



Christchurch City Council

LDRP097 Multi-Hazard Baseline Modelling Joint Risks of Pluvial and Tidal Flooding

Rev 0

February 2021

Executive summary

High flood levels on the lower Avon River in Christchurch can be caused by high river discharges, and high sea levels or both. High sea levels can be caused by astronomical tide conditions, and other influences of sea level collectively referred to here as storm surge. Storm surge is predominantly driven by low barometric pressure weather events but can be also driven by wind and long wave conditions as well as secondary influences such as seasonal temperature variation.

More upstream the rainfall will be more important. This importance gradually shifts to high sea water levels at more downstream locations. High discharges are caused by large rainfall. These rainfall events often coincide with a low barometric pressure, which causes the sea water level to be elevated. When both these meteorological conditions occur at the same moment the water levels along the river might be elevated such that it results in a combined flood risk for the surrounding areas, higher than would be expected from only the rainfall or sea water level. Assuming independence between these two boundary conditions would result in an underestimate of the water levels on Avon River. To date expert engineering judgement has been used to estimate dependence between these two conditions and determine rules for hydraulic modelling of flood risks.

A more substantiated quantification of the probability of coincidence will lead to a better rule for setup of hydraulic modelling and better estimate of flood risks along the Avon River. It is expected that there is some correlation between the two; the question is however the magnitude of this correlation and its effect on the river flood level risks?

To answer this question the statistical relation between rainfall and surge is analysed. Surge is defined by the difference between measured water level and (predicted) astronomical tide. With this quantified relationship boundary conditions are estimated, for combinations of high sea water level and extreme rainfall that would occur once in 10, 50 or 200 years.

To get these results, a number of analyses are carried out: Data availability and quality, sea level rise, wind effects on the estuary and non-stationary extreme sea water level analyses, are all required prior to correlation between rainfall and surge, joint event ARI analysis and finally the resulting boundary conditions. Each of the following paragraphs gives a short overview of the actions and conclusions.

Data. What data are available, and what is their quality? In quality checks for (sea) water level, rainfall, wind, and flow data are analysed and corrected if appropriate. Especially the sea water levels are important for this study, so additional effort is put into selecting the best and enhancing the quality of these measurements in order to get reliable results for sea level rise, extreme sea levels, surge and tide.

Relative sea level rise. The measured sea level rises every year a bit, the question is however what the long-term rate of sea level rise is? This is partly important for adjusting historical measurements to a comparable 'if they happened today' basis for the extreme value analysis. It also provides insight in how the boundary conditions (and subsequently the flood risk) might change in the coming years. This result also enables comparison of recent history with global projection models near future predictions. The best estimate from the analysis is that the rise in mean sea level from 1990 to 2020 has been 120 mm in total, resulting an average of 4 mm per year. This is illustrated in the SLR1 trend in Figure 1 below.

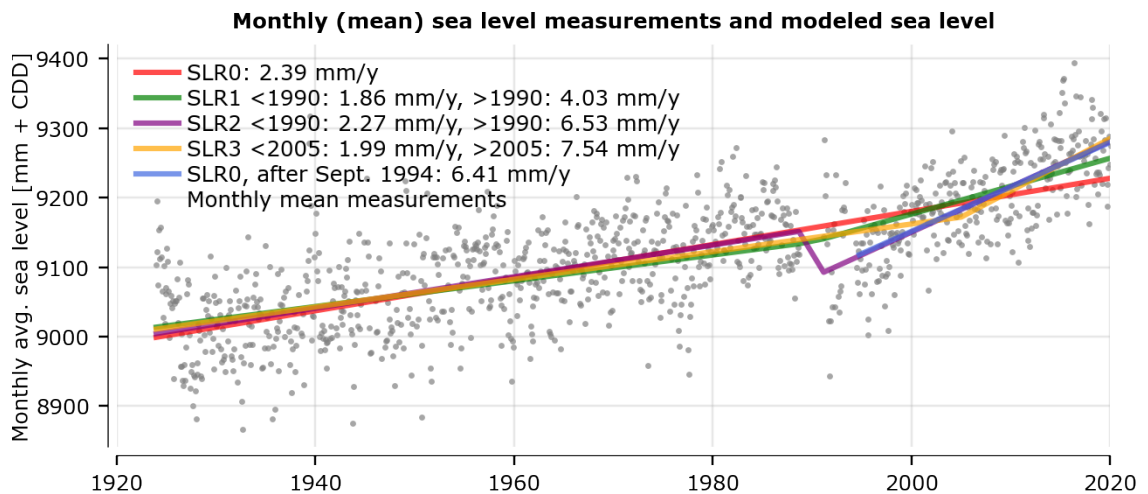


Figure 1 Linear trends in relative sea level rise in monthly averaged water levels at Lyttelton

Wind effects on estuary. Between Bridge Street and Ferrymead Bridge the observed water level differences seem unexplained by just the wind measurements. In this analysis a correlation analysis on the water level, wind and river flow measurements is performed. The main result is that both gauges seem to be shifted vertically from time to time, which disturbs the physical relation. When these shifts are corrected the water level differences seem to be well explained by a mainly south-southwest component of the wind.

Extreme value analysis of the sea water levels. For determining the magnitude of extreme sea water levels in the boundary conditions, it is necessary to have an estimate for the height of infrequent sea water levels. We need at least an estimate for a 200 ARI sea water level to derive extreme boundary conditions. To get these, extreme value distributions are fitted to the measured data and extrapolated. An exponential distribution fitted to the observed peaks gives the best estimate for extreme sea levels. This is the first-time historic sea level events have been adjusted for sea level rise trend. The new results are materially different to precedent work and consistency across the various sites is improved.

The figures below (Figure 2) compare the final results with the results of the extreme value analysis with the previous work. In general, the extreme value distributions derived in this study have slightly flatter slopes, because of the sea level rise correction. This correction also results in increased sea levels, at lower ARIs, but the magnitude differs at each site.

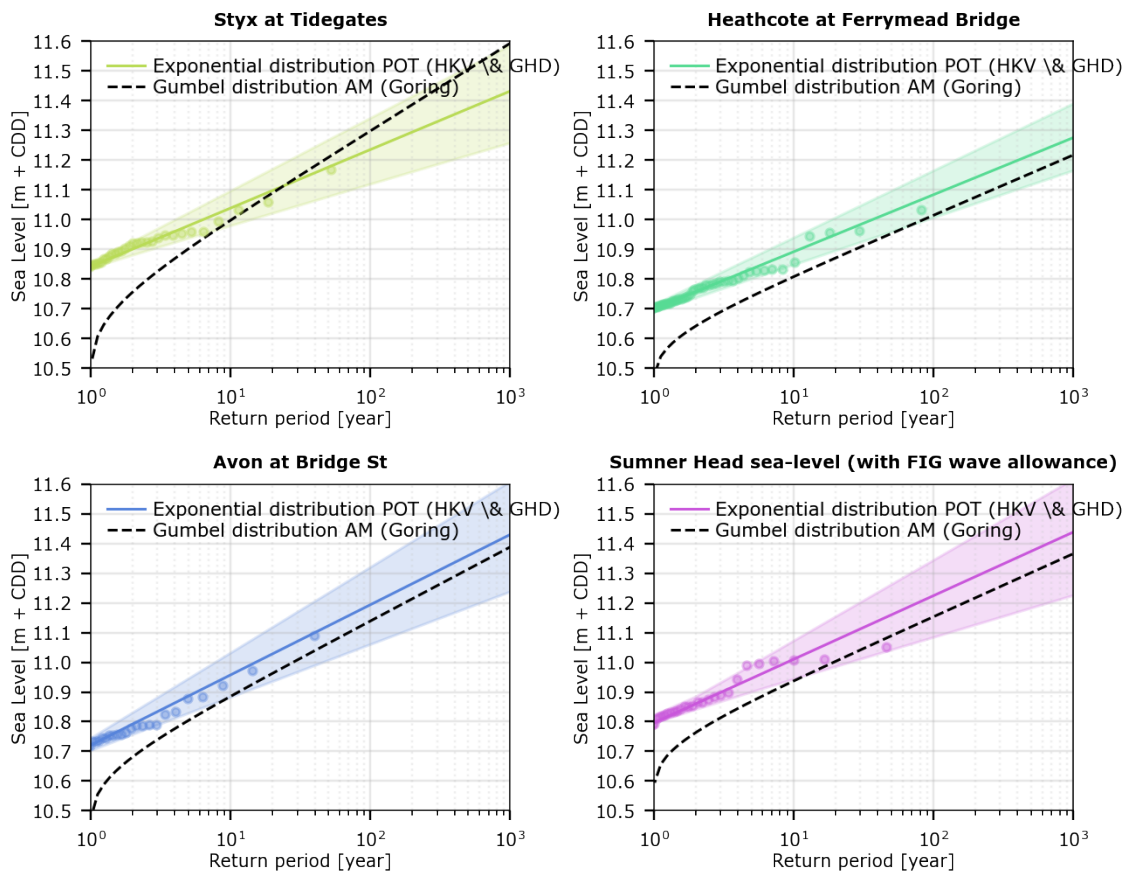


Figure 2 Comparison of Peaks over Threshold Sea Level results with Goring

Correlation analysis rainfall and surge. The correlation between rainfall and surge is quantified with observations and classifications of Kidson weather patterns. The resulting correlation model estimates the probability distribution of the surge for any given rainfall event depth, for a selected rainfall duration. The surge distribution has the highest values for rain events in the 2-10yr ARI region, dropping close to a small positive mean surge for larger ARI rain events. The surge distribution effect is strongest in the longer 24 hrs rainfall duration and weakest in the short 2 hrs rainfall duration. For 24 hrs rainfall duration at 80mm rainfall depth the mean surge is 200mm and the 95th% surge is 400mm (80mm rainfall is depth where the surge distribution is most elevated).

Joint Event ARI. Extreme joint event ARIs are estimated based on the three components: extreme sea levels, extreme rainfall and the correlation between rainfall and surge. Since two variables are classified for the joint event ARI (sea water level and rainfall), there is no single combination that is relatable to a 10, 50 or 200 ARI events. Rather, there are an infinite number of combinations that are 10, 50 or 200 ARI probable events. These combinations are depicted by the lines of equal probability in this chapter. The lines show a convex shape. If there would be no correlation, between rainfall and surge the lines would all be straight diagonal lines. The correlation causes the lines to be convex as illustrated by example on Figure 3 below.

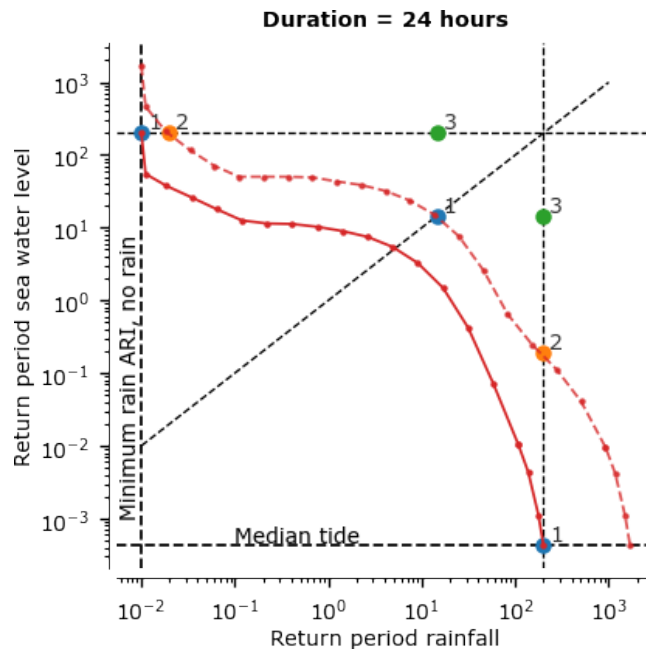


Figure 3 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines

Boundary Conditions. Having quantified the frequency of occurrence of joint rain and sea level events, challenges remain in selection of model boundary conditions as the ideal strategy would require a high number of model runs, to cover the full range of joint events that could occur below a certain ARI threshold. Being practical requires selection of preferably two or three model run setups, and this simplification involves significant compromises in the results and choices as to which compromises are more acceptable.

From the joint probability results, choice of symmetry and avoidance of any unconservatism (allowing conservatism) we pick a new pairing rule recommended to replace the current two point rule. The recommended rain/sea event ARI paired values for long duration rain events (6hr duration and above) are shown in Table 1, and compared to the previous pairings.

Table 1 Recommended rain/sea event ARI paired values for long duration rain events, compared to previous pairings

Joint Event ARI (years)	10 year	50 year	100 year	200 year
Historic (1/10 th rule)	1	5	10	20
Now recommended for long duration (rain >= 6 hrs)	2	7	10	13

Our recommended rule will increase some 10yr flood levels and reduce some 200yr flood levels, with no change in the 100yr flood levels. We also discuss options to adopt three-point rules where greater accuracy and less compromise is justified, as well as rules which are not fully conservative but rather have some mixed conservatism.

There is significant potential still ahead for Council in developing further understanding from this initial study. After we conclude with our formal recommendations for adoption, we list several opportunities for such further work in chapter nine of this report.

Table of contents

Executive summary.....	i
1. Introduction.....	1
1.1 Introduction.....	1
1.2 Purpose of this scope of work.....	1
1.3 Goal of the project.....	2
1.4 Structure of the report.....	2
1.5 Terminology.....	3
1.6 Assumptions and limitations.....	5
2. Data.....	6
2.1 Introduction.....	6
2.2 Overview of data.....	6
2.3 Data availability.....	8
2.4 Sea level data quality and corrections.....	8
2.5 Conclusions for Lyttelton and Sumner.....	10
2.6 Rainfall data.....	10
2.7 Estuary level data quality.....	12
3. Relative sea level rise.....	13
3.1 Approach and results of relative sea level rise.....	13
3.2 Results and conclusions on relative sea level rise.....	13
3.3 Absolute Sea Level Rise.....	15
4. Wind effect estuary.....	19
5. Extreme value sea level statistics.....	20
5.1 Introduction.....	20
5.2 Data and comparison to Goring.....	20
5.3 Sea level rise adjustment.....	21
5.4 Statistical methods.....	21
5.5 Extreme value analysis.....	23
5.6 Sensitivity results.....	27
5.7 Conclusions.....	28
6. Correlation analysis.....	30
6.1 Summary of analysis.....	30
6.2 Conclusions.....	32
7. Joint event ARI analysis.....	33
7.1 Summary of analysis.....	33
7.2 Conclusions.....	36
8. Model boundary conditions.....	37
8.1 Rule options - assessment.....	37
8.2 Rule specifications.....	39

8.3	Tabulated values.....	39
8.4	Discussion.....	40
8.5	Recommendations	41
8.6	Boundary condition conclusions.....	41
8.7	Time series development.....	42
9.	Conclusion and recommendations.....	43
9.1	Overall conclusions	43
9.2	Recommendations for adoption	44
9.3	Potential future refinements.....	44
10.	References	47
11.	Disclaimer	50

Table Index

Table 1 Recommended rain/sea event ARI paired values for long duration rain events, compared to previous pairings.....	iv
Table 2 Structure of the report and appendices	2
Table 3 Terminology	3
Table 4 Data sites list.....	7
Table 5 NSEV Parameters and Levels	27
Table 6 Curve intercept rain/sea event ARI values on the 1:1 line.....	40
Table 7 Curve intercept rain/sea event ARI ratios on the 1:1 line	40
Table 8 Recommended rain/sea event ARI values on the 1:1 line for adoption	41

Figure Index

Figure 1 Linear trends in relative sea level rise in monthly averaged water levels at Lyttelton.....	ii
Figure 2 Comparison of Peaks over Threshold Sea Level results with Goring.....	iii
Figure 3 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines.....	iv
Figure 4 Example of bad data. The grey part has been removed from the time series.....	9
Figure 5 Sea water levels for Lyttelton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.....	10
Figure 6 Rainfall gauges around Christchurch for which historic time series are available. The red lines indicate the edge of the Avon catchment	11
Figure 7 Weights in the averaging of the rainfall data.....	11
Figure 8 Linear trends in relative sea level rise in monthly averaged water levels at Lyttelton.....	13
Figure 9 Linear trends in relative sea level rise in monthly averaged high tide level at Lyttelton	14
Figure 10 Relative versus Absolute sea-level rise (Source: NIWA, after Figure 16, MfE (2017)	16
Figure 11 Contour Map showing uplift and subsidence in mm/year (Source: Pearson et al, 2019)	16
Figure 12 Quadratic fit of the wind speed through the observed water level differences.	19
Figure 13 Data availability of all stations.....	20
Figure 14 Comparison between annual maxima (AM) and Peak over Threshold (PoT) for Lyttelton.	22
Figure 15 Extreme value distribution for Styx.....	24
Figure 16 Extreme value distribution for Sumner	24
Figure 17 Extreme value distribution for Heathcote.....	25
Figure 18 Extreme value distribution for Avon.....	26

Figure 19 Exponential fit Peaks over Threshold.....	28
Figure 20 Comparison of Peaks over Threshold Sea Level results with Goring.....	29
Figure 21 Measured water level (blue) with the predicted tide (red) and the difference, the surge (black). The bottom figure is part of the period shown in the top figure.....	30
Figure 22 Observations of 24-hour average surge with 2-, 9- or 24-hour rainfall sum. The coloured markers indicate the Kidson weather pattern where this is available.	31
Figure 23 Rainfall observation with the correlation models, for which the 5 percentile lines are shown. The coloured crosses indicate extreme historical events.	32
Figure 24 Lines of equal probability (Joint ARI).....	34
Figure 25 Scaled contour lines defining regions of the graph domain which will be exceeded with 10/50/200 year ARI (based on scaling factors).....	35
Figure 26 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines - The coloured dots labelled 1, 2, 3 are reference points used for discussion in the following chapter)	36
Figure 27 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines.....	38

Appendices

Appendix A – QA for Lyttelton and Sumner

Appendix B – Relative sea level rise at Lyttelton

Appendix C – Water level differences on estuary

Appendix D – Rainfall data

Appendix E – Extreme value analysis SWL around Christchurch

Appendix F – Correlation analysis surge and rain on Avon river

Appendix G – Combining statistics to boundary conditions

1. Introduction

1.1 Introduction

Christchurch City Council has been working for many years on developing their understanding of flood risks in Christchurch. As part of this effort, GHD has been providing Council with advice and support by developing a range of numerical models. The quality of the output of the models is improving over time and the results have supported Council in numerous ways, ranging from developing policies to understanding the complex tidal systems that influence significant areas of Christchurch.

Currently GHD is working on the Multi-hazard study, assessing a range of mainly coastal focused flooding hazards that pose a risk at present and in the future to the residents of Christchurch. As part of this study, it was important to consider how the risks of river and tidal flooding coincided so as to better understand and ultimately model using hydraulic models, the joint flood risks in transition areas significantly subject to both risks. Some initial work was carried out and reported by NIWA (ref 1.). This study significantly advances on that initial work.

1.2 Purpose of this scope of work

The scope for this project was divided in various packages to enable Council to proceed if the results of a package were satisfying in terms of evidence that correlation could be proven to be significant.

The scope of this report is to present the findings of the various steps that were taken during this project. The steps were divided into separate work packages and finished by presenting the findings during workshops with Council.

As part of these work packages GHD and HKV:

- Analysed historical measurements of wind, rainfall, barometric pressure and water levels from various locations in Christchurch and in Lyttelton.
- Analysed correlations and patterns in the available data to present trends
- Analysed historical weather patterns influencing short- and longer-term variations in those measurements
- Analysed available literature describing similar analysis and liaising with the authors of the papers to fully understand their assessment
- Drafted a statistical (correlation) model that describes the joint probability of rain and storm surge events for the Christchurch area
- Determined extreme value relationships for various gauged sea level locations to inform the probabilistic model
- Determined extreme value relationships for rain events in some gauged locations to inform the probabilistic model
- Determined long-term sea-level movements to determine potential sea level rise and acceleration
- Determined the correlation between rain and storm surge events
- Determined the individual and joint probabilities of these events
- Determined combinations of rainfall and tide conditions for various locations in Christchurch which satisfy a given statistical probability (for 10, 50, 200-year ARI).

1.3 Goal of the project

The goal of this project is to prove that there is a correlation between the pluvial and tidal flood events in Christchurch and to present a method and results that can be used for future assessment of those events.

1.4 Structure of the report

Each of the chapters in this report are supported by more detailed descriptions, calculation methodology and graphs in the appendices. There is a one to one, chapter to appendix relationship as tabulated below.

Table 2 Structure of the report and appendices

Chapter	Name	Appendix
2	Data	A, D
3	Relative sea level rise	B
4	Wind effect estuary	C
5	Extreme Value Sea Level Statistics	E
6	Correlation analysis	F
7	Joint event ARI analysis	G
8	Model boundary conditions	-

Readers are recommended to read the main body of the report sequentially or according to interest, noting that later sections often rely on conclusions presented in the earlier sections. If there is a particular area where a more detailed understanding is sought, then refer into the associated appendix for the additional reading.

Note that primary authorship of Chapters 2-7 is led by HKV, (except for Section 3.3 Absolute Sea Level Rise). Primary authorship of all other parts of the report is led by GHD. GHD and HKV have each provided assistance and feedback on sections which they are not the primary author. Respective report QA signoffs from each organisation apply to the areas of primary authorship. GHD signoff as contractual lead also confirms our satisfaction that HKV have completed suitable QA of their contribution.

For HKV this report has been reviewed and signed off by Prof. dr. ir. Matthijs Kok. Matthijs is Professor of Flood Risk at Delft University and holds a MSc degree of Applied Mathematics from the Twente University (1981) and a PhD degree of Operations Research from the Delft University of Technology (1986).

Additionally, the report has been reviewed by Christchurch City Council's Peer Review Panel for the Multi Hazards study. We thank the panel for their cooperation, enthusiasm and contributions during the project*.

* This paragraph had a minor update on 9 June 2021 after the reports initial release. This does not impact the intent of this section.

1.5 Terminology

Table 3 Terminology

Term (NZ)	NL	Description
ARI	T	Average Recurrence Interval is the average time period in years between events of a certain size e.g.: flood flow, flood level, rainfall depth, peak tide level)
AEP		Annual Exceedance Probability is the probability of a certain size event being equalled or exceeded in a single year – in common applications AEP is the inverse of ARI
HIRDS		High Intensity Rainfall Design System developed by NIWA and able to be used to estimate the magnitude and frequency of high intensity rainfall at any point in NZ, currently in version 4.
WWDG		Waterways, Wetlands and Drainage Guide outlining the Christchurch City Council’s waterways and wetlands philosophy and stormwater quality and quantity design guidelines
CDD		Christchurch Drainage Datum
LVD37		Lyttleton Vertical Datum (MSL) 1937
MSL		Mean Sea Level
Boundary conditions	Load variables	Input conditions for a hydraulic model, often time and space varying, such as rainfall, sea level, groundwater level, wind speed and direction
Result location point		A location at which hydraulic model results are selected and exported to the statistical database.
Statistical database		A database primarily containing hydraulic model results along with associated model setup parameters and reference information such as x, y of result locations.
Hydraulic model		A deterministic computational model simulating real life water flows based on numerical approximations to fundamental equations describing the flow of water. For example, the Avon/Estuary hydraulic model.
Statistical model		A mathematical model that provides a relationship between a set of random and non-random variables
Design point(s)		In plural design points are a set of boundary conditions which cause the same flood level. Each condition has a measure of probability from which the most probable condition(s) can be selected as a particular design point for further analyses.
Extreme value analysis		A statistical modelling tool used to estimate the likelihood of occurrence of an extreme event based on observed/measured data

Term (NZ)	NL	Description
EV		Extreme Value
NSEV		“Non-Stationary”. A more general extreme value analysis where the underlying variable is not assumed to be stationary (statistically the same mean and variance) over time.
RMSE		Root Mean Squared Error – statistical measure of goodness of fit for a trend relationship fitted to measured or observed real life data
POT (PoT)		Peak over threshold. One method of selecting peak values from a time series for extreme value analyses.
Location parameter		For POT extreme value analysis this is the value at the ARI = 1 yr position (in this report meaning sea level, in CDD datum)
Scale parameter		For POT extreme value analysis this is the slope of the fitted extreme value trend line defined using the natural log of the ARI value
SLR		Sea Level Rise
Sea level rise (relative)		Sea level rise as measured relative to the adjacent land mass (typically a cluster of survey benchmarks). If the land mass is not moving, then this equates to absolute sea level rise.
Sea level rise (absolute)		Sea level rise as measured relative to a known stationary datum. In modern times this has typically been established from satellites which are independent of any land movement.
Bi-linear sea level rise analysis	Trend break analysis	Linear sea level rise analysis where the slope (rate of rise) is allowed to change at some point in history
Astronomical tide		The contribution of planetary motions to the sea water level. The astronomical tide is the sum of the tidal constituents.
Fitted tide		The fitted or predicted tide is determined by fitting a number of known tidal constituents to the measured sea water level. If this fitting process is done correctly (which depends on the quality of the measured data), it results in a number of tidal constituents than can be summed towards the predicted astronomical tide.
Predicted tide		Similar to fitted tide. The term is more used for future tide predictions.
Surge		The surge is the difference between the sea water level and the astronomical tide. In practice this becomes the difference between the measures sea water level and the fitted astronomical tide. For Christchurch the main component in the surge will be the barometric pressure, but it can also be caused by for example seasonal temperature differences in the sea water level.

Term (NZ)	NL	Description
EQ		Earthquake
VLM		Vertical Land Movement
cGPS		continuous Global Positioning System – used to monitor ground level continuously over long periods to identify VLM
FIG waves		Far Infra-gravity waves. An ocean water level phenomenon typically with a wave period of 1-20 minutes, in our study significant at Sumner Head.

1.6 Assumptions and limitations

There are a number of limitations and assumptions inherent in this work which the reader should understand and for which the author does not accept responsibility.

The analyses presented herein are variously dependant on the quality of measured data collected by Council and others. We have carried out limited data quality inspection and, in many cases, found cause to adjust, delete, fill gaps or otherwise improve the data quality. The inspection processes are not exhaustive, and some types of data quality issues are not discoverable from our inspection processes. In many cases corrections are inferred, different inferences are often possible, and it is not practicable to quantify confidence in our chosen inferred understanding in adjusting the raw data. From our inspection we did discover data quality issues which reduce confidence in certain results or have influenced decisions about which data to select for our analyses.

From our inspections, we did not find any material issues or concerns with the data quality that was used to derive the Lyttelton sea level rise findings, Sumner extreme value findings, rainfall data or the resulting correlation model and conclusions on joint event probability for the Avon catchment.

Throughout evaluations presented in this report we assume stationarity of rainfall frequency, wind distribution and stationarity of sea level variability (variations from mean seal level). Assumed stationarity means that these variables do not change over time and we do not consider the possibility nor implications should they exhibit change over time trends.

In the joint probability analysis, the above also implies an assumption of stationary (non-time varying) multivariable correlations.

Similarly, HIRDSv4 rainfall statistics which are used herein assume stationary statistics and because it uses data periods of different lengths for different sites its conclusions could be sensitive to any actual non-stationary rainfall frequency characteristics.

We also rely somewhat on the assumption of historical non-time varying (stationarity) of flood water level frequencies at Gloucester St and at PS205. This level data is only used to identify dates of interest and so the sensitivity of our findings to anthropogenic changes in flood level response to rainfall is expected to be minimal.

2. Data

2.1 Introduction

The statistical analyses in this report are based on measured data. This chapter outlines the data used. Section 2.1 shows the water level, wind and rainfall data. Section 2.4 describes the analysis of the data quality and any corrections made to the data, the water level measures for Lyttelton and Sumner in particular. Section 2.5 describes the data sources used in the analysis for Lyttelton and Sumner. Section 2.6 discusses rainfall data.

2.2 Overview of data

This analysis uses water level (sea, estuary and river levels), rainfall, wind and barometric pressure data from various locations across Christchurch:

- Sea water level measurement at Lyttelton and Sumner Head are the primary data source. These measurements are used in the Sea Level Rise analysis (Chapter 2.6), the extreme value analysis (Chapter 5) and the Correlation analysis (Chapter 6). The data are provided by NIWA, LINZ and John Hannah. These data are extensively checked and corrected such that the quality is improved for use in this study (also see section 2.1, 2.4 and Appendix A).
- Additional water level measurements at Styx, Avon (Bridge Street) and Heathcote (Ferrymead Bridge) are used in the extreme value analysis (also see Chapter 5).
- Wind speed and direction data from Living Earth, discharge and water level data at Bridge Street and Ferrymead Bridge are used in the analysis of wind effect in the estuary (Chapter 3.3).
- Rainfall data from seven stations in and around Christchurch are analysed (also see section 2.6). The data and statistics from the station of Christchurch Botanical Gardens are used in the correlation analysis (Chapter 6).

The following tabulates all the data locations and types used in this project. Most data was supplied to GHD by Council, from NIWA (Kathy), unless otherwise noted.

Table 4 Data sites list

Site no.	Name	Type	Start date	Source	Data availability [%]
66440	Styx River at Tide Gates	Level	10/7/1990	NIWA	93
66602	Avon at Gloucester St Br	Level & Flow	7/7/1980	NIWA	87
66605	Heathcote at Ferrymead Bridge	Level	1/1/1974	NIWA	94
66611	Avon at PS205	Level	21/8/1987	NIWA	93
66634	Avon at Bridge St	Level	18/9/1997	NIWA	93
66670	Avon at Fitzgerald Ave	Level	16/5/2016	NIWA	100
66699	Sumner Head NIWA	Sea-level	3/6/1994	NIWA	98
GHD3	Sumner LINZ	Sea-level	11/8/2010	LINZ	99
	Lyttelton Port			LINZ	98
	Lyttelton Port			John Hannah (monthly)	94 (monthly values)
66680	Lyttelton Port	Sea-level	30/6/1903	NIWA	89
324610	Firestone Factory, Papanui	Rainfall	1/1/1981	NIWA	94
325616	Christchurch Botanical Gardens	Rainfall	1/1/1962	NIWA	100
325617	Avon at PS205	Rainfall	1/1/1987	NIWA	100
4843	Christchurch Aero	Rainfall	1/1/1955	Cliflo	100
324607	Styx at Lower Styx Road	Rainfall	1/1/1987	NIWA	100
325603	Shirley rain gauge	Rainfall	1/1/1940	Cliflo	96 (daily values)
325507	College of Education	Rainfall	1/6/1964	NIWA	100
37654	Brighton Pier	Wind	27/8/2009	NIWA	93
43967	Bromley	Wind	23/10/2019	Cliflo	95

Site no.	Name	Type	Start date	Source	Data availability [%]
GHD1	Living Earth, located at the Eco Drop metro transfer station	Wind	1/4/2015	Living Earth (via Graham H)	99
66697	Sumner Head barometric pressure	Baro	8/3/1996	NIWA	100
4843	Christchurch Aero	Baro	1/1/1960	Cliflo	100

2.3 Data availability

The most critical data for our analysis was sea level and estuary level data. The distribution of sites was generally satisfactory however it would have been advantageous to have had longer histories of measured data to better support estimation of the probabilities of high extreme sea levels.

Rainfall data was relatively plentiful for this analysis. The number of gauges available for the earlier historic period used in the correlation analysis was a minor weakness, but given the typically high uniformity of rain on the catchment and the moderate sensitivity to individual event measurements in this process, the rainfall data posed no material concerns.

Barometric pressure data seemed to be suitable for this analysis with coverage back to before 1962 which was chosen as the beginning of the key correlation analysis. A longer record would have helped marginally with the sea level rise analysis.

Wind data was somewhat sparse, fragmented over time and with varying significant influences of local topography. While only the Living Earth wind data was used for this preliminary assessment, other potentially useful wind data is available and we have tabulated the two highest relevance sites for future reference. We also note potentially useful records from www.windfinder.com including the Mt Pleasant Yacht Club as well as private wind gauges. The NZ Met service also collects and presents private wind data using WOW (Weather Observations Website) at www.metservice.com/maps-radar/weather-stations/nz.

We note that the Living Earth wind site has a short mast and is close to significant buildings and vegetation so its measurement may not well reflect local clear air wind conditions. We recommend further work on the wind assessment to consider and compare multiple sites to enable data QA and to use a longer data time period. Eventually it would be desirable for the joint water level risks at Bridge St to include an explicit statistical allowance for wind setup.

2.4 Sea level data quality and corrections

The Lyttelton and Sumner sea level data are the most important data sources for this study. The water levels are important for the correlation analyses and for the estimates of sea level rise. Several sources of (partly overlapping) data are available. All have some errors in them. Therefore, the overlapping parts have been visually checked and corrected. Data sources for other stations (wind, water level, discharge, rainfall) have also been checked before the application in the analysis. This process is described in the appendix.

The following data sources are available for Lyttelton and Sumner:

1. Sumner tide gauge measurements from NIWA, from mid-1994 onward.
2. Sumner tide gauge measurements from LINZ, from mid-2010 onward.
3. Lyttelton tide gauge measurements from NIWA and LINZ, from 1924 on an hourly basis. The data is measured by the Lyttelton Port Company, stored and processed by LINZ. The pre-1924 data is not used in the analysis as the data in period is considered to unreliable (and considerations on the data are unknown).
4. Lyttelton monthly averages from John Hannah.

The data of John Hannah is considered as the most reliable source in terms of datums and corrections as this data has already undergone quality analysis and correction by John Hannah (which is logged). This data is however not present in the most recent years and higher resolution (hourly) data is available from LINZ. For the hourly and higher frequency data, LINZ is the official custodian and reliable data source with NIWA only holding some data on the request of CCC and having no associated records of reliability nor correction status.

The process of data checks involves comparison and corrections to the data using the three sources of data for Lyttelton (LINZ, John Hannah and NIWA) and aims to construct a single data series for the time period from 1962 to date. The comparisons and corrections includes (1) Removal and interpolation of missing values, (2) visually inspection and removal of invalid data from the time series - see Figure 4 for an impression, (3) corrections basis based on John Hannah's log and (4) evaluation the different data sources in recent years. The full procedure is described in Appendix A.

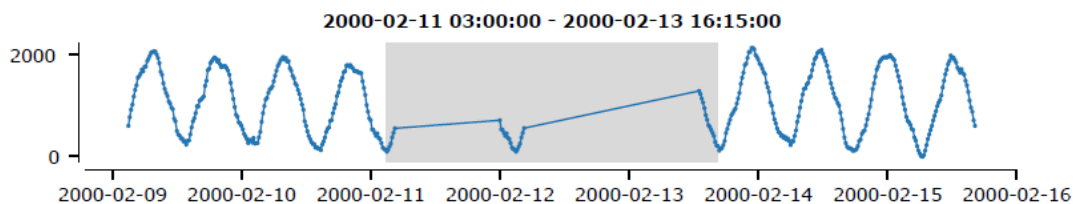


Figure 4 Example of bad data. The grey part has been removed from the time series.

After corrections the three data sources are compared (1) Sumner [NIWA data], (2) Sumner [LINZ data] and (3) Lyttelton. The comparison in Figure 5 shows:

- In the left plot Sumner (NIWA) is compared to Lyttelton. We observe that Sumner often has lower values than Lyttelton. These are the periods in which Sumner has consistently lower values.
- In the middle plot Sumner (LINZ) is compared to Lyttelton. The deviations are small, Lyttelton seems to be a bit higher on average.
- In the right plot Sumner (LINZ) is compared to Sumner (NIWA). Here we see the same effect of the lower values for Sumner (NIWA) as in the first plot.

Lyttelton and Sumner (LINZ) gives the best agreement in data when considering all sea level observations.

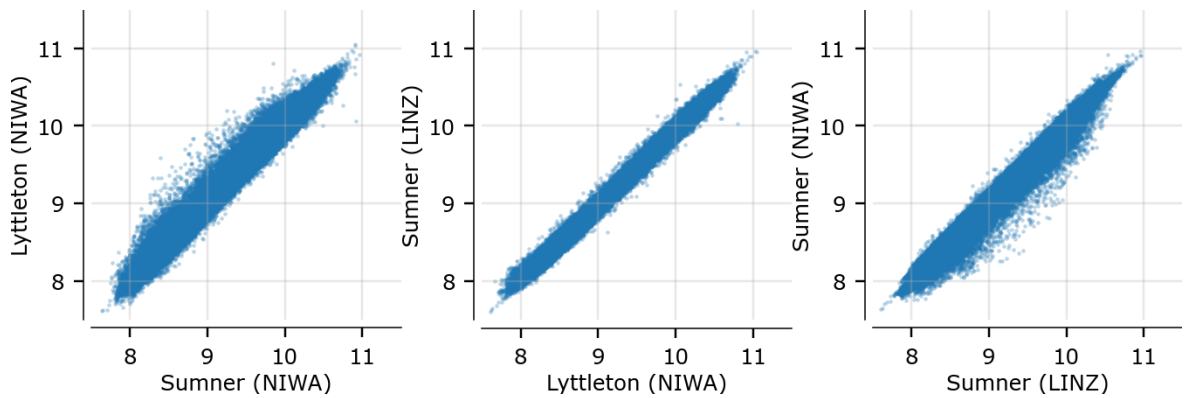


Figure 5 Sea water levels for Lyttelton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.

2.5 Conclusions for Lyttelton and Sumner

The data at Lyttelton and Sumner are used for the extreme value analysis, the correlation analysis and the sea level rise analysis. For Lyttelton the data resulting for the corrections and checks described in section 2.4 is used.

For the extreme value analysis, we need data that has good quality in the maxima. The temporary lowering of Sumner (NIWA) data is a point of concern in the analysis. However, the LINZ data is only available for a short time period. We therefore use Sumner (NIWA) for EV analysis at Sumner. The temporary lowering of data is likely also excluded in the peak selection process in the EV analysis.

For the correlation analysis we mainly need consistent data (for estimating tide) with minimal bias (for estimating surge). Since Lyttelton data is available all the way back to 1960, the source has our preference for this purpose. If needed, missing periods could be filled by Sumner (LINZ) if available and Sumner (NIWA) otherwise. Continuous data with no gaps (gaps filled) is desirable for deriving the surge, but a modest frequency of data gaps is acceptable and has no material impact for the analysis of the tidal constituents, so a single data source with gaps unfilled is used for that.

For the purpose of estimating sea level rise we need data with minimal biases. Sumner (NIWA) data has consistent lower periods and is therefore not preferred. This can partly be solved by using the Sumner (LINZ) data, but this is only available for 2010 and onwards. The Sumner results therefore have a higher level of uncertainty (also see Appendix B).

2.6 Rainfall data

Rainfall (or precipitation) data is used as a basis for the correlation analysis in Chapter 6. Rainfall stations in the Avon River catchment are therefore considered (Figure 6). These are:

- Firestone Factory, Papanui
- Christchurch Botanical Gardens
- Avon at PS205
- Christchurch Aero
- College of Education
- Styx at Lower Styx Road
- Shirley rain gauge.

The data is aggregated and missing values are filled. The measurement from the seven stations are subsequently combined using an area weighted average for rainfall in the Avon river catchment (Figure 7). This weighted time series is applied in the correlation analysis.

Estimated catchment aggregate hourly rainfall back to 1962 is used for the correlation analyses. The only gauge which collects hourly data since 1962 is Christchurch Botanic Gardens, with some early data collected daily, and most sites data collection start more recently than 1962. The processes by which best estimates of catchment rainfall are derived from the available data are detailed in Appendix D.

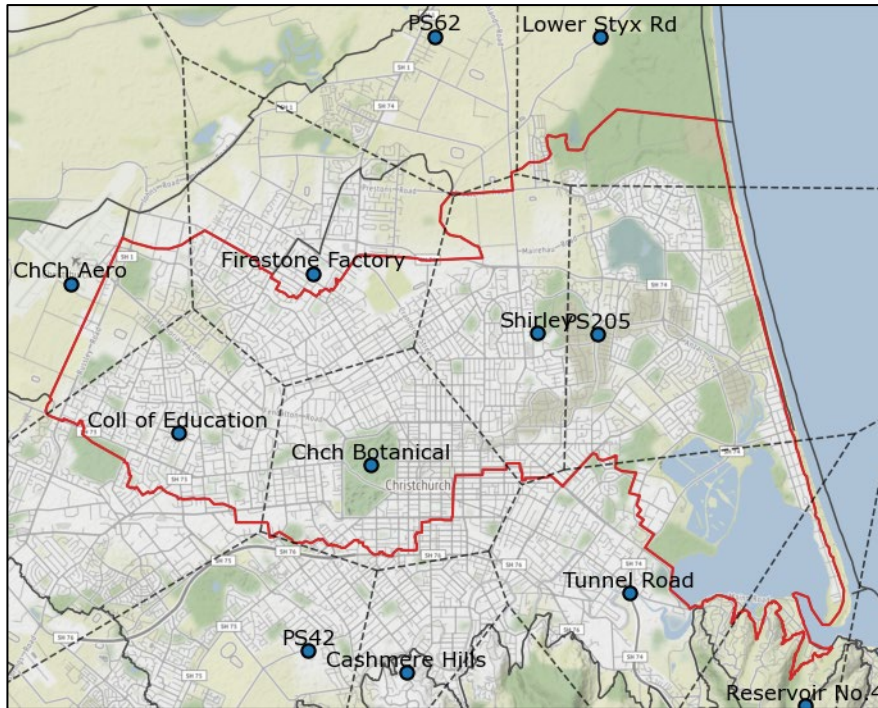


Figure 6 Rainfall gauges around Christchurch for which historic time series are available. The red lines indicate the edge of the Avon catchment

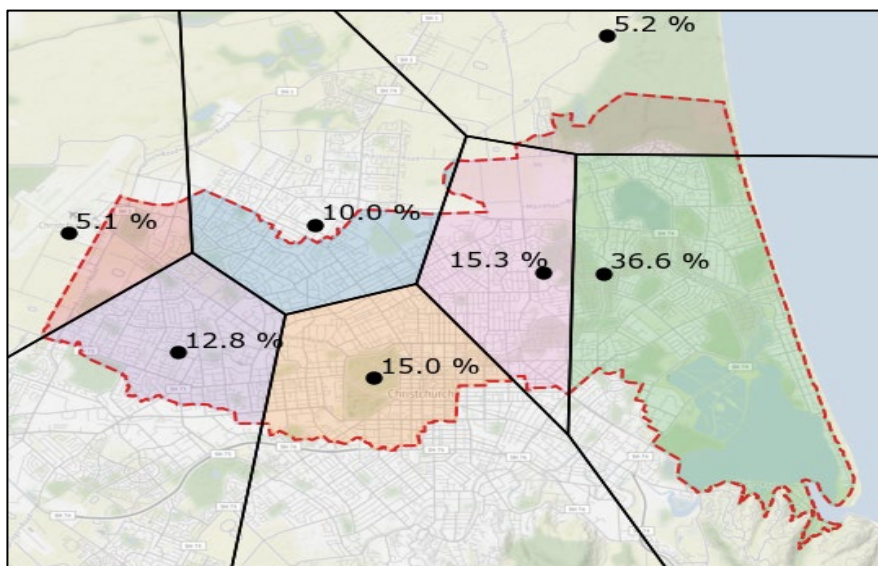


Figure 7 Weights in the averaging of the rainfall data.

Given the large number of rain events used in this analysis, the conclusions here are relatively tolerant of possible rainfall data imperfections and details of the gauging.

2.7 Estuary level data quality

During high tides, water level at the two-estuary water level monitoring sites (Avon Bridge Street and Heathcote Ferrymead) are expected to be similar other than during intermittent periods of wind setup. The estuary water level data supplied for these two sites however contained numerous exceptions, which were concluded to being evidence of some level measurement error.

Council undertook a data review process and estimated a series of corrections to the data as described in Appendix section C.1. In the majority of this report, the raw data was used in draft reporting, but this has been replaced with corrected data in the final report. The uncorrected data remains is mainly relevant to Appendix section E.3, where it is used to demonstrate data consistency with Goring's previous work on extreme sea levels.

3. Relative sea level rise

3.1 Approach and results of relative sea level rise

To assess the current flood risk to the city of Christchurch and to make substantiated predictions on future flood risk, it is important to know the historic rate of sea level rise. This information is used to correct for historic sea level rise in flood risk assessments of the current situation. It also provides an indication for the likely short to medium term prediction of future sea level for comparison with sea level rise projections.

This analysis investigates if there is historic sea level rise, provides insight in the rate of rise and explores whether the rate of sea level rise is potentially accelerating. The analysis makes use of sea level measurement of the past century from the Lyttelton Port tide gauge. The *relative* sea level changes are investigated, this therefore also includes potential land settlement (captured in the measurements). Settlement should be subtracted from the result to find the absolute sea level rise.

The analysis applies statistical techniques from the Dutch Sea Level monitor. This approach filters temporal physical components from the measurements to improve estimates of the average sea level for any given time period as much as possible. These physical components include i.e. tidal fluctuations, barometric surge, decadal oscillation patterns (IPO and ENSO) and season effects. Linear regression is applied to the residual signal to investigate potential trends. As the analysis focusses on long term patterns of sea level rise, we apply monthly averaging to further reduce noise.

3.2 Results and conclusions on relative sea level rise

The typical simple linear sea level rise trend is evaluated. Based on international literature three piecewise linear trend models are also evaluated (see Appendix B.4 for details of the trend models):

- SLR0: A linear increase of the relative sea water level for the period 1920 – 2020.
- SLR1: A trend break in 1990, resulting in accelerated sea level rise.
- SLR 2: A trend break in 1990 and a shift, resulting in accelerated sea level rise.
- SLR 3: A trend break in 2005.

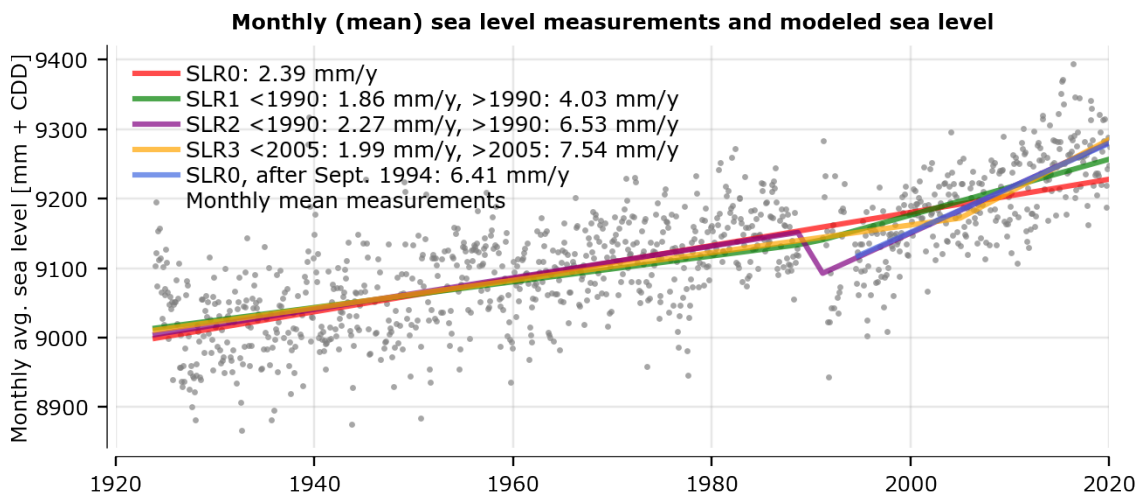


Figure 8 Linear trends in relative sea level rise in monthly averaged water levels at Lyttelton

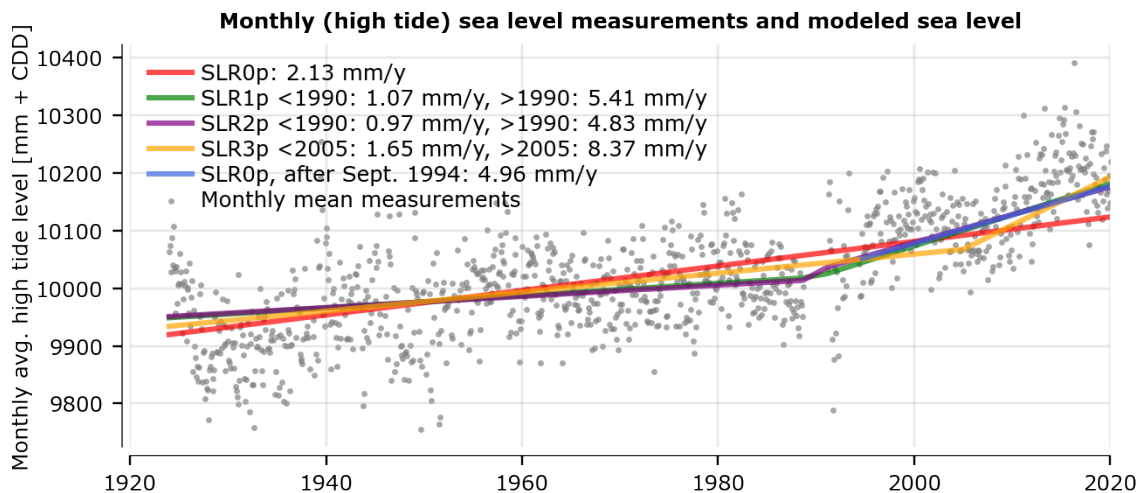


Figure 9 Linear trends in relative sea level rise in monthly averaged high tide level at Lyttelton

The analysis on sea level rise shows:

- All applied models show an increasing relative sea level over the period 1920 – 2020.
- The relative sea level rise rate of about 2 mm /year (1.85 – 2.25 mm/year in the analysis) in the 1920 – 1990 period, while in recent years (1990 to date) the rate is about 5 mm/year (4-7 mm/year in the analysis).
- For Lyttelton the rate of mean sea level rise is about 70% higher than the rate of high tide rise over a 70-year period (based on SLR1: 1.07 mm year vs 1.86 mm/year). This difference implies that the tidal range may be increasing (as well as the mean sea level). Further analysis would be required to validate and understand cause(s) of this difference. Within this report we have little motivation to investigate further because the rate of sea level rise is only used to correct the historic sea level data for the extreme value statistics. We show in Section 5.6 (and Appendix section E.4) that our conclusions with respect to extreme values of sea level are insensitive to the choices of sea level rise trend (SLR1, SLR2, SLR3). This means for the purposes of this report, we do not need high confidence or a more exact conclusion on the sea level rise questions.
- The ‘1990 shift’ in SLR2 is slightly positive in the high tide results, in contrary to the mean level results where it is negative. Physically this model was already questionable, but due to this inconsistency between mean and high tide we will not use SLR2 for the remainder of this analysis.
- The SLR3 result is of interest due to its high rate of recent sea level rise, however the 15 year duration of the trend analysis is too short to be considered reliable and we do not recommend it’s use for subsequent analyses either.
- The significance of the increased rate of sea level rise in the second portion of SLR1 is compelling so we do not recommend use of SLR0. This leaves SLR1 recommended as the preferred SLR trend.
- The most likely explanation for the differences between mean and high tide is that over the course of time the method of measuring changed a bit, which could lead to a gradual change in the measurements. In one of the earlier figures, we saw a drop in the low tide values after the data gap around 1990, which might support this possibility.
- Regardless of the significant and unquantified concerns with the Sumner NIWA sea level data sea level rise trends on this data were calculated and showed generally satisfactory

consistency with Lyttelton for the same period. This indirectly adds some level of confidence to the Sumner NIWA data for this purpose, but this does not diminish our reliance on Lyttelton for confident understanding of sea level rise trends.

- As the best estimate of sea level rise trend, we conclude to report the mean tide results for Lyttelton as value for the sea level rise. It is a very long time series and agrees with Sumner for the post 1994 data. So, use the rates of sea level rise from scenario SLR1, with a trend break in 1990 and no vertical shift in the data.

3.3 Absolute Sea Level Rise

Paul Denys and John Hannah provided valuable information regarding measurement and knowledge of vertical land movements. This information allowed data corrections as discussed in 2.4 and contributed to developing an understanding of the local sea level rise around Christchurch. Most of the available data focusses on Lyttelton, but more recent data (post EQ) is available at or close to other measuring points (see also Figure 11).

When land movement occurs over a wide area and slowly over long timeframes, then any traditional terrestrial survey approaches, (as distinct from satellite technology), are typically unable to detect the movement. Historically sea level was assumed to be a constant and with sufficient observation thereof then relative land movement could be detected. Vertical movement of both land and sea is however now of interest and it is generally only practicable to observe this recently using satellite survey technology, including high spec GPS survey (with temporary occupancy times) and continuous GPS. Continuous GPS collects vastly more data and through averaging is able to estimate levels with much lower error bounds.

In Christchurch a continuous GPS unit (LYTT) was established in 2000 near the Lyttelton tide gauge, to understand land movements and absolute sea level rise and land movement. In the first ten years this site collected steady trend data. In 2010 this was disrupted semi continuously by earthquake events. Denys (2020) analysis of the data until 2015 understands the post seismic displacements modelled using a log decay function continue to be significant through to 2015, however the aggregate post seismic movement is consistent with the same linear pre-earthquake trend plus reasonable estimates for the post seismic effects. As part of this project a regional cGPS unit (MQZG) was also established apparently somewhere in the vicinity of Gebbies Pass.

A more recent continuous GPS unit (SMNT) was also established in 2015 near the Sumner Head tide gauge.

If sufficient data is available, the relative sea level measurements can be translated to absolute sea level movements which include vertical land movements (See Figure 10).

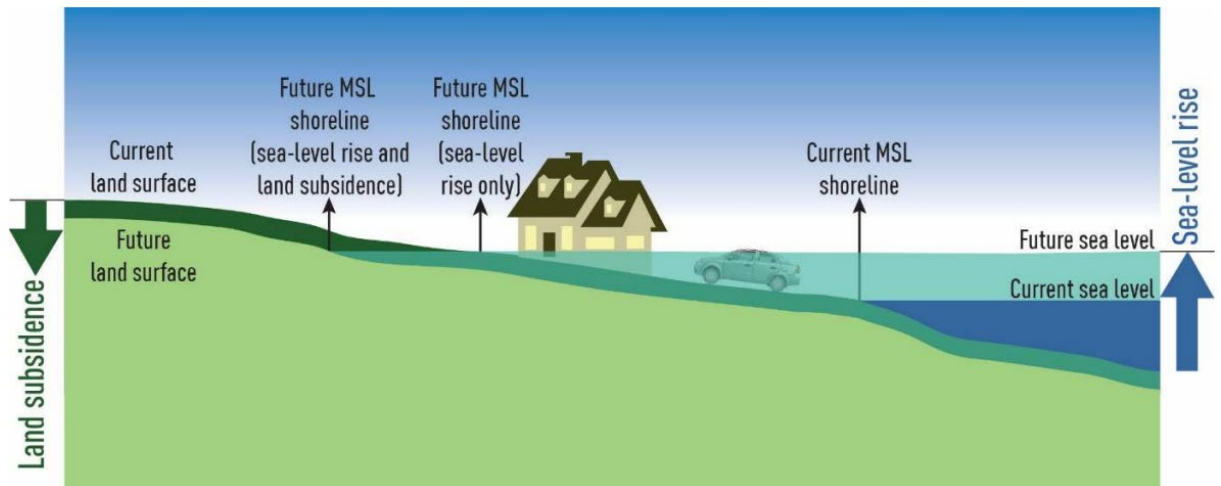


Figure 10 Relative versus Absolute sea-level rise (Source: NIWA, after Figure 16, MfE (2017))

Denys et al (2019) describes the “late 20th century sea level acceleration” that is also observed in chapter 3.2. Vertical land movements (VLM) need to be included in sea level rise discussions in order to inform relevant stakeholders on the actual impact of absolute sea level rise.

In Pearson et al (2019) the summary of the vertical land movements since the two large Christchurch earthquakes are presented. Yaldhurst (YALD) is the only station where uplift is measured (approx. 1.64 ± 0.47 mm/year) and Sumner (SMNT) showing a close to zero subsidence/uplift measured for 4 years (2015-2019) of measurements.

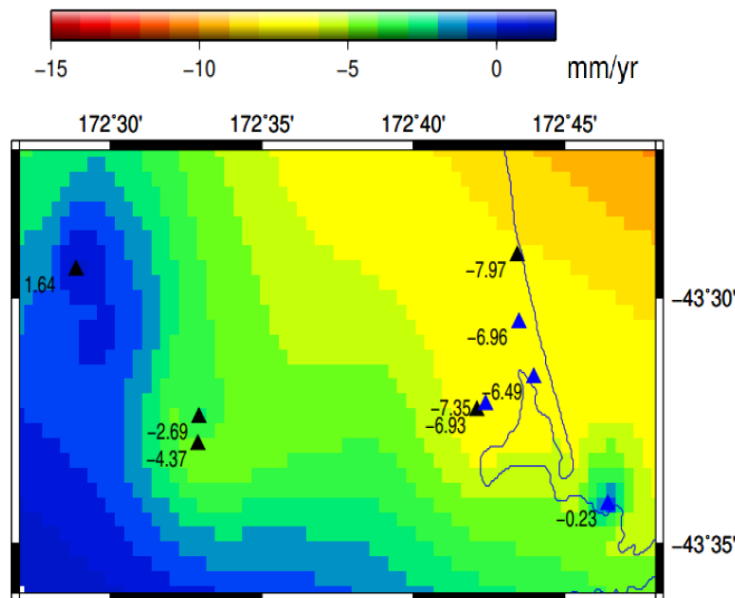


Figure 11 Contour Map showing uplift and subsidence in mm/year (Source: Pearson et al, 2019)

In the 2019 annual report from Otago University School of Surveying to ECan regarding vertical ground motions at the Sumner tide gauge, it is concluded that the estimated -0.23 mm/yr vertical land subsidence at the Sumner Tide gauge is not yet statistically different to zero. With time additional years of data will reduce the confidence interval at SMNT and may demonstrate a significant but low rate of subsidence. That study also reports that the Sumner gauge (referred to as SMNT) is located on a relatively dense layer of sediments compared to the looser packed sand in the north/eastern areas of Christchurch.

The Sumner tide gauge was once lowered in order to improve measurements at lower water levels. It is unlikely that vertical land movements are incorporated in updated reference levels of the tide gauge. If the limited years of VLM data at Sumner are taken to indicate long term stability (zero land movement), then the sea levels measured at Sumner would provide an indication of absolute Sea levels at this location.

Information on Lyttelton land movements are found in Denys et al (2020). Lyttelton's tide gauge has a continuous GPS (cGPS) measurement that provides continuous measurements of the location and height of the tide gauge. The effects of the reception of this GPS related to the measurement accuracies are unknown. The GPS data used is from 2000 until 2015. Besides the 3 known corrections of the tide gauge due to incorrect settings and changing reference levels no corrections are known. The 2010 and 2011 earthquakes resulted in an accumulated uplift of 111 mm of the tide gauge until Dec 2012. The uplift was corrected 'stage by stage' by LINZ, therefore influencing the sea level measurements.

Denys (2020) states that the cGPS data at LYTT in Lyttelton shows an average subsidence of -0.47 mm \pm 0.2 mm per year over the measured 15.7-year period. His conclusion is that the best estimate for the rate of rise of absolute sea level over the 1903 – 2012 period are 0.27 mm/yr less than the relative sea level rise. This difference is relatively small compared to the recent relative sea level rise trend at Lyttelton determined in this study.

As a result of the various changes in the tide gauge measurements, it is unclear how much of the continuous (/yearly) vertical land movements are incorporated into the water level measurements. Estimating historic absolute sea level rise at Lyttelton based on these measurements therefore requires further investigation in order to be conclusive.

Further assessment focussing on satellite sea level data and more in-depth assessment of the vertical land movements versus the corrected water level measurements can contribute towards a clearer understanding of absolute sea level rise around Christchurch.

Two other key sites at Bridge St and Ferrymead warrant comment. These sites are not used for estimating sea level rise and so comment on vertical land movement there is not directly relevant to this study, but it is of general interest to due to their water level records.

We found no land movement data within a useful distance from Ferrymead and are unable to draw conclusions there in relation to absolute sea level rise. Pearson 2019 shows a cGPS location near to Bridge St with a large rate of subsidence, however we understand this measurement location is on soft ground away from Bridge St and while representative of the general land in the area, will not be representative of movement of the Bridge St Bridge which has deep foundations. The rate of subsidence of the bridge is expected to be much less and might be approximately zero so we are unable to draw conclusions there in relation to absolute sea level rise. CCC have begun a terrestrial survey programme to relate the Bridge level to adjacent cGPS levels in order to answer this question.

Our conclusions are that the significance of historically unaccounted for land movement at Sumner and Lyttelton may be in the order of -0.25mm/yr, meaning absolute sea level rise would be marginally less than relative sea level rise. If this is confirmed, then these influences are minor in comparison to the understanding of historic sea level rise for both Sumner and Lyttelton presented here in the earlier sections of this Chapter.

While much detail is provided in associated reports, some key details of the cGPS monitoring project at Lyttelton remain obscure. In Denys 2020 we find several reported rates of change where the benchmark against which the reported change is measured is unclear, or the time period for the rate of change is unclear.

Most critical would be getting better understanding of the data sequence which starts with semi continuous monitoring levels of the land-based network of benchmarks, monitoring level over

time of the tide gauge top level, measuring from this top level down to sea level and thus establishing the monitored sea level. This data is collected by Lyttelton Port Company and communicated to LINZ as custodians. Throughout history the physical details and characteristic risks in this data sequence do change. And any part of this sequence at any time can suffer from accidental errors or inaccuracies, which are sometimes discovered later. The intermittent nature of survey checks and intermittent batch data reporting to LINZ create wrinkles in time where the best estimate is no longer the 'at the time' estimate and reporting protocols between LPC and LINZ are vital regarding reporting of raw and corrected data distinctly as is coordination between them when further corrections are discovered after the data is communicated. While LINZ rely on LPC for communicated data, in some instances LPC and LINZ may not have a common view on the best interpretation of corrections.

Further general investigation would be required to gain a more thorough understanding of the cGPS data implications particularly for recent Absolute Sea Level Rise near Christchurch. Further specialist analysis of the Lyttelton cGPS data since 2015 would also be recommended.

4. Wind effect estuary

The correlation analysis in Chapter 6 focuses on predicted tide plus surge and rainfall as the primary explanatory parameters in determining extreme water levels in and around coastal Christchurch. The wind, its magnitude and direction, also has an influence on the water level. This is especially applicable to the water levels in the Avon Heathcote Estuary. This chapter investigates to what extent water level differences between Bridge Street (Avon) and Ferrymead Bridge (Heathcote) can be explained by wind.

The peak tidal levels at Bridge Street (Avon) and Ferrymead Bridge (Heathcote) are compared in order to minimise shallow water effects on the comparison. The difference between the peak tides is corrected based on a 30-day moving average to account for likely datum type inaccuracies in the measurements. A linear regression model is applied to explain the water level differences across the estuary. The model uses (1) Heathcote discharges at Buxton Terrace, (2) Avon discharges at Gloucester Street, (3) wind speed and direction as the explanatory parameters and (4) the average water level at the estuary. The model is based on 4 years of data as this the period in which the Living Earth wind data was available.

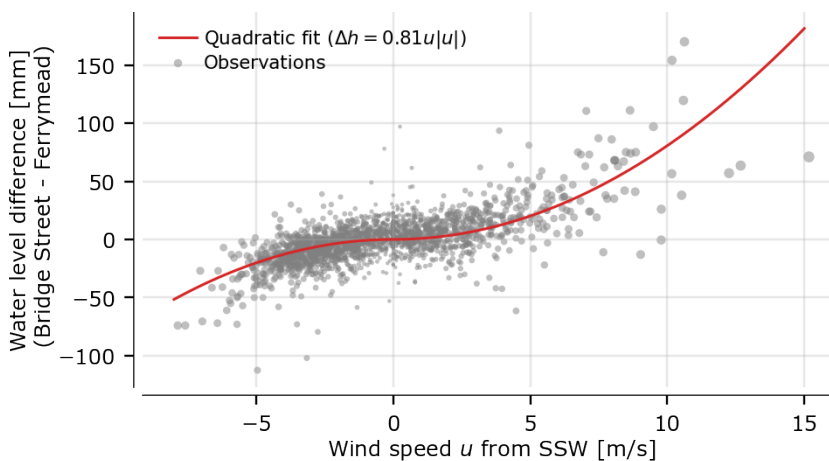


Figure 12 Quadratic fit of the wind speed through the observed water level differences.

The analysis shows that 46% of the variation in the water level difference between Bridge Street and Ferrymead Bridge is explained by the regression model. From the four explanatory parameters in the regression model, the SSW component gives the clearest explanation of the water level difference. A velocity squared relation is assumed between the water level difference and the SSW-wind (Figure 12). The full analysis and all details are elaborated on in Appendix C.

The results show that we would expect an 81 mm water level difference during a 10 m/s wind from the SSW direction. This compares reasonably well with findings from LDRP044 Citywide Modelling, Avon Model Status Report, Section 6.3 which adopts a wind of 10 m/s from the SW direction for design runs producing an expected water level difference of 100 mm.

