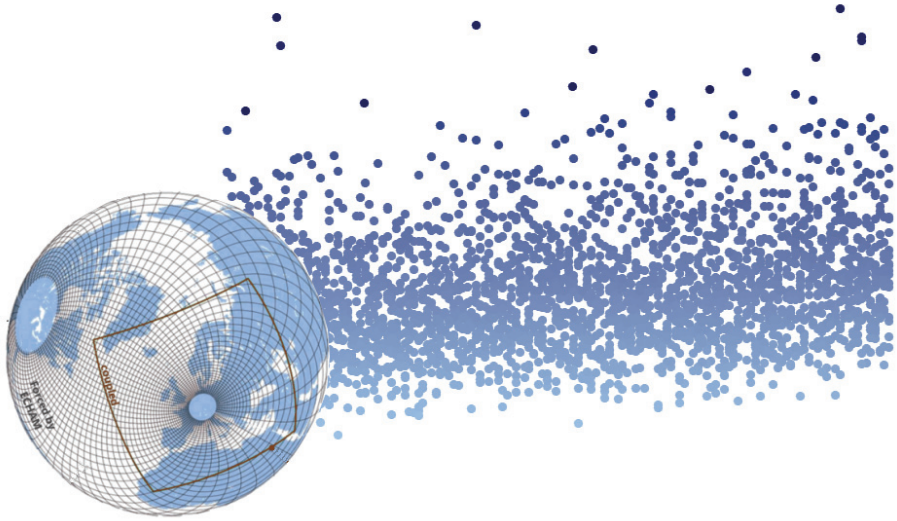
Max-Planck-Institut
für Meteorologie

MAX PLANCK
GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Extreme High Sea Levels in the German Bight

Past Variability and Future Changes



Andreas Lang

Hamburg 2020

Berichte zur Erdsystemforschung
Reports on Earth System Science

228
2020

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Extreme High Sea Levels in the German Bight

Past Variability and Future Changes

Andreas Lang

Hamburg 2020

Andreas Lang
aus Gräfelfing, Deutschland

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Geowissenschaften
Meteorologisches Institut
Bundesstr. 55
20146 Hamburg

Tag der Disputation: 26. Mai 2020

Folgende Gutachter empfehlen die Annahme der Dissertation:
Dr. Uwe Mikolajewicz
Prof. Dr. Gerhard Schmiedl

Vorsitzender des Promotionsausschusses:
Prof. Dr. Dirk Gajewski

Dekan der MIN-Fakultät:
Prof. Dr. Heinrich Graener

Illustration on title page: Sketch of the evolution of sampled extreme high sea levels in a changing climate together with the employed Downscaling Model system.

Typeset using the classicthesis template developed by André Miede, available at:
https://bitbucket.org/amiede/classicthesis/

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 228
Reports on Earth System Science / Max Planck Institute for Meteorology 2020

ISSN 1614-1199

To my parents

— Monika & Herbert (1951-2015)

ABSTRACT

In this dissertation, I investigate the variability of extreme high sea levels (ESLs) in the German Bight. Downscaling global climate simulations with a regionally coupled climate model, I (1) quantify the long-term variability of ESLs over the last millennium and (2) project future changes in their statistics using a large ensemble approach. Compared to previous studies, I generate a larger sample of extreme values that allows a more robust estimation of the policy-relevant *high-impact-low-probability* events. At the same time, the dynamical downscaling with high-frequency ocean-atmosphere coupling provides both a sufficiently high resolution to adequately simulate high-frequency sea level variations and a globally consistent simulation of associated climate states.

In the first part, I downscale a transient global climate simulation of the past millennium to quantify the long-term ESL variability. I find that simulated ESL statistics vary substantially on interannual to centennial timescales but without following preferred oscillation periods. Due to this large variability, ESL variations mask any signals from background sea level variations or from external natural variability such as solar or volcanic forcing. Nevertheless, periods of high ESLs can be linked to an atmospheric mode of variability, namely a sea level pressure dipole between northeastern Scandinavia and the Gulf of Biscay. The high ESL variability emphasizes the inherent uncertainties related to traditional extreme value estimates based on short data subsets, which fail to account for such long-term variations. This further suggests that uncertainties related to high-impact-low-probability ESL estimates are higher than previously assumed.

In the second part, I investigate future changes in ESL statistics under climate change which occur irrespective of a shift due to gradual mean sea level rise. To account for the large ESL variability detected in the first part, I downscale a large ensemble of global *1pctCO2* scenario simulations. While individual realizations exhibit highly different responses, the full ensemble reveals that ESLs statistically increase with rising atmospheric CO_2 levels, particularly along the western coastlines of Schleswig-Holstein and Denmark. Here, the magnitude of ESL change reaches around half of the existing projections of the regional background sea level rise (BSLR) until the end of the century. This ESL response is related to an enhanced large-scale activity along the North Atlantic storm belt, more predominant storms of the major West-Northwest track-type, and a subsequent local increase in mainly westerly wind speed extremes. The response is seasonally opposite, as summer ESLs and their drivers decrease in magnitude, contrasting the response of the higher winter ESLs which govern the annual response.

These results have important implications for coastal protection. First, the high ESL variability shows that estimates of high-impact-low-probability ESLs – and thus also flood protection standards – strongly depend on the state of long-term variability and are thus associated with deep uncertainty. Second, ESLs in the German Bight do not only scale with BSLR but increase additionally. Statistically, ESLs may rise stronger than previously assumed; however, as manifestations of the large internal variability, (multi-)decadal deviations from the statistical long-term trend are to be expected. These deep uncertainties in estimating high ESLs may thus demand even further safety measures.

ZUSAMMENFASSUNG

In dieser Dissertation untersuche ich die Variabilität extremer Hochwasserstände (*extreme sea level*, ESL) in der Deutschen Bucht. Durch Regionalisierung (*Downscaling*) globaler, einen langen Zeitraum umfassender Klimasimulationen mit einem regional gekoppelten Klimamodell, quantifiziere ich deren langzeitliche Variabilität während des letzten Jahrtausends (1), und projiziere zukünftige Änderungen in deren Statistik mithilfe eines großen *Ensembles* (2). Verglichen mit bisherigen Studien generiere ich eine weitaus größere Anzahl an Extremwerten, was eine statistisch robuste Abschätzung der ESLs von größtem Schadensrisiko (*high-impact-low-probability*) ermöglicht. Zugleich bietet das dynamische Downscaling mit hochfrequenter Kopplung zwischen Ozean und Atmosphäre einerseits eine für die adäquate Simulation von ESLs hinreichend hohe Auflösung, und andererseits eine global konsistente Simulation damit assoziierter Klimazustände.

Im ersten Teil regionalisiere ich eine globale Klimasimulation des letzten Jahrtausends, um die langzeitliche ESL-Variabilität zu quantifizieren. Demnach weisen simulierte ESLs eine hohe Variabilität auf Zeitskalen von Jahren bis Jahrhunderten auf, ohne jedoch systematischen Schwankungen zu folgen. Dadurch sind Signale durch Schwankungen des mittleren Meeresspiegels oder durch externe Variabilität solaren oder vulkanischen Ursprungs nicht feststellbar. Nichtsdestotrotz können Zeiten besonders hoher ESLs mit Moden atmosphärischer Variabilität verknüpft werden: So ist das Zirkulationsregime bestehend aus einem Luftdruck-Dipol zwischen der Biskaya und Nordost-Skandinavien für erhöhte ESLs in der Region verantwortlich. Die hohe ESL-Variabilität unterstreicht hierbei die inhärenten Unsicherheiten der Extremwertanalyse auf der Basis kurzer Datenreihen, da diese keine langzeitlichen Variationen abbilden können. Dies deutet darauf hin, dass die Unsicherheit von high-impact-low-probability ESLs höher ist als bislang angenommen, und dass bisherige ESL Abschätzungen stark vom Zustand der langzeitlichen Variabilität abhängen.

Im zweiten Teil untersuche ich solch zukünftige Änderungen in der ESL-Statistik in einem sich verändernden Klima, die unabhängig von einer generellen Verschiebung durch den Anstieg des Meeresspiegels auftreten. Um der hohen Variabilität der im ersten Teil simulierten ESLs Rechnung zu tragen, regionalisiere ich ein Ensemble aus 32 Läufen globaler *1pctCO2* Klimasimulationen. Während einzelne Realisationen unterschiedlich auf eine Erhöhung der CO_2 Konzentrationen reagieren, zeigt das volle Ensemble einen statistisch signifikanten Ansteig der ESLs, vor allem an der Westküste Schleswig-Holsteins und Dänemarks. Die Änderung der ESLs macht damit stellenweise die Hälfte des erwarteten mittleren Meeresspiegelanstiegs in

der Region bis zum Ende des 21. Jahrhunderts aus. Dieses Änderungsmuster hängt mit einer erhöhten großskaligen Sturmaktivität im Nordatlantik, einer Häufung des für ESLs in der Deutschen Bucht wichtigsten Sturmtyps, und einem lokalen Anstieg an extremen Westwinden zusammen. Die klimabedingte Änderung der ESLs ist hierbei saisonal entgegengesetzt, mit einer Zunahme im Winter und einer Abnahme im Sommer. Da Winterwerte insgesamt höher liegen, dominiert über das ganze Jahr ein ESL-Anstieg.

Die Ergebnisse beider Teile haben wichtige Implikationen für den Küstenschutz. Zum einen zeigt die hohe ESL-Variabilität, dass Angaben von high-impact-low-probability ESLs – und damit auch Richthöhen im Hochwasserschutz – mit tiefer Unsicherheit verknüpft sind. Zum anderen skalieren ESL in der Deutschen Bucht nicht nur mit dem erwarteten mittleren Meeresspiegelanstieg, sondern steigen statistisch gesehen zusätzlich. So muss zwar mit einem stärkeren Anstieg an ESLs als bisher angenommen gerechnet werden; aufgrund der hohen internen Variabilität einzelner Realisationen können jedoch durchaus (multi-)dekadische Abweichungen von dem langzeitlichen, statistischen Trend auftreten. Diese nichtreduzierbaren Unsicherheiten in der Abschätzung hoher ESLs könnten dadurch noch strengere Sicherheitsmaßnahmen erfordern.

PUBLICATIONS

This cumulative thesis is based on the following articles:

Lang, A. and Mikolajewicz, U. (2019): The long-term variability of extreme sea levels in the German Bight, *Ocean Sci.*, 15, pp. 651–668. DOI: 10.5194/os-15-651-2019.

Lang, A. and Mikolajewicz, U. (under review): Rising extreme sea levels in the German Bight under enhanced CO_2 levels – a regionalized large ensemble approach for the North Sea, *Clim. Dynamics*.

CONTENTS

LIST OF FIGURES

ACRONYMS

ESL	Extreme Sea Level
BSL	Background Sea Level
BSLR	Background Sea Level Rise
MHW	Mean High Water
GHG	Greenhouse Gas
GMST	Global Mean Surface Temperature
SLP	Sea Level Pressure
NAO	North Atlantic Oscillation
SCA	Scandinavian Pattern
EAP	East Atlantic Pattern
AMO	Atlantic Multidecadal Oscillation
GCM	General Circulation Model
RCM	Regional Climate Model
IPCC	Intergovernmental Panel on Climate Change
AR	Assessment Report
SROCC	Special Report on the Ocean and Cryosphere in a Changing Climate
RCP	Representative Concentration Pathway
CMIP	Climate Model Intercomparison Project
PMIP	Paleo Model Intercomparison Project
POT	Peak-Over-Threshold
GEV	Generalized Extreme Value
EOF	Empirical Orthogonal Function
PC	Principal Component

UNIFYING ESSAY

1

ON EXTREME SEA LEVEL VARIABILITY IN THE GERMAN BIGHT

This dissertation is about high water level extremes. Compared to the mean state, it is primarily the short-lived upper extremes which are responsible for coastal impacts and thus most relevant from a societal perspective. With its low-lying coasts, the North Sea is an area that is particularly exposed to extreme sea levels in form of storm floods, while anthropogenic climate change is expected to exacerbate these threats (IPCC, 2019). Adequate and timely coastal protection thus relies on estimates of expected heights and frequencies of such extreme sea level events. These build on both information of the recent past and estimates of future changes from climate change simulations. Yet, since by definition, extreme events occur only rarely, their very nature complicates a robust estimation of trends and variability in past and future.

In this dissertation, I address this signal-to-noise problem for extreme high water levels in the German Bight. Using downscaled coupled climate simulations, I (1) quantify the long-term variability of upper-end extreme sea levels and identify their main drivers, and, motivated by this information, (2) use a large ensemble approach to project future changes in their statistics. I show that extreme sea levels are characterized by a large internal variability component which complicates the estimation of upper-end statistics based on the relatively short instrumental record. Consequently, the estimation of potential changes in the future requires the use of large ensembles to tackle the uncertainties due to the high variability. By downscaling a set of global climate change simulations from the Max Planck Institute Grand Ensemble (MPI-GE), I show that the increase in extreme sea level heights that can be expected from mean sea level rise in a warming climate could be significantly magnified by an additional rise in upper-end extreme sea levels.

With this unifying essay, I give a more general introduction into the problem of investigating extreme sea levels (1.1-1.3), motivate and pose the main research questions of the thesis (1.4 & 1.5) and briefly present the key findings. These are addressed in detail in the two journal papers (Chapter 2 on past extreme storm flood variability & Chapter 3 on its future change). Finally, I close with a unifying synthesis (1.6).

1.1 INTRODUCTION

Home to currently around 680 million people, low-lying coastal zones are the most densely populated in the world (IPCC, 2019). With higher than average population growth and urbanization rates (Small & Nicholls, 2003), this number is expected to continue to rise in the future, with projections to reach more than one billion by mid-century (Neumann et al., 2015).

Naturally, coastal zones are also vulnerable to natural disasters. As one of the major hazards for coastal regions, storm floods provoke high water level extremes that can lead to inundation of the coastal zone and the hinterland, and subsequently to saltwater intrusion or coastal erosion. Particularly heavy floods can destroy the local environment and infrastructure, and pose danger to ecosystems and human life.

Once described as a "splendid sea for storm surges" (Heaps, 1967), the *German Bight*, with its low-lying coasts of the Netherlands, Germany and Denmark, is particularly exposed to storm floods. As the German Bight is located in the zone of the major northern hemispheric storm belt and the northwest European shelf with strong tides and an open connection with the North Atlantic, these are all contributing factors to frequent and particularly strong storm floods with often devastating consequences (see Lamb & Frydendahl (1991), Buisman (2006) and Jensen & Müller-Navarra (2008) for a historic overview). Already as early as a couple of hundred years B.C., a period of severe storm surges have forced the Cimber people to retreat from the area of Holstein and Jutland (Arends, 1974). In 1362, the great storm flood *Grosse Manndrenke* ('Great Drowning of Men') led to high death tolls including the disappearance of the legendary town *Rungholdt* (Heimreich, 1819) and even shifts in the Wadden Sea coastline (Hadler et al., 2018). The *Second Manndrenke* or *Burchardi flood* in 1634 (Fig. 1, left) changed the landscape to the current coastline (Arends, 1974). A couple of particularly strong floods in the 18th and 19th century, foremost the *Christmas flooding* of 1717 and the *February flood* in 1825, had a wide impact from Danish to Dutch coasts and resulted in the drowning of tens of thousands of people (Jakubowski-Tiessen, 1992; Kempe, 2006). Finally, as one of the strongest storm floods of the twentieth century, the flood of 1962 has shown that the impact of extreme storm floods is not only felt at the coast; it has also led to vast economic losses in bigger estuaries up to some 100 km further inland (MunichRe, 2012), particularly in the city of Hamburg (Fig. 1, right).

Nowadays, North Sea adjoining countries seem generally well-adapted. In Germany, massive investments in coastal defense have been made after the flood in 1962. In Hamburg for instance, the minimum required flood protection height is 8.10 meter above *Normal Null* (FHH, 2019), about 1.5 meter higher than the locally highest flood on the instrumental record (Schirmer, 2018).

Figure 1 | **Extreme storm floods and their consequences over time.** Left: Contemporary engraving of the *Burchardi flood* of 1634 (Picture credit: Public Domain). Right: Flood-damaged Fährstraße in Wilhelmsburg, Hamburg, after the storm flood of 1962 (Picture credit: G. Pietsch; Public Domain).

Due to historic exposure and some of the world's longest high-frequency tide gauge records dating back to the early 1900s, the North Sea is with respect to storm floods one of best studied regions in the world. For a review of past storm surge statistics and projected changes in the North Sea region see von Storch & Woth (2008), Weisse et al. (2012) and Schrum et al. (2016).

Despite the vast amount of literature on past, present and future storm floods, though, our understanding is not exhaustive. One of the most renowned experts in the field, Hans v. Storch writes about a "fragmentation of research" (von Storch & Woth, 2008), as the field is divided into regional research communities, from a more general oceanography variant to coastal engineering. With this division comes an important fragmentation: the temporal and spatial discrepancy between the nature of a local storm flood event and the long-term and large-scale character of climate variability and change.

This problem is most obvious for extreme values. As they happen by definition only rarely, it is their very nature that complicates the detection of a robust trend or variability. Yet, it is those *high-impact-low-probability* extremes that often are of particular interest for the public. Such estimates of, for instance, 50-year or 100-year flood levels ('century-floods') are often also required in coastal protection for the design of flood defense standards. Estimating these extreme levels on the basis of the comparably rather short instrumental tide gauge record in itself is not straight forward as this implies a statistical extrapolation of the data. However, this is further complicated by the fact that the observational record points to a considerable multi-decadal variability in sea level, both in the mean and in higher per-

centiles (e.g. Dangendorf et al., 2013a,b; Weisse et al., 2014). When records are short, one risks sampling different states of long-term variability rather than a significant signal.

Besides estimating the heights of such extreme storm floods, it is therefore vital to understand their temporal variations on different timescales. Are there regularities in their occurrence on different time scales? And what mechanisms in the climate system govern their variability? These questions stand in the focus of the first article (Chapter 2, brief summary in section 1.4). Additionally, anthropogenic climate change is expected to increase the flood risk in the future through a rise in mean sea level. It is uncertain though if extreme sea level statistics also change on top of such a baseline shift, and what drives such potential changes. These questions build the focus of the second article (Chapter 3, brief summary in section 1.5).

Together, these are some of the problems coastal engineers and adaptation planners face. They rely on combining information from past and present heights and variability as well as from future projections, which are all subject to high uncertainty. Climate model simulations present a powerful tool to investigate questions of this kind, as we can artificially increase the sample size required for robust extreme value analysis, while they create a consistent picture of associated climate states. A climate model approach to study the variability of extreme sea levels by utilizing a long-term transient simulation or a large initial conditions ensemble can thus shed new light on an old problem.

Before addressing these problems in detail in sections 1.4 and 1.5, I will first introduce concept and modelling of extreme sea levels.

1.2 EXTREME SEA LEVELS: THEORY

The total sea level can be expressed as the sum of the time-mean background sea level, tidal elevation and atmospheric surge (Pugh, 1987).

The *background sea level* (BSL) – or often referred to as (time-) mean sea level – builds the state the other components act upon. It represents the residual longer-term average level, where the short-term fluctuations are averaged out. The remaining trends or low-frequency variations are caused by ocean thermodynamics, mass changes, vertical land movement, or a result of the redistribution of water by various oceanographic and atmospheric processes such as coastally trapped waves or alongshore winds (Calafat et al., 2012; Dangendorf et al., 2014b; Sturges & Douglas, 2011). Such variations are typically not accounted for in conventional regional surge models, although they affect total sea level heights.

Tidal sea level variations are caused by gravitational forces of moon and sun and generally regular and predictable. Strongest and most important constituent in the North Sea is the semi-diurnal M2 tide. In the German

Bight, the tides are comparably strong, with tidal ranges of around one to four meters, depending on the location. As Kelvin waves, tides in the northern hemisphere move counter clockwise along the coastlines. Furthermore, as a result of the superposition of the different constituents, tides vary on longer time scales, such as the 18.6-year nodal cycle. As tidal amplitudes and phases depend on the bathymetry, they can also be subject to anthropogenic influence (Talke & Jay, 2020). For instance, due to the deepening of estuaries for the shipping industry, the tidal range in the German North Sea coast has been reported to have risen by up to 5% between 1950 and 2015 (Jensen & Mudersbach, 2007).

Finally, the atmospheric *surge* is the most important component for extreme sea levels in the German Bight (Lowe et al., 2010). Surges can be described by the shallow-water equations and the conservation of mass and momentum; their severity depends on the one hand on wind speed, direction and duration, and on the other hand on the coastline geometry and bathymetry (von Storch & Woth, 2008). They are a result of atmospheric processes and can consist of (i) the *local wind surge* that pushes water against the coast, (ii) the direct effect of air pressure acting on the water surface, called the *inverse barometric effect*, and (iii) the *external surge* (Schmitz et al., 1988). The latter is provoked by cyclones and air pressure variations in the North Atlantic that travel into the North Sea, pushing additional water masses into the basin (Gönnert et al., 2012).

Together, these components drive regional extreme sea levels, particularly when coinciding in time. However, their net effect is influenced by various interactions between them, especially in shallow waters. The non-linear tide-surge interaction, for instance, causes storm surges to cluster on the rising tide (Horsburgh & Wilson, 2007). Furthermore, the interaction of tide and BSL can lead to shifts of the *amphidromic points* and changes in the tidal range due to a rise in BSL (Kauker & von Storch, 2000; Plüß, 2004). The interaction of BSL and surge component, though, has found to be small: Although theory suggests that surges diminish with a rise in sea level due to the decreased effect of surface wind stress on the water column (Arns et al., 2015), model experiments support the approximation of a linear superposition of surge and BSL (Lowe & Gregory, 2005; Sterl et al., 2009; Howard et al., 2010). That is, BSL and surge can simply be added together and the changes in the two components studied separately.

1.2.1 *Extreme sea level nomenclature*

Multiple different terms and definitions for describing extreme high water levels can be found in the literature, making it difficult to compare their findings. In its basic sense, 'extreme' simply refers to the tails of the distribution. However, since low extremes are usually not of interest, the term *extreme sea level* typically just refers to the upper tail. While some authors target the extreme sea level in terms of the total absolute value (Mudersbach

& Jensen, 2010; Feng et al., 2015; Wahl & Chambers, 2016; Marcos & Woodworth, 2017), many separate the atmospheric component from background sea level and/or tide and use terms such as *surge component* (e.g. Weisse & Plüß, 2006; Woodworth et al., 2007; Marcos et al., 2015), *storm surge* (e.g. Barriopedro et al., 2010; Cid et al., 2016), *storm-tide* (Gönnert, 2004), *non-tidal residual* (e.g. Pugh & Vassie, 1978; Mudersbach et al., 2013; Marcos & Woodworth, 2017) or increasingly, *skew surge* (e.g. de Vries et al., 1995; Mawdsley & Haigh, 2016). Although an ambiguous term, the term *storm flood* is in literature on North Sea sea levels sometimes used synonymously to the total extreme sea level.

Following the recommendations of Gregory et al. (2019), I define *extreme sea level* (ESL) as the total water level of an exceptionally high local sea surface height. Note that these local long-term mean values do not correspond to meters above *Normal Null*, which is often used by public authorities in coastal protection. The definition using the total water level instead of surge residual or skew surge, avoids a decomposition between tidal and surge parts and their nonlinear interaction (Horsburgh & Wilson, 2007).

1.2.2 *How extreme is extreme?*

A closely related simple but not trivial question emerges: what water levels constitute an extreme of 'exceptionally high' sea surface height? Again, many different methods and practices on how to sample extreme values exist. ESL can either refer to high percentiles of the data or, more directly, to the largest values in a particular record, sampled in terms of frequency (*r-largest*) or magnitude (*peak-over-threshold*). Unlike percentiles, which produce probabilistic estimates depending on the entire distribution, the latter two methods yield a sample of 'real' events. Both *r-largest* and *peak-over-threshold* methods have advantages and disadvantages: While the former results in an equidistant sample of extremes and is thus robust to temporal variations, it might also include levels that may not necessarily be considered 'extreme' during other periods, and vice versa. The choice of the sampling method thus always includes a trade-off between bias (too high r or too low threshold) and variance (too low r or too high threshold) of the selected extremes (Coles et al., 2001). Here, I sample ESL with the simplest form of the *r-largest* method, in terms of annual maximum sea level. However, as storm floods mainly occur in winter, I compute annual statistics in terms of years defined from July to June. As a result, storm flood seasons are not split up or individual storm floods double-counted.

From the sampled extreme values, different target variables can be computed. Coastal protection is often interested in probabilistic estimates of exceptionally high and thus statistically speaking rather rare events, so-called "high-impact-low-probability" events (e.g. Wahl et al., 2017). To also allow probabilistic estimates of sea levels of return periods longer than the underlying record, an often-used concept is that of return levels (e.g. Coles et

al., 2001), representing the water level expected to be exceeded on average once every x years. Such probability-based exceedance levels of assigned return periods are often required for designing coastal defense structures, such as the 100-year return level in Schleswig-Holstein (MFLR, 2012). As the instrumental tide gauge data record or simulated data from time slice experiments rarely cover more than a couple of decades, these are typically estimated based on parametric extreme value analysis, resulting in a considerable extrapolation of the data. That is, an extreme value distribution is fitted to the sampled extreme values, with the chosen sampling method determining the type of extreme value distribution.

The use of transient long-term or ensemble simulations, as done here, maximizes the number of sampled extremes. This has the advantage that it allows to infer upper-end extreme value statistics without relying on parametric methods.

1.3 MODELLING EXTREME SEA LEVELS

1.3.1 *Previous research*

For many decades, mainly barotropic models (e.g. Flather, 1987) have been employed to study the dynamics of extreme sea levels. For a review see von Storch & Woth (2008). In the early days, most research effort has gone into the reproduction of individual storms and their surge heights (e.g. Davies & Flather, 1977; Proctor & Flather, 1989) using storm surge prediction schemes (e.g. Flather, 1979; Peeck et al., 1983) in order to study involved processes or test forecast tools. Since the report of the *Waves and Storms in the North Atlantic Group* (WASA) was published in 1998 (WASA-group, 1998), trends and variability on longer time scales moved more into the research focus.

On the one hand, to understand the variations of past storm floods, multi-decadal *hindcasts*, driven by observations or reanalysis products for the second half of the twentieth century have been employed (e.g. Flather & Smith, 1998; Langenberg et al., 1999; Woth, 2005; Weisse & Plüß, 2006; Woodworth et al., 2007). These studies point towards a long-term variability with an increase in storm flood heights at the end of the century. This temporal behavior has partly been linked to atmospheric phenomena, namely changes in the North Atlantic Oscillation (NAO, Wakelin et al., 2003; Woodworth et al., 2007), the dominant pattern of atmospheric variability over the North Atlantic (Hurrell, 1995).

On the other hand, climate change experiments have been performed to investigate potential future changes in extreme sea levels (For a review see Schrum et al., 2016). These studies rely on similar hydrodynamic surge models, with the exception that the forcing here stems from wind and pressure fields simulated by global general circulation models (GCMs; Flather &

Smith, 1998; Lowe et al., 2001; Sterl et al., 2009; Vousdoukas et al., 2017) or regional climate models (RCMs; Lowe & Gregory, 2005; Woth, 2005; Debernard & Røed, 2008). To investigate the impact of climate change, in these studies the forcing from climate simulations typically covered two time slices of a couple of decades under a current and a future state.

However, statistically speaking, high-frequency observational records or hindcasts driven by observed wind fields are not sufficiently long to quantify the variability of highest extremes or to derive robust relationships between ESL and climate states on decadal and longer scales. Similarly, unless a large ensemble is used, future time slice experiments only cover a couple of decades and can thus only produce reliable statements about rather moderate extremes. Long-term or large ensemble climate model simulations can tackle this sampling issue. Yet, currently available simulations of this kind do not simulate tides and have an insufficient resolution to adequately represent storm surges in the North Sea with its complex coastline and bathymetry. Regional climate models, driven by forcing fields from global climate simulations, can provide a better representation of small-scale processes of marginal seas, topographic influences and land-sea contrasts, making them better suited for the simulation of extreme events (Rockel & Woth, 2007). However, due to their lateral boundaries, regional models cannot account for the consistent propagation of climate signals from the open Atlantic into the North Sea. Barotropic signals for instance, such as alongshore winds or coastally trapped waves (Calafat et al., 2012), or remote steric effects have been shown to influence sea level variations in North Sea (Dangendorf et al., 2013a; Chen et al., 2014). Furthermore, the regional approach complicates the linking of sea level variations to associated drivers in the global climate system. A robust investigation of regional sea level extremes over long timescales with consideration of large-scale climate variability requires a different approach.

1.3.2 *The downscaling model REMO-MPIOM*

I here employ a regionally coupled climate model with a zoom on the North Sea (Mikolajewicz et al., 2005; Elizalde et al., 2014; Sein et al., 2015) to dynamically downscale global climate simulations by the Max Planck Institute Earth System Model (MPI-ESM). The downscaling model consists of the global ocean model MPIOM (Marsland et al., 2004) and the regional atmospheric model REMO (Roeckner et al., 2003), and combines the respective advantages of global and regional models. The hydrological budget in the study domain is closed using a hydrological discharge model (Hagemann & Dümenil, 1999). In order to maximize the grid resolution in the study focus areas, the ocean model MPIOM is run on a stretched grid configuration, with poles shifted to Europe and North America, respectively. This approach yields a resolution in the German Bight of under 10 km; the atmospheric model REMO has a uniform horizontal resolution of 50 km.

Over the wider European domain (Jacob et al., 2014), REMO and MPIOM are interactively coupled to capitalize from the benefits of a high-resolution atmospheric forcing and to account for ocean-atmosphere feedback mechanisms. The regionally high resolution and the inclusion of the full lunisolar ephimeridic tidal potential (Thomas et al., 2001) as well as the high-frequency hourly coupling allow the adequate simulation of shelf processes that are important for the simulation of sea level and their extremes. At the same time, the global ocean grid approach allows the consistent simulation of short-term sea level signals propagating from the open Atlantic onto the North West European shelf. Other than previous studies, this global approach thus also accounts for external surges (Schmitz et al., 1988; Gönnert & Sossidi, 2011; see section 1.2). This might be even more important in the future, as a strengthened North Atlantic storm belt would result in more frequent external surges (Woth et al., 2006). For a more detailed description of the model components as well as a validation using the instrumental record, see chapter 2.

This model system is used for both parts of the analysis (long-term variability and future climate change impact) and only differs by the respective forcing data. The parent GCM simulations are (i) a *Last Millennium* simulation from MPI-ESM-P (Moreno-Chamarro et al., 2017a) for chapter 2, and (ii) an ensemble of the *1pctCO2* simulations from the MPI Grand Ensemble (Maher et al., 2019) for chapter 3. Both data sets comprise a plethora of simulation years, either in terms of a long-term simulation or through the ensemble approach. A large sample of extreme values is vital to quantify the variability of high extremes and to link them to climate states – or for simulations of future climate, to allow the detection of changes in the upper-end extreme values. A more detailed description of the respective experiments is given in chapters 2 and 3. Net sea level changes through the melting of land ice, vertical land movement, or – since employing a Boussinesq model – a change in ocean volume through the thermosteric effect, are not accounted for in either of the two scenario simulations. Only in chapter 3, section 3.4, I also provide projections of the combined effect of changes in ESL statistics and BSLR with the model data extended by literature estimates of the missing terms (see section 3.3.2) to set the future ESL change into context with the background sea level rise (BSLR) under future climate change.

1.4 UNDERSTANDING THE LONG-TERM VARIABILITY OF ESL

1.4.1 *Current knowledge and research gap*

In sea level research, studies on long-term variability and on the driving mechanisms of the observed variability mostly focused on variations in mean sea level (Church et al., 2013). The fewer studies on the long-term behavior of extreme sea levels have been based on the instrumental record from local tide gauges or on multi-decadal hindcasts of dynamical models and mainly centered around three core questions: (1) How do ESLs vary in time and is a trend emerging? (2) Do extreme values mainly follow the background sea level? Or if not, (3) are their variations dependent on specific modes of climate variability? These questions are not only important for understanding the past but are also vital to project a future change in ESL with climate change.

Regarding the first question, observational studies in the North Sea have pointed to a considerable multi-decadal variability in both BSL as well as in extremes, typically in terms of high percentiles (see Fig. 2; e.g., Weisse & Plüß, 2006; Weisse et al., 2009; Dangendorf et al., 2013b; Mudersbach et al., 2013; Marcos et al., 2015), with an increase towards the end of the century.

Regarding the second question if ESLs have mainly followed BSL variations, several studies using global tide gauge records have suggested that past ESL variability has primarily followed variations in regional BSL (Woodworth & Blackman, 2004; Marcos et al., 2009; Menéndez & Woodworth, 2010; Woodworth et al., 2011). However, analyzing sea level data from 13 gauges in the German Bight, Dangendorf et al. (2013b) found that linear extreme sea level trends in terms of high percentiles exceeded BSL trends in the second half of the twentieth century. Similarly, analyzing the Cuxhaven tide gauge record, Mudersbach et al., 2013 found that ESL and BSL have shown different trends during the second half of the twentieth century. That is, there is an indication that variations in storminess or local winds have influenced ESL variability on top of a gradual baseline shift in the recent past.

Regarding the third question, North Sea time-mean sea level variations have often been related to the NAO (Wakelin et al., 2003; Woodworth & Blackman, 2004; Woodworth et al., 2007; Marcos et al., 2009; Dangendorf et al., 2012; Ezer et al., 2016; Marcos & Woodworth, 2017). However, such a link to climate variability is more difficult to draw with ESL variations since they occur more rarely. A couple of studies have related ESLs in terms of high percentiles to climate variability along the U.S. coast (Park et al., 2010; Marcos et al., 2015; Wahl & Chambers, 2016) or with solar pacing in the Mediterranean (Barriopedro et al., 2010; Kaniewski et al., 2016). In the North Sea, links to the NAO (Woodworth et al., 2007) or NAO-like situations (Dangendorf et al., 2014c) have been proposed.

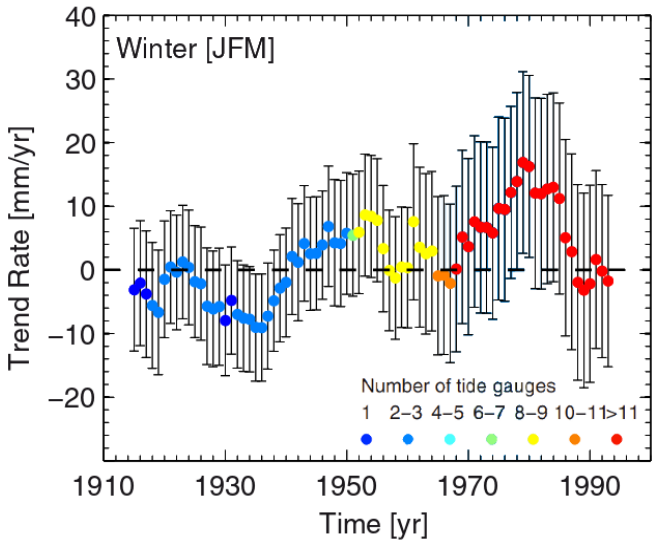


Figure 2 | **Variability of observed ESL trends in the German Bight.** Wintertime 30-year moving trends of the 99th percentile, reduced to BSL, for sea levels from a combination of tide gauge records in the German Bight. Colors indicate the number of tide gauges used for the trend computation. Adapted from Dangendorf et al. (2013b).

The core of the problem, though, is that the long-term variability of ESL has only been assessed (i) from the instrumental record which, depending on the location, covers a couple of decades up to 100 years, or (ii) from model hindcasts of the second half of the twentieth century. The limited record challenges a robust and spatially consistent quantification of long-term variability, detection of trends, and relation to climate modes. Thus, conclusions on multidecadal variability regarding the above questions are either uncertain or limited to relatively moderate extremes, such as high percentiles of annual sea levels, and to shorter modes of variability. For more extreme sea levels that have occurred only a couple of times during the instrumental record, there is the danger of under-sampling of the variability component.

The presence of multidecadal variability in ESL and/or BSL can also affect high-impact-low-probability extreme sea level estimates. Non-stationary extreme value analysis (e.g. Menéndez & Woodworth, 2010; Mudersbach & Jensen, 2010) can tackle this problem, but the resulting estimates are still only based on the variability of the last decades of the instrumental record. Low-frequency variability may thus not be fully sampled and, as a result, extreme value estimates biased by the state of long-term variability during the baseline period. This can lead to dangerous "unexpected" extreme events given the baseline record. For instance, Hurricane Katrina levels at New Orleans were by far the highest recorded, and previous 100-year return levels were up to 3 meters lower than the updated ones after the Hurricane (FEMA, 2005).

Thus, despite the wide literature on past and present ESLs and their variability up to several decades, there is still high uncertainty about variations in the margin of the distribution and their associated drivers. Specifically, what is lacking is (1) a quantification of ESL variability over long timescales and (2) a relation of ESL variations to large-scale climate variability. Identification of such relationships with variability in the climate system could not only improve our understanding of past and present-day extreme sea levels, but furthermore improve predictability of future ESLs.

1.4.2 *Research objectives*

The study in chapter 2 aims to close the above outlined gaps. Specifically, I pose the following research questions:

1. *How do ESL vary over longer timescales?*
2. *What are the driving mechanisms of variability?*
3. *Given the results from 1., how well can we estimate large return levels?*

To quantify the long-term variability of ESLs in the German Bight, I downscale a global Earth System Model simulation covering the last 1000 years with full transient forcing (*past1000* and *historical*, Moreno-Chamarro et al., 2017b). This approach provides an unprecedented long high-resolution sea level data record that allows to robustly quantify high extreme value statistics without relying on extreme value distribution fits and to identify associated large-scale forcing mechanisms in the climate system. At the same time, the regionally coupled model setup with a global ocean model (see section 1.3.2) allows for a consistent simulation of oceanic signals from the North Atlantic into the North Sea.

This study marks the first coupled downscaling focusing on sea level extremes and long-term behavior, and it is the first to simulate ESL variability during the entire last Millennium.

1.4.3 *Brief summary of results*

How do ESL vary over longer timescales?

The downscaling of the *past1000* simulation shows that, in its entirety, simulated ESLs reasonably capture observed ESL statistics from the Cuxhaven tide gauge record. I find that ESLs exhibit large variations on interannual to centennial timescales but on a white power spectrum. That is, the spectral amplitude is frequency-independent, and ESL do not cluster around preferred oscillation periods. This large variability is most pronounced for high-impact-low-probability events, leading to considerable scatter in corresponding return level estimates (see Fig. 3).

What are the driving mechanisms of variability?

The large ESL variability masks any potential influence of external volcanic or solar forcing during the last millennium and leads for the most part to a decoupling from BSL variations. However, on decadal timescales, periods of high ESL activity could be linked to patterns of internal atmospheric variability. The associated pressure pattern resembles a superposition of two known atmospheric modes of variability, namely the NAO and the *Scandinavian pattern*, creating a sea level pressure (SLP) dipole between the Gulf of Biscay and Northeastern Scandinavia and leading to northwesterly winds over the North Sea. This pattern is distinct to the NAO which has been found to govern the long-term variability of the time-mean sea level.

How well can we estimate large return levels?

Any extreme value estimate is associated with uncertainties. Traditional extreme value analysis addresses this by providing confidence bounds for the goodness of fit of an assumed extreme value distribution. However, such uncertainty bounds depend on the state of long-term variability of the underlying baseline record. The large scatter in high-impact-low-probability events, as found here, suggests that the uncertainty related to high-end extreme value estimates are larger than previously assumed. Thus, assessing extreme sea level estimates from short records leads to a considerable under-sampling of internal variability and potentially results in an underestimation of the upper-end uncertainty. For instance, the uncertainty of a 1000-year return level estimated from a combination of 100-year segments of the full simulated data (grey bar in Fig. 3) is considerably larger than an extreme value distribution fit based on one random 100-year segment such as the observational record (green bar) would suggest. This stresses the necessity to consider the large ESL variability in coastal protection. Likewise, a robust analysis of changes in future ESL statistics (Chapter 3) can only be achieved by using model simulations of large ensemble sizes.

These results are presented in detail in article I (Chapter 2). In short, the key findings are:

1. **Simulated ESLs show large variations on interannual to centennial timescales without preferred oscillation periods.**
2. **Periods of exceptionally high ESLs are associated with a SLP dipole between northeastern Scandinavia and the Gulf of Biscay.**
3. **ESL variations and existing high-impact-low-probability estimates are strongly influenced by internal variability. Large sample sizes are necessary to infer robust estimates of future changes in high ESLs amid the large variability.**

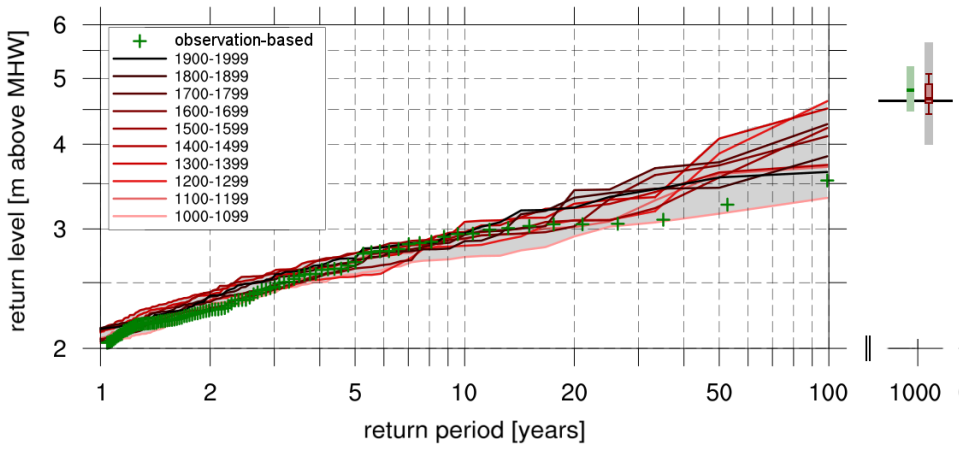


Figure 3 | **Variations of extreme sea levels in terms of return levels based on the downscaled *Last Millennium* simulation.** Return levels were produced using non-parametric plotting positions of simulated sea level at Cuxhaven. Colored lines represent 100-year long subsets of the full 1000 years; green crosses represent observations from tide gauges. The bars on the right mark the corresponding 1000-year return level estimates using a Gumbel fit to the observations (dark green, 95% confidence interval in light green) and to each 100-year subset of the simulation (red Box-Whisker, range of the 95% confidence interval in grey). The horizontal black line shows the 1000-year return level directly inferred from the full simulation. For details see Chapter 2.

1.5 ESTIMATING FUTURE CHANGES IN ESL STATISTICS

1.5.1 *Uncertainties in future extreme value estimates*

How will extreme sea level statistics change under increasing atmospheric greenhouse gas concentrations? Trivially, a rising BSL directly affects ESL heights by shifting the entire distribution and thus raising the baseline for extremes to act on (upper panels in Fig. 4). This already results in a substantial reduction of return period: Due to the log-linear relationship between flood height and its recurrence interval (see section 1.2), even a small increase in the background state can drastically augment flood frequency. This effect is strongest in regions where the sea level variability is low. As a result, many coastal protection standards that aim to accommodate future extreme sea level conditions have adopted a *climate buffer* (*Klimazuschlag*; MFLR, 2012), which simply adds a fixed level to the current probabilistic estimate of a high-impact flood return level.

However, while ESL changes are often handled by such a simple scaling of the sea level distribution with global or regional sea level rise estimates (e.g. Buchanan et al., 2017; IPCC, 2019), this mere shift of extreme values does not a priori hold for the future. Moreover, relative changes in the upper tail of the sea level distribution, e.g. through changes in storminess in a warmer climate (e.g. Woth et al., 2006; Rockel & Woth, 2007), can alter extreme value statistics on top of a simple shift to a higher baseline (see lower panels in Fig. 4). This might be especially important for coastal regions bounded by shallow continental shelf areas (Arns et al., 2017).

Both types of changes are associated with various uncertainties. First, constraining the rate of mean sea level rise remains a challenge due to uncertainties about (i) the relative contributions and interplay of different components (Miller & Douglas, 2004; Rietbroek et al., 2012), (ii) their regional importance, as sea level does not rise uniformly globally (Perrette et al., 2013; Slangen et al., 2014; Kopp et al., 2015), and (iii) the internal variability of these rates (Hu & Deser, 2013). While only volume changes (e.g. due to thermal expansion) and changes through the redistribution of water masses can be derived from state-of-the-art GCMs, other components like ocean mass changes (e.g. through melting of ice sheets and glaciers), resulting gravitational changes through the loss of ice mass, or geological changes, e.g. through the post-glacial rebound effect (also known as glacial isostatic adjustment, GIA) need to be estimated by complementary models. The estimated regional sea level change is thus often a result of the combination of different models and expert judgement. Especially the Antarctic ice sheet has been subject to considerable uncertainty (Hanna et al., 2013; DeConto & Pollard, 2016; Pattyn et al., 2018) and the bulk of the different upper limits of sea level rise estimates are a result of how ice sheet instabilities are handled (Slangen et al., 2017; Bamber et al., 2019; IPCC, 2019). The relative

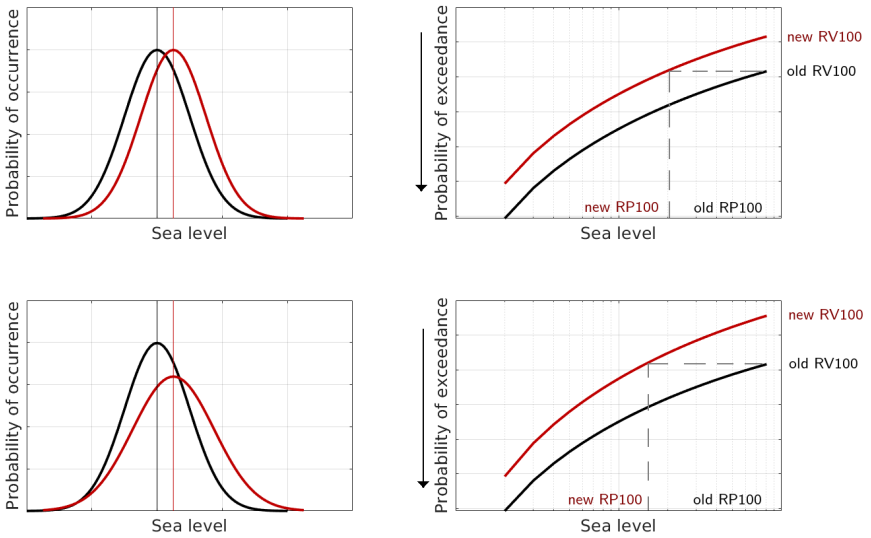


Figure 4 | **Extreme sea levels and sea level rise.** Sketch of the impact of internal variability changes on sea level extremes in a changing climate assuming no change (top) or an increase (bottom) in internal variability. Left panels show a probability distribution of sea level for current (black) a future (red) climate, the right panels show the corresponding return value plots.

contributions of the components also vary regionally, particularly regarding GIA associated with the last major deglaciation (Peltier, 2001; Steffen & Wu, 2011), and regarding ice loss from Greenland. In the southern North Sea, for instance, the imprint of the Greenland ice sheet is relatively low, as the addition of meltwater and the lowered gravitational effect due to the corresponding mass loss nearly cancel each other out (Cazenave & Llovel, 2010).

There is even less certainty about relative changes in the sea level distribution on top of a mean shift, e.g. through changes in the storm climate. Although most climate models point to a poleward shift of the major storm tracks (Beersma et al. (1997) and Fischer-Bruns et al. (2005); see Feser et al. (2015) for a review), uncertainties remain concerning regional patterns (Pinto et al., 2007; Rockel & Woth, 2007; Donat et al., 2011; Lehmann et al., 2014). Drivers of potential changes in extra-tropical storminess can be an enhanced amount of water vapor (Chang et al., 2002), changes in the meridional temperature gradient (Bengtsson et al., 2006) or changing sea surface temperatures (Bengtsson et al., 2009). Resulting changes in storm climatology, pathways, frequency, or duration can all affect ESLs and thus have consequences for coastal protection. Overall, due to the large spread in model simulations, the 5th Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) specifies only *low confidence* in regional projections of storm surges and storminess (Church et al., 2013).

1.5.2 *Existing studies and research gap*

Two main approaches have been used to simulate future changes in ESL statistics independent of sea level rise: the statistical and dynamical method.

The *statistical approach* is based on observed or simulated relationships between large-scale driving meteorology and local ESLs over the last decades. Changes in these large-scale drivers, obtained from GCM simulations, are then used to estimate future ESLs based on these relationships (von Storch & Reichardt, 1997; Langenberg et al., 1999), assuming that the relationships between predictor and predictand remain constant in past and future.

The widely used *dynamical approach*, on the other hand, does not rely on the assumption of linearity with future climate change. Typically, regional hydrodynamic models are employed. These are driven by the output from GCM or RCM simulations under future greenhouse gas (GHG) conditions; meteorological conditions and tides are prescribed at the open boundaries. The studies are designed as time slice experiments, typically covering 30 years of a present and future climate. Traditionally, these time slices comprise a single realization, although more effort has been made recently into better sampling internal variability, thus employing an ensemble of a couple of simulations as boundary forcing (e.g. Sterl et al., 2009; Howard et al., 2010; Gaslikova et al., 2013).

Studies with a global approach suggest that changes in ESLs differ by region (Feng et al., 2015; Marcos et al., 2015; Vousdoukas et al., 2017). For the North Sea, there is no clear answer either: While some studies found indications of a climate change-induced increase in storm surge extremes, primarily at the southeastern North Sea coasts (Langenberg et al., 1999; Lowe et al., 2001; Lowe & Gregory, 2005; Woth et al., 2006; Vousdoukas et al., 2016), other studies found only minor or no significant changes for the German and Dutch North Sea coastline (von Storch & Reichardt, 1997; Debernard & Røed, 2008; Sterl et al., 2009; Gaslikova et al., 2013). The usage of different terminology and metrics of ESLs, however, results in a low consistency between different studies and impedes direct comparison.

Moreover, this approach has important limitations because the design of the surge model with lateral boundaries does not adequately account for sea level variations outside the model domain that may travel into the modelling domain. Furthermore, the use of time slices combined with only one or a couple of realizations poses problems to the robustness of extreme value estimates and their changes, as it risks sampling states of long-term internal variability. Previous climate change experiments of limited length do not adequately sample internal variability of high ESL. Given the large scatter in extreme values (see Chapter 2), these studies can only yield robust conclusions for relatively moderate extremes. A large ensemble, however, would enable robust statements on changes of these higher ESLs. Given the quantified ESL uncertainty of chapter 2, the use of 30 independent realizations instead of a single 30-year period can lead to an effective increase in precision of the 100-year return value estimate by a factor of 10.

1.5.3 *Research objectives*

In Chapter 3, I thus pose the following research questions:

1. How do ESL statistics change under global warming?
2. What drives these changes?
3. How does this change relate to time-mean sea level rise?

To address these problems, I use the same regionally coupled climate model (Chapter 1.3.2) to downscale parts of the Max Planck Institute for Meteorology Grand Ensemble (MPI-GE, Maher et al., 2019). I rely on 32 realizations of the *1pctCO2* scenario simulations, since these compose the largest available set of simulations with the necessary high-frequency atmospheric output data required for downscaling. In the *1pctCO2* scenario, the atmospheric CO_2 concentration gradually increases over 150 years by 1% each year, from preindustrial levels to an approximate quadrupling. It is an idealized and, concerning current emission pathways, a rather drastic scenario, yet for the twenty-first century of "plausible magnitude" (Gregory et al., 2015). The global mean surface temperature (GMST) response lies with more than four degrees just above the end-of-the-century-response of the more commonly used high-forcing scenario of the IPCC's *Representative Concentration Pathways* (RCP) family *RCP 8.5* (Maher et al., 2019).

This study marks the largest ensemble study using a coupled climate model to investigate future changes in the distribution of sea level extremes. Such a high ensemble size substantially reduces the uncertainty of high ESLs and thus allows a more robust estimation of their changes than in previous studies. The focus is set on changes in ESL statistics on top of the expected gradual BSLR. Note that, as a Boussinesq model, MPIOM does not explicitly represent the thermosteric sea level rise and does not account for net ocean mass gain through melting of glaciers or land ice nor for relative changes due to vertical land movement. However, to separate the change in ESL statistics from the remaining dynamic background sea level changes, I here subtract the 30-year moving average. The choice of a long-term average rather than the more commonly used annual mean is motivated in order to still account for interannual to decadal low-frequency variability such as described above.

1.5.4 *Brief summary of results*

How do ESL change in the future?

My analysis shows that simulated ESLs increase for a wide range of return periods. The increase is highest along the western coastlines of the German and Danish North Sea coasts (see Fig. 5). Here, the climate change response of the pooled ensemble is up to +50 cm for high-impact return periods; single members can widely differ, with negative changes to changes

of over one meter. The direction of change and the spatial pattern are qualitatively consistent with the results of other studies using the traditional setup of a barotropic surge model forced by GCM data, but relying on less realizations (Vousdoukas et al., 2016) and targeting more moderate extremes, e.g. the annual 99th percentile (Woth, 2005). The here applied approach using a coupled model setup and a large ensemble sets the results in context of stochastic internal variability and large-scale climate variability. Through the large ensemble approach, the signal-to-noise ratio increases and the impact of climate change on ESLs can be seen in their pooled ensemble statistics, even for the typically rather uncertain high-impact-low-probability events. The transient character of the simulation furthermore suggests that multi-decadal variability is evident throughout the simulation, both within a single member as well as within the pooled ensemble. The ensemble spread is hereby so large that for individual members no or even negative trends can be found over essentially all timescales and return periods. That is, while the pooled ensemble reveals a tendency for a future increase in ESL statistics, individual members, or our world for that matter, may well follow different pathways.

What drives these changes?

A benefit from the global climate simulation approach is that changes in ESL statistics can be linked with states of variability in the climate system. The simulations have shown that the rise in ESLs is related to changes in atmospheric drivers with an enhanced storm activity over the North Atlantic and a subsequent increase of the local extreme wind speeds from westerly directions. This is consistent with the spatial pattern of ESL changes in the North Sea with largest climate change responses along the eastern coasts of the German Bight. Furthermore, results from a lagged Empirical Orthogonal Function (EOF) analysis (Weare & Nasstrom, 1982) suggest a strengthening of the main storm track type that is associated with the historically highest storm floods in the German Bight.

The change patterns in ESLs and their atmospheric drivers, though, have a unique seasonal character. While winter storminess and local westerly wind extremes increase, the opposite holds for the summer months. Again, this seasonally opposing response is consistent with the changes in ESLs and a corresponding confinement of ESLs to the main storm flood season in winter.

How does this change relate to time-mean sea level rise?

As global mean sea level is expected to continue to rise within the next decades to centuries (IPCC, 2019), a climate change induced rise in ESLs as shown above will act additionally on top of a gradual BSLR. Due to substantially different rates of BSLR and their individual components in different regions, conventional global mean estimates are not suitable and local estimates need to be considered.

Since MPIOM does not does not account for net sea level changes due to the melting of land ice or vertical land movement, missing components are added from gridded estimates of RCP8.5 end-of-the-century sea level rise from the Integrated Climate Data Center (ICDC, icdc.cen.uni-hamburg.de; Carson et al., 2016). Computed ESL changes from the *1pctCO2* scenario are here translated into RCP8.5 equivalents by scaling with GMST. Together, the different components amount to around 80-95 cm of BSLR in the German Bight, depending on the location. At the coastline, the magnitude of ESL-changes alone (without considering BSLR) reaches up to half of the region's end-of-the-century BSLR estimate. As a result, BSLR and changes in ESL together result in a strong increase in return levels of well above one meter and a corresponding reduction of the associated return periods of extreme events. The change factor is hereby – depending on the location – up to twice as high as if only BSLR were considered. For instance, what used to happen once every 50 years is projected to become an almost annual event under future climate change. This stresses the importance of considering both BSLR and changes in ESL statistics in the German Bight. A simple upward shift of the past sea level distribution could thus result in large and potentially costly underestimation of design heights.

These results are presented in detail in article II (Chapter 3). In short, the key findings are:

1. **ESLs increase over a wide range of return periods and especially along the western coastlines of Schleswig-Holstein and Denmark. This, however, is only indicated by the full pooled ensemble; individual members show a high variability.**

2. **ESL changes are seasonally opposite with a wintertime increase and a summertime decrease in ESL statistics.**

3. **The ESL changes are related to a strengthened North Atlantic storm belt and stronger westerly winds in the German Bight.**

4. **The magnitude of ESL changes at the German Bight coasts reaches up to half of the regional BSLR until the end of the century, thus greatly exacerbating the joint flood risk. ESL changes and BSLR should both be considered when estimating future flood heights and adaptation measures.**

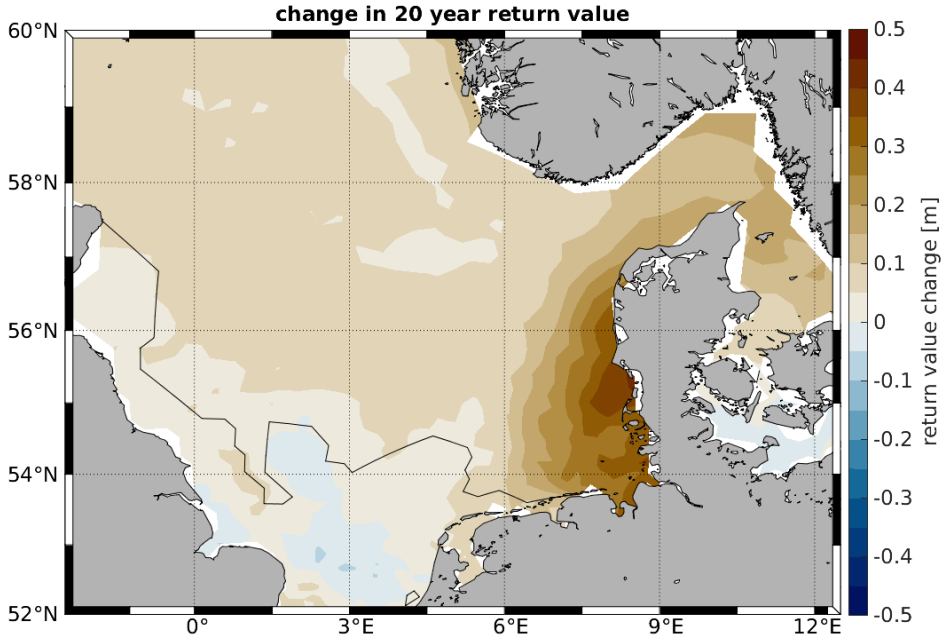


Figure 5 | **Climate change signal of ESLs in terms of 20-year return levels.** Colors show the change between high and low CO_2 states based on the pooled ensemble of downscaled *1pctCO2* simulations from the MPI-GE. The contour line marks areas significant on the 95% confidence level. See Chapter 3 for details.

1.6 SYNTHESIS & OUTLOOK

1.6.1 *Novelty of the study*

This study has contributed to the field with a more robust quantification of past and future extreme sea level variability and its atmospheric drivers based on a novel approach using long-term simulations from a regionally coupled climate model setup.

Compared to previous modelling studies using regional barotropic storm flood models, the use of a regionally coupled climate system model with a global ocean domain to downscale long-term transient or ensemble climate simulations has closed the gap between estimates of local surge extremes and aspects of large-scale climate variability as well as stochastic internal variability. This has been achieved by combining the respective advantages of global and regional models. At the same time, performing a long-term simulation or ensemble simulations has created a large sample of extreme values that allows a statistically more robust estimation of upper-end extreme sea level variability and their future changes than in previous dynamical modelling studies.

Downscaling a *Last-Millennium* simulation (Chapter 2), I have shown that ESLs vary substantially on interannual to centennial time-scales. Correspondingly, associated uncertainties of high-end extreme value estimates based on short records with a couple of decades length are likely to be underestimated. Furthermore, due to the large internal variability, externally induced natural fluctuations in the past or trends from anthropogenic climate change are masked and can only be statistically detected with a large ensemble approach. Employing such a large ensemble approach for climate change experiments (Chapter 3) of the *1pctCO2* scenario has revealed that despite largely differing signals of individual members, ESLs statistically increase with rising CO_2 emissions, with local signals of up to half of current estimates of end-of-the-century mean sea level rise. That is, not only a rise in time-mean BSLR but also a change in ESLs is responsible for an increase in absolute extreme sea levels which is often not accounted for. Thus, to project future changes in absolute sea level heights, changes in ESL and BSLR should thus be jointly examined. I could further link the long-term characteristics of extreme sea levels to variability in the climate system on scales from local, instantaneous wind forcing to more general large-scale patterns that lead to an enhanced ESL activity. Finally, the analysis of future changes in these drivers has drawn a more complete and consistent picture of the associated changes in ESL statistics.

A detailed discussion on the limitations of the study concerning the experimental design is already given in chapters 2 and 3. In the following, I thus focus on discussing some overarching considerations and implications of my results.

1.6.2 *Added value of a climate modelling perspective for coastal protection*

With its large-scale modelling approach, this work is not designed as a local coastal protection study. The, for local impacts, comparably coarse resolution cannot represent the finer site-specific bathymetry and geography of the coastline and harbor basin. Moreover, wind waves, especially relevant from an impact perspective, are not accounted for but are – since BSLR leads to a reduction of the frictional damping effect (Niemeyer, 1983) – also projected to increase under climate change (e.g. Arns et al., 2017).

Nonetheless, the benefits of this dissertation rather reflect its unique climate modelling point of view and a focus on longer time scales as well as a mechanistic understanding of ESL variability. Only the large sample of physically consistent ESLs allows the quantification of their long-term variability, the metric which ultimately suggested that the associated uncertainties of current ESL estimates may be substantially larger than previously assumed. This also implies that coastal protection design heights may be too conservative. Furthermore, only through the hybrid climate model approach, the local ESL variability could be set in context with large scale climate modes. With these aspects, it becomes evident that the climate modelling perspective can produce important information that can be of benefit for a local coastal protection perspective. That is, the findings can be applied to smaller scales: The model output can be further regionalized to e.g. scales of individual estuaries or harbor facilities by other models which do account for the missing coast-relevant processes. Finally, as the *1pctCO2* scenario is a rather strong but not necessarily implausible emission scenario, it can even represent a valuable upper bound estimate.

1.6.3 *Planning with deep uncertainties*

Concerning coastal protection and the estimation of ESL for the design of flood standards, two important results can be drawn from this study: First, due to a large ESL variability, the uncertainty of a single extreme value estimate based on the instrumental record may be biased low (1). And second, although showing a large scatter in their response, ESLs are statistically expected to increase with climate change, additional to a gradual increase related to BSLR (2).

Concerning (1), the quantification of the internal variability component of ESLs from 1000 simulation years in chapter 2 has highlighted the importance of a large sampling size: Due to the large scatter in high-impact-low-probability events, studies relying on small samples risk under-sampling the internal variability of extreme values. Consequently, they risk misinterpretations, as corresponding extreme value statistics can strongly depend on the state of long-term variability. That is, traditional extreme value estimates may suffer from an underestimated uncertainty component and, as a result, flood protection design heights may be rather conservative.

Concerning (2), no robust statement on changes in future ESL statistics amid the large variability would be possible without the use of large ensemble simulations. Relying on 32 ensemble members instead of the traditional single 30-year time slices improves the signal-to-noise ratio and allows to robustly detect changes in upper extreme value statistics. Concerning such future changes, I have shown that not only BSLR but also changes in ESL statistics are responsible for a rise in absolute sea level heights, which are often not considered when assessing climate related flood impacts. This strongly suggests that changes in *both* ESL and BSL need to be considered when estimating future flood heights and adaptation measures.

However, BSLR and changes in ESL are both associated with deep uncertainties. The rate of future BSLR and ESLs is not only dependent on the emission scenario but also related to our understanding of the relevant mechanisms on global and local scale. Concerning BSLR, high uncertainties are i.a. associated with the contribution of Antarctica, especially the magnitude of the *marine ice sheet instability* (IPCC, 2019). Thus, estimates of BSLR alone can differ substantially, especially concerning upper end BSLR. Concerning ESL changes, the response of the large-scale wind climate is particularly uncertain and has shown to be model dependent (Donat et al., 2010a; Mölter et al., 2016).

In flood protection, the use of a *climate buffer* is a measure of addressing such uncertain future changes in sea level statistics. For instance, the federal state of Schleswig-Holstein, Germany, adds half a meter *Klimazuschlag* (MFLR, 2012) to the ESL estimates to account for a change in ESL statistics due to BSLR based on global numbers of the IPCC AR5 (Church et al., 2013). This might be enough as a first-step adaptation, but the combined effect of BSLR and changing ESL statistics of more than one meter at the end of the century or even higher values for individual realizations, as found in the pooled ensemble, greatly surpasses this estimate.

Finally, an important albeit irreducible source of uncertainty arises from internal variability. These fluctuations are related to the chaotic nature of the climate system; even with a perfect model, this variability component remains due to the impossibility of accurately measuring the current states of all relevant components of the climate system. While the results from the full ensemble do reveal a tendency towards an enhanced upper tail of the sea level distribution which significantly amplifies the expected change due to BSLR, the response of each individual realization can differ greatly from the ensemble response. This uncertain response has important implications for adaptation planning, since the future pathway of our world's climate can essentially be viewed as one of the different realizations. As manifestations of the large internal variability, (multi-)decadal deviations from the statistical long-term trend are possible and may in fact mask the statistically expected increase in ESL in the short term, or worse, exceed an expected ESL pathway. This is especially evident for the upper tail of the ESL distribution, i.e. for floods of return periods of decades and longer.

As a result, coastal protection inevitably has to deal with deep uncertainty. This uncertainty related to ESL statistics is expected to be most important in the near future, when the internal climate variability exceeds the uncertainty of the emission scenarios (Wahl et al., 2017). In the far future and for sufficiently strong emission scenarios, these uncertainties will most likely decrease relative to those related to the BSLR, especially the contribution from ice sheet dynamics and, ultimately, the emission pathway. However, in the very near-term, additional predictability may be introduced through the relation to large-scale atmospheric modes that enhance the likelihood of ESL occurrence. How predictable such atmospheric patterns can ultimately be, though, remains contentious (e.g. Weisheimer et al., 2019).

In the light of these uncertainties, adaptive planning will need to be able to tackle a wide range of possible ESL pathways. As Weisse et al. (2014) notes, optimal adaptation strategies comprise those that either cost-efficiently work under many different scenarios (*robust strategies*) or which can easily be adopted and modified (*flexible strategies*) when climatic conditions or our scientific understanding change.

JOURNAL ARTICLES

The next chapters present the results of the two research questions on past and future storm flood variability. They comprise the following articles:

Lang, A. and Mikolajewicz, U. (2019): The long-term variability of extreme sea levels in the German Bight. *Ocean Science*, 15, pp.651–668. DOI: 10.5194/os-15-651-2019.

Lang, A. and Mikolajewicz, U. (under review): Rising extreme sea levels in the German Bight under enhanced CO2 levels – a regionalized large ensemble approach for the North Sea. *Climate Dynamics*.

2

LONG-TERM VARIABILITY DURING THE LAST MILLENNIUM

This chapter has been published as:

Lang, A. and Mikolajewicz, U. (2019): The long-term variability of extreme sea levels in the German Bight. *Ocean Science*, 15, pp.651–668. DOI: 10.5194/os-15-651-2019.

Abstract. Extreme high sea levels (ESLs) caused by storm floods constitute a major hazard for coastal regions. We here quantify their long-term variability in the southern German Bight using simulations covering the last 1000 years. To this end, global Earth System Model simulations from the PMIP3 *past1000* project are dynamically scaled down with a regionally coupled climate system model focusing on the North Sea. This approach provides an unprecedented long high-resolution data record that can extend the knowledge of ESL variability based on observations, and allows for the identification of associated large-scale forcing mechanisms in the climate system.

While the statistics of simulated ESLs compare well with observations from the tide gauge record at Cuxhaven, we find that simulated ESLs show large variations on interannual to centennial timescales without preferred oscillation periods. As a result of this high internal variability, ESL variations appear to a large extent decoupled from those of the background sea level, and mask any potential signals from solar or volcanic forcing. Comparison with large-scale climate variability shows that periods of high ESL are associated with a sea level pressure dipole between northeastern Scandinavia and the Gulf of Biscay. While this large-scale circulation regime applies to enhanced ESL in the wider region, it differs from the North Atlantic Oscillation pattern that has often been linked to periods of elevated background sea level.

The high internal variability with large multidecadal to centennial variations emphasizes the inherent uncertainties related to traditional extreme value estimates based on short data subsets, which fail to account for such long-term variations. We conclude that ESL variations as well as existing estimates of future changes are likely to be dominated by internal variability rather than climate change signals. Thus, larger ensemble simulations will be required to assess future flood risks.

2.1 INTRODUCTION

Inundation due to storm floods is one of the major geophysical risks in coastal regions and bears high damage potential for coastal environments, in both natural and socio-economic terms. This is especially important for low-lying regions such as coasts and estuaries of the Southern North Sea and, in particular, the German Bight. Situated between Denmark to the North and the Netherlands to the West, the German Bight is a shallow shelf sea under the influence of the major northern hemispheric storm-track paths. At the same time, the shallow water depths of under approximately 40 meter (see Fig. 2.1) in combination with the basin's geometry lead to relatively strong tides, with a maximum tidal range of around 4.5 m. Thus, storms can generate particularly high floods in this region. The region has seen many devastating storm floods in the past; one of the most severe, the great storm flood 'Grosse Manndrenke' (the Great Drowning of Men) in 1362 resulted in death tolls of tens of thousands, the destruction of numerous settlements including the disappearance of the legendary town 'Rungholdt' (Heimreich, 1819) and shifts in the Wadden Sea coastline (Hadler et al., 2018). More recently, the great flood in 1962, resulted in a high death toll and vast economic loss along the coastal regions of Germany and in particular the city of Hamburg. Disasters like these have driven extensive research in the field; yet, most studies focused either on individual events, on the observed trend during the last couple of decades, or a projection of future storm flood exposure. Variations on longer time scales and their underlying mechanisms have received less attention. However, a deeper understanding of the long-term variability of strength and occurrence of extreme storm floods can be of great importance for coastal planning and risk assessment. Here we assess the longer-term variability of high sea level extremes in the German Bight using a novel regionally coupled model approach over a 1000-year long simulation period.

Following Pugh (1987), the sea level at a certain point and time can be decomposed into three factors: Meteorological surge, astronomical tide and background sea level (BSL). The *surge* is the "dynamic response of the sea surface to forcing by the atmosphere" (Mawdsley & Haigh, 2016) and can consist of (i) the *local wind surge* that pushes water against the coast, (ii) an *external surge* generated in the North Atlantic by fast bypassing cyclones and air pressure variations that travel as a Kelvin wave counterclockwise in the North Sea (Gönnert et al., 2012), and (iii) the direct effect of air pressure acting on the water surface (the *inverse barometric effect*). The total sea level is thus a product of interactions between the surge component, the astronomical tide and the underlying longer term change in BSL, the latter of which depends on various oceanographic and atmospheric processes such as coastally trapped waves, local steric effects or longshore winds (Calafat et al., 2012; Dangendorf et al., 2014b; Sturges & Douglas, 2011). Extreme sea

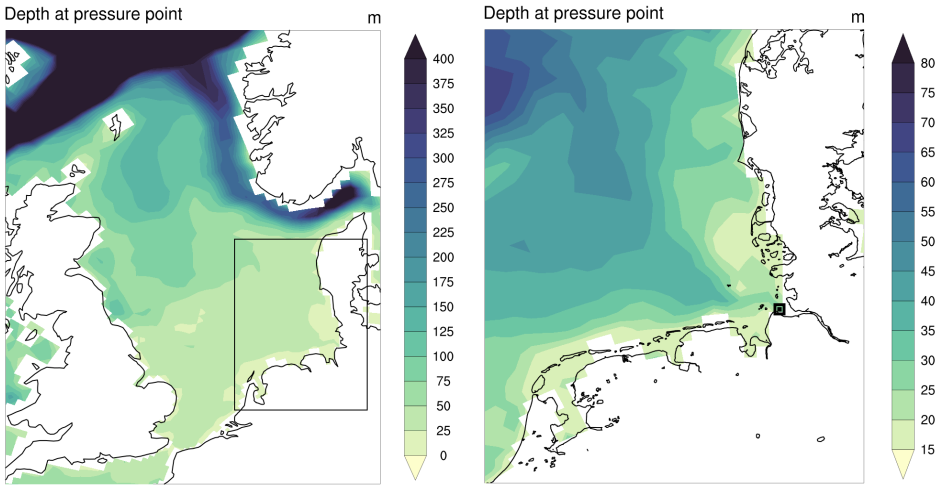


Figure 2.1 | Bathymetry of the North Sea (left) and the German Bight (right; the main study location Cuxhaven is marked by the black square) as represented in MPIOM. The model's land mask is shown in white, the present-day coastline is shown in black.

levels particularly arise when these components are in superposition, e.g. if a strong wind surge occurs concomitantly with a tidal maximum, or – as a result of the tide-surge interaction – a few hours before on the rising tide (Horsburgh & Wilson, 2007). Topographic features, such as water depth, sand bars or reefs can further affect their height locally. These components interact non-linearly (e.g., Kauker & von Storch, 2000; Plüß, 2004) and are non-stationary, with variations occurring on multiple timescales. Long term changes in any of those components, e.g. due to internal variability or external climate change, may substantially alter the risk associated with sea level extremes. Since we are solely interested in high water extremes, we refer with the term *extreme sea level* (ESL) only to the upper end of the distribution.

Many studies have investigated the recent dynamics of ESL both regionally and globally using data from tide gauges (e.g., Marcos et al., 2009; Menéndez & Woodworth, 2010; Mudersbach et al., 2013; Tsimplis & Woodworth, 1994; Wahl & Chambers, 2015; Weisse et al., 2014; Woodworth & Blackman, 2004; von Storch & Reichardt, 1997) or using barotropic surge models for hindcasts of storm-surges (e.g., Kauker & von Storch, 2000; Langenberg et al., 1999; Weisse & Plüß, 2006; Woth, 2005). For a review of past storm surge statistics and projected changes in the North Sea region see Weisse et al. (2012). While most studies agreed on an overall increase in storm surge activity along the German Bight since the 1960s (e.g., Mudersbach et al., 2013; WASA-group, 1998; Weisse & Plüß, 2006; von Storch & Woth, 2008), both observations and hindcast simulations over a longer time span have set this recent trend into the perspective of a marked multidecadal variability during the last century (e.g., Dangendorf et al., 2013a;

Mudersbach et al., 2013; WASA-group, 1998; Weisse & Plüß, 2006) with higher values at the beginning and end of the century.

The data record, though, is limited and only few high-frequency tide gauge records (e.g. Cuxhaven) date back more than a couple of decades. Thus, conclusions on multidecadal to centennial variability as well as the separation of longer-term fluctuations from the transient sea level rise are difficult from a statistical point of view. Concerning the latter, different studies on German Bight sea level reported similar ESL and BSL trends (e.g. Kauker & Langenberg, 2000) or trends at rather different rates (e.g., Dangendorf et al., 2014c; Mudersbach et al., 2013) – a question of great importance for estimations of future ESL behavior on top of a gradual sea level rise. Further, as ESL are by definition rare, the attribution to modes of climate variability, which often operate on similar or even longer timescales, is hampered by the short instrumental record. While many studies have related mean sea level variations to modes of climate variability, esp. the North Atlantic Oscillation (NAO) (e.g., Dangendorf et al., 2012; Ezer et al., 2016; Wakelin et al., 2003), the dominant pattern of atmospheric variability over the North Atlantic (Hurrell, 1995), which showed coherent trends during the last decades, only a few have set this in context with ESL variations (Marcos & Woodworth, 2017; Marcos et al., 2009; Woodworth et al., 2007; Woodworth & Blackman, 2004). Finally, such long-term ESL fluctuations can also have important implications for storm flood protection measures. The design of coastal defense structures in Germany is based on deterministic or statistical approaches (e.g. MFLR, 2012). For the latter, water levels with assigned return periods are needed, which are typically based on parametric extreme value analysis of observed sea level data. Yet, due to the relatively short tide gauge records, the quality of the estimation of return periods longer than the investigated sea level data series depend on the type of extreme value distribution and its goodness-of-fit to the data. Additionally, any longer-term variability in ESL further complicates the estimation of high return levels, as they depend on the state of long-term variability during the underlying baseline period. Here we argue that significant centennial variations in high-impact return levels entail a large source of uncertainty for parametric ESL estimates. The standard approach using a typically short baseline period for such sea level estimates thus fails to reflect the possible range of most extreme events that happen only once or twice during that period.

A longer, high-frequent data series as obtained from a climate simulation can offer a statistically more robust assessment of these problems, as the full time series can be treated as an ensemble of data series comparable in size to the observational record. However, currently available long-term climate simulations do not include tides and have an insufficient resolution to realistically represent storm surges in the North Sea. Dynamical surge models or regional climate models, driven by global climate model simulations, can provide a better representation of small-scale processes, topographic influences and land-sea contrasts, and are thus better suited for the simulation

of extreme events. Due to their open lateral boundaries, however, they cannot account for a consistent propagation of external signals into the study region, which has been shown to affect North Sea sea level variability (e.g., Chen et al., 2014). Here we employ a global ocean model with regionally high horizontal resolution, which allows a consistent simulation of signals propagating from the open Atlantic onto the North West European shelf, coupled to an atmospheric regional model to dynamically downscale the climate variations from a *Last Millennium* simulation from MPI-ESM (Jungclaus et al., 2014; Moreno-Chamarro et al., 2017b). With a long-term simulation, this study allows a non-parametric approach in estimating such high return levels and can thus give insight into the uncertainties in extreme value analysis when based on short records. The goal is to describe and understand the (multi)decadal variability of ESL in the German Bight as well as their relationship with BSL and large-scale climate variability, which to our knowledge has not yet been investigated in a model study over such long timescales.

This article is structured as follows: Section 2.2 introduces the model system and experimental setup. In section 2.3, we present results on ESL variability, including the validation of the model with respect to observations from tide gauges (2.3.1), the relation to BSL (section 2.3.2) and climate variability (section 2.3.3). The results on ESL variability as well as some implications for observation-based extreme value estimates are discussed in section 2.4. Finally, in section 2.5 we close with a summary and conclusions for coastal defense measures.

2.2 METHODS

2.2.1 *Model system and experimental design*

This study employs a regionally coupled climate system model, consisting of the global ocean model MPIOM (Jungclaus et al., 2013; Marsland et al., 2004) and the regional atmospheric model REMO (Jacob & Podzun, 1997). Both models are interactively coupled over the wider EURO-CORDEX domain (e.g., Jacob et al., 2014) with the coupler OASIS-3 (Valcke, 2013). The coupled model system has been described in Elizalde et al. (2014) and Mikolajewicz et al. (2005) and Sein et al. (2015); a sketch is shown in Fig. 2.2.

REMO is run with a 0.44 degree setup, corresponding to approx. 50 km grid spacing, and with 27 vertical levels covering Europe, northern Africa and the northeast Atlantic. MPIOM is run on a stretched grid configuration with a nominal horizontal grid resolution of 1.5° and 30 vertical layers. In order to maximize the grid resolution in the study focus area, the model's poles are shifted to Central Europe and North America, respectively. This results in a maximum grid resolution of under 10 km in the German Bight and thus enables a more realistic simulation of small-scale shelf processes.

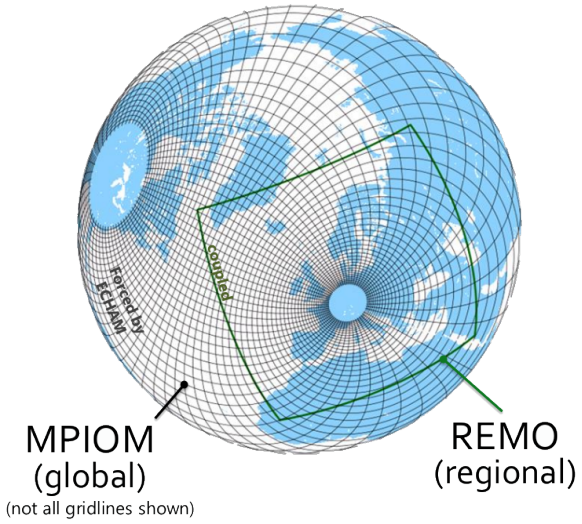


Figure 2.2 | Coupled model setup, consisting of MPIOM (black) and REMO (green).

Further, it includes the full luni-solar ephimeridic tidal potential according to Thomas et al. (2001). At the same time, the ocean model's global domain without lateral boundaries allows the full simulation of signals propagating from the open Atlantic into the North Sea. In order to prevent ocean gridpoints to fall dry due to strong tidal sea level variations, as for example in the English channel, MPIOM's uppermost layer thickness is set to 16 meter. This model setup is identical to the one used in Mathis et al. (2018).

We employ this model setup to downscale transient coupled climate simulations performed with the paleo-version of the global Earth System Model MPI-ESM (Max-Planck-Institute Earth System Model) in its low resolution (LR) version (Giorgetta et al., 2012), corresponding to 1.875° or approx. 200 km grid spacing in the atmosphere. The parent global simulations cover the period 900–2000AD and comprise parts of the PMIP3 simulation *Last-Millennium* (850–1850AD, Jungclaus et al., 2014; Moreno-Chamarro et al., 2017b), extended with the corresponding CMIP5 "historical" simulation (1850-2005AD; Taylor et al., 2012), including all relevant transient forcings. Greenhouse gases (GHG) follow PMIP3 protocol (Schmidt et al., 2012), solar irradiance is prescribed after Wang et al. (2005) and volcanic eruptions in terms of radiation imbalance after Crowley et al. (2008) (see Supplementary Fig. A3). 6-hourly atmospheric forcing derived from the atmosphere component of the driving GCM is used as lateral boundary conditions for REMO, or as surface forcing for MPIOM outside the coupling domain, respectively. Topography and coastlines as well as ice sheets are immutable, and thus transient sea level modulations due to ice sheet melt or changes in coastal morphology including land sinking or lifting are not accounted for. Furthermore, as a Boussinesq model, MPIOM conserves volume rather than mass and does therefore not explicitly represent the thermosteric sea level rise.

The downscaling was performed as one continuous run with hourly coupling started in year 900AD, the first 100 years are used as a spin-up. Sea level and selected atmospheric fields are stored at hourly resolution. Additionally, in order to quantify the variability from the downscaling process, as well as the contributions of natural variability and external forcing, we performed two further downscalings, one of the same global realization and one of another ensemble member of the global *Last-Millennium* simulations, each starting in 1400AD, with the first 100 years again used as spin-up (see Supplementary Material A3). The implications of these additional downscalings are discussed in the respective sections; yet, for simplicity, we only show results from the 1000-year continuous downscaling.

2.2.2 *Extreme value sampling*

Several techniques have been used to characterize extreme sea level, e.g. the use of high percentiles (e.g., Dangendorf et al., 2013b; Woodworth & Blackman, 2004), the selection of *r-largest* maxima over a block of time (e.g., Araújo & Pugh, 2008; Marcos et al., 2009; Méndez et al., 2007), and the selection of peaks over a certain threshold (*POT*) (e.g., Méndez et al., 2006). While the choice of the respective percentile, threshold, block length or number of block maxima is essentially arbitrary, the resulting events are sensitive to the choice of extreme value sampling index which represents a trade-off between bias (too high r or too low threshold) and variance (too low r or too high threshold) of the estimates.

Here, we have chosen the annual maximum sea level as an index representing ESL. Due to its relative definition of what constitutes an extreme it is robust to temporal variations. The use of a 'direct' method for ESL (instead of e.g. *surge residual* or *skew surge*) is chosen in order to avoid a decomposition between tidal and surge parts and their nonlinear interaction. Since storm floods primarily occur in winter, annual statistics are computed for years defined as starting in July and ending in June in order to not split up one storm flood season. If not specified otherwise, these definitions are used when referring to ESL in the text.

For the design of coastal defense structures, policy makers and adaptation planners often require statistics of water levels of a certain assigned return period (e.g. MFLR, 2012), especially those of high impact but low probability, i.e. the upper tail. The return periods and associated exceedance probabilities are typically estimated based on parametric extreme value analysis of the available instrumental data record. As data records rarely date back more than a couple of decades, this implies a substantial extrapolation of the data. That is, in order to obtain estimates for large return periods, different extreme value distributions are fitted to the comparably short data series. The choice of extreme value distribution depends on the considered extreme value sampling method and ultimately represents a tradeoff between bias and variance of selected extremes. While the *POT* method is

linked to the Generalized Pareto distribution, the r-largest samples follow approximately a three-parameter generalized extreme value (GEV) distribution (e.g., Coles et al., 2001). Its cumulative distribution function is

$$F(z;\mu,\sigma,k) = \begin{cases} e^{-(1+k(\frac{z-\mu}{\sigma}))^{-1/k}} & k \neq 0 \\ \\ e^{-e^{-\frac{z-\mu}{\sigma}}} & k = 0 \end{cases} \tag{1}$$

where F is the probability that a water level z will not be exceeded, while μ, σ, and k are the location, scale and shape parameters, respectively. The special case for $k = 0$, $k < 0$ and $k > 0$ represent three extreme value families, namely the Gumbel (type I), Weibull (type II) and Frechet (type III) distribution (Coles et al., 2001). From Eq. 1, the probability of exceedance is $E = 1 - F$, where $E(z)$ represents the expected frequency of events exceeding z. The expected time-interval between events of level z or greater, the return period RP, is

$$RP(z) = 1/E(z) \tag{2}$$

The advantage of our long simulation period is that it allows us to infer high-end extreme value statistics without the use of parametric methods. These non-parametric estimates have been inferred by first ranking the data points of the sea level time series and associating a cumulative probability to each value. The probability of exceedance is $P = \frac{m}{(N+1)}$, where m is the rank of N observations ordered in decreasing order. Following Eq. 2, return periods are again defined as the reciprocal of the respective probability of exceedance.

2.3 RESULTS: EXTREME SEA LEVEL VARIABILITY

The simulated timeseries of ESL at Cuxhaven (see Fig. 2.1) over the last 1000 years is shown in Fig. 2.3 (black curve). For comparison, we show the frequency of storm floods as events binned per decade (blue), following the storm surge definition from the Federal Maritime and Hydrographic Agency (BSH) (Müller-Navarra et al., 2003). Heavy (extreme) storm floods correspond to elevations above 2.5 (3.5) meter, relative to the long-term mean high water level (MHW). The BSL as a reference in terms of winter median is shown in green. Sea level is given in meter above the long-term mean, the model location of Cuxhaven refers to its nearest gridpoint.

Simulated ESL range between 2 and 5.5 m above the long-term mean. With a standard deviation of 50 cm and a maximum year-to-year amplitude of roughly 3 m, ESL exhibit large interannual variations and pronounced variability on various timescales. Yet, the highest events occur without pronounced clustering throughout the full 1000 years. This variability is analyzed in more detail below.

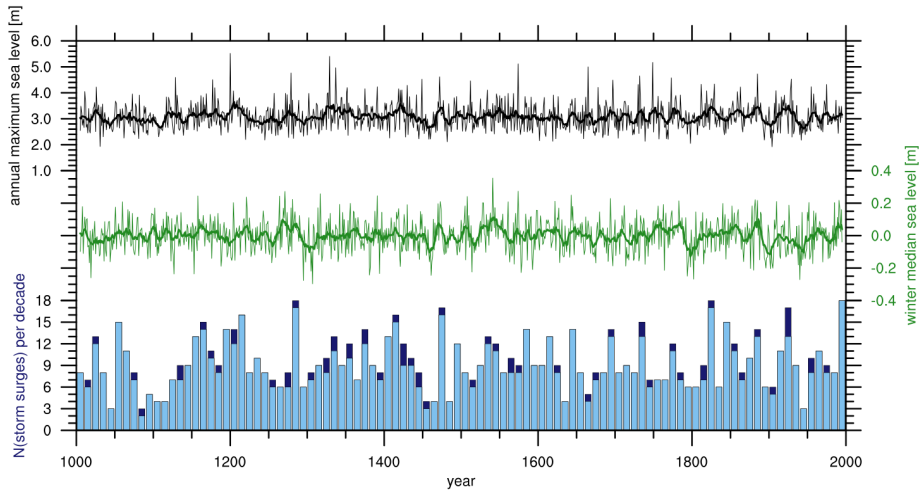


Figure 2.3 | Simulated ESL (black), winter median sea level (green) as well as number of heavy (blue bins) and extreme (dark blue bins) storm surges per decade at Cuxhaven. Thick lines denote the 11-year running mean.

2.3.1 *Validation of simulated storm surge statistics*

As the reality can be viewed as only one realization of the climate system one should not compare individual historic events in simulation and observation, but rather their statistics. To validate the model's performance considering storm flood statistics, we compare the simulated ESL with observations from the tide gauge record in the German Bight (data from AMSeL project, see Jensen et al., 2011; Wahl et al., 2011), and specifically show results at Cuxhaven which, starting in 1900 marks one of the longest reaching records of the German Bight tide gauge stations. The long-term (linear) trend in the tide gauge data has been removed.

While the general North Sea tidal oscillations are well reflected in the model, the tidal range is underrepresented at Cuxhaven (see Supplementary Fig. A1 for a comparison of the broader highest water level from observations and model simulation, respectively). Accordingly, this results in a lower long-term MHW (defined as the time-mean over tidal maximum values). Relative to the respective long-term MHW, however, simulated and observed ESL compare well (see the quantile plot in Fig. 2.4 for ten 100 year-long segments of the simulation against the Cuxhaven tide gauge record.), and we therefore express ESL in such relative terms in the remainder of the study.

Figure 2.5 shows return values of ESL compared to the Cuxhaven tide gauge record. In order to better compare the different data record lengths, the simulated 1000 years of data are again split into ten 100 year-long segments, roughly the length of the Cuxhaven tide gauge record. Both slope and magnitude above MHW are well-captured; the return values inferred from observations lie within the spread of the model simulation at all peri-

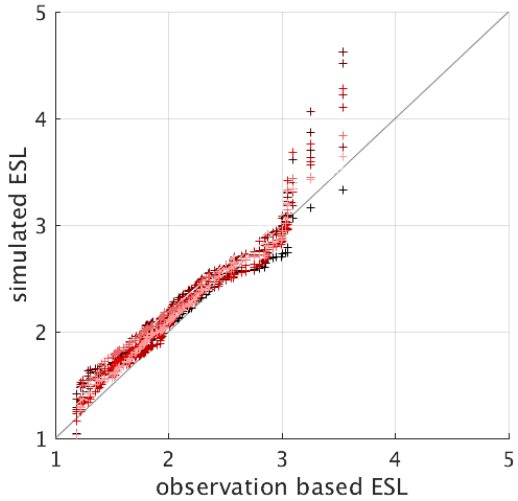


Figure 2.4 | Quantile-quantile plot of simulated extreme sea level at Cuxhaven against the 100-year long ESL from the tide gauge record. Colors following the gradient from light red to black represent ascending 100-year segments from 1000-2000.

ods. Yet, other than in the simulation, observations tend to level off for return periods above 20 years, before they rise again for return periods higher than 50 years, while simulated return levels rather increase steadily. As a consequence, the 100-year return levels (RL_{100}) in all but the 11th century exceed the corresponding return water level inferred from the instrumental record. The large scatter of about 1.2 m between the highest simulated sea levels of each century has important implications for extreme value analysis (see section 2.4).

The spatial structure of simulated return values (Fig. 2.6) shows lowest values for open waters which increase towards the coast, especially the inner German Bight. Most points along the German Bight coast (circles represent Cuxhaven, Husum, Norderney and Delfzijl (Netherlands)) compare well with the respective tide gauge records. Yet, while the return values at Cuxhaven lie slightly higher than the observed ones, ESL along the coastline of Lower Saxony and the Netherlands are rather low compared to the observation-based estimates. Furthermore, the model's uppermost layer thickness is with 16 meter well above the shallow shorelines of the Wadden Sea (see Fig. 2.1) and thus likely leads to lower surge heights. Note that for Husum and Norderney the observational record does not date back the full 100 years, so these points are only shown for the 50-year return period.

The seasonality is well-captured, with strongest and most frequent storm floods in winter (Supplementary Fig. A2), especially in the months of October to January. However, the distribution is slightly shifted towards autumn and early winter.

In agreement with observational studies (Gerber et al., 2016), simulated storm floods at Cuxhaven stem from predominantly west-north-westerly di-

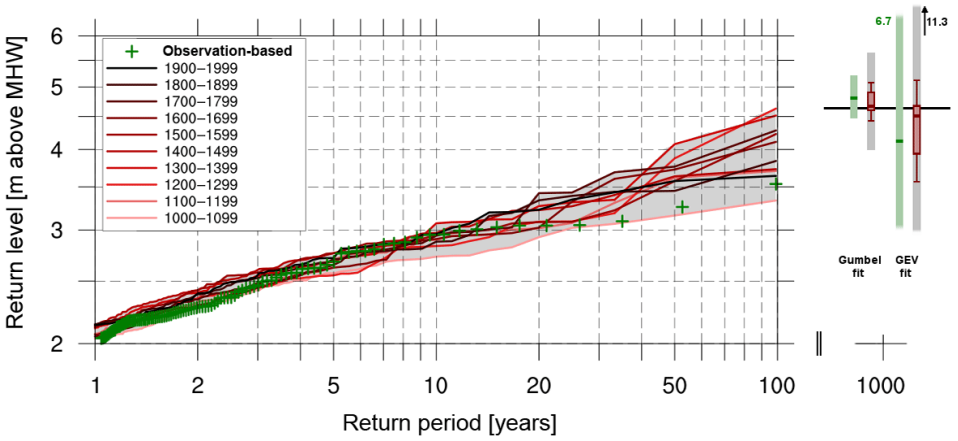


Figure 2.5 | Return value plot using non-parametric plotting positions of simulated sea level at Cuxhaven (colored lines represent 100-year long subsets of the full 1000 years) against observations from tide gauges (green crosses). The bars on the right mark the corresponding RL_{1000} estimates using a Gumbel (left) and GEV (right) fit to the observations (dark green, 95% confidence interval in light green) and to each 100-year subset of the simulation (red Box-Whiskers, range of the 95% confidence intervals in grey). The horizontal black line shows the RL_{1000} directly inferred from the full simulation.

rections, while their associated daily pressure anomaly patterns (not shown) are similar to observations of storm flood weather situations (Dangendorf et al., 2014a; Donat et al., 2010b).

We thus conclude that the model reasonably reproduces storm flood physics and statistics. Due to the good skill in reproducing storm surge statistics as well as the comparability owing to its relatively long instumental record, the remainder of this study will focus on Cuxhaven only. However, the temporal variability of other gridpoints along the German Bight is similar (see Supplementary Fig. A6) and the here discussed results qualitatively agree irrespective of the exact location.

2.3.2 *Relation to the background state*

An important question concerning ESL variability as well as future ESL projections is the relation to time-averaged sea levels, i.e. the background state. Separation into the different components of extreme water levels, such as the subtraction of mean and tide are useful methods to investigate this question (e.g. Woodworth & Blackman, 2004). Reviewing literature on recent ESL trends globally, Woodworth et al. (2011) concluded that the majority of studies suggest an increase over the last century, but at most locations at rates comparable to those observed in BSL. However, analyzing tide gauge records in the German Bight, Mudersbach et al. (2013) found differences in linear trends in high sea level percentiles from those in mean sea level. Are

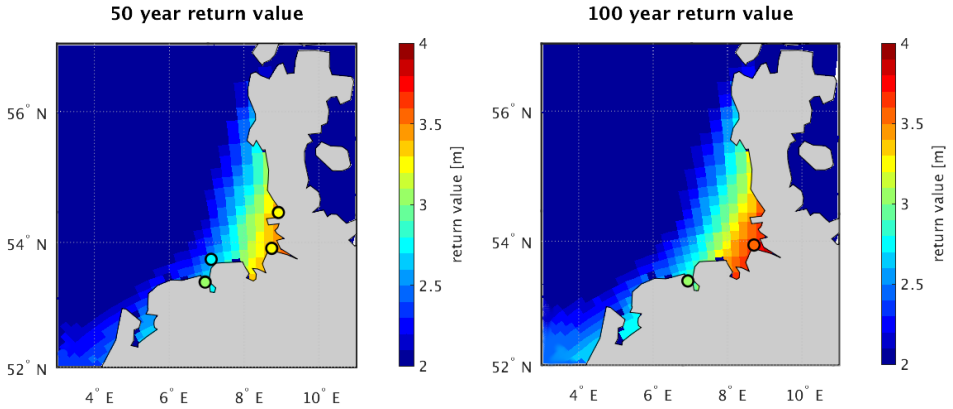


Figure 2.6 | Gridded 50- and 100-year return levels and the corresponding water values from tide gauge observations at selected locations (circles). All values given in m above mean high water.

these differences representative for this period only or can the finding be extended to a general statement?

Fig. 2.3 shows both ESL in terms of annual maxima (black) and BSL in terms of winter median sea level (green), and their respective 11-year running mean. We choose the median instead of the more simple mean in order to not obtain a skewed value due to an exaggerated influence of the very maxima. As the predominant storm surge season we only average over an extended winter period (October - March). Neither ESL nor BSL exhibit long-term trends, but show high interannual to multidecadal variability. Yet, their modulations are not always coherent: As the histogram at the bottom of Fig. 2.3 shows, years with one *extreme storm surge* do not necessarily coincide with a greater occurrence of more moderate *heavy storm floods* or elevated BSL. The correlation between BSL and ESL is comparably low ($r = 0.36$) and highly variable over time (see black curve in Fig. 2.7 for a 100 year running correlation), while the different magnitudes of variances lead to a low explained variance ($R^2 = 0.12$). While there are periods of significant positive correlation between BSL and ESL after 1400, lower insignificant correlations are prevailing during the 1st half of the last millennium. That is, the coherent behavior between mean and upper-end extreme sea level states varies on decadal to centennial timescales. After subtraction of the annual median from the ESL, the correlation between the resulting atmospheric surge residual and BSL further reduces and is insignificant over a large fraction of the last 1000 years (blue curve in Fig. 2.7, $r = 0.10$, $R^2 = 0.01$). In fact, significant coherence only applies in the 15th century and towards the end of the past millennium on timescales of several decades. This further indicates that the similar trends between ESL and BSL during the last century that have often been described (Kauker & Langenberg, 2000; Menéndez & Woodworth, 2010) might merely be an unusual state if compared to a longer time horizon as obtained from our long-term simulation.

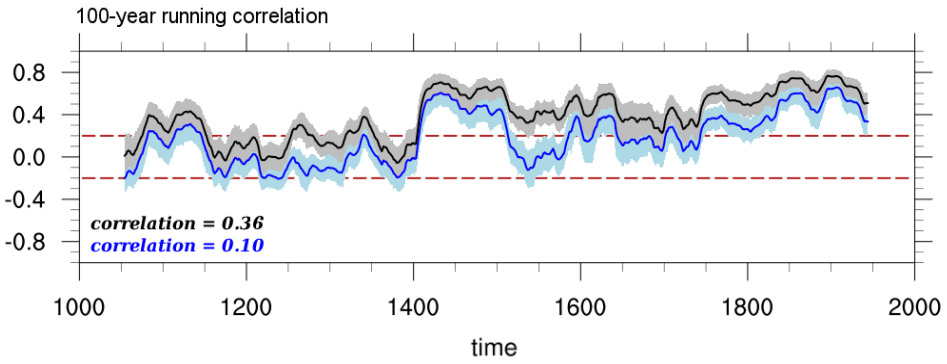


Figure 2.7 | 100-year running correlation between BSL and ESL (black) and between BSL and median-reduced ESL (blue) with shading marking the uncertainty of the correlation using bootstrapping. Time series have been smoothed with a 11y moving window prior to calculating the running correlations. Red dashed lines mark significant correlation on the 95th percent confidence level. The long-term correlation is given in the bottom left corner.

Are the temporal modulations of coherent behavior between ESL and BSL due to different modes of variability and are there any systematic variations in ESL? Spectral analysis (Fig. 2.8a) shows that ESL exhibit white power spectra across all resolved periods ($p = 0.57$ in Ljung-Box Q test) and do – except for a minor spectral peak around 8 years – not show preferred modes of variability (Fig. 2.8a). There is no significant difference between sites located along the coast of Lower Saxony and on sites at the coast of Schleswig-Holstein (not shown). On the other hand, BSL in terms of annual median sea level (Fig. 2.8b, separately shown for both winter and summer seasons) exhibits a red spectrum with more power on multidecadal and centennial timescales. The predominance of lower frequencies in the power spectrum stresses the influence of the ocean which carries the "memory" of the system. In summer and further off the coast (not shown), the high-frequency variability is reduced and thus lower frequencies are dominating the power spectrum. In winter, the higher frequencies are not as damped and the spectrum appears flatter, as the wind stress effect on the water surface is more pronounced during winter when winds are strongest. Coherent variability of BSL and ESL is only visible on multidecadal timescales, at higher frequencies the random ESL variations outweigh the BSL ones and do thus not result in coherent variations for periods shorter than several decades (Supplementary Fig. A7).

2.3.3 *Relation to climate variability*

Several mechanisms have been related to sea level variations, which mainly focused on those of the mean state. These range from large-scale atmospheric circulation patterns (e.g., Chafik et al., 2017; Wakelin et al., 2003; Woodworth & Blackman, 2004) over longshore winds and resulting Kelvin

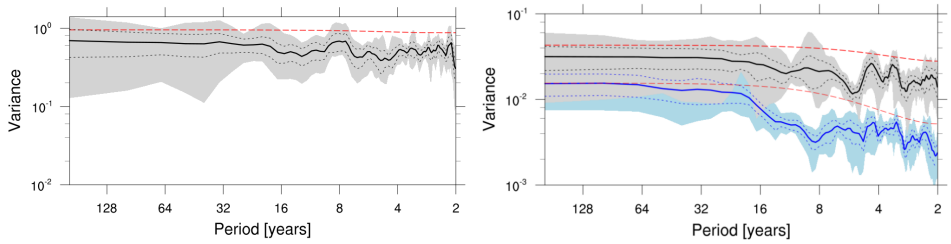


Figure 2.8 | Power spectrum of sea level at Cuxhaven: Annual maximum sea level (left) and annual median sea level (right), split into winter (black) and summer season (blue). Spectra have been smoothed over 7 spectral estimates using a Daniell window (Daniell, 1946). The thick lines correspond to an average over 9 overlapping 200-year subsets of the full time series, the shadings mark their range. Red lines indicate the 95 % confidence bounds using a theoretical Markov spectrum (red noise), black dashed lines the 95% confidence bounds derived from the 9 realizations.

waves (Calafat et al., 2012; Sturges & Douglas, 2011) to steric variations due to temperature oscillations (Frankcombe & Dijkstra, 2009). Mechanisms leading to longer-term ESL variations, however, remain more uncertain as data is limited, which challenges the robustness of statistical relationships between sea level extremes and other variables in the climate system. Yet, the patterns of large-scale climate variability over the North Atlantic that potentially influence ESL may be different to those responsible for BSL variations. In order to investigate large-scale climate patterns associated with enhanced storm surge activity, we relate such periods with multi-decadal climate variability, both internally generated as well as externally forced.

2.3.3.1 *Internal variability*

The NAO has often been linked to BSL variations in the North Sea (e.g., Dangendorf et al., 2012; Wakelin et al., 2003; Woodworth & Blackman, 2004), both through baroclinic as well as barotropic processes (Chen et al., 2014). Correlating observed ESL from tide gauges with climate reanalysis data, some authors found the same large-scale patterns responsible for high ESL as the pattern persists after removing the annual median (e.g., Marcos & Woodworth, 2017; Woodworth et al., 2007). Yet, the standard NAO is not necessarily the most indicative index: Kolker & Hameed (2007) have shown that the location of the centers of action comprising the NAO is affecting observed mean sea level trends and variability. Introducing a "tailored NAO index", Dangendorf et al. (2014a) showed that slightly different pressure constellations than the standard NAO can better describe the observed ESL variability in the German Bight. Additionally, other teleconnection patterns such as the East Atlantic Pattern (EAP) or the Scandinavian pattern (SCA) have been shown to exert some influence on North Sea storminess (Seierstad et al., 2007) and mean sea level (Chafik et al., 2017). However, as some of the modes of climate variability are operating on similar timescales as the

high-resolution instrumental sea level record, it is difficult to obtain robust conclusions.

As an indicator of the large-scale circulation, we compute positive composite maps of winter (Oct-Mar) mean sea level pressure (SLP) during times of high sea level at Cuxhaven, both in terms of ESL and, for comparison, BSL. Composites have been calculated as the average over periods where the simulated sea level time series at Cuxhaven exceeds its mean μ plus 1.5 times its standard deviation σ. The choice of the threshold is arbitrary, but the character of the pattern remain robust to minor changes in its value, specifically the range of $\pm 0.25\sigma$ around the chosen threshold, especially for the BSL case. In the case of the ESL composites, the spatial variability of associated SLP patterns is large and single years can differ in shape and magnitude. The broader spatial character of the averaged anomaly pattern, however, is less affected. Again, we restrict the analysis to the extended winter season as storm surges primarily occur during these months. Further, SLP has been low-pass-filtered with a 3 year moving average to better investigate variability on longer timescales, but to still be able to account for the slightly pronounced variations with periods of around 8 years.

The SLP pattern associated with enhanced BSL (Fig. 2.9a) deviates from the long-term winter mean pattern as the meridional pressure gradient over the North Atlantic is strengthened, comprising a negative SLP anomaly East of Iceland and a positive one over the Iberian Peninsula. This dipole with a meridional axis is similar to a positive NAO. The resulting correlation coefficients between (i) associated SLP index and NAO as well as (ii) the BSL timeseries and the NAO, computed as the leading principal component of the North Atlantic SLP, are significant ($r = 0.9$ and 0.5, respectively), marking a qualitative agreement with the literature outlined above.

The SLP constellation favoring high ESL (Fig. 2.9b), however, differs slightly. As for BSL, it comprises a dipole over the northeast Atlantic, yet its centers of action are shifted to northeastern Scandinavia and the Gulf of Biscay, leading to a further eastward stretching Icelandic Low and a clockwise turned dipole favoring a more northwesterly wind component. This pattern is different from the meridional NAO-like dipole as in the case of elevated BSL. Due to the large ESL variability and the considerable noise in the extreme value time-series, the composite pattern is weaker pronounced than the one associated with high BSL and exhibits a higher variance. Yet, despite this large spatial variance, there is an overall tendency towards a more zonal axis of the associated SLP anomaly pattern. Note that the SLP composites are averaged over the winter season while the ESL time series is based on winter maximum values; the SLP pattern do therefore not represent the situation during individual extreme storm floods but rather reflect a general circulation pattern during times of enhanced storm surge activity. The anomaly structure resembles the dipole described by Dangendorf et al. (2014c) for cross correlations of SLP with observations of daily wind surges at Cuxhaven and agrees well with the mean weather situation triggering strong storm surges found by Heyen et al. (1996) in the wider region us-

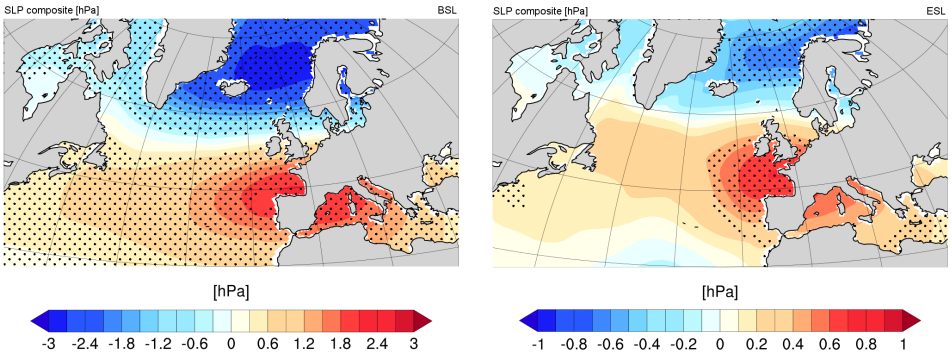


Figure 2.9 | Composite gridded winter SLP anomaly for periods of high BSL (left) and ESL (right) at Cuxhaven. Stippling marks areas significant at the 5 % confidence level.

ing statistical downscaling. Its spatial structure points to an influence of the Scandinavian Pattern in its negative phase (SCA−) onto the NAO centers of action, such as described by Chafik et al. (2017) for North Sea sea level variability. The pattern is also in agreement with a complementary analysis performed on the 2-6 day band-pass-filtered pressure variance (not shown), which indicates enhanced storm track activity over the northeast Atlantic and North Sea.

As the SLP pattern associated with enhanced storm surge activity at Cuxhaven differs from the NAO, we use the shifted centers of action (Biscay – Northeast Scandinavia) to define a SLP index based on the normalized SLP difference between those points of the dipole, to directly investigate the influence of this circulation pattern on storm surge activity in the wider region. Compared to the pattern associated with high BSL (correlation with NAO of $r = 0.9$), this SLP pattern has a lower correlation with the NAO ($r = 0.67$). As a result, the ESL time series at Cuxhaven as well shows a weaker correlation with the NAO ($r = 0.19$) than with the newly defined SLP pattern ($r = 0.31$). Only towards the end of the millennium, BSL and ESL related SLP patterns evolve rather coherently. This feature may also explain why ESL and BSL show a higher coherency (see Fig. 2.7) during the last centuries. Regressing the time series of this SLP index onto the ESL field shows highest values not only in the German Bight (Fig. 2.10a), but also in the southern Baltic Sea where the regressed annual maximum sea levels are of similar magnitude. This suggests that periods of enhanced ESL activity in both German Bight as well as the southern Baltic Sea are linked via the same large-scale circulation pattern. The spatial coherence of long-term ESL variations (see also Supplementary Fig. A6) is in agreement with the study by Marcos et al. (2015) using data from tide gauges globally. The wavelet coherence – which can illustrate both timescales and timing of coherent behavior of both data series – between this index and ESL at Cuxhaven shows that this relationship acts mainly on multidecadal to centennial timescales (Fig. 2.10b).

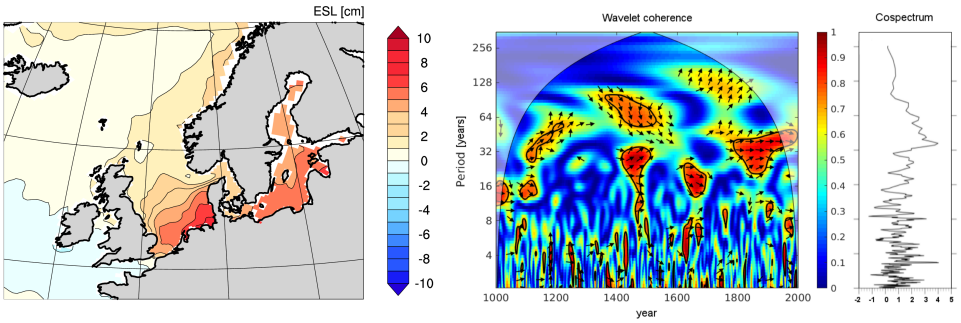


Figure 2.10 | *Left:* Pointwise regression of the SLP index derived from Fig. 2.9 onto annual maximum sea level (color shading, [m per unit of SLP index]). *Right:* Wavelet coherence and cospectrum between annual maximum sea level at Cuxhaven and the tailored SLP index. Arrows to the right (left) indicate a positive (negative) correlation and upward (downward) arrows indicate a lag (lead) of ESL at Cuxhaven. Thick contours designate the 5% significance level against red noise, the cone of influence is shown in a lighter shade.

The second downscaling (1500-2000) of the same GCM simulation yields a similar spatio-temporal variability. This indicates that the large-scale pattern responsible for high ESL activity is also a feature of the parent global simulation, which determines the temporal ESL variability on larger scales – the regionalization however can give added value in the development of dynamic systems such as blockings and is more important for the specific surge heights and finer regional differences that are related to the exact wind direction and strength.

2.3.3.2 *External forcing*

It has been suggested that external influence, such as solar variations or large volcanic eruptions can have an impact on magnitude and phasing of various climate phenomena such as the Atlantic Multidecadal Oscillation (AMO) (Knudsen et al., 2014; Otterå et al., 2010), longer-term anomalous temperature regimes (e.g., Miller et al., 2012) or atmospheric variability patterns such as the NAO (Swingedouw et al., 2011; Zanchettin et al., 2013), and can even trigger periods of enhanced storminess and coastal flooding (Barriopedro et al., 2010; Kaniewski et al., 2016; Martínez-Asensio et al., 2016). For instance, using geological proxy data of the central Mediterranean Sea, Kaniewski et al. (2016) argue for long-term correlations on cycles of around 2200-yr and 230-yr between storminess and solar activity; periods of lower solar activity will intensify the risk of frequent flooding in coastal areas. For the same region, Barriopedro et al. (2010) and Martínez-Asensio et al. (2016) also found coherent decadal changes in solar activity and autumn sea level extremes from tide gauges linked to the 11-year solar cycle through modulation of the atmospheric variability, namely a large-scale wave train pattern, implying an indirect role of solar activity in the decadal modulation of storm flood frequency.

The last millennium has seen substantial variations in solar irradiation, that have affected surface temperatures and lead to various longer-term temperature regimes such as the Late Maunder Minimum (1675-1710) or Dalton Minimum (1790-1840). Additionally, volcanic eruptions can alter the radiation balance substantially for a shorter time and clearly outweigh the variations of solar irradiance alone (see Supplementary Fig. A3). Due to the different timescales, magnitude and expected lag of a potential response to these external variations, it can be useful to separately investigate both forcings and their potential relationship with ESL variations through atmospheric variability.

However, a relation to extreme storm surge activity in the German Bight is not evident in our simulations. Wavelet coherence between total solar irradiance and ESL at Cuxhaven (Supplementary Fig. A8a) and a 'superposed epoch' analysis between volcanic eruptions and ESL (Supplementary Fig. A8b) do not show a consistent significant relationship. Furthermore, the different temporal variations of ESL found in the downscalings of the two different global *Last Millennium* simulations (Supplementary Fig. A4) stresses the dominance of natural variability in the timing of ESL variations. Due to the high internal variability of ESL, any signal from a potential external influence is masked and there is no evidence of coherent variability between German Bight storm surge activity and insolation variations – with or without the inclusion of volcanic forcing – during the last millennium. That is, extreme storm floods have occurred independently of major changes in forcing mechanisms or resulting long-term anomalous temperature regimes. This is in agreement with the findings by Fischer-Bruns et al. (2005) for multi-century simulations of mid-latitude storms. The above described link to atmospheric modes rather stresses the internal component of storm flood variability; this is in accordance with Gómez-Navarro & Zorita (2013) who have shown that the decadal variability of atmospheric modes such as the wintertime NAO is mainly unforced in CMIP5 *Last Millennium* simulations.

2.3.4 *Relation to ESL components*

Which components are governing ESL variability? As described above, the high-end extreme sea levels arise as a combination of three components: the tide, longer-term base level variations and the surge residual, comprising all faster meteorological and oceanographic influences, from both local as well as remote forcing.

Following Woodworth & Blackman (2004), we investigate the surge residual via removal of the two other components, namely tide and background state in terms of the winter median. Removal of tides (using matlab program "t-tide" Pawlowicz et al., 2002) alters the extreme storm surge time series by up to 1 m, depending on the tidal phase during the storm surge. Yet, after removal of tides and median, the general features of variability

and spectra remains similar (not shown), stressing that the variability of extreme storm surges mainly stems from the atmosphere. This may not be surprising, as wind stress is expected to be the most important factor in shallow seas. The unchanged variability also implies that the absolute ESL index using annual maxima is a reasonable indicator of storm flood variability. Furthermore, the large-scale circulation pattern associated with high ESL qualitatively persists if the annual median is subtracted, although weaker. The comparably large spatial variance in the ESL pattern, however, leads to slight shifts in the location of the centers of action of the corresponding SLP dipole; yet, both ESL and ESL-BSL share a tendency towards a more zonal character of the associated SLP patterns. This similarity emphasizes that the SLP pattern associated with high ESL is linked to the surge residual variations. BSL variations are of much smaller amplitude than ESL variations and thus become marginal amid the strong variability of the latter. This is in accordance with findings by Woodworth & Blackman (2004) and Marcos & Woodworth (2017) who concluded that relationships with larger scale climate variability, and specifically the NAO, remain even after removal of the other storm surge components.

2.4 DISCUSSION

The pronounced ESL variability on various time scales found in our simulation has important implications for the interpretation of the instrumental record, including trends as well as estimates of present and future storm floods. With a single realization of limited length such as the instrumental ESL record, statements about potential correlation with BSL, climate variability or alleged trends are statistically problematic and should therefore be treated with caution. For instance, setting recent ESL trends from the observational record (5.7 ± 4.3 cm/decade for the 99.9th percentile of hourly sea level at Cuxhaven from 1953 to 2008, see Mudersbach et al., 2013) in context with the simulated ESL variability shows that the trends lie within the internal variability obtained from the long-term simulation: Using a running trend with a window length similar to the aforementioned observational data (55 years) over the 1000 years of simulated data, a trend of the same or higher magnitude occurs in roughly 10% of the segments. A trend higher than the given upper uncertainty limit (10cm/decade) still occurs 10 times during the 1000 years, or in around 1% of the segments.

Further, the large variations in high return values (around 1 meter for RL_{100}) illustrate the peril of using the standard approach of estimating extreme sea levels for flood protection standards. Typically, such estimates are based on parametric extreme value theory which is based on only a couple of decades of data (typically 30-50 years), be it the observational record or simulated data at end of the century. The effect of fitting different extreme

value distributions onto the same baseline period alone can be significant: A recent study by Wahl et al. (2017) quantified the uncertainties related to different extreme value estimates and showed that the "high-impact-low-probability" sea level states can vary substantially depending on both extreme value distribution and sampling technique. If for example a GEV distribution is fitted to a sample with more weight on more moderate extremes it might lead to a mismatch in higher return levels. For instance, Arns et al. (2013) have shown that for Cuxhaven, the use of r-largest order statistics with $r > 1$ value per year leads to an overestimation of return water levels.

Together with the considerable spread of high-impact return level statistics however, this results in even larger uncertainties: The error bars in Fig. 2.5 illustrate the uncertainties related to extrapolation from shorter subsets. The 1000-year return level (RL_{1000}) estimated from the 100-year long instrumental record yields error bars of 0.8 (3.2) meter for a Gumbel (GEV) fit using maximum-likelihood estimation. The uncertainty is further stressed by the ensemble of ten 100-year long segments, which in itself scatter around 1.2 meter for 100-year return levels. This means in turn that a flood protection standard based on the highest observed sea level from a 100 year data set could as well correspond to a 30-year return level, only if another 100 year period were considered. That is, the observational record is not necessarily representative of the distribution and likelihood of the uppermost ESL. The spread of the simulated RL_{100} from the 10 subsets is around twice as large as the 95% confidence interval of the estimated RL_{1000} using a Gumbel distribution fit onto the observed 100-year data (green bar). This uncertainty range of the Gumbel fit doubles though if the spread in RL_{100} is taken into account (grey bar). The non-parametrically obtained RL_{1000} lies with 3.7 meter within both distribution ranges, but closer to the median of the Gumbel distribution fits. Considering the large variations and corresponding uncertainties it is obvious that high-impact return levels (or flood protection standards for that matter) cannot reliably be inferred from short time series. The sample size considered as a base for parametric extreme value analysis affects not only the likely range of high-impact return events, but may also lead to a problematic negligence of potential scenarios.

How the data record combined with the choice of extreme value distribution impacts the estimation of high-impact return levels is illustrated in Fig. 2.11. For this, we fit the GEV distribution and its special case of a Gumbel distribution to shorter subsets (33 x 30 year and 16 x 60 year segments, respectively) of the full 1000 annual maximum sea levels and compare the RL_{100} estimates to the non-parametrically obtained RL_{100} (Fig. 2.11). A doubling of the data length from 30 to 60 years, for instance, roughly reduces the range of the RL_{100} from each segment's distribution fit to half and the uncertainty range to 1/3 (GEV fit) or 2/3 (Gumbel fit), respectively. The flexible GEV distribution gives more weight to the strongly varying tails of the distribution and results in a larger range of estimated RL_{100}, which is mostly due to the variations in the shape parameter k which strongly depends on

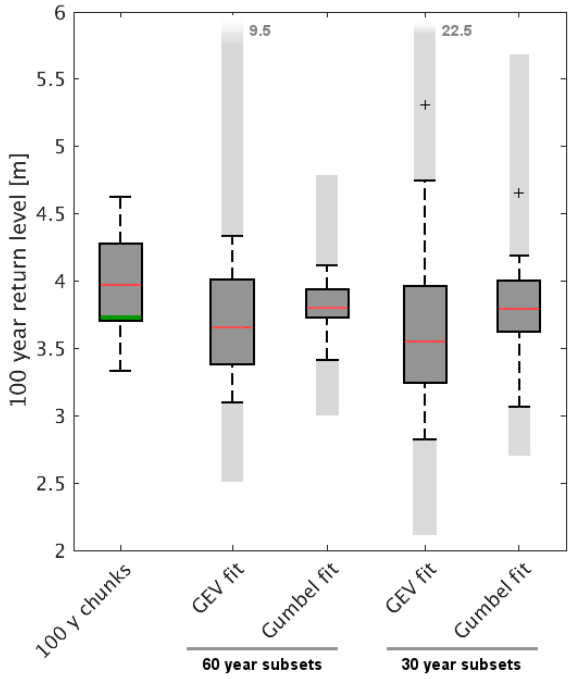


Figure 2.11 | Box-Whisker-plots of RL_{100} non-parametrically obtained from 10×100-year segments of the full simulation against RL_{100} estimates based on GEV and Gumbel distributions fitted to 60-year and 30-year subsets using maximum likelihood. The range of the dashed whiskers represents the total range of the fits, light grey shading represents the maximum range of the 95% confidence limits of the respective fits. The red lines indicate the median of the realizations, the green lines mark the RL_{100} directly inferred from the full 1000-year long simulation.

the respective subset. If k is held constant (which is often assumed in non-stationary extreme value analysis, e.g. in Mudersbach & Jensen, 2010), the uncertainty range reduces considerably. In contrast, the Gumbel fit with a constant zero shape parameter leads to a narrower estimate of RL_{100}. Yet, both distribution fits tend to favor the lower end of simulated 100-year return levels, although the RL_{100} inferred non-parametrically from the full simulation (green bar) lies within the inter-quartile range of both distributions' estimates. This discrepancy can also manifest itself in the temporal variations: The parametric return value estimates can further exhibit different temporal behavior in comparison to the non-parametric plotting positions of the full simulation, and may even lie outside the associated uncertainty range. That is, even such non-stationary extreme value analysis (e.g. Méndez et al., 2007; Mudersbach & Jensen, 2010) does not necessarily reflect the 'real' variations in high return levels. Thus, existing ESL estimates based on limited data lengths do therefore not reflect the full range of ESL variability, e.g. if data stems from a period of unusual state of a driving mechanism such as the SLP dipole outlined above.

Putting these findings into the context of climate change, mean sea level rise will – following a shift of the entire distribution – accordingly translate into a higher probability of occurrence of a particular level and lower return periods. However, while the probability distribution shifts to the right, the large noise in the simulated ESL time series and the low explained variance by BSL imply that in the short term, German Bight ESL might in reality show a much stronger or even a reversed trend.

Yet, additional to a simple shift, a change in the shape of the distribution with increasing atmospheric GHG concentrations can further complicate the picture. Such a change may arise from changes in storm surge climate (e.g. due to an intensification and poleward shift of storm activity; e.g. Fischer-Bruns et al., 2005; Yin, 2005), or from a damping effect on the wind surge in a deeper sea, although this effect has been found to be comparably small (Lowe et al., 2001). This uncertainty in future storm surge projections is expressed by differing findings in existing studies that show a spectrum from no (e.g., Sterl et al., 2009) over little (e.g., Langenberg et al., 1999; Woth, 2005) to considerable change (Lowe & Gregory, 2005). As the variations of ESL are of one order of magnitude higher than the corresponding BSL ones (see Fig.2.3), the variance explained by the latter is low and the detection of changes in the distribution difficult from a statistical point of view. With the strong, but random fluctuations of ESL on timescales of years to decades, we expect existing estimates of ESL changes to be dominated by natural variability rather than climate change signals. Large ensemble simulations will be necessary to detect any significant change in ESL statistics in the presence of the high natural variability found in our simulation. Sterl et al. (2009) have already addressed this issue by using a large ensemble (17 members) of the SRES A1b climate change scenario run with a regional storm surge model and found no signal from climate change on $RL_{10\,000}$ along the Dutch coast. Yet, most studies evaluating future storm risks on data shorter than the estimated return periods, either from observations or from scenario simulations, do not account for such large ensembles and thus systematically disregard ESL variability on timescales longer than the baseline periods. The ESL variation can be substantial though as the large spread of simulated upper-end return levels during the last millennium has shown. For instance, assuming a sea level rise of 0.5 m until the end of the century and given the here quantified ESL variability, more than roughly 200 (350) years of data would be necessary for RL_{100} estimates (95% confidence bounds) using a Gumbel fit to range over less than the estimated signal. Without using large ensembles, ESL projections may be biased by the respective baseline period for extreme value analysis; and even worse, with a small ensemble size or one realization only (e.g. the instrumental record) we cannot say whether they are biased or not. On the other hand, the high internal variability which is essentially irreducible (Fischer et al., 2013) also implies that even perfect models cannot provide well-constrained information on local ESL changes from one realization that might be desirable for adaptation planners.

A couple of caveats that may have an influence on our results are worth discussing. Simulated ESL are – relative to the long-term mean – biased low, which is most likely related to an under-representation of the tidal range in the German Bight (see Supplementary Fig. A1) and to simplifications in the model bathymetry: the ocean model's uppermost layer thickness of 16 meter exceeds the real depth of the shallow waters in many coastal areas of the German Bight and thus may lead to lower wind surges than observed. At Cuxhaven, this effect should be smaller than at points along the flatter Wadden Sea where the tidal oscillations and the shallow waters lead to coastline changes which cannot be represented here. The influence of the bottom topography is expected to play a subordinate role compared to the rather complex horizontal coastline geometry. Yet, in terms relative to mean high waters, simulated ESL statistics agree well with observations, and the temporal variability is not affected by this.

Additionally, the model does not allow for changes in bathymetry, shoreline or coastal management that may influence relative sea levels. This may benefit the homogeneity of the simulated sea level variations, but may hamper the comparability to observations. Processes such as changes in local bathymetry, wave interference at ports or simply inconsistencies in the data record can obstruct the homogeneity of observations from tide gauge records. Additionally, discrepancies between the exact tide gauge position and the nearest-neighbor grid-box can further complicate the picture. A direct comparison between simulated and observed sea levels should therefore be treated with caution. Furthermore, a transient sea level rise due to melting of ice sheets, post-glacial isostatic rebound or the thermosteric effect is not accounted for in the model and a potential increase in ESL with a gradual rise in the BSL base could not be investigated. Such transient sea level changes can further impact ESL on longer time scales, since the sea level distribution shifts with changes in BSL and may potentially also change in shape.

Finally, the results were obtained by downscaling simulations from one GCM only. Potential biases in the parent GCM can thus feed into the downscaled results. For instance, both Northern Hemispheric storm tracks as well as the North Atlantic Current have been found to be too zonal in ECHAM and MPIOM, respectively (Jungclaus et al., 2013; Sidorenko et al., 2015).

As the downscaling of another ensemble member of the parent GCM simulations has shown, the temporal ESL variations of the two different downscalings differ significantly (see Supplementary Fig. A4), albeit their long-term statistics are comparable. That is, any external influence on long-term ESL variability is negligible and it is the natural variability of the parent GCM which determines the temporal variability on a larger scale. The regionalization, however, can offer more detailed dynamical patterns such as blockings. With more precise wind speeds and directions as well as the consideration of regional shelf dynamics it is thus more important for individual surge heights and finer regional differences. The variability due to the

downscaling has been addressed by performing an additional downscaling of the same global simulation which has yielded similar spatio-temporal variability (see Supplementary Fig. A5), indicating that the temporal ESL variability directly linked to the downscaling is negligible. This is in accordance with Woth et al. (2006) who, comparing a number of regional climate models, concluded that the added uncertainty from the downscaling step of ESL variations from global to regional models was comparably small.

2.5 SUMMARY AND CONCLUSIONS

Our study has provided the first coupled downscaling simulation focusing on storm surges and sea level evolution, which gives an unprecedentedly long high-resolution data record that can extend the knowledge of long-term ESL variability based on observations from tide gauge data which are limited in time and space. This simulation renders non-parametric extreme value analysis possible and has the advantage of not relying on extreme value distributions that are typically applied to short data series to provide information about return periods longer than the original time series, as well as their associated uncertainties. The special setup of coupling a high-resolution regional atmospheric model to a global ocean model including tides combines their respective advantages of (i) a consistent simulation of signals both inside and outside the region of interest, and (ii) a sufficiently high resolution in the region of interest to properly account for regional ocean-atmosphere dynamics and other shelf processes. At the same time, the continuous global simulation allows for setting the ESL variability into the context of simulated climate states.

The variability of extreme storm floods has been investigated through the means of annual maximum sea levels at Cuxhaven; the results obtained from different extreme value indices and other gridpoints along the German Bight coastline however do not differ significantly, suggesting that the qualitative behavior and variability are robust and do not depend on the extreme value sampling method or exact location.

Our results suggest that

1. The model reproduces observed storm surge statistics at Cuxhaven, both in terms of seasonality as well as magnitude above mean high waters.

2. ESL variations are large, but operate on a white spectrum and do not exhibit significant oscillatory modes beyond the seasonal cycle.

3. High-impact extreme events vary substantially on timescales longer than the typically available base period for return period estimates. Estimates of ESL obtained via the standard parametric approach based

on short data records are therefore not representative for the full ESL variations.

4. Long-term ESL variations in the German Bight are regionally consistent, indicating a common large-scale forcing. Large-scale circulation regimes that favor periods of enhanced ESL in the German Bight are similar to those associated with elevated BSL, but the location of the respective centers of action of the governing SLP dipole differs. While BSL variations correlate well with the wintertime NAO, ESL variations are rather associated with a shifted NAO+/SCA- like pressure pattern leading to a stronger local northwesterly wind component.

5. Any potential links to BSL fluctuations as well as external influence through solar variability or volcanic activity are masked by the strong internal variability of ESL. This is in accordance with the findings by Fischer-Bruns et al. (2005) who, using coupled *Last Millennium* simulations of the last five centuries, concluded that the natural variability of mid-latitude storms is not related to solar, volcanic or GHG forcing nor to anomalous climate states such as the Maunder Minimum. Similar conclusions have been made for North Atlantic summer storm tracks over Europe by combining observations, simulations and reconstructions of the last millennium (Gagen et al., 2016).

We thus conclude that the magnitude of ESL and existing estimates of changes thereof are dominated by natural variability rather than forced signals. Given the large variability from our simulation, large ensemble simulations are required to detect a potential future change in ESL statistics with respect to climate change induced BSL.

Nevertheless, the obtained information on the statistics of ESL variability together with the here established links to large-scale climate variability may be used to better explore future pathways of extreme sea levels. Owing to the large ESL variability, a responsible adaptation strategy should therefore reflect the range of possible developments rather than solely being designed to a forced signal. At the end, uncertainties in both SLR projections as well as ESL estimates need to be better understood and combined to fully assess potential impacts and required adaptation measures.

Code and data availability. Primary data and scripts used in the analysis that may be useful in reproducing the work are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/os-15-651-2019-supplement.

Author contributions. UM came up with the idea for the paper and developed the model code. AL and UM jointly designed the experiments; simulation and analysis were carried out by AL. The paper was written by AL with contributions from UM.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. The research was supported by the Max Planck Society and by the German Research Foundation (DFG) funded project SEASTORM under the umbrella of the Priority Programme SPP-1889 "Regional Sea Level Change and Society" (grant no. MI 603/5-1). The authors acknowledge Eduardo Moreno-Chamorro and Johann Jungclaus for producing the past1000 global simulations and Sönke Dangendorf and Wilfried Wiechmann for providing the observational data from selected tide gauges. We thank Johann Jungclaus and Moritz Mathis, as well as editor Markus Meier and three anonymous reviewers, for their constructive comments and remarks which helped to improve this paper. We further acknowledge the German Climate Computing Center (DKRZ) for providing the necessary computational resources.

Financial support. This research has been supported by the German Research Foundation (DFG) (grant no. MI 603/5-1). The article processing charges for this open-access publication were covered by the Max Planck Society.

Review statement. This paper was edited by Markus Meier and reviewed by three anonymous referees.

3

EXTREME SEA LEVELS UNDER CLIMATE CHANGE

The content of this chapter has been submitted to *Climate Dynamics*:

Lang, A. and Mikolajewicz, U. (under review): Rising extreme sea levels in the German Bight under enhanced CO2 levels – a regionalized large ensemble approach for the North Sea. *Clim. Dynamics*.

Abstract. We quantify the change in extreme high sea level (ESL) statistics in the German Bight under rising CO_2 concentrations by downscaling a large ensemble of global climate model simulations using the regionally coupled climate system model REMO-MPIOM. While the model setup combines a regionally high resolution with the benefits of a global ocean model, the large ensemble size of 32 members allows the estimation of high return levels with much lower uncertainty. We find that ESLs increase with atmospheric CO_2 levels, even without considering a rise in the background sea level (BSL). Local increases of up to 0.5 m are found along the western shorelines of Germany and Denmark for ESLs of 20-50 year return periods, while higher return levels remain subject to sampling uncertainty. This ESL response is related to a cascade of an enhanced large-scale activity along the North Atlantic storm belt to a subsequent local increase in predominantly westerly wind speed extremes, while storms of the major West-Northwest track type gain importance. The response is seasonally opposite: summer ESLs and the strength of its drivers decrease in magnitude, contrasting the response of the higher winter ESLs, which governs the annual response.
These results have important implications for coastal protection. ESLs do not only scale with the expected BSL rise, but become even more frequent, as preindustrial 50-year return levels could be expected to occur almost every year by the end of the century. The magnitude of the relative change in ESL statistics is hereby up to half of the expected rise in BSL, depending on the location. Changes in the highest extremes are subject to large multidecadal variations and remain uncertain, thus potentially demanding even further safety measures.

3.1 INTRODUCTION

Extreme high sea levels caused by storm floods constitute a major geophysical hazard for low-lying coastal regions such as the southeastern North Sea, known as the German Bight. Strong tidal oscillations due to the shallow shelf sea and the basin's geometry (Fig. 3.1), combined with its situation along the major northern hemispheric storm-track path can lead to particularly high storm floods. While these conditions have lead to several severe disasters such as the flood of 1962, they have also triggered various counteractions and regulations in coastal protection (e.g. MELUR, 2012).

Anthropogenic climate change, however, may further alter the risks for coastal areas (Hinkel et al., 2015; IPCC, 2019). While the expected rise in the background sea level (BSL) is – at least until the end of the century – relatively well-constrained (e.g. IPCC, 2019; Slangen et al., 2014) and directly affects extreme sea level heights by shifting the entire distribution, there is less certainty about the relative changes in the sea level distribution, e.g. through changes in the storm climate (Feser et al., 2015). In particular, there is only little confidence on the change of upper-end extreme values (Wahl et al., 2017; Weisse et al., 2012), which by definition occur only rarely. Yet, it is those 'high-impact-low-probability' extremes that are typically required for flood defense standards. While such height changes are often handled by simply scaling the sea level distribution with the regional or even global mean sea level rise (e.g. Buchanan et al., 2017; IPCC, 2019), relative changes in the upper tail of the sea level distribution may substantially alter the risk for flooding. We here investigate the regional change in those extreme sea level statistics in the German Bight following an increase in atmospheric CO_2 by downscaling a large ensemble of global climate model simulations. A rise in the background state due to sea level rise is not taken into account.

Several studies have investigated how storm induced water levels in the North Sea will change under global warming (e.g. Flather & Smith, 1998; Gaslikova et al., 2013; Howard et al., 2010; Langenberg et al., 1999; Lowe & Gregory, 2005; Lowe et al., 2001; Sterl et al., 2009; Weisse et al., 2014; Woth et al., 2006; von Storch & Reichardt, 1997). They traditionally employ a regional barotropic storm surge model forced by atmospheric conditions from climate models, to compare time-slices of a present-day climate with one at the end of the century based on certain future climate change scenarios. While most studies point towards no or only small increases in the frequency or magnitude of storm floods, the rather short data samples, naturally limited by the length of the time slices (typically 30 years), limit the analysis to relatively frequent extremes and impede statistically robust conclusions on upper-end extreme events that happen only once per couple of years. Further, North Sea extreme sea levels have been shown to exhibit a large interannual to multidecadal variability (Dangendorf et al., 2013a), especially when dealing with more extreme and thus rare events (Lang &

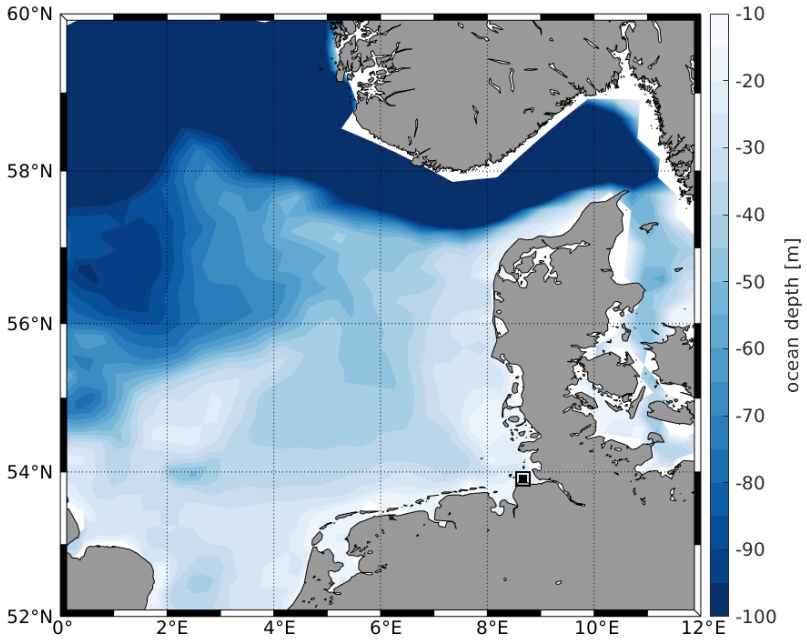


Figure 3.1 | Bathymetry of the German Bight with main study location Cuxhaven.

Mikolajewicz, 2019), additionally stressing the need for large samples of extremes to separate a climate change signal from internal variability. Only a few studies have addressed this signal-to-noise issue for more high-impact events: Using a 17-member initial-conditions ensembles and the SRES A1B scenario, Sterl et al., 2009 found no changes in upper-end North Sea storm flood return levels above natural variability, although the specific focus was on Dutch flood defense standards and selected locations.

Yet, due to the lateral boundaries of the underlying storm surge models, none of the above studies employ a globally consistent climate scale simulation with full atmosphere-ocean coupling. Hence, they do not account for climate-related sea level variations external to the North Sea, which have been shown to impact regional North Sea sea level on different scales (e.g. Calafat et al., 2012; Chen et al., 2014), nor for external surges (e.g. Gönnert & Sossidi, 2011) that travel from the North Atlantic into the basin and can build an important component of the total storm surge induced water level, especially under climate change (Woth et al., 2006). Further, studies with only a regional domain can only indirectly relate ESL variations to large-scale drivers.

This study attempts to close these gaps as we (i) adequately sample the internal variability of ESLs by using a large ensemble of general circulation model (GCM) experiments, and (ii) downscale this ensemble with a coupled climate model with a global ocean, but a regional zoom on the North Sea (see Lang & Mikolajewicz, 2019; Mathis et al., 2018). This allows us to robustly investigate potential changes in high-impact-low-probability

events following an increase in atmospheric CO_2. By employing a total of 32 members, this marks to our knowledge the largest ensemble study using a coupled climate system model to investigate changes in the statistics of regional sea level extremes. The comparably large number of ensemble members leads to much smaller confidence intervals and allows the detection of changes in sea levels of even larger return periods.

This article is structured as follows: First, the model and downscaling setup as well as forcing and scenario are introduced (3.2); Second, results are presented for (i) changes between high and low CO_2 worlds (3.3.1), and for (ii) the transient response of extremes (3.3.2), followed by a seasonal analysis (3.3.3) and an investigation of the associated drivers in the climate system (3.3.4). We close with a discussion of the relevance of these results in the light of current emission pathways (3.4).

3.2 METHODS

3.2.1 *Model system and experimental design*

We employ the regionally coupled climate system model REMO-MPIOM (Lang & Mikolajewicz, 2019; Mathis et al., 2018; Mikolajewicz et al., 2005; Sein et al., 2015), which consists of the global ocean model MPIOM (Jungclaus et al., 2013; Marsland et al., 2004) and the regional atmospheric model REMO (Jacob & Podzun, 1997), centered over Europe (Fig. 3.2). Inside the REMO domain, MPIOM and REMO are interactively coupled with the coupler OASIS-3 (Valcke, 2013); outside, the ocean model is forced by output from the parent atmospheric model. This model setup is identical to the one used in Lang & Mikolajewicz, 2019, with the exception that he bottom friction has been modified to produce more realistic tidal amplitudes in the southeastern North Sea.

The regional atmospheric model covers the wider European region, parts of northern Africa and the northeast Atlantic, and is run with a 0.44° setup, corresponding to approx. 50 km grid spacing, and with 27 vertical levels. The ocean model is run on a stretched grid configuration with a nominal horizontal grid resolution of 1.5° and 30 vertical layers, and includes the full luni-solar ephimeridic tidal potential according to Thomas et al., 2001. In order to maximize the grid resolution in the North Sea, the model's poles are shifted to Central Europe and North America, resulting in a horizontal grid resolution of up to 9 km in the German Bight, thus enabling a more realistic simulation of small-scale shelf processes. At the same time, the ocean model's global domain allows for a consistent simulation of signals across the lateral boundaries defined by REMO. This could be important for coastal sea levels through teleconnections, e.g. modulations in the Gulf stream or the Atlantic Multidecadal Oscillation (Ezer et al., 2016; Turki et al., 2019). To avoid the problem of dry-falling ocean grid-points due to strong tidal fluctuations, MPIOM's uppermost layer thickness is set to 16 meter.

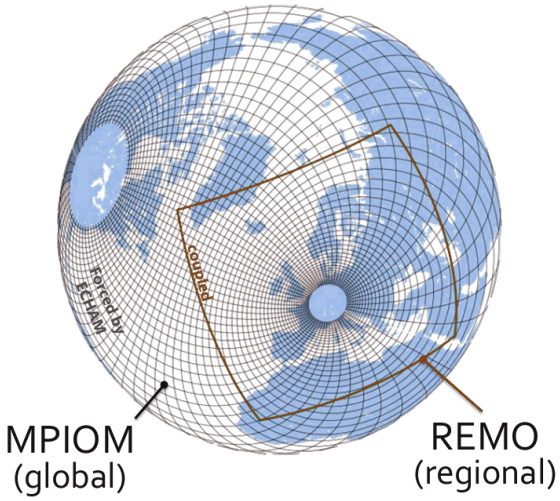


Figure 3.2 | Sketch of the regionally coupled climate system model used for downscaling. Not all grid-lines are shown.

For a discussion of the suitability in analysing storm floods including a validation using the Cuxhaven tide gauge record, see Lang & Mikolajewicz, 2019.

As forcing, we use output from the *1pctCO2* simulations of the Max Planck Institute Grand Ensemble (MPI-GE, Maher et al., 2019), model version MPI-ESM1.1. The model is run in its low resolution configuration (resolution T63) with 47 vertical levels in the atmosphere model ECHAM (Giorgetta et al., 2012) and 1.5° resolution and 40 vertical layers in the ocean model MPIOM (Jungclaus et al., 2013). The *1pctCO2* scenario covers 150 years, with the atmospheric CO_2 concentration increasing by 1% every year, from preindustrual levels to approximately a quadrupling at the end of the simulation period. The transient character of the simulation provides additional insights into the evolution of long-term variability. To tackle the large internal variability in sea level extremes (Lang & Mikolajewicz, 2019), we downscale a set of 32 ensemble members (the maximum number of ensemble members with 6-hourly output that is required for downscaling), resulting in $32 \times 150 = 4800$ years of simulation. Initial conditions of the individual runs stem from different years of the corresponding downscaled PI-Control run. As we employ a boussinesq model and mainly target changes in ESL statistics, a transient rise in BSL is not taken into account. Since atmosphere-related sea level variations have been shown to act to a first approximation independently of the water depth (e.g. Howard et al., 2010; Lowe & Gregory, 2005; Sterl et al., 2009), this should however not affect the relative variations in sea level extremes.

3.2.2 *Extreme value sampling*

Sea level can be described as the sum of a longer-term mean sea level, the tide and the atmospheric surge, as well as their non-linear interactions (e.g. Pugh, 1987). Instead of treating the terms separately, we here investigate extreme sea levels in terms of total sea level above a long-term reference state. Since we are solely interested in high-water extremes, we use the term extreme sea level (ESL) to refer to the upper end of the distribution only.

We sample ESLs in terms of annual (or seasonal, respectively, for section 3.3.3) maxima. In order to not split up the winter storm flood season, we define a year as starting in July. To remove trends in the background sea level, but to still account for low-frequency variability, the centered 30-year running mean of the annual median sea level has been subtracted from each annual maximum value.

To characterize ESLs based on their probabilities, we rely on the commonly used concept of return levels (e.g. Coles et al., 2001), representing the water level expected to be exceeded on average once every x years. Such probability-based exceedance levels of assigned return periods are often required for designing coastal defense structures. As simulated data from the instrumental record or from the above mentioned future time slice experiments rarely cover more than a couple of decades, ESLs are typically estimated based on parametric extreme value analysis, resulting in a considerable extrapolation of the data. The advantage of a large ensemble approach is that it allows to infer upper-end extreme value statistics without the use of parametric methods. These non-parametric estimates have been inferred by first ranking the data points of each time slice of the full sea level ensemble and associating a cumulative probability to each value. The respective return periods are inferred from the sampled ESLs using the reciprocal of the probability of exceedance P_E, defined as $P_E = \frac{m}{N+1}$, where m is the rank of N events ordered in decreasing order. To increase the robustness of lower return periods, the underlying sample of extreme values has been increased to the highest 10 per year for this type of analysis.

3.2.3 *Change in ESL statistics*

For the analysis of climate-related ESL changes, we analyse pooled ESL statistics of all 32 ensemble members. We analyse the climate-related change in ESL statistics in two ways: First, we follow the common time-slice approach and compare two 30-year periods at the beginning and end of the 150-year period (climate change response, section 3.3.1). The two time slices comprise the years 5-34 ("low CO_2", 305-403 ppm) and 120-149 ("high CO_2", 957-1278 ppm), corresponding to a global mean surface temperature (GMST) change of a bit more than four degrees. To avoid potential artefacts after the branching-off from the control run, we exclude the first five years of each

run for the analysis. We define 'climate change response' as the difference of high and low CO_2 states. Secondly, we investigate the 150-year transient response (section 3.3.2) by analysing pooled ensemble extreme value statistics of running 30-year periods. This analysis is followed by a seasonal analysis (section 3.3.3) and an investigation of associated drivers in the climate system (section 3.3.4).

3.3 RESULTS

The simulated evolution of ensemble statistics of ESL together with a display of the two time slices used for the analysis of the climate change response is shown in Fig. 3.3 exemplary for Cuxhaven. With values between 2.5 and 6.5 m above the 30-year mean, the individual members exhibit a high spread. For comparison, we include the range of observation-based ESL using the same definition from the last 100 years of tide gauge record (green bar, data from the AMSeL project, see Jensen et al., 2011). For a more detailed validation of simulated extreme value statistics, see Lang & Mikolajewicz (2019).

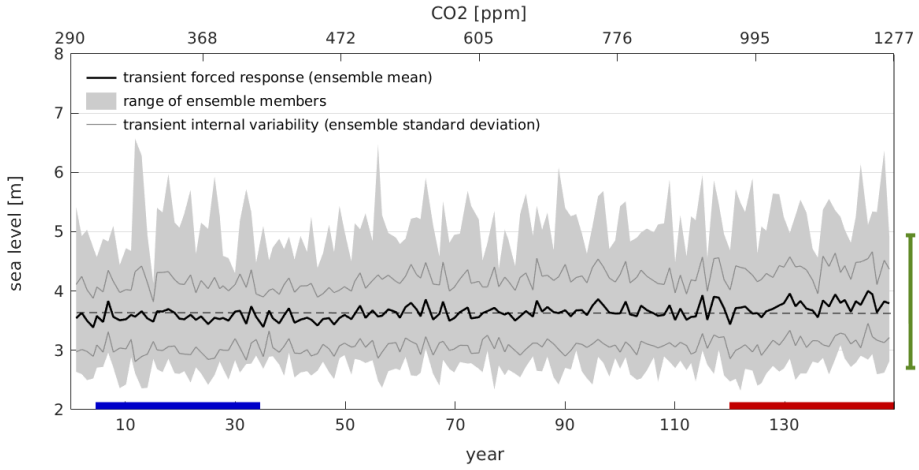


Figure 3.3 | Ensemble statistics of extreme sea level at Cuxhaven, sampled in terms of annual maxima relative to the 30-year running median (see methods). As a reference, the overall mean ESL (i.e. the 1-year return level) is indicated by the dashed line. The range of ESL from the tide gauge record is shown by the green bar on the side. Time slices used for the climate change signals in low and high CO_2 worlds are marked on the x-axis.

3.3.1 *Climate change response*

The traditional approach to assess changes in sea level extremes under climate change is to compare the statistics of typically 30-year periods at the beginning and end of a certain scenario run (usually RCP8.5 or SRES A1B).

However, with a single or only a few ensemble members one cannot robustly identify changes in the tails of the extreme value distribution. With the 32 members a 30 year length, a total of $32 \times 30 = 960$ years is available for each time slice. Hence, a comparison of 30-year segments of low and high CO_2 states can yield much higher levels of significance.

Figure 3.4 shows ESL statistics at Cuxhaven in the two periods in terms of their return levels. The individual ensemble members of 30 year length result in a large spread of more than 2 meters for 30-year return periods. With 32 ensemble members and effectively 960 years for each period, however, we can reduce the associated uncertainties, so that much higher return periods can be calculated without the need of fitting an extreme value distribution to the data. While individual ensemble members (thin lines) show a large spread and non-uniform climate change signals, the full ensemble (thick lines) shows that return levels generally increase with rising CO_2. The change in ESLs is largest for return periods of around 20-30 years, where

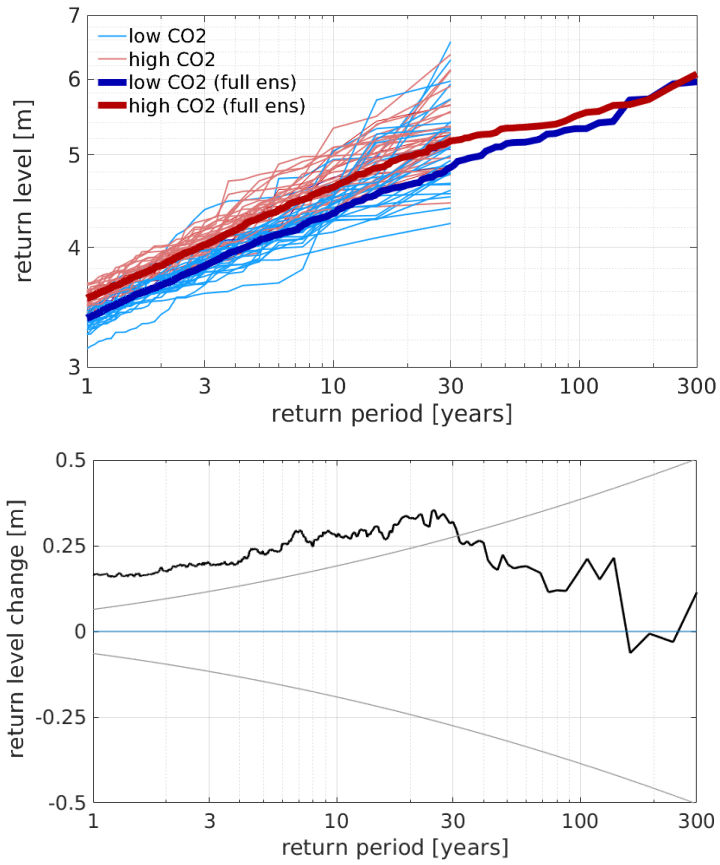


Figure 3.4 | Non-parametric extreme value analysis: Return value plots of high and low CO_2 states for individual members and the full pooled ensemble (top), and full ensemble difference between high and low CO_2 states (bottom) at Cuxhaven. Grey lines in the lower panel show the associated 95% confidence limits using a GEV fit to the full ensemble via least squares.

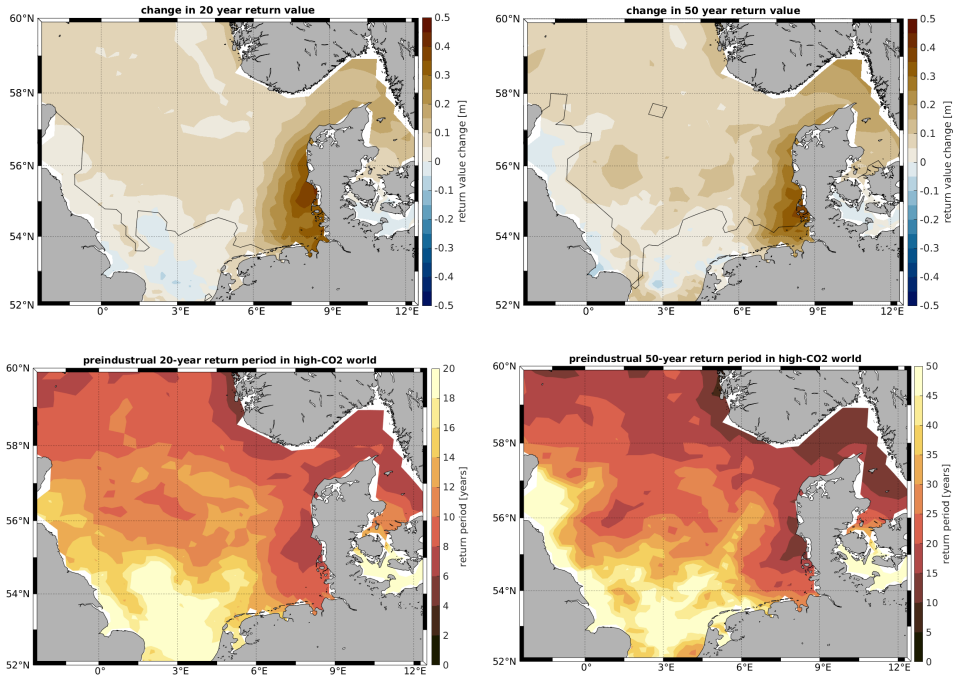


Figure 3.5 | Spatial ESL change in 20-year (left) and 50-year (right) return levels (top, significance on the 95% confidence level in black contours) and in corresponding return periods (bottom).

the higher CO_2 concentrations lead to a shift of around 30 cm relative to the mean state. The change is significant for extreme values up to return periods of around 33 years, based on both a two-sided student t-test ($\alpha = 0.05$) and the 95% confidence intervals from a parametric extreme value analysis using a Generalized Extreme Value (GEV) distribution. For higher return periods, the confidence range of the extreme value estimates surpasses the climate change signal.

Is this climate change response sight-specific or is there a more general regional pattern visible? The spatial climate change response pattern of ESLs is shown in Fig. 3.5 for return periods of 20 and 50 years. While there is a slight increase in both ESL estimates in the entire German Bight, changes are most pronounced along the western Danish and German coasts, with maximum values of 0.4-0.5 meter at around 55°N close to the German/-Danish border. As a consequence, ESLs that correspond to a certain return period in preindustrial times will thus also become more frequent over the wider area, with preindustrial 50 (20) year return levels occurring up to every 10-15 (6-8) years in a high-CO_2 world. This marks an increase by the factor three to five. Note that this analysis only considers the relative changes in extreme values with the long-term mean sea level subtracted. If a BSL rise due to thermal expansion or melting of ice sheets were included as well, these numbers would change even more (see section 3.4).

3.3.2 *Transient response*

Additionally to comparing the statistics of two time-slices of low and high CO_2 states, the transient character of the simulation allows a more detailed analysis of temporal changes in extreme value statistics, including the detection of possible nonlinearities in the system or other forms of long-term variability, as a function of both time and atmospheric CO_2 concentrations.

The internal ESL variability, represented by the 30-year running mean of the ensemble standard deviation (see Supplementary Fig. B1) increases with rising CO_2 concentrations, but also exhibits strong multidecadal fluctuations by about 10 cm for the pooled ensemble. This long-term ESL variability is also manifest in the more extreme ESL measures, as the temporal evolution of different return level estimates based on pooled ESL statistics show (Fig. 3.6). As already seen in the time-slice approach above, ESLs increase accordingly with atmospheric CO_2 concentrations. Although the 20-year return level shows a statistically significant trend over the 150 years, the transient approach reveals a more detailed behavior: a period of low ESL variability between years 30 and 50 leads to a marked 'dint' in the gradual response, especially in the upper-end return levels, where it is of similar magnitude as the overall 150 year response. Pooled ESLs in this period rarely surpass the 5 meter threshold (compare Fig. 3.3), thus leading to the described opposed signal in ESLs of higher return periods. Here, the signal emerges from variability already at around 800 ppm or around 3 degrees of warming, when the ESL change exceeds the limits of the 95% confidence bounds. This impact of ESL variability on the robustness of the ESL signal for different ESL heights is in line with the climate change signal

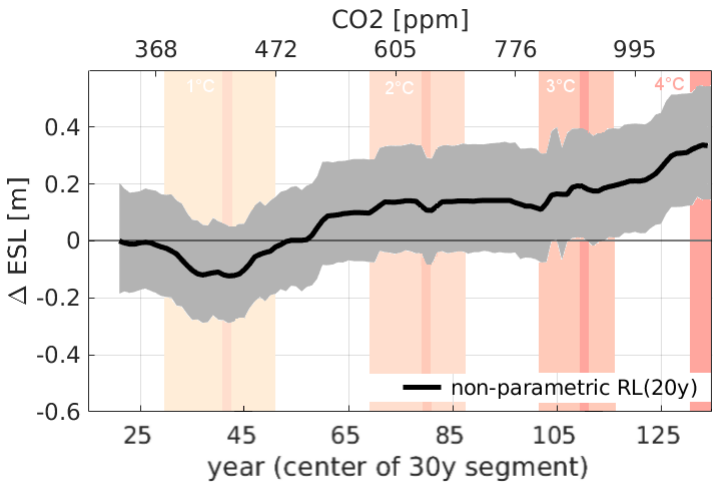


Figure 3.6 | Time-dependent extreme value estimates for 20-year return levels based on moving 30-year segments of the pooled ensemble at Cuxhaven. Grey shading marks the 95% confidence bounds using a GEV fit. Colored shading indicates different warming levels.

in Fig. 3.4. The presence of such marked multidecadal variability of individual members as well as the pooled ensemble suggests that the temporal variations of ESLs, especially of return periods above a couple of decades can be of similar magnitude as the climate change response.

3.3.3 *Seasonality*

Since annual maximum sea levels in the German Bight would typically occur during the main storm flood season in winter, responses in other seasons are thus masked in the above analysis. Yet, the individual seasonal responses might still be of interest. Wild birds, for instance, nesting on open land in front of the dikes in spring and summer are vulnerable to flooding (Camphuysen & Leopold, 1994), even if we might not expect strong storm floods in those seasons.

Fig. A.2 shows histograms of the frequency of the extreme sea levels occurring in a certain month for low and high CO_2 levels, sampled as annual maxima (i.e. 1-year return levels, top) and 30-year return levels (bottom). As expected, the bulk of ESLs occurs in the winter months, although the strongest annual floods can also occur in the summer months in individual years. This strong seasonality is even more evident for more extreme return levels (lower panel). Over time, however, a change in the seasonal distribution from low to high CO_2 state is evident, with the timing of ESLs becoming more confined to the winter months December and January. In fact, the most extreme storm floods in a high CO_2 world occur in our simulations only in the months from November to March.

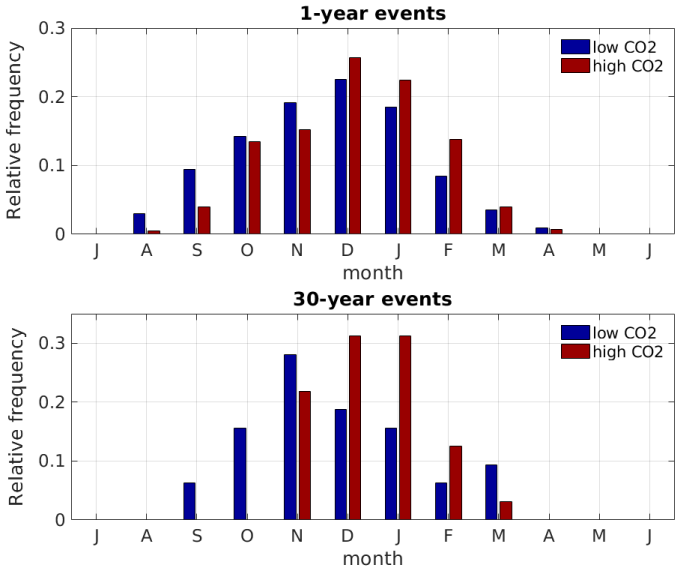


Figure 3.7 | Histograms of seasonality of ESLs at Cuxhaven in terms of 1-year return periods (top) and 30-year return periods (bottom).

This confinement to the winter season has implications for the seasonal climate change signal in return levels. Fig. 3.8 shows an equivalent analysis of return levels to Fig.3.4, but split up into the 4 different seasons. The change in seasonality results in a stronger climate change signal in winter-only return levels than if based on annual maximum sea levels. If treating the four seasons individually, maximum winter ESL changes reach up to 60 cm for return periods of around 40 years. This is most likely due to the fact that a 'washing-out' of the return level changes due to the annual maximum appearing in different seasons is no longer possible with a seasonal analysis. Interestingly, the summer response shows opposing trends, with a reduction of storm flood heights in summer, with maximum change signals for ESLs of 60-80 year return periods. Autumn and Spring responses are only minor. That is, the seasonally opposing signals in winter and summer together with a shift of the annual maximum to the winter-months lead to a dampening in the year-long signal and accordingly to smaller increases in return levels.

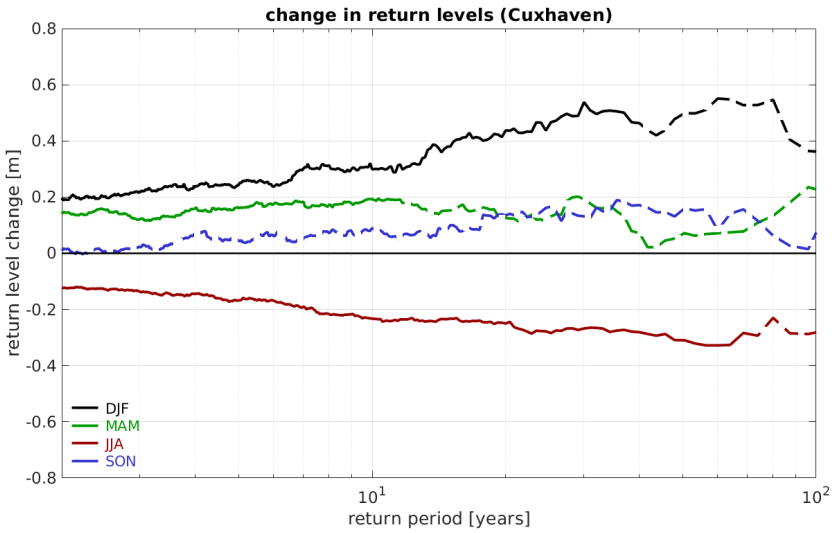


Figure 3.8 | Seasonal climate change signals for ESLs of different return levels at Cuxhaven. Lines are dashed for ESL responses outside the 95% confidence ranges equivalent to Fig. 3.4b.

3.3.4 *Atmospheric drivers*

What drives the change in ESL statistics? With the long-term background state removed, it is the meteorological conditions such as changes in wind speed or direction, the frequency of severe storms, both locally and large-scale, and pressure patterns that govern the atmospheric surge component of sea level.

Several studies have investigated the changes in European wind climate (for a review see Feser et al., 2015). Although an overall tendency for a poleward shift in storm activity is emerging, regional and model-dependent deviations from this large scale picture stress that the confidence in future changes in wind climate in Europe remains "relatively low" (Christensen et al., 2007). As a main driver of surge levels, we here investigate changes in the regional 10m wind climate as well as the large-scale storminess and the prevailing storm track categories.

Regional wind climate

The change in regional wind climate, expressed as the ensemble mean winter maximum (i.e. RV_{1y}) wind speed change between high and low CO_2 worlds based on annual maximum 10m wind speeds is shown in Fig. 3.9 (top left). Predominant westerly winds in the North Sea are stronger in the high-CO_2 world, while the opposite is true for parts of Southern Europe and the Mediterranean. The Northern Atlantic and parts of central Europe do not show significant signals. The annual maximum corresponds to approx. 22 m/s, i.e. around nine Beaufort, and should thus be a good measure of storminess. The 30-year return levels (RV_{30y}) of winter maximum wind speeds show a similar picture with an even stronger signal, although the

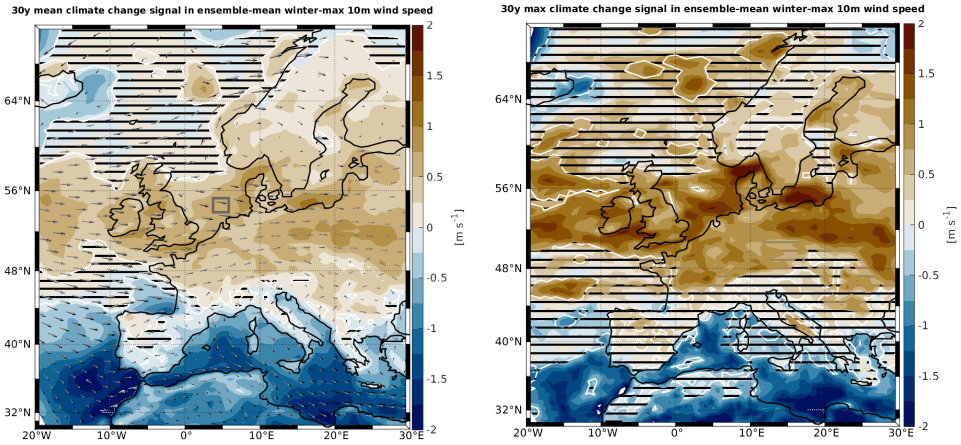


Figure 3.9 | Regional change (high - low CO_2 world) in winter maximum 10-m wind speeds with wind vectors for 1-year return periods (left) and 30-year return periods (right). Areas not significant at the 95% confidence level are striped out.

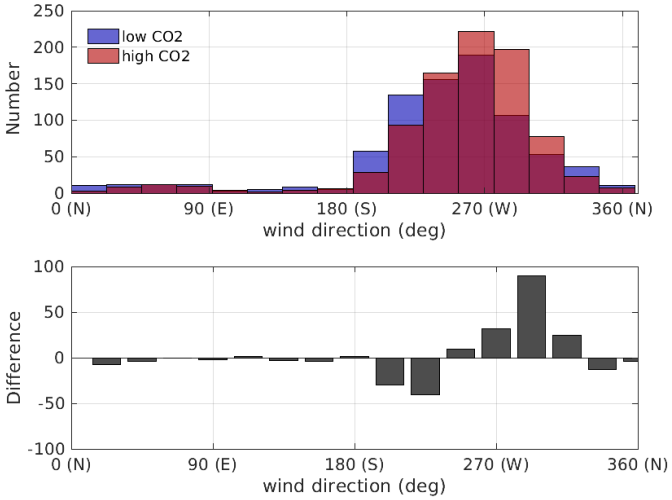


Figure 3.10 | Wind directions of winter maximum 10m winds in a low and high CO_2 world (top) and their difference (bottom) at the location indicated in Fig. 3.9.

statistical significance is more regionally confined. Yet, in the North Sea, statistically significant positive climate change signals in the regional wind climate are evident. With an increase in magnitude, also the predominant winter maximum wind direction changes slightly with rising atmospheric CO_2 concentrations, from WNW towards more more zonal W winds (Fig. 3.10). That is, the westerly winds which are responsible mainly for ESLs along the eastern part of the German Bight, i.e. the western German and Danish coasts, are becoming more frequent. This feature is consistent with the spatial pattern of the ESL changes, which show strongest climate change signals along these coasts. In summer, the simulation shows a regionally consistent and significant reduction in extreme wind speeds, both for RV_{1y}. and RV_{30y} (see Supplementary Fig. B.2). The predominant wind direction during summer does also change slightly, yet other than in winter rather towards more northerly directions (see Supplementary Fig. B.3). The winter shift to stronger winds in the southern North Sea is consistent with results from studies by, e.g. Beniston et al., 2007, Pinto et al., 2007 or Weisse et al., 2014 who, using a set of regional climate models (RCM) driven by ECHAM4 and HadAM3H, and ECHAM5 output respectively, found an increase in storm activity and in the wind climate over North and Baltic Seas. Similar increases of higher percentiles of westerly wind directions in the Southern North Sea have also been reported by Gaslikova et al., 2013 and Ganske et al., 2016 for different RCMs and the A1B scenario, although most changes in the latter were statistically insignificant. Concerning the more extreme wind speeds, however, De Winter et al., 2013 found no significant change in the annual maximum wind speed in the North Sea from output of 12 CMIP5 models.

Large-scale storminess

The large-scale storminess has been investigated using 2-5-day bandpass-filtered (Blackmon, 1976) sea-level-pressure data. The resulting change between high and low CO_2 worlds shows an intensification of the North Atlantic storm belt in winter, esp. between $45-60°$N in the northeastern Atlantic, as well as a shift towards a more zonal storm belt (Fig. 3.11). Again, the summer response is of opposite sign, with a broad reduction in storminess, although the change is smaller in absolute numbers (see Supplementary Fig. B.4). This seasonally opposing pattern is consistent with the seasonality of the ESL change and the increase in storm flood magnitude in (late) winter.

Generally, there is a consistent cascade from changes in the large-scale storm climate towards changes in regional winds, which eventually produces a response in the statistics of ESLs. The response patterns of winter and summer seasons oppose each other; however, since North Sea wind speeds and ESLs are highest in the winter months, the yearly response of maximum wind speeds broadly follows the winter response.

Storm tracks

High ESLs in the German Bight are a result of storm floods induced by cyclones that travel from the North Atlantic into the North Sea and thereby generate large onshore winds that pile up water against the coast. Several authors have classified such storm-flood inducing strong cyclones according to their tracks. Analysing strong cyclones that led to severe storm floods registered at the Cuxhaven gauge since 1900, Kruhl (1978) categorized them

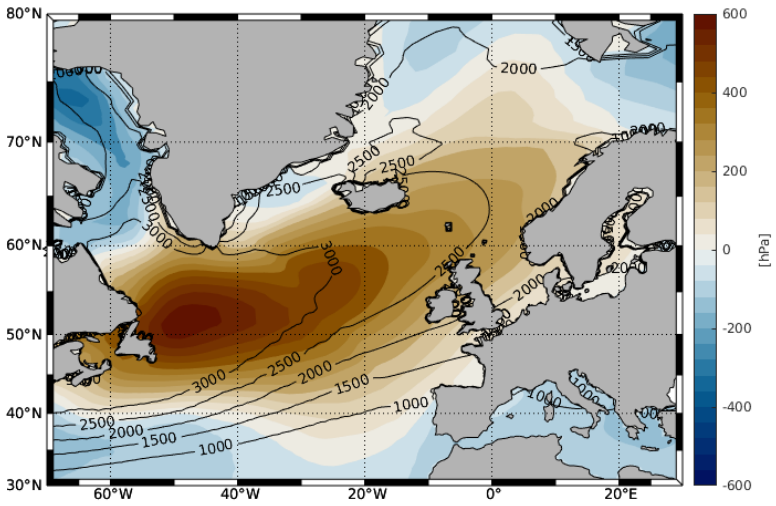


Figure 3.11 | Large-scale change of winter storminess (2-5-day Blackmon-filtered SLP). Black lines show the climatology in the low CO_2 world.

into two main types, the "North-West Type" and "West and South-West Type".

To investigate potential changes in the dominant storm track types, we perform a lagged Empirical Orthogonal Function (EOF) analysis (Weare & Nasstrom, 1982) on SLP data for each ESL event over the entire simulation period at Cuxhaven, with lead times of zero, 24 and 48 hours. The advantage of a lagged EOF (or often 'extended EOF') analysis is that it incorporates both spatial and temporal patterns and can thus unravel dynamical structures such as propagating structures or oscillations.

The first EOF (Fig. 3.12) explains 60% of the variance and shows the predominant storm track with a path from Iceland over the southern Norwegian Sea towards southern Scandinavia. This track fits the "North-West Type" category and has been found to be responsible for most of the storm floods in the German Bight (e.g. Ganske et al., 2018; Gerber et al., 2016). The second and third EOFs show deviations from this track, e.g. a more northern or southern track (see Supplementary Figures B.5 and B.6).

By analysing the distribution of the corresponding principal components during both low and high CO_2 states, one can identify temporal changes in their relative frequency, and hence potentially shifts in the importance of each track type. Fig. 3.12 shows histograms of the corresponding first principal component (PC1) for both low and high CO_2 states. It can be seen that the distribution shifts towards more positive values (significant on the 95% confidence level using a two-sided student-t test). The quantile-quantile plot suggests that this shift is partly due to an increase in the frequency of highest absolute values increases. This suggests that storms of this North-West-Type become more severe with rising atmospheric CO_2 levels. The statistics of the next PCs remains similar, indicating that there is – other than an intensification in the dominating track type – no change in the relative predominance of storm track types.

3.4 DISCUSSION

The increase in ESL heights with rising atmospheric CO_2 is qualitatively consistent with the results from, e.g., Lowe et al. (2001), Woth et al. (2006), Debernard & Røed (2008), Gaslikova et al. (2013) or Vousdoukas et al. (2017). Yet, these studies differ in the experimental design as they used barotropic storm surge models driven by atmospheric GCM data, in contrast to the coupled GCM-RCM setup described here. Further, given their smaller ensemble size, the increase in ESL heights is (i) statistically not distinguishable from internal variability (e.g. Debernard & Røed, 2008; Flather & Smith, 1998; Gaslikova et al., 2013; Langenberg et al., 1999; Sterl et al., 2009; von Storch & Reichardt, 1997, or (ii) only evident for more moderate ESLs with return periods of the order of magnitude of one year (Lowe et al., 2001; Woth et al., 2006). Here, we find statistically significant climate change signals of up to 30-40 cm in ESLs of return periods of up to approx. 30 years, with largest

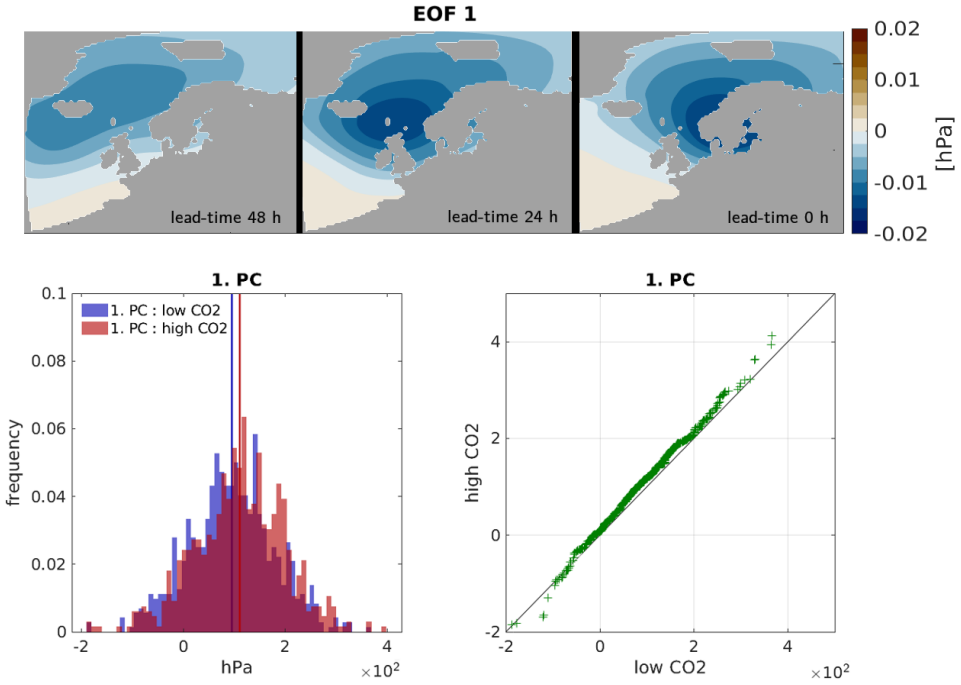


Figure 3.12 | Top: First lagged EOF of the hourly SLP field during ESL events with lead-time 48, 24 and 0 hours. Bottom: Histograms (left, mean indicated as vertical line) and quantile-quantile plot of the corresponding first principal component (1.PC) for low and high CO_2 states.

differences confined to the eastern German Bight (North Frisian Islands) and the Danish coast. These increases are in agreement with Vousdoukas et al., 2017 who used a hydrodynamic model forced by an 8-member climate model ensemble for RCP8.5. The spatial climate change response pattern of ESLs found here is consistent with the prevailing westerly wind directions during storms and their increase in magnitude in the high-CO_2 world.

The simulated increase in ESLs, independent of any climate-related background sea level rise, stresses the importance of foresighted planning of coastal defense structures and other adaptation measures. At the same time, the large ESL variability of individual members suggests that interannual to multidecadal deviations from the pooled ensemble signal are possible and may in fact mask an expected increase in ESL heights in the short term, or worse, overshoot an expected ESL pathway. This is especially evident for the upper tail of the ESL distribution, i.e. for floods of return periods of 50 years and more. Such multidecadal variability in ESLs might be related to variations in the large-scale atmospheric modes such as the NAO (Hurrell, 1995) or other NAO-like patterns that have been linked to ESLs in the German Bight (Dangendorf et al., 2014a; Lang & Mikolajewicz, 2019).

3.4.1 *ESL changes in the context of sea level rise*

The results presented above only consider the relative change of ESL statistics and do not account for a sea level rise due to changes in volume, through, e.g. thermal expansion, the melting of ice caps or vertical land movement. Inclusion of these processes will, additionally to the presented changes in ESL statistics, lead to a shift of the entire sea level distribution, and thus amplify the rise in ESL heights.

To illustrate what such a superposition of background sea level rise (SLR) could effectively mean for probabilistic flood events in the German Bight region, we extend the analysis of return period changes in section 3.3.1 by including estimates of SLR until the end of the century. With atmospheric CO_2 concentrations of about 1200 ppm and a change in GMST of more than four degrees at the end of the 150 years of simulation, the here underlying 1%CO_2 scenario is an idealized, yet – considering current emission pathways – rather drastic emission scenario. To set the results of the above analysis into the context of a more commonly used scenario, we translate the estimates of our *1pctCO2* simulation into the IPCC's RCP8.5 pathway by scaling them with GMST. That is, we calculate the change in ESLs not only until the end of the 150-year simulation period, but until the year when the GMST of 1pctCO2 ensemble mean equals the one from the corresponding RCP8.5 ensemble simulations. As sea level is expected to change differently over the globe, we use gridded estimates of end-of-the-century sea level rise for each component from Carson et al., 2016, which is partly based on Slangen et al., 2014. These different components of the SLR comprise thermosteric and dynamic sea level (both taken from our simulation in year 125), as well as contributions from ice sheets of Antarctica and Greenland, melting of glaciers, glacial isostatic adjustment (GIA) and groundwater depletion (see table 3.1). These estimates are similar to the IPCC AR5 (Church et al., 2013), but with a considerable upward correction concerning the dynamic ice contribution from Antarctica (see IPCC, 2019). Together they re-

Table 3.1 | Magnitude of SLR components in the German Bight at the end of the century under RCP8.5

SLR component	SLR [meter]	Source
Thermosteric	0.25 m	REMO-MPIOM
Dynamic	0.15 m	REMO-MPIOM
Antarctica	0.17 m	Carson et al. (2016)
Greenland	0.01 m	Carson et al. (2016)
Groundwater storage	0.04 m	Carson et al. (2016)
GIA	0.14 m	Carson et al. (2016)
Glaciers	0.16 m	Carson et al. (2016)

sult in around 0.9 m of SLR in the southern North Sea at the end of the century. Note though that all contributions, especially the GIA estimates, show considerable regional differences in the exact numbers. Comparing this number with the climate change signals from section 3.3.1, shows that the magnitude of ESL changes alone without considering a SLR lies, at least at the coast, at around half of the region's end-of-the-century SLR estimate.

Without considering a change in ESL statistics, i.e. by simply shifting the distribution by the SLR estimate as often done for extreme sea level projections, for example in the IPCC SROCC (IPCC, 2019), a 50-year event would be reached every ten to five years along the German Bight coast (Fig. 3.13a). Note that the SLR estimate corresponds to the IPCC's rather conservative "best estimate". Using the upper estimate with the highly uncertain contribution from the Antarctic ice sheet would yield an even stronger reduction in return periods.

However, by considering SLR *together* with the above described changes in ESL statistics, the climate change impact on return periods becomes even more drastic. Since it has been shown that a rise in BSL can, to a first approximation, be added linearly onto the atmosphere-induced water levels (e.g. Lowe & Gregory, 2005; Sterl et al., 2009), we add the gridded SLR estimates onto the change in return period of a preindustrial 50-year return level (compare Fig. 3.5) until the RCP8.5 end-of-century equivalent year 125.The resulting change in the return period of a preindustrial 50-year return level is shown in Fig. 3.13b. Along the German Bight coast, a preindustrial 50-year event could thus be reached every five to three years at the end of the century. Note that the change factors are strongly determined by the local ESL variability: Locations along the coast where absolute ESLs are highest will experience rather moderate relative changes in return periods. Further offshore, SLR will exceed the here relatively low annual maximum values, hence leading to comparably high change factors. Yet, even along the coast, the return periods increase by a factor of 10-20, thus leading to twice as high change factors as for SLR only. Accordingly, the 50-year return level would rise by more than 1 meter if both BSL and ESL changes are considered. This underlines the importance of considering both ESL and BSL changes. A simple shift of the past sea level distribution to a higher mean could thus result in large and potentially costly errors.

3.4.2 *Limitations*

Two main caveats that add uncertainty to the presented results are introduced by (1) the model system and (2) the extreme value sampling.

Concerning (1), the ocean model used in this study has with a horizontal resolution of up to 9 km in the German Bight a relatively high resolution compared to traditional global ocean models. Yet, it is still too coarse to resolve local details of coastline and bottom topography. Similarly, its upper-

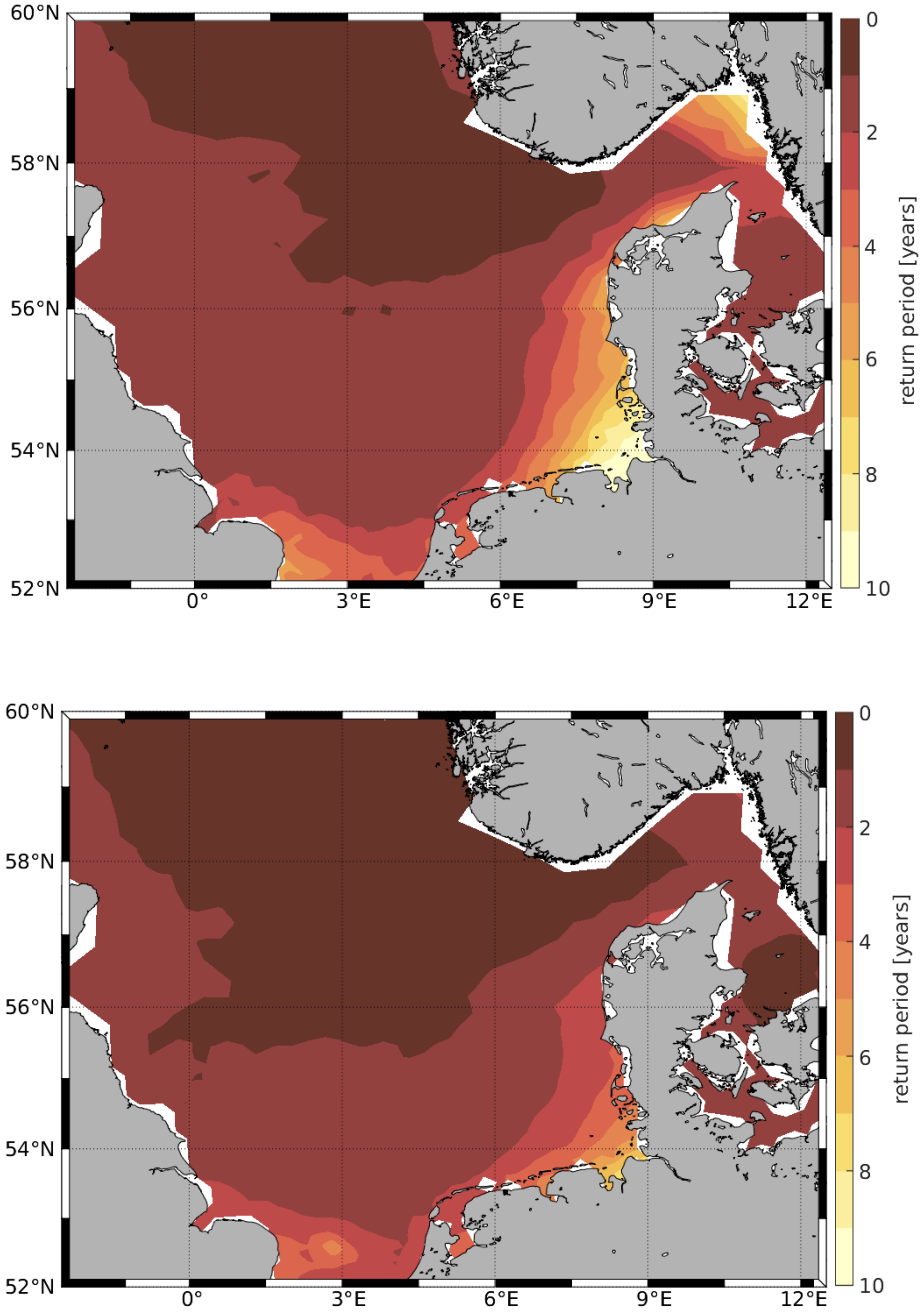


Figure 3.13 | End-of-the-century RCP8.5-equivalent return period of the preindustrial 50-year return level, estimated from considering SLR only (top), and from both BSLR and ESL change (bottom).

most layer thickness of 16 meters cannot resolve the finer vertical structures of the predominantly shallow waters in most coastal parts of the German Bight, especially along the flatter Wadden Sea coast. Here, the shallow waters and the strong tidal oscillations lead to coastline changes that are not accounted for in this model study. A direct comparison between simulated and observed sea levels should therefore be treated with caution. However, using data from the tide gauge record at Cuxhaven, Lang & Mikolajewicz, 2019 have shown that the regionally coupled climate system model REMO-MPIOM adequately simulates extreme sea level variations. Furthermore, although the large ensemble size enables an adequate sampling of the large internal variability, the choice of a single GCM-RCM model chain implies that the here presented results may be model dependent and thus that potential biases in the parent GCM could feed into the downscaled results. Finally, we have not considered the contribution to extreme sea levels from coastal waves, which add an additional hazard to coastal ecosystems.

Concerning (2), it has been shown that different measures of extreme events are sensitive to the choice of extreme value sampling index (e.g. Wahl et al., 2017). Several techniques have been applied to characterize ESL, from the use of block maxima (e.g. Marcos et al., 2009; Méndez et al., 2007) over percentiles (e.g. Dangendorf et al., 2013b; Woodworth & Blackman, 2004), to the selection of events over a given threshold (e.g. Méndez et al., 2006). The choice of the respective selection definition essentially represents a trade-off between bias (too many sampled extremes) and variance (too few sampled extremes) for the estimates. Here, we have tested several approaches, leading to no qualitative difference to the presented results. The choice of the block maxima method to represent ESLs has the advantage that it is robust to temporal variations in their magnitude. Instead of relative definitions such as 'surge residual' or 'skew surge' we here use a 'direct' method for sampling ESLs, because it eliminates the need to differentiate between tidal and surge parts and their nonlinear interaction. Further, as most studies with small ensembles rely on much fewer sampled events, an extrapolation by fitting an assumed extreme value distribution is common practice. Yet again, the choice of different extreme value distribution fits onto the same data can lead to substantial differences in the estimated extreme value indices (Wahl et al., 2017). In contrast, the large sample of extreme values due to the ensemble approach here allows the direct computation of extreme indices without the need to rely on such parametric methods.

Finally, it should be noted that irreducible uncertainties in future sea level extremes will remain regarding emission scenario (rather important in the long-term) and initial conditions representing the internal variability of both BSL and atmosphere-induced ESL variations as well as their regional variations (rather important in the short term).

3.5 SUMMARY & CONCLUSION

In this study we quantified the changes in North Sea extreme sea level statistics using a large ensemble of transient climate change simulations, downscaled with a coupled climate system model focusing on the southern North Sea. The 32 ensemble members have produced a large enough sample for a more robust analysis of higher extreme values, which can give new insights into future changes of the often desired high-impact-low-probability events. At the same time, the coupled climate model approach with a global ocean allows the free propagation of signals from the North Atlantic to the North Sea as well as a consistent analysis of related drivers in the climate system.

Specifically, we have found that ESL heights increase in terms of return levels, with significant changes for return periods of up to approx. 30 years. The change is most pronounced in the winter season, while the summer shows an opposite response as ESL heights reduce with increasing atmospheric CO_2 concentrations. Analysing changes in the underlying atmospheric drivers, we found that these changes in ESL statistics are mainly driven by a cascade of a *large-scale*, hemispheric response (increased activity along the North Atlantic storm belt) to a *regional* increase in wind speed maxima, eventually leading to changes in surge heights and ESL statistics. The seasonality of changes in ESL statistics is also evident in the response of atmospheric drivers.

Extreme high sea levels will thus not only scale with the expected change in BSL, but may become even more extreme due to a widening of the sea level distribution. The resulting change in ESL alone accounts hereby for around half of what we expect the background sea level to rise in the region. A high temporal variability on interannual to multidecadal scales, particularly of the highest extreme values, though may complicate the picture in reality, as the variations in a single realization – or the real world for that matter – may temporally exceed or counteract the climate change signal. Opposing seasonal patterns suggest that the time of largest ESLs will become more confined to the winter months in a future warmer world.

Even though the *1pctCO2* scenario can be regarded as an idealized scenario, the qualitative conclusions can still hold for present-day emission pathways and thus be of use for decision-making in coastal protection. The combined effect of mean and extreme sea level rise could surpass traditional estimates based on a simple shift of the sea level distribution. This implies that ESL conditions as during preindustrial times may occur almost every year at the end of the century and thus stresses the necessity of timely adaptation and coastal protection measures.

Author contributions. AL and UM jointly designed the experiments; simulation and analysis were carried out by AL. The paper was written by AL with contributions from UM.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. The research was supported by the Max Planck Society and by the German Research Foundation (DFG) funded project SEASTORM under the umbrella of the Priority Programme SPP-1889 "Regional Sea Level Change and Society" (grant no. MI 603/5-2). The authors thank Sebastian Milinski for his constructive comments and remarks which helped to improve this paper. We further acknowledge the German Climate Computing Center (DKRZ) for providing the necessary computational resources.

Financial support. This research has been supported by the German Research Foundation (DFG) (grant no. MI 603/5-2).

APPENDIX

The following Appendices comprise the Supplementary Material of the above articles.

A

SUPPLEMENTARY MATERIAL FOR CHAPTER 2

The following has been published as Supplementary Material for the article "The long-term variability of extreme sea levels in the German Bight" (Lang & Mikolajewicz, 2019), available under https://doi.org/10.5194/os-15-651-2019-supplement.

COMPARISON WITH THE INSTRUMENTAL RECORD

Sea level observations used in this study stem from the tide gauge record at Cuxhaven (courtesy S. Dangendorf & W. Wiechmann). Comparison of simulated and observed extreme sea level at Cuxhaven in terms of the respective highest event (Fig. A.1) and their seasonal statistics relative to the long-term tidal mean high water (Fig. A.2).

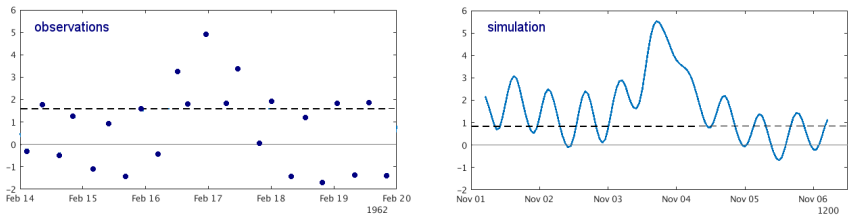


Figure A.1 | Time series of highest storm surge from observations (left) and model simulation (right). The dashed black line indicates the long-term mean high water. The respective long-term mean has been removed from both time series (grey line).

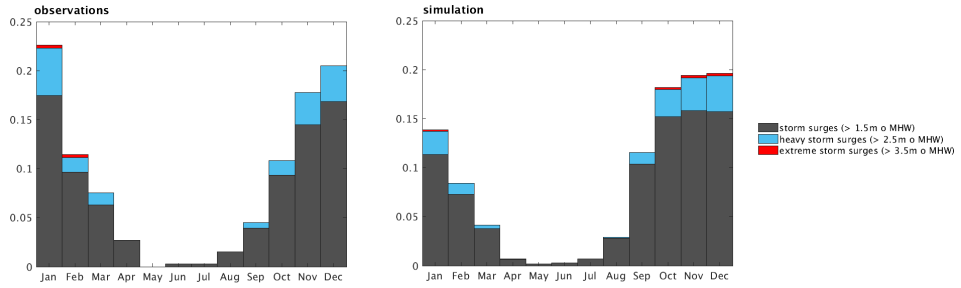


Figure A.2 | Relative storm surge frequency per month for observations (left) and model simulation (right) for storm surge classes following the definitions by the Federal Maritime and Hydrographic Agency (BSH).

EXTERNAL FORCING FOR SIMULATIONS

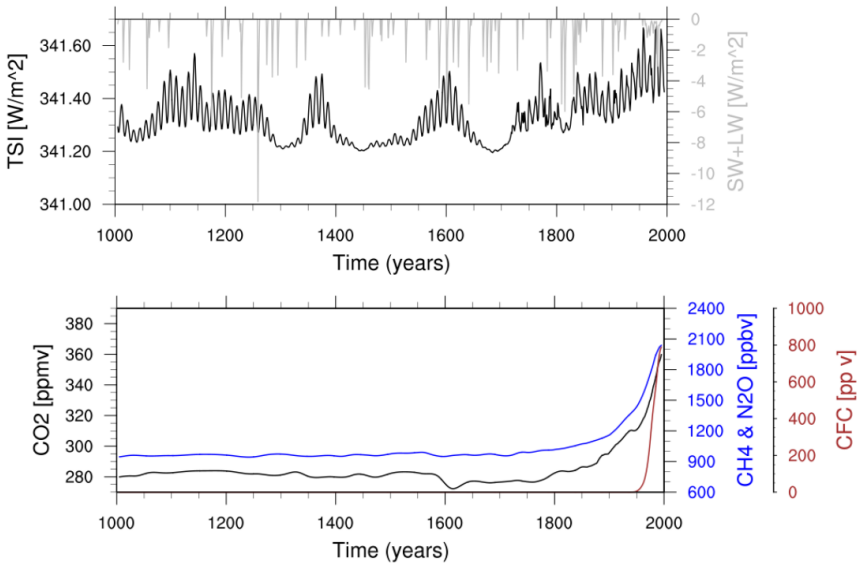


Figure A.3 | *Top:* Forcing input of annual solar irradiance [Wang et al. (2005), black line] and volcanic eruptions [Crowley et al. (2008), grey line]. *Bottom:* Greenhouse gas forcing, including CO_2 (black), CH_4 and N_2O (blue) and Chlorofluorocarbons (CFC, dark red).

DOWNSCALING SIