

---

## Preface

This thesis is the culmination of the course *TMA4905 Statistics, Master Thesis*, and has been researched for and written during the course of 20 weeks in the Spring semester of 2011. It completes my fifth and final year of the Master of Technology program at the Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU) in Trondheim.

Theoretical introductions are given to all featured methods, but the style of the text is suited for readers with a basic knowledge of statistical and mathematical concepts.

The statistical software package R has been used for most calculations, but MATLAB has been used for the estimates in the section of the text covering the ACER method.

I would like to thank my supervisor Professor Arvid Næss for helpful advice and general guidance during the semester, Even Haug at the Norwegian Hydrographic Service for providing the data sets and finally the residents of office number 393C for extra-curricular distractions in times of stress.

Trondheim, June 14, 2011  
Morten Skjong



## Abstract

In this Master's thesis we have used data from eight different locations in Norway to estimate the distributions of extreme water levels. The locations used are Oslo, Heimsjø, Honningsvåg, Narvik, Harstad, Tregde, Andenes and Viker, and four different methods have been used for the estimation. We use two established and widely used methods, the *Annual Maxima Method* (AMM) and the *Peaks-over-Threshold* (POT) method. The AMM uses the observed annual maxima from each location, while the POT method uses the exceedances of some high threshold in relation to the data. In addition, we also use two approaches that are less known and less used, the *Revised Joint Probability Method* (RJPM) and the *Average Conditional Exceedance Rate* (ACER) method. In the former, a distribution for extreme surges is found and used together with information from numerical tide predictions. We also take into account the interaction between tides and surges where applicable. The ACER method uses only the series of measured sea levels, but tries to account for dependence by a cascade of conditioning probabilities. We find that for return periods up to 20 years the return levels are usually very similar for all methods except RJPM, which seems to underestimate these values. The data intensive methods of POT, RJPM and ACER seem to agree for the 200 year levels, while particularly the Gumbel method seems to overshoot the levels in comparison to the other methods.



# Contents

<b>1</b>	<b>Data sets</b>	<b>6</b>
1.1	Oslo . . . . .	6
1.2	Heimsjø . . . . .	7
1.3	Honningsvåg . . . . .	7
1.4	Narvik . . . . .	8
1.5	Harstad . . . . .	8
1.6	Tregde . . . . .	8
1.7	Andenes . . . . .	8
1.8	Viker . . . . .	8
<b>2</b>	<b>Annual maxima method</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	Parameter estimation . . . . .	9
2.3	Model checking . . . . .	10
2.4	Application to water level measurements . . . . .	11
2.4.1	Oslo . . . . .	11
2.4.2	Heimsjø . . . . .	14
2.4.3	Honningsvåg . . . . .	17
2.4.4	Narvik . . . . .	20
2.4.5	Harstad . . . . .	23
2.4.6	Tregde . . . . .	26
2.4.7	Andenes . . . . .	29
2.4.8	Viker . . . . .	32
<b>3</b>	<b>The peaks-over-threshold (POT) method</b>	<b>35</b>
3.1	Introduction . . . . .	35
3.2	Introducing stationarity . . . . .	35
3.3	Identification of clusters . . . . .	36
3.4	Threshold selection . . . . .	36
3.5	Return level estimation . . . . .	37
3.6	Parameter estimation . . . . .	38
3.7	Model checking . . . . .	39
3.8	Application to water level measurements . . . . .	40
3.8.1	Oslo . . . . .	40
3.8.2	Heimsjø . . . . .	44
3.8.3	Honningsvåg . . . . .	47
3.8.4	Narvik . . . . .	50
3.8.5	Harstad . . . . .	53
3.8.6	Tregde . . . . .	56

3.8.7	Andenes . . . . .	59
3.8.8	Viker . . . . .	62
<b>4</b>	<b>Revised Joint Probabilities Method</b>	<b>65</b>
4.1	Introduction . . . . .	65
4.2	A distribution for surges . . . . .	65
4.3	Finding the distribution for annual sea-level maxima . . . . .	66
4.4	Including interaction between tide and surge . . . . .	66
4.4.1	Testing for interaction . . . . .	66
4.4.2	Modelling the tide-surge interaction . . . . .	67
4.4.3	Estimating the normalisation functions $a(X)$ and $b(X)$ . . . . .	67
4.5	Estimating return levels with RJPM . . . . .	68
4.5.1	Estimation of surge distribution parameters . . . . .	68
4.5.2	Estimation of extremal indices . . . . .	68
4.5.3	Estimation of return levels . . . . .	69
4.6	Application to water level measurements . . . . .	70
4.6.1	Oslo . . . . .	70
4.6.2	Heimsjø . . . . .	73
4.6.3	Honningsvåg . . . . .	76
4.6.4	Narvik . . . . .	78
4.6.5	Harstad . . . . .	80
4.6.6	Tregde . . . . .	82
4.6.7	Andenes . . . . .	84
4.6.8	Viker . . . . .	86
<b>5</b>	<b>The ACER method</b>	<b>88</b>
5.1	Cascade of conditioning approximations . . . . .	88
5.2	Numerically estimating the ACER functions . . . . .	90
5.3	Polynomial smoothing of confidence intervals . . . . .	91
5.4	Estimating return levels with the ACER method . . . . .	91
5.5	Application to water level measurements . . . . .	93
5.5.1	Oslo . . . . .	93
5.5.2	Heimsjø . . . . .	95
5.5.3	Honningsvåg . . . . .	97
5.5.4	Narvik . . . . .	98
5.5.5	Harstad . . . . .	99
5.5.6	Tregde . . . . .	100
5.5.7	Andenes . . . . .	101
5.5.8	Viker . . . . .	103

---

<b>6</b>	<b>Discussion of results</b>	<b>105</b>
6.1	Oslo . . . . .	105
6.2	Heimsjø . . . . .	106
6.3	Honningsvåg . . . . .	107
6.4	Narvik . . . . .	108
6.5	Harstad . . . . .	109
6.6	Tregde . . . . .	110
6.7	Andenes . . . . .	111
6.8	Viker . . . . .	112
6.9	Conclusion . . . . .	113
	<b>Bibliography</b>	<b>116</b>

## List of Figures

1.1	Map with locations of all water level measuring stations analysed in this report . . . . .	7
2.1	Probability and quantile plots for the annual maxima methods, Oslo.	12
2.2	Return level plots for the annual maxima methods, Oslo . . . . .	13
2.3	Density plots for the annual maxima methods, Oslo . . . . .	13
2.4	Probability and quantile plots for the annual maxima methods, Heimsjø. . . . .	15
2.5	Return level plots for the annual maxima methods, Heimsjø. . . . .	16
2.6	Density plots for the annual maxima methods, Heimsjø. . . . .	16
2.7	Probability and quantile plots for the annual maxima methods, Honningsvåg. . . . .	18
2.8	Return level plots for the annual maxima methods, Honningsvåg. . . . .	19
2.9	Density plots for the annual maxima methods, Honningsvåg. . . . .	19
2.10	Probability and quantile plots for the annual maxima methods, Narvik. . . . .	21
2.11	Return level plots for the annual maxima methods, Narvik. . . . .	21
2.12	Density plots for the annual maxima methods, Narvik. . . . .	22
2.13	Probability and quantile plots for the annual maxima methods, Harstad. . . . .	24
2.14	Return level plots for the annual maxima methods, Harstad. . . . .	24
2.15	Density plots for the annual maxima methods, Harstad. . . . .	25
2.16	Probability and quantile plots for the annual maxima methods, Tregde	27
2.17	Return level plots for the annual maxima methods, Tregde . . . . .	27
2.18	Density plots for the annual maxima methods, Tregde . . . . .	28
2.19	Probability and quantile plots for the annual maxima methods, Andenes. . . . .	30
2.20	Return level plots for the annual maxima methods, Andenes. . . . .	31
2.21	Density plots for the annual maxima methods, Andenes. . . . .	31
2.22	Probability and quantile plots for the annual maxima methods, Vikar. . . . .	33
2.23	Return level plots for the annual maxima methods, Vikar. . . . .	34
2.24	Density plots for the annual maxima methods, Vikar. . . . .	34
3.1	Threshold selection plots for the POT model, Oslo. . . . .	40
3.2	Change in extremal index and 200 return level as calculated by a POT model, for Oslo. . . . .	41
3.3	Model diagnostics for the POT model, Oslo. . . . .	43
3.4	Threshold selection plots for the POT model, Heimsjø. . . . .	44
3.5	Change in extremal index and 200 return level as calculated by a POT model, for Heimsjø. . . . .	45
3.6	Model diagnostics for the POT model, Heimsjø. . . . .	46



3.7	Threshold selection plots for the POT model, Honningsvåg. . . . .	47
3.8	Change in extremal index and 200 return level as calculated by a POT model, for Honningsvåg. . . . .	48
3.9	Model diagnostics for the POT model, Honningsvåg. . . . .	49
3.10	Threshold selection plots for the POT model, Narvik. . . . .	50
3.11	Change in extremal index and 200 return level as calculated by a POT model, for Narvik. . . . .	51
3.12	Model diagnostics for the POT model, Narvik. . . . .	52
3.13	Threshold selection plots for the POT model, Harstad. . . . .	53
3.14	Change in extremal index and 200 return level as calculated by a POT model, for Harstad. . . . .	54
3.15	Model diagnostics for the POT model, Harstad. . . . .	55
3.16	Threshold selection plots for the POT model, Tregde. . . . .	56
3.17	Change in extremal index and 200 return level as calculated by a POT model, for Tregde. . . . .	57
3.18	Model diagnostics for the POT model, Tregde. . . . .	58
3.19	Threshold selection plots for the POT model, Andenes. . . . .	59
3.20	Change in extremal index and 200 return level as calculated by a POT model, for Andenes. . . . .	60
3.21	Model diagnostics for the POT model, Andenes. . . . .	61
3.22	Threshold selection plots for the POT model, Vikar. . . . .	62
3.23	Change in extremal index and 200 return level as calculated by a POT model, for Vikar. . . . .	63
3.24	Model diagnostics for the POT model, Vikar. . . . .	64
4.1	Tide-surge interaction plots for Oslo. . . . .	72
4.2	Tide-surge interaction plots for Heimsjø. . . . .	75
4.3	Tide-surge interaction plots for Honningsvåg. . . . .	77
4.4	Tide-surge interaction plots for Narvik. . . . .	79
4.5	Tide-surge interaction plots for Harstad. . . . .	81
4.6	Tide-surge interaction plots for Tregde. . . . .	83
4.7	Tide-surge interaction plots for Andenes. . . . .	85
4.8	Tide-surge interaction plots for Vikar. . . . .	87
5.1	ACER decision plots, with $k = 1$ to 6, for Oslo. . . . .	94
5.2	Log plot of estimated ACER functions for Heimsjø, with $k = 1, 2, 3, 4, 5$ and 6 . . . . .	95
5.3	ACER curves for high levels, Heimsjø . . . . .	96
5.4	ACER decision plots, with $k = 1$ to 6, for Honningsvåg. . . . .	97
5.5	ACER decision plots, with $k = 1$ to 6, for Narvik. . . . .	98
5.6	ACER decision plots, with $k = 1$ to 6, for Harstad. . . . .	99
5.7	ACER decision plots, with $k = 1$ to 6, for Tregde. . . . .	100

5.8	ACER decision plots, with $k = 1$ to 6, for Andenes. . . . .	101
5.9	ACER decision plots, with $k = 1$ to 6, for Vikar. . . . .	103
6.1	Comparison between return levels in Oslo for methods applied in the thesis . . . . .	105
6.2	Comparison between return levels in Heimsjø for methods applied in the thesis . . . . .	106
6.3	Comparison between return levels in Honningsvåg for methods ap- plied in the thesis. . . . .	107
6.4	Comparison between return levels in Narvik for methods applied in the thesis. . . . .	108
6.5	Comparison between return levels in Harstad for methods applied in the thesis. . . . .	109
6.6	Comparison between return levels in Tregde for methods applied in the thesis. . . . .	110
6.7	Comparison between return levels in Andenes for methods applied in the thesis. . . . .	111
6.8	Comparison between return levels in Vikar for methods applied in the thesis. . . . .	112

## List of Tables

2.1	Estimates of return levels and confidence intervals in for Oslo, Gum- bel model. Values in cm. . . . .	11
2.2	Estimates of return levels and confidence intervals for Oslo, GEV model. Values in cm. . . . .	12
2.3	Estimates of return levels and confidence intervals for Heimsjø, Gumbel model. Values in cm. . . . .	14
2.4	Estimates of return levels and confidence intervals for Heimsjø, GEV model. Values in cm. . . . .	14
2.5	Estimates of return levels and confidence intervals for Honningsvåg, Gumbel model. Values in cm. . . . .	17
2.6	Estimates of return levels and confidence intervals for Honningsvåg, GEV model. Values in cm. . . . .	17
2.7	Estimates of return levels and confidence intervals for Narvik, Gum- bel model. Values in cm. . . . .	20
2.8	Estimates of return levels and confidence intervals for Narvik, GEV model. Values in cm. . . . .	20
2.9	Estimates of return levels and confidence intervals for Harstad, Gumbel model. Values in cm. . . . .	23

2.10	Estimates of return levels and confidence intervals for Harstad, GEV model. Values in cm. . . . .	23
2.11	Estimates of return levels and confidence intervals for Tregde, Gumbel model. Values in cm. . . . .	26
2.12	Estimates of return levels and confidence intervals for Tregde, GEV model. Values in cm. . . . .	26
2.13	Estimates of return levels and confidence intervals for Andenes, Gumbel model. Values in cm. . . . .	29
2.14	Estimates of return levels and confidence intervals for Andenes, GEV model. Values in cm. . . . .	30
2.15	Estimates of return levels and confidence intervals for Vikar, Gumbel model. Values in cm. . . . .	32
2.16	Estimates of return levels and confidence intervals for Vikar, GEV model. Values in cm. . . . .	33
3.1	Return level estimates and confidence intervals for the POT model, Oslo. Values in cm. . . . .	42
3.2	Return level estimates and confidence intervals for the POT model, Heimsjø. Values in cm. . . . .	45
3.3	Return level estimates and confidence intervals for the POT model, Honningsvåg. Values in cm. . . . .	48
3.4	Return level estimates and confidence intervals for the POT model, Narvik. Values in cm. . . . .	51
3.5	Return level estimates and confidence intervals for the POT model, Harstad. Values in cm. . . . .	54
3.6	Return level estimates and confidence intervals for the POT model, Tregde. Values in cm. . . . .	57
3.7	Return level estimates and confidence intervals for the POT model, Andenes. Values in cm. . . . .	60
3.8	Return level estimates and confidence intervals for the POT model, Vikar. Values in cm. . . . .	63
4.1	Ratio of tidal versus observed total maxima and ratio of means of tidal and observed annual maxima. . . . .	70
4.2	Return levels and 95% confidence intervals for the RJPM model, Oslo. Values in cm. . . . .	71
4.3	Return levels and 95% confidence intervals for the RJPM model, Heimsjø. Values in cm. . . . .	74
4.4	Return levels and 95% confidence intervals for the RJPM model, Honningsvåg. Values in cm. . . . .	76
4.5	Return levels and 95% confidence intervals for the RJPM model, Narvik. Values in cm. . . . .	78

---

4.6	Return levels and 95% confidence intervals for the RJPM model, Harstad. Values in cm. . . . .	80
4.7	Return levels and 95% confidence intervals for the RJPM model, Tregde. Values in cm. . . . .	82
4.8	Return levels and 95% confidence intervals for the RJPM model, Andenes. Values in cm. . . . .	84
4.9	Return levels and 95% confidence intervals for the RJPM model, Viker. Values in cm. . . . .	86
5.1	Return level estimates and confidence intervals for the ACER method, Oslo. Values in cm. . . . .	94
5.2	Return level estimates and confidence intervals for the ACER method, Heimsjø. Values in cm. . . . .	96
5.3	Return level estimates and confidence intervals for the ACER method, Honningsvåg. Values in cm. . . . .	97
5.4	Return level estimates and confidence intervals for the ACER method, Narvik. Values in cm. . . . .	98
5.5	Return level estimates and confidence intervals for the ACER method, Harstad. Values in cm. . . . .	99
5.6	Return level estimates and confidence intervals for the ACER method, Tregde. Values in cm. . . . .	100
5.7	Return level estimates and confidence intervals for the ACER method, Andenes. Values in cm. . . . .	102
5.8	Return level estimates and confidence intervals for the ACER method, Viker. Values in cm. . . . .	104
6.1	Estimated shape parameters for the POT models. . . . .	115

# 1 Data sets

The data sets for the methods used in this Master’s thesis are water level measurements and tidal predictions for eight locations on the Norwegian coast. Water level measurements are collected by automated equipment, while the tidal predictions are based on numerical models.

We have three data sets from the south of Norway; Oslo, Tregde and Vikar. Tide isn’t as large an influence here as it is further north; the locations are called *surge-dominant*. The total height of sea levels is also lower here than further north. We have one data set in the middle of Norway, Heimsjø, and four from the north: Narvik, Harstad, Andenes and Honningsvåg. Sea levels are more influenced by tides here; the locations are *tide-dominant*. The data sets are not freely available, but were provided by Even Haug at the Norwegian Hydrographic Service, a division of the Norwegian Mapping Authority.

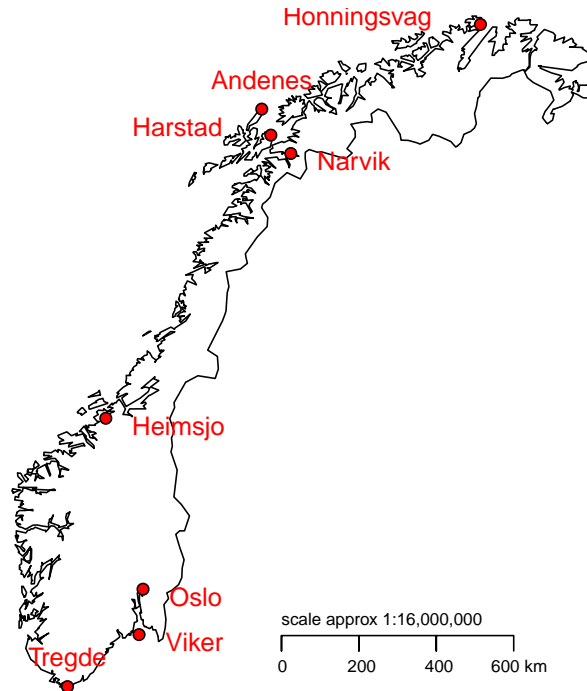
The sea level measurements are done either hourly or every 10 minutes. For the 10 minute data, we have extracted hourly measurements for usage in the POT, RJPM and ACER methods. There are several reasons for this: Firstly, the RJPM method uses tidal predictions which were provided as hourly sea levels. Secondly, many locations have hourly data for older periods and ten minute measurements for the recent years. In order to use both data sets, we must extract values every hour from the more frequent observations. Thirdly, the literature used in the theoretical studies have hourly measurements. Methodological comparisons are therefore made simpler.

If not otherwise noted, all references to the height of the sea level are in centimetres. All return periods are, if not otherwise noted, measured in years.

## 1.1 Oslo

For the measurement station in Oslo, we have hourly measured sea levels from December 10 1914 to December 31 1991 and 10 minute interval data from October 1 1991 to September 16 2010. These are uncorrected for post-glacial rebound, which is an important factor in the Oslo area. To correct for this, 4 mm/year have been added to data points for years after 1988 and the same amount has been subtracted per year before 1988, according to the formula  $currentValue + 0.4 \cdot (currentYear - 1988)$ . This was found in a previous report [3], and is based on the fact that 1988 was the base year to calculate mean sea level (MSL) in Oslo.

For the same periods, tidal sea level predictions are available. These are based on a numerical model, the details of which are not publicly available.



**Figure 1.1:** Map with locations of all water level measuring stations analysed in this report

## 1.2 Heimsjø

Heimsjø is located on the coast of Sør-Trøndelag, and the measurement station is found at latitude  $63^{\circ}26'$  N and longitude  $09^{\circ}07'$  E. Data are available as hourly measurements from November 1 1928 to December 31 1990, and as 10-minute interval measurements from November 1 1990 to September 16 2010. Tide predictions are also available for the same periods and corresponding measurement intervals.

## 1.3 Honningsvåg

Honningsvåg is found at the very north of Norway, in Finnmark. The location of the measurement station is latitude  $70^{\circ}59'$  N and longitude  $25^{\circ}59'$  E. Measured

values exist in hourly form from from June 5 1970 to December 31 1988, and as 10 minute data from June 1 1988 to September 16 2010. Corresponding tidal predictions are also available.

## **1.4 Narvik**

Narvik, in Nordland fylke, has its water level measurement station at latitude  $68^{\circ}26'$  N and longitude  $17^{\circ}25'$  E. Available are hourly measurements from August 6 1931 to December 31 1989, and 10 minute measurements from October 10 1989 to April 11 2011. Tidal predictions are available for the same period.

## **1.5 Harstad**

The measurement station in Harstad, Troms fylke, is located at latitude  $68^{\circ}48'$  N and longitude  $16^{\circ}33'$  E. Hourly measurements are available from March 1 1952 to December 31 1988, and 10 minute data from June 1 1988 to April 11 2011. Tidal predictions are available for the same period.

## **1.6 Tregde**

Tregde, in Vest-Agder, has a water level station at latitude  $58^{\circ}00'$  N and longitude  $07^{\circ}34'$  E. There are hourly measurements from October 5 1927 to December 31 1987, and 10 minute measurements from December 10 1987 and April 11 2011. Tidal predictions are available for the same period.

## **1.7 Andenes**

Andenes in Nordland, at latitude  $69^{\circ}19'$  N and longitude  $16^{\circ}09'$  E, has data of poor quality from April 1 1940 to October 23 1986 - too poor to be used for analysis. 10 minute data are available from October 9 1991 to April 11 2011, however. Tidal predictions are available for the same period.

## **1.8 Vikør**

Vikør in Østfold, located at latitude  $59^{\circ}02'$  N and longitude  $10^{\circ}57'$  E, had its water level gauge installed as late as in 1990, and 10 minute interval data are available from October 16 that year to April 11 2011. Tidal predictions are available for the same period.

## 2 Annual maxima method

### 2.1 Introduction

The Annual maxima method (AMM) builds on the extremal types theorem, which says that, for some sequence of independent and identically distributed random variables  $X_1, \dots, X_n$ ,  $X \sim F$ , with maximum  $M_n = \max\{X_1, \dots, X_n\}$ , if the probability

$$Prob\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z), \quad (2.1)$$

for some sequences of constants  $\{b_n\}$  and  $\{a_n > 0\}$ , converges to some non-degenerate distribution function  $G$  for  $n \rightarrow \infty$ , then this distribution belongs to either the Gumbel, Fréchet or Weibull families (denoted type I, type II and type III extreme value distributions). Through a rescaling of parameters, all families may be collected in a Generalized Extreme Value (GEV) distribution,

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (2.2)$$

defined on  $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$ .  $\xi = 0$  (taken as the limit  $\xi \rightarrow 0$ ) corresponds to Gumbel,  $\xi > 0$  to Fréchet and  $\xi < 0$  to Weibull.

Since we use annual maxima,  $n$  is the number of observations in a year. Equation (2.2) is solved for the return level  $z_p$  by setting  $G(z_p) = 1 - p$ , where  $p$  is the probability of a level  $z_p$  in a given year - i.e. the reciprocal of the return period. This means that, say, a 5 year return period is represented by  $p = 1/5$ . With parameter estimates, this inversion provides the formula

$$\hat{z}_p = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}[1 - \{-\log(1 - p)\}^{-\hat{\xi}}] & \text{for } \xi \neq 0 \\ \hat{\mu} - \hat{\sigma} \log[-\log(1 - p)] & \text{for } \xi = 0 \end{cases} \quad (2.3)$$

It is worth noting that the extremal types theorem providing the GEV distribution relies on the  $X_i$  to be independent and identically distributed. There is often considerable dependence in a real data set, however. This problem is overcome by the fact that the dependence only affects the values of estimated parameters. Since they are to be estimated anyway, this can mostly be ignored [1].

### 2.2 Parameter estimation

For the published work done on Norwegian water levels, the shape parameter is assumed and fixed at  $\xi = 0$ . This means that the Gumbel distribution is used, and the unknown parameters to be estimated are  $\mu$  and  $\sigma$  in

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z - \mu}{\sigma}\right)\right]\right\}. \quad (2.4)$$



These parameters are estimated by maximum likelihood fitting, in practice by the `fgev` function in the R package `evd`. This is also the case for  $\xi \neq 0$ , and results for both cases will be presented below.

### 2.3 Model checking

The empirical distribution function for an ordered sample  $z_{(1)} \leq \dots \leq z_{(n)}$  of independent observations is given by

$$\tilde{G}(z_{(i)}) = \frac{i}{n+1}, \quad i = 1, \dots, n$$

where  $G$  is the underlying distribution for the population. Since  $\tilde{G}$  is an estimate of  $G$ , it should be in good agreement with the estimated model  $\hat{G}$ . This is commonly checked by two graphical techniques, the *probability plot* and the *quantile plot*.

The probability plot is made by plotting  $\hat{G}$  for the sorted observations  $z_{(1)} \leq \dots \leq z_{(n)}$  against the corresponding empirical distribution function,

$$\left\{ \left( \frac{i}{k+1}, \hat{G}(z_{(i)}) \right); i = 1, \dots, k \right\}.$$

The points should be close to the unit diagonal. A problem with this plot is that both  $i/(k+1)$  and  $\hat{G}(z_{(i)})$  approach 1 for values of  $i$  close to  $n$ , which is exactly where the accuracy of the model is of greatest concern. This can be rectified by applying  $\hat{G}^{-1}(\cdot)$  to both x- and y-axis values. This gives the quantile plot,

$$\left\{ \left( \hat{G}^{-1}\left(\frac{i}{k+1}\right), z_{(i)} \right); i = 1, \dots, k \right\},$$

where linearity again is the desired quality.

For both probability and quantile plots, it is common to draw a straight line from the origin sloping upwards at a 45° angle. For a perfect fit, the points should be exactly on this line.

In addition, it is informative to look at a *return level plot*, where return period  $R = 1/p$  is plotted against return level  $\hat{z}_p$  for the estimated model,

$$\{(1/p, \hat{z}_p) : 0 < p < 1\},$$

with a logarithmic scale on the x-axis. Empirical point estimates of the return level function and 95 % confidence intervals are also added. The point estimates should be within the confidence intervals and ideally on the return level line.

Finally, a comparison between the GEV probability density function and a histogram of the observed maxima are combined in the *density plot*. This is not as informative as the other three plots, since the shape of the histogram often varies considerably between different choices of group intervals.

## 2.4 Application to water level measurements

### 2.4.1 Oslo

Oslo has data available from all years from 1914 to 2010 - except 1939, from which there are no available measurements. In addition, some of the years with available data have important data missing. For instance, there are only 312 data points in 1914, all from the month of December. This means that there is a very real possibility that the real annual maximum is excluded. In 1915, there is much missing data from the important autumn and winter months, where annual maxima are often found. This year is therefore also excluded from consideration. In 1972, there are no measurements in February, July, August, September and October, enough missing that the probability that the year's maximum is left out is large. 1974 lacks data for July, August and September, and 1991 only has data for October, November and December. Both these are excluded. Finally, measurements for 2010 only go to September 16, so its maximum is also dropped from consideration.

With the data as specified above, estimates together with maximum likelihood confidence intervals are published in tables 2.1 and 2.2, for Gumbel and GEV respectively. The published values in the second column are collected from a report by the Norwegian Map Authority [3], since they are more accurately calculated than those on the Vannstand.no web pages. Furthermore, values exist for return periods up to 200 years.

For the Gumbel model, parameter estimates are  $\hat{\mu} = 167.31(2.0735)$  and  $\hat{\sigma} = 18.630(1.4974)$ . For the GEV model,  $\hat{\mu} = 168.21(2.2325)$ ,  $\hat{\sigma} = 19.030(1.5792)$  and  $\hat{\xi} = -0.0892(0.0700)$ . Standard errors are shown in parentheses.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	194.1	188.4	195.2	202.1
10	207.5	200.4	209.2	218
20	220.3	211.9	222.6	233.4
100	249.3	237.7	253	268.2
200	261.7	248.7	265.9	283.2

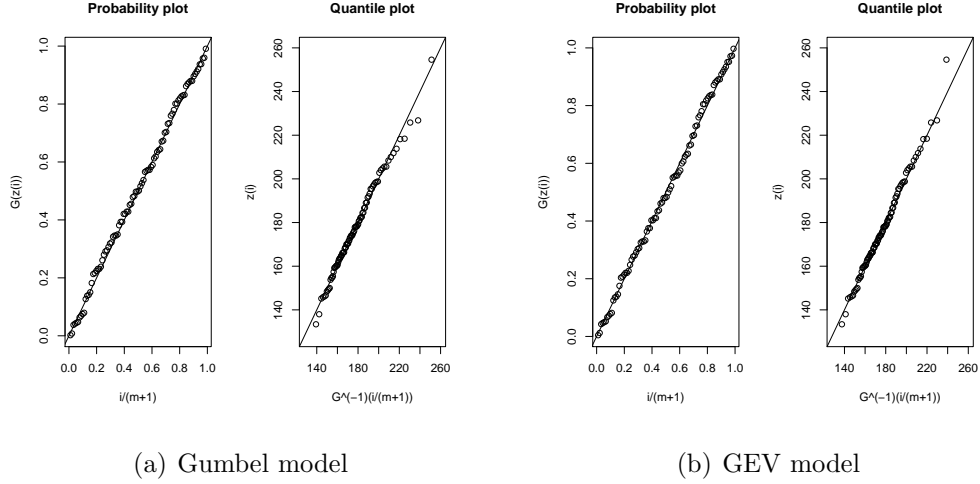
**Table 2.1:** Estimates of return levels and confidence intervals in for Oslo, Gumbel model. Values in cm.

As would be expected, the new estimates are very close to the pre-existing values. The differences come from the years used in estimation as well as the method of estimation; the work in [3] uses Probability Weighted Moments (PWN) for the parameter estimates and annual maxima from 1950 to 1999, excluding

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	195	188.7	194.9	201.1
10	207.3	199.1	207	214.9
20	218.6	207.4	217.9	228.3
100	241.8	220.5	240.1	259.7
200	250.8	223.9	248.8	273.8

**Table 2.2:** Estimates of return levels and confidence intervals for Oslo, GEV model. Values in cm.

1980-1982. The probability and quantile plots in figure 2.1 do not reveal any

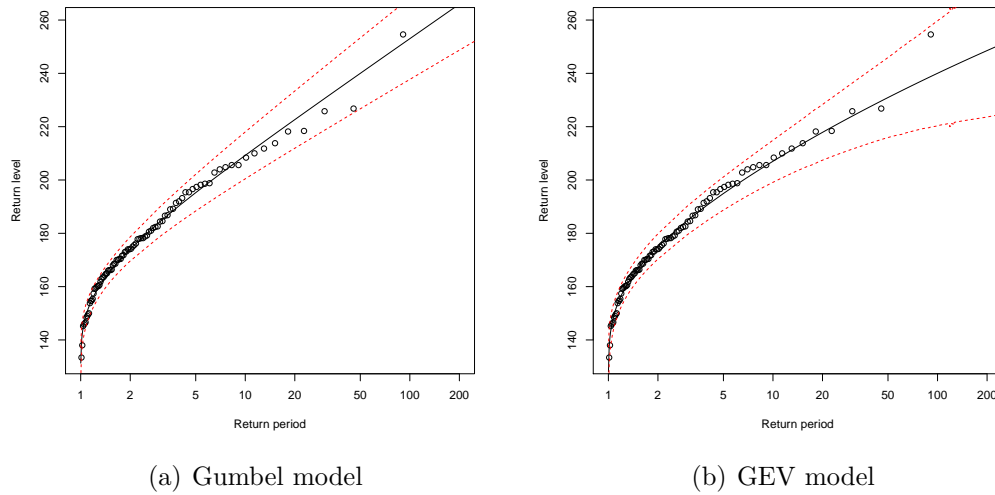


**Figure 2.1:** Probability and quantile plots for the annual maxima methods, Oslo.

fallacies with neither of the models, and points stay mostly on the line. For the Gumbel model in figure 2.1(a), the two leftmost points and a couple of the points most to the right are off the model line. For the GEV model in figure 2.1(b), the two leftmost points and the rightmost point are off the line.

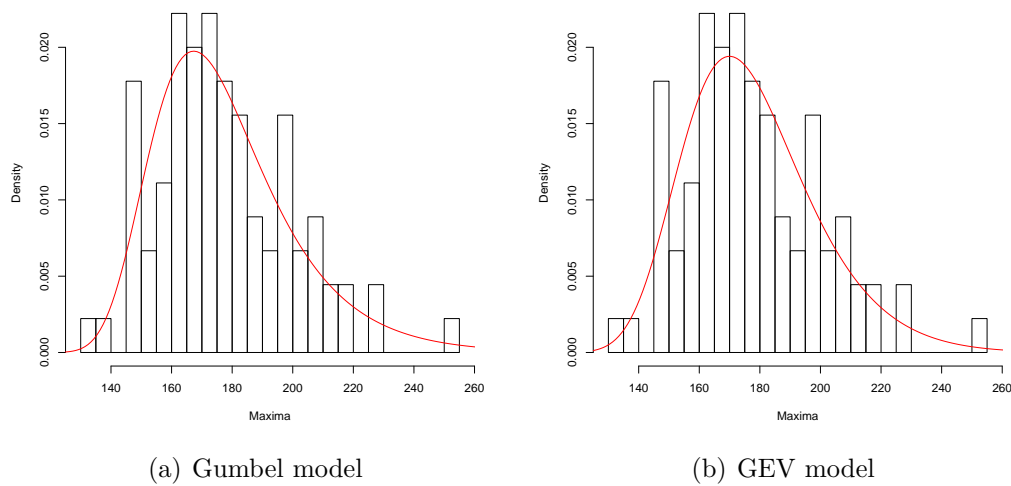
The return level plot in figure 2.2 shows that all points stay within the 95 % confidence intervals for both models. The two curves follow slightly different paths; the Gumbel model in figure 2.2(a) fits generally OK to all points, while the GEV model in figure 2.2(b) fits generally better to most points but poorly to the rightmost point. Since the estimated shape parameter is negative, the return levels of the GEV model are bounded at  $\hat{\mu} - \hat{\sigma}/\hat{\xi} = 381.7$  (cm), while the Gumbel

model return levels are unbounded. Finally, we look at the density plots in figures



**Figure 2.2:** Return level plots for the annual maxima methods, Oslo

2.3(a) and 2.3(b), where the general trend of the distribution seems to be followed in both cases.



**Figure 2.3:** Density plots for the annual maxima methods, Oslo

### 2.4.2 Heimsjø

Heimsjø has data for all years from 1928 to 2010. As with Oslo, some years have little enough data that it is likely that the true annual maximum has been left out. One example is 1934, where data for all the first seven months are missing. As such, we leave this year out. 1938 has missing data for January, February and December; important months where the real annual maximum is likely to be. This year is therefore also left out. 1943 lacks data for months January to April, and is also left out. 1959 lacks any data from September to December, and has little from February and August. Finally, 2010 lacks data for the last months of the year.

The published values are collected from the Vannstand.no web pages [9]. The web pages have values only up to 20 years, and they are based on a Gumbel model fitted to annual maxima.

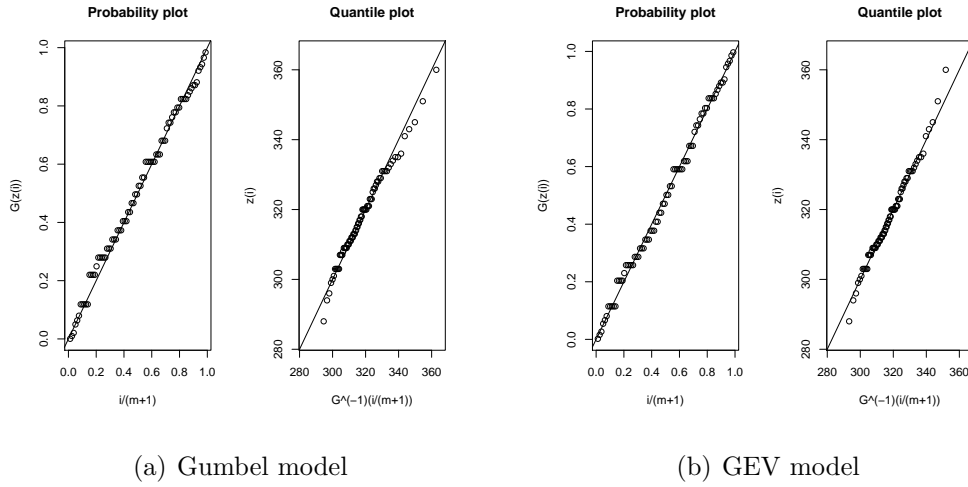
For the Gumbel model, parameter estimates are  $\hat{\mu} = 311.84(1.399)$  and  $\hat{\sigma} = 11.685(0.97756)$ . For the GEV model,  $\hat{\mu} = 312.69(1.4798)$ ,  $\hat{\sigma} = 11.889(1.0219)$  and  $\hat{\xi} = -0.13655(0.066255)$ .

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	329	324.8	329.4	333.9
10	337	332.3	338.1	343.9
20	345	339.5	346.5	353.6
100	-	355.5	365.5	375.6
200	-	362.4	373.7	385

**Table 2.3:** Estimates of return levels and confidence intervals for Heimsjø, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	329	325	328.8	332.7
10	337	331	335.7	340.4
20	345	335.8	341.7	347.6
100	-	343.2	353.3	363.3
200	-	345.2	357.5	369.9

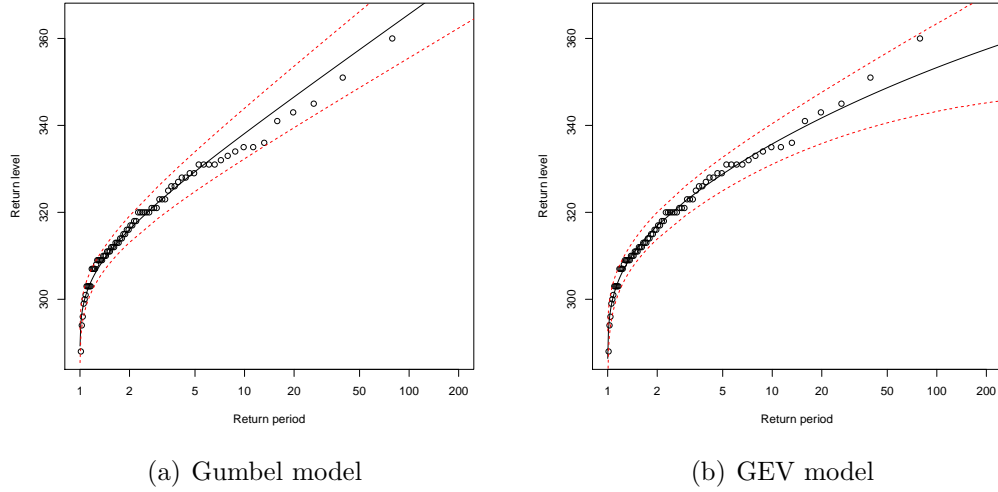
**Table 2.4:** Estimates of return levels and confidence intervals for Heimsjø, GEV model. Values in cm.



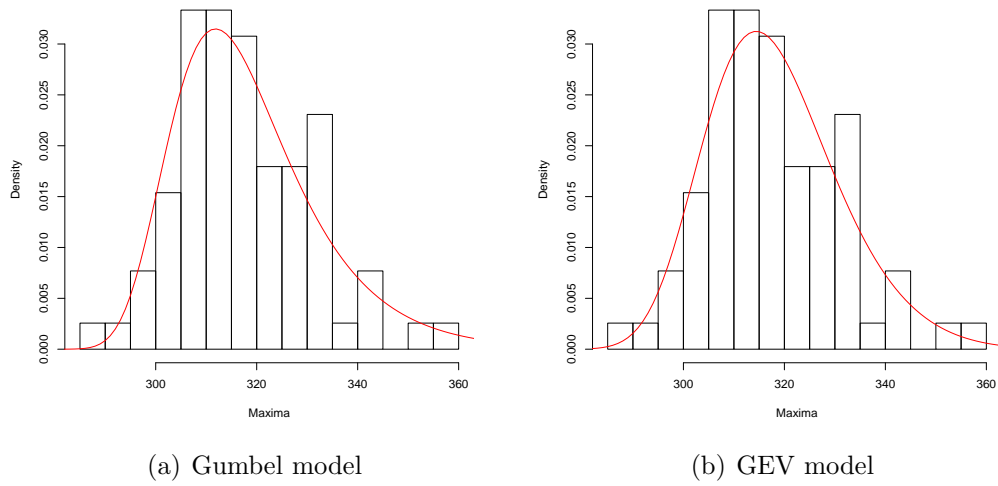
**Figure 2.4:** Probability and quantile plots for the annual maxima methods, Heimsjø.

The probability plots in figure 2.4 shows points mostly on the line, but the fit is seemingly not as good as for the Oslo data. This notion is also reflected in the quantile plots, with the Gumbel quantile plot having a poor fit for the three first and 7-9 of the last points. The GEV quantile plot in figure 2.4(b) has a relatively poor fit for the the first and two last points.

The return level plot in figure 2.5 shows the curves of the Gumbel and GEV models, together with confidence intervals and observed points. The Gumbel model in figure 2.5(a) seems to somewhat overestimate the points at high levels, while the GEV model underestimates them. All points are within confidence intervals, however. The GEV model return levels are bounded at 399.7 cm. Finally, we look at the density plots in figures 2.6(a) and 2.6(b), where the general trend of the distribution seems to be followed in both cases.



**Figure 2.5:** Return level plots for the annual maxima methods, Heimsjø.



**Figure 2.6:** Density plots for the annual maxima methods, Heimsjø.

### 2.4.3 Honningsvåg

Honningsvåg has data for all years from 1970 to 2010, except 1985. Most of the years have acceptable amounts of data, but four years have quite little: 1970 has no data from January to May, 1988 nothing in February and very little in March, 1989 nothing in August and September and 2010 nothing from October to December. All these are excluded from the model fitting.

The published values are collected from the Vannstand.no web pages [9]. The web pages have values only up to 20 years, and they are based on a Gumbel model fitted to annual maxima.

For the Gumbel model, parameter estimates are  $\hat{\mu} = 331.08(2.087)$  and  $\hat{\sigma} = 11.812(1.4575)$ . For the GEV model,  $\hat{\mu} = 332.85(2.2848)$ ,  $\hat{\sigma} = 12.344(1.6255)$  and  $\hat{\xi} = -0.27711(0.11502)$ .

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	348	342	348.8	355.6
10	357	349	357.7	366.3
20	366	355.6	366.2	376.7
100	-	370.4	385.4	400.4
200	-	376.7	393.6	410.6

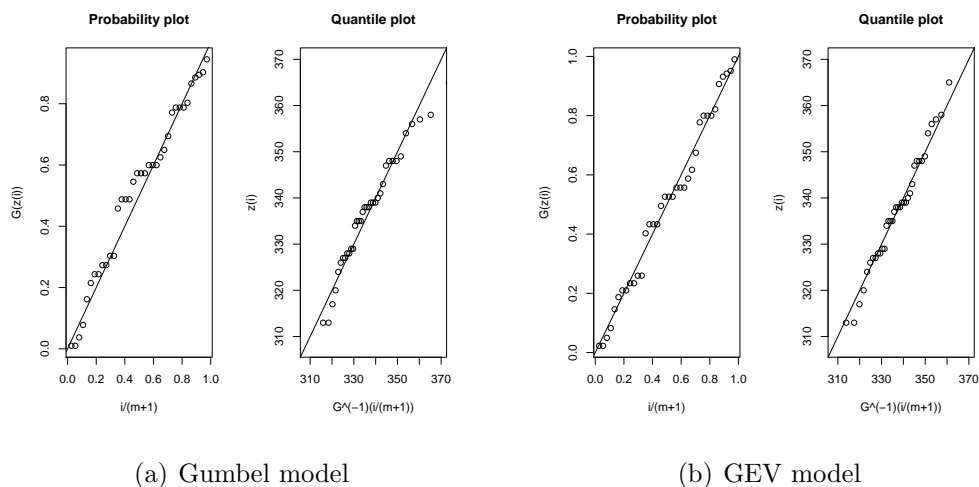
**Table 2.5:** Estimates of return levels and confidence intervals for Honningsvåg, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	348	343.2	348	352.8
10	357	348.3	353.5	358.7
20	366	351.7	357.8	363.9
100	-	355.3	365	374.7
200	-	355.6	367.2	378.7

**Table 2.6:** Estimates of return levels and confidence intervals for Honningsvåg, GEV model. Values in cm.

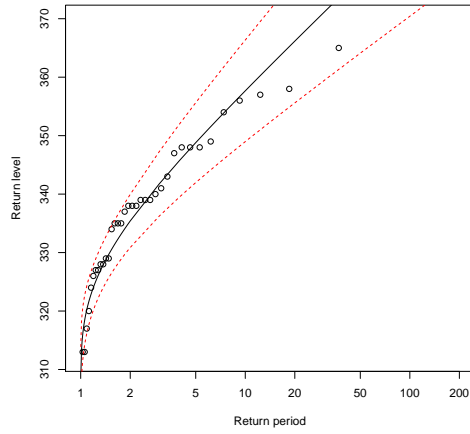
The probability and quantile plots in figure 2.7 show that there are fewer points available for the model estimation than for Oslo and Heimsjø. The fit is also quite as good as for these two locations - even though a few points are on the line, the points seem to follow a sort of oscillating S shape.



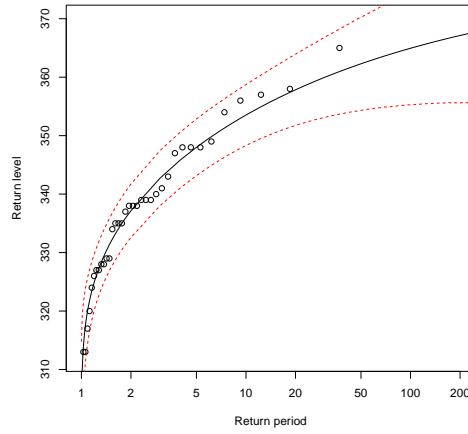


**Figure 2.7:** Probability and quantile plots for the annual maxima methods, Honningsvåg.

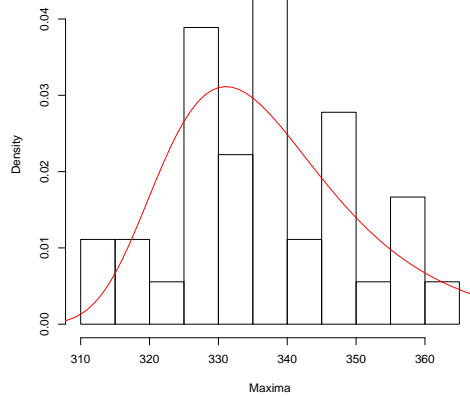
The return level plots in figure 2.8 show the trend of the Gumbel and GEV models. Again we see that the fit isn't very impressive, although points stay within confidence bounds. As for Oslo and Heimsjø, the estimated shape parameter is negative,  $\hat{\xi} = -0.2771$ , meaning that the Gumbel model gives higher estimates for the return levels for long periods. Thus the higher return levels seem to be overestimated by the Gumbel model in figure 2.8(a) and somewhat underestimated in figure 2.8(b). The GEV return levels are bounded at 377.4 cm, quite far below even the 200 year return level of the Gumbel model. Finally, we look at the density plots in figures 2.9(a) and 2.9(b), where the general trend of the distribution seems to be followed in both cases. As usual, it is hard to glean any new and revealing information from this type of diagnostic plot.



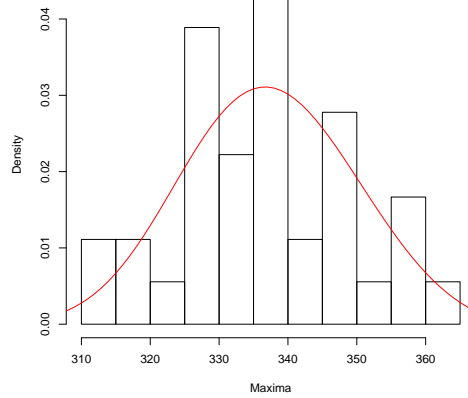
(a) Gumbel model



(b) GEV model

**Figure 2.8:** Return level plots for the annual maxima methods, Honningsvåg.

(a) Gumbel model



(b) GEV model

**Figure 2.9:** Density plots for the annual maxima methods, Honningsvåg.

#### 2.4.4 Narvik

69 annual maxima are used for the calculation of the Gumbel and GEV return level estimates in Narvik. The years 1932 to 2010 are used, excluding 1936, 1940-1947 and 1989.

For the Gumbel model, parameter estimates are  $\hat{\mu} = 401.32(2.1016)$  and  $\hat{\sigma} = 16.499(1.4916)$ . For the GEV model,  $\hat{\mu} = 402.91(2.2978)$ ,  $\hat{\sigma} = 17.087(1.6149)$  and  $\hat{\xi} = -0.17672(0.084322)$ . GEV return levels are therefore bounded at 499.6 cm. With these estimated parameters, results are found in tables 2.7 and 2.8 for the Gumbel and GEV models, respectively.

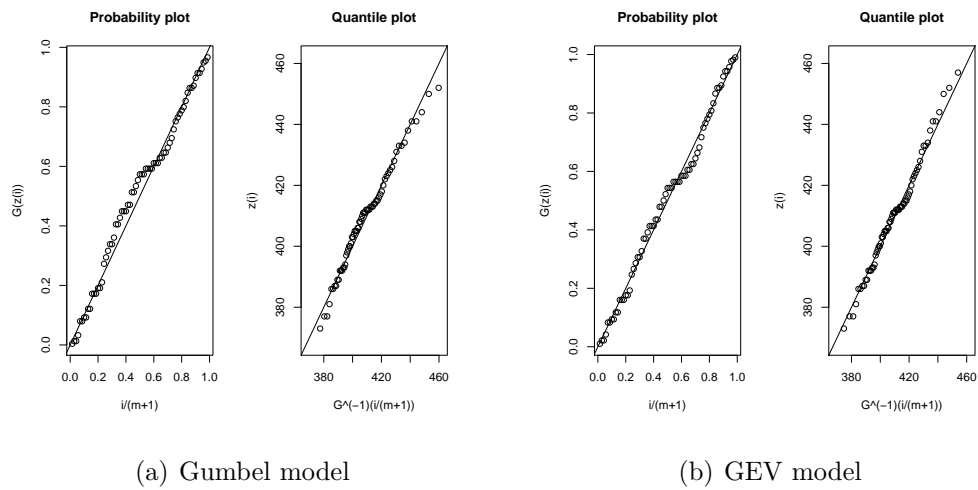
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	427	419.1	426.1	433
10	439	429.6	438.5	447.3
20	451	439.6	450.3	461.1
100	-	461.9	477.2	492.5
200	-	471.4	488.7	506

**Table 2.7:** Estimates of return levels and confidence intervals for Narvik, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	427	419.8	425.4	431
10	439	428	434.6	441.3
20	451	434	442.4	450.8
100	-	442	456.7	471.5
200	-	443.6	461.8	480

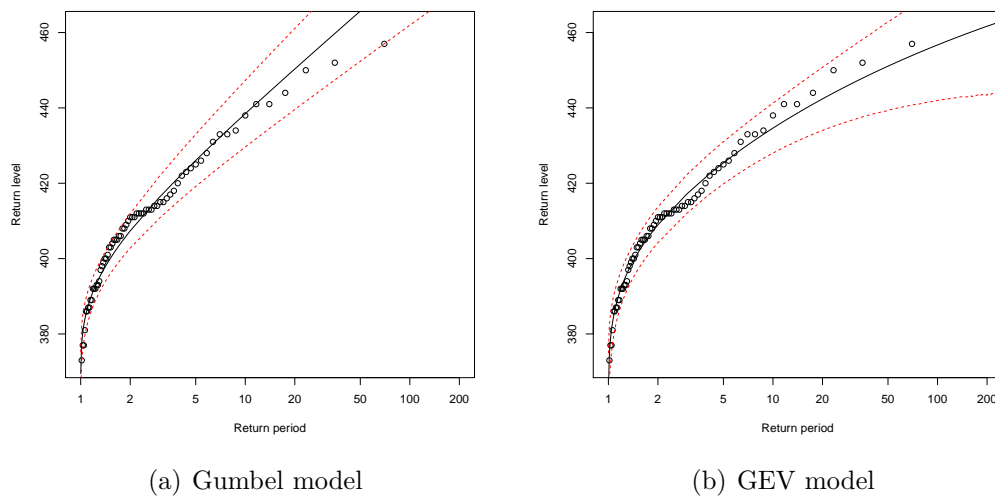
**Table 2.8:** Estimates of return levels and confidence intervals for Narvik, GEV model. Values in cm.

The probability and quantile plots in figure 2.10 show that most points stay on the line, but there is a slight indication of an S shape centred in the middle of the plots. The probability plot for the Gumbel model reveals some problems with the fit of the points to the left of this centre. The Gumbel quantile plot reveals some further problems with the points at the very highest levels, where they are overestimated. For the GEV model, the probability and quantile plots in figure 2.10(b) shows some poor fit to the right of the centre of the S, as well as some underestimation of points at the highest levels.



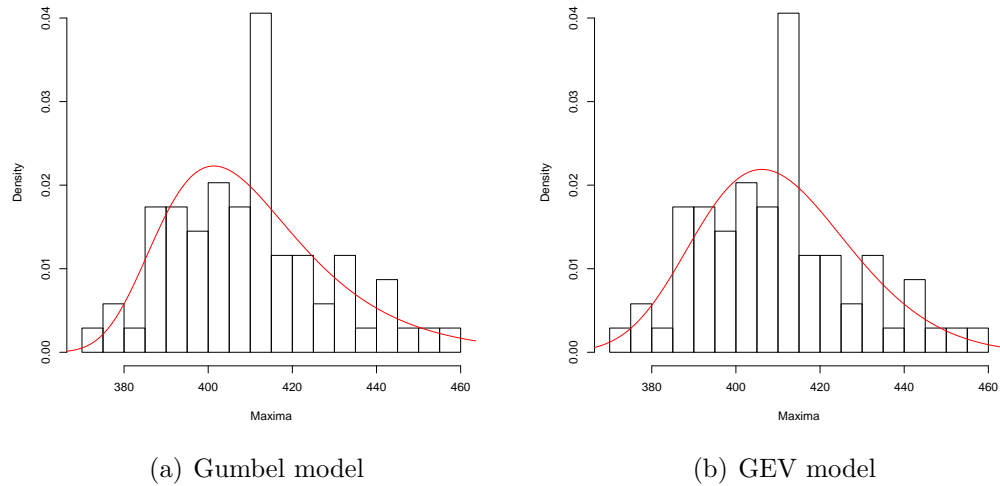
**Figure 2.10:** Probability and quantile plots for the annual maxima methods, Narvik.

The return level plot in figure 2.11 shows better the overestimation of the Gumbel model at high levels, as well as the underestimation of high-level points for the GEV model. All points are safely within the confidence intervals, with the exception of the rightmost point in figure 2.11(a) which sits on the lower confidence bound. Finally, we look at the density plots in figures 2.12(a) and 2.12(b). The



**Figure 2.11:** Return level plots for the annual maxima methods, Narvik.

general trend is followed, but we see a massive spike a little before the 420 mark, indicating a high density of points at this level. This area coincides with the centre of the S mentioned when discussing the probability and quantile plots.



**Figure 2.12:** Density plots for the annual maxima methods, Narvik.

### 2.4.5 Harstad

Harstad has annual maxima from 1952 to 2010, excluding 1966, 1976, 1980, 1988 and 1990. This gives a total of 54 data points for the estimation of parameters. For the Gumbel model, parameter estimates are  $\hat{\mu} = 281.68(1.8033)$  and  $\hat{\sigma} = 12.471(1.1983)$ . For the GEV model,  $\hat{\mu} = 284.68(1.9095)$ ,  $\hat{\sigma} = 12.746(1.4552)$  and  $\hat{\xi} = -0.46541(0.099296)$ .

For the GEV model,  $\hat{\xi}$  is quite close to  $-0.5$ , which corresponds to a distribution with a very short bounded upper tail. This is rarely encountered in practical applications of extreme value modeling [1], and could indicate a poor model. The short bound of the upper tail is illustrated by the upper limit of  $\hat{\mu} - \hat{\sigma}/\hat{\xi} = 284.68 - 12.746/(-0.465) = 312.091$ , which is below the 20 year published value of 316.

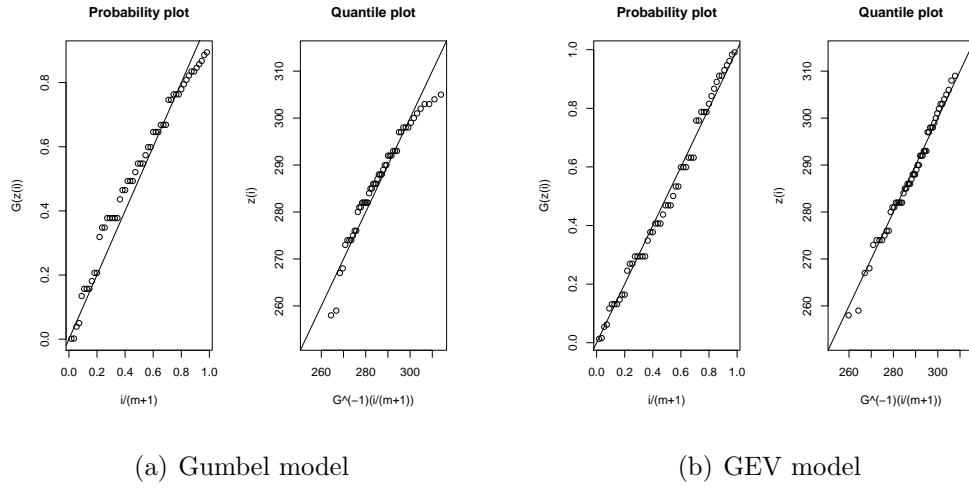
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	299	294.6	300.4	306.1
10	308	302.5	309.7	317
20	316	309.9	318.7	327.5
100	-	326.6	339	351.4
200	-	333.7	347.7	361.7

**Table 2.9:** Estimates of return levels and confidence intervals for Harstad, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	299	295.3	298.4	301.6
10	308	299.6	302.5	305.3
20	316	302.4	305.2	308
100	-	305.1	308.9	312.6
200	-	305.5	309.7	314

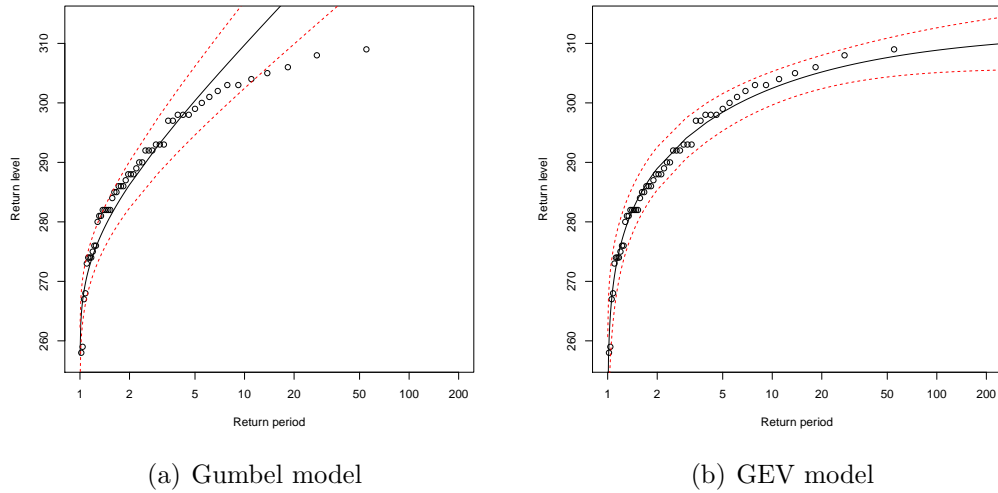
**Table 2.10:** Estimates of return levels and confidence intervals for Harstad, GEV model. Values in cm.

The probability and quantile plots of figure 2.13 show that the GEV model is better fitted to the annual maxima than the Gumbel model, with both probability and quantile plots in figure 2.13(a) showing a substantial amount of points off the diagonal. The return level plot in figure 2.14 shows the behaviour of both models



**Figure 2.13:** Probability and quantile plots for the annual maxima methods, Harstad.

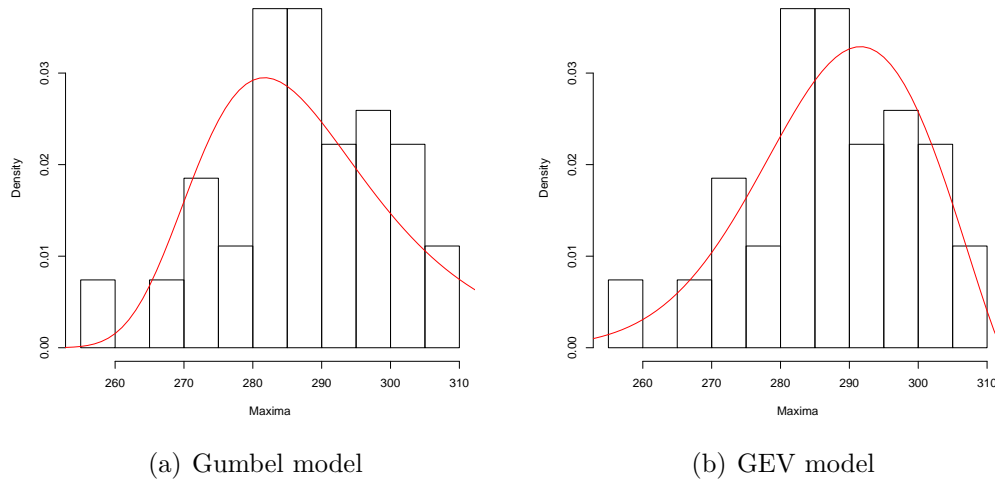
compared to the data, and the GEV model in figure 2.14(b) has most points on or close to the model line and all points within confidence intervals. For the Gumbel model, figure 2.14(a) shows that the model significantly overestimates the highest maxima, with four points outside confidence intervals.



**Figure 2.14:** Return level plots for the annual maxima methods, Harstad.

The density plots in figures 2.15(a) and 2.15(b) illustrate the large difference

between the Gumbel and GEV models created by the large negative  $\hat{\xi}$  in the latter model. The GEV model seems to better fit the histogram of real data.



