ENHANCED EXTREME SEA LEVEL INFORMATION FOR FLOOD RISK ASSESSMENTS ALONG THE GERMAN BALTIC COAST

Dissertation

zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen-Fakultät der
Christian-Albrechts-Universität zu Kiel

vorgelegt von

Leigh Richard MacPherson

Kiel, September 2024

Erster Gutachter:
Prof. Dr. Athanasios T. Vafeidis

Zweiter Gutachter:
Prof. Dr. Arne Arns

Datum der mündlichen Prüfung:
13.02.2025

gez. Prof. Dr. Robert Hassink

ABSTRACT

Accurate assessments of flood risk are vital for efficient adaptation to extreme sea level (ESL) induced coastal floods. A major source of uncertainty surrounding coastal flood risk are those associated with the estimation of ESLs and their probabilities. Given limited sea level observations, uncertainty is an inherent feature of ESL estimation — leading to misallocated resources, higher adaptation costs, reduced flood risk mitigation, or even counterproductive outcomes if ignored. Nevertheless, statistical and numerical techniques can enhance available sea level information, adding greater value to flood risk assessments and enabling more efficient adaptation planning. This dissertation explores a number of techniques to provide enhanced ESL information for coastal flood risk along the German Baltic Sea coast. Specifically, this thesis (1) models ESL hydrographs, simulating not only the peak water level of extreme events but also their temporal evolution, (2) incorporates historical information into the analysis of ESLs to improve estimates of ESLs, and (3) demonstrates how these enhanced information can be used to perform probabilistic flood modelling, providing a comprehensive assessment of flood exposure.

This cumulative dissertation consists of two parts. The first is a unifying essay which serves as a foundation for the entire body of work. In this essay, the main topic is introduced in Chapter 1 along with the research aim and specific objectives. Chapter 2 describes the theoretical framework underpinning the work presented. Research papers are then summarised in Chapter 3, outlining the motivation, data and methods, results and key findings of each of the main articles. Finally, a synthesis for the entire work is provided in Chapter 4, highlighting the main findings and avenues for future research. Part two of this dissertation contains the published articles as individual chapters.

Chapter 5 describes a stochastic ESL model developed for the German Baltic Sea coast. The model employs multivariate statistics and Monte Carlo simulations to generate artificial ESL hydrographs. Each hydrograph is a time-series of sea levels, allowing for flood risk assessments to account for the temporal variability of ESLs in addition to peak water levels, which have been shown to affect flood exposure. The model is computationally inexpensive, producing many thousands of events for locations where observational data exists. The main benefit of the stochastic model is the extremely large number of artificial events it can produce despite limited observational data, providing flood risk assessments with realistic information to explore flooding scenarios.

Chapter 6 investigates how historical information can be combined with systematic tide-gauge data to enhance estimates of ESLs. Bayesian extreme value analysis (EVA) is employed at seven sites along the German Baltic Sea coast where measurements of historical events have been recorded. The implementation of Bayesian techniques provides a natural framework for the handling of uncertainties, regarding those surrounding the measurement of historical events and the interpretation of results. I show that the inclusion of historical information greatly enhances estimates of ESL probabilities, including reduced uncertainties and better incorporation of outliers. Further, long-term dependence in the series of ESLs at three sites along the German Baltic coast is discovered, highlighting a recent period of low ESL activity for the region. This period of low activity suggests that estimates of ESL made using tide-gauge data alone are prone to underestimations.

Finally, Chapter 7 demonstrates the added value provided by the first two studies by performing probabilistic flood modelling at Lübeck on the German Baltic coast. Here, flooding under a large number of ESL scenarios generated using the stochastic model described in Chapter 5 is modelled, and flood extents are quantified. Changes in flood extent due to changes in peak water levels and ESL temporal variability are explored and results are presented as probabilistic flood maps. Flood extents differ by up to 60% when accounting for ESL temporal variability.

ZUSAMMENFASSUNG

Eine genaue Bewertung des Hochwasserrisikos ist entscheidend für eine effiziente Anpassung an durch extreme Wasserstände verursachte Küstenüberschwemmungen. Eine wesentliche Ursache der Unsicherheit hierbei entsteht durch die notwendige Bestimmung von extremen Wasserständen und deren Wahrscheinlichkeit. Aufgrund begrenzter Verfügbarkeit von Beobachtungsdaten des Meeresspiegels, ist die Bestimmung von extremen Wasserständen immer mit einer gewissen Unsicherheit verbunden. Werden diese Unsicherheiten nicht in die Bewertung des Hochwasserrisikos miteinbezogen, kann dies zu einer Fehlallokation von Ressourcen, höheren Anpassungskosten, einer verringerten Minderung des Hochwasserrisikos oder sogar kontraproduktiven Maßnahmen führen. Das Anwenden statistischer und numerischer Methoden kann jedoch dazu beitragen, die Nutzung der verfügbaren Informationen aus Beobachtungsdaten des Meeresspiegels zu verbessern. Dies verleiht der Bewertung des Hochwasserrisikos eine größere Aussagekraft und ermöglicht eine effizientere Anpassungsplanung. In dieser Dissertation werden verschiedene Methoden untersucht, mit dem Ziel, verbesserte Informationen über extreme Wasserstände zur Bewertung des Küstenhochwasserrisikos entlang der deutschen Ostseeküste bereitzustellen. Dies beinhaltet (1) die Modellierung von Hydrographen extremer Wasserstände, die nicht nur den maximalen Wasserstand extremer Ereignisse simulieren, sondern auch deren zeitlichen Verlauf, (2) die Einbeziehung historischer Daten in die Analyse extremer Wasserständen zur Verbesserung der Bestimmung von extremen Wasserständen, und (3) das Aufzeigen, wie diese verbesserten Informationen zur Durchführung probabilistischer Hochwassermodellierung verwendet werden können, um eine umfassende Bewertung der Hochwasserexposition zu ermöglichen.

Diese kumulative Dissertation gliedert sich in zwei Teile. Der erste Teil ist ein umfassendes Essay, in dem die Grundlagen der gesamten Arbeit erläutert werden. Das übergreifende Thema der Dissertation sowie das Forschungsziel und die spezifischen Zielen werden in Kapitel 1 vorgestellt. Kapitel 2 beschreibt die theoretischen Hintergründe, auf dem die vorgelegte Arbeit basiert. In Kapitel 3 folgt eine Zusammenfassung, in der die Motivation, Daten und Methoden, sowie Ergebnisse der verschiedenen Forschungsarbeiten dargelegt werden, die dieser Dissertation zu Grunde liegen. Abschließend bietet Kapitel 4 eine Synthese der gesamten Arbeit, die nocheinmal die wichtigsten Forschungsergebnisse hervorhebt und mögliche Fragestellungen für zukünftige Forschung aufzeigt. Der zweite Teil dieser Dissertation besteht aus zuvor veröffentlichten Artikeln als jeweils einzelnes Kapitel.

Kapitel 5 beschreibt ein stochastisches Modell extremer Wasserstände, das für die deutsche Ostseeküste entwickelt wurde. Das Modell verwendet Methoden von multivariater Statistik und Monte-Carlo-Simulationen, um künstliche Hydrographen extremer Wasserstände zu simulieren. Jeder Hydrograph ist eine Zeitreihe von Meeresspiegelmesswerten, die es ermöglichen, bei Hochwasserrisikobewertungen nicht nur die maximalen Wasserstände, sondern auch die zeitliche Variabilität der extremen Wasserstände zu berücksichtigen, was nachweislich die Hochwasserexposition beeinflusst. Das Modell ist rechnerisch kostengünstig und erzeugt mehrere tausend Ereignisse an Orten, an denen bereits Beobachtungsdaten vorliegen. Der Hauptvorteil des stochastischen Modells besteht in der extrem großen Anzahl künstlicher Ereignisse, die es trotz begrenzter Beobachtungsdaten erzeugen kann und die realistische Informationen für die Hochwasserrisikobewertungen bereitstellen, um Überschwemmungsszenarien zu analysieren.

Kapitel 6 untersucht, wie historische Informationen mit systematischen Pegeldaten kombiniert werden können, um die Bestimmung extremer Wasserstände zu verbessern. Die Bayessche Extremwertanalyse wird an sieben Standorten entlang der deutschen Ostseeküste eingesetzt, an denen historische Ereignisse aufgezeichnet wurden. Der Einsatz von bayesschen Methoden bietet einen Rahmen für den Umgang mit Unsicherheiten, sowohl im Hinblick auf die Messung historischer Ereignisse als auch auf die Interpretation der Ergebnisse. Ich zeige, dass die Einbeziehung historischer Informationen die Bestimmung der Wahrscheinlichkeiten extremer Wasserstände erheblich verbessert, einschließlich der Reduzierung von Unsicherheiten und einer besseren Berücksichtigung von Sonderfällen. Darüber hinaus wird an drei Standorten entlang der deutschen Ostseeküste eine langfristige Abhängigkeit in den Messreihen extremer Wasserstände festgestellt, die hervorhebt, dass die Region sich zuletzt in einer Periode geringer Aktivität extremer Wasserstände befand. Diese Periode geringer Aktivität deutet darauf hin, dass Schätzungen extremer Wasserstände, die ausschließlich auf Pegeldaten basieren, zu Unterschätzungen des Hochwasserrisikos neigen.

Schließlich wird in Kapitel 7 die zusätzliche Bedeutung der ersten beiden Studien durch die Durchführung probabilistischer Hochwassermodellierung in Lübeck an der deutschen Ostseeküste gezeigt. Hier wird die Überschwemmung unter einer Vielzahl verschiedener Szenarien von extremen Wasserständen modelliert, die mit dem in Kapitel 5 beschriebenen stochastischen Modell erzeugt wurden, und die Überflutungsgebiete werden quantifiziert. Änderungen im Überflutungsgebiet aufgrund von Änderungen der maximalen Wasserstände und der zeitlichen Variabilität extremer Wasserstände werden untersucht, und die Ergebnisse werden als probabilistische Hochwasserkarten dargestellt. Die Überflutungsgebiete unterscheiden sich um bis zu 60%, wenn die zeitliche Variabilität extremer Wasserstände berücksichtigt wird.

CONTENTS

ACRONYMS

ESL	Extreme Sea Level
EVA	Extreme Value Analysis
MSL	Mean Sea Level
SLR	Sea Level Rise
i.i.d.	independent and identically distributed
GEV	Generalized Extreme Value
GP	Generalized Pareto
AMAX	annual maxima
BM	block maxima
POT	Peaks-over-Threshold
JPM	joint probability method
RFA	regional frequency analysis
HW200	Hochwasser 200
NHN	Normalhöhennull
PDF	probability density function
MLE	maximum likelihood estimation
MCMC	Markov-Chain Monte Carlo
RMSE	root mean square error
STM	stability threshold method
MRL	mean residual life
BIC	Bayesian information criteria
GOF	goodness-of-fit
KS	Kolmogorov-Smirnov

Part I

UNIFYING ESSAY

This dissertation is the culmination of my published work, and is structured in two parts. This first part consists of a unifying essay, which serves as the foundation for the entire body of work. In this essay, I introduce the overarching research theme — *enhancing extreme sea level information for use in flood risk assessments*. I then outline the theoretical framework that underpins my studies and summarise each of the papers included in the second part of the dissertation. The key findings from these papers are synthesised, highlighting their collective impact and relevance to the field. The second part of the dissertation presents my published work, offering detailed insights and evidence of the research conducted. I have structured my dissertation in this manner to allow the unifying essay to be read as a complete and cohesive work. This approach ensures that the reader can fully grasp the research context and synthesis without needing to refer to the individual papers immediately.

Whenever a large
sample of chaotic elements are taken in hand
and marshalled in the order of their magnitude,
an unsuspected and most beautiful form of
regularity proves to have been latent all along.

— Sir Francis Galton (Galton, 1894)

1

INTRODUCTION

Living by the sea has historically offered numerous benefits — coastal areas provide access to fertile land for agriculture and the opportunity to harvest food directly from the sea, which has spurred the establishment of coastal settlements worldwide (Pugh, 2004). Moreover, proximity to water facilitates easier transportation of goods between neighbouring regions and countries, fostering enhanced commerce and driving economic growth to levels unattainable inland (Griggs, 2017). This increased economic activity generated by trade attracts supporting industries and spurs the development of infrastructure, thereby offering a wide array of commodities, services, and employment opportunities (Merkens et al., 2016; Kummu et al., 2016). Additionally, the milder climates and unique lifestyle associated with coastal living appeal to many individuals (Griggs, 2017).

However, coastal living comes with its own set of disadvantages, as seaside communities confront a range of unique challenges and risks. Among these, flooding triggered by extreme sea levels (ESLs) stands out as particularly devastating (IPCC, 2022; Tebaldi et al., 2021). Recent instances, such as Hurricane Maria in the northeastern Caribbean (Rivera, 2019), Typhoon Haiyan in Southeast Asia (Santos et al., 2016), and Hurricane Katrina in the US (Bleemer and Van der Klaauw, 2017), illustrate the severity of the resulting floods. Coupled with the dense population and extensive development of the affected regions, these events have led to widespread damage and significant loss of life, as well as displacing many individuals from their homes.

Although these events may be considered exceptional, numerous examples can be found throughout history (e.g., Jensen and Töppe, 1990). Furthermore, even non-exceptional events may be damaging

if a coastal community is exposed and vulnerable. For these reasons, the study of extreme events at the coast has garnered considerable interest. The amplifying effects of climate change on ESLs via sea level rise (SLR) (Nerem et al., 2018; Slangen et al., 2022; Vitousek et al., 2017) and storm intensification (IPCC, 2022; Walsh et al., 2016; Knutson et al., 2010; Wolski and Wiśniewski, 2021), in combination with socioeconomic development (Reimann et al., 2023; MacManus et al., 2021; Kummu et al., 2016), leads to a constantly evolving level of coastal flood risk, and there is a need for coastal communities to evolve with it (Hinkel et al., 2013).

However, there is no single solution to address the threat of coastal flooding (Hinkel et al., 2019; de Bruijn, 2004). With the potential for significant damages and loss of life and only finite resources to implement coastal protection measures, the final decision on adaptation carries with it a large responsibility. An accurate assessment of risk can support the decision making process, where adaptation options may be compared quantitatively based on a benefit-cost ratio (e.g. van der Pol et al., 2021; Hinkel et al., 2014).

A key component of any flood risk assessment is the estimation of probabilities (Stewart and Melchers, 1997). By assigning probabilities to ESLs, analysts may explore a range of possible outcomes and quantify risk under different scenarios. This aids in identifying sensitive areas that may require close attention, enabling efficient allocation of resources for risk management and damage mitigation. It is important to realise however, that ESL estimates are inherently uncertain (Coles, 2001; Dangendorf et al., 2016), a factor that can impede effective planning for adaptation or even yield counterproductive outcomes if ignored (Rohmer et al., 2021; Le Cozannet et al., 2015). As a consequence, uncertainties surrounding ESLs and their associated flood extents have been the focus of much research, with a particular emphasis on exploring analytical methods that can minimise these values.

In my research, I have developed and applied statistical and numerical techniques to improve current estimates of ESLs along the German Baltic coast, and provide valuable insights and useful information for flood risk assessments. My work has been published in a number of research papers, the culmination of which I describe herein. While I have focused primarily on ESLs, the ability to provide accurate sea level data for flood risk assessments has lead to a number of collaborations. These demonstrate how the outcomes of my work may be used to enhance studies on coastal flood risk.

Research aim and objectives

The overarching aim of my research is to *enhance available sea level*

information along the German Baltic coast for use in flood risk assessments. My primary objectives are:

1. Model ESL hydrographs along the German Baltic coast.
2. Investigate how historical information can be incorporated into EVA to improve estimates of ESLs.
3. Demonstrate how these improved information can be used to conduct probabilistic flood modelling.

In this unifying essay, I summarise my doctoral research and provide a context for my motivations. The essay is organised as follows: I begin with a brief background of my research, detailing the context from which my motivations stem. Next, I outline the theoretical framework and key concepts that underpin my research, discussing previous work and highlighting gaps in knowledge. Following this, I summarise my work by describing each of the research papers I have authored. Finally, I highlight the main findings of my combined research and suggest directions for future research.

1.1 BACKGROUND AND CONTEXT

On the 23rd of October 2007, the European Parliament adopted a directive to manage and reduce the risks that floods pose to human health, the environment, cultural heritage and economic activities in the EU (EU, 2007). The directive requires member states to perform flood risk assessments to identify river basins and coastal areas at risk of flooding. While the main aim of the directive is to enhance the EU's resilience to floods, it also promotes efficient use of resources when tackling the issue of flooding.

Similar regulations and guidelines can be found in other states, including but not limited to the National Flood Insurance Program in the US (FEMA, 1968) and the Flood and Water Management Act 2010 in the UK (UK, 2010). While details vary, each approach includes elements of risk assessment, mapping, planning and public engagement to mitigate the impact of floods. These illustrate the global recognition of the importance of implementing comprehensive flood risk management strategies.

At the very foundation of these programs, a solid understanding of flood risk is required. Although flooding may occur for a number of

reasons (e.g. pluvial, fluvial, dam and levee failure, etc.), the main topic of this thesis involves coastal flooding induced by ESLs. In this context, flood risk management requires accurate information on the likelihood of ESLs, in combination with data on exposed populations.

In this section, I will define the main concepts of my research and provide a historical context for understanding their development and relevance within the scope of my study. First, I will discuss extreme sea levels (ESLs) and how these events may be modelled statistically. Second, I will provide a brief outline of coastal flood risk assessments, highlighting the importance of accurate information regarding ESLs. Last, I will describe the study region - the German Baltic coast.

1.1.1 *Extreme sea levels*

Sea level is dynamic. Changes in sea level may extend just a few meters or across the entire globe, and the processes which drive these changes act over time scales ranging from seconds to millenia (Pugh, 2004). What we observe as sea level is a complex and ever-changing phenomenon, resulting from the combination of a multitude of forces acting on Earth's oceans (Pugh and Woodworth, 2014). In the context of ESLs, it is useful to regard the observed sea level as the combination of mean sea level (MSL), tides and storm surges. Taking only these three components, the resulting sea level is given the misnomer *still water level*, due to the absence of wind-generated waves (Pugh, 2004).

Mean sea level (MSL)

Before describing ESLs, it is helpful to understand the three main components listed above. First, MSL refers to the average height of the ocean's surface over a specific period of time (Pugh, 2004). The exact period used for MSL calculation may be chosen to eliminate unwanted short-term fluctuations. For example, a period of one year may be chosen to remove seasonal effects. Alternatively, a longer period of approximately 18.6 years accounts for the lunar nodal cycle which influences the tides (Pugh, 1987). In any case, MSL represents a base level over which tides and storm surges act.

Tides

Tides are the regular and predictable variations in sea level driven by gravitational forces of the moon and sun (Pugh and Woodworth, 2014). These forces grow and subside with the movement of the two bodies relative to the earth. For example, tides follow a *spring-neap* cycle influenced by the alignment of the Earth, moon, and

sun. Higher high tides (spring tides) occur during full and new moons, and lower high tides (neap tides) during the first and third quarters of the moon. As tides propagate across the Earth's oceans as waves, they are influenced by oceanographic factors such as currents, coastal morphology and water depth. Additionally, local features like estuaries, bays, and channels can significantly modulate tidal patterns, leading to complex tidal behaviours unique to each region (Pugh, 2004).

Despite their complexity, predicting tides at single locations is mathematically straightforward owing to the predictable movements of the astronomical bodies influencing them (Pugh and Woodworth, 2014). Here, harmonic analysis decomposes tides into simpler periodic components based on known frequencies (see Codiga, 2011). By applying harmonic analysis to tidal observations, dominant tidal constituents can be identified and their characteristics quantified, including amplitudes, phases, and periods. Tidal reconstructions may then be performed to determine the effect of each constituent on the water level, and final tidal predictions may be calculated by summing these effects.

Storm surges

Acting over MSL and tides are storm surges. In theory, storm surges are the sea level response to atmospheric forcing (Pugh and Woodworth, 2014). In practice, it is difficult to isolate only this signal. Instead, storm surges are typically found by removing MSL and tides from the total observed sea level, leaving only the *residual water level* (Pugh, 2004). Although changes in the residual water level may occur for many reasons, at the vast majority of locations, the dominant signal is due to atmospheric forcing (Pugh, 2004). Along coastlines, the elevation of residual water levels is often attributed to the combined effect of two key atmospheric factors: low atmospheric pressure and strong winds, both of which are present during storms. Firstly, low atmospheric pressure, commonly associated with approaching storms, results in a "bulge" of water beneath the center of the storm due to the reduced pressure pushing down on the ocean surface, creating a localized increase in sea level. Secondly, strong winds can push large volumes of water towards the coast, where it accumulates and augments the height of the sea level. As both of these factors occur during storms, their combined effect on the residual water level is referred to as *storm surge* (Pugh and Woodworth, 2014).

Extreme sea levels (ESLs) defined

Pugh (2004) states that ESLs occur when large storm surges coincide with high spring tides. While this definition provides a straightforward understanding of ESLs, it does not explain all cases. For example, tides along the German Baltic coast are small (0.1-0.2 m) and thus contribute little to ESLs (Feistel et al., 2008; Sterr, 2008). The

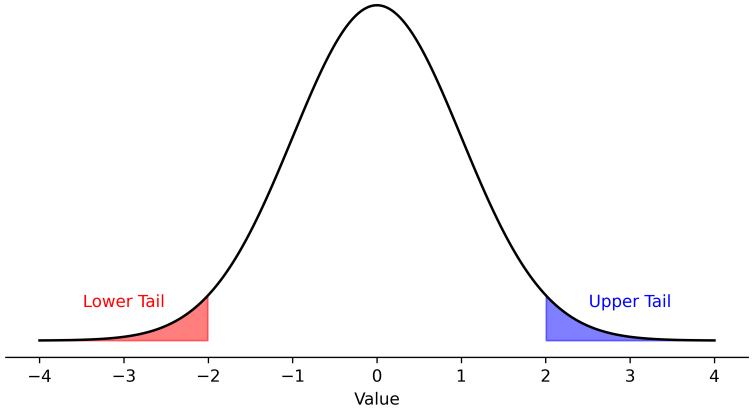

Figure 1.1: Schematic of a sea level distribution with the lower and upper tails highlighted. The vast majority of sea levels occur near the middle of the "bell curve", and ESLs occur within the tails.

occurrence of ESLs in this region is therefore influenced almost solely by storm surges. Furthermore, the term "large" is subjective, and provides little information on how to identify ESLs. From a statistical point of view, ESLs are those events which fall within the tails of the sea level distribution (Figure 1.1). Whether these events occur during a high spring tide or not is irrelevant to their classification as extreme. As extreme *low* waters are rarely of interest, especially in the context of coastal flooding, the term ESL typically applies to events in the right distribution tail only.

Return periods

Similar to other natural hazards, it is useful to classify ESLs in terms of their magnitude. Whereas earthquakes are scaled based on their seismic moments, and tropical cyclones are categorised according to wind speeds, ESLs are quantified by their peak water levels. However, ESLs exhibit considerable spatial variability (Pugh, 2004), making any measurement of peak water level only meaningful locally. Therefore, the conventional approach is to classify ESLs according to their *return period*, which is the average amount of time expected between exceedances of some specific peak water level (Arns et al., 2013; Haigh et al., 2010). For example, if sea levels exceed 1.5 m on average once every 10 years, an ESL with a peak water level of 1.5 m would be assigned a return period of 10 years. As return periods define an expected number of occurrences over some duration, they are a measure of probability. Calculating probabilities of ESLs is a major part of flood risk assessments, and falls under the domain of extreme value analysis (EVA). An in-depth description of EVA is given in Section 2.1.

A note on nomenclature

There is a wide array of terms used in scientific literature to refer to

ESLs, and I have taken care to use the most appropriate form. Most typical is the term *storm surge*, which has not been used in this thesis to avoid confusion with the sea level component described above. Similarly, *storm tide* has also been excluded. There is some debate regarding the use of *extreme water level* or *extreme coastal water level* instead of ESL, however I prefer the latter partly due to its common use in literature and partly out of habit. Finally, the term *return water level* is used often in literature and refers to ESLs with specific return periods. I find this term useful as it conveys more meaning than simply ESL, and has therefore been used in this thesis.

1.1.2 *Coastal flood risk*

Definition of risk

In its simplest form, risk can be defined as a function of probability and consequence (EU, 2007). With the innovation of EVA (see Section 2.1), extreme events of specific magnitudes could be assigned probabilities and systematic flood risk assessments became possible. Early studies focused on the management of floods, rather than flood *risk* (Bruijn et al., 2007), which limited coastal protection decisions to *hard* solutions (or structural solutions: dikes, surge barriers, drainage systems, etc.) aimed at preventing flooding. Later, the concept of risk became more formally incorporated into flood management and a number of *soft* solutions (non-structural or natural solutions: dune and wetland restoration, floodplain zoning, operational forecasting, etc.) were developed to reduce the *impact* of floods (Bruijn et al., 2007).

The shift from managing floods to managing flood risk expanded focus to include not only flood hazards and the exposure of populations and assets but also their vulnerability to floods (White et al., 2001; Kron, 2005). This broader understanding enabled comprehensive flood risk assessments that consider both the physical characteristics of flood hazards and socioeconomic factors, incorporating flood resilience alongside flood resistance (de Bruijn, 2004).

The concept of resilience has since become prominent in discussions on flood risk (see White et al., 2018; Ashley et al., 2011; Connelly et al., 2015; Garvin et al., 2013; McClymont et al., 2020). By enhancing the capacity of communities to adapt and recover from flood events, a resilience strategy is better equipped to cope with uncertainty compared to a resistance strategy (de Bruijn, 2004). Considering the high level of uncertainty in any one factor of future flood risk, be it physical (sea level rise, storm intensity and frequency, subsidence, sediment supply, erosion, etc.) or socioeconomic (population growth,

urbanisation, coastal migration, etc.) (see Rohmer et al., 2021; Le Cozannet et al., 2015; Hinkel et al., 2021), the adoption of coastal resilience in the field of flood risk is not surprising. Given future flood risk depends on a combination of these factors, comprehensive flood risk assessments must deal with compounded deep uncertainty.

Uncertainties in flood risk

A major component of flood risk uncertainty in the short term (before 2050) arises due to uncertainties in the estimates of ESLs (Rohmer et al., 2021; Le Cozannet et al., 2015). By improving understanding of local ESL processes, uncertainty surrounding near-term flood risk is reduced and more efficient adaptation planning becomes possible. Beyond this point, uncertainties in future human activities and global SLR become the dominant factors. To deal with these uncertainties, Le Cozannet et al. (2015) suggest a need for continuous improvement in the probabilistic modelling of coastal floods, which is better suited to handling and communicating uncertainties than deterministic approaches.

Another factor of uncertainty in the assessment of coastal flood risk is the temporal evolution of the forcing ESLs (Santamaria-Aguilar et al., 2017; Quinn et al., 2014). Each ESL event begins near MSL, rises to a peak water level and then subsides again. Whereas two events may both reach the same peak water level, their evolution over time can be vastly different (MacPherson et al., 2019; Wahl et al., 2011). The time-series of water levels which describes each event is hereafter referred to as a *hydrograph*. Since no two hydrographs are identical, it is possible for events of equal magnitude (same peak water level) to produce different flood characteristics (Santamaria-Aguilar et al., 2017; Quinn et al., 2014; Höffken et al., 2020). However, very little research into this phenomenon has been conducted. Instead, the typical approach is to define some standardised *design* hydrograph based on an observed event or set over some standard duration (e.g., Dawson et al., 2005; Wadey et al., 2015; MacPherson et al., 2011; Kiesel et al., 2022).

1.1.3 *The German Baltic coast*

The German Baltic coast is a highly developed and populated region located along the southwest Baltic Sea. Despite a long history of devastating coastal floods (Jensen and Müller-Navarra, 2008; Jensen et al., 2022), the current level of coastal protection in the region is far inferior to that along the German North Sea coast (Sterr, 2008). This is somewhat justified, as ESL activity along the North Sea coast is typically much higher than that at the Baltic Sea (Sterr, 2008).

However, ESL induced flooding in the German Baltic region is not uncommon. Additionally, several large cities in the region with significant economic, social and cultural value are situated directly at the sea. These factors combined; high population and development, relatively low levels of flood protection, and a history of large ESLs; suggest high levels of coastal flood risk (Sterr, 2008).

Due to its location in the Baltic Sea, the German Baltic coast experiences ESLs during strong northeasterly winds (Feistel et al., 2008). However, as the Baltic Sea is a semi-enclosed basin, water levels are affected by large-scale atmospheric forcing over the entire sea (Andersson, 2002; Jevrejeva et al., 2005). Furthermore, seiches act over the Baltic sea and can contribute several decimeters to local water levels (Jensen and Müller-Navarra, 2008). These are influenced not only by the large-scale atmospheric forcing, but also by certain temporal sequences of regional wind patterns. Tides in the region are small (0.1-0.2 m) (Feistel et al., 2008), being effectively filtered from the water level signal by the narrow entrance of the Kattegat and the danish islands. Consequently, the effect of tides on the occurrence of ESLs is negligible.

Coastal protection in the region is designed according to the water level of an ESL with a return period of 200 years (see MELUND, 2022; MLUV, 2009), and is referred to in technical reports as Hochwasser 200 (HW200). Since 2012, design heights were set equal to HW200 with an additional 0.5 m climate surcharge (Klimazuschlag) to account for uncertainties arising from future changes in sea level and storm surges. The climate surcharge has since been increased to 1.0 m (MELUND, 2022).

The exceptional ESL of 1872

Until recently, there has been a large disparity between estimated return water levels and past observed events along the German Baltic coast. In November of 1872, an exceptional storm over the Baltic Sea caused widespread flooding along much of the German and Danish coastlines, and in some areas of Sweden (Hallin et al., 2021). At Travemünde, water levels reached a height of approximately 3.4 m above MSL (Jensen et al., 2022). In 2012, the federal state of Schleswig-Holstein published a HW200 value for Travemünde of 2.24 m above Normalhöhennull (NHN), the standard vertical datum used in Germany (MELUND, 2022). The difference between current MSL and NHN at Travemünde is small (less than 5 cm). With the climate surcharge, design heights at Travemünde between 2012 to 2022 were 2.74 m, approximately 66 cm less than the event observed in 1872.

Attempts to classify the 1872 ESL have resulted in very large return periods (Gräwe and Burchard (2012) estimate a return period exceeding 100,000 years) and its validity in traditional EVA has been questioned (Hofstede and Hamann, 2022; Mudersbach and Jensen, 2009). The argument here is that the 1872 event is a statistical outlier, and thus does not belong to the distribution of ESLs which have been observed since and its inclusion would therefore bias any analysis of ESLs. However, there exists evidence of other events, comparable in magnitude to that of 1872 at Travemünde and other locations along the German Baltic coast (Jensen and Müller-Navarra, 2008; Jensen et al., 2022), which have occurred as far back as 1044. Consequently, the 1872 event may not be such a singular occurrence, and it could provide valuable information for the estimation of ESLs. However, this is yet to be shown.

2

THEORETICAL FRAMEWORK

In this section, I describe the key concepts upon which my research is based. Paramount to my research is the *statistical modelling of extreme sea levels*. Here I describe the main methods of statistical analysis and data processing required to model ESLs. I outline the quantification of uncertainties, and discuss some indirect methods of EVA developed to reduce these uncertainties. Finally, I provide an overview of *probabilistic flood modelling*.

2.1 STATISTICAL MODELLING OF EXTREME SEA LEVELS

Annual maxima (AMAX) sampling and the GEV distribution

Modeling ESLs falls within the realm of extreme value analysis (EVA), a specialized branch of statistics dedicated to understanding and estimating rare and extreme events. A foundation for EVA can be found in the work by Fisher and Tippett (1928). Here, it was proven that the distribution of maximum (or minimum) values in a large sample of independent and identically distributed (i.i.d.) random variables, converges to one of three possible types. The second of which was described at the same time by Fréchet (1927), with the others characterised later by Gumbel (1958) and Weibull (1951), lending their names to the respective distributions. Jenkinson (1955) would unify these three types into a single Generalized Extreme Value (GEV) distribution. The GEV distribution has since found wide application in the field of hydrology to model the occurrence of extremes, such as precipitation, flooding, river flows and ESLs among others.

The GEV distribution was developed for, and thus best suited to modelling extremes sampled using a block maxima (BM) approach (Fisher and Tippett, 1928; Jenkinson, 1955). Here, extreme events are sampled from data by dividing the data into blocks of equal size and extracting only the maximum (or minimum) values from each block (Coles, 2001). This remains a popular approach in hydrology, where the most common block size is one year, producing a sample of annual maxima (AMAX), to which the method owes its name.

One of the main limitations of the AMAX method however, is that it does not utilise all available data efficiently (Coles, 2001). As only the largest event within any one year will be sampled, other extreme events which may occur within the same year are disregarded. Furthermore, moderate events may be included within the analysis if no extreme events occur within any one block, leading to inaccuracies. These issues in combination with the limited sample size produced when only one event per year is considered, can lead to large uncertainties, especially at the right tail of the distribution (e.g., Hamdi et al., 2014).

Peaks-over-Threshold (POT) sampling and the GP distribution

To address these limitations, Balkema and Haan (1974) proposed a new Peaks-over-Threshold (POT) approach, which focuses on modelling exceedances over a high threshold rather than BM. They demonstrated that these exceedances converge to the Generalized Pareto (GP) distribution. The theoretical basis of which was later developed by Pickands (1975). As the POT method does not limit the sampling of data to blocks (sometimes arbitrarily defined), it does not suffer the same inefficiency limitations of AMAX sampling (a comparison of the two sampling methods is shown in Figure 2.1). As a result, the POT approach typically provides estimates of extremes with reduced uncertainties at the right tail (Hamdi et al., 2014), at the expense of a more complicated methodology.

Both the GEV and GP distributions are characterised by three parameters: location (μ), scale (σ) and shape (ξ) (Jenkinson, 1955; Pickands, 1975). To simplify the probability density function (PDF) of both distributions, consider a standardised variable $z = (x - \mu)/\sigma$, where x is the i.i.d. random variable to be modelled. The PDF of the GEV distribution is:

$$f(z;\xi) = \begin{cases} \exp(-z)\exp(-\exp(-z)) & \text{for } \xi = 0 \\ (1+\xi z)^{-(1+1/\xi)}\exp\left(-(1+\xi z)^{-1/\xi}\right) & \text{for } \xi \neq 0 \text{ and } \xi z > -1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

and the PDF of the GP distribution is:

$$f(z;\xi) = \begin{cases} (1+\xi z)^{-\frac{\xi+1}{\xi}} & \text{for } \xi \neq 0 \\ \exp(-z) & \text{for } \xi = 0. \end{cases} \tag{2.2}$$

The GEV distribution is valid for $x > \mu - 1/\xi$ where $\xi > 0$, and $x < \mu - 1/\xi$ where $\xi < 0$. Where $\xi = 0$, support for the distribution is found for all real values of x. The GP distribution is valid for all $x > \mu$, unless $\xi < 0$, in which case $\mu \leq x \leq \mu - \sigma/\xi$.

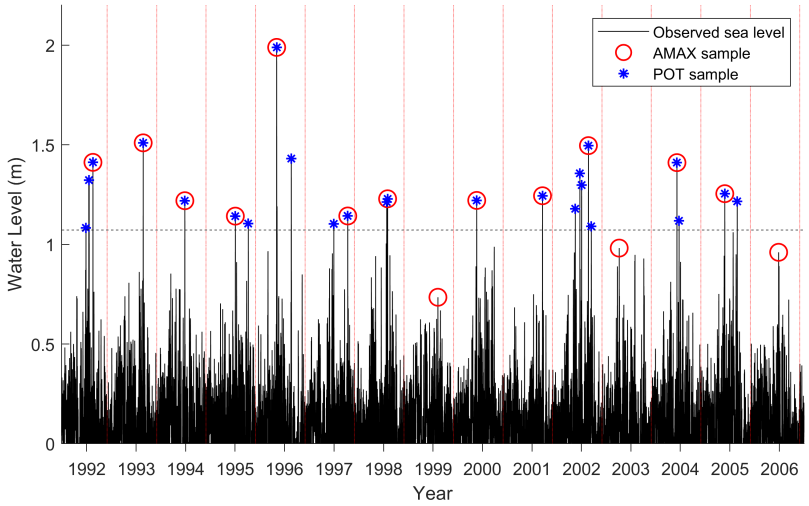

Figure 2.1: A comparison of AMAX and POT sampling on a time series of water levels measured at the Kiel-Holtenau tide gauge in Germany. The AMAX approach samples single events from blocks 1-year in length, starting July 1st and ending June 30th the following year (vertical dotted lines). This approach includes three moderate events and misses several larger events in years where multiple ESLs are present. The horizontal dashed black line represents the threshold used in the POT method, which samples the data more efficiently than the AMAX approach.

2.1.1 *Best practices*

A direct approach to EVA is one that applies the concepts described in Section 2.1 directly to observational data. This is in contrast to indirect methods where data is pooled from multiple sources, or manipulated before analysis, such as separating it into its component parts. Arns et al. (2013) provide an overview of direct methods for the statistical modelling of ESLs, specifically focusing on the AMAX and POT approaches. Here, three main steps are outlined: 1) detrending, 2) sampling, and 3) parameter estimation, and best practices for their implementation are explored. I will briefly describe each of these steps.

Detrending

A fundamental assumption of traditional EVA is that the underlying random variable is independent and identically distributed (i.i.d.) (Coles, 2001). The second term, *identically distributed*, indicates that each realisation of the random variable to be modelled comes from the same probability distribution. A sequence of identically distributed random variables is considered *stationary* if their statistical properties remain constant over time. However, sea level processes are inherently non-stationary because they are influenced by a va-

riety of dynamic factors (see Pugh, 2004). Dixon and Tawn (1999) identify four sources of non-stationarity which affect ESLs: 1) changes in MSL, 2) tides, 3) surge seasonality, and 4) tide-surge interaction. They assess that non-stationarity due to tides leads to significant underestimations when using the AMAX approach, recommending the use of indirect methods such as the regional joint probability method (JPM) instead. However, in regions with small tidal ranges, both methods yield similar results. Allamano et al. (2011) found similar underestimations with the POT approach when examining surge seasonality.

Dixon and Tawn (1999) account for non-stationarity due to changes in MSL by simply removing a linear trend. This technique has often been applied when employing direct EVA approaches (AghaKouchak et al., 2012; Haigh et al., 2014b). However, changes in MSL exhibit periods of acceleration and deceleration which are not captured by simple linear regression. Instead, Arns et al. (2013) recommend the use of a 1-year moving average of sea levels to account for both long-term and seasonal trends. Indeed, the use of low-pass filters such as running means have been shown to transform non-stationary data sets into stationary series (Mentaschi et al., 2016; Vousdoukas et al., 2018b).

Sampling

Next, it is necessary to sample the extreme values from the sea level record. The main approaches (AMAX and POT) and their limitations are outlined in Section 2.1. To summarise, the AMAX method samples the largest event for each year-long block of data, whereas the POT approach selects all events that exceed some threshold. The POT approach is generally preferred as it utilises the available data more efficiently (Arns et al., 2013; Coles, 2001), at the expense of a more complicated methodology. However, the POT approach is more prone to bias due to the subjective nature of threshold selection. Although methods for threshold selection have been proposed (see Coles, 2001; Arns et al., 2013; Watson, 2023; Bader, 2016), their applicability and performance from site to site differs, making them difficult to implement systematically. Nevertheless, thresholds are often selected based on some percentile of peak values (Arns et al., 2013; Takbash and Young, 2019).

Although the POT approach is generally preferred over the AMAX approach, AMAX sampling is still commonly practiced due to its much simpler implementation. Indeed, for long series of data, limitations of the AMAX method are reduced (Hamdi et al., 2014). At this point, implementation of the POT approach is less desirable due to its complicated and somewhat subjective methodology. Here, there is a bias-variance trade-off when selecting an appropriate threshold, and

methods to assist in the process leave room for interpretation (Coles, 2001). Setting a threshold low will introduce bias as more moderate events are included, whereas setting a threshold high increases variance and uncertainty as the sample size is reduced.

Declustering

Care must be taken when sampling extremes to ensure that each event is *independent*, fulfilling the first term of the i.i.d. requirement. This process of omitting events that are clustered together is referred to as *declustering* (Arns et al., 2013). The most common method employed when dealing with ESLs is to define some standard storm duration (Tawn, 1988), where events that occur within this time-frame are considered dependent and only the largest event is sampled (e.g., Zachary et al., 1998; Ward et al., 2018; Mathiesen et al., 1994; Haigh et al., 2014b). A similar approach, the *runs* method (see Acero et al., 2011; Leadbetter et al., 1989), distinguishes independent events by setting a water-level threshold, with the requirement that a specified number of observations fall below this threshold between consecutive events. The implementation of both these methods however rely on a subjective choice of either a standard storm duration or threshold, and these decisions can greatly affect the analysis (Arns et al., 2013). Systematic approaches such as the use of the *extremal index* (Ferro and Segers, 2003) or through correlation analysis (Martín et al., 2024) offer a more objective methodology of declustering, at the expense of simplicity.

Parameter estimation

The final step required to model ESLs directly involves parameter estimation of the underlying probability distribution, where the choice of distribution is dependent on the sampling method (Coles, 2001). The most commonly used approach is maximum likelihood estimation (MLE). Here, the goal is to find the parameters of the distribution (parameter vector θ; which in the case of the GEV or GP distributions is (μ, σ, ξ)) that maximises the likelihood function. For a set of s systematic ESL observations (x), the likelihood function $(\mathcal{L})$ is equal to the product of the relevant PDF (f_θ) (Coles, 2001), evaluated at each sample:

$$\mathcal{L}(x \mid \theta) = \prod_{i=1}^{s} f_\theta(x_i). \tag{2.3}$$

However, as this typically involves multiplying numerous small terms, the product can be cumbersome and prone to numerical issues. Thus, the log-likelihood is preferred, allowing for a summation of terms instead:

$$\log \mathcal{L}(x \mid \theta) = \sum_{i=1}^{s} \log f_\theta(x_i). \tag{2.4}$$

Recently, the application of Bayesian techniques for parameter estimation has become increasingly common. In the past, Bayesian models have often proved intractable, but the development of simulation-based techniques such as Markov-Chain Monte Carlo (MCMC) have made their use possible (Coles and Tawn, 1996). Whereas frequentist approaches, such as MLE, rely on long-run frequency properties of estimators, treating model parameters as fixed but unknown quantities and deriving estimates based on repeated sampling from the data, Bayesian statistics aim to determine the posterior distribution of the parameters, given available data and prior information. According to Coles and Tawn (2005), this provides two major benefits over frequentist EVA approaches of parameter estimation: 1) Bayesian techniques allow for a modular incorporation of different data types, providing a single method to assess a range of complex models; and 2) Bayesian inference provides a natural framework for managing uncertainties, leading to probabilistic estimates of return water levels.

2.1.2 *Uncertainties*

The extreme value models described in Sections 2.1 and 2.1.1 provide best estimates given limited information. In other words, the true distribution of extremes cannot be known for certain without infinite data. Thus, EVA is inherently uncertain (e.g., Dangendorf et al., 2016). Dealing with these uncertainties is a vital component of EVA, especially in the context of coastal flood risk, where accurate predictions are essential for effective mitigation and planning.

Confidence intervals

The most common approach to quantifying uncertainties in the estimates of return water levels is to define *confidence intervals*. These describe a range of values within which the true estimate is assumed to lie, according to a specified level of confidence (Coles, 2001). These are calculated using standard errors (Efron and Hinkley, 1978), which measure the variability of a model's parameters. For models such as the GEV and GP distributions, standard errors may be derived from the Fisher Information Matrix, which quantifies how much information the observable data provides about the unknown parameters (Davison and Smith, 1990).