With water levels already directly measured at Bridge St and Ferrymead, understanding the physical drivers of these water levels (including the importance of wind setup) has a somewhat secondary importance to understanding of flood risks. However, a better understanding of the actual estuary wind effects would help to improve modelling of those effects and enable consideration of the joint probability of extreme wind events to identify statistically appropriate design conditions for wind on the estuary.

5. Extreme value sea level statistics

5.1 Introduction

The extreme value analysis has been performed for the four stations which are also analysed by Goring [Reference 2 in Chapter 10]. The station at Lyttelton has been added to the analysis as this station has the longest data series (which is likely to reduce uncertainties of the analysis). The analysis comprises of the following steps:

1. Replication of and comparison to Goring's results to ensure we have similar base data for the analysis.
2. Selection of the statistical method applied for the extreme value analysis.
3. Extreme value analysis and sensitivity of the results.

The full analysis is shown in Appendix E. This chapter shows the main results.

5.2 Data and comparison to Goring

In the analysis water level measurements with an interval of 15 minutes from NIWA are used for all stations except Ferrymead. At Ferrymead we have data that consists of a combination high/low data, hourly, 15 minute and 5-minute data. The datum used for all sites (except for Lyttelton) was Christchurch Drainage Datum (CDD) which is 9.043 above Lyttelton Vertical Datum (1937). The measured water levels at Lyttelton are converted from LVD37 to CDD. Figure 13 shows the various station have different periods of measurements.

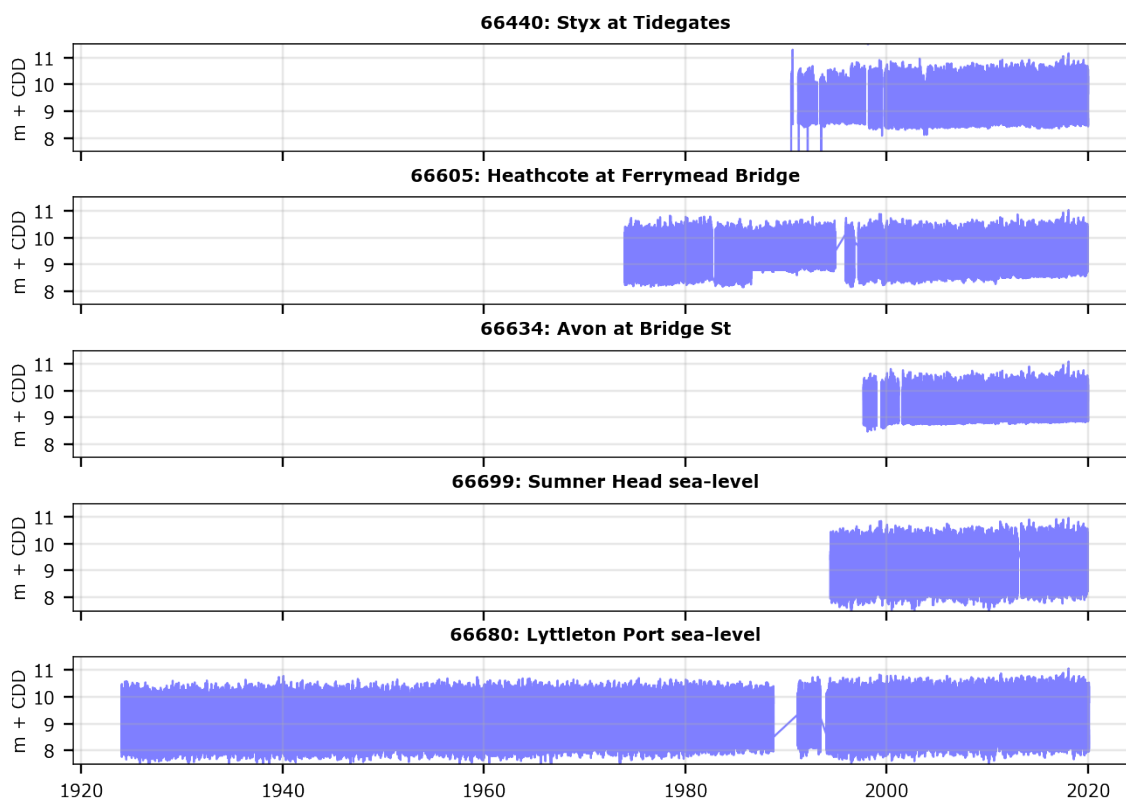


Figure 13 Data availability of all stations

To have a similar starting point of the data the data corrections of Goring are reproduced (to the extent possible based on his report) and the statistics are recomputed, using his statistical method (annual maxima and GEV 1 distribution). The comparison shows the stations at Bridge St and Styx have a nearly exact match.

At Ferrymead our finding is 10-20 mm higher than Goring's. This is likely due to data corrections which are not listed in Goring's report or that he may have obtained his data from a different source. The final statistics at Ferrymead may have the potential to be a little lower if Goring's data was used (the above mentioned 10-20 mm). Figure E.3 in Appendix E shows the comparison.

To investigate this the Ferrymead annual maxima data which Goring used in 2018 was obtained and cross checked with the NIWA data supplied for this study. A large number of minor differences were found, with an average of 10mm and with no apparent meaningful patterns of differences. Further details are given in Appendix section E.3.

These checks confirm that the differences between Goring 2018 and NIWA 2020 maximum's data are minor and are not sufficiently material to provoke further interest or investigation and that confidence in use of the NIWA 2020 data is established.

The Ferrymead estuary water level data is not used in this study other than for the extreme value analysis discussed herein and the wind effect analysis (in section 4).

At Sumner our comparable finding is 50-100mm lower than Goring's. The difference is because FIG waves are not included in our analysis and in the results of Goring we understand that these are included. The full overview of the comparison is shown in Appendix E. The appendix also shows (1) a time series of the 12 highest events (48 hours around the peak water levels), (2) a numerical comparison of the three highest events in this and Goring's study and (3) a sensitivity analysis when the results of Goring are corrected for historic sea level rise.

5.3 Sea level rise adjustment

In all the analyses that follow, measured sea levels have been adjusted to a 2020 basis. This means that early sea level events are adjusted upward by the sea level rise trend difference, making them all comparable with 2020 sea level conditions, thus providing equal weight to the earlier events.

This is a difference from Goring's work which made no such adjustment. In that work, earlier events typically had lower levels and ranked lower than recent events, which did not make the best use of this early data in estimating current risks.

The SLR1 trend from Lyttelton mean sea levels as presented in section 3.2 was selected for the adjustments (during a workshop with Council) due to it being the best substantiated trend (and the least controversial). Sensitivity to the other SLR trend findings was also tested and found to be relatively insensitive.

We chose to use the trend from the mean tide measurements for adjustment of historic sea levels, since the high measurements are more susceptible to measuring errors so the mean tide data is accordingly typically more reliable.

5.4 Statistical methods

The reliability of the analysis increases when longer time series are available. At least, when the boundary conditions such as the climate are more or less stationary or can be corrected for changes. The choice for the statistical approach is therefore based on Lyttelton data as that is the longest time series available. Two choices are required to decide between four common statistical methods:

- 1) The method for data sampling, either annual maxima (AM) or peak over threshold (PoT);
- 2) The type of distribution, a distribution with an exponential tail (EV1 or exponential) or a distribution in which an additional shape factor can be fitted, such that tail events are more likely or less likely compared to the exponential variant (GEV for AM or Pareto for PoT). The EV1 or

exponential distribution are thus special cases of GEV and Pareto, for which the shape factor is 0, resulting in an exponential tail.

The figures below show the results for the various statistical methods at Lyttelton which has the longest time series and is the site least affected by shallow water wind and wave conditions.:

- Left: annual maxima (red dots) with EV1 distribution (red line – similar distribution to Goring) and GEV (blue line)
- Right: Peak of Threshold (green dots) with Exponential distribution (green line) and Pareto (yellow line)

It is noted that the plotting position of the measurements differ in the range of the low return periods (high frequencies), when comparing peak over threshold to annual maxima. The red dots in Figure 14 for low return periods, are lower than the green dots in the right figure. This is because when selecting annual maxima, you might also encounter years with a lower maximum than the threshold value used in a peak over threshold analysis. See Appendix E for the details. The green and red dots therefore should not be compared in the high frequency range.

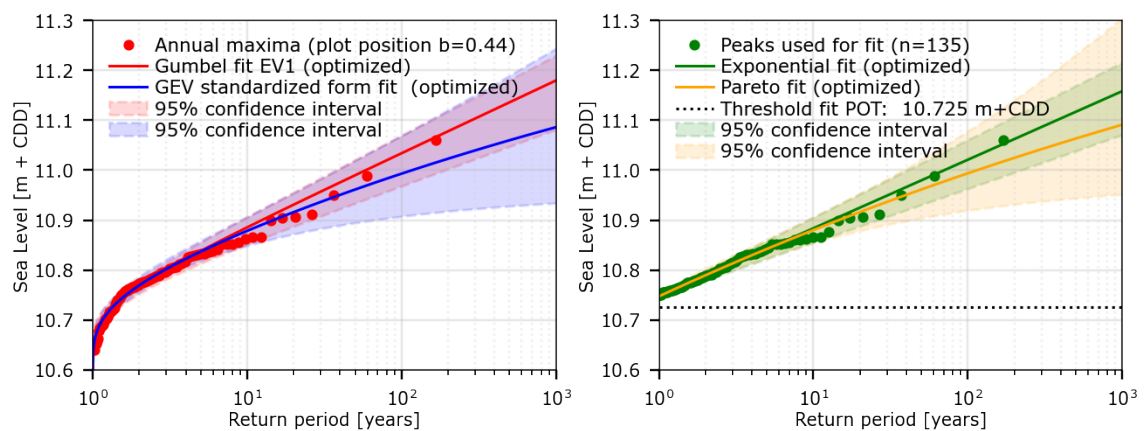


Figure 14 Comparison between annual maxima (AM) and Peak over Threshold (PoT) for Lyttelton.

The PoT approach uses a larger number of data points which results in increased confidence. This method is therefore used to determine the extreme value distribution for the other four locations. To the understanding of the project team, the Peak over Threshold method is a widely accepted method for extreme value analysis (i.e. by the government / hydraulic research institute in the Netherlands, Belgium and abroad [32]) and was initially preferred by Goring in 2018.

The Gumbel distribution (AM) and Exponential distribution (POT) are very similar. They have an exponential tail and implicitly assume no physical upper limit. The Pareto and GEV allow for an upper limit or long tail and would make the most sense if a physical limit can be substantiated. Both the PoT and AM data points show no sign of an upper limit in the Lyttelton measurements (red and grey dots). Also, a physical upper limit is not expected based on the physical processes around Christchurch leading to extreme water levels¹. For those reasons an exponential distribution POT is used (ie: no upper limit). This is again similar to what is done for coastal sites in the Netherlands and Belgium (currently implemented by the Belgium Hydraulic Research Institute).

Note that the EV1 and exponential distribution in Figure 14 are a straight line, which might make it seem that they are the default. However, every distribution can be plotted as a straight line

¹ Up to the extreme sea levels which we are considering. At some point the amount of water that can appear in front of or in the estuary will be limited by physical factors. But for the first meters, we assume that these limitations do not play a significant role.

against its reduced variate. The reduced variate of an exponential function just happens to be a log function, which is also the scale we use for plotting ARI's. There is however another reason which makes the variants with a fixed shape parameter (EV1 and exponential) the most convenience choice: since the shape parameter is fixed, it has one degree of freedom less, which usually leads to a more consistent fit between several datasets or locations.

The choice of the threshold for POT can be done (1) iterative by eye or (2) determine the RMSE of the fit to the data points for various threshold levels and then pick the level with the lowest RMSE. In this study the threshold is based on the minimum RMSE. This process has been designed together with the Belgium Hydraulic Research Institute [32] to make the fitted distributions less sensitive for the (manual) choice of the threshold levels. Sensitivity to threshold levels with a range between $T = 0.25$ years (4 times per year on average) and $T = 1$ year are also analysed.

5.5 Extreme value analysis

The results of the extreme value analysis are shown in Figure 15 to Figure 18. The sensitivity of the results is investigated by sensitivity analyses (Appendix E and summarized in section 5.6). In these analyses the effects of (1) the choice of the threshold level, (2) data sampling method – annual maxima vs. peak over threshold, (3) measurement data and (4) corrections for sea level rise are investigated. 2020 based extreme sea level parameters and sea level results are tabulated in

Table 5

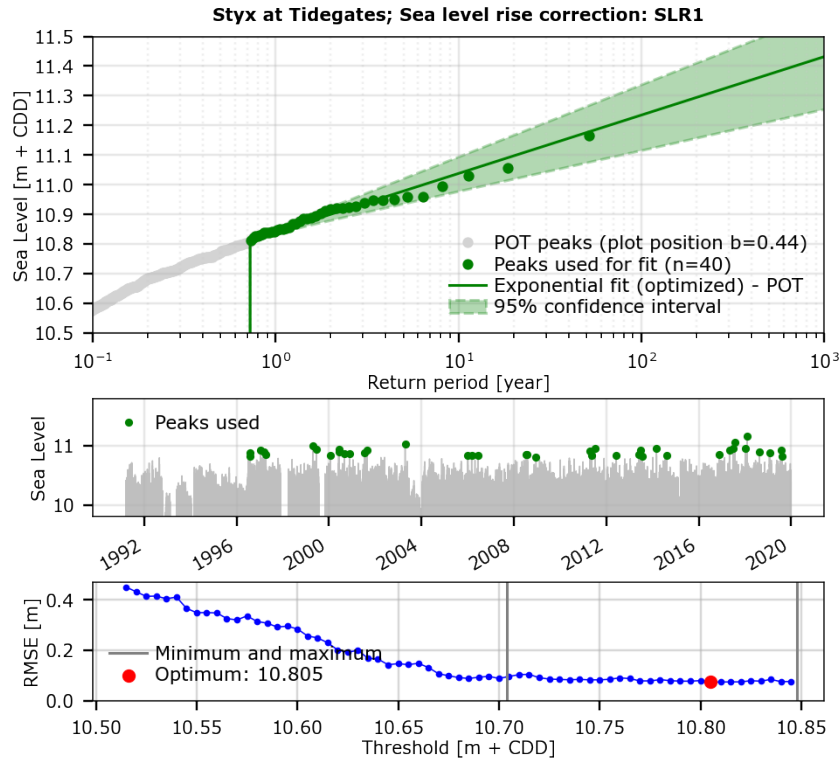


Figure 15 Extreme value distribution for Styx

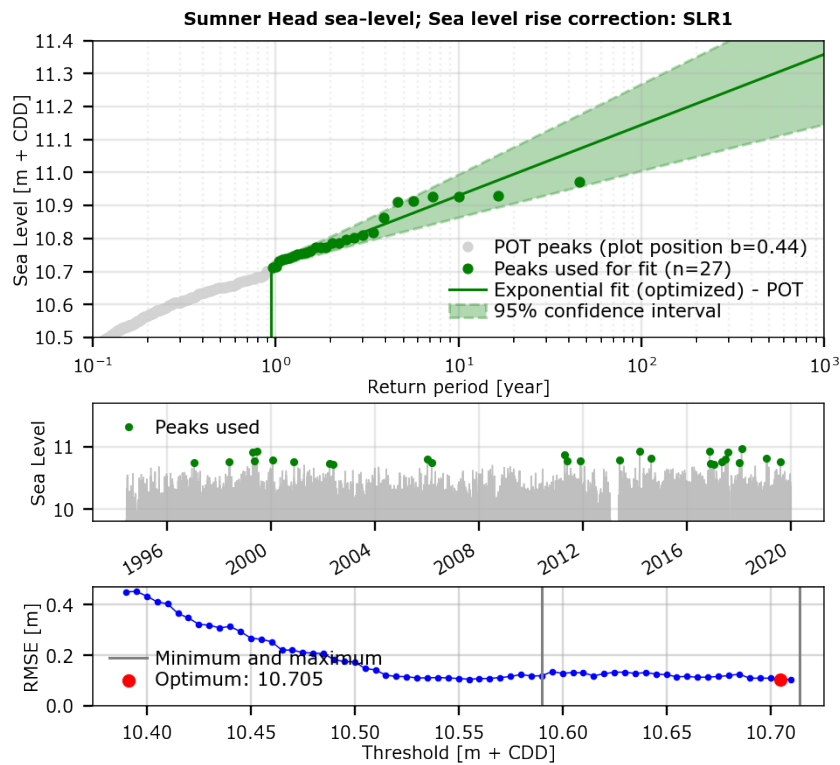


Figure 16 Extreme value distribution for Sumner

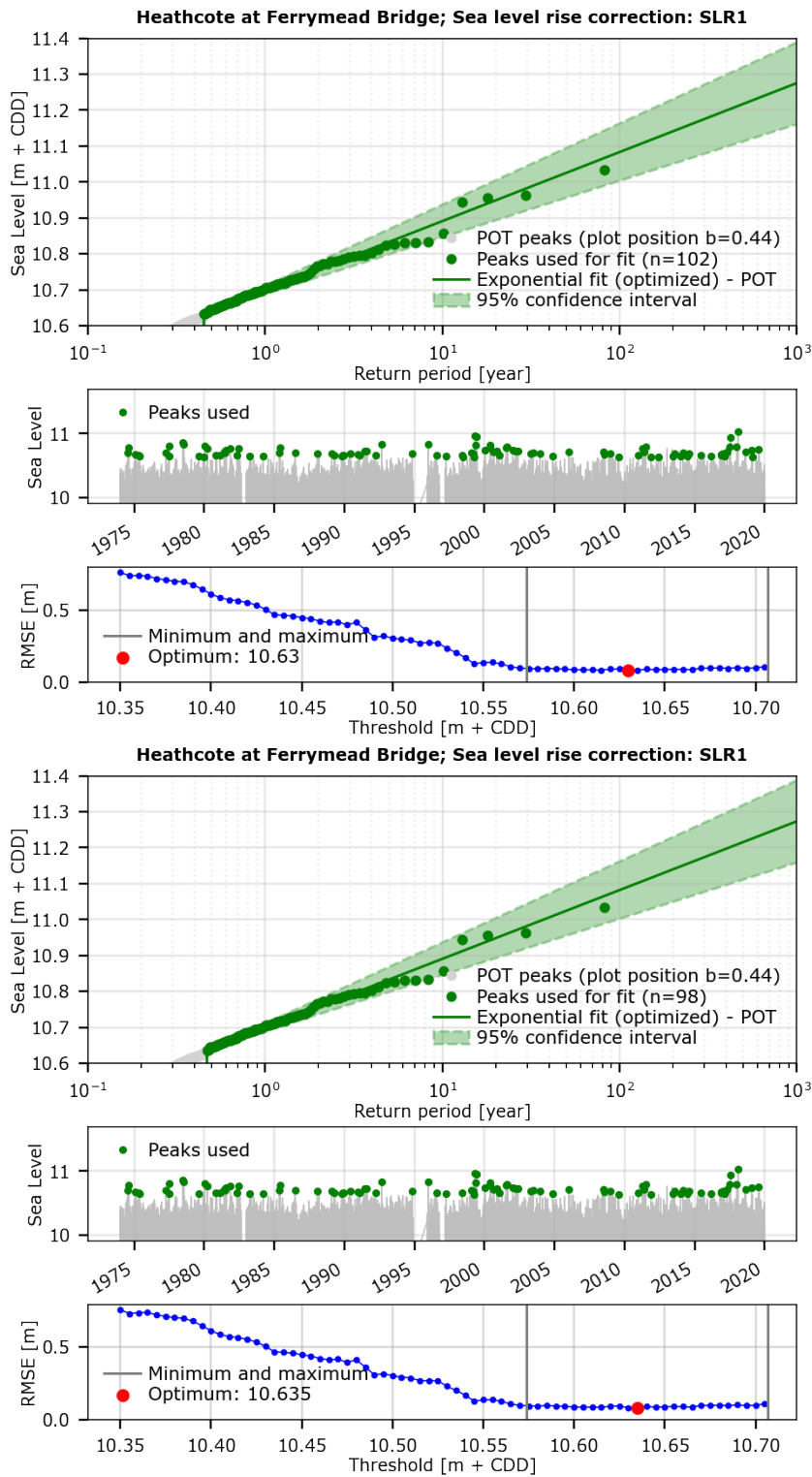


Figure 17 Extreme value distribution for Heathcote

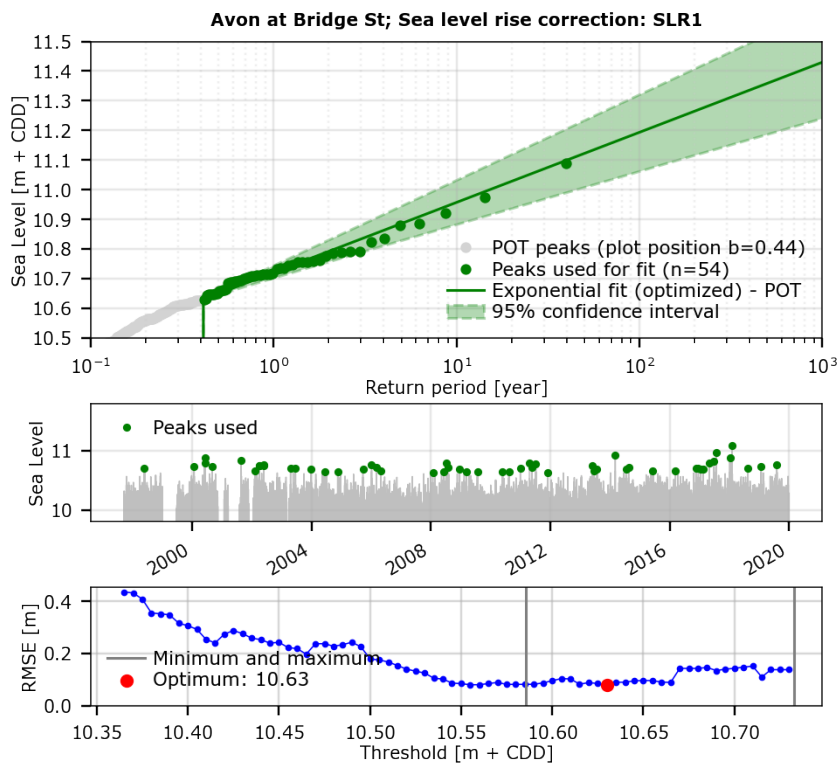


Figure 18 Extreme value distribution for Avon

Table 5 NSEV Parameters and Levels

	Bridge St	Ferrymead	Lyttelton	Styx	Sumner excl FIG	Sumner + FIG
Location parameter	10.721m	10.699m	10.747m	10.841m	10.716m	10.796m
Scale parameter	0.1026	0.0834	0.0585	0.0855	0.0930	0.0930
ARI	Tide level	Tide level	Tide level	Tide level	Tide level	Tide level
1	10.721	10.699	10.747	10.841	10.716	10.796
2	10.792	10.757	10.788	10.900	10.780	10.860
5	10.886	10.833	10.841	10.979	10.866	10.946
10	10.957	10.891	10.882	11.038	10.930	11.010
20	11.028	10.949	10.922	11.097	10.995	11.075
50	11.122	11.025	10.976	11.176	11.080	11.160
100	11.194	11.083	11.016	11.235	11.144	11.224
200	11.265	11.141	11.057	11.294	11.209	11.289
500	11.359	11.217	11.110	11.373	11.294	11.374
1000	11.430	11.275	11.151	11.432	11.358	11.438

Note: A FIG wave allowance of 80 mm is included (adopted from Goring's reporting) with the intention that future work including hydraulic modelling can determine whether to use extreme values with FIG waves, similar to Goring's previous work or without.

5.6 Sensitivity results

The conclusions from the sensitivity analyses are shown in Appendix E. The sensitivity analyses show:

- The sensitivity of the extreme value relationship results to the choice of sea level rise trend are generally low. This is important as it means our conclusions are robust regardless of uncertainty over which sea level rise trend is adopted.
- The sensitivity to the choice for the threshold level is generally low. The difference in the 1/1,000 per year water level is a few centimetres. This is well within the uncertainty range of the analysis.
- For the shorter time series (Sumner, Styx, Bridge St) the extreme value distribution differs 3-10 cm for the 1/1,000 per year event if they are based on annual maxima instead of peak of threshold. When the time series of the measurements is longer (Lyttelton and Ferrymead), the differences are small.
- The measurements at Styx are clearly lower in the first 5 years of the 30 years of available data. Sensitivity checks showed that removing the early data from Styx barely influences the results.

- Due to sea level rise corrections the slope of the extreme value statistics is expected to reduce somewhat. This is evident for all locations except for Ferrymead as the first few years dominate the extreme value statistics there. For Heathcote the extreme value statistics shifts upward with sea level rise correction.

5.7 Conclusions

We used POT in combination with the exponential distribution and an optimised threshold selection for five locations (Lyttelton, Styx, Heathcote - Ferrymead, Avon - Bridge St and Sumner) to derive extreme sea level value statistics. This chapter presented the main results, but Appendix E shows a more elaborate analysis for the interested reader.

The sensitivity analyses show that the resulting fitted distributions are in general not sensitive to choices made (threshold value, sea level rise correction and non-homogeneities in time series) and gives us confidence that the results are stable. The final results can be seen in Figure 19. The details of the distribution are shown in the Appendix.

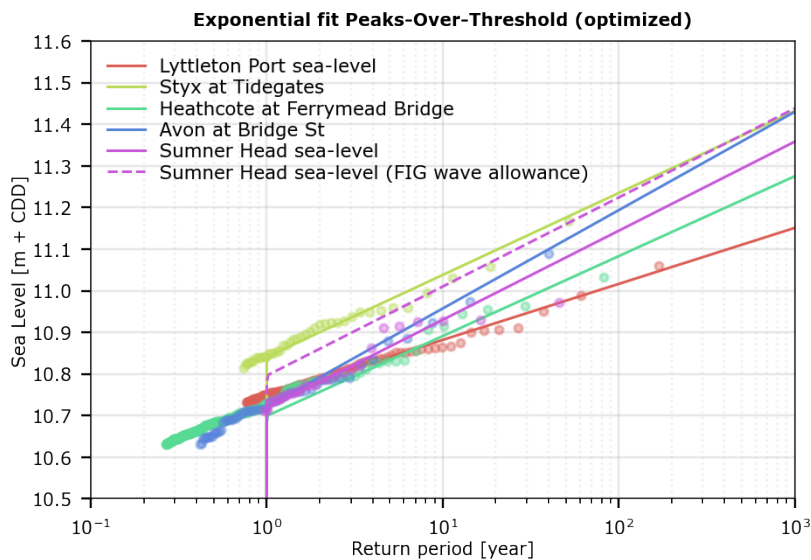


Figure 19 Exponential fit Peaks over Threshold

The figures below (Figure 20) compare the final results with the results of the Extreme Value Analysis by Goring. Note that we are comparing a POT result to an AM result. These result from two different distributions, which both have an exponential tail. Therefore, when comparing the two curves, consider the part with a return period larger than 10 years.

In general, the extreme value distributions derived in this study have slightly flatter slopes compared to Goring, because of the sea level rise correction. The correction of sea level rise also results in increased sea levels for the more common events, but the magnitude differs per station.

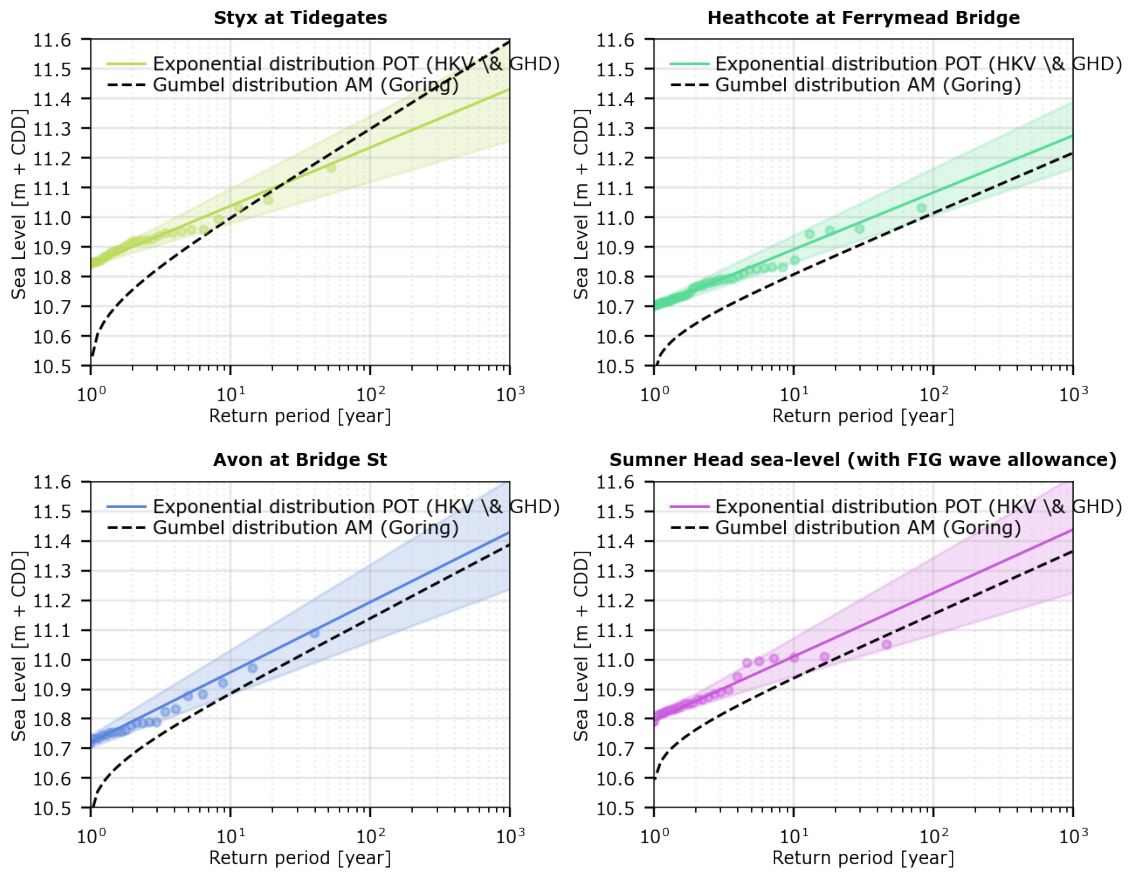


Figure 20 Comparison of Peaks over Threshold Sea Level results with Goring

The difference in steepness between Goring and the new conclusion is largest for Styx at Tidegates. This seems to be caused the early years of Styx data with low peaks. These early year data are used when fitting to the annual maxima but are not selected when using peak over threshold. Appendix section E.9 discusses this further and Figure E.24 shows this difference in event selection between AM and PoT.

6. Correlation analysis

6.1 Summary of analysis

This chapter gives a description of the correlation model that is derived to quantify the relation between rainfall and surge. With surge we mean the difference between the observed sea water level and the predicted tide. This difference is the result of atmospheric conditions, like wind and barometric pressure (Figure 21 shows the difference between surge, water level and tide). For Sumner and Lyttelton the main component in the surge is the barometric pressure. Surge is also an important aspect in extreme water levels. The astronomical tide will never be higher than the sum of all coinciding tidal constituents, where the surge is in principle not limited. Since the surge is assumed to be correlated to the rainfall, it can have a large effect on high water levels on Avon river. In modelling flood risk (in extreme condition) a good description of the boundary conditions is essential. The relation between rainfall and surge is therefore quantified.

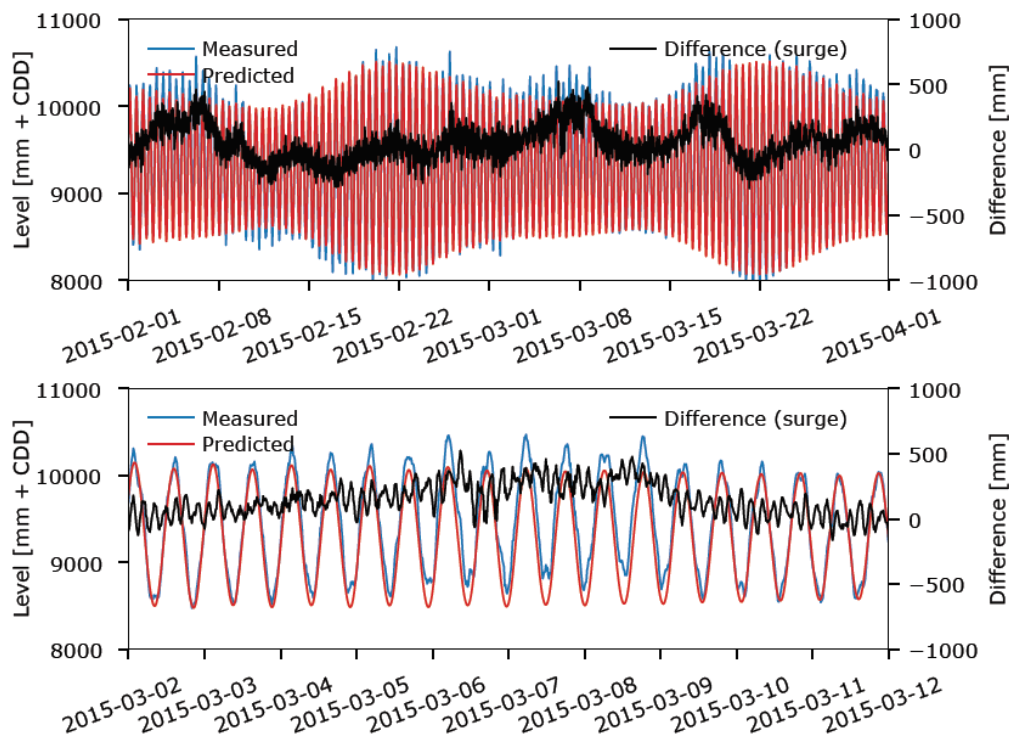


Figure 21 Measured water level (blue) with the predicted tide (red) and the difference, the surge (black). The bottom figure is part of the period shown in the top figure.

The time series shown in the figure show that the surge has significant amplitude of change in shorter frequencies such as four hourly periods. A 24-hour average surge is used in the analysis. This is for two reasons: 1) we are looking for the relation between rainfall and surge on a time scale in the order of a day. This is a characteristic duration of a rain induced hydrograph, and such a period contains at least a full tide cycle. 2) it cancels some of the hour-to-hour fluctuations in the surge data, so the resulting correlation is clearer.

The Kidson weather patterns indicate a relation between type of weather pattern and barometric pressure. This relation turns out to be useful for the correlation analysis, since it gives a physical substantiation for the surge correlation. Figure 22 shows the rainfall depth and surge derived from measurements. The coloured markers indicate the Kidson weather patterns classifications [Reference 1 from Chapter 10]. They have only been obtained from NIWA “Coincident Inclement Weather Study for climate change planning “2018 (Ref 1). These observations cover

selected historical events with large rainfall depth or high-water level, which are the type of events considered in this study.

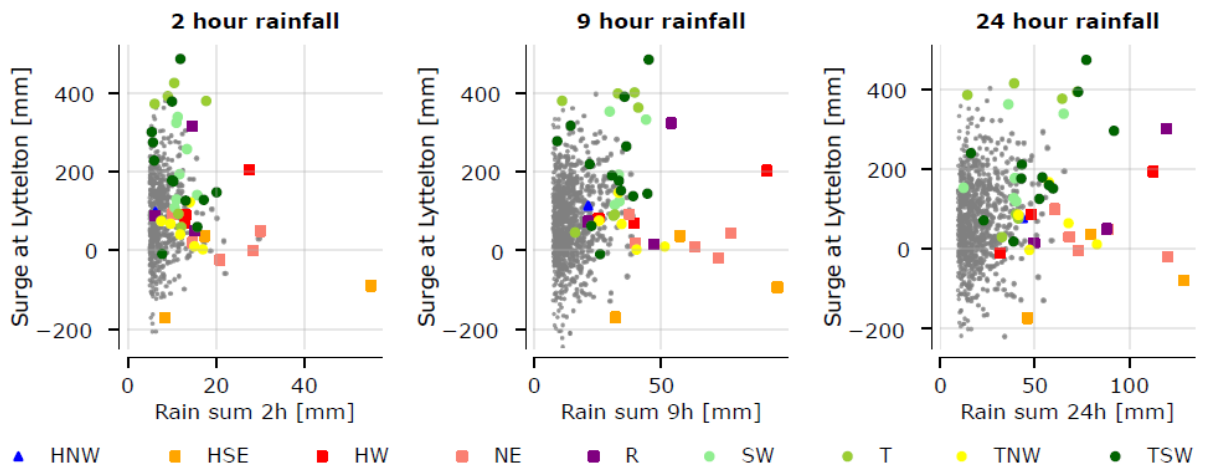


Figure 22 Observations of 24-hour average surge with 2-, 9- or 24-hour rainfall sum. The coloured markers indicate the Kidson weather pattern where this is available.

We understand that additional Kidson pattern data is held by the Met service since 2007 and that daily barometric pressure map records could be used to derive patterns. Having a greater proportion of the event records classified by Kidson patterns would have improved the clarity of findings in Figure 22, however the patterns have only a secondary value in helping to define the method for our correlation model methodology and additional Kidson pattern data would be unlikely to produce any material influence on subsequent methodology or findings.

The round and greenish markers are the trough patterns. They often coincide with high surge. The square and reddish markers represent the blocking patterns, without high surge (on average). The most extreme large rainfall events occur during blocking type weather patterns. Note that the 2, 9 and 24-hour durations are the duration of the period used to calculate the rainfall depth. It is the maximum rainfall sum in any continuous 2, 9 or 24-hour period during the storm.

About 60 years of data are available. The correlation model however needs to be able to estimate the surge for more extreme events. It needs to cover the range for which no measurements are available (i.e. 1/100 per year conditions). In order to determine trends and enable extrapolation a correlation model has been developed based on binary consideration of blocking and trough patterns as described in Appendix F. The resulting fits are shown in Figure 23. The black lines show the percentiles: Given any depth of rainfall (seen on the horizontal axis) there is a 90 per cent chance that the surge will be between the upper and lower black line above it. The main lesson is that the correlation is particularly strong for 24-hour rain events with a return interval of around 10 years (80 mm), 9-hour rain events with a return interval of 5 years (45 mm). For 2-hour rain events the correlations are weak, short intense rain events have only a minor 'effect' on the surge. The physical explanation is in the Kidson weather patterns. A positive surge is mainly caused by trough patterns, and these cause relatively large rain sums over longer durations.

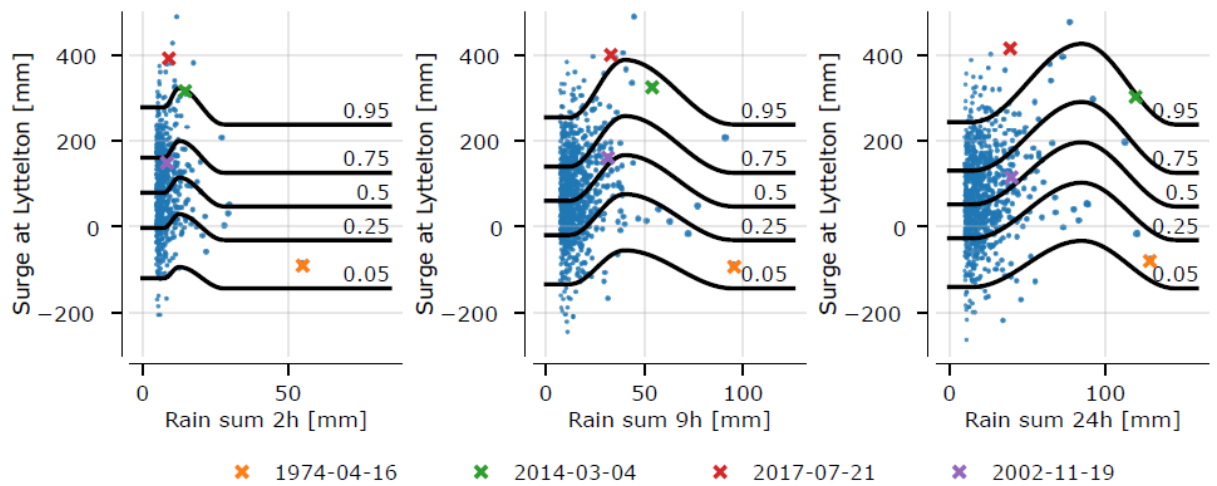


Figure 23 Rainfall observation with the correlation models, for which the 5 percentile lines are shown. The coloured crosses indicate extreme historical events.

6.2 Conclusions

The correlation between rainfall and surge is quantified with observations and classifications of Kidson weather patterns. The resulting correlation model estimates the probability distribution of the surge for any given rainfall event depth, for a selected rainfall duration. A more detailed explanation of the used method and results is presented in Appendix F.

Especially during longer rainfall events caused by the “trough” type Kidson weather patterns the surge is higher. The causal (physical) relation between longer rainfall events and surge is that they are both caused by these kind of barometric pressure fields, which Kidson classified and named. For a rainfall event with a high 24-hour rainfall sum, the correlation is highest for rainfall events with 80 mm of rain: on average 200 mm of surge. For even larger rainfall depths, the expected surge is lower again, to the base level (low surge) for > 130 mm rainfall. For lower intensity rain events (lower than 80 mm in 24 hours) the expected surge is also lower. For shorter duration rainfall events (2 and 9 hours) the correlation is weaker, the expected surge will be 100 or 150 mm for the rainfall depths with the strongest correlation.

7. Joint event ARI analysis

7.1 Summary of analysis

All the statistical analyses done so far can be combined into sets of boundary conditions for the downstream sea water level and the rainfall intensity. This chapter outlines how a combination of events (sea water level and rainfall) can be selected as boundary conditions. Appendix G provides a more detailed description. For some background on a similar approach to assigning probabilities to correlated extreme events, we refer to Chapter 7 from the Hydra-Zoet probabilistic model description [Reference 19 from Chapter 10].

Previous chapters presented the extreme value statistics for the sea water level, rainfall and the correlation model. These can be combined into joint events with certain return intervals. The general steps to establish the ARI for the joint occurrence of say a 10-year rain event with a 1.5-year extreme sea level is:

1. Find the rainfall depth with a 10 ARI for a 24-hour duration. From the HIRDS statistics for Christchurch Botanical Garden, historical data with no areal reduction factors this is 84.1 mm.
2. Get the distribution of the surge for the rainfall depth and duration. The right figure in Figure 23 with the correlation model shows the expected surge is relatively high for this rainfall depth, with a median of 200 mm.
3. Combine the surge distribution with the tide statistics to get a conditional sea level distribution. From this distribution, pick sea level probability that matches the 1.5-year sea level (10.75 m), in this case 1/10th probability.
4. From this we understand that the joint event is expected to occur in 1/10th of the 10-year ARI rain events, resulting in a preliminary joint event ARI of 100 years.
5. However, we know from experience and analyses that a line of equal probability -line for a certain joint ARI does *not* generally give the bounds of the area with an exceedance probability of 1/ARI which we want. This is illustrated by an example in Appendix G and this same example is used to derive suitable ARI scaling factors to determine the final joint event ARI. The scaled final ARI is then 20 years.
6. This process can also be carried out in reverse sequence if starting with the joint event ARI and rainfall then seeking to find the associated sea level.
7. Through repetition lines of equal probability (ARI) are developed as illustrated below.

To further understand this example, it is simple to hand calculate the joint event ARI if the rainfall and sea levels had been uncorrelated, based on the one-day window of coincidence. The rain event is expected to occur on average once every 3,650 days. The sea level event is expected to occur on average once every 549 days. The two events would therefore occur together once every 2,003,850 days which is a 5,490 ARI. Without correlation these two events would almost never occur on the same day. With correlation considered these two events would be expected occur with a 20 ARI. This difference illustrates the high significance of this correlation.

Figure 24 shows the lines of equal probability for 10, 50 and 200 joint event ARIs. Look for example at the red line. On the left end, the 200 ARI boundary condition is a 200 ARI sea water level with no rain. On the right end this is a 200 ARI rainfall event, with a median sea water level. In the middle the lines show a circular shape. If there would be no correlation, between rainfall and surge the coloured lines would all be more or less straight diagonal lines. The

correlation leads to the curved lines; since surge and rain coincide an event with a certain ARI has more extreme rain and tide values than if there had been no correlation.

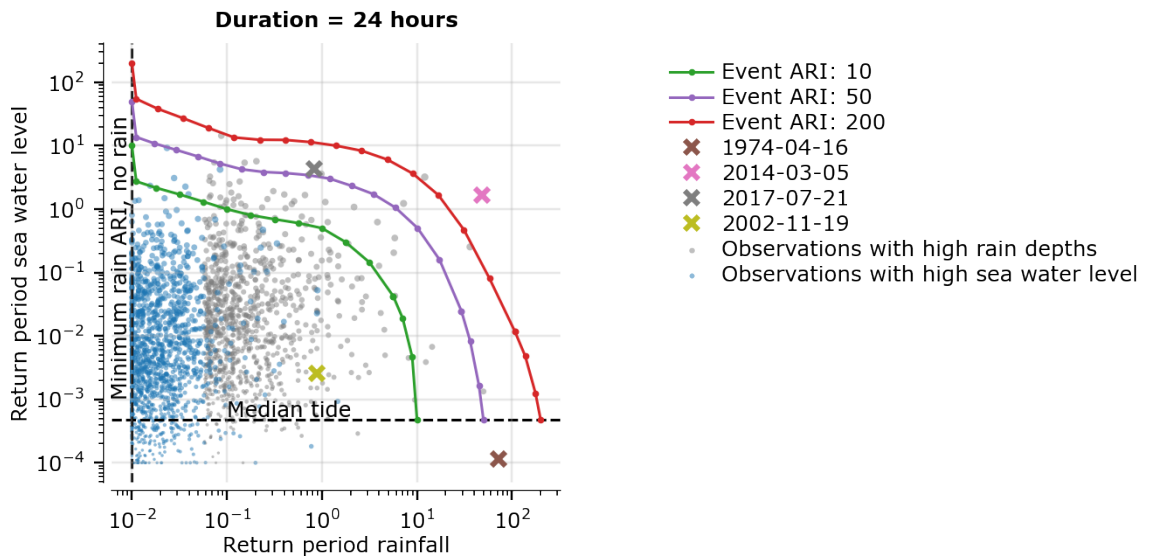


Figure 24 Lines of equal probability (Joint ARI)

The result in Figure 24 shows that there are two observations above the 200 ARI line while there is only 58 years of data. This could be a coincident, but the tens of observations above a 10 ARI line are extremely unlikely to occur within 58 years (the difference is statistically significant). This shows that these preliminary lines of equal probability do not bound an area with an exceedance probability of $1/\text{ARI}$. For our purposes we want to define a contour line that gives a better representation of a specific ARI, such as 50 years or 200 years.

To determine the required position of the line to match the searched exceedance probability of a flood level, the lines of equal probability are scaled using a simplified model with two uncorrelated exponential functions. Figure 25 shows the results, the details are elaborated in Appendix G. There it is shown that a line drawn on the contour of equal probability with $\text{ARI} = 1686$, will be exceeded with the desired 200-year ARI.

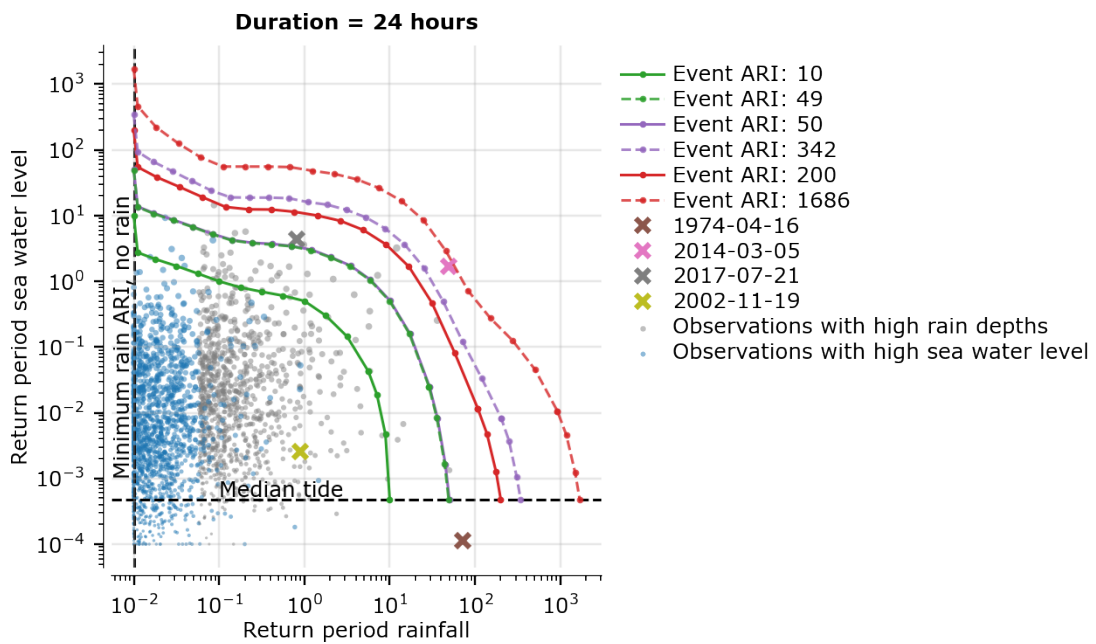


Figure 25 Scaled contour lines defining regions of the graph domain which will be exceeded with 10/50/200 year ARI (based on scaling factors)

The set of three dash line types above are almost the final conclusion with respect to the joint ARI event analysis. Selection of a large number of points from for example the 200 year contour line (1686 year line of equal probability), and running models with these as boundary conditions and then taking the max of the set of results would almost produce a reasonable estimate of the 200 year flood risk.

There is however one more consideration that requires further adjustment. The scaled contour lines in either top left or bottom right extremities now exceed the 200-year ARI on both axis intercepts. These would clearly overestimate flood levels if these were used for model boundary conditions. This is solved by limiting the scaled contour lines on the top and right-hand side to the desired ARI. These final adjustments are illustrated in Figure 26 below. The red dashed line, which is the scaled contour line, should not exceed either of the black dashed lines, since that would lead to a more extreme rain or sea water level than the searched once per 200 year.

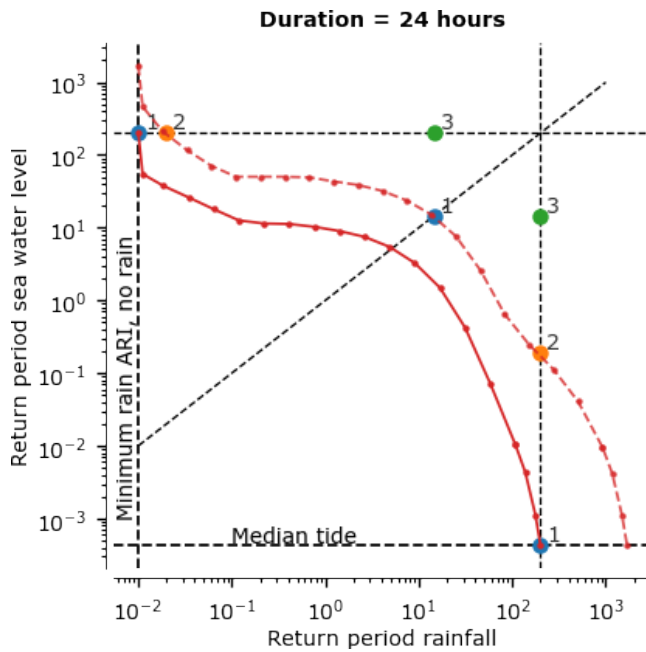


Figure 26 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines - The coloured dots labelled 1, 2, 3 are reference points used for discussion in the following chapter)

7.2 Conclusions

This chapter gives a concise explanation of the method for deriving boundary conditions. A more detailed explanation is given in Appendix G. Extreme boundary conditions are estimated based on the three components: extreme sea levels, extreme rainfall and the correlation between rainfall and surge. Since two variables are classified for the boundary condition (sea water level and rainfall), there is no single combination that is relatable to a 10, 50 or 200 ARI events. In fact, there are an infinite number of combinations that are 10, 50 or 200 ARI probable events. These combinations are depicted by the scaled contour lines in this chapter. The lines show a circular shape. If there has been no correlation between rainfall and surge the lines would all be straight diagonal lines, with a negative 1:1 slope on the log-log plots. The correlation leads to the curved lines; since the surge and rain that coincide in events with a certain ARI will then have more extreme rain and tide values than if there had been no correlation.

8. Model boundary conditions

8.1 Rule options - assessment

Even with the superior understanding of risks presented in the previous chapter, the choice of model boundary conditions remains challenging. The two key themes to be considered are:

1. Risks of conservatism and
2. Quality vs cost

These are each discussed below.

Preamble

In a general type of hydraulic flood model (such as the Avon model), there exists an infinite spectrum of model flood locations with differing characteristics in terms of sensitivity to sea level as opposed to sensitivity to rainfall. Many locations will be 100% sensitive to only one of the two variables (e.g. western Christchurch is only sensitive to rain not sea level).

Given the wide significance of such singular locations it is important that any boundary condition strategy to serve a particular flood level ARI must have at least two model runs, with a rainfall run containing a 'full ARI' rainfall and the tidal run containing a 'full ARI' sea level.

When contemplating a trade-off boundary line such as that shown below in Figure 27, for any given geographic location in the flood study area, there is likely to be one point on the curve for which modelling generates a maximum flood level at that location, with diminishing flood levels produced from any other point on the curve. For example, if a particular location is sensitive to rain and sea level flooding, then the optimal point on the curve is likely close to the 1:1 line (same ARI rainfall and ARI tide level) and then any flood result from the curve, but distant from the optimum, would be expected to produce a lower flood level.

With this realisation it is apparent that no reasonably finite number combination of boundary conditions selected on the line will generate desirable flood results for all possible model locations. If a collection of points on the curve are chosen, then for locations where the optimum point falls into a gap, its flood level may be underestimated for all the collection of points and hence also from the max of results, but the flood level would not either be overestimated at any location in the model.

Risks of Conservatism

In this discussion we use the phrase conservative analysis to describe when a simplified analysis is used, and that analysis produces a flood level that is higher or the same as a benchmark comprehensive analysis. In the opposite sense we use non-conservative to describe when a simplified analysis produces a flood level that is lower or the same as a benchmark comprehensive analysis. Fully conservative and fully non-conservative have the same meanings but apply across the entire domain of flood level results. Mixed conservatism means analyses which will produce some conservative results and some non-conservative results in different locations within the domain of flood level results.

The above preamble has identified the initially most obvious strategy for selection of boundary conditions, namely, to picking 2, 3 or more points 'on the curve'. We would characterise that strategy as fully non-conservative. This is because only in a few fortunate geographic locations where its optimal point happens to be one of the selected points on the curve then the flood result should be just right. But for everywhere else in the model domain the flood levels will contain varying and unquantifiable amounts of underestimation. Examples of this strategy on

Figure 27 below would be either the set of points labelled 1, or the set labelled 2. Non-conservative strategies are typically not favoured in most Engineering applications and hence this strategy is not recommended.

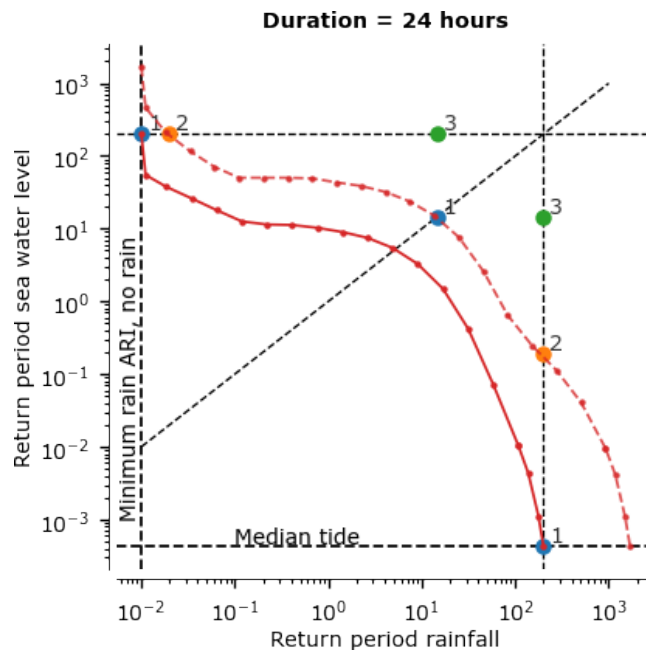


Figure 27 Diagram for 200-year ARI 24 hr - joint event contour lines and scaled contour lines

The next question is then how to pick a strategy that is fully conservative, meaning that for anywhere that the flood level was not just-right, that the flood level would be overstated. In order to clearly meet this requirement, the boundary conditions must be chosen generously so that in combination they are bigger, than any of the infinite number of points on the curve. An example of this in Figure 27 above would be the two points labelled 3, which in this example is a 200/14, 14/200-year ARI combination. If these boundary conditions are used then it is impossible to find any point on the curve which could plausibly produce higher flood levels than the chosen boundary conditions, so conservatism is reasonably assured.

This (our second) strategy is simple, and most comparable to the previous modelling boundary rule strategy and we recommend it primarily for those reasons. However, there are two key weaknesses to acknowledge with this simple conservative method. First there is unlikely to be anywhere interesting in the model that is just right, but rather there will be varying degrees of conservatism everywhere. Consider for example a geographic location with balanced sensitivity such that its optimum boundary condition was 14/14 (point 1 on the 1:1 line). By running 200/14 and 14/200 it is inevitable that both results will produce higher flood levels than that which the optimum 14/14 model result would produce, and they might be significantly higher flood levels. For locations either side of the 1:1 line the conservatism is likely to be even worse. In some locations the estimated flood levels could be overstated by as much as an order of magnitude in ARI terms (i.e. a nominally 200-year flood level might actually report back a 2000-year flood level result). This strategy clearly still leaves much to be desired.

There is a third more general strategy that targets middle ground between the two extremes discussed above. The third strategy will mitigate extremes of both non-conservatism and conservatism, but inevitably will also allow moderate non-conservatism and moderate conservatism (mixed conservatism) in different parts of the results. This strategy is simply to pick boundary condition-points anywhere between (such as midway between) the simpler strategies. In Figure 27 above a two-model run example would be new boundary condition

points each midway between points 2 and 3 along each of the limiting 200-year lines. This approach would likely produce non-conservative results for a balanced location (point 1 on the 1:1 line), but would equally produce conservative results for any locations with optimum points on the curve that fall below and to the left of either of the chosen boundary points. This strategy is variable because there are any number of points between the extremes which could be chosen and a value judgement is required as to the importance of conservatism and non-conservatism to influence and largely arbitrary final choice of point.

We would recommend the general strategy only with caution for particular applications where being conservative was no more desirable than being non-conservative. Pilot study work would also be recommended to test this strategy against higher quality benchmark outcomes in order to form some quantified mapped understanding of conservative and non-conservative level differences. The high-quality benchmark outcome would use a large number of model runs (maybe 9 or 12) so as to closely approximate the trade-off line. These would be summarised into a 'max of runs' overall flood level. This process could then be repeated for simpler 2- or 3-point strategies (ie: 2 or 3 model runs) and the max of runs for each of those strategies calculated. Differences in flood level between the different strategies would provide understanding of the significance of adopting any of the practical simpler strategies.

Quality vs Cost

The second and rather simpler theme is essentially how many model runs to do for each ARI and rainfall duration condition. Regardless of the strategy on conservatism, more runs will enable a better choice of boundary conditions and produce better results.

In the case of the recommended simple strategy the degree of conservatism will be substantially reduced by introduction of a third point. In Figure 27 above, this might result in boundary conditions such as 200/1, 50/50, 1/200. The benefit of having fourth or fifth points is equally obvious but so is the trend of diminishing returns on the extra run costs as the approximations improve.

In the case of the complex strategy both conservatism and non-conservatism can both be reduced by introduction of a third point (or more). In Figure 27 above, this might result in boundary conditions such as 200/1, 20/20, 1/200.

8.2 Rule specifications

Explained by way of continuing the 200-year ARI event example above, any two-point rule will have pairings of 200/X, X/200. For the simple rule X = 14. For the complex rule X might be anywhere around 5. In special circumstances asymmetry may be beneficial resulting in a 200/X, Y/200 type situation.

Any three-point rule, with symmetry, will have pairings of 200/X, Y/Y, X/200. In either simple or complex rule types there is a trade-off between choice of X and Y, with higher values of X enabling lower values of Y or vice versa. Examples are given above. Asymmetry may result in a 200/W, X/Y, Z/200 type of rule.

8.3 Tabulated values

The only type of boundary condition rule where the outcome is fully defined by the curve is the simple two-point rule (symmetric). Definition of this rule only then requires knowledge of the curve intercept with the 1:1 line, in the above example at the 14-year ARI. The following table illustrates these rules.

Table 6 Curve intercept rain/sea event ARI values on the 1:1 line

Joint Event ARI	10 year	50 year	200 year
2 hr rain	1.9	4.7	7.4
9 hr rain	2.4	6.4	12.2
24 hr rain	2.1	7.1	14.5

The following table converts the above individual ARI values into an ARI ratio by dividing by the joint event headline ARI.

Table 7 Curve intercept rain/sea event ARI ratios on the 1:1 line

Joint Event ARI	10 year	50 year	200 year
2 hr rain	1: 5.3	1: 10.6	1: 27
9 hr rain	1: 4.1	1: 7.8	1: 16
24 hr rain	1: 4.9	1: 7.1	1: 14

8.4 Discussion

The incumbent hydraulic modelling approach uses a two-point rule, with a 1/10th ratio, without knowledge as to whether it is fully conservative. The findings in Table 7 now show that the incumbent rule is not fully conservative for 10 year ARI modelling, approximately fully conservative for 50 year ARI modelling, and excessively conservative for 200 year ARI modelling.

Continued use of a two-point rule is justifiable on cost grounds but we would recommend standardising the approach such that all modelling is approximately fully conservative, and certainly none is over conservative (more than fully conservative). A simple conservative two-point rule is well defined herein and continues current practice in terms of paired model runs, which is why we recommend Council adopt it as a minimum modelling approach.

For projects requiring few model results, then a three-point strategy is recommended unless the project team is confident that a two point strategy will deliver sufficient result quality for the particular project purposes, as the marginal costs of a third point will be low relative to the costs of assessing quality differences with a two point approach.

For large scale projects then some pilot sensitivity work is recommended to evaluate performance of at least two and three-point rules (as described in section 8.1, risks of conservatism, pilot study work). The project team would then be able to make an informed decision as to whether the additional costs of a three-point rule is justified by the improved outcomes.

Projects such as floor level setting, where absolute flood level results are the primary objective would benefit most from a three-point rule. Other projects evaluating change and focused on differencing between flood levels results from for example sea level rise will likely benefit least from using a three-point rule.

8.5 Recommendations

All that is required to define the simple conservative two-point approach is to know the individual rain or sea (equal) ARI value on the 1:1 line for the given joint event ARI as provided in Table 6. There is a notable similarity between the ARI values for 9 and 24 hr durations, but a more significant difference to the 2 hr duration values. Accordingly, we recommend that Council adopt a simplification and generalisation of the above results such that any short duration rain events, with duration less than 6 hrs use the above 2 hrs result, and that any long duration rain events (with duration of 6 hours or more) use an average of the 9 hr and 24 hr results, as follows.

Table 8 Recommended rain/sea event ARI values on the 1:1 line for adoption

Joint event ARI	10 year	50 year	100 year	200 year	500 year
Short duration (rain < 6 hrs)	2	5	6	7	10
Long duration (rain >= 6 hrs)	2	7	10	13	20

These values are recommended as the new values for tidal pairing to replace the current 1/10th ratio rule. To estimate pairing values for other joint event ARIs on the 1:1 line, we recommend:

1. For any joint event ARI between 10 and 200 year ARI, taking the log of the Joint Event ARI and log of the rain/sea event ARI, then using linear interpolation between those adjacent data points to determine log(rain/sea event ARI) and from that the estimated rain/sea event ARI (ie. a log-log-linear interpolation).
 - i) This method has been applied to generate the 100 and 500yr data tabulated above.
2. For any joint event ARI above 200-year ARI we recommend extrapolation using the above log-log-linear method based on the 50 year and 200-year data above. An electronic deliverable (xls) is supplied with this report to facilitate this calculation.
3. For any joint event with 5-year ARI and below, we recommend pairing with a 1-year ARI. In the unlikely event that a joint event between 5 and 10-year ARI is required we would recommend using an ARI ratio of 5. For example, an event pairing of 8:1.6 has the recommended ratio of 5.
4. For the purposes of batch generation of rainfall input files for modelling we would also recommend generating the 1-year ARI rainfall as it is likely to be used often in three-point rules.