**Figure 2.15:** Density plots for the annual maxima methods, Harstad.



### 2.4.6 Tregde

For Tregde, we have annual maxima available for 1928 to 2010, excluding 1941, 1944 and 1956. This means that 80 maxima are available for the model fitting.

For the Gumbel model, parameter estimates are  $\hat{\mu} = 114.96(1.1428)$  and  $\hat{\sigma} = 9.716(0.85818)$ . For the GEV model,  $\hat{\mu} = 114.83(1.2216)$ ,  $\hat{\sigma} = 9.6311(0.89878)$  and  $\hat{\xi} = 0.025046(0.088029)$ . The estimated shape parameter for the GEV model is slightly positive, meaning that the model has no upper bound.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	129	125.7	129.5	133.4
10	136	131.9	136.8	141.8
20	143	137.8	143.8	149.9
100	-	151	159.7	168.4
200	-	156.6	166.4	176.3

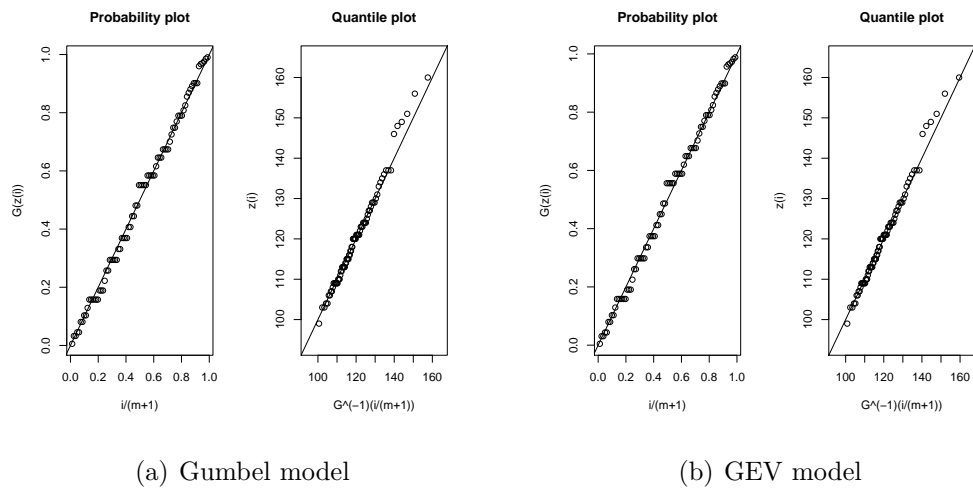
**Table 2.11:** Estimates of return levels and confidence intervals for Tregde, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	129	125.6	129.6	133.5
10	136	131.5	137.1	142.8
20	143	136.3	144.5	152.8
100	-	143.6	161.7	179.9
200	-	145.3	169.1	193

**Table 2.12:** Estimates of return levels and confidence intervals for Tregde, GEV model. Values in cm.

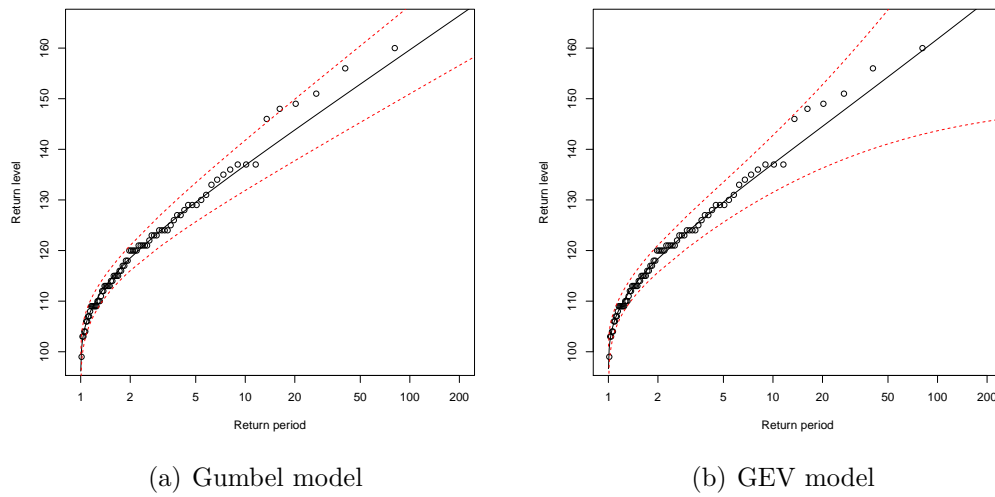
Figures 2.16 and 2.17 show that both models are quite similar, not at all surprising of course, given the similarity of scale and shape parameters and the GEV shape parameter quite close to zero. The actual difference on return levels is also small, with a 2.7 cm difference between 200 year return levels as shown in tables 2.11 and 2.12.

As for the fit, the diagnostic plots in figures 2.16 and 2.17 show that some of the topmost are above the model line, with two points slightly above the upper Gumbel confidence bound. The highest maximum is reasonably well fitted in both models.

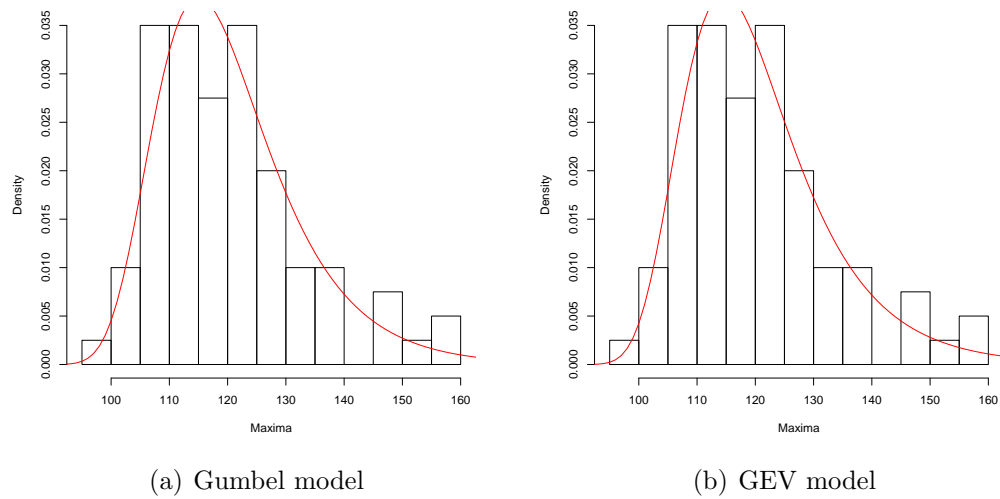


**Figure 2.16:** Probability and quantile plots for the annual maxima methods, Tregde

Density plots in figure 2.18 seem to indicate a general good fit, but perhaps some slight underestimation at the top.



**Figure 2.17:** Return level plots for the annual maxima methods, Tregde



**Figure 2.18:** Density plots for the annual maxima methods, Tregde

### 2.4.7 Andenes

For Andenes, we only have annual maxima from 1992 to 2010, i.e. 19 data points. We are therefore not very likely to get a very good estimate of the true distribution of annual maxima here. Due to the small amount of available data, there are no available published values either.

For the Gumbel model, parameter estimates are  $\hat{\mu} = 279.3(3.1463)$  and  $\hat{\sigma} = 12.871(2.0883)$ . For the GEV model,  $\hat{\mu} = 284.54(3.2207)$ ,  $\hat{\sigma} = 12.902(3.0972)$  and  $\hat{\xi} = -0.82098(0.21323)$ .

We see that the estimated shape parameter for the GEV model has a very large negative value. Since it is below -0.5, the maximum likelihood estimators are no longer regular [1], and lack the usual asymptotic properties. And both table 2.14 and figure 2.20(b) show that confidence intervals inexplicably narrow in as we near the longer return periods. This is certainly unrealistic. Furthermore, the model is bound at 300.3 cm, 0.2 cm above the 200 year return level for the GEV model.

The probability and quantile plots in figure 2.19 and the return level plots in figure 2.20 show that the GEV model is better fitted to the maxima than the Gumbel model, but since the data set we should be careful in trusting either.

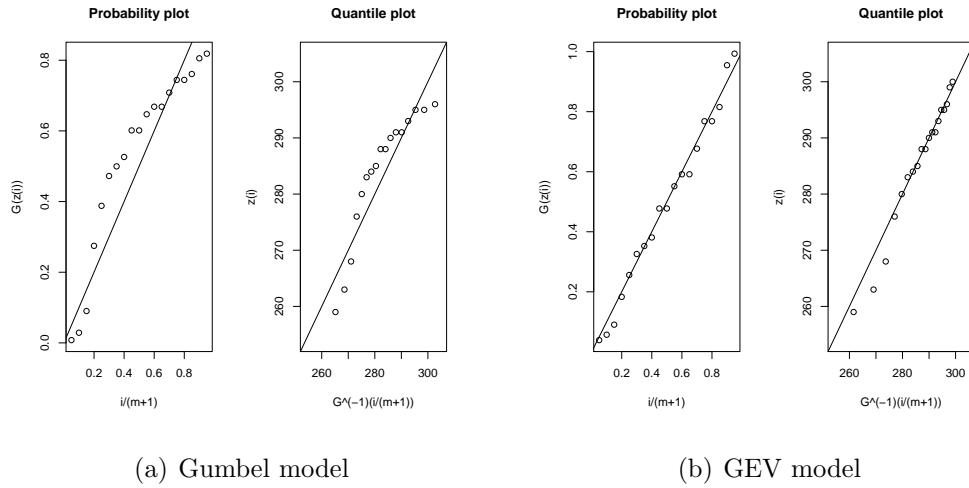
In the density plots in figure 2.21, the GEV model also seems to better fitted.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	-	288.5	298.6	308.7
10	-	295.5	308.2	321
20	-	302.1	317.5	332.9
100	-	316.7	338.4	360.1
200	-	322.9	347.3	371.7

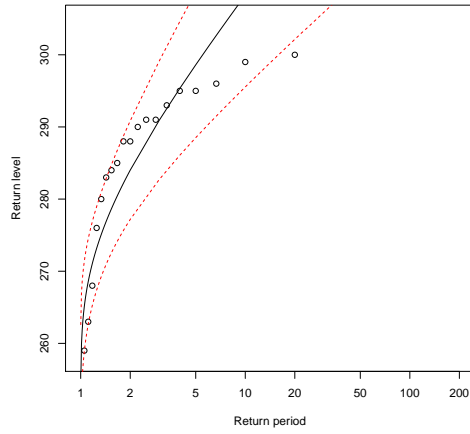
**Table 2.13:** Estimates of return levels and confidence intervals for Andenes, Gumbel model. Values in cm.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	-	292.2	295.7	299.1
10	-	295.5	297.8	300
20	-	297.4	298.9	300.3
100	-	298.8	299.9	301
200	-	298.8	300.1	301.3

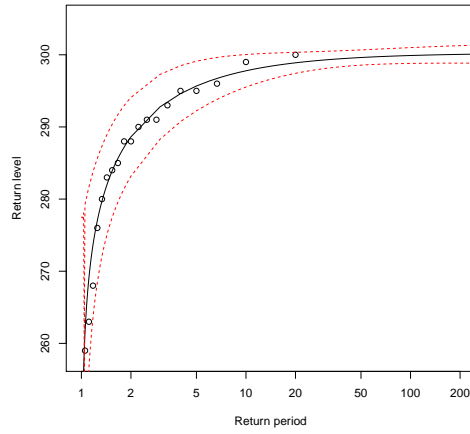
**Table 2.14:** Estimates of return levels and confidence intervals for Andenes, GEV model. Values in cm.



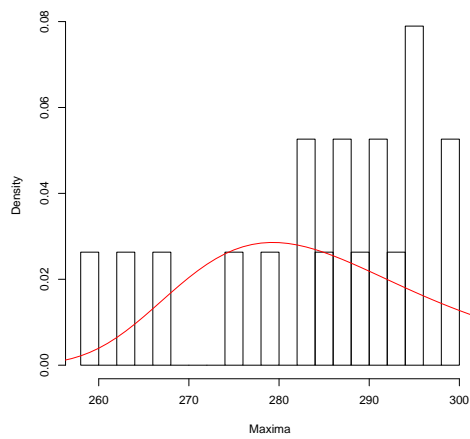
**Figure 2.19:** Probability and quantile plots for the annual maxima methods, Andenes.



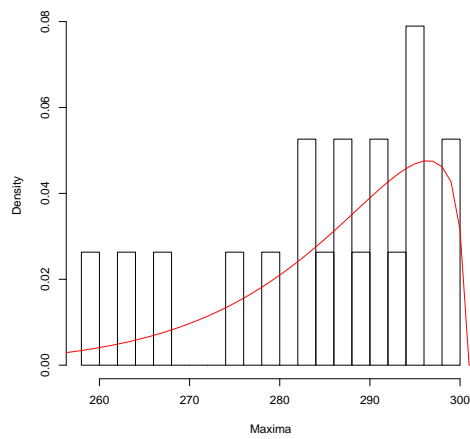
(a) Gumbel model



(b) GEV model

**Figure 2.20:** Return level plots for the annual maxima methods, Andenes.

(a) Gumbel model



(b) GEV model

**Figure 2.21:** Density plots for the annual maxima methods, Andenes.

### 2.4.8 Viker

Viker has, like Andenes, few years of observations available for the estimation of the Gumbel and GEV parameters. We have maxima for the years 1991 to 2010, i.e. 20 years and thus one year more than Andenes. For the Gumbel model, parameter estimates are  $\hat{\mu} = 147.03(3.1069)$  and  $\hat{\sigma} = 13.198(2.3857)$ . For the GEV model,  $\hat{\mu} = 146.61(3.5362)$ ,  $\hat{\sigma} = 12.857(2.7606)$  and  $\hat{\xi} = 0.058715(0.26323)$ . Tables 2.15 and 2.16 show that return levels for both models are very similar, as we expect from the similarity in location and scale parameters and small GEV shape parameter. The 5, 10 and 20 year return levels disagree with their corresponding published values, indicating that some other method than the Gumbel annual maxima method has been used to estimate them.

Figure 2.22 shows that the model is quite well fitted to the annual maxima data for Viker, but we stress again that the amount of data is likely too little to trust the models. We see in 2.23(b) that problems arise for the longer return periods, where the lower bound of the confidence interval behaves erratically. We see this in table 2.16, where the lower bound of the 200 year confidence interval is actually lower than the 5 year lower confidence bound.

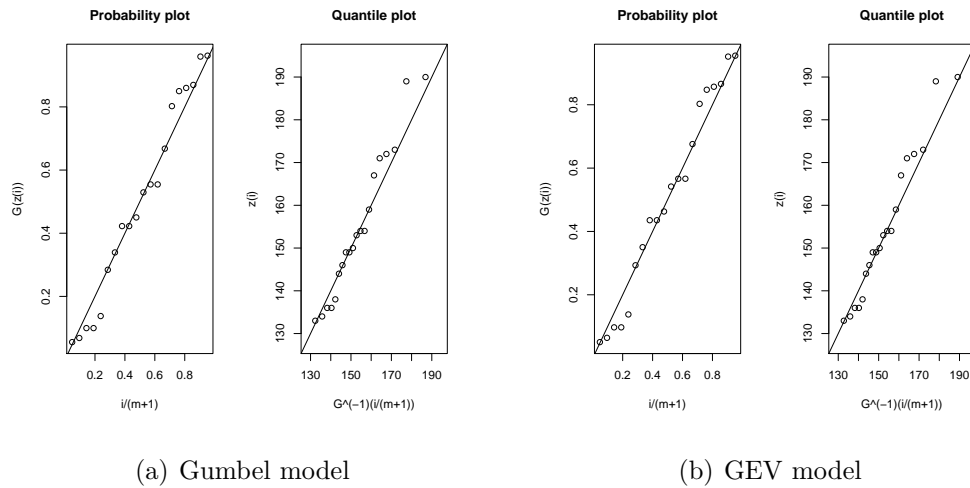
Density plots in figure 2.24 seems to indicate a quite poor fit, with several bars either much higher or much lower than the model line.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	173	156.2	166.8	177.5
10	185	163	176.7	190.4
20	196	169.4	186.2	203.1
100	-	183.7	207.9	232.1
200	-	189.7	217.1	244.4

**Table 2.15:** Estimates of return levels and confidence intervals for Viker, Gumbel model. Values in cm.

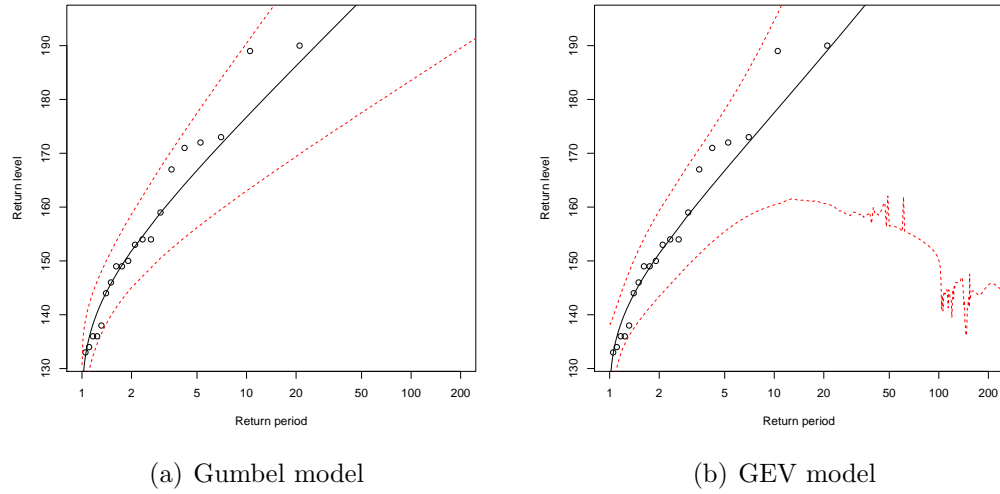
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	173	155.5	166.8	178.1
10	185	160.4	177.5	194.7
20	196	160.5	188.2	215.9
100	-	147.6	212	276.4
200	-	139.7	220.7	301.6

**Table 2.16:** Estimates of return levels and confidence intervals for Viker, GEV model. Values in cm.

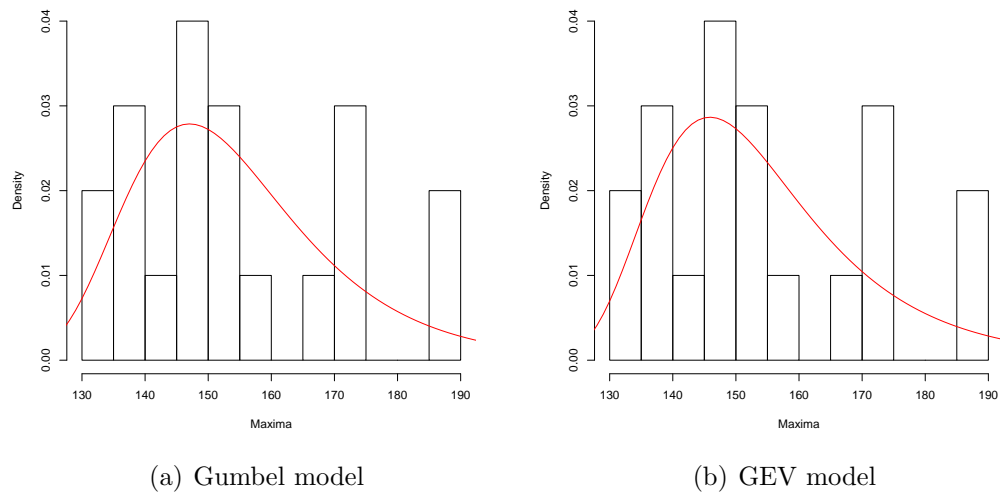


**Figure 2.22:** Probability and quantile plots for the annual maxima methods, Viker.





**Figure 2.23:** Return level plots for the annual maxima methods, Viker.



**Figure 2.24:** Density plots for the annual maxima methods, Viker.

### 3 The peaks-over-threshold (POT) method

#### 3.1 Introduction

A problem with the annual maxima method is the amount of discarded data, which is very wasteful when more frequent observations are available. A more data efficient method is by defining a threshold  $u$ , above which all exceeding observations are deemed extreme.

If the distribution of maxima of some sequence of independent and identically distributed random variables  $X_1, X_2, \dots$  converges asymptotically to a generalized extreme value distribution with parameters  $\mu$ ,  $\sigma$  and  $\xi$ , in the manner shown in (2.1) and (2.2), then the exceedances given by  $Y = X - u$  of some high threshold  $u$ , conditional on  $X > u$ , are approximately independently distributed as

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi} \quad (3.1)$$

defined on  $\{y : y > 0, (1 + \xi y/\sigma) > 0\}$  and called the Generalized Pareto (GP) distribution. The parameters of the GP distribution are determined by the corresponding distribution of block maxima, with  $\tilde{\sigma} = \sigma + \xi(u - \mu)$  and  $\xi$  equal to that of the corresponding GEV distribution [1]. From here on,  $\tilde{\sigma}$  will be referred to as  $\sigma$  for the sake of notational convenience.

#### 3.2 Introducing stationarity

For many physical processes, the assumption of temporal independence is unrealistic. Stationarity is usually a more plausible assumption, and says that even though variables may be dependent upon one another, their stochastic properties are temporally homogeneous. GEV remains an appropriate model for block maxima of stationary series, and the GP distribution can also be shown to remain appropriate for threshold excesses [1].

The amount of stationarity is in some sense quantified in the *extremal index*  $\theta$ , defined on  $0 < \theta \leq 1$ . This can be seen as the tendency for the process to cluster at the extreme levels, and we can informally say the inverse of the extremal index is the limiting mean cluster size. This means that if  $\theta = 0.5$ , then extreme values should approximately arrive in groups of two.

For the block maxima case, stationarity is mostly absorbed in the parameters, which have to be estimated anyway. Some change in practice is necessary for threshold excess modelling, however. A common method to overcome this issue is *declustering*, where the generalized Pareto distribution is instead fitted to the maxima of clusters. These clusters are identified by some empirical rule.

### 3.3 Identification of clusters

The easiest method of cluster identification is taking a threshold and saying that subsequent observations must be above this threshold to be part of the current cluster. A modification can be made by allowing one or more subsequent observations to be below the threshold before the current cluster is left.

Selecting the amount of subsequent observations allowed before a cluster is left is done by selecting a number for  $r$ , the allowed distance between observations above the threshold. For instance, this means that if observation number 50 and observation number 55 are both above the threshold  $u$ , they would be in the same cluster for  $r \geq 5$  since the distance between the observations is 5. For  $r < 5$  they would be considered to be from different clusters.

### 3.4 Threshold selection

To be able to use a POT model in practice, the threshold  $u$  must somehow be located. To do this, we use that if  $Y$  has a GP distribution with scale parameter  $\sigma$  and shape parameter  $\xi$ , then

$$E(Y) = \frac{\sigma}{1 - \xi}, \quad (3.2)$$

for  $\xi < 1$  and infinite mean otherwise. Equation (3.2) tells us that if the GP distribution is a valid model for the excesses of a threshold  $u_0$  by elements in the series  $X_1, \dots, X_n$ , then

$$E(X - u_0 | X > u_0) = \frac{\sigma_{u_0}}{1 - \xi}, \quad (3.3)$$

given that  $\xi < 1$ . The parameter  $\sigma_{u_0}$  is the scale parameter corresponding to excesses of  $u_0$ . A GP distribution valid for excesses of  $u_0$  should also be valid for excesses of  $u > u_0$ , given suitable changes to the scale parameter [1]. So for  $u > u_0$ ,

$$\begin{aligned} E(X - u | X > u) &= \frac{\sigma_u}{1 - \xi} \\ &= \frac{\sigma_{u_0} + \xi(u - u_0)}{1 - \xi}. \end{aligned} \quad (3.4)$$

This means that  $E(X - u | X > u)$  is a linear function of  $u$  where  $u > u_0$ . Equation (3.4) therefore says that the expectation should change linearly with  $u$  where the generalized Pareto distribution is appropriate. This can be investigated for a data set by taking the sample mean for different  $u$ :

$$\left\{ \left( u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{(i)} - u) \right) : u < x_{max} \right\}, \quad (3.5)$$

with  $x_{(1)}, \dots, x_{(n_u)}$  the observations exceeding  $u$  and  $x_{max}$  the largest of the  $X_i$ . Plotting this gives the *mean residual life plot*, and should be investigated for linear trends.

Another method is available, based on fitting the model to a wide range of thresholds. In equation (3.4) we used that there is a linear relation between scale parameters  $\sigma_u$  above a valid threshold  $u_0$ . Furthermore, shape parameters  $\xi_u$  should be constant. By reparametrizing the scale parameter to  $\sigma_u^* = \sigma_u - \xi_u u$ , we get a parameter which should also be constant above  $u_0$ . The plot of  $(u, \xi_u)$  and  $(u, \sigma_u^*)$  is hereby called the *stability plot*.

### 3.5 Return level estimation

If the GP distribution is suitable for the exceedances of  $u$  by  $X$ , then for  $x > u$ ,

$$P(X > x | X > u) = \left[ 1 + \xi \left( \frac{x - u}{\sigma} \right) \right]^{-1/\xi}.$$

If we denote  $\zeta_u = P(X > u)$ , we get the level  $x_m$  exceeded on average once every  $m$  observations by solving

$$\zeta_u \left[ 1 + \xi \left( \frac{x_m - u}{\sigma} \right) \right]^{-1/\xi} = \frac{1}{m}. \quad (3.6)$$

Rearranging equation (3.6) gives

$$x_m = u + \sigma [(m\zeta_u)^\xi - 1]/\xi, \quad (3.7)$$

as long as  $\xi \neq 0$ . It can also be shown that for  $\xi = 0$ ,

$$x_m = u + \sigma \ln(m\zeta_u). \quad (3.8)$$

A slight change must be made to (3.7) and (3.8) to accomodate stationarity. Since we fit the model to the  $n_c$  maxima of clusters instead of all exceeding observations  $n_u$ , we have reduced the number of observations used in the model fitting. We now have  $n_u \theta$  effective observations instead of  $n_u$ . The  $m$ -observation return level is therefore

$$x_m = u + \sigma [(m\zeta_u \theta)^\xi - 1]/\xi, \quad (3.9)$$

for the case of  $\xi \neq 0$ , and

$$x_m = u + \sigma \ln(m\zeta_u \theta), \quad (3.10)$$

for  $\xi = 0$ . Finally, we also make a small notational adjustment. It is much more convenient to talk about return periods in terms of years  $N$  than number of observations, so we say that  $m = N \cdot n_y$ .  $n_y$  is then the (average) number of

observations of year. For the sake of hourly observations  $n_y = 24 \cdot 365.25 = 8766$ , leap years then being taken into account. This gives

$$x_N = u + \sigma[(Nn_y\zeta_u\theta)^\xi - 1]/\xi, \quad (3.11)$$

for the case of  $\xi \neq 0$ , and

$$x_N = u + \sigma \ln(Nn_y\zeta_u\theta), \quad (3.12)$$

for  $\xi = 0$ .

### 3.6 Parameter estimation

The probability of an observation  $x$  to be larger than the threshold  $u$ ,  $\zeta_u = P(X > u)$ , is estimated by the ratio of number of points exceeding  $u$  divided by the number of observations,

$$\hat{\zeta}_u = \frac{n_u}{n},$$

while the extremal index is estimated by the numbers of clusters above  $u$  divided by the number of points above  $u$ ,

$$\hat{\theta} = \frac{n_c}{n_u}.$$

An estimate of  $\zeta_u\theta$  is therefore  $n_c/n$ .

The parameters of the generalized Pareto distribution are estimated by maximum likelihood. From (3.1) we derive the log-likelihood

$$l(\sigma, \xi) = -k \ln \sigma - (1 + \frac{1}{\xi}) \sum_{i=1}^{n_c} \ln(1 + \xi \frac{y_i}{\sigma}), \quad (3.13)$$

defined where  $(1 + \sigma^{-1}\xi y_i) > 0$  and with  $y_1, \dots, y_{n_c}$  the maxima of the  $n_c$  clusters of exceedances of  $u$ . For  $\xi = 0$  we get the log-likelihood

$$l(\sigma) = -k \ln \sigma - \sigma^{-1} \sum_{i=1}^{n_c} y_i. \quad (3.14)$$

In practice, the parameters are estimated with the maximum likelihood procedure in the `fpot` function in the `evd` library in R. Standard errors for the GP parameters come from this R function as well, and are extracted from a numerical approximation of the observed information. Confidence intervals for the parameters are then calculated by using the approximate normality of the maximum likelihood estimator. The uncertainty in  $\zeta_u$  is ignored since it is usually small compared to the errors of the other parameters [1].

### 3.7 Model checking

There are several available plots to check the quality of the fit of the data to a generalized Pareto model. First, we can look at the probability plot, which consists of the points

$$\{(i/(k+1), \hat{H}(y_{(i)})); i = 1, \dots, k\},$$

with

$$\hat{H}(y) = 1 - \left(1 + \frac{\hat{\xi}y}{\hat{\sigma}}\right)^{-1/\hat{\xi}}$$

for the case of  $\hat{\xi} \neq 0$  and

$$\hat{H}(y) = 1 - \exp\left(-\frac{y}{\hat{\sigma}}\right)$$

for  $\hat{\xi} = 0$ .

The quantile plot consists of points

$$\{(\hat{H}^{-1}(i/(k+1)), y_{(i)}); i = 1, \dots, k\},$$

where

$$\hat{H}^{-1}(y) = u + \frac{\hat{\sigma}}{\hat{\xi}}[y^{-\hat{\xi}} - 1],$$

for  $\hat{\xi} \neq 0$ , and

$$\hat{H}^{-1}(y) = u + \hat{\sigma} \ln(1/y),$$

for  $\hat{\xi} = 0$ . Both probability and quantile plot should be approximately linear.

The return level plot consists of points  $\{(N, \hat{x}_N)\}$ , where  $\hat{x}_N$  is (3.11) with estimated parameters when  $\hat{\xi} \neq 0$  and (3.12) with parameter estimates when  $\hat{\xi} = 0$ .

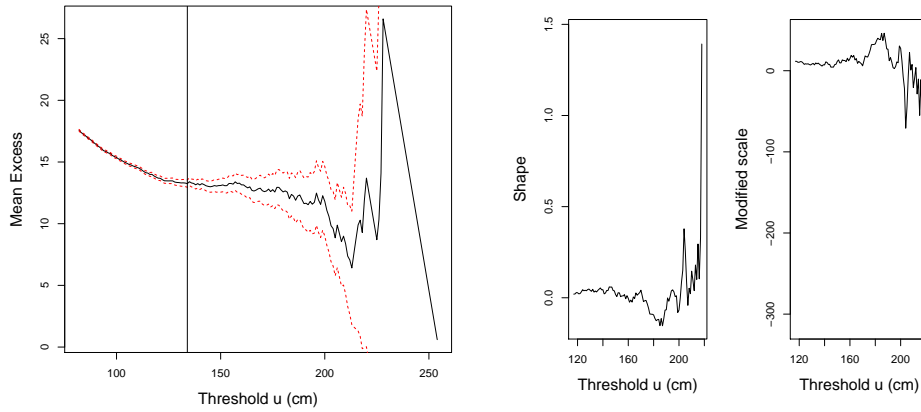
We can also look at the density function of the fitted model for the range of the threshold exceedances, and compare with a histogram of the exceedances.

### 3.8 Application to water level measurements

#### 3.8.1 Oslo

Figure 3.1(a) shows the mean residual life plot for Oslo. To select an appropriate threshold, we search for approximate linearity within confidence bounds. The vertical line shows where we mean to have found this, at 134 cm. The figure doesn't paint a completely clear picture, as is the case with most real data sets, but from around the indicated spot and up to around 200 cm there is a certain level of linearity.

