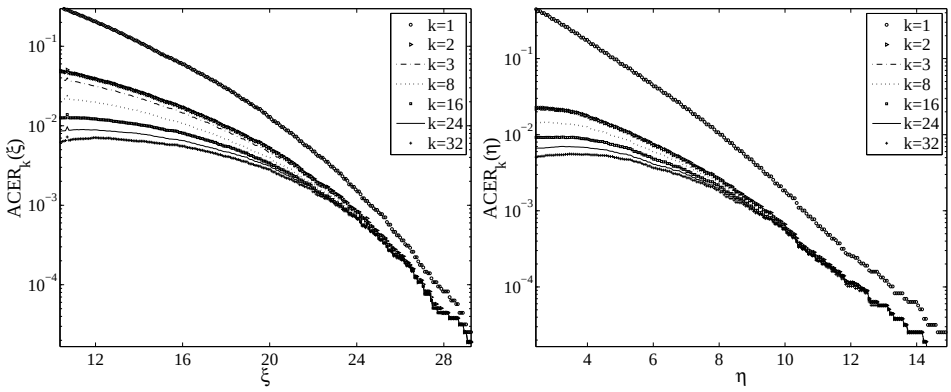


Figure 5.20: Coupled observations of Wind speed data (ξ axes) and significant wave height data (total sea, η axes).



(a) Wind speed data

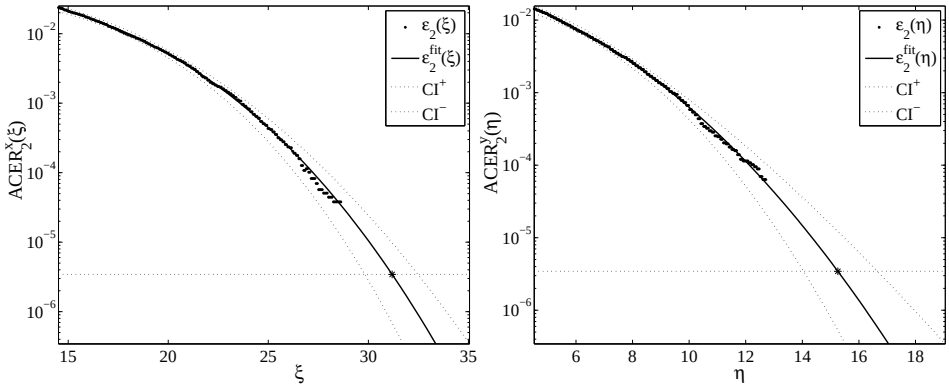
(b) Hs (total sea) data

Figure 5.21: Comparison between univariate ACER estimates for different degrees of conditioning: a) wind speed data; b) significant wave height (total sea) data.

also understood that this dependence effect is largely accounted for by $k = 16$ because there is a marked degree of convergence in the tail of $\hat{\varepsilon}_k$ for $k \geq 16$ in both cases. Clearly, $k = 16$ corresponds to exceedances separated by at least two days of non-exceedances for three-hour observations. For $k \geq 32$, that is, four-day declustered data, full convergence has been achieved. Figures (5.21a) and (5.21b) also demonstrate that for extreme value estimation, $\hat{\varepsilon}_2$ can be used because the ACER functions for $k \geq 2$ all converge in the far tail. This, inter alia, clearly demonstrates the power of an ACER function plot as a diagnostic tool to determine the value

of k required for extreme value estimation in a particular case. Although there are significant dependence effects for the WS and Hs data with lower magnitudes, for the extreme values, these effects are largely absent. This makes it possible to choose $k = 2$, which makes considerably more data available for estimation, with a possible reduction of uncertainty in the estimation as a result.

Figures (5.22a) - (5.22b) present plots of the optimized parametrical fit to the data for $\hat{\varepsilon}_k$ for $k = 2$ for both time series. In particular, the 100-year return level value and its 95% CI are estimated parametrically and plotted. For the wind speed data, the optimal parameters are $q = 0.05$, $b = 0.1$, $a = 1.9 \cdot 10^{-4}$, and $c = 3.14$, whereas for the Hs data, the parameters of the optimal curve are $q = 0.04$, $b = -2.27$, $a = 0.02$, and $c = 2.23$.



(a) Wind speed data (b) Hs (total sea) data

Figure 5.22: Plot of $\hat{\varepsilon}_2$ versus the observed data on a logarithmic scale for the optimized parameter values: a) wind speed data; $\xi_1 = 14.5$; b) significant wave height (total sea) data; $\eta_1 = 4.5$.

Figure (5.23) presents the empirically estimated bivariate ACER surfaces $\hat{\mathcal{E}}_k(\xi, \eta)$ for different values of k on a logarithmic scale. $\hat{\mathcal{E}}_k(\xi, \eta)$ with $k = 1$ is the upper most, then the following surfaces match in the tail for $k \geq 2$. As shown in this figure, the cross-section of the surfaces at the high level of wave height η provides the univariate ACER functions of the wind speed data, whereas the cross-section at a high wind speed level represents the univariate ACER of the Hs time series.

The same arguments as in the univariate case are applied to determine the bivariate ACER surface to be used in the analyses. That is, as long as the surfaces for $k \geq 2$ all converge in the tail and estimation of $\hat{\mathcal{E}}_2(\xi, \eta)$ is more accurate due to the availability of more data, we would choose the surface with the degree of conditioning $k = 2$.

The optimal parameters of the asymmetric logistic fit were found to be $m_A = 7$, $\theta = 1$ and $\phi = 0.91$. Because the additional parameters θ and ϕ are close to one, it would appear reasonable to consider the logistic model defined by Eq.(5.51). The same optimization procedure as in the asymmetric logistic case was applied

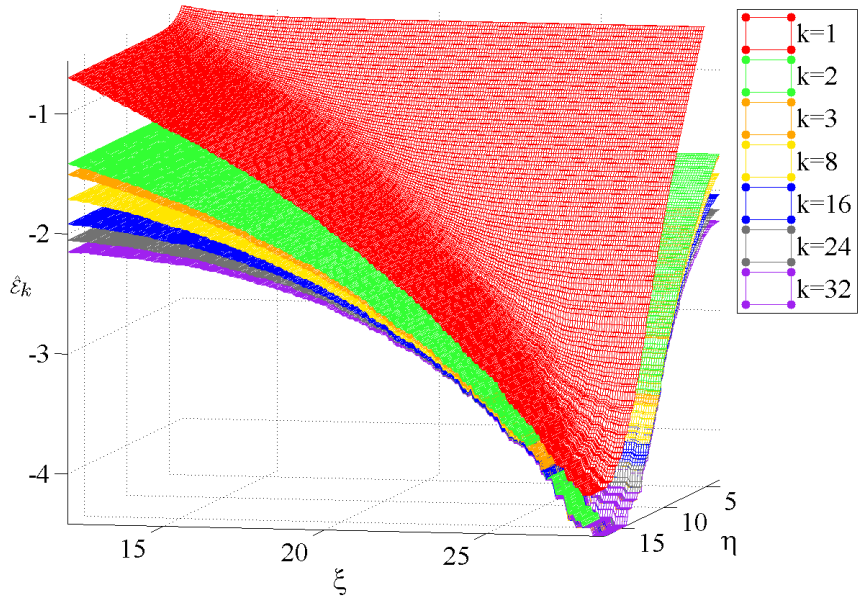


Figure 5.23: Comparison between Bivariate ACER surface estimates for different degrees of conditioning. $\hat{\mathcal{E}}_k(\xi, \eta)$ surfaces are plotted on a logarithmic scale

to obtain the optimal dependence parameter of the logistic copula $m_G = 4.78$.

Figures (5.24) and (5.25) show the contour plots of the optimized asymmetric logistic fit, $\mathcal{A}_2(\xi, \eta)$, to the data for the $\hat{\mathcal{E}}_2(\xi, \eta)$ surface and the contour plots of the optimized Gumbel logistic surface, $\mathcal{G}_2(\xi, \eta)$.

In Figure (5.24), contour lines of three surfaces are plotted for those levels of ξ and η, where the bivariate ACER surface $\hat{\mathcal{E}}_2(\xi, \eta)$ has been empirically estimated. Contour lines that correspond to the return period levels are presented in Figure (5.25).

These figures reveal that the empirical bivariate ACER surface $\hat{\mathcal{E}}_2$ captures high correlation between the data, as do the optimally fitted $\mathcal{G}_2$ and $\mathcal{A}_2$ surfaces. Note that the contour lines of the bivariate ACER surface of fully correlated data would show up as lines that consist of only horizontal and vertical line segments. This happens because for such data, $H(\xi, \eta) = \min\{F(\xi), G(\eta)\}$, which implies that $\mathcal{E}_k(\xi, \eta) = \max\{\varepsilon_k^x(\xi), \varepsilon_k^y(\eta)\}$, cf. Eq. (5.77), Section 5.5.1. It is observed that the level of agreement between the estimated bivariate ACER and the optimized asymmetric logistic and Gumbel logistic surfaces is equally significant. Thus, the optimized logistic model, $\mathcal{G}_k$, with asymptotically consistent marginals obtained from the optimized univariate ACER can be used as the parametric representative of the bivariate ACER surface estimated from the given data set.

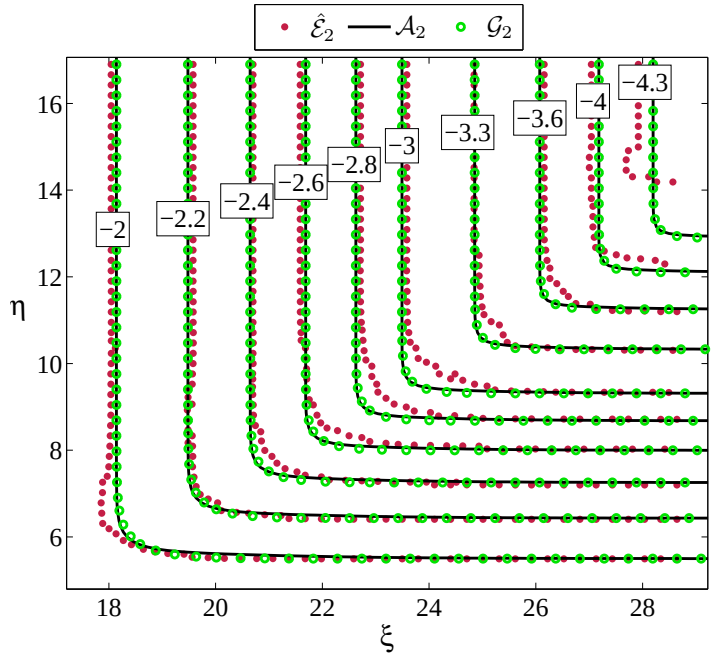


Figure 5.24: Contour plot of the empirically estimated $\hat{\mathcal{E}}_2(\xi, \eta)$ surface (••); optimized asymmetric logistic model, $\mathcal{A}_2(\xi, \eta)$, (—); optimized logistic model, $\mathcal{G}_2(\xi, \eta)$, (○ ○). Boxes indicate levels on a $\log_{10}$ scale.

5.7 Measured Wind Speed Data. Sula vs. Nordøyan.

Efficient calculation of the safety index for reliability analysis requires estimation of the extreme quantiles of structural loads as e.g. wind, wave etc., data based on the statistical information. In this section, application of the bivariate ACER method for simultaneous wind speed measurements from two separate locations will be demonstrated.

In the work by Naess and Karpa (2015a), the simultaneous wind speed data measured along the Norwegian coast at the Sula and Nordøyan Fyr weather stations (station numbers are 65940 and 75410, respectively) were analysed to obtain numerical estimates of bivariate extreme wind speeds. Figure (5.26) shows the geographical locations of the measurement sites. The hourly maximum of the three-second wind gust (10 meters above the ground) were recorded during 13 years (1999 - 2012).

Figure (5.27) presents the scatterplot of the observed data. This plot reveals a rather strong dependence between the two time series.

The Pearson product-moment correlation coefficient is found to be $\rho = 0.73$. The Kendall's rank correlation coefficient (Kendall, 1938) is $\tau = 0.5$, and Spearman's ρ (Spearman, 1904) is equal to 0.68, which also indicates nonlinear agreement between the Sula and Nordøyan wind speeds.

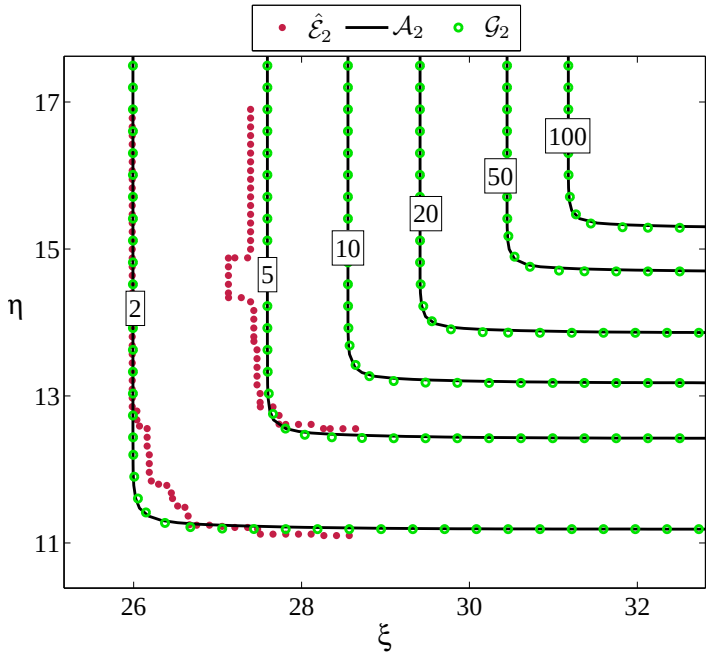


Figure 5.25: Contour plot of the return period levels for $\hat{\mathcal{E}}_2(\xi, \eta)$ surface, (••); optimized asymmetric logistic model, $\mathcal{A}_2(\xi, \eta)$, (—); optimized logistic surface, $\mathcal{G}_2(\xi, \eta)$, (○○). Boxes indicate return period levels in years.

It was decided to divide the data series into 13 one-year records for the analysis. In this way, the standard deviation of the ACER function estimates can be calculated fairly accurately.

The univariate ACER functions were estimated first using the Matlab-based standalone downloadable application (Karpa, 2012). In Figures (5.28a) and (5.28b), the cascades of $\hat{\varepsilon}_1 \ldots \hat{\varepsilon}_{96}$ are plotted versus different wind speed levels.

Both figures reveal that there is significant temporal dependence between consecutive data. It is also observed that this dependence effect is largely accounted for by $k = 24$ because there is a marked degree of convergence in the tail of $\hat{\varepsilon}_k$ for $k \geq 24$ in both cases. Here, for $k = 96$, which corresponds to conditioning on data recorded up to 4 days earlier, $\hat{\varepsilon}_{96}$ is considered to represent the final converged results because $\hat{\varepsilon}_{96} \approx \hat{\varepsilon}_k$ for $k > 96$ in the tail. Therefore, there is no need to consider conditioning of an even higher order than 96. So, effectively, $k_e = 96$ for our data. Also note that 4 days is a typical separation of wind speed data adopted in the declustering process to achieve independence between the data used in extreme value or peaks-over-threshold analyses. Figures (5.28a) and (5.28b) also demonstrate that for extreme value estimation, $\hat{\varepsilon}_1$ can be used because the ACER functions all coalesce in the far tail. This makes it possible to choose $k = 1$, which makes considerably more data available for estimation, with a possible reduction of uncertainty in the estimation as a result.

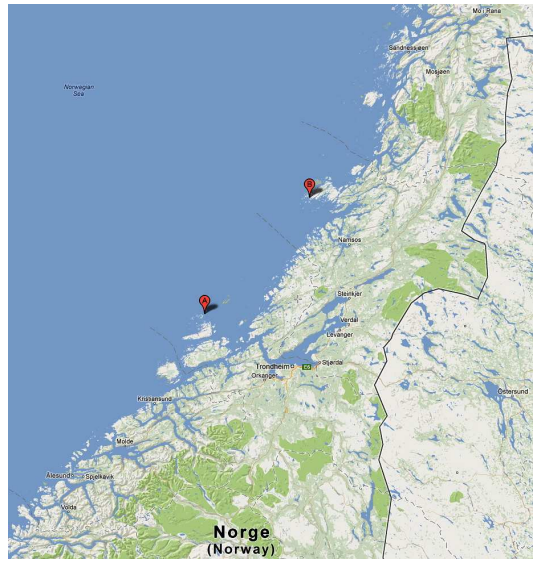


Figure 5.26: Map of the part of Norway with marked weather stations: A – Sula station, B – Nordøyan Fyr station.

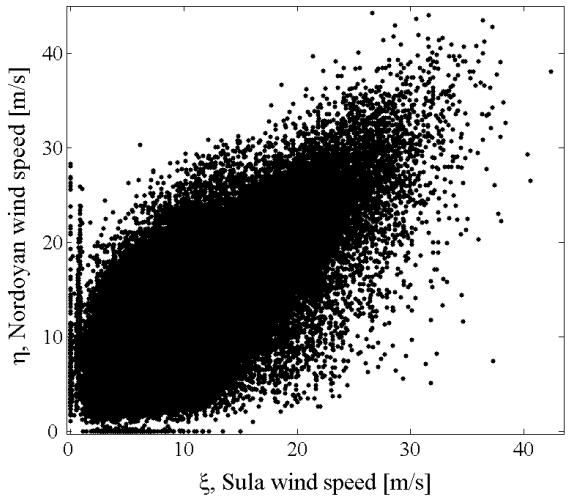
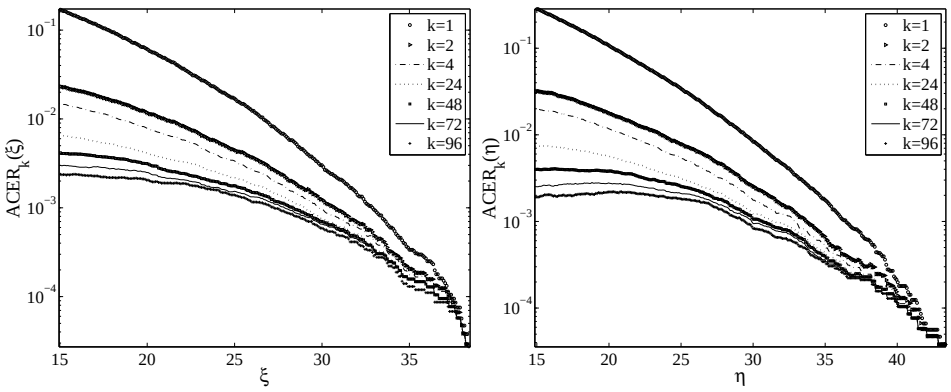


Figure 5.27: Coupled observations of Wind speed data observed at the Sula station (ξ axes) and at the Nordøyan Fyr station (η axes).

The cascade of estimated bivariate ACER surfaces $\hat{\mathcal{E}}_k(\xi, \eta)$ is shown in Figure (5.29). $\hat{\mathcal{E}}_k(\xi, \eta)$ with $k = 1$ is the upper most, then the following surfaces that match in the tail for all k. The ACER surface for $k = 96$ is very close to the surface obtained by taking the logarithm of the exact bivariate extreme value distribution.



(a) Sula (b) Nordøyan Fyr

Figure 5.28: Comparison between univariate ACER estimates for different degrees of conditioning: a) wind speed data from the Sula station; b) wind speed data from the Nordøyan Fyr station.

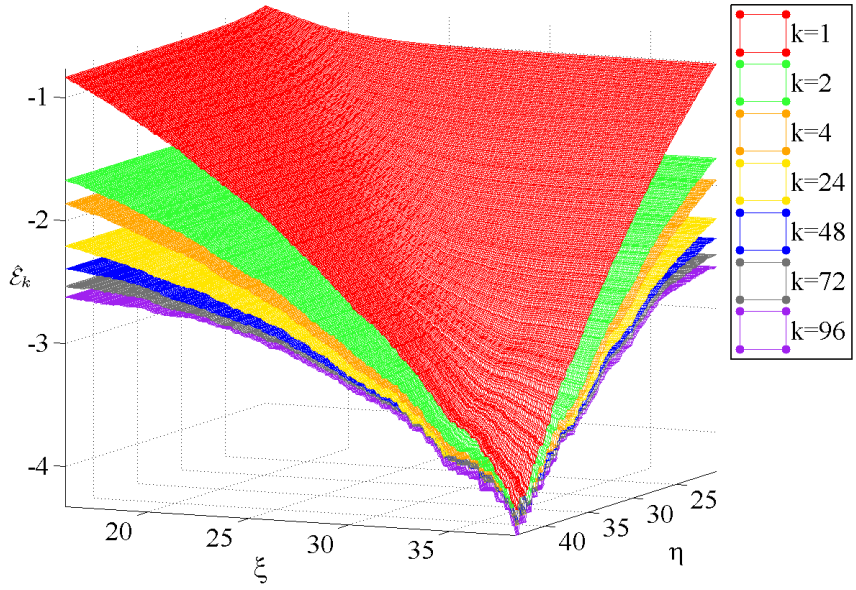


Figure 5.29: Comparison between Bivariate ACER surface estimates for different degrees of conditioning.

As shown in this figure, the cross-section of the surfaces at the high value of wind level η provides the univariate ACER functions of the wind speed data from the

Sula station, whereas the cross-section at a high level of ξ represents the univariate ACER of the time series from the Nordøyan Fyr.

The parameters of the optimal asymmetric logistic and Gumbel logistic surfaces are presented in Table 5.4.

k	$\mathcal{A}_k$	$\mathcal{G}_k$
1	$m_A = 2.44, \quad \theta = 0.86, \quad \phi = 0.92$	$m_G = 2.01$
96	$m_A = 4.53, \quad \theta = 0.97, \quad \phi = 0.97$	$m_G = 3.87$

Table 5.4: Optimal parameters of $\mathcal{A}_k$ and $\mathcal{G}_k$ fits.

Figures (5.30) and (5.31) show the contour plots of the optimized asymmetric logistic $\mathcal{A}_k(\xi, \eta)$ and optimized logistic $\hat{\mathcal{G}}_k(\xi, \eta)$ fits to the data for the $\hat{\mathcal{E}}_k(\xi, \eta)$ surface for $k = 1$ and $k = 96$, respectively. The contour lines of three surfaces are plotted for those levels of ξ and η, where the bivariate ACER surface $\hat{\mathcal{E}}_k(\xi, \eta)$ have been empirically estimated.

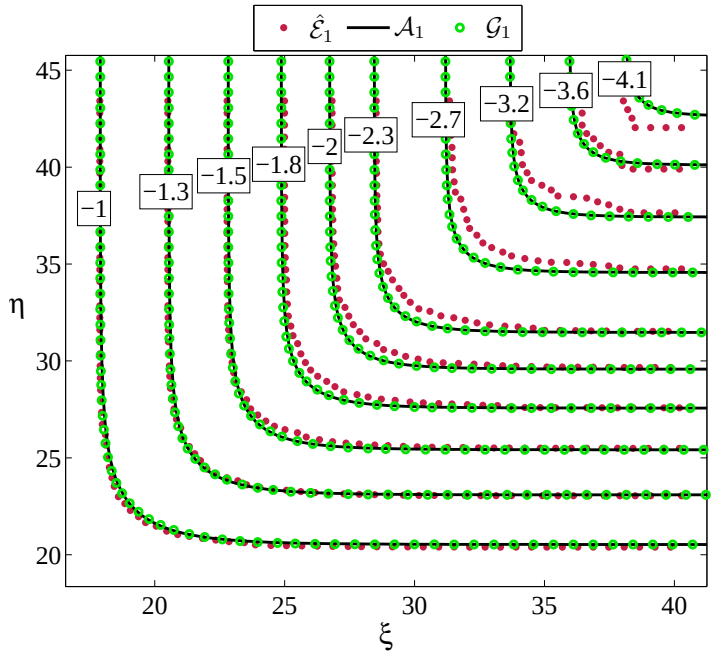


Figure 5.30: Contour plot of the empirically estimated $\hat{\mathcal{E}}_1$ surface, (••), and the optimized asymmetric logistic $\mathcal{A}_1$, (—), and optimized Gumbel logistic $\mathcal{G}_1$, (○○), surfaces based on marginal univariate ACER. Boxes indicate levels on a logarithmic scale.

These figures reveal that the empirical bivariate ACER surface $\hat{\mathcal{E}}_k$ captures high correlation between the data, as do the optimally fitted $\mathcal{G}_k$ and $\mathcal{A}_k$ surfaces. It is also observed that the behaviour of the estimated ACER surface in the case

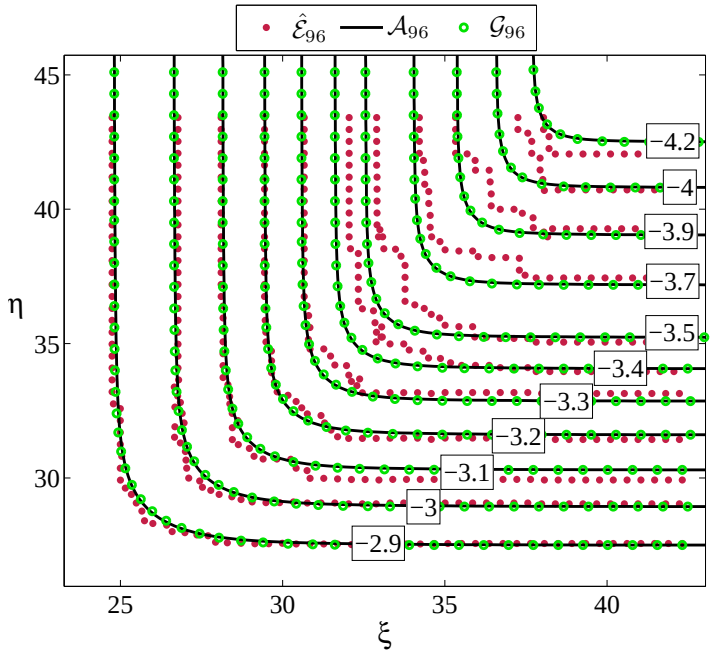


Figure 5.31: Contour plot of the empirically estimated $\hat{\mathcal{E}}_{96}$ surface, (• •), and the optimized asymmetric logistic $\mathcal{A}_{96}$, (—), and optimized Gumbel logistic $\mathcal{G}_{96}$, (○ ○), surfaces based on marginal univariate ACER. Boxes indicate levels on a logarithmic scale.

of $k = 96$ in Figure (5.31) affirms high uncertainty due to the deficiency of data. However, the optimal surfaces $\mathcal{G}_{96}$ and $\mathcal{A}_{96}$ capture the statistical properties of the bivariate observations.

It is easily seen that the level of agreement between the estimated bivariate ACER and the optimized asymmetric logistic and Gumbel logistic surfaces is equally significant. Yet it is also important to keep in mind that the empirical bivariate ACER $\hat{\mathcal{E}}_k$ is the only discrete surface. The Matlab (2009) built-in routine contourc that has been used to obtain the figures, calculates the contour lines by producing a regularly spaced grid determined by the dimensions of a surface. Therefore, it evidently generates a certain spacing inaccuracy of the bivariate ACER surface $\hat{\mathcal{E}}_k$ level lines plot. In addition, the figures ascertain that the optimized asymmetric logistic and Gumbel logistic surfaces conform at a level sufficient to affirm that they actually coincide. Thus, the optimized logistic model $\mathcal{G}_k$ with asymptotically consistent marginals obtained from the optimized univariate ACER functions can be used as the parametric representative of the bivariate ACER surface estimated from the given data set.

Finally, a comparison of the contour lines of $\mathcal{G}_1$ and $\mathcal{G}_{96}$ that correspond to the same return period levels shows fairly good agreement considering the high uncertainty for the case $k = 96$. To highlight the results that would be obtained

by adopting the common approach of assuming Gumbel marginal extreme value distributions combined with a suitable copula model, in Figure (5.32) are plotted the 50 and 100 year return period levels obtained by using the asymmetric logistic model with asymptotically consistent marginals obtained from the optimized univariate ACER marginals together with the corresponding return levels obtained by using the Gumbel logistic model with Gumbel marginals fitted by the method of moments, $\mathcal{G}_{MM}$. It is clear that the discrepancy is significant, which is primarily caused by the use of asymptotic Gumbel marginals.

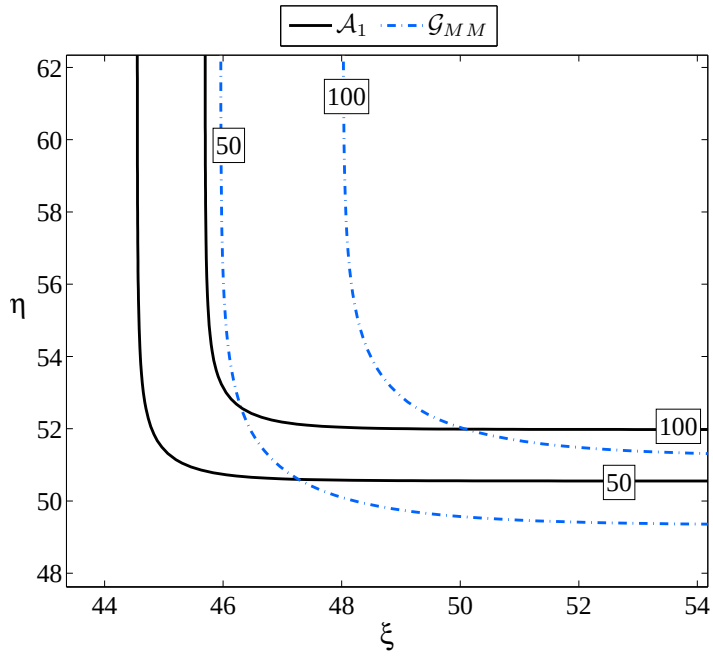


Figure 5.32: Contour plot of the return period levels for the optimized asymmetric logistic $\mathcal{A}_1$ surface, (—), and the Gumbel logistic model with Gumbel marginals $\mathcal{G}_{MM}$, (-·-). Boxes indicate return period levels in years.

Chapter 6

Conclusions

In this final chapter, the main research contributions and conclusions of this dissertation will be reviewed. This chapter also lists possible directions for future research.

6.1 Summary of Achievements

The following are the main research contributions of this dissertation. They are listed here in the order they are presented in the thesis text. (Some minor contributions are omitted.)

6.1.1 The univariate ACER

- **A preliminary method for estimating the tail marker η_1** is introduced, see Section 3.4.1. This value corresponds to the beginning of regular tail behaviour of the ACER function on a logarithmic scale. The method consists in basic statistical analysis of sampled data, and in close scrutiny of the estimated components $\hat{a}_k(\eta)$ and $\hat{b}_k(\eta)$ of the ACER function $\hat{\varepsilon}_k(\eta)$. It is emphasised that unlike the threshold level used in POT method, if the value of the tali marker η_1 has been chosen with some care, it does not affect the predicted return value significantly.
- **Robustness of the results of the optimization problem is improved** by implementation of the nonlinearly constrained optimization. In fact, several algorithms are used, notably the interior-point algorithm for finding minimum of constrained nonlinear multivariable function; the trust-region approach for solving nonlinear least-squares constrained minimization problems; the Levenberg-Marquardt algorithm for solving nonlinear least-squares unconstrained optimization. Then the results of various estimation approaches are compared and the optimal one defines the solution of the specific problem at hand.
- **Explicit expressions for the T-year return level by the ACER method** are presented by Eqs. (3.59) and (3.60) (see page 44). Although it is not a significant discovery, availability of these formulas promote general

clarification of the ACER method. In addition, expression in Eq. (3.58) is useful as it provides a level, the optimally fitted ACER function should be extrapolated to.

- **A reasonable extended in-depth analysis of the ACER method by application to synthetic data** is performed, see Section 3.5.1. 200 independent Monte Carlo-simulated wind data with known 100-year return period extreme value were used. It is demonstrated that the results provided by the ACER method are, on average, considerably more accurate as compared with those obtained by the POT and Gumbel methods. Furthermore, this study reveals that the extrapolation of confidence bands using the ACER scheme can secure a reasonable estimation of the 95% confidence interval.
- **Substantial analysis of real data**, that is, measured wind speeds at five locations, and measured tether tension obtained from model tests of an offshore platform for oil production, by the ACER method is performed in chapter 3. An efficient sequence of operations for extreme value analysis by the ACER method is described. In addition, it is demonstrated that the results of the ACER analysis are not sensitive to the outliers that can often be present in raw field measurements. Finally, the results of the analysis empirically prove the fact that when the ACER functions all converge in the far tail, the first ACER function ($k = 1$), which has been estimated with the least uncertainty, can be used for most practical applications. Thus, based on the overall performance, it is concluded that the ACER method may indeed be the preferred choice over the other two methods tested.
- **Development and introduction of the robust and straightforward standalone application** for the univariate ACER method (Karpa, 2012) is also a notable contribution of this research. When a novel method is introduced, it appears to be crucial to propagate it with easy-to-use tool available. The Matlab-based standalone downloadable application is designed for this purpose.
- **Bootstrapping for ACER** is presented in Section 3.5.3. It is an important contribution to the ACER method in regard to estimating the 95% CI. Particularly, we discuss the non-parametric bootstrap on the basis of the observed real environmental data. The results of the analysis reveal that estimation of the 95% CI through the extrapolation of confidence bands using the ACER scheme can generally provide a reasonable estimate. Although this method is a first-stage estimation, and in some cases the right endpoint of the estimated 95% CI can be shifted to the left, it is not as time consuming as the bootstrap method and provides 95% CI estimation for any degree of conditioning k.

  The parametric bootstrap is also discussed in Section 3.5.3. This bootstrapping procedure was carried out using two partially similar methods. It is ascertained that the considered parametric bootstrap approaches provide the 95% CI estimate, the consistency of which is, however, questionable. In addition, the considered parametric bootstrap can be used only with the data assumed to be independent, and as for any bootstrap approach, its execution requires considerable time.

The main conclusion of the study is that the adoption of the method of extrapolating the confidence bands using the ACER fit can remain the main approach for estimating the 95% CI of the T-year return level. The main advantages of this approach are the prompt and straightforward implementation, applicability to any degree of data conditioning k, and credibility of the confidence interval estimated by this method.

- **Analysis of synthetic data governed by the heavy tail distribution** is presented in Section 3.5.4. It is the first effort in investigating the performance of the general ACER approach (3.51). Based in 1000 independent 30-year Monte Carlo simulations, it is established that, on average, the general ACER method provides considerably more accurate predictions of the 100 - year return level than the other methods.
- **Analyses of narrow-banded time series** by the ACER method is discussed in Section 3.5.6. Peak values of many response processes and oceanographic time series have a tendency to occur in groups, which is a common characteristic of a narrow-banded time series. It is demonstrated that the ACER method solves this correlation problem efficiently and elegantly in a statistical sense. In addition, the positional relationship of the ACER functions of the densely sampled narrow-banded time series is discussed in detail.

6.1.2 The bivariate ACER

- **The bivariate ACER surface** $\mathcal{E}_k(\xi,\eta)$ is defined by Eq. (5.24). This definition differs from the one given by Naess (2011). Although this proposal is a minor contribution, it has predetermined further development of the bivariate ACER analysis.
- A crucial contribution is the concept of **representing the bivariate ACER as a function of two marginal univariate ACER functions**. Particularly, the following representation is introduced:

$$\mathcal{E}_k(\xi,\eta) = \left(\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)\right) \mathcal{D}\left(\frac{\varepsilon_k^x(\xi)}{\varepsilon_k^x(\xi) + \varepsilon_k^y(\eta)}\right), \tag{6.1}$$

  cf. Eq. (5.49) on page 121, where $\mathcal{D}$ is the Pickands dependence function. In this way, the temporal dependence characteristic of the bivariate time series is defined by the univariate marginal ACER functions, whereas the spatial dependence structure is given by the function $\mathcal{D}$.
- **A two-leveled optimization approach is used to define optimal parameters** of the functional representation of the bivariate ACER surface. First, it is proposed to employ the dependence function of logistic and asymmetric logistic extreme value copulas. Then the optimally fitted univariate ACER function are used to define optimal parameters of the bivariate ACER fit. The latter procedure recalls the best result from the interior-point algorithm for finding minimum of constrained nonlinear multivariable function and the trust-region approach for solving nonlinear least-squares constrained minimization problems. The possibility of using the Hausdorff distance for optimal surface matching is also discussed in Section 5.4.

- **A thorough analysis of the bivariate ACER based on synthetic Monte Carlo-simulated data** is addressed in Section 5.5. First, behaviour of the empirically estimated bivariate ACER function subject to correlation between components of the process is discussed in detail. Followed by this, an extended analysis of the bivariate ACER method by application to three types of synthetic data is performed. The main conclusions of the overall performance of the bivariate ACER method are as follows: 1) The degree of concordance between the empirical ACER surface $\hat{\mathcal{E}}_k$ and the optimized asymmetric logistic and logistic surfaces, $\mathcal{A}_k$ and $\mathcal{G}_k$, respectively, is equally high; 2) The level of agreement between the contour lines of $\mathcal{A}_k$ and $\mathcal{G}_k$ allows to infer that these surfaces actually coincide entirely. Moreover, in some cases the optimized one-parameter logistic model $\mathcal{G}_k$ appears to provide better fit than the three-parameter asymmetric logistic model $\mathcal{A}_k$; 3) The distinction between contours of the empirical ACER surface $\hat{\mathcal{E}}_k$ (or the contours of the true bivariate EVD) and contours of estimated Gumbel logistic model with Gumbel marginals is rather significant. Finally, the quantitative index for measuring the level of agreement between contour lines of the considered surfaces is introduced.
- **Study of the bivariate ACER method by application to environmental data** is performed in Section 5.6 and Section 5.7. An efficient sequence of operations for bivariate extreme value analysis by the ACER method is described. It is illustrated that the exact extreme value distribution given by the data can be captured within the inherent statistical uncertainty by using the ACER surfaces.

6.2 Further Work

There are several directions for future research. Some of them are listed here:

- As mentioned in Section 3.4.1, the used Levenberg-Marquardt algorithm was implemented for unconstrained optimization problem. It is advisable, however, to consider the constrained optimization as the one that can provide more reasonable values for the parameters. Hence, a natural direction for future research involves employing the Levenberg-Marquardt algorithm for constrained optimization.
- In the same way, a scientific interest is shown in using a sequential quadratic programming (SQP) method incorporated in the NAG Numerical Library (Numerical Algorithms Group, 2010). Research on the efficiency of this toolbox for the optimization problem can be performed in the future.
- The analysis presented in Section 3.5.4 is the initial study of the ACER method for general case when the data are not asymptotically Gumbel. Moreover, it appears that there is no robust method that would provide a reasonable estimation of the 95% CI for the general ACER. Thus, a thorough study of the general ACER is required, as well as its further practical application to financial data, for example.

- It is emphasised that adopted bivariate extreme value copula approach is only the first effort in investigating the functional representation of the empirically estimated bivariate ACER surface. In this way, the considered representation given by Eq. (6.1) cannot be considered as final and absolutely certain. This is particularly because the initial joint cumulative distribution function (5.45) (see page 120), with the Pickands dependence function $\mathcal{D}$, can be treated as asymptotic, whereas the sub-asymptotic representation for the bivariate ACER surface is required. Therefore, further in-depth study of the possible sub-asymptotic functional form of the ACER surface is required. A natural direction for future research is thorough analysis of possible sub-asymptotic dependence function that interconnect two univariate ACER marginals.
- The Hausdorff distance if often used in applications to measure difference between similar 3D objects. Thus, employing the Hausdorff distance for the optimization purposes (see Section 5.4) is a possible course of investigation.
- In the present research the common approach of assuming Gumbel marginal extreme value distributions combined with a logistic copula, that is, the Type B bivariate EVD, is considered for comparison purposes. It appears to be imperative to also consider other approaches together with the ACER method in application to more bivariate time series of different kind.
- It is necessary to develop a standalone application for the bivariate ACER method (or, in fact, to upgrade the existing Matlab routine to the level of downloadable standalone application). This task is slightly beyond the research activities, yet a crucial one.
- Last, but definitely not least, particularly interesting and necessary scientific activity should include analysis of the conditional extreme value distribution based on the bivariate ACER. For example, it is interesting to verify whether the bivariate ACER is able to improve the predicted T-year return period level for the process observed during shorter period of time than the other one. In this case, the conditional extreme value distribution of the process with less measurements given the coupled process is expressed as follows:

$$\begin{aligned} G^{1\mathrm{yr}}(\eta \,|\, \xi) &= \frac{H^{1\mathrm{yr}}(\xi,\eta)}{F^{1\mathrm{yr}}(\xi)} \\ &\approx \exp\left\{ -\left(\frac{N_1 - k + 1}{n_{y_1}} \mathcal{E}_k(\xi,\eta) - \frac{N_2 - k + 1}{n_{y_2}} \varepsilon_k(\xi) \right) \right\}, \end{aligned} \tag{6.2}$$

  where $N_2 > N_1$ and $n_{y_2} \geq n_{y_1}$, cf. Eq. (5.26) on page 116, Eq. (5.42) on page 119 and Eq. (3.57) on page 43. The T-year return period level $\eta^{\mathrm{Tyr}}|\xi_0$ is obtained as the solution of the equation

$$G^{1\mathrm{yr}}(\eta^{\mathrm{Tyr}}|\xi_0) \;=\; 1 - \frac{1}{T}\,. \tag{6.3}$$

References

Alt, H., B. Behrends, and J. Blömer (1995). Approximate matching of polygonal shapes. *Annals of Mathematics and Artificial Intelligence 13*, 251–265.

Alt, H. and L. J. Guibas (2000). Chapter 3 - discrete geometric shapes: Matching, interpolation, and approximation*. In J.-R. S. Urrutia (Ed.), *Handbook of Computational Geometry*, pp. 121–153. Amsterdam: North-Holland.

Balakrishnan, N. and C.-D. Lai (2009). *Continuous Bivariate Distributions*. New York: Springer Science+Business Media.

Balkema, A. A. and L. de Haan (1974, 10). Residual life time at great age. *The Annals of Probability 2*(5), 792–804.

Braunstein, S. L. (1992, July). How large a sample is needed for the maximum likelihood estimator to be approximately Gaussian? *Journal of Physics A Mathematical General 25*, 3813–3826.

Byrd, R. H., J. C. Gilbert, and J. Nocedal (2000). A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming 89*(1), 149–185.

Carroll, R. J. and D. Ruppert (1988). *Transformation and Weighting in Regression*. Chapman and Hall/CRC Monographs on Statistics and Appl. Probability Series. Chapman and Hall.

Cartwright, D. E. (1958). On estimating the mean energy of sea waves from the highest waves in a record. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 247*(1248), 22–48.

Cignoni, P., C. Rocchini, and R. Scopigno (1998). Metro: Measuring error on simplified surfaces. *Computer Graphics Forum 17*(2), 167–174.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. London: Springer-Verlag.

Coles, S., J. Heffernan, and J. Tawn (1999). Dependence measures for extreme value analyses. *Extremes 2*, 339–365.

Coles, S. G. (1994). A temporal study of extreme rainfall. In V. Barnett and K. F. Turkman (Eds.), *Statistics for the Environment 2: Water Related Iissues*, Chapter 4, pp. 61–78. Chichester: John Wiley & Sons.

Coles, S. G. and J. A. Tawn (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological) 53*(2), 377–392.

Coles, S. G. and J. A. Tawn (1994). Statistical methods for multivariate extremes: An application to structural design. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 43*(1), 1–48.

Cook, N. J. (1982). Towards better estimation of extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics 9*(3), 295–323.

Cook, N. J. (1985). *The designer's guide to wind loading of building structures Part 1*. London: Butterworths.

Cook, N. J. and R. I. Harris (2004). Exact and general FT1 penultimate distributions of extreme wind speeds drawn from tail-equivalent Weibull parents. *Structural Safety 26*, 391–420.

Danziger, Z. (2009). File Exchange – Matlab Central: Hausdorff Distance by Zachary Danziger. http://www.mathworks.com/matlabcentral/fileexchange/26738-hausdorff-distance. [online; accessed May 14, 2012.].

de Haan, L. (1976). Sample extremes: an elementary introduction. *Statistica Neerlandica 30*(4), 161–172.

de Haan, L. and J. de Ronde (1998). Sea and wind: Multivariate extremes at work. *Extremes 1*, 7–45.

de Haan, L. and S. Resnick (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 40*(4), 317–337.

Devroye, L. (1986). *Non-Uniform Random Variate Generation*. New York, NY: Springer-Verlag.

Draper, N. R. and H. Smith (1998). *Applied Regression Analysis* (3rd ed.). Wiley Series in Probability and Statistics: Texts and References Section. New York, NY: John Wiley & Sons.

Eastoe, E., S. Koukoulas, and P. Jonathan (2013). Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. *Ocean Engineering 62*(0), 68–77.

Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap*. Monographs on statistics and applied probability. London: Chapman & Hall.

Embrechts, P., A. McNeil, and D. Straumann (1999). Correlation and dependence in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, pp. 176–223. Cambridge University Press.

Ewans, K. and P. Jonathan (2014). Evaluating environmental joint extremes for the offshore industry using the conditional extremes model. *Journal of Marine Systems 130*(0), 124–130.

Falk, M. and F. Marohn (1993). Von mises conditions revisited. *The Annals of Probability 21*(3), 1310–1328.

Falk, M. and R.-D. Reiss (2003). Efficient estimators and LAN in canonical bivariate POT models. *Journal of Multivariate Analysis 84*(1), 190–207.

Ferro, C. A. T. and J. Segers (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 65*(2), 545–556.

Fisher, R. A. and L. H. C. Tippett (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society 24*, 180–190.

Forst, W. and D. Hoffmann (2010). *Optimization - Theory and Practice*. New York, NY: Springer.

Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*.

Frees, E. W. and E. A. Valdez (1998). Understanding relationships using copulas. *North American Actuarial Journal 2*, 1–25.

Genest, C., K. Ghoudi, and L.-P. Rivest (1998). "Understanding Relationships Using Copulas," by Edward Frees and Emiliano Valdez, January 1998. *North American Actuarial Journal 2*(3), 143–149.

Gill, P., W. Murray, and M. H. Wright (1981). *Practical Optimization*. London: Academic Press.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics 44*(3), 423–453.

Gudendorf, G. and J. Segers (2010). Chapter 6. Extreme-Value Copulas. In P. Jaworski, F. Durante, W. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications: Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*. Berlin: Springer-Verlag.

Gumbel, E. J. (1958a). Distributions à plusieurs variables dont les marges sont données. *Comptes Rendus de l'Académie des Sciences, Paris 246*, 2717–2719.

Gumbel, E. J. (1958b). *Statistics of extremes*. New York, NY: Columbia University Press.

Gumbel, E. J. (1960a). Bivariate exponential distributions. *Journal of the American Statistical Association 55*(292), 698–707.

Gumbel, E. J. (1960b). Multivariate extremal distributions. *Bulletin de l'Institut International de Statistique 37*(2), 471–475.

Gumbel, E. J. (1961). Bivariate logistic distributions. *Journal of the American Statistical Association 56*(294), 335–349.

Gumbel, E. J. (1965). Two systems of bivariate extremal distributions. *Bulletin de l'Institut International de Statistique 41*(2), 749–763.

Gumbel, E. J. and N. Goldstein (1964). Analysis of empirical bivariate extremal distributions. *Journal of the American Statistical Association 59*(307), 794–816.

Gumbel, E. J. and C. K. Mustafi (1967). Some analytical properties of bivariate extremal distributions. *Journal of the American Statistical Association 62*(318), 569–588.

Harris, R. (2001). The accuracy of design values predicted from extreme value analysis. *Journal of Wind Engineering and Industrial Aerodynamics 89*(2), 153–164.

Heffernan, J. E. and J. A. Tawn (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 66*(3), 497–546.

Hougaard, P. (1986). A class of multivanate failure time distributions. *Biometrika 73*(3), 671–678.

Hsing, T. (1987). On the characterization of certain point processes. *Stochastic Processes and their Applications 26*(0), 297–316.

Hsing, T. (1991). Estimating the parameters of rare events. *Stochastic Processes and their Applications 37*(1), 117–139.

Hutchinson, T. and C. Lai (1990). *Continuous Bivariate Distributions, Emphasising Applications*. Rumsby Scientific Publishing.

Huttenlocher, D., G. Klanderman, and W. Rucklidge (1993, Sep). Comparing images using the hausdorff distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 15*(9), 850–863.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society 81*(348), 158–171.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

Jonathan, P., J. Flynn, and K. Ewans (2010). Joint modelling of wave spectral parameters for extreme sea states. *Ocean Engineering 37*(11–12), 1070–1080.

Kanzow, C., N. Yamashita, and M. Fukushima (2004). Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *Journal of Computational and Applied Mathematics 172*(2), 375–397.

Karpa, O. (2012, October). ACER User Guide. http://folk.ntnu.no/karpa/ACER/ACER_User_guide.pdf. [PDF file available online; accessed October 04, 2012.].

Karpa, O. and A. Naess (2013). Extreme value statistics of wind speed data by the ACER method. *Journal of Wind Engineering and Industrial Aerodynamics 112*, 1–10.

Kaufmann, E. and R.-D. Reiss (1995). Approximation rates for multivariate exceedances. *Journal of Statistical Planning and Inference 45*(1–2), 235–245.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika 30*(1/2), 81–93.

Kotz, S. and S. Nadarajah (2000). *Extreme Value Distributions: Theory and Applications*. Imperial College Press.

Leadbetter, M. (1983). Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 65*(2), 291–306.

Leadbetter, M. (1995). On high level exceedance modeling and tail inference. *Journal of Statistical Planning and Inference 45*(1–2), 247–260.

Leadbetter, M. R., G. Lindgren, and H. Rootzen (1983). *Extremes and Related Properties of Random Sequences and Processes*. New York, NY: Springer-Verlag.

Leadbetter, M. R., I. Weissman, L. de Haan, and H. Rootzen (1989). On clustering of high levels in statistically stationary series. In J. Samson (Ed.), *Proceedings of the fourth International Meeting on Statistical Climatology: [27-31 March, 1989, Rotorua, New Zealand]*, Wellington. New Zealand Meteorological Service.

Lourakis, M. (2004). levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++. http://www.ics.forth.gr/~lourakis/levmar. [online; accessed September 11, 2013.].

Löw, P. (2015). Natural Catastrophes in 2012 Dominated by U.S. Weather Extremes. http://www.worldwatch.org/natural-catastrophes-2012-dominated-us-weather-extremes-0. [online; accessed April 20, 2015.].

Marshall, A. W. and I. Olkin (1967). A Multivariate Exponential Distribution. *Journal of the American Statistical Association 62*(317), 30–44.

Marshall, A. W. and I. Olkin (1988). Families of Multivariate Distributions. *Journal of the American Statistical Association 83*(403), 834–841.

Matlab (2009). *version 7.9.0.529 (R2009b)*. Natick, Massachusetts: The MathWorks Inc.

Melchiori, M. R. (2006, April). Tools for Sampling Multivariate Archimedean Copulas. YieldCurve, http://dx.doi.org/10.2139/ssrn.1124682.

Montgomery, D. C., E. A. Peck, and G. G. Vining (2001). *Introduction to Linear Regression Analysis* (3rd ed.). Wiley series in probability and statistics: Texts, references, and pocketbooks section. New York, NY: John Wiley & Sons.

Moré, J. and D. Sorensen (1983). Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing 4*(3), 553–572.

Naess, A. (1984). On the long-term statistics of extremes. *Applied Ocean Research 6*(4), 227–228.

Naess, A. (1985). The joint crossing frequency of stochastic processes and its application to wave theory. *Applied Ocean Research 7*(1), 35–50.

Naess, A. (1990). Approximate first-passage and extremes of narrow-band Gaussian and non-Gaussian random vibrations. *Journal of Sound and Vibration 138*(3), 365–380.

Naess, A. (1998a). Estimation of long return period design values for wind speeds. *Journal of Engineering Mechanics, ASCE 124*(3), 252–259.

Naess, A. (1998b). Statistical extrapolation of extreme value data based on the peaks over threshold method. *Journal of Offshore Mechanics and Arctic Engineering, ASME 120*, 91–96.

Naess, A. (2011). A Note on the Bivariate ACER Method. *Department of Mathematical Sciences, Norwegian University of Science and Technology. Preprint Statistics* (01/2011).

Naess, A. and P. H. Clausen (2001). Combination of peaks-over-threshold and bootstrapping methods for extreme value prediction. *Structural Safety 23*(4), 315–330.

Naess, A. and O. Gaidai (2008). Monte Carlo methods for estimating the extreme response of dynamical systems. *Journal of Engineering Mechanics, ASCE 134*(8), 628–636.

Naess, A. and O. Gaidai (2009). Estimation of extreme values from sampled time series. *Structural Safety 31*(4), 325–334.

Naess, A., O. Gaidai, and O. Batsevych (2009). Extreme value statistics of combined load effect processes. *Structural Safety 31*, 298–305.

Naess, A., O. Gaidai, and O. Batsevych (2010). Prediction of extreme response statistics of narrow-band random vibrations. *Journal of Engineering Mechanics 136*(3), 290–298.

Naess, A., O. Gaidai, and S. Haver (2007). Efficient estimation of extreme response of drag dominated offshore structures by Monte Carlo simulation. *Ocean Engineering 34*(16), 2188–2197.

Naess, A., O. Gaidai, and O. Karpa (2013). Estimation of extreme values by the average conditional exceedance rate method. *Journal of Probability and Statistics 2013*, 15 pages.

Naess, A., O. Gaidai, and P. S. Teigen (2007). Extreme response prediction for nonlinear floating offshore structures by Monte Carlo simulation. *Applied Ocean Research 29*(4), 221–230.

Naess, A. and E. Haug (2010). Extreme value statistics of wind speed data by the POT and ACER methods. *Journal of Offshore Mechanics and Arctic Engineering, ASME 132*(4), 041604–1 – 7.

Naess, A. and O. Karpa (2015a). Statistics of bivariate extreme wind speeds by the ACER method. *Journal of Wind Engineering and Industrial Aerodynamics 139*(0), 82–88.

Naess, A. and O. Karpa (2015b). Statistics of extreme wind speeds and wave heights by the bivariate ACER method. *Journal of Offshore Mechanics and Arctic Engineering 137*(2), 021602 (2015) (7 pages).

Naess, A. and T. Moan (2012). *Stochastic Dynamics of Marine Structures*. Cambridge University Press.

Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer Series in Statistics. New York: Springer-Verlag.

Nelsen, R. B. (2006). *An Introduction to Copulas, 2nd edition*. Springer Series in Statistics. New York: Springer Science+Business Media.

Newland, D. (1993). *An introduction to random vibrations, spectral and wavelet analysis*. Longman Scientific & Technical.

NOAA (2015). Billion-Dollar Weather and Climate Disasters: Summary Stats. http://www.ncdc.noaa.gov/billions/summary-stats. [online; accessed April 20, 2015.].

Norwegian Meteorological Institute (2012). Climate Data Web Services. http://eklima.met.no/wsKlima/start/start_en.html. [online; accessed December 29, 2011].

Numerical Algorithms Group (2010). *NAG Toolbox for Matlab*. Oxford, UK: NAG Ltd.

Ochi, M. K. (1990). *Applied probability and stochastic processes in engineering and physical sciences*. Wiley Series in Probability and Statistics. New York, NY: Wiley.

Palutikof, J. P., B. B. Brabson, D. H. Lister, and S. T. Adcock (1999). A review of methods to calculate extreme wind speeds. *Meteorological Applications 6*, 119–132.

Perrin, O., H. Rootzen, and R. Taesler (2006). A discussion of statistical methods used to estimate extreme wind speeds. *Theoretical and Applied Climatology 85*(3-4), 203–215.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics 3*(1), 119–131.

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute 49*, 859–878.

Rees, D. (2001). *Essential Statistics* (4th ed.). Chapman and Hall / CRC.

Reiss, R.-D. and M. Thomas (2007). *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields* (3rd ed. ed.). Basel: Birkhäuser.

Resnick, S. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer-Verlag.

Robert, C. Y. (2009). Inference for the limiting cluster size distribution of extreme values. *Annals of Statistics 37*(1), 271–310.

Robinson, M. E. and J. A. Tawn (2000). Extremal analysis of processes sampled at different frequencies. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 62*(1), 117–135.

Ryan, T. (2008). *Modern Regression Methods* (2nd ed.). Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons.

Salvadori, G., C. D. Michele, N. T. Kottegoda, and R. Rosso (2007). *Extremes in Nature: An Approach Using Copulas*. Water Science and Technology Library. Dordrecht: Springer.

Schlather, M. and J. A. Tawn (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika 90*(1), 139–156.

Segers, J. (2005). Approximate distributions of clusters of extremes. *Statistics and Probability Letters 74*, 330–336.

Sempi, C. (2003). *Copulæ and Their Uses*, Chapter 6, pp. 73–85.

Sim, D.-G., O.-K. Kwon, and R.-H. Park (1999, Mar). Object matching algorithms using robust hausdorff distance measures. *Image Processing, IEEE Transactions on 8*(3), 425–429.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

Smith, A. B. and R. W. Katz (2013). Us billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Natural Hazards 67*(2), 387–410.

Smith, R. L. (1992). The extremal index for a markov chain. *Journal of Applied Probability 29*(1), 37–45.

Smith, R. L., J. A. Tawn, and S. G. Coles (1997). Markov chain models for threshold exceedances. *Biometrika 84*(2), 249–268.

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology 15*(2), 201–292.

Sun, W. and Y. Yuan (2006). Trust-region methods for constrained problems. In *Optimization Theory and Methods*, Volume 1 of *Springer Optimization and Its Applications*, pp. 561–595. Springer US.

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika 75*(3), 397–415.

Tiago de Oliveira, J. (1962). Structure theory of bivariate extremes: Extension. *Estudos de Matematica, Estatistica, e Economicos 7*, 165–195.

Tiago de Oliveira, J. (1974). Regression in the nondifferentiable bivariate extreme models. *Journal of the American Statistical Association 69*(347), 816–818.

Tiago de Oliveira, J. (1982). *Bivariate Extremes: Models and Statistical Decision.* North Carolina, USA: Center for Stochastic Processes, University of North Carolina, Chapel Hill. Technical report No. 14.

Tiago de Oliveira, J. (1984). Bivariate models for extremes; Statistical decision. In *Statistical Extremes and Applications*, pp. 131–153. Dordrecht: Reidel.

Valberg, M. (2010). Aspects of Applying the ACER method to Ocean Wave Data. Master's thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Vanmarcke, E. H. (1975). On the distribution of the first-passage time for normal stationary random processes. *Journal of Applied Mechanics 42*(1), 215–220.

von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Revue de l'Union Interbalkanique 1*, 1–20.

Waal, D. and P. Gelder (2005). Modelling of extreme wave heights and periods through copulas. *Extremes 8*, 345–356.

WAFO-group (2000). *WAFO - A Matlab Toolbox for Analysis of Random Waves and Loads - A Tutorial.* Lund, Sweden: Math. Stat., Center for Math. Sci., Lund University.

Waltz, R., J. Morales, J. Nocedal, and D. Orban (2006). An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical Programming 107*(3), 391–408.

Watson, G. S. (1954). Extreme values in samples from m-dependent stationary stochastic processes. *The Annals of Mathematical Statistics 25*(4), 798–800.

Wu, S. and F. Wu (1991). Computation of a trust region step. *Acta Mathematicae Applicatae Sinica 7*(4), 354–362.

Yue, S. (2000). The Gumbel mixed model applied to storm frequency analysis. *Water Resources Management 14*, 377–389.

Yue, S. (2001a). A bivariate extreme value distribution applied to flood frequency analysis. *Nordic Hydrology 32*(1), 49–64.

Yue, S. (2001b). The Gumbel logistic model for representing a multivariate storm event. *Advances in Water Resources 24*(2), 179–185.

Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (1999). The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology 226*(1–2), 88–100.

Yue, S. and C. Y. Wang (2004). A comparison of two bivariate extreme value distributions. *Stochastic Environmental Research and Risk Assessment 18*, 61–66.

Yun, S. (1998). The extremal index of a higher-order stationary Markov chain. *Annals of Applied Probability 8*, 408–437.

Yun, S. (2000). The distribution of cluster functionals of extreme events in a d'th-order Markov chain. *Journal of Applied Probability 37*, 29–44.

Zachary, S., G. Feld, G. Ward, and J. Wolfram (1998). Multivariate extrapolation in the offshore environment. *Applied Ocean Research 20*(5), 273–295.

Previous PhD theses published at the Department of Marine Technology (earlier: Faculty of Marine Technology) NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Report No.	Author	Title
	Kavlie, Dag	Optimization of Plane Elastic Grillages, 1967
	Hansen, Hans R.	Man-Machine Communication and Data-Storage Methods in Ship Structural Design, 1971
	Gisvold, Kaare M.	A Method for non-linear mixed-integer programming and its Application to Design Problems, 1971
	Lund, Sverre	Tanker Frame Optimalization by means of SUMT-Transformation and Behaviour Models, 1971
	Vinje, Tor	On Vibration of Spherical Shells Interacting with Fluid, 1972
	Lorentz, Jan D.	Tank Arrangement for Crude Oil Carriers in Accordance with the new Anti-Pollution Regulations, 1975
	Carlsen, Carl A.	Computer-Aided Design of Tanker Structures, 1975
	Larsen, Carl M.	Static and Dynamic Analysis of Offshore Pipelines during Installation, 1976
UR-79-01	Brigt Hatlestad, MK	The finite element method used in a fatigue evaluation of fixed offshore platforms. (Dr.Ing. Thesis)
UR-79-02	Erik Pettersen, MK	Analysis and design of cellular structures. (Dr.Ing. Thesis)
UR-79-03	Sverre Valsgård, MK	Finite difference and finite element methods applied to nonlinear analysis of plated structures. (Dr.Ing. Thesis)
UR-79-04	Nils T. Nordsve, MK	Finite element collapse analysis of structural members considering imperfections and stresses due to fabrication. (Dr.Ing. Thesis)
UR-79-05	Ivar J. Fylling, MK	Analysis of towline forces in ocean towing systems. (Dr.Ing. Thesis)
UR-80-06	Nils Sandsmark, MM	Analysis of Stationary and Transient Heat Conduction by the Use of the Finite Element Method. (Dr.Ing. Thesis)
UR-80-09	Sverre Haver, MK	Analysis of uncertainties related to the stochastic modeling of ocean waves. (Dr.Ing. Thesis)
UR-81-15	Odland, Jonas	On the Strength of welded Ring stiffened cylindrical Shells primarily subjected to axial Compression
UR-82-17	Engesvik, Knut	Analysis of Uncertainties in the fatigue Capacity of Welded Joints
UR-82-18	Rye, Henrik	Ocean wave groups
UR-83-30	Eide, Oddvar Inge	On Cumulative Fatigue Damage in Steel Welded Joints
UR-83-33	Mo, Olav	Stochastic Time Domain Analysis of Slender Offshore Structures
UR-83-34	Amdahl, Jørgen	Energy absorption in Ship-platform impacts
UR-84-37	Mørch, Morten	Motions and mooring forces of semi submersibles as determined by full-scale measurements and theoretical analysis