8.6 Boundary condition conclusions

Even after the joint event ARI curve has been determined, numerous choices of model boundary conditions remain. We have described both simple conservative and complex mixed conservatism approaches and also the quality benefits of using a third model run for each ARI event.

The exact ARI values needed for a simple conservative symmetrical two-point rule have been provided for the three ARIs and three rain event durations in this study, together with recommended methods for other rainfall durations and interpolation and extrapolation to other event ARIs.

Detailed tabulated curve values for the three ARIs and three rain event durations in this study are provided in Appendix G and are supplied to Council in XLS electronic format with this report.

These values can be used to determine other rules, including three or more points, complex mixed conservatism and asymmetric rule types when these are of interest.

We expect for typical future projects that the simple conservative two-point symmetrical rule will remain popular, but that three-point rules will be used for more specialised projects and that sensitivity work will likely lead to adoption of complex mixed conservatism rules on some projects. It is possible that asymmetric rules may be adopted for some specific purposes, but it is unlikely that Council will need to consider four-point rules (or more) in the foreseeable future.

8.7 Time series development

For the purposes of hydraulic modelling a timeseries of sea levels is required at the downstream end of the models. Until recently a timeseries developed by Goring in 2018 was used. This was based on a standard sinusoid (HAT, highest astronomical tide) with a superimposed storm surge represented as a sech^2 curve which is symmetrical and decays to 1% of maximum after 2 days after its peak (this equates mathematically to a characteristic duration of 8 hours). The Goring timeseries were hard coded and provide a limited ARI range of peak sea levels.

A spreadsheet to generate future timeseries has been developed as part of our work reported here. It follows the Goring approach but is generalised with live calculations that enable creation of a time series with any desired sea level. Technical details are recorded within the spreadsheet. "Generate sea water level boundary conditions v2.xlsx".

9. Conclusion and recommendations

9.1 Overall conclusions

Significant progress has been made for Christchurch in better understanding of the coastal environment in respect of risk of inundation from high sea levels and river floods, and in particular simultaneous combinations of those two risks.

Measured data around Christchurch has its fair share of anomalies and difficulties, but in general Christchurch data was sufficient for this analysis and we have certainly benefitted from the foresight of Council in the 1980's and 1990's in establishing an effective network of monitoring sites which now provide valuable intelligence about current and future risks.

This study concludes with recommendations on boundary conditions for hydraulic modelling scenarios. Applying these results will answer the question, "What is the modelled flooding that results from the statistically combined ARI boundary conditions?" but not answer the question, "What is the flood risk?" Clear communication of the limitations in applying the findings will be necessary in subsequent work undertaken by Council.

This study is first to characterise recent historic relative sea level rise in the Christchurch area. While uncertainty remains between 4-7 mm/year, either figure is materially higher than past assessments of historic relative sea level rise rates. It is also the first to demonstrate material acceleration in relative sea level rise when recent years are compared to earlier historic periods. Reports from monitoring of land movement seem to indicate that the differences between relative sea level rise and absolute sea level rise at the benchmark locations in Sumner and Lyttleton is small.

This intelligence has been used to refine and materially improve on understanding of the probabilities of extreme high values for sea level. This work has resulted in much better consistency and comparability of conclusions across the four primary sites. It has generally shown that relatively common (2-10-year ARI) high sea levels are materially higher (circa 100 mm) than previously estimated, and that the relationship slope is less than previously estimated meaning the high sea levels occurring in the 100-1000 year ARI range are similar to previous estimates.

A statistically sophisticated analysis of hundreds of historic rain and high sea level events from 1962 to 2020 has been completed. Significant correlation was confirmed between rainfall depth and sea levels elevated above their normal tidal patterns. This correlation is strongest for the mid-sized (1-10-year ARI) rain events materially increasing the risks of abnormally high sea levels coinciding with such rain events. A synthetic correlation trend type relationship has been established from this data which provides a consistent basis for characterising events within the range of the measured input data as well as a means to extrapolate understanding to more extreme events beyond the range of the measured data.

This analysis has then been further developed to understand the probabilities of coincident occurrence of any combinations of rainfall event and high sea level event with curves developed for 10/50/200-year ARI events based on rainfall and sea level ARI. The March 2014 event is shown to be the most outstanding historic joint event on record with 70-year rainfall (for 24 hr duration) and 2-year high sea level and a predicted joint event ARI of 150-200 years.

The choice of hydraulic model boundary conditions is then reviewed based on this new understanding. Refinements in the current simple modelling approach to boundary conditions are recommended. The changes are pleasingly modest given that the current approach was determined by expert judgement during times when measured data was rare and statistical analyses impracticable. More sophisticated modelling approaches to boundary conditions are

also identified and circumstances in which it may be appropriate to use them characterised. We anticipate a more diverse future in terms of model boundary conditions with choices made dependant on the nature and objectives of the project which motivates the modelling activity.

Further to the above achievements, there is more work to be done and a series of recommendations for further work are listed in section 9.3. Key recommendations from this study for formal adoption by Council authority are presented next in section 9.2.

Finally, the effect of wind on water levels in the estuary has undergone an initial statistical evaluation. This aligns well with preliminary evaluation work done previously by GHD.

9.2 Recommendations for adoption

Adopt the SLR1 trend analysis from the Lyttelton mean sea level analysis as being representative of historic relative sea level conditions in Christchurch for purposes including Extreme Value Sea Level Analyses and any historic modelling or historic flood risk assessments. This is consistent with our evaluation of the various mean SLR trends discussed in 3.2.

Replace WWDG advice on static sea level data with above dynamic historic data and reference to MfE Guidance for future sea levels, (including updates of Table 22-3 and the first table of sea levels in Appendix 1).

Adopt the recommended extreme values for sea level determined herein from the SLR1 historic sea level trend and from the Exponential Peak over Threshold method with minimum RMSE for selection of threshold, as superseding the Goring 2018 values, with both versions of the Sumner Head values (with and without FIG waves). Discussion in section 5.6 highlights that our conclusions have low sensitivity to the choice of SLR trend and are therefore robust regardless of the uncertainty as to which trend might best represent sea level history.

Adopt the simple conservative two-point method (described in sections 8.1 and 8.2) as the minimum approach for specifying rain and tide boundary conditions for hydraulic modelling on the Avon / Estuary catchment. Adopt the pairing ratios recommended in section 8.4 as the values to be used for rainfall / tidal pairings in this simple conservative two-point rule.

Acknowledge that where greater accuracy of results is important that use of the three-point method (and/or complex mixed conservatism methods) should be considered and adopted ahead of the simple conservative two-point method. This should be considered on a case-by-case basis by Council guided by the conditions discussed in section 8.4.

9.3 Potential future refinements

Opportunities for further analyses to further refine outcomes in a future iteration of this work are set out below. These may become motivated either by new funding priorities or when significant additional years of data become available. In general, we consider the current findings provide a suitable foundation to continue with Avon catchment in the Multi hazard project scope and we do not consider that further refinements should be required in order to continue with that work.

Advanced Joint Probability Method: Council has prepared a comprehensive scope of work for advanced JPM considering various ways it could be carried out and to varying levels of detail and thoroughness. It would build on the analyses here and changes the question from finding the best approximate modelling setup (boundary conditions) to what the actual statistically appropriate flood level risks are.

Other geographies: This study has only considered the Avon Estuary catchment. Similar work would be desirable in the Sumner, Heathcote, Styx and Halswell catchments. Differing conclusions are possible in these different geographies.

Wind: With respect to wind on the estuary, three further analyses are suggested.

- Firstly, comparison of additional wind data sources to the one used in this study, including consideration of data from proximate but privately operated wind gauges. From that we would have longer data records and better understand wind speed and direction variability in different locations around the estuary and which gauge(s) provide the best prediction of wind setup between Ferrymead and Bridge St.
- Secondly comparison along the other two sides of the Estuary triangle would be useful to understand wind setup differences from Sumner Head which is the ultimate model boundary condition. By this we mean comparison of Sumner Head with Bridge St as well as comparison of Sumner Head with Ferrymead.
- Thirdly the joint probability of wind speed, direction and setup with the currently assessed variables of rainfall and sea level, would enable recommendations for modelling wind to be developed in a similar method herein and/or in an advanced joint probability study.
- We note that Council have recently commenced collection of higher quality wind data near the estuary from a new NIWA climate station at Bromley. In future once this site has gathered some annual periods of measurements then Council's ability to understand wind on the estuary and its impacts on water levels and flood risks will be significantly enhanced.

Antecedent soil moisture and groundwater conditions: Wind setup and antecedent conditions are commonly felt to be the two most significant drivers of flood risk which have been excluded from this study. There are various measures and indicators of antecedent soil moisture and groundwater conditions. Initial work should explore the available measures and indicators and study which of them should provide the best estimate of antecedent conditions across the catchment and whether correlation between these measures explain variations in river flow at Gloucester St from otherwise similar rain events. Further work would then enable recommendations to select antecedent conditions for future modelling.

Styx level datum: The Styx extreme sea levels data at 1-year ARI is approximately 100 mm higher than the levels from all other sites. This suggests either the datum survey at Styx has been erroneous or imprecise, or that there is a significantly higher risk of extreme water levels at Styx due to its unique geography and greater distance from the open ocean. We would recommend analysis of both mean sea level and peak tide levels comparing Styx with the other sites, with the data filtered to exclude days with high wind speeds. In parallel with this we would recommend an audit of the survey work used to establish and periodically validate the Styx datum with respect to the distant locations.

Council have further advised that “the higher level of the Styx recordings has been noted and examined in the past with benchmarks re-surveyed. No really satisfactory explanation has been found. Council (Graham Harrington) have suggested that “it is due to the slope of the lower Waimakariri river from Brooklands lagoon to the sea however no cross checking of that has been done. If there is in fact a benchmark datum error in that area, then all measurements are on the same datum so infrastructure will not be compromised. It is curious.”

FIG waves: Council will need to decide, and may wish to seek specialist advice, as to whether to continue to use extreme sea levels at Sumner Head with an allowance for FIG (Far Infra-gravity) waves.

It has been the practice of the Council to date to include a FIG wave allowance in the downstream boundary of flood models which use the Sumner Head statistics. With the current Avon / estuary model this will lead to higher estimates of high tide within the Avon Heathcote Estuary and estimates of flooding around the estuary compared to what would result if no allowance was made for FIG waves.

The FIG wave difference from recommendations in this report is 80mm (adopted from Goring 2018). GHD expect this difference would fully propagate through the model water level results around the estuary. GHD understand that FIG waves are unlikely to exist materially within the estuary and that there could therefore be an 80mm overestimation of model flood levels in the estuary if Avon modelling continues to include a FIG wave allowance in its downstream boundary at Sumner Head.

Further consideration of FIG waves is not intended in the scope of this report. Accordingly, the GHD / HKV project team note this as an apparent concern but without further assessment are not technically prepared to make a recommendation to Council on this subject.

For Avon modelling work results around the estuary are not required, then model could be reconfigured and the downstream sea level statistics for the Bridge St used thereby negating the relevance of any Sumner Head FIG wave allowance.

Kidson patterns: A relatively small number of events in the study had Kidson pattern data available. This data was useful to the study and some parts of the analysis could be improved if all of the event dates were classified by Kidson pattern and this analysis updated thereon. GHD consider this does not have urgent priority and we understand Council considers this low priority commensurate with taking no further action.

10. References

The following references are sorted with the most important and relevant references listed first.

1. "Coincident Inclement Weather Study for climate change planning", Graeme Smart for Christchurch City Council, May 2018
(GHD/HKV OneDrive References-NZ)
2. Goring, D.G. 24/7/2018. Extreme Sea Levels at Christchurch Sites: EV1 Analysis. Mulgor Consulting Ltd report for Christchurch City Council.
3. Goring April 2018 "Factors affecting High Water-Levels in Christchurch Estuaries"
4. Robinson, B., Bell, R.G. (2018) Letter Report: "Comparing the NIWA and LINZ/GeoNet Sumner sea-level gauges", Version: 4 -July-2018. Prepared for Christchurch City Council and Environment Canterbury, NIWA, 20 p.
5. "Sumner Sea-Level Station Biennial Report for 2017-2018", NIWA Rob Bell, for Environment Canterbury, January 2020
6. Sea Level Rise in New Zealand: The Effect of Vertical Land Motion on Century-Long Tide Gauge Records in a Tectonically Active Region, Denys et al, January 2020,
<https://doi.org/10.1029/2019JB018055>
Supporting information
<https://agupubs.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1029%2F2019JB018055&file=jgrb539535-sup-0001-2019JB01805-SI.pdf>
7. Update to 2018 of the annual MSL series and trends around New Zealand, Hannah (2019)
<https://www.mfe.govt.nz/publications/marine/update-2018-of-annual-msl-series-and-trends-around-new-zealand>
8. "Regional sea level trends in New Zealand", John Hannah and Robert G. Bell, Jan 2012, Journal of Geophysical Research, Vol. 117
(GHD/HKV OneDrive References-NZ)
9. Hannah, J. (2004), An updated analysis of long-term sea level change in New Zealand, Geophys. Res. Lett., 31, L03307
<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2003GL019166>
(GHD/HKV OneDrive References-NZ)
10. Hannah, J. (1990), Analysis of mean sea level data from New Zealand for the period 1899–1988, J. Geophys. Res., 95(B8), 12,399–12,405
(GHD/HKV OneDrive References-NZ)
<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB095iB08p12399> (abstract only)
11. Hannah, J. (1988), Analysis of mean sea level trends in New Zealand from historical tidal data, Report No. 2, 41 pp., Dept. of Survey and Land Information, Wellington, N. Z.
12. "The effect of sea-level rise on the frequency of extreme sea levels in New Zealand", 2015
<https://www.pce.parliament.nz/media/1382/the-effect-of-sea-level-rise-on-the-frequency-of-extreme-sea-levels-in-new-zealand-niwa-2015.pdf>
13. "Spatial and temporal analysis of extreme storm-tide and skew-surge events around the coastline of New Zealand" Scott A. Stephens, Robert G. Bell, and Ivan D. Haigh, Nat. Hazards Earth Syst. Sci., 20, 783–796, 2020

14. A conditional approach for multivariate extreme values, Heffernan and Tawn 2004, Lancaster University, UK, J. R. Statist. Soc. B 66, Part 3, pp. 497–546
<https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1467-9868.2004.02050.x>
15. Mapping Dependence Between Extreme Rainfall and Storm, WU et al. 2018 (Australia)
<https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1002/2017JC013472>
16. "National Report of New Zealand Prepared for Global Sea Level Observing System (GLOSS) Experts 15 Meeting" July 2017, Rob Bell, Glen Rowe
<https://www.gloss-sealevel.org/sites/gloss/files/publications/documents/national-report-new-zealand-2017.pdf>
17. "Fundamentals of Flood Defences", Netherlands Ministry of Infrastructure and the Environment, Dec 2016"
<http://publicaties.minienm.nl/documenten/fundamentals-of-flood-protection>
18. Dutch sea level monitor (rise calculation)
References provided by Fedor Baart (author) following Guus inquiry
The method description (without wind, but that does not really matter for NZ):
<https://bioone.org/journals/journal-of-coastal-research/volume-28/issue-2/JCOASTRES-D-11-00169.1/The-Effect-of-the-186-Year-Lunar-Nodal-Cycle-on/10.2112/JCOASTRES-D-11-00169.1.short>
And some general considerations for trend forecasting:
<https://doi.org/10.2112/JCOASTRES-11A-00024.1>
This notebook described the Dutch sea level monitor (rise calculation):
<https://nbviewer.ipython.org/github/openearth/sealevel/blob/master/notebooks/dutch-sea-level-monitor.ipynb>
19. "Hydra-Zoet for the freshwater systems in the Netherlands, Probabilistic model for the assessment of dike heights", HKV consultants Dec 2011 (Hydra-zoet model development report and software user manual)
20. "Assessing trends in extreme precipitation events using non-stationary POT analysis", International Journal of Climatology, Sep 2010
<https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.2218>
21. "Development of Practical Guidance for Coincidence of Catchment Flooding and Oceanic Inundation", Duncan McLuckie and Grantley Smith, 2014 Floodplain Management Australia National Conference
<https://www.floodplainconference.com/papers2014/Duncan%20McLuckie%20and%20Grantley%20Smith.pdf>
22. www.mfe.govt.nz/publications/climate-change/coastal-hazards-and-climate-change-guidance-local-government/coastal-hazards-guide-final_2017.pdf
(In this table 10 has quite aggressive estimates for 2005-2020)
23. www.mfe.govt.nz/publications/climate-change/climate-change-projections-new-zealand
(Sep 2018)
24. CCC rainfall depth duration frequency standards
<https://ccc.govt.nz/assets/Documents/Environment/Water/waterways-guide/21.RainfallAndRunoff.pdf>
<https://ccc.govt.nz/assets/Documents/Environment/Water/waterways-guide/Appendices2.pdf>
25. Tom Parsons email "Goring 2018 methodology – bias" 20200305 (job no: 12506449)

26. Graham Harrington raw and smoothed graphical analyses and comparison of four gauges in Christchurch and two gauges in Dunedin, personal discussion presentations in Feb 2020 and email "Southshore stopbanks" 26/3/2020 (project: 12519849)
27. Christchurch City Ground Monitoring — Vertical ground motion at the Sumner tide gauge, Otago University School of Surveying, October 2019
N:\NZ\Christchurch\Projects\51\12519849\01.Document Transfer\Incoming\20200212 Otago land settlement survey data GrahamH to TimP\2019 Annual Report on Christchurch post seismic survey land settlement Otago University School of Surveying Sumner Head.PDF
28. Beavan RJ, Litchfield NJ. 2012. Vertical land movement around the New Zealand coastline: Implications for sea-level rise. Lower Hutt: GNS Science.
www.gns.cri.nz/static/pubs/2012/SR%202012-029.pdf
29. IPO - Interdecadal Pacific Oscillation: (also PDO)
http://archive.stats.govt.nz/browse_for_stats/environment/environmental-reporting-series/environmental-indicators/Home/Atmosphere-and-climate/interdecadal-pacific-oscillations.aspx
<https://www.ncdc.noaa.gov/teleconnections/pdo/> ??
30. ENSO - El Niño–Southern Oscillation Index (ENSO, SOI)
<http://www.bom.gov.au/climate/current/soi2.shtml>
https://niwa.co.nz/climate/information-and-resources/el_nino
<https://www.climate.gov/news-features/understanding-climate/climate-variability-oceanic-ni%C3%B1o-index>
<http://www.tideman.kiwi.nz/SOI/SOI.html>
31. Denys, P et al. (2019), The ups and downs of coastal regions: the implications of vertical land motion on coastal hazards, Paper, 7 pp., Otago University, Dunedin, N. Z.
32. Vuik, V.; Kuijper, B.; Geerse, C.P.M.; Strijker, B.; Gautier, C. (2020). Hydraulic boundary conditions Belgian Coast. (Official citation – in Dutch: Vuik, V.; Kuijper, B.; Geerse, C.P.M.; Strijker, B.; Gautier, C. (2020). Het hydraulische randvoorwaardenboek 2020. Versie 1.0. WL Rapporten, 18_000_1. Waterbouwkundig Laboratorium: Antwerpen.)

11. Disclaimer

This report: has been prepared by GHD and for Christchurch City Council and may only be used and relied on by Christchurch City Council for the purpose agreed between GHD and the Christchurch City Council as set out in section 1.2 of this report.

GHD otherwise disclaims responsibility to any person other than Christchurch City Council arising in connection with this report. GHD also excludes implied warranties and conditions, to the extent legally permissible.

The services undertaken by GHD in connection with preparing this report were limited to those specifically detailed in the report and are subject to the scope limitations set out in the report.

The opinions, conclusions and any recommendations in this report are based on conditions encountered and information reviewed at the date of preparation of the report. GHD has no responsibility or obligation to update this report to account for events or changes occurring subsequent to the date that the report was prepared.

The opinions, conclusions and any recommendations in this report are based on assumptions made by GHD described in this report, particularly those set out in section 1.6. GHD disclaims liability arising from any of the assumptions being incorrect.

GHD has prepared this report on the basis of information provided by Christchurch City Council and others who provided information to GHD (including Government authorities)], which GHD has not independently verified or checked beyond the agreed scope of work. GHD does not accept liability in connection with such unverified information, including errors and omissions in the report which were caused by errors or omissions in that information.

Appendices

Contents

A QA for Lyttelton and Sumner	3
A.1 Data uses	3
A.2 Lyttelton	4
A.3 Sumner	10
A.4 Comparison between data sources	16
A.5 Conclusions	19
B Relative sea level rise at Lyttelton	21
B.1 Sea level rise components	21
B.2 Measurements and other data	22
B.3 Tide	26
B.4 Sea level rise models	31
B.5 Lyttelton trends	33
B.6 Sumner trends	37
B.7 Conclusions	38
B.8 Full results SLR1	39
C Water level differences on estuary	48
C.1 Method and data	48
C.2 Peak tide selection	51
C.3 Differences between peak tides	52
C.4 Explaining the variation	53
C.5 OLS model	54
C.6 Results	55
C.7 Conclusions	60
D Rainfall data	61
D.1 Catchment and rain gauges	61
D.2 Data aggregation	62
D.3 HIRDS tables and formulas	66

E	Extreme value analysis SWL around Christchurch	68
E.1	Structure notebook	68
E.2	Data used	68
E.3	Comparison with Goring’s study	69
E.3.1	Plotting positions	70
E.4	Correction sea level rise scenario’s	77
E.5	Confidence bounds	79
E.6	Peaks-Over-Threshold vs. Annual maxima and fitted distribution (based on Lyttelton)	80
E.7	POT and Exponential distribution	83
E.8	Influence of different threshold values	86
E.9	Comparison with Annual Maxima	89
E.10	Effect of inhomogeneous time series and data quality issues on fitted distributions	92
E.11	Conclusions	94
E.12	Tabular results	96
F	Correlation analysis surge and rain on Avon river	98
F.1	Elements for the correlation model	98
F.1.1	The choice for a 24-hour period	100
F.2	Creating a correlation model	105
F.3	Sea water level with correlation	109
F.4	Conclusion	110
G	Combining statistics to boundary conditions	111
G.1	Introduction	111
G.2	Different statistics within the event	112
G.3	Different events for a required ARI	118
G.4	Picking boundary conditions	123

A QA for Lyttelton and Sumner

This appendix describes the quality assurance (QA) for Lyttelton and Sumner data. It takes the reader through the steps involved in cleaning the data and fixing the vertical references. For the QA several data are available from different sources:

1. Sumner tide gauge measurements from NIWA, from halfway 1994 onward.
2. Sumner tide gauge measurements from LINZ, from halfway 2010 onward.
3. Lyttelton tide gauge measurements from NIWA and LINZ, from 1924 on an hourly base. The data are measured by the Lyttelton port company (LPC), stored and processed by NIWA and for recent years (1995 onwards) stored and processed by LINZ too.
4. Lyttelton monthly averages from John Hannah. These values are derived from the NIWA data sources, but gone through QA by John Hannah to guarantee the quality.

Every of these sources provide some information on the 'true' data, and helps to establish data sources to be used for further analysis.

A.1 Data uses

There are three uses for the data, each requiring a different quality of the data.

1. Sea level rise.

Sea level rise is estimated from the monthly average values. Fluctuations in the data are likely averaged out. Biases should be avoided. Biases could be caused by:

- A (temporary) shift of the tide gauge;
- Large outliers;
- Interpolation of NaN¹ values between two consecutive low or high waters.

2. Extreme value analysis.

For selecting annual maxima, fluctuations in the data can cause a momentary higher water level to be measured, so these should be avoided if possible. Small biases are less relevant than for sea level rise. A 10 mm bias has a significant effect on the estimated rate of sea level rise, while it is of lesser importance for selecting extremes.

3. Tidal constituents analysis.

To calculate storm surge we filter the predicted tide from the observed sea level. For this the pattern should be as clean as possible:

- Long term (> 1 year) deviations from the mean, like sea level rise, should be removed.
- Filling NaN values by interpolation can disturb the tidal analysis, and should be removed.

¹NaN is an abbreviation for "Not a Number", often used as a fill value for missing data.

For different uses, different data sources are suitable: For example for sea level rise John Hannah's data is the best source, since it has already gone through QA. Recent months are however not present, so this should be filled by NIWA or LINZ data. For the tidal and surge analysis we want long (many years) high-frequency data filtered for errors, so the cleaned Lyttelton NIWA source is best.

A.2 Lyttelton

Data sources

Three data-sources are available for Lyttelton:

- The NIWA raw data from the measurements, on an hourly scale for the original measurements, switching to 15 minute data in 1995;
- John Hannah annual and monthly averages. These are derived from the NIWA data, but have an extensive QA administration, making it a more reliable source;
- LINZ hourly measurements from 2010.

To correct the data for errors, the following adjustments are made:

- Remove NaN values by interpolation;
- Visually remove invalid data from the time series;
- Decide what data to use in recent years;
- Correct biases based on John Hannah's log. The relevant points from his log are shown below.

The first two points are only for the tidal analysis.

Remove NaN-values from interpolation

Relevant for: tidal constituents, SLR

The provided data has a 15 minute frequency, also for periods where there was originally no 15 minute data available. We can filter this by identifying where the change in slope is smaller than 1 mm. This indicates that subsequent linear points have a linear pattern. Actually we want to look for a change in slope of 0 mm, but due to the 1 mm resolution of the data the 1 mm step is the minimum. As equation: $\frac{d^2y}{dt^2} < 1\text{mm}$. Note that we might be filtering data which fit the criterion while they are not interpolated. To reduce this filtering we only filter two or more consecutive points. Note that accidental filtering of such point (which might occur more often during high water than low water) can create a bias. This is however not a problem since we'll be only using this data for analysis of tidal constituents and storm surge.

After removing the interpolated values we calculate the percentage available data available per month. This is an indication of the data quality. The figure below shows the results for both Lyttelton high frequent data sources. The availability is similar in the post 1994 period, which is unsurprising given that it comes from the same source.

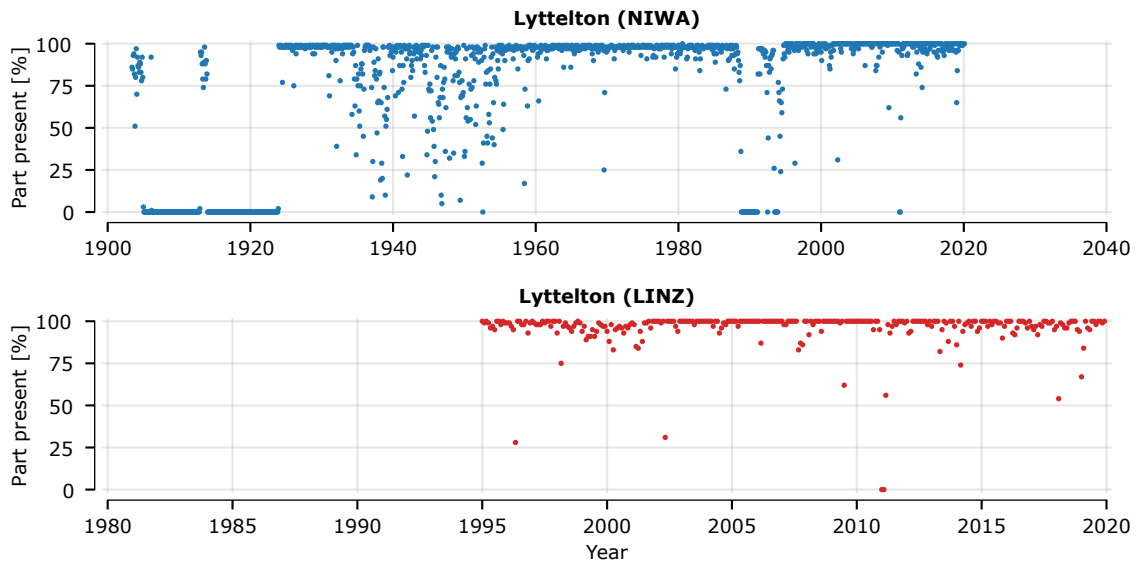


Figure A.1: Part of Lyttelton data that is present per month, for both the NIWA and LINZ source.

Visually remove outliers

Relevant for: tidal constituents

Bad parts in the measurement series can most easily be removed visually. If both Sumner sources and the Lyttelton data are plotted on top of each other, it is mostly obvious which of the measurements are wrong. Cleaning it up is a tedious job but it is the most practical way if fixing the data, so the data are corrected visually.

Some examples of removed sections for the Lyttelton NIWA data are shown below:

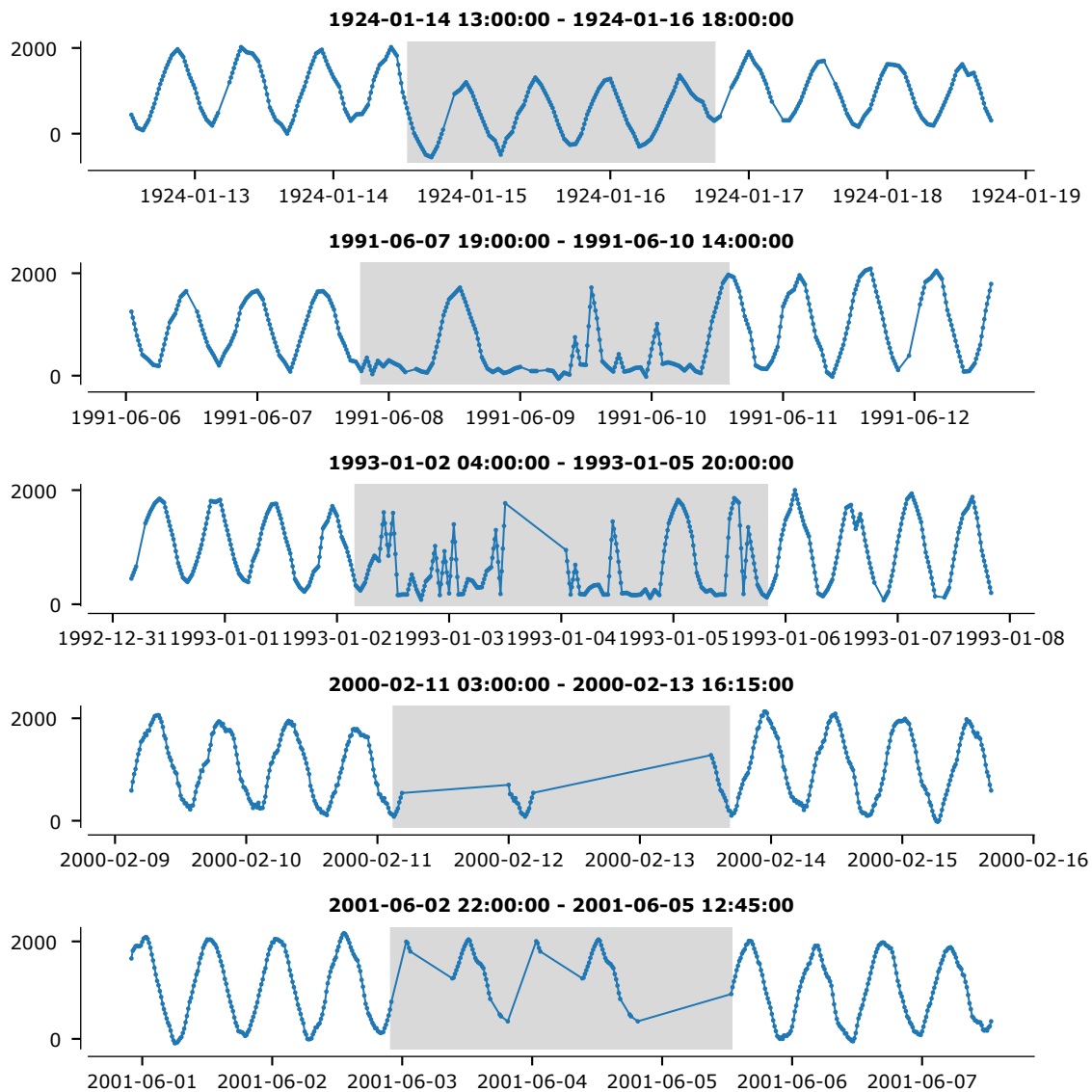


Figure A.2: Examples of bad parts of the data that have been removed manually.

Level correction (bias)

Relevant for: sea level rise, extreme values, tidal constituents

We have two means for shifting the measured data to the right vertical datum. The first is John Hannah's log of adjustments, by following the steps he did to get the monthly averaged data. The second is the LINZ source for Lyttelton data, in which the datum corrections are already processed.

John Hannah log for adjustments:

1. The tide pole was moved in 1934 but care taken to ensure that continuity of datum was maintained. However, leveling in 1970 shows the tide pole as being -0.04 ft lower than in 1940. In the absence of information as to when this change

occurred, the period has been split in two with no correction being applied to the MSL data from 1940 -1955 inclusive, but -0.012 m being applied to the data from 1956 – 1970. Given the evidence for tide pole stability from 1970 – 1980, this correction has continued to be applied to MSL data up to and including 1980.

Note: 0.012 m has been applied to the second half of the period, while a time varying shift might be a more suitable assumption. The difference (12 mm) is however small compared to the magnitude of noise in the data.

2. A new metric tide pole was installed in 1980. Leveling in 1981 showed this pole as being 0.08 ft too low. In 1987, the local Port Company returned the pole to its pre-1981 position. A correction of -0.024 m has thus been applied to all the MSL data from 1981-1987 inclusive. A correction of -0.012 m has continued to be applied to all MSL data thereafter.
3. In 2001, the tide pole was moved and a new tide pole installed. At that stage the datum was lowered by 0.96 ft (0.293 m). Thus, all raw MSL data from 2001 onwards must first be corrected by (-0.293 m) and then again corrected by -0.012 m in order to ensure consistency with the pre-1956 datum.

These adjustments are defined per year, and not per month. The biggest one, the 293 mm, can easily be spotted in the data. In John Hannah’s monthly data (in the yearly data it has already been processed) this shift should be applied to the first month (January 2001). The other corrections, 12 mm and 24 mm (an additional 12) are minor corrections, so the exact month of correction is less relevant.

The shift in the NIWA data is observed in between 2003-7-28 09:00 and 2003-7-29 11:00, two and a half year later than the correction in John Hannah’s data. The 12 and 24 mm are also corrected, from 1956-01-01 00:00 (12 mm) and 1981-01-01 00:00 to 1987-01-01 00:00 (24 mm, an additional 12 mm).

For the LINZ data we find an average difference to Hannah’s source (which is the best reference given the QA) of 305 mm. The expected difference was 293 mm, but an extra 12 mm has to be subtracted. This is the long term 12 mm correction that John Hannah mentioned in the second point above.

The resulting shifts are shown in the figure below. By applying these shifts all data are converted to LVD-37 (Lyttelton vertical datum 1937). To get the values to CDD Christchurch Drainage Datum we add 9.043 meters.

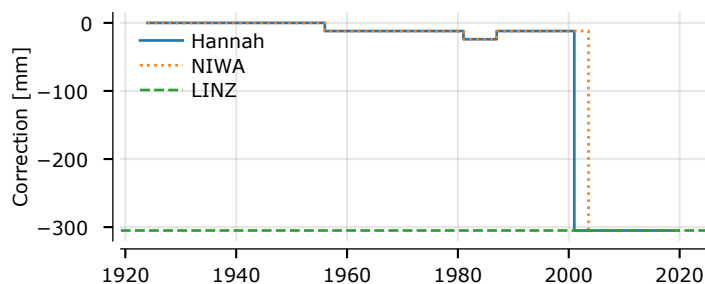


Figure A.3: Vertical corrections on Lyttelton data from NIWA

To check if we did the corrections correctly, we compare these data to the Lyttelton data provided by LINZ. Note that we do not use the LINZ data directly, since we only have them from 1995 and onwards, while the data provided by NIWA goes back to 1923 for high frequency values. If the correction are similar, the result should be an on average 0 mm difference. The frequency of the two sources is also different. LINZ provided hourly values, where NIWA provided 15 minute values for recent years. By plotting them on top of each other we show that the hourly values agree with a measurement every 15 minutes, so they are not average values over the last hour:

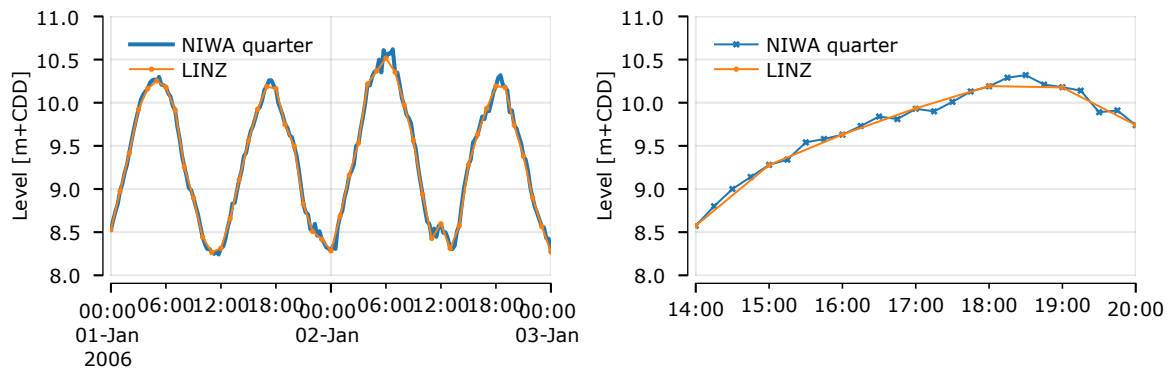


Figure A.4: Time series comparison between Lyttelton NIWA 15 minute data and Lyttelton LINZ hourly data.

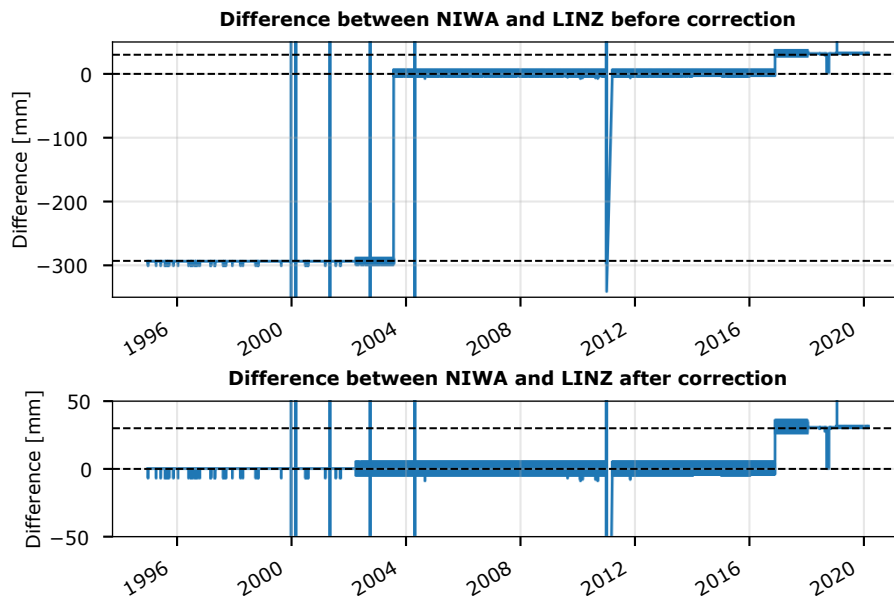


Figure A.5: Difference between vertical NIWA and LINZ data before correction.

The observed difference give confidence in the vertical corrections we applied. Some fluctuations of a few millimeters remain, especially in the period 2002 onwards. Some larger fluctuations are also present. These may be due to some errors in the LINZ data, that seems to be the results of a wrongly interpreted decimal separator, perhaps in the Python processint or in the source data. The resulting outliers are corrected manually. Somewhere end 2016 there seems be a 30 mm

shift, before as well as after the correction. This shift is corrected in the next section:

Data source for recent years

Relevant for: sea level rise, extreme values, tidal constituents

The verified Hannah source has data up to 2019. The sea level rise pattern is sensitive to data at the ends of the data ranges, so we want to add the missing years 2019 and 2020. For this we have two options, the NIWA data and the LINZ data. To determine the best match, we plot all three data series from 2008 to 2020, and try to determine based on the plot which source to use for recent years.

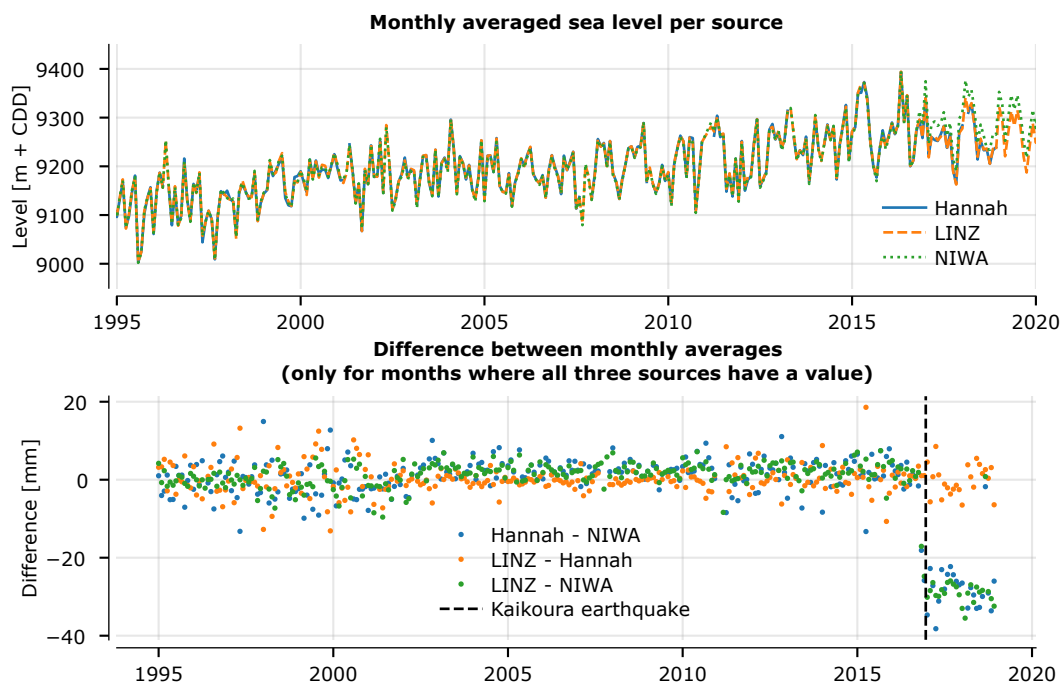


Figure A.6: Hannah, LINZ and NIWA monthly averaged sea levels above, with differences between them below.

Based on the above comparison, LINZ and Hannah seem to give the best match. It looks like a vertical shift of at least 30 mm of the gauge (perhaps due to Kaikoura earthquake) was not corrected in NIWA's data. This has quite a big effect on sea level rise estimates. We use the LINZ data for the 2019 and 2020 months.

A difference of 30 mm per month is corrected in added to the corrections in the NIWA series. 30 mm is the value mentioned by John Hannah:

The daily raw data (as recorded by the tide gauge) was been corrected, in stages, between 4 Sept. 2010 and 7 Dec. 2012 to account for the regional tectonic movement caused by the various Christchurch earthquakes. On 8 Dec. 2012, the accumulated uplift of 111 mm was applied directly by the Port Company to the recorded data. At that time the adjustment made by LINZ was discontinued. However, the November

2016 Kaikoura earthquake produced another required adjustment of 30 mm. Since 00.00 hours on 14 Nov. 2016, 30 mm has been deducted by LINZ from the recorded data. The daily data files are thus already free of these datum inconsistencies, as are the monthly and annual means.

In contrary to the other level corrections, the 30 mm was already corrected in Hannah's data, while it was not the case for the NIWA data. This is what caused the mismatch. The additional correction will be added to NIWA's data. So it will not affect the analysis of tidal constituents.

Over the periode 2003 to Kaikoura earthquake, the NIWA data on average seems to be a bit higher, around 3 mm higher. However if we take the exact same hourly values for both data sources the difference is gone. This indicates that hourly values give on average slightly higher values than 15 minute values. Perhaps the low tides are more peaked than the high tides.

Saving data

The data is saved again in two separate files:

- The monthly mean sea levels;
- The cleaned NIWA record, which can be used for analysis of tidal constituents, after being detrended for sea level rise.

The Lyttelton data measured by LPC and provided by NIWA will be used throughout this study. The data are verified and compared to the data provided by LINZ, with which we can guarantee the quality of the data for 1995 and onwards. By using the NIWA data, we have data that go back to 1923 with high frequency (hourly) values.

A.3 Sumner

For Sumner we run the same analyses as for Lyttelton to determine to suitability of the data for sea level rise, tidal constituent and extreme value analysis.

Data sources

Three data-sources are available for Sumner:

- The NIWA raw data from the measurements, with a 1 minute and 15 minute frequency;
- LINZ 1 minute frequency data;
- The NIWA Lyttelton data, corrected as described in the previous section.

To correct the data, the following adjustments are made:

- Remove interpolated values, replace with NaN;
- Visually remove bad areas, for the NIWA as well as the LINZ data;
- Shift the LINZ data in time.

Since we have three sources with more or less the same observations), we first try to determine which source to use for what purpose:

- Extreme value analysis;
- Analysis of tidal constituents;
- Sea level rise.

For this purpose we plot several peculiar time ranges for which the data are available.

Remove interpolated values

Similar to Lyttelton data, we remove parts of the data where the second derivative of the level in time is zero.

Vertical reference

To check the vertical reference of the data, we shift the data to LVD-37, so we can compare it with the results from Bell, Robinson 2018. To do so:

- we shift the Sumner NIWA data 9.043 meters down. This distance is also the difference between the data sources with data id 966699 (LVD-37) and 66699 (CDD).
- we shift the Sumner LINZ data 1.745 meters. There is no clear reference for this shift since the port company doesn't vertically calibrate the gauge, but it seems to be the mean MSL difference.
- we shift the Lyttelton data 1.241 - 0.293 meters. Bell, Robinsin 2018 names 1.236 as the pre-2012 correction and 1.241 as the post 2012 correction. Since it is unclear where the 5 mm comes from, we don't use the 1.236 shift but apply 1.241 to the whole range. The 293 is to compensate the early 2000's shift.

LVD-37 is used for consistency with Bell, Robinson 2018. After this check we will convert the data back to CDD (+9.043 m) for consistency with the other data sources.

Based on these shifts we reproduce the figures from Bell, Robinson 2018:

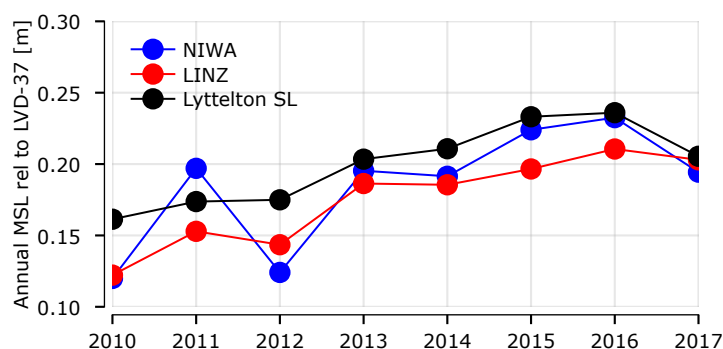


Figure A.7: Similar figure as made by Bell and Robinsin, 2018. Shows the annual mean sea level for the three data sources at Sumner.

Year	Sumner (NIWA)	Sumner (LINZ)	Lyttelton (NIWA)	Difference at Sumner (NIWA - LINZ)
2010	0.12	0.12	0.16	-0.00
2011	0.20	0.15	0.17	0.04
2012	0.12	0.14	0.17	-0.02
2013	0.20	0.19	0.20	0.01
2014	0.19	0.19	0.21	0.01
2015	0.22	0.20	0.23	0.03
2016	0.23	0.21	0.24	0.02
2017	0.19	0.20	0.21	-0.01

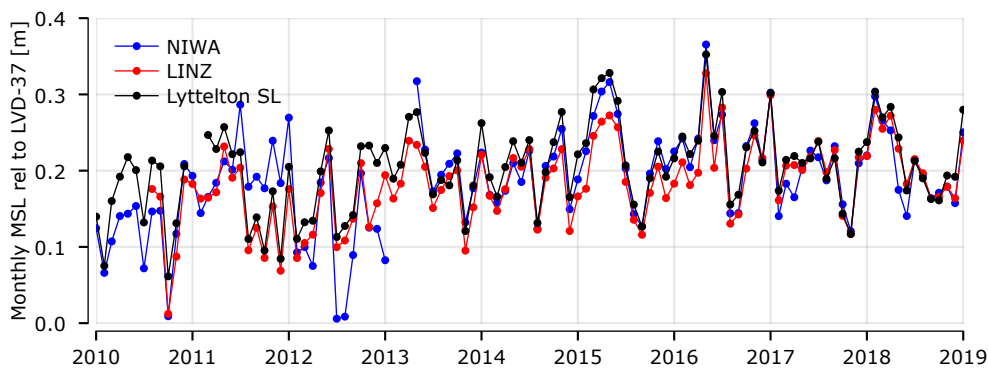


Figure A.8: Similar figure as made by Bell and Robinsin, 2018. Shows the monthly mean sea level for the three data sources at Sumner.

Now we shift the data back to CDD, the vertical reference in which we put the final data.

Time shift of LINZ data

The first thing we note is that the LINZ data seems to be shifted back by 12 hours. This is clearly visible for example during the 2016 Kaikoura earthquake on the 14th of November, an event which naturally causes fluctuations in the measurements. The earthquake happened a few minutes after midnight, so we expect the oscillations to be just after the start of the 14th.

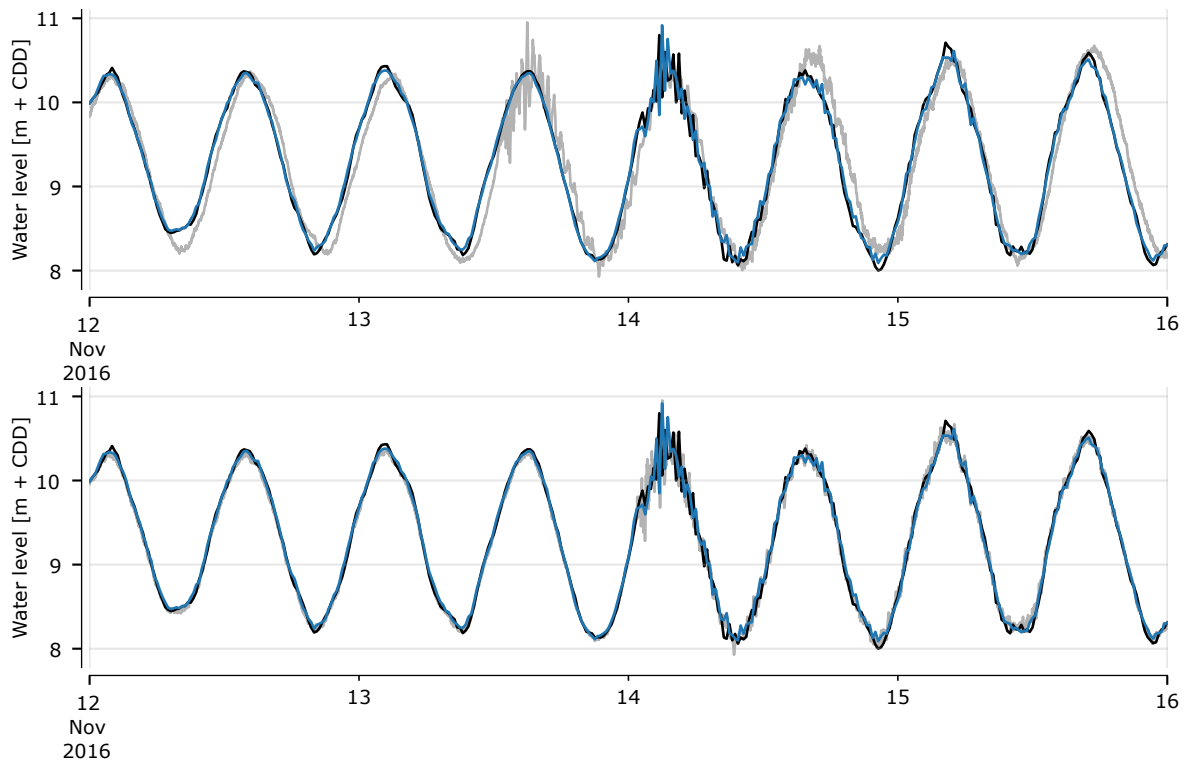


Figure A.9: Part of the time series where the 12 difference for the LINZ data is clearly visible. Below, after correcting.

After shifting the data 12 hours, they seem to match much better. If we look at the figures in Bell, Robinson 2018, the differences in the figures seem to make much more sense if we take this 12 hour shift into account.

Observed differences Now we plot a few time periods to show some observed differences or errors.

During some period, the Sumner (NIWA) data seems to be consistently lower than the Sumner (LINZ) and Lyttelton (NIWA) data. This can take several days or week, causing lower mean sea levels on a monthly and maybe annual scale:

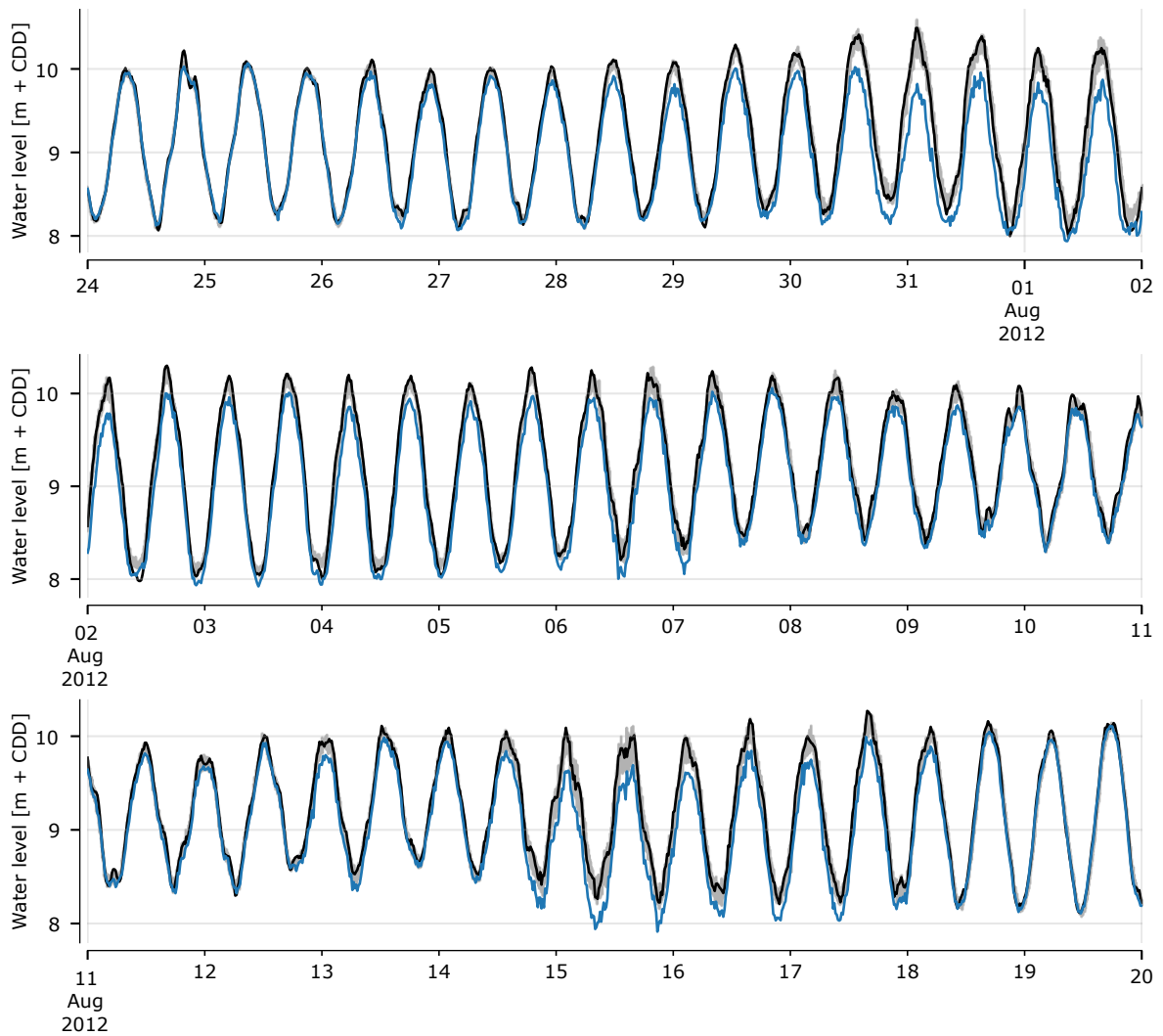


Figure A.10: Part of the time series where Sumner NIWA is consistently lower than Sumner LINZ and Lyttelto

During some period, all three data sources seem to give bad results. This actually gives confidence in the measurements, since it is unlikely that all three have erroneous measurement. It probably a stormy period of a period with large waves.

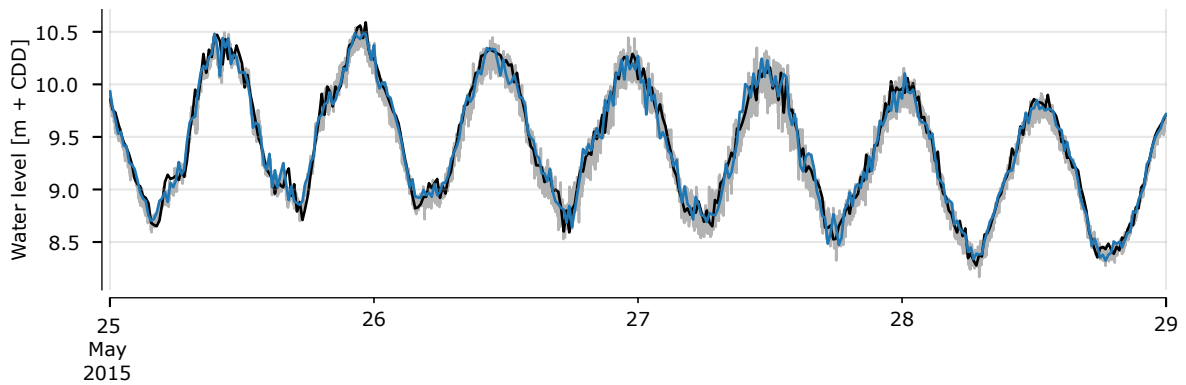


Figure A.11: Seemingly bad part of the data, which actually looks good if compared to the other sources.

During some periods the LINZ data is shifted in time for a few hours or days:

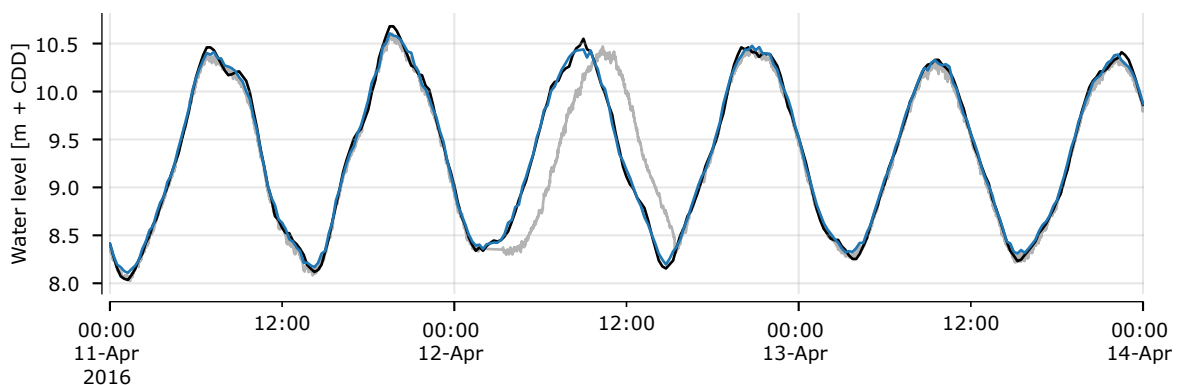


Figure A.12: LINZ time shifted data for a short period of time.

The parts of the data in which the data (LINZ or NIWA) seems bad, are corrected visually in a way similar to Lyttelton. This produces two data series where bad data are removed.

Relation between 1 minute and 15 minute data

For LINZ the provided data frequency is one minute. We choose to average these values to 15 minute values, which we can subsequently use as complementary data source for missing Lyttelton or bad Sumner data. The figure below shows the effect of using 15 minute values instead of the once per minute values: averaging does not undesirably remove peak values. In the displayed period the NIWA data is a bit higher than the LINZ data, this is the result of it being measured by a different gauge. The two series vary from time to time, sometimes one is higher, sometimes the other.

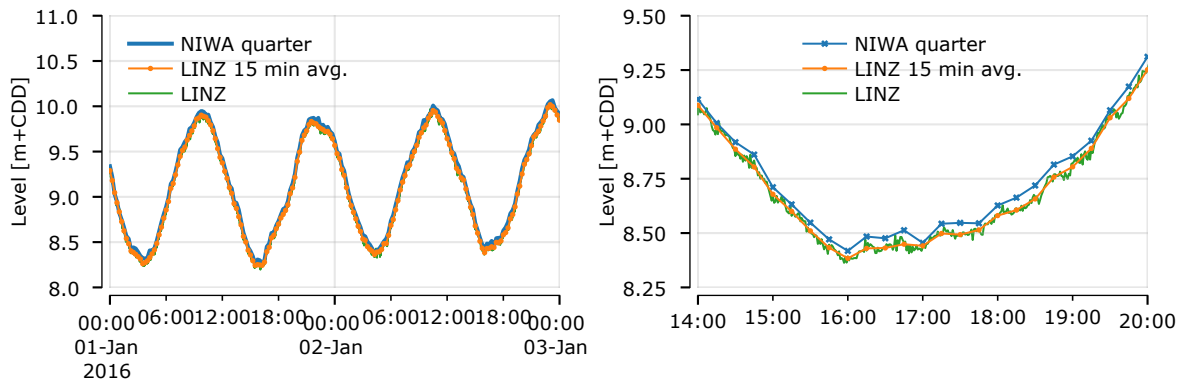


Figure A.13: Time series comparison between Sumner NIWA 15 minute data and Sumner LINZ data.

A.4 Comparison between data sources

After removing the bad data we can compare the three data sources:

1. Sumner (NIWA) data
2. Sumner (LINZ) data
3. Lyttelton (NIWA) data

By plotting the overlapping data in scatterplots, we can observe the effect of these differences on the complete dataset. The first plot shows the comparison between the three sources for all overlapping values. This means that the pre 1994 data is not present in the comparison, since we only have Lyttelton data for the period.

Overall comparison of values

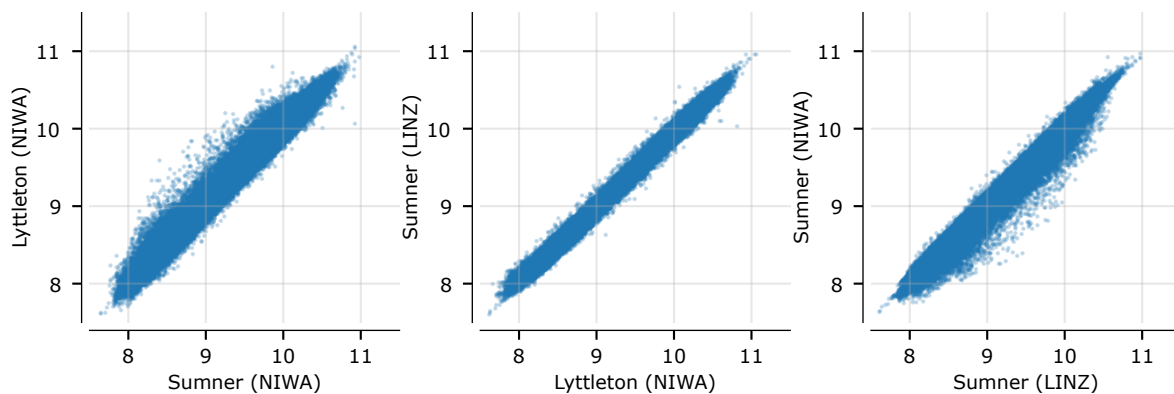


Figure A.14: Momentary sea water levels for Lyttelton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.

- In the first plot Sumner (NIWA) is compared to Lyttelton (NIWA). We observe that Sumner often has lower values than Lyttelton. These are the periods in which Sumner has consistently lower values.

- In the second plot Sumner (LINZ) is compared to Lyttelton (NIWA). The deviations are small, Lyttelton seems to be a bit higher on average.
- In the third plot Sumner (LINZ) is compared to Sumner (NIWA). Here we see the same effect of the lower values for Sumner (NIWA) as in the first plot.