Figure 3.1(b) shows the effect of threshold selection on the model parameters. Ideally, the shape and modified scale parameters should be invariant to threshold change as long as we are above the minimum threshold. We can see that this is basically the case from the beginning of the plot and up to about 170-175 cm. A threshold of 134 cm is within this range, and cannot be rejected based on figure 3.1(b).



(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

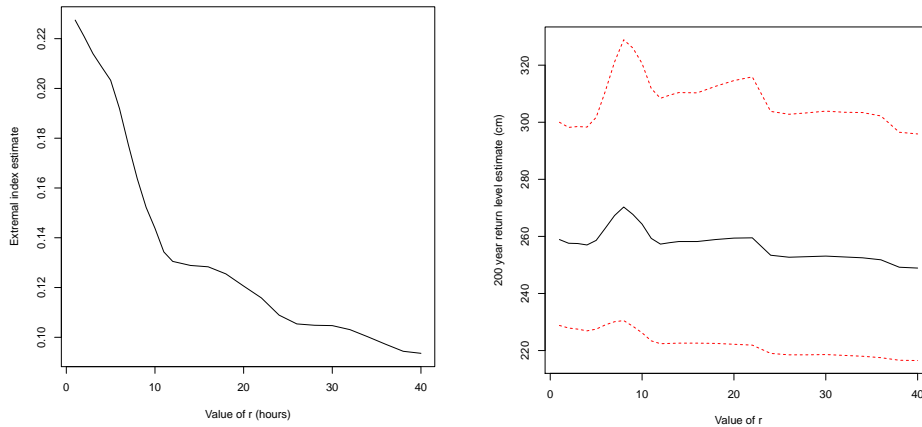
**Figure 3.1:** Threshold selection plots for the POT model, Oslo.

Having selected a threshold, we go on to select the allowed distance between points in a cluster, our  $r$ . Figure 3.2(a) shows the behaviour of the extremal index as we increase the allowed distance between observations in a cluster. The extremal index declines quite sharply until about  $r = 12$ , then almost flattens out before declining faster again. A similar tapering in the steepness of the curve is seen at about  $r = 24$ , and possibly at about  $r = 36$ .

Since each observation is spaced an hour between, these changes in steepness coincide with the 12 hour cycle of the lunar tidal component. It is therefore safe

to assume that extreme storms above  $u = 134$  may last from one high tide to another. This suggests selecting a value of  $r$  larger than 12 to encapsulate the full length of such storms. But it doesn't seem plausible to have  $r$  much larger, since if the tide was the dampening factor, then the subsequent rising tide should bring the storm with it up to extreme levels again.

More insight into the choice of  $r$  is granted by looking at figure 3.2(b), which shows the development of the 200 year return level as we increase  $r$  from 1 to 40. The return level stays approximately constant from  $r = 1$  to  $r = 5$ , but increases relatively sharply from there and up to  $r = 8$ . It then decreases just as sharply, before staying relatively constant from  $r = 12$  up to about  $r = 22$ . It then drops sharply, before staying constant up to  $r = 36$  again. For  $r \in (1, 5)$  and  $r \in (12, 22)$ , the return level is approximately the same. Having argued that  $r$  shouldn't be much larger than 12, we therefore choose  $r = 1$  since this gives approximately the same result while using more data.



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

**Figure 3.2:** Change in extremal index and 200 return level as calculated by a POT model, for Oslo.

With the threshold at 134 cm, the model parameters were estimated to be  $\hat{\xi} = 0.01923(0.02547)$  and  $\hat{\sigma} = 13.993(0.4939)$  by the `fpot` function in the `evd` package in *R*. From the 778199 points in the data series, 7364 points are above the chosen threshold, giving an estimate of  $\zeta$  of  $\hat{\zeta} = 0.00946$ . From the 7364 points above the threshold, 1675 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.22746$ . All this gives the return level estimates shown in table 3.1.

The listed published values are not those featured on the Vannstand.no [9], but rather those collected from a previous GEV analysis performed by the authorities



[3]. We see that values are quite similar, but the difference increases with longer return periods.

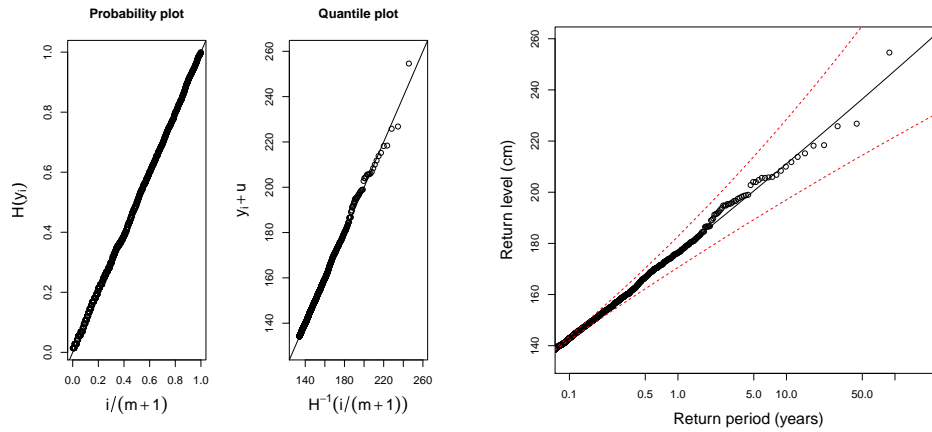
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	179	170.6	176.3	182.7
5	195	189.3	200.5	213.9
10	207.3	197	211.1	228.5
20	218.6	204.7	221.9	243.8
100	241.8	221.7	247.6	282.2
200	250.8	228.8	258.9	300

**Table 3.1:** Return level estimates and confidence intervals for the POT model, Oslo. Values in cm.

Figure 3.3(a) shows probability and quantile plots for the model compared with actual data. The points represent the maxima of the identified clusters. The probability plot doesn't reveal any problems with the model, since the points and the line seem to coincide perfectly. The quantile plot also shows points that are mostly on the line, but a couple of points are not fully located on it. In summary, figure 3.3(a) shows that the model with estimated parameters fits very well to the maxima of the identified clusters.

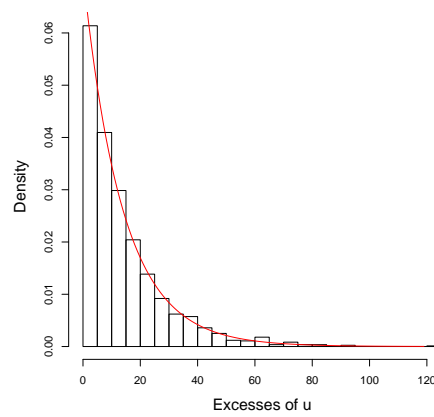
The next plot is the return level plot in figure 3.3(b). All points are contained within confidence bounds, and most of them stay on or very close to the line. Points above 185 cm behave somewhat more erratically than those below, but not to a dramatic extent. This level corresponds approximately to the level in which the stability of parameters, as displayed in figure 3.1(b), begins to wear off.

Finally, the density plot in figure 3.3(c) shows a good fit for the model to the excesses of the threshold  $u = 134$ .



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).

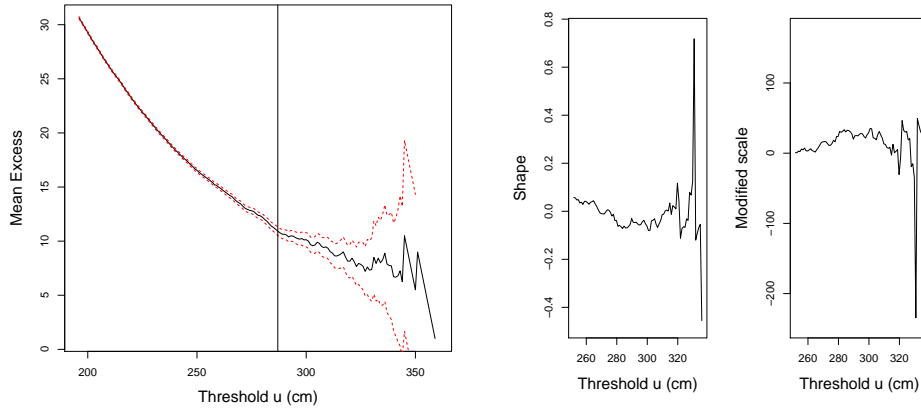


(c) Histogram of excesses versus model density.

**Figure 3.3:** Model diagnostics for the POT model, Oslo.

### 3.8.2 Heimsjø

Figure 3.4(a) shows the mean residual life plot for the Heimsjø data. Approximate linearity is found around the vertical line, which is placed at  $u = 287$ . The stability plot in figure 3.4(b) shows the effect on the model parameters of threshold change around the 287 cm mark. Around and above 280 cm up to around 325-330 cm, parameters are approximately constant.



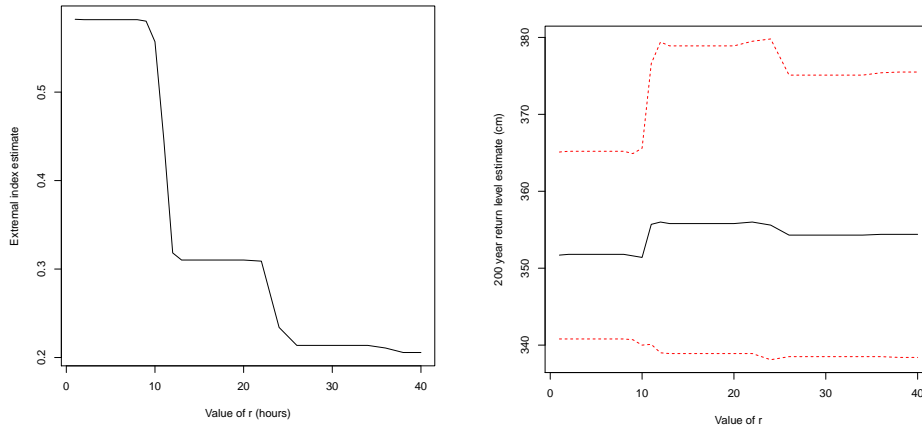
(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

**Figure 3.4:** Threshold selection plots for the POT model, Heimsjø.

We continue by finding an appropriate value of  $r$  for the identification of independent clusters and the estimation of the extremal index. Figure 3.5(a) shows how the extremal index estimate behaves. It is close to piecewise constant, with significant drops at around  $r = 10$  to  $12$  and at  $r = 24$ . Figure 3.5(b) shows that there is quite little difference between the estimate for  $r = 1$  and larger  $r$ . The difference is somewhat larger than we found in Oslo, however, and we choose an  $r$  to include storms crossing from one tide to another. We therefore set  $r = 13$ .

With the 287 cm threshold, model parameters are estimated as  $\hat{\sigma} = 13.709(0.65293)$  and  $\hat{\xi} = -0.11634(0.030375)$ . From the 686867 points in the Heimsjø data set, 2363 points are above 287 cm, giving an estimate of  $\hat{\zeta} = 0.0034403$ . From the points above the threshold, 733 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.3102$ .

Return level estimates are presented in table 3.2. They go from slightly overshooting to slightly undershooting the published return levels, but all are certainly within confidence bounds.



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

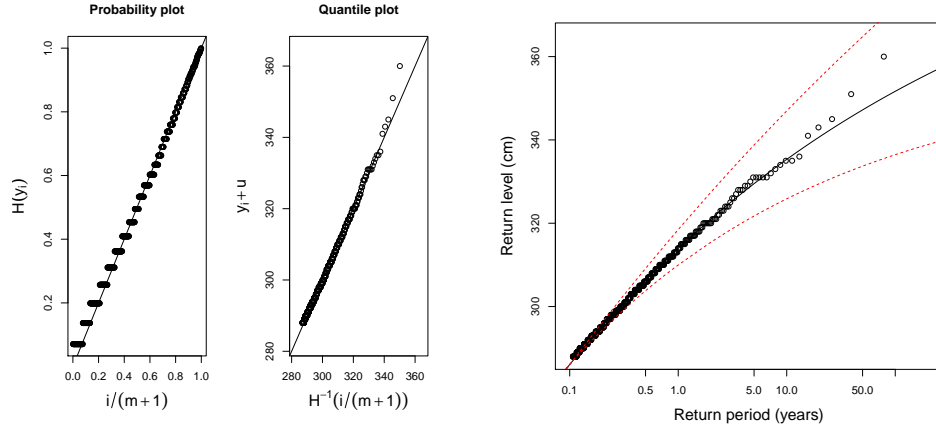
**Figure 3.5:** Change in extremal index and 200 return level as calculated by a POT model, for Heimsjø.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	312	310	314	318.5
5	329	321.7	329.5	338.8
10	337	325.9	335.3	347
20	345	329.5	340.7	354.8
100	-	336.5	351.7	372
200	-	338.9	355.8	378.9

**Table 3.2:** Return level estimates and confidence intervals for the POT model, Heimsjø. Values in cm.

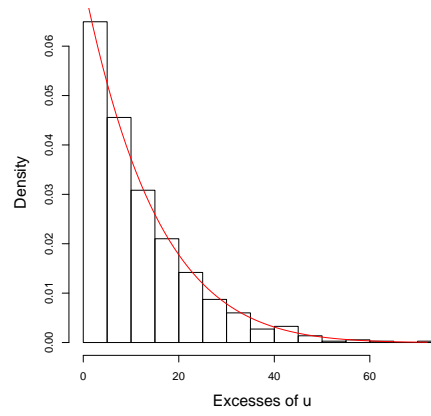
To see how well the model fits to the maxima of the identified clusters, we first look at figure 3.6(a). For the highest sea levels, there are some issues with the model fit. This is better seen by looking at the return level plot in figure 3.6(b), which shows irregularities after about 330 cm - corresponding to the level after which parameters in figure 4(b) are no longer near-constant. The five highest points are above the model line, and it seems possible that return levels for long return periods can be underestimated.

The density plot in figure 3.6(c) doesn't reveal any new bits of information, but shows an overall good fit as far as density is concerned.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).

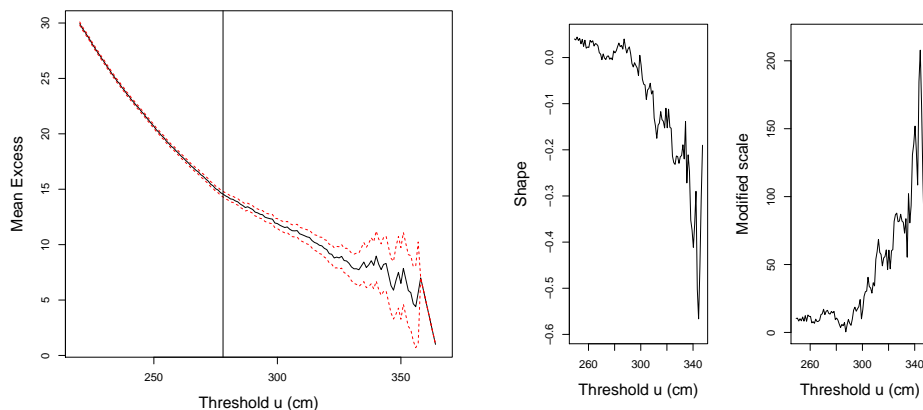


(c) Histogram of excesses versus model density.

**Figure 3.6:** Model diagnostics for the POT model, Heimsjø.

### 3.8.3 Honningsvåg

The mean residual plot in figure 3.7(a) shows tendencies of linearity within confidence bounds slightly before the 280 cm mark. We select a threshold of 278 cm, indicated by the vertical black line. The stability plot in figure 3.7(b) is far from constant up to around 275-280 cm, but from there it is reasonable to call the parameters near-constant up to perhaps 310 cm. The threshold  $u = 278$  is therefore barely within the acceptable area as far as the stability plot is reckoned.



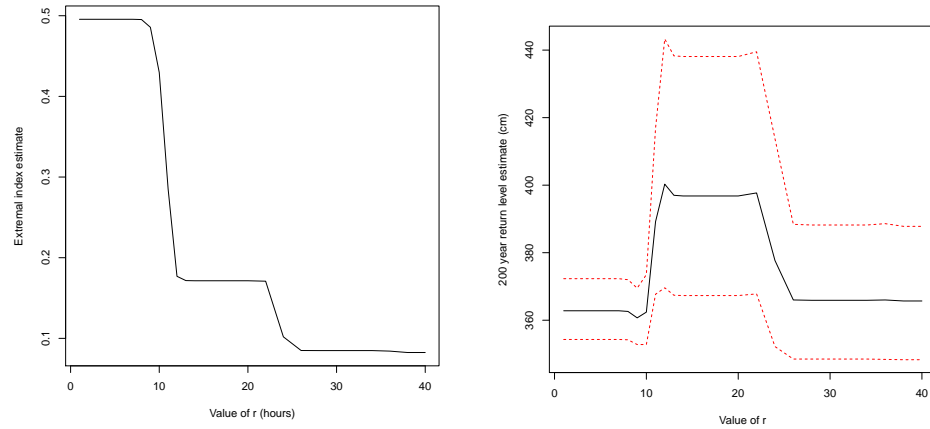
(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

**Figure 3.7:** Threshold selection plots for the POT model, Honningsvåg.

To find an appropriate level of  $r$ , we first turn to figure 3.8(a). Honningsvåg is tide-dominant like Heimsjø, and the same pattern is shown here as there; the extremal index plot is approximately piecewise constant. Figure 3.8(b) shows a similar pattern as that for Heimsjø as well, but the differences in return levels are much larger. The values for  $r \leq 10$  and  $r > 24$  are about the same, however. The difference is large enough to not be within confidence intervals, and we choose  $r = 13$  since we earlier argued for a  $r \geq 12$ .

A threshold of 278 cm gives estimated model parameters of  $\hat{\sigma} = 14.692(0.52962)$  and  $\hat{\xi} = -0.024785(0.025902)$ . There are 330883 sea level measurements in the Honningsvåg data, 9264 of which above 278 cm. This makes for an estimate of  $\hat{\zeta} = 0.027998$ . From the points above the threshold, 1590 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.17163$ .

Return levels are featured in table 3.3, where there is little difference between the POT estimates and published values.



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

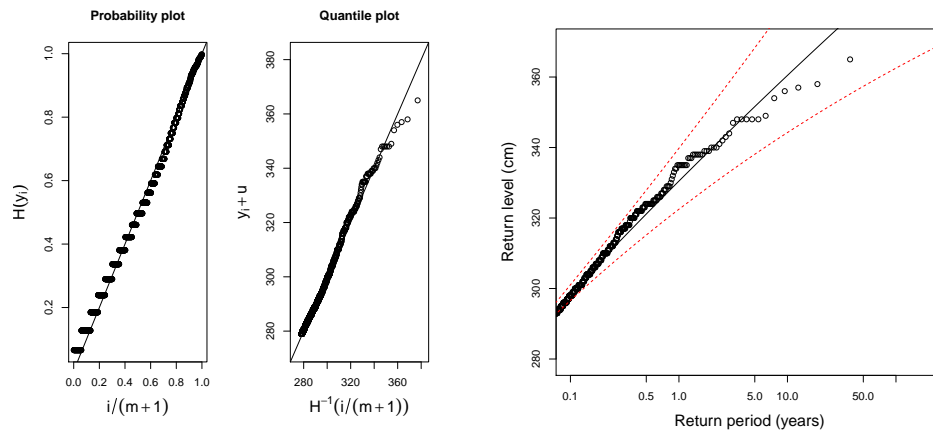
**Figure 3.8:** Change in extremal index and 200 return level as calculated by a POT model, for Honningsvåg.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	330	322.5	330.5	339.8
5	348	338.1	351.6	368.3
10	357	344.2	360.5	380.9
20	366	350.1	369.2	393.8
100	-	362.5	388.8	424.6
200	-	367.4	397	438.3

**Table 3.3:** Return level estimates and confidence intervals for the POT model, Honningsvåg. Values in cm.

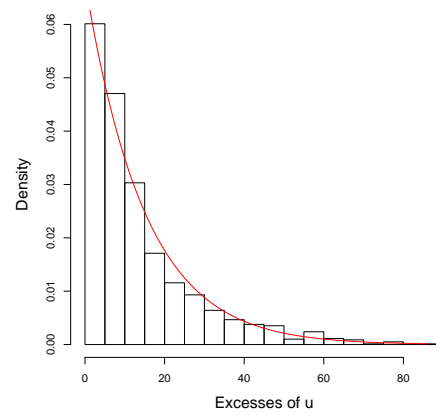
The quantile plot to the right in figure 3.9(a) shows that a few points at the top are off the line, with figure 3.9(b) giving a clearer impression. It looks likely that higher return levels may be overestimated, although all points are within confidence intervals.

The density plot in figure 3.9(c) shows a relatively good fit.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).



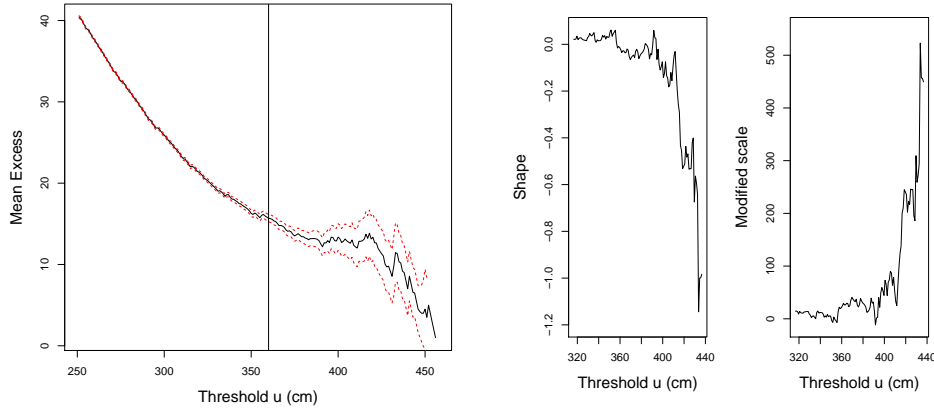
(c) Histogram of excesses versus model density.

**Figure 3.9:** Model diagnostics for the POT model, Honningsvåg.



### 3.8.4 Narvik

In the mean residual life plot in figure 3.10(a), we locate the start of approximate linearity within confidence bounds to about  $u = 360$ . The stability plot in figure 3.10(b) shows that the shape and modified scale parameters are approximately constant around the threshold and up to around 400, and it supports a threshold of 360 cm as reasonable.



(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

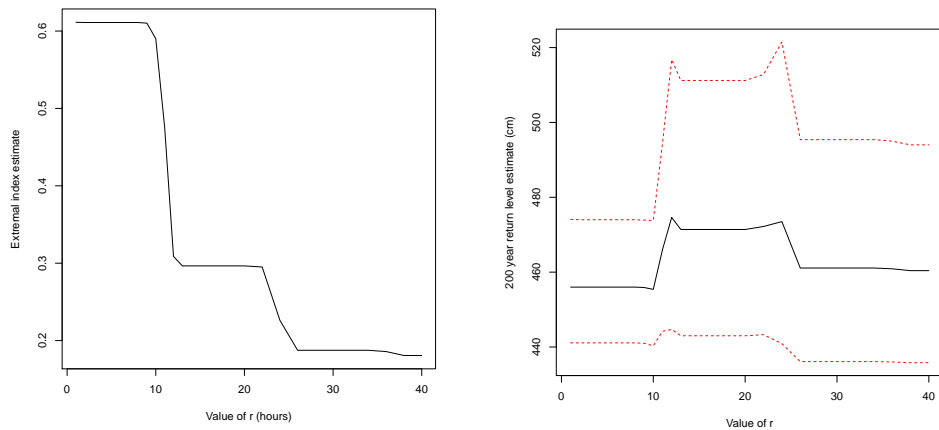
**Figure 3.10:** Threshold selection plots for the POT model, Narvik.

Narvik is tide-dominant, and the plots in figure 3.11 show the same tendencies as those for Heimsjø and Honningsvåg. The difference in return levels isn't as large as in Honningsvåg, and is within confidence intervals. We choose  $r = 13$  like in the previous tide-dominant locations.

Estimates of the model parameters are  $\hat{\sigma} = 17.725(0.76052)$  and  $\hat{\xi} = -0.063107(0.029946)$  with  $u = 390$ . There are 614738 data points available for Narvik, with only 481 above the threshold of 390. This makes for an estimate of  $\zeta$  of  $\hat{\zeta} = 0.0057878$ . From the points above the threshold, 1055 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.29651$ .

The resulting return level estimates are provided in table 3.4. All published values are within their corresponding estimated confidence intervals, with the 5 and 10 year estimates particularly close to official values.

Probability and quantile plots in figure 3.12(a) show that the fit is relatively good, but with some issues at the highest levels as usual. The return level plot in figure 3.12(b) fortifies this notion with a clearer view of the behaviour of points for



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

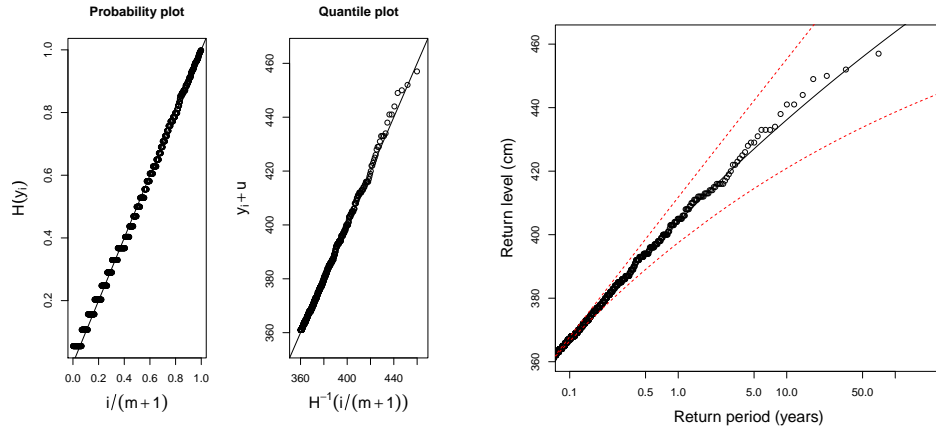
**Figure 3.11:** Change in extremal index and 200 return level as calculated by a POT model, for Narvik.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	402	397.5	404.2	411.8
5	427	414.5	427	442.2
10	439	420.9	436.2	455.3
20	451	426.8	444.9	468.3
100	-	438.6	463.9	498.3
200	-	443	471.4	511.2

**Table 3.4:** Return level estimates and confidence intervals for the POT model, Narvik. Values in cm.

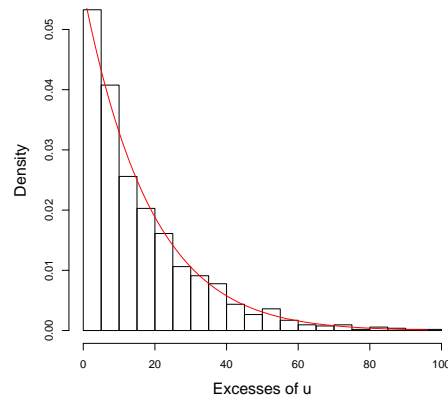
long return periods. Quite a few of the points highest up are off the model line, but they stay well within confidence bounds.

The density plot in figure 3.12(c) completes the picture of a relatively good fit to the maxima of the identified clusters.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).

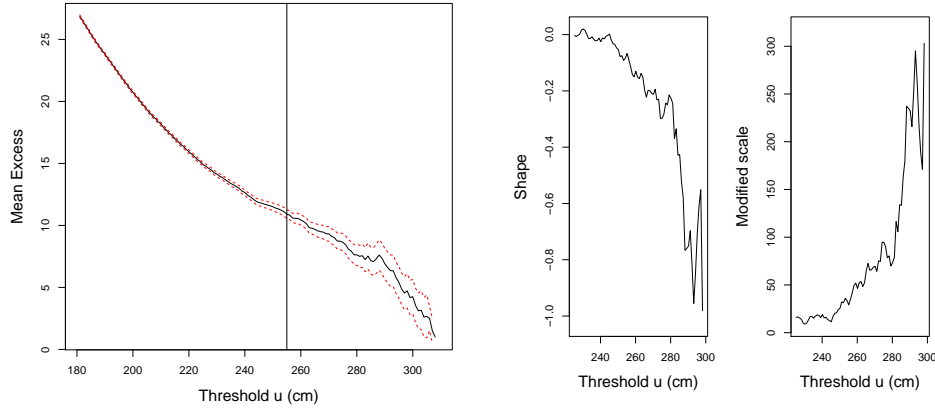


(c) Histogram of excesses versus model density.

**Figure 3.12:** Model diagnostics for the POT model, Narvik.

### 3.8.5 Harstad

Figure 3.13(a) shows the mean residual life plot for the Harstad data. Linearity is a reasonable assumption from the vertical line, located at  $u = 255$ . The stability plot in figure 3.13(b) agrees with this assumption.



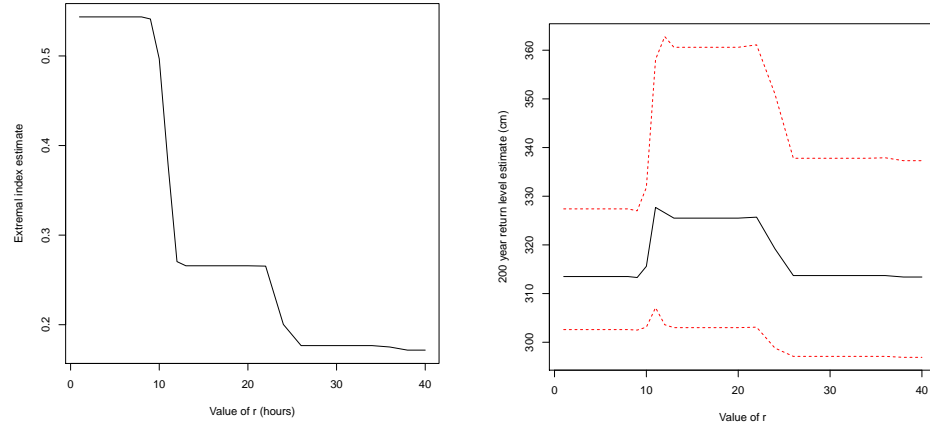
(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

**Figure 3.13:** Threshold selection plots for the POT model, Harstad.

The plots in figure 3.14 show the same pattern as other tide-dominant locations, and we choose  $r = 13$ . With the threshold at 255 cm, the parameter estimates are through maximum likelihood calculated to be  $\hat{\sigma} = 12.342(0.69501)$  and  $\hat{\xi} = -0.084862(0.041516)$ . There are 485855 sea level measurements in the Harstad data set used here, with 2581 points above the threshold  $u = 255$ . This corresponds to  $\hat{\zeta} = 0.0053123$ . From the points above the threshold, 686 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.26579$ .

Return level estimates are found in table 3.5. The 1 and 5 year return levels are very similar to the published values, and the confidence bounds of the 10 and 20 year estimates comfortably contain the corresponding published values.

Figure 3.15(a) shows a couple of poorly fitted points at the top of the range. The return level plot in figure 3.15(b) shows better the behaviour of the model curve at high levels, and we see that it somewhat underestimates the points highest up. The density plot in figure 3.15(c) seems to indicate a good model fit.

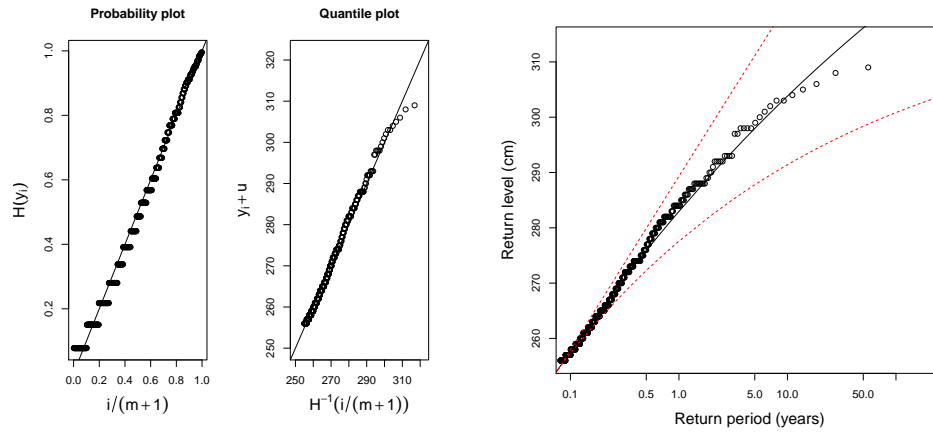


(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

**Figure 3.14:** Change in extremal index and 200 return level as calculated by a POT model, for Harstad.

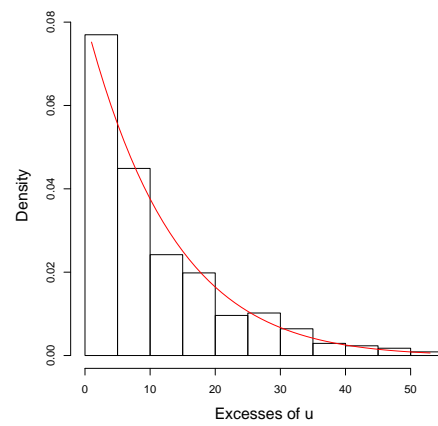
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	281	277.6	283	289.3
5	299	287.8	298	311.1
10	308	291.4	303.8	320.5
20	316	294.6	309.3	329.8
100	-	300.8	321	351.4
200	-	303	325.5	360.6

**Table 3.5:** Return level estimates and confidence intervals for the POT model, Harstad. Values in cm.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).



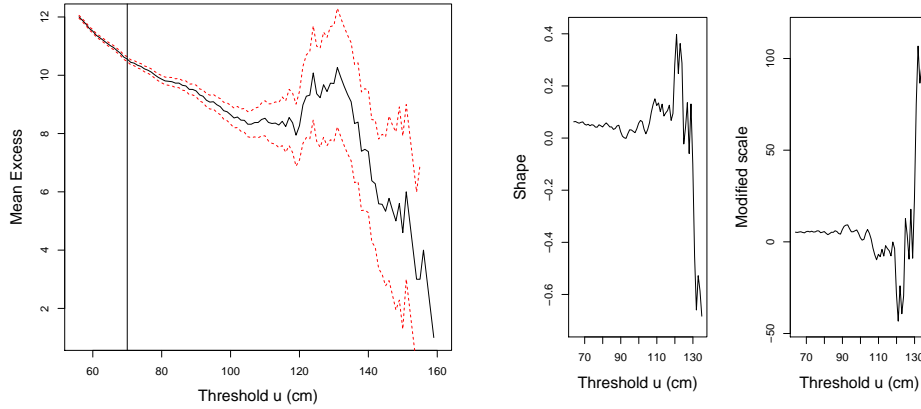
(c) Histogram of excesses versus model density.

**Figure 3.15:** Model diagnostics for the POT model, Harstad.

### 3.8.6 Tregde

For Tregde, the mean residual life plot in figure 3.16(a) is more or less linear within confidence intervals from about  $u = 70$  up to  $u = 125$ . After this there are few points available and the plot behaves erratically.

The stability plot in figure 3.16(b) shows that the parameters are near-constant from right after  $u = 60$  and up to around  $u = 100$ . A threshold of  $u = 70$  is acceptable.



(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

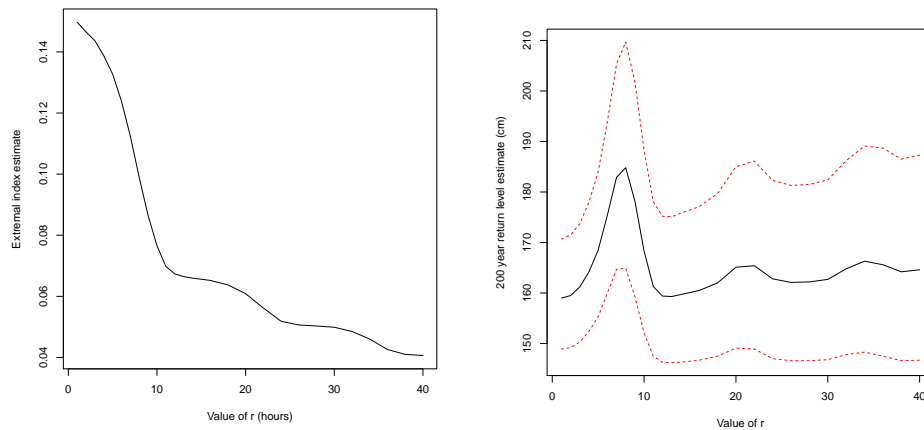
**Figure 3.16:** Threshold selection plots for the POT model, Tregde.

Surge-dominant Tregde has plots in figure 3.17 looking similar to those for Oslo, with more smooth transitions in value of the extremal index than the corresponding plots for tide-dominant locations. Like for Oslo, we see that the 200 year return level for  $r = 1$  is quite similar to those after the spike between  $r = 3$  and  $r = 12$ . We therefore choose  $r = 1$  like we did in Oslo.

A threshold at 70 cm is quite low when looking at the percentage of points still above  $u$ : 51136 points from a total of 700491 points, meaning 7.3 % above and  $\hat{\xi} = 0.07300$ . Estimated parameters are  $\hat{\xi} = -0.04554(0.01027)$  and  $\hat{\sigma} = 11.206(0.1721)$ , while the number of clusters is 7655 and therefore  $\hat{\theta} = 0.14970$ .

With the parameters as presented above, return level estimates are provided in table 3.6. Low levels are very close to their published counterparts.

The quantile plot in figure 3.18(a) shows that the six highest cluster maxima are poorly fitted by this model. This can be a result of fitting with a slightly low threshold, so that there are too many points at relatively low levels. This means



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

**Figure 3.17:** Change in extremal index and 200 return level as calculated by a POT model, for Tregde.

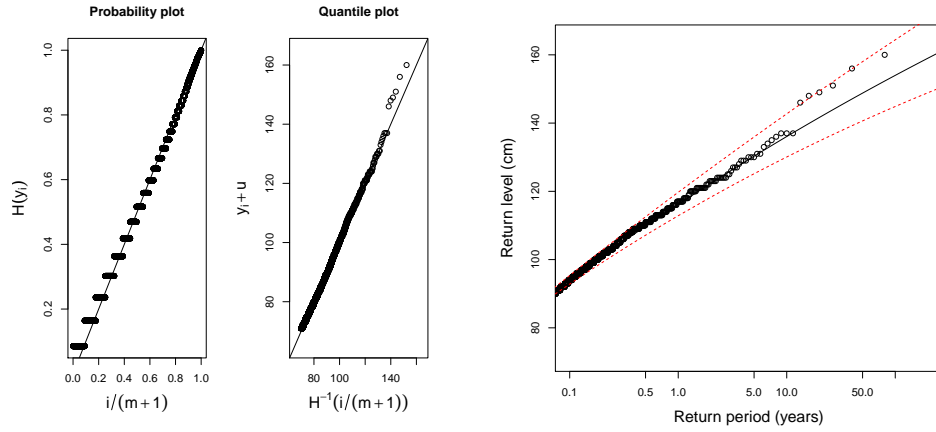
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	114	112.8	116.2	119.7
5	129	125.2	130.3	135.9
10	136	130.1	136.1	142.7
20	143	134.8	141.7	149.4
100	-	144.9	154	164.4
200	-	148.9	159	170.7

**Table 3.6:** Return level estimates and confidence intervals for the POT model, Tregde. Values in cm.

that long period return levels are likely to be somewhat underestimated, at least if the trend of the six topmost points is to be believed.

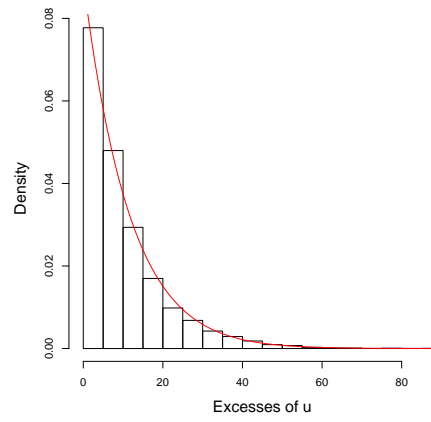
This notion is further supported by the return level plot, shown in figure 3.18(b), where the six aforementioned points touch the upper boundary of the confidence interval. Although no points are definitely outside, they are too close for comfort. The density plot in figure 3.18(c) doesn't offer any new information.





(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).

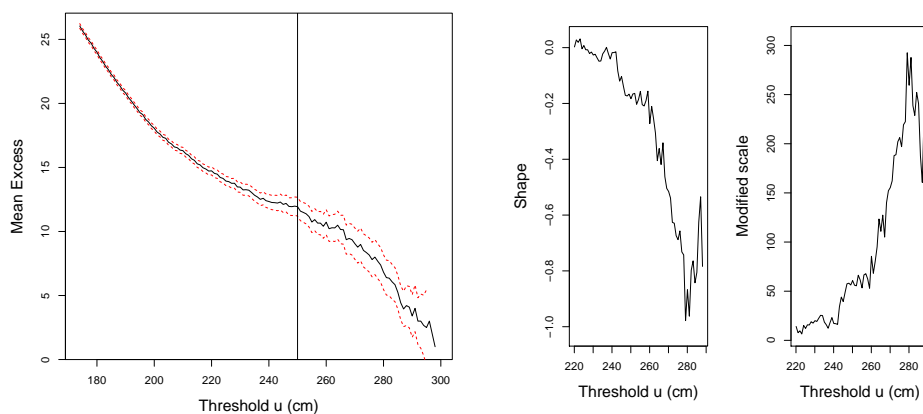


(c) Histogram of excesses versus model density.

**Figure 3.18:** Model diagnostics for the POT model, Tregde.

### 3.8.7 Andenes

In the mean residual life plot for Andenes, shown in figure 3.19(a), a certain level of linearity is found from about  $u = 240$ . The stability of the threshold choice is shown in figure 3.19(b), which shows that parameters are near-constant from about just before 250 to slightly after 260. From 240 to 250 there is a sharp change in parameters, the same is true slightly after 260 cm. A threshold of  $u = 250$  seems likely, and we select this. This threshold is marked by a vertical line in figure 3.19(a).



(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

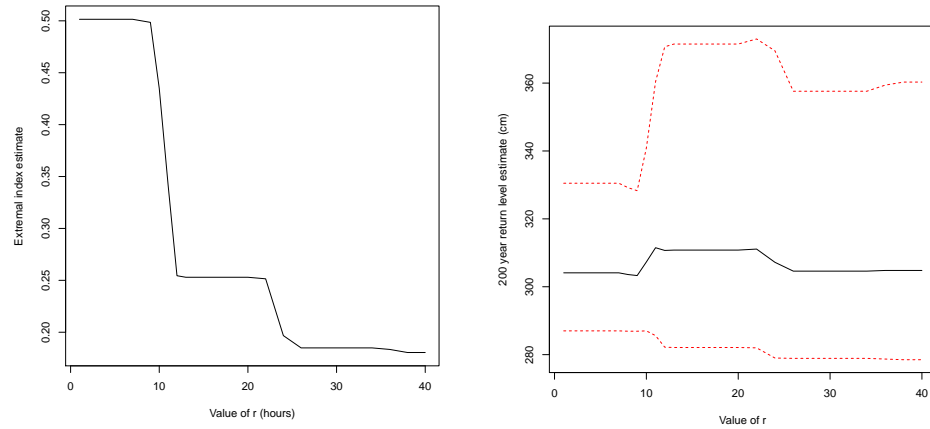
**Figure 3.19:** Threshold selection plots for the POT model, Andenes.

We now wish to select the appropriate  $r$ . Andenes is tide-dominant, and the usual pattern for such locations is seen in the plots in figure 3.20. We select  $r = 13$  as usual for such locations.

With the threshold at 250 cm, the parameters are estimated to be  $\hat{\sigma} = 15.205(1.6561)$  and  $\hat{\xi} = -0.18948(0.07891)$ . There are 169556 sea level measurements in the Andenes data set used here, with 676 points above the threshold. This corresponds to  $\hat{\zeta} = 0.0039869$ . From the points above the threshold, 171 clusters were identified, giving an extremal index of  $\hat{\theta} = 0.25296$ .

Return levels estimated by the above-mentioned parameters are shown in table 3.7. Because of the limited data available at the time that return level analyses were made, no published values are available for comparison. Since there is still not much data available, caution should be applied in trusting the estimates.

The probability and quantile plots in figure 3.21(a) show a generally good fit, though with some points off the line at the top. The return level plot in figure



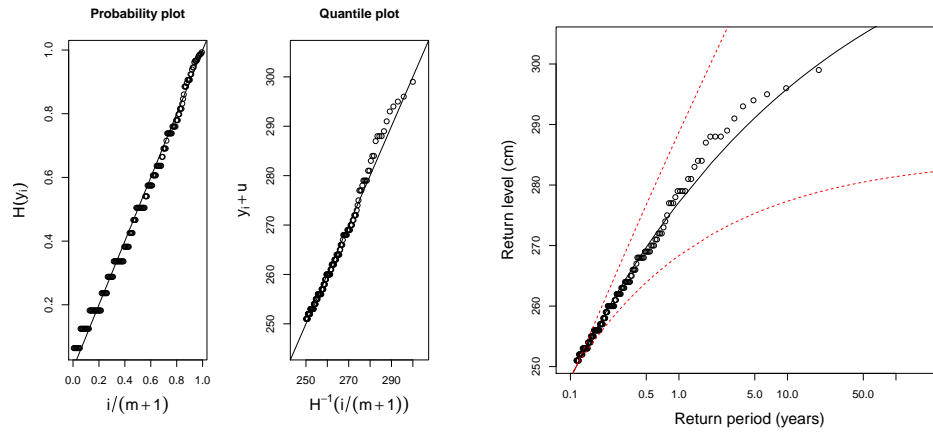
(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

**Figure 3.20:** Change in extremal index and 200 return level as calculated by a POT model, for Andenes.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	-	268.3	277.1	288.7
5	-	275.3	291.1	315.5
10	-	277.3	295.9	326.6
20	-	278.9	300.1	337.4
100	-	281.4	308.1	361.5
200	-	282.1	310.8	371.5

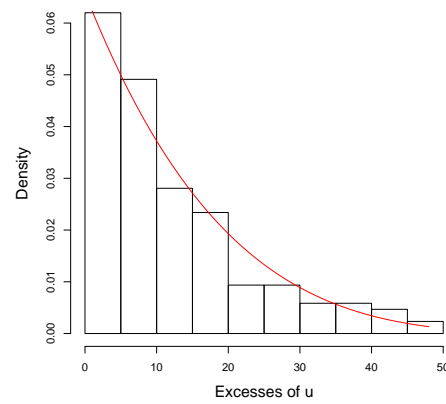
**Table 3.7:** Return level estimates and confidence intervals for the POT model, Andenes. Values in cm.

3.21(b) shows the same thing, but points are well within confidence interval and most points are on the model line. The density plot in figure 3.21(c) shows an overall good fit.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).

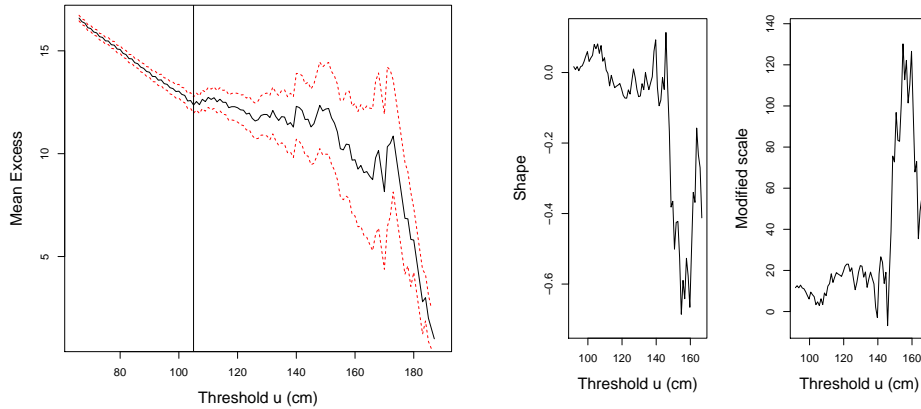


(c) Histogram of excesses versus model density.

**Figure 3.21:** Model diagnostics for the POT model, Andenes.

### 3.8.8 Viker

The mean residual life plot for Viker is approximately linear, at least within confidence bounds, from about  $u = 105$  up to  $u = 175$ . This is shown in figure 3.22(a), where  $u = 105$  is marked by the black line. We review this choice by looking at figure 3.22(b), which shows that there is some fluctuation in the parameters around the 105 cm mark. Still, the parameters are approximately the same when compared to the large fluctuations starting at around 140 cm.



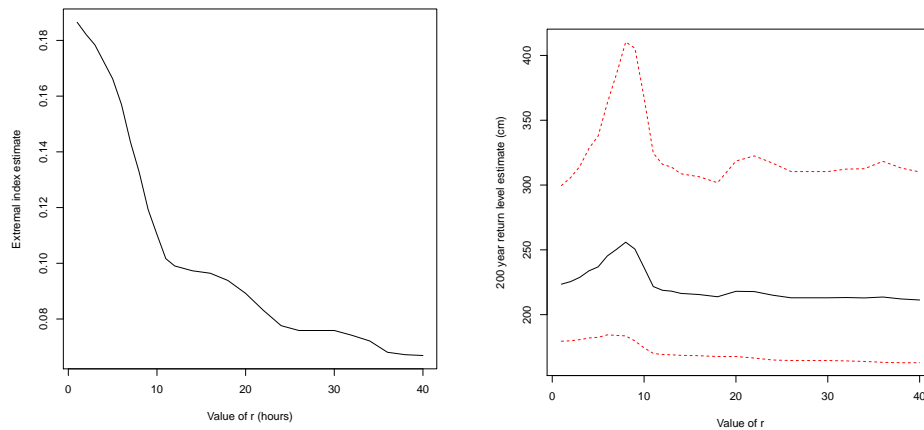
(a) Mean of exceedances of increasing thresholds (cm). (b) Stability of parameters for increasing threshold (cm).

**Figure 3.22:** Threshold selection plots for the POT model, Viker.

As Oslo and Tregde, and as opposed to Heimsjø, Honningsvåg, Narvik, Harstad and Andenes, Viker is surge-dominant. This is reflected in figure 3.23, which shows that 200 year return estimates are similar for  $r = 1$  and  $r > 12$ . As we did in Oslo and Tregde, we select  $r = 1$ . With the 105 cm threshold, model parameters are estimated as  $\hat{\xi} = 0.00594(0.04230)$  and  $\hat{\sigma} = 13.171(0.7614)$ . From the 175873 points in the Viker data, 3453 points are above 105 cm, giving a  $\zeta$  estimate of  $\hat{\zeta} = 0.01963$ . From the points above the threshold, 644 clusters are identified, meaning an extremal index of  $\hat{\theta} = 0.1865$ .

Table 3.8 shows the resulting return level estimates, presented alongside published, official values. The values agree very closely, but it is worth mentioning that the official estimates are based on data for only 1990-2000 and need not be the most dependable estimates.

To see how well the model is fitted to the cluster maxima, we first look at the probability and quantile plots in figure 3.24(a). The fit is very good, with no points far off the line. Secondly we look at figure 3.24(b), which shows that there



(a) Extremal index for  $r$  between 1 and 40. (b) 200 year return level for  $r$  between 1 and 40.

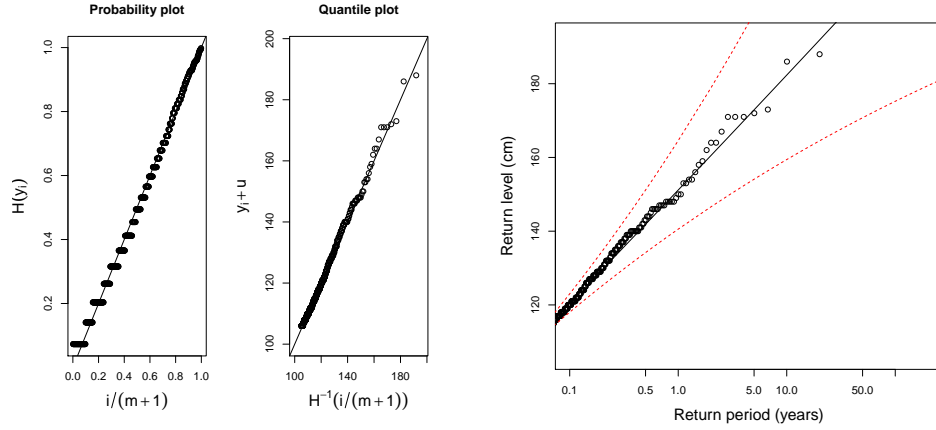
**Figure 3.23:** Change in extremal index and 200 return level as calculated by a POT model, for Viker.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
1	149	140.6	151.2	164.6
5	173	154.1	172.9	199.1
10	185	159.4	182.3	215.6
20	196	164.5	191.8	233.1
100	-	175.2	213.9	278.1
200	-	179.5	223.5	299.6

**Table 3.8:** Return level estimates and confidence intervals for the POT model, Viker. Values in cm.

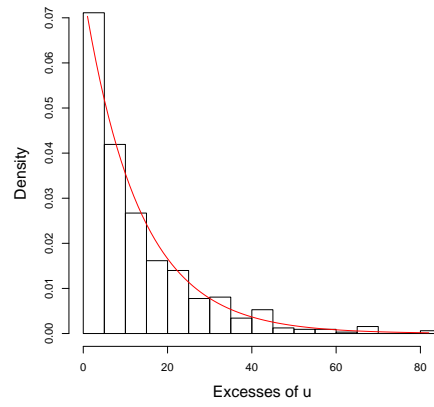
are a few points at the high levels that are off the line, but they are all well within confidence bounds. The density plot in figure 3.24(c) agrees that the fit is good.

In summary, the fit of the model to the cluster maxima is good, even though there are some stability issues around the chosen threshold. Furthermore, the data set is quite limited.



(a) Probability and quantile plot.

(b) Return levels (cm) for increasing return periods (years).



(c) Histogram of excesses versus model density.

**Figure 3.24:** Model diagnostics for the POT model, Viker.

## 4 Revised Joint Probabilities Method

### 4.1 Introduction

The revised joint probabilities method (RJPM) is an attempt by Tawn [7] to improve the joint probabilities method (JPM) employed by Pugh and Vassie. Whereas the latter assume that hourly surge levels are independent, Tawn argues that this is a clearly false assumption. Instead he uses that, provided a weak mixing condition holds, then for a stationary sequence  $X_1, \dots, X_n$ ,

$$P(\max\{X_1, \dots, X_n\} < x) \approx [P(X_1 < x)]^{n^\theta} \quad (4.1)$$

for large  $x$ , and  $0 \leq \theta \leq 1$ .  $\theta$  is the extremal index, and  $\theta^{-1}$  is defined as the limit of  $\theta^{-1}(x)$  as  $x$  tends towards the upper end point of the distribution of  $X$ .  $\theta^{-1}(x)$  is defined as the mean of the distribution of cluster sizes. The extremal index can be equal to 1 for both dependent and independent sequences, when the sequence behaves like an independent sequence at high levels.

The RJPM method relies on componentwise analysis of the sea level  $Z$ . It is divided into three components,

$$Z_t = M_t + X_t + Y_t,$$

where  $M_t$  is the mean sea level,  $X_t$  the tidal level and  $Y_t$  the surge level. For the purposes of this introduction,  $M_t = 0$  for all  $t$ . The tidal level  $X_t$  is estimated by a numerical model.

### 4.2 A distribution for surges

If  $N$  is the number of hours in a year, surge levels  $Y_1, \dots, Y_N$  are taken to be a realization of a stationary sequence. Extreme value theory [1] then dictates that

$$\max\{Y_1, \dots, Y_N\} \sim \text{GEV}(\mu_s, \sigma_s, \xi_s), \quad (4.2)$$

where GEV denotes the Generalized Extreme Value distribution in equation (2.2). Applying the approximation in equation (4.1) to surges, we get

$$P(\max\{Y_1, \dots, Y_N\} \leq y) = F_s(y)^{N^{\theta_s}}. \quad (4.3)$$

for large values of  $y$ , say  $y > u$ .  $F_s$  is the marginal distribution for surges, while  $\theta_s, 0 \leq \theta_s \leq 1$  is the extremal index for surges. Combining equations (4.2) and (4.3) then gives that

$$F_s(y) = \exp \left\{ -(N^{\theta_s})^{-1} \left[ 1 + \xi_s \frac{y - \mu_s}{\sigma_s} \right]^{-1/\xi_s} \right\} \quad \text{for } y > u, \quad (4.4)$$

on  $\{y : 1 + \xi_s(y - \mu_s)/\sigma_s > 0\}$ , with  $\sigma_s > 0$  and arbitrary  $\xi_s$  and  $\mu_s$ . Below the level  $u$ , the empirical distribution function  $\tilde{F}(s)$  is taken as a sufficiently accurate estimate.



### 4.3 Finding the distribution for annual sea-level maxima

If  $T$  is the tidal cycle length in hours, we have[7] that

$$P(\max\{Z_1, \dots, Z_T\} \leq z) = P(Y_t \leq z - X_t \text{ for } t = 1, \dots, T), \quad (4.5)$$

which would equal  $\prod_{t=1}^T F_s(z - X_t)$  if hourly surges were independent. But since they are dependent, the product is modified by an hourly sea level extremal index  $\theta$ :

$$P(\max\{Z_1, \dots, Z_T\} \leq z) = \left\{ \prod_{t=1}^T F_s(z - X_t) \right\}^{\theta}, \quad (4.6)$$

where  $\theta_s \leq \theta \leq 1$ . Combining equations (4.6) and (4.4) gives the RJPM distribution for annual maximum sea levels,

$$G(z) = \exp \left\{ -\theta(T\theta_s)^{-1} \sum_{t=1}^T \left[ 1 + \xi_s \frac{z - \mu_s - X_t}{\sigma_s} \right]^{-1/\xi_s} \right\}, \quad (4.7)$$

where  $z > u + \max\{X_1, \dots, X_T\}$ .

### 4.4 Including interaction between tide and surge

Equation (4.7) shows the case where the surge distribution  $F_s(\cdot)$ , and therefore its parameters  $\mu_s, \sigma_s$  and  $\xi_s$ , is independent of the concomitant tidal level. Unfortunately, the assumption that tide and surge are independent processes is poor in shallow water areas, where turbulent frictional processes on the sea bed cause the tide and surge components to interact. This causes effects such as surge values at high tides being damped and surges on the rising tide being amplified [2]. These effects vary from site to site however, so we will attempt to model the interaction on the residuals from the observed surges.

#### 4.4.1 Testing for interaction

The tidal range from lowest observed tide (LAT) to highest observed tide (HAT) is split into  $n_b$  equi-probable bands, i.e. each band has an equal amount of observed measurements. For each of the tidal observations, there is a concurrent surge observation. This means that if there are  $n_b$  tidal bands and  $n_{obs}$  observations in total, there are  $n_{obs}/n_b$  observations of both the tide and surge in each band.

If the tide and surge were independent, an equal amount of points should also be expected to exceed a given level  $u$ . But if there is interaction, then the least number of points should be in the top band where surges are damped. Similarly, the largest number should be in the middle bands where surges are magnified. The

amount of discrepancy between bands will therefore be a quantifiable measure of the tide-surge interaction.

We choose  $u$  to be a high empirical quantile of the surge distribution,  $z_q$  for  $q = 0.9975$ . For the independent case, there should then be  $(1 - q) \cdot n_{obs}/n_b = e$  observations in each band. We choose  $n_b = 5$  as in [2], meaning  $n_{obs} \cdot 0.0025/5 = e$  observations per tidal band. In actuality, since independence is a flawed assumption, we observe  $N_i$  surges per band, for  $i = 1, \dots, n_b$ .

To put the interaction into a quantifiable setting, we use a standard  $\chi^2$  test statistic,

$$\chi^2 = \sum_{i=1}^5 \frac{(N_i - e)^2}{e}. \quad (4.8)$$

If  $N_i \approx e$ , then  $\chi^2$  will be small. Tide and surge are deemed to interact with 95 % confidence if the test statistic is above the associated 4 degrees of freedom table value of  $\chi_{4,0.95}^2 = 9.488$ .

#### 4.4.2 Modelling the tide-surge interaction

We adopt the method used in [2], where the surge series is location-scale normalised by

$$S_t^* = \frac{Y_t - a(X_t)}{b(X_t)}, \quad (4.9)$$

where  $a(X_t)$  and  $b(X_t)$  are some tide-dependant functions.  $\{S_t^*\}$  is then supposed stationary, and we can use established methods to estimate the associated model parameters  $\mu_{s^*}$ ,  $\sigma_{s^*}$  and  $\xi_{s^*}$ . The parameter estimates for the original surge series are then given by

$$\mu_s(X) = \mu_{s^*}b(X) + a(X), \quad \sigma_s(X) = \sigma_{s^*}b(X), \quad \text{and } \xi_s(X) = \xi_{s^*}. \quad (4.10)$$

Equation (4.7) is then modified to

$$G(z) = \exp \left\{ -\theta(T\theta_s)^{-1} \sum_{t=1}^T \left[ 1 + \xi_{s^*} \frac{z - \mu_{s^*}b(X_t) - a(X_t) - X_t}{(\sigma_{s^*}b(X_t))} \right]^{-1/\xi_{s^*}} \right\}, \quad (4.11)$$

for  $z > u + \max\{X_1, \dots, X_T\}$ . Below this threshold the empirical distribution  $\tilde{F}_{s^*}([y - a(X)]/b(X))$  is a sufficient approximation.

#### 4.4.3 Estimating the normalisation functions $a(X)$ and $b(X)$

As in the interaction test in section 4.4.1, the tidal range is divided into  $n_b$  bands, with equal numbers of tidal and concomitant surge observations in each. These bands are separated by points  $x_0, x_1, \dots, x_{n_b}$ , such that each band is given

by  $[x_0, x_1), [x_1, x_2), \dots, [x_{n_b-1}, x_{n_b}]$ . A point  $X_t$  is then in the  $i$ th tidal band if  $x_{i-1} \leq X_t < x_i$ . LAT corresponds to  $x_0$  and HAT to  $x_{n_b}$ .

We then define  $a_{1,i}$  and  $a_{2,i}$  for  $i = 1, \dots, n_b$  as the 98% and 99% empirical quantiles of the empirical surge distribution in band  $i$ . Finally, we denote the midpoints of each band as  $m_i = (x_{i-1} + x_i)/2$  for  $i = 1, \dots, n_b$ .

With the above definitions, we define estimates of  $a(X)$  and  $b(X)$  as  $\hat{a}(X) = \hat{a}_1(x)$  and  $\hat{b}(X) = \hat{a}_2(x) - \hat{a}_1(x)$ , where

$$\hat{a}_j(X) = \begin{cases} a_{j,1} & \text{for } LAT \leq X < m_1 \\ a_{j,i} \frac{X-m_{i-1}}{m_i-m_{i-1}} + a_{j,i-1} \frac{X-m_i}{m_{i-1}-m_i} & \text{for } m_{i-1} \leq X < m_i \\ a_{j,n_b} & \text{for } m_{n_b} \leq X \leq HAT \end{cases}, \quad (4.12)$$

with  $i = 2, \dots, n_b$ . In this sort of interaction modelling, we need to select a suitable number of bands  $n_b$ . As baseline number of bands we take  $n_b = 5$ , as adopted in [2]. We will review this choice for each data set, however. The selection of number of bands is a classic case of bias versus variance, as we need  $n_b$  to be large to reduce bias where as variance then increases.

## 4.5 Estimating return levels with RJPM

### 4.5.1 Estimation of surge distribution parameters

In the 1992 and 1994 reports[7][2] the  $r$ -largest method is applied to estimate the parameters of the surge distribution. This method is a modification of the GEV annual maxima method, and uses the  $r$  largest values per year to estimate model parameters.

We choose instead to use the threshold used in the estimation of  $\theta_s$  and apply the POT method to estimate model parameters, a method which has the advantage of using much more data. We use the fact that for a POT model,  $\xi_{POT} = \xi_{GEV} = \xi$  for the corresponding GEV model, with  $\sigma_{POT} = \sigma_{GEV} + \xi \cdot (u - \mu_{GEV})$  [1].

In practice, the shape and scale are estimated with a maximum likelihood procedure by the function `fspot` in the *R* package `evd`, while the location parameter is calculated by the `fgev` function. This is also the case for the confidence intervals of model parameters. We fit the model to the maxima of clusters, where the clusters are identified by the same rule that governs the estimation of  $\theta$ .

### 4.5.2 Estimation of extremal indices

The extremal indices  $\theta$  and  $\theta_s$  require two choices each; a threshold above which clusters are counted and an empirical clustering rule which says how many non-exceedances are allowed before we have left the current cluster. For the first choice, we use the quantile that was found in the POT analysis in section 3. There a

thorough analysis was done on where the threshold should be placed, and it makes sense using this same threshold to estimate  $\theta$ . Applying the procedure of using the same quantiles for both extremal indices, as was done in [2], we get corresponding thresholds for the surge series and the estimation of  $\theta_s$  as well.

For the second choice, we need to decide what constitutes an independent storm. In [2], they found that  $r = 30$  was a good choice, and that the ratio of extremal indices wasn't too dependent on this choice in any case. In this analysis the results are more sensitive to this choice, since we opt to use the POT method with clustering to estimate model parameters. We still choose to use this  $r$  value of 30.

### 4.5.3 Estimation of return levels

A practical challenge with the RJPM method is that (4.11) is impossible to solve analytically for  $z$ . We start out by defining  $G(z_p) = 1 - p$  and get the relation in (4.13)

$$-T\theta_s\theta^{-1}\ln(1-p) = \sum_{t=1}^T \left[ 1 + \xi_{s^*} \frac{z_p - \mu_{s^*}b(X_t) - a(X_t) - X_t}{(\sigma_{s^*}b(X_t))} \right]^{-1/\xi_{s^*}}, \quad (4.13)$$

but we cannot go further by algebraic manipulation. We therefore develop a numerical procedure where the right side of the equation is calculated for a range of relevant  $z$  and then matched to the left side of the equation.

The tidal cycle length is 18.61 years, meaning that  $T = 18.61 \cdot 8766 = 163135.3 \approx 163135$ . Any span of consecutive 163135 observations should therefore approximately contain a full tidal cycle, and in our numerical procedure we take the last 163135 observations in the tidal series.

The numerical procedure is written in R, and the relevant algorithms and functions are found in the appendix of this thesis.

## 4.6 Application to water level measurements

Location	Ratio of tidal and observed maxima	Ratio of mean annual maxima
Oslo	0.424	0.562
Heimsjø	0.822	0.904
Honningsvåg	0.871	0.925
Narvik	0.832	0.906
Harstad	0.861	0.903
Tregde	0.434	0.567
Andenes	0.849	0.873
Viker	0.430	0.514

**Table 4.1:** Ratio of tidal versus observed total maxima and ratio of means of tidal and observed annual maxima.

### 4.6.1 Oslo

We begin by observing some characteristics of the data sets, seen in table 4.1. We see that the maximum of the tidal values is only about 42% as large as the maximum of the measured values. We also see that the mean of annual maxima for tidal values divided by the mean of annual maxima for measured values is 0.562. These values show that a large percentage of the observed sea levels arise from surges.

Locations with such characteristics are called surge-dominant, as opposed to tide-dominant where most of an observed sea level stems from the current tidal level. Areas where surge dominates usually do not have as much tide-surge interaction as tide-dominant areas, but we still use the  $\chi^2$  test in (4.8) to quantify the amount.

By using 5 tidal bands and cutting the data at the 99.75% quantile, we expect 389 surges per band. Instead we get 346, 380, 398, 431 and 386 points per band, in order from lowest tide level to the highest. This is not a very bad result, but there is still evidence of significant tide-surge interaction. The test statistic has a result of  $\chi^2 = 9.778$ , compared to the table value of  $\chi_{4,0.95}^2 = 9.488$ .

Since the  $\chi^2$  test shows a significant level of interaction, we try to model it using 5 tidal bands. Figure 4.1(a) shows the tide against surge data to the left, and tide against transformed surge data to the right, both showing only points above the 99.75% quantile. There seems to be very little difference between the left and the right plots except for the scales on the  $x$  axes. The left hand side of figure 4.1(b) shows the reason for this, where the correction vectors  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$  are almost vertical. The former varies between 51 and 52.38, while the

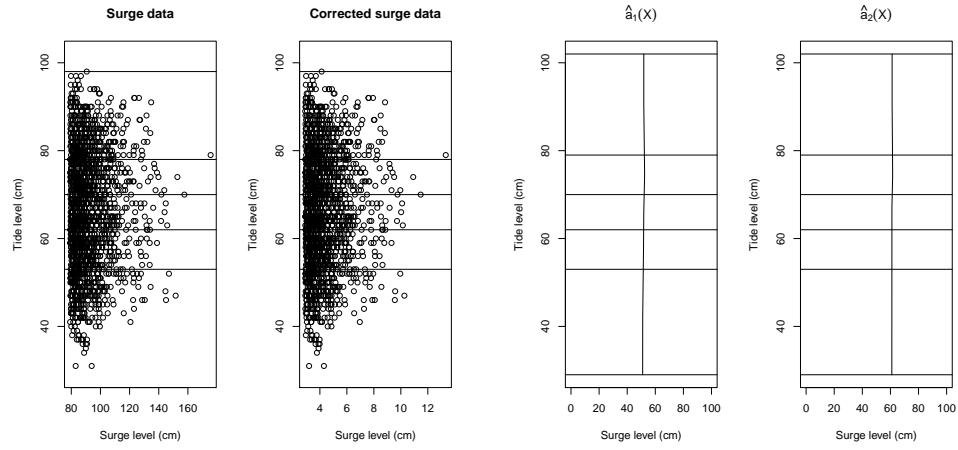
latter varies between 61 and 61.59. And in fact, the corrected surge performs even poorer on the  $\chi^2$  test, with a score of  $\chi^2 = 14.270$ .

Figure 4.1(c) shows hows the ratio between the  $\chi^2$  test statistic and the corresponding  $\chi^2_{n_b-1,0.95}$  value develops after 2-30 bands have been used to transform the surge data.  $\chi^2/\chi^2_{n_b-1,0.95} = 1$  is indicated by the horizontal line, and ratios above this correspond to  $\chi^2$  tests showing significant interaction. We see that for  $n_b = 2$  and  $n_b = 15$  the interaction is insignificant, but for  $n_b = 2$  the interaction is insignificant for the untransformed data as well, with a test statistic of 1.85 versus the corresponding table value of 3.84. Transforming the data with  $n_b = 15$  seems needlessly complex for a model where there hardly was significant interaction in the original test. Furthermore, there is little actual difference in return levels. The same is true if we compare them to the return levels achieved using no tide-surge correction. For parsimony we therefore select the model without correction, based on (4.7) instead of (4.11). Estimated parameters are  $\hat{\mu}_s = 95.655(2.3134)$ ,  $\hat{\sigma}_s = 17.576(1.0531)$ ,  $\hat{\xi}_s = -0.0488(0.0389)$ ,  $\hat{\theta} = 0.0788$  and  $\hat{\theta}_s = 0.0679$ .

Resulting return levels and confidence intervals are shown in table 4.2. They are very similar to the published values.

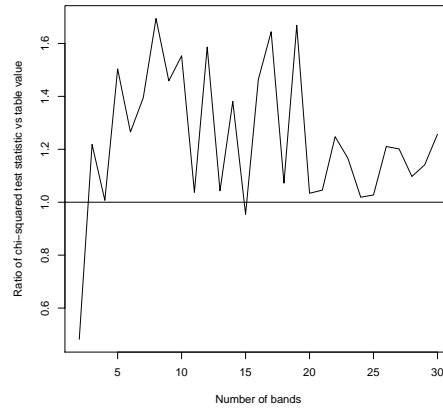
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	195	184.1	194.5	207.6
10	207.3	192.5	206.7	225.2
20	218.6	200	217.9	242.5
100	241.8	215	242	282.9
200	250.8	220.7	251.8	300.7

**Table 4.2:** Return levels and 95% confidence intervals for the RJPM model, Oslo. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1,0.95}$ .

**Figure 4.1:** Tide-surge interaction plots for Oslo.

### 4.6.2 Heimsjø

Heimsjø is tide-dominant, and we see in table 4.1 that its characteristics show that there is a quite small distance between the maximum of its tidal series and the maximum of observed sea levels. This also holds for the mean of annual maxima from tide calculations and observations. Tide-dominance usually implies larger tide-surge interaction than surge-dominance, and the  $\chi^2$  test for Heimsjø certainly lives up to this expectation. With 5 tidal bands and cutting at the 99.75% quantile, we expect about 343 surge observations per band. Instead we get 790, 479, 288, 112 and 75 points per band. As expected for a tide-dominant location, the least number of points are in the highest tidal band, where the high tide dampens surges. By far the most points are in the bottom band.

The huge discrepancy between expected and observed number of surges in each band leads to a massively large value for the  $\chi^2$  test statistic, with  $\chi^2 = 1007.3$  compared to the table value of  $\chi_{4,0.95}^2 = 9.488$ . We try transforming the data with 2 to 30 bands, the results of which are shown in figure 4.2(c). The ratio never drops below 1 as we would wish, but instead we choose  $n_b = 12$ , which gives a relatively low ratio compared to other choices. It gives  $\chi^2 = 89.865$  versus the quantile value of  $\chi_{11,0.95} = 19.675$ . Figure 4.2(a) shows that the number of points in each band is much more similar in the transformed case to the right of the figure compared to the left, uncorrected side.

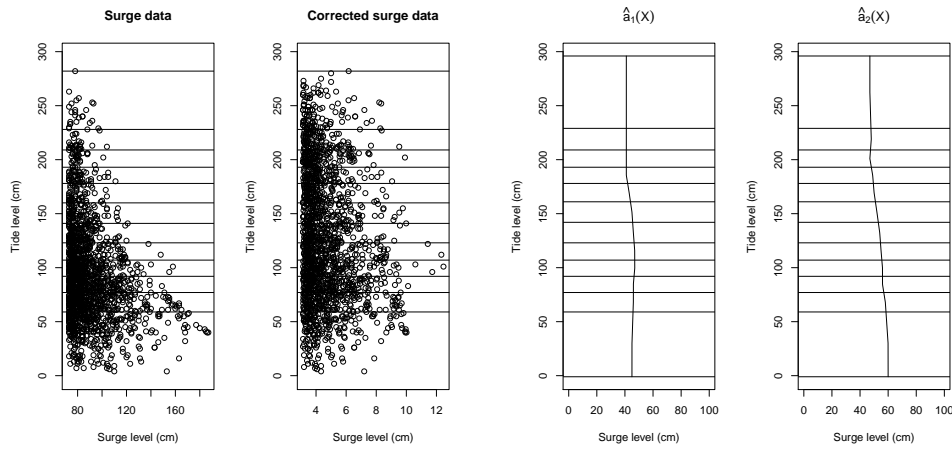
Estimated parameters are  $\hat{\mu}_{s*} = 3.4942(0.2826)$ ,  $\hat{\sigma}_{s*} = 1.6708(0.1604)$ ,  $\hat{\xi}_{s*} = 0.0736(0.0793)$ ,  $\hat{\theta} = 0.1060$  and  $\hat{\theta}_{s*} = 0.0627$ . Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.2(b).

Return level and confidence intervals are shown in table 4.3. The short period return levels are quite different from their published counterparts, but the upper confidence bound of the 20 year return level estimate contains the corresponding published value. This could perhaps be explained by the tide-surge interaction still present, but estimation of return levels without any tidal bands also show this apparent underestimation of short period return levels. The standard error of the estimated  $\xi_s$  larger than the estimate itself, and the estimate being positive means very large upper bounds for long return periods.



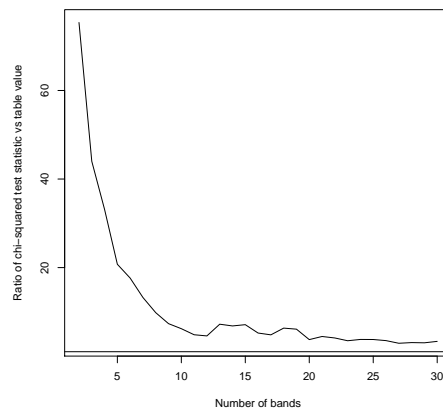
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	329	314.6	318.2	323.5
10	337	321.1	325.4	336.1
20	345	327	333	354
100	-	339.3	352.5	422.4
200	-	344.1	362.1	466.5

**Table 4.3:** Return levels and 95% confidence intervals for the RJPM model, Heimsjø. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1,0.95}$ .

**Figure 4.2:** Tide-surge interaction plots for Heimsjø.

### 4.6.3 Honningsvåg

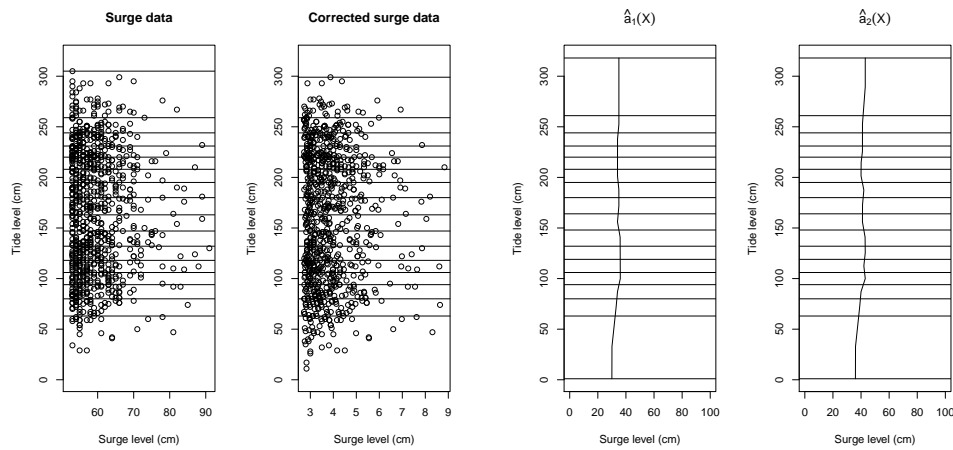
Honningsvåg is tide-dominant like Heimsjø, but the  $\chi^2$  test shows much less interaction than in Heimsjø. It is still significant on a 95% confidence level, with  $\chi^2 = 23.652$ . Figure 4.3(c) shows how the  $\chi^2$  ratio develops. After  $n_b = 2$  it doesn't go below 1 until  $n_b = 25$ , but for  $n_b = 16$  it is quite close and we choose this. Here we have that  $\chi^2 = 26.106$  against  $\chi_{15,0.95}^2 = 24.996$ . Figure 4.3(a) shows the surge data before and after transformation with 16 tidal bands.

Estimated parameters are  $\hat{\mu}_{s*} = 2.9430(0.2738)$ ,  $\hat{\sigma}_{s*} = 1.6169(0.1238)$ ,  $\hat{\xi}_{s*} = -0.0408(0.0495)$ ,  $\hat{\theta} = 0.0848$  and  $\hat{\theta}_{s*} = 0.0428$ . Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.3(b).

Return levels and confidence intervals are shown in table 4.4. As was the case for Heimsjø, the return level estimates for Honningsvåg are lower than the published values.

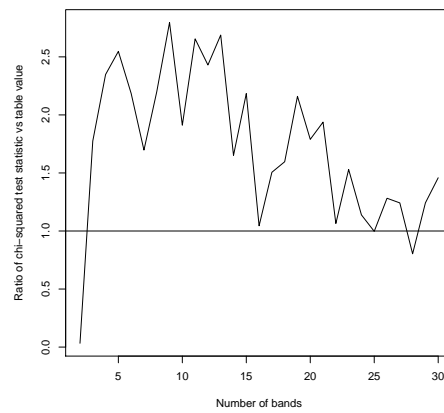
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	348	331.3	332.4	335.6
10	357	339.1	341.9	347.9
20	366	346.1	350.9	360.4
100	-	360.2	370.6	390.7
200	-	365.6	378.7	404.4

**Table 4.4:** Return levels and 95% confidence intervals for the RJPM model, Honningsvåg. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.3:** Tide-surge interaction plots for Honningsvåg.

#### 4.6.4 Narvik

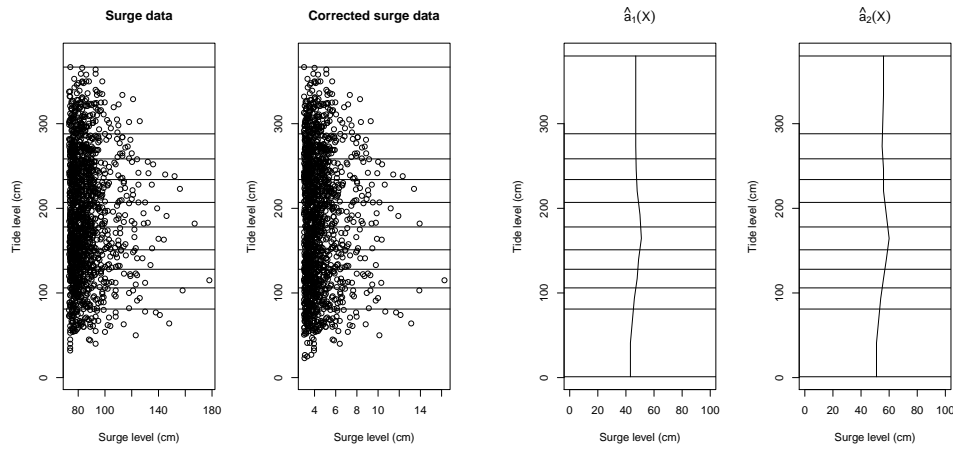
Narvik is tide-dominant, and the  $\chi^2$  test shows a large amount of tide-surge interaction. With 5 tidal bands as usual, we get  $\chi^2 = 71.100$ . Figure 4.4(c) shows the ratio between test and corresponding table value after transforming with 2-30 tidal bands. The ratio drops below 1 at  $n_b = 10$ , and we choose this for the transformation of the surge series. Figure 4.4(a) shows the series before and after correction with 10 bands.

Estimated parameters are  $\hat{\mu}_{s^*} = 4.2648(0.3184)$ ,  $\hat{\sigma}_{s^*} = 1.5218(0.2326)$ ,  $\hat{\xi}_{s^*} = 0.13344(0.1102)$ ,  $\hat{\theta} = 0.3742$  and  $\hat{\theta}_{s^*} = 0.2100$ . Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.4(b).

Table 4.5 shows return levels for 5, 10, 20, 100 and 200 year return periods. Again the return levels seem to be underestimated for low return levels, but the 20 year published value is contained in the confidence intervals of the corresponding estimate. The estimated shape parameter  $\hat{\xi}_{s^*}$  is positive, and the standard error is about the same size, making the upper confidence bound very large.

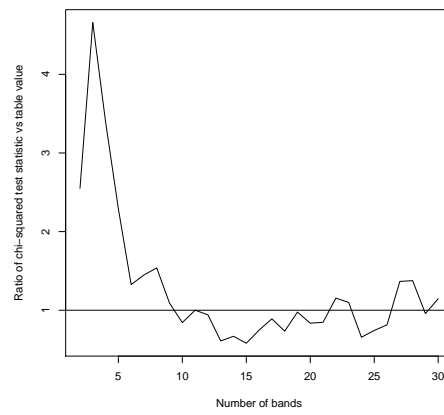
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	427	408.2	416.3	424.1
10	439	417.3	424.8	437.4
20	451	425.6	434.2	461
100	-	443	460.8	568.9
200	-	449.9	474.7	645.9

**Table 4.5:** Return levels and 95% confidence intervals for the RJPM model, Narvik. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.4:** Tide-surge interaction plots for Narvik.

#### 4.6.5 Harstad

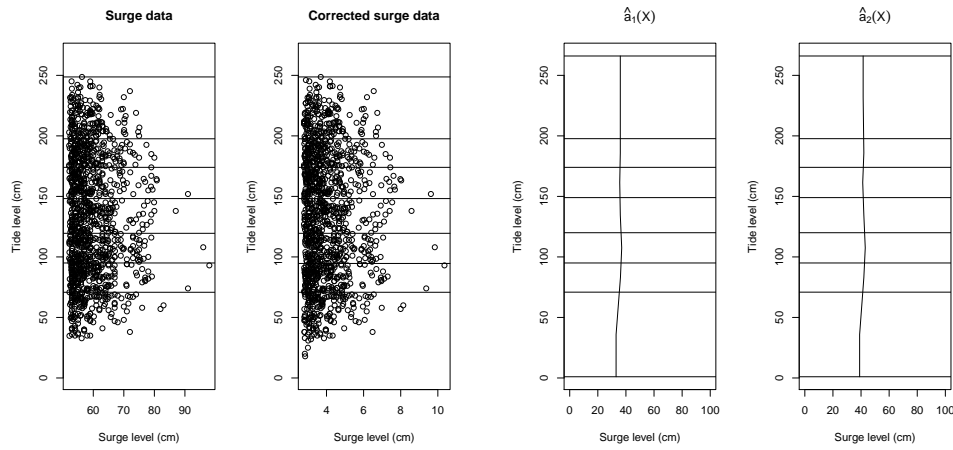
Harstad is another tide-dominant location, with tide-interaction at approximately the level of Honningsvåg. The  $\chi^2$  test statistic has a result of 21.658, and is thoroughly significant. We use figure 4.5(c) for guidance in selection of number of bands, and end up with  $n_b = 7$ . We see that we stay below 1 for all choices of band number after this, but select the smallest  $n_b$  for simplicity of model as well as because it has the lowest ratio. Figure 4.5(a) shows the surge series before and after transformation with 7 tidal bands.

Estimated parameters are  $\hat{\mu}_{s^*} = 2.9972(0.2613)$ ,  $\hat{\sigma}_{s^*} = 1.6034(0.1837)$ ,  $\hat{\xi}_{s^*} = -0.0394(0.0825)$ ,  $\hat{\theta} = 0.1767$  and  $\hat{\theta}_{s^*} = 0.0690$ . Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.5(b).

Table 4.6 shows the return level estimates, which as usual for tide-dominant locations are below those published by the authorities for applicable return periods. The standard error for the shape parameter is very large here, being about the twice the size of the absolute value of the estimate itself.

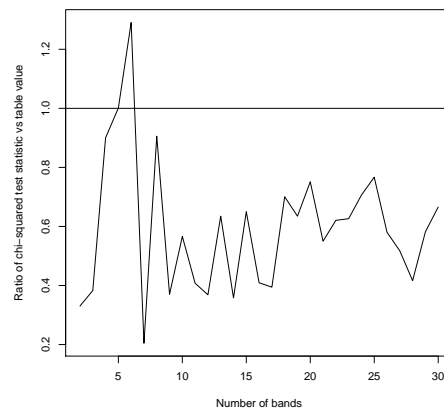
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	299	283.6	286	290.4
10	308	289.5	292.8	300.1
20	316	294.7	299.1	311
100	-	304.9	312.8	341.1
200	-	308.7	318.8	356.5

**Table 4.6:** Return levels and 95% confidence intervals for the RJPM model, Harstad. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.5:** Tide-surge interaction plots for Harstad.



#### 4.6.6 Tregde

Tregde is surge-dominant, and we expect less tide-surge interaction than for the tide-dominant locations. This is certainly not what we find, as with the 5 tidal bands the test statistic is  $\chi^2 = 160.09$ . We expect about 350 points per band under independence, but get 150, 361, 356, 436 and 440. This is not what we would have expected if Tregde was tide-dominant, since the least amount of points then should be at the top. Since Tregde isn't as dominated by tide, we see that large surges coincide more with large tides.

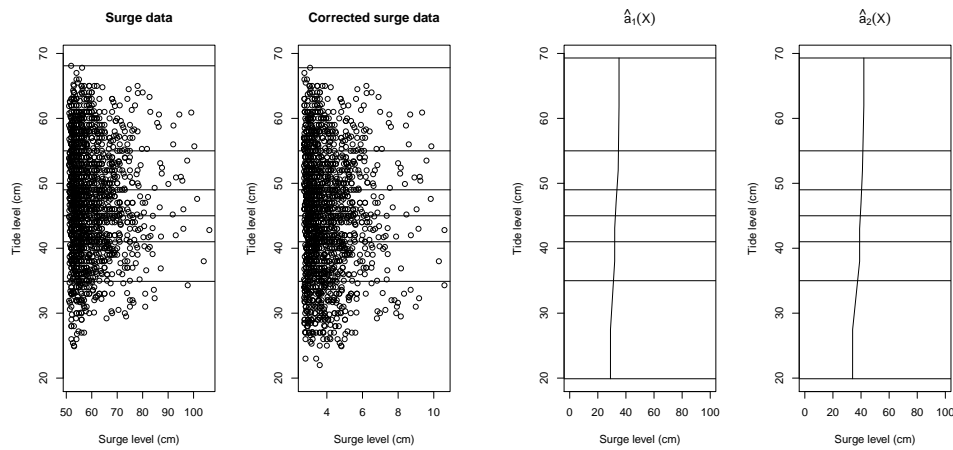
In any case, we see that the lowest band has almost less than half as many points as ideally expected, and need to transform the surge series to get a more even distribution of points. Figure 4.6(c) shows how the transformed surge series performs on the  $\chi^2$  test for different  $n_b$ , and we see a sharp drop to well below 1 at  $n_b = 6$ . We therefore select this, where we get  $\chi^2 = 5.896$  versus  $\chi^2_{5,0.95} = 11.071$ . Figure 4.6(a) clearly show that a number of points are moved to the bottom band after transformation.

Estimated parameters are  $\hat{\mu}_{s*} = 4.3257(0.2053)$ ,  $\hat{\sigma}_{s*} = 1.6746(0.0568)$ ,  $\hat{\xi}_{s*} = -0.05134(0.0204)$ ,  $\hat{\theta} = 0.0499$  and  $\hat{\theta}_{s*} = 0.0485$ . Tregde cuts the data at the lowest quantile of the data sets, and standard errors are small in comparison to the other locations. Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.6(b).

Return levels and confidence intervals are shown in table 4.7. As for the other surge-dominant location, Oslo, the return levels are very similar to their published counterparts.

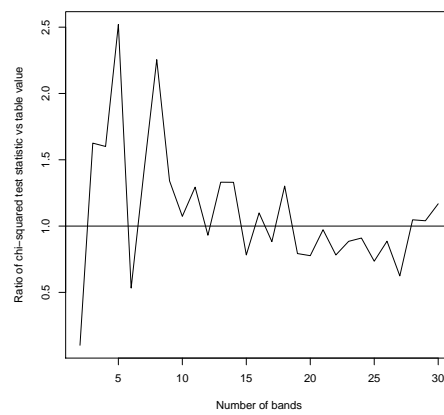
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	129	123.9	129	134.9
10	136	130.1	137	145
20	143	135.7	144.5	154.7
100	-	147.2	160	176.3
200	-	151.6	166.4	185.4

**Table 4.7:** Return levels and 95% confidence intervals for the RJPM model, Tregde. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.6:** Tide-surge interaction plots for Tregde.

#### 4.6.7 Andenes

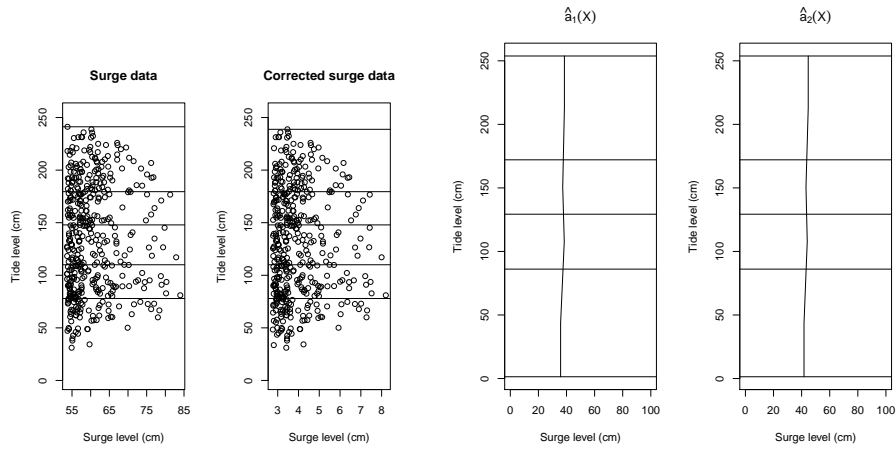
Andenes is tide-dominant, and the test statistic shows significant interaction of  $\chi^2 = 14.860$ . Along with Viker, Andenes doesn't have very many years of data, however, so a small value here might be a little misleading. At any rate, we wish to take the interaction into account, and figure 4.7(c) shows the usual ratio between test statistic and table value for transformed surge series with different numbers of tidal bands. It shows that as long as we take a measure of interaction into account, we get an insignificant  $\chi^2$  test statistic for all  $n_b$  between 2 and 30. We select  $n_b = 4$ , which gives  $\chi^2 = 2.547$  measured against  $\chi_{3,0.95}^2 = 7.815$ . Figure 4.7(a) shows the effect of the transformation of the series in terms of surges versus concomitant tides.

Estimated parameters are  $\hat{\mu}_{s^*} = 3.9036(0.5017)$ ,  $\hat{\sigma}_{s^*} = 1.7209(0.4415)$ ,  $\hat{\xi}_{s^*} = -0.2164(0.1582)$ ,  $\hat{\theta} = 0.1849$  and  $\hat{\theta}_{s^*} = 0.0740$ . Andenes has the most extreme shape parameter, that is farthest away from 0, of all the surge series in the RJPM analyses. Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.7(b).

Return levels and confidence intervals are shown in table 4.8. Due to the scarcity of data at the location, there are no published values with which to compare the estimates.

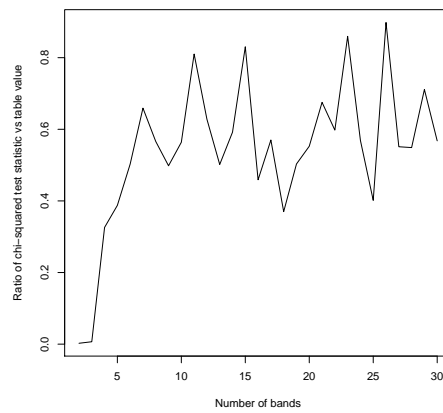
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	-	278.3	282.5	296.3
10	-	284.7	290.7	313.2
20	-	290	297.8	330.7
100	-	299.4	311.5	375.4
200	-	302.5	316.5	396.9

**Table 4.8:** Return levels and 95% confidence intervals for the RJPM model, Andenes. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.7:** Tide-surge interaction plots for Andenes.

#### 4.6.8 Viker

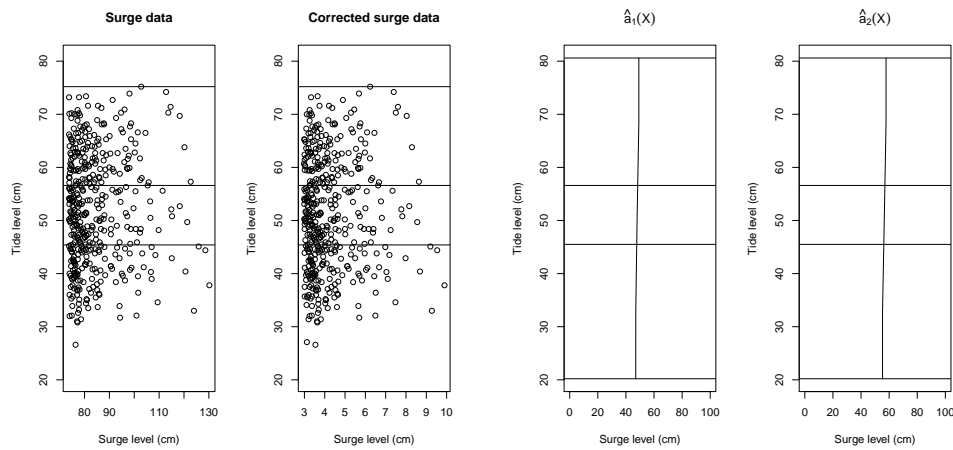
Viker is, like Oslo and Tregde, surge-dominant. We expect little tide-surge interaction, and a test statistic for 5 tidal bands of  $\chi^2 = 6.830$  against the usual  $\chi^2_{4,0.95} = 9.488$ . Although this isn't significant, we still opt to model the interaction to some extent to get even less interaction. By using  $n_b = 3$ , we get  $\chi^2 = 0.1886$ . Figure 4.8(c) shows that transformation for all  $n_b$  between 2 and 30 produces insignificantly interactive surge series, further sementing the notion that tides and surges are more or less independent in Viker. With 3 tidal bands we get 148, 151 and 144 points per band, compared to the expected 146.56 points. Figure 4.8(a) shows both the original series and the transformed one.

Estimated parameters are  $\hat{\mu}_{s^*} = 4.8551(0.5198)$ ,  $\hat{\sigma}_{s^*} = 1.9037(0.2083)$ ,  $\hat{\xi}_{s^*} = -0.0356(0.0740)$ ,  $\hat{\theta} = 0.0759$  and  $\hat{\theta}_{s^*} = 0.0631$ . Details of the correction vectors  $\hat{a}_1(X)$ ,  $\hat{a}_2(X)$  are omitted, but visually shown in figure 4.8(b).

Return levels and confidence intervals are shown in table 4.9. As for the other surge-dominant locations, the results are very similar.

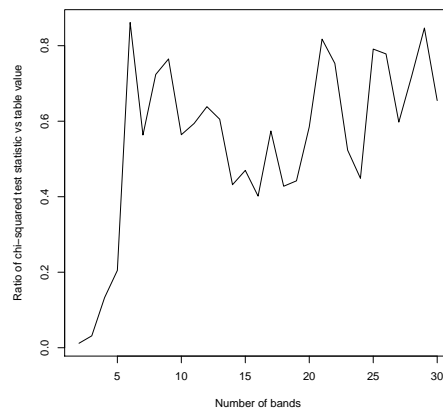
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	173	153.1	172	201.3
10	185	158.4	183.5	226.3
20	196	162.9	194.2	252.1
100	-	171.8	217.6	269.8
200	-	175	227.2	269.8

**Table 4.9:** Return levels and 95% confidence intervals for the RJPM model, Viker. Values in cm.



(a) Tide levels and concurrent surges above the 99.75% quantile.

(b)  $\hat{a}_1(X)$  and  $\hat{a}_2(X)$ .



(c) Ratio between  $\chi^2$  test statistic and table value  $\chi^2_{n_b-1, 0.95}$ .

**Figure 4.8:** Tide-surge interaction plots for Viker.

## 5 The ACER method

A major problem in using both the generalized extreme value distribution and the generalized Pareto distribution is the difficulty in establishing the extent to which the model in question is applicable for the observed data. Using such asymptotic extreme value distributions therefore requires a 'leap of faith' in selecting the appropriate distribution.

The approach described in this section is an attempt to address this issue by a more flexible approach than the ones based on asymptotic theory. It is able to capture subasymptotic behaviour, but does have a shortcoming in the fact that it is restricted to cases where the Gumbel distribution is the appropriate extreme value distribution (i.e. for  $\xi \rightarrow 0$  in the generalized extreme value distribution) [5].

### 5.1 Cascade of conditioning approximations

Let  $M(T) = \max\{X(t) : 0 \leq t \leq T\}$  be the extreme value of the historical recorded time series  $X(t)$  on the time interval  $(0, T)$ . If the process has an average number of peaks  $\bar{N}$  per unit of time, it can be assumed that it has  $N = \bar{N}T$  peaks if the time span  $(0, T)$  is sufficiently long. The points in time corresponding to these maxima are denoted  $0 \leq t_1 < \dots < t_N \leq T$ , and  $A_k = A(t_k)$ ,  $k = 1, \dots, N$  is defined as the heights of the peaks in question. The goal is to estimate the distribution of  $M(T)$  accurately, specifically  $P(\eta) = \text{Prob}(M(T) \leq \eta)$  for some high (in relation to the data) threshold  $\eta$ . Obviously it follows that

$$\begin{aligned} P(\eta) &= \text{Prob}(A_N \leq \eta, \dots, A_1 \leq \eta) \\ &= \text{Prob}(A_N \leq \eta | A_{N-1} \leq \eta, \dots, A_1 \leq \eta) \cdot \text{Prob}(A_{N-1} \leq \eta, \dots, A_1 \leq \eta) \quad (5.1) \\ &= \prod_{j=2}^N \text{Prob}(A_j \leq \eta | A_{j-1} \leq \eta, \dots, A_1 \leq \eta) \cdot \text{Prob}(A_1 \leq \eta) \end{aligned}$$

An approximation is made to simplify the expression in (5.1), by saying that for each  $k = 2, 3, \dots (\ll N)$ ,

$$\text{Prob}(A_j \leq \eta | A_{j-1} \leq \eta, \dots, A_1 \leq \eta) \approx \text{Prob}(A_j \leq \eta | A_{j-1} \leq \eta, \dots, A_{j-k+1} \leq \eta), \quad k \leq j \leq N \quad (5.2)$$

In addition to this cascade of conditioning approximations, the function  $\epsilon_k(\eta)$  is defined by

$$\epsilon_k(\eta) = \begin{cases} 1 - p_1(\eta) & , k = 1 \\ 1 - \frac{p_k(\eta)}{p_{k-1}(\eta)} & , k = 2, \dots, N \end{cases} \quad (5.3)$$

where  $p_k = \text{Prob}(A_1 \leq \eta, \dots, A_k \leq \eta)$ . The  $\epsilon_k(\eta)$ ,  $k = 1, 2, \dots$  functions are the average conditional exceedance rate (ACER) functions, and  $\epsilon_k(\eta)N$  can be

interpreted as the expected number of peaks above some high threshold  $\eta$  during the time  $(0, T)$  - given that  $k - 1$  preceding peaks are below this threshold.

For the case of  $k = 2$ , combining (5.1) and (5.2) gives the following approximate relation:

$$P(\eta) \approx \frac{p_2(\eta)^{N-1}}{p_1(\eta)^{N-2}} \quad (5.4)$$

Assuming statistical independence between peaks would mean a possible rewriting of  $P(\eta)$  on the form  $P(\eta) = [1 - \epsilon_1(\eta)]^N$ , approximated by

$$P(\eta) \approx \exp\{-\epsilon_1(\eta)N\} \quad (5.5)$$

In this Poisson approximation, the mean upcrossing rate at time  $t$ ,  $\nu^+(\eta; t)$ , is the key parameter.  $P(\eta)$  could then be written  $P(\eta) = \exp\{-\int_0^T \nu^+(\eta; t)dt\}$ . This can be rewritten as

$$P(\eta) = \exp\{-\bar{\nu}^+(\eta)T\}, \quad (5.6)$$

where  $\bar{\nu}^+(\eta) = 1/T \int_0^T \nu^+(\eta; t)dt$ .

Such an independence assumption is not always very accurate, however, since narrow-band characteristics of the time series  $X(t)$  are likely to lead to noticeable dependence between neighbouring peaks [6]. But having introduced the conditioning cascade in (5.3), the relation in (5.4) can instead be used for the  $k = 2$  case. This leads to:

$$\begin{aligned} P(\eta) &\approx [1 - \epsilon_2(\eta)]^{N-1} p_1(\eta) \\ &\approx \exp\{-\epsilon_2(\eta)(N-1)\} p_1(\eta) \\ &= \exp\{-\epsilon_2(\eta)(N-1) - \epsilon_1(\eta)\}, \end{aligned} \quad (5.7)$$

since  $p_1(\eta) \approx \exp\{-\epsilon_1(\eta)\}$ . This can be extended to a general  $k$  (for  $k \ll N$ ), yielding

$$\begin{aligned} P(\eta) &\approx \exp\{-\epsilon_k(\eta)(N-k+1) - \sum_{j=1}^{k-1} \epsilon_j(\eta)\} \\ &\approx \exp\{-\epsilon_k(\eta)N\} \end{aligned} \quad (5.8)$$

Equation (5.2) represents a refinement of the Poisson assumption. It yields quite similar results for  $k = 2$  and  $k = 3$  for a purely narrow-band process, but there is little increase in numerical effort when three or more preceding peaks are included [6]. Which  $k$  to use is therefore a balance between a large number of conditioning events - which decreases the accuracy of the estimates, and a small number - which may not always sufficiently account for dependence effects in the time series.

The ACER functions  $\epsilon_k(\eta)$ ,  $k = 1, 2, \dots$  behave quite similarly to the Poisson assumption mean upcrossing rate  $\nu^+(\eta; t)$  for high thresholds  $\eta$  and long intervals  $(0, T)$ . For example, large  $\eta$  gives  $\epsilon_1(\eta) \approx \nu^+(\eta)/\bar{N}$ . This means that equations (5.6) and (5.5) are equivalent, given a reasonable approximation of  $\bar{\nu}^+(\eta)$ .



## 5.2 Numerically estimating the ACER functions

To examine how to empirically estimate the ACER functions  $\epsilon_k(\eta)$ ,  $k = 2, 3, \dots$ , two random indicator functions are defined:

$$Q_{kj}(\eta) = \mathbf{1}\{A_j > \eta, A_{j-1} \leq \eta, \dots, A_{j-k+1} \leq \eta\}, \quad j = k, \dots, N, \quad k = 2, 3, \dots \quad (5.9)$$

$$R_{kj}(\eta) = \mathbf{1}\{A_{j-1} \leq \eta, \dots, A_{j-k+1} \leq \eta\}, \quad j = k, \dots, N, \quad k = 2, 3, \dots \quad (5.10)$$

where  $\mathbf{1}(\mathcal{A})$  is the indicator function corresponding to some random event  $\mathcal{A}$ . Stationarity and the fact that  $E[\mathbf{1}(\mathcal{A})] = \text{Prob}(\mathcal{A})$  gives

$$\epsilon_k(\eta) = \frac{E[Q_{kj}(\eta)]}{E[R_{kj}(\eta)]}, \quad j = k, \dots, N, \quad k = 2, \dots \quad (5.11)$$

It may be assumed that

$$\epsilon_k(\eta) = \lim_{N \rightarrow \infty} \frac{\sum_{j=k}^N Q_{kj}(\eta)}{\sum_{j=k}^N R_{kj}(\eta)} \quad (5.12)$$

Clearly,  $\lim_{N \rightarrow \infty} \sum_{j=k}^N R_{kj}(\eta) = N - k + 1 \approx N$ , which means that  $\lim_{N \rightarrow \infty} \tilde{\epsilon}_k(\eta) / \epsilon_k(\eta) = 1$ , where the following modified ACER function is used:

$$\tilde{\epsilon}_k(\eta) = \frac{1}{N - k + 1} \lim_{N \rightarrow \infty} \sum_{j=k}^N Q_{kj}(\eta) \quad (5.13)$$

which applies to stationary as well as non-stationary processes. Empirically, this function is used for  $k \geq 2$  whenever convenient. Since the focus is on values of an ACER function at the extreme levels, any function providing a correct estimate at these levels may be used. This is true for  $\tilde{\epsilon}_k(\eta)$ , which is also easier to apply for non-stationary or long-term statistics than its counterpart of  $\epsilon_k(\eta)$ . Some care should be exercised when choosing the tail marker however, since this modified version of the ACER function shows a strong decrease for lower response levels [6].

The sample estimate of  $\tilde{\epsilon}_k(\eta)$  is

$$\hat{\epsilon}_k(\eta) = \frac{1}{m} \sum_{r=1}^m \hat{\epsilon}_k^{(r)}(\eta), \quad (5.14)$$

where  $m$  is the number of time series samples and

$$\hat{\epsilon}_k^{(r)}(\eta) = \frac{\sum_{j=k}^N Q_{kj}^{(r)}(\eta)}{N - k + 1} \quad (5.15)$$

The additional index  $(r)$  refers to the realization number.

For a wide range of dynamical systems, it is a reasonable approximation [4][5][6] that the tail of the ACER function can be modelled as

$$\tilde{\epsilon}_k(\eta) \approx q_k(\eta) \exp\{-a_k(\eta - b_k)^{c_k}\}, \quad (5.16)$$

for constants  $a_k, b_k$  and  $c_k$ , with a function  $q_k(\eta)$  slowly varying compared with the exponential function it is multiplied with. It is generally slow enough in its variation in the tail region that it might be replaced with a constant value, call this  $q_k$ .

Because this assumption fails on the lower levels, we need to specify a tail marker  $\eta_0$ . This is found by visual inspection of the ACER plot, and corresponds to where the function starts behaving on the form in equation (5.16) with a constant  $q_k(\eta) = q_k$ . Since this form should hold above the chosen  $\eta_0$ , choosing a higher tail marker should not affect estimates to a large extent, but variance increases since the number of points decreases. Choosing the tail marker low, on the other hand, can put too much emphasis on low levels.

For further details on the optimization process with regard to these parameters, see [6].

### 5.3 Polynomial smoothing of confidence intervals

The 95 % confidence interval of  $\hat{\epsilon}_k(\eta)$  is approximated well for  $m \geq 20$  by

$$(\hat{\epsilon}_k(\eta) - 1.96\hat{s}_k(\eta)/\sqrt{m}, \hat{\epsilon}_k(\eta) + 1.96\hat{s}_k(\eta)/\sqrt{m}), \quad (5.17)$$

where

$$\hat{s}_k^2(\eta) = \frac{1}{m-1} \sum_{r=1}^m (\hat{\epsilon}_k^{(r)}(\eta) - \hat{\epsilon}_k(\eta))^2 \quad (5.18)$$

It is assumed that confidence intervals in the tail behave similarly to the mean ACER function as in equation (5.16), i.e.

$$CI^+(\eta) = q_{CI^+} \exp\{-a_{CI^+}(\eta - b_{CI^+})^{c_{CI^+}}\}, \quad (5.19)$$

for the curve of the upper bound and equivalently for  $CI^-$  [8].

### 5.4 Estimating return levels with the ACER method

To estimate return levels, it is first necessary to find an expression for the level  $\eta_p$  for the desired probability  $p$ . This probability is defined by  $P(\eta_p) = 1 - p$ , since  $P(\eta)$  is the probability that all preceding peaks are below the threshold  $\eta$ . This yields the approximate relation

$$1 - p \approx \exp\{-N \cdot q_k \exp\{-a_k(\eta_p - b_k)^{c_k}\}\}, \quad (5.20)$$

where the function  $q_k(\eta)$  from (5.16) is assumed constant. To simplify notation, the subscripts will be dropped on the parameters. Additionally, the approximation symbol is replaced by an equality sign. We also define  $d = \ln q$ . This gives

$$1 - p = \exp\{-N \cdot \exp\{d - a(\eta_p - b)^c\}\}, \quad (5.21)$$

which can be rearranged for  $\eta_p$ :

$$\eta_p = b + \left[ \frac{d - \ln(-1/N \cdot \ln(1 - p))}{a} \right]^{1/c} \quad (5.22)$$

The parameters  $a, b, c, d$  are estimated using data with one year block sizes, meaning that every year of observations is viewed as a realization of the same process.  $N$  therefore corresponds to the number of peaks from one year. This is taken to be the total number of peaks in the period divided by the number of years used in estimation. The return probability  $p$  is the reciprocal of the return period, meaning that for instance a desired return period of 5 years corresponds to a probability  $p$  of  $1/5$ .

Something to be wary of is that even if the empirical ACER functions merge in the tail of the data, we extrapolate to values beyond those which we have data for. This means that the estimated ACER functions may diverge at higher levels. This will be illustrated with real data.

All calculations are performed with the ACER package for Matlab.

## 5.5 Application to water level measurements

### 5.5.1 Oslo

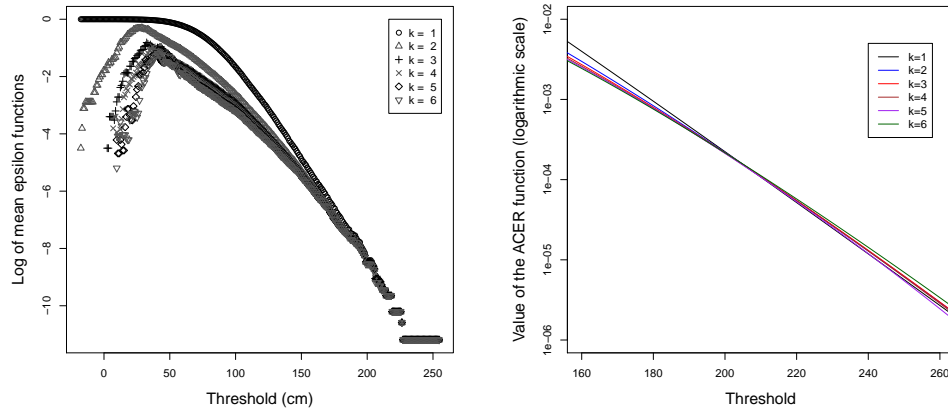
As for the POT method, we firstly wish to find a tail marker  $\eta_0$  that represents a sufficiently high threshold. For the ACER method, this is found after where the curves of the estimated ACER functions  $\hat{e}_k(\eta)$  start behaving well-behaved in the sense of (5.16). After locating an approximate such point, we have to go further into the tail to where an appropriate amount of data is left - enough to get a good basis for the curve estimation but not so much that too little weight is put on the points far out in the tail. Finding such a tail marker is not a formulaic procedure, but some leeway is given. As a general rule for these hourly measurements, we have found that a place where remaining data amounts to about 0.5 % to 1 % of the total data gives stable and similar numerical estimates of the ACER functions for high levels and different values of  $k$ . This is desirable, since the ACER functions should be identical for very high levels.

The curves in figure 5.1(a) detail the effect dependence has on the ACER function estimates, and we can see that the independent case of  $k = 1$  has a curve which stays significantly above the rest until around  $\eta = 140$  to  $\eta = 150$ . This tells us that peaks above this level rarely come in sequence - they are mostly from independent storms.

As for well-behavedness in the curves, evidence of such behaviour is seen at around the 120 cm mark. At this level there are still noticeable dependence effects, and this affects the parameter estimates. To see the effect this has on return level estimates at high levels, we turn to figure 5.1(b). It shows that estimates are very similar, indicating that the dependence still existing at 120 cm is not crucial. The dependence probably does have a little impact, since the curves are somewhat different even at the highest levels. Equality of estimates is ideally what we would expect; for such high levels, peaks should definitely be independent. This means that conditioning on one or more peaks shouldn't introduce differences in estimates - not unless the number of previous peaks included is very large.

Since the curves in theory should follow identical paths, we don't worry too much about which  $k$  is the best choice. We extrapolate to levels beyond which we have data, and simply select the ACER function which seems to predict the most reasonable curve. In this case there is very little difference between the estimates, and no choices of  $k$  are easily discarded. We therefore select one of the curves in the middle at the far right of figure 5.1(a), and come up with  $k = 2$ .

With the selected estimated ACER function, i.e.  $\hat{e}_2(\eta)$ , the optimized parameter estimates are  $d = 5.5774$ ,  $b = -172.836$ ,  $a = 0.000398439$ ,  $c = 1.76718$ . The return level estimates are presented in table 5.1.



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.1:** ACER decision plots, with  $k = 1$  to 6, for Oslo.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	195	190.383	198.133	204.453
10	207.3	199.834	209.392	217.488
20	218.6	208.712	219.958	229.721
100	241.8	228.197	243.115	256.538
200	250.8	236.273	252.702	267.642

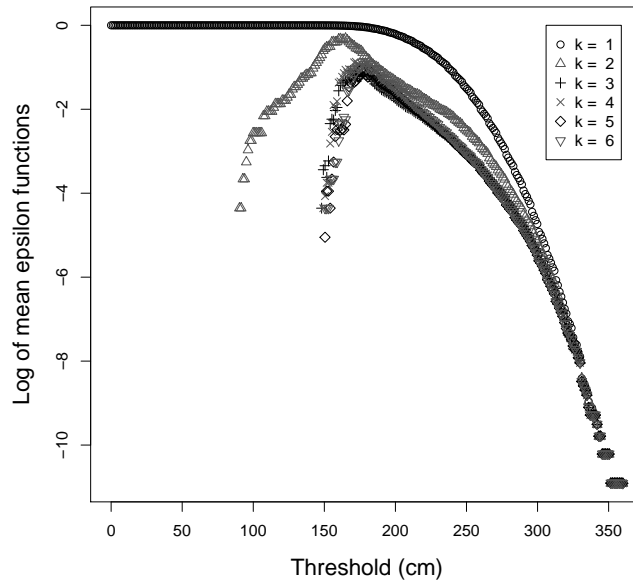
**Table 5.1:** Return level estimates and confidence intervals for the ACER method, Oslo. Values in cm.

### 5.5.2 Heimsjø

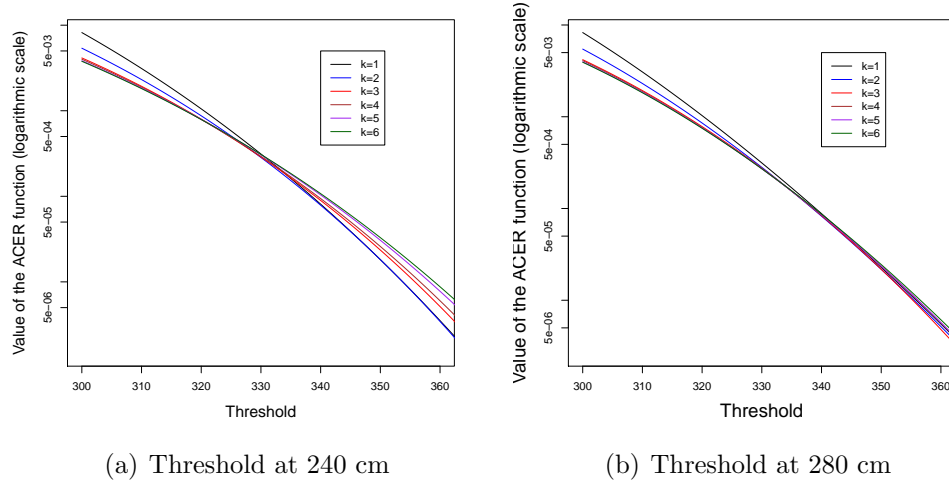
We start by looking at figure 5.2, which shows well-behaved functions somewhere around 240 to 300 cm. As much as 6% of the data are above 240 cm, so this is probably not a good tail marker. Less than 0.1% are above 300 cm, however, so this is probably too far out in the tail. We therefore tentatively select 280 cm, with around half a percent of data above, as our  $\eta_0$ . Figure 5.3 illustrates the differences between models created with thresholds at 240 cm and 280 cm, respectively. The latter model estimates the high values much more consistently, and is preferable to the former model.

With the threshold  $\eta_0 = 280$ , we select  $k = 2$  for the estimation of return levels. These are found in table 5.2. Published values are much in agreement with the corresponding ACER estimates.

Optimal values of parameters are:  $d = -3.35941$ ,  $b = 261.256$ ,  $a = 0.00369969$ ,  $c = 1.69398$ .



**Figure 5.2:** Log plot of estimated ACER functions for Heimsjø, with  $k = 1, 2, 3, 4, 5$  and 6



**Figure 5.3:** ACER functions for high levels in Heimsjø, with  $k = 1, 2, 3, 4, 5$  and  $6$

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	329	324.938	328.958	332.897
10	337	329.622	335.185	341.015
20	345	333.811	340.834	348.527
100	-	342.461	352.702	364.703
200	-	345.874	357.451	371.308

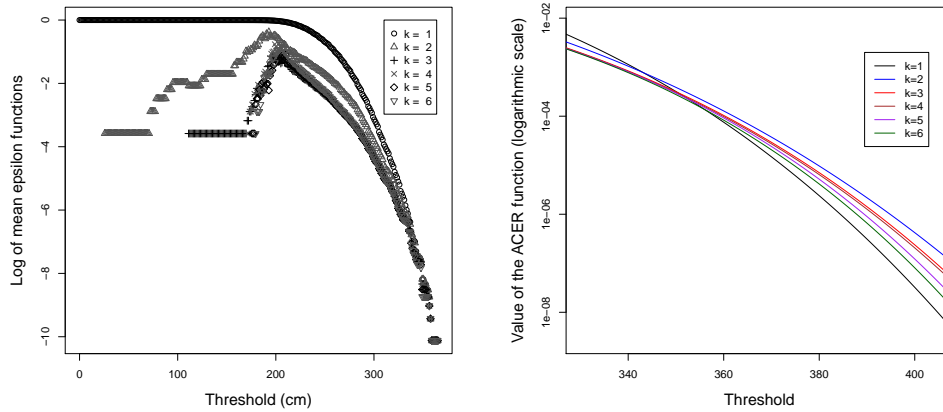
**Table 5.2:** Return level estimates and confidence intervals for the ACER method, Heimsjø. Values in cm.

### 5.5.3 Honningsvåg

Figure 5.4(a) shows the ACER functions for  $k = 1$  to  $k = 6$ . Consistent, smooth behaviour in the curves is seen from around 270 cm, where there is about 4.3 % left of the data. This percentage is too high, and we set the tail marker a little higher;  $\eta_0 = 300$  is more appropriate, with about 0.6% of the data above.

Figure 5.4(b) shows the behaviour of the estimated ACER functions for different  $k$  with the tail marker at 300 cm. We choose a curve in the middle, and go for  $k = 3$ . With this ACER function, the results are found in table 5.3.

Optimal values of parameters are:  $d = -2.03381$ ,  $b = 113.477$ ,  $a = 1.1368e - 009$ ,  $c = 4.09641$ .



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.4:** ACER decision plots, with  $k = 1$  to 6, for Honningsvåg.

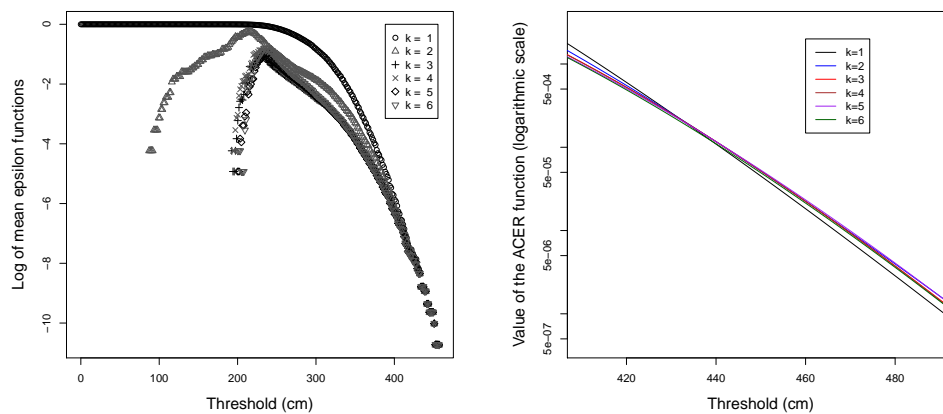
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	348	344.491	349.746	355.744
10	357	348.675	356.642	364.804
20	366	352.222	362.732	372.967
100	-	359.06	375.035	389.872
200	-	361.603	379.781	396.527

**Table 5.3:** Return level estimates and confidence intervals for the ACER method, Honningsvåg. Values in cm.



### 5.5.4 Narvik

Well-behavedness in the curves is seen in figure 5.5(a) from around 320 cm, where 4.2 % of the data remain. Dependence wears off around 400 cm. Here there are about 0.03 % left of the data. Using the general rule of 0.5-0.6 % data left at the tail marker, we select the threshold to be  $\eta_0 = 360$ . 0.57 % of the data remain above this level. With the tail marker at 360 cm, estimated ACER functions are shown in figure 5.5(b). Values for high levels are very similar for all values of  $k$ , particularly above  $k = 1$ . We choose  $k = 2$ , and the results are shown in table 5.4. Optimal parameter values are  $d = -3.71051$ ,  $b = 357.429$ ,  $a = 0.0224224$ ,  $c = 1.24027$ . Estimates for the 5-20 year periods are very similar to published values.



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.5:** ACER decision plots, with  $k = 1$  to 6, for Narvik.

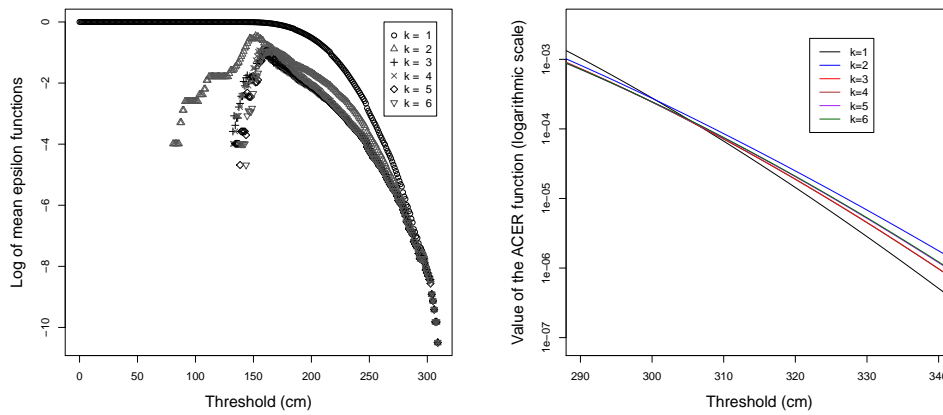
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	427	419.973	427.147	433.3
10	439	426.726	436.726	445.645
20	451	432.865	445.66	457.36
100	-	445.813	465.145	483.492
200	-	451.015	473.2	494.498

**Table 5.4:** Return level estimates and confidence intervals for the ACER method, Narvik. Values in cm.

### 5.5.5 Harstad

Harstad has well-behaved curves at somewhat above 200 cm, say approximately 225 cm. About 4.5 % of the data are left here. Dependence wears off at around 280 cm, where 0.04 % are left. We set the tail marker at 255 cm, where 0.53 % remain. The estimated ACER functions are shown in figure 5.6(b), and we choose  $k = 3$  for the estimation of return levels. Again, the 5-20 year return level values are very similar to those that are published, as can be seen in table 5.5.

Optimal values of parameters are:  $d = -4.03566$ ,  $b = 242.56$ ,  $a = 0.00753561$ ,  $c = 1.56615$ .



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

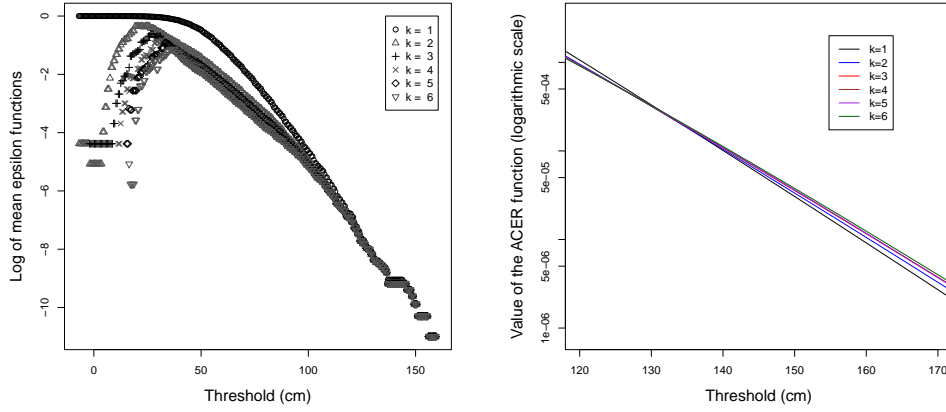
**Figure 5.6:** ACER decision plots, with  $k = 1$  to 6, for Harstad.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	299	293.528	297.518	302.671
10	308	298.121	303.892	311.818
20	316	302.213	309.672	320.362
100	-	310.642	321.828	338.992
200	-	313.966	326.701	346.681

**Table 5.5:** Return level estimates and confidence intervals for the ACER method, Harstad. Values in cm.

### 5.5.6 Tregde

Figure 5.7(a) shows empirical ACER functions. Curves for  $k > 1$  behave erratically at first, but start acting well-behaved a little after 50 cm. Dependence wears off completely at approximately the 120 cm mark, where 0.03 % of the data remain. Figure 5.7(a) shows that curves are already quite close to each other at around 100 cm. We choose the threshold to be  $\eta_0 = 95$ , where 0.58 % of the data are left. Figure 5.7(b) shows estimated ACER functions for high levels with this threshold. We select  $k = 3$  for the return level predictions. These are found in table 5.6. Estimates agree with published levels where the latter are available. Optimal values of parameters are:  $d = -4.58839$ ,  $b = 95.5187$ ,  $a = 0.0768807$ ,  $c = 1.0749$ .



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.7:** ACER decision plots, with  $k = 1$  to 6, for Tregde.

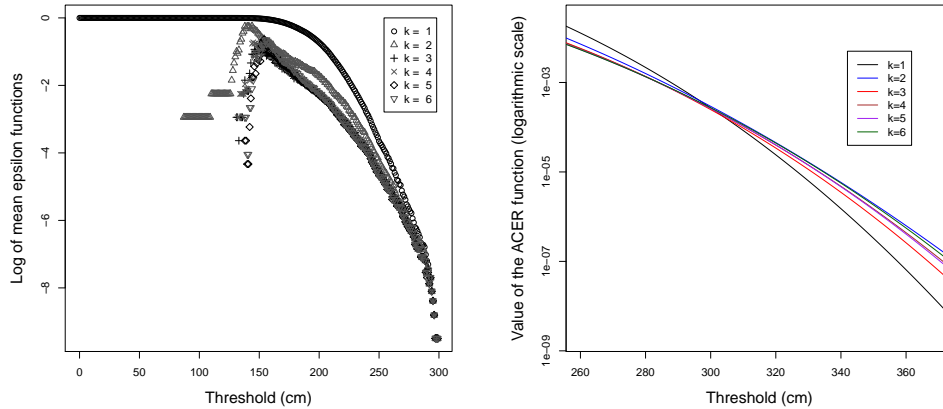
Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	129	127.364	130.221	133.678
10	136	131.839	137.135	143.014
20	143	135.833	143.686	152.15
100	-	144.079	158.288	173.38
200	-	147.337	164.44	182.636

**Table 5.6:** Return level estimates and confidence intervals for the ACER method, Tregde. Values in cm.

### 5.5.7 Andenes

Andenes has relatively little data available for analysis, with only 19 usable years of observations. To estimate 200 year return levels we therefore have to extrapolate more than 10 times as long into the future as we have data. With the Andenes station having little data available, this also means that no previous return level analyses have been performed. We therefore do not have any previous values to compare with, other than those calculated previously in this report.

Figure 5.8(a) shows the behaviour of the empirical estimates of the ACER functions. The curves act well-behavedly from and above around 220 cm, where 3.5 % data remain. The figure shows that the  $k = 1$  curve (black circles), stays above the rest very far out into the tail. Dependence doesn't wear off until after 280 cm, where about 0.02 % data remain. With the small amount of data available at Andenes, this only constitutes slightly more than 30 data points before extraction of peaks. To select the tail marker, we experimented to find a level that produced relatively similar results for the different  $k$  at high levels. We finally select  $\eta_0 = 240$ , with resulting curves in figure 5.8(b). About 0.9 % of the data remain above this tail marker, which is the largest percentage so far. We select  $k = 3$  for the return level estimates, with resulting return levels shown in table 5.7. Optimized parameter values are  $d = -0.0676168$ ,  $b = 24.8026$ ,  $a = 2.63863e - 007$  and  $c = 3.07189$ .



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.8:** ACER decision plots, with  $k = 1$  to 6, for Andenes.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	-	289.764	297.273	304.838
10	-	296.435	305.355	314.536
20	-	302.472	312.679	323.337
100	-	315.051	327.972	341.753
200	-	320.03	334.038	349.071

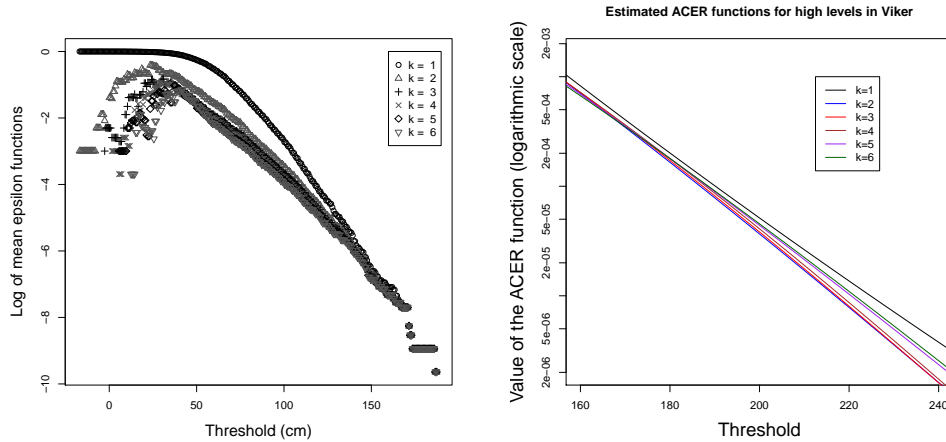
**Table 5.7:** Return level estimates and confidence intervals for the ACER method, Andenes. Values in cm.

### 5.5.8 Viker

Figure 5.9(a) shows the empirical estimated ACER functions for  $k = 1, 2, 3, 4, 5, 6$  in Viker. The figure shows well-behavedness in the curves from around the 75 cm mark, where 14.6 % of the data still are left above. Dependence wears off at around 140 cm, with 0.1 % data left. We set the tail marker at  $\eta_0 = 120$ , where 0.55 % of the data amount is left.

We select a  $k$  by looking at figure 5.9(b), which firstly shows that values are quite similar for the high levels as we would wish. Secondly, we see that the curves for  $k = 5$  and  $k = 6$  are in the middle of the range of curves for the entirety of the plot. We select  $k = 5$  for the estimation of the return levels.

Table 5.8 shows the results of the return level estimation, with optimal values of parameters as  $d = -4.78553$ ,  $b = 117.669$ ,  $a = 0.0376866$ ,  $c = 1.11914$ . The maximum of the approximately 20 years of data used in estimation is 188 cm, which corresponds well with the 20 year return level in table 5.8. The return levels also agree with the published values in the second column of table 5.8. Caution needs to be taken in interpreting the longer period return levels however, because of the short span of available data.



(a) Log plot of estimated ACER functions. (b) Estimated ACER functions for high levels.

**Figure 5.9:** ACER decision plots, with  $k = 1$  to 6, for Viker.

Return period (years)	Published value	CI lower bound	Estimate	CI upper bound
5	173	156.564	172.699	197.55
10	185	161.426	183.614	220.298
20	196	165.746	193.884	242.873
100	-	174.625	216.573	296.239
200	-	178.121	226.061	319.826

**Table 5.8:** Return level estimates and confidence intervals for the ACER method, Viker.  
Values in cm.

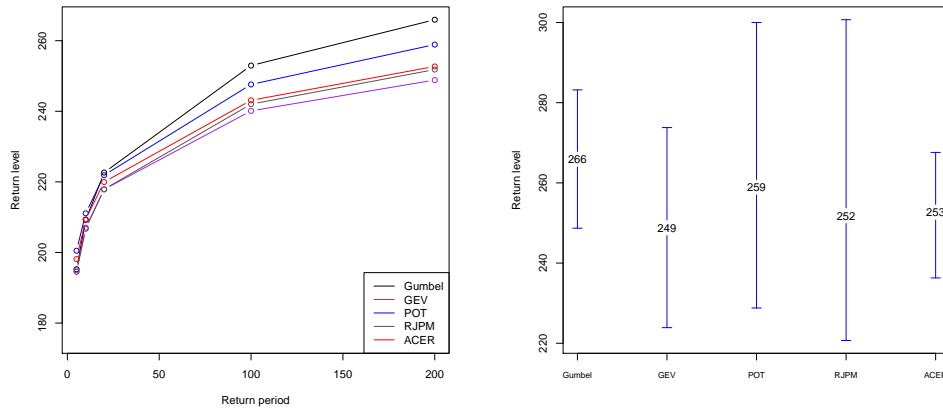
## 6 Discussion of results

### 6.1 Oslo

Oslo is surge-dominant and has the largest amount of data of the locations featured in this thesis, both these factors mean that all the used methods should perform well here. Surge-dominance means that the non-stochastic tidal component of the measured sea levels doesn't interfere to a large degree, and a large amount of data is naturally desirable. As for possible sources of error, the removal of the post-glacial rebound trend is certainly worth mentioning, although the same procedure as in the report by Hansen [3] was followed.

Figure 6.1(a) shows that return levels are very similar for all methods for 5, 10 and 20 year periods, while larger differences arise for the 100 and 200 year periods. Note that only the five aforementioned return level estimates have been calculated, meaning that the lines between the indicated points are not the true model curves.

Figure 6.1(b) shows the difference between the 200 year estimates, with the Gumbel model producing the largest estimate, at 266 cm, and the GEV model the smallest, at 249 cm. The 200 year Gumbel estimate is 6.8% larger than the GEV estimate. This difference is not huge, and all methods contain the results from the other methods in their confidence intervals.



(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

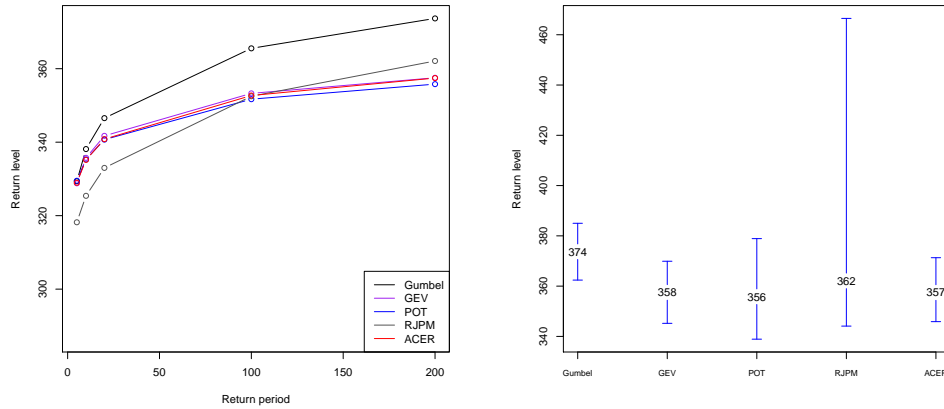
**Figure 6.1:** Comparison between return levels in Oslo for methods applied in the thesis



## 6.2 Heimsjø

For Heimsjø, the non-stochastic tidal component dominates, and estimating the extreme value distribution becomes more difficult. We see this in the larger differences between methods. RJPM, in particular, disagrees markedly for the three first periods of 5, 10 and 20 years. This is shown in figure 6.2(a). Interestingly enough, the RJPM method is not as controversial for the 100 and 200 year levels, where it basically is in line with all models but the Gumbel one.

Figure 6.2(b) shows better the difference between the 200 year estimates, and highlights the extreme upper bound of the RJPM estimate, which was discussed in section 4. The other confidence bounds look minuscule in comparison, but this is because of the scale of RJPM band. The Gumbel model produces a significantly higher return level compared to the other methods, with no other estimates than the RJPM one within its lower 95% confidence bound. The difference between highest estimate (Gumbel) and lowest (POT) for the longest return period is only 5.1%, however.



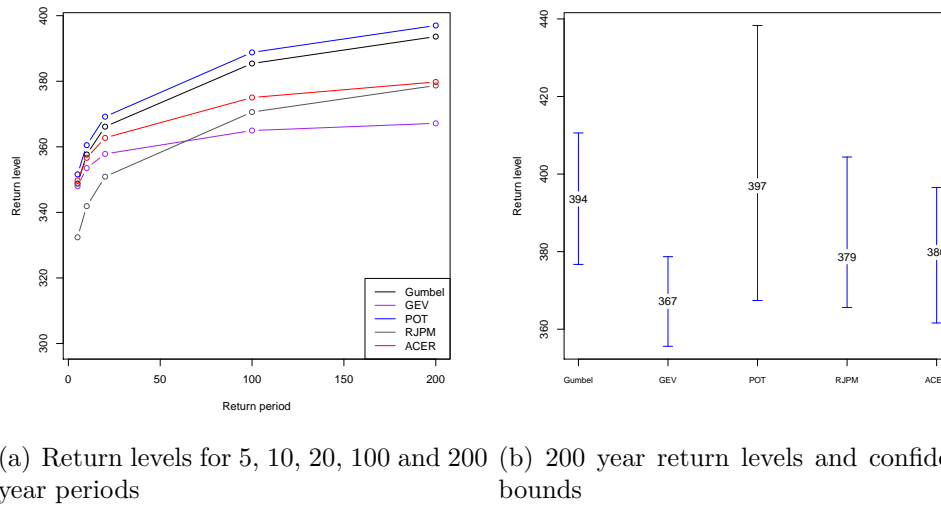
(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

**Figure 6.2:** Comparison between return levels in Heimsjø for methods applied in the thesis

### 6.3 Honningsvåg

Mostly the same pattern as in Heimsjø is shown for Honningsvåg. All methods except RJPM agree at first, but diverge from there. RJPM starts out estimating lower return levels, but ends up in the same region as the ACER estimate for the 200 year period. The Gumbel and POT methods give the by far largest estimates, while the GEV method produces by far the lowest.

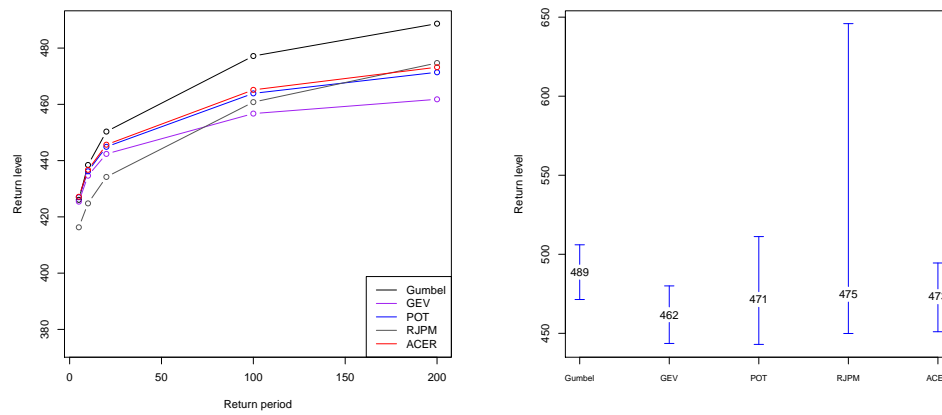
Figure 6.3(a) compares the 5, 10, 20, 100 and 200 year estimates for all methods, while figure 6.3(b) shows the comparatively large difference between 200 year estimates. The POT estimate for that period is 8.2% larger than the corresponding GEV estimate.



**Figure 6.3:** Comparison between return levels in Honningsvåg for methods applied in the thesis.

## 6.4 Narvik

Figure 6.4(a) shows the return levels for all methods, again showing the trend of underestimation in the RJPM method for low levels. Figure 6.4(b) shows that the difference between highest estimate (Gumbel) and lowest (GEV) is smaller here than in Honningsvåg - at 6.1%, and all methods excepting the Gumbel model include the estimates of one another in their confidence bounds. The Gumbel model intervals contain the RJPM and ACER estimates, and is itself included in the bounds of the POT, RJPM and ACER methods.

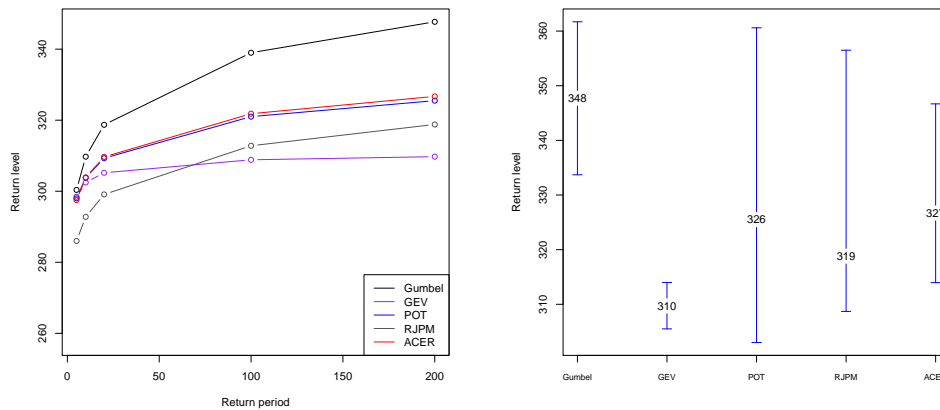


(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

**Figure 6.4:** Comparison between return levels in Narvik for methods applied in the thesis.

## 6.5 Harstad

The plot showing return level estimates for Harstad, in figure 6.5(a), shows a similar development as Narvik, Honningsvåg and to some extent Heimsjø. RJPM underestimates the short period return levels when compared to the other estimates, but ends up in the middle of the range for 200 year return levels. The difference between highest and lowest 200 year estimate is higher here, however, and comes to 12.3%. This is because of the large difference between the Gumbel and GEV models, with the three other models giving quite similar values. The GEV model gave a shape parameter estimate of  $\hat{\xi} = -0.465$ , meaning a very short bounded upper tail. This is an uncommonly large negative value, and could be indicative of a poor model. The POT model, using much more data, estimated  $\xi$  to be  $-0.085$ , which is more reasonable.

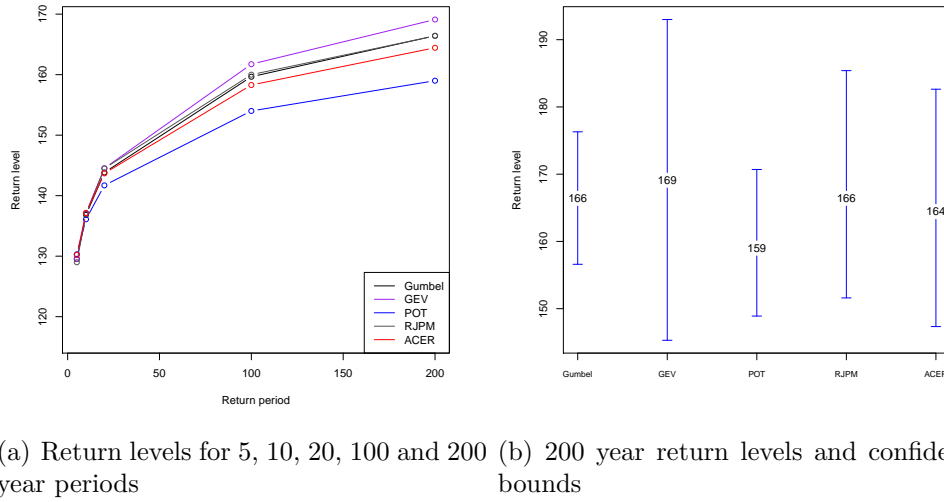


(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

**Figure 6.5:** Comparison between return levels in Harstad for methods applied in the thesis.

## 6.6 Tregde

Tregde is dominated by surges, and this makes for a plot in figure 6.6(a) dissimilar from those in tide-dominant locations but similar to figure 6.1(a) for Oslo. Figure 6.6(b) shows the difference between the 200 year return level estimates, with the GEV estimate being the highest. This is unusual compared to the previously discussed locations, but is explained by the fact that the shape parameter is slightly positive -  $\hat{\xi} = 0.0251$ . The standard error for the GEV model is very large however, at 0.0880, or at 3.5 times the size of the estimate itself. The 200 year return level estimate from the GEV model is approximately 6.3% larger than that from the POT model.

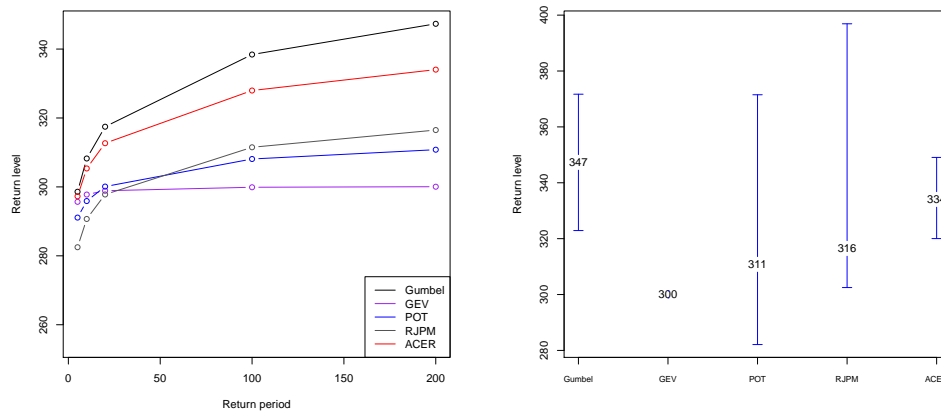


**Figure 6.6:** Comparison between return levels in Tregde for methods applied in the thesis.

## 6.7 Andenes

Andenes is a difficult location to estimate return levels for. Little data is available, and the location itself is tide-dominant. The GEV model produced an estimated shape parameter of  $\hat{\xi} = -0.821$ , corresponding to an extremely short tail and a is large enough negative  $\hat{\xi}$  that normal likelihood properties no longer apply. This produces extremely narrow (and unrealistic) confidence intervals, and are too short to even be plotted in figure 6.7(b).

Figure 6.7(a) shows that all methods except RJPM agree on the 5 year return level, while larger differences start to arise already for the 10 year period. For the 200 year return levels, as shown in figure 6.7(b), estimates are quite different. The problem of estimating the return levels here is reflected in the difference between highest and lowest 200 year estimate, at 15.7%.

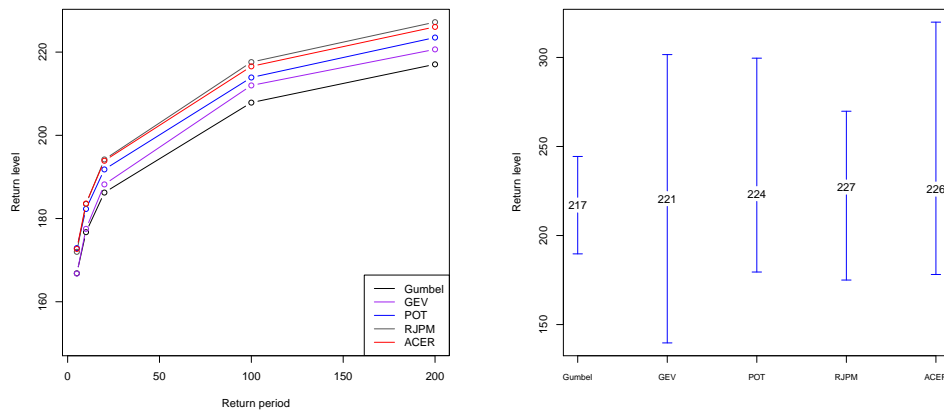


(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

**Figure 6.7:** Comparison between return levels in Andenes for methods applied in the thesis.

## 6.8 Viker

Viker, like Andenes, has little available data. It is, however, dominated by surge and not tide, causing the RJPM results to more or less agree with the rest for the shortest return periods. We see this both in figures 6.8(a) and 6.8(b). The difference isn't very large between smallest and largest estimated return level, at about 4.6%.



(a) Return levels for 5, 10, 20, 100 and 200 year periods (b) 200 year return levels and confidence bounds

**Figure 6.8:** Comparison between return levels in Viker for methods applied in the thesis.

## 6.9 Conclusion

The Annual Maxima Method and the Peaks-over-Threshold method are both widely known and much applied methods. They both have some possible defects, however. As mentioned, the AMM throws away much data, and may fit a poor model since all locations have less than 100 years of observed maxima to draw from. The POT method uses a great deal more data, but is subject to two individual choices: the extreme threshold and the empirical clustering rule. The former of the choices has available supportive literature [1], but the selection of clustering rule has little general theory. Here it requires some knowledge of tides, and we are not very accustomed with that field other than through research for this thesis. The choices that are made are therefore subject to possibly flawed logic. Furthermore, neither of the two methods have been investigated for trends in the data. Such considerations could possibly have improved the analysis.

The Revised Joint Probability Method is less widely applied than the first two methods, but has seen application to a host of British locations [7][2] by the creators and proponents of the method. But because the literature is more sparse, and from fewer sources, it is more difficult to assess the level with which we can rely on this method. Furthermore, it makes it more difficult to check if we have fully grasped its ins and out and fitted models in the way that was intended by Tawn and Dixon.

As for the models estimated by RJPM, a notable feature is the seemingly underestimated short period return levels in the tide-dominant locations of Heimsjø, Honningsvåg, Narvik, Harstad and Andenes. This is not an artefact from the tide-surge interaction modelling, since corresponding return levels were equally low when the estimation was performed without such modelling. We use the POT method to estimate surge parameters, something not done by Dixon and Tawn in earlier work. This could affect results.

We do, however, adopt their methods in declustering the data using  $r = 30$ , which is quite different from the choice done in the POT analysis. It was done to emulate the method as used by Dixon and Tawn [2], but may have benefited from more careful consideration. Other choices that need to be made are the functions with which tide-surge interaction is corrected for and the quantiles that are deemed extreme. The correction functions were made to the specifications in [2], while the quantiles used were taken from the preceding POT chapter which provides a generally thorough explanation of the threshold selection.

In total, it is difficult to recommend a method that has the likely defect of underestimating short period return levels. Although it agrees more with the other methods for higher return levels, this is an area where there is a large margin of error anyway. The method is also somewhat difficult to implement, since we need to solve (4.11) numerically. We also need to have tide measurements available.



Using such data is, however, a strength of the model, if used right. We then use more data in the estimation. This attempt to incorporate sea level specific data and methodology into the return level estimation is arguably the most desirable quality of the method.

The Average Conditional Exceedance Rate method is the most recently developed method, and no published articles exist where it is applied to water levels. It is therefore interesting to see that it produces results similar to those by more conventional methods. In the return level plots in this chapter it is seen that the ACER method produces 200 year estimates in the middle of the range of the methods used here.

Two choices need to be made when estimating return levels with the ACER method; the tail marker  $\eta_0$  and the ACER function  $\epsilon_k$  used for the parameter estimation. We found that as long as we went quite far out into the tail, at around or above the 99% quantile, we got quite consistent results and similar results for the different ACER functions.

The ACER method seems to be a viable alternative for water level return estimation, and with the developed methodology for parameter estimation and creation of confidence interval it is quite easy to implement. More study need perhaps to be done to the choice of ACER function as it relates physically to the actual dynamic system. Since we extract peaks, it is more difficult to see from the time series what  $k$  needs to be to sufficiently take dependence into account. We have used a pragmatic approach where we plot the behaviour of the estimated  $\hat{\epsilon}_k$  functions for high levels and decide on a  $k$  in the middle range. This is a similar argument to the one used by Dixon and Tawn [2] to choose their  $r$ , but may not be the most physically sound choice.

In summary, the ACER method seems to be an attractive method, with its quite easily implemented methodology and attractive statistical properties. When looking at 200 year return levels, it usually agrees with the other data-intensive methods, that is RJPM and POT. We mentioned in section 5 that the ACER is restricted to cases where the Gumbel distribution is the appropriate extreme value distribution, i.e. when  $\xi \rightarrow 0$ . This seems to mostly be the case for water levels, since all POT models had estimated shape parameters quite close to 0, as shown in table 6.1.

---

Location	$\hat{\xi}_{POT}$
Oslo	0.0192
Heimsjø	-0.1153
Honningsvåg	-0.0248
Narvik	-0.0631
Harstad	-0.0849
Tregde	-0.0455
Andenes	-0.1895
Viker	0.00595

**Table 6.1:** Estimated shape parameters for the POT models.

## References

- [1] Coles, S. *An introduction to Statistical Modeling of Extreme Values*, Springer-Verlag London Limited, 2001
- [2] Dixon, M.J.; Tawn, J.A. *Extreme sea-levels at the UK A-class sites: site-by-site analyses.*, Proudman Oceanographic Laboratory Internal Doc. No. 65, 1994
- [3] Hansen, H.; Roald, L. *Ekstremvannsanalyse i sj ved utvalgte stasjoner*, Norwegian Water Resources and Energy Directorate, Doc. No. 11, 2000
- [4] Næss, A.; Gaidai, O. *Estimation of extreme values from sampled time series*, Structural Safety 31, 2009, pp. 325-334
- [5] Næss, A; Stansberg, C.T.; Batsevych, O. *Prediction of extreme tether tension for a TLP*, Proceedings of OMAE 2009, (J)OMAE-09-1123
- [6] Næss, A.; Gaidai, O.; Batsevych, O. *Prediction of extreme response statistics of narrow-band random vibrations*, Journal of Engineering Mechanics 3, 2010, pp. 290-298
- [7] Tawn, J.A. *Estimating probabilities of extreme sea-levels*, Appl. Statist. 41, No. 1, 1992, pp. 77-93
- [8] Sichani, M. T. *ACER User's guide*, 2010
- [9] <http://vannstand.no/index.php/nb/statistikk/gjentaksintervall>, Statens Kartverk, 2011-02-21