Report No.	Author	Title
UR-84-38	Soares, C. Guedes	Probabilistic models for load effects in ship structures
UR-84-39	Aarsnes, Jan V.	Current forces on ships
UR-84-40	Czujko, Jerzy	Collapse Analysis of Plates subjected to Biaxial Compression and Lateral Load
UR-85-46	Alf G. Engseth, MK	Finite element collapse analysis of tubular steel offshore structures. (Dr.Ing. Thesis)
UR-86-47	Dengody Sheshappa, MP	A Computer Design Model for Optimizing Fishing Vessel Designs Based on Techno-Economic Analysis. (Dr.Ing. Thesis)
UR-86-48	Vidar Aanesland, MH	A Theoretical and Numerical Study of Ship Wave Resistance. (Dr.Ing. Thesis)
UR-86-49	Heinz-Joachim Wessel, MK	Fracture Mechanics Analysis of Crack Growth in Plate Girders. (Dr.Ing. Thesis)
UR-86-50	Jon Taby, MK	Ultimate and Post-ultimate Strength of Dented Tubular Members. (Dr.Ing. Thesis)
UR-86-51	Walter Lian, MH	A Numerical Study of Two-Dimensional Separated Flow Past Bluff Bodies at Moderate KC-Numbers. (Dr.Ing. Thesis)
UR-86-52	Bjørn Sortland, MH	Force Measurements in Oscillating Flow on Ship Sections and Circular Cylinders in a U-Tube Water Tank. (Dr.Ing. Thesis)
UR-86-53	Kurt Strand, MM	A System Dynamic Approach to One-dimensional Fluid Flow. (Dr.Ing. Thesis)
UR-86-54	Arne Edvin Løken, MH	Three Dimensional Second Order Hydrodynamic Effects on Ocean Structures in Waves. (Dr.Ing. Thesis)
UR-86-55	Sigurd Falch, MH	A Numerical Study of Slamming of Two-Dimensional Bodies. (Dr.Ing. Thesis)
UR-87-56	Arne Braathen, MH	Application of a Vortex Tracking Method to the Prediction of Roll Damping of a Two-Dimension Floating Body. (Dr.Ing. Thesis)
UR-87-57	Bernt Leira, MK	Gaussian Vector Processes for Reliability Analysis involving Wave-Induced Load Effects. (Dr.Ing. Thesis)
UR-87-58	Magnus Småvik, MM	Thermal Load and Process Characteristics in a Two-Stroke Diesel Engine with Thermal Barriers (in Norwegian). (Dr.Ing. Thesis)
MTA-88-59	Bernt Arild Bremdal, MP	An Investigation of Marine Installation Processes – A Knowledge-Based Planning Approach. (Dr.Ing. Thesis)
MTA-88-60	Xu Jun, MK	Non-linear Dynamic Analysis of Space-framed Offshore Structures. (Dr.Ing. Thesis)
MTA-89-61	Gang Miao, MH	Hydrodynamic Forces and Dynamic Responses of Circular Cylinders in Wave Zones. (Dr.Ing. Thesis)
MTA-89-62	Martin Greenhow, MH	Linear and Non-Linear Studies of Waves and Floating Bodies. Part I and Part II. (Dr.Techn. Thesis)
MTA-89-63	Chang Li, MH	Force Coefficients of Spheres and Cubes in Oscillatory Flow with and without Current. (Dr.Ing. Thesis
MTA-89-64	Hu Ying, MP	A Study of Marketing and Design in Development of Marine Transport Systems. (Dr.Ing. Thesis)

Report No.	Author	Title
MTA-89-65	Arild Jæger, MH	Seakeeping, Dynamic Stability and Performance of a Wedge Shaped Planing Hull. (Dr.Ing. Thesis)
MTA-89-66	Chan Siu Hung, MM	The dynamic characteristics of tilting-pad bearings
MTA-89-67	Kim Wikstrøm, MP	Analysis av projekteringen for ett offshore projekt. (Licenciat-avhandling)
MTA-89-68	Jiao Guoyang, MK	Reliability Analysis of Crack Growth under Random Loading, considering Model Updating. (Dr.Ing. Thesis)
MTA-89-69	Arnt Olufsen, MK	Uncertainty and Reliability Analysis of Fixed Offshore Structures. (Dr.Ing. Thesis)
MTA-89-70	Wu Yu-Lin, MR	System Reliability Analyses of Offshore Structures using improved Truss and Beam Models. (Dr.Ing. Thesis)
MTA-90-71	Jan Roger Hoff, MH	Three-dimensional Green function of a vessel with forward speed in waves. (Dr.Ing. Thesis)
MTA-90-72	Rong Zhao, MH	Slow-Drift Motions of a Moored Two-Dimensional Body in Irregular Waves. (Dr.Ing. Thesis)
MTA-90-73	Atle Minsaas, MP	Economical Risk Analysis. (Dr.Ing. Thesis)
MTA-90-74	Knut-Ariel Farnes, MK	Long-term Statistics of Response in Non-linear Marine Structures. (Dr.Ing. Thesis)
MTA-90-75	Torbjørn Sotberg, MK	Application of Reliability Methods for Safety Assessment of Submarine Pipelines. (Dr.Ing. Thesis)
MTA-90-76	Zeuthen, Steffen, MP	SEAMAID. A computational model of the design process in a constraint-based logic programming environment. An example from the offshore domain. (Dr.Ing. Thesis)
MTA-91-77	Haagensen, Sven, MM	Fuel Dependant Cyclic Variability in a Spark Ignition Engine – An Optical Approach. (Dr.Ing. Thesis)
MTA-91-78	Løland, Geir, MH	Current forces on and flow through fish farms. (Dr.Ing. Thesis)
MTA-91-79	Hoen, Christopher, MK	System Identification of Structures Excited by Stochastic Load Processes. (Dr.Ing. Thesis)
MTA-91-80	Haugen, Stein, MK	Probabilistic Evaluation of Frequency of Collision between Ships and Offshore Platforms. (Dr.Ing. Thesis)
MTA-91-81	Sødahl, Nils, MK	Methods for Design and Analysis of Flexible Risers. (Dr.Ing. Thesis)
MTA-91-82	Ormberg, Harald, MK	Non-linear Response Analysis of Floating Fish Farm Systems. (Dr.Ing. Thesis)
MTA-91-83	Marley, Mark J., MK	Time Variant Reliability under Fatigue Degradation. (Dr.Ing. Thesis)
MTA-91-84	Krokstad, Jørgen R., MH	Second-order Loads in Multidirectional Seas. (Dr.Ing. Thesis)
MTA-91-85	Molteberg, Gunnar A., MM	The Application of System Identification Techniques to Performance Monitoring of Four Stroke Turbocharged Diesel Engines. (Dr.Ing. Thesis)
MTA-92-86	Mørch, Hans Jørgen Bjelke, MH	Aspects of Hydrofoil Design: with Emphasis on Hydrofoil Interaction in Calm Water. (Dr.Ing. Thesis)

Report No.	Author	Title
MTA-92-87	Chan Siu Hung, MM	Nonlinear Analysis of Rotordynamic Instabilities in Highspeed Turbomachinery. (Dr.Ing. Thesis)
MTA-92-88	Bessason, Bjarni, MK	Assessment of Earthquake Loading and Response of Seismically Isolated Bridges. (Dr.Ing. Thesis)
MTA-92-89	Langli, Geir, MP	Improving Operational Safety through exploitation of Design Knowledge – an investigation of offshore platform safety. (Dr.Ing. Thesis)
MTA-92-90	Sævik, Svein, MK	On Stresses and Fatigue in Flexible Pipes. (Dr.Ing. Thesis)
MTA-92-91	Ask, Tor Ø., MM	Ignition and Flame Growth in Lean Gas-Air Mixtures. An Experimental Study with a Schlieren System. (Dr.Ing. Thesis)
MTA-86-92	Hessen, Gunnar, MK	Fracture Mechanics Analysis of Stiffened Tubular Members. (Dr.Ing. Thesis)
MTA-93-93	Steinebach, Christian, MM	Knowledge Based Systems for Diagnosis of Rotating Machinery. (Dr.Ing. Thesis)
MTA-93-94	Dalane, Jan Inge, MK	System Reliability in Design and Maintenance of Fixed Offshore Structures. (Dr.Ing. Thesis)
MTA-93-95	Steen, Sverre, MH	Cobblestone Effect on SES. (Dr.Ing. Thesis)
MTA-93-96	Karunakaran, Daniel, MK	Nonlinear Dynamic Response and Reliability Analysis of Drag-dominated Offshore Platforms. (Dr.Ing. Thesis)
MTA-93-97	Hagen, Arnulf, MP	The Framework of a Design Process Language. (Dr.Ing. Thesis)
MTA-93-98	Nordrik, Rune, MM	Investigation of Spark Ignition and Autoignition in Methane and Air Using Computational Fluid Dynamics and Chemical Reaction Kinetics. A Numerical Study of Ignition Processes in Internal Combustion Engines. (Dr.Ing. Thesis)
MTA-94-99	Passano, Elizabeth, MK	Efficient Analysis of Nonlinear Slender Marine Structures. (Dr.Ing. Thesis)
MTA-94-100	Kvålsvold, Jan, MH	Hydroelastic Modelling of Wetdeck Slamming on Multihull Vessels. (Dr.Ing. Thesis)
MTA-94-102	Bech, Sidsel M., MK	Experimental and Numerical Determination of Stiffness and Strength of GRP/PVC Sandwich Structures. (Dr.Ing. Thesis)
MTA-95-103	Paulsen, Hallvard, MM	A Study of Transient Jet and Spray using a Schlieren Method and Digital Image Processing. (Dr.Ing. Thesis)
MTA-95-104	Hovde, Geir Olav, MK	Fatigue and Overload Reliability of Offshore Structural Systems, Considering the Effect of Inspection and Repair. (Dr.Ing. Thesis)
MTA-95-105	Wang, Xiaozhi, MK	Reliability Analysis of Production Ships with Emphasis on Load Combination and Ultimate Strength. (Dr.Ing. Thesis)
MTA-95-106	Ulstein, Tore, MH	Nonlinear Effects of a Flexible Stern Seal Bag on Cobblestone Oscillations of an SES. (Dr.Ing. Thesis)
MTA-95-107	Solaas, Frøydis, MH	Analytical and Numerical Studies of Sloshing in Tanks. (Dr.Ing. Thesis)

Report No.	Author	Title
MTA-95-108	Hellan, Øyvind, MK	Nonlinear Pushover and Cyclic Analyses in Ultimate Limit State Design and Reassessment of Tubular Steel Offshore Structures. (Dr.Ing. Thesis)
MTA-95-109	Hermundstad, Ole A., MK	Theoretical and Experimental Hydroelastic Analysis of High Speed Vessels. (Dr.Ing. Thesis)
MTA-96-110	Bratland, Anne K., MH	Wave-Current Interaction Effects on Large-Volume Bodies in Water of Finite Depth. (Dr.Ing. Thesis)
MTA-96-111	Herfjord, Kjell, MH	A Study of Two-dimensional Separated Flow by a Combination of the Finite Element Method and Navier-Stokes Equations. (Dr.Ing. Thesis)
MTA-96-112	AEsøy, Vilmar, MM	Hot Surface Assisted Compression Ignition in a Direct Injection Natural Gas Engine. (Dr.Ing. Thesis)
MTA-96-113	Eknes, Monika L., MK	Escalation Scenarios Initiated by Gas Explosions on Offshore Installations. (Dr.Ing. Thesis)
MTA-96-114	Erikstad, Stein O., MP	A Decision Support Model for Preliminary Ship Design. (Dr.Ing. Thesis)
MTA-96-115	Pedersen, Egil, MH	A Nautical Study of Towed Marine Seismic Streamer Cable Configurations. (Dr.Ing. Thesis)
MTA-97-116	Moksnes, Paul O., MM	Modelling Two-Phase Thermo-Fluid Systems Using Bond Graphs. (Dr.Ing. Thesis)
MTA-97-117	Halse, Karl H., MK	On Vortex Shedding and Prediction of Vortex-Induced Vibrations of Circular Cylinders. (Dr.Ing. Thesis)
MTA-97-118	Igland, Ragnar T., MK	Reliability Analysis of Pipelines during Laying, considering Ultimate Strength under Combined Loads. (Dr.Ing. Thesis)
MTA-97-119	Pedersen, Hans-P., MP	Levendefiskteknologi for fiskefartøy. (Dr.Ing. Thesis)
MTA-98-120	Vikestad, Kyrre, MK	Multi-Frequency Response of a Cylinder Subjected to Vortex Shedding and Support Motions. (Dr.Ing. Thesis)
MTA-98-121	Azadi, Mohammad R. E., MK	Analysis of Static and Dynamic Pile-Soil-Jacket Behaviour. (Dr.Ing. Thesis)
MTA-98-122	Ulltang, Terje, MP	A Communication Model for Product Information. (Dr.Ing. Thesis)
MTA-98-123	Torbergsen, Erik, MM	Impeller/Diffuser Interaction Forces in Centrifugal Pumps. (Dr.Ing. Thesis)
MTA-98-124	Hansen, Edmond, MH	A Discrete Element Model to Study Marginal Ice Zone Dynamics and the Behaviour of Vessels Moored in Broken Ice. (Dr.Ing. Thesis)
MTA-98-125	Videiro, Paulo M., MK	Reliability Based Design of Marine Structures. (Dr.Ing. Thesis)
MTA-99-126	Mainçon, Philippe, MK	Fatigue Reliability of Long Welds Application to Titanium Risers. (Dr.Ing. Thesis)
MTA-99-127	Haugen, Elin M., MH	Hydroelastic Analysis of Slamming on Stiffened Plates with Application to Catamaran Wetdecks. (Dr.Ing. Thesis)

Report No.	Author	Title
MTA-99-128	Langhelle, Nina K., MK	Experimental Validation and Calibration of Nonlinear Finite Element Models for Use in Design of Aluminium Structures Exposed to Fire. (Dr.Ing. Thesis)
MTA-99-129	Berstad, Are J., MK	Calculation of Fatigue Damage in Ship Structures. (Dr.Ing. Thesis)
MTA-99-130	Andersen, Trond M., MM	Short Term Maintenance Planning. (Dr.Ing. Thesis)
MTA-99-131	Tveiten, Bård Wathne, MK	Fatigue Assessment of Welded Aluminium Ship Details. (Dr.Ing. Thesis)
MTA-99-132	Søreide, Fredrik, MP	Applications of underwater technology in deep water archaeology. Principles and practice. (Dr.Ing. Thesis)
MTA-99-133	Tønnessen, Rune, MH	A Finite Element Method Applied to Unsteady Viscous Flow Around 2D Blunt Bodies With Sharp Corners. (Dr.Ing. Thesis)
MTA-99-134	Elvekrok, Dag R., MP	Engineering Integration in Field Development Projects in the Norwegian Oil and Gas Industry. The Supplier Management of Norne. (Dr.Ing. Thesis)
MTA-99-135	Fagerholt, Kjetil, MP	Optimeringsbaserte Metoder for Ruteplanlegging innen skipsfart. (Dr.Ing. Thesis)
MTA-99-136	Bysveen, Marie, MM	Visualization in Two Directions on a Dynamic Combustion Rig for Studies of Fuel Quality. (Dr.Ing. Thesis)
MTA-2000-137	Storteig, Eskild, MM	Dynamic characteristics and leakage performance of liquid annular seals in centrifugal pumps. (Dr.Ing. Thesis)
MTA-2000-138	Sagli, Gro, MK	Model uncertainty and simplified estimates of long term extremes of hull girder loads in ships. (Dr.Ing. Thesis)
MTA-2000-139	Tronstad, Harald, MK	Nonlinear analysis and design of cable net structures like fishing gear based on the finite element method. (Dr.Ing. Thesis)
MTA-2000-140	Kroneberg, André, MP	Innovation in shipping by using scenarios. (Dr.Ing. Thesis)
MTA-2000-141	Haslum, Herbjørn Alf, MH	Simplified methods applied to nonlinear motion of spar platforms. (Dr.Ing. Thesis)
MTA-2001-142	Samdal, Ole Johan, MM	Modelling of Degradation Mechanisms and Stressor Interaction on Static Mechanical Equipment Residual Lifetime. (Dr.Ing. Thesis)
MTA-2001-143	Baarholm, Rolf Jarle, MH	Theoretical and experimental studies of wave impact underneath decks of offshore platforms. (Dr.Ing. Thesis)
MTA-2001-144	Wang, Lihua, MK	Probabilistic Analysis of Nonlinear Wave-induced Loads on Ships. (Dr.Ing. Thesis)
MTA-2001-145	Kristensen, Odd H. Holt, MK	Ultimate Capacity of Aluminium Plates under Multiple Loads, Considering HAZ Properties. (Dr.Ing. Thesis)
MTA-2001-146	Greco, Marilena, MH	A Two-Dimensional Study of Green-Water Loading. (Dr.Ing. Thesis)

Report No.	Author	Title
MTA-2001-147	Heggelund, Svein E., MK	Calculation of Global Design Loads and Load Effects in Large High Speed Catamarans. (Dr.Ing. Thesis)
MTA-2001-148	Babalola, Olusegun T., MK	Fatigue Strength of Titanium Risers – Defect Sensitivity. (Dr.Ing. Thesis)
MTA-2001-149	Mohammed, Abuu K., MK	Nonlinear Shell Finite Elements for Ultimate Strength and Collapse Analysis of Ship Structures. (Dr.Ing. Thesis)
MTA-2002-150	Holmedal, Lars E., MH	Wave-current interactions in the vicinity of the sea bed. (Dr.Ing. Thesis)
MTA-2002-151	Rognebakke, Olav F., MH	Sloshing in rectangular tanks and interaction with ship motions. (Dr.Ing. Thesis)
MTA-2002-152	Lader, Pål Furset, MH	Geometry and Kinematics of Breaking Waves. (Dr.Ing. Thesis)
MTA-2002-153	Yang, Qinzheng, MH	Wash and wave resistance of ships in finite water depth. (Dr.Ing. Thesis)
MTA-2002-154	Melhus, Øyvin, MM	Utilization of VOC in Diesel Engines. Ignition and combustion of VOC released by crude oil tankers. (Dr.Ing. Thesis)
MTA-2002-155	Ronæss, Marit, MH	Wave Induced Motions of Two Ships Advancing on Parallel Course. (Dr.Ing. Thesis)
MTA-2002-156	Økland, Ole D., MK	Numerical and experimental investigation of whipping in twin hull vessels exposed to severe wet deck slamming. (Dr.Ing. Thesis)
MTA-2002-157	Ge, Chunhua, MK	Global Hydroelastic Response of Catamarans due to Wet Deck Slamming. (Dr.Ing. Thesis)
MTA-2002-158	Byklum, Eirik, MK	Nonlinear Shell Finite Elements for Ultimate Strength and Collapse Analysis of Ship Structures. (Dr.Ing. Thesis)
IMT-2003-1	Chen, Haibo, MK	Probabilistic Evaluation of FPSO-Tanker Collision in Tandem Offloading Operation. (Dr.Ing. Thesis)
IMT-2003-2	Skaugset, Kjetil Bjørn, MK	On the Suppression of Vortex Induced Vibrations of Circular Cylinders by Radial Water Jets. (Dr.Ing. Thesis)
IMT-2003-3	Chezhian, Muthu	Three-Dimensional Analysis of Slamming. (Dr.Ing. Thesis)
IMT-2003-4	Buhaug, Øyvind	Deposit Formation on Cylinder Liner Surfaces in Medium Speed Engines. (Dr.Ing. Thesis)
IMT-2003-5	Tregde, Vidar	Aspects of Ship Design: Optimization of Aft Hull with Inverse Geometry Design. (Dr.Ing. Thesis)
IMT-2003-6	Wist, Hanne Therese	Statistical Properties of Successive Ocean Wave Parameters. (Dr.Ing. Thesis)
IMT-2004-7	Ransau, Samuel	Numerical Methods for Flows with Evolving Interfaces. (Dr.Ing. Thesis)
IMT-2004-8	Soma, Torkel	Blue-Chip or Sub-Standard. A data interrogation approach of identity safety characteristics of shipping organization. (Dr.Ing. Thesis)
IMT-2004-9	Ersdal, Svein	An experimental study of hydrodynamic forces on cylinders and cables in near axial flow. (Dr.Ing. Thesis)

Report No.	Author	Title
IMT-2005-10	Brodtkorb, Per Andreas	The Probability of Occurrence of Dangerous Wave Situations at Sea. (Dr.Ing. Thesis)
IMT-2005-11	Yttervik, Rune	Ocean current variability in relation to offshore engineering. (Dr.Ing. Thesis)
IMT-2005-12	Fredheim, Arne	Current Forces on Net-Structures. (Dr.Ing. Thesis)
IMT-2005-13	Heggernes, Kjetil	Flow around marine structures. (Dr.Ing. Thesis
IMT-2005-14	Fouques, Sebastien	Lagrangian Modelling of Ocean Surface Waves and Synthetic Aperture Radar Wave Measurements. (Dr.Ing. Thesis)
IMT-2006-15	Holm, Håvard	Numerical calculation of viscous free surface flow around marine structures. (Dr.Ing. Thesis)
IMT-2006-16	Bjørheim, Lars G.	Failure Assessment of Long Through Thickness Fatigue Cracks in Ship Hulls. (Dr.Ing. Thesis)
IMT-2006-17	Hansson, Lisbeth	Safety Management for Prevention of Occupational Accidents. (Dr.Ing. Thesis)
IMT-2006-18	Zhu, Xinying	Application of the CIP Method to Strongly Nonlinear Wave-Body Interaction Problems. (Dr.Ing. Thesis)
IMT-2006-19	Reite, Karl Johan	Modelling and Control of Trawl Systems. (Dr.Ing. Thesis)
IMT-2006-20	Smogeli, Øyvind Notland	Control of Marine Propellers. From Normal to Extreme Conditions. (Dr.Ing. Thesis)
IMT-2007-21	Storhaug, Gaute	Experimental Investigation of Wave Induced Vibrations and Their Effect on the Fatigue Loading of Ships. (Dr.Ing. Thesis)
IMT-2007-22	Sun, Hui	A Boundary Element Method Applied to Strongly Nonlinear Wave-Body Interaction Problems. (PhD Thesis, CeSOS)
IMT-2007-23	Rustad, Anne Marthine	Modelling and Control of Top Tensioned Risers. (PhD Thesis, CeSOS)
IMT-2007-24	Johansen, Vegar	Modelling flexible slender system for real-time simulations and control applications
IMT-2007-25	Wroldsen, Anders Sunde	Modelling and control of tensegrity structures. (PhD Thesis, CeSOS)
IMT-2007-26	Aronsen, Kristoffer Høye	An experimental investigation of in-line and combined inline and cross flow vortex induced vibrations. (Dr. avhandling, IMT)
IMT-2007-27	Gao, Zhen	Stochastic Response Analysis of Mooring Systems with Emphasis on Frequency-domain Analysis of Fatigue due to Wide-band Response Processes (PhD Thesis, CeSOS)
IMT-2007-28	Thorstensen, Tom Anders	Lifetime Profit Modelling of Ageing Systems Utilizing Information about Technical Condition. (Dr.ing. thesis, IMT)
IMT-2008-29	Berntsen, Per Ivar B.	Structural Reliability Based Position Mooring. (PhD-Thesis, IMT)
IMT-2008-30	Ye, Naiquan	Fatigue Assessment of Aluminium Welded Box-stiffener Joints in Ships (Dr.ing. thesis, IMT)
IMT-2008-31	Radan, Damir	Integrated Control of Marine Electrical Power Systems. (PhD-Thesis, IMT)

Report No.	Author	Title
IMT-2008-32	Thomassen, Paul	Methods for Dynamic Response Analysis and Fatigue Life Estimation of Floating Fish Cages. (Dr.ing. thesis, IMT)
IMT-2008-33	Pákozdi, Csaba	A Smoothed Particle Hydrodynamics Study of Two-dimensional Nonlinear Sloshing in Rectangular Tanks. (Dr.ing.thesis, IMT/ CeSOS)
IMT-2008-34	Grytøyr, Guttorm	A Higher-Order Boundary Element Method and Applications to Marine Hydrodynamics. (Dr.ing.thesis, IMT)
IMT-2008-35	Drummen, Ingo	Experimental and Numerical Investigation of Nonlinear Wave-Induced Load Effects in Containerships considering Hydroelasticity. (PhD thesis, CeSOS)
IMT-2008-36	Skejic, Renato	Maneuvering and Seakeeping of a Singel Ship and of Two Ships in Interaction. (PhD-Thesis, CeSOS)
IMT-2008-37	Harlem, Alf	An Age-Based Replacement Model for Repairable Systems with Attention to High-Speed Marine Diesel Engines. (PhD-Thesis, IMT)
IMT-2008-38	Alsos, Hagbart S.	Ship Grounding. Analysis of Ductile Fracture, Bottom Damage and Hull Girder Response. (PhD-thesis, IMT)
IMT-2008-39	Graczyk, Mateusz	Experimental Investigation of Sloshing Loading and Load Effects in Membrane LNG Tanks Subjected to Random Excitation. (PhD-thesis, CeSOS)
IMT-2008-40	Taghipour, Reza	Efficient Prediction of Dynamic Response for Flexible amd Multi-body Marine Structures. (PhD-thesis, CeSOS)
IMT-2008-41	Ruth, Eivind	Propulsion control and thrust allocation on marine vessels. (PhD thesis, CeSOS)
IMT-2008-42	Nystad, Bent Helge	Technical Condition Indexes and Remaining Useful Life of Aggregated Systems. PhD thesis, IMT
IMT-2008-43	Soni, Prashant Kumar	Hydrodynamic Coefficients for Vortex Induced Vibrations of Flexible Beams. PhD thesis, CeSOS
IMT-2009-43	Amlashi, Hadi K.K.	Ultimate Strength and Reliability-based Design of Ship Hulls with Emphasis on Combined Global and Local Loads. PhD Thesis, IMT
IMT-2009-44	Pedersen, Tom Arne	Bond Graph Modelling of Marine Power Systems. PhD Thesis, IMT
IMT-2009-45	Kristiansen, Trygve	Two-Dimensional Numerical and Experimental Studies of Piston-Mode Resonance. PhD thesis, CeSOS
IMT-2009-46	Ong, Muk Chen	Applications of a Standard High Reynolds Number Model and a Stochastic Scour Prediction Model for Marine Structures. PhD-thesis, IMT
IMT-2009-47	Hong, Lin	Simplified Analysis and Design of Ships subjected to Collision and Grounding. PhD-thesis, IMT
IMT-2009-48	Koushan, Kamran	Vortex Induced Vibrations of Free Span Pipelines, PhD thesis, IMT
IMT-2009-49	Korsvik, Jarl Eirik	Heuristic Methods for Ship Routing and Scheduling. PhD-thesis, IMT

Report No.	Author	Title
IMT-2009-50	Lee, Jihoon	Experimental Investigation and Numerical in Analyzing the Ocean Current Displacement of Longlines. Ph.d.-Thesis, IMT.
IMT-2009-51	Vestbøstad, Tone Gran	A Numerical Study of Wave-in-Deck Impact usin a Two-Dimensional Constrained Interpolation Profile Method, Ph.d.thesis, CeSOS.
IMT-2009-52	Bruun, Kristine	Bond Graph Modelling of Fuel Cells for Marine Power Plants. Ph.d.-thesis, IMT
IMT-2009-53	Holstad, Anders	Numerical Investigation of Turbulence in a Sekwed Three-Dimensional Channel Flow, Ph.d.-thesis, IMT.
IMT-2009-54	Ayala-Uraga, Efren	Reliability-Based Assessment of Deteriorating Ship-shaped Offshore Structures, Ph.d.-thesis, IMT
IMT-2009-55	Kong, Xiangjun	A Numerical Study of a Damaged Ship in Beam Sea Waves. Ph.d.-thesis, IMT/CeSOS.
IMT-2010-56	Kristiansen, David	Wave Induced Effects on Floaters of Aquaculture Plants, Ph.d.-thesis, CeSOS.
IMT-2010-57	Ludvigsen, Martin	An ROV-Toolbox for Optical and Acoustic Scientific Seabed Investigation. Ph.d.-thesis IMT.
IMT-2010-58	Hals, Jørgen	Modelling and Phase Control of Wave-Energy Converters. Ph.d.thesis, CeSOS.
IMT-2010-59	Shu, Zhi	Uncertainty Assessment of Wave Loads and Ultimate Strength of Tankers and Bulk Carriers in a Reliability Framework. Ph.d. Thesis, IMT/ CeSOS
IMT-2010-60	Shao, Yanlin	Numerical Potential-Flow Studies on Weakly-Nonlinear Wave-Body Interactions with/without Small Forward Speed, Ph.d.thesis, CeSOS.
IMT-2010-61	Califano, Andrea	Dynamic Loads on Marine Propellers due to Intermittent Ventilation. Ph.d.thesis, IMT.
IMT-2010-62	El Khoury, George	Numerical Simulations of Massively Separated Turbulent Flows, Ph.d.-thesis, IMT
IMT-2010-63	Seim, Knut Sponheim	Mixing Process in Dense Overflows with Emphasis on the Faroe Bank Channel Overflow. Ph.d.thesis, IMT
IMT-2010-64	Jia, Huirong	Structural Analysis of Intect and Damaged Ships in a Collission Risk Analysis Perspective. Ph.d.thesis CeSoS.
IMT-2010-65	Jiao, Linlin	Wave-Induced Effects on a Pontoon-type Very Large Floating Structures (VLFS). Ph.D.-thesis, CeSOS.
IMT-2010-66	Abrahamsen, Bjørn Christian	Sloshing Induced Tank Roof with Entrapped Air Pocket. Ph.d.thesis, CeSOS.
IMT-2011-67	Karimirad, Madjid	Stochastic Dynamic Response Analysis of Spar-Type Wind Turbines with Catenary or Taut Mooring Systems. Ph.d.-thesis, CeSOS.
IMT-2011-68	Erlend Meland	Condition Monitoring of Safety Critical Valves. Ph.d.-thesis, IMT.
IMT-2011-69	Yang, Limin	Stochastic Dynamic System Analysis of Wave Energy Converter with Hydraulic Power Take-Off, with Particular Reference to Wear Damage Analysis, Ph.d. Thesis, CeSOS.

Report No.	Author	Title
IMT-2011-70	Visscher, Jan	Application of Particla Image Velocimetry on Turbulent Marine Flows, Ph.d.Thesis, IMT.
IMT-2011-71	Su, Biao	Numerical Predictions of Global and Local Ice Loads on Ships. Ph.d.Thesis, CeSOS.
IMT-2011-72	Liu, Zhenhui	Analytical and Numerical Analysis of Iceberg Collision with Ship Structures. Ph.d.Thesis, IMT.
IMT-2011-73	Aarsæther, Karl Gunnar	Modeling and Analysis of Ship Traffic by Observation and Numerical Simulation. Ph.d.Thesis, IMT.
IMT-2011-74	Wu, Jie	Hydrodynamic Force Identification from Stochastic Vortex Induced Vibration Experiments with Slender Beams. Ph.d.Thesis, IMT.
IMT-2011-75	Amini, Hamid	Azimuth Propulsors in Off-design Conditions. Ph.d.Thesis, IMT.
IMT-2011-76	Nguyen, Tan-Hoi	Toward a System of Real-Time Prediction and Monitoring of Bottom Damage Conditions During Ship Grounding. Ph.d.thesis, IMT.
IMT-2011-77	Tavakoli, Mohammad T.	Assessment of Oil Spill in Ship Collision and Grounding, Ph.d.thesis, IMT.
IMT-2011-78	Guo, Bingjie	Numerical and Experimental Investigation of Added Resistance in Waves. Ph.d.Thesis, IMT.
IMT-2011-79	Chen, Qiaofeng	Ultimate Strength of Aluminium Panels, considering HAZ Effects, IMT
IMT-2012-80	Kota, Ravikiran S.	Wave Loads on Decks of Offshore Structures in Random Seas, CeSOS.
IMT-2012-81	Sten, Ronny	Dynamic Simulation of Deep Water Drilling Risers with Heave Compensating System, IMT.
IMT-2012-82	Berle, Øyvind	Risk and resilience in global maritime supply chains, IMT.
IMT-2012-83	Fang, Shaoji	Fault Tolerant Position Mooring Control Based on Structural Reliability, CeSOS.
IMT-2012-84	You, Jikun	Numerical studies on wave forces and moored ship motions in intermediate and shallow water, CeSOS.
IMT-2012-85	Xiang, Xu	Maneuvering of two interacting ships in waves, CeSOS
IMT-2012-86	Dong, Wenbin	Time-domain fatigue response and reliability analysis of offshore wind turbines with emphasis on welded tubular joints and gear components, CeSOS
IMT-2012-87	Zhu, Suji	Investigation of Wave-Induced Nonlinear Load Effects in Open Ships considering Hull Girder Vibrations in Bending and Torsion, CeSOS
IMT-2012-88	Zhou, Li	Numerical and Experimental Investigation of Station-keeping in Level Ice, CeSOS
IMT-2012-90	Ushakov, Sergey	Particulate matter emission characteristics from diesel enignes operating on conventional and alternative marine fuels, IMT
IMT-2013-1	Yin, Decao	Experimental and Numerical Analysis of Combined In-line and Cross-flow Vortex Induced Vibrations, CeSOS
IMT-2013-2	Kurniawan, Adi	Modelling and geometry optimisation of wave energy converters, CeSOS

Report No.	Author	Title
IMT-2013-3	Al Ryati, Nabil	Technical condition indexes doe auxiliary marine diesel engines, IMT
IMT-2013-4	Firoozkoohi, Reza	Experimental, numerical and analytical investigation of the effect of screens on sloshing, CeSOS
IMT-2013-5	Ommani, Babak	Potential-Flow Predictions of a Semi-Displacement Vessel Including Applications to Calm Water Broaching, CeSOS
IMT-2013-6	Xing, Yihan	Modelling and analysis of the gearbox in a floating spar-type wind turbine, CeSOS
IMT-7-2013	Balland, Océane	Optimization models for reducing air emissions from ships, IMT
IMT-8-2013	Yang, Dan	Transitional wake flow behind an inclined flat plate – Computation and analysis, IMT
IMT-9-2013	Abdillah, Suyuthi	Prediction of Extreme Loads and Fatigue Damage for a Ship Hull due to Ice Action, IMT
IMT-10-2013	Ramìrez, Pedro Agustìn Pèrez	Ageing management and life extension of technical systems. Concepts and methods applied to oil and gas facilities, IMT
IMT-11-2013	Chuang, Zhenju	Experimental and Numerical Investigation of Speed Loss due to Seakeeping and Maneuvering. IMT
IMT-12-2013	Etemaddar, Mahmoud	Load and Response Analysis of Wind Turbines under Atmospheric Icing and Controller System Faults with Emphasis on Spar Type Floating Wind Turbines, IMT
IMT-13-2013	Lindstad, Haakon	Strategies and measures for reducing maritime CO2 emissons, IMT
IMT-14-2013	Haris, Sabril	Damage interaction analysis of ship collisions, IMT
IMT-15-2013	Shainee, Mohamed	Conceptual Design, Numerical and Experimental Investigation of a SPM Cage Concept for Offshore Mariculture, IMT
IMT-16-2013	Gansel, Lars	Flow past porous cylinders and effects of biofouling and fish behavior on the flow in and around Atlantic salmon net cages, IMT
IMT-17-2013	Gaspar, Henrique	Handling Aspects of Complexity in Conceptual Ship Design, IMT
IMT-18-2013	Thys, Maxime	Theoretical and Experimental Investigation of a Free Running Fishing Vessel at Small Frequency of Encounter, CeSOS
IMT-19-2013	Aglen, Ida	VIV in Free Spanning Pipelines, CeSOS
IMT-1-2014	Song, An	Theoretical and experimental studies of wave diffraction and radiation loads on a horizontally submerged perforated plate, CeSOS
IMT-2-2014	Rogne, Øyvind Ygre	Numerical and Experimental Investigation of a Hinged 5-body Wave Energy Converter, CeSOS
IMT-3-2014	Dai, Lijuan	Safe and efficient operation and maintenance of offshore wind farms, IMT
IMT-4-2014	Bachynski, Erin Elizabeth	Design and Dynamic Analysis of Tension Leg Platform Wind Turbines, CeSOS
IMT-5-2014	Wang, Jingbo	Water Entry of Freefall Wedged – Wedge motions and Cavity Dynamics, CeSOS

Report No.	Author	Title
IMT-6-2014	Kim, Ekaterina	Experimental and numerical studies related to the coupled behavior of ice mass and steel structures during accidental collisions, IMT
IMT-7-2014	Tan, Xiang	Numerical investigation of ship's continuous-mode icebreaking in leverl ice, CeSOS
IMT-8-2014	Muliawan, Made Jaya	Design and Analysis of Combined Floating Wave and Wind Power Facilities, with Emphasis on Extreme Load Effects of the Mooring System, CeSOS
IMT-9-2014	Jiang, Zhiyu	Long-term response analysis of wind turbines with an emphasis on fault and shutdown conditions, IMT
IMT-10-2014	Dukan, Fredrik	ROV Motion Control Systems, IMT
IMT-11-2014	Grimsmo, Nils I.	Dynamic simulations of hydraulic cylinder for heave compensation of deep water drilling risers, IMT
IMT-12-2014	Kvittem, Marit I.	Modelling and response analysis for fatigue design of a semisubmersible wind turbine, CeSOS
IMT-13-2014	Akhtar, Juned	The Effects of Human Fatigue on Risk at Sea, IMT
IMT-14-2014	Syahroni, Nur	Fatigue Assessment of Welded Joints Taking into Account Effects of Residual Stress, IMT
IMT-1-2015	Bøckmann, Eirik	Wave Propulsion of ships, IMT
IMT-2-2015	Wang, Kai	Modelling and dynamic analysis of a semi-submersible floating vertical axis wind turbine, CeSOS
IMT-3-2015	Fredriksen, Arnt Gunvald	A numerical and experimental study of a two-dimensional body with moonpool in waves and current
IMT-4-2015	Jose Patricio Gallardo Canabes	Numerical studies of viscous flow around bluff bodies
IMT-5-2015	Vegard Longva	Formulation and application of finite element techniques for slender marine structures subjected to contact interactions
IMT-6-2015	Jacobus De Vaal	Aerodynamic modelling of floating wind turbines
IMT-7-2015	Fachri Nasution	Fatigue Performance of Copper Power Conductors
IMT-8-2015	Oleh Karpa	Development of bivariate extreme value distributions for applications in marine technology