Lyttelton (NIWA) and Sumner (LINZ) seem to give the best agreement in data if we look at all sea level observations. We do the same comparison for the high tide values:

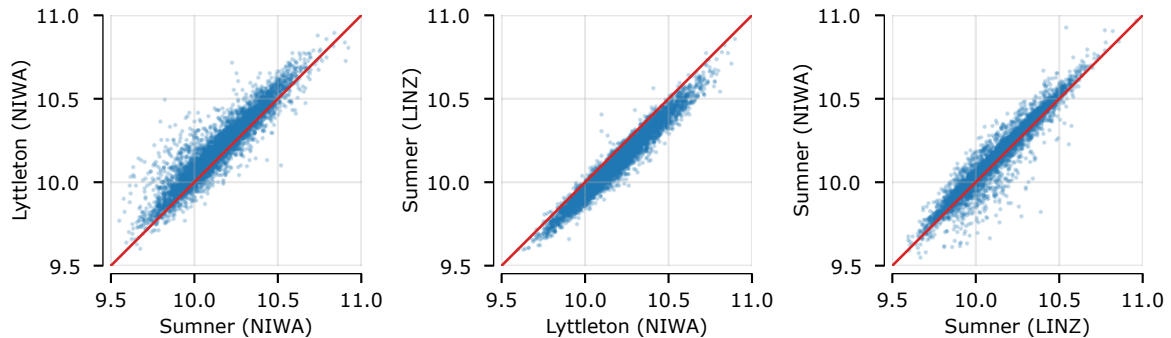


Figure A.15: High tide levels for Lyttelton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.

The results are very similar to the first comparison:

- A good agreement between Lyttelton and Sumner (NIWA), however Lyttelton tends to be a bit higher.
- Some noise due to lower periods for Sumner (NIWA). This effect is present as well in the full tide pattern as the high waters.

A last comparison is between the surge for the three locations. For this we analyze the tide from the data that was cleaned in this notebook, and subtract it from the measured sea level. The tidal constituents from the fitting process are compared to Goring's work, as a verification of the fitting process. The amplitudes and phases of the most important tidal constituents are shown in the following table. The last two columns are from Goring, and can be compared to the first two columns (Sumner). The fitting results are very similar. The most important components (largest amplitude) have very similar values for amplitude and phase.

Table A.1: Tidal constituents analysed from the measured sea water levels. The right two columns are adapted from Goring (2018).

	Fitted		Fitted		Fitted		Goring	
	Sumner (NIWA)		Lyttleton		Sumner (LINZ)		Sumner	
	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase
Mm	4	106	3	133	7	98	4	151
MSF	3	146	3	126	8	163	4	148
Q1	8	46	8	43	8	39	8	47
O1	26	53	27	51	27	53	26	52
P1	15	98	15	98	14	93	16	92
K1	46	95	46	93	46	93	43	97
N2	192	90	197	88	191	90	192	90
M2	834	125	856	124	831	125	834	125
S2	55	151	53	150	56	150	55	152
K2	18	122	18	119	19	120	18	121
MN4	2	80	2	123	2	125	-	-
M4	5	73	3	99	2	79	6	72
MS4	4	132	3	135	4	137	-	-
2MN6	-	-	-	-	-	-	-	-
M6	6	84	6	95	5	89	6	83

When the fitted tide is subtracted from the measured sea water levels, the remaining part is the momentary surge. These surge values are compared between the different data sources:

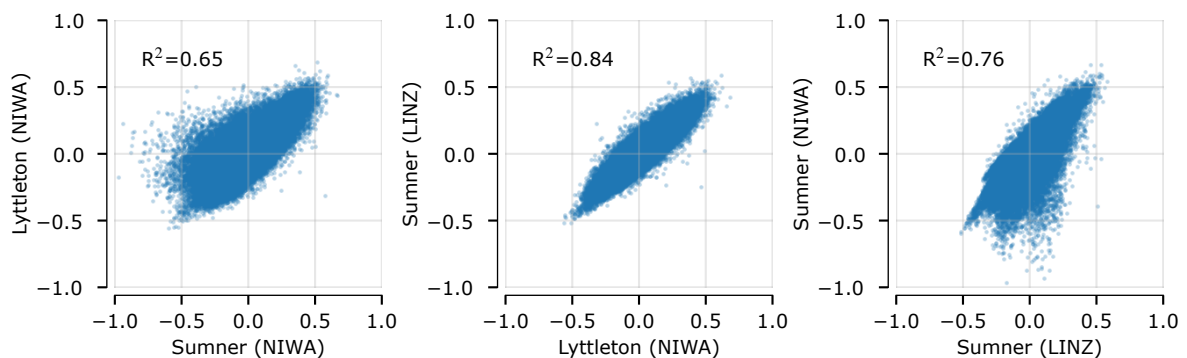


Figure A.16: Momentary surge differences for Lyttleton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.

In the statistical model we will be using 24 hour average values of the surge, which we also plot with scatters:

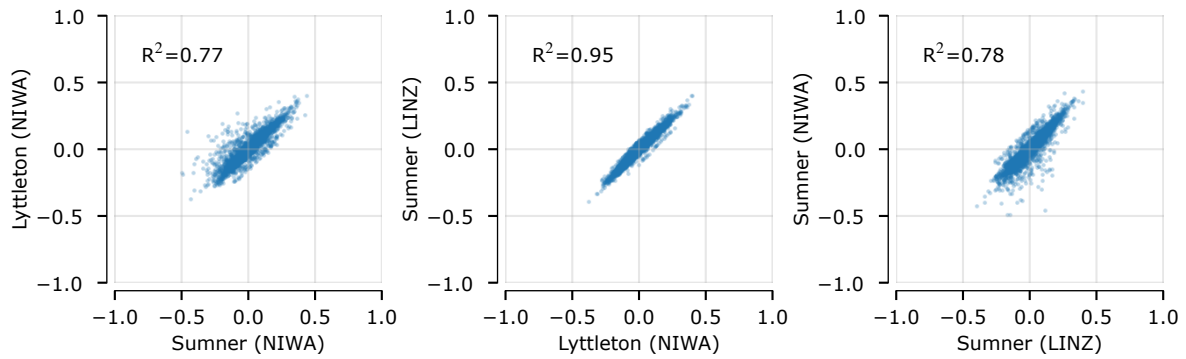


Figure A.17: 24 hour surge differences for Lyttelton NIWA, Sumner LINZ and Sumner NIWA, plotted with scatters to each other.

The relation between the two surge patterns is stronger when comparing the 24 hrs value than when comparing the instantaneous measurements, which makes sense as the surge is more sensitive to small outliers. The agreement especially between Lyttelton (NIWA) and Sumner (LINZ) is very good, an R^2 of 0.84 for the momentary values and an R^2 of 0.95 for the 24 hour average values. This is good news since it enables us to use the Lyttelton surge as a proxy for the Sumner surge, and the Lyttelton data goes all the way back to the start of the rain measurements in 1960.

A.5 Conclusions

We looked at several data sources for Sumner and Lyttelton, to assess to usefulness of the data for several goals:

1. Extreme value analysis
2. Analysis of tidal constituents
3. Sea level rise

Extreme value analysis

For the Extreme value analysis we need data that has good quality in the maxima. The temporary lowering of Sumner (NIWA) might be a problem for this use. For the period of 2010 and onwards we have different data available, the LINZ data. This however still leaves the pre 2010 period, for which there is no LINZ data. We could use the Lyttelton data here, but these tend to have slightly higher peaks. In that case it would be good to just determine the EV statistics at Lyttelton, and use these statistics. The advantage is the longer time period for which these data are available. We advice to derive to extreme value statistics with the cleaned data:

1. Use Sumner (NIWA) for EV analysis at Sumner. The LINZ data after 2010 are better, but since it only covers part of the time range it doesn't solve the issue. Therefore as a check:
2. Use Lyttelton (NIWA) for EV analysis at Lyttelton. These data are better for at least the recent period for which we could compare it to the other data sources. For the earlier years

there is no check available, but deriving both statistics might give some more confidence in both EV results.

Tidal constituents and surge

For the analysis of the tidal constituents and surge we mainly need consistent data (for estimating tide) with minimal bias (for estimating surge). Since Lyttelton data is available all the way back to 1960, the source has our preference for this purpose. If needed, missing periods can be filled by Sumner (LINZ) if available and Sumner (NIWA) otherwise. This could be desirable for deriving the surge, but a continuous data series is no requirement for the analysis of the tidal constituents.

Sea level rise

For the purpose of estimating sea level rise we need data with minimal biases. The problem with the Sumner (NIWA) data are the consistent lower periods. These can cause a temporary bias up to 300 mm. If these occur more often in the early years, it might give higher sea level rise estimates, or vice versa. This can partly be solve by using the Sumner (LINZ) data, but these are only available for 2010 and onwards. The differences with Lyttelton vary from time to time, sometimes Sumner is higher, sometimes Lyttelton. This makes it hard to classify the low periods before 2010. For this reason we advise to use the Lyttelton dataset for estimations of sea level rise, and apply these rates to Sumner too.

B Relative sea level rise at Lyttelton

This appendix contains an analysis on the sea level rise measured at Lyttelton and Sumner over the past decades (Sumner) and century (Lyttelton).

Given the short time period of data at Sumner and quality issues with much of its early period, the level data at Sumner is not considered suitable for primary sea level rise analysis. The short data which is available is assessed for comparison but confidence from this data is low.

B.1 Sea level rise components

First we describe the different components that together cause the sea level to be the level we can measure. We also explain which of the components will cause the sea level to rise over the long term.

The monthly averaged sea level is affected by a lot of environmental components, such as:

- **Tide**

That the tide affects the mean sea level is no surprise. However that it does also change quite a bit on a longer than monthly scale may be less self-evident. When the sun and moon align in certain ways, this can affect the gravitational pull on the sea water, causing higher or lower than usual tides. The most important example is the 18.6 year cycle in which the orbit of the moon around the earth intersects the ecliptic.

- **Seasonal changes**

The average sea level temperature changes throughout the year, which results in expansion of the water, affecting the sea level.

- **Barometric pressure**

A high barometric pressure pushes the sea level down locally. Theoretically this relation is 10 cm per 10 hPa, but the measured effect is a bit less due to environmental circumstances.

- **Interdecadal Pacific Oscillation (IPO)**

The IPO leads to fluctuations in the sea level temperature with a period of a few decades. Just as the temperature itself, this leads to density differences.

- **El Nino Southern Oscillation (ENSO)**

Fluctuations in the sea level temperature with a period of multiple years.

- **Glacial isostatic adjustment**

Glaciers caused by ice ages are so heavy that they push down tectonic plates. In warmer periods these plates slowly spring back, causing a rotation of the tectonic plates that can cause subsidence or rising land.

- **Gravity effects of melting land ice**

Land ice can be so thick and heavy that it has a noticeable gravitational effect on the sea water. Greenland for example attracts water, so if it melts more of the sea water will move to the southern hemisphere. For Antarctica the effect is opposite, the water will move to the northern hemisphere

- **Global warming**

As already described, temperature differences causes density changes in the sea water. Due to global warming the average sea temperature rises, causing a rising sea level.

- **Other land storage** For example dams, cutting down rainforest and groundwater depletion all have an effect on the water balance between the volume of water on land and in the sea. On the long term this also affects the sea level.

The first five effects are of a periodic nature, which means on the long term they won't have a lasting effect on the sea level. This is debatable for the episodic IPO and ENSO. Both might be affected by climate change, but we'll leave the possible effect of that for now since we are considering the past sea level rise. The last four components have a permanent (or very long term) nature, they result in a rate of sea level rise. We will try to quantify this by filtering the known periodic tidal components and fitting the residual with statistical trends. One could debate that the all these climate change affected components are not linear (or whatever statistical trend is chosen). That is true, but if the long term pattern adds up to an on average linear trend, that is a fact that might help us quantifying the observed rate of change.

John Hannah and Robert G. Bell have published several papers on this subject in the past. They concluded that the past sea level rise was around 1.5 to 2.0 mm per year. This analysis will focus on a present day rate of sea level rise, using tide gauge data to estimate *relative* sea level rise.

B.2 Measurements and other data

This section describes the different sources of data. Both the measured data we try to explain, such as the tide gauge measurements at Sumner and Lyttelton, and the components to use in explaining the data (barometric pressure, IPO, ENSO).

Lyttelton sea level data

Based on QA of the Lyttelton data we concluded that John Hannah's data is the most reliable data source for Lyttelton monthly average sea levels. For 2019 and 2020 (not present in Hannah's source) we extended this with data from the LPC gauge, governed by LINZ. The substantiation for using these data is given in the notebook on data QA. However since these data are only monthly averages, we do not have high frequency data for the period, which we need in order to determine sea level rise by analysing high tide levels. That's why we use the LPC gauge data from NIWA for that analysis. These data are the base for John's monthly averages, but are not assessed and filtered based on quality (by John). We did go through the data ourselves, to remove periods that seemed erroneous (also described in the data QA notebook).

As a check if there are no visually clear errors in the NIWA data, we plot the monthly averages for both sources (NIWA and Hannah/LINZ) of the LPC gauge data:

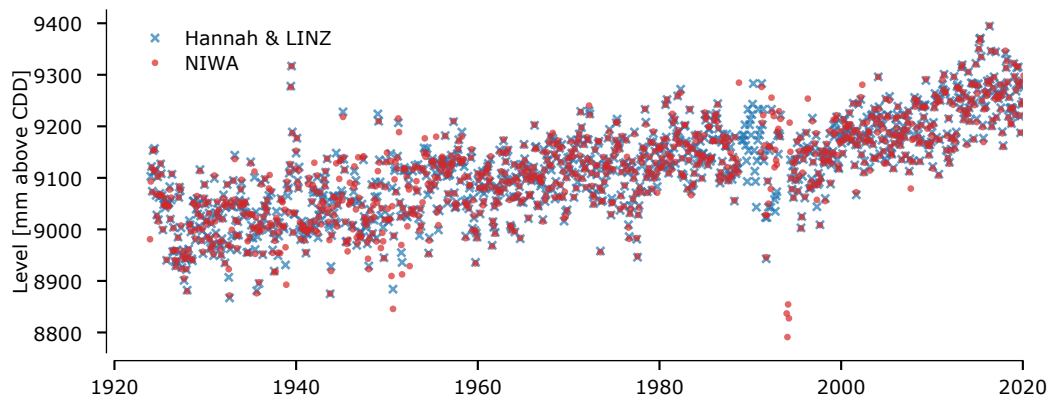


Figure B.1: Monthly average sea level from 1920 to 2020 at Lyttelton

As we can see, the monthly averages are mostly similar. There is no reason to assume we're using 'wrong' data.

Selecting high (and low) tide

The next step is to pick the high tide values. This is done by calculating the moving maximum in a 13 hour period. Centered, so 6.5 hours before and 6.5 hours after the value. By picking the values where they match, we will get the high (and low) tide values:

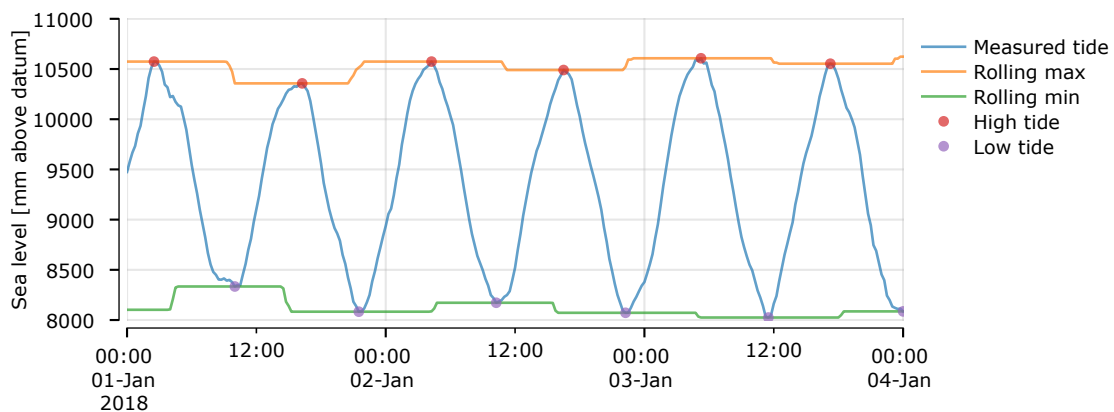


Figure B.2: Illustration of high and low tide selection

Note that fluctuations in the measured sea level might not always return the high tide level, but some value near high tide. Also the fact that the data frequency increases from 1 hour to 15 minutes in 1995 might lead to more fluctuations after 1995. Additionally an hourly value might give a slightly lower value, because you're less likely to find the exact peak. By fitting the tide to the observed values this effect is compensated, but in the figures below the post 1994 data might seem to give higher peaks and lower troughs.

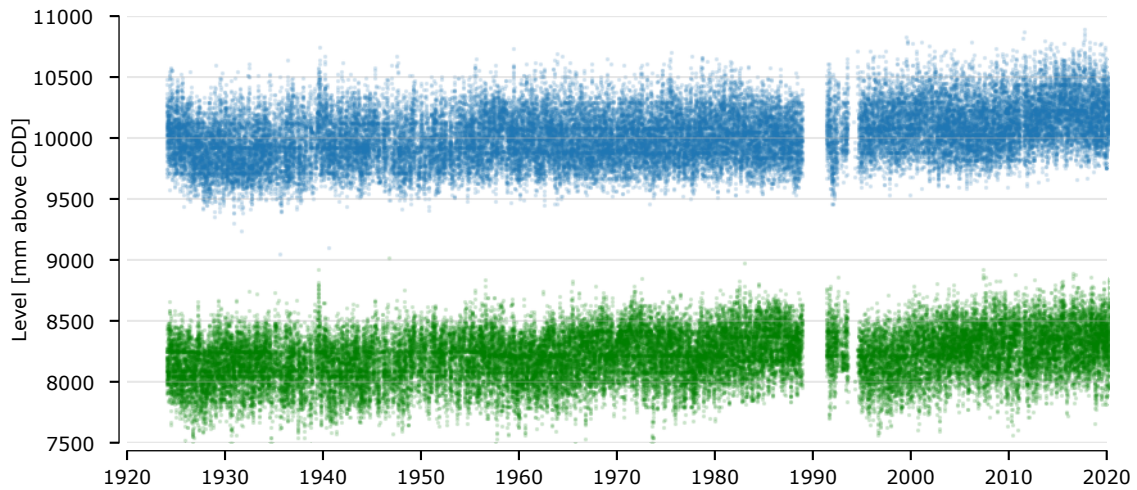


Figure B.3: High and low tides at Lyttelton from 1923 to 2020

The average low tide after 1990 is relatively low, compared to before 1990, while the high tide shows a more consistent pattern.

Summer sea level data

Similarly the data for Summer are important, since they are the boundary condition of our area of interest. The data is shorter and of lesser quality, but we can still analyse long term patterns such as sea level rise with it, as long as we're aware of the data shortcomings.

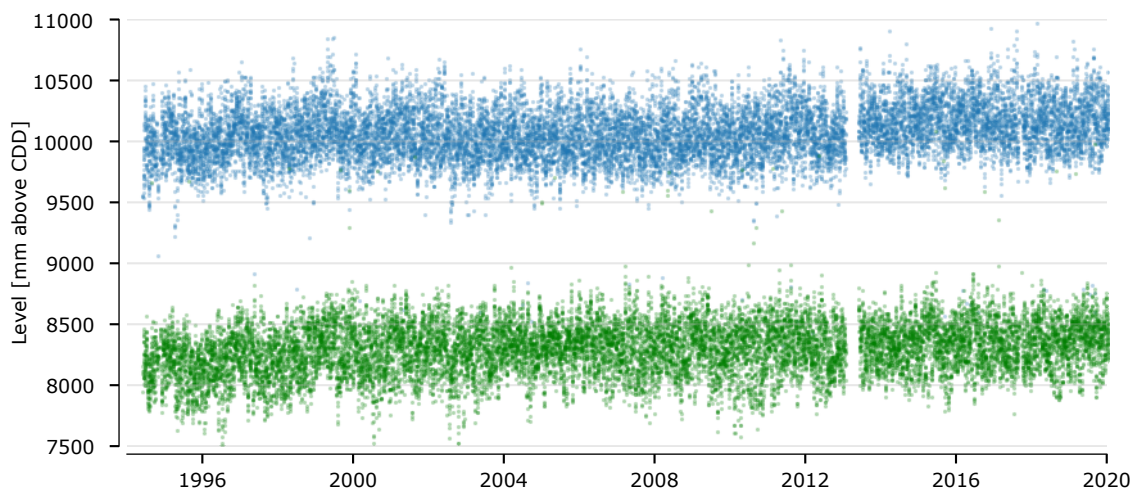


Figure B.4: High and low tides at Sumner (NIWA) from 1995 to 2020

Sea level explaining components

By quantifying the contribution of separate components, the pattern can be refined.

The following components are plotted (refer to the first section for a brief description of the components):

- The fitted astronomical tide.
- Interdecadal Pacific Oscillation (IPO)
- El Nino Southern Oscillation (ENSO)
- Barometric pressure data from Christchurch Aero and Kyle street. The Christchurch Aero data is not complete up to present day, so we use the Kyle Street data for the last years. These data are compensation for the altitude differences, and show a very good overlap: $R\text{-squared} = 0.999$.

Each of these components will be fitted to the model, to explain part of the 'noise' in the measurements.

Note that the IPO and ENSO might have an effect on the barometric pressure, so we might be overfitting a bit. However the barometric pressure has a very local effect, so we want the local data. The IPO and ENSO have a more regional effect, and explain some of the ocean temperature variations, for which we have no other good data source. That's why we use all three sources.

The different components are visualized in the next figure, together with the monthly sea level to which the components will be fitted. Note that for the barometric pressure the mean is subtracted, such that the pressure varies around 0. Otherwise it would affect the intercept, which is now only the offset for the sea level itself.

The monthly average and high tide values seem to have quite a lot of outliers, something you would not expect from a predicted pattern consisting of periodics. The fluctuations are caused by missing data, which can leave a month with just a few data points. If these are by coincidence a few low or high values, the monthly average tide (or level) is deviating from the mean. Since the difference of the tide and level is the factor we try to fit, these missing values are no problem for the least squares fitting.

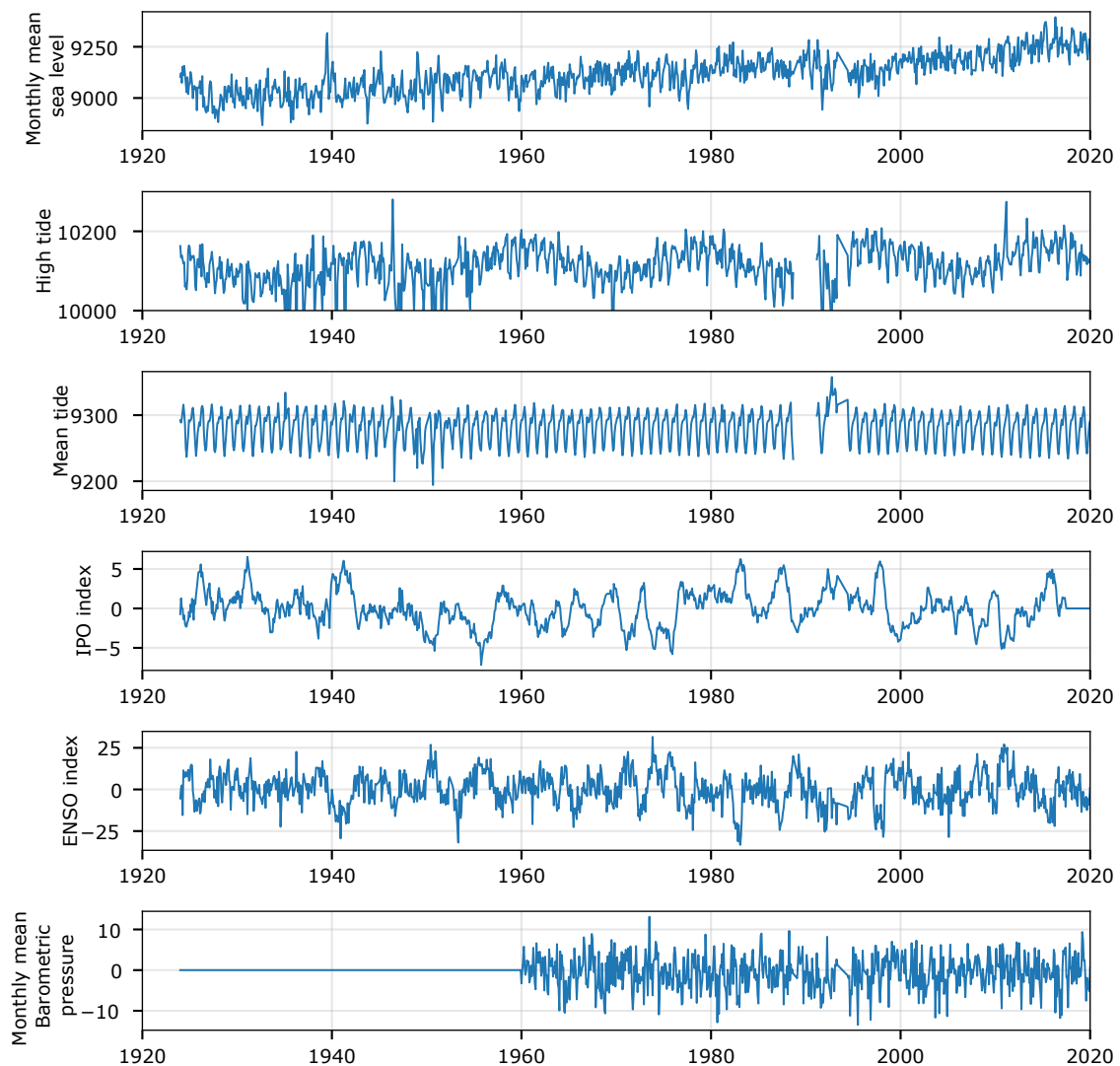


Figure B.5: Components that are fitted to the sea level in the least squares analysis

B.3 Tide

The daily sea levels are mainly determined by the astronomical tide. Good tide data are thus important for analysing patterns in the remaining factors influencing the sea levels. These (predicted) tide data are subtracted from the monthly sea level data, to produce an unexplained residual to which we can fit the trends.

To estimate the effect of the tidal constituents on the sea level rise, we've chosen to fit the tide as a whole. This has some advantages compared to fitting the constituents separately, namely that we don't need to know the exact components before fitting. Especially for the shorter data series (Sumner) the longer tidal constituents like the lunar node is hard to fit, and we might easily be fitting it wrongly.

The disadvantage of fitting the tide as a whole is that we lose track of which components are

important. To solve this we do a frequency analysis on the *fitted* tide. After this we fit the found frequencies to the tide with an ordinary least squares approach. This way we still have a quantification of the important components.

The harmonics in the least square fit are defined with a sin and cos term, which can be combined in with coefficients that they fit any sinusoidal with amplitude A and phase ϕ . So:

$$\text{periodic} = b_u \cos\left(2\pi \frac{t}{T}\right) + b_v \sin\left(2\pi \frac{t}{T}\right) \quad (1)$$

The amplitude A and phase ϕ can be calculated with:

$$A = \sqrt{b_u^2 + b_v^2} \quad (2)$$

$$\phi = y_0 + \tan^{-1}\left(\frac{b_u}{b_v}\right) \cdot \frac{T}{2\pi} \quad (3)$$

Lyttelton

The next four figures show the frequency analysis of the monthly high tide and mean tide, for Lyttelton and Sumner. The frequency analysis is aimed at low frequency components, so we only show the result for > 3 month harmonics. This means that tidal constituents like N2, M2, S2 are not present in the FFT results, expect for the lower frequency components that are the results of the coinciding of the phases of such components. Refer to appendix A for the fitted amplitude and phase values for these constituents.

The found harmonics are fitted to the *predicted* tide with an ordinary least squares fit. From this the most important components are listed ($p < 0.025$), as well as the R^2 .

High tide:

R^2 value for the fit: 0.86

Significant periodics:

- Lunar Node (18.6 y)
- Lunar Perigee (0.56 y)
- Semi-annual (0.5 y)
- Annual (1.0 y)

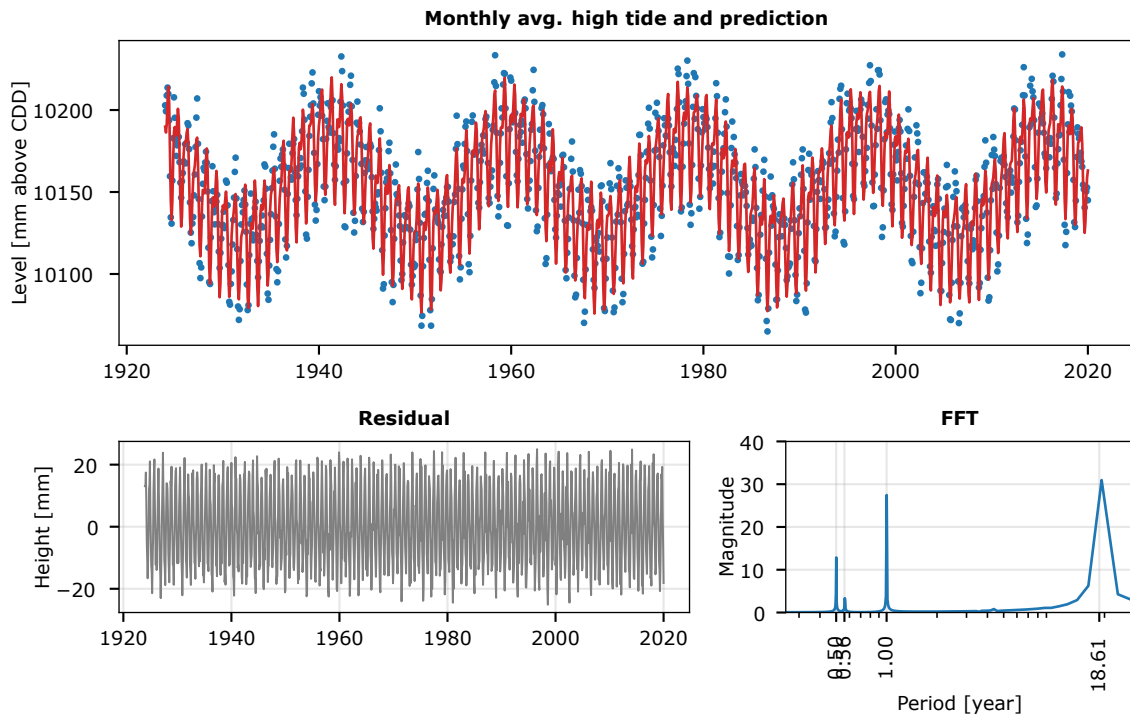


Figure B.6: FFT analysis of low frequency tidal components in Lyttelton high tide levels

Mean tide:

R^2 value for the fit: 1.00 (rounded to 2 decimals)

Significant periodics:

- Semi-annual (0.5 y)
- Annual (1.0 y)

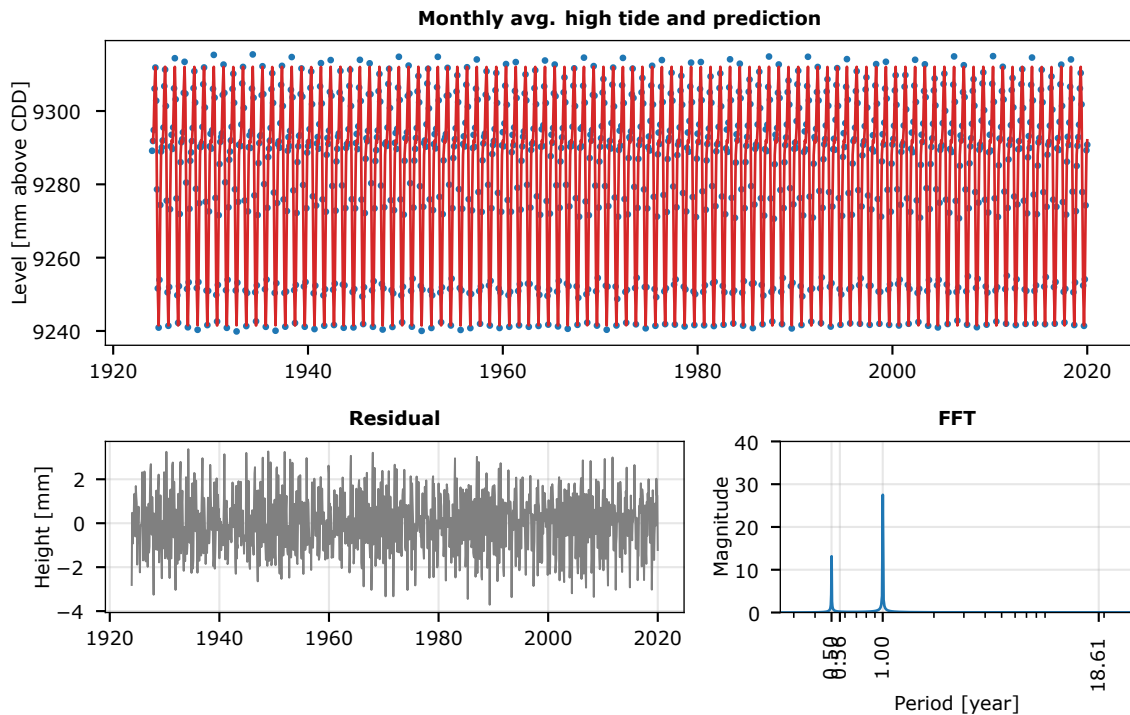


Figure B.7: FFT analysis of low frequency tidal components in Lyttelton mean tide levels

For Lyttelton there is a clear difference between the components that are analysed from the high tide, and the components present in the mean tide. The 18.613 year lunar nodal cycle catches the eye directly in the figure above. Apparently this effect is not distinguishable in the average sea level. This could be compensated by the low tide in the tidal prediction, which can be equally lower. Note that we are analysing the predicted tide here, not the measured tide. So if the prediction does not cover a certain physical effect (the shallow water components of the tide *are* fitted, the question is how good) it will not be discovered in this analysis.

Three shorter duration components are:

- annual effect, seasonal (1 y);
- semi-annual effect, seasonal (0.5 y);
- the phase in which the synodic month (period between different new or full moons viewed from earth) and the anomalistic month (period between perigees, the moment the varying distance between moon and earth is shortest) coincide (0.56 y). In other words, the moment the moon and sun align, and the month is closest to earth at the same time.

This last component is also not present in the average sea level frequency analysis, because it will effect the low tide too. The 0.5 and 1 year periods are however present, since these are seasonal effects (like temperature differences of the sea water) that will effect low and high tide alike.

Sumner

For the Sumner data from NIWA we did a similar analysis of the frequencies. Note that the Sumner (NIWA) data have some problems with consistently lower values for extended periods of time. For tidal constituent analysis the data are however good enough. The amplitudes of the constituents might be a bit off, but this will not prevent us from discovering the main constituents with their phases, which is the information we need for this FFT analysis.

High tide:

R^2 value for the fit: 0.86

Significant periodics:

- Lunar Node (18.6 y)
- Lunar Perigee (0.56 y)
- Semi-annual (0.5 y)
- Annual (1.0 y)

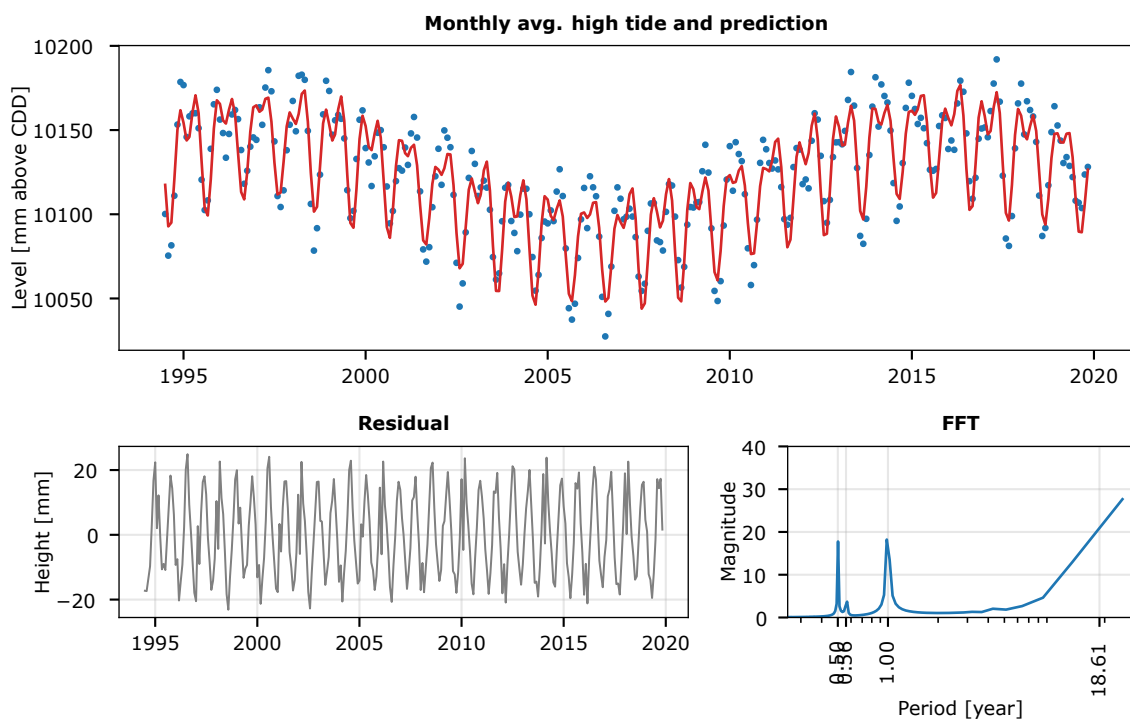


Figure B.8: FFT analysis of low frequency tidal components in Sumner high tide levels

Mean tide:

R^2 value for the fit: 1.00 (rounded to two decimals)

Significant periodics:

- Semi-annual (0.5 y)
- Annual (1.0 y)

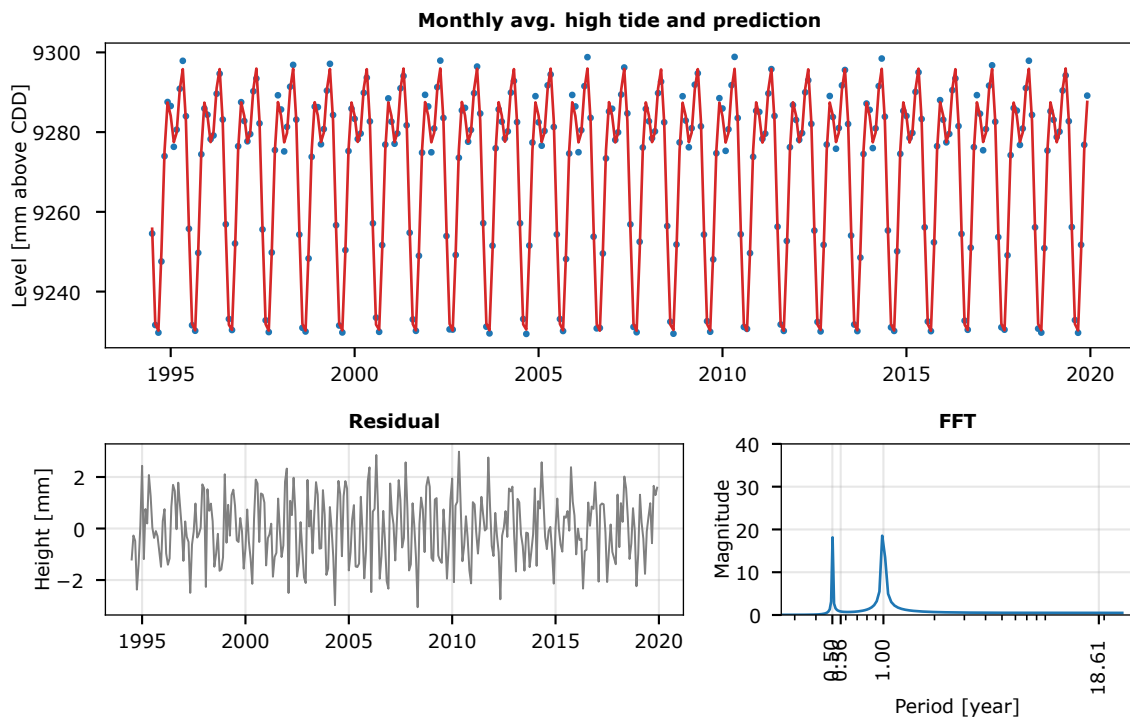


Figure B.9: FFT analysis of low frequency tidal components in Sumner mean tide levels

For Sumner the results are similar as for Lyttelton, except that the 18.613 year cycles can not clearly be found with a FFT analysis for the high tide. The rising end of the curve however suggest that it could more clearly be distinguished with a longer time series; the available data period is too short to find components in this low frequency range.

B.4 Sea level rise models

In this chapter the different models with which we'll try to model the sea level are explained. It contains mainly formulas and an explanation of the components in the sea formulas. Four of these trend models are defined:

0. A single linear trend The first model is a single linear trend. It is the most basic variant, without any presumed acceleration of the rate of sea level rise.

The model can also be used for the comparison between Sumner and Lyttelton, by selecting only the data after september 1994 for the last location.

The equation for the model is:

$$H_0(t) = H_{tide} + a + b_{trend}(t) + b_p p(t) + b_{IPO} I_{IPO}(t) + b_{ENSO} I_{ENSO}(t) + \varepsilon(t-1) + \varepsilon(t) \quad (4)$$

In which H_{tide} is the level of the fitted (predicted) tide. Note that we do not fit the tide separately in the least squares model. a is the intercept and b_{trend} is the linear trend coefficient. Similarly b_p is the factor for the barometric pressure, $p(t)$. b_{IPO} is the factor for the interdecadal pacific oscillation index, $I_{IPO}(t)$, and b_{ENSO} is the factor for the El Nino Southern Oscillation index, $I_{ENSO}(t)$. Finally $\varepsilon(t)$ is the error term. By including $\varepsilon(t-1)$ the autocorrelation in the error is taken into account. By taking this autocorrelation into account we are actually fitting a GLSAR model (Generalized Least Squares with AR covariance structure) instead of a OLS (Ordinary Least Squares) model.

1. A trend break in 1990 This scenario is based on the assumption that the rate of sea level rise over the past decades is higher that it was before. On itself, this is not a strange assumption: we only need to look at the IPCC climate change scenario's to see that expected rate of sea level rise go up to 1 meter per 100 years. Modelling this as a break is however a crude model for the acceleration. In reality the contributing components are too complex to be modeled with one or two linear components. But it is a convenient model since the fitted components directly return the rate of sea level rise per year. Since the early '90s had a few colder years with a slightly lower sea level, a trend break in the early '90s gives usually a good fit. Additionally, global coverages for sea level measurements from satellites started in 1993. So when using these data only the trend after the '93 is reported, which will be more or less a similar results as looking at the trend after the '90s break. Since the Lyttelton data around 1990 is largely missing, it is hard to determine which year would be best for the breakpoint. We therefore use 1990, halfway the missing data. For Sumner this does not matter, since there is only data after 1994 anyway.

$$H_1(t) = H_0(t) + b_{break}(t > 1990) \cdot (t - 1990) \quad (5)$$

In which b_{break} is the break coefficient, which is the extra linear trend after the trend break.

2. A trend break and shift in 1990 In this scenario we add a shift in 1990. This shift causes an instantaneous drop in the sea level rise. This model has even less physical substantiation than a break. It however enables us to get a more accurate estimate of the rate of sea level rise over the last 30 years. This is often the only period we for which we have data, and it will get the low and high tides and the extremes more relative to the mean sea level. For fitting tidal constituents these 'more relative' differences is what we need.

The formula is adjusted by adding a coefficient b_{shift} , that is non-zero for all years after the break year:

$$H_2(t) = H_0(t) + b_{break}(t > 1990) \cdot (t - 1990) + b_{shift}(t > 1990) \quad (6)$$

3. A trend break in 2005 This scenario is introduced since it fits the MfE assessment, in which the rate is presented before and after 2005. The formula is similar to $H_1(t)$, except that the breakyear is 2005:

$$H_3(t) = H_0(t) + b_{break}(t > 2005) \cdot (t - 2005) \quad (7)$$

We will denote these four models as SLR0, SLR1, SLR2 and SLR3. In most of the cases the models are used to describe the mean sea water level rise. A 'p' (for peak) is added when we are using the model specifically for *high tide* rise

B.5 Lyttelton trends

In this chapter the trend models are fitted to the data for Lyttelton. For one of the fits the results are explained elaborate, for the other only the fitted trends are shown to keep the results concise. The appendix contains the full results for the final chosen model.

Linear fit

First we fit the linear model (SLR0), and present the results the model returns:

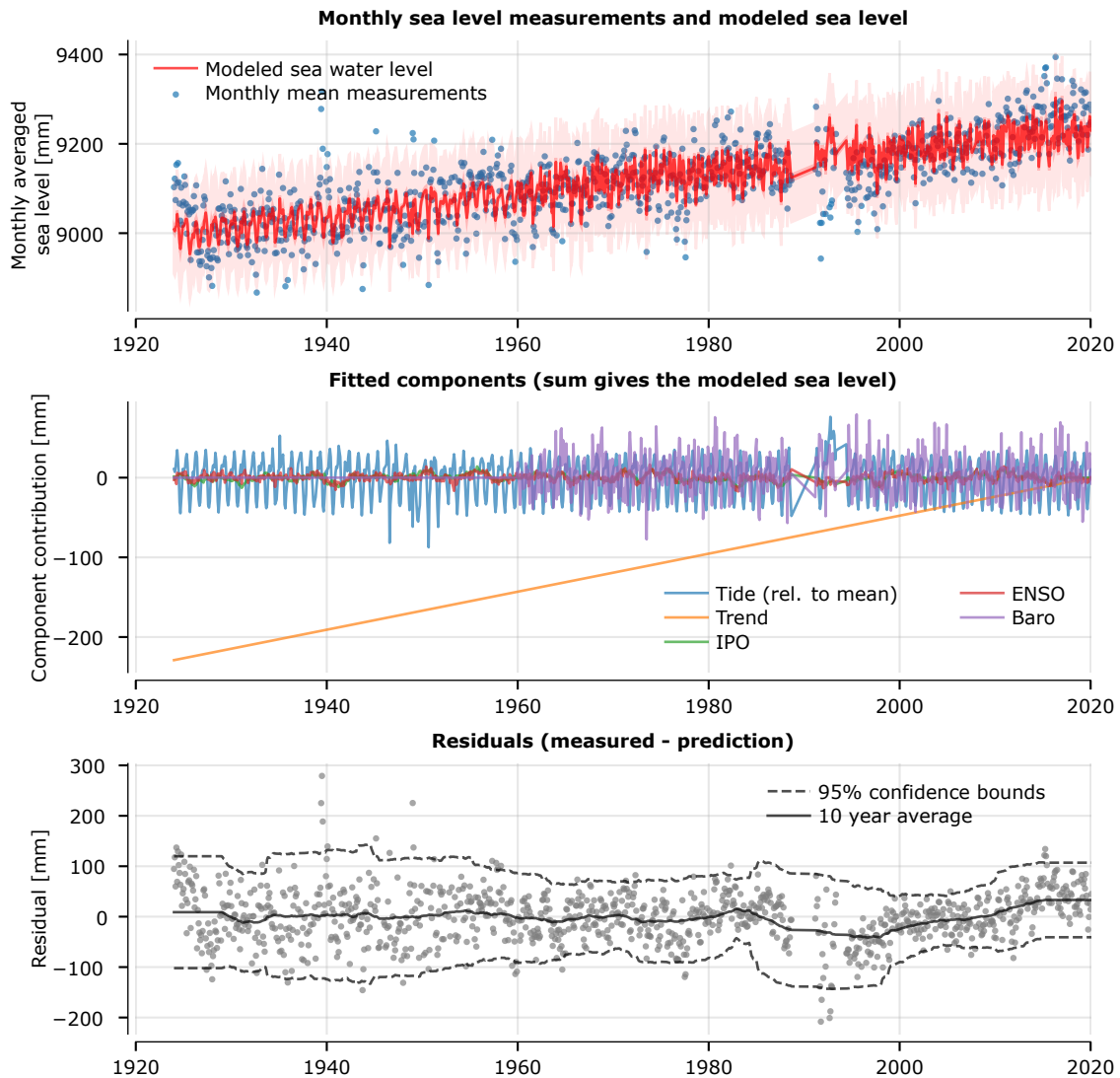


Figure B.10: Linear trend in relative sea water level rise for the monthly mean level at Lyttelton

The top graph shows the measured sea level (blue dots) and the predicted sea level from the model (red line). The pink area is the 95% prediction interval, the limits within which 95% of the monthly mean measurements are expected from the model. Closer to the red line, there is also a narrow light red band. This is the 95% confidence interval, which is the interval with which we expect our model to be in 95% of the cases, under varying data (blue dots). The plot below shows the contribution of each component. We can see that the trend is the most important component, followed by the barometric pressure and tide. The IPO and ENSO are least important.

The last figure shows the residual, we can distinguish three regions:

1. Up to 1960 there is relatively much noise. This is the period for which there is no barometric pressure data available. This is the main component of the large noise.
2. From 1960 up to 1990 the noise is less. There seems to be some autocorrelation; the noise

seems to follow a pattern. This indicates that there might be a physical component we're missing, but it can just as well be an error in the measurements.

- In the period from 1990, the 10 year average residual seems to rise gradually. This indicates that the rate in sea level rise might be higher in that period than it was before 1990. Considering the climate change scenario this is not strange, these scenarios have a tendency to give increasingly high estimates for the sea level in 2050, 2085 or 2100. Rates that are closer to 10 mm/year for the worst-case scenarios rather than 2.5 mm/year. The problem is however that the rise seems to come from below average. The early 1990's were relatively cold ², which might have caused lower sea levels due to higher water densities, but the average residual is negative until 2010, while the cold years were only a few.

Next we will look at the numerical results for the GLSAR model (an abbreviation for Generalized Least Squares with AR covariance structure). This model is similar to an Ordinary Least Squares (OLS) model, except that it takes an autocorrelation in the error into account. We do this because the error is likely not only noise, but also contains some (by the model) unexplained physics. These physical process will likely be consistent over several months, so this gives an autocorrelation in the error, which is added to the OLS model.

Table B.1: GLSAR Regression Results for Lyttelton linear trend

Dep. Variable:	y	R-squared:	0.665			
Model:	GLSAR	Adj. R-squared:	0.664			
Method:	Least Squares	F-statistic:	493.3			
Date:	Fri, 10 Jul 2020	Prob (F-statistic):	1.40e-239			
Time:	11:33:38	Log-Likelihood:	-5651.1			
No. Observations:	1058	AIC:	1.131e+04			
Df Residuals:	1053	BIC:	1.134e+04			
Df Model:	4					
Covariance Type:	HC0					
	coef	std err	z	P> z 	[0.025	0.975]
Constant	-53.5652	2.939	-18.223	0.000	-59.326	-47.804
Trend	2.3852	0.059	40.441	0.000	2.270	2.501
IPO	-1.9713	0.836	-2.359	0.018	-3.609	-0.333
ENSO	0.5224	0.196	2.670	0.008	0.139	0.906
Baro	-5.9014	0.437	-13.513	0.000	-6.757	-5.045
Omnibus:	55.733	Durbin-Watson:	0.864			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	181.936			
Skew:	0.127	Prob(JB):	3.11e-40			
Kurtosis:	5.015	Cond. No.	112.			

²http://archive.stats.govt.nz/browse_for_stats/environment/environmental-reporting-series/environmental-indicators/Home/Atmosphere-and-climate/temperature-time-series.aspx

The tables describe:

1. The first table shows some general characteristics.
2. The second table shows the importance of each coefficient. A small $P > |z|$ shows that the term is significant. Z is the number of standard deviations the mean (coefficient) is from zero. Note that the periodic factors consist of a U and V component, from which the amplitude and phase can be calculated as described earlier.
3. The third table shows some additional data on the quality of the model, by doing statistical tests on the residuals.

Last but not least: By simply fitting a linear model to the pattern, the rate of mean sea level rise is 2.4 mm per year. This is higher than rates presented in literature. This is because the years 1900-1923 are not included. For these years we do not have the monthly data (only annual), so we can not use them in this fit.

The next figures show the fitted trends (all described models) to the Lyttelton data, for mean sea level and high tide level. The rates of sea level rise are shown in the legend.

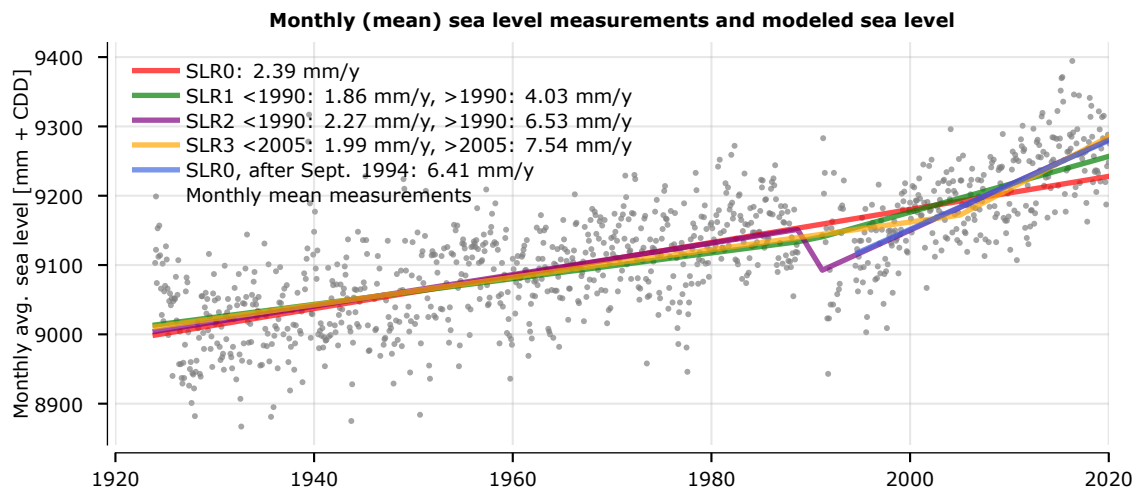


Figure B.11: Trends in relative sea water level rise for the monthly mean level at Lyttelton.

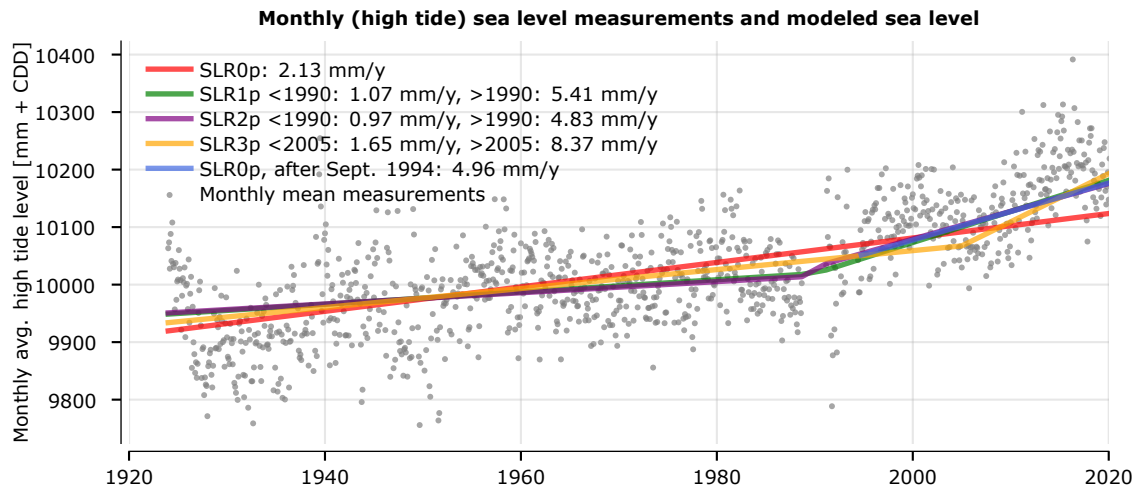


Figure B.12: Trends in relative sea water level rise for the monthly high tide level at Lyttelton.

B.6 Sumner trends

In this section the trends are fitted to Sumner data. Sumner data are available from September 1994 onwards. Therefore any scenario with a break or shift in 1990 is not relevant. We only use the linear trend (no break or shift), which is SLR1, and the break in 2005, SLR3.

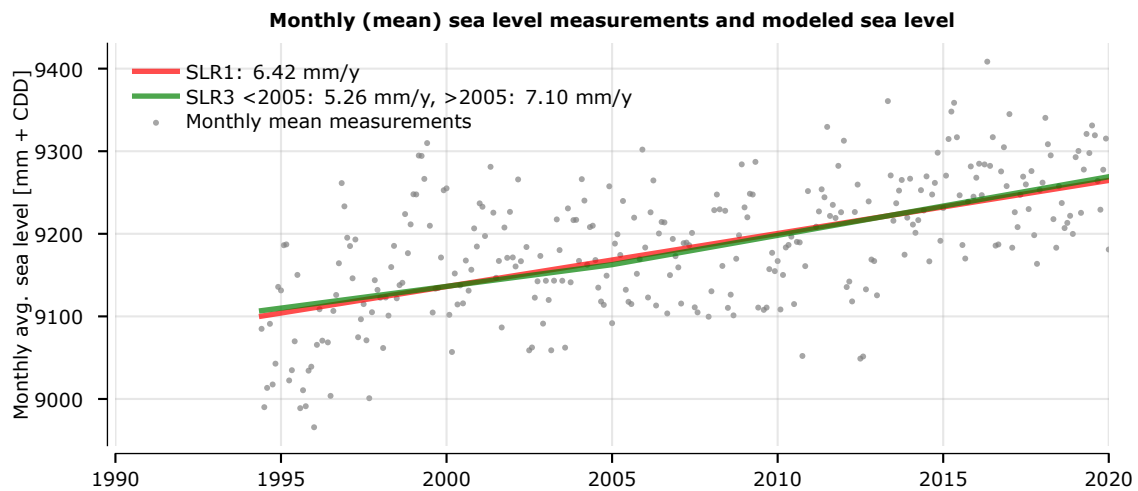


Figure B.13: Trends in relative sea water level rise for the monthly mean level at Sumner

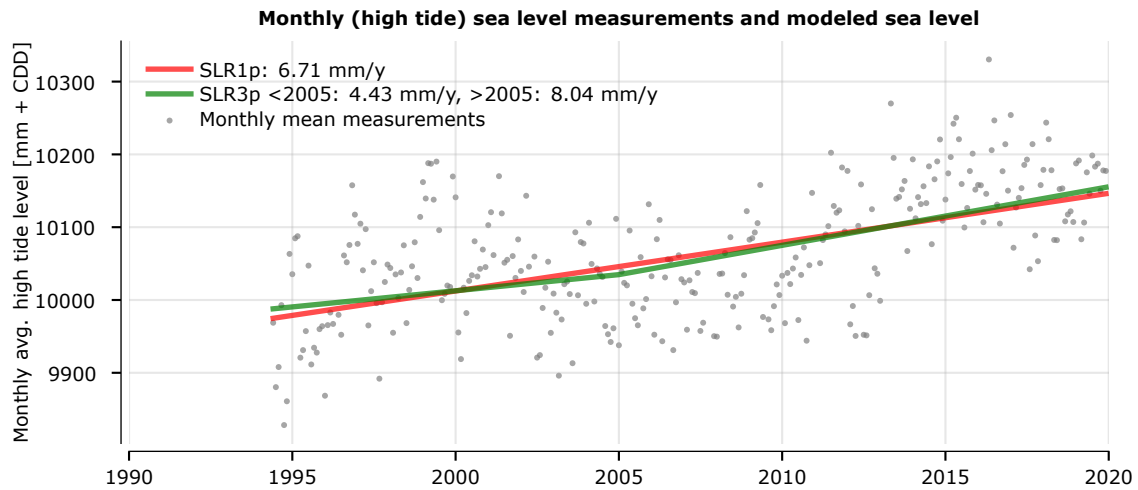


Figure B.14: Trends in relative sea water level rise for the monthly high tide level at Sumner

B.7 Conclusions

The following tables summarize the trends values before and after the break.

Table B.2: Lyttelton rates of sea level rise

	Mean level Trend	Mean level Trend after break	Mean level Shift in 1990	High tide Trend	High tide Trend after break	High tide Shift in 1990
SLR1(p)	1.86 ± 0.10	$4.03 \pm 0.26^*$		1.07 ± 0.12	$5.41 \pm 0.28^*$	
SLR2(p)	2.27 ± 0.10	$6.53 \pm 0.34^*$	-69 ± 7	0.97 ± 0.13	$4.83 \pm 0.35^*$	16 ± 7
SLR3(p)	1.99 ± 0.07	$7.54 \pm 0.44^*$		1.65 ± 0.09	$8.37 \pm 0.51^*$	
SLR0(p), after Sept. 1994	6.41 ± 0.25			4.96 ± 0.29		

Table B.3: Sumner rates of sea level rise

	Mean level Trend	Mean level Trend after break	High tide Trend	High tide Trend after break
SLR1(p)	6.42 ± 0.43		6.71 ± 0.43	
SLR3(p)	5.26 ± 1.42	$7.10 \pm 1.86^*$	4.43 ± 1.37	$8.04 \pm 1.82^*$

* Note that trend after the break consists of the trend before the break plus the additional rate. The standard error for the trend after the break is only for the additional rate of sea level rise.

- For Lyttelton the rate of *mean sea level* rise is significantly higher than the rate of *high tide* rise. This difference implies that the tidal range is increasing (as well as the mean sea level rise). We saw that the nodal cycle (18.6 y) is much more present in the high tide fit than the average sea level, but the same tidal constituents are used in fitting the model. A reason could be a change in the bathymetry around the gauge in time, which had a different effect on the measures high waters than the average measurements.

A possible explanation for the differences between sea level rise from mean and high tide is that over the course of time the method of measuring changed a bit, which could lead to a gradual change in the measurements. In figure B.3 we noticed a drop in the low tide values after the data gap around 1990, which might support this possibility. Also the fact that the Sumner trends show a similar value for high tide and mean sea level supports the idea. Although the Sumner data has some problems, it is probably a more consistently measured source given the shorter time span it covers.

Further analysis would be required to validate and understand the true cause of this difference.

- The '1990 shift' in SLR2 is slightly positive in the high tide results, in contrary to the mean level results where it is negative. Physically this model was already questionable, but due to this inconsistency between mean and high tide we will not use SLR2 for the final analysis.
- The sea level rise (SLR1 for Sumner, comparable to Lyttelton "SLR0, after Sept. 1994") are similar for the mean tide: 6.4 mm per year. For the high tide there is a difference, 6.7 for Sumner to 5.0 mm per year for Lyttelton.
- We advise to use the mean tide results for Lyttelton as value for the sea level rise. It is a very long time series and agrees with Sumner for the post 1994 data. So use the rates of sea level rise from scenario SLR1, with a trend break in 1990 and no vertical shift in the data.
- The standard errors are shown with the \pm sign in the tables. The standard error is an estimate of the standard deviation of the fitted coefficient. Note that this is not equal to the 95% confidence interval. If the standard errors are assumed to be normally distributed, the 95% confidence interval is given by $[\mu - 1.96 \sigma, \mu + 1.96 \sigma]$.
- Furthermore, if the standard errors are indeed normally distributed, OLS is the maximum likelihood estimator. If this is not the case, a maximum likelihood estimation gives different results, and therefore also different standard errors. The OLS approach is however relatively convenient and accessible, which is why we apply it.

B.8 Full results SLR1

This last section gives the full regression results for SLR1. Both Lyttelton and Sumner with high tide and mean sea water level are given.

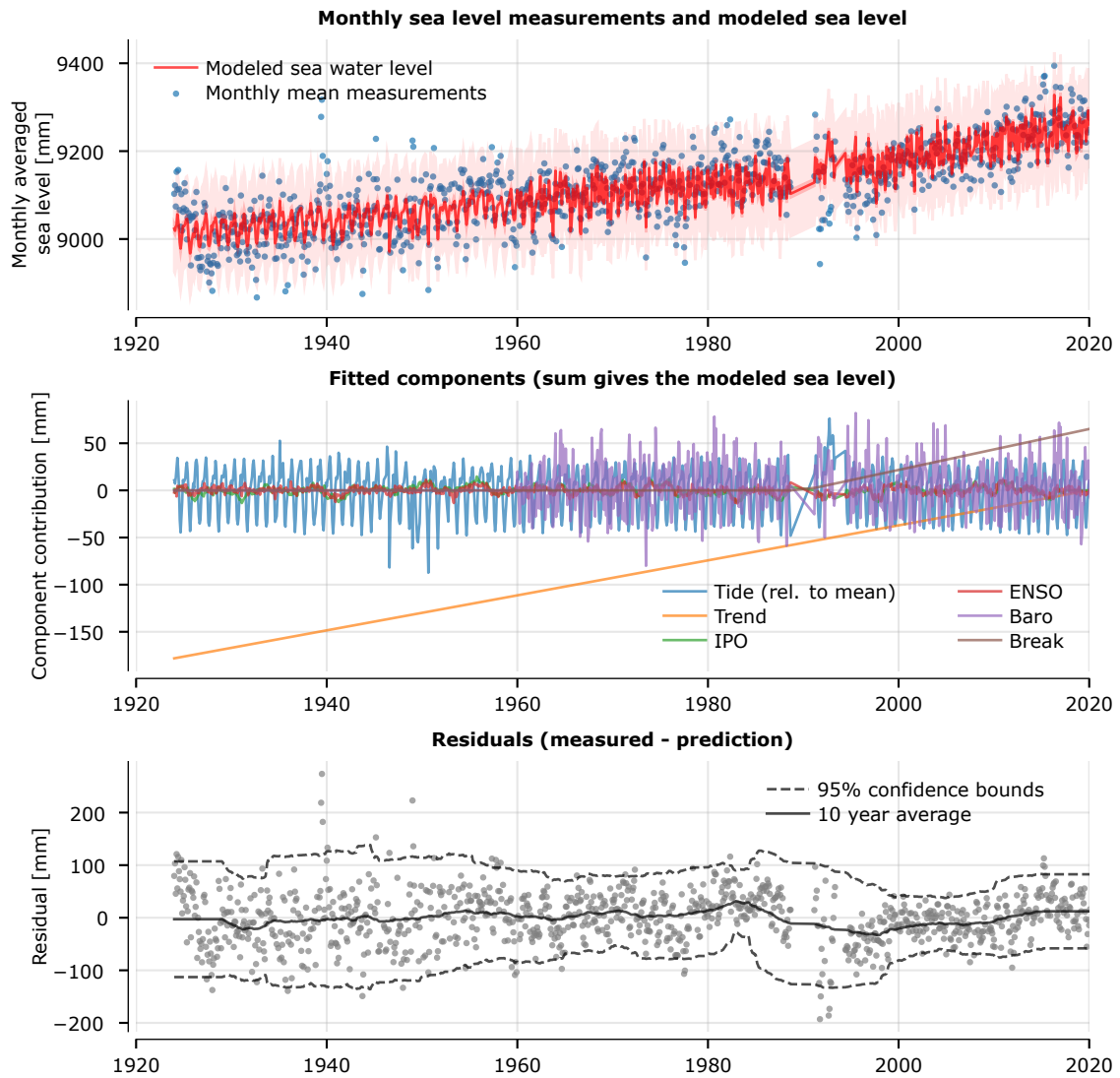


Figure B.15: SLR1 trend fit of mean sea level at Lyttelton

Table B.4: GLSAR Regression Results for Lyttelton mean level SLR1

Dep. Variable:	y	R-squared:	0.684			
Model:	GLSAR	Adj. R-squared:	0.682			
Method:	Least Squares	F-statistic:	570.5			
Date:	Fri, 10 Jul 2020	Prob (F-statistic):	1.47e-296			
Time:	11:40:03	Log-Likelihood:	-5620.9			
No. Observations:	1058	AIC:	1.125e+04			
Df Residuals:	1052	BIC:	1.128e+04			
Df Model:	5					
Covariance Type:	HCO					
	coef	std err	z	P> z 	[0.025	0.975]
Constant	-89.8026	5.755	-15.605	0.000	-	-78.523
					101.082	
Trend	1.8557	0.097	19.189	0.000	1.666	2.045
IPO	-2.0605	0.821	-2.510	0.012	-3.669	-0.452
ENSO	0.4196	0.190	2.209	0.027	0.047	0.792
Baro	-6.0985	0.404	-15.093	0.000	-6.891	-5.307
Break	2.1701	0.261	8.307	0.000	1.658	2.682
Omnibus:	51.499	Durbin-Watson:	0.909			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	161.221			
Skew:	0.106	Prob(JB):	9.80e-36			
Kurtosis:	4.901	Cond. No.	203.			

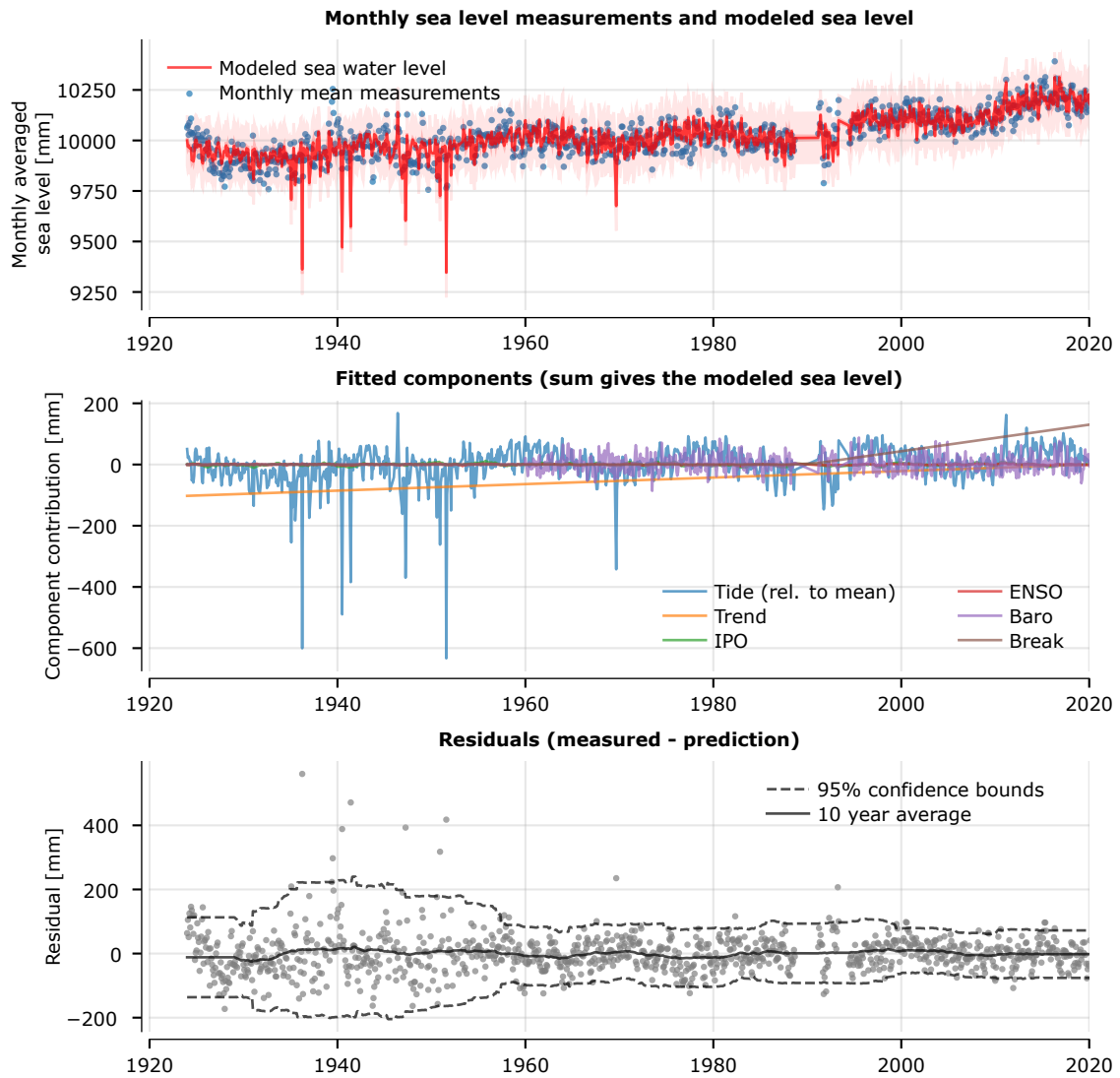


Figure B.16: SLR1 trend fit of high tide level at Lyttelton

Table B.5: GLSAR Regression Results for Lyttelton high tide SLR1p

Dep. Variable:	y	R-squared:	0.547			
Model:	GLSAR	Adj. R-squared:	0.545			
Method:	Least Squares	F-statistic:	444.5			
Date:	Fri, 10 Jul 2020	Prob (F-statistic):	2.15e-256			
Time:	11:40:14	Log-Likelihood:	-5877.8			
No. Observations:	1058	AIC:	1.177e+04			
Df Residuals:	1052	BIC:	1.180e+04			
Df Model:	5					
Covariance Type:	HC0					
	coef	std err	z	P> z 	[0.025	0.975]
Constant	-61.1064	6.212	-9.837	0.000	-73.282	-48.931
Trend	1.0667	0.117	9.121	0.000	0.837	1.296
IPO	-1.6575	1.265	-1.311	0.190	-4.136	0.821
ENSO	0.1619	0.328	0.494	0.621	-0.481	0.804
Baro	-6.5535	0.397	-16.521	0.000	-7.331	-5.776
Break	4.3400	0.275	15.758	0.000	3.800	4.880
Omnibus:	586.077	Durbin-Watson:	1.366			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	9774.308			
Skew:	2.172	Prob(JB):	0.00			
Kurtosis:	17.243	Cond. No.	203.			

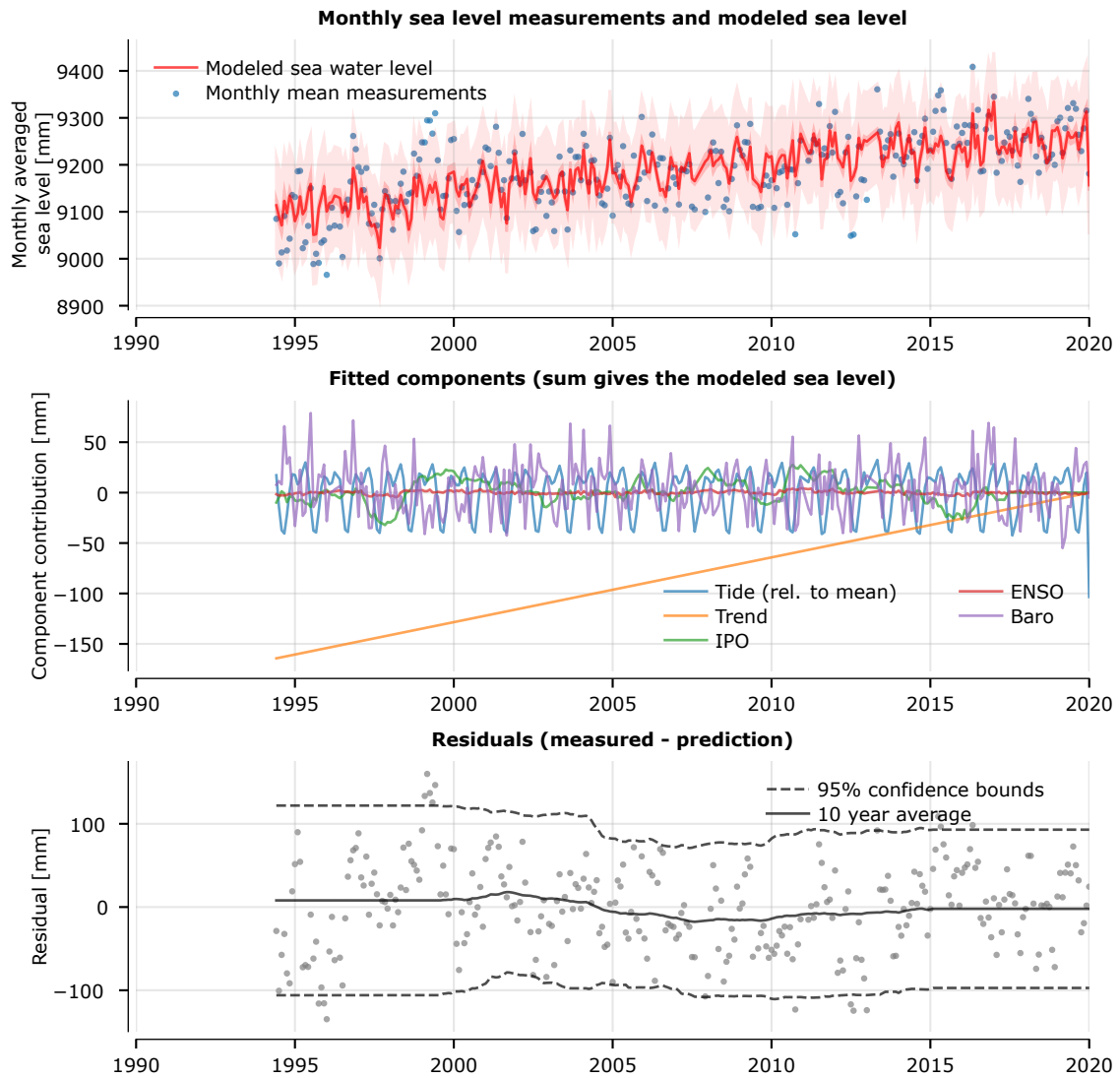


Figure B.17: SLR1 trend fit of mean sea level at Sumner

Table B.6: GLSAR Regression Results for Summer mean level SLR1

Dep. Variable:	y	R-squared:	0.523			
Model:	GLSAR	Adj. R-squared:	0.516			
Method:	Least Squares	F-statistic:	86.78			
Date:	Fri, 10 Jul 2020	Prob (F-statistic):	2.18e-48			
Time:	11:40:22	Log-Likelihood:	-1587.3			
No. Observations:	295	AIC:	3185.			
Df Residuals:	290	BIC:	3203.			
Df Model:	4					
Covariance Type:	HCO					
	coef	std err	z	P> z 	[0.025	0.975]
Constant	-4.6836	5.641	-0.830	0.406	-15.740	6.372
Trend	6.4219	0.434	14.801	0.000	5.571	7.272
IPO	-5.3958	1.844	-2.926	0.003	-9.010	-1.782
ENSO	0.1604	0.362	0.444	0.657	-0.548	0.869
Baro	-5.8665	0.720	-8.143	0.000	-7.279	-4.454
Omnibus:	0.008	Durbin-Watson:	0.759			
Prob(Omnibus):	0.996	Jarque-Bera (JB):	0.024			
Skew:	0.010	Prob(JB):	0.988			
Kurtosis:	2.961	Cond. No.	29.6			

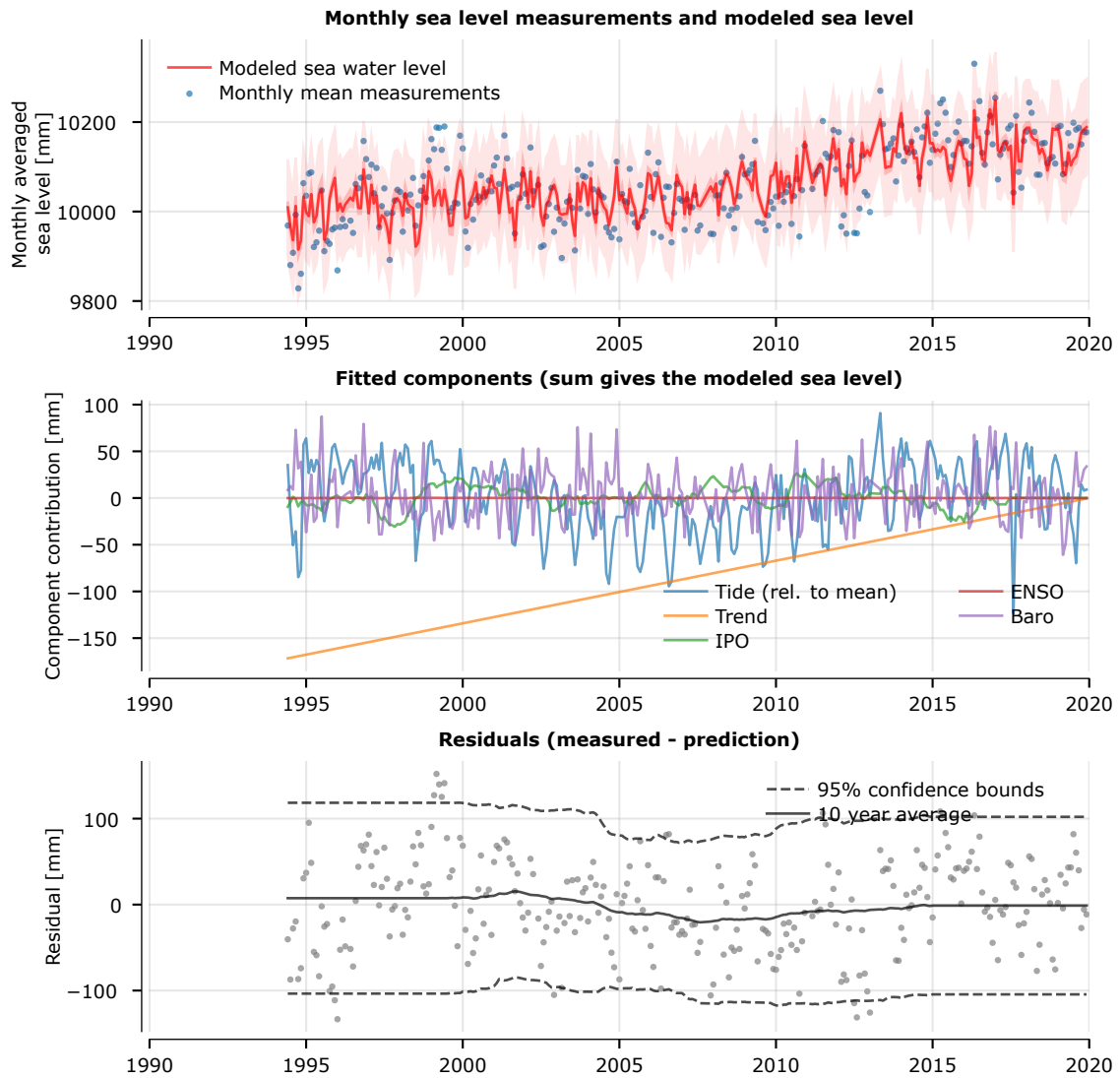


Figure B.18: SLR1 trend fit of high tide level at Sumner

Table B.7: GLSAR Regression Results for Sumner high tide SLR1p

Dep. Variable:	y	R-squared:	0.536			
Model:	GLSAR	Adj. R-squared:	0.529			
Method:	Least Squares	F-statistic:	88.03			
Date:	Fri, 10 Jul 2020	Prob (F-statistic):	7.99e-49			
Time:	11:40:27	Log-Likelihood:	-1587.7			
No. Observations:	294	AIC:	3185.			
Df Residuals:	289	BIC:	3204.			
Df Model:	4					
Covariance Type:	HCO					
	coef	std err	z	P> z 	[0.025	0.975]
Constant	43.9230	5.925	7.413	0.000	32.311	55.535
Trend	6.7085	0.434	15.465	0.000	5.858	7.559
IPO	-5.1544	1.917	-2.689	0.007	-8.911	-1.398
ENSO	0.0095	0.358	0.026	0.979	-0.692	0.711
Baro	-6.4943	0.716	-9.068	0.000	-7.898	-5.091
Omnibus:	0.432	Durbin-Watson:	0.760			
Prob(Omnibus):	0.806	Jarque-Bera (JB):	0.562			
Skew:	0.009	Prob(JB):	0.755			
Kurtosis:	2.787	Cond. No.	29.8			

Note that all the coefficients for barometric pressure change are closely grouped around 6 cm / hPa. This compares reasonably with classical physics which would expect 10cm/hPa. It is common to see such differences in real life due to secondary environmental factors.

C Water level differences on estuary

This appendix is about the water level differences observed between Ferrymead and Bridge Street. These water levels gauges measure the water level at the mouth of the Heathcote and Avon River, and are located at opposite sides of the estuary. Because the estuary is relatively shallow, the wind could cause a small slope on the water level, which can be measured from the difference between the tide gauges. Other physical factors like the river discharge could also be affect the measures water levels.

C.1 Method and data

In this analysis we try to explain the observed water level differences between the two gauges. To do so, we follow the next steps:

1. Selecting peak tide values

The tidal pattern between Ferrymead and Bridge Street differs, particularly during low water. This is due to shallow water effects. The high water seem to be consistent between the location, so they are a good source of comparison.

2. Peak tide differences

By plotting the differences between the peak tide values we can visually distinguish some effects, and maybe explain them as well.

3. Comparing with other data

The next step is trying to explain the differences based on the wind and discharge data. We're trying to explain the variation. For this we use an ordinary least squares (OLS) model. The advantage of this method is that it shows the coefficient of each factor and the coefficient of determination, R^2 , that can be interpreted as the fraction of the variance that is explained by the model.

During this project, HKV and GHD notified Council that there were apparent inconsistencies between the sea level measurements at Bridge St and Ferrymead sites. During high tides, water level at the two sites is expected to be similar (generally considered to be within 25mm) other than during intermittent periods of wind setup. The sea level data supplied for these sites contained numerous departures from this expectation. Council then undertook a data review process and estimated a series of corrections to the data as shown in the next two tables. A part of this process used a synthetic tide peaks timeseries for Sumner Heads which was often useful to discriminate which site should be adjusted when a difference arose between the two sites. A reasonable level of care was applied by Council as indicated by the resolution both vertically and in timing of the adjustments. In some cases the adjustments were clearly motivated and relatively self evident in time, but in some cases the data wandered slowly and in order to use simple time specific adjustments it was necessary to pick somewhat arbitrary timings near the midpoint of a loosely defined range of time. Attempts to connect the adjustments with known historic events from NIWA site maintenance records, to understand and explain cause were found to be generally unproductive, with the exception of adjustments on 24/2/2011 and 13/6/2011 which coincided with major Christchurch earthquake events. GHD has reviewed the Council

assessment process and conclusions and so as to be satisfied that the conclusions reached are reasonable. But Council remain responsible for the detail of data and corrections.

Table C.1 and C.2 below show the data corrections applied to both time series. Note that the last record from the first table concerns a part of the data that is deleted. This is due to anomalies at the time due to bridge work. For the analysis in the appendix this does not matter much, since the disturbances are only of short duration. These corrections are however also applied to the data before doing the extreme value analysis described in Appendix E. For the EV analysis it is an important part of the time series, since the disturbances cause high peak values that might affect the analysis.

Table C.1: Applied corrections to water levels at Ferrymead

Start	End	Correction
1974-01-01 05:20	1981-04-01 12:00	-100 mm
1981-04-01 12:00	1983-01-01 00:00	-150 mm
1983-01-01 00:00	1994-11-23 00:00	-50 mm
1997-12-12 03:15	1998-06-25 06:00	-100 mm
1998-06-25 18:45	1998-07-18 00:30	-150 mm
1998-07-18 13:00	1998-12-31 16:00	-100 mm
1999-06-24 14:45	1999-08-01 20:15	-200 mm
2000-01-01 01:15	2004-04-08 20:30	25 mm
2004-05-10 10:15	2006-05-15 19:30	25 mm
2011-02-24 10:00	2011-06-13 02:45	-25 mm
2011-06-13 15:15	2012-05-28 23:30	50 mm
2012-05-29 12:15	2012-11-12 03:45	-25 mm
2012-11-12 15:45	2014-04-08 12:45	50 mm
2013-01-08 00:00	2013-01-10 00:00	<i>Removed</i>

Table C.2: Applied corrections to water levels at Bridge Street

Start	End	Correction
1998-09-11 09:30	1998-11-04 05:00	-50 mm
2011-02-24 10:00	2011-06-13 02:45	75 mm
2011-12-23 16:30	2012-05-28 23:30	25 mm
2013-10-08 20:15	2014-01-04 07:15	50 mm

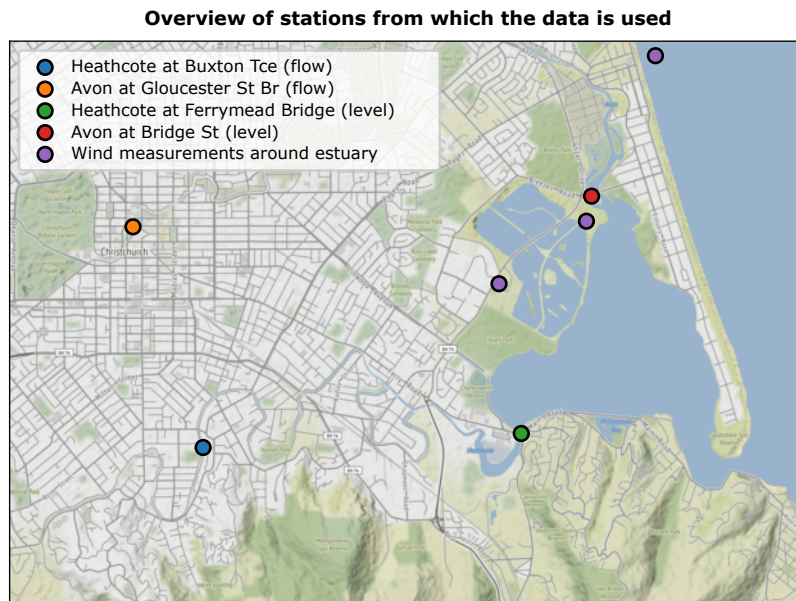


Figure C.1: Map of Christchurch with geographical location from the measurements.

In figure C.1 the location of each of the station is shown. Living Earth, located at the Eco Drop Metro transfer system, is the most southern of the three wind locations. We import the following data sources:

- Ferrymead water levels
- Bridge street water levels
- Heathcote discharges, measured at Buxton Terrace, where the influence of the tide is negligible.
- Avon discharges, measured at Gloucester Street, where the influence of the tide is negligible.
- Wind data: From the three stations the “Living Earth” is used. This station is located in between Bridge Street and Ferrymead so it gives the best estimate for the wind on the estuary. There is approximately 4 years of overlap between the wind data and the water level measurements, which is long enough to make an estimate of the relation between them.

The plot below shows the wind rose for the wind speed data at Living Earth. The ENE direction is most prevalent, followed by the WSW direction. The SSW direction however shows the largest wind speeds, during the 5 years of measured data.

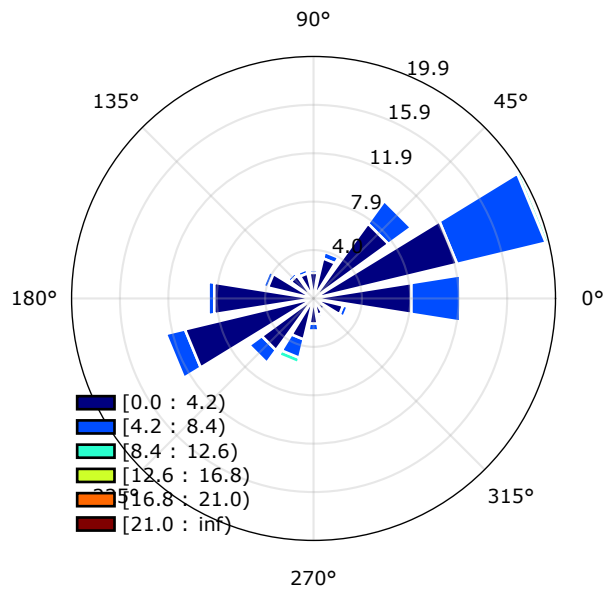


Figure C.2: Wind rose for Christchurch wind measured at Living Earth.

C.2 Peak tide selection

Since both measurement series have different tidal components, mostly due to the shallow water effects, it makes most sense to compare the high water levels. To do so, we determine the maximum value of the measurements, with should be the peak tide. The figures below show the results for Ferrymead and Bridge Street:

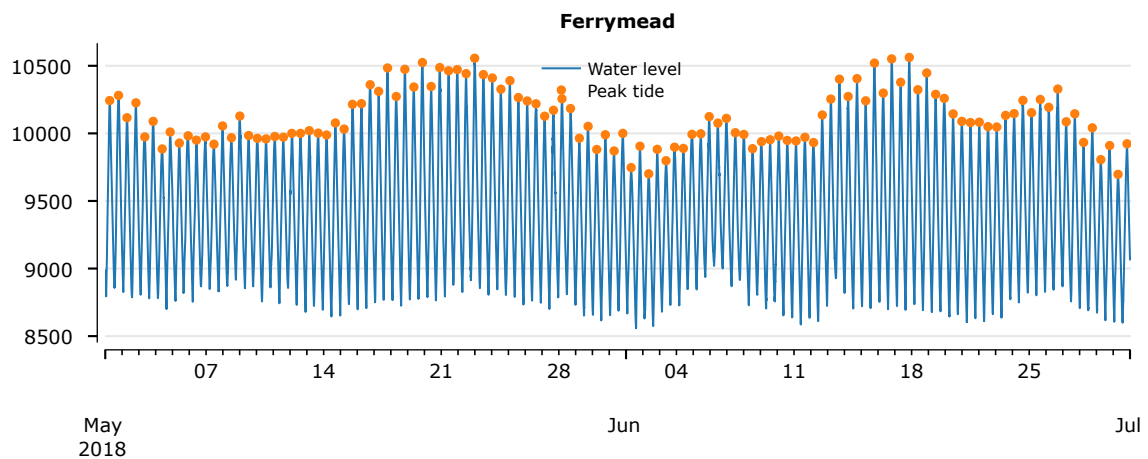


Figure C.3: Part of measured time series at Ferrymead with high tide values indicated.

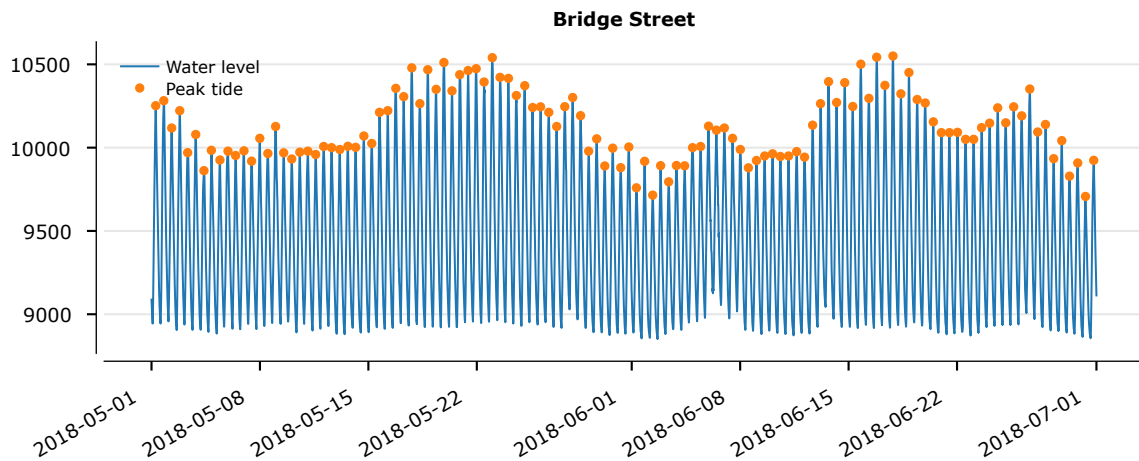


Figure C.4: Part of measured time series at Bridge Street with high tide values indicated.

The two figures look quite alike each other for the top part. At the bottom Bridge Street seems to be cut off, which is the shallow water effect.

The peak tide values are mostly not observed at exactly the same moment. We search the nearest peak tide for Ferrymead within a two hour range around Bridge Street. This should result in a match between peak tide values.

C.3 Differences between peak tides

The differences between the high tide values are shown in figure C.5. Also the 30 day moving averages are shown with the gray and black line. Before correcting the data, there used to be periods with a consistent shift in difference. After correction the difference is mostly gone: the gray and black line are more or less on top of each other. There is still some oscillatory motion visible, but it happens at both stations, and will therefore not affect the water level *difference* that we analyse.

The reason that the average water level is not constant is the reason for comparing the peak tides, and not comparing the surges at every moment of the tide cycle such as low tide or in between. For this we would have to model the tide and subtract it to get non-tidal effects, the surge.

A 30-day moving average is calculated to take into account some remaining inconsistencies in the measurements. After correcting the data this is no longer essential, but we still do it to take into account the remaining differences. The red line in the lower Figure C.5 shows the 30-day moving average.



Figure C.5: Above: Moving average of high tide values with a monthly (30 day) moving window. Below: The difference between the high tide values with a moving average of this difference.

C.4 Explaining the variation

So part of the variation in water level can be explained by these variations with a characteristic duration of months. By subtracting this moving average we will try to explain the remaining variation. For this we add these components in an Ordinary Least Squares model, from which we can determine the coefficient of their contribution, and the total explained variation.

An equation for the wind set-up on a lake is (assuming we the estuary is reasonably well modeled as a lake):

$$\Delta h = 0,5\kappa \frac{u^2}{gh} F \cos(\phi) \quad (8)$$

In which κ is a constant for the drag resistance, u is de wind speed, g the constant of gravitational acceleration, h the water depth, F the fetch and ϕ the angle between wind and direction in which

the fetch is measured. We don't use the formula directly (it doesn't seem to give that good result in this case) but use it to indicate important components:

- Wind speed and wind direction. Looking at the position of the water level gauges we expect a dominant influence of the wind in the South to North direction. We assume a quadratic effect of the wind speed.
- The average high tide level for Ferrymead and Bridge Street. The effect of the wind on the estuary is higher for shallow water (the h in the formula).

The other components are implicitly included in the least squares model, by the constant or fitted coefficients.

Additionally the discharge of Avon and Heathcote River are added. High discharges will give a higher water level at the mouth. Since we calculated the difference as Bridge St. minus Ferrymead/Buxton, we expect a positive coefficient for the Avon River, and a negative coefficient for Heathcote River. The relation between discharge and water level is not linear but rather quadratic, so we fit the square root of the observed discharges. This mathematical background for this is a standard rating curve relation.

C.5 OLS model

This results in the following model for estimating the surge:

$$\Delta h = a + b_1 u_{main} |u_{main}| + b_2 u_{lat} |u_{lat}| + b_3 \sqrt{Q_{Avon}} + b_4 \sqrt{Q_{Heathcote}} + b_5 \frac{(h_{BrSt} + h_{Ferrymead})}{2} \quad (9)$$

In which Q represent the Avon and Heathcote discharges. a is the constant, which is added because h_{BrSt} , $h_{Ferrymead}$ are water levels, and not water depths. The set-up is affected by depth, not level. The notation $u_{main} |u_{main}|$ instead of u_{main}^2 is so that the direction can still be negative.

In this model we need the main and lateral component of the wind. We want to choose this direction such that the lateral component has coefficient zero. By choosing this direction we get the best quantification of the effect of the main direction on the water level difference. Note that this direction does not affect the goodness of fit. For each direction two components (b_1 and b_2) can be found such that the difference is described optimally.

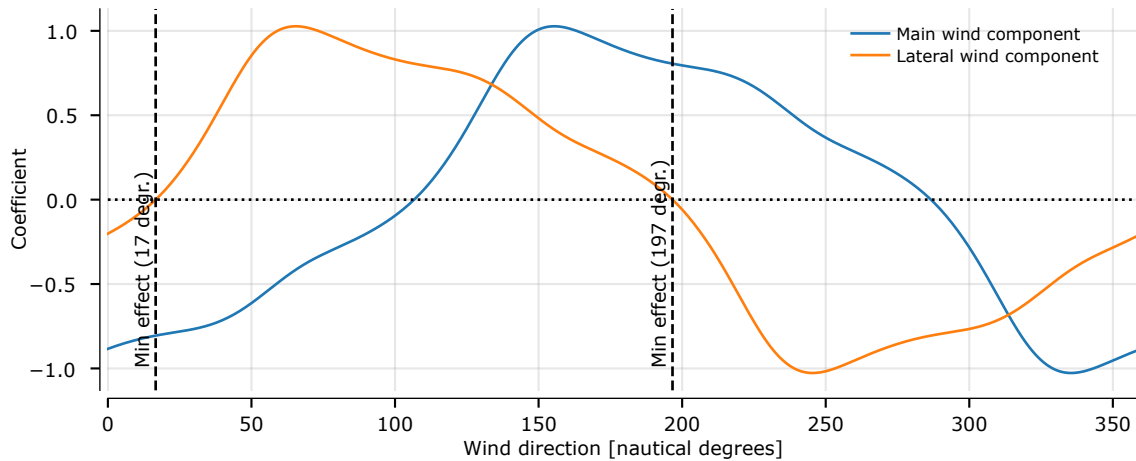


Figure C.6: Fitted main and lateral wind component per wind direction.

We see that the minimal coefficient of the lateral component is at 16 and 196 degrees from North. The coefficient of the main wind component is not maximal for this direction, which might seem counterintuitive. However this doesn't have to be the case, it only means that the water level difference is highest in comparison to the *combined* main and lateral wind speed for a direction of 155 degrees.

The wind direction 196 degrees (approx. SSW) we apply to the main dataset, to convert u_x and u_y to the main and perpendicular component. With this we fit the OLS model.

C.6 Results

The results from the Ordinary Least Squares model is shown in the three tables below. The first table presents the general model characteristics. The second table is most important, it gives the coefficients that result from the model fit. In the third table some statistical tests to describe the model quality are shown.

Table C.3: OLS Regression Results

Dep. Variable:	y	R-squared:	0.455			
Model:	OLS	Adj. R-squared:	0.453			
Method:	Least Squares	F-statistic:	363.4			
Date:	Fri, 11 Dec 2020	Prob (F-statistic):	8.59e-284			
Time:	17:41:06	Log-Likelihood:	-8927.1			
No. Observations:	2186	AIC:	1.787e+04			
Df Residuals:	2180	BIC:	1.790e+04			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P > t 	[0.025	0.975]
const	-164.5819	16.513	-9.967	0.000	-196.965	-132.199
$\sqrt{Q_{Avon}}$	1.7851	3.047	0.586	0.558	-4.190	7.760
$\sqrt{Q_{Heathcote}}$	-8.5357	3.484	-2.450	0.014	-15.368	-1.704
h	0.0168	0.002	10.286	0.000	0.014	0.020
$u_{perp} u_{perp} $	7.444e-05	0.025	0.003	0.998	-0.048	0.048
$u_{main} u_{main} $	0.8055	0.022	37.252	0.000	0.763	0.848
Omnibus:	562.044	Durbin-Watson:	1.624			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	7901.772			
Skew:	-0.813	Prob(JB):	0.00			
Kurtosis:	12.171	Cond. No.	5.45e+05			

Next we remove the insignificant components. Note that we also remove 9.8 meters from the average level h , since the wind set-up is affected by the water *depth* not the water *level*. This 9.8 meter seems to be the value for which the **const** becomes 0, so we can remove that too.

Table C.4: OLS Regression Results

Dep. Variable:	y	R-squared (uncentered):	0.455			
Model:	OLS	Adj. R-squared (uncentered):	0.454			
Method:	Least Squares	F-statistic:	606.8			
Date:	Fri, 11 Dec 2020	Prob (F-statistic):	8.44e-287			
Time:	17:41:06	Log-Likelihood:	-8927.3			
No. Observations:	2186	AIC:	1.786e+04			
Df Residuals:	2183	BIC:	1.788e+04			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P > t 	[0.025	0.975]
$\sqrt{Q_{Heathcote}}$	-6.5799	0.669	-9.830	0.000	-7.893	-5.267
h	0.0168	0.001	11.627	0.000	0.014	0.020
$u_{main} u_{main} $	0.8062	0.020	40.400	0.000	0.767	0.845
Omnibus:	562.572	Durbin-Watson:	1.626			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	7934.448			
Skew:	-0.814	Prob(JB):	0.00			
Kurtosis:	12.190	Cond. No.	867.			

The model returns significant factors for all components except the lateral component of the wind speed, and the Avon River discharge. The insignificance of the lateral wind component agrees with our expectation, since that is how we chose the direction. The fact that the Avon discharge is not statistically significant while the Heathcote discharge is, is more surprising since a first guess would be to assign both variables equal value. Perhaps the Gloucester Street gauge is too far upstream to give a good estimate of higher water levels all the way down at Bridge Street.

In total half the variation (46%) is explained by the factors we included, which is good for the simple model we made and considering the shifts in the water level measurements. The main wind component $u_{main} |u_{main}|$ is most important, which can be seen from the high value t in the second table (it indicates the number of standard deviations the coefficient is from 0). The high tide level h and the Heathcote discharge are less important.

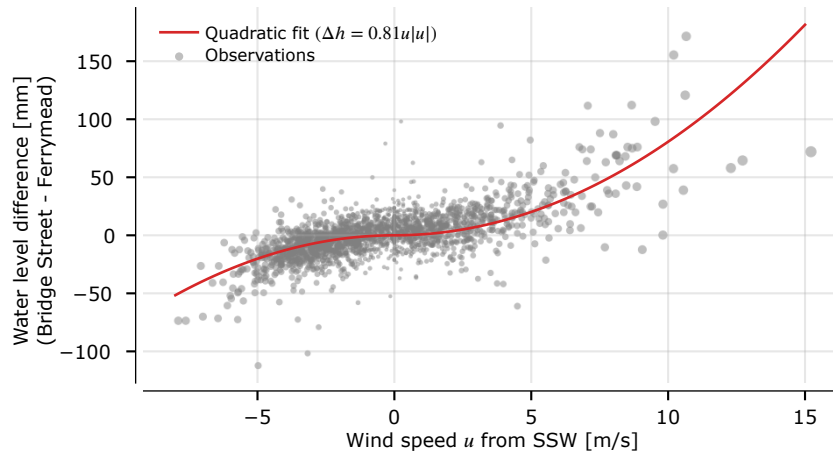


Figure C.7: Quadratic fit of the wind speed effect through the observed water level differences.

From this graph we read that we can expect a 81 mm water level difference in case of a 10 meters per seconds wind from SSW. By adding the other coefficients, such as the Heathcote discharge and the high water level, the model improves a bit, but the main factor is the wind speed in SSW direction.

To visualize the performance of the model the next figures show:

1. The observed and predicted water level
2. Similar, but only one year so the differences are easier to distinguish
3. The residuals, so the unexplained part of the water levels (54%).

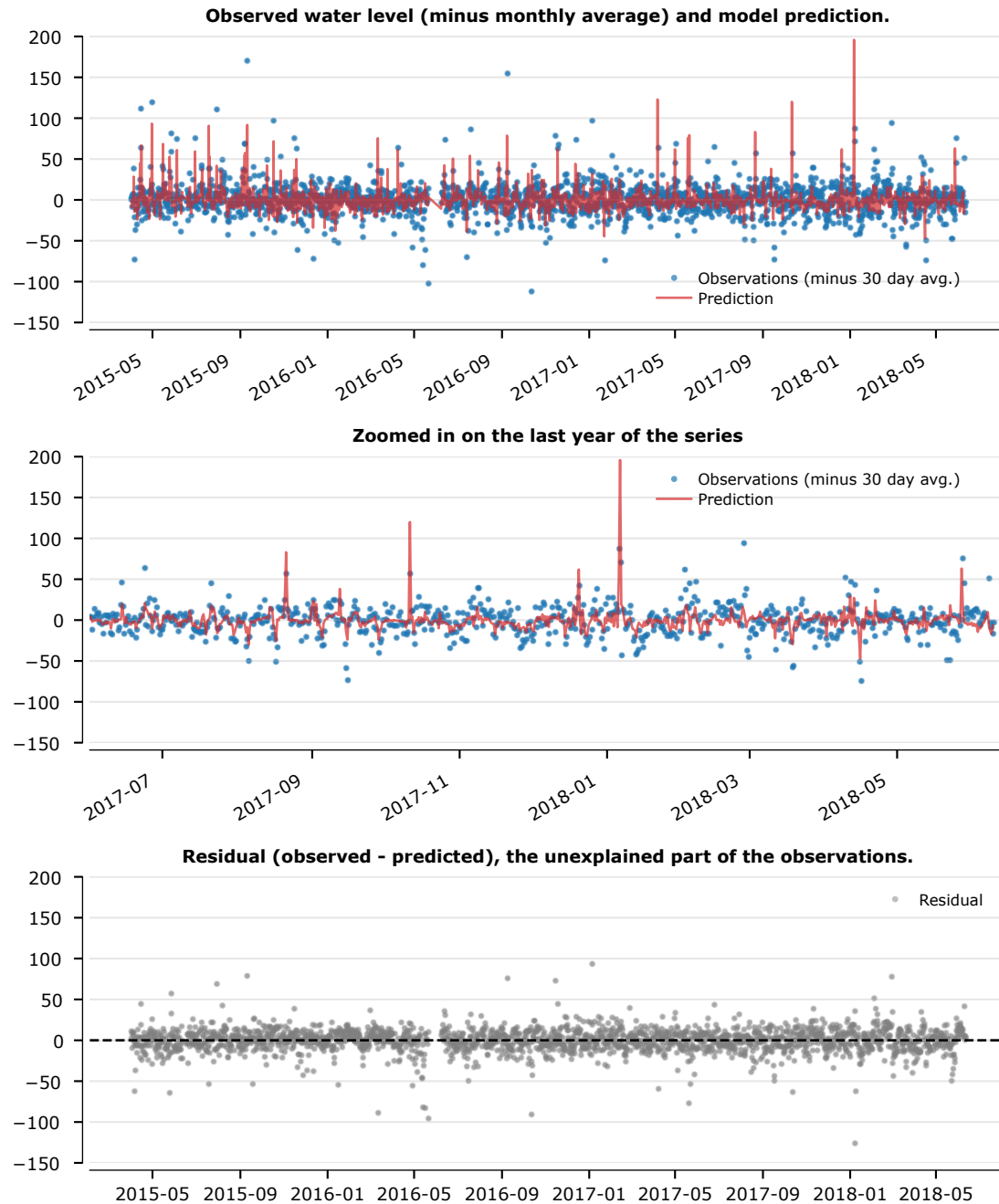


Figure C.8: Observed water levels and predicted water levels with OLS model. The plot below shows the difference between the two, the residu.

By looking at the water level peaks, we see some are overestimated by the quadratic wind speeds while others are underestimated. Apparently wind, water level and Heathcote discharge are only part of the explanation. These three components however give a good indication of the water level difference. If the jumps in the measured water level differences would be explained better, the results might be better. It is however unlikely that the conclusions would be different. The same applies to using longer data series.

C.7 Conclusions

- We find that the variation in water level differences can for 46% be explained by the wind, water level and the Heathcote discharge. The SSW component of the wind fits best to the water level differences. This is also approximately the orientation of the two water level gauges. The estuary water level (depth) and Heathcote discharge are a significant coefficient in the OLS fit of the water levels, but they reduce the unexplained variation only marginally.
- The raw water level measurements for both Bridge St. and Ferrymead contain errors that result in a consistent increase or decrease of the water levels for time periods with a length in the order of months. The most obvious of these errors were removed manually based on the observed differences between high tide values. Additionally the data are corrected with a 30-day moving average with which we aim to remove the remaining differences. Note that this is not strictly necessary after manually correction the data, but is done anyway since it still increases the quality of the fit a bit. It is however unlikely that further data correction will the explained variance.

D Rainfall data

This appendix describes the available rainfall data for the Avon catchment and how to process these data. With processing is meant filling of gaps and aggregating the different stations into a single time series. Note that we speak about rain throughout the appendix for convenience, but actually we're talking about precipitation.

The rainfall series do not all span the full period from 1960 to now, which is the period we're interested in. To solve this we can fill the missing data with rain data from adjacent stations. For this we determine a scale factor, and fill the missing value by filling with data from a near station with this determined multiplication factor.

For one of the rain gauges we only have daily data. To use this data in a similar manner as the data from the other stations, the data needs to be converted to 24 hour data. Here we can again use the nearby stations, by spreading the measured daily rainfall over these 24 hours.

D.1 Catchment and rain gauges

The figure below shows the Avon catchment boundaries (red) and the rain gauges around it (blue). The Voronoi diagram around the rain gauges is also shown. We're interested in the rain gauges that represent an area that intersects the red outlined area.

For this we select the following stations:

- Firestone Factory, Papanui
- Christchurch Botanical Gardens
- Avon at PS205
- Christchurch Aero
- College of Education
- Styx at Lower Styx Road
- Shirley rain gauge

Tunnel Road also crosses part of the catchment, but none of it is area that will lead to runoff on the Avon river when it rains.

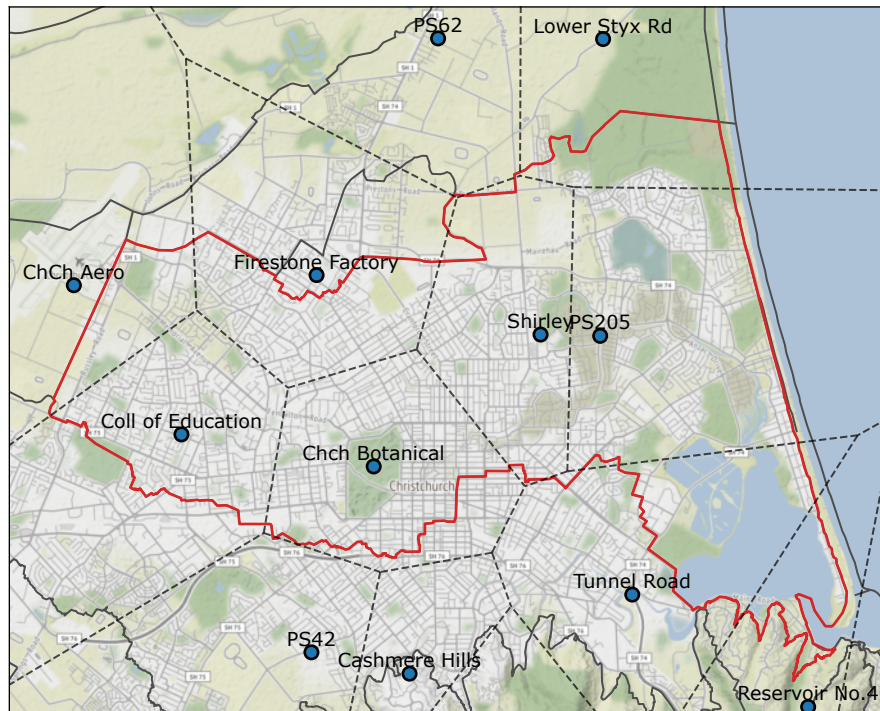


Figure D.1: Rain gauges around Christchurch for which historic time series are available. The red lines indicates the edge of the Avon catchment.

D.2 Data aggregation

To get rainfall data for the full period and for each station, we need to do two processing actions. First we need to generate hourly values for Shirley which is available as daily rain depth values. Then we fill the missing values for all stations based on the ratio of rain intensity in the year.

The next figure shows the average rain intensity per day for each year. We calculate the intensity per day, and not per year, so we can take into account the missing values. The ratio between the different station in the figure are used for scaling missing data between stations.

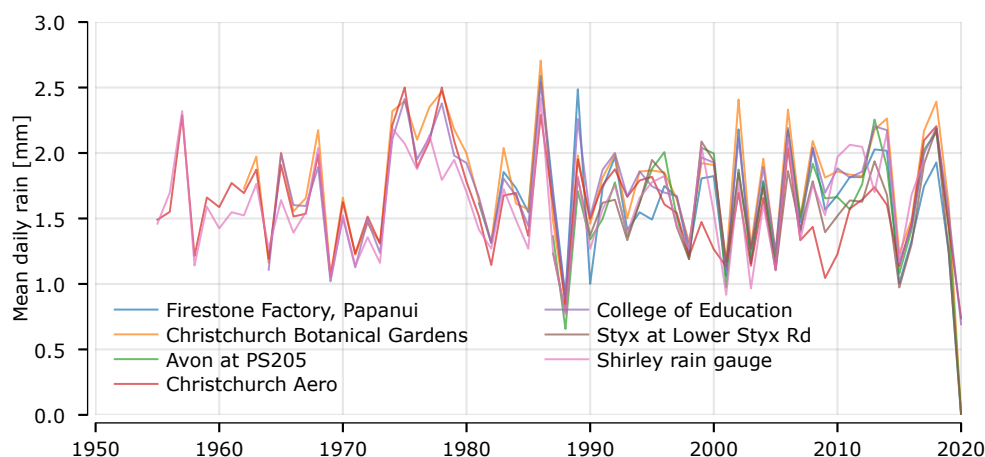


Figure D.2: Mean daily rainfall for different stations and all years for which the source data is available.

Over the full period the average rain intensity per day is given in the table below:

Daily average rainfall [mm]	
Firestone Factory, Papanui	1.650
Christchurch Botanical Gardens	1.789
Avon at PS205	1.621
Christchurch Aero	1.617
College of Education	1.730
Styx at Lower Styx Rd	1.559
Shirley rain gauge	1.595

There is little variation, max 10 %, but still significant given the small spatial scale. This variation can just as well be caused by missing (dry or wet) years as actual differences in intensity. To investigate this further, we calculate the ratio for overlapping periods.

	Firestone Factory	Chch Bot.	PS205	ChCh Aero	Coll of Educ.	Lower Styx Rd	Shirley
Firestone Factory		0.930	1.000	1.071	0.954	1.044	1.040
Chch Botanical	1.075		1.087	1.106	1.031	1.129	1.159
PS205	1.000	0.920		1.055	0.935	1.038	1.097
ChCh Aero	0.933	0.904	0.948		0.928	0.984	1.045
Coll of Education	1.048	0.970	1.070	1.078		1.110	1.128
Lower Styx Rd	0.958	0.886	0.963	1.017	0.901		1.057
Shirley	0.962	0.863	0.911	0.957	0.887	0.946	

A factor of 1.07 means that the station on the index has a 1.07 times as high rainfall for the overlapping period as the station on the column. To get the rain intensity for the index name,

multiply the other stations series with this value.

We note the following:

- Christchurch Aero and Shirley have on average a lower rain intensity than the other station.
- Christchurch Botanical Gardens has the highest rain intensity.

Geographically these differences are hard to explain. For example Shirley and PS205 are located very close to each other. We will however use them in scaling the data.

Generating hourly values from daily shirley values

To generate hourly values for Shirley, we follow the next steps:

1. For every other station, from close to far away:
 - Avon at PS205
 - Christchurch Botanical Gardens
 - Firestone Factory, Papanui
 - Styx at Lower Styx Rd
 - College of Education
 - Christchurch Aero
2. Select the part where the other station has data, and fill in. So everything from PS205, the remaining missing part from Botanical Gardens, than again the missing part from Firestone Factory, etc.
3. Determine the ratio between daily rainfall for the two stations. This is the factor with which the pattern from the last step is multiplied.
4. Fill the hourly rainfall for Shirley by multiplying the pattern with the factors.

The following figure shows the year sum for Shirley, with 24 hour values (blue) and hourly values (orange). In the early years the lines are on top of each other. In the later years there is a difference, which is because the daily values have some missing days, which will lead to a lower sum. This is compensated when creating the hourly pattern.

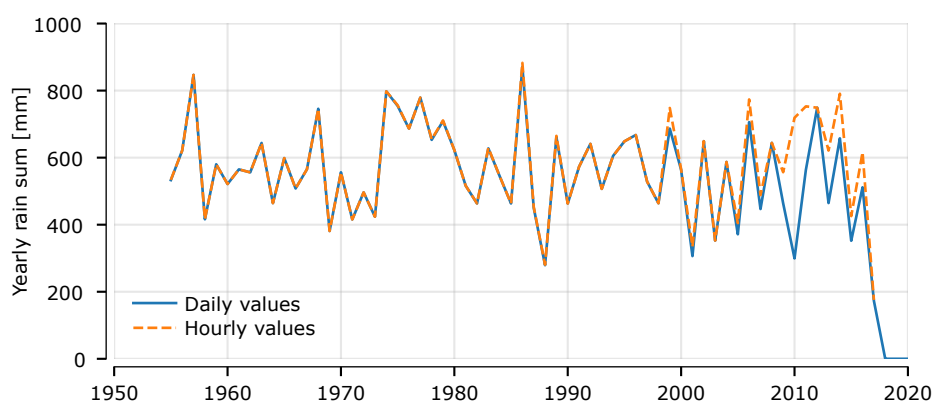


Figure D.3: Yearly rain sum from daily and hourly values for Shirley. The difference is the result from filling the gaps within the data aggregation process.

Filling the missing values

Now we fill the missing values for the other stations:

1. For every other station, from close to far away:
2. Select the part where the other station has data, and fill in.
3. Determine the ratio between daily rainfall for the overlapping period between the two stations. This is the factor with which the pattern from the last step is multiplied.

The table below shows the average rain depth per day before and after hourly values for Shirley and filling the NaN values.

	With daily Shirley	Before filling missing	After filling missing
Firestone Factory, Papanui	1.650	1.650	1.658
Christchurch Botanical Gardens	1.789	1.789	1.782
Avon at PS205	1.621	1.621	1.691
Christchurch Aero	1.617	1.617	1.616
College of Education	1.730	1.731	1.742
Styx at Lower Styx Rd	1.559	1.559	1.629
Shirley rain gauge	1.595	1.601	1.600

We see that the average rain intensity for Avon at PS205 has increased a few per cents. The nearest station for PS205 is Shirley, so this means that the early period data have been derived from Shirley.

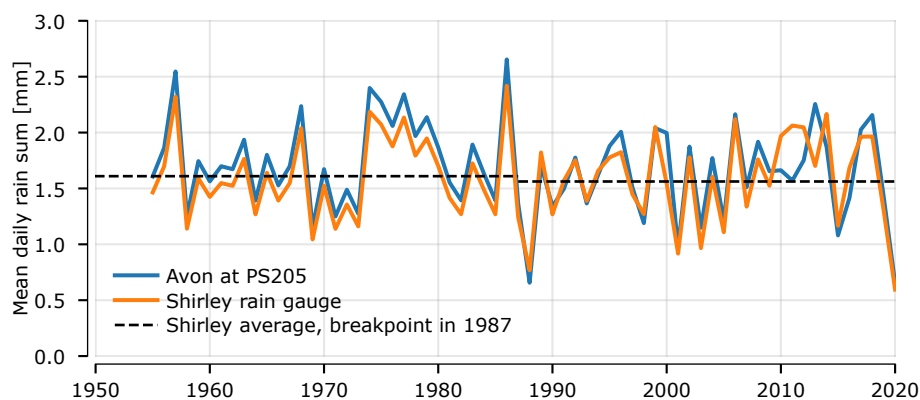


Figure D.4: Daily mean rain sum for Shirley and Avon at PS205

Apparently the average rainfall per hour was higher before 1987, which leads to higher values for PS205 after adding the pre 1987 values. The difference is not significant, so there is no reason to revise our approach in which we assume the rainfall intensity to be stationary over the years.

Creating a weighted average

To combine the different rainfall pattern into one rain series for the Avon catchment. We calculate the intersecting area between the Thiessen polygons around each station and the Avon catchment boundaries. The different overlapping areas with their percentage of coverage are shown in the figure below.

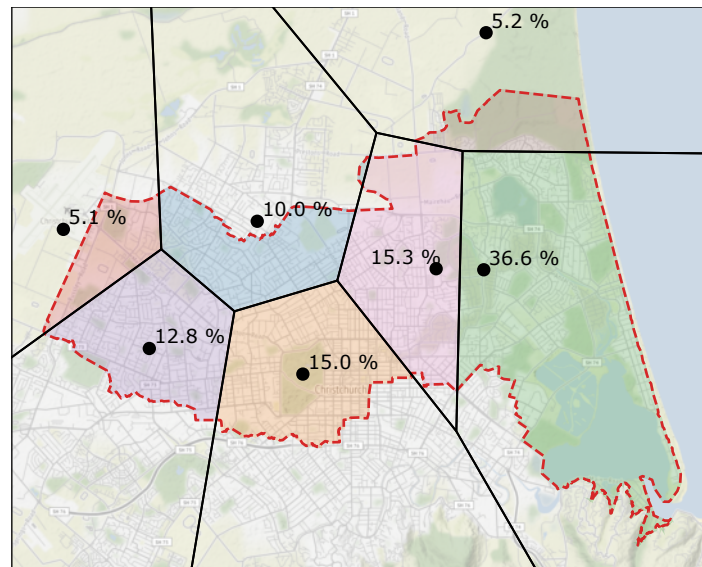


Figure D.5: Overview of catchment with the different used rain gauges and the percentage of the catchment they cover.

These percentages are used for weighing the patterns of each station.

D.3 HIRDS tables and formulas

HIRDS (High Intensity Rainfall System) contains functions for calculating the rain depth (or intensity) given a duration. The derived statistical relations are useful, since they contain formulas for calculating the expected rainfall depth for a X hour duration event with an Y ARI. These relations can be used to determine the relation between depth/intensity and ARI. However 7 gauges give 7 relations:

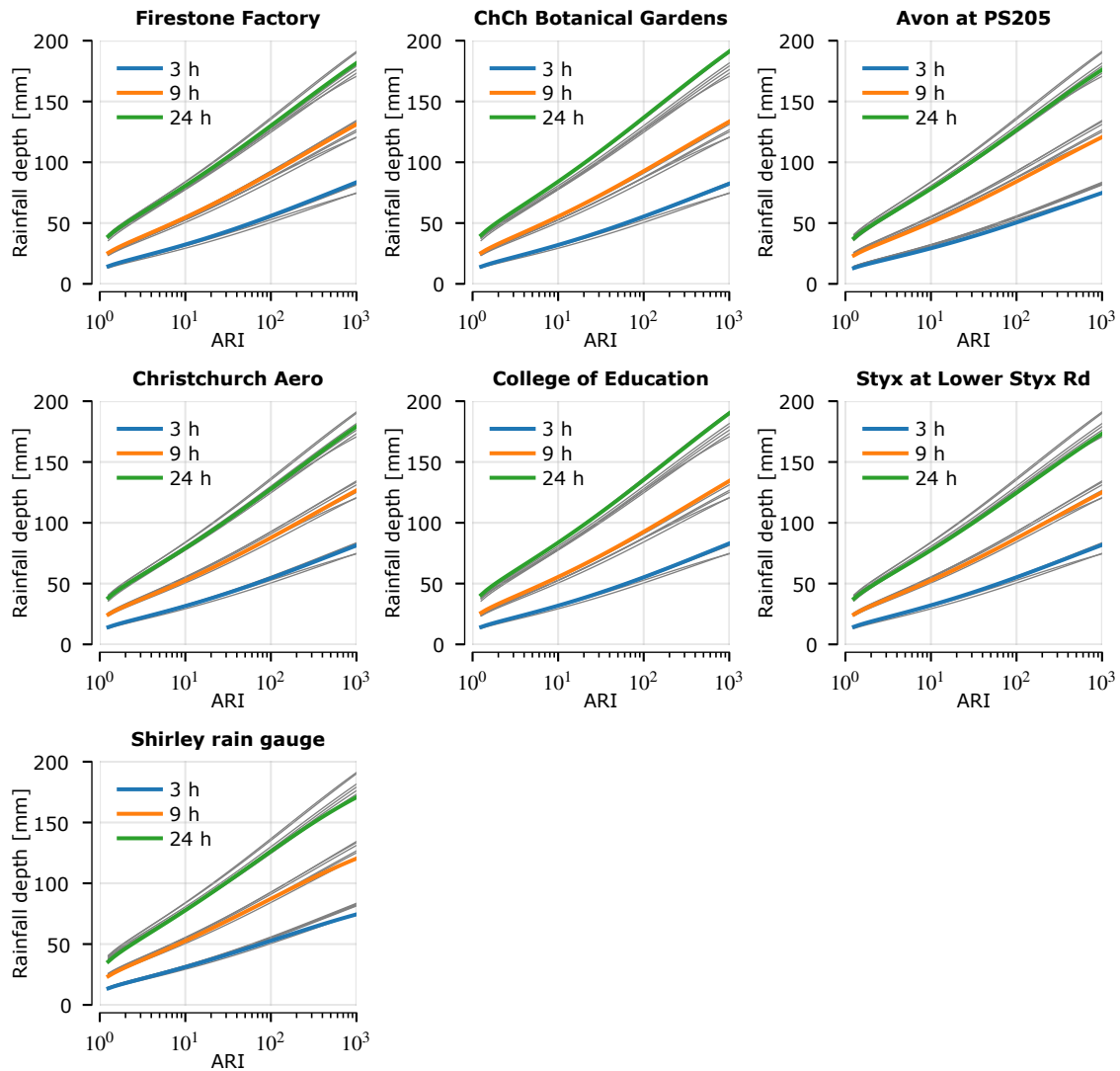


Figure D.6: HIRDS rainfall statistics for the different gauges on the Avon catchment. The grey lines on the background are the statistics from the other gauges.

The variation between different depths for certain return periods is not too large, around 10% between the lowest and highest curve. To get a single 100 year ARI intensity/depth for a certain duration, we advice to use curve that gives the highest value. This is because the scenario's also have a standard error, which is not taken into account explicitly. This standard error would likely (assuming the errors are somewhat symmetric) lead to more conservative values if integrated. Therefore we advice to use the maximum statistic, which is Christchurch Botanical Gardens.