Bootstrap sampling, introduced by Efron (1979) as a non-parametric approach to statistical inference, offers an alternative method for deriving confidence intervals. This technique makes minimal assumptions about the underlying data distribution, making it applicable to a broad spectrum of data types and statistical models. The method is particularly useful in the context of uncertainties as it provides a

way to quantify the variability and confidence in statistical estimates when analytical methods are impractical or impossible. The basic idea behind bootstrap sampling is to generate multiple *bootstrap samples* from the original dataset by sampling with replacement. Each bootstrap sample is of the same size as the original dataset but may include some observations multiple times while omitting others. By repeatedly resampling and calculating the statistic of interest for each bootstrap sample, uncertainties are defined as an empirical distribution of the statistic.

Credibility intervals

Lastly, Bayesian statistics offer a natural approach to managing uncertainties (Coles and Tawn, 2005). Here, uncertainties are quantified and interpreted differently than in frequentist statistics. Whereas MLE provides point estimates of parameter values, and uncertainties are defined based on the variability of these estimates under repeated sampling (as reflected in confidence intervals), Bayesian methods produce a posterior distribution for each parameter. These posterior distributions integrate prior information and the likelihood of the observed data, allowing for a direct probabilistic interpretation of uncertainty. Bayesian credibility intervals, derived from the posterior distribution, indicate the range within which the true parameter value lies with a specified probability, given the observed data and prior beliefs.

Framing uncertainties in terms of probabilities is beneficial in the context of flood risk, because it provides a clear and understandable measure of uncertainty, enhancing decision-making processes, improving communication with non-statisticians, and aligning with the inherent probabilistic nature of risk assessments (Coles and Tawn, 2005). A discussion on the use of Bayesian statistics in lieu of frequentist approaches is given in Section 4.2.2.

2.1.3 *Indirect methods*

While direct methods of EVA are relatively straightforward, they often produce estimates with significant uncertainties due to limited data, especially for low-frequency events (Haigh et al., 2010). To address these issues, several methods have been developed, which leverage broader data sources and statistical or numerical techniques to enhance the reliability of extreme value predictions. As one or more intermediate steps are required, these methods are referred to as *indirect methods*.

Hydrodynamic numerical modelling

A widely used approach to improve estimates of return water lev-

els is to extend available data both temporally and spatially using hydrodynamic numerical models (see Gräwe and Burchard, 2012; Muis et al., 2016; Vousdoukas et al., 2016; Haigh et al., 2014b; Muis et al., 2020). These models provide useful data at ungauged sites or locations with only short sea level records, thereby enhancing the effectiveness of direct EVA methods by extending the underlying dataset. However, their implementation is highly complex and simulations require significant computational resources. Moreover, their accuracy depends heavily on the availability and quality of atmospheric and bathymetric data.

Historical information

Another approach is to employ historical information in combination with systematic data to enhance studies on ESLs (Coles and Tawn, 2005; Payrastre et al., 2011; Bulteau et al., 2015; Hamdi et al., 2015). Although a number of methods exist (see Prosdocimi, 2018), the use of Bayesian techniques has become popular due to its natural framework for handling uncertainties, and flexibility when including a wide range of data sources (Coles and Tawn, 2005). Here, the likelihood function is updated to include terms for historical data, and estimates including credibility intervals are derived using a MCMC algorithm.

A number of other indirect methods exist but are not covered in this thesis. The two most common approaches are: the JPM, where the contribution of tides and storm surges are assessed separately (Pugh and Vassie, 1980; Tawn et al., 1989a; Tawn, 1992); and regional frequency analysis (RFA), which combines data sets within homogenous regions to yield larger samples of ESLs (Wiltshire, 1985; Weiss et al., 2014b). Application of the JPM along the German Baltic coast was not considered due to the small tides in the region. Alternatively, RFA offers several potential benefits, but was not prioritised. This method is briefly discussed as an avenue for future research in Section 4.2.4.

2.2 PROBABILISTIC FLOOD MODELLING

Flood modelling is a critical component of disaster management, enabling stakeholders to predict flood extents, assess risks, and implement effective mitigation strategies (Hall and Solomatine, 2008; Merz et al., 2010). The increasing frequency and intensity of flood events, driven by climate change and urbanisation, underscore the importance of flood modelling. However, the accuracy of flood models is often hindered by significant uncertainties (Merz and Thieken, 2009; Beven and Hall, 2014; Wagenaar et al., 2016; Winter et al., 2018). The growing recognition of uncertainties in flood risk mod-

elling has led to an increased focus on uncertainty analysis (see Hall and Solomatine, 2008; Merz and Thieken, 2009; Hinkel et al., 2015). Beven and Hall (2014) highlight that incorporating uncertainty analysis into flood models enhances their robustness and reliability. Understanding and quantifying uncertainties allow for better risk communication and more informed decision-making.

Probabilistic flood modelling involves the use of statistical methods to account for uncertainties and provide a range of possible flood scenarios (Hinkel et al., 2015). This approach contrasts with deterministic models, which typically produce a single flood extent prediction (e.g., Muis et al., 2016; Wadey et al., 2015). Merwade et al. (2008) and Bales and Wagner (2009) advocate for the use of uncertainty zones in inundation mapping. These zones represent areas with varying probabilities of inundation, providing a more comprehensive view of potential flood extents. Communicating these uncertainty zones is crucial for effective risk management and public awareness.

To quantify uncertainty in flood extent, it is recommended to consider a full range of return water levels, from high-probability low-magnitude events to low-probability high-magnitude events (Hinkel et al., 2015; Merz and Thieken, 2009; Ward et al., 2011). This comprehensive approach captures the full spectrum of possible flood scenarios, enhancing the predictive capability of flood models. Despite these advantages, flood risk assessments often rely on a limited number of ESL events (see Muis et al., 2016; Purvis et al., 2008; Wadey et al., 2012). This practice can lead to underestimation of uncertainties and potentially flawed risk assessments. Hauer et al. (2021) provide an example of how accounting for a full range of ESL return periods can improve the robustness of flood risk assessments.

3

PAPER SUMMARIES

As part of my doctoral research, I have published a number of papers, each addressing a specific objective stated in Section 1. These papers collectively contribute to the overarching aim of *enhancing available ESL information for coastal flood risks assessments along the German Baltic coast*. In this chapter, I summarise these papers, outlining their motivations, the specific data and methods used, and present their key outcomes and findings. Each paper and their publication details are provided in their entirety in Chapters 5, 6 and 7 of this dissertation.

3.1 A STOCHASTIC EXTREME SEA LEVEL MODEL FOR THE GERMAN BALTIC SEA COAST

Motivation

The vast majority of studies which assess coastal flood risk consider only the peak water level of the forcing ESL hazard, or single realisations of the relevant events (see Le Cozannet et al., 2015; Vousdoukas et al., 2020; Vousdoukas et al., 2017; Lambert et al., 2020; Hinkel et al., 2021; MacPherson et al., 2011; Kirezci et al., 2020; Muis et al., 2017). Consequently, differences in flood extent, and by extension flood risk, caused by variance in other ESL parameters (e.g. storm surge duration) are disregarded. In fact, only two studies could be found where the effect of ESL temporal variability on coastal inundation is quantified (Quinn et al., 2014; Santamaria-Aguilar et al., 2017).

In the study by Quinn et al. (2014), the number of inundated buildings predicted during an ESL with a return period of 200 years in Portsmouth in the UK, differed by approximately 30% when comparing simulations using the upper and lower bounds of time series variability. It was also shown that temporal variability of the ESL time series leads to larger differences where tides are small, suggesting the effect may be larger along the German Baltic coast. Further, Santamaria-Aguilar et al. (2017) show for a case study at St. Peter Ording on the German North Sea coast, that overflow volumes can be up to three times higher when ESL temporal variability is considered.

Although the studies by Quinn et al. (2014) and Santamaria-Aguilar et al. (2017) highlight significant differences in flood characteristics due to ESL temporal variability, it remains uncommon to integrate the temporal variability of ESL time series into flood risk assessments. Instead, an idealised *design* curve is typically used, usually based on some observed event and stretched to the desired height (e.g., Dawson et al., 2005; Purvis et al., 2008; Batstone et al., 2013; Gallien et al., 2011; Quinn et al., 2013; MacPherson et al., 2011; Kiesel et al., 2022). Such methods ignore ESL temporal variability completely, and may therefore underestimate flood extent.

One of the main reasons flood risk assessments consider only single events, is that obtaining a sufficiently large number of ESL events is difficult. Directly acquiring ESL events from tide-gauge measurements is often impractical due to the insufficient length of observational records (see Muis et al., 2016; Batstone et al., 2013; Kiesel et al., 2022; MacPherson et al., 2011; Haigh et al., 2014a). Alternative methods, such as hydrodynamic numerical modeling (e.g., Haigh et al., 2014b; Arns et al., 2015a) and empirical approaches that analyse individual contributions to ESL and their interactions (e.g., Gönnert and Sossidi, 2011), are limited in the number of scenarios they can generate due to their computational expense and time-consuming nature.

To address this, Quinn et al. (2014) quantify ESL temporal variability using an ensemble of the largest 1% of peak water levels. By measuring water levels over time proportional to the peak, lower and upper bounds based on the 5^{th} and 95^{th} percentile values could be defined and transferred to a design curve with a 200-year return period. A similar method is used by Santamaria-Aguilar et al. (2017). However, these methods are only applicable where a reasonably large sea level data set is available. Quinn et al. (2014) use approximately 20 years of observations from tide-gauge records whereas Santamaria-Aguilar et al. (2017) use 39 years of simulated water levels produced using a hydrodynamic numerical model (Arns et al., 2015a). In both studies, very few large ESL events (e.g. 100 year return period) are present within the available data, and the applicability of the computed upper and lower bounds of temporal variability on such an event is questionable. This issue is further exacerbated by a lack of tides, which regulate the water level signal (Quinn et al., 2014). Thus, the validity of such a method applied along the microtidal German Baltic coast is tenuous.

An alternative method by Wahl et al. (2011) proposes a stochastic storm surge generator which is capable of producing large numbers of hydrographs for flood risk assessments. Here, ESL events are

identified within sea level records and parameterised. A number of probability distribution functions are used to model these parameters and synthetic ESLs can be constructed using the modelled values. These events are then filtered and adjusted based on interdependencies between the event parameters, resulting in a large set of realistic ESL hydrographs. As the distribution of each parameter is modelled separately, the method is capable of extrapolating beyond the length of data to produce large ESL events. Furthermore, as the events are modelled statistically rather than numerically, it is computationally inexpensive.

Research aim objectives

However, modelling hydrographs using the method of Wahl et al. (2011) is not suited to the microtidal German Baltic coast. Although it may be transferred to other locations (the study focuses on the Elbe estuary), it relies on a strong tidal signal to define the general shape of individual ESL events. Therefore, the aim of this paper is to address the first key objective of my doctoral research: to *model ESL hydrographs along the German Baltic coast*. To achieve this, I focus on two specific objectives for the study region:

1. Provide accurate estimates of ESLs.

2. Simulate the temporal variability of ESL hydrographs.

Data and methods

In this study, I utilise high-resolution (sampled hourly) observational sea level data collected from 45 tide-gauge records situated along the German Baltic Sea coast. These data provide between 14 and 66 years of sea level information, and are referenced to the German ordnance datum NHN. To address the first objective, a thorough analysis of ESLs in the region was conducted, and best practices for detrending, sampling and parameter estimation were employed, as suggested by Arns et al. (2013) (see Section 2.1.1 for an overview).

Next, a stochastic ESL model was developed following the methodology of Wahl et al. (2011), where ESL hydrographs are identified, parameterised and then modelled stochastically using Monte Carlo simulations. While this general framework remains unchanged, I implemented several large modifications to improve the methodology and adapt it to the microtidal environment of the German Baltic coast. These modifications are:

1. **A new method to identify ESL hydrographs**. While Wahl et al. (2011) define a hydrograph as encompassing three tidal cycles, this approach is not feasible along the German Baltic coast due to the region's small tidal range. Therefore, a new iterative

method was developed that tests the low waters surrounding each ESL event to identify the beginning and end of the hydrograph.

2. **A new parameterisation scheme**. In contrast to Wahl et al. (2011), who characterise hydrographs according to the dominant shape of the tidal cycle, I defined each hydrograph based on the general shape of their flow and ebb curves (the leading and trailing water level curves surround the peak water level). Differences between this general shape and the observed hydrographs were accounted for using an autoregressive model. In total, 17 parameters are used.

3. **A new method to model parameter interdepencies separate from parameter values**. In the method by Wahl et al. (2011), probability distibution functions are fitted to the parameters of the ESL hydrographs and sampled using Monte Carlo simulations. Interdependencies between observed parameters are transferred to the set of synthetic parameters by adjusting their distributions based on linear regression. Later, filter functions are used to remove any deformed events. Although this approach produces realistic individual ESLs, both the adjustment of parameter distributions and filtering of results leads to a set of events which no longer represent the set of observed hydrographs. In other words, the distributions of artificial parameters differ from those of the observed parameters. To address this, I model parameter interdependencies separate from parameter values by employing a Gaussian copula. This multivariate distribution models the dependence structure of the parameters over the unit hypercube $[0, 1]^d$, which may be used as sampling points for the individual parameter distributions. As a result, parameter interdependencies are maintained throughout the process and no adjustments of the underlying distributions are needed.

The performance of the model is measured by comparing two key parameters of the observed and artificial sets of ESL hydrographs, peak water level and intensity. The latter is a measure of the fullness of the hydrograph, defined as the time-distance area beneath the water level curve. Thus, a higher intensity event will maintain a higher water level for a longer duration. It is useful to validate the model using event intensity for two reasons: 1) the parameter is not specifically modelled in the set of 17 parameters, and 2) the intensity of any ESL event is affected by all 17 modelled parameters.

Results

Results show that the stochastic ESL model is highly capable of pro-

ducing realistic ESL hydrographs along the German Baltic Sea coast. Differences between the peak water level of observed and artificial events are small (<2 cm) at all sites. However, this is to be expected as peak water levels are derived directly from a parametric distribution function fitted to the observational data. These results however, confirm the accuracy of the underlying copula which models parameter interdepencies. Differences between observed and artificial event intensities are larger, but still show good agreement with observational data. The largest errors are found at sites along the bodden coasts, where local bathymetry restricts the flow of water resulting in ESLs with large durations. Nevertheless, model predictions of ESL intensities remain relatively accurate and fall within the 95% confidence bounds for most tide-gauges, up to a return period of 10,000 years. The key outcomes can be summarised as follows:

Key outcomes

1. The stochastic ESL model developed in this study is capable of producing large numbers of hydrographs along the German Baltic coast.

2. Model output is useful for quantifying uncertainties in flood extent due to temporal variability of ESL events.

3. The method used to identify the full extent of ESL hydrographs is more robust than the typical method of declustering when performing EVA.

Conclusions

The artificial ESL hydrographs generated by the model are valuable for flood risk analyses, allowing for the assessment of uncertainties in flood extent due to temporal variability. The practice of identifying individual ESL hydrographs offers an alternate, and more robust approach to declustering, especially in microtidal environments, ensuring independence between sampled events. The model is computationally efficient, enabling the production of large datasets covering a wide range of peak water levels and intensities. Therefore, the model is particularly useful when conducting probabilistic flood risk assessments, which require large numbers of events at high return periods. Furthermore, the model can account for changes in the distribution of ESLs, producing sets of artificial hydrographs that include the effects of climate change or are informed by historical information.

3.2 BAYESIAN EXTREME VALUE ANALYSIS OF EXTREME SEA LEVELS ALONG THE GERMAN BALTIC COAST USING HISTORICAL INFORMATION

Motivation

The German Baltic Sea coast has a long history of ESL induced coastal flooding. Of particular note is the storm surge of 1872, which caused widespread flooding in Germany, Denmark and parts of Sweden (see Hallin et al., 2021). Along the German Baltic coast, water levels rose to a maximum height of approximately 3.4 m above NHN at Travemünde (Jensen et al., 2022). Such was the devastation of the resulting floods, that the former Prussian states were ordered to compile available information regarding the event. The final report by Baensch (1875) details the atmospheric conditions and water levels along much of the current German Baltic coast.

Until recently, there has been a large disparity between estimated ESLs and historical events along the German Baltic coast (see MELUR, 2012; Jensen and Töppe, 1990). Coastal protection in the region is designed according to the water level of an ESL with a return period of 200 years (MELUND, 2022), and is referred to in technical reports as HW200. An additional 0.5 m climate surcharge (Klimazuschlag) is included to account for uncertainties arising from future changes in sea level and storm surges (MELUR, 2012). Although, this value has recently been increased to 1.0 m (MELUND, 2022). The official HW200 value for Travemünde in 2012 was given to be 2.24 m above NHN, the standard vertical datum used in Germany (MELUR, 2012). Including the climate surcharge at the time, design heights at Travemünde between 2012 to 2022 were 2.74 m, 66 cm less than the event observed in 1872.

Despite the exceptional nature of the 1872 event, its inclusion in the analysis of ESLs has rarely been attempted. This is due to two main challenges: first, incorporating historical information into traditional EVA is not straightforward (Prosdocimi, 2018); second, the event has often been regarded as an outlier (Hofstede and Hamann, 2022). Regarding the first point, traditional EVA relies on a well-defined *period of observation* (Prosdocimi, 2018) during which all extremes have occurred and are known. For tide-gauge records, sea levels are measured systematically, capturing all events within a known period. However, historical records lack a formal definition of what constitutes an extreme, and their documentation has therefore been subjective rather than objective (i.e. floods that resulted in deaths were documented whereas other events were not). It is therefore difficult to combine historical measurements together with systematic data.

To the second point, classification of the 1872 event in terms of return periods has often produced extremely large estimates that cannot be reliably computed. Gräwe and Burchard (2012) use observations and simulation data to assign a probability to the 1872 event at three sites along the German Baltic coast. At each site, return periods in excess of 100,000 years are computed. Mudersbach and Jensen (2009) extend the available sea level records at several locations to cover a period which includes 1872. Here, a correlation analysis between neighbouring gauges was used to fill in missing data values. The expanded record covers 180 years, and provides a return period estimate of between 3,000 and 10,000 years for the 1872 event (3,400 years for the preferred method).

The study by Gräwe and Burchard (2012) illustrates why the 1872 event is often disregarded as an outlier in assessments of ESL. However, they rely solely on systematic data for the estimation of return periods. In contrast, Mudersbach and Jensen (2009) show that by utilising both systematic and historical data, the 1872 event is more effectively integrated into the overall dataset. Although their methodology is practical and feasible, it can only provide data as far back as 1826 due to the availability and reliability of sea level records. However, there exists evidence of other historical events along the German Baltic coast, comparable in magnitude to that of 1872 (Jensen and Müller-Navarra, 2008; Jensen et al., 2022), the earliest of which occurred in 1044. Incorporating these events may provide further benefits in terms of improved estimate accuracy, reduced uncertainties, and the reclassification of 1872 as an exceptional, but statistically important non-outlier.

Research aim objectives

The aim of this paper is to address the second key objective of my doctoral research: to *investigate how historical information can be incorporated into EVA to improve estimates of ESLs*. To achieve this, I define three specific objectives:

1. Incorporate historical information into EVA of ESLs along the German Baltic coast.

2. Quantify the effect of historical information on return water level estimates.

3. Investigate changes in ESL activity for the region.

Data and methods

In this study, I assembled a number of systematic and historical data at seven sites along the German Baltic coast. To incorporate both data types into EVA, I used a Bayesian framework for both the GEV and GP

distributions. The Bayesian method (see Coles et al., 2003; Coles and Tawn, 2005; Payrastre et al., 2011; Hamdi et al., 2015; Bulteau et al., 2015) incorporates historical information into the likelihood function as censored data, offering a natural approach to handle uncertainties. Parameter estimation for the underlying distribution is performed using a MCMC approach based on the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). Results of the MCMC may be used to provide ESL quantiles, including a maximum likelihood estimate and uncertainties in terms of credibility intervals.

Changes in ESL activity were analysed using the long (195 years) record of combined systematic and historical AMAX at Travemünde. Return water level estimates and their uncertainties were assessed using the Bayesian MCMC approach for a moving 70-year window of data. Finally, long-range dependence was assessed using the Hurst exponent (Hurst, 1951), which can indicate the presence of long-term variability.

Results

The incorporation of historical information into EVA provides several benefits, including improved accuracy, reduced uncertainties and better representation of outliers. At all locations, return water level estimates are significantly increased when historical information is included. At Travemünde, Wismar and Warnemünde, ESL activity exhibits long-range dependence, with clear decreases in the level of ESL activity during the latter part of the records. This has large implications for coastal adaptation in the region as this period covers the full extent of systematic sea level measurement. Consequently, estimates made using systematic tide-gauge data alone may be prone to severe underestimations. The key findings are summarised as:

Key findings

1. Historical information provides several benefits when estimating return water levels, including reduced uncertainties and better incorporation of outliers.

2. The German Baltic region is currently experiencing a period of low ESL activity due to some long-term variability. Consequently, return water levels derived from systematic tide-gauge data alone are prone to underestimations.

Conclusions

The application of Bayesian techniques allows for the incorporation of several data types when modelling ESLs. This method provides a natural framework to account for historical information and their inherent uncertainties. Despite significant uncertainties surrounding the true value of historical events, their inclusion in EVA provides

significant benefits. At the German Baltic coast, the computation of return water levels using systematic data alone is prone to underestimations, however the inclusion of one exceptional historical event (i.e. 1872) can enhance estimates similar to the those made using much longer tide-gauge records.

3.3 A COMPREHENSIVE PROBABILISTIC FLOOD ASSESSMENT ACCOUNTING FOR HYDROGRAPH VARIABILITY OF ESL EVENTS

Motivation

In the face of accelerating climate change and rising sea levels, coastal regions worldwide are increasingly vulnerable to ESL events (IPCC, 2022; Kummu et al., 2016; Tebaldi et al., 2021). Traditional flood risk assessments often consider only one or few events (e.g., Dawson et al., 2005; Purvis et al., 2008; Batstone et al., 2013; Gallien et al., 2011; Quinn et al., 2013; MacPherson et al., 2011), failing to capture the full spectrum of potential flooding scenarios and leading to gaps in our understanding and preparedness. The effects on flooding from a wide range of potential ESL events and their associated probabilities can be quantified through probabilistic flood risk assessments. Such assessments are crucial for developing robust coastal defense strategies and ensuring the resilience of coastal communities.

Unlike deterministic methods, which may focus on a single or a few scenarios, probabilistic assessments incorporate large ensembles of potential events. This approach allows for a more accurate representation of flooding due to the uncertainty and variability inherent in ESLs (Le Cozannet et al., 2015). By considering numerous potential scenarios, probabilistic assessments can provide a more detailed and nuanced understanding of flood risks, enabling better-informed decision-making for coastal protection (Ward et al., 2011; Hinkel et al., 2015).

One of the critical aspects that probabilistic flood risk assessments can address is the temporal variability of ESL events. As discussed in Section 2.2, traditional assessments often rely on static peak water levels, overlooking the dynamic nature of ESLs. However, the temporal evolution of ESL hydrographs can be a significant determinant of flood impacts. (Quinn et al., 2014; Santamaria-Aguilar et al., 2017). Probabilistic assessments that account for hydrograph variability can thus provide a more accurate depiction of flood risks, capturing the full range of possible outcomes.

Conducting probabilistic flood risk assessments presents several significant challenges however, spanning data acquisition, method-

ological complexities and computational demands. Regarding data acquisition, there are two main challenges: 1) accurate estimates of return water levels are required to accurately assess flood risk, and 2) large numbers of ESL scenarios are needed to fully explore their effects on flooding, including the influence of ESL temporal variability.

Research aim and objectives

Following the overall aim of my doctoral research: *to enhance available sea level information along the German Baltic coast for use in flood risk assessments*, both of these issues have been addressed. In my second paper (see Section 3.2, I incorporated historical information to improve estimates of return water levels, and in my first paper (see Section 3.1), I developed a stochastic ESL model capable of generating large numbers of hydrographs for flood risk assessments. The aim of this paper therefore corresponds to my final key objective of my doctoral research: to *demonstrate how these improved information can be used to perform probabilistic flood risk assessments*. The specific objectives of the paper are:

1. Develop a framework in which large numbers of ESL events may be simulated in a hydrodynamic inundation model.

2. Utilise the return water level estimates and ESL scenarios from my first two papers in combination to perform a probabilistic flood risk assessment

Data and methods

In this study, a number of data were used for the development and forcing of the inundation model. Bathymetric and topographic data were used to construct the high-resolution (10 m) model grid, with bottom roughness specified according to land usage. The model was validated using observational tide-gauge data, and flooding was simulated under ESL hydrographs generated by the stochastic model developed in my first paper (Section 3.1). Return water level estimates were taken from my second paper (Section 3.2).

The hydrodynamic inundation model was developed in Delft3D for a case study of Lübeck, situated on the river Trave connected to the Baltic Sea. The model domain covers an area of 305 km^2, and includes the historic city center of Lübeck and Travemünde at the coast. The FLOW-module of the fully integrated open-source hydrodynamic model Delft3D (Lesser et al., 2004) is applied in a (2D) depth-averaged mode.

Flood simulations using the numerical model described above are computationally expensive. To allow for a large number of scenarios to be simulated, a framework was developed in which model setup, execution and analysis could be performed in an automated procedure. In combination with high-powered computing facilities provided by the Christian-Albrechts-University of Kiel, a full range of ESL scenarios could be tested.

In total, 311 simulations were conducted. First, the effects of flooding under 37 return water levels were assessed, ranging from 1.6 m (∼4-year return period) to 3.4 m (∼3,650-year return period) in 5 cm increments. The temporal variability of the ESL hydrographs at each water level was accounted for by simulating events that correspond to the 5^{th}, 50^{th} and 95^{th} percentile of intensity. These were created by generating 100 hydrogaphs at each water level using the stochastic ESL model from my second paper, and extracting those events which correspond to the desired intensity level.

Next, a more in-depth analysis of the effect of ESL temporal variability on flooding was conducted. Here, 100 ESL hydrographs were generated at peak water levels corresponding to the 50-year return water level (2.24 m) and the 200-year return water level (2.55 m) as specified by MELUND (2022). Flood simulations were conducted for all 200 events, and an analysis was conducted based on flood extent and hydrograph intensity.

Results

Results were assembled into probabilistic flood maps which detail flood extents according to probability. These maps are especially useful for flood management, as they highlight a large range of flooding possibilities due to variability in both return water levels and hydrograph intensity. Using these maps, we identify *tipping points* where flooding can increase substantially due to slight changes in peak water levels or hydrograph intensity. Presenting the results as maps allows us to identify areas where flood protection can address these issues. We find a large spatial variability in flood extents, driven by local topographies and dependent on the peak water levels of the forcing ESLs. This is also true when considering the temporal variability of ESL events, however the relationship between flood extent and hydrograph intensity is not linear. The key findings are:

Key findings

1. Sensitivity of flood extent to ESL temporal variability changes depending on location and the return water level of the forcing event.

2. Considering only the lower and upper bounds of hydrograph intensity does not capture the full potential of

FLOOD EXTENT AT ANY GIVEN RETURN WATER LEVEL, AS THE RELATIONSHIP BETWEEN FLOOD EXTENT AND HYDROGRAPH INTENSITY IS NOT LINEAR.

3. PROBABILISTIC FLOOD MAPS PROVIDE VALUABLE INFORMATION AND INSIGHTS INTO FLOODING UNDER A LARGE RANGE OF EVENTS. THIS DATA CAN ADDRESS VARYING LEVELS OF RISK AVERSION, PROVIDING COASTAL MANAGEMENT WITH SUBSTANTIAL SUPPORT FOR MAKING INFORMED ADAPTATION DECISIONS.

Conclusions

In this paper, we developed a framework to conduct large numbers of flooding simulations. In combination with hydrographs from my first paper (see Section 3.1; MacPherson et al., 2019) and return water level estimates from my second paper 3.2;][]macphersonBayesianExtremeValue2023, we quantify uncertainties in flood extent due to uncertainties in return water level estimates and hydrograph temporal variability. The results are presented as probabilistic flood maps, which present a comprehensive overview of flooding. In contrast to deterministic assessments, these highlight the range of possible flood extents, allowing for a more nuanced understanding of flood risk. By incorporating variability in both return water levels and hydrograph characteristics, our approach provides a more robust evaluation of flood hazards. This probabilistic perspective is crucial for risk-informed decision-making, offering stakeholders valuable insights into the likelihood and potential severity of flooding events under different scenarios. Such detailed assessments can better inform flood mitigation strategies and urban planning, ultimately contributing to more resilient infrastructure and communities.

3.4 HYDRODYNAMIC MODEL OF THE WESTERN BALTIC SEA

I have been involved in a number of collaborations during my doctoral studies where I have provided *enhanced sea level information along the German Baltic coast*, following the main aim of my work. Although these do not directly contribute to my specific objectives, they have been a substantial amount of work and provided valuable insights and supplementary data. In this section, I outline these contributions and summarise the research and key findings.

As discussed in Section 2.1.3, direct methods of EVA are not always possible due to insufficient observational data, both temporally and spatially. To overcome this limitation, hydrodynamic numerical models may be employed to simulate water propagation across a defined model domain, generating water level data at ungauged locations.

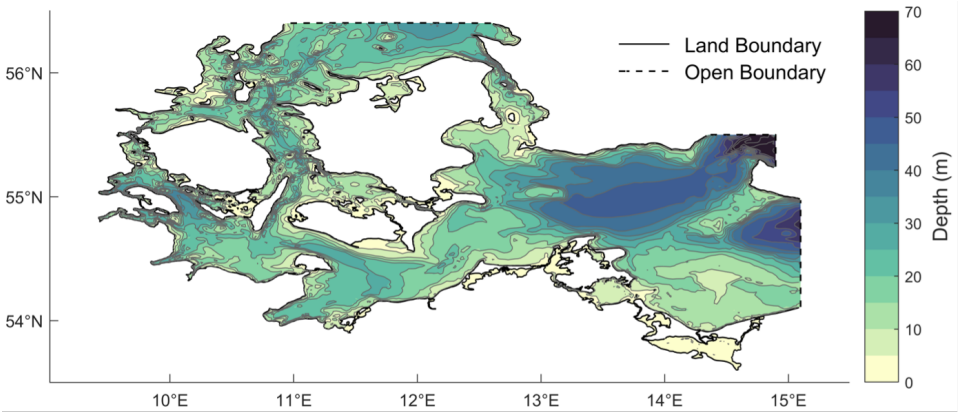

Figure 3.1: Hydrodynamic model domain and bathymetry.

As part of my research, I developed a hydrodynamic model of the western Baltic Sea to extend available sea level information along the German Baltic coast. This model was used to provide enhanced information on ESLs for two research papers (van der Pol et al., 2021; Beckmann et al., 2022). In this section, I describe the hydrodynamic model and outline how this work has contributed to the papers mentioned above.

Model setup

Water levels in the western Baltic Sea were simulated using the model system SCHISM (Semi-implicit Cross-scale Hydroscience Integrated System Model; Zhang et al., 2016), a derivative product built from the original SELFE (Semi-implicit Eulerian–Lagrangian Finite-Element; Zhang and Baptista, 2008). Although developed to improve the performance of unstructured grid models in baroclinic regimes cross-scale, SCHISM has a well-documented record as a barotropic tide-surge model (e.g., Fernández-Montblanc et al., 2019; Fortunato et al., 2017; Krien et al., 2017).

The model grid was constructed using bathymetry data supplied by the European Marine Observation and Data Network (EMODnet, 2022) with a resolution of 1/8-minute (0.0021°) (see Figure 3.1). Grid resolution ranges from 2 km in the open Baltic Sea to ∼300 m along the German coast. Atmospheric pressure gradients and wind velocities used for model forcing were taken from the EU project "Uncertainties in Ensembles of Regional Re-Analyses" (UERRA), which provides hourly data with an approximate resolution of 11 km (Ridal et al., 2024). Water levels and velocities at open boundaries were extracted from a regional ocean model of the Baltic Sea (Gräwe et al., 2019).

Model calibration

The model was calibrated by simulating several ESL events from January 1987, November 1995, February 2002 and November 2006, and comparing model output with available observational data. A

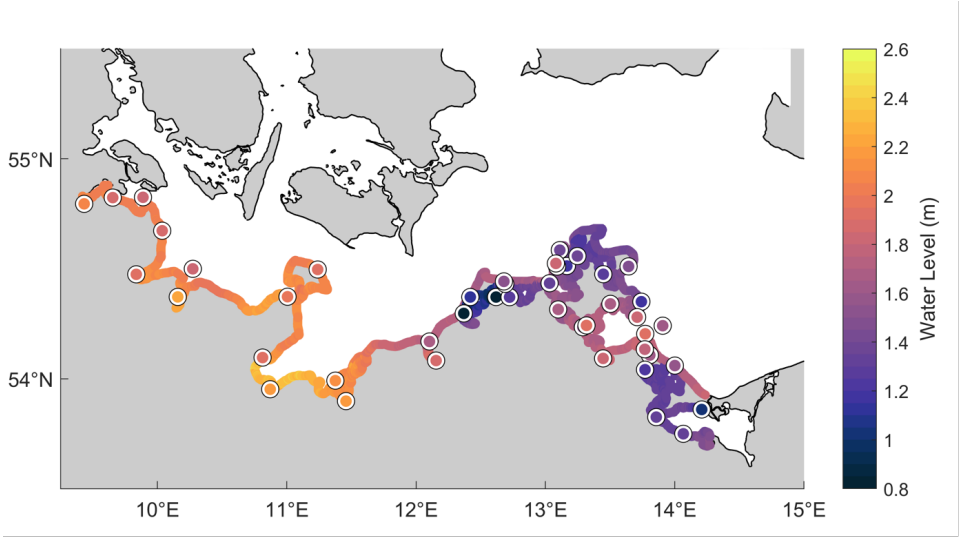

Figure 3.2: Estimated 200-years return water levels for the entire German Baltic coast. Estimates made using observational data are overlaid as circles for comparison.

spatially constant bottom roughness parameter was modified to provide the best results at available tide-gauge records. The performance of the model at these sites was calculated using root mean square error (RMSE), r-squared and by measuring the difference in peak water levels. I provide a detailed description of the model setup and validation in Beckmann et al. (2022).

3.4.1 *Hindcast simulations and return water level estimates*

A 36-year hindcast of sea levels was performed, providing hourly sea level data for each grid point within the model domain. A bias correction featuring both parametric and non-parametric components was developed for the German Baltic coast, which assessed differences between modelled and simulated sea levels at available tide-gauge locations. This correction was transferred to un-gauged sites using inverse distance weighted interpolation. Finally, EVA was conducted along the coast using hindcast data, giving return water level estimates and associated 95% confidence intervals. Due to the high-resolution of the model grid, these estimates are provided every ∼300 m along the coast. Estimates of the 200-years return water level for the entire German Baltic coast are shown in Figure 3.2.

Summary of van der Pol et al. (2021)

Return water level estimates based on the hindcast data are employed in the study by van der Pol et al. (2021). In this paper, a number of flood protection strategies for the German Baltic Sea coast

are explored using a combination of cost-benefit analysis and robust decision-making methods. The paper deals with deep uncertainty in future flooding due to the compounded uncertainties surrounding ESLs and future SLR. This required highly resolved estimates of the 200-years return water level along the entire German Baltic coast, which was derived using the hindcast sea level data described above.

Cost-benefit analyses were conducted to identify efficient flood protection strategies for the region under a large number of scenarios. Efficient strategies are those that reduce costs relating to the implementation and maintenance of flood protection, as well as any potential damages if flooding occurs. Next, robust optimisation methods were applied to investigate how each strategy performs across scenarios. These methods assessed the performance of each strategy according to five decision criteria for loss or regret aversion. The key findings are:

Key findings of van der Pol et al. (2021)

1. Combining cost-benefit analysis with robust decision-making methods aids in identifying a smaller number of flood mitigation solutions that are both potentially efficient and robust.

2. Up to ∼10% of the German Baltic coast may currently be underprotected from a social welfare perspective.

3.4.2 *Simulating the ESL of 1872*

The hydrodynamic numerical model was further employed to simulate water levels in the western Baltic Sea during the ESL of 1872. Atmospheric forcing was taken from the 20th Century Reanalysis Project (Compo et al., 2011) which provides data from 1836 onwards. Here, the three hour ensemble mean was used, which represents a *very likely* state of the global atmosphere (Slivinski et al., 2019). However, the data is not sufficiently resolved, neither spatially (1°) nor temporally (three hours), for accurate simulations of ESLs. Furthermore, no data is available for the open boundaries of the hydrodynamic model. These factors combined result in water level predictions that severely underestimate available observations.

Summary of Beckmann et al. (2022)

Despite the significant differences between modelled and observed water levels during the 1872 ESL, Beckmann et al. (2022) employ a statistical bias correction to account for model inaccuracies. Here,

modelled sea levels are adjusted based on a difference function derived at locations where observational data is available (see Baensch, 1875), and interpolated spatially using an inverse distance weighted approach similar to the method I employed in van der Pol et al. (2021). The performance of the method was assessed by omitting certain observational data when calculating the bias correction. These data could then be used for validation. Here, the corrected modelled water levels show good agreement with observational data, suggesting that the application of a non-parametric bias correction provides good results at un-gauged sites. The key findings are:

Key findings of Beckmann et al. (2022)

1. Simulating historical ESLs using hydrodynamic models can provide valuable information, but are severely limited by the availability of high-quality, high-resolution atmospheric data.

2. Bias correction methods provide a simple and effective approach to account for inaccuracies in hydrodynamically simulated water levels.

3. Accounting for model biases can provide valuable information to inform coastal flood risk mitigation.

4

SYNTHESIS

In this section, I summarise the outcomes and key findings of my work, highlighting their relevance to the broader field and identifying avenues for future research. By synthesising the research findings, this section offers a comprehensive overview of the study's contributions, underscoring their importance in advancing knowledge and informing practice within the domain of ESL modelling and coastal flood risk management.

4.1 SUMMARY OF MAIN RESEARCH FINDINGS

The work presented in this thesis was done with the aim of enhancing available ESL information for use in flood risk assessments along the German Baltic coast. For my first two objectives, I developed a stochastic ESL model to generate large numbers of hydrographs and incorporated historical information in EVA to improve estimates of return water levels. As part of my third objective, I demonstrated the value of this enhanced information when conducting a probabilistic flood risk assessment. In addition to these, I collaborated on a number of studies which utilised enhanced ESL information.

The main findings of my research can be summarised into the following three categories: 1) Statistical and numerical techniques can enhance available extreme sea level information, 2) the German Baltic coast is experiencing an extended period of low ESL activity, and 3) the temporal variability of extreme sea levels can lead to substantial differences in flood exposure. In this chapter, I summarise these findings and discuss challenges that remain, providing recommendations for future research.

4.1.1 *Statistical and numerical techniques can enhance available extreme sea level information*

At the outset of my doctoral research, I identified two key limitations in current flood risk assessments that could be addressed with enhanced ESL information: (1) a lack of high-end ESL scenarios, and (2) large uncertainties in the estimates of return water levels.

In regards to the first limitation, I developed a stochastic ESL model capable of generating large numbers of ESL hydrographs (Section 3.1; Chapter 5; MacPherson et al., 2019). Here, an existing modelling framework from Wahl et al. (2011) was adapted and improved for the microtidal German Baltic coast. My enhancements allowed the model to produce ESL scenarios that cover a full range of return water levels and event intensities at sites along the German Baltic coast. This capability is particularly valuable for use in flood risk assessments, as it provides a comprehensive set of potential scenarios to be explored. The effectiveness of this model is demonstrated in detail in my third paper (Section 3.3; Chapter 7; Kupfer et al., 2024).

In my second paper (Section 3.2; Chapter 6; MacPherson et al., 2023), I tackled the issue of large uncertainties in the estimates of return water levels. Here, I utilised systematic tide-gauge data in combination with historical information to compute return water level estimates with increased accuracy. I found that the inclusion of historical information can enhance estimates of return water levels, resulting in reduced uncertainties. Furthermore, I discovered that incorporating even a single historical event (i.e. the 1872 ESL) can significantly enhance estimates made using short tide-gauge records, aligning them with estimates derived using much longer data sets.

As part of my collaboration with other researchers (van der Pol et al., 2021; Beckmann et al., 2022), I developed a hydrodynamic numerical model of the western Baltic Sea (see Section 3.4). This model was used to generate water levels at a resolution of $\sim$300 m along the German coast, far exceeding the spatial resolution of observational data. The model was used to generate a 36-year hindcast of sea levels (van der Pol et al., 2021, see Section 3.4.1), allowing for the computation of high-resolution spatially variable return water level estimates and a regional analysis of flood defences. Further, coastal ESLs were simulated for the exceptional 1872 event and corrected by Beckmann et al. (2022, see Section 3.4.2). These simulations produced sea level information with a far greater spatial coverage provided by the observational network.

4.1.2 *The German Baltic coast is experiencing an extended period of low ESL activity*

In my second paper (Section 3.2; Chapter 6; MacPherson et al., 2023), I discovered that the occurrence of ESLs along the German Baltic coast exhibits long-term variability which may affect current estimates of return water levels. Analyses of long-range dependence in the records of AMAX water levels at Travemünde, Wismar and Warnemünde reveal significant patterns of fluctuation over extended periods. The cause of this variability was not investigated as it was considered outside the scope of my paper.

In recent decades, the German Baltic coast appears to be experiencing a period of low ESL activity, with larger ESLs occurring significantly more often in the first half of the long AMAX records. This has large implications for the assessment of coastal flood risk in the region. First, estimates of return water levels made using short tide-gauge records may not reflect the true distribution of ESLs. Second, the period of relatively low ESL activity covers the period of systematic tide-gauge observations entirely. Therefore, estimates of return water levels made solely using tide-gauge data may be prone to underestimations.

4.1.3 *Probabilistic flood modelling provides a comprehensive view of flooding probabilities*

Flood models typically only consider uncertainties in the estimates of return water levels. Following the findings of Quinn et al. (2013) and Santamaria-Aguilar et al. (2017), we show that the temporal variability of ESLs can be a major source of uncertainty in flood extent (Section 3.3; Chapter 7; Kupfer et al., 2024). These uncertainties are quantified and communicated using probabilistic flood modelling, where large numbers of scenarios are tested, covering a range of return water levels and ESL temporal variability.

We find that the effect of temporal variability changes depending on the return water level and local topographic features. There exists *tipping points* where longer duration events may result in significantly more flooding. This study highlights the benefit of probabilistic flood models, which provide a comprehensive view of flooding probabilities and aid in identifying vulnerable areas.

4.2 RECOMMENDATIONS FOR FUTURE RESEARCH

4.2.1 *Non-stationary extreme value analysis of extreme sea levels along the German Baltic Sea coast*

In my second paper (Section 3.2; Chapter 6; MacPherson et al., 2023), I identify long-term variability in the occurrence of ESLs in the region. Despite this, the methods employed to assess ESLs were based on the i.i.d. assumption of classical EVA, and the ESL data was made stationary by detrending the data (Arns et al., 2013; Dixon and Tawn, 1999). However, a number of non-stationary approaches to EVA exist (Mentaschi et al., 2016; Muis et al., 2023; Muis et al., 2020), and could offer further insight into how the distribution of ESLs in the region is changing.

Further, the cause of long-term variability in the occurrence of ESL for the region remains unknown. A thorough investigation into the climatic and oceanographic processes which affect ESLs in the region was considered outside the scope of the original study, but may provide insights into the mechanism forcing these events. Futhermore, understanding these mechanisms may elucidate possible future changes in ESLs, providing useful information for future flood risk assessments.

4.2.2 *Bayesian vs frequentist methods of extreme value analysis*

Frequentist approaches to EVA have long been applied in the field of hydrology (see MacPherson et al., 2019; Arns et al., 2013; Haigh et al., 2010) in lieu of Bayesian methods due to their simplicity, well-developed theory and computational efficiency (Coles and Tawn, 2005). However, Bayesian methods offer several significant benefits which make them increasingly attractive when studying natural hazards such as coastal floods. Coles et al. (2003) adopt a Bayesian approach to EVA as it provides a more coherent framework for the handling of uncertainties. They also note that Bayesian approaches offer more flexibility, allowing for the analysis of data and models of greater complexity.

I have employed both frequentist (Section 3.1; Chapter 5; MacPherson et al., 2019) and Bayesian methods (Section 3.2; Chapter 6; MacPherson et al., 2023) in my research when analysing ESLs. Following Coles et al. (2003), I see several benefits of using Bayesian approaches to

EVA when considering ESL induced coastal flooding. Most beneficial is the natural framework for handling uncertainties provided by Bayesian methods. This framework provides a comprehensive methodology for interpreting and presenting uncertainties surrounding both observations and predictions of ESLs. As such, analyses may incorporate a broad range of data and their associated uncertainties, while providing probabilistic estimates of return water levels — which are extremely valuable when conducting uncertainty analyses of coastal flood risk. The computational limitations of Bayesian approaches are no longer an issue due to advancements in computing and simulation-based techniques such as MCMC. Therefore, I suggest that future research into ESLs and coastal flooding should adopt Bayesian approaches to EVA.

4.2.3 *Combined numerical and statistical assessment of extreme sea levels*

As discussed above, the German Baltic region appears to be experiencing a period of low ESL activity, driven by some long-term variability. Consequently, return water levels computed using tide-gauge data only are prone to underestimations. In my second paper (Section 3.2; Chapter 6; MacPherson et al., 2023) I show that including historical information in EVA can provide stable estimates of return water levels despite short tide-gauge records. However, tide-gauge data and information on historical ESLs are spatially limited, reducing the number of locations where such an analysis is feasible.

To address this, hydrodynamic numerical models can be used to extend available information spatially. In my research, I show that hydrodynamic models are capable of producing high-resolution sea level data (van der Pol et al., 2021, Section 3.4.1) and reproducing historical events (Beckmann et al., 2022, Section 3.4.2), filling the observational gaps in both systematic and historical data. Combining these two data in an analysis following the methodology of my second paper, would allow for more accurate and stable estimates of return water levels along the entire German Baltic coast.

4.2.4 *Regional analysis of Baltic Sea extreme sea level*

Although this dissertation focus on the German Baltic Sea coast, all analyses conducted are essentially local — focusing on individual

locations. As mentioned in Section 2.1.3, indirect methods of EVA can enhance the precision of local assessments. Specifically, regional frequency analysis (RFA) offers significant benefits in terms of improved return water level estimates. While studies have been conducted to identify regional differences in ESLs in the Baltic Sea (e.g., Wolski and Wiśniewski, 2021; Wolski et al., 2014; Ribeiro et al., 2014; Marcos and Woodworth, 2017; Wolski and Wiśniewski, 2020), none have employed this knowledge to improve local estimates of return water levels by applying RFA. Therefore, this remains an avenue for future research.

Part II

PUBLICATIONS

5

A STOCHASTIC EXTREME SEA LEVEL MODEL FOR THE GERMAN BALTIC SEA COAST

This chapter is published as:

MacPherson, L. R., Arns, A., Dangendorf, S., Vafeidis, A. T., and Jensen, J. (2019). A Stochastic Extreme Sea Level Model for the German Baltic Sea Coast. *Journal of Geophysical Research: Oceans*, 124, pp. 2054–2071. DOI: 10.1029/2018JC014718..

ABSTRACT

This paper describes a framework in which artificial extreme sea levels (ESLs) can be generated for use in flood risk analyses. Such analyses require large numbers of events to accurately assess the risk associated with certain return water levels and quantify uncertainties surrounding the temporal variability of ESL events. Stochastic models satisfy this requirement as they are computationally inexpensive, and thus, many thousands of events may be generated over a very short period of time. As a case study, we have developed a stochastic model for the German Baltic Sea coast capable of simulating the temporal behaviour of ESLs. To do this, high-resolution water level data from 45 tide gauges have been used as model input. At each location, observed ESLs are identified and parameterised. Artificial ESLs are generated using Monte Carlo simulations based on the parametric distribution functions fitted to the parameterised observed ESLs. We show that the method outlined here provides an accurate representation of ESLs at all tide-gauges tested. However, the model is limited by the availability, length, and quality of high-resolution water level data. Due to the rarity of ESLs in the German Baltic Sea, including historical measurements into the stochastic procedure allows for the generation of artificial ESLs more in-line with past extremes.

5.1 INTRODUCTION

Coastal flooding events caused by extreme sea levels (ESLs) are a major threat to coastal systems with both direct and indirect consequences for human populations. Currently, an estimated 310 million people and US$11 trillion of assets are exposed to a 100-year coastal flood event globally (Hinkel et al., 2014). This is expected to increase throughout the 21st century due to a combination of several factors, including an increased frequency of ESL events as a result of rising mean sea levels (MSLs) (Church et al., 2006; Dangendorf et al., 2017; Haigh et al., 2011), along with population growth, urbanization, and socioeconomic development (Hanson et al., 2011; Jongman et al., 2012; Merkens et al., 2016; Neumann et al., 2015). Under no adaptation, losses due to coastal flooding may reach 10% of global gross domestic product by 2100 (Hallegatte et al., 2013). For much of the world's developed coastlines, providing coastal defenses against ESLs is a more economically viable option than inaction (Hinkel et al., 2013).

To ensure efficient planning of coastal defenses and avoid costly overdesign, an accurate assessment of risk must be conducted to inform policy decisions. In the case of coastal flooding, risk is determined by quantifying potential monetary losses associated with specific return water levels (Stewart and Melchers, 1997). This is typically done by means of inundation modeling, where the forcing water levels are estimated using extreme value analysis (EVA; Coles, 2001). However, a number of uncertainties may arise during this process and it is important that these are quantified and properly communicated (Teng et al., 2017). In regard to the model forcing data, uncertainties emerge as a consequence of the statistical methods used to estimate return water levels and a lack of empirical data, in both time and space. When these data are extended using storm surge models, further uncertainties are introduced due to the specific model used, model parameters, and grid resolution (Wahl et al., 2017). Additionally, coastal flood risk analyses typically consider only one event for each specific return water level; however, the temporal evolution of the storm surge can be a major factor in flood impact due to overflow volumes alone (Quinn et al., 2013; Santamaria-Aguilar et al., 2017). Quantifying these uncertainties in terms of flood risk is typically handled via sensitivity analyses (Teng et al., 2017). Therefore, an accurate assessment of the risk associated to coastal flooding, including uncertainties arising from ESL estimates and temporal variability, requires a large number of ESL scenarios, each containing information on the temporal evolution of water level with respect to the local conditions.

Deriving a sufficiently large number of ESL events for risk analyses presents several challenges depending on the method employed. Directly obtaining ESL events from tide-gauge measurements is infeasible due to the lack of observational records of sufficient length. Other methods, such as hydrodynamic numerical modeling (Haigh et al., 2014b) and empirical approaches where individual contributions to ESL and their interactions are analyzed (Gönnert and Sossidi, 2011), limit the number of scenarios generated as they can be computationally expensive and time-consuming. Wahl et al. (2011) outline a method to stochastically model large numbers of ESL events at two tide-gauge locations in the German North Sea, Cuxhaven, and Hörnum. The statistical model identifies ESL events within tide-gauge records, characterizes each event through a parameterization scheme, and generates artificial storm surge curves based on parametric distributions functions fitted to each parameter. While the model produces good results at the two tide-gauge locations mentioned above, further development is required to apply the model along larger coastline stretches such as the entire German Baltic Sea coast. One key requirement is to adapt the model for the microtidal environment of the Baltic Sea, as the parameterization scheme used by Wahl et al. (2011) is highly dependent on the distinct tidal signal present at the two North Sea tide-gauges. Through this stochastic method, large numbers of storm surge curves may be generated while remaining computationally inexpensive.

The overall aim of this study is to develop a framework in which large numbers of ESL events may be generated stochastically for use in flood risk analyses. First, the main focus is to provide accurate estimates of return water levels and second simulate the temporal variability of ESL time series using Monte Carlo simulations. As a case study, we have developed a stochastic ESL model for the German Baltic Sea coast, capable of generating artificial ESL events where each is a time series of water levels, hereafter referred to as hydrographs. The model has been used to generate artificial ESL events at 45 locations along the German Baltic Sea coast. Information on the study area and data used during this study is given in section 2. Methods used to stochastically simulate ESL events are presented in section 3. Results of the stochastic simulations and validation techniques are given in section 4 with final conclusions given in section 5.

5.2 BACKGROUND

5.2.1 *Study Area*

The Baltic Sea is a semienclosed basin where the water level is influenced strongly by large-scale atmospheric forcing (Andersson, 2002; Jevrejeva et al., 2005). Salt water inflows from a connection to the North Sea via the Kattegat and freshwater river discharges result in a highly stratified water column and weak estuarine circulation within the sea (Feistel et al., 2008). Tides in the Baltic Sea are small (0.1–0.2 m) and thus contribute little to ESLs. Due to its location in the western Baltic Sea, ESLs occur along the German coast under strong northeasterly winds. Seiches acting over the entire Baltic Sea may also contribute several decimeters to the total water level. Hence, ESLs are influenced not only by local winds but by certain temporal sequences of regional wind patterns (Jensen and Müller-Navarra, 2008). This was most evident during the storm surge of 1872, which resulted in the highest recorded ESL along the German Baltic Sea coast. Here sustained southwesterly winds caused high waters in the eastern Baltic Sea and allowed for large volumes of water to enter the western Baltic Sea through the Kattegat. The forcing storm then reversed direction and intensified (Rosenhagen and Bork, 2009). Strong northeasterly winds and restricted water flow at the North Sea-Baltic Sea border resulted in ESLs along much of the German Baltic Sea coast, with a maximum water level of 3.4 m above MSL recorded at Travemünde. Such events are further influenced by local bathymetry; thus, there is a tendency for higher ESLs in the Bays of Kiel and Mecklenburg (Jensen and Töppe, 1990).

The German Baltic Sea coast was formed during the last glacial period and consists mostly of shallow sandy coasts, marshlands, and beach ridges. Around half of the coastline belongs to the Bodden coast (semienclosed inlets), where shallow bays are formed by small islands and peninsulas. Of the total 2,110 km of coastline, 1,150 km is experiencing erosion at an average rate of approximately 40 cm/year (Sterr, 2008). Design heights for ESL defenses in Germany are defined by government departments in federal states. The German Baltic Sea coast is bordered by two federal states, Schleswig-Holstein and Mecklenburg-Vorpommern. As the methods employed to determine these design water levels differ between states, the level of coastal protection is difficult to assess (Arns et al., 2013). Approximately 497,400 people and assets equivalent to 48 billion Euros lie below 5 m above MSL along the German Baltic Sea coast (Sterr, 2008). However, first-grade dikes protect only 13% of the coast, with dunes and other

Table 5.1: List of Tide-Gauge Records Used in This Study

		Location		Duration			Location		Duration
Tide gauge		lon	lat	(years)	Tide gauge		lon	lat	(years)
1	Flensburg	09°26	54°48	61	24	Neuendorf Ostsee	13°05	54°31	16
2	Schleswig	09°34	54°31	24	25	**Neuendorf Hafen**	13°05	54°31	31
3	Langballigau	09°39	54°49	24	26	**Stralsund**	13°06	54°19	54
4	Eckernförde	09°50	54°28	26	27	**Kloster**	13°07	54°35	46
5	LT Kalkgrund	09°53	54°49	26	28	**Wittower Fähre**	13°15	54°33	63
6	Kappeln	09°56	54°40	24	29	Stahlbrode	13°17	54°14	40
7	Schleimünde SP	10°02	54°40	25	30	Greifswald Eldena	13°27	54°06	52
8	Kiel-Holtenau	10°09	54°22	51	31	Greifswald Wieck	13°27	54°06	50
9	LT Kiel	10°16	54°30	26	32	**Ralswiek**	13°27	54°29	34
10	Lübeck Bauhof	10°42	53°54	25	33	Lauterbach	13°30	54°20	40
11	Neustadt	10°49	54°06	24	34	Sassnitz	13°39	54°31	61
12	Travemünde	10°52	53°57	66	35	Thiessow	13°43	54°17	32
13	Heiligenhafen	11°00	54°22	26	36	Göhren	13°45	54°21	19
14	Marienleuchte	11°14	54°30	24	37	Peenemünde	13°46	54°08	26
15	Timmendorf	11°23	53°60	54	38	**Wolgast**	13°46	54°03	50
16	Wismar Baumhaus	11°27	53°54	58	39	Ruden	13°46	54°12	50
17	Warnemünde	12°06	54°10	62	40	**Karlshagen**	13°48	54°06	23
18	Rostock	12°09	54°05	47	41	**Karnin**	13°51	53°50	60
19	**Althagen**	12°25	54°22	62	42	Greifswalder Oie	13°54	54°14	32
20	Zingst	12°41	54°27	16	43	Koserow	14°00	54°04	43
21	**Zingst Bodden**	12°41	54°26	40	44	**Ueckermuende**	14°04	53°45	50
22	**Barth**	12°43	54°22	56	45	**Kamminke**	14°13	53°52	14
23	**Barhöft**	13°02	54°26	40					

smaller hard structures covering approximately 9%. The remaining coast is unprotected. This is in stark comparison to the German North Sea coast where 85% is dike-protected (Sterr, 2008). The disparity between coastal protection levels is an indicator of ESL rarity along the German Baltic Sea coast.

5.2.2 *Data*

In this study, we drive the stochastic model using high-resolution water level data measured at tide-gauges along the German Baltic Sea coast. We use a subsample of 45 hourly tide-gauge records provided by Schmidt et al. (2017), who compiled observations of total still water levels for the entire Baltic Sea. These records vary substantially in length, with the shortest record being approximately 14 years at Kamminke, and the longest is 66 years at Travemünde. Each data set is referenced to the German ordnance datum Normalhöhennull. A list of tide gauges is given in Table 5.1 with locations shown in Figure 5.1.

Assessing ESLs in the Baltic Sea is made difficult by the rarity of such events combined with the relatively short sea level records available.

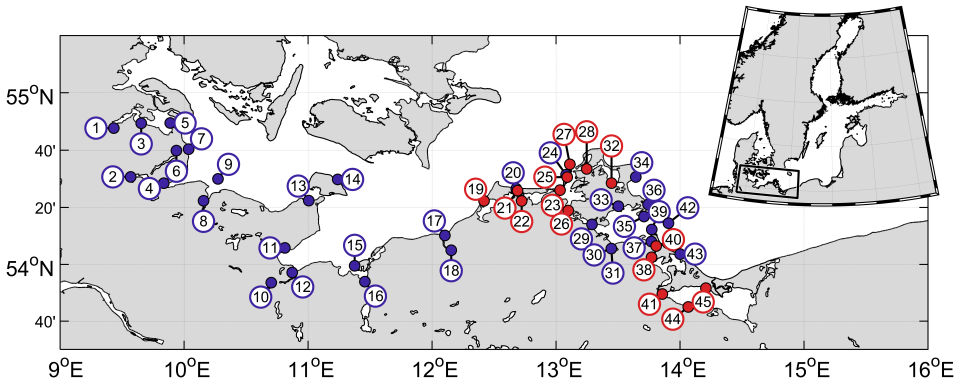

Figure 5.1: Study area in the Baltic Sea and locations of the 45 tide-gauge records. Tide-gauges located on the open coast are shown as blue with bodden tide-gauges shown as red.

For example, hourly sea level data are available for the tide-gauge of Travemünde from November 1949, where the maximum water level above MSL is 2.05 m. However, several historic storms have resulted in ESLs in Travemünde and nearby Lübeck, which greatly exceed this value (Jensen and Töppe, 1990). Using hourly tide-gauge data, Gräwe and Burchard (2012) estimate a return period greater than 100,000 years for the ESL event of 1872. They therefore conclude that a reliable estimate of return periods for such events cannot be computed using present data. Thus, to provide more robust estimates of return water levels along the German Baltic Sea coast, more digitization efforts at sites with long tide-gauge records are required.

5.3 METHODS

5.3.1 *Overview*

This section outlines the methodology used to stochastically simulate ESLs. We first define several relevant terms using Figure 5.2, which shows an ESL hydrograph extracted from the Timmendorf tide-gauge data. These hydrographs describe the temporal evolution of the water level curve. Section 3.1 describes the method used to identify these events within the observational records. To model the observed hydrographs, each event is parameterized in order to define all events by a set of common parameters. Here we define the general shape of each event using seven parameters, which describe the shape of the water level curves to the left and right of the peak water level, hereafter referred to as the flow and ebb curves (Figure 5.2a). To simulate variations in water level along these curves, resid-

ual water levels (Figure 5.2b) are modeled using an autoregressive process, requiring a further 10 parameters. Thus, the final parameterization scheme consists of 17 parameters and is described in detail in section 3.2. Finally, section 3.3 describes the generation of artificial ESL hydrographs using Monte Carlo simulations based on the distributions of extracted parameters.

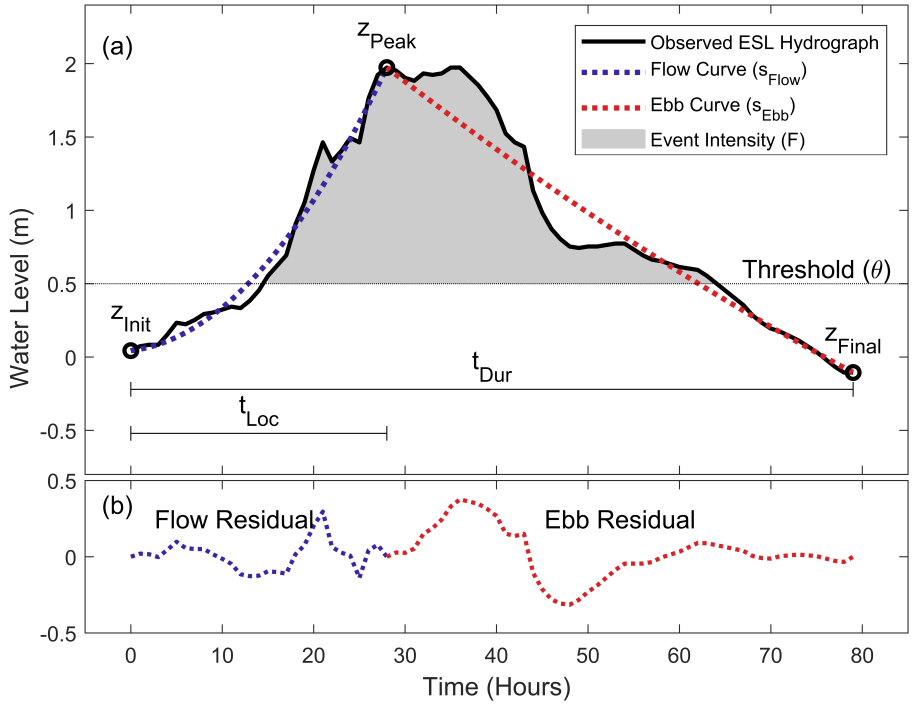

Figure 5.2: (a) An extreme sea level hydrograph (black line) extracted from the Timmendorf tide-gauge record. Parameters 1–7 (zInit, zPeak, zFinal, tDur, tLoc, sFlow, and sEbb) describe the basic shape of the hydrograph in terms of the flow and ebb curves (blue and red dashed lines). The shaded area (F) represents event intensity. (b) Residual water levels along the flow and ebb curves equal to the difference between the observed event and the basic shape described by parameters 1–7. These residuals are modeled using an autoregressive process.

All artificial events are produced as hydrographs, which contain information on the entire water level curve in addition to the peak water value. This is advantageous for flood risk analyses as it has been shown that the temporal variability of the ESL time series can account for large uncertainties in the impact of flood events (Quinn et al., 2013; Santamaria-Aguilar et al., 2017). To quantify this variability, the intensity of an ESL hydrograph is defined as the area between the water level curve and some threshold (θ; Figure 5.2a), with equation:

$$F = \int_{t1}^{t2} z(t) - \theta dt \tag{5.1}$$

where F is intensity, z(t) describes the ESL water level time series, and t1 and t2 mark where z(t) is greater than the threshold θ (Wahl et al., 2011). In this paper, intensity is calculated using θ equal to mean high water (MHW) and is simply used as a diagnostic.

5.3.2 *Identifying Extreme Events*

The first step in generating artificial ESL hydrographs requires the identification of such events within the observational records and is done using EVA based on peak water levels. To do this, the water level time series are first made stationary by removing any long-term trends. These trends are identified using a 1-year moving average of high waters as recommended by Arns et al. (2013). The high-water trend is preferred as it includes changes in MSL as well as trends in tidal range and storminess. Next, a Peaks-over-Thresholds (POTs) method is used to identify extreme events within the detrended water level data. We use this method over the block maxima method as it better represents the tails of the water level distribution, thus providing more reliable estimates of ESLs, particularly if records are short (Arns et al., 2013; Wahl et al., 2017). Furthermore, the number of events extracted is typically higher than common block maxima intervals (annual maxima) when using POT. This provides a larger number of hydrographs for the stochastic model input, thus contributing to a more robust representation of local ESL characteristics in the artificial ESL hydrographs.

Care must be taken to ensure an appropriate threshold is selected as this may bias the statistical assessment at the very beginning. For each observational record, a threshold is selected using the stability threshold method (STM) and mean residual life (MRL) plot, as defined by Coles (2001). A description of this method is given in Appendix A including a sensitivity analysis, which assesses the stability of return water level estimates under various threshold selection techniques.

For each observational record, ESLs are identified as those peak water levels that exceed the selected thresholds. As each event is extracted as a hydrograph, declustering of the tide-gauge data is not required. However, identifying the start and end position of these hydrographs is challenging. This is due to the microtidal environment of the Baltic Sea, which lacks a regular and clear variation in water level. Hydrographs identified in Wahl et al. (2011) were bounded by low waters, which are distinct along the German North Sea coast due to the significant influence of tides. For the microtidal German Baltic Sea coast, where water levels may not fall below MHW for extended

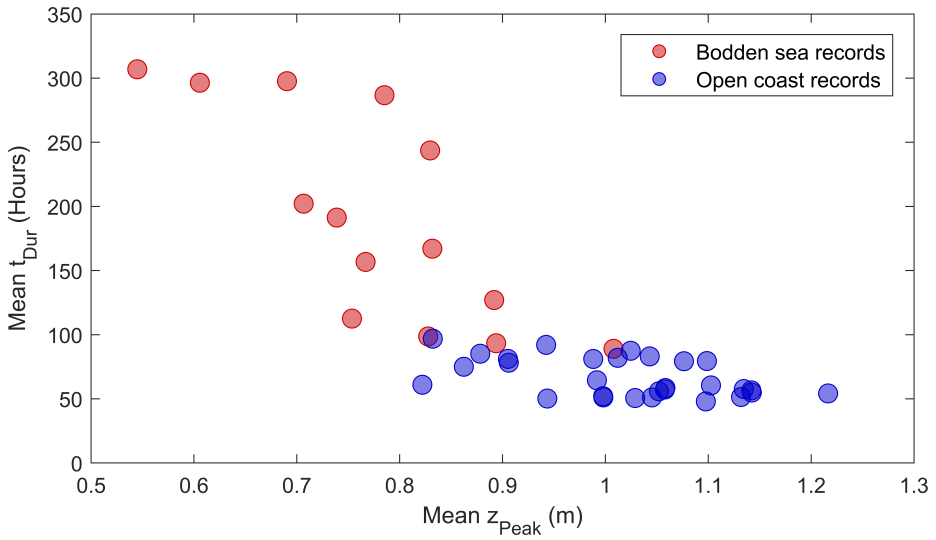

Figure 5.3: Scatter plot of hydrograph mean durations and peak water levels for each tide-gauge record. Tide-gauges located within bodden seas are shown as red with other records shown as blue. Typically, hydrographs extracted from bodden tide-gauges have longer durations with lower peak water values due to the restricted flow of water to the open Baltic Sea.

periods, we developed an iterative process, which tests the low waters surrounding the peak water levels of each hydrograph. Working along the flow and ebb curves, outward from the peak water level, each trough in the water level time series is tested until start and end points are found. A tested trough is selected as the start or end point of the hydrograph if it lies below MHW or is the lowest point of five consecutive tested troughs, with at least two troughs exceeding it on the outer side. The second point is required to ensure hydrograph durations do not become excessive. However, this is unavoidable for some tide-gauge records, specifically those located within bodden seas. Here restricted flow to the open Baltic Sea results in more stable water levels. Hence, ESL events in these areas are typically smaller with longer durations. This can be seen in Figure 5.3, which shows mean durations and peak water levels of ESLs extracted from all tide-gauge records.

5.3.3 *Parameterization*

Parameterization of the hydrographs identified in section 3.1 is undertaken to characterize each event. This provides a basis for all hydrographs to be described by a set of common parameters. The parameterization scheme outlined in Wahl et al. (2011) relies on a uniform hydrograph shape due to a strong tidal influence. Again,

Table 5.2: List of Stochastic Model Parameters

Parameter	Description
z_{Init}	Initial water level (m)
z_{Peak}	Peak water level (m)
z_{Final}	Final water level (m)
t_{Dur}	Event duration (hours)
t_{Loc}	Peak location (factor of t_{Dur})
s_{Flow}	Flow curve shape
s_{Ebb}	Ebb curve shape
ϕ_{8-11}	Flow curve residuals, AR coefficients
ϕ_{12-15}	Ebb curve residuals, AR coefficients
ϕ_{16-17}	Variance estimates for flow and Ebb residuals

Note. A schematic of these parameters is shown in Figure 5.2. AR, autoregressive.

this is not possible for the German Baltic Sea given the small tidal signal. Therefore, we developed a new parameterization scheme consisting of 17 parameters listed in Table 5.2. A schematic of the first seven parameters, which describe the basic hydrograph shape can be seen in Figure 5.2a. A detailed description of each parameter is provided in the following.

To identify an appropriate parameterization scheme, a comparison of all hydrographs extracted from two tide-gauge records was conducted to highlight any distinguishing ESL features. Figure 5.4 shows all ESL hydrographs extracted from the Timmendorf and Althagen tide-gauge records. The Althagen record is included as it contains the longest duration events and lowest peak water levels due to its location in the Saaler Bodden. By normalizing the duration of the flow and ebb curves of each event, it can be seen that the mean hydrograph shape along the open coast is triangular. The parameterization scheme is therefore developed based on this general feature. However, the nonlinear shape of the flow and ebb curves seen in bodden seas (Figure 5.4b) must also be accounted for.

First, each event is outlined by its initial, peak, and final water levels. Their locations in time are defined by the first two parameters, which describe the total event duration in hours (tDur) and location of the peak water level as a factor of the duration (tLoc). The values of these water levels are defined in meters by zInit, zPeak, and zFinal. Using these five parameters, the general triangular shape of an ESL event can be constructed (see Figure 5.2a). Next, as the flow and ebb

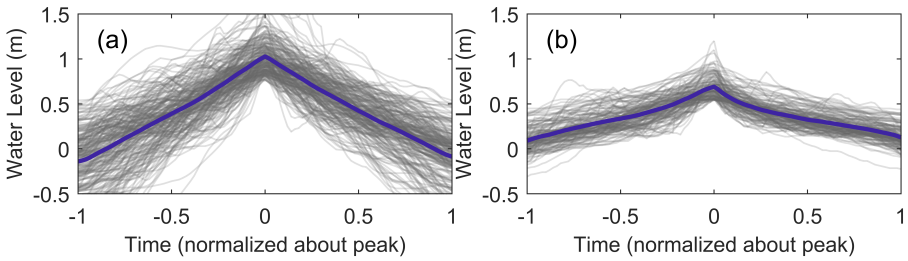

Figure 5.4: Observed extreme sea level hydrographs extracted from the (a) Timmendorf and (b) Althagen tide-gauge records. Both plots show the ensemble average (blue line) of all observed extreme sea level events on record (gray line), with time normalized about the peak (water level is not normalized). The different shapes are due to the location of the tide-gauges, where Timmendorf is located along the open coast and Althagen is located within the Saaler Bodden.

curves are not necessarily linear, as seen clearly within bodden seas (Figure 5.4b), parameters sFlow and sEbb are introduced to describe their shape. We use a simple second-degree polynomial to define this shape, with equation:

$$z'\left(t'\right) = s \cdot t^2 + (1 - s) \cdot t' \tag{5.2}$$

where z' and t' are nondimensional forms of water level and time, respectively, and s is the parameter value. sFlow and sEbb are calculated using a linear least squares regression where time and water level are normalized (nondimensional). The basic shape of the hydrograph is thus constructed by stretching the flow and ebb water levels, as described by $z'(t')$, to the points defined by the first five parameters. Equality constraints are applied when calculating flow and ebb shape parameters, which restrict water levels at $z'(0) = 0$ and $z'(1) = 1$. Parameter values are also constrained to ensure water levels remain between $z' = [0,1]$. Therefore, when the flow and ebb curves are stretched to the positions defined by the first five parameters, water levels do not exceed zPeak, nor do they fall below zInit or zFinal on the flow and ebb faces, respectively. If this occurs, the reconstructed ESL hydrograph will not adhere to the shape specified by the first five parameters and is thus nonrepresentative of the observed event.

Lastly, residual water levels between the basic hydrograph shape described by parameters 1–7 and the hydrograph extracted from the observational data are accounted for. This is done to provide the model a mechanism in which water level deviations can be generated, such as secondary peak events and small tidal signatures, which are present in the observational data. However, considerable variability exists in the amplitude and behavior of residual water

levels between events. Furthermore, the behavior of the residual water level is independent of event duration, thus ruling out the use of further parameters, which describe a water level at specific points along the flow and ebb curves, as these are inherently connected to event duration. Therefore, a procedure is required which is able to generate a range of residual water levels, where their behavior is not influenced by the duration of the event. We consider the use of an autoregressive (AR) model to generate these residual water levels.

An AR(n) model can be described by the following equation:

$$z_{\mathrm{Res}(t)} = \sum_{i=1}^{n} \varphi_i \cdot z_{\mathrm{Res}(t-i)} + \varepsilon_t \tag{5.3}$$

where zRes(t) is the residual water level at time t, ϕi, ..., ϕn are AR model coefficients, ε is white noise, and n represents the model order. Thus, the AR process describes a time-varying signal, where each value is equal to a weighted sum of the past n values and a random white noise element. Consequently, an artificial residual water level time series may be generated using an AR model with coefficients estimated from an observed event and random white noise as input. AR models have been used previously to model water level (Irvine and Eberhardt, 1992; Vaziri, 1997), and while they have been somewhat superseded by other techniques such as AR neural networks (Seo et al., 2015), which more accurately model nonlinear systems, they are ideal for our requirements. Namely, they can produce signals of varying length, where the signal behaviour is independent of length, and they are easily applied with low computational cost.

Finally, we integrate an AR model of order 4 into the stochastic model framework. Here 10 parameters are introduced to account for the residual water levels along the flow and ebb curves. These are four AR(4) coefficients (ϕ) for both the flow ($\phi_{8\tilde{}11}$) and ebb ($\phi\tilde{}12\tilde{}15$) curve residual water levels and the variance of a white noise signal (ε) which is used as AR model input (flow curve: ϕ_{16}; ebb curve: ϕ_{17}). These parameters are estimated from the observed flow and ebb curve residuals using the Yule-Walker equations by means of Levinson-Durbin recursion (Hayes, 1996). An AR model of order 4 was chosen based on sensitivity analyses, which considered the time dependence of the residual water level signal. Reflection coefficients, calculated for an AR(12) model fitted to the data, decay significantly after a lag of 4, indicating that an AR(4) model is sufficient in describing the residual signal.

Several reconstructions of the largest ESL event on record at the Timmendorf and Althagen tide-gauges were made using this parameter scheme and are shown in Figure 5.5. Exact reconstructions are not

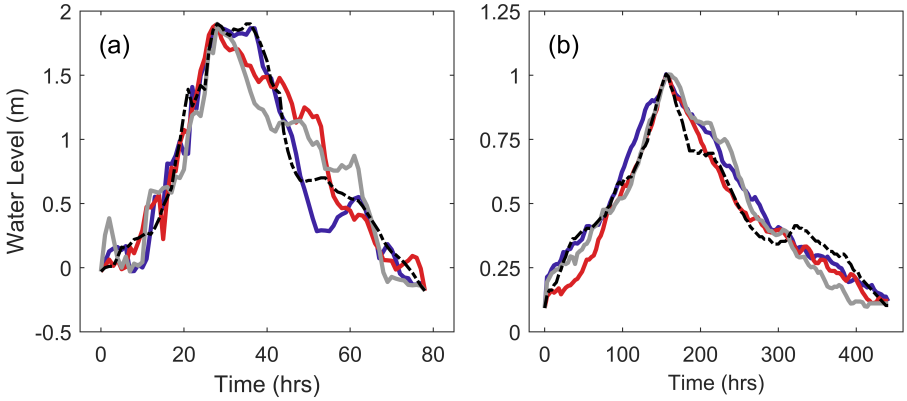

Figure 5.5: Reconstructions of the largest extreme sea level events on record at the (a) Timmendorf and (b) Althagen tide-gauges using the parameter scheme described in section 3.2. The observed event (black dashed line) is compared to three reconstructions (red, blue, and gray). Reconstructions differ from the original extreme sea level time series due to the stochastic process of the AR model, which generates residual water levels along the flow and ebb curves.

possible as the AR model used to generate residual water levels is a stochastic process, driven by random white noise. Despite this, the general structure of each event is preserved.

5.3.4 *Monte Carlo Simulations*

To construct a large number of stochastic ESL hydrographs, artificial parameter sets are generated by means of Monte Carlo simulations. This involves random sampling from parametric distribution functions fitted to the observed parameter data. The use of parametric distribution functions allows for the generation of artificial parameters, which lie outside the observed data; hence, the model is able to produce ESL events larger than those observed. To ensure that the artificial hydrographs are realistic, it is vital that the interdependencies between observed parameters are preserved. Thus, to maintain this dependence structure, sampling locations are taken from a uniform multivariate distribution of the observed parameters.

The multivariate cumulative distribution function of a d-dimensional data set may be expressed in terms of its marginals and a copula (C, Sklar, 1959). From a modeling perspective, the copula may be used to generate random values from a multivariate distribution over the unit hypercube [0,1]d. For our purposes, these random uniform values represent sets of sampling points that maintain the

dependence structure of the underlying data. (Li et al., 2014) consider the use of Archimedean and elliptical copulas when modeling a number of storm parameters for the Dutch coast. While lacking the capability to model tail-dependence, elliptical copulas such as the Gaussian and t-Student copulas are more easily applied to multidimensional data sets (Wahl et al., 2016). Furthermore, the performance of Archimedean copulas decreases in higher dimensions (Díaz, 2014). Using a goodness-of-fit (GOF) test, Li et al. (2014) concluded that the elliptical Gaussian copula was most suited for their study. Further, Wahl et al. (2016) examined the use of the Gaussian and t-Student elliptical copulas in modeling six hydrodynamic parameters in the northern Gulf of Mexico. A comparison of the rank correlation matrix and nonparametric tail-dependence coefficients derived from observations and simulations showed that the t-Student copula performed slightly better than the Gaussian copula. In this study, we consider the use of both the Gaussian and t-Student elliptical copulas. A comparison of both copulas is conducted later in the study, where we compare observed and simulated parameter values to identify the best suited copula for our requirements.

For a given linear correlation matrix ρ from the uniform d-dimensional data u, an elliptical copula has the form:

$$C_\rho(u) = F_\rho\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right) \tag{5.4}$$

where Fρ is the multivariate distribution function and F_j^{-1} are the inverse marginal distribution functions of the parameters uj for j = 1, ..., d. In the place of F, Gaussian and t-Student copulas are constructed by using the standard normal and the t-Student distributions, respectively. The linear correlation matrix ρ is calculated using the following:

$$\rho = \sin\left(\frac{\pi}{2} \cdot \tau\right) \tag{5.5}$$

where τ is Kendall's rank correlation, which represents a nonparametric measure of dependence. For parameter pair (x,y) with sample size N, τ is calculated using

$$\tau = \begin{pmatrix} N \\ 2 \end{pmatrix}^{-1} \sum_{j=1}^{N} \sum_{i=1}^{N} \text{sign}\left[\left(x_i - x_j\right)\left(y_i - y_j\right)\right] \tag{5.6}$$

where $sign = 1$ if $[(xi - xj).(yi - yj)] > 0$ or $sign = -1if[(xi - xj).(yi - yj)] < 0$ (Karmakar and Simonovic, 2009). Rank correlation is preferred here over linear correlation as it is able to capture nonlinear relationships that may exist between parameter pairs.

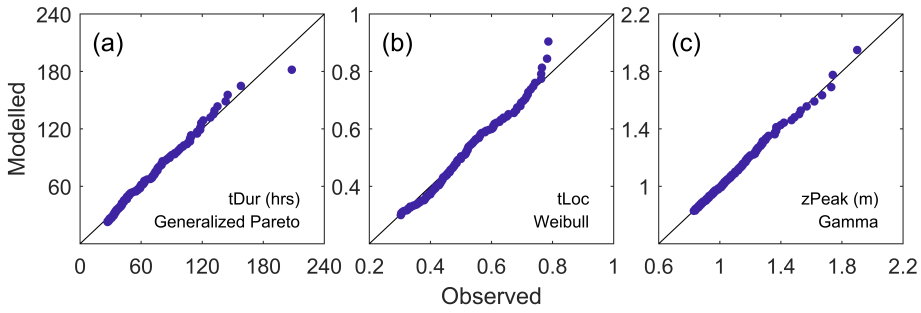

Figure 5.6: Observed versus theoretical values for parameters (a) tDur, (b) tLoc, and zPeak (c) zPeak at the Timmendorf tide-gauge record. Parameter units and best fit distributions are given.

Once a copula has been fitted to the observed data, large numbers of random values may be generated. These uniform values belong to the cumulative probability distribution defined by the copula and maintain the dependence structure of the underlying data. To transfer these uniform values into artificial parameter sets, inverse cumulative distribution functions are used. As the dependence structure expressed within the copula is separate from the marginal distributions of the individual parameters, greater flexibility is gained when fitting these distribution functions to the observed data. Hence, we consider a number of distribution functions commonly used in coastal hydrology; these are Generalized Extreme Value (GEV), Generalized Pareto (GP), Weibull, normal, lognormal, gamma, Rayleigh, inverse Gaussian, logistic, log-logistic, beta, and exponential. During each simulation, a best fit distribution is identified for each parameter by comparing the empirical data with the theoretical values corresponding to each distribution. Best fit is determined by calculating the Bayesian information criteria (BIC) for each distribution fitted to the empirical data. Figure 5.6 shows three separate distributions fitted to the parameters tDur, tLoc, and zPeak measured at the Timmendorf tide-gauge.

Finally, to identify the most appropriate copula for our purposes, parameters extracted from the observed ESL events of all tide-gauge records were compared to parameters simulated using the Gaussian and t-Student copulas. GOF for parameters 1–7 was measured using statistical energy (Aslan and Zech, 2005) due to its good performance in multidimensional applications. Further, the maximum distance between observed and simulated cumulative distributions was measured using the two-sample two-dimensional Kolmogorov-Smirnov(KS) test (Peacock, 1983) for each parameter pair. We found that the Gaussian copula outperforms the t-Student copula in both tests albeit only slightly. Statistical energy measured between observed and artificial parameters is lower for the Gaussian copula

compared to the t-Student copula (mean values of 0.0036 and 0.0044, respectively, for all tide-gauges). Kolmogorov-Smirnov (KS) tests return similar results, where maximum distances between the bivariate cumulative probability distributions are small at all tide-gauges (<0.1), for all parameter pairs of both copulas. While differences between copulas are small, the Gaussian copula performs better in both of these measures at 33 of the 45 tide-gauge records. Accordingly, we use a Gaussian copula to model the parameter dependence. Figure 5.7 shows a comparison between observed and simulated parameters using a Gaussian copula for the Timmendorf tide-gauge record. In several cases there exists a shift in the mode of certain parameters. This is a result of the stochastic process and does not necessarily occur in each simulation.

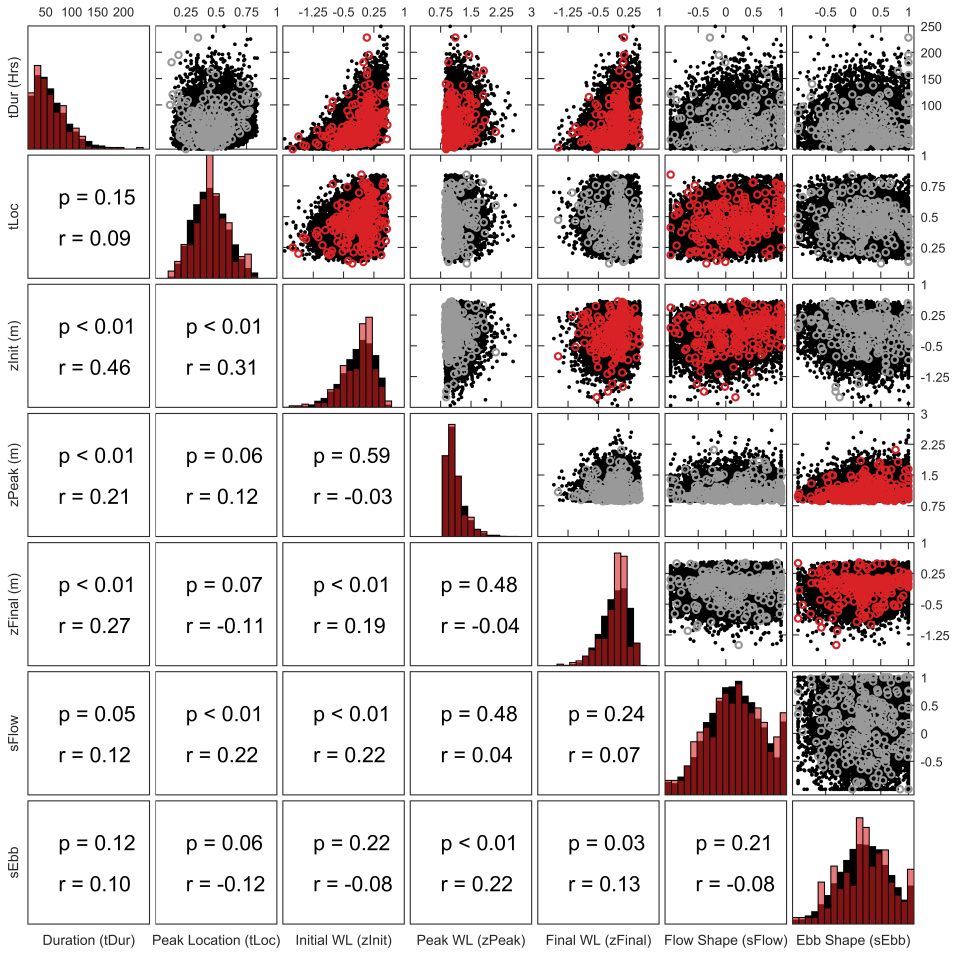

Figure 5.7: Kendall's rank correlation for parameters 1–7 of the extreme sea level hydrographs extracted from the Timmendorf tide-gauge record. Where the correlation is significant (95% confidence), observed parameter values are shown as red, otherwise, they are shown as gray. Parameters modeled using a Gaussian copula are shown as black.

Generating all parameters by means of Monte Carlo simulations leads to an issue where the residual water level along the flow and ebb curves may become divergent, thus resulting in highly deformed ESL hydrographs. This is a result of the AR process, which is highly sensitive to its coefficients, represented by parameters 8–17 in the stochastic model. For this reason, randomly generated parameters 8–17 occasionally result in divergent residual water levels. To avoid this issue, only parameters 1 to 7 are generated using Monte Carlo simulations. Suitable values for the AR coefficients are extracted from the observational data, thereby ensuring residual water levels are nondivergent. Hence, AR coefficients for both the flow and ebb curves of an artificial ESL event are taken from an observed event where the duration of the respective curves most closely match. As duration values are selected randomly, the selection of the AR coefficients remains a stochastic process. A comparison of durations was chosen as there is a significant correlation between the length of the ebb and flow curves and the AR coefficients. This is in accordance with observations where longer durations are likely to see a more variable flow and ebb curve.

Finally, simulated parameter sets are assembled into artificial ESL hydrographs. Care must be taken to ensure all artificial events adhere to the general ESL shape specified by their parameters 1–5, as the stochastic process of the AR model may generate residual water levels, which significantly deform this shape (i.e., water levels along the flow and ebb curves may exceed the peak water level [zPeak] or drop below the start and end water levels [zInit, zFinal]). To ensure the modeled events adhere to the observed statistics, such deformed events must be discarded. However, as the residual water levels are randomly generated, the process is repeated until the assembled event fits within the parameter bounds.

5.4 RESULTS AND DISCUSSION

5.4.1 *Model Validation*

For each observational data set, 10,000 artificial ESL hydrographs have been generated. Differences between the artificial and observational ESL data sets were quantified using root-mean-square error (RMSE) between simulated and theoretical peak water level and intensity values. Theoretical values are derived from best fit distributions

fitted to the observational data and evaluated at locations calculated using Gringorten plotting positions:

$$R = \frac{i - 0.44}{N + 0.12} \tag{5.7}$$

where R is the exceedance probability, i is rank from largest to smallest, and N is the total number of observations (Gringorten, 1963). For all locations, RMSE between simulated and theoretical peak water levels is small (<2 cm). This is to be expected, as simulated peak water levels are derived directly from the parametric distribution functions fitted to the observational data. RMSE between simulated and theoretical intensities provide a better measure of model performance, as intensity is dependent on the entire ESL hydrograph, and thus all model parameters. A comparison of simulated and observed intensities is shown in Figure 5.8 at eight different tide-gauge locations. Each plot shows event intensities from observed (red dots) and simulated (blue dots) ESL events. Theoretical return water levels are also shown (black line, 95% confidence bounds included) and are calculated using a best fit distribution fitted to observed intensities.

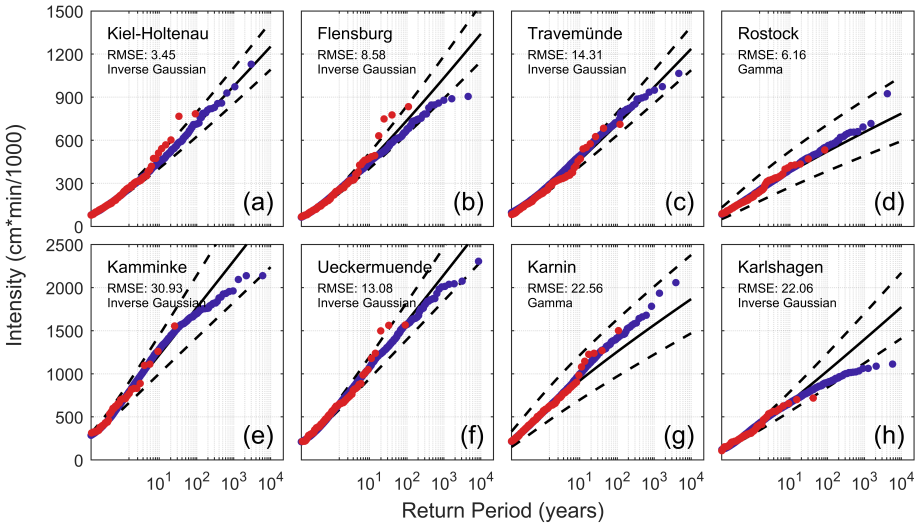

Figure 5.8: Comparison of observed (red) and simulated (blue) extreme sea level intensities at eight tide-gauge locations. Included are tide-gauges from open coastlines (a-d) and bodden seas (e-h). Best fit distributions are shown (black lines) with 95% confidence bounds and listed beneath the values for RMSE between simulated and theoretical intensities.

Overall, there is a good fit between observed and artificial intensities at all tide-gauges. The largest errors occur at locations along the bodden coast, where ESL durations can be long due to the restricted water flow caused by local bathymetry. As event durations increase, the accuracy of the model to predict water level changes over this

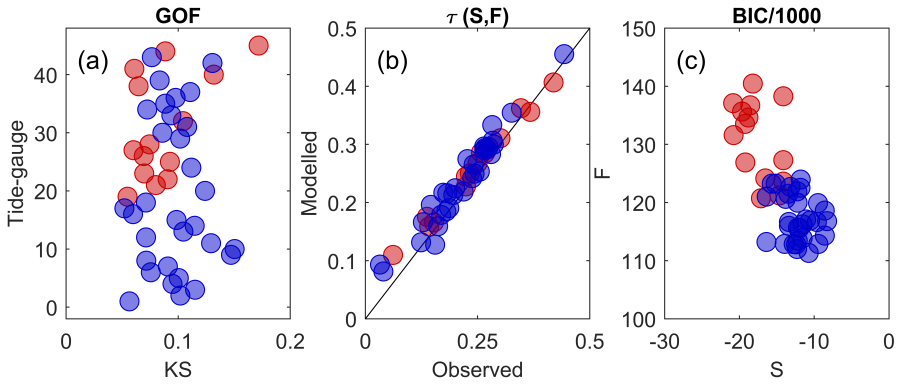

Figure 5.9: Model performance based on the parameters of peak water level (zPeak) and intensity (F) for all tide-gauge locations. The blue dots represent tide-gauges along the open coast, and the red dots represent tide-gauges in bodden seas. (a) Goodness-of-fit of the modeled zPeak and F values using the two-sample, two-dimensional KS test. (b) Q-Q plot of rank correlation between zPeak and F for all observed and modeled events. (c) Scatter plot of Bayesian Information Criterion (BIC) corresponding to the log likelihood of modeled zPeak and F. A lower BIC indicates a better goodness-of-fit.

period decreases. Despite this, model predictions of ESL intensities remain relatively accurate and within the 95% confidence bounds for most tide-gauges, below a return period of 10,000 years. However, a good fit between observed and artificial intensities by itself does not necessarily suggest that simulations reflect real-life conditions, as the dependence between peak water level and intensity may not be captured. Therefore, we calculate the bivariate GOF using the two-sample, two-dimensional KS test (Figure 5.9a) and compare the rank correlation between peak water level and intensity values of observed and modeled ESLs (Figure 5.9b). We also consider the GOF of peak water level and intensity values individually using BIC (Figure 5.9c).

For all locations, the bivariate distribution of artificial peak water levels and intensities fit well to observed values, with a mean KS statistic of 0.091. Rank correlation between artificial peak water level and intensity values also matches well to observational data with no noticeable differences between locations along the open coast or within bodden seas. However, when comparing GOF of the individual parameters of peak water level and intensity, there is a clear difference between open coast and bodden tide-gauges. Figure 5.9c shows that the model reproduces peak water levels more accurately in bodden locations and intensity values more accurately along the open coast.

While the calculation of ESL intensities is somewhat time-consuming for the entire data set, as it requires a full ESL time series, we find that the area defined by parameters zPeak and tDur are a good approximation for intensity (F), such that

$$F \approx \frac{1}{2} \cdot z_{\text{Peak}} \cdot t_{\text{Dur}} \tag{5.8}$$

However, this method overestimates intensity at open coast tide-gauges by approximately 10% and at bodden records by approximately 15% when compared to our method (Equation 5.1). These figures are a good proxy for the impact of the five remaining parameters (tLoc, zInit, zFinal, sFlow, and sEbb), excluding those which describe the residual water levels along the flow and ebb curves. As the AR process, which models these residuals, is stochastic, their influence on this figure when considering large numbers of events is negligible. The value of the AR model however, is in producing individual events with realistic water level curves.

5.4.2 *Limitations*

Artificial ESL hydrographs generated by the model fit well to observational data; however, care must be taken when considering observational records with short lengths. As mentioned earlier, a number of the available tide-gauge records are relatively short (17 tide-gauges with less than 30 years of data). While there are several studies, which consider the effect of short records on estimates of return water levels (Arns et al., 2013; Haigh et al., 2010), we consider a number of storm surge parameters, which have not been analyzed before. As such, it is not known how the extreme values of these parameters should be treated, given the length of records available. In addition, the German Baltic coastline is a very specific location given the generation of extreme events, which are comparable to the situation in tropical-cyclone prone areas (e.g. Haigh et al., 2014a; Lin and Emanuel, 2016)(e.g., Haigh, MacPherson, et al. (2014); Lin and Emanuel, 2016), and as such, are very rare. For instance, individual extreme events of very large magnitude require prefilling of the entire basin (typically driven by westerly winds over the North Sea) followed by strong winds from the northeast (Jensen and Töppe, 1990). Such situations have not been observed very often in recent decades but are visible in records extending back into the 19th century. We therefore tested the model's ability to produce artificial ESL hydrographs given a range of record lengths.

For all tide-gauge records with more than 50 years of data, 10,000 artificial ESLs were generated using subsamples of the available data

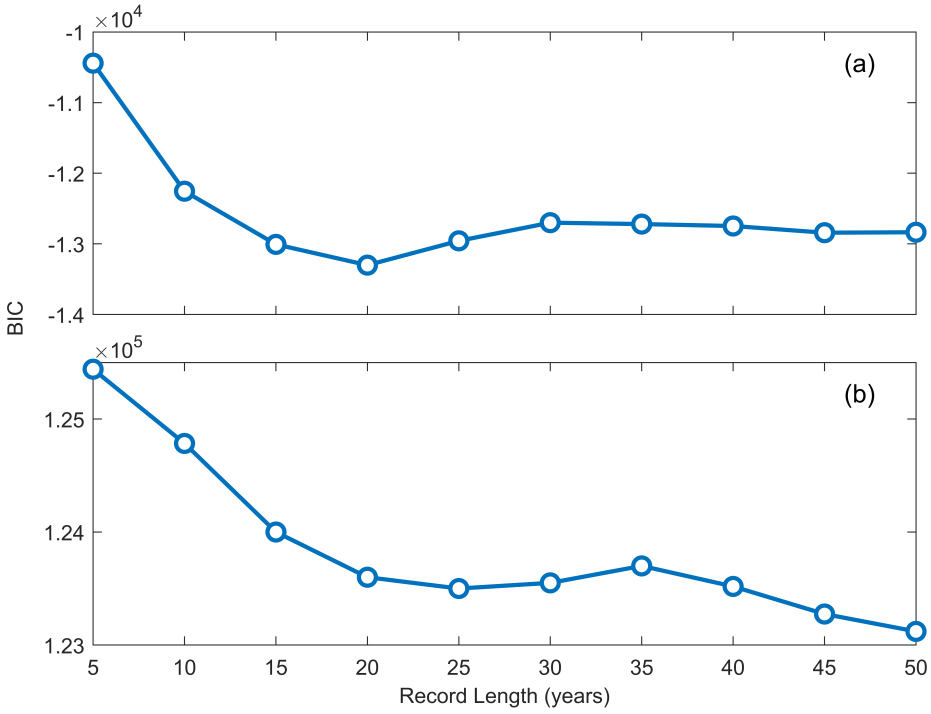

Figure 5.10: Comparison of model performance based on the length of input data. Both plots show goodness-of-fit between artificial and observed extreme sea level calculated using Bayesian Information Criterion (BIC). (a) Peak water level and (b) intensity are shown. A lower BIC indicates a better goodness-of-fit.

as input. These subsamples ranged from 5 to 50 years in length with 5-year intervals. On average, each interval represents approximately 23 events. At each tide-gauge location, the resulting ESLs were compared to the complete ESL observational record to determine the difference in model performance given the length of input data. GOF for each simulation was measured using BIC. Figure 5.10 shows mean BIC values at each tested record length for all tested tide-gauge records. Interestingly, the best fit between modeled and observed peak water levels occurs when using 20 years of data. However, GOF does not change considerably for record lengths of 15 years and longer. This is also true for ESL intensity, suggesting that 15 years of data, or approximately 70 events, is sufficient to simulate events with a return period of 50 years.

5.4.3 *Accounting for Additional Information*

For a number of locations along the German Baltic Sea coast, large historical ESLs have been recorded before the introduction of high-resolution measurements. Figure 5.11 shows the peak water levels

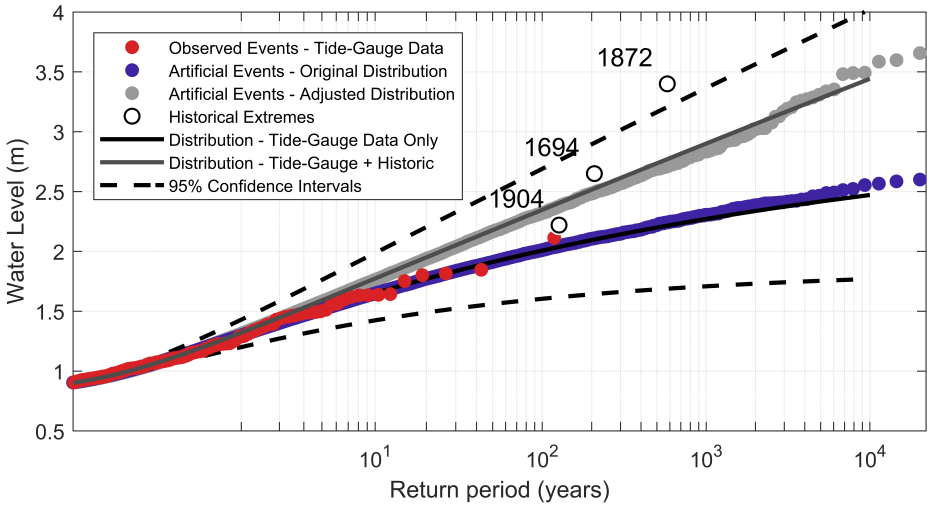

Figure 5.11: Return water level estimates at Travemünde. The peak water level of events recorded in the high-resolution tide-gauge data is shown as red dots. Three large historical events (black circles) are also shown. The model was used to generate artificial events according to the distribution of high-resolution data (blue dots) and an adjusted distribution, which accounts for the historical events (gray dots). Distributions are shown as black lines with the 95% confidence bounds of the high-resolution data also shown (dashed black lines).

of all ESL events taken from the high-resolution tide-gauge record of Travemünde and all artificial events produced by the model at this site. Three historical events are also shown, where the water level reached 2.65 m in 1694, 3.4 m in 1872, and 2.22 m in 1904 (Jensen and Töppe, 1990). The return period of these events was estimated using Gringorten plotting positions (Equation 5.7), where they were ranked as the three highest events, and the number of observations (N) was taken as the period in years between the first event and the present day (324). Despite generating events with peak water levels that follow the distribution of observed events well, the model is unable to produce an event where the peak water level nears that of the 1872 event. This is due to the bounded distribution of peak water levels, which does not consider historical events as they are not present within the high-resolution tide-gauge data.

While the model is limited by the availability of high-resolution water level data, including historical ESLs into the stochastic process may result in a more robust simulation of artificial ESL events. Due to the model's dependence on parametric distribution functions fitted to observed ESL parameters, adjusting the distribution of peak water levels to better fit historical data will allow for the generation of larger ESL events, more in-line with historical measurements. This is

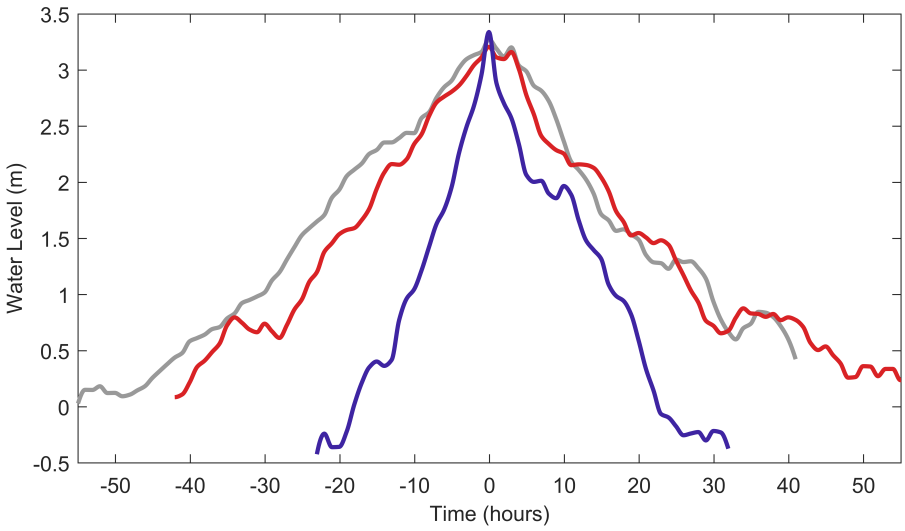

Figure 5.12: Three reconstructions of the extreme sea level event of 1872. To produce such an extreme event, the stochastic model requires an adjustment to the parametric distribution function fitted to the observed peak water levels. This can be seen in Figure 5.11.

shown in Figure 5.11, where the distribution of zPeak was adjusted through the inclusion of the three historical events mentioned earlier. The parametric distribution fitted to zPeak is adjusted to maximize the likelihood function, which is simply the product of likelihood functions calculated for each complete set of observations, here that is the high-resolution tide-gauge data and the historical records. While this is not necessarily the most robust method of accounting for historical events (see Frau et al., 2018), it is useful as a demonstration. Using this adjusted distribution, three artificial ESLs produced by the model with peak water levels approximately equal to the 1872 event are shown in Figure 5.12.

One issue with the inclusion of historical events is that typically only peak water level is recorded. Hence, only the distribution of peak water levels is affected by this new information. While the model relies heavily on an accurate assessment of parameter interdependencies, there is concern that changing one parameter distribution would affect the accuracy of the model to simulate ESL behavior. However, as the dependence structure between ESL parameters is modeled before any parametric distributions are fitted to the data, parameter dependencies observed from the high-resolution records remain unchanged. That being said, care must be taken to ensure any adjusted distribution functions continue to fit the observed high-resolution events well, in addition to any included historical extremes. Interestingly, future sea levels may also be considered using the same method. Here the distribution functions fitted to the water level parameters (zInit, zPeak, and zFinal) may be shifted

without influencing the other parameters. As such, larger ESLs may be generated where the temporal structure remains the same. This however assumes that dependencies between ESL parameters do not change under different sea level conditions.

5.4.4 *Application*

Artificial ESL events produced by the stochastic model may be used for a wide range of applications, most evidently, for use in statistical analyses and contributing to the approximation of risk in coastal flooding scenarios. As all artificial events are generated as hydrographs, analyses may consider several ESL parameters rather than just peak water level. As an example, a bivariate analysis of artificial ESLs is described in Appendix B. Here ESLs are classified based on their combined peak water level and intensity values. Furthermore, artificial ESL hydrographs may be used as input for flood inundation models, where large numbers of ESL events are useful when quantifying uncertainties in flood extent. As the model is easily applied and computationally inexpensive, it is able to produce significantly larger data sets of ESL events than what is possible using empirical or numerical modeling methods. While the simulation of all artificially generated events in flood inundation models is not possible due to the computation costs of such models, the stochastic ESL model remains useful as it is capable of producing large numbers of events for a certain design height to be tested. Therefore, only a select number of artificial events are needed to model specific ESL heights and quantify uncertainty.

In comparison to the stochastic model developed by Wahl et al. (2011), the framework described in this paper is well adapted for use in microtidal environments, due to the new parameterization scheme. While developed specifically for the German Baltic Sea coast, the model performs well under the various conditions experienced at bodden and open coast locations. As such, we believe that the model is well suited to most microtidal locations and is not restricted to sites within the Baltic Sea.

5.5 CONCLUSIONS

We describe a method to stochastically simulate ESL events in microtidal environments for use in flood risk analyses. The model is capable of generating artificial ESL hydrographs that represent ob-

served events along the German Baltic Sea coast well; however, the proposed approach is applicable to other microtidal locations and requires only high-resolution water level data. The ability of the model to generate many thousands of ESL hydrographs while remaining computationally inexpensive is useful for flood risk analyses, which require large numbers of events to accurately assess risk associated with certain return water levels and quantify uncertainty surrounding the temporal variability of the ESL time series. Furthermore, by identifying the full extent of each ESL hydrograph, the assumption of independent peaks is justified. As such, the application of EVA becomes more valid through our approach compared to methods where a constant declustering time or standard storm surge length is used. The model requires high-resolution total water level data as input and contains three main steps to stochastically simulate artificial ESL events. These are the following: (i) identify ESLs within the input data and extract event hydrographs, (ii) parameterize each event, and (iii) generate artificial events by means of Monte Carlo simulations.

The model has been applied to 45 tide-gauge records located along the German Baltic Sea coast. ESL characteristics are modeled using a Gaussian copula, which preserves the dependence structure between observed ESL parameters. Artificial ESL hydrographs were compared to observed events in a number of statistical analyses to ensure model results reproduce local conditions well. These analyses showed that the model produces realistic and reliable ESL events at all locations tested, based on the available high-resolution tide-gauge data. However, as the model requires high-resolution data, it is therefore limited by the availability, length, and quality of such records. This is apparent when considering large historical ESLs measured before the introduction of high-resolution records. Nevertheless, it is possible to include such historic events in the model by adjusting the parametric distribution fitted to the peak water levels. It is therefore important that information on such past events are made available through digitization, allowing for more robust estimates of return water levels.

APPENDIX A: THRESHOLD SELECTION METHOD AND ANALYSIS

ESLs are identified using the POT method where thresholds are selected using the STM and MRL plot as described by Coles (2001). For both methods, a range of thresholds above MHW are tested. In the STM, a GP distribution is fitted to the high-water peaks exceeding each threshold. Where the shape parameter remains approximately

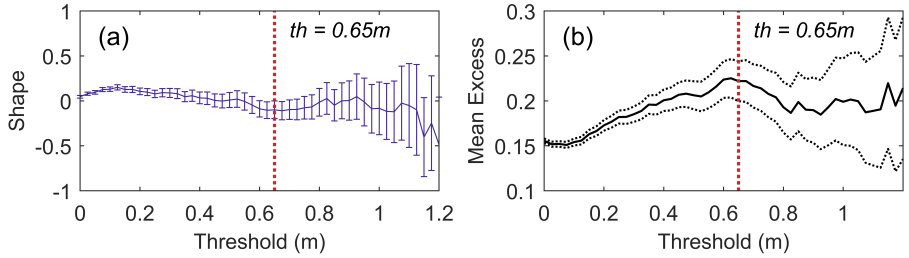

Figure 5.A1: Threshold selection using the (a) Stability Threshold Method and (b) Mean Residual Life methods for the water level data set measured at Timmendorf. A threshold is selected where the shape parameter of a generalized par- eto distribution remains constant and where the mean of the peaks above the threshold become linear. Here a threshold of 0.65 m (dashed red line) above mean high water was identified. The 95% confidence bounds are shown for each plot.

constant over consecutive thresholds tested, a threshold is selected. MRL plots the mean excess for each threshold tested. A threshold is selected where the function becomes linear. Figure 5.A1 shows both methods for the tide-gauge of Timmendorf, where an optimal threshold value of 0.65 m was obtained. It can be seen that an objective interpretation is difficult and care must be taken when selecting a threshold as this may bias the statistical assessment at the very beginning. Therefore, we conduct a sensitivity analysis for all tide-gauge records to identify the effect of threshold selection.

A comparison of return water level estimates made under a range of thresholds was conducted to measure the performance of the threshold selection technique proposed by Coles (2001). First, peaks within the water level data were declustered using an interval of 3 days. Next, thresholds were calculated based on the method proposed by Coles (2001) and by using the 95th, 97th, and 99th percentiles of declustered high water peaks. GP distributions were then fitted to the data to provide estimates of return water levels. An objective measure of the stability of return water level estimates was calculated using the Index of Return Period Stability (IRPS) outlined by Arns et al. (2013), which is equal to the product of the variance of the return water level curve, its maximum distance to the reference truth, and the mean distance to the reference truth, all as a function of time. Here the reference truth is given as the 100-year return water level, estimated using a GEV distribution fitted to the annual maximum water levels of the entire record (Arns et al., 2013). Figure 5.A2 shows the IRPS values calculated at all tide gauges.

When considering the median IRPS values, the greatest stability of return period estimates is provided when the threshold is selected

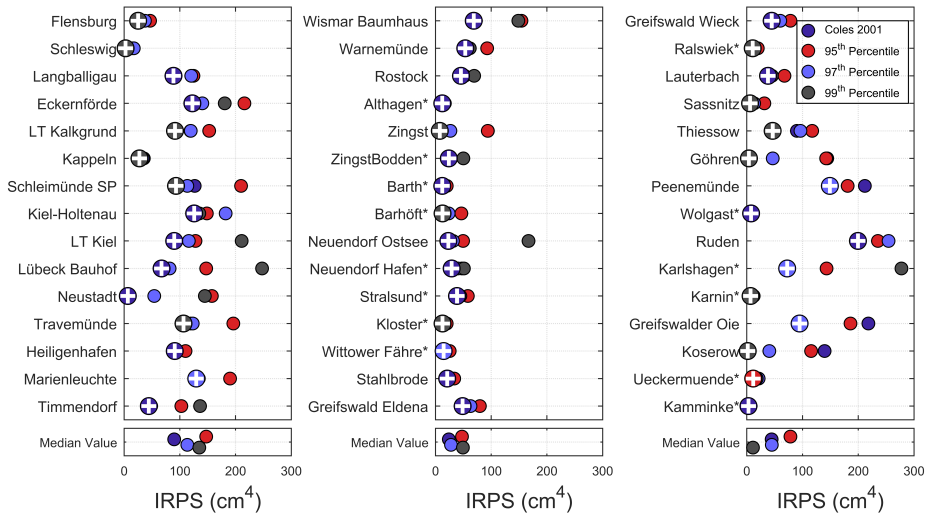

Figure 5.A2: Stability of the 100-year return water level based on thresholds selected using the Stability Threshold Method and Mean Residual Life methods (Coles, 2001) and by selecting only the 95th, 97th, and 99th percentile peak water levels. Values represent the Index of Return Period Stability and are calculated using the return period curve given by a generalized pareto. Techniques that provide the most stable estimates at each tide-gauge are highlighted with a white cross.

using the STM and MRL method as proposed by Coles (2001). Here a median IRPS of 44 cm^4 is calculated. This differs only slightly from the value provided by estimates where only the 97th percentile of high-water peaks are used (47 cm^4). However, in 34 of the 45 tide gauge records, the STM and MRL techniques provide more stable return period estimates than the 97th percentile method. Return periods estimated with 95th and 99th percentile high water peaks are less stable (80 and 51 cm^4 median values) and less consistent between tide-gauge records. Thus, we consider thresholds selected via the STM and MRL methods.

APPENDIX B: BIVARIATE ANALYSIS OF ESL EVENTS

We use Kendall's rank correlation to conduct a bivariate statistical analysis of ESL events at all tide-gauge locations, according to their values of peak water level and intensity. For the combined observed and artificial ESL data, we fit a bivariate distribution to the values of peak water level and intensity, which allows for the calculation of joint nonexceedance probabilities. To identify an appropriate multivariate distribution for this task, we test the ability of three Archimedean copulas (Frank, Clayton, and Gumbel) as well as the el-

liptical Gaussian and t-Student copulas to model the underlying data. First, each copula is fitted to the bivariate data using Kendall's rank correlation. Through Monte Carlo simulations, we model the peak water level and intensity of 100,000 events using each copula, fitted to the observed and simulated peak water level and intensity values. A two-dimensional KS test is used to select the most suitable copula based on distance between bivariate cumulative distributions of the initial and generated data. We find the Gumbel copula best suited to model this dependence at most tide-gauge locations. Overall, the mean KS statistic for all tide-gauges is lowest for the Gumbel copula (0.0301), followed by the Gaussian (0.0326), t-Student (0.0326), Frank (0.0338), and Clayton (0.0481) copulas. Variance in the KS statistic between tide gauges is also lowest for the Gumbel copula. Therefore, we calculate joint nonexceedance probabilities for peak water level and intensity data using the Gumbel copula.

Joint nonexceedance probabilities for all observed and artificial events are calculated for peak water level and intensity values using a Gumbel copula with function:

$$C_\rho(u,v) = \exp\left[((-\log(u))^\rho + (-\log(v))^\rho)^{\frac{1}{\rho}}\right], \tag{5.B1}$$

where C_ρ is the cumulative probability of the bivariate copula, u and v are the marginal cumulative probabilities of the two variables (in this case, peak water level and intensity), and ρ is the linear correlation coefficient, derived from Kendall's rank correlation via:

$$\rho = (1-\tau)^{-1}. \tag{5.B2}$$

By subtracting C_ρ from the marginals (u,v), joint nonexceedance probabilities are found.

Results of the stochastic simulation are shown in Figure 5.B1 for the tide-gauge of Travemünde. Here all observed and artificial events are plotted according to their peak water level and intensity. Return periods are shown as contour lines and were estimated using a Gumbel copula. Three events with a bivariate return period equal to 200 years have been highlighted. Despite sharing an equal likelihood of occurrence, each of these events shows different peak water levels and intensities. These factors are likely to affect potential damages along the coast and thus lead to different levels of risk.

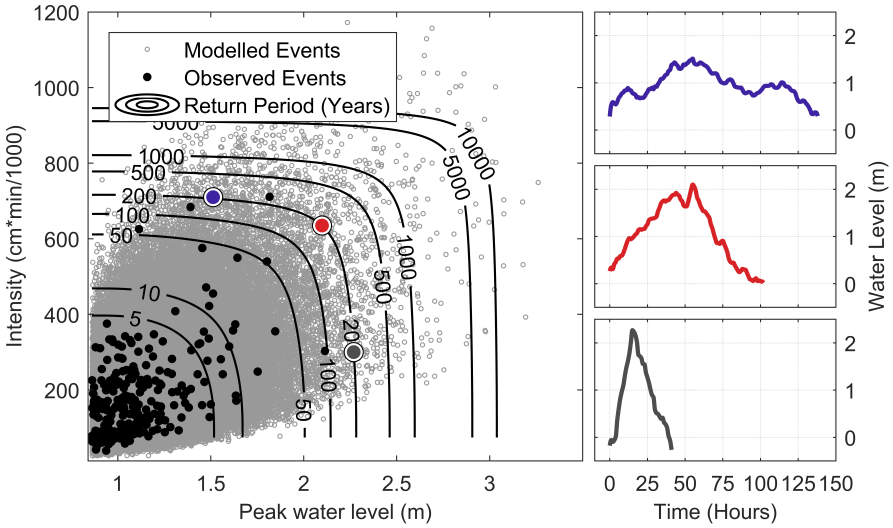

Figure 5.B1: Distribution of observed (black dots) and modeled (grey circles, 200,000 events) extreme sea level events based on peak water level and intensity at Travemünde. Contours highlight the bivariate return periods for all events. Three events with a return period of 200 years are highlighted and plotted on the right.

6

BAYESIAN EXTREME VALUE ANALYSIS OF EXTREME SEA LEVELS ALONG THE GERMAN BALTIC COAST USING HISTORICAL INFORMATION

This chapter is published as:

MacPherson, L. R., Arns, A., Fischer, S., Méndez, F. J., and Jensen, J. (2023). Bayesian Extreme Value Analysis of Extreme Sea Levels along the German Baltic Coast Using Historical Information. *Natural Hazards and Earth System Sciences*, 23, pp. 3685–3701. DOI: 10.5194/nhess-23-3685-2023..

ABSTRACT

Developed coastlines require considerable investments into coastal protection measures to mitigate the effects of flooding caused by extreme sea levels (ESLs). To maximise the effectiveness of these measures, accurate estimates of the underlying hazard are needed. These estimates are typically determined by performing extreme value analysis on a sample of events taken from tide-gauge observations. However, such records are often limited in duration and the resulting estimates may be highly uncertain. Furthermore, short records make it difficult to assess whether exceptionally large events within the record are appropriate for analysis or should be disregarded as outliers. In this study, we explore how historical information can be used to address both of these issues for the case of the German Baltic coast. We apply a Bayesian Markov-chain Monte-Carlo approach to assess ESLs using both systematic tide-gauge observations and historical information at seven locations. Apart from the benefits provided by incorporating historical information in extreme value analysis, which include reduced estimate uncertainties and the reclassification of outliers into useful samples, we find that the current tide-gauge records in the region alone are insufficient for providing accurate estimates of ESLs for the planning of coastal protection. We find long-range dependence in the series of ESLs at the site of Travemünde, which suggests the presence of some long-term variability affecting events in the region. We show that ESL activity over the full period of systematic observation has been relatively low. Consequently, analyses which consider only this data are prone to underestimations.

6.1 INTRODUCTION

Extreme sea levels (ESLs) and their associated probabilities of exceedance have long been studied due to their role in driving coastal flooding. The application of extreme value analysis (EVA) in this field is thus a well-developed science (Coles, 2001) and best practices based on observations from tide-gauge data have been suggested (Arns et al., 2013; Haigh et al., 2010). However, direct approaches to EVA require sufficiently long records of systematic data to maintain manageable uncertainties at high return periods (Pugh, 2004). In fact, uncertainties in the estimates of ESLs are a major source of uncertainty in expected flood damages in the short term (before 2040; Rohmer et al., 2021), leading to inefficient coastal adaptation. Furthermore, concerns regarding the sensitivity of estimates to extraordinarily large events have been raised (Dangendorf et al., 2016; MacPherson et al., 2019). Due to the difficulty of including such events in EVA using direct approaches, they are often treated as outliers and disregarded (Hofstede and Hamann, 2022; Jensen et al., 2022). However, excluding errors in reporting or measurement, or realisations of different random processes, these events offer important information on the underlying distribution, especially at the tails, and should thus be kept (Mazas and Hamm, 2011).

Alternatives to direct approaches include the joint probability method (JPM) and regional frequency analysis (RFA). The former was introduced to address the main limitations of the direct methods (Pugh and Vassie, 1980; Tawn et al., 1989b; Tawn, 1992) and involves analyses of the astronomical tide and non-tidal residual water level separately, whereby the final probability distribution is obtained based on the joint probabilities of the two components (Haigh et al., 2010). The latter increases the available sample of extremes by combining records within homogenous regions (Weiss et al., 2014a; Arns et al., 2015b; Bardet et al., 2011) and dealing with local characteristics using a scaling factor. Although both methods address the limitations of direct EVA approaches, concerns regarding their use remain. Principally, they are still constrained by the observation period of the tide-gauges used.

Tide-gauge data offers the most valuable information on ESLs, owing to their widespread implementation, providing systematic measurements often over many decades. In addition to this data, information on ESLs that occurred prior to the introduction of tide-gauges is available at many locations. Despite this, historical events are rarely considered in EVA due to difficulties in reconciling the historical information with systematic data. The main cause of these diffi-

culties is that historical information does not have a well-defined period of observation (Prosdocimi, 2018). That is, traditional EVA depends on a known period in which all sampled extremes have occurred, and as historical measurements are isolated data points, a duration of observation is not defined (Frau et al., 2018). Despite this, several statistical methods exist to combine historical information and systematic data in the field of hydrology (Benito et al., 2004). Of these methods, Bayesian techniques offer a natural framework for handling uncertainties in an extreme value setting (Coles et al., 2003).

In response to gross underestimations of predicted rainfall events in comparison to historical measurements, Coles et al. (2003) introduced a Bayesian approach as an alternative to standard statistical tools for the prediction of extreme events. This was later adapted for flood frequency analysis (Reis and Stedinger, 2005) and the modelling of ESLs (Coles and Tawn, 2005). Here, historical information is treated as censored observations and incorporated into the model likelihood. Standard techniques to maximise the likelihood of such models are intractable, therefore stochastic algorithms such as Markov chain Monte Carlo (MCMC) are employed. This method of modelling extremes, particularly for the incorporation of historical information, has been applied regularly in the field of hydrology (Bulteau et al., 2015; Isikwue et al., 2015; Payrastre et al., 2011; Gaume, 2018; Gaál et al., 2010).

In this study, we apply a Bayesian MCMC algorithm to incorporate historical information in the analysis of ESLs along the German Baltic coast. The region has a long history of ESLs, which includes an extraordinary event occurring in November of 1872 (Jensen et al., 2022). Due to the exceptional magnitude of the event, which lead to widespread flooding, it is often disregarded as a statistical outlier (Hofstede and Hamann, 2022). With the recent devastating floods in western Germany of July 2021 (Mohr et al., 2023; Ludwig et al., 2023), consideration of historical information has grown in relevance (Jensen et al., 2022). We assess ESL quantiles at seven sites along the German Baltic coast where historical information in combination with systematic water level records are available. In particular, we compare estimates of design water levels where historical information is both included and omitted. Lastly, we study recent trends in ESL activity in the region and examine how historical information can provide stable ESL estimates despite short systematic records.

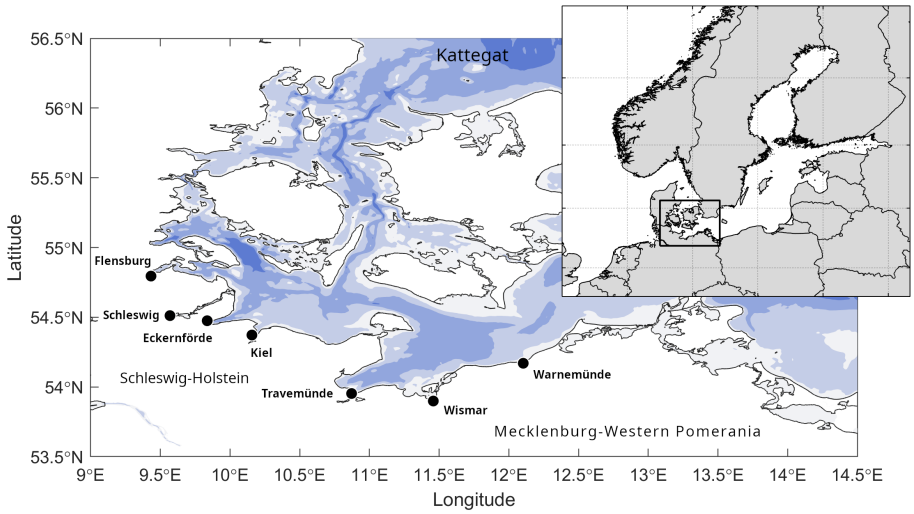

Figure 6.1: Sites considered in this study along the German Baltic coast. Black dots mark the locations of tide-gauges from which systematic records have been used. The two federal states that share the German Baltic Sea coast, Schleswig-Holstein and Mecklenburg-Western Pomerania, have a border at Travemünde.

6.2 BACKGROUND

6.2.1 *Study site and data*

The German Baltic coast located in the southwest Baltic Sea (Figure 6.1) has a long history of ESLs. Due to its location, ESLs in the region occur mainly during periods of strong northeasterly winds. However, as the Baltic Sea is a semi-enclosed basin, water levels at the German coast are affected by seiches acting over the entire sea. These seiches are influenced not only by large-scale atmospheric winds and pressure, but by specific sequences of regional wind patterns (Jensen and Müller-Navarra, 2008). Rosenhagen and Bork (2009) and Bork et al. (2022) describe one such sequence in November of 1872, when strong southwesterly winds drove large volumes of water into the Baltic Sea via the Kattegat and caused high water levels in the eastern Baltic Sea. The storm than reversed direction and intensified, causing high water levels and widespread flooding along much of the western Baltic Sea coast. The resulting ESLs along the German coast remain the highest on record, registering approximately 3.4 m above mean sea level (MSL) at Travemünde.

Coastal defence heights along the German Baltic coast are defined by the two federal states which share the coastline, Schleswig-Holstein (MELUND 2022) and Mecklenburg-Western Pomerania

Table 6.1: List of sea level data used in this study, including information on data type, sampling rate and source.

Data	Type	Sample Rate	Source
Tide-gauge records	Systematic	hourly	GESLA 3 (Haigh et al., 2022), Schmidt et al. (2017)
AMAX water levels	Systematic/Historical	annual	MLUV (2009)
Historical information	Historical	-	Jensen and Töppe (1990), Jensen et al. (2022)

(MLUV 2009). Current design heights are based on an ESL with a return period of 200 years (hereafter referred to as HW200) and include an additional 50 cm to account for future climate induced changes. Values of HW200 are determined based on the statistical analysis of past observations, whose accuracy are thus dependent on the length and quality of the sea level records used. Unfortunately, the exact data and methods used to derive the official return water levels are not published. MLUV (2009) have stated that the Gumbel distribution generally results in the best fit which suggests the use of AMAX data as input.

At present, more than 45 high-resolution (at least hourly samples) tide-gauge records cover the approximately 2,110 km length of the German Baltic coastline (MacPherson et al., 2019). The longest high-resolution (hourly sampling) systematic record was installed at Travemünde at the end of 1949. In addition to these, records of annual maxima (AMAX) water levels are available at 14 sites, which were compiled by the ministry of Agriculture and Environment in Mecklenburg-Western Pomerania (MLUV 2009) from both systematic tide-gauge data and historical information (see Figure 6.3). These data cover only the coastline of Mecklenburg-Western Pomerania but also at Travemünde, which is a special case due to its location at the border of both states. Here, values for AMAX water levels are available for 184 years between 1826 and 2009.

Lastly, historical measurements can be found for a number of past ESLs at sites located along the Schleswig-Holstein coast. These measurements have been compiled and summarised by Jensen and Töppe (1990). Additionally, Jensen et al. (2022) conducted and in-depth review of historical water levels at Travemünde, providing best estimates of historical ESLs including measurement uncertainties. Where available, we employ all data mentioned above at the sites of Flensburg, Schleswig, Eckernförde, Kiel, Travemünde, Wismar and Warnemünde (see Figure 6.1 for locations). Sources for the data used in this study are summarised in Table 6.1, and the extent of the data at each site is illustrated in Figure 6.2.

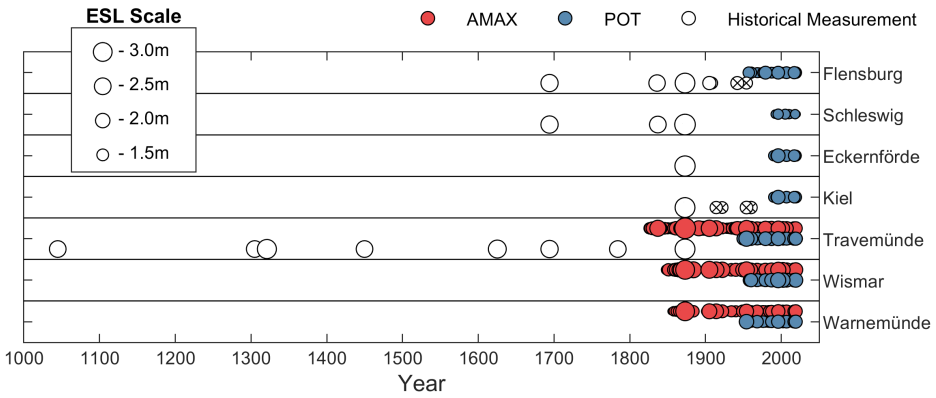

Figure 6.2: The extent of data available at all locations. Each circle denotes a sampled ESL with its size proportional to the event's magnitude (height above NHN). All data has been detrended using MSL. Historical events which lie below the perception threshold and are thus disregarded in the final analysis are shown with a black cross.

6.2.2 *Extreme value models*

Observed ESLs exhibit an asymptotic behaviour that can be modelled using EVA (Coles, 2001). The choice of extreme value model is dependent on the behaviour of the distribution, which is influenced by how the extreme events were sampled. The two most common extreme value models used in hydrology are the generalized extreme value (GEV) and generalized Pareto (GP) distributions. The former is suited to modelling extremes sampled using the block maxima method, where maxima are taken from individual blocks of data equal in length. For ESLs, which are influenced by seasonal trends, a block length of 1 year is most appropriate, and hence annual maxima (AMAX) samples are preferred. A special case of the GEV distribution occurs when the shape parameter (ξ) is equal to 0. Here, the distribution becomes a Gumbel distribution, which is mentioned by MLUV (2009) as the best fit for ESLs along the Mecklenburg-Western Pomeranian coast. An alternative approach is to sample ESLs using the peaks-over-threshold (POT) method, where all events above a certain threshold are selected and modelled using the GP distribution. This approach is generally preferred over the simpler AMAX method as it addresses two main limitations. First, the AMAX method can be wasteful, discounting extremes if multiple large events lie within any one block. Second, it is possible that the analysis becomes biased by the inclusion of moderate values if the data contains long periods of non-extremes.

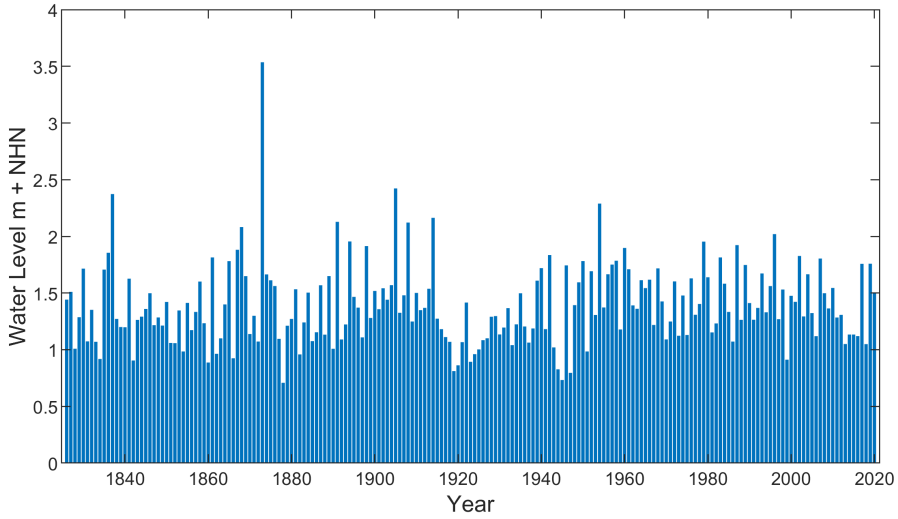

Figure 6.3: Annual Maxima (AMAX) sea levels recorded at Travemünde detrended using MSL.

In this study, we make use of both GEV and GP distributions to model the occurrence of ESLs. Both distributions employ a shape (ξ), scale (σ) and location (μ) parameter. To simplify the probability density function (PDF) and cumulative distribution function (CDF) of both models, we consider the standardized variable $z = (x - \mu)/\sigma$, where x is the independent and identically distributed (iid) random variable to be modelled. The PDF of the GEV distribution is:

$$f(z;\xi) = \begin{cases} \exp(-z)\exp(-\exp(-z)) & \text{for } \xi = 0 \\ (1+\xi z)^{-(1+1/\xi)} \exp\left(-(1+\xi z)^{-1/\xi}\right) & \text{for } \xi \neq 0 \text{ and } \\ & \quad \xi z > -1 \\ 0 & \text{otherwise,} \end{cases} \tag{6.1}$$

and the CDF is thus:

$$F(z;\xi) = \begin{cases} \exp(-\exp(-z)) & \text{for } \xi = 0 \\ \exp\left(-(1+\xi z)^{-1/\xi}\right) & \text{for } \xi \neq 0 \text{ and } \xi z > -1 \\ 0 & \text{for } \xi > 0 \text{ and } \xi z \leq -1 \\ 1 & \text{for } \xi < 0 \text{ and } \xi z \leq -1. \end{cases} \tag{6.2}$$

The GEV distribution is valid for $x > \mu - 1/\xi$ where $\xi > 0$, and $x < \mu - 1/\xi$ where $\xi < 0$. Where $\xi = 0$, support for the distribution is found for all real values of x. The GP distribution is valid for all

$x > \mu$, unless $\xi < 0$, in which case $\mu \leq x \leq \mu - \sigma/\xi$. The PDF of the GP distribution is:

$$f(z;\xi) = \begin{cases} (1+\xi z)^{-\frac{\xi+1}{\xi}} & \text{for } \xi \neq 0 \\ \exp(-z) & \text{for } \xi = 0, \end{cases} \tag{6.3}$$

and the CDF of the GP distribution is:

$$F(z;\xi) = \begin{cases} 1-(1+\xi z)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ 1-\exp(-z) & \text{for } \xi = 0. \end{cases} \tag{6.4}$$

6.3 METHODS

6.3.1 *Bayesian framework*

If we consider a sample of ESL observations x, the distribution of events within this sample can be represented using an extreme value model with parameter vector θ. Due to the finite sample of extremes, the exact parameters cannot be known for certain. However, Bayes' theorem relates the posterior distribution of the parameters θ given the sample of events x, to the likelihood function $\mathcal{L}(x \mid \theta)$:

$$f(\theta \mid x) = \frac{\mathcal{L}(x \mid \theta) f(\theta)}{f(x)}, \tag{6.5}$$

where $f(\theta)$ is the prior distribution and $f(x)$ is a normalising constant dependent on the sampled extremes only. While including information in the prior distribution can improve inference results and reduce uncertainties (Reis and Stedinger, 2005), El Adlouni and Ouarda (2009) cautions that the choice of prior distribution can introduce bias and should be made with care. Where little to no information on the prior distribution can be found, a non informative prior should be used (Payrastre et al., 2011). In this case, $f(\theta) \propto 1$ and the posterior distribution becomes proportional to the likelihood. By sampling the posterior distribution using a stochastic MCMC algorithm (described in Section 3.2), calculation of the normalising constant is not required.

Formulation of the likelihood function is dependent on the characteristics of the observations, and can be split into two key parts, separating the periods of systematic and historical observations. For

the set of s systematic observations $X = \{x_1, x_2, ..., x_s\}$ and h historical observations $Y = \{y_1, y_2, ..., y_h\}$, the likelihood function is:

$$\mathcal{L}(X, Y \mid \theta) = \underbrace{\mathcal{L}(X \mid \theta)}_{\text{syst. likelihood}} \cdot \underbrace{\mathcal{L}(Y \mid \theta)}_{\text{hist. likelihood}} . \tag{6.6}$$

The likelihood function for the systematic data is simply equal to the product of the probability density (f_θ) of each observation:

$$\mathcal{L}(X \mid \theta) = \prod_{i=1}^{s} f_\theta(x_i). \tag{6.7}$$

Formulation of the historical likelihood function is more complex and is taken from Payrastre et al. (2011) and Bulteau et al. (2015) for the GEV and GP distributions respectively. Both formulations require a perception threshold (X_0), above which the sample of historical information must be exhaustive. The key difference between the two formulations relates to the handling of sampling frequencies. Whereas Payrastre et al. (2011) deal entirely with AMAX data, Bulteau et al. (2015) adapt the methodology for POT observations and must therefore deal with an unknown number of censored events. In general, the likelihood function for the historical information is:

$$\mathcal{L}(Y \mid \theta) = \underbrace{\prod_{j=1}^{h_1} f_\theta(y_j)}_{\text{(a)}} \cdot \underbrace{\prod_{k=1}^{h_2} F_\theta(y_k^{ub}) - F_\theta(y_k^{lb})}_{(b)} \cdot \underbrace{(1 - F_\theta(X_0))^{h_3}}_{\text{(c)}} \cdot \underbrace{P(h \mid \theta)}_{\text{(d)}}, \tag{6.8}$$

where f_θ and F_θ is the probability density and cumulative distribution functions respectively. The terms in Eq. 6.8 give general expressions for different types of historical observations in Y, where Y consists of $Y = h_1 + h_2 + h_3$ events. These terms describe (a) h_1 events with precise measurements, (b) h_2 events with upper (y^{ub}) and lower bounds (y^{lb}), and (c) h_3 events known to have exceeded the perception threshold but with no known upper limit. The term in (d) relates to the probability of observing h events exceed X_0 for the period of historical observation (ny) in years. For the GEV distribution and AMAX data, where there are a known number of missing observations, this term becomes:

$$P(h \mid \theta) = F_\theta(X_0)^{(n_y - h)}. \tag{6.9}$$

For the POT data and GP distribution, where observations have been sampled at an average frequency of λ events per year, the number of exceedances of X_0 follow a Poisson process:

$$P(h \mid \theta) = \frac{(\lambda n_y [1 - F_\theta(X_0)])^h}{h!} \exp(-\lambda n_y [1 - F_\theta(X_0)]). \tag{6.10}$$

6.3.2 *Markov chain Monte Carlo*

Sampling of the posterior distribution $f(\theta \mid X, Y)$ is done using the Metropolis-Hastings MCMC algorithm (Metropolis et al., 1953; Hastings, 1970). Starting with an arbitrary parameter vector θ, an iterative process is conducted where candidate vectors θ' are tested and either accepted or rejected based on the ratio of likelihoods:

$$\alpha = \frac{\mathcal{L}(X, Y \mid \theta')}{\mathcal{L}(X, Y \mid \theta_t)}, \tag{6.11}$$

where θ_t is the last accepted parameter vector. For a uniform random number $r \in [0, 1]$, the candidate vector θ' is accepted if $\alpha \geq r$, and thus $\theta_{t+1} = \theta'$, otherwise it is rejected and $\theta_{t+1} = \theta_t$. Candidate vectors are chosen based on a Gaussian distribution centred on the last accepted vector. The scale of this distribution is controlled so that the acceptance rate is around 25%.

The output of the MCMC is a large set of parameter vectors θ with densities $f(\theta|X, Y)$. For each vector, an extreme value model is defined, and quantiles of ESLs can be computed. Maximum likelihood estimates are associated with the mode of the set of vectors, and credibility intervals may be calculated based on the quantiles computed from the whole set of vectors. Naturally, the mode of the parameter vectors is that which maximises the likelihood function.

6.3.3 *Data preparation, sampling and simulations*

Before the Bayesian MCMC simulations can be performed, the sea level data must first be prepared and sampled. Data preparation involves removing long-term trends such as changes in MSL and ensuring observed events are measured relative to the standard vertical datum used in the region, normalhöhennull (NHN). For the AMAX data, this is done using monthly MSL values taken from the Permanent Service for Mean Sea Level (PSMSL, 2018). Unfortunately, these records do not begin at the inception of the AMAX measurements (30 years short at Travemünde, 22 years at Wismar and 29 years at Warnemünde), instead we extend the PSMSL trends linearly to cover this shortfall. This method may not be the most appropriate due to accelerations in the rate of sea level rise. Although a quadratic trend results in differences of less than 1 cm at Travemünde and Warnemünde, at Wismar a maximum difference of approximately 5 cm is found. The use of a linear trend over a

quadratic trend results in an increase to the AMAX samples not covered by the PSMSL data, which in turn leads to a positive bias of the final ESL estimates. However, it is unclear whether the quadratic trend would be better suited to the data, and in combination with the minor differences seen at Travemünde and Warnemünde, a linear trend is considered suitable for our purposes.

For the high-resolution tide-gauge data, MSL is calculated as a 1-year moving average of sea levels as suggested by Arns et al. (2013). Thankfully, most historical measurements are recorded relative to MSL at the time, and thus no correction is needed. Historical events measured to NormalNull (NN), the old standard vertical datum used in Germany, are detrended by transferring the measurements to NHN using local adjustment values and removing MSL using the available PSMSL data.

Given the availability of AMAX and high-resolution data, we employ two approaches to sample ESLs which are described in Section 2.2. First, AMAX sampling is an obvious choice as long records of AMAX data are available at Travemünde, Wismar and Warnemünde. These records provide the largest water levels for each hydrological year, starting on the 1st of November and ending on the 31st of October the following year. While no further sampling is required, each of these records ends in 2009, 11 years before the end of the high-resolution tide-gauge data. To maximise the available data for our analyses, each AMAX set is extended by sampling from the tide-gauge data over the missing years (2009-2020). As a result, the AMAX data consists of 195 ESLs at Travemünde, 155 at Wismar and 147 at Warnemünde. For the high-resolution data, ESLs are sampled using the POT method. We follow the approach outlined by MacPherson et al. (2019) for threshold selection, who conducted an analysis of EVA techniques for the German Baltic coast. Here, a threshold equal to the 98th percentile of high-water peaks is chosen. Before selecting the threshold, peak water levels are declustered using an interval of 3 days to ensure independence. All ESLs that exceed this threshold are extracted from the record to form the ESL sample.

As mentioned in Section 3.1, it is a necessary condition that the available historical information is exhaustive above a perception threshold. That is, the only events which have exceeded the perception threshold for the duration of historical observation exist within the historical record. Therefore, the perception threshold should be set high enough to ensure this assumption is true. At first, a systematic approach to setting a perception threshold was attempted based on the systematic data and the period of historical information. Here, ESLs were estimated using systematic data only for return pe-

riods dependent on the number of historical events available and the length of the historical record. For example, the perception threshold might be set to a height equivalent to a 1-in-100 year event, where a 200 year long historical record is available which contains 2 events. However, due to large differences in the magnitude of systematic and historical observations, relying on the systematic data alone was not sufficient, and no one method could be applied at all sites. Instead, perception thresholds were chosen on a site-by-site basis, using all available data for each case.

Given the lack of a clear physical threshold at any of the tested locations (e.g. a sea wall where all exceeding events are recorded), a threshold selection process was conducted based simply on the author's intuitive reasoning. Factors that influenced the selection process include the magnitude and occurrence of ESLs in both the systematic and historical records and the length of the historical record in question. Keeping in mind the assumption that the historical record is exhaustive, and due to the subjective nature of this method, final perception thresholds were set conservatively high at 2.3 m at Flensburg, 2 m at Schleswig, 2.25 m at Eckernförde, 2.25m at Kiel, 2.6 m at Travemünde, 2.25 m at Wismar and 2 m at Warnemünde. Historical ESLs that do not exceed the perception threshold cannot be used in the analysis, and are thus disregarded. These events are highlighted in Figure 6.2.

ESLs at each site were modelled using the Bayesian MCMC approach described in Section 3.1 and 3.2. Although traditional EVA may be performed using the systematic data alone, the Bayesian MCMC approach is used in all cases for comparison purposes. We use the GP distribution to model POT data (POT-GP) and the GEV distribution to model AMAX data (AMAX-GEV). Depending on data availability at each site, we perform four separate analyses using: 1) POT samples only, 2) POT samples with historical measurements, 3) AMAX samples only, and 4) AMAX samples with historical measurements. At Wismar and Warnemünde, as no historical records are available, ESLs from the AMAX records are used in lieu of historical measurements for the second analysis. For the POT-GP analyses, the sampling frequency (λ) and location parameter (μ) are assumed to be constant for the combined period of historical and systematic observation. At each site, μ is equal to the threshold used to sample the ESLs. Hence, only the shape (ξ) and scale parameters (σ) of the GP distribution are considered. In contrast, all three parameters of the GEV distribution (ξ, σ, μ) are sampled.

6.3.4 *ESL stationarity, long-range dependence and variability*

From the AMAX record at Travemünde (Figure 6.3), a decrease in the number of large ESL events over the past century can be seen. Of these events (>95th percentile), 7 occurred prior to 1920 with only 3 occurring afterwards. We perform several analyses to assess the stationarity and long-range dependence of ESLs at Travemünde and to identify periods of low and high ESL activity.

Stationarity refers to the property of a stochastic process or time series where the statistical properties of the process remain constant over time. More specifically, a stationary process has a constant mean, constant variance, and constant autocovariance over time. As the methods used to estimate ESLs in this study assume the underlying data is stationary, we first perform KPSS (Kwiatkowski et al., 1992) tests on samples of ESLs at all sites for both POT and AMAX samples to confirm this. The KPSS test evaluates the cumulative sum of the deviations from the estimated trend in the time series. If this sum exceeds some critical value based on the sample size, the null hypothesis of stationarity is rejected, indicating the presence of a unit root and non-stationarity. On the other hand, if the sum is below the critical value, the null hypothesis cannot be rejected, suggesting the series is stationary.

Next, we check for long-range dependence within the AMAX record which describes the persistence or correlation between distant observations. We quantify long-range dependence using the Hurst exponent (Hurst, 1951), which is a measure of the strength of the correlation between observations that are far apart in time. For a series of N observations, the Hurst exponent is calculated using the rescaled range (R/S) analysis. This involves creating a mean adjusted series and segmenting the observations into smaller subsets of size $n < N$. For each segment, the range (R) between the maximum and minimum values is calculated. The rescaled range (S) is then determined by averaging the range of all segments and dividing by the standard deviation of the entire series. This is repeated for a number of segment sizes to account for both short and long periods of observations. Given the 195 observations available at Travemünde, we chose to use segment sizes of $n = 3, 6, 12, 24, 48, 96$. The Hurst exponent is found by plotting the log of the rescaled range ($\log(S)$) against the log of the segment size ($\log(n)$) and fitting a straight line. The Hurst exponent is thus the slope of the fitted line. A Hurst exponent greater than 0.5 indicates long-range dependence, while a value less than 0.5 indicates short-range dependence or anti-persistence.

To identify periods of low or high ESL activity, EVA is applied using an iterative process where AMAX data is limited to those events occurring within a 70-year moving window. A window size of 70 years was chosen to match the length of the high-resolution tide-gauge record at Travemünde (71 years in length). For each iteration, the Bayesian MCMC is applied, and the window is then moved forward one year. Maximum likelihood estimates of HW200 and their corresponding 95% credibility intervals are computed at each step from the resulting set of GEV parameter vectors. The analysis was performed twice, once with systematic data only and again with historical information included. For the second case, all available historical information is used, even those events which occur outside of the 70-year window.

Lastly, we consider how the sample size of the systematic data affects ESL estimates. Here, probability density estimates of HW200 were computed for samples of AMAX observations ranging in size from 70 to 195 events. At each tested sample size, 10,000 sets of AMAX observations were generated using bootstrap sampling, from which an equal number of HW200 maximum likelihood estimates were made. Probability density estimates of HW200 at each sample size are calculated for comparison.

6.4 RESULTS

6.4.1 *ESL estimates*

Incorporating historical information in the analysis of ESLs along the German Baltic coast results in significant changes to ESL estimates. Estimates of HW200 and HW1000 (1-in-1000 year ESL event) at all tested locations are given in Table 6.2, including the upper and lower bounds of the 95% credibility intervals. Maximum likelihood estimates from the POT-GP analyses increase at all sites when historical information is included, by as little as 16.5% (33 cm) at Warnemünde and as much as 47.9% (75 cm) at Schleswig. There are decreases in the range of the 95% credibility intervals at four of the seven tested locations. This is most evident at Travemünde where a decrease of 51.4% occurs at the HW200 level. At sites where AMAX data was considered, the effect of incorporating historical information can only be examined at Travemünde. Here, changes in the maximum likelihood estimation are negligible (< 1%) at both HW200 and HW1000. However, uncertainties in the form of 95% credibility

Table 6.2: Estimates of HW200 and HW1000 at all sites (rounded to the nearest cm). The first figure provides the maximum likelihood estimate while the range of the 95% credibility intervals are given in square parentheses with lower and upper bounds in round parentheses below. Changes to maximum likelihood estimates are given as percentages, with changes to the range of the 95% uncertainty bounds are shown in square parentheses.

		HW200			HW1000		
	Site	**Syst. Only**	**Syst. + H.I.**	**Change**	**Syst. Only**	**Syst. + H.I.**	**Change**
POT-GP	**Flensburg**	1.95 [0.60] (1.84-2.44)	2.49 [0.66] (2.28-2.93)	27.9% [9.8%]	2.04 [0.93] (1.90-2.83)	2.93 [1.20] (2.57-3.77)	43.3% [28.7%]
	Schleswig	1.56 [0.67] (1.51-2.18)	2.30 [0.78] (2.05-2.84)	47.9% [16.6%]	1.59 [1.01] (1.53-2.54)	2.86 [1.60] (2.42-4.02)	79.7% [58.4%]
	Eckernförde	2.07 [1.77] (1.90-3.67)	2.52 [1.19] (2.22-3.40)	21.7% [-33.0%]	2.24 [3.23] (2.00-5.23)	3.02 [2.35] (2.54-4.89)	35.1% [-27.3%]
	Kiel	2.07 [1.67] (1.90-3.57)	2.50 [1.13] (2.21-3.34)	20.4% [-32.4%]	2.25 [3.02] (2.00-5.02)	2.98 [2.19] (2.52-4.71)	32.7% [-27.6%]
	Travemünde	2.22 [1.14] (2.03-3.17)	2.72 [0.55] (2.52-3.07)	22.7% [-51.4%]	2.43 [2.01] (2.15-4.16)	3.29 [1.08] (2.92-4.01)	35.4% [-46.0%]
	Wismar	2.21 [0.97] (2.02-2.99)	2.66 [0.98] (2.39-3.37)	20.5% [1.3%]	2.40 [1.64] (2.13-3.77)	3.14 [1.93] (2.69-4.62)	30.8% [17.2%]
	Warnemünde	1.99 [0.95] (1.82-2.77)	2.32 [0.85] (2.08-2.93)	16.5% [-10.9%]	2.20 [1.72] (1.94-3.66)	2.77 [1.64] (2.37-4.01)	25.9% [-4.5%]
AMAX-GEV	**Travemünde**	2.67 [0.66] (2.47-3.13)	2.68 [0.38] (2.53-2.92)	0.6% [-42.1%]	3.07 [1.10] (2.77-3.86)	3.10 [0.63] (2.86-3.49)	0.8% [-42.7%]
	Wismar	2.68 [0.74] (2.46-3.20)	- -	- -	3.07 [1.24] (2.72-3.96)	- -	- -
	Warnemünde	2.27 [0.66] (2.10-2.76)	- -	- -	2.62 [1.09] (2.35-3.45)	- -	- -

intervals decrease by approximately 42%, which is equivalent to 28 cm at HW200 and 48cm at HW1000.

To illustrate the benefits of incorporating historical information in EVA, Figure 6.4 shows the results of the MCMC method at Travemünde for both POT-GP and AMAX-GEV analyses. While the high-resolution tide-gauge record is not considered short and is in fact the longest available record of its type along the German Baltic coast, there are large differences between estimates of ESLs made using the POT sample taken from the tide-gauge record and the much longer AMAX data set (71 vs. 195 years). However, by incorporating historical information in EVA, these discrepancies are substantially reduced. For example, maximum likelihood estimates of HW200 and HW1000 differ by 45 cm and 64 cm respectively between the POT-GP and AMAX-GEV analyses. These values are reduced to 5 cm at HW200 and 19 cm at HW1000 when historical information is included.

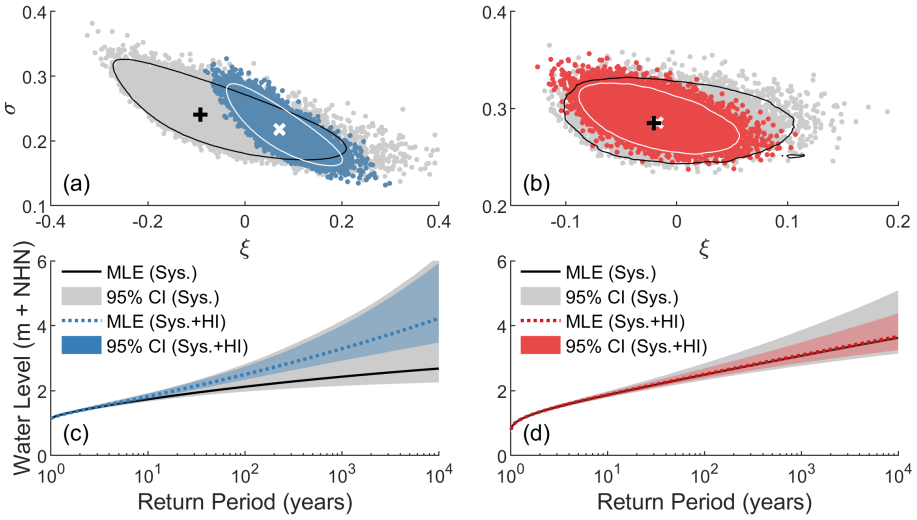

Figure 6.4: Results of the MCMC at Travemünde (blue: POT-GP, red: AMAX-GEV): **(a)** and **(b)** scatter plots of the shape (ξ) and scale (σ) parameters. For analyses conducted using systematic data only, sampled parameter pairs are shown as grey dots with the mode of the distribution shown as a black plus. Analyses conducted using both systematic data and historical information show sampled parameter pairs as coloured dots with the mode of the distribution shown as a white cross. 95th percentile contour lines are shown as black for analyses of systematic data only and white for systematic data and historical information. Return water level plots are shown in **(c)** and **(d)** for the POT-GP and AMAX-GEV analyses respectively, including maximum likelihood estimates (MLE) and 95% credibility intervals (CI).

While changes to the maximum likelihood estimates in the POT analysis is stark, perhaps more apparent is the reduction in uncertainties. For instance, the range of estimates within the 95% credibility intervals decreases by approximately 54% and 42% at HW200 for the POT-GP and AMAX-GEV analyses respectively, when historical information is included. Similar values can also be seen at HW1000. Interestingly, most of the reduction in uncertainty occurs at the lower bound for the POT-GP analysis and the upper bound for the AMAX-GEV analysis.

Similar benefits can also be seen at sites other than Travemünde. Figure 6.5 shows maximum likelihood estimates and 95% credibility intervals for HW200 at all sites considered in this study. Although uncertainties increase at Flensburg, Schleswig and Wismar, these also coincide with large increases to the maximum likelihood estimates. Large reductions in uncertainties can be seen at Eckernförde and Kiel, with a slight reduction at Wismar. Where only systematic data is considered, a large proportion of uncertainty lies above the

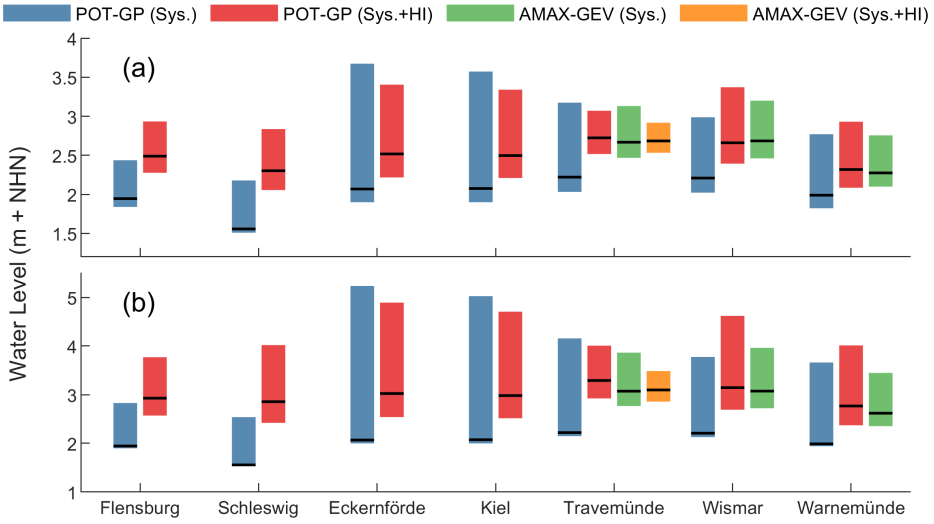

Figure 6.5: Comparison of **(a)** HW200 and **(b)** HW1000 estimates at all sites. Maximum likelihood estimates are shown as black horizontal lines. 95% credibility intervals are shown as colored bars. Where historical information is included, ESL estimates increase at all sites and credibility intervals are generally reduced. This occurs for both the POT-GP and AMAX-GEV analyses.

maximum likelihood estimates (Flensburg: 82.2%, Schleswig: 92.8%, Eckernförde: 90.5%, Kiel: 89.5%, Travemünde: 83.5%, Wismar: 80.6% and Warnemünde: 82.3%). Understandably, the inclusion of large historical events reduces uncertainty in the upper bound of the credibility intervals, proportionally if not quantitatively (Flensburg: 67.6%, Schleswig: 68.3%, Eckernförde: 74.7%, Kiel: 74.6%, Travemünde: 62.6%, Wismar: 72.7% and Warnemünde: 72.3%).

Looking only at the sites where long AMAX records are available (Travemünde, Wismar and Warnemünde), we see much better agreement between the POT-GP analyses including historical information and the AMAX-GEV analyses of systematic data only, despite the significantly shorter tide-gauge records. Differences in HW200 estimates decrease from 45cm, 48 cm and 28 cm at Travemünde, Wismar and Warnemünde respectively, to 5 cm, 2 cm and 5 cm.

Including historical information also allows for a more reasonable representation of the 1872 event. Table 6.3 shows the estimated return period in years of the 1872 event at each site and for each analysis. Given only high-resolution tide-gauge data, return period estimates of 1872 are not realistic, suggesting that the event is an outlier. At travemünde, 1872 is estimated to have a return period of more than 500 billion years. Furthermore, no estimates could be made at Flensburg, Schleswig, Eckernförde or Kiel, as the 1872 value is not defined within the resulting maximum likelihood distributions.

Table 6.3: Return period estimates (years) of the 1872 event at all sites. Estimates that are 'Undefined' represent values which lie outside the range of the maximum likelihood distribution. Values that could not be found due to insufficient data are shown as dashes (-).

Site	POT-GP (Syst.Only)	POT-GP (Syst. + H.I.)	AMAX-GEV (Syst.Only)	AMAX-GEV (Syst. + H.I.)
Flensburg	Undefined	1.57×10^3	-	-
Schleswig	Undefined	2.86×10^3	-	-
Eckernförde	Undefined	1.63×10^3	-	-
Kiel	Undefined	1.37×10^3	-	-
Travemünde	5.59×10^{11}	1.90×10^3	6.80×10^3	5.82×10^3
Wismar	3.99×10^6	0.73×10^3	0.90×10^3	-
Warnemünde	4.23×10^6	2.29×10^3	6.76×10^3	-

At Wismar and Warnemünde, estimated return periods are also high at approximately 4 million years. When historical information is included, return periods of between 700 and 2860 years are assigned to the 1872 event. These estimates are in the same order of magnitude provided by the AMAX-GEV analyses, which include the 1872 event within the systematic data.

6.4.2 *ESL variability*

Before considering ESL variability, stationarity was confirmed at each site using the KPSS test for POT samples and AMAX samples where available. In regards to the AMAX data at Travemünde, we find that the series exhibits strong long-range dependence with a Hurst exponent of 0.69. Similarly high Hurst exponents were found at Wismar (0.77) and Warnemünde (0.62). This suggests that there is persistency in the series of ESLs at Travemünde, Wismar and Warnemünde which can be seen as some long-term variability. This is likely to result in fluctuations in the occurrence of ESLs with periods of many large events and periods of many small events.

The discovery of long-range dependence in the series of AMAX ESLs at Travemünde in combination with large increases in ESL estimates due to the inclusion of historical information raises questions about the stability of estimates made using only short tide-gauge records. As described in Section 3.4, we assess changes in HW200 estimates over time by (a) performing traditional AMAX analyses based on a moving window of 70 years and (b) comparing probability density estimates from bootstrap AMAX samples of different sizes. For both analyses, AMAX observations from Travemünde were used. Results are shown in Figure 6.6.

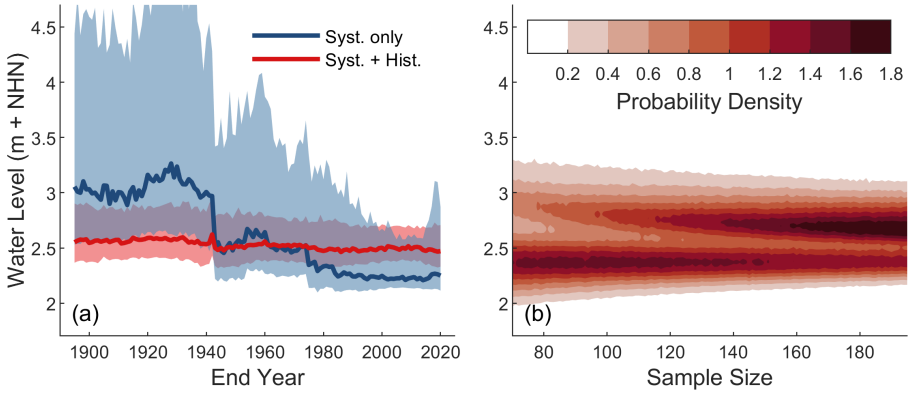

Figure 6.6: **(a)** Maximum likelihood estimates and 95% credibility intervals of HW200 at Travemünde, estimated using a 70 year moving window of AMAX values, both omitting (blue) and including (red) historical information. **(b)** Probability density estimates of HW200 made using bootstrap samples of AMAX data of different sizes.

Considering only the systematic data in analysis (a), there is a clear downward trend in the maximum likelihood estimates of HW200, and general decrease in the range of 95% credibility intervals. The largest change occurs in 1943, when the AMAX window no longer includes the exceptionally large 1872 event. The largest HW200 estimate of 3.21 m occurs for the period of 1858 to 1927, with the smallest estimate of 2.22 m occurring for the period of 1943 to 2012, a difference of approximately 1 m. In contrast, HW200 estimates made using systematic data in combination with historical information differ by a maximum of only 14 cm. The large variability of HW200 estimates made using systematic data alone, and by contrast the generally stable estimates made with historical information included, is also reflected in the range of 95% credibility intervals. Before 1943, the difference between upper and lower bounds of the 95% credibility intervals range from 1.70 m in 1907 to 2.63 m in 1929, for estimates made using systematic data only. Afterwards, this range generally decreases to a minimum of 0.49 m in 2008. For estimates made using both systematic data and historical information, these ranges are more stable, decreasing somewhat steadily from a maximum of 0.58 m in 1896 to 0.32 m in 2010.

For the second analysis (b), the resulting probability density contours of the HW200 estimates (Figure 4b) show two clear regions of high probability density. The lower region exists at a HW200 value of approximately 2.4 m and is dominant for sample sizes < 140. At higher sample sizes, the upper region becomes dominant and corresponds to the HW200 estimate of 2.66m derived using the full systematic AMAX record. The upper region of high probability is clearly a

result of the presence of the exceptionally large 1872 event in the data. When the 1872 event is removed from the analysis (results not shown), the upper region disappears entirely, and the probability density of the lower region grows substantially.

6.5 DISCUSSION

Estimates of current ESLs can be dramatically improved by incorporating historical information in EVA. For the German Baltic coast, estimates made using limited data from high resolution tide-gauges in combination with historical information show good agreement with estimates made using much longer AMAX records. Furthermore, we show that even long records of systematic data can benefit from the inclusion of historical information in terms of reduced estimate uncertainties. These results support the conclusions of several studies which consider the benefits of including historical information in EVA (Bulteau et al., 2015; Payrastre et al., 2011; Benito et al., 2004; Hamdi et al., 2015; Coles and Tawn, 2005; Reis and Stedinger, 2005; Coles et al., 2003).

At all sites tested along the German Baltic coast, combining historical information with tide-gauge records results in large increases to the estimates of ESLs. Of the seven sites tested, four benefit in terms of reduced estimate uncertainties. At all sites, including those where uncertainties increase, the proportion of uncertainty in the upper bound decreases. This is to be expected given that historical measurements provide valuable information in the upper range.

The results presented in this study raise concerns regarding the estimates of ESLs made using limited data. It is generally accepted that estimative return periods should not exceed four times the length of the record considered (Pugh, 2004), in order to maintain manageable uncertainties. Despite this, we show that even long records can lead to severely underestimated ESLs despite fulfilling this criterion. Indeed, at most sites tested in this study, incorporating historical information results in HW200 estimates that remain within the upper uncertainty bounds of the initial estimates, highlighting the importance of quantifying uncertainties in the first place. However, at Flensburg and Schleswig, updated HW200 estimates exceed the 95% credibility intervals computed using systematic data only.

For the case of Schleswig, comparisons between events recorded at the two nearest sites of Flensburg and Eckernförde show that within the tide-gauge data, ESL events are typically much smaller.

However, these large differences are not reflected in the historical record. Changes in the local bathymetry, including development of the sand spit at the mouth of the Schlei inlet, likely explains the reduced size of ESLs at Schleswig between the two records. Consequently, the use of historical information in EVA at this site would not be applicable, as the systematic and historical ESLs would be drawn from different distributions. ESL estimates provided in this paper for Schleswig should thus be approached with caution, and not considered accurate or reliable.

Like Schleswig, Flensburg lies at the end of an inlet connected to the Baltic Sea, however the Flensburg Firth is much larger than the narrow Schlei. Given the larger size and lack of development at the mouth of the Flensburg Firth, it is unlikely that changes in the local bathymetry would result in significant changes in the attenuation of ESLs. Instead, it appears that the available systematic data is simply insufficient for the estimation of large ESLs, such as HW200. Given the tide-gauge record at Flensburg is one of the longest available along the German Baltic coast, spanning 66 years, and ESLs in the region are influenced by the same large-scale atmospheric forcing, we must question the validity of HW200 estimates along the entire coastline that rely solely on tide-gauge data. This finding is supported by the results at Travemünde, where large changes occur due to the inclusion of historical information, despite a relatively long tide-gauge record.

Long-range dependence in the series of AMAX ESLs at Travemünde may explain the underestimation of ESLs when only limited tide-gauge data is used. The discovery of persistence in the data suggests the presence of some long-term variability. While the cause of this variability is not known and considered outside the scope of this paper, it would suggest that available tide-gauge records along the German Baltic coast cover a period of relatively low ESL activity. In such a case, direct EVA methods which rely on high-resolution tide-gauge data, in addition to JPM and RFA, would be prone to underestimations due to bias within the available data.

Based on the extensive AMAX record at Travemünde, a period of high ESL activity took place prior to 1927 and has since been followed by a period of lower extremes. Further research is necessary to establish if this trend is the result of permanent changes in the mechanisms driving ESLs in the region, or simply some long-term variability not yet defined. However, research conducted by Jensen and Töppe (1990) and Jensen et al. (2022), in combination with the high Hurst exponent found in this study suggest the latter. The occurrence of large events before the introduction of systematic records,

not equal in height to 1872 but of comparable magnitude, argue against the notion that the decrease in ESL activity is permanent. Furthermore, we show that the inclusion of historical information in EVA leads to stable estimates of ESLs regardless of the systematic period analysed. Thus, we believe that some long-term variability in ESLs along the German Baltic coast is responsible for the apparent low ESL activity over the past century. If the current phase of low ESL activity is indeed the result of long-term variability, this has significant implications for the management of coastal areas along the German Baltic coast given the lack of long sea level records. In such a case, the method described herein is a valuable tool where historical information is available.

The largest influence on the results of this study is the event of November 1872, which caused exceptionally high water levels along much of the German Baltic coast. Differences in estimates produced using systematic tide-gauge data and AMAX records are largely due to this event. In fact, 1872 was the only event considered in combination with tide-gauge data at three of the seven sites tested, however changes to the maximum likelihood estimates due to its inclusion are not indifferent to results at other sites where multiple historical events are considered. Due to the exceptional magnitude of the event, it has until recently been treated as an outlier and thus excluded from statistical analyses. Hofstede and Hamann (2022) argue that based on the series of AMAX ESLs at Travemünde, 1872 is indeed an outlier given that it is more than 50% higher than the second largest event. Indeed, Mudersbach and Jensen (2009) assessed the return period of a corrected 1872 event at about 10,000 years. However, they concluded that the event could not be well defined statistically given the limited sample population, and suggested extending the available data using historical information. Jensen et al. (2022) highlight the occurrence of events within historical records of similar magnitudes to 1872. Given these events, Jensen et al. (2022) argue that 1872 should not be considered an outlier and that the systematic records are not sufficiently long to deal with events of these magnitudes.

When considering the full historical record at Travemünde in combination with high-resolution tide-gauge data, we assess the return period of the 1872 event to be approximately 1,900 years. Combining the long AMAX record with historical information provides a return period estimate of approximately 5,800 years. Given the length of the historical record at Travemünde (approximately 980 years), and the occurrence of other large ESLs within it (1320, 1625, 1694), we agree with the arguments of Jensen et al. (2022) that the 1872 event should not be considered an outlier, but rather an exceptional realisation of

the underlying ESL distribution, and we would recommend for its use in EVA. While sea level records that cover a period that includes 1872 can provide very good ESL estimates using traditional EVA methods, only few sufficiently long records exist along the German Baltic coast (including the AMAX records described herein). Despite this, we show that even short tide-gauge records (approximately 30 years in our case) with one measurement of 1872 can provide similar results. Therefore, reconstructions of past extremes offer valuable information with which to improve EVA.

Large differences exist in the estimates of ESLs made using either the POT-GP or AMAX-GEV approaches. While the incorporation of historical information reduces these differences, it does not provide any insight into which method performs best. Indeed, the POT-GP approach is generally preferred in the literature (Arns et al., 2013; Wahl et al., 2017), but this does not necessarily apply to the case of the German Baltic Sea coast. We find that when both methods are constrained to the same record length (see Appendix A), the POT-GP method generally performs better with lower uncertainties at the distribution tails. At all sites, the AMAX-GEV provides larger ESL estimates at high return periods. Interestingly, including historical information in the analysis produces a different result, with the AMAX-GEV analysis providing lower uncertainties at high return periods. One possible explanation for this involves the sampling threshold for the POT method. We assume that this threshold is constant for the full duration of historical and systematic observation, following the study by Bulteau et al. (2015), but this may not be the case. Indeed, large differences in results due to the inclusion of historical information suggests this assumption may be false. Thus, an advantage of the AMAX-GEV approach is that no sampling threshold is required. Given a single sea level record with no historical information, we would recommend the POT-GP approach over the AMAX-GEV due to the reduced estimate uncertainties. However, the AMAX-GEV approach may provide more precise results when historical information is available. Where a longer AMAX record is available, such as in this study, the AMAX-GEV approach provides clearly better results due to the increased data.

We confirm that exceptionally large events commonly considered to be outliers can be placed within classical EVA using Bayesian techniques (Bulteau et al., 2015; Coles and Tawn, 2005; Reis and Stedinger, 2005), and demonstrate this for the case of the 1872 event along the German Baltic coast. Much discussion surrounding the suitability of 1872 and other supposed outliers in statistical analyses exists (Jensen et al., 2022; Hofstede and Hamann, 2022; MacPherson et al., 2019; Hamdi et al., 2015). While scientific and coastal man-

agement perspectives on its importance may differ, arguments tend to revolve around the consequences of such an event occurring or not. That is, what are the costs on one hand to implement coastal defences for an event that might not occur again, and on the other hand, what are the potential damages. Due to the exceptional height of the 1872 ESL event, we agree with (Jensen et al., 2022) that design water levels do not necessarily have to be based on the largest past event, but all available relevant information should be used for their derivation.

There exists a gap between the design heights of coastal defences and what is actually possible. The German Baltic coast, and Travemünde in particular, is a good example of this fact. The 1872 event produced exceptionally high water levels as a result of a rare sequence of atmospheric forcing. Design water levels in the region, based on a return period of 200 years, are thus much lower than what was observed during this event. Even after including the 1872 measurements, estimates of HW200 remain well below the level of the 1872 ESL. Although this can be addressed by raising design heights to that of the largest event on record, this would require significantly large investments which may be of better use elsewhere. Furthermore, this strategy of coastal defence planning is reactive rather than proactive. Instead, integrated flood risk management should be implemented with an aim to reduce vulnerability to coastal floods via appropriate urban planning based on current and future ESLs, and investment into advanced warning systems, evacuation planning and emergency response networks in exposed regions.

6.6 CONCLUSIONS

In this study, we combine systematic tide-gauge data with historical information using a Bayesian MCMC method to assess ESLs at seven sites along the German Baltic coast. In general, we find the inclusion of historical information in EVA is beneficial, resulting in reduced estimate uncertainties and the incorporation of exceptionally large events that would otherwise be considered outliers.

At the German Baltic coast, the incorporation of historical information in EVA results in large increases in the estimates of ESLs. At all sites, even those with long systematic records, (Flensburg: 66 years, Travemünde: 71 years, Wismar: 63 years and Warnemünde: 67 years), estimates for HW200 increase by between 16-28%. We find the presence of long-range dependence in the series of ESLs at Travemünde, which suggests that long-term variability is affecting

extremes in the region. Further, we show that recent ESL activity has been relatively low. As this period of low ESL activity covers the full period of systematic observation, we caution against the use of systematic data alone when assessing ESLs in the region.

Estimates of ESLs in the region are largely affected by an exceptional event that took place in 1872. Although it is commonly dismissed as a statistical outlier, we argue against this notion and advocate for its inclusion, partly due to its substantial impact on estimates of ESLs. Where available, incorporating the 1872 event into analyses of ESLs based on limited systematic data provides similar results to analyses which employ much longer sea level records. Sourcing further information on this event may allow for improved ESL estimates at other sites where only limited systematic data is available. Such information may be found in the field, by consulting historical records or by conducting re-analyses of the event using numerical models.

APPENDIX A: ESL SAMPING

In this study, we employ two different approaches for sampling ESLs. The first technique, POT is generally preferred in literature for reasons explained in Section 2.2. Wahl et al. (2017) also note that AMAX sampling may result in larger uncertainties at the tails of the distribution when sea level records are short. While we are constrained by the records sourced from MLUV (2009), which are only available as AMAX samples, this is not the case for the high-resolution tide-gauge data. Thus, we decided to employ POT sampling for these records due to their preference in literature, and so that we may demonstrate the use of the Bayesian MCMC EVA method for both POT-GP and AMAX-GEV approaches. To directly compare the two approaches, we also performed an AMAX-GEV analysis using the high-resolution tide-gauge data. Figure 6.A1 shows ESL estimates including 95% credibility intervals estimated at each site using the high-resolution tide-gauge data only. At all sites, the AMAX-GEV method results in larger estimates of ESLs at high return periods. Also of note are the larger uncertainties at the tails of the distribution at all sites except for Warnemünde, which supports the findings of Wahl et al. (2017). In general, the POT-GP appears to produce more reliable results given the same record duration based purely on the reduced uncertainties.

Next, we considered how these estimates are affected by the addition of historical information. We performed the analysis again, but in-

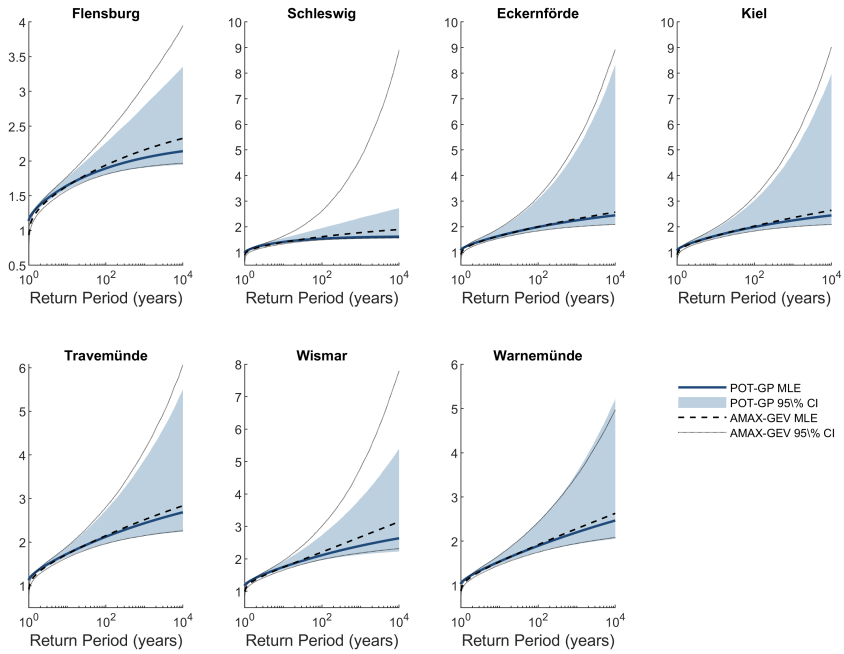

Figure 6.A1: Comparison of POT-GP and AMAX-GEV approaches to the Bayesian MCMC method for estimating ESLs. At each site, estimates of ESLs are made using high-resolution tide-gauge data only. In general, the AMAX-GEV approach results in higher estimates of ESLs at high return periods and larger uncertainties at the tails of the distributions.

cluded historical information with measurement uncertainties given by Jensen et al. (2022). Results are shown in Figure 6.A2. As with the first analysis, both the POT and AMAX samples are taken from the same high-resolution tide-gauge record. For both POT-GP and AMAX-GEV analyses, the introduction of historical information is beneficial in terms of reduced estimate uncertainties. Interestingly, we find that the AMAX-GEV approach performs better in terms of reduced uncertainties at all sites except Schleswig. Differences in the maximum likelihood estimates between the two analyses are much reduced.

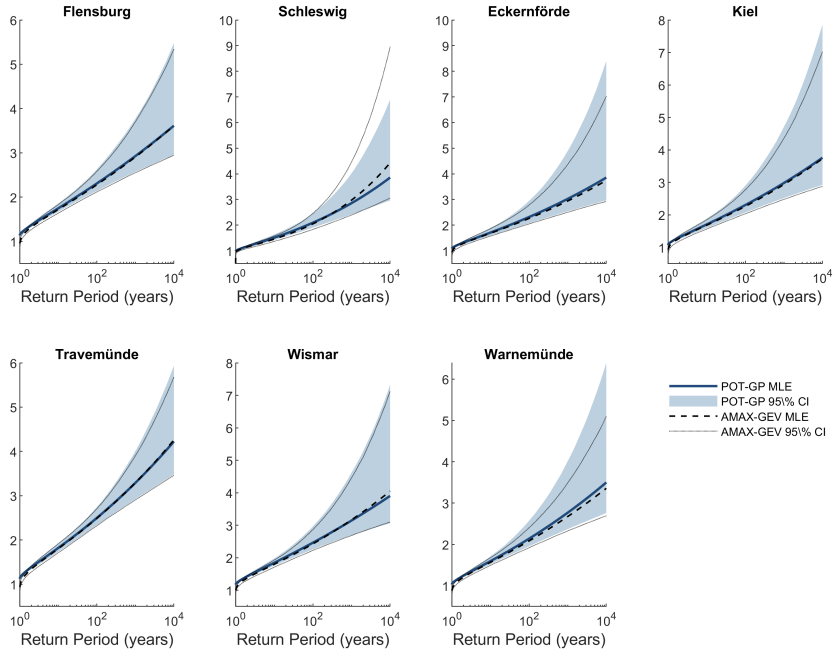

Figure 6.A2: Comparison of POT-GP and AMAX-GEV approaches to the Bayesian MCMC method for estimating ESLs including historical information. At each site, estimates of ESLs are made using the same high-resolution tide-gauge record in combination with historical information. In contrast to results where historical information is omitted, the AMAX-GEV approach performs somewhat better than the POT-GP approach in terms of reduced uncertainties at the distribution tails. Differences in the maximum likelihood estimates between the two methods are much reduced.

7

A COMPREHENSIVE PROBABILISTIC FLOOD ASSESSMENT ACCOUNTING FOR HYDROGRAPH VARIABILITY OF EXTREME SEA LEVEL EVENTS

This chapter is published as:

Kupfer, S., MacPherson, L. R., Hinkel, J., Arns, A., and Vafeidis, A. T. (2024). A Comprehensive Probabilistic Flood Assessment Accounting for Hydrograph Variability of ESL Events. *Journal of Geophysical Research: Oceans*, 129, e2023JC019886. DOI: 10.1029/2023JC019886..

ABSTRACT

Flood characteristics caused by extreme sea level (ESL) events depend largely on the magnitude of peak water levels and their temporal evolution. However, coastal flood risk is generally assessed based on only a limited number of potential peak water levels and a selection of past events or a design hydrograph. We address this gap and systematically estimate (a) spatial annual and (b) event-based flood probabilities by comprehensively accounting for both a wide range of peak ESLs and their temporal evolution, herein referred to as hydrograph intensity. We simulate flooding at the German Baltic Sea coast with the hydrodynamic model Delft3D. We produce probabilistic flood maps, which detail flood exposed areas together with annual probability of flooding. Additionally, we show how the flood extent changes, when accounting for upper, median, and lower quantiles of hydrograph intensities. Our results demonstrate that the relevance of the intensity is site and ESL dependent. While flood extents of some ESLs of the upper and lower intensity bounds indicate no differences, others differ by up to 45%. Further, we consider two ESLs (2.24 and 2.55 m) and simulate 100 intensities for each. Compared to intensity quantiles, this results in flood extents of up to 60% difference. Hence, we find that quantiles of intensity do not cover the full range when addressing uncertainty due to hydrograph variability. We, therefore, recommend accounting for a wide range of hydrograph intensities in addition to using a wide range of ESLs in future flood risk assessments.

7.1 INTRODUCTION

Flooding from extreme sea levels (ESLs) is one of the greatest threats that low-lying coastal areas are facing in the 21st century (IPCC, 2022; Tebaldi et al., 2021). These areas are densely populated which exacerbates their exposure to flooding (Kummu et al., 2016; MacManus et al., 2021; Reimann et al., 2023). An estimated 148 million people and total assets worth $US7.76 trillion are currently exposed to a 1-in-100-year ESL event globally (Kirezci et al., 2020). With climate change induced accelerating sea level rise (SLR) (Nerem et al., 2018; Slangen et al., 2022; Vitousek et al., 2017) the frequency of ESLs and, as a direct result, coastal flooding is expected to increase (Vousdoukas et al., 2018b). Without adaptation, annual global flood losses are expected to grow from approximately US$6 billion in 2005 to US$1 trillion or more by 2050 (Hallegatte et al., 2013). For much of the world's developed coastline, the cost of inaction outweighs the cost of adaptation (Hinkel et al., 2014). With finite resources to implement and maintain defenses against coastal flooding, it is important that coastal protection agencies are well informed about current and future flood exposure.

Flood exposure is typically determined using inundation models, and has been analyzed at global (Brown et al., 2018; Muis et al., 2017; Ward et al., 2013), regional (Hauer et al., 2021; Muis et al., 2015; Vousdoukas et al., 2018a) and local scales (Hallegatte et al., 2011; Höffken et al., 2020; Kumbier et al., 2018). While global flood assessments typically use simple static GIS-based models, local and regional flooding is often simulated using hydrodynamic numerical models (Lewis et al., 2013; Vousdoukas et al., 2018a). Static approaches have been shown to overestimate flood extent (Didier et al., 2019; Vousdoukas et al., 2016) as they do not account for flood propagation processes, including flood duration and the temporal evolution of the forcing ESL event (Bates et al., 2005; Ramirez et al., 2016). Hydrodynamic numerical models account for physical processes (i.e., atmospheric-ocean-land interactions) as well as the temporal variability of ESL events (Ramirez et al., 2016). While these models have been widely employed in flood inundation modelling, uncertainties in the estimates of flood exposure still exist.

One major source of uncertainty surrounds the estimation of return water levels, which according to (Rohmer et al., 2021) is most relevant for the short-term (before 2040). Return water levels describe not only the magnitude of the forcing ESL event but also their specific probability of exceedance. Return water levels are used extensively when implementing coastal defenses (see MELUR, 2022), and many

coastal protection measures are designed to withstand ESLs up to a specific return water level. These design heights are typically computed using extreme value analysis based on a sample of observed ESL events (Coles, 2001). Although extreme value statistics have been widely employed for the study of natural disasters, and best practices have been suggested for the analysis of ESLs (Arns et al., 2013; Haigh et al., 2010), uncertainties remain (Dangendorf et al., 2016).

Another major source of uncertainty in flood exposure arises due to the fact that no two ESL events are alike, and may have very different temporal evolutions of the water level curve (referred to hereafter as a hydrograph). Despite this, most regional-to-local flood assessments use a single standardized design hydrograph, for each return level (e.g., Dawson et al., 2005; Wadey et al., 2015), without accounting for their variability. In fact, studies have shown that ESLs of the same return level but with different hydrographs lead to substantially different flood exposure (Quinn et al., 2014). Santamaria-Aguilar et al. (2017), for example, show that such variability can result in a threefold increase of estimated overflow volumes. But only few coastal flooding studies consider the variability of hydrographs (Höffken et al., 2020; Quinn et al., 2014; Santamaria-Aguilar et al., 2017), and those quantify their effects in different ways but for only a maximum of five hydrograph shapes. The reasons for this are that first, this uncertainty cannot be quantified using static models as these only account for peak water levels. Moreover, where hydrodynamic models are employed, sourcing a sufficiently high number of ESL events of equal magnitude for comparison is difficult, due to the rarity of the events and lack of long sea level records. In the study of Dullaart et al. (2023), the authors present their developed method HGRAPHER, which provides location-specific hydrographs, based on storm surge and tidal time series, along the global coastline. Another solution presented by Höffken et al. (2020) is to make use of a stochastic ESL event generator, which simulates ESL events based on measured hydrographs at a site of interest (MacPherson et al., 2019; Wahl et al., 2011). Here, a large range of ESLs of a certain return level may be generated, each with a different hydrograph. To quantify the uncertainty of hydrograph variability, Wahl et al. (2011) define hydrograph intensity, which describes the fullness of the hydrograph, or the area beneath the water level curve. This includes the duration of the storm event, in combination with the shape of the hydrograph.

A common method in flood exposure or risk assessments is to express uncertainties in terms of probability. To quantify uncertainty in flood extent, it is recommended to account for the full range of event return periods from high-probability low-magnitude events, to low-probability high-magnitude events (Hinkel et al., 2015; Merz

and Thieken, 2009; Ward et al., 2011). However, in flood exposure or risk assessments, it is still common practice to use a limited number of ESL events and SLR scenarios (Muis et al., 2016; Wadey et al., 2012), and only few studies quantify uncertainties by accounting for the full range of ESL return periods. One example is Hauer et al. (2021), who employ a static modeling approach to estimate the annual probability of exposed population to flooding along the US coast. To our knowledge however, due to the high computational cost no study has systematically mapped flood probabilities using a full range of ESL magnitudes, in order to quantify the current likelihood of flooding. Moreover, no study has additionally estimated the effect of a large number of hydrograph intensities on flood extents.

In this study, we generate probabilistic flood maps for the city of Lübeck on the German Baltic Sea coast: first, by estimating which areas are exposed to flooding from ESLs of 37 different return periods; and second, by analyzing how flood exposure is affected by hydrograph variability. To achieve these objectives, we use hydrodynamic modeling and conduct 311 simulations in total.

7.2 STUDY AREA

The city of Lübeck is the second largest city of Germany's northernmost state Schleswig-Holstein. It is located in the southeast of Schleswig-Holstein, at the Lübeck Bight of the Baltic Sea. Most of the region's coastal areas are low-lying. Greater Lübeck mainly consists of its historic city center, surrounding industrial areas, and the largest district of Lübeck, Travemünde, which developed along the Trave River mouth. The historic city center is located on an island, located 16 km inland from the coast. The island is surrounded by the Trave River, which connects the city center with the Baltic Sea. In 2022 the city of Lübeck counted 221,225 inhabitants (Hansestadt Lübeck, 2022). The historic city center of Lübeck has outstanding cultural and touristic value and belongs to the UNESCO world heritage (UNESCO WHC, 2022), which may lead to challenges in managing and planning flood protection measures (Orr et al., 2021; Phillips, 2014; Reimann et al., 2018). Travemünde is directly located at the coastline and has banks along the Trave River. For this reason, a high number of hotels and other touristic infrastructure has developed in Travemünde, particularly along the northern shore. The Priwall peninsula on the southern shoreline is populated by holiday apartments, a marina, and associated infrastructure. Between the historic city island of Lübeck and Travemünde, large industrial as well as agricultural areas surround the Trave.

The Baltic Sea is a semi-enclosed system, connected to the North Sea via the narrow Kattegat. As a result, tides in the region are small (0.1–0.2 m) (Jensen and Müller-Navarra, 2008). Due to its location in the Baltic Sea, ESLs along the German Baltic coast are driven mainly by strong north-easterly winds. However, seiches acting over the entire Baltic Sea can contribute significantly to ESLs in the region (Jensen and Müller-Navarra, 2008). In comparison with the German North Sea coast, ESLs along the German Baltic coast are less common and generally of lower magnitudes, however, this fact is accompanied by a significantly lower level of coastal protection (Sterr, 2008).

The region has experienced several flooding events in the past, most recently in 2017 and 2019 with peak water levels of 1.74 and 1.76 m, respectively (BSH, 2019). These events resulted in minor flooding in Travemünde and Lübeck's historic city center. The largest recorded coastal flood event, however, occurred in 1872, reaching a maximum height of approximately 3.4 m above mean sea levels (MSLs) (Jensen et al., 2022). Of the total 2,110 km of coastline of the German Baltic Sea coast only 13% is protected by first-grade dikes and 9% is protected by dunes. Despite this, approximately 497,000 people live below 5 m above MSL, and 48 billion Euros of assets are located in this low-lying region (Sterr, 2008). The situation is representative of that at Lübeck and Travemünde, where no engineered flood protection measures exist (Hofstede, 2008). Sand dunes along the beach at Travemünde are the only coastal protection in the region.

7.3 DATA AND METHODS

7.3.1 *Data*

Our methods, that is, estimation of return periods, hydrograph generation, and hydrodynamic modeling require a large volume of data, such as water level observations, topography, and bathymetry, land cover data serving as bottom roughness, boundary conditions (e.g., input water levels, or tidal data), and observations for model calibration and validation. The data employed in this study are summarized in Table 7.1.

Table 7.1: Data Used for Hydrodynamic Model Input

Data Set	Horizontal resolution	Temporal resolution	Vertical accuracy	Application	Reference
Bathymetry	50 m	-	-	Model input topography	BSH (2023)
Elevation	1 m	-	<15-30 cm	Model input topography	LvermGEO (2022)
Land Cover	100 m	-	-	Manning coefficient	CLC (2018)
Manning Coefficient	-	-	-	Bottom roughness	Several references[a,b,c]
Water level (Travemünde)	-	Hourly	-	Boundary conditions for model validation and calibration	WSA (2022)
Water level (Lübeck)	-	Hourly	-	Observations for calibration and validation	WSA (2022)
ESL hydrographs	-	Hourly	-	Hydrographs for boundary conditions	MacPherson et al. (2019)
Annual Maxima water level (Travemünde)	-	Annual	-	EVA	MacPherson et al. (2023)
Monthly MSL (Travemünde)	-	Monthly	-	EVA	PSMSL (2018)

7.3.2 *Extreme Sea Levels*

Extreme sea levels occur as episodic increases in water levels driven by shoreward wind-driven water circulation and atmospheric pressure (Muis et al., 2016). They are commonly referred to by their return period, which describes an average period between events of equal magnitude (Pugh, 2004), typically in years. Return periods are the inverse of the exceedance probability, and thus define the likelihood that a specific threshold, called the return level, will be exceeded. Defining ESLs by their probability is useful in flood risk analyses, due to the relationship between risk and probability.

To force our inundation model, we produced ESL hydrographs using the method developed by MacPherson et al. (2019). They stochastically generate artificial events based on observations from the tide-gauge record at Travemünde. Here, ESL events are identified within the sea level record, and a parameterization scheme is used to characterize the individual events. Interdependencies between these parameters are modeled using a Gaussian copula, and artificial events are generated via Monte Carlo simulations. Thus, large numbers of artificial ESLs may be generated for specific return water levels, preserving the characteristics of the observed events. This allows us to consider not only the peak water level of ESLs at Travemünde but also the temporal evolution of those events.

Exceedance probabilities for the simulated ESLs are estimated using best practices recommended by Arns et al. (2013). Observations of annual maxima (AMAX) sea levels at Travemünde are modeled using the Generalized Extreme Value (GEV) distribution. First, the data is detrended using monthly mean sea levels from the Permanent Service for Mean Sea Level (PSMSL, 2018). We must point out that this data does not cover the full extent of the AMAX observations, missing the first 30 years from 1826 to 1856. Here, the MSL signal was assumed by extrapolating the PSMSL data linearly. Next, the AMAX values were adjusted to reflect current sea levels by adding the mean value of water levels taken from the Travemünde tide gauge over the final available year of data. While this does not account for interannual variability, this value (4.9 cm) corresponds well to the mean value of water levels over the final 20 years of tide gauge data (4.79 cm) and is thus a good representation of current MSL. Lastly, parameters of a GEV model were derived using maximum likelihood estimation. Probabilities of the tested ESLs were then inferred from this probability distribution. It should be noted that these values differ from the official values used as design heights for coastal defenses. These differences are due to differences in data and methods used. However, official values are available only for a few ESLs (MELUR, 2022), and are thus inappropriate for use in our study.

7.3.3 *Hydrodynamic Model Setup*

We simulate flooding from ESLs using the hydrodynamic model Delft3D for the larger area of the lower Trave, including the city of Lübeck with its largest district Travemünde (see Figure 7.1). We apply the FLOW-module of the fully integrated open-source hydrodynamic model Delft3D (Lesser et al., 2004), which has been extensively used in coastal applications (Kumbier et al., 2018; Luijendijk et al., 2017; Lyddon et al., 2019). FLOW calculates non-steady flow and transport phenomena that result from tidal and meteorological forcing.

Simulations are performed on a rectangular grid with a 10 m grid cell size in a depth-averaged (2D) mode. The 2D mode has been successfully applied in numerous hydrodynamic modeling studies (Bates et al., 2005; Kumbier et al., 2018; Skinner et al., 2015). Underlying topography and bathymetry data are resampled from their original horizontal resolution (Table 1) to a resolution of 10 m to match the model grid. The model domain covers an area of 305 km², reaching from south-west to north-east from the river gauge station Lübeck Moisling to the Trave mouth in Travemünde and beyond that

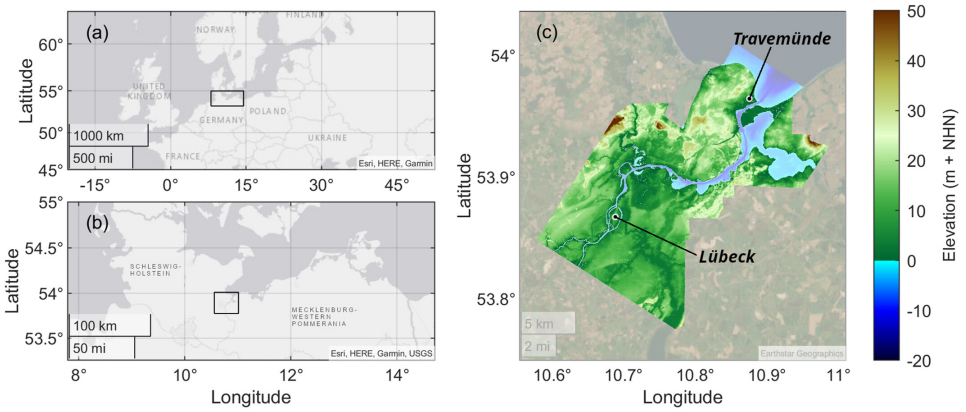

Figure 7.1: **(a)** Location of the western Baltic Sea coast in Europe, **(b)** location of our study site along the German Baltic Sea coast, and **(c)** our study site of greater Lübeck. The topographic map in panel **(c)** shows the model domain and locations of Lübeck and Travemünde are specified.

into the Baltic Sea. The model contains two open boundaries, the Baltic Sea boundary and the Trave boundary. The Baltic Sea boundary faces north-east, which is also the main direction of which ESLs are generated at Lübeck (Jensen and Müller-Navarra, 2008). We force the model at this open boundary using hourly measured water level data from the tide gauge station at Travemünde, for the calibration and validation (Appendix A). For the probabilistic flood analysis, we use synthetic hydrographs as boundary conditions, which are generated using the method of MacPherson et al. (2019), described in Section 3.2. The Trave boundary was placed 5 km southwest of the historic city center at the south-western edge of the model domain. This boundary is simply kept open to allow water outflow and to prevent water accumulation upstream.

We specify spatially varying manning bottom roughness values based on land cover data (Table 1) via literature review and from a sensitivity analysis which we perform during the model calibration process (Table 7.A2, Appendix A). Physical and numerical parameters are tested during the model calibration and verified during model validation. We calibrate the model by changing physical and numerical parameters in individual simulations of a storm surge event, occurring in October 1995, obtained from the tide gauge Travemünde. We validate the model with the best fitting parameter set-up derived from the model calibration using a second storm surge event, which occurred in January 2017. A detailed description of the calibration and validation processes can be found in Appendix A.

7.3.4 *Probabilistic Flood Maps*

To assess the potential flood extents due to (a) ESL magnitudes and (b) hydrograph intensities at greater Lübeck, we conduct a large number of model runs on a high-performance computer using an automated procedure for the model set-up, hydrograph implementation, as well as flood extent and flood depth extraction. The combination of high-performance computing, automated model setup, and data extraction makes the simulation of many events feasible. For each simulation, the open boundary near Travemünde is forced with hydrographs generated using the stochastic ESL model described in Section 3.2. Each simulation produces hourly water levels at all pixels in the model domain.

To create the probabilistic flood maps, we simulate 111 different ESL events ranging in magnitude from 1.6 to 3.4 m in 5 cm increments. The lower limit of 1.6 m corresponds to a return period of approximately 4 years (HW4) at Travemünde and was chosen as an arbitrary lower limit to ensure flooding under small ESLs is considered. The upper limit of 3.4 m has a return period of approximately 3,650 years (HW3650), and was chosen as it represents the highest ESL recorded at Travemünde, which occurred in November of 1872. We chose an increment of 5 cm as a compromise between high resolution results and a reduced number of simulations. To account for differences in flood extent due to hydrograph intensity, at each tested magnitude, three hydrographs are selected which represent the 5^{th}, 50^{th} (median), and 95^{th} percentile of intensity values. We define hydrograph intensity according to MacPherson et al. (2019) and Wahl et al. (2011), who describe intensity as the area below the water level curve, above a specified threshold θ. This threshold is not fixed and is simply applied as a diagnostic. It can be adjusted depending on the research objective and case study. In this case, we set the threshold to 0, being equal to the local vertical datum Normalhöhennull (NHN), approximately 5 cm below current MSL. Changes in this threshold were examined with no significant changes to the results represented herein. Figure 7.2 demonstrates examples of hydrographs at each intensity and our used terminology. We use the parameter of hydrograph intensity as a measure to estimate flood potential from selected ESL magnitudes, but we do not consider it as the only relevant parameter causing flood extent variability, as other factors additionally appear to affect flood extents. However, exploring this in detail is beyond the scope of our study. These hydrographs are obtained by generating 100 events at each magnitude, sorted according to intensity. As the probability distribution of hydrograph intensities is handled within the stochastic ESL model used

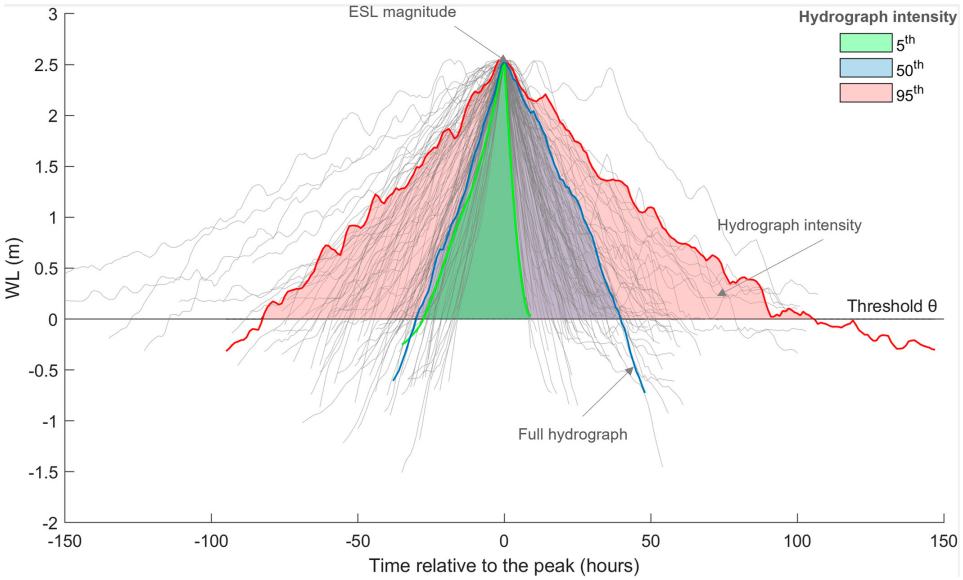

Figure 7.2: Examples of hydrographs and their intensity, shown by the 5th (blue), 50th (green), and 95th (red) hydrograph intensity, above the threshold θ. Gray hydrographs represent all 100 considered hydrographs of the 200-year event (2.55 m).

to generate the events, simply selecting the "5th," "50th," and "95th" events that were generated provides us with the desired intensity values.

Flood probabilities are assigned to flooded pixels based on the lowest magnitude event at which flooding occurs. For example, at pixels that experience flooding only when ESL magnitudes are 2.10 m and greater, a flood probability of 0.0425 is assigned, which corresponds to the exceedance probability of an ESL event equal to 2.10 m. Pixels are considered flooded if the water level rises above 0 m at any time in the simulation. Areas in the model domain that are already considered water bodies (topography elevation <0 m) are disregarded. At pixels where no flooding occurs, flood probabilities are set to 0. This is done for all pixels to create probabilistic flood maps. We assess flood probabilities at each of the three tested intensity levels individually, resulting in separate probabilistic flood maps for each one of the 5th, 50th, and 95th percentile intensity levels.

To better visualize and analyze the results, for this and the following analyses, we divide the study area of greater Lübeck into two localities: (a) the historic city center of Lübeck and its closest surroundings (further referred to as "Lübeck") and (b) Travemünde. During the analysis, we consider both separately as two locations.

7.3.5 *Sensitivity Analysis of Hydrograph Intensity*

As a second step, we explore the effect of ESL intensity on flood exposure in further detail by simulating all 100 sampled ESL events at magnitudes of 2.24 and 2.55 m. The first value of 2.24 m was initially chosen as it represented the official design height (HW200) of coastal defenses at Travemünde (MELUR, 2012) with an official return period of 200 years. However, this value was recently changed to 2.55 m in 2022 (MELUR, 2022) and is thus also included in our analysis. According to MELUR (2022), an ESL of 2.24 m now corresponds to a return period of approximately 50 years.

In contrast to the first analysis, flood probabilities in this second step are based on the proportion of events under which flooding occurs, rather than the exceedance probability of the tested ESL event. The reason for this is that we are comparing flooding under events of equal magnitude, and thus equal probability. Thus, the resulting flood probabilities are conditional on the tested ESL magnitude. This differs from the first analysis where probabilities can be thought of as absolute. Here, pixels are assigned probabilities based on the percentage of events which result in flooding. For example, a pixel that experiences flooding in 25 of the 100 simulations is assigned a probability of 0.25. For comparison purposes, we additionally calculate flood extents using the bathtub method, which provides a first-order estimate of the maximum potential floodplain at the tested ESLs of 2.24 and 2.55 m. This method assumes that all areas below the height of the specified ESL and connected to the sea become flooded, and does not account for dynamic processes, such as storm duration (Ramirez et al., 2016).

7.4 RESULTS

7.4.1 *Probabilistic Flood Maps*

We produce probabilistic flood maps for Lübeck and Travemünde, which illustrate the flood extents and likelihoods at Lübeck and Travemünde due to coastal flooding. This includes events of high probability with negligible consequences to extremely rare events that cause greater flooding. Figure 7.3 shows flood probabilities at Lübeck and Travemünde taken from simulations of ESLs with median intensities. We produced this map for each of the three analyzed intensity levels, but only show the median intensity, as

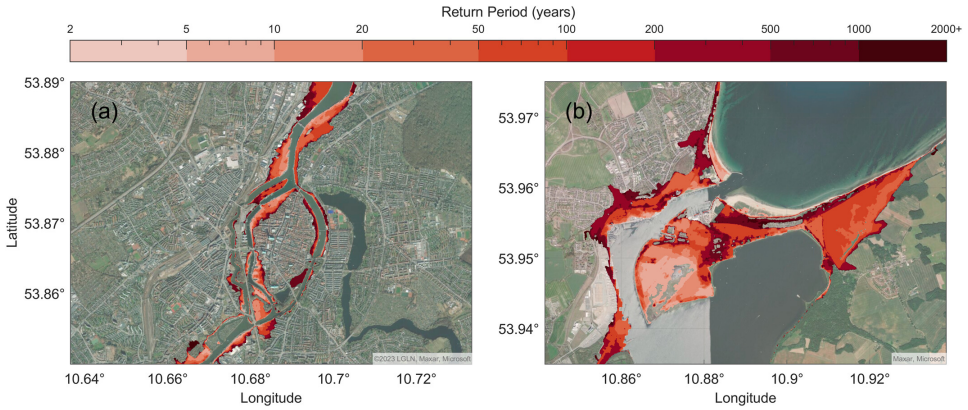

Figure 7.3: Probabilistic flood maps of **(a)** Lübeck and **(b)** Travemünde, taken from simulations of extreme sea levels with median intensities. Lighter shades show areas flooded from low return periods (high probability) and darker shades show areas of high return period flooding (low probability).

differences are difficult to recognize. Moreover, to better visualize the results, we consider only the localities of Lübeck and Travemünde, rather than the entire model domain. Both Lübeck and Travemünde experience large variability in flood extents resulting from different return water levels, with maximum differences of 1.37 and 4.58 km², respectively.

Total flood extent under each tested ESL and intensity is shown in Figure 7.4. It should be noted that the return periods given are calculated using data and methods described in Section 3.2, rather than taken from official sources. This was done as official values from MELUR (2022) are only available for a few discrete return water levels. Thus, some differences in ESL values exist due to differences in data and methods used to derive them.

The increase in flood extent at Lübeck (Figure 7.4a) accelerates once water levels exceed HW10 (1.85 m) and flood extents expand somewhat linearly, albeit at a lower rate after HW200 (2.65 m). Differences in flood extent resulting from simulations with different hydrograph intensity quantiles are small. At Travemünde (Figure 7.4b), flood extents increase linearly with water levels until approximately HW50 (2.30 m), where higher water levels lead to larger flooding thereafter. This continues until approximately HW500 (2.85 m), where afterward flood extends at a slower rate. Unlike at Lübeck, hydrograph intensities have a more apparent effect at Travemünde, especially for ESLs that exceed HW50 up to HW1000 (2.30–3.05 m). Indeed, flood extent at 2.3 m (∼HW50) for the 95th percentile event exceeds that of the median intensity event 10 cm higher (2.4 m ∼ HW80).

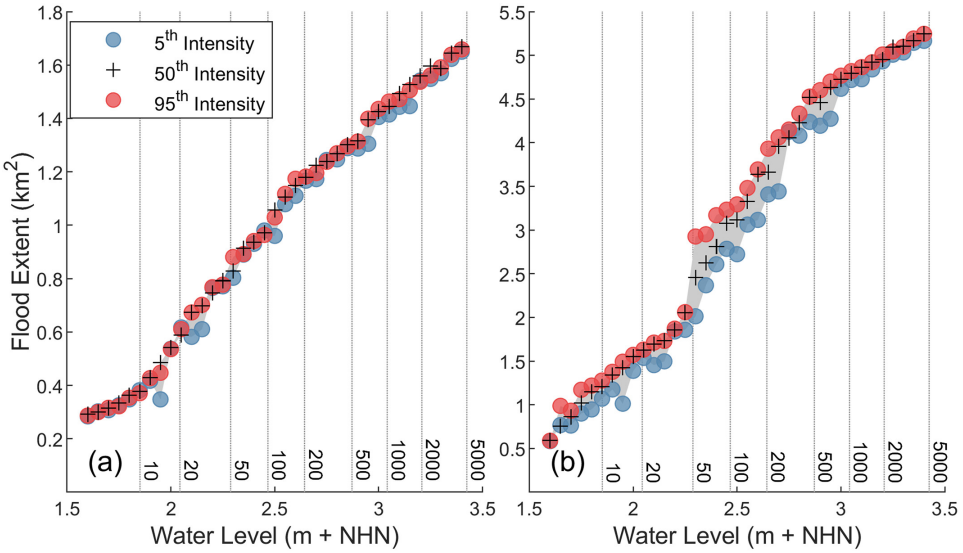

Figure 7.4: Flood extent at **(a)** Lübeck and **(b)** Travemünde at all return water levels and including the 5^{th} (blue circle), 50^{th} (black plus), and 95^{th} (red circle) percentiles of hydrograph intensity. The gray shading demonstrates the range of flood extent differences caused by intensity.

A closer comparison of differences in flood extent associated with hydrograph intensity is shown in Figure 7.5 for both Lübeck and Travemünde. As mentioned above, hydrograph intensity appears to have a clearer influence on flood extent at Travemünde than at Lübeck. The large and positive differences between the 95^{th} and the 5^{th} percentile at Travemünde suggest that large intensities tend to lead to larger flood extents. This is particularly true for ESLs between 2.3 and 2.7 m. The overall larger positive differences between the median and tsntensity (Figure 7.5d) compared to the $95^{th}-50^{th}$ differences (Figure 7.5e) suggest a more pronounced influence at lower intensity ranges. At Lübeck however, differences in flood extent are typically small, compared to differences at Travemünde. Even though overall larger intensities are associated with larger flood extents, which is particularly true for lower intensity ranges ($50^{th}-5^{th}$, Figure 7.5a), larger intensity ranges ($95^{th}-50^{th}$, Figure 7.5b) cause negative differences in many ESL cases. Therefore, no clear relationship between hydrograph intensity and flood extent can be seen at the historic city center of Lübeck. To support the finding of an unclear relationship, particularly at Lübeck, we have further described Figure 7.5 in Appendix C.

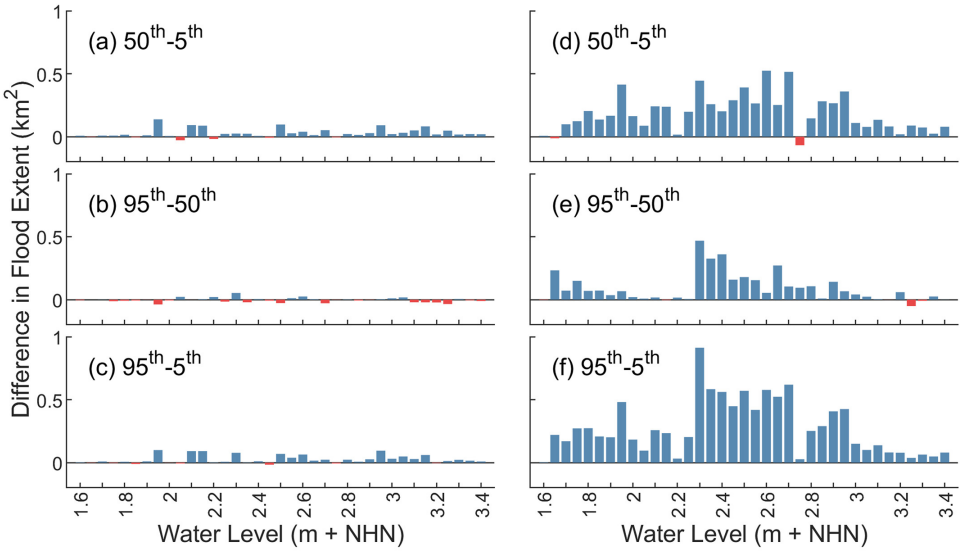

Figure 7.5: Differences in flood extent at Lübeck due to differences in hydrograph intensity (left column: **(a–c)**) and Travemünde (right column: **(d–f)**).

7.4.2 *Effect of Full Range of Hydrograph Intensity on Flood Extent*

Although hydrograph intensity can lead to large differences in flood extent, from Figures 4 and 5 we can see that this influence is dependent on location and ESL magnitude. To further explore the effect of hydrograph intensity on flood extent, we simulated 100 events with a range of intensities at two different water levels. Results of these simulations confirm that flood extent at Travemünde is much more sensitive to changes in hydrograph intensity than at Lübeck, and this sensitivity is affected by the magnitude of the ESL. Flood extents for each simulation are shown in Figure 7.6.

At Travemünde we find substantial variability in flood extent when considering a large range of ESLs (Figure 7.6b), with a maximum difference of approximately 0.97 km² (60%) for the 2.24 m ESL. However, this variability is reduced when considering the larger 2.55 m ESL, with a maximum difference of 0.69 km² (24%). We find a positive relationship between flood extent and hydrograph intensity at Travemünde with statistically significant correlation coefficients of 0.68 at 2.24 m and 0.83 at 2.55 m. There is a steep increase in flood extent at low to medium intensities which quickly flattens out as intensities grow. This shows that differences in flood extent cannot be explained fully by changes in hydrograph intensity, as the relationship between these parameters is not perfect. Thus, the full range of flood extents shown in Figure 7.6 must also be influenced by some other factor not yet identified. At Lübeck, however, the

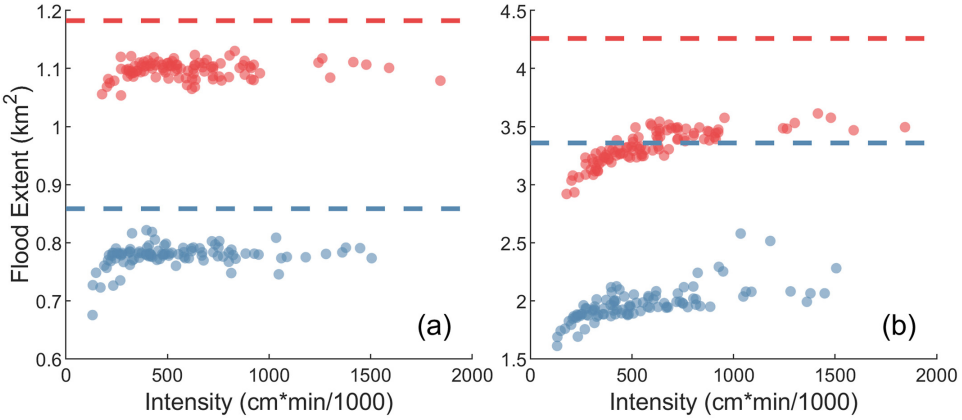

Figure 7.6: Flood extent sensitivity to hydrograph intensity at **(a)** Lübeck and **(b)** Travemünde. Blue and red dots show the results of simulations with an extreme sea level magnitude of 2.24 and 2.55 m, respectively. Maximum flood extent modeled using the bathtub method are shown as blue and red dashed lines for the 2.24 and 2.55 m events, respectively.

variability in flood extent is almost negligible and there is no clear relationship between intensity and flood plain, as shown in Figure 7.6a. This is further demonstrated by the correlation coefficients of 0.1 at 2.24 m and 0.12 at 2.55 m. While several of the low intensity events produce low flood extents in comparison to the other simulations, no statistically significant relationship between flood extent and hydrograph intensity is found.

For comparison purposes, flood extent at 2.24 and 2.55 m was calculated using the bathtub method and is shown in Figure 7.6 as dashed lines. It is commonly known that the bathtub method tends to overestimate flood extents (Didier et al., 2019; Vousdoukas et al., 2016), as it accounts for topography, but it does not consider bottom roughness or other dynamic processes, of which an important parameter is storm duration. For comparison, we also include floodplains produced with the bathtub method, which provides a first-order estimate of the potential floodplain for the respective ESL height, when storm duration, for example, is disregarded. The flood extent from the bathtub method exceeds the mean flood extent of all 100 simulations at Travemünde by approximately 71% at 2.24 m and 28% at 2.55 m, and at Lübeck by approximately 10% and 8% at 2.24 and 2.55 m, respectively. Although it is expected that the bathtub method overestimates flood extent compared to hydrodynamic simulations, these differences are substantial at Travemünde, especially at 2.24 m. Interestingly, even the largest flood extents determined using the hydrodynamic model at Travemünde are much lower than those calculated using the bathtub method. This would suggest that large areas of the floodplain begin flooding only after substantial

pre-filling, and that these requirements are not met by the events modeled in this study. As we consider low-probability events with very large intensities, it appears physically impossible for an event to result in flood extents given by the bathtub method at Travemünde. We include maps showing flood extents from the bathtub method in comparison with the maximum flood extent from the hydrodynamic model in Appendix D (Figure D1).

In addition to the calculation of flood extents, we generated probabilistic flood maps based on different hydrograph intensities at both ESL magnitudes, 2.24 and 2.55 m. Figure 7.7 shows flood probabilities in comparison to flood extent caused by the median intensity events at Travemünde. The comparison of the probabilistic flood plain resulting from hydrograph intensities with the flood extent of the median intensity enables the visualization of disregarded flood extents, when only accounting for a median intensity event. Flood probabilities at Lübeck are not shown, as differences between flood extents are small. At Travemünde, there are large areas of flooding not covered by the flood extent of the median event at the 2.24 m water level (Figure 7.7a). The largest area is located at the unpopulated and low-lying flood plain east of the Priwall peninsula, with probabilities ranging from 1% to 20%. However, two other flood plains additionally cover populated areas of Travemünde with a probability of approximately 35%. For the 2.55 m event (Figure 7.7b), differences in flood extent are less due to the flooding that already occurs from the median intensity in the flood plain east of the Priwall peninsula. Nevertheless, differences between flood extents caused by the median intensity and other hydrographs do exist, and again we find these mostly in this area, east of the Priwall. Unlike for the 2.24 m event, flooding east of the Priwall starts to encroach on houses in the region.

7.5 DISCUSSION

In this study, we assess flooding at Lübeck on the German Baltic coast over a large range of ESL magnitudes and hydrograph intensities. Whereas other studies commonly focus on few ESLs (Hallegatte et al., 2011; Purvis et al., 2008; Wadey et al., 2015), we employ high powered computing and automated procedures for model setup and analysis to simulate large numbers of events. In total, 311 individual events were modeled. We combined results from these simulations to generate probabilistic flood maps of the study site, and explore the effect of hydrograph intensity on estimates of flood extent.

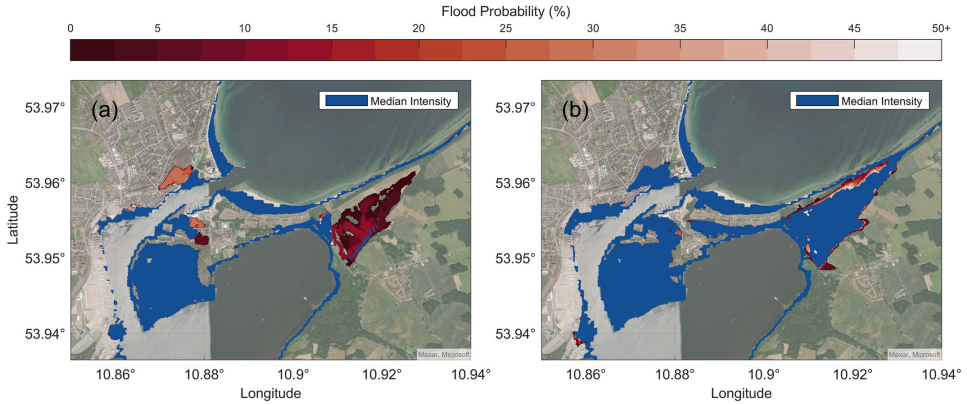

Figure 7.7: Flood extent not captured by the simulation of one median intensity extreme sea levels (ESLs) event. Flood probabilities (red shades) are based on the simulation of 100 ESL hydrographs varying in intensity, at magnitudes of **(a)** 224 and **(b)** 255 cm. For comparison purposes, flood extent due to the median intensity ESL hydrograph is overlaid in blue.

The first outcomes of this study are the probabilistic flood maps of Lübeck, which detail not only the areas exposed to flooding but also the likelihood of flooding in terms of annual probability. These maps provide useful and broad information to coastal managers and decision makers. Studies that estimate exposure or damage over one (e.g., 100-year floodplain) (Hallegatte et al., 2011; Neumann et al., 2015) or only a few floodplains (e.g., Muis et al., 2015; Wadey et al., 2012) provide important information but only at the assessed return water levels. Indeed, Ward et al. (2011) show that the selection of return periods can affect estimates of flood risk. Studies which do account for large ranges of return water levels, assessing either exposure (e.g. Hauer et al., 2021) or damage (e.g., Ward et al., 2011), express their results in terms of expected annual exposed population or expected annual damages. In contrast, our results provide information on the annual probability that a specific area is exposed to flooding. This broad information is useful for long-term adaptation decision-making, including adaptation location, and height, and risk communication. By considering a large range of ESLs, irrespective of their probability or impact, these maps reveal areas that lie outside of the standard protection level. This may be especially relevant for evacuation planning of very densely populated areas, or for the design of critical infrastructure, in places where risk aversion is high (de Bruijn et al., 2019). As it is often of interest for coastal decision-makers to minimize the costs of adaptation (Hinkel et al., 2019), probabilistic flood maps can be used to identify areas which can be left unprotected, and at the same time to communicate risk to low probability zones under the consideration of the selected ESLs (Koerth et al., 2017).

In Germany, adaptation decisions (e.g., implementing coastal defenses such as dikes) are based on a design height equal to an ESL with a return period of 200 years (HW200, MELUR, 2022). However, no probabilistic risk assessment is required. Beyond this standard protection level, there are large uncertainties regarding the exposure of populations to flooding, and further uncertainty when considering the possibility of failure such as dike breaching (Apel et al., 2008). Further exacerbating the problem, climate change induced SLR leads to an increased frequency of ESLs over time (Church et al., 2006) and by extension, higher probabilities of coastal flooding (Vousdoukas et al., 2018b). While the effect of SLR would undoubtedly lead to changes in the annual flood probabilities presented in this study, this falls beyond the scope of this paper and should be addressed in future research. Nevertheless, the method used in this study allows for flexibility when assigning probabilities, due to the number of ESLs simulated. As the floodplains derived in this study are associated with the heights of specific ESLs rather than their probabilities, the information may be adapted to include the most current distribution of ESLs available.

While the magnitude of ESLs is the most important parameter for determining flood extent, previous studies have shown that variability in the temporal evolution of ESLs can affect flood exposure (Quinn et al., 2014; Santamaria-Aguilar et al., 2017). This is often ignored in flood risk analyses due to the increased number of simulations required and difficulties in sourcing a large range of ESL hydrographs to force the model. Thus, most studies rely on a standard ESL shape or "design curve," based on a single event within the observational record. In contrast, we employ ESL hydrographs generated using a stochastic ESL model from MacPherson et al. (2019). This allows us to consider very large numbers of events with specific magnitudes, and select events for simulations based on both ESL magnitude and hydrograph intensity. We therefore considered three levels of hydrograph intensity, selected at the 5^{th}, 50^{th}, and 95^{th} percentiles.

Our results support the findings of Quinn et al. (2014) and Santamaria-Aguilar et al. (2017), and show that temporal variability in hydrographs, in our case applied as hydrograph intensity can cause considerable variations in flood extent. It must be mentioned, however, that both studies were conducted in environments with a relatively strong tidal signal, while our study is located in the Baltic Sea Region, where the tidal signal is negligible. In tidal areas, the effect of the hydrograph intensity on the flood extent may differ, depending on the timing of the storm event with respect to the tidal phase. This is also shown by Quinn et al. (2014) who find that flooding is greater due to storm surge variability where tides are low. We further show

that differences in flood extent depend on both location and ESL magnitude. While at Lübeck, these are negligible, Travemünde can experience substantial changes in flood extent due to hydrograph intensity (up to approximately 50% increase), albeit at specific ESL magnitudes. Differences in the sensitivity of flood extent to hydrograph intensity between Lübeck and Travemünde may be explained by the fact that the flood plain at Lübeck is generally restricted by steep banks, resulting in relatively small differences at all tested ESLs. In contrast, flood plains at Travemünde are large with relatively flat or slowly sloping banks to the Trave River. Thus, these flood plains are often only filled under ESL events with high intensities. Comparisons of the ESL hydrographs at Lübeck and Travemünde suggest little to no reduction in flood potential at Lübeck as the flood wave propagates upriver. That is, the general shape of the hydrograph remains the same over the full length of the Trave considered.

In contrast to studies by Quinn et al. (2014), Santamaria-Aguilar et al. (2017), and Höffken et al. (2020), who consider only the upper and lower bounds of hydrograph intensity, we further examine the effect of hydrograph intensity on flooding by considering a complete distribution of events. At the current official HW200 value of 2.55 m, we find that flood extent differs by approximately 24% at Travemünde, based on 100 simulations of ESL hydrographs. These results are similar to those shown by Quinn et al. (2014), where the number of buildings exposed at Portsmouth in the UK differed by more than 30% during an ESL event with a return period of 200 years. However, as discussed above, the sensitivity of flood extent to hydrograph intensity is dependent on ESL magnitude and local topographic characteristics. In fact, we see a much greater difference in flood extent of around 60% at Travemünde when considering the old official HW200 value of 2.24 m (now approximately the official HW50) (Figure 7.6).

At both 2.24 m (see 2.25 m in Figure 7.5) and 2.55 m, the increase in flood extent at Travemünde from simulations of the 5^{th} and 95^{th} percentile intensity hydrographs is about 45% and 14% respectively, much lower than the 60% and 24% increases seen when 100 events are simulated (Figure 7.6). Flood extent and hydrograph intensity show some correlation; however, this relationship is not linear. Therefore, simulating ESLs based on some quantile of hydrograph intensity does not necessarily produce a flood extent in the same quantile. This is highlighted in Figure 7.5, demonstrating that higher intensities can cause lower flood extents, even at Travemünde where there are larger low-lying areas and where thus the sensitivity of flood extent is generally high. As a consequence, simulating the upper and lower bounds of hydrograph intensity may give a general idea

about flood extent sensitivity to hydrograph intensity, but it does not capture the full picture. We must also note that by choosing the 5^{th}, 50^{th}, and 95^{th} quantiles, we covered a considerably broad range of intensity quantiles. We would expect that covering a narrower quantile range (e.g., 10th/90th or 17th/83rd) may produce even less clear results. Studies that consider only a limited number of hydrograph intensities may therefore underestimate its effect on flood extent and consequently underestimate flood risk.

However, changes in flood extent are not fully attributable to differences in hydrograph intensity. As seen at Travemünde, the relationship between flood extent and hydrograph intensity is not perfect (Figure 7.6). Instead, some other unidentified factors must also influence flooding. We compared a number of ESL parameters with flood extents (Appendix B, Table 7.B1, and Figure 7.B1) in an effort to identify the most important factors that influence flooding. We find that hydrograph intensity is indeed the most relevant parameter, likely due to its calculation based on the entire ESL hydrograph. Differences in flood extent are thus likely explained by a combination of ESL parameters and dynamics relating to local bathymetry. Accounting for all of these factors in a complex system is difficult. Instead, the measurement of hydrograph intensity offers a reasonable proxy for flooding potential while remaining simple to calculate.

Last, as previously mentioned, many coastal flood risk assessments apply a so-called standard design hydrograph (a standard ESL curve) to simulate flooding. Such a curve can be comparable to our median intensity event. In Figure 7.7 we show, in particular for the 2.24 m ESL, that a large amount of flood extent would be disregarded and thus underestimated in the case of Travemünde, when only considering one average hydrograph shape. Therefore, specifically for localities with gently sloping and large low-lying areas, we recommend taking a large number of hydrographs into consideration, when modeling floods. Last, as a result of progress in high-performance computing, hydrodynamic models of lower complexity can be successfully employed to dynamically simulate flooding at larger scale (Kiesel et al., 2023; Vousdoukas et al., 2016). Nevertheless, at this scale, flood plains are often obtained from a single event with a specific return-period. As our results suggest high variability in flood extents when accounting for a large number of hydrographs at locations with large flood plains, we propose that also regional-to-global flood modeling assessments can be greatly improved using (a) the probabilistic approach and (b) a wider number of hydrographs for the selected events. There are first attempts to account for hydrograph variability, as it was done in a simplified way by Le Gal et al.

(2023), who considered a small number of storm durations at the European scale.

7.6 CONCLUSIONS

In this study, we systematically estimate annual flood probabilities at Lübeck on the German Baltic Sea coast. We simulate flooding over a large range of ESLs and explore the effect of hydrograph intensity on flood extents. Using automated model setup and analysis procedures combined with a stochastic ESL generator, we perform 311 simulations to comprehensively assess flood exposure.

The main outcome of the study are probabilistic flood maps, which reveal areas exposed to flooding and at what probabilities. Accounting for a large range of ESLs, we provide annual flood probabilities from a return period of approximately 4 years up to approximately 3,600 years. In contrast to studies that consider one or few ESLs, we provide broad information that details areas exposed to different types of flooding, from nuisance flooding up to once-in-a-generation floods and beyond. This is useful for coastal planners to not only show which areas are below the current safety level (200-year return period in Germany), but to identify which areas may be suitable or unsuitable for critical infrastructure, where risk aversion is higher.

We provide probabilistic flood maps for ESLs with hydrograph intensities in the 5^{th}, 50^{th}, and 95^{th} percentile. We find that the sensitivity of flood extent to hydrograph intensity is dependent on both location and ESL magnitude. For example, flood extents differ substantially more at Travemünde than upriver at Lübeck, particularly when ESLs are between 2.3 and 3.05 m. We attribute these differences to local topography at the two locations. Whereas Lübeck is defined by steep banks, Travemünde's floodplains are largely flat which require sufficiently large hydrograph intensities to fill these areas.

Although simulations of the upper and lower bounds of hydrograph intensity show differences in flood extent, we find that this does not capture the full picture. Simulations of 100 different hydrographs equal in magnitude show large differences in flood extent compared to those seen when only the 5^{th} and 95^{th} percentile events are considered. Although flood extent and hydrograph intensity are correlated, this relationship is not perfect. Thus, selecting ESLs based purely on quantiles of hydrograph intensity does not necessarily produce flood extents within the same quantiles.

We consider current expected annual probabilities of flooding, without accounting for climate change induced SLR or adaptation. Despite this, our results are based on ESL magnitudes rather than probabilities. This provides flexibility in assigning probabilities to floodplains and may be updated with the most current information as required. Nevertheless, the effect of SLR on the frequency of ESLs in the region is likely to change the results presented herein. Thus, to improve the information provided within this study, future studies should examine how ESLs are affected by SLR. As with this study, we propose a methodological framework for a comprehensive flood assessment without accounting for damage, an analysis of assets within the presented floodplains, and by extension, potential damages, would assist coastal managers in adaptation planning.

APPENDIX A: MODEL CALIBRATION AND VALIDATION

Physical and numerical parameters of the hydrodynamic model Delft3D, such as the bottom roughness, time step, and the α-reflection parameter (Deltares, 2024) are tested during the model calibration and verified during model validation. Model calibration and validation are performed on a 25 × 25 m² grid for computational reasons. The relevance of the horizontal resolution was previously tested by changing the horizontal resolution from 50 to 25 m, which only slightly improved model performance, according to the time series comparison (Table 7.A2). We calibrate the model by changing the parameters time step, the α-coefficient, and the manning bottom roughness in individual simulations of a storm surge event with a peak water level of 2 m, occurring in October 1995. Therefore, we use input water levels, obtained from the tide gauge Travemünde. For the manning roughness, we test different settings, using a uniform roughness value of 0,03 (Quinn et al., 2014), a binary manning roughness, differentiating between land (0.03) and water (0.02) (LW) only (Garzon and Ferreira, 2016), or implementing varying values according to land cover. We derive the variable manning roughness corresponding to land cover classes from Bunya et al. (2010), Garzon and Ferreira (2016), Hossain et al. (2009), Kalyanapu et al. (2009), Wamsley et al. (2009), Liu et al. (2013), Dorn et al. (2014), Bellos et al. (2017), and Quinn et al. (2014). However, depending on the reference, studies use higher or lower values for specific land cover classes. Thus, we test land use corresponding manning roughness using lower and higher values (Table 7.A1). We compare the model output to observed water levels at a second tide gauge station Lübeck Bauhof. To evaluate the model performance, we calculate the goodness-of-fit parameters root mean square error (RMSE) and the

correlation coefficient r^2 between the model output and observed water level time series. The changed parameters and corresponding goodness-of-fit estimates are summarized in Table 7.A2. The effect of changed parameters is relatively small. Nevertheless, the best fitting parameters result in a variable manning roughness with low manning values, a time step of 0.125 s, and $\alpha = 750$. Parameters not mentioned here are kept as their default values as suggested by Deltares (2024).

Table 7.A1: Land Use Classes and Corresponding Manning Roughness Values According to the Literature

Land use class	**High**	**Low**	
Agriculture	0.06[a]	0.03[c]	a. Bunya et al. (2010)
Forest	0.2[a]	0.1[d]	b. Garzon and Ferreira (2016)
Urban	0.15[a]	0.015[e, f]	c. Wamsley et al. (2009)
Wetland	0.08[a]	0.035[a]	d. Liu et al. (2013)
Sea and ocean	0.03[b]	0.012[b]	e. Bellos et al. (2017)
Inland waterbodies/courses	0.06[b]	0.02[b]	f. Hossain et al. (2009)
Green urban areas	0.12[a]	0.035[a, g]	g. Dorn et al. (2014)
Natural grasslands	0.042[a]	0.034[d]	h. Kalyanapu et al. (2009)
Traffic	0.032[a]	0.012[h]	
Unvegetated coastal sediment	0.09[d]	0.03[a] /0.025[f]	

We validate the model using a second storm surge event, which occurred in January 2017 with the best fitting parameter set-up derived from the model calibration and again compare the output time series with observed WL time series from the tide gauge Lübeck Bauhof. Goodness-of-fit (GOF) values estimated between the two WL curves demonstrate reasonable results, with a RMSE of 0.06 and r^2 of 0.99.

Table 7.A2: Tested Parameters of Model Calibration and GOF Estimates

GOF	**Manning coeff**				**Time step (s)**				**Alpha**			**Horizontal resolution (m)**	
Tested values	Uniform	High	Low	LW	0.02	0.05	0.1	0.125	1000	750	500	50	25
RMSE	0.09	0.1	0.085	0.087	0.1	0.11	0.12	0.09	0.1	0.08	0.1	0.1	0.09
r^2	0.97	0.96	0.97	0.98	0.87	0.98	0.96	0.98	0.96	0.98	0.97	0.97	0.98

APPENDIX B: HYDROGRAPH CHARACTERISTICS

Table 7.B1: Relationship of Flood Extent and Other Parameters of Hydrograph Characteristics, Taken From MacPherson et al. (2019)

Correlation	t_{Dur}	t_{Dur}*t_{loc}	Intensity	Duration (hours) > 2 m
ρ	0.69	0.45	0.72	0.7
ρ_{val}	0.001	0	0	0

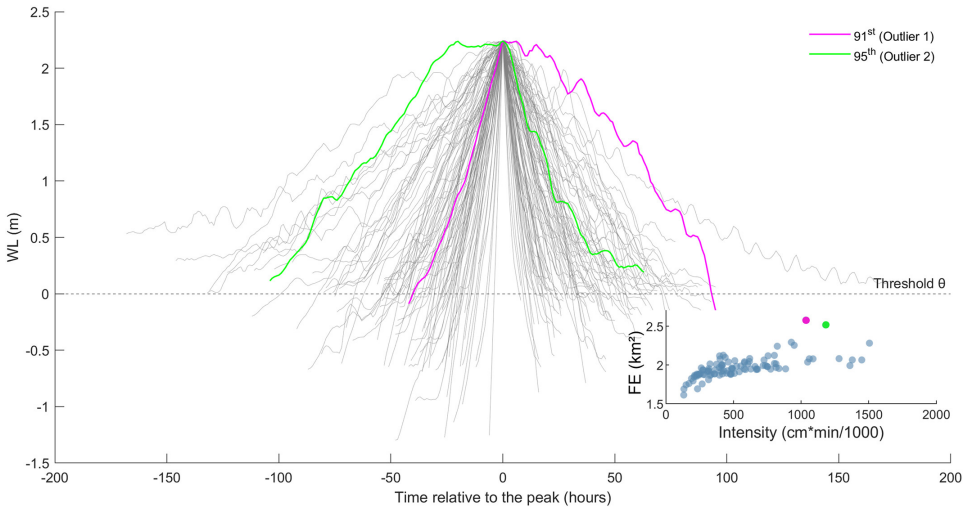

Figure 7.B1: Hydrographs of all 100 intensities (gray) for the 50 yr-event (2.24 m). The hydrographs highlighted in green and pink represent the hydrograph intensities causing two outliers of the relationship flood extent (FE)-intensity, as we have shown in Figure 7.6b (subfigure in this figure), at Travemünde.

APPENDIX C: HYDROGRAPH INTENSITY QUANTILES

We have shown in Figure 7.5 that while positive differences in flood extent are large at Travemünde, those are comparably small at the historic city center of Lübeck. Looking closer into the differences between median and low intensities ($50^{th}-5^{th}$, Figure 7.5d) and median and high intensities (95^{th}~50^{th}, Figure 7.5e) it becomes apparent that positive differences between intensities and flood extent are more pronounced at lower intensity ranges. However, higher hydrograph intensities do not always result in larger flood extents. For example, flood extent caused by the median intensity event exceeds that of the 95^{th} percentile intensity at 5 of the 37 tested WLs at Travemünde (e.g., Figure 7.5e at 1.65, 2.15, and 3.25 m). Interestingly, this occurs even more often at Lübeck, where flood extent during the median

intensity exceeds that of the 95th percentile intensity event 23 times (Figure 7.5b). In fact, 5th percentile intensity events result in larger flood extents at Lübeck than the corresponding 95th percentile intensity events on eight occasions. It should be noted, however, that differences in flood extent are typically small when an event with a lower intensity produces a larger flood plain than a more intense event of the same return water level.

APPENDIX D: BATHTUB METHOD

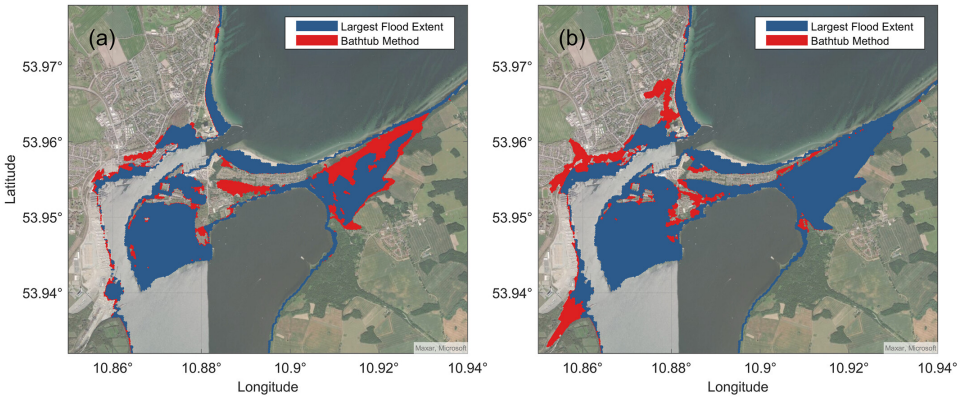

Figure 7.D1: Largest flood extent produced from 100 hydrograph intensities (blue) on top of the flood extent from the bathtub method (red), for the **(a)** 2.24 and the **(b)** 2.55 m extreme sea level magnitudes, at Travemünde.

BIBLIOGRAPHY

Acero, F. J., García, J. A., and Gallego, M. C. (2011). Peaks-over-Threshold Study of Trends in Extreme Rainfall over the Iberian Peninsula. *Journal of Climate*, 24, pp. 1089–1105. DOI: 10.1175/2010JCLI3627.1.

AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., and Sorooshian, S. (2012). Extremes in a Changing Climate: Detection, Analysis and Uncertainty. *Springer Science & Business Media, New York, NY*.

Allamano, P., Laio, F., and Claps, P. (2011). Effects of Disregarding Seasonality on the Distribution of Hydrological Extremes. *Hydrology and Earth System Sciences*, 15, pp. 3207–3215. DOI: 10.5194/hess-15-3207-2011.

Andersson, H. C. (2002). Influence of Long-Term Regional and Large-Scale Atmospheric Circulation on the Baltic Sea Level. *Tellus A*, 54, pp. 76–88. DOI: 10.1034/j.1600-0870.2002.00288.x.

Apel, H., Merz, B., and Thieken, A. H. (2008). Quantification of Uncertainties in Flood Risk Assessments. *International Journal of River Basin Management*, 6, pp. 149–162. DOI: 10.1080/15715124.2008.9635344.

Arns, A., Wahl, T., Dangendorf, S., and Jensen, J. (2015). The Impact of Sea Level Rise on Storm Surge Water Levels in the Northern Part of the German Bight. *Coastal Engineering*, 96, pp. 118–131. DOI: 10.1016/j.coastaleng.2014.12.002.

Arns, A., Wahl, T., Haigh, I. D., Jensen, J., and Pattiaratchi, C. (2013). Estimating Extreme Water Level Probabilities: A Comparison of the Direct Methods and Recommendations for Best Practise. *Coastal Engineering*, 81, pp. 51–66. DOI: 10.1016/j.coastaleng.2013.07.003.

Arns, A., Wahl, T., Haigh, I. D., and Jensen, J. (2015). Determining Return Water Levels at Ungauged Coastal Sites: A Case Study for Northern Germany. *Ocean Dynamics*, 65, pp. 539–554.

Ashley, R., Blanksby, J., Maguire, T., and Leahy, T. (2011). Frameworks for Adapting to Floodrisk: Experiences from the EU> s Flood Resilient City Project. *Proceedings of the 1st IAHR European Congress*, pp. 4–6.

Aslan, B. and Zech, G. (2005). Statistical Energy as a Tool for Binning-Free, Multivariate Goodness-of-Fit Tests, Two-Sample Comparison and Unfolding. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 537, pp. 626–636. DOI: 10.1016/j.nima.2004.08.071.

Bader, B. (2016). *Automated, Efficient, and Practical Extreme Value Analysis with Environmental Applications*. arXiv: 1611.08261 [math, stat]. URL: http://arxiv.org/abs/1611.08261 (visited on 06/25/2024). Pre-published.

Baensch, O. (1875). Die Sturmfluth Vom 12./13. November 1872 an Den Ostseeküsten Des Preussischen Staates. *Zeitschrift für Bauwesen*, 25, pp. 156–220.

Bales, J. and Wagner, C. (2009). Sources of Uncertainty in Flood Inundation Maps. *Journal of Flood Risk Management*, 2, pp. 139–147. DOI: 10.1111/j.1753-318X.2009.01029.x.

Balkema, A. A. and Haan, L. de (1974). Residual Life Time at Great Age. *The Annals of Probability*, 2, pp. 792–804. DOI: 10.1214/aop/1176996548.

Bardet, L., Duluc, C.-M., Rebour, V., and L'Her, J. (2011). Regional Frequency Analysis of Extreme Storm Surges along the French Coast. *Natural Hazards and Earth System Sciences*, 11, pp. 1627–1639. DOI: 10.5194/nhess-11-1627-2011.

Bates, P. D., Dawson, R. J., Hall, J. W., Horritt, M. S., Nicholls, R. J., Wicks, J., and Mohamed Ahmed Ali Mohamed Hassan (2005). Simplified Two-Dimensional Numerical Modelling of Coastal Flooding and Example Applications. *Coastal Engineering*, 52, pp. 793–810. DOI: 10.1016/j.coastaleng.2005.06.001.

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., and Hunt, T. (2013). A UK Best-Practice Approach for Extreme Sea-Level Analysis along Complex Topographic Coastlines. *Ocean Engineering*, 71, pp. 28–39. DOI: 10.1016/j.oceaneng.2013.02.003.

Beckmann, S., Arns, A., MacPherson, L., and Jensen, J. (2022). Possibility of Statistical Correction of Hydrodynamic-Numerical Model Results Using the Example of the Storm Surge of 1872. *Die Küste*, 92, DOI: 10.18171/1.092105.

Bellos, V., Kourtis, I. M., Moreno-Rodenas, A., and Tsihrintzis, V. A. (2017). Quantifying Roughness Coefficient Uncertainty in Urban Flooding Simulations through a Simplified Methodology. *Water*, 9, p. 944. DOI: 10.3390/w9120944.

Benito, G. et al. (2004). Use of Systematic, Palaeoflood and Historical Data for the Improvement of Flood Risk Estimation. Review of Scientific Methods. *Natural Hazards*, 31, pp. 623–643. DOI: 10.1023/B:NHAZ.0000024895.48463.eb.

Beven, K. J. and Hall, J. (2014). Applied Uncertainty Analysis For Flood Risk Management. *Imperial College Press, London, UK*.

Bleemer, Z. and Van der Klaauw, W. (Staff Report) (2017). Disaster (over-) Insurance: The Long-Term Financial and Socioeconomic Consequences of Hurricane Katrina. URL: https://www.econstor.eu/handle/10419/175215.

Bork, I., Rosenhagen, G., and MÜller-Navarra, S. (2022). Modelling the Extreme Storm Surge in the Western Baltic Sea on November 13, 1872, Revisited. *Die Küste*, DOI: 10.18171/1.092103.

Brown, S., Nicholls, R. J., Lázár, A. N., Hornby, D. D., Hill, C., Hazra, S., Appeaning Addo, K., Haque, A., Caesar, J., and Tompkins, E. L. (2018). What Are the Implications of Sea-Level Rise for a 1.5, 2 and 3 °C Rise in Global Mean Temperatures in the Ganges-Brahmaputra-Meghna and Other Vulnerable Deltas? *Regional Environmental Change*, 18, pp. 1829–1842. DOI: 10.1007/s10113-018-1311-0.

Bruijn, K. D., Green, C., Johnson, C., and McFadden, L. (2007). Evolving Concepts in Flood Risk Management: Searching for a Common Language. In: *Flood Risk Management in Europe: Innovation in Policy and Practice*. *Springer Netherlands*, *Dordrecht, NL*.pp. 61–75.

BSH (Bundesamt für Seeschifffahrt und Hydrographie) (2019). Sturmflut Vom 02.01.2019. URL: https://www.bsh.de/DE/THEMEN/Wasserstand_und_Gezeiten/Sturmfluten/_Anlagen/Downloads/Ostsee_Sturmflut_20170104.pdf;jsessionid=EA4C46DE05D4153900C304EC4F236ED1.live11291?__blob=publicationFile%26v=6.

BSH (Bundesamt für Seeschifffahrt und Hydrographie) (2023). GeoSeaPortal - Bathymetrie. URL: https://data.bsh.de/SpatialData/Main/ELC_INSPIRE/.

Bulteau, T., Idier, D., Lambert, J., and Garcin, M. (2015). How Historical Information Can Improve Estimation and Prediction of Extreme Coastal Water Levels: Application to the Xynthia Event at La Rochelle (France). *Natural Hazards and Earth System Sciences*, 15, pp. 1135–1147. DOI: 10.5194/nhess-15-1135-2015.

Bunya, S. et al. (2010). A High-Resolution Coupled Riverine Flow, Tide, Wind, Wind Wave, and Storm Surge Model for Southern Louisiana and Mississippi. Part I: Model Development and Validation. *Monthly Weather Review*, 138, pp. 345–377. DOI: 10.1175/2009MWR2906.1.

Church, J. A., Hunter, J. R., McInnes, K. L., and White, N. J. (2006). Sea-Level Rise around the Australian Coastline and the Changing Frequency of Extreme Sea-Level Events. *Australian Meteorological Magazine*, 55, pp. 253–260.

CLC (CORINE Land Cover) (2018). Copernicus Land Monitoring Service. URL: https://land.copernicus.eu/en/products/corine-land-cover/clc2018.

Codiga, D. L. (2011). Unified Tidal Analysis and Prediction Using the UTide Matlab Functions. URL: URL: https://www.researchgate.net/profile/Daniel-Codiga/publication/280722790_Unified_tidal_analysis_and_prediction_using_the_UTide_Matlab_functions/links/55ff3e5808aeba1d9f8404fb/Unified-tidal-

analysis - and - prediction - using - the - UTide - Matlab - functions.pdf.

Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. *Springer, London, UK*.

Coles, S., Pericchi, L. R., and Sisson, S. (2003). A Fully Probabilistic Approach to Extreme Rainfall Modeling. *Journal of Hydrology*, 273, pp. 35–50. DOI: 10.1016/S0022-1694(02)00353-0.

Coles, S. and Tawn, J. (2005). Bayesian Modelling of Extreme Surges on the UK East Coast. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 363, pp. 1387–1406. DOI: 10.1098/rsta.2005.1574.

Coles, S. G. and Tawn, J. A. (1996). A Bayesian Analysis of Extreme Rainfall Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45, pp. 463–478. DOI: 10.2307/2986068.

Compo, G. P. et al. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137, pp. 1–28. DOI: 10.1002/qj.776.

Connelly, A., Gabalda, V., Garvin, S., Hunter, K., Kelly, D., Lawson, N., O'Hare, P., and White, I. (2015). Testing Innovative Technologies for Flood Resilience. *Institution of Civil Engineers. Proceedings. Water Management*, 168, pp. 66–73. DOI: 10.1680/wama.14.00063.

Dangendorf, S., Arns, A., Pinto, J. G., Ludwig, P., and Jensen, J. (2016). The Exceptional Influence of Storm 'Xaver' on Design Water Levels in the German Bight. *Environmental Research Letters*, 11, p. 054001. DOI: 10.1088/1748-9326/11/5/054001.

Dangendorf, S., Marcos, M., Wöppelmann, G., Conrad, C. P., Frederikse, T., and Riva, R. (2017). Reassessment of 20th Century Global Mean Sea Level Rise. *Proceedings of the National Academy of Sciences*, 114, pp. 5946–5951. DOI: 10.1073/pnas.1616007114.

Davison, A. C. and Smith, R. L. (1990). Models for Exceedances Over High Thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52, pp. 393–425. DOI: 10.1111/j.2517-6161.1990.tb01796.x.

Dawson, R. J., Hall, J. W., Bates, P. D., and Nicholls, R. J. (2005). Quantified Analysis of the Probability of Flooding in the Thames Estuary under Imaginable Worst-case Sea Level Rise Scenarios. *International Journal of Water Resources Development*, 21, pp. 577–591. DOI: 10.1080/07900620500258380.

De Bruijn, K. (2004). Resilience and Flood Risk Management. *Water Policy*, 6, pp. 53–66. DOI: 10.2166/wp.2004.0004.

De Bruijn, K. M., Maran, C., Zygnerski, M., Jurado, J., Burzel, A., Jeuken, C., and Obeysekera, J. (2019). Flood Resilience of Critical Infrastructure: Approach and Method Applied to Fort Lauderdale, Florida. *Water*, 11, p. 517. DOI: 10.3390/w11030517.

Deltares (2024). Delft3D-FLOW: Simulation of Multidimensional Hydrodynamic Flows Ad Transport Phenomena, Including Sedi-

ments, User Manual. URL: URL: https://content.oss.deltares.nl/delft3d4/Delft3D-FLOW_User_Manual.pdf.

Díaz, G. (2014). A Note on the Multivariate Archimedean Dependence Structure in Small Wind Generation Sites. *Wind Energy*, 17, pp. 1287–1295. DOI: 10.1002/we.1633.

Didier, D., Baudry, J., Bernatchez, P., Dumont, D., Sadegh, M., Bismuth, E., Bandet, M., Dugas, S., and Sévigny, C. (2019). Multihazard Simulation for Coastal Flood Mapping: Bathtub versus Numerical Modelling in an Open Estuary, Eastern Canada. *Journal of Flood Risk Management*, 12, e12505. DOI: 10.1111/jfr3.12505.

Dixon, M. J. and Tawn, J. A. (1999). The Effect of Non-Stationarity on Extreme Sea-Level Estimation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 48, pp. 135–151. DOI: 10.1111/1467-9876.00145.

Dorn, H., Vetter, M., and Höfle, B. (2014). GIS-Based Roughness Derivation for Flood Simulations: A Comparison of Orthophotos, LiDAR and Crowdsourced Geodata. *Remote Sensing*, 6, pp. 1739–1759. DOI: 10.3390/rs6021739.

Dullaart, J. C. M., Muis, S., de Moel, H., Ward, P. J., Eilander, D., and Aerts, J. C. J. H. (2023). Enabling Dynamic Modelling of Coastal Flooding by Defining Storm Tide Hydrographs. *Natural Hazards and Earth System Sciences*, 23, pp. 1847–1862. DOI: 10.5194/nhess-23-1847-2023.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7, pp. 1–26. DOI: 10.1214/aos/1176344552.

Efron, B. and Hinkley, D. V. (1978). Assessing the Accuracy of the Maximum Likelihood Estimator: Observed Versus Expected Fisher Information. *Biometrika*, 65, pp. 457–482. DOI: 10.2307/2335893.

El Adlouni, S. and Ouarda, T. B. M. J. (2009). Joint Bayesian Model Selection and Parameter Estimation of the Generalized Extreme Value Model with Covariates Using Birth-Death Markov Chain Monte Carlo. *Water Resources Research*, 45, DOI: 10.1029/2007WR006427.

EMODnet (2022). Bathymetry. URL: URL: https://sextant.ifremer.fr/record/ff3aff8a-cff1-44a3-a2c8-1910bf109f85.

EU (European Parliament) (2007). Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment and Management of Flood Risks. URL: https://eur-lex.europa.eu/eli/dir/2007/60/oj.

Feistel, R., Nausch, G., and Wasmund, N. (2008). State and Evolution of the Baltic Sea, 1952-2005: A Detailed 50-Year Survey of Meteorology and Climate, Physics, Chemistry, Biology, and Marine Environment. *Wiley, London, UK*.

FEMA (Federal Emergency Management Agency) (1968). National Flood Insurance Act of 1968. URL: https://www.fema.gov/national-flood-insurance-program.

Fernández-Montblanc, T., Vousdoukas, M. I., Ciavola, P., Voukouvalas, E., Mentaschi, L., Breyiannis, G., Feyen, L., and Salamon, P. (2019). Towards Robust Pan-European Storm Surge Forecasting. *Ocean Modelling*, 133, pp. 129–144. DOI: 10.1016/j.ocemod.2018.12.001.

Ferro, C. A. T. and Segers, J. (2003). Inference for Clusters of Extreme Values. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65, pp. 545–556. DOI: 10.1111/1467-9868.00401.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24, pp. 180–190. DOI: 10.1017/S0305004100015681.

Fortunato, A. B., Freire, P., Bertin, X., Rodrigues, M., Ferreira, J., and Liberato, M. L. R. (2017). A Numerical Study of the February 15, 1941 Storm in the Tagus Estuary. *Continental Shelf Research*, 144, pp. 50–64. DOI: 10.1016/j.csr.2017.06.023.

Frau, R., Andreewsky, M., and Bernardara, P. (2018). The Use of Historical Information for Regional Frequency Analysis of Extreme Skew Surge. *Natural Hazards and Earth System Sciences*, 18, pp. 949–962. DOI: 10.5194/nhess-18-949-2018.

Fréchet, M. (1927). Sur La Loi de Probabilité de l'écart Maximum. *Ann. de la Soc. Polonaise de Math.*

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., and Viglione, A. (2010). Inclusion of Historical Information in Flood Frequency Analysis Using a Bayesian MCMC Technique: A Case Study for the Power Dam Orlík, Czech Republic. *Contributions to Geophysics and Geodesy*, 40, pp. 121–147. DOI: 10.2478/v10126-010-0005-5.

Gallien, T. W., Schubert, J. E., and Sanders, B. F. (2011). Predicting Tidal Flooding of Urbanized Embayments: A Modeling Framework and Data Requirements. *Coastal Engineering*, 58, pp. 567–577. DOI: 10.1016/j.coastaleng.2011.01.011.

Galton, F. (1894). Natural Inheritance. *Macmillan and Company, London, UK*.

Garvin, S., White, I., O'Hare, P., Connelly, A., and Lawson, N. (2013). Integrating Innovative Flood Resilience Measures: Lessons from Europe. *Proceedings of the International Conference on Flood Resilience: Experiences in Asia and Europe ICFR. Exeter, United Kingdom*, pp. 5–7.

Garzon, J. L. and Ferreira, C. M. (2016). Storm Surge Modeling in Large Estuaries: Sensitivity Analyses to Parameters and Physical Processes in the Chesapeake Bay. *Journal of Marine Science and Engineering*, 4, p. 45. DOI: 10.3390/jmse4030045.

Gaume, E. (2018). Flood Frequency Analysis: The Bayesian Choice. *WIREs Water*, 5, e1290. DOI: 10.1002/wat2.1290.

Gönnert, G. and Sossidi, K. (2011). A New Approach to Calculate Extreme Storm Surges. *Irrigation and Drainage*, 60, pp. 91–98. DOI: 10.1002/ird.681.

Gräwe, U. and Burchard, H. (2012). Storm Surges in the Western Baltic Sea: The Present and a Possible Future. *Climate Dynamics*, 39, pp. 165–183. DOI: 10.1007/s00382-011-1185-z.

Gräwe, U., Klingbeil, K., Kelln, J., and Dangendorf, S. (2019). Decomposing Mean Sea Level Rise in a Semi-Enclosed Basin, the Baltic Sea. *Journal of Climate*, 32, pp. 3089–3108. DOI: 10.1175/JCLI-D-18-0174.1.

Griggs, G. (2017). 1. Human Settlement of the Coastal Zone. In: *1. Human Settlement of the Coastal Zone. University of California Press, Oakland, CA*.pp. 1–20.

Gumbel, E. J. (1958). Statistics of Extremes. *Columbia University Press, New York, NY*.

Haigh, I., Nicholls, R., and Wells, N. (2011). Rising Sea Levels in the English Channel 1900 to 2100. *Proceedings of the Institution of Civil Engineers - Maritime Engineering*, 164, pp. 81–92. DOI: 10.1680/maen.2011.164.2.81.

Haigh, I. D., MacPherson, L. R., Mason, M. S., Wijeratne, E. M. S., Pattiaratchi, C. B., Crompton, R. P., and George, S. (2014). Estimating Present Day Extreme Water Level Exceedance Probabilities around the Coastline of Australia: Tropical Cyclone-Induced Storm Surges. *Climate Dynamics*, 42, pp. 139–157. DOI: 10.1007/s00382-012-1653-0.

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., Arns, A., Bradshaw, E., and Thompson, P. (2022). GESLA Version 3: A Major Update to the Global Higher-Frequency Sea-Level Dataset. URL: URL: https://gesla787883612.wordpress.com.

Haigh, I. D., Nicholls, R., and Wells, N. (2010). A Comparison of the Main Methods for Estimating Probabilities of Extreme Still Water Levels. *Coastal Engineering*, 57, pp. 838–849. DOI: 10.1016/j.coastaleng.2010.04.002.

Haigh, I. D., Wijeratne, E. M. S., MacPherson, L. R., Pattiaratchi, C. B., Mason, M. S., Crompton, R. P., and George, S. (2014). Estimating Present Day Extreme Water Level Exceedance Probabilities around the Coastline of Australia: Tides, Extra-Tropical Storm Surges and Mean Sea Level. *Climate Dynamics*, 42, pp. 121–138. DOI: 10.1007/s00382-012-1652-1.

Hall, J. and Solomatine, D. (2008). A Framework for Uncertainty Analysis in Flood Risk Management Decisions. *International Journal of River Basin Management*, 6, pp. 85–98. DOI: 10.1080/15715124.2008.9635339.

Hallegatte, S., Green, C., Nicholls, R. J., and Corfee-Morlot, J. (2013). Future Flood Losses in Major Coastal Cities. *Nature Climate Change*, 3, pp. 802–806. DOI: 10.1038/nclimate1979.

Hallegatte, S., Ranger, N., Mestre, O., Dumas, P., Corfee-Morlot, J., Herweijer, C., and Wood, R. M. (2011). Assessing Climate Change

Impacts, Sea Level Rise and Storm Surge Risk in Port Cities: A Case Study on Copenhagen. *Climatic Change*, 104, pp. 113–137. DOI: 10.1007/s10584-010-9978-3.

Hallin, C. et al. (2021). A Comparative Study of the Effects of the 1872 Storm and Coastal Flood Risk Management in Denmark, Germany, and Sweden. *Water*, 13, p. 1697. DOI: 10.3390/w13121697.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V. (2014). Extreme Storm Surges: A Comparative Study of Frequency Analysis Approaches. *Natural Hazards and Earth System Sciences*, 14, pp. 2053–2067. DOI: 10.5194/nhess-14-2053-2014.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V. (2015). Use of Historical Information in Extreme-Surge Frequency Estimation: The Case of Marine Flooding on the La Rochelle Site in France. *Natural Hazards and Earth System Sciences*, 15, pp. 1515–1531. DOI: 10.5194/nhess-15-1515-2015.

Hansestadt Lübeck (2022). *Statistische Nachrichten Nr. 50*. URL: https://www.luebeck.de/de/rathaus/verwaltung/statistik/index.html (visited on 08/05/2024).

Hanson, S., Nicholls, R., Ranger, N., Hallegatte, S., Corfee-Morlot, J., Herweijer, C., and Chateau, J. (2011). A Global Ranking of Port Cities with High Exposure to Climate Extremes. *Climatic Change*, 104, pp. 89–111. DOI: 10.1007/s10584-010-9977-4.

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications.,

Hauer, M. E., Hardy, D., Kulp, S. A., Mueller, V., Wrathall, D. J., and Clark, P. U. (2021). Assessing Population Exposure to Coastal Flooding Due to Sea Level Rise. *Nature Communications*, 12, p. 6900. DOI: 10.1038/s41467-021-27260-1.

Hayes, M. H. (1996). Statistical Digital Signal Processing and Modeling. 1st edition. *Wiley, New York, NY*.

Hinkel, J. et al. (2021). Uncertainty and Bias in Global to Regional Scale Assessments of Current and Future Coastal Flood Risk. *Earth's Future*, 9, e2020EF001882. DOI: 10.1029/2020EF001882.

Hinkel, J., Church, J. A., Gregory, J. M., Lambert, E., Le Cozannet, G., Lowe, J., McInnes, K. L., Nicholls, R. J., van der Pol, T. D., and van de Wal, R. (2019). Meeting User Needs for Sea Level Rise Information: A Decision Analysis Perspective. *Earth's Future*, 7, pp. 320–337. DOI: 10.1029/2018EF001071.

Hinkel, J., Jaeger, C., Nicholls, R. J., Lowe, J., Renn, O., and Peijun, S. (2015). Sea-Level Rise Scenarios and Coastal Risk Management. *Nature Climate Change*, 5, pp. 188–190. DOI: 10.1038/nclimate2505.

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J., Marzeion, B., Fettweis, X., Ionescu, C., and Levermann, A. (2014). Coastal Flood Damage and Adaptation Costs under 21st Century Sea-Level Rise. *Proceedings of the National Academy of Sciences*, 111, pp. 3292–3297. DOI: 10.1073/pnas.1222469111.

Hinkel, J., van Vuuren, D. P., Nicholls, R. J., and Klein, R. J. T. (2013). The Effects of Adaptation and Mitigation on Coastal Flood Impacts during the 21st Century. An Application of the DIVA and IMAGE Models. *Climatic Change*, 117, pp. 783–794. DOI: 10.1007/s10584-012-0564-8.

Höffken, J., Vafeidis, A. T., MacPherson, L. R., and Dangendorf, S. (2020). Effects of the Temporal Variability of Storm Surges on Coastal Flooding. *Frontiers in Marine Science*, 7, DOI: 10.3389/fmars.2020.00098.

Hofstede, J. (2008). Coastal Flood Defence and Coastal Protection along the North Sea Coast of Schleswig-Holstein. *Die Küste*, 74, pp. 134–142.

Hofstede, J. and Hamann, M. (2022). The 1872 Catastrophic Storm Surge at the Baltic Sea Coast of Schleswig-Holstein; Lessons Learned? *Die Küste*, DOI: https://doi.org/10.18171/1.092101.

Hossain, A. K. M. A., Jia, Y., and Chao, X. (2009). Estimation of Manning's Roughness Coefficient Distribution for Hydrodynamic Model Using Remotely Sensed Land Cover Features. *2009 17th International Conference on Geoinformatics*, pp. 1–4.

Hurst, H. E. (1951). Long-Term Storage Capacity of Reservoirs. *Transactions of the American society of civil engineers*, 116, pp. 770–799.

IPCC (Intergovernmental Panel on Climate Change) (2022). The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change. URL: https://www.cambridge.org/core/books/ocean-and-cryosphere-in-a-changing-climate/A05E6C9F8638FA7CE1748DE2EB7B491B.

Irvine, K. N. and Eberhardt, A. J. (1992). Multiplicative, Seasonal Arima Models for Lake Erie and Lake Ontario Water Levels. *JAWRA Journal of the American Water Resources Association*, 28, pp. 385–396. DOI: 10.1111/j.1752-1688.1992.tb04004.x.

Isikwue, M. O., Onoja, S. B., and Naakaa, D. S. (2015). Classical and Bayesian Markov Chain Monte Carlo (MCMC) Modeling of Extreme Rainfall (1979-2014) in Makurdi, Nigeria. *International Journal of Water Resources and Environmental Engineering*, 7, pp. 123–131. DOI: 10.5897/IJWREE2015.0588.

Jenkinson, A. F. (1955). The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements. *Quarterly Journal of the Royal Meteorological Society*, 81, pp. 158–171. DOI: 10.1002/qj.49708134804.

Jensen, J., Habib, M., and Beckmann, S. (2022). Best Estimates for Historical Storm Surge Water Level and MSL Development at the Travemünde/Baltic Sea Gauge over the Last 1,000 Years. *Die Küste*, 92, DOI: https://doi.org/10.18171/1.092102.

Jensen, J. and Müller-Navarra, S. H. (2008). Storm Surges on the German Coast. *Die Küste, 74 ICCE*, pp. 92–124.

Jensen, J. and Töppe, A. (1990). Untersuchungen Über Sturmfluten an Der Ostsee Unter Spezieller Berücksichtigung Des Pegels Travemünde. *Deutsche Gewässerkundliche Mitteilungen*, 34, pp. 29–37.

Jevrejeva, S., Moore, J. C., Woodworth, P. L., and Grinsted, A. (2005). Influence of Large-Scale Atmospheric Circulation on European Sea Level: Results Based on the Wavelet Transform Method. *Tellus A: Dynamic Meteorology and Oceanography*, 57, p. 183. DOI: 10.3402/tellusa.v57i2.14609.

Jongman, B., Ward, P. J., and Aerts, J. C. J. H. (2012). Global Exposure to River and Coastal Flooding: Long Term Trends and Changes. *Global Environmental Change*, 22, pp. 823–835. DOI: 10.1016/j.gloenvcha.2012.07.004.

Kalyanapu, A. J., Burian, S. J., and McPherson, T. N. (2009). Effect of Land Use-Based Surface Roughness on Hydrologic Model Output. *Journal of Spatial Hydrology*, 9,

Karmakar, S. and Simonovic, S. (2009). Bivariate Flood Frequency Analysis. Part 2: A Copula-Based Approach with Mixed Marginal Distributions. *Journal of Flood Risk Management*, 2, pp. 32–44. DOI: 10.1111/j.1753-318X.2009.01020.x.

Kiesel, J., Lorenz, M., König, M., Gräwe, U., and Vafeidis, A. T. (2023). Regional Assessment of Extreme Sea Levels and Associated Coastal Flooding along the German Baltic Sea Coast. *Natural Hazards and Earth System Sciences*, 23, pp. 2961–2985. DOI: 10.5194/nhess-23-2961-2023.

Kiesel, J., MacPherson, L. R., Schuerch, M., and Vafeidis, A. T. (2022). Can Managed Realignment Buffer Extreme Surges? The Relationship Between Marsh Width, Vegetation Cover and Surge Attenuation. *Estuaries and Coasts*, 45, pp. 345–362. DOI: 10.1007/s12237-021-00984-5.

Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D., and Hinkel, J. (2020). Projections of Global-Scale Extreme Sea Levels and Resulting Episodic Coastal Flooding over the 21st Century. *Scientific Reports*, 10, p. 11629. DOI: 10.1038/s41598-020-67736-6.

Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J. P., Srivastava, A. K., and Sugi, M. (2010). Tropical Cyclones and Climate Change. *Nature Geoscience*, 3, pp. 157–163. DOI: 10.1038/ngeo779.

Koerth, J., Vafeidis, A. T., and Hinkel, J. (2017). Household-Level Coastal Adaptation and Its Drivers: A Systematic Case Study Review. *Risk Analysis*, 37, pp. 629–646. DOI: 10.1111/risa.12663.

Krien, Y. et al. (2017). Towards Improved Storm Surge Models in the Northern Bay of Bengal. *Continental Shelf Research*, 135, pp. 58–73. DOI: 10.1016/j.csr.2017.01.014.

Kron, W. (2005). Flood Risk = Hazard • Values • Vulnerability. *Water International*, DOI: 10.1080/02508060508691837.

Kumbier, K., Carvalho, R. C., and Woodroffe, C. D. (2018). Modelling Hydrodynamic Impacts of Sea-Level Rise on Wave-Dominated Australian Estuaries with Differing Geomorphology. *Journal of Marine Science and Engineering*, 6, p. 66. DOI: 10.3390/jmse6020066.

Kummu, M., Moel, H. de, Salvucci, G., Viviroli, D., Ward, P. J., and Varis, O. (2016). Over the Hills and Further Away from Coast: Global Geospatial Patterns of Human and Environment over the 20th–21st Centuries. *Environmental Research Letters*, 11, p. 034010. DOI: 10.1088/1748-9326/11/3/034010.

Kupfer, S., MacPherson, L. R., Hinkel, J., Arns, A., and Vafeidis, A. T. (2024). A Comprehensive Probabilistic Flood Assessment Accounting for Hydrograph Variability of ESL Events. *Journal of Geophysical Research: Oceans*, 129, e2023JC019886. DOI: 10.1029/2023JC019886.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of econometrics*, 54, pp. 159–178.

Lambert, E., Rohmer, J., Cozannet, G. L., and Wal, R. S. W. van de (2020). Adaptation Time to Magnified Flood Hazards Underestimated When Derived from Tide Gauge Records. *Environmental Research Letters*, 15, p. 074015. DOI: 10.1088/1748-9326/ab8336.

Le Cozannet, G., Rohmer, J., Cazenave, A., Idier, D., van de Wal, R., de Winter, R., Pedreros, R., Balouin, Y., Vinchon, C., and Oliveros, C. (2015). Evaluating Uncertainties of Future Marine Flooding Occurrence as Sea-Level Rises. *Environmental Modelling & Software*, 73, pp. 44–56. DOI: 10.1016/j.envsoft.2015.07.021.

Le Gal, M., Fernández-Montblanc, T., Duo, E., Montes Perez, J., Cabrita, P., Souto Ceccon, P., Gastal, V., Ciavola, P., and Armaroli, C. (2023). A New European Coastal Flood Database for Low–Medium Intensity Events. *Natural Hazards and Earth System Sciences*, 23, pp. 3585–3602. DOI: 10.5194/nhess-23-3585-2023.

Leadbetter, M. R., Weissman, I., De Haan, L., and Rootzén, H. (1989). On Clustering of High Values in Statistically Stationary Series. *Proc. 4th Int. Meet. Statistical Climatology*, 16, pp. 217–222.

Lesser, G. R., Roelvink, J. A., van Kester, J. A. T. M., and Stelling, G. S. (2004). Development and Validation of a Three-Dimensional Morphological Model. *Coastal Engineering*, 51, pp. 883–915. DOI: 10.1016/j.coastaleng.2004.07.014.

Lewis, M., Bates, P., Horsburgh, K., Neal, J., and Schumann, G. (2013). A Storm Surge Inundation Model of the Northern Bay of Bengal Using Publicly Available Data. *Quarterly Journal of the Royal Meteorological Society*, 139, pp. 358–369. DOI: 10.1002/qj.2040.

Li, F., van Gelder, P. H. A. J. M., Ranasinghe, R., Callaghan, D. P., and Jongejan, R. B. (2014). Probabilistic Modelling of Extreme Storms along the Dutch Coast. *Coastal Engineering*, 86, pp. 1–13. DOI: 10.1016/j.coastaleng.2013.12.009.

Lin, N. and Emanuel, K. (2016). Grey Swan Tropical Cyclones. *Nature Climate Change*, 6, pp. 106–111. DOI: 10.1038/nclimate2777.

Liu, H., Zhang, K., Li, Y., and Xie, L. (2013). Numerical Study of the Sensitivity of Mangroves in Reducing Storm Surge and Flooding to Hurricane Characteristics in Southern Florida. *Continental Shelf Research*, 64, pp. 51–65. DOI: 10.1016/j.csr.2013.05.015.

Ludwig, P. et al. (2023). A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe – Part 2: Historical Context and Relation to Climate Change. *Natural Hazards and Earth System Sciences*, 23, pp. 1287–1311. DOI: 10.5194/nhess-23-1287-2023.

Luijendijk, A. P., Ranasinghe, R., de Schipper, M. A., Huisman, B. A., Swinkels, C. M., Walstra, D. J. R., and Stive, M. J. F. (2017). The Initial Morphological Response of the Sand Engine: A Process-Based Modelling Study. *Coastal Engineering*, 119, pp. 1–14. DOI: 10.1016/j.coastaleng.2016.09.005.

LvermGEO (2022). *ATKIS®-Digitale Geländemodelle (DGM und bDOM)*. schleswig-holstein.de. URL: https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/LVERMGEOSH/Service/serviceGeobasisdaten/geodatenService_Geobasisdaten_DGM (visited on 08/05/2024).

Lyddon, C. E., Brown, J. M., Leonardi, N., Saulter, A., and Plater, A. J. (2019). Quantification of the Uncertainty in Coastal Storm Hazard Predictions Due to Wave-Current Interaction and Wind Forcing. *Geophysical Research Letters*, 46, pp. 14576–14585. DOI: 10.1029/2019GL086123.

MacManus, K., Balk, D., Engin, H., McGranahan, G., and Inman, R. (2021). Estimating Population and Urban Areas at Risk of Coastal Hazards, 1990–2015: How Data Choices Matter. *Earth System Science Data*, 13, pp. 5747–5801. DOI: 10.5194/essd-13-5747-2021.

MacPherson, L. R., Haigh, I. D., and Pattiaratchi, C. (2011). Coastal Flooding in the Peel Harvey Estuary and the Effects of Mean Sea Level Rise. *Coasts and Ports Conference, Perth, Australia (28-30 Sep 2010)*,

MacPherson, L. R., Arns, A., Dangendorf, S., Vafeidis, A. T., and Jensen, J. (2019). A Stochastic Extreme Sea Level Model for the German Baltic Sea Coast. *Journal of Geophysical Research: Oceans*, 124, pp. 2054–2071. DOI: 10.1029/2018JC014718.

MacPherson, L. R., Arns, A., Fischer, S., Méndez, F. J., and Jensen, J. (2023). Bayesian Extreme Value Analysis of Extreme Sea Levels along the German Baltic Coast Using Historical Information.

Natural Hazards and Earth System Sciences, 23, pp. 3685–3701. DOI: 10.5194/nhess-23-3685-2023.

Marcos, M. and Woodworth, P. L. (2017). Spatiotemporal Changes in Extreme Sea Levels along the Coasts of the North Atlantic and the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 122, pp. 7031–7048. DOI: 10.1002/2017JC013065.

Martín, A., Wahl, T., Enriquez, A. R., and Jane, R. (2024). Storm Surge Time Series De-Clustering Using Correlation Analysis. *Weather and Climate Extremes*, 45, p. 100701. DOI: 10.1016/j.wace.2024.100701.

Mathiesen, M., Goda, Y., Hawkes, P. J., Mansard, E., Martín, M. J., Peltier, E., Thompson, E. F., and Van Vledder, G. (1994). Recommended Practice for Extreme Wave Analysis. *Journal of Hydraulic Research*, 32, pp. 803–814. DOI: 10.1080/00221689409498691.

Mazas, F. and Hamm, L. (2011). A Multi-Distribution Approach to POT Methods for Determining Extreme Wave Heights. *Coastal Engineering*, 58, pp. 385–394. DOI: 10.1016/j.coastaleng.2010.12.003.

McClymont, K., Morrison, D., Beevers, L., and Carmen, E. (2020). Flood Resilience: A Systematic Review. *Journal of Environmental Planning and Management*, 63, pp. 1151–1176. DOI: 10.1080/09640568.2019.1641474.

MELUND (Ministerium für Energiewende, Landwirtschaft, Umwelt, Natur und Digitalisierung des Landes Schleswig-Holstein) (2022). Generalplan Küstenschutz Des Landes Schleswig-Holstein. URL: https://www.schleswig-holstein.de/DE/fachinhalte/K/kuestenschutz/Downloads/Generalplan.pdf?__blob=publicationFile&v=3.

MELUR (Ministerium für Ener- giewende, Landwirtschaft, Umwelt und ländliche Räume des Landes Schleswig-Holstein) (2012). Generalplan Küstenschutz Des Landes Schleswig-Holstein. URL: https://www.schleswig-holstein.de/DE/landesregierung/themen/kueste-wasser-meer/wasserstarkSH/_documents/_pdf/vorl_Generalplan_Kuestenschutz.pdf?__blob=publicationFile&v=1.

Mentaschi, L., Vousdoukas, M., Voukouvalas, E., Sartini, L., Feyen, L., Besio, G., and Alfieri, L. (2016). The Transformed-Stationary Approach: A Generic and Simplified Methodology for Non-Stationary Extreme Value Analysis. *Hydrology and Earth System Sciences*, 20, pp. 3527–3547. DOI: 10.5194/hess-20-3527-2016.

Merkens, J.-L., Reimann, L., Hinkel, J., and Vafeidis, A. T. (2016). Gridded Population Projections for the Coastal Zone under the Shared Socioeconomic Pathways. *Global and Planetary Change*, 145, pp. 57–66. DOI: 10.1016/j.gloplacha.2016.08.009.

Merwade, V., Olivera, F., Arabi, M., and Edleman, S. (2008). Uncertainty in Flood Inundation Mapping: Current Issues and Future Directions. *Journal of Hydrologic Engineering*, 13, pp. 608–620. DOI: 10.1061/(ASCE)1084-0699(2008)13:7(608).

Merz, B., Hall, J., Disse, M., and Schumann, A. (2010). Fluvial Flood Risk Management in a Changing World. *Natural Hazards and Earth System Sciences*, 10, pp. 509–527. DOI: 10.5194/nhess-10-509-2010.

Merz, B. and Thieken, A. H. (2009). Flood Risk Curves and Uncertainty Bounds. *Natural Hazards*, 51, pp. 437–458. DOI: 10.1007/s11069-009-9452-6.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The journal of chemical physics*, 21, pp. 1087–1092.

MLUV (Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg-Vorpommern) (2009). Regelwerk Küstenschutz Mecklenburg-Vorpommern. URL: https://www.stalu-mv.de/serviceassistent/download?id=1639581.

Mohr, S. et al. (2023). A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe – Part 1: Event Description and Analysis. *Natural Hazards and Earth System Sciences*, 23, pp. 525–551. DOI: 10.5194/nhess-23-525-2023.

Mudersbach, C. H. and Jensen, J. Ü. R. G. E. N. (2009). Extremwertstatistische Analyse von Historischen, Beobachteten Und Modellierten Wasserständen an Der Deutschen Ostseeküste. *Die Küste*, 75, pp. 131–161.

Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., Yan, K., and Verlaan, M. (2020). A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections. *Frontiers in Marine Science*, 7, DOI: 10.3389/fmars.2020.00263.

Muis, S., Güneralp, B., Jongman, B., Aerts, J. C. J. H., and Ward, P. J. (2015). Flood Risk and Adaptation Strategies under Climate Change and Urban Expansion: A Probabilistic Analysis Using Global Data. *Science of The Total Environment*, 538, pp. 445–457. DOI: 10.1016/j.scitotenv.2015.08.068.

Muis, S., Verlaan, M., Nicholls, R. J., Brown, S., Hinkel, J., Lincke, D., Vafeidis, A. T., Scussolini, P., Winsemius, H. C., and Ward, P. J. (2017). A Comparison of Two Global Datasets of Extreme Sea Levels and Resulting Flood Exposure. *Earth's Future*, 5, pp. 379–392. DOI: 10.1002/2016EF000430.

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H., and Ward, P. J. (2016). A Global Reanalysis of Storm Surges and Extreme Sea Levels. *Nature Communications*, 7, p. 11969. DOI: 10.1038/ncomms11969.

Muis, S. et al. (2023). Global Projections of Storm Surges Using High-Resolution CMIP6 Climate Models. *Earth's Future*, 11, e2023EF003479. DOI: 10.1029/2023EF003479.

Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., and Mitchum, G. T. (2018). Climate-Change–Driven Accelerated Sea-Level Rise Detected in the Altimeter Era. *Proceedings of the*

National Academy of Sciences, 115, pp. 2022–2025. DOI: 10.1073/pnas.1717312115.

Neumann, B., Vafeidis, A. T., Zimmermann, J., and Nicholls, R. J. (2015). Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment. *PLOS ONE*, 10, e0118571. DOI: 10.1371/journal.pone.0118571.

Orr, S. A., Richards, J., and Fatorić, S. (2021). Climate Change and Cultural Heritage: A Systematic Literature Review (2016–2020). *The Historic Environment: Policy & Practice*,

Payrastre, O., Gaume, E., and Andrieu, H. (2011). Usefulness of Historical Information for Flood Frequency Analyses: Developments Based on a Case Study. *Water Resources Research*, 47, DOI: 10.1029/2010WR009812.

Peacock, J. A. (1983). Two-Dimensional Goodness-of-Fit Testing in Astronomy. *Monthly Notices of the Royal Astronomical Society*, 202, pp. 615–627. DOI: 10.1093/mnras/202.3.615.

Phillips, H. (2014). Adaptation to Climate Change at UK World Heritage Sites: Progress and Challenges. *The Historic Environment: Policy & Practice*, DOI: 10.1179/1756750514Z.00000000062.

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

Prosdocimi, I. (2018). German Tanks and Historical Records: The Estimation of the Time Coverage of Ungauged Extreme Events. *Stochastic Environmental Research and Risk Assessment*, 32, pp. 607–622. DOI: 10.1007/s00477-017-1418-8.

PSMSL (2018). Mean Sea Level. URL: URL: https://psmsl.org/.

Pugh, D. (2004). Changing Sea Levels: Effects of Tides, Weather and Climate. *Cambridge University Press, Cambridge, UK*.

Pugh, D. and Woodworth, P. (2014). Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes. *Cambridge University Press, Cambridge, UK*.

Pugh, D. T. (1987). Tides, Surges and Mean Sea-Level: A Handbook for Engineers and Scientists., p. 472.

Pugh, D. and Vassie, J. (1980). Applications of the Joint Probability Method for Extreme Sea Level Computations. *Proceedings of the Institution of Civil Engineers*, 69, pp. 959–975. DOI: 10.1680/iicep.1980.2179.

Purvis, M. J., Bates, P. D., and Hayes, C. M. (2008). A Probabilistic Methodology to Estimate Future Coastal Flood Risk Due to Sea Level Rise. *Coastal Engineering*, 55, pp. 1062–1073. DOI: 10.1016/j.coastaleng.2008.04.008.

Quinn, N., Bates, P. D., and Siddall, M. (2013). The Contribution to Future Flood Risk in the Severn Estuary from Extreme Sea Level Rise Due to Ice Sheet Mass Loss. *Journal of Geophysical Research: Oceans*, 118, pp. 5887–5898. DOI: 10.1002/jgrc.20412.

Quinn, N., Lewis, M., Wadey, M. P., and Haigh, I. D. (2014). Assessing the Temporal Variability in Extreme Storm-Tide Time Series for Coastal Flood Risk Assessment. *Journal of Geophysical Research: Oceans*, 119, pp. 4983–4998. DOI: 10.1002/2014JC010197.

Ramirez, J. A., Lichter, M., Coulthard, T. J., and Skinner, C. (2016). Hyper-Resolution Mapping of Regional Storm Surge and Tide Flooding: Comparison of Static and Dynamic Models. *Natural Hazards*, 82, pp. 571–590. DOI: 10.1007/s11069-016-2198-z.

Reimann, L., Vafeidis, A. T., Brown, S., Hinkel, J., and Tol, R. S. J. (2018). Mediterranean UNESCO World Heritage at Risk from Coastal Flooding and Erosion Due to Sea-Level Rise. *Nature Communications*, 9, p. 4161. DOI: 10.1038/s41467-018-06645-9.

Reimann, L., Vafeidis, A. T., and Honsel, L. E. (2023). Population Development as a Driver of Coastal Risk: Current Trends and Future Pathways. *Cambridge Prisms: Coastal Futures*, 1, e14. DOI: 10.1017/cft.2023.3.

Reis, D. S. and Stedinger, J. R. (2005). Bayesian MCMC Flood Frequency Analysis with Historical Information. *Journal of Hydrology*, 313, pp. 97–116. DOI: 10.1016/j.jhydrol.2005.02.028.

Ribeiro, A., Barbosa, S. M., Scotto, M. G., and Donner, R. V. (2014). Changes in Extreme Sea-Levels in the Baltic Sea. *Tellus A: Dynamic Meteorology and Oceanography*, 66, DOI: 10.3402/tellusa.v66.20921.

Ridal, M. et al. (2024). CERRA , the Copernicus European Regional Reanalysis System. *Quarterly Journal of the Royal Meteorological Society*, qj.4764. DOI: 10.1002/qj.4764.

Rivera, J. (2019). A Content Analysis on the Phases of Emergency Management for Hurricane Maria in Puerto Rico. *The Pegasus Review: UCF Undergraduate Research Journal*, 11, p. 2.

Rohmer, J., Lincke, D., Hinkel, J., Le Cozannet, G., Lambert, E., and Vafeidis, A. T. (2021). Unravelling the Importance of Uncertainties in Global-Scale Coastal Flood Risk Assessments under Sea Level Rise. *Water*, 13, p. 774. DOI: 10.3390/w13060774.

Rosenhagen, G. and Bork, I. (2009). Rekonstruktion Der Sturmwetterlage Vom 13. November 1872. *Die Küste*, 75, pp. 51–70.

Santamaria-Aguilar, S., Arns, A., and Vafeidis, A. T. (2017). Sea-Level Rise Impacts on the Temporal and Spatial Variability of Extreme Water Levels: A Case Study for St. Peter-Ording, Germany. *Journal of Geophysical Research: Oceans*, 122, pp. 2742–2759. DOI: 10.1002/2016JC012579.

Santos, C. T., Toda, L., Orduña, J. R., Santos, F. D., and Ferrão, J. (2016). The Impacts of Typhoon Haiyan in the Philippines: Implications to Land Use Planning. *Climate, Disaster and Development Journal*, 1, pp. 57–66.

Schmidt, J., Dangendorf, S., Mir Calafat, F., Patzke, J., and Jensen, J. (2017). A Novel Tide Gauge Dataset for the Baltic Sea - Part

1: Spatial Features and Temporal Variability of the Seasonal Sea Level Cycle., p. 4407.

Seo, Y., Kim, S., Kisi, O., and Singh, V. P. (2015). Daily Water Level Forecasting Using Wavelet Decomposition and Artificial Intelligence Techniques. *Journal of Hydrology*, 520, pp. 224–243. DOI: 10.1016/j.jhydrol.2014.11.050.

Skinner, C. J., Coulthard, T. J., Parsons, D. R., Ramirez, J. A., Mullen, L., and Manson, S. (2015). Simulating Tidal and Storm Surge Hydraulics with a Simple 2D Inertia Based Model, in the Humber Estuary, U.K. *Estuarine, Coastal and Shelf Science*, 155, pp. 126–136. DOI: 10.1016/j.ecss.2015.01.019.

Sklar, M. (1959). Fonctions de Répartition à N Dimensions et Leurs Marges. *Annales de l'ISUP*, VIII, pp. 229–231.

Slangen, A. B. A., Haasnoot, M., and Winter, G. (2022). Rethinking Sea-Level Projections Using Families and Timing Differences. *Earth's Future*, 10, e2021EF002576. DOI: 10.1029/2021EF002576.

Slivinski, L. C. et al. (2019). Towards a More Reliable Historical Reanalysis: Improvements for Version 3 of the Twentieth Century Reanalysis System. *Quarterly Journal of the Royal Meteorological Society*, 145, pp. 2876–2908. DOI: 10.1002/qj.3598.

Sterr, H. (2008). Assessment of Vulnerability and Adaptation to Sea-Level Rise for the Coastal Zone of Germany. *Journal of Coastal Research*, 24, pp. 380–393. DOI: 10.2112/07A-0011.1.

Stewart, M. and Melchers, R. E. (1997). Probabilistic Risk Assessment of Engineering Systems. *Springer Netherlands, Dordrecht, NL*.

Takbash, A. and Young, I. R. (2019). Global Ocean Extreme Wave Heights from Spatial Ensemble Data. *Journal of Climate*, 32, pp. 6823–6836. DOI: 10.1175/JCLI-D-19-0255.1.

Tawn, J., Vassie, J., and Gumbel, E. (1989). Extreme Sea Levels; the Joint Probabilities Method Revisited and Revised. *Proceedings of the Institution of Civil Engineers*, 87, pp. 429–442. DOI: 10.1680/iicep.1989.2975.

Tawn, J., Vassie, J., and Gumbel, E. (1989). Extreme Sea Levels; the Joint Probabilities Method Revisited and Revised. *Proceedings of the Institution of Civil Engineers*, 87, pp. 429–442. DOI: 10.1680/iicep.1989.2975.

Tawn, J. A. (1988). An Extreme-Value Theory Model for Dependent Observations. *Journal of Hydrology*, 101, pp. 227–250. DOI: 10.1016/0022-1694(88)90037-6.

Tawn, J. A. (1992). Estimating Probabilities of Extreme Sea-Levels. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 41, pp. 77–93. DOI: 10.2307/2347619.

Tebaldi, C., Ranasinghe, R., Vousdoukas, M., Rasmussen, D. J., Vega-Westhoff, B., Kirezci, E., Kopp, R. E., Sriver, R., and Mentaschi, L. (2021). Extreme Sea Levels at Different Global Warming Levels.

Nature Climate Change, 11, pp. 746–751. DOI: 10.1038/s41558-021-01127-1.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D., and Kim, S. (2017). Flood Inundation Modelling: A Review of Methods, Recent Advances and Uncertainty Analysis. *Environmental Modelling & Software*, 90, pp. 201–216. DOI: 10.1016/j.envsoft.2017.01.006.

Tolkien, J. R. R. (1955). The Return of the King: J.R.R. Tolkien. 1st edition. *HarperCollins, London, UK*.

UK (UK Parliament) (2010). Flood and Water Management Act 2010, Chapter 29. URL: https://www.legislation.gov.uk/ukpga/2010/29/contents.

UNESCO WHC (United Nations Educational, Scientific and Cultural Organization World Heritage Centre) (2022). Interactive Map. URL: https://whc.unesco.org/en/interactive-map/.

Van der Pol, T., Hinkel, J., Merkens, J., MacPherson, L., Vafeidis, A. T., Arns, A., and Dangendorf, S. (2021). Regional Economic Analysis of Flood Defence Heights at the German Baltic Sea Coast: A Multi-Method Cost-Benefit Approach for Flood Prevention. *Climate Risk Management*, 32, p. 100289. DOI: 10.1016/j.crm.2021.100289.

Vaziri, M. (1997). Predicting Caspian Sea Surface Water Level by ANN and ARIMA Models. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 123, pp. 158–162. DOI: 10.1061/(ASCE)0733-950X(1997)123:4(158).

Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., and Storlazzi, C. D. (2017). Doubling of Coastal Flooding Frequency within Decades Due to Sea-Level Rise. *Scientific Reports*, 7, p. 1399. DOI: 10.1038/s41598-017-01362-7.

Vousdoukas, M. I., Mentaschi, L., Hinkel, J., Ward, P. J., Mongelli, I., Ciscar, J.-C., and Feyen, L. (2020). Economic Motivation for Raising Coastal Flood Defenses in Europe. *Nature Communications*, 11, p. 2119. DOI: 10.1038/s41467-020-15665-3.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Bianchi, A., Dottori, F., and Feyen, L. (2018). Climatic and Socioeconomic Controls of Future Coastal Flood Risk in Europe. *Nature Climate Change*, 8, pp. 776–780. DOI: 10.1038/s41558-018-0260-4.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., and Feyen, L. (2017). Extreme Sea Levels on the Rise along Europe's Coasts. *Earth's Future*, 5, pp. 304–323. DOI: 10.1002/2016EF000505.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., and Feyen, L. (2018). Global Probabilistic Projections of Extreme Sea Levels Show Intensification of Coastal Flood Hazard. *Nature Communications*, 9, p. 2360. DOI: 10.1038/s41467-018-04692-w.

Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., and Feyen, L. (2016). Projections of Extreme Storm Surge Levels

along Europe. *Climate Dynamics*, 47, pp. 3171–3190. DOI: 10.1007/s00382-016-3019-5.

Wadey, M. P., Cope, S. N., Nicholls, R. J., McHugh, K., Grewcock, G., and Mason, T. (2015). Coastal Flood Analysis and Visualisation for a Small Town. *Ocean & Coastal Management*, 116, pp. 237–247. DOI: 10.1016/j.ocecoaman.2015.07.028.

Wadey, M. P., Nicholls, R. J., and Hutton, C. (2012). Coastal Flooding in the Solent: An Integrated Analysis of Defences and Inundation. *Water*, 4, pp. 430–459. DOI: 10.3390/w4020430.

Wagenaar, D. J., de Bruijn, K. M., Bouwer, L. M., and de Moel, H. (2016). Uncertainty in Flood Damage Estimates and Its Potential Effect on Investment Decisions. *Natural Hazards and Earth System Sciences*, 16, pp. 1–14. DOI: 10.5194/nhess-16-1-2016.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., and Slangen, A. B. A. (2017). Understanding Extreme Sea Levels for Broad-Scale Coastal Impact and Adaptation Analysis. *Nature Communications*, 8, p. 16075. DOI: 10.1038/ncomms16075.

Wahl, T., Mudersbach, C., and Jensen, J. (2011). Assessing the Hydrodynamic Boundary Conditions for Risk Analyses in Coastal Areas: A Stochastic Storm Surge Model. *Natural Hazards and Earth System Sciences*, 11, pp. 2925–2939. DOI: 10.5194/nhess-11-2925-2011.

Wahl, T., Plant, N. G., and Long, J. W. (2016). Probabilistic Assessment of Erosion and Flooding Risk in the Northern Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 121, pp. 3029–3043. DOI: 10.1002/2015JC011482.

Walsh, K. J. et al. (2016). Tropical Cyclones and Climate Change. *WIREs Climate Change*, 7, pp. 65–89. DOI: 10.1002/wcc.371.

Wamsley, T. V., Cialone, M. A., Smith, J. M., Ebersole, B. A., and Grzegorzewski, A. S. (2009). Influence of Landscape Restoration and Degradation on Storm Surge and Waves in Southern Louisiana. *Natural Hazards*, 51, pp. 207–224. DOI: 10.1007/s11069-009-9378-z.

Ward, P. J., de Moel, H., and Aerts, J. C. J. H. (2011). How Are Flood Risk Estimates Affected by the Choice of Return-Periods? *Natural Hazards and Earth System Sciences*, 11, pp. 3181–3195. DOI: 10.5194/nhess-11-3181-2011.

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I. E., Winsemius, H. C., and Wahl, T. (2018). Dependence between High Sea-Level and High River Discharge Increases Flood Hazard in Global Deltas and Estuaries. *Environmental Research Letters*, 13, p. 084012. DOI: 10.1088/1748-9326/aad400.

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., Beek, R. van, Bierkens, M. F. P., Ligtvoet, W., and Winsemius, H. C. (2013). Assessing Flood Risk at the Global Scale: Model Setup, Results, and Sensitivity. *Environmental Research Letters*, 8, p. 044019. DOI: 10.1088/1748-9326/8/4/044019.

Watson, P. J. (2023). Extreme Value Analysis of Ocean Still Water Levels along the USA East Coast—Case Study (Key West, Florida). *Coasts*, 3, pp. 294–312. DOI: 10.3390/coasts3040018.

Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, 18, pp. 293–297. DOI: 10.1115/1.4010337.

Weiss, J., Bernardara, P., and Benoit, M. (2014). Formation of Homogeneous Regions for Regional Frequency Analysis of Extreme Significant Wave Heights. *Journal of Geophysical Research: Oceans*, 119, pp. 2906–2922. DOI: 10.1002/2013JC009668.

Weiss, J., Bernardara, P., and Benoit, M. (2014). Modeling Intersite Dependence for Regional Frequency Analysis of Extreme Marine Events. *Water Resources Research*, 50, pp. 5926–5940. DOI: 10.1002/2014WR015391.

White, G. F., Kates, R. W., and Burton, I. (2001). Knowing Better and Losing Even More: The Use of Knowledge in Hazards Management. *Global Environmental Change Part B: Environmental Hazards*, 3, pp. 81–92. DOI: 10.3763/ehaz.2001.0308.

White, I., Connelly, A., Garvin, S., Lawson, N., and O'Hare, P. (2018). Flood Resilience Technology in Europe: Identifying Barriers and Co-producing Best Practice. *Journal of Flood Risk Management*, 11, DOI: 10.1111/jfr3.12239.

Wiltshire, S. E. (1985). Grouping Basins for Regional Flood Frequency Analysis. *Hydrological Sciences Journal*, 30, pp. 151–159. DOI: 10.1080/02626668509490976.

Winter, B., Schneeberger, K., Huttenlau, M., and Stötter, J. (2018). Sources of Uncertainty in a Probabilistic Flood Risk Model. *Natural Hazards*, 91, pp. 431–446. DOI: 10.1007/s11069-017-3135-5.

Wolski, T. and Wiśniewski, B. (2020). Geographical Diversity in the Occurrence of Extreme Sea Levels on the Coasts of the Baltic Sea. *Journal of Sea Research*, 159, p. 101890. DOI: 10.1016/j.seares.2020.101890.

Wolski, T. and Wiśniewski, B. (2021). Characteristics and Long-Term Variability of Occurrences of Storm Surges in the Baltic Sea. *Atmosphere*, 12, p. 1679. DOI: 10.3390/atmos12121679.

Wolski, T., Wiśniewski, B., Giza, A., Kowalewska-Kalkowska, H., Boman, H., Grabbi-Kaiv, S., Hammarklint, T., Holfort, J., and Lydeikaitė, Ž. (2014). Extreme Sea Levels at Selected Stations on the Baltic Sea Coast*. *Oceanologia*, 56, pp. 259–290. DOI: 10.5697/oc.56-2.259.

WSA (Wasserstraßen- und Schifffahrtsamt Ostsee) (2022). Tide-Gauge Data. URL: https://www.wsa-ostsee.wsv.de/Webs/WSA/Ostsee/DE/1_Startseite/startseite_node.html.

Zachary, S., Feld, G., Ward, G., and Wolfram, J. (1998). Multivariate Extrapolation in the Offshore Environment. *Applied Ocean Research*, 20, pp. 273–295. DOI: 10.1016/S0141-1187(98)00027-3.

Zhang, Y. and Baptista, A. M. (2008). SELFE: A Semi-Implicit Eulerian–Lagrangian Finite-Element Model for Cross-Scale Ocean Circulation. *Ocean Modelling*, 21, pp. 71–96. DOI: 10.1016/j.ocemod.2007.11.005.

Zhang, Y. J., Ye, F., Stanev, E. V., and Grashorn, S. (2016). Seamless Cross-Scale Modeling with SCHISM. *Ocean Modelling*, 102, pp. 64–81. DOI: 10.1016/j.ocemod.2016.05.002.

I can't carry it for you, but I can carry
you and it as well. So up you get.

— **Samwise Gamgee** (Tolkien, 1955)

ACKNOWLEDGMENTS

Much like Frodo and bringing the ring to Mount Doom, I would not have been able to complete this dissertation without the help of others. As the final part of this work I would like to thank those colleagues, friends and family, who have supported me on this adventure.

First, I find it difficult to separate my two supervisors, Arne Arns and Nassos Vafeidis, but I will do so based on alphabetic and chronological order. When I first came to Germany, it was Arne who offered me a position looking at extreme sea levels along the German Baltic Sea coast. When I said that a position in Siegen was not possible as I was moving to Kiel, it was Arne who organised for me a place in Nassos' working group. Thankyou Arne, you have helped me along this entire journey, and I followed you as far as I could go! Before you were my supervisor you were a great mate, and I am happy to say that this remains the case.

Nassos, thanks for accepting me into your group all that time ago, and bearing with me throughout the years. I find it amazing how you only ever offer good advice. Recently I have started to understand how life and work can combine to be absolutely chaotic, and I look to you for inspiration on how to deal with this! Although I know you do not manage this effortlessly, you make it seem that way, always making time for those who need it. You once made me an espresso which changed my mind on coffee, and now I have a habit which helps me work while raising young children. I can't thank you enough!

To the both of you, I consider myself extremely lucky to have had not one but two exceptional supervisors. Although I consider you both brilliant teachers and scientists, you have been tremendous friends that I have relied on not only for professional support, but advice in life as well. Thank you so much for all the help you have given me over the years.

Jürgen Jensen, thankyou for giving me the opportunity to complete my PhD at Siegen. Although it did not work out that way, I remember my trips to Siegen very fondly and enjoyed our conversations regarding my work and discussions on historical extreme sea levels. Your expertise in this regard was essential for my work on incorporating historical information in extreme value analysis. In 2011 or 2012 (I don't quite remember which), I lent you my car so that you could pick up your wife from the airport in Perth. I consider this favour well and truly repaid! I hope to catch you for a drink sometime, as I know from past conferences how enjoyable a beer with you can be! Cheers!

Special appreciations go to my co-authors: Sunna Kupfer, it was great working with you and I hope to do so again! Jochen Hinkel, Svenja Fischer and Fernando Méndez, thanks for your expertise and guidance in the work we completed together.

Thanks to my mates in the Coastal Risks and Sea-Level Rise working group. Special thanks to Sara Santamaria-Aguilar and Jan Merkens, who tolerated me as their office mate. I know Australians and Spanish have a reputation of being loud, so well done to Jan who never once showed any annoyance towards us. Claudia Wolff, thanks for always being their for a chat, whether its about work, life, bread or if I need something printed! To Joshua Kiesel, I did it! Maybe a scotch is in order? To the rest: Sunna, Lena, Jana, Doro, Hedda, Saskia, John and Jacob for making the CRSLR office more than a workplace.

To my colleagues in Siegen, I wish now that I managed to spend more time with you all. Sönke Dangendorf, you were a great help at the start of my PhD. You gave me direction in the early days, and you are a big reason why I decided to continue my research after my first paper. Jessica, Sebastian Gürke, Mai, Andra, Sebastien Niehüser, Marius, Leon, Jens, Sandra, Andre and Felix, thanks for making my visits memorable and making me feel part of the team. Special thanks to Matthias Hirt for hosting me when I visited and providing food and beer!

To all my friends, thanks for taking my mind off my work and helping me recharge whenever possible! To my Australian friends, thanks for staying in touch and having time for me when I return home!

To my family, Mum and Dad for bearing with my absence. I miss you both every day and I can't wait for our next catch up! To my big sister Alana, you work so hard to help others and I have benefited greatly because of this. Thanks so much! To my children, Lily and

Noah, you are the best motivation to get up every day and try my best. You have given me a boat load of perspective and I think about you two in everything that I do.

Finally, Mia — I can say with all honesty that I would never have achieved this without you. You are responsible for so much good in my life and I will forever be grateful.

ERKLÄRUNG

Hiermit erkläre ich, dass ich diese Dissertation, abgesehen von der Beratung durch meine Betreuer, sowohl inhaltlich als auch formal selbstständig verfasst habe und keine anderen Quellen oder Hilfsmittel als die angegebenen verwendet habe. Diese Arbeit wurde weder ganz noch teilweise im Rahmen eines anderen Dissertationsverfahrens eingereicht. Als kumulative Dissertation wurden die Kapitel 5 bis 7 in den zu Beginn der jeweiligen Kapitel angegebenen Zeitschriften veröffentlicht. Ich bestätige, dass diese Arbeit gemäß den Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft erstellt wurde. Darüber hinaus erkläre ich, dass mir bisher kein akademischer Grad entzogen wurde.

Kiel, September 2024

Leigh Richard MacPherson