E Extreme value analysis SWL around Christchurch

E.1 Structure notebook

This appendix describes the extreme value analysis of sea water levels at stations near Christchurch. The structure of this notebook is as follows:

- Section E.2 provides an overview of the data used in this notebook
- Section E.3 describes the comparison with Goring's study. We reproduce the Extreme value analysis of Goring to make sure we use a similar basis for this Extreme value analysis. These analyses are not corrected for sea level rise. The subsequent analysis in this notebook does correct for sea level rise.
- Section E.4 to section E.5 substantiates some method choices, such as the sea level rise correction, confidence bounds, the difference between Peaks-Over-Threshold and Annual Maxima and the choice of the fitted distribution.
- Section E.7 shows the resulting fitted extreme value distributions.
- Section E.8 to section E.10 contain some sensitivity analysis carried out to investigate the reliability of the results.
- Finally section E.11 summarizes the conclusions and recommendations.

E.2 Data used

We consider five stations located near Christchurch. These are the four stations evaluated by Goring, and additionally Lyttelton too. We analyse Lyttelton because it has the longest time series of measurements. Usage of these longer time series reduces the uncertainties in the analysis.

- For Sumner and Lyttelton the data described in Appendix A are used.
- We used the measurement series received from NIWA for Bridge Street, Ferrymead Bridge and Styx which consists of water level measurements with an interval of 15 minutes except for Ferrymead. At Ferrymead we have instantaneous data, which consist of a combination of 15 and 5 minute data. We used the 5 minute data when it was available as this seemed to correlate better with Goring's data.
- The datum used for all sites (except for Lyttelton) was Christchurch Drainage Datum (CDD) which is 9.043 above Lyttelton Vertical Datum (1931). The measured water levels at Lyttelton are converted from LVD37 to CDD.
- For Bridge Street and Ferrymead a number of corrections are done on the data, as listed in table C.1 and table C.2. These corrections are used throughout this appendix, except for section E.3. For clarity we will call both datasets the "corrected data" and "as supplied data", the last being the uncorrected original data.

The corresponding time series are shown below (as supplied data up to 1-1-2020 are used):

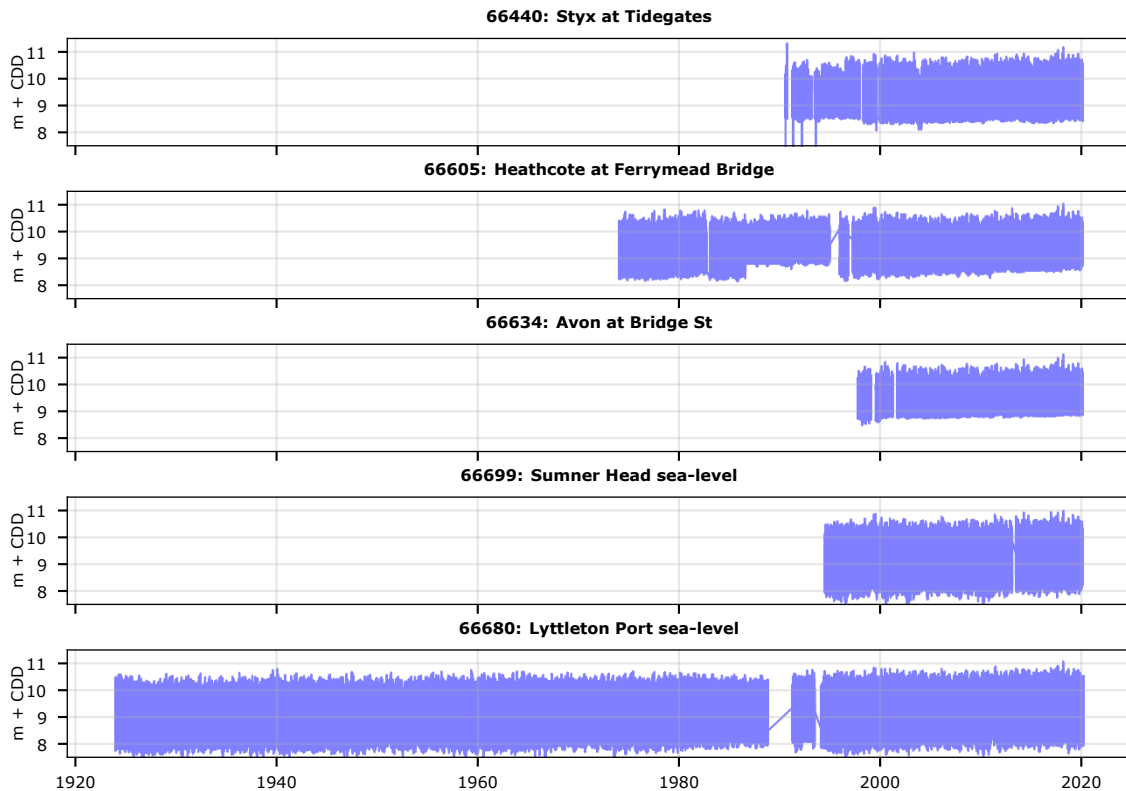


Figure E.1: Time series of measured water levels for the 5 gauges in this analysis.

E.3 Comparison with Goring's study

Extreme value theory

In estimating exceedance probabilities of events, the Annual Maxima / Generalized Extreme Value distribution (AM/GEV) and Peaks over Threshold / Generalized Pareto distribution (POT/GPD) approaches can be used. In this notebook we discuss both methods and derive extreme value statistics for five stations around Christchurch.

Annual Maxima: The extremal types theorem gives rise to the annual maxima (AM) method of modelling extremes, in which the GEV distribution is fitted to a sample of block maxima. The GEV distribution has three parameters: location, scale and shape parameters. Distinguish is made between three different types of GEV distribution, which is determined by the sign of the shape parameter; Gumbel distribution (shape param = 0), Fréchet distribution (shape parameter > 0) and Weibull distribution (shape param < 0).

Peaks-Over-Threshold: The POT approach fits the generalized Pareto distribution (GPD) to the peaks that exceed a chosen threshold. The GPD distributions are described by two parameters: the scale and shape parameters. When the scale parameter of the GPD is equal to zero, the GPD has a type I tail and is equal to the exponential distribution. This exponential tail corresponds to the tail of the Gumbel distribution in the Annual Maxima approach with GEV fit.

E.3.1 Plotting positions

In this study the Gringorten plotting positions are used to calculate the exceedance frequency of an observation. For annual maxima, the frequency is calculated as follows:

$$P_{exc} = \frac{(k - b)}{N + a} \quad (10)$$

In which a and b are the coefficients for calculating the plotting positions. For Gringorten these are $b = 0.44$ and $a = 0.12$. These numbers mostly affect the positions of the most extreme and least extreme events. k is the rank number and N the number of observations.

For Peaks-Over-Threshold the exceedance probability is calculated with the next formula:

$$P_{exc} = \frac{N * (k + a + b - 1)}{(N + a) * T} \quad (11)$$

Additionally T is the total period in which the observations are observed. The exceedance probability P_{exc} is related to a year in the formula for annual maxima, and to T/N (years) for the POT formula.

Analysis

First we reproduce the extreme value analysis of Goring (to the extent possible based on the correction listed in his report) to make sure we use a similar basis for this Extreme value analysis. Extreme value analysis were executed for four different stations around Christchurch: Styx, Ferrymead, Bridge St and Sumner. The Gumbel distribution is fitted to the annual maxima selected from the time series. The Gumbel distributions is used in accordance with the study of Goring (Extreme Sea Levels at Christchurch Sites: EV1 Analysis).

We made the following adjustments to the time series to correct for errors or improve the fitting process:

- Styx : We removed the lowest annual maxima which improves the fit at Styx (1993-08-19: 10.39 m+CDB). Compared to Goring, the received 15 minutes data is already corrected for an error in the extreme event in 2003 (this is in line with Goring). Furthermore, we removed two outliers (27-2-1998: 11.5 m+CDB and 1-9-1990: 11.29) which were also not found in Goring's study.
- Ferrymead: The measured water level from 8-1-2013 until 10-1-2013 are excluded in this analysis. In this period, the data show some anomalies which have probably to do with bridge building work. There was a temporary recorder on site at the time which apparently would have been more subject to such disturbances than the present permanent device. Note that Goring in his 2018 work adjusted this event 10.862 m+CBD event to 10.763 m+CBD without commenting on the adjustment, but we assess the measured data as too unreliable and that it is more appropriate to ignore this potential event.

The figures below shows the 1) the year maxima and their return periods, 2) the fitted Gumbel distribution and the fitted distribution of Goring. Next to each figure the results of Goring are shown. The following notes relate to the comparison between Goring's analysis and the analysis executed:

- The fits of the extreme value distributions in this study and Goring's result match almost exactly for two stations; Styx and Bridge street. There are small differences, which can be caused by different fitting procedures and slightly different data.
- At Ferrymead our finding is 10-20 mm higher than Goring's. This is likely due to data corrections which are not listed in Goring's report or that he may have obtained his data from a different source. The final statistics at Ferrymead may have the potential to be a little lower if Goring's data was used (the mentioned 10-20 mm). Figure E.3 in Appendix E shows the comparison, which is a good measure for the possible lowering. The Ferrymead annual maxima data Goring used in 2018 has been obtained and cross checked with the NIWA data supplied for this study. Only 20% of the data was found to match. In the remainder the NIWA data was higher than Derek's data in all instances, with differences ranging from 2 – 36mm and there was no evident time trend in the differences. Three specific years were more significant as follows;

1. 1999; 17/4/99 10.879 m, Derek's data was 11 mm lower (rank 3 event)
2. 1978; 24/6/78, 10.814 m, Derek's data was 24 mm lower on 20/7/1978 (rank 4 event)
3. 2013; 14/1/13 10.714 m, Derek's data was 10.763 m on 8/1/2013 (rank 7 event)

The average of all the differences (excluding 2013) was 10mm which is consistent with the gap observed between the current findings and Goring 2018. These checks confirm that the differences between Goring 2018 and NIWA 2020 maximum's data are minor and are not sufficiently material to provoke further interest or investigation and that confidence in use of the NIWA 2020 data is established.

- The frequency line derived at Sumner deviates significantly from Goring's fit. Goring EV at Sumner Heads appears to be inclusive of an allowance for FIG waves. We did not allow in our scope to assess FIG waves and our scope is to deliver EV analyses based on 15 minute average data which removes the FIG waves. As agreed with Council in discussions 30/6/2020 (Parsons / Harrington / Preston), we have added an allowance for FIG-waves to our final result in order to provide comparison with Goring's work.
- Council will need to consider whether Sumner extreme sea level values, with or without FIG waves are appropriate for future work.

Note that the results in this section are based on the data without the data corrections for Bridge Street and Ferrymead listed in appendix C.

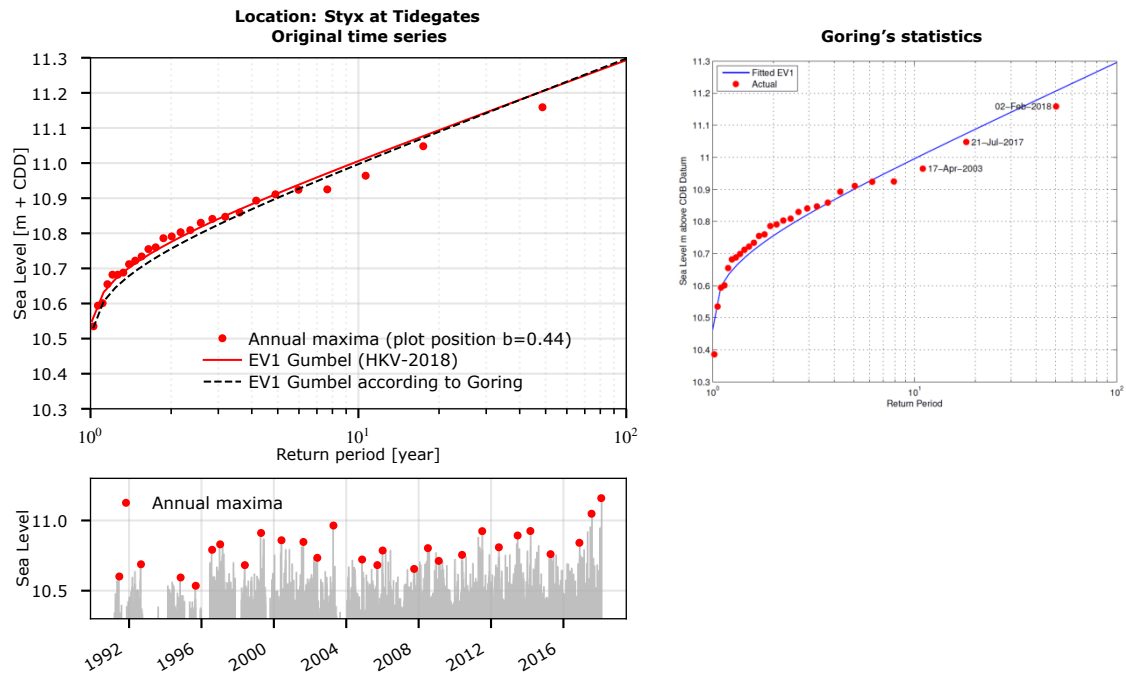


Figure E.2: EV1 (Gumbel) fits through annual maxima for Styx at Tidegates

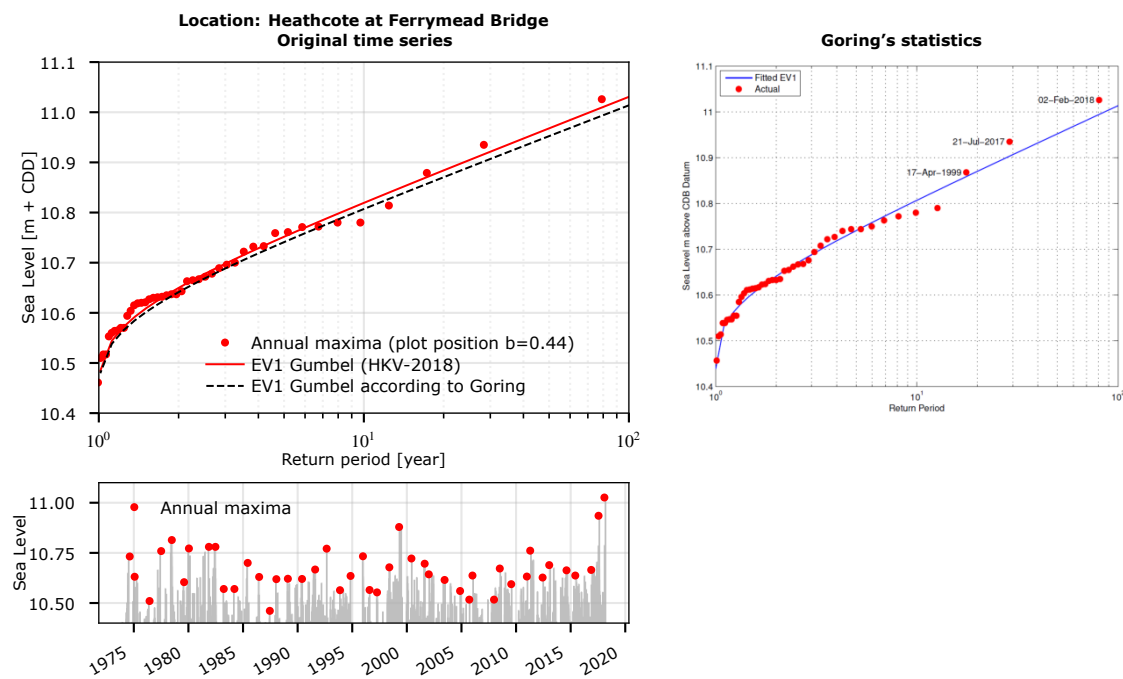


Figure E.3: EV1 (Gumbel) fits through annual maxima for Heathcote at Ferrymead Bridge

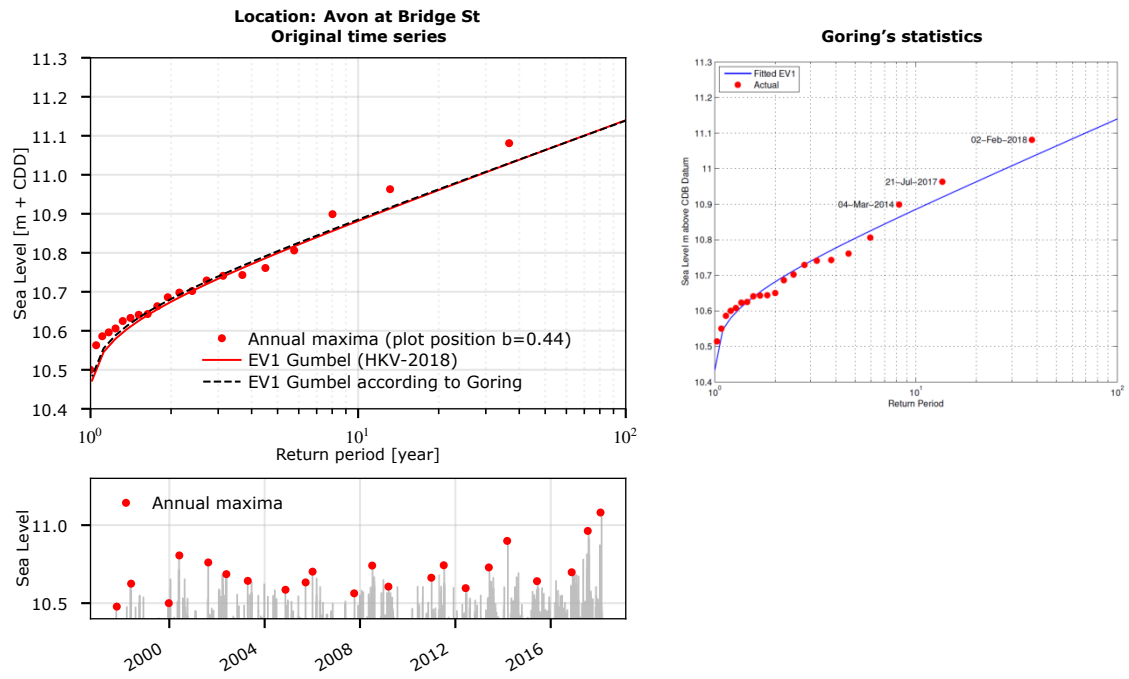


Figure E.4: EV1 (Gumbel) fits through annual maxima for Avon at Bridge Street

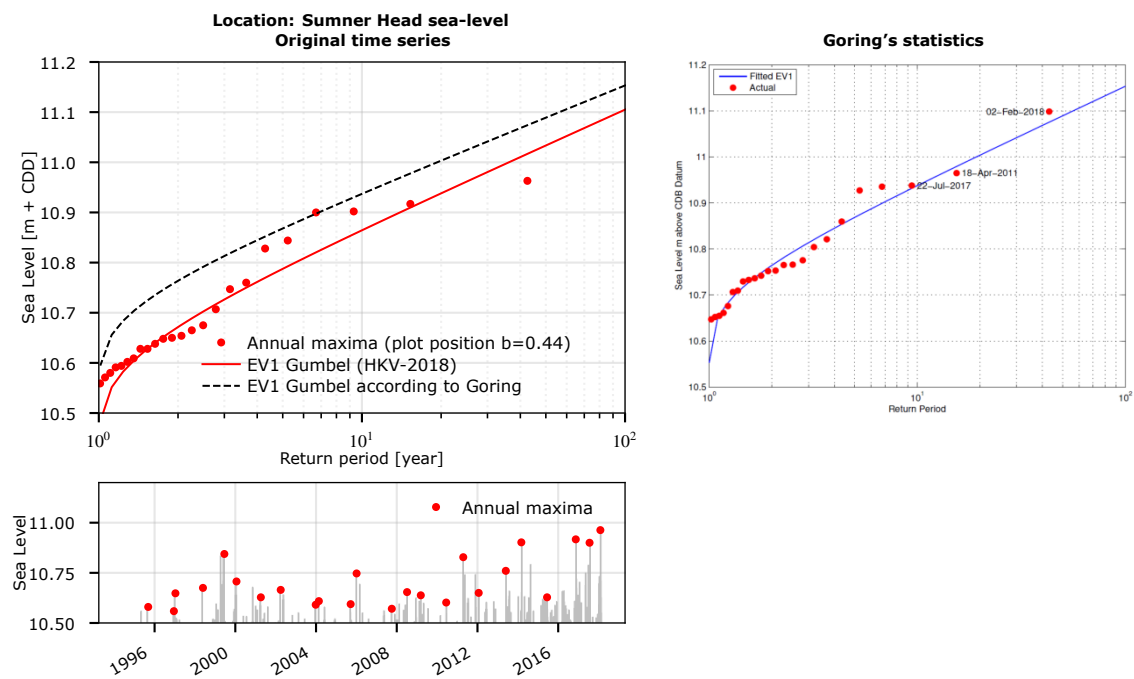


Figure E.5: EV1 (Gumbel) fits through annual maxima for Styx at Sumner Head

Using corrected data

Note that from this point on, we will shift our attention away from matching Goring's results, and towards establishing a new best estimate. Accordingly the used data change from the "as supplied" data to the corrected data hereafter.

Highest annual maxima events

The twelve highest peaks are shown for every station in the figures below.

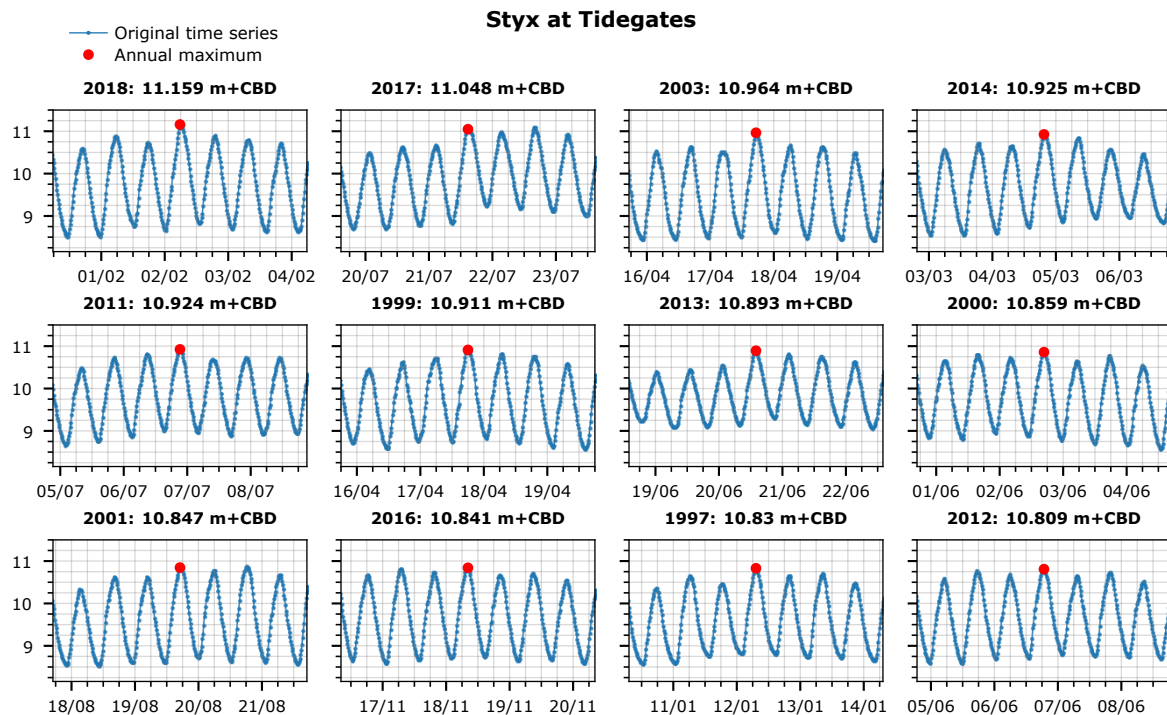


Figure E.6: Time series (corrected data) during annual maximum for Styx at Tidegates

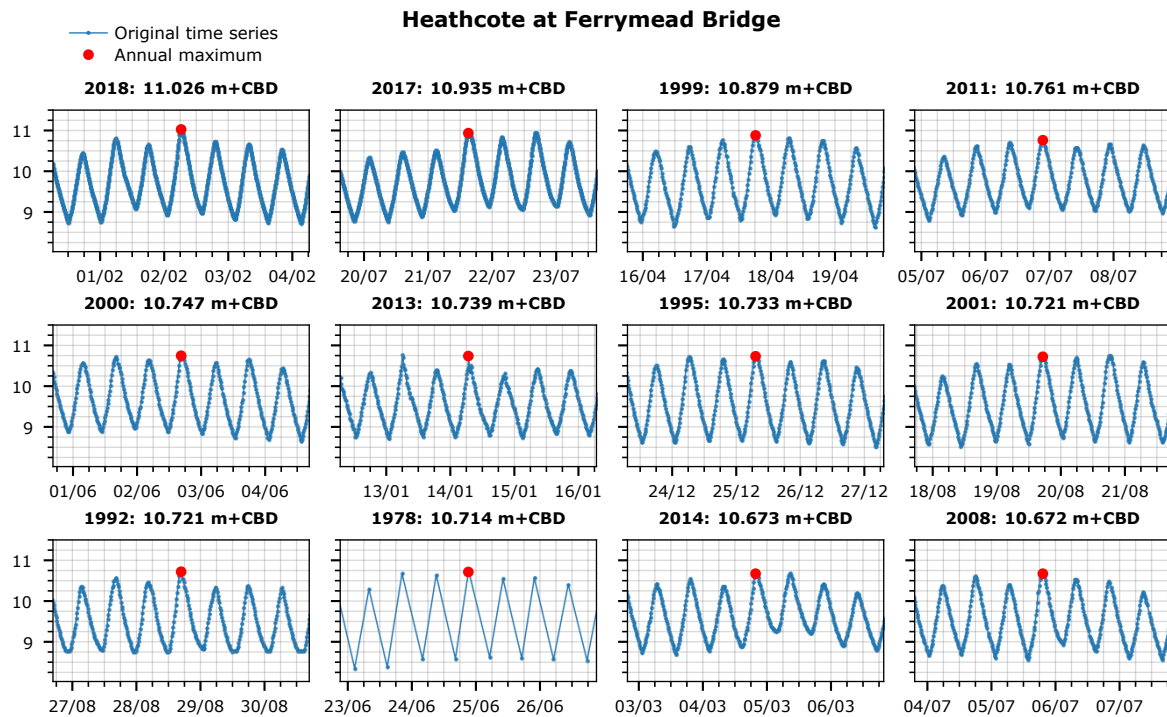


Figure E.7: Time series (corrected data) during annual maximum for Heathcote at Ferrymead Bridge

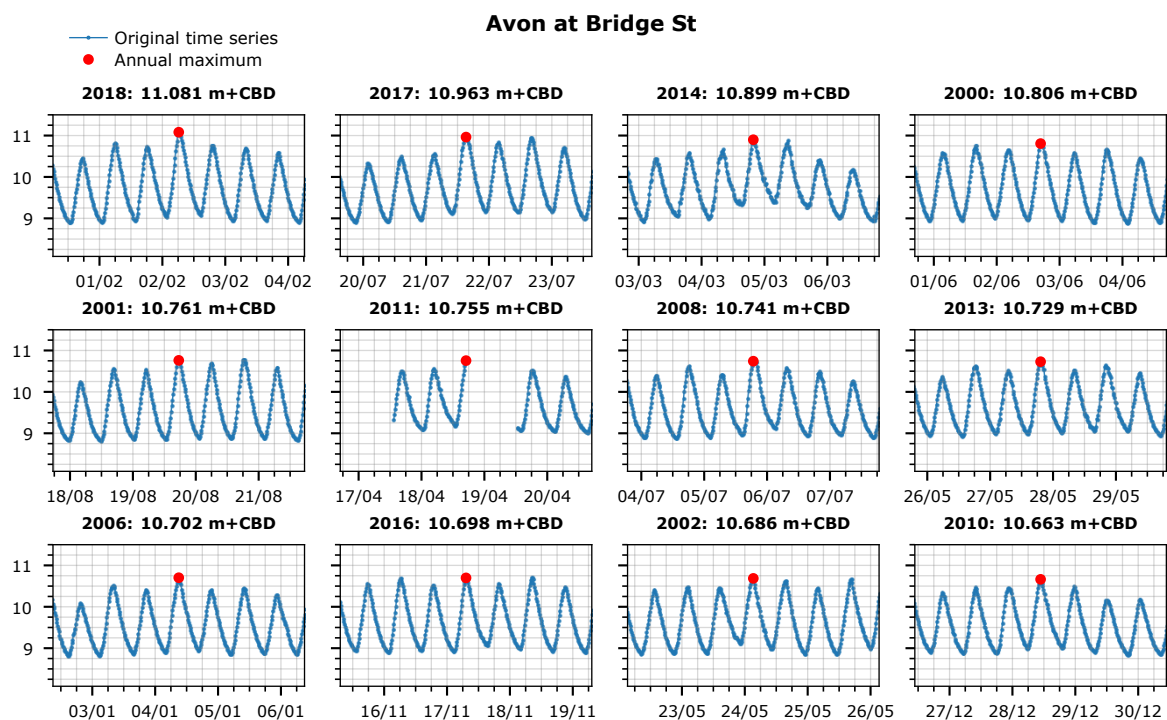


Figure E.8: Time series (corrected data) during annual maximum for Avon at Bridge Street

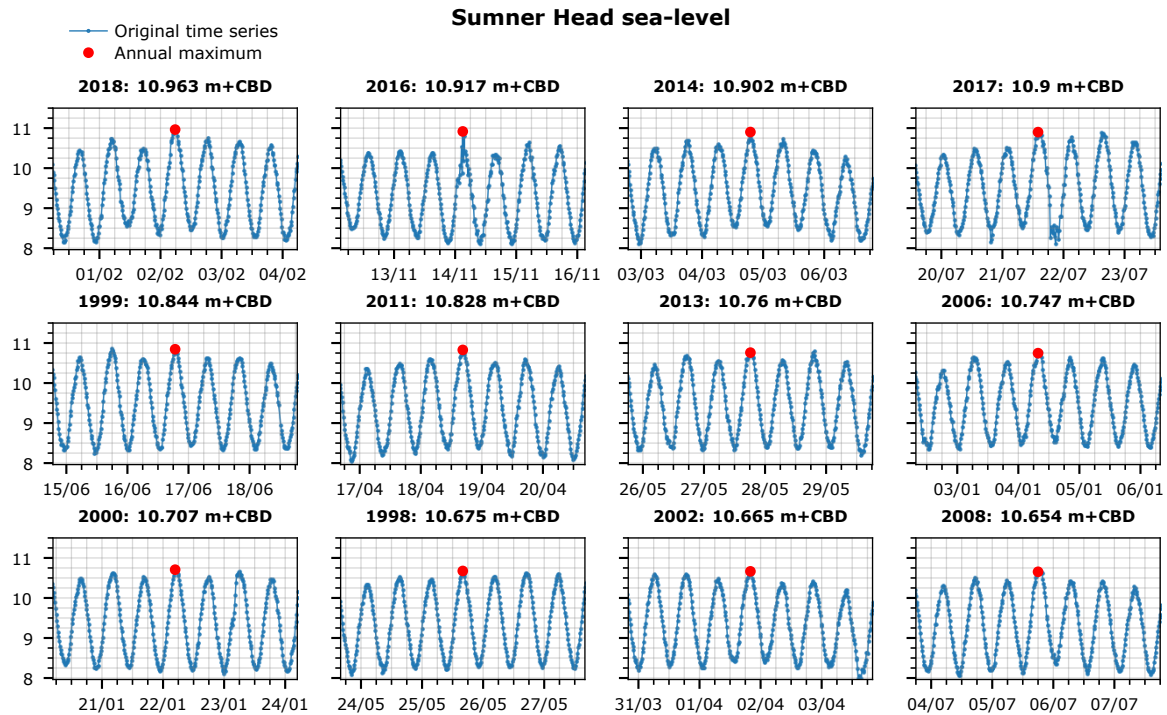


Figure E.9: Time series (corrected data) during annual maximum for Styx at Sumner Head

Comparison top three events

The top three events of Goring are compared to the events on the same date in the time series used in this analysis. The table below shows the differences in data used in both studies. The top three events of Goring's study and this study matches for all locations, except Sumner. At Sumner the second and third highest water level in our study occurred in 2016 and 2014 (see the figures above). As already mentioned, this is caused by different time series used in both studies.

Sumner	2-Feb-18	18-Apr-11	22-Jul-17
maximum time series	10.963	10.828	10.836
maximum Goring:	11.098	10.965	10.938
difference:	-0.135	-0.137	-0.102
Ferrymead	2-Feb-18	21-Jul-17	17-Apr-99
maximum time series	11.026	10.935	10.879
maximum Goring:	11.026	10.935	10.868
difference:	0	0	0.011
Brigte St	2-Feb-18	21-Jul-17	04-Mar-14
maximum time series	11.081	10.963	10.899
maximum Goring:	11.081	10.963	10.899
difference:	0	0	0
Styx	2-Feb-18	21-Jul-17	17-Apr-03
maximum time series	11.159	11.048	10.964
maximum Goring:	11.159	11.048	10.965
difference:	0	0	-0.001

E.4 Correction sea level rise scenario's

The above extreme value analysis is based on the original time series, in which the occurred sea level rise is included; measured water levels in the past occurred at a lower mean sea level than today. To obtain an extreme value distribution of today, the water levels in the past need to be corrected by sea level rise. Since the mean sea level of today is higher than 20 years ago. Therefore the measured sea levels in the past have to be heightened to correct for sea level rise. We considered three sea level rise (SLR) scenario's (details of this analysis are in the notebook on sea level rise):

- SLR1: 1.86 mm per year before 1990, 4.03 mm per year after 1990
- SLR2: 2.27 mm per year before 1990, -69 shift in 1990, 6.53 mm per year after 1990
- SLR3: 1.99 mm per year before 2005, 7.54 mm per year after 2005

The corrections made can be seen in the figure below.

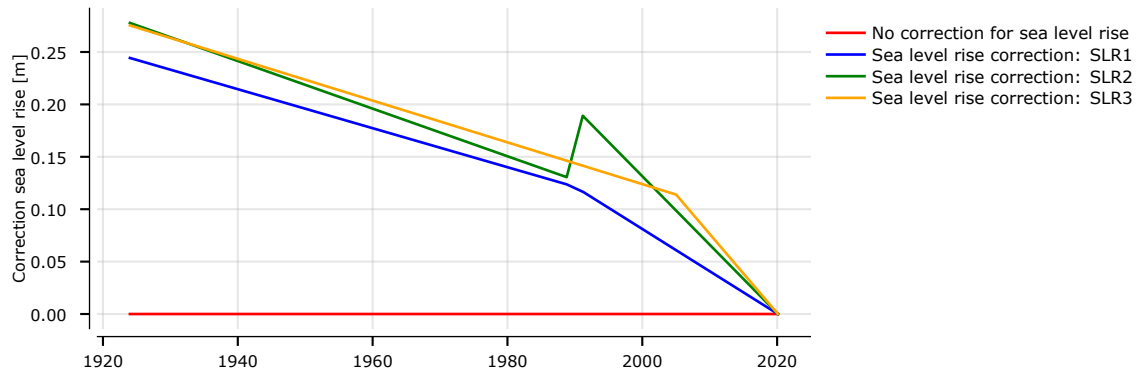


Figure E.10: Sea level rise corrections for the 4 used scenario's

Effect of sea level rise on extreme value distributions

The effect of the different sea level rise scenarios is evaluated for the four stations considered. The time series are corrected for sea level rise and the annual maxima are selected. The Gumbel distribution is fitted to these maxima and the resulting extreme value distributions (together with the annual maxima) of the different sea level scenarios are shown in the figures below. In this analysis, the water level time series are used until 1/1/2020 and therefore the results deviate from the analysis shown earlier.

For three of the four stations, the frequency line of the sea levels become flatter. The highest peaks in the original time series occurred in recent years and the SLR corrections for these events are small. The measured sea levels further back in the past have a larger correction and therefore the more frequent sea levels extremes become higher. This caused the flatter frequency line, if the time series is corrected for sea level rise. For Ferrymead this is not the case and the whole frequency line is sifted. This is because the first few years of the data is dominant in the analysis (also see the sensitivity analysis in Chapter 4).

There is no scientific basis to prefer one SLR scenario over the other. The upcoming analyses includes the sea level rise corrections SLR1, which has been ratified with GHD and CCC during the project workshops.

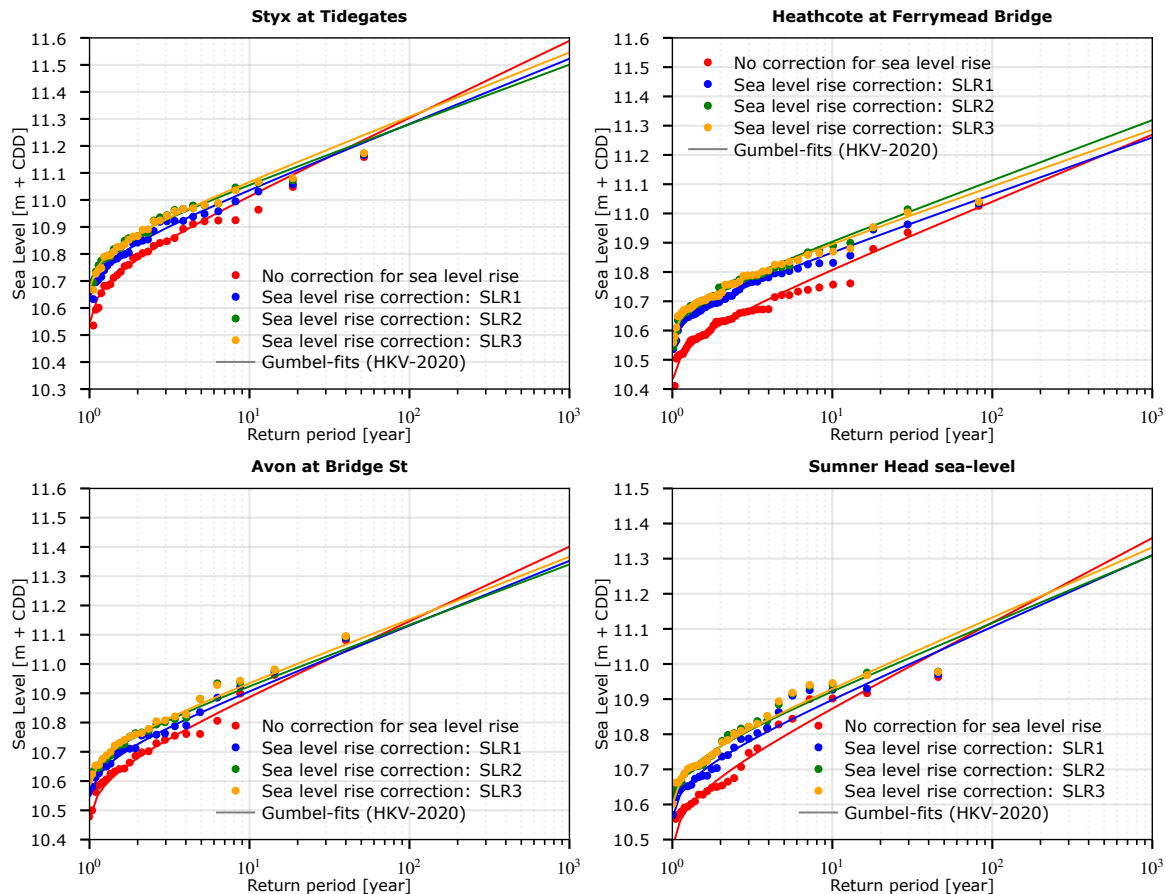


Figure E.11: Gumbel fits (Annual Maxima) on the corrected data, adjusted for sea level rise. The adjustments are done for each of the 4 used scenario's

E.5 Confidence bounds

Confidence bounds are a measure to express the sensitivity of the model to the observations it is fitted to. They give a range that likely (with $X\%$ confidence) contain a population parameter, for example the mean. In our case, confidence bounds are used to express the confidence of the fitted extreme value distribution. The range will be given by a lower and higher curve, which express the bandwidth around the values that are exceeded within a Y ARI.

In this study, the confidence bounds are calculated with the bootstrapping method. This is a test that uses random resampling with replacement of the measurements, and subsequently fits the distribution to this sample. So for example:

1. Say we have measured 100 peaks over the past 30 years. In case of the Annual Maxima method, this would obviously be 30 peaks, but the steps are the same.
2. Draw 1000 combinations of 100 realizations from these 100 peaks, in which it is possible to draw a single value twice or more. Some values will also not be drawn.
3. Fit the chosen extreme value distribution to each of these samples.

4. Calculate the value (sea level) that is exceeded with a for example 100 ARI, for each of the fitted curves. This results in 1000 values for the 100 ARI.
5. Calculate the 2.5th and the 97.5th percentile of these 1000 values. This is the 95% confidence interval for the searched 100 ARI sea water level.
6. Repeat the previous two steps for all requested ARI's.

The advantage of this method is that it is relatively simple, and doesn't need any assumptions of the distribution of the errors. The method does assume that the samples are independent and stationary. Since we corrected the data for sea level rise, and the peaks are selected within a sufficiently large time window, this assumption seems to be no problem.

The confidence bounds are presented when comparing different distributions in the next section, and for the definite results. For these last, the values are also presented in tabulated form. The confidence bounds are calculated based on 10000 samples, so the 95% confidence interval is based on the lowest 250 and highest 9750 values in these 10000.

E.6 Peaks-Over-Threshold vs. Annual maxima and fitted distribution (based on Lyttelton)

To decrease the uncertainty of the statistical analysis, a long time series is generally preferred. The analysis is therefore first executed for the Lyttelton data to make a choice on (1) the method for data sampling and (2) the type of distribution. As this data and site have the longest available time series. We assume the choices we make for Lyttelton are then applicable to the other stations around Christchurch. The differences between the five sites do not ratify other choices for either the method for peak selection nor the type of distribution fitted. All data used are the corrected data, adjusted for the SLR1 scenario. Figure E.12 shows the results, with on the:

- Left: annual maxima (red dots) with EV1 distribution (red line – similar distribution to Goring) and GEV (blue line)
- Right: Peak of Threshold (grey dots) with Exponential distribution (green line) and Pareto (yellow line)

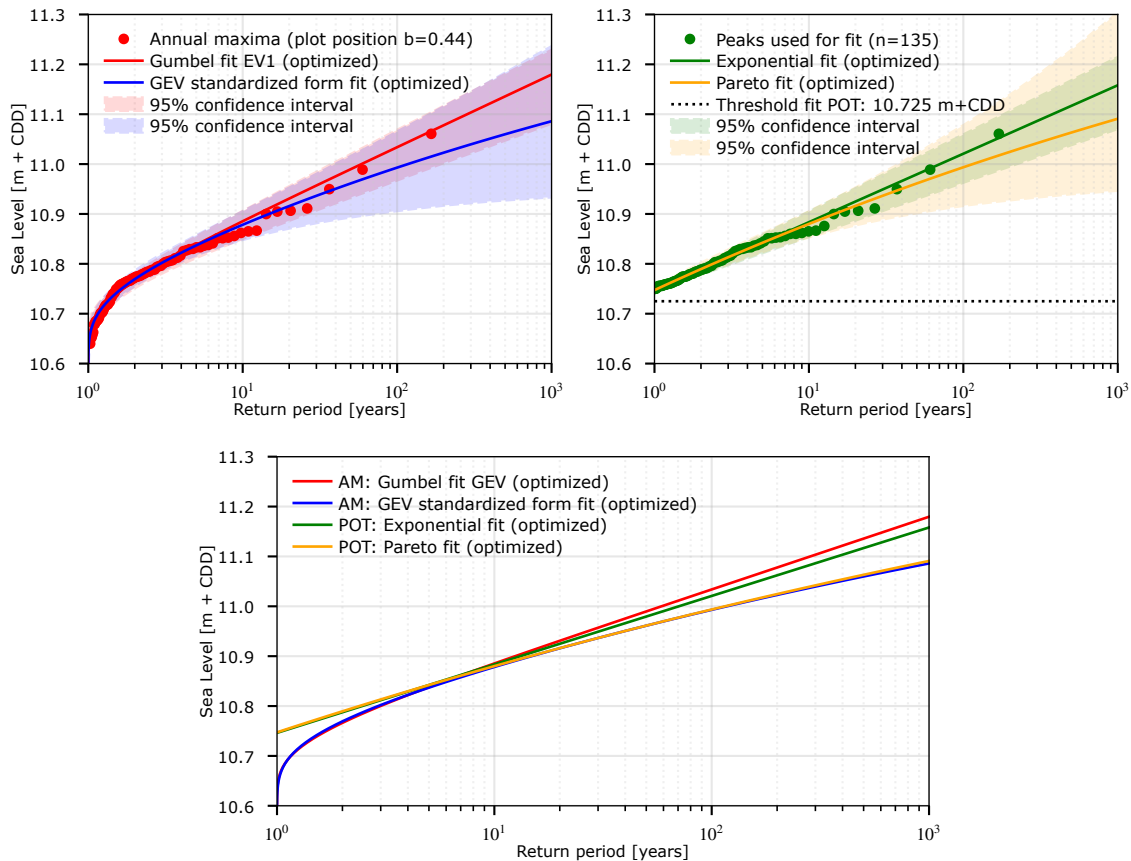


Figure E.12: Comparison between annual maxima and peak over threshold for Lyttelton port. The sea level rise is corrected with scenario SLR1

The peaks for the PoT method are selected by choosing a threshold value and window size. In this first analyses, the threshold is chosen based on an iterative process (see section E.7 for the optimize process). A rule of thumb is used to choose the threshold level around the level at the 1yr ARI (10^0). The threshold then is 10.725. Based on expert judgement the window size for event separation is chosen to be 50 hours, based on the duration of storms / atmospheric pressure fields.

Frequency vs. probability

The occurrence of an event can be expressed in different ways. In general it can be said that an event with level s has a certain exceedance frequency $F_S(s)$ (per year) or an exceedance probability $P(S > s)$ (dimensionless). A return period is often used in frequency analysis (an estimated average time between events to occur), which is the inverse of exceedance probability per year or exceedance frequency. The exceedance probability per year of selected Annual Maxima can not be larger than one, because only one event is selected every year. In the POT-method it is possible to select multiple peaks a year and therefore the exceedance frequency can be larger than once per year. This is a different sample, to which subsequently a different distribution is fitted. These distributions (GEV or GPa) differ for frequent events ($T < 10$ year), while they are almost similar for extreme events ($T > 10$ year), because the highest peaks will likely also be annual maxima.

The Gumbul distribution (AM) and Exponential distribution (POT) both have an exponential tail. The shape parameter is 0, which means they do not have an upper limit or a lower bound. The Pareto and GEV, of which the exponential distribution and Gumbel distribution are a specific case, do not have a fixed shape parameters. This can result in an upper limit (like the type III Weibul distribution) or a long tail (like the type II Frechet distribution). An upper limit would make the most sense if a physical limit can be substantiated. A long tail is related to higher frequencies for increasingly extreme events, than the exponential tail would give.

In E.12 the observations and distribution are all plotted on a logarithmic scale. The best method for comparing observations to a fitted distribution, is plotting the distributions and the observations on a standardized scale: $y = \frac{x-loc}{scale}$. In which the loc and scale of the chosen distribution are close to the mean and sigma, but not similar. Plotting in this manner is especially useful for the to annual maxima related distributions, since the high frequent events can be distinguished better. The result of this is shown in figure E.13 below. Visually the the left distributions with shape parameter 0 seem to fit the extreme events better than the right distributions with a fitted shape parameter.

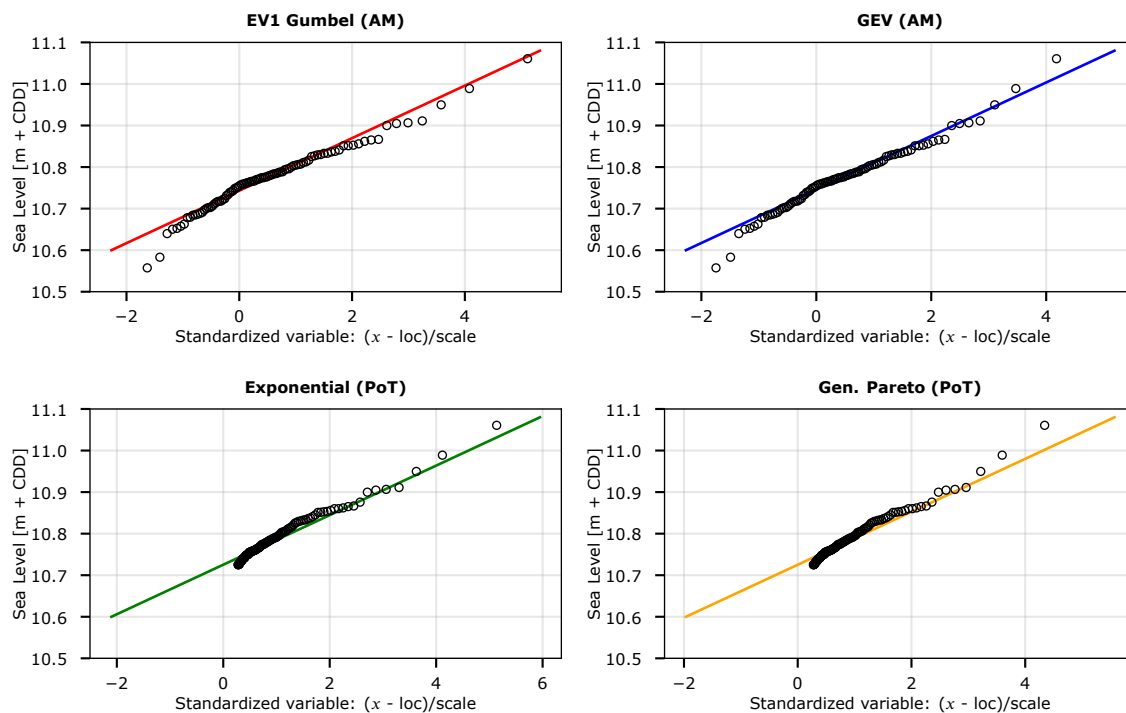


Figure E.13: Distributions and observations plotted on their standardized scale.

We consider using more data points to be better, so we prefer peaks over threshold (POT). HKV uses this in other statistical analyses for government / hydraulic research institute (in the Netherlands and abroad). Both POT and AM data show no indication of an upper limit in the measurements (red and grey dots). Additionally we can't substantiate a physical reason why you would expect a physical upper limit, at least not for the sea level range we consider. For that reason we would advise an exponential or EV1 type of distribution. This is again similar to what we have always done for coastal sites in the Netherlands and are now doing in Belgium as well (in cooperation with the Belgium Hydraulic Research Institute).

The choice of the threshold can be done (1) iterative by eye or (2) determine the RMSE of the fit to the data points for various threshold levels and then pick the level with the lowest RMSE. In the following paragraph, we choose the threshold based on the minimum RMSE. This process has been designed together with the Belgium Hydraulic Research Institute to make the fitted distributions less sensitive for the (manual) choice of the threshold levels.

E.7 POT and Exponential distribution

We have chosen to use the POT method and fit an exponential distribution to the selected peaks of the stations around Christchurch. The choice of the threshold value is substantiated by minimizing the RMSE of the fitted distribution and data points for peaks that have a higher return period than once a year. This means a range of threshold values are chosen and the one with the lowest RMSE is said to be the optimized value. The maximum threshold value is set at the once per year sea level and the minimum is set at the four times per year sea level. The results are shown in the figures below.

The following paragraphs describe the sensitivity analysis carried out to quantify the effect of choices made in this analysis.

1. Effect of threshold
2. Comparison between Peak of Threshold and Annual Maxima
3. Effect of data
4. Effect of Sea Level Rise

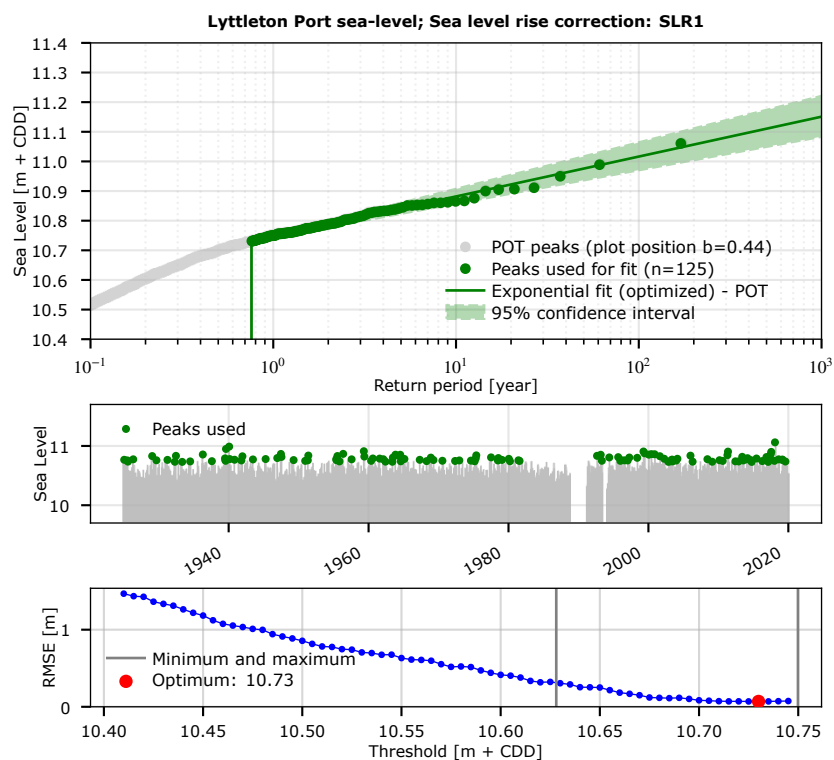


Figure E.14: Peak over threshold fit optimization for Lyttelton Port

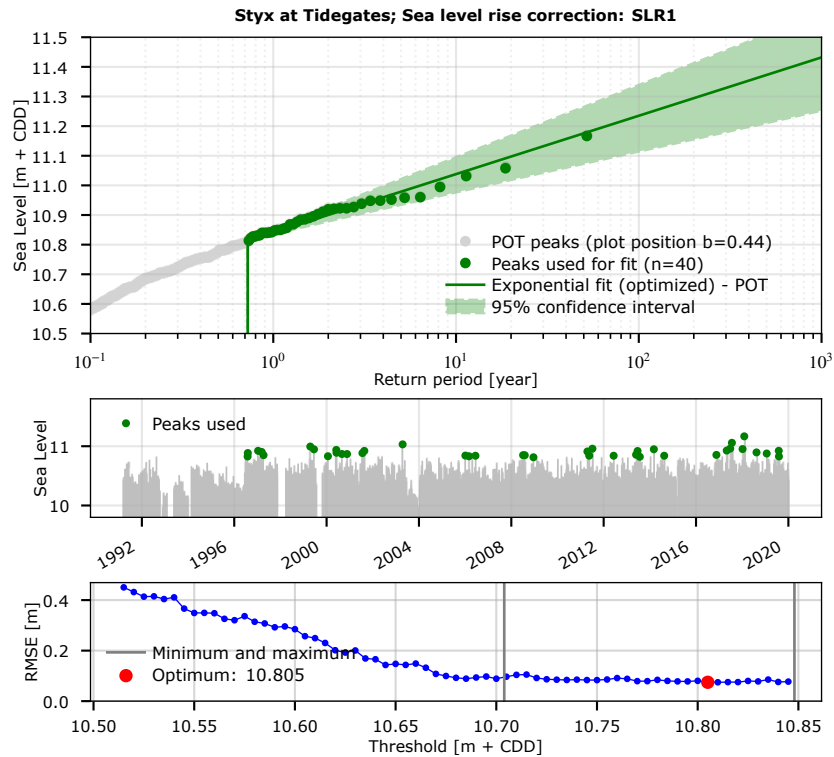


Figure E.15: Peak over threshold fit optimization for Styx at Tidegates

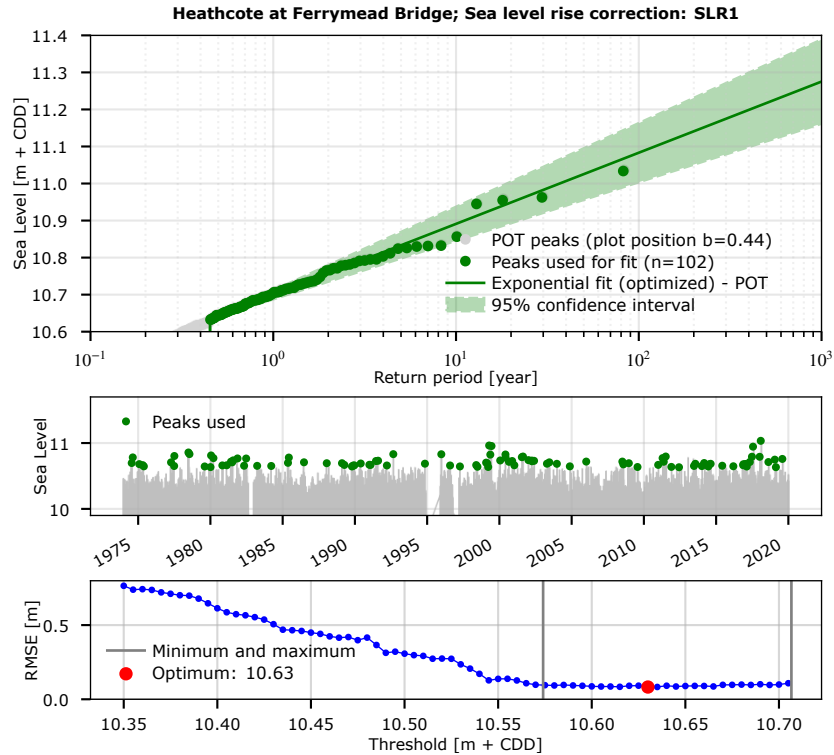


Figure E.16: Peak over threshold fit optimization for Heathcote at Ferrymead Bridge

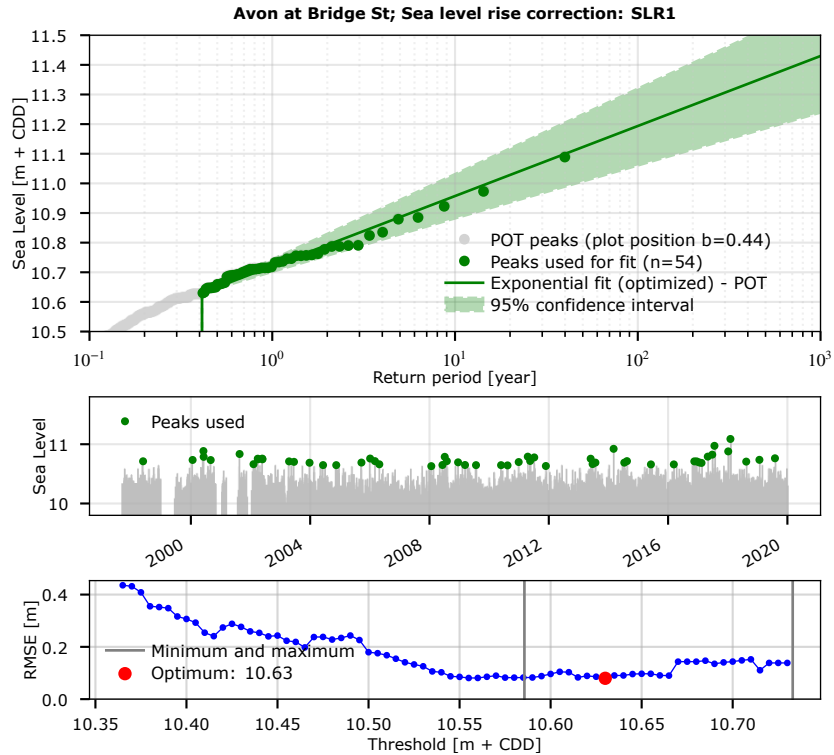


Figure E.17: Peak over threshold fit optimization for Avon at Bridge Street

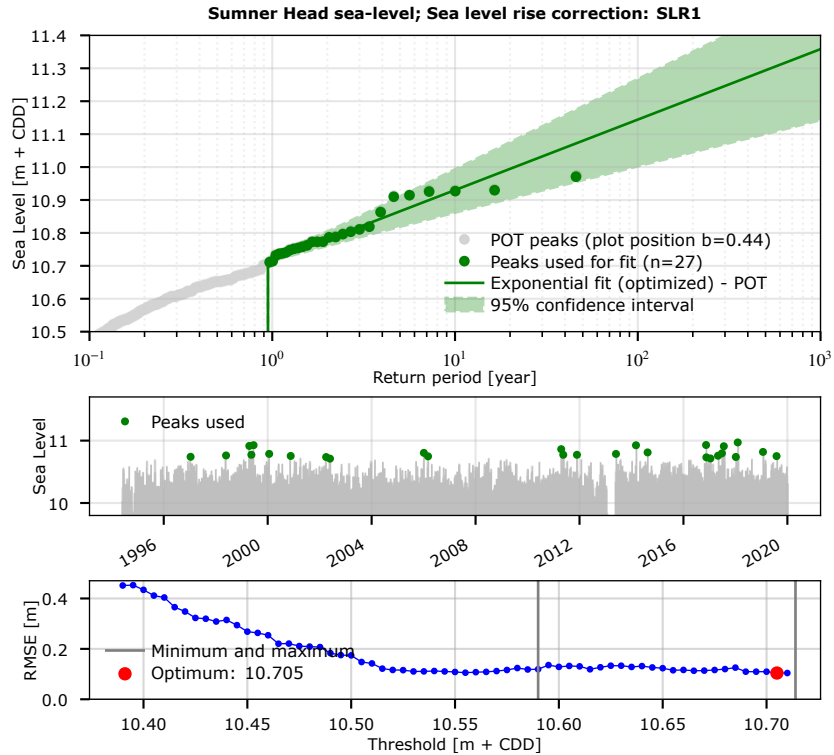


Figure E.18: Peak over threshold fit optimization for Styx at Sumner Head

E.8 Influence of different threshold values

The influence of different threshold values are investigated. In addition to the optimized threshold, we have considered threshold values with the following frequency of occurrence:

- $T=0.25$ (level that is exceeded 4 times per year)
- $T=0.5$ (level that is exceeded 2 times per year)
- $T=1.0$ (level that is exceeded once a year)

It is debatable whether a level that is exceeded 2 or 4 times per year should be considered a peak. In the following method we determine the threshold of peaks by looking for the level that provides the best fit for the extreme value distribution. The underlying assumption is that lower peaks that improve the fit, should be considered peaks.

The figures below show the results for 5 stations. In general, the variations in the threshold values have a maximum effect in the order of 10 centimetres for sea levels at $T=1000$ years. If the threshold value is chosen lower ($T=0.25$), the extreme sea levels statistic generally become higher. The observations with a frequency of higher than once per year shows a modestly downward bend. If you include these peaks in the fit, the left tail is also pulled to these points and tilts the right tail (in which we are interested).

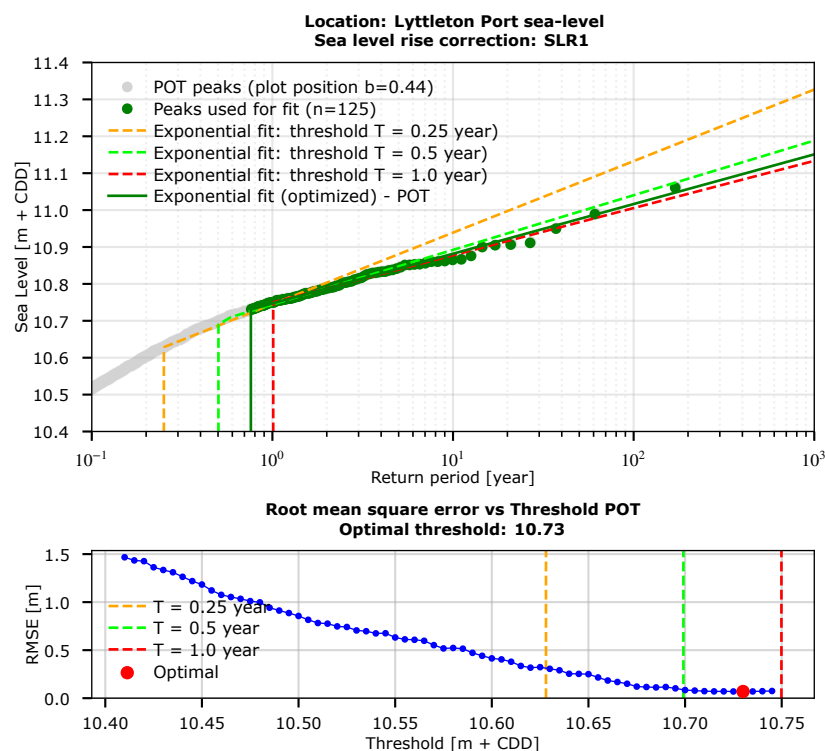


Figure E.19: Peak over threshold influence for Lyttelton Port

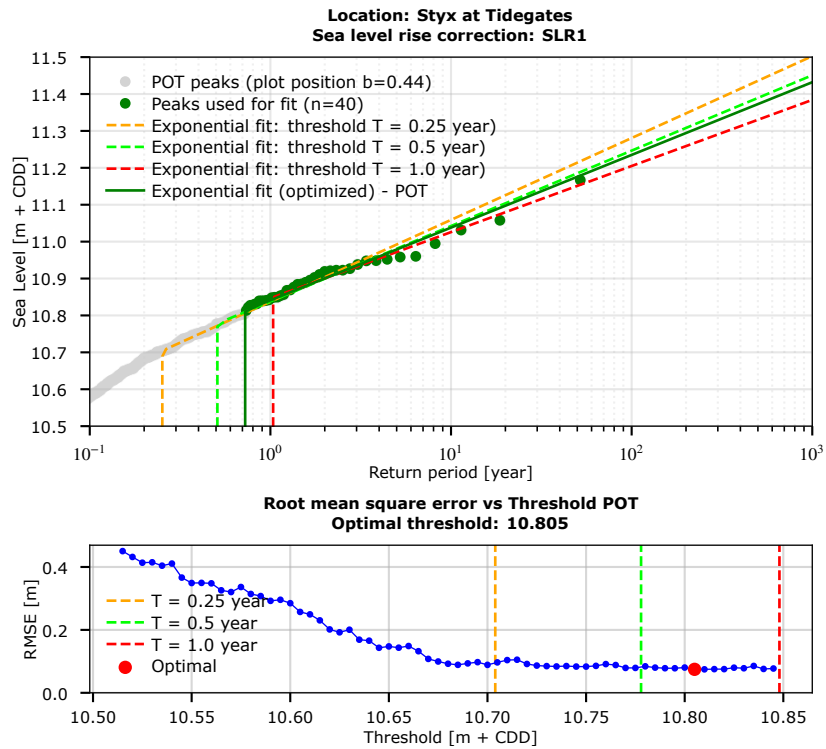


Figure E.20: Peak over threshold influence for Styx at Tidegates

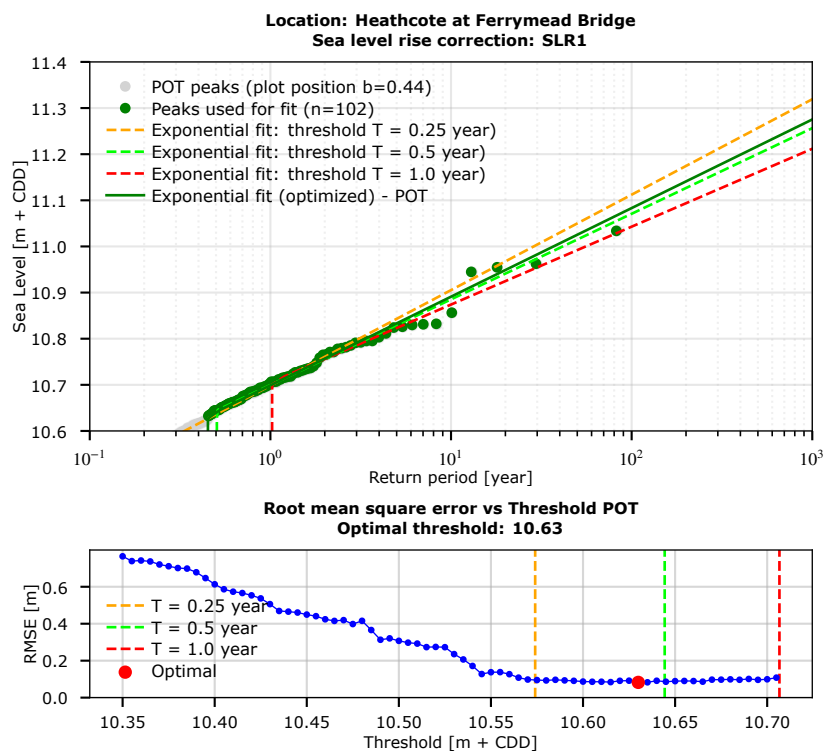


Figure E.21: Peak over threshold influence for Heathcote at Ferrymead Bridge

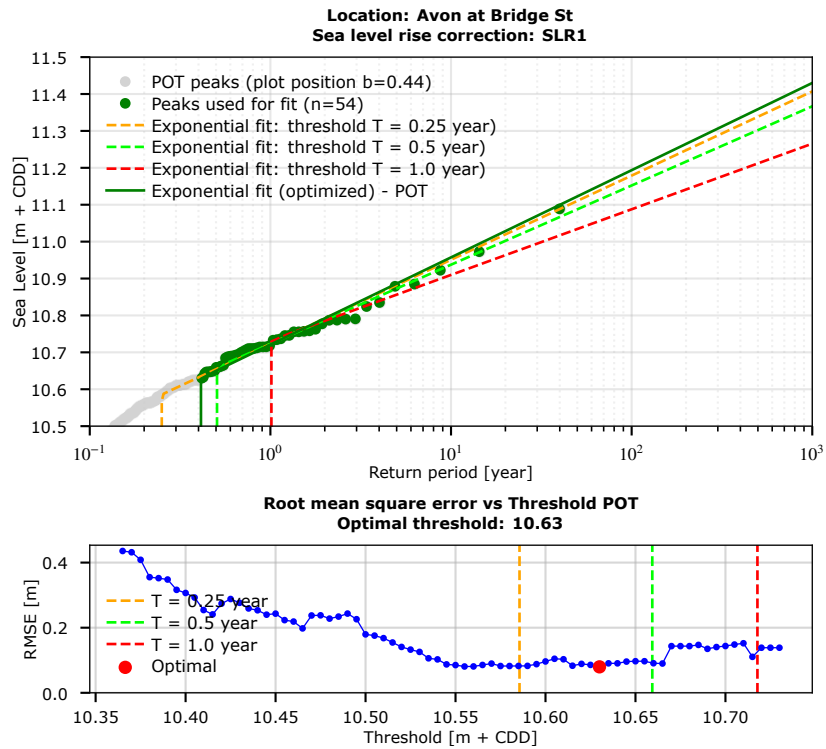


Figure E.22: Peak over threshold influence for Avon at Bridge Street

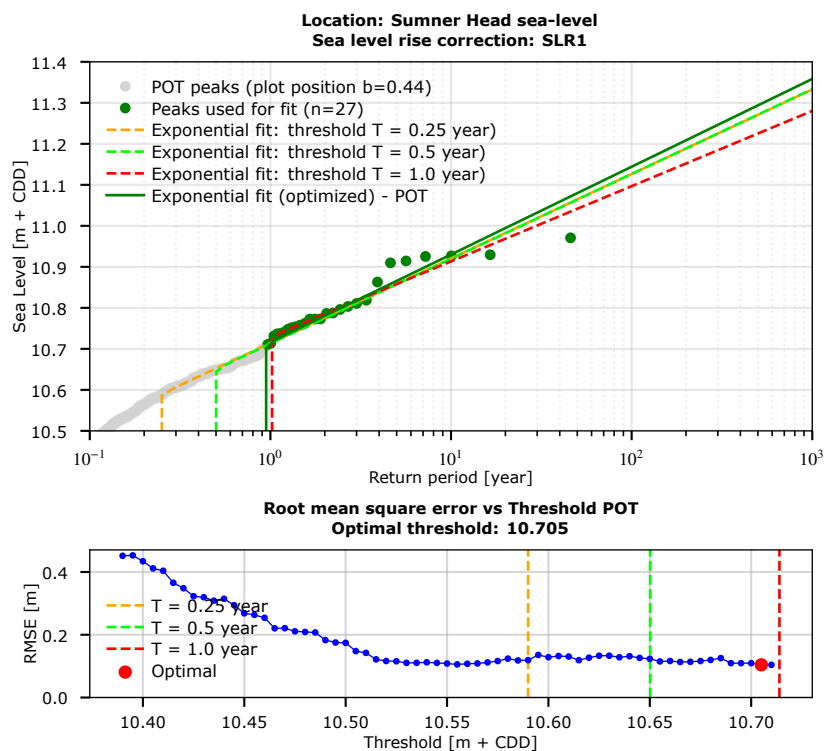


Figure E.23: Peak over threshold influence for Styx at Sumner Head

E.9 Comparison with Annual Maxima

We substantiated the choice to use PoT earlier, but for completeness we show the differences between POT/GPD and AM/GEV in the figures below. It can be observed that at some locations the fitted Gumbel distribution (AM) results in a much steeper line. This is because the AM-method also selects peaks from years with only minor storms. The Gumbel distribution also has a lower tail, which is just as important in the fit as the upper tail. The POT-method (GPD) only includes peaks above an optimised threshold and ignores the lower years. In case of a dataset from a stationary source, this should not matter. However, if the dataset is not stationary resulting in a period with particularly lower measurements, the Gumbel distribution with its lower tail is 'stretched' towards these lower years, resulting in a steeper line. The GPD at the same time just ignores these lower values, resulting in less peaks over the full period and thus a milder slope. So both methods are affected by a non stationary data source, but the effect is more clearly visible for the AM-fitting.

For Sumner and Avon the slope of the fitted distributions are similar. At Styx the slopes of AM are steeper. This is caused by the first few years of the data having an influence on the statistic (the data quality is questionable in the first few years of the time series – also see the next paragraph).

Note: the plot positions on the return period axis, and fitted distribution of AM/GEV and POT/GPD are expressed in different ways (as described in section 2 of this appendix). We assume that the exceedance probability of the AM are equal to the exceedance frequency, which is not the case for frequencies $T < 10$ year. The two approaches should therefore be compared in the range of $T > 10$ years.

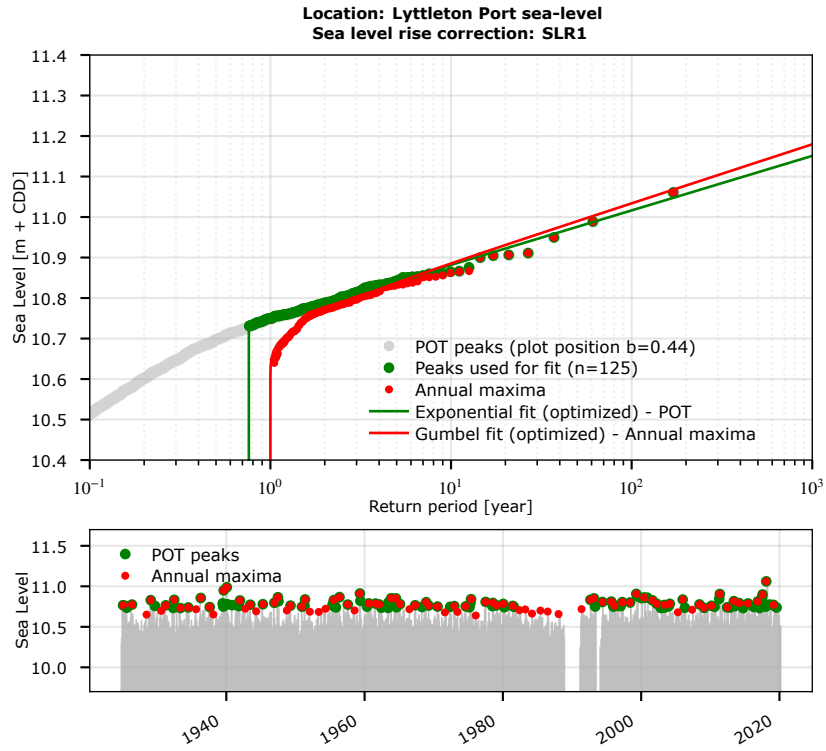


Figure E.24: Comparison PoT and AM for Lyttelton Port

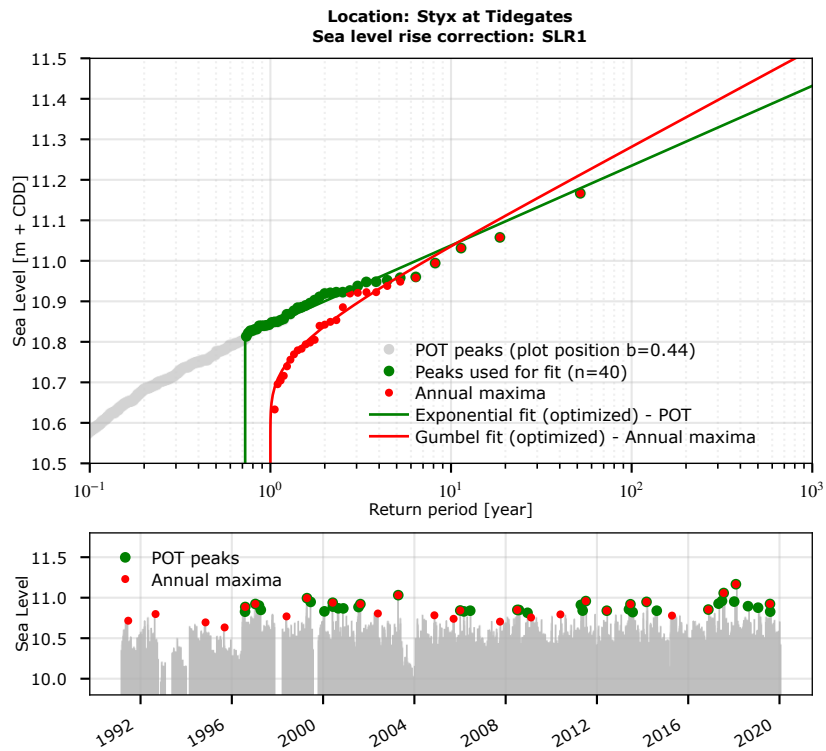


Figure E.25: Comparison PoT and AM for Styx at Tidegates

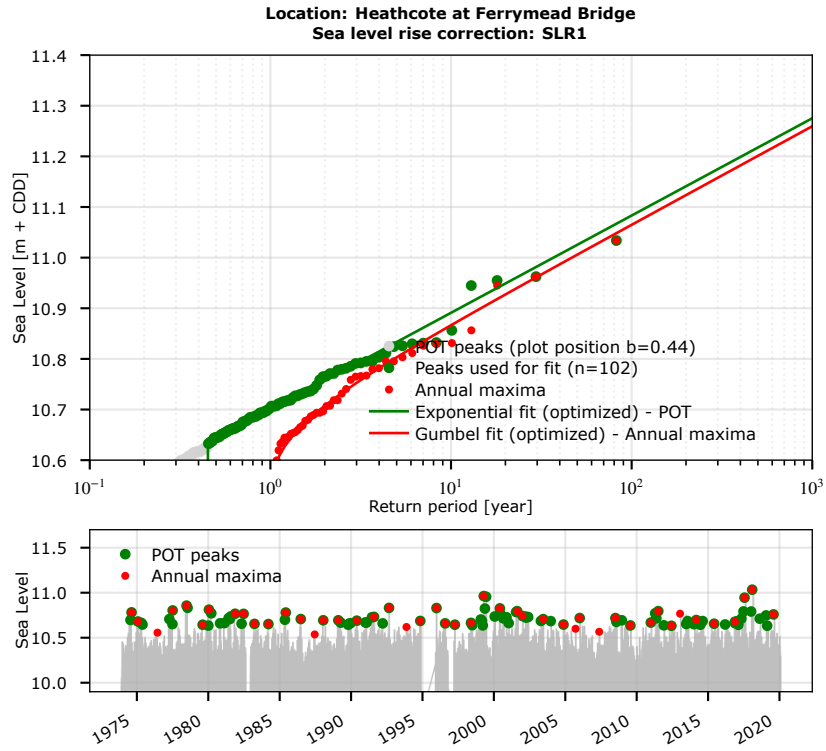


Figure E.26: Comparison PoT and AM for Heathcote at Ferrymead Bridge

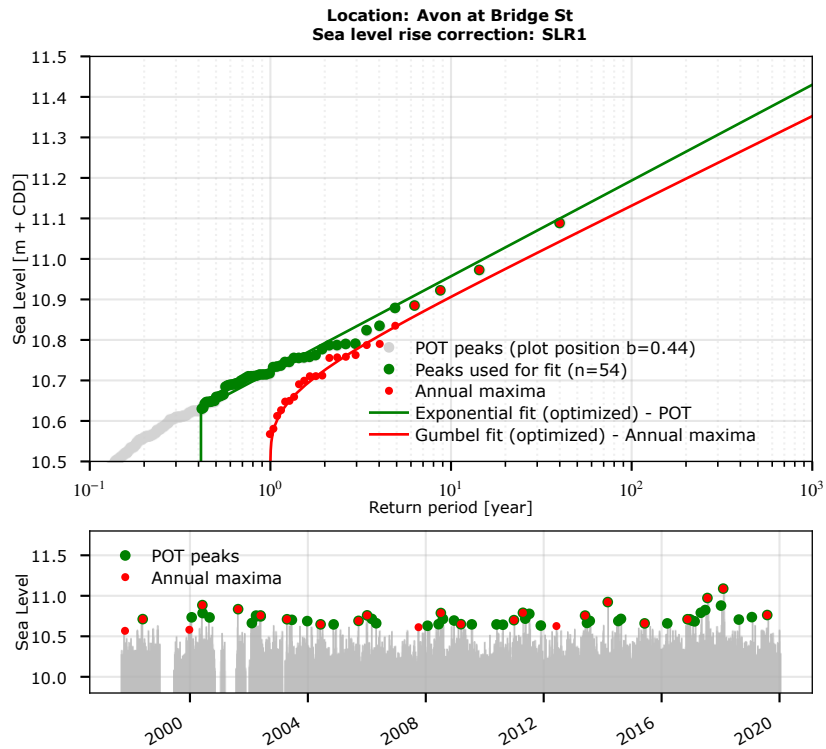


Figure E.27: Comparison PoT and AM for Avon at Bridge Street

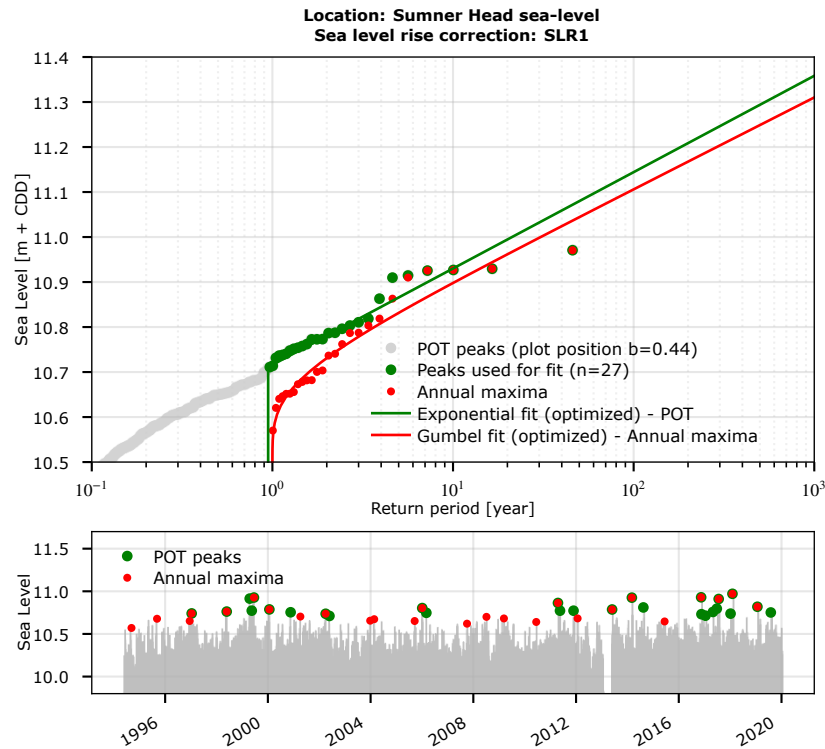


Figure E.28: Comparison PoT and AM for Styx at Sumner Head

E.10 Effect of inhomogeneous time series and data quality issues on fitted distributions

The selected sea level peaks are expected occur more or less evenly distributed over time (especially after the correction for sea level rise). There can always be years in which less high sea levels occur or clusters due to natural oscillations, but if the peaks are clearly not evenly distributed, or there is an unnatural abrupt change, the time series may not be homogeneous. An inhomogeneous time series could lead to misleading results and conclusions. The station Styx shows unevenly distributed POT-peaks. Here the water levels in the early years appear too low. Therefore two sensitivity analyses are executed related to:

1. The rate of sea level rise and the corrections made
2. The selected period of the time series

Sea level rise

We investigated the effect of 25% more sea level rise which means that sea levels in the past are corrected/raised to a greater extend. This percentage is chosen roughly and high compared to the rates of sea level rise found in other SLR analyses. The effect on the correction is graphically shown in the figure below:

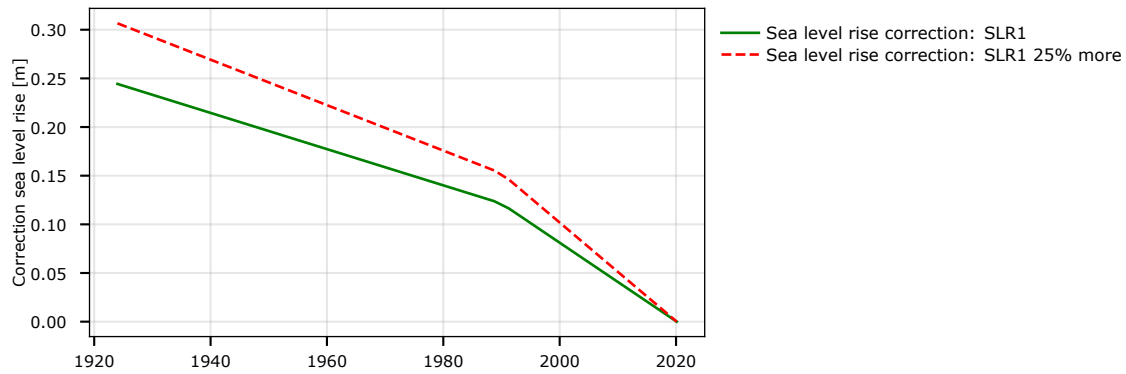


Figure E.29: Sea level rise correction, SLR1 and SLR1 + 25%

In the figures below, it can be seen that 25% more sea level rise does not significantly influence the results. The water levels increase up to 5 centimetres for the most extreme events.

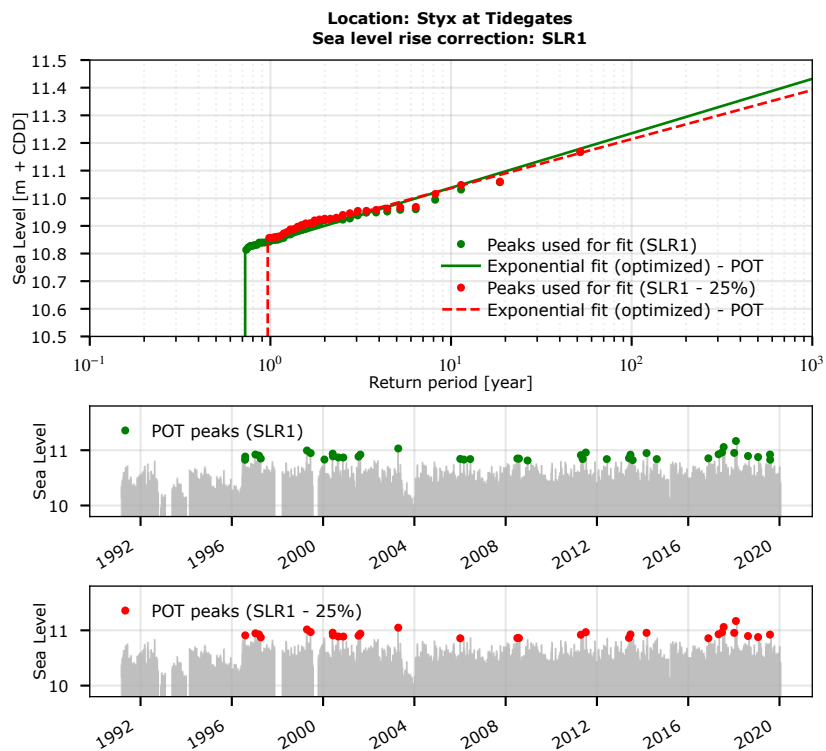


Figure E.30: Effect of additional sea level rise for Styx at Tidegates

Shortening time series

We exclude several years at the beginning of the time series for the station Styx. This period is based on visual inspection of selected peaks in the time series. The start of the shortened time series is 1-1-1996 instead of 1-1-1991.

The effect on the extreme value analysis can be observed in the figures below. The sea levels

at Styx in the period that is excluded were lower than other years. Therefore, the time series become shorter, while no POT-peaks are excluded. Therefore, all the POT-peaks occur a bit more frequently, but this adjustment barely influences the result. Note that this is not the case when using AM, as demonstrated in figure E.25. The PoT method is less susceptible for a few years with consistently lower data. Since we use the PoT method, and the effect of the shortened series is very small, we will use the full time series in the concluding fit and tabulated results.

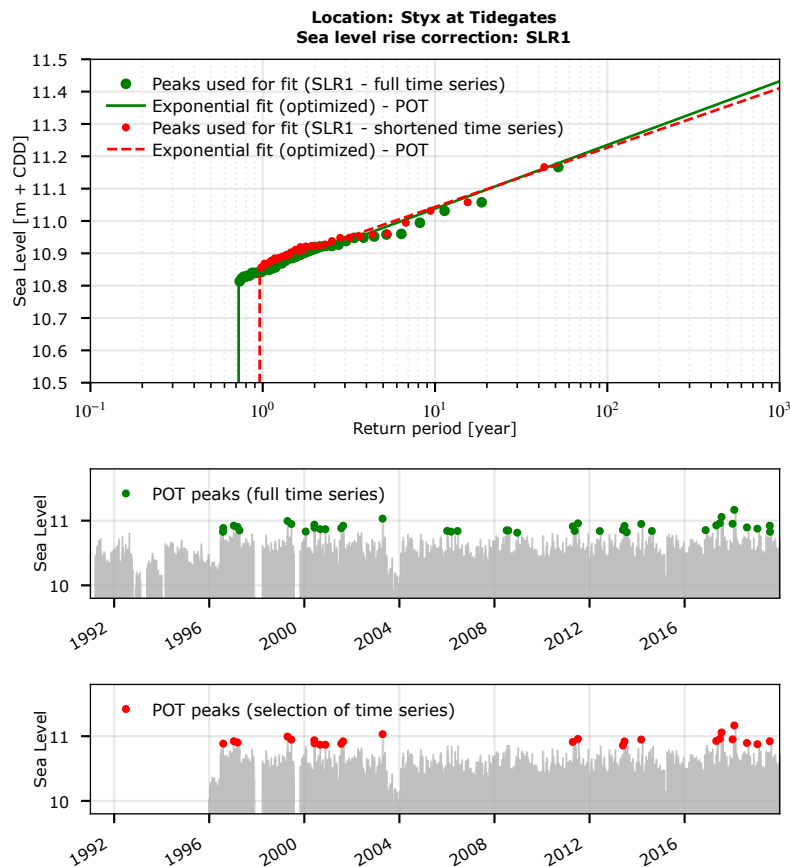


Figure E.31: Effect of shorter time series for Styx at Tidegates

E.11 Conclusions

In this appendix we first reproduced the extreme value analysis of sea levels around Christchurch executed by Goring. By doing so, we ensure the same starting point of our analyses. The subsequent analyses were based on time series in which we corrected for sea level rise. This is necessary to estimate the extreme value statistics of the sea level nowadays, since the extreme events in the past occurred at a lower mean sea level than nowadays. We compared the Annual Maxima (AM) / Generalized Extreme Value (GEV) distribution and the Peaks-Over-Threshold (POT) / Generalized Pareto Distribution (GPD) for the Lyttelton station where a long time series is available. We advise to use POT in combination with the exponential distribution, because it allows to select multiple peaks per year. Especially for shorter time series this results in more data and increased reliability of the extreme value distributions. This method is used for the other four locations (Styx, Ferrymead, Bridge St and Sumner) to derive extreme value statistics.

The threshold is optimised by using the RMSE of observations and the fitted distribution. The sensitivity analyses show that the resulting fitted distributions are in general not very sensitive to choices made (threshold value, sea level rise correction and inhomogeneities in time series) and gives us confidence that the results are relative stable.

The final results can be seen in the figures below. The parameters corresponding to the fitted exponential distributions can also be found in the table below. The exceedance frequency of level x described by an exponential distribution is defined as follows:

$$F_X(x) = \exp(-(x - \mu)/\sigma) \quad (12)$$

where x is the sea level, μ is the location parameter and σ is the scale parameter. This equation can be derived in such a way that it expresses the sea level as function of the return period ($T = 1/F(X)$):

$$x = \sigma \ln(T) + \mu \quad (13)$$

Note that a **FIG wave allowance** of 80 mm is included (adopted from Goring's reporting) with the intention that this new EV work be the new basis for future modelling, similar to Goring's previous work.

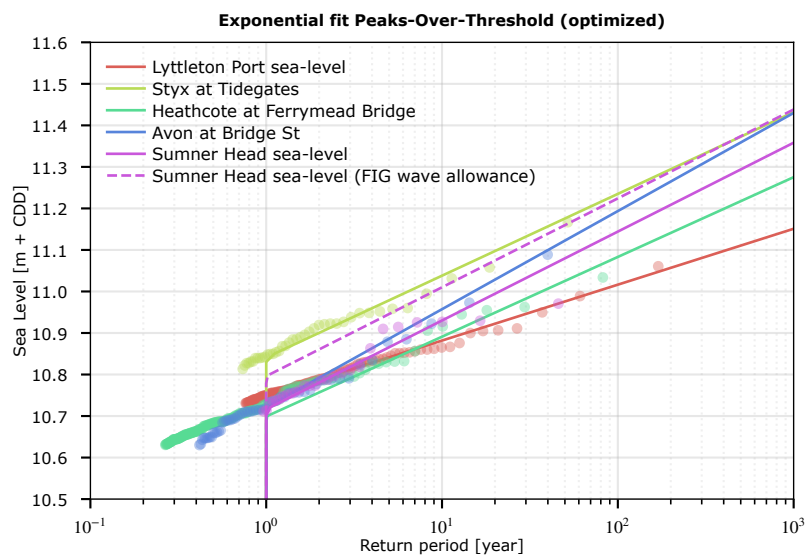


Figure E.32: Final results of the extreme sea level analysis.

The figures below compare the final results with the results of the Extreme Value Analysis by Goring. In general the extreme value distributions derived in this study have slightly flatter slopes compared to Goring, because of the sea level rise correction. The correction of sea level rise also results in increased sea levels, but the magnitude differs per station.

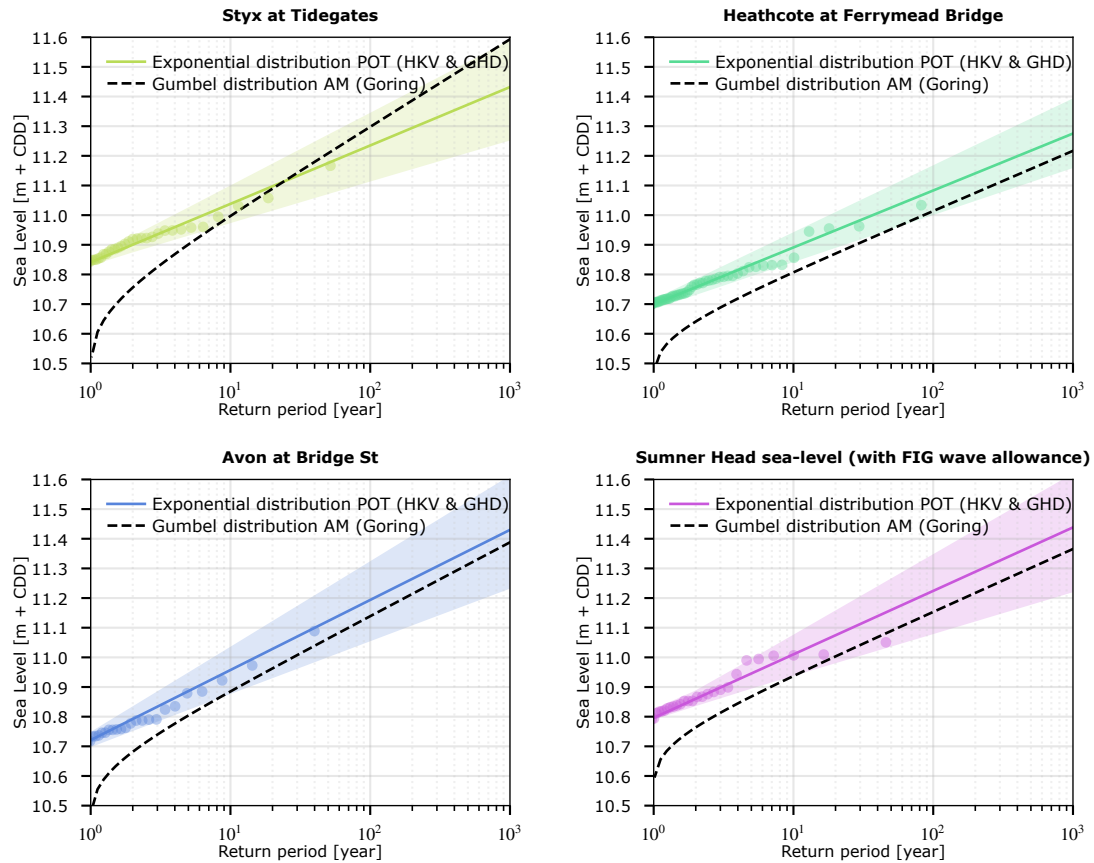


Figure E.33: Final results of the extreme sea level analysis, compared to the results from Goring

E.12 Tabular results

This section contains the results from the fitted extreme value distributions in tabulated formula. The next table contains the parameters of the fitted exponential distributions. The parameters are chosen such that they result in a distribution from a once per year return period. For the exact values see the second table, that also shows the values that are exceeded on average once per ARI years. The lower and upper bound of the 95% confidence interval is shown, as well as the values from the fit itself.

	Location parameter	Scale parameter	Threshold value
Avon at Bridge St	10.721	0.1026	10.63
Heathcote at Ferrymead Bridge	10.699	0.0834	10.63
Lyttleton Port sea-level	10.747	0.0585	10.73
Styx at Tidegates	10.841	0.0855	10.81
Sumner Head sea-level	10.716	0.0930	10.71
Sumner Head sea-level - FIG wave allowance	10.796	0.0930	10.79

ARI	Avon at Bridge St			Heathcote at Ferrymead Br			Lyttleton Port			Styx at Tidegates			Sumner Head		
	2.5%	Fit	97.5%	2.5%	Fit	97.5%	2.5%	Fit	97.5%	2.5%	Fit	97.5%	2.5%	Fit	97.5%
1	10.701	10.721	10.744	10.688	10.699	10.711	10.745	10.747	10.750	10.834	10.841	10.856	10.715	10.716	10.741
2	10.756	10.792	10.830	10.736	10.757	10.779	10.779	10.788	10.797	10.879	10.900	10.924	10.761	10.780	10.805
5	10.827	10.886	10.944	10.799	10.833	10.869	10.824	10.841	10.860	10.935	10.979	11.021	10.821	10.866	10.912
10	10.881	10.957	11.031	10.847	10.891	10.937	10.858	10.882	10.907	10.976	11.038	11.094	10.865	10.930	10.995
20	10.935	11.028	11.118	10.894	10.949	11.005	10.892	10.922	10.954	11.018	11.097	11.168	10.908	10.995	11.078
50	11.005	11.122	11.233	10.958	11.025	11.095	10.937	10.976	11.016	11.073	11.176	11.265	10.965	11.080	11.188
100	11.059	11.194	11.320	11.005	11.083	11.163	10.970	11.016	11.064	11.115	11.235	11.338	11.008	11.144	11.271
200	11.113	11.265	11.407	11.053	11.141	11.232	11.004	11.057	11.111	11.156	11.294	11.412	11.050	11.209	11.355
500	11.183	11.359	11.522	11.116	11.217	11.322	11.049	11.110	11.173	11.211	11.373	11.509	11.107	11.294	11.465
1000	11.237	11.430	11.609	11.164	11.275	11.390	11.083	11.151	11.220	11.253	11.432	11.583	11.149	11.358	11.548
2000	11.290	11.501	11.696	11.211	11.333	11.458	11.117	11.192	11.268	11.294	11.491	11.656	11.192	11.423	11.632
5000	11.361	11.595	11.811	11.274	11.410	11.548	11.162	11.245	11.330	11.349	11.570	11.754	11.249	11.508	11.742
10000	11.414	11.666	11.897	11.322	11.467	11.616	11.196	11.286	11.377	11.390	11.629	11.827	11.292	11.572	11.825
20000	11.468	11.737	11.984	11.370	11.525	11.684	11.230	11.326	11.424	11.432	11.688	11.901	11.334	11.637	11.908
50000	11.538	11.831	12.099	11.433	11.602	11.774	11.275	11.380	11.487	11.487	11.767	11.998	11.392	11.722	12.019
100000	11.592	11.903	12.186	11.481	11.659	11.842	11.309	11.420	11.534	11.528	11.826	12.072	11.434	11.786	12.102

F Correlation analysis surge and rain on Avon river

This appendix describes the derivation of the correlation between rainfall and surge for Avon river. Long duration rainfall events are likely to be caused by a low barometric air pressure, which will raise the sea water level a bit. Theoretically this relation is 10 hPa barometric pressure difference to 10 cm sea levels, but this can differ a bit due to environmental conditions.

To determine the relation between rainfall and surge, we follow the following steps:

1. Select periods with a high 2, 9 or 24 hour rainfall sum.
2. Find the phase between the peak of the rainfall events, and the moment we expect the hydrograph caused by the rainfall event to reach the sea. For this the measured water levels on Avon and Heathcote River are used.
3. Determine the average surge in the 24 hour period around this moment, the period in which the flood wave reaches the sea. A time average is used because the phase from the last bullet can differ per event. Also an average is less susceptible for outliers. The 24 hour period is chosen as a characteristic time scale for a weather event.
4. Fit a correlation model to the results.

This correlation model can later be used to find a distribution for the surge to expect during a given rainfall event.

F.1 Elements for the correlation model

First the different source elements for the correlation model are described. This is more or less the analysis of the different data, before synthesizing a correlation model.

Data

We use the following data sources:

- **Surge** The surge, defined as the measured sea level minus the astronomical tide, has been determined by fitting the tidal constituents (table A.1) to the observed sea level. For this the Lyttelton data series are used, since this is the longest series and after QA we are confident that the quality is good enough for this analysis. Everything that contributes to the sea level which is not the tide, is surge.
- **Water level** The water levels from two gauges on Avon river are used. The first is Gloucester Street, a bit more upstream. The second one is PS205 halfway down the river from Gloucester Street to the estuary.
- **Rain** For the rain the rain from 7 rain gauges in and around the Avon catchment are used. The rainfall depths are weighted with the area of the Thiessen polygon around the gauge that intersects the Avon catchment. This is explained in the appendix on rainfall data.

Surge

The following plots shows parts of the time series of sea water level, tide and surge for Lyttelton. The difference between measured and predicted tide (black line) we call the surge.

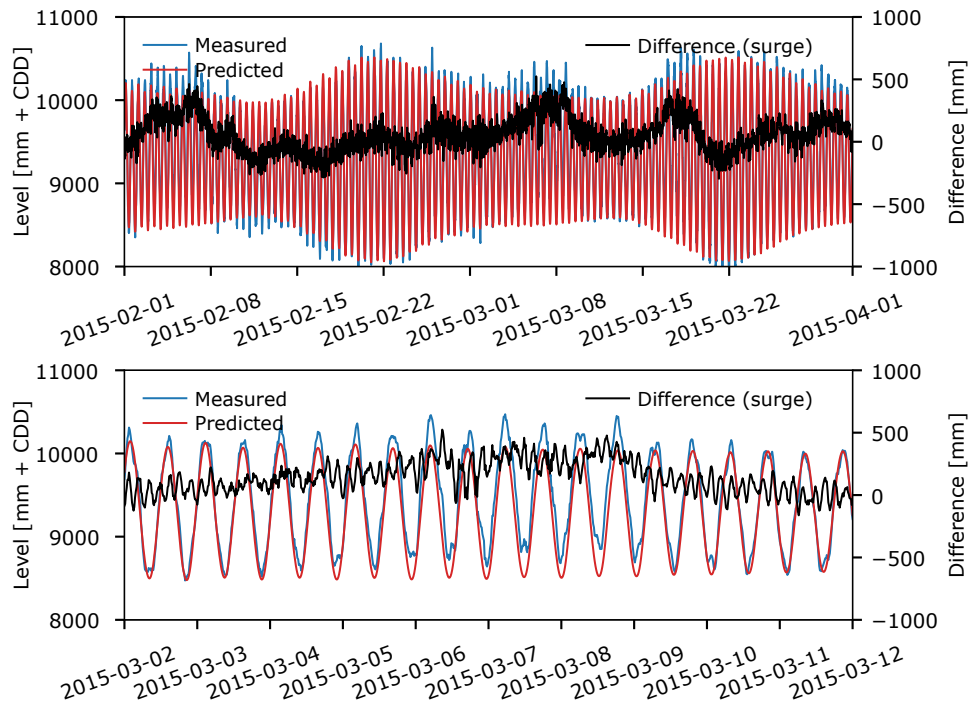


Figure F.1: The above plot shows the measured sea water level and the predicted tide. The difference (on the right axis) we call the surge. The lower plot shows the same, but is zoomed in on a 10 day period.

The closest sea-side tide gauge for Avon river is Sumner, but the data here are shorter and of lower quality. With a data analysis presented in section A.4 we showed that the relation between Sumner and Lyttelton for the overlapping (good) part is strong: $R^2 = 0.84$ for the momentary values and $R^2 = 0.95$ for the 24 hour average surge values.

In the plot below the histogram of the surge values is shown. The blue histogram consists of all surge values, so once per hour for the period up to 1995, and once per 15 minutes for the period after 1-1-1995. In the analysis we determine the surge within a 24 hour period (12 hours before and 12 hours after) around the expected hydrograph peak. This is a more stable value than the highest surge.

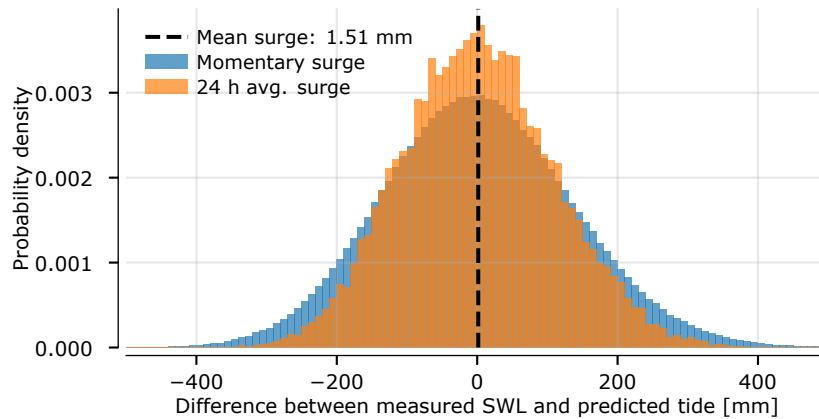


Figure F.2: Surge distribution for Lyttelton, both the momentary values and the 24 hour averages.

F.1.1 The choice for a 24-hour period

In this analysis we choose a 24 duration for averaging the surge. The choice for this time period is based on several aspects. The final goal is to relate rainfall and surge in such a manner, that we can model the effect of coincidence of the two on the water levels. So would we pick a very short duration like 1 hour, we would probably pick a surge value that does not coincide with the arriving of the resulting hydrograph peak at the estuary. If we would on the other hand pick a very long duration like 4 days, the hydrograph peak would surely fall in this period. There will however probably be no surge peak left after averaging over 4 days.

24 hours seems a good period, because:

- Current hydraulic modelling of the Avon River using design rain events of a certain ARI (eg: 50 years), shows that the downstream river level (and flow) are impacted most by the 24 hour rain event duration.
- We note on figure F.1 upper plot a typical 5-7 day wavelength in the surge pattern. This gives confidence that for most events the difference between the 24 hour average and a 12 or 48 hour average will be small.
- Early work tested 48-hour and 24-hour averaging and found negligible sensitivity of the results to this duration.
- We note on figure F.1 lower plot an apparent high amplitude of noise with a 3-4 hour wavelength. This is unexplained in case it is a fiction without a physical cause, then this encourages use of a long enough averaging window to smooth out this noise.

To describe the sensitivity qualitatively, if we picked a longer duration (eg: 48hr averaging) then we would reduce the noise and therefore expect to see a stronger correlation (higher correlation coefficient) but with lower surge values (less effect of the correlation). If we picked shorter (eg: 12hr averaging) then we would expect to see a weaker correlation (lower coefficient) due to more noise, but with higher surge values (larger effect).

Selection of peak events

The peak selection is based on the 2, 9 or 24 hour rainfall sum. To get the maximum rainfall sums, for every hour we sum the rainfall over the last 2, 9 or 24 hours. The selected peaks are thus the moment after this rain sum has fallen. We use the peak over threshold method to find these peaks. We look for all peaks within a 4 day time period (2 days before, 2 days after), to distinguish separate rain events. The threshold value is 5 mm, 7.5 mm, 10 mm for a 2, 9 and 24 hour rain sum. This approximates a 1 month ARI threshold for rainfall event depth.

Determine the time shift between rain and water level peaks

The selected peaks represent the moment at which the amount of rain fallen over the last 2, 9 or 24 hours is maximal. This does not mean that the rain intensity itself was maximal too at this final moment. So from this peak we search the time difference for the highest rain intensity within 2 days from the peak. Also we are interested in the expected moment the peak of the hydrograph reaches the river mouth and the moment of maximum surge, so we select the water levels and surge within 2 days from the peak too. The results are plotted below, per duration.

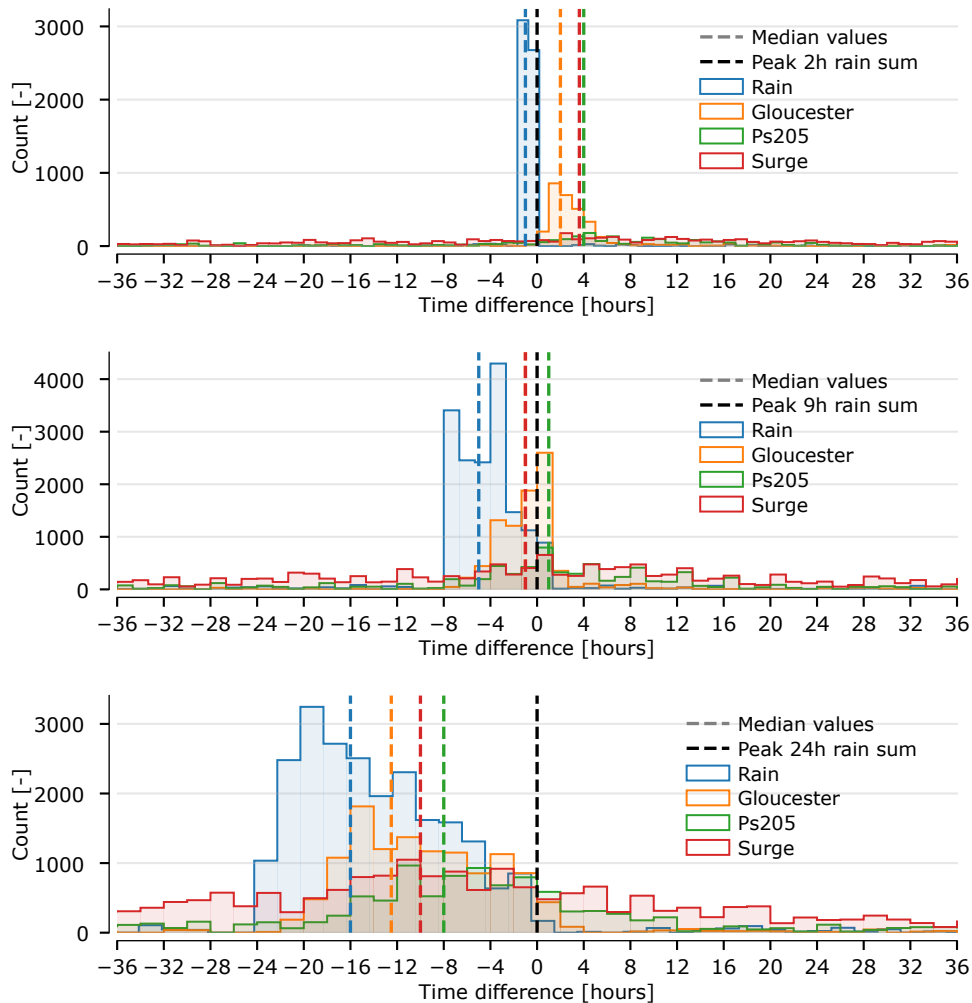


Figure F.3: Histograms of peaks of rainfall, surge and water levels relative to the peak of the rain sum. The top graph is for 2 hour, the middle one for 9 hours and the lower for the 24 hour rain sum.

For each of the figures, the 0 hour time difference position is the moment after the 2, 9 or 24 rain sum has fallen. It therefore makes sense to expect the highest rainfall intensity in a period within this length before 0. The vertical dashed lines show the median time difference for which the rain intensity, water level or surge is highest.

For the 2 hour case the peak rain intensity is at -1 hour. The max water level at Gloucester is at +2 hours, and at PS205 at +4 hours, as well as the surge. This surge and water level peak at PS205 are very spread out. The relation between rain and surge seems to be small.

For the 9 hour case the surge and water level at PS205 become more clear. Apparently these long events have a stronger relation with low air pressure, and the water level at PS205 is raised due to the prolonged rain.

The 24 hour case shows the clearest distribution of peak moments. The rain peak is usually early in the 24 hours, the median is -16 hours. The peak at Gloucester is around -12 hours, at Avon around -8. If we know the time the discharge wave reaches the estuary, we can look for the average barometric surge around this period.

To determine the timing of the hydrograph peak at Bridge Street, we extrapolate the time difference from Gloucester to PS205. Knowing that Bridge St is a similar distance further downstream, we get the following moment for the flow peak to arrive at Bridge St (Note that deriving this from the flow data at Bridge street itself, which might seem more logical, is difficult since the tide dominates the flow). For the 24 hour duration, the time difference between Gloucester (-12) and PS205 (-8) is 4 hours. We therefore expect the peak of the flow to arrive at the estuary at $-8 + 4 = -4$ hours. Note that this is still 12 hours after the median maximum rain intensity.

- 2 hour rain: +6 hours
- 9 hour rain: +3 hours
- 24 hour rain: -4 hours

A negative time difference (for 24 hour durations) might seem counterintuitive, but keep in mind that it is the time difference between the *end* of the rain event and the expected moment of the peak flow reaching the estuary. During a 24 hour period of large rainfall, the peak intensity is likely somewhere in the beginning of the event, such that the highest flow intensity reaches the sea before the end of the event. We derive the average barometric surge in a 1 day time window (before and after) around the moment the flow reaches the sea.

Kidson weather patterns

The Kidson weather patterns (Kidson, 2000)³, shown in F.4, are classifications of local atmospheric circulation patterns over New Zealand. These classifications provide some additional information of the analyzed peak events. The figure has been adapted from Smart et al. (2018).⁴

³Kidson, J.W. (2000) An analysis of New Zealand synoptic types and their use in defining weather regimes. *International Journal of Climatology*, 20, 299-316.

⁴Smart, G., Mullan, B. and Henderson, R. (2018) Coincident Inclement Weather Study for climate change planning. Prepared for Christchurch City Council. May 2018.

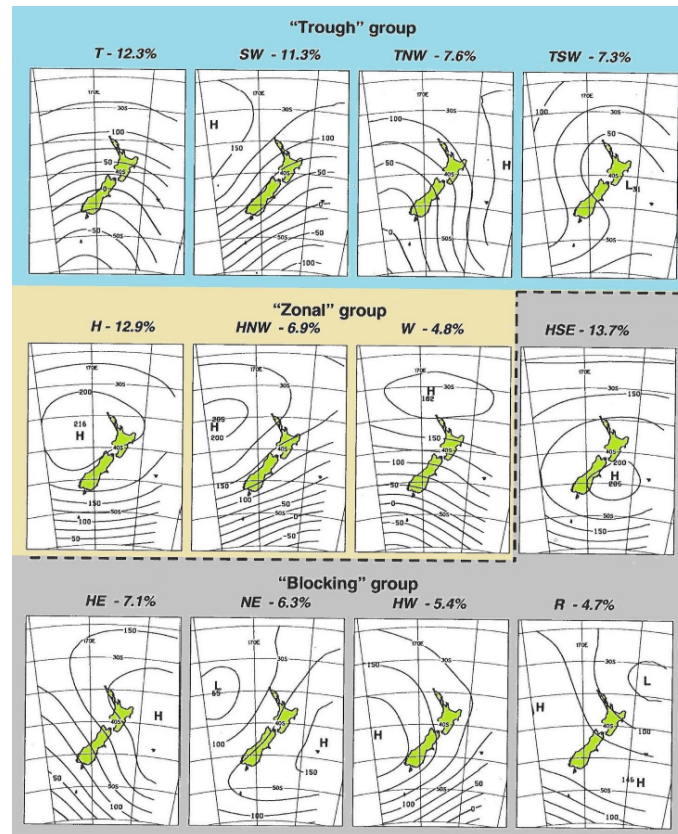


Figure F.4: Kidson weather patterns.

Unfortunately we don't have the classification of the Kidson pattern for all events, but for a lot of the most extreme events the Kidson pattern is known. In the plots below, the classification is shown for the largest events for 2, 9 and 24 hours.

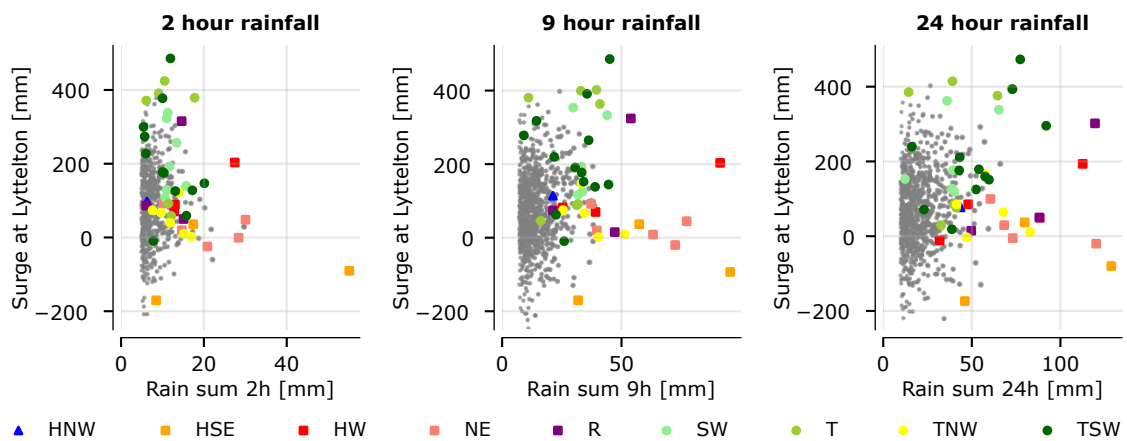


Figure F.5: Observed surge and rainfall values with Kidson classification for the event where a classification is available.

The scatters show a clear difference between blocking patterns (square markers), with on average

close to no surge, and the trough group (circle markers) that cause high surges. The trough leads to high rainfall amounts for long durations, while the blocking group can lead to very high intensities for short durations.

F.2 Creating a correlation model

Now we have gathered all information to determine the correlations between the rainfall and surge. The scatter plots of rain sum and average surge are shown below with the blue dots. By comparing the three plots you can see the relation between rainfall and surge becomes larger for long duration, which can be explained from the rain events that cause this.

To use the found relation in a model, we need to schematize the relation somehow. Our first guess is fitting a distribution to each part of the scatter, by calculating the percentile lines.

- Calculating the 5th, 25th, 50th, 75th and 95th percentile for a moving window, these are the grey lines on the background. For the large events these percentiles get more jagged since there are only few observations. In the left figure there is only a single observation left at the right end, so all 'percentiles' converge to this point.
- Fitting a third degree polynomial function through the percentile line for smoothing (black lines).

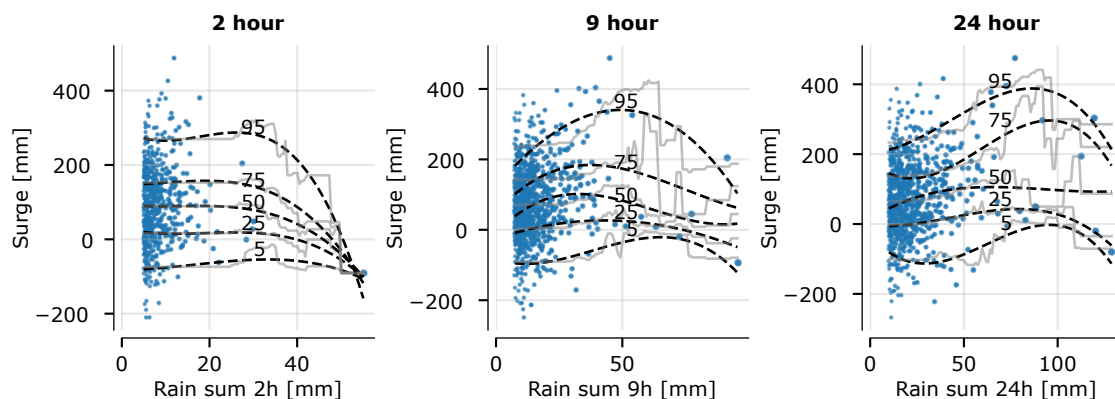


Figure F.6: Percentile fit to the observed surge and rainfall values. The five lines are polynomial fits through the grey lines on the background. The percentiles are labeled.

The fitted lines give a general understanding of the pattern:

- For low rain depths the surge is relatively low, but the average is still above zero.
- For intermediate rain depths, the surge is relatively high, particularly in the 24hr rain duration relationship.
- For the highest rain depths the surge is low again, with the average appearing to go negative in the 2hr plot, but this could be due to the small number of available points.

The problem with this approach is that we aim to derive boundary conditions for larger than measured rainfall depths. Extrapolating the polynomials would lead to unrealistically low surges.

Also the few observations for large events are not sufficient for estimating band widths. Therefore we try a different approach based on the Kidson weather patterns.

Ideally the correlation model has to fulfill certain requirements:

1. Fit the surge related to blocking and trough patterns;
2. Fit the total distribution of surge;
3. Reproduce the extreme value statistics for sea level rise, derived for Sumner Head.

For the rain events with the highest rainfall depths the pattern seems to be blocking in all the cases, so the model should fit blocking pattern for the highest rain depths. For low rain depth it should fit the total set ⁵. In between it can fit the trough patterns to some extent. So we create a model that divides the probability density between trough and blocking, and ends with all blocking for high rainfall depths.

The Kidson patterns marked in the figures above give us a small dataset with which we can derive the standard deviations for the blocking and trough events. We also add the fit of the total event set (events with rainfall above threshold) in the background. The concept of the correlation model is that for each rain intensity/depth a different ratio between blocking and trough gives the distribution of 24 hour surge values.

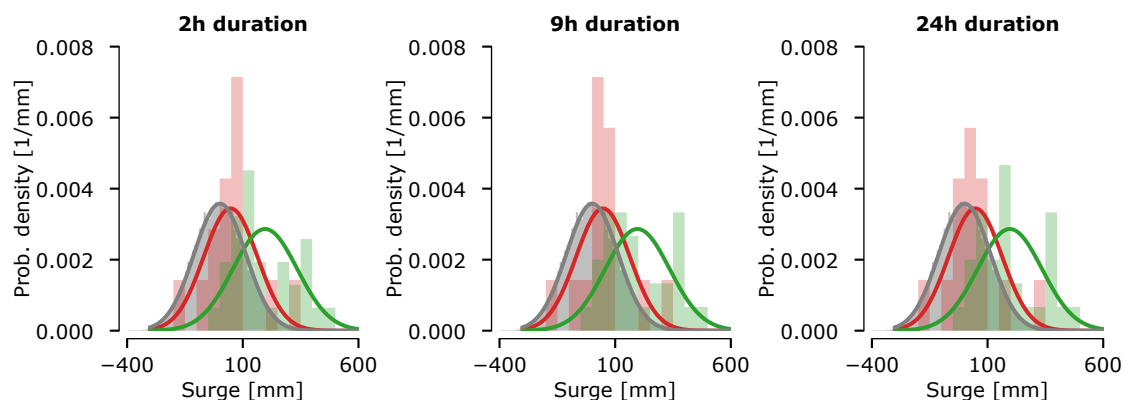


Figure F.7: Histograms of the surges for the three durations, conditional to the Kidson weather patterns. Red is blocking, green is trough. On the background the grey histogram shows the surge distribution for all events with rainfall above threshold.

First of all, there are very little data points in the set for which we have the Kidson patterns. It is therefore an assumption to use a normal distribution for the surge. It seems to be a good distribution for fitting the complete set, so we assume that would also be the case in which we have a classification for each rain event.

Because the difference between the calculated means and standard deviations is small, we use the average of the three for all durations.

⁵This means a surge distribution as shown in figure F.2, centered around zero. We created model will always have a above zero average. When deriving the boundary conditions we will see that this does not affect the results, since the extreme sea water level statistics are dominant in that range.

The next step is to fit this model to every bin of data. To do so we fit a bimodal normal distribution to every bin. The following figure shows the fits and the resulting weight for the 24 hour duration observations. The title of every subplot shows the range of rainfall depths for the bin, and the weight for the blocking and trough group.

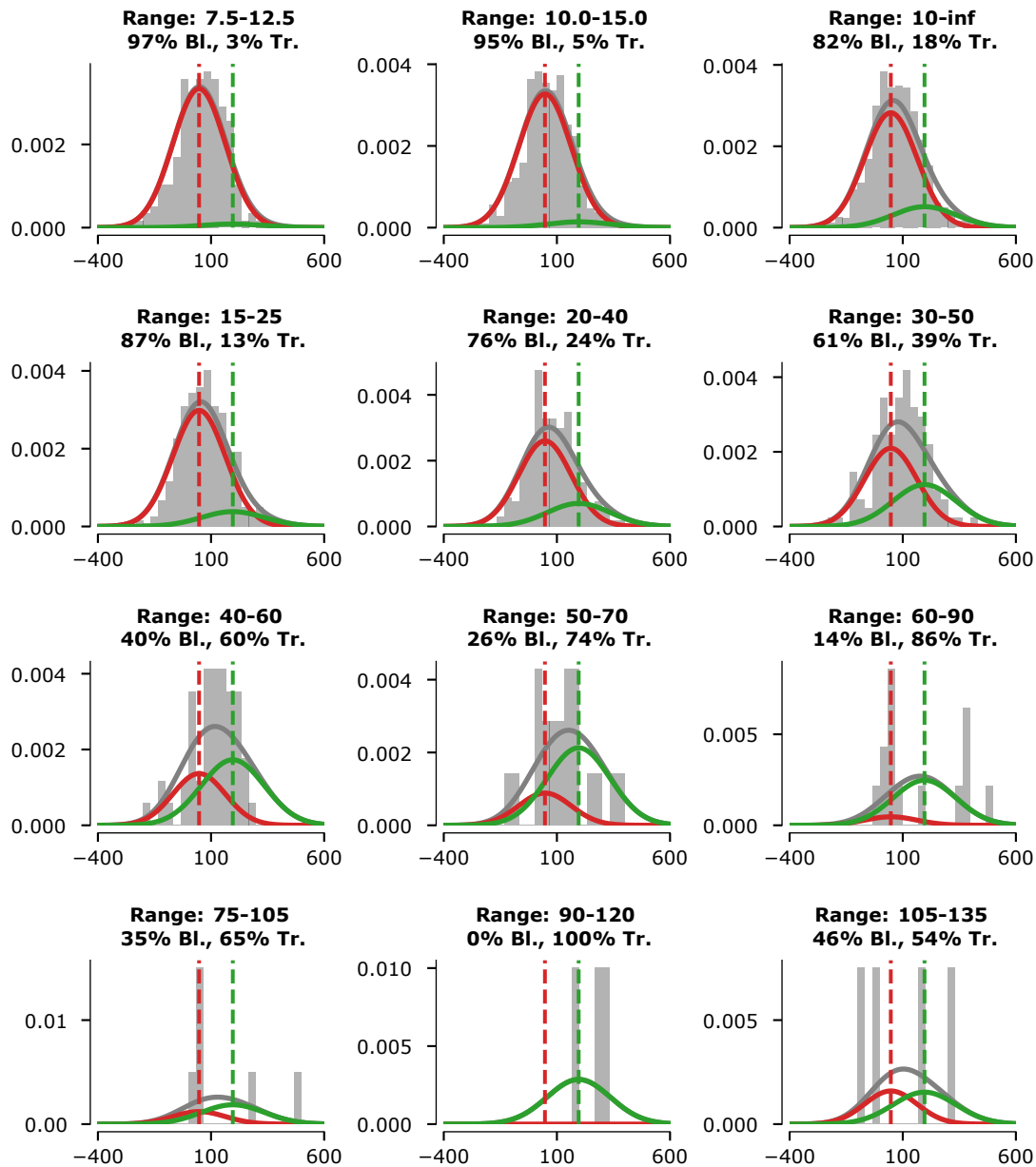


Figure F.8: Fits for the different bins of the 24 hour duration. The fitted bimodal distribution (grey) is the sum of the red and the green curve.

The fitted weights can be a bit irregular. For example the third last is all more blocking than the bin before and after. Also the last bin does not go back to all blocking, which is something we decided on for extrapolation. To solve both issues we fit a smooth curve consisting of two sine

curves to the weight pattern. The following figure shows the result, where the green area is the trough weight, and the red area is the blocking weight. The interface between the two planes is smoother than the weights.

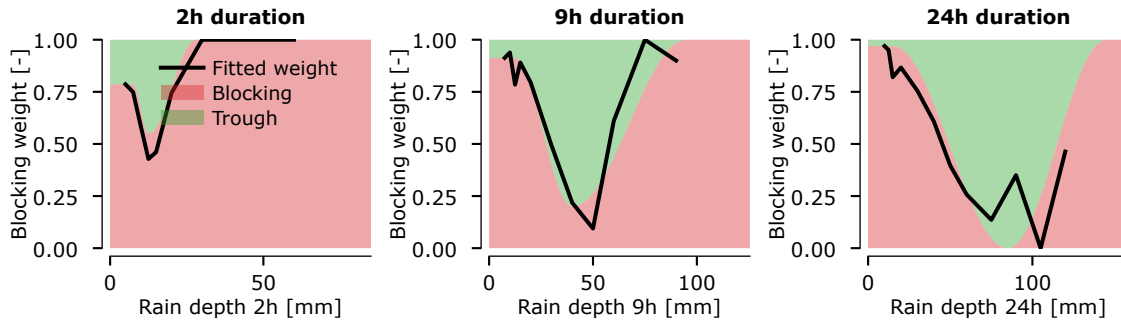


Figure F.9: Fitted weights for the three durations. The black line shows the calculated weights. The red and green zones are the smoothed contribution of blocking or trough per rainfall depth.

For all rain depths the expected surge distribution can now be derived. The result of this is shown in the figures below, by the black lines which indicate the 5th, 25th, 50th, 75th and 95th percentiles. By drawing a vertical line for a given rain depth you would get 5 percentiles. These 5 percentiles are the percentiles of the bimodal distribution of surge for the chosen rainfall depth.

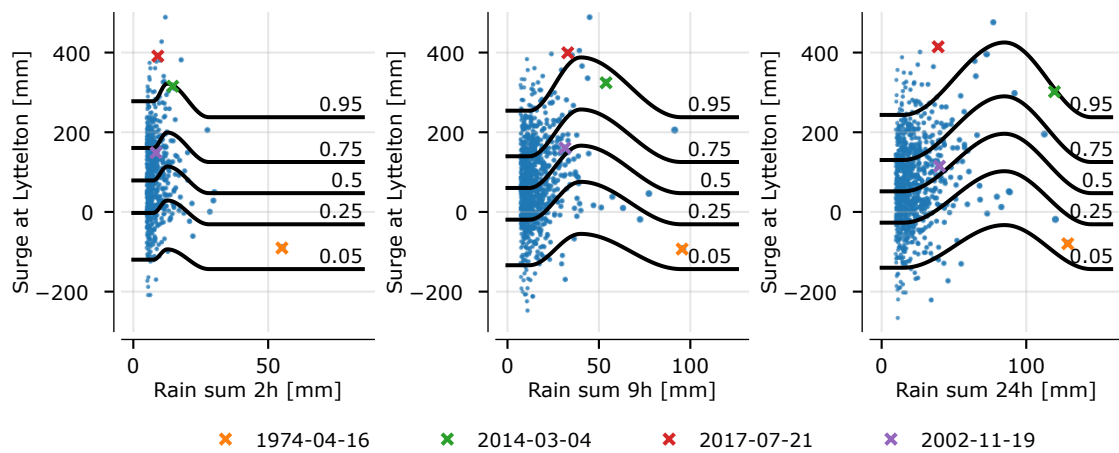


Figure F.10: Scatter plot with observations and fitted correlation model. The 5 black lines show the 5th, 25th, 50th, 75th and 95th percentiles of the fitted distributions.

The percentile lines are actually a modeled version of the 5 polynomial percentile lines we derived at first. The differences are also clear: the curves are much smoother, the lines retain a more or less equal distance and go back to the blocking distribution for high rainfall depths.

F.3 Sea water level with correlation

The model we just derived gives a correlation between the 24 hour average surge and rainfall. This needs to be translated to sea water level, by adding the tide and the difference between 24 hour average surge and peak surge.

To do so, we use the derived tidal constituents for Sumner LINZ A.1, which gives an empirical distribution for the high tide levels. Note that we use Sumner here, instead of Lyttelton for determining the surge; we are interested in the high tide levels at Sumner, and the Sumner NIWA data is good enough for analysing tidal constituents. After determining high tide levels the distribution for the difference between 24h average surge and peak surge is derived and added. The convolution integral of the three variables (24h surge, high tide, peak surge) is the distribution for the sea water level. The convolution process implies independence between avg. surge, peak surge and tide. For tide and 24h surge this is no issue, but there likely is a correlation between peak surge and 24h average surge. This correlation is however not determined or implemented in this model. This choice might give an underestimation of the total surge up to tens of millimeters, given that the standard deviation is 35 mm. We assume this is permissible in the context of the other model assumptions, such as the zero phase difference between peak surge and peak tide.

For Sumner we already have derived an extreme value distribution for the sea level, but this does not explicitly contain the correlation between rainfall and surge. By combining these three variables we produce sea level statistics that do contain this correlation, but for work of this nature it is typical that we will end up with a different probability distribution for the sea water level compared to the extreme value distribution.

In the range where the sea water level dominates the event (the rain is less important and so the surge correlation too) we can use the sea water level distribution directly. This we explain in more detail in the next notebook where the boundary conditions themselves are derived.

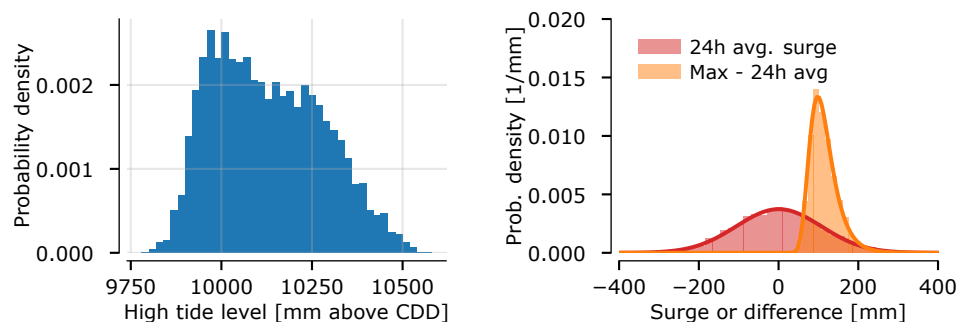


Figure F.11: On the left: histogram of high tide values for Sumner (LINZ). On the right: histogram with 24 hour average surge and difference between average surge and max surge.

The high tide values need to be combined with the 24 hour average surge and the difference between mean and max within these 24 hours. Note that the surge values in the figure above are all surge values in the measured series. There is no difference between blocking and trough patterns. Using the correlation model, the red distribution would shift to the right.

By combining the tide, the 24 hour surge (red) and the max surge difference (orange) into a convolution integral, we get the sea water level. For the situation without positive surge correlation (all events), the sea water level statistics are shown in the figure below. The second figure shows the comparison between this statistic and the derived sea water level statistics at Sumner.

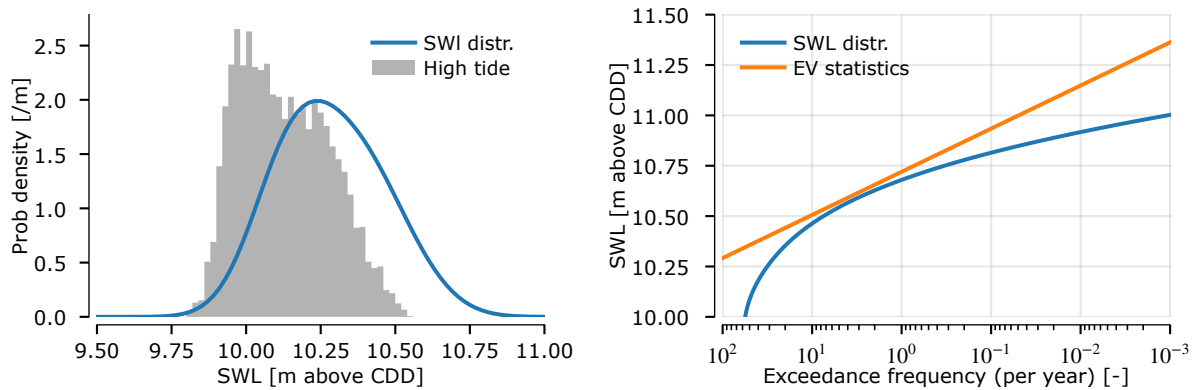


Figure F.12: On the left the histogram of the tide with the EV-statistics for no correlation (no rain, average surge). On the right the comparison between the EV statistic for this surge, and the separately derived EV-statistic for Sumner without FIG waves.

Both the statistics do not give the same extreme sea levels. Part of this can be explained by the fact that there is no correlation in the recently derived blue distribution, while in the total set (orange) there will be some correlated events. Another part is explained that the full sea water level can not only be explained with this model of tide and surge. By definition the sum of tide and surge is the sea water level, but that does not mean that the model of tide, avg. surge and peak surge assign the correct probability to high sea water levels: The blue line should be used to determine sea water levels while rainfall with surge correlation is dominant, the orange line when the sea water level itself is dominant.

F.4 Conclusion

This appendix described the derivation of the correlation model. We saw that we had to make some schematization choices to get a model that is applicable throughout the whole range of rainfall. The correlation is strongest for longer duration rain events that happen about once every decade. For these events the expected surge is on average highest. The surge from the correlation model however is a distribution: for every rainfall depth also lower and higher values are possible, however less likely. In the appendix about the boundary conditions this model is used to actually derive boundary conditions, by combining rainfall and sea water levels.

G Combining statistics to boundary conditions

All the statistical analyses done so far can be combined into sets of boundary conditions. This appendix goes through this process. First the definition of a boundary condition in this method is introduced. Second a single example is given of how to derive one such boundary condition. After this the method for deriving a total set of combinations for boundary conditions given a required ARI is presented. Finally we present some examples of boundary conditions that are derived with this method.

G.1 Introduction

A boundary condition is related to an annual return interval (ARI)⁶, which is equal to one divided by the annual exceedance frequency. For high return intervals (>10 year) this is equal to the annual exceedance probability (AEP).

For a boundary condition that results from a single forcing variable the meaning of this is clear; it is for example the rainfall that is exceeded once per 10 year, which will likely result in a once per 10 year exceeded water level. When we combine multiple random variables into a boundary condition it becomes a bit more complicated. First of all, two 0.1 AEP events (both rain and tide) occurring at the same time, will be more extreme than a 0.01 AEP event, because they also need to occur at the same time. Therefore we will define a 0.01 AEP event as a once per 10 year (0.1 AEP) event for rain, and a tide (or vice versa) that happens once every 10 such events. So the tide does not have a 0.1 AEP, but a 0.1 exceedance probability during the rain event.

Now let's say we have this specific combination that is exceeded once per 100 years. The water level will probably be more sensitive for one of the two variables. The water levels at upstream locations are more affected by rainfall, downstream location more by sea water level. Perhaps another combination that also had a product of exceedance probabilities of 0.01 leads to a higher water level. To get a more accurate estimate of a 100 ARI boundary condition, we need to integrate the probability of all events that cause a higher water level. This also means calculating the resulting water levels for all combinations, which can be a lot of work, depending on the used model. In this appendix we therefore try to estimate these conditions from the marginal variables and the correlation model. This is statistically more complicated, and it is based on the assumption that one of the 0.01 AEP combinations will give a 0.01 AEP water level (as discussed in G.4), but it does not require to calculate the resulting water levels.

Our definition of a X ARI boundary condition in this method is:

An X ARI boundary condition we call the combination of a Y ARI rainfall event combined with a sea water level that is exceeded during Y/X of the events, or vice versa.

So for example a 100 ARI boundary condition could be:

1. 1. A 20 year rainfall event combined with a tide that is exceeded during the $20/100 = 0.2$ highest fraction (20%) of these events.

⁶Note that we will also use the term return period from time to time. With this we mean exactly the same as with ARI.

1. A 5 year sea water level event combined with a rainfall that is exceeded during the 5/100 = 0.05 highest fraction (5%) of these events.

To derive the boundary conditions we follow these steps:

1. Pick a desired return period for the total event.
2. Pick the return period for the rainfall, and the corresponding duration.
3. From the chosen rainfall statistic, get the distribution of the surge.
4. Combine the surge distribution with the tide statistics to get a conditional sea level statistic.
5. Pick sea level that matches the remaining probability.
6. Use these data to get the time series that matches the peak values resulting from the previous steps.

These steps are described one by one in this appendix. Note that we always start with rainfall, since we derived a correlation model in which the surge is a consequence of the rainfall.

G.2 Different statistics within the event

The starting point for deriving the boundary condition is choosing the rain statistics. We choose to use the Christchurch Botanical Gardens rain statistics from HIRDS as representative for the catchment. See the appendix on rainfall for an explanation for this choice. In this example we illustrate a **10 ARI** rain event with a **24 hour duration**, which is part of a **100 ARI total** event.

Rainfall statistics

The rain statistics provided by HIRDS are only valid for return periods larger than 1 year, the formula cannot handle return periods of less than a year. Since we are also interested in events that occur for example twice a year, we extend this statistic by selecting the peaks from the rainfall statistics we derived for this study. This is the weighted statistic for the Avon catchment. The used station or combination of stations does however not really matter, since we scale the Avon catchment statistic to align HIRDS at an ARI of 3 years. Note that we use the HIRDS statistics from Botanical Gardens because it is the most conservative (the differences are small). For the < 3 year ARI we use the catchment wide rainfall pattern, because it is the most complete and representative data source. Also the derived statistics are scaled to join the HIRDS statistics at a 3 year ARI, so the used location makes little difference.

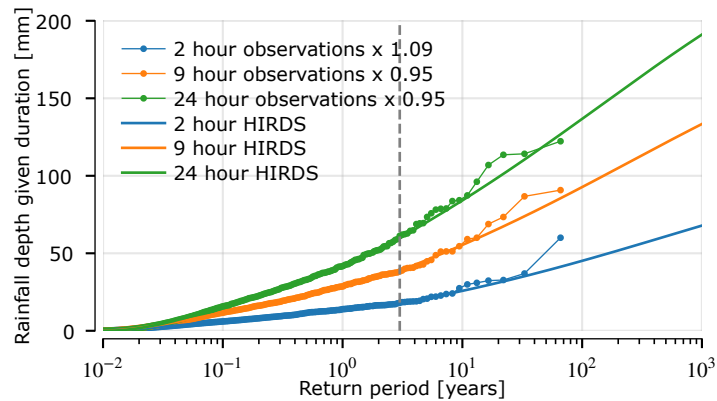


Figure G.1: The combined rainfall statistics. The smooth right part is the HIRDS statistic. The dots on the left are the statistics derived with peak over threshold for lower return periods.

The scale factors are shown in the legend.

For return periods smaller than once per 3 years we interpolate between the observations. Since there are so many realizations there is no need for curve fitting in this range. For larger than 3 year return periods we use the HIRDS formula.

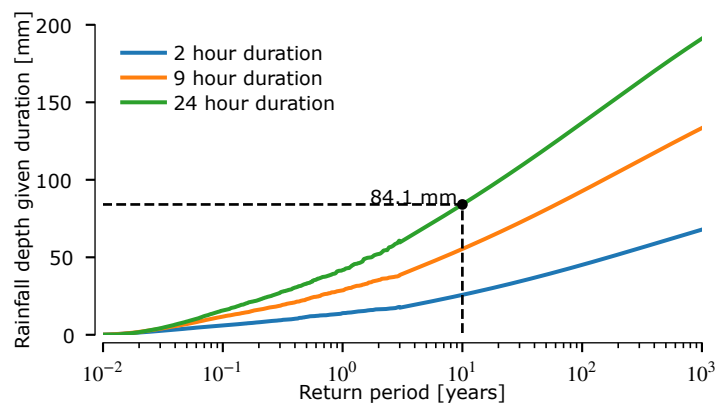


Figure G.2: Reading a 10 year ARI event from the rainfall statistic. The resulting rainfall depth for this duration is 84.1 mm over 24 hours.

The resulting rainfall depth in the 24 hour duration is 84 mm. We can read this from the graph by drawing a vertical line from the 10 year ARI (10^1), and looking for the intersection with the green curve, that represents the 24 hour statistic.

Surge statistic

From this rainfall depth/ARI we can read the corresponding surge statistic from the correlation model. The figure below shows the percentile lines from the correlation model for the 24 hour rain duration. The red line in the 84 mm rain depth. The distribution of the surge given the 24

hour duration and 84 mm rainfall depth is shown in the right figure. The black dots correspond with the percentile line.

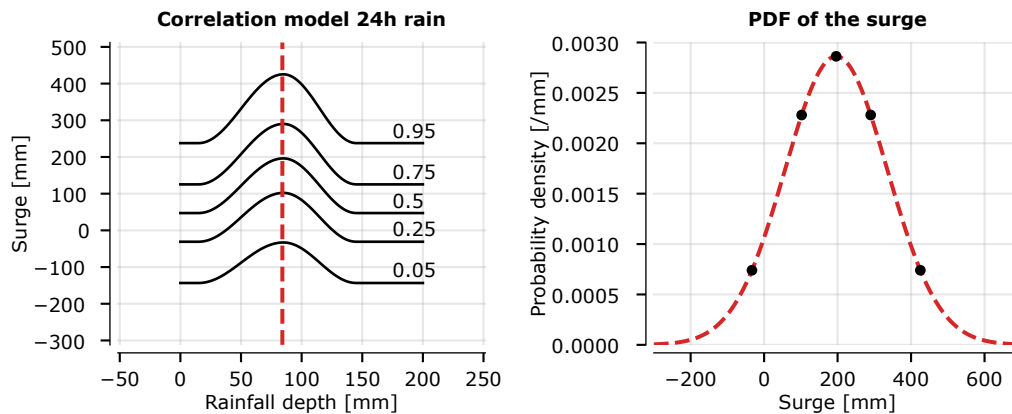


Figure G.3: Deriving surge for a given rainfall depth. The left figure shows the correlation model for a 24 hour rain duration. The right figure shows the resulting PDF when the surge is derived for the red line in the left figure.

The next question is how to combine the surge pattern with the tide in the boundary condition. The next figure shows both tidal patterns and a surge patterns, and some examples of how they can be combined.

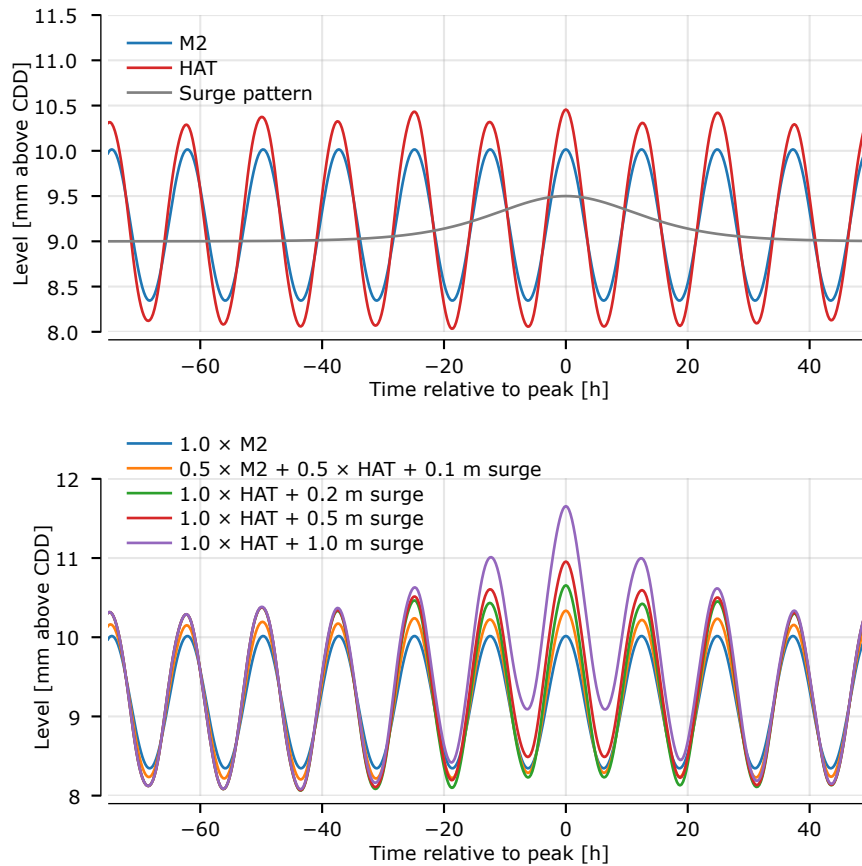


Figure G.4: Templates for surge patterns. The top graph shows the three base components. The bottom graph shows some combination of the top components, for different sea water levels.

The surge is modeled with a standard pattern (sech^2 , a logistic distribution), shown with the grey line in the first of the above figures. Furthermore we also have the M2 tide as the lowest realization of the tide, and the HAT (Highest Astronomical Tide) as the highest realization of tide without surge. Depending on the required sea water level, for a SWL below HAT + 0.2 meters and above M2 + 0.2, then, a linear combination of M2 and HAT is used as the base sinusoid with the combination factor in proportion to proximity to either limit. For higher than HAT + 0.2 meters, the base sinusoid is HAT. In the hypothetical case of sea level below M2 + 0.2m then the base sinusoid is M2. In all cases the necessary surge height is added to the base sinusoid to generate the final time series. The second of the two figures above shows some generated combinations for different sea water levels.

Tide statistics

Surge is the component in the sea level that is not explained by the astronomical tide. So to get the extreme value for the sea level itself, we combine the surge with the probability distribution of predicted high tide values. A histogram of the high tide levels at Sumner (LINZ) is shown in the figure below. These are *predicted* values, so not measured. Sumner LINZ is chosen (not NIWA) because of its better data quality, which gives a better estimate of the tidal constituents. We do not use Lyttelton because we need the boundary conditions at Sumner, and the data are

sufficient for estimating high tide values. We use the histogram as a probability distribution for the high tide value.

Sea water level statistics

By making combinations of high tide and surge on a regular grid, and assigning probabilities to it, we can derive the exceedance frequency of the sea water level, *given* the rain depth and duration. The figure below shows the joint probability distribution of tide and surge. The orange line is the one we use to model the sea water level, which is the convolution integral of high tide, 24 hour surge and surge peak. More information on this is given in the appendix on correlation models. Note that the surge is positive for this rainfall intensity, that is why the blue distribution (high tide + 24 hour surge) is on the right of the green distribution (high tide).

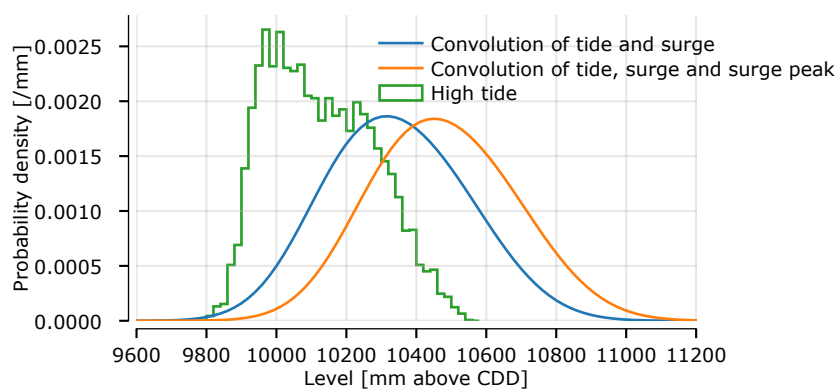


Figure G.5: High tide, convoluted with surge and convoluted with surge and surge peak.

Now the sea water level statistics are known for the given surge distribution, we can read the value we need for the given exceedance probability. The blue lines result from an integration of the orange line in Figure 5 above. The red dashed line shows the (remaining) exceedance probability for the sea water level. In our example this is 0.1, or 10%, since we the total event is 1/100 year and the rain event is 1/10 year. This 10% is not a once per 10 year return period, but the once per *this specific rain event* return period. The left and right figures are similar, except for the log scale on the x-axis.

Now the sea water level statistics are known for the given surge, we can read the value we need for the given exceedance probability. The red dashes line shows the (remaining) exceedance probability for the sea water level. In our example this is 0.1, or 10%, since the total event is 1/100 year and the rain event is 1/10 year. This 10% is not a once per 10 year return period, but the once per *this specific rain event* return period, so it is a conditional probability. The left and right figure are similar, except for the scale on the x-axis.

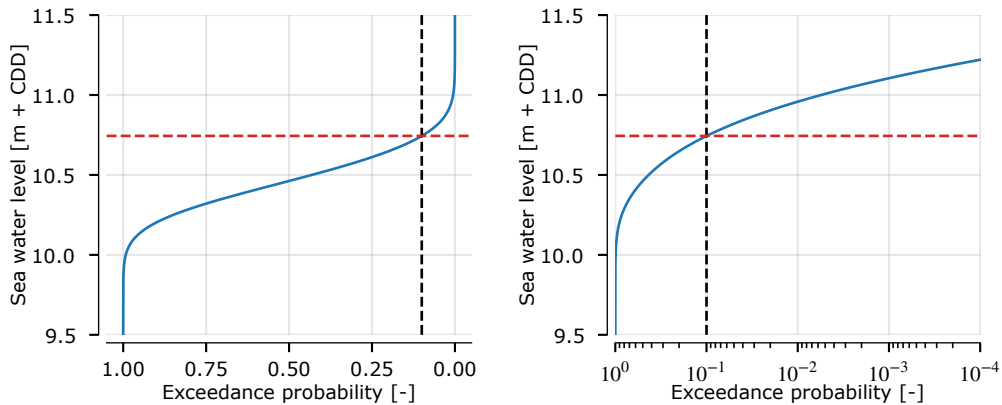


Figure G.6: Reading the sea water level with the remaining exceedance probability from the derived sea water level statistics. The right graph is limited to 10^{-4} , so the sea water levels do not continue up to 11.5 m+CDD.

The resulting sea water level, approximately 10.75 m above CDD is, modeled in the same pattern as currently used by CCC, which were already shown above. The surge is calculated by subtracting the highest sea water level by the highest standard tide. The standard surge pattern is multiplied with this surge value, and added to the standard tide.

In the correlation analysis the time differences between the peaks of rainfall, surge and confluence were analyzed. We can use these values to set the timing of the rainfall event in relation to the peak of the surge. However there is a lot of spread in these timings, particularly between the surge and rain timing for the shorter rain durations. Therefore we advice to set the timings such that the resulting peak of the water levels reach the estuary during high tide and high surge, the most conservative (but not unlikely) case. We added the rainfall period for which we expect this peak coincidence to the plot, but it is up to the modeler to choose these phases, taking into account the shape of the rain event.

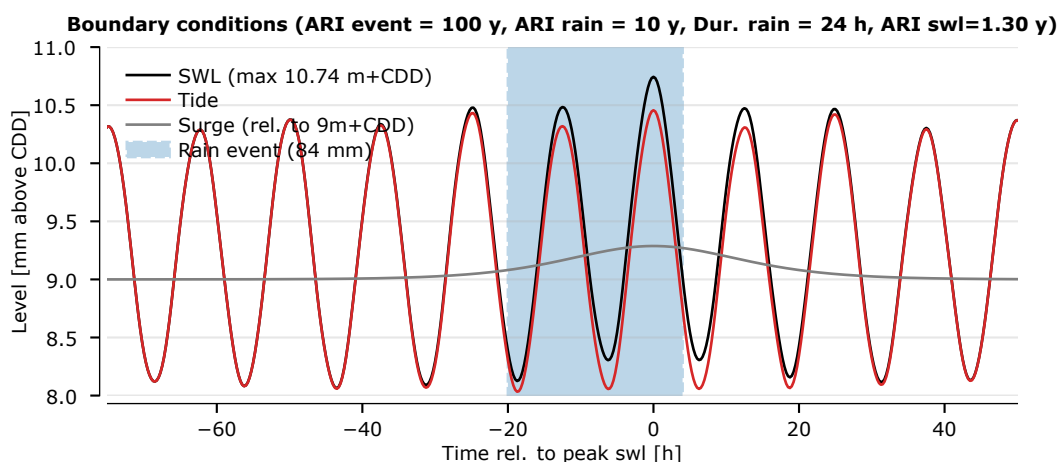


Figure G.7: Boundary conditions for a 100 year ARI event, with a 10 year rain depth and a once per 1.3 year sea water level/

The calculated sea water level is roughly equal to a once per year sea water level. The reason

this is not a once per 10 year sea water level, is that we determined the statistics given the rain characteristics (giving a correlation to surge), and for the period in which the rain event takes place. An ARI represents the frequency with which an event is observed in a year, which is naturally higher than the frequency the sea water level is observed during the period of a rain event.

G.3 Different events for a required ARI

In the last chapter an example was given of a single 100 year ARI boundary condition. There are however multiple combinations that we would refer to as a 100 year boundary condition. Also remember that we derived sea water level statistics given the expected amount of rainfall and the corresponding surge, but if the sea water level is much more extreme than the rain event we need to use the separately derived sea water level EV-statistics themselves: it is a better representation of extreme sea water levels. In this section we describe these two ways of deriving the conditions, and what we need to do to merge them into a single set of boundary conditions.

In figure G.8 different combinations of rain events and high sea water level events are shown for a duration of 24 hours.

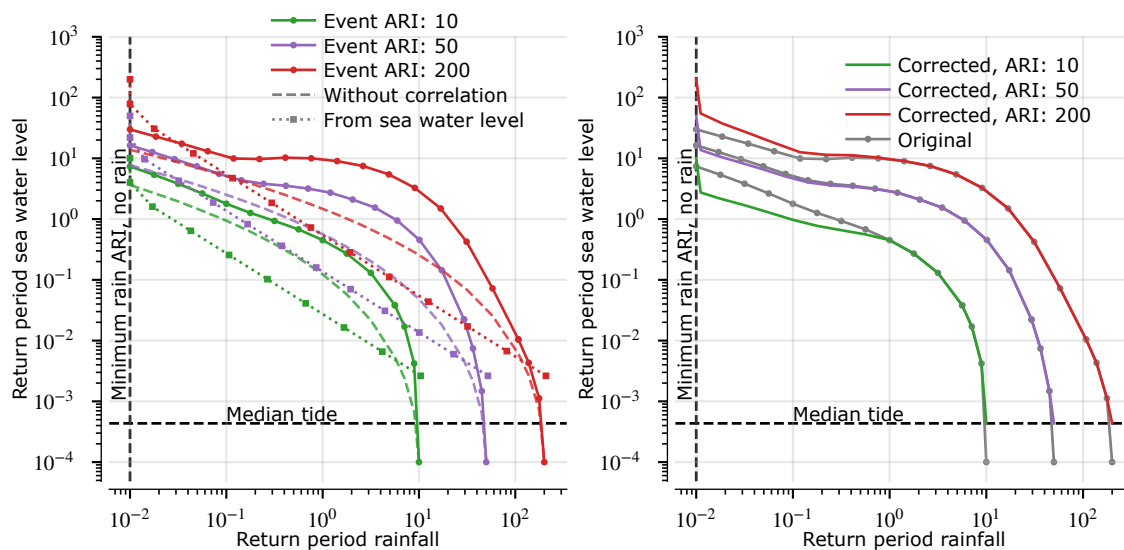


Figure G.8: Different possibilities for deriving boundary conditions for 10, 50 and 200 ARI events.

First focus on the left figure. The solid line with the round markers is closest to the end results, but to explain how it is created we start with explaining the other lines.

- The dashed lines without markers in the background show events that are derived from rainfall return period by finding a matching sea water level *without* correlation. Pick a rainfall return period and a total event probability, assign the remaining probability to the sea water level *during* the event, and you get the resulting lines. There is however no correlation and the sea water level statistics during the event do not necessarily match

with the EV statistics derived at Sumner. This we can see on the left axis, where the 200 year event for example does not lead to a 200 year sea water level. That the sea water level statistics during a rain event do not match the sea level statistics at Sumner is the reason that the uncorrelated lines are not straight.

- The solid line with round markers is similar to just described, but with rainfall-surge correlation. This leads to higher sea water levels particularly around the once per 10 year rain events, because the correlation is strongest in this range. The line however also does not match the sea water level statistics at the left axis.
- To solve this we also derived the conditions from starting with the EV sea water level statistics and find a matching rainfall event for this. This is represented with the dotted line with the square markers. It shows some peculiar things. First of all the left ends of the curves are vertical. This is because any rain event that happens more than 100 times a year is *no* rain at all. On the right end the lines do not end on the lowest tide. This is because the EV statistics for Sumner gives a higher value for an everyday tide than the statistic in the correlation model.

For the largest part the solid line is fine. However it doesn't match the EV statistics for Sumner in case of a dominant sea water level. Also the line isn't necessarily monotonously decreasing (or increasing, depending on the side you are looking from), and the lowest sea water levels are below median tide. Therefore the following corrections are made to the curves:

1. Log-linear shift the correlated statistic for rainfall return periods lower than once per year towards the EV sea level statistic for Sumner. The once per year is chosen because from that point the sea level becomes (much) more important (extreme) than the rainfall.
2. Make sure the graphs are monotonously decreasing. Otherwise, a lower rain return period cannot be combined with a lower sea water level than the higher return period. If not monotonously decreasing, increase each value until it is larger or equal to the next value.
3. Shift all values below median (M2) tide, 10 m above CDD, towards median tide.
4. The minimum sea water level that corresponds to no rain gets some slight rain, (90 days per year instead of a 100 days per year rain event) mainly in order to clarify the graph.

After doing these correction we end up with figure G.8 on the right in which the corrected lines (colored) are plotted on top of the original lines (grey). Doing these corrections for all durations gives us the following lines with possible boundary conditions:

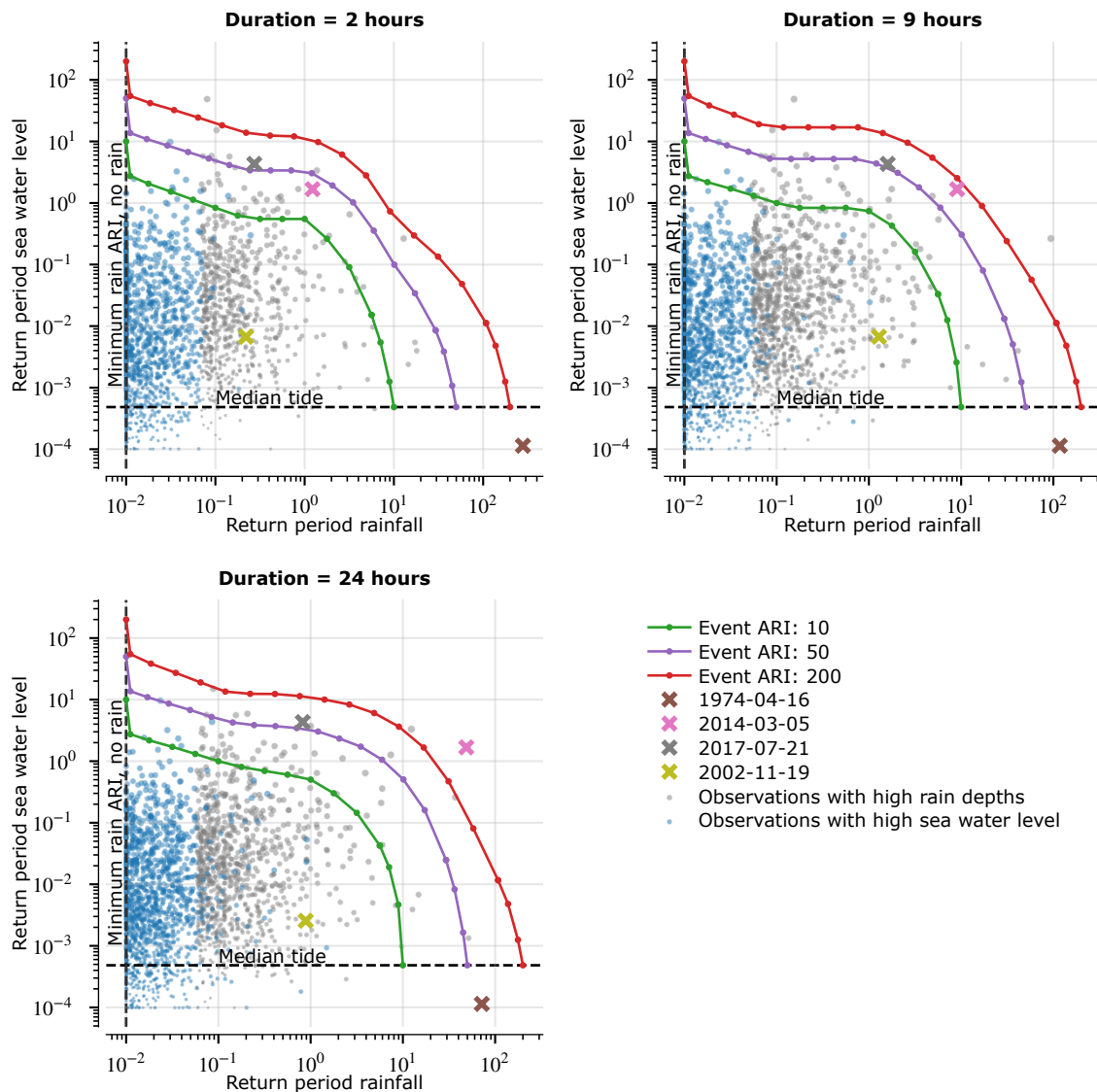


Figure G.9: Possible boundary conditions for 10, 50 and 200 ARI for a 2-hour, 9-hour and-24 hour duration rain event.

The gray dots within the line are the rain events selected for the correlation analysis with depth exceeding threshold and a threshold in each case of approximately 0.05 ARI. The blue dots the measured high sea water level events, that are not already described with a grey dot. The few blue dots in the grey area can be explained because the time windows with which the events are selected and compared are not exactly the same; the starting point (the moment of a high sea water level or high rain sum) defines the period in which the other is searched. Also the surge (grey) has been determined at Lyttelton (long data series), while the extreme sea water levels have been selected at Sumner (the location for the EV statistics).

Figure G.9 also shows four observed past events with a high rainfall depth, a high surge or both. The event observed in March 2014 was by far the most extreme, where a high rainfall, high surge and high tide coincided. One peculiar point is the April 1974 event, which gave a very high rainfall depth, combined with a very low tide and no surge. The four events are the same as the

ones shown in figure F.10 (but with different colors).

Since we have about 60 year of rain measurements, we wouldn't expect an event higher than the 200 ARI line. However we see 2 or 3, depending in the rain duration. This is still possible, the nature of a probability is uncertain. But we also see tens of events above the 10 ARI line, which is very (extremely) unlikely. This underestimation of combined events is caused by the fact that the resulting observation are a combination of two random variables. To illustrate this, consider the more simple case with two exponential distributions below:

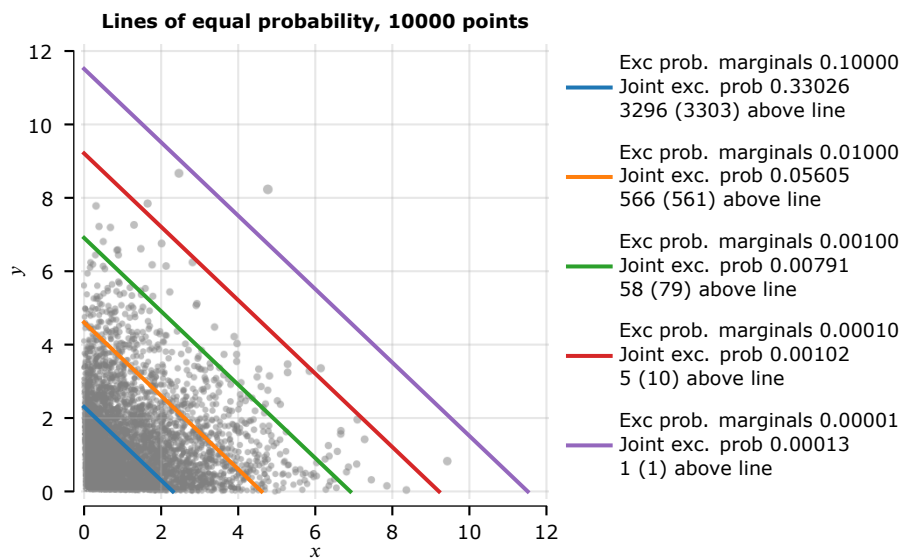


Figure G.10: Example of exceedances for an uncorrelated joint exponential distribution. The legend shows the exceedance probability as product of the marginals. The second line shows the exceedance probability of the joint distribution function, and the third line shows the number of realizations above the line, with between parentheses the expected number based on the joint exceedance probability.

The joint probability density function is:

$$f(x, y) = e^{-(x+y)} \quad (14)$$

The exceedance probability can be calculated with:

$$P(X + Y > x + y) = \frac{1}{T} + \frac{1}{T} \cdot \log(T) \quad (15)$$

It goes beyond this appendix to give the derivation, but the 'Exc. prob' values in the figure legend are calculated with this formula, and we can see that they match the fraction of points above the line. This formula leads to 5.6 times as much exceedances for a 100 ARI case, in relation to a single exponential random variable. For a single random variable it would be $1/T = 1/100$, 1%, but the additional $1/T \cdot \log(T) = 0.046$, an additional 4.6%. This is due to the exponential (non-linear) scale of the PDF's. The exceedance probability of two moderately extreme events is larger than the exceedance probability of a single twice as extreme event.

The example shows the error we would make in case of two exponential distributions without a correlation, by multiplying the two marginals instead of calculating the exceedance probability of the joint probability distribution. We can also calculate what "ARI" level to use, in order to achieve an exceedance probability that matches the searched ARI. The example with two exponential distributions is used as a proxy the probabilistic model we created, in which rain and tide are correlated. Both the extreme rainfall and extreme sea water levels have more or less exponential tails (more or less since the HIRDS statistics consists of several exponential terms). Ignoring the correlation might lead to a small overestimation.

If we apply the proxy to the three ARI's under consideration, we find the following required ARI's on the third row that match the searched exceedance probabilities:

Searched ARI for boundary conditions	10	50	200
Corresponding exc. prob. in the joint exponential model	0.330	0.098	0.031
Required ARI to match the exceedance probability	49	342	1686
Resulting exc. prob. for the required ARI in the joint exponential model	.100	0.020	0.005

Note that the scaling factors are derived from the (assumingly representative) case with two exponential distributions. If we would (and could) derive those for our statistical model, they would be a different. This would be the result of the correlation which would lead to bit lower scaling, and the different marginals which would lead to a bit lower of higher result.

If we add the lines corresponding to the required ARI's to the 24 hour duration plot, we get the following figure:

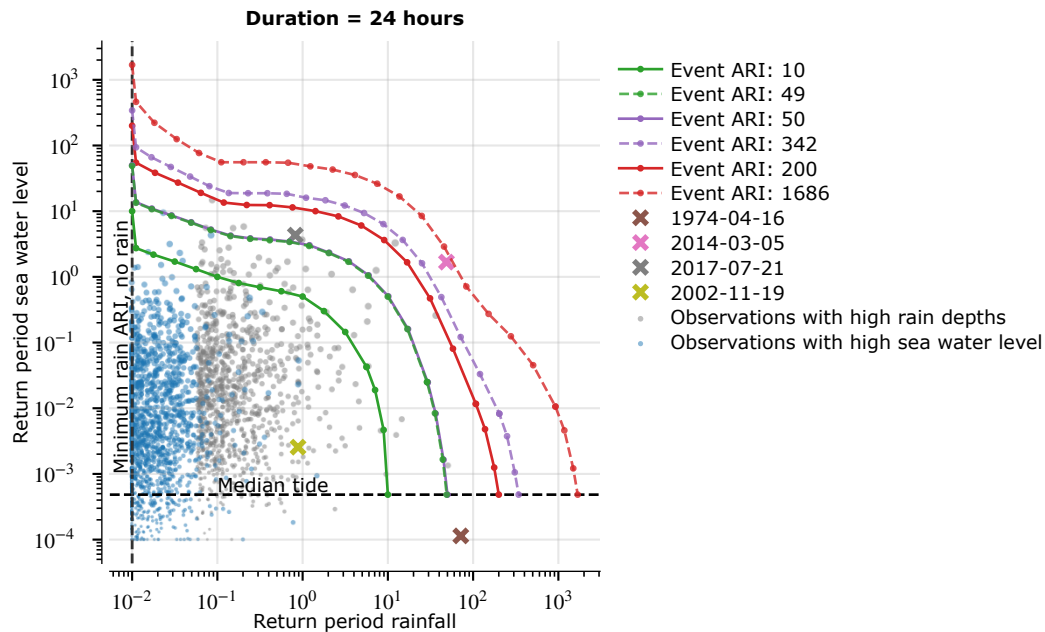


Figure G.11: Possible boundary conditions for 10, 50 and 200 ARI for a 24 duration rain event, including scaled lines.

The dashed lines are the scaled version of the solid line with the same color. Note that the green dashed line and purple solid line are almost on top of each other, since their ARIs are very close, 49 and 50 years. If we look at the dashed lines, the number of points below the line makes more sense. There are no points above the scaled 200 year line, only one above the scaled 50 year line and approximately 10 above the once per 10 year line. This agrees better with the 60 years of data we have. We will therefore use the scaled lines to pick boundary conditions when both boundary conditions are important. When only a single boundary is of importance, we do not use the scaled lines.

The numerical values for the 6 lines in figure G.11 are presented in tabular form at the end of this appendix, also for the 2-hour and 9-hour durations.

G.4 Picking boundary conditions

There isn't a single combination of rainfall and sea water level that defines a 200 ARI. In fact, we can pick infinite points on the 200 ARI lines. Each of these points gives a different combination of sea water level and rain. In this section we present a few thoughts on how to pick points on this line.

As the example at the start of this appendix showed, the time series of the boundary conditions are chosen such that they correspond with the maximum sea water level and rainfall depth. A ratio of tide and surge is chosen, but this is not directly related to the surge and tide from the correlation model, so that the approach is similar to the currently used approach in Christchurch. This current approach is to pick two boundary conditions with an equal ARI, calculate the water levels and pick the maximum of these water levels.

In this section we present a few thoughts on how to pick from the infinite choice of points on this line. The figure below shows a diagram for picking boundary conditions with a 200 year return interval. The thin black dashes lines indicate the 200 year sea water level, 200 year rain fall, and the 1:1 ARI line. If we are looking for a 200 ARI event, the realization of both marginals (sea water level and rainfall) should not be higher than 200 years. We did however see that we should pick a more extreme event in the joint range. With these thoughts in mind we present three options, indicated with the numbered markers in the figure:

1. The 200 ARI single driver event on the marginals (rain and sea water level), and the combined event on the intersection of the scaled line and the 1:1 line in the joint range.
2. The 200 ARI event of each marginal, combined with a realization of the other variable on the scaled line.
3. A more conservative approach. The 200 ARI event of each marginal combined with the realization of the other variable that is equal to the third point in the first option.

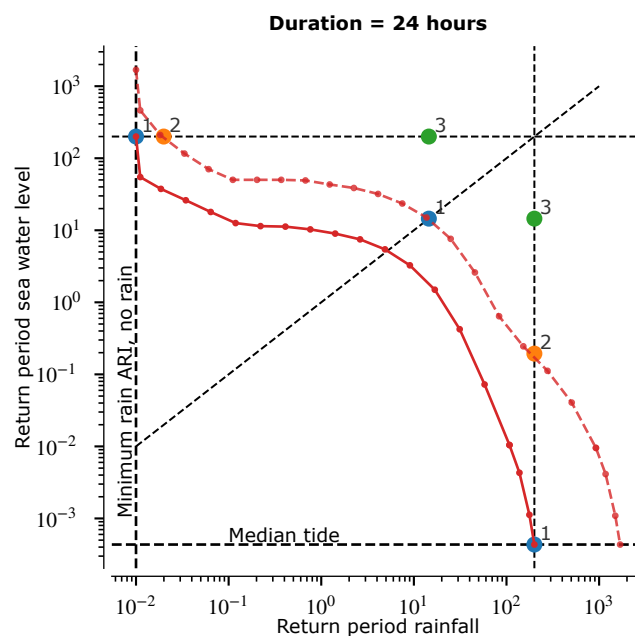


Figure G.12: Diagram for picking 200 year ARI boundary conditions.

The first option results in three combinations with varying rain and sea water levels, which is why we plot the time series in the figures below.

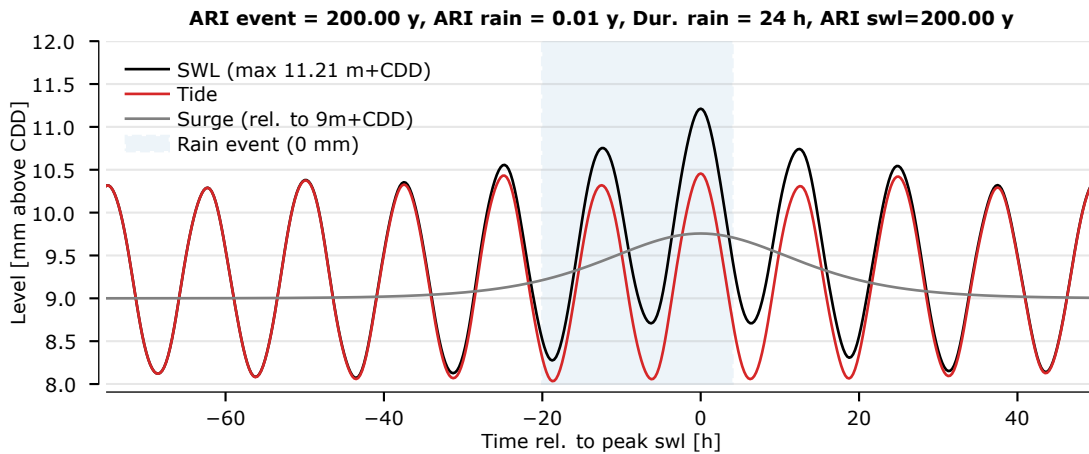


Figure G.13: Boundary condition with an everyday rain and a 200 year sea water level

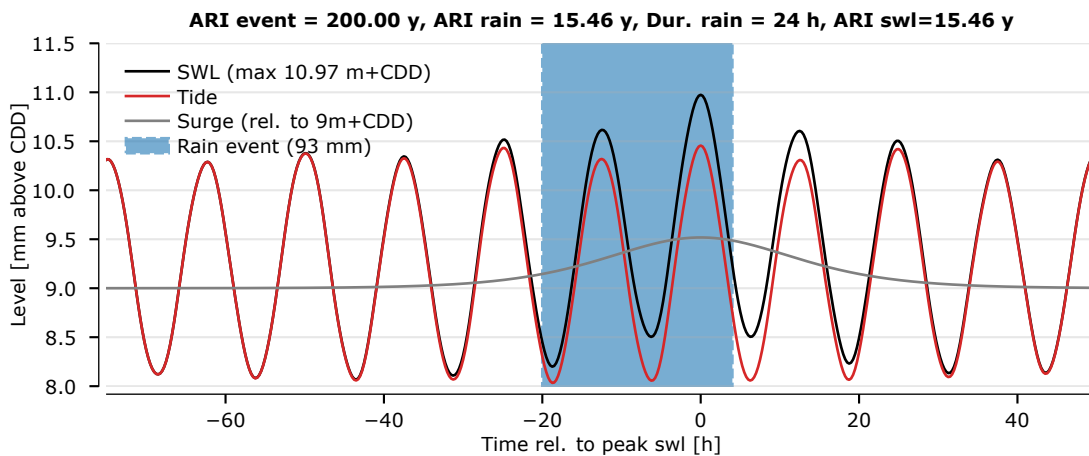


Figure G.14: Boundary condition with a 14.5 ARI rain and a 14.5 ARI sea water level

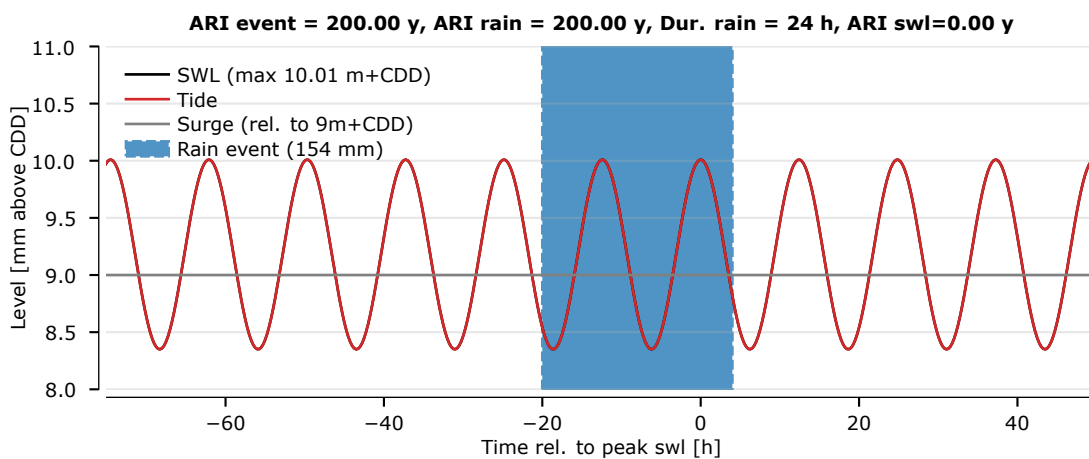


Figure G.15: Boundary condition with a 200 ARI rain and an everyday sea water level

The depth of the blue color is an indication of the rain intensity. As reported earlier the rain period is only an indication from the data. It is up to the modeler to set the timing such that the peak of the resulting water levels coincides with the highest sea water level. By using this method of picking boundary conditions that agree with a certain return period, and subsequently using the maximum water level, we hope to get a good indication of an T ARI water level. The method does however not guarantee that this will indeed be the result. For a more accurate estimate of an event that exceeds the resulting X ARI water level, the probability of the combinations that lead to higher water levels need to be integrated. This requires calculating these water levels on beforehand (a more advanced methodology).

Table G.1: Numerical ARI values of the contour lines for a 2-hour rain duration.

	10 ARI		50 ARI		200 ARI	
	Rain	SWL	Rain	SWL	Rain	SWL
0	1.00e-02	1.00e+01	1.00e-02	5.00e+01	1.00e-02	2.00e+02
1	1.11e-02	1.00e+01	1.11e-02	5.00e+01	1.11e-02	2.00e+02
2	1.70e-02	1.00e+01	1.69e-02	5.00e+01	1.83e-02	2.00e+02
3	2.89e-02	8.40e+00	2.84e-02	5.00e+01	3.33e-02	2.00e+02
4	4.92e-02	6.58e+00	4.79e-02	4.70e+01	6.08e-02	1.74e+02
5	8.37e-02	5.17e+00	8.07e-02	3.69e+01	1.11e-01	1.22e+02
6	1.42e-01	4.06e+00	1.36e-01	2.83e+01	2.03e-01	8.60e+01
7	2.42e-01	3.33e+00	2.29e-01	2.26e+01	3.70e-01	6.81e+01
8	4.12e-01	3.32e+00	3.86e-01	1.99e+01	6.75e-01	6.02e+01
9	7.00e-01	3.32e+00	6.51e-01	1.96e+01	1.23e+00	4.81e+01
10	1.19e+00	3.00e+00	1.10e+00	1.70e+01	2.25e+00	3.54e+01
11	2.02e+00	1.90e+00	1.85e+00	1.27e+01	4.11e+00	2.21e+01
12	3.44e+00	1.02e+00	3.12e+00	8.41e+00	7.50e+00	7.63e+00
13	5.86e+00	3.58e-01	5.25e+00	4.13e+00	1.37e+01	2.30e+00
14	9.96e+00	9.96e-02	8.85e+00	1.30e+00	2.50e+01	1.49e+00
15	1.00e+01	3.39e-02	1.49e+01	5.91e-01	4.56e+01	9.13e-01
16	1.00e+01	8.50e-03	2.51e+01	3.44e-01	8.32e+01	5.22e-01
17	1.00e+01	8.50e-03	4.24e+01	1.84e-01	1.52e+02	2.73e-01
18	1.00e+01	3.84e-03	5.00e+01	8.62e-02	2.00e+02	1.25e-01
19	1.00e+01	1.07e-03	5.00e+01	3.25e-02	2.00e+02	4.52e-02
20	1.00e+01	4.84e-04	5.00e+01	8.26e-03	2.00e+02	1.06e-02
21			5.00e+01	8.26e-03	2.00e+02	1.06e-02
22			5.00e+01	3.75e-03	2.00e+02	4.60e-03
23			5.00e+01	1.06e-03	2.00e+02	1.21e-03
24			5.00e+01	4.84e-04	2.00e+02	4.84e-04
1:1 pt	1.97e+00	1.97e+00	4.88e+00	4.88e+00	7.57e+00	7.57e+00

Table G.2: Numerical ARI values of the contour lines for a 9-hour rain duration.

	10 ARI		50 ARI		200 ARI	
	Rain	SWL	Rain	SWL	Rain	SWL
0	1.00e-02	1.00e+01	1.00e-02	5.00e+01	1.00e-02	2.00e+02
1	1.11e-02	1.00e+01	1.11e-02	5.00e+01	1.11e-02	2.00e+02
2	1.70e-02	1.00e+01	1.69e-02	5.00e+01	1.83e-02	2.00e+02
3	2.89e-02	8.46e+00	2.84e-02	4.71e+01	3.33e-02	1.25e+02
4	4.92e-02	6.68e+00	4.79e-02	3.37e+01	6.08e-02	7.69e+01
5	8.37e-02	5.16e+00	8.07e-02	2.56e+01	1.11e-01	7.52e+01
6	1.42e-01	5.11e+00	1.36e-01	2.56e+01	2.03e-01	7.52e+01
7	2.42e-01	5.11e+00	2.29e-01	2.56e+01	3.70e-01	7.52e+01
8	4.12e-01	5.11e+00	3.86e-01	2.56e+01	6.75e-01	7.52e+01
9	7.00e-01	5.11e+00	6.51e-01	2.56e+01	1.23e+00	6.33e+01
10	1.19e+00	4.29e+00	1.10e+00	2.28e+01	2.25e+00	5.00e+01
11	2.02e+00	3.06e+00	1.85e+00	1.85e+01	4.11e+00	3.47e+01
12	3.44e+00	1.78e+00	3.12e+00	1.31e+01	7.50e+00	2.12e+01
13	5.86e+00	8.36e-01	5.25e+00	8.26e+00	1.37e+01	1.11e+01
14	9.96e+00	3.06e-01	8.85e+00	4.56e+00	2.50e+01	4.57e+00
15	1.00e+01	8.04e-02	1.49e+01	2.13e+00	4.56e+01	1.50e+00
16	1.00e+01	1.32e-02	2.51e+01	8.07e-01	8.32e+01	5.54e-01
17	1.00e+01	1.32e-02	4.24e+01	2.72e-01	1.52e+02	2.73e-01
18	1.00e+01	5.08e-03	5.00e+01	9.37e-02	2.00e+02	1.25e-01
19	1.00e+01	1.23e-03	5.00e+01	3.25e-02	2.00e+02	4.52e-02
20	1.00e+01	4.84e-04	5.00e+01	8.26e-03	2.00e+02	1.06e-02
21			5.00e+01	8.26e-03	2.00e+02	1.06e-02
22			5.00e+01	3.75e-03	2.00e+02	4.60e-03
23			5.00e+01	1.06e-03	2.00e+02	1.21e-03
24			5.00e+01	4.84e-04	2.00e+02	4.84e-04
1:1 pt	2.57e+00	2.57e+00	6.74e+00	6.74e+00	1.27e+01	1.27e+01

Table G.3: Numerical ARI values of the contour lines for a 24-hour rain duration.

	10 ARI		50 ARI		200 ARI	
	Rain	SWL	Rain	SWL	Rain	SWL
0	1.00e-02	1.00e+01	1.00e-02	5.00e+01	1.00e-02	2.00e+02
1	1.11e-02	1.00e+01	1.11e-02	5.00e+01	1.11e-02	2.00e+02
2	1.70e-02	1.00e+01	1.69e-02	5.00e+01	1.83e-02	2.00e+02
3	2.89e-02	8.46e+00	2.84e-02	4.71e+01	3.33e-02	1.25e+02
4	4.92e-02	6.68e+00	4.79e-02	3.37e+01	6.08e-02	7.69e+01
5	8.37e-02	5.16e+00	8.07e-02	2.40e+01	1.11e-01	5.56e+01
6	1.42e-01	4.17e+00	1.36e-01	1.89e+01	2.03e-01	5.56e+01
7	2.42e-01	3.80e+00	2.29e-01	1.88e+01	3.70e-01	5.56e+01
8	4.12e-01	3.61e+00	3.86e-01	1.88e+01	6.75e-01	5.47e+01
9	7.00e-01	3.41e+00	6.51e-01	1.84e+01	1.23e+00	4.81e+01
10	1.19e+00	2.95e+00	1.10e+00	1.60e+01	2.25e+00	4.30e+01
11	2.02e+00	2.30e+00	1.85e+00	1.47e+01	4.11e+00	3.56e+01
12	3.44e+00	1.70e+00	3.12e+00	1.22e+01	7.50e+00	2.62e+01
13	5.86e+00	1.04e+00	5.25e+00	9.38e+00	1.37e+01	1.68e+01
14	9.96e+00	5.01e-01	8.85e+00	6.36e+00	2.50e+01	8.47e+00
15	1.00e+01	1.61e-01	1.49e+01	3.65e+00	4.56e+01	2.89e+00
16	1.00e+01	2.51e-02	2.51e+01	1.61e+00	8.32e+01	7.13e-01
17	1.00e+01	2.51e-02	4.24e+01	4.92e-01	1.52e+02	2.73e-01
18	1.00e+01	8.40e-03	5.00e+01	1.21e-01	2.00e+02	1.25e-01
19	1.00e+01	1.66e-03	5.00e+01	3.31e-02	2.00e+02	4.52e-02
20	1.00e+01	4.84e-04	5.00e+01	8.26e-03	2.00e+02	1.06e-02
21			5.00e+01	8.26e-03	2.00e+02	1.06e-02
22			5.00e+01	3.75e-03	2.00e+02	4.60e-03
23			5.00e+01	1.06e-03	2.00e+02	1.21e-03
24			5.00e+01	4.84e-04	2.00e+02	4.84e-04
1:1 pt	2.22e+00	2.22e+00	7.50e+00	7.50e+00	1.55e+01	1.55e+01

GHD

Level 3

138 Victoria Street

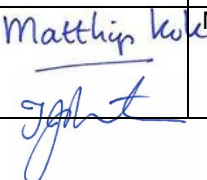

T: 64 3 378 0900 E: chcmal@ghd.com

© GHD 2021

This document is and shall remain the property of GHD. The document may only be used for the purpose for which it was commissioned and in accordance with the Terms of Engagement for the commission. Unauthorised use of this document in any form whatsoever is prohibited.

https://projectsportal.ghd.com/sites/pp02_01/jointprobabilitystud/12531791 - IConnect - Delivery/Main report/JointProbability_Final_20210204.docx

Document Status

Revision	Author	Reviewer		Approved for Issue		
		Name	Signature	Name	Signature	Date
DRAFT Rev A	G. van Rongen, G. Pleijter, T. Preston, R. Lenselink	Tim Preston, Mattijs Kok		T Preston		16/08
DRAFT Rev B	G. van Rongen, G. Pleijter, T. Preston, R. Lenselink	Tim Preston, Mattijs Kok		T Preston		1/09/2020
Final Rev 0	G. van Rongen, G. Pleijter, T. Preston	Tim Preston, Mattijs Kok		M Dasler		4/2/2021

www.ghd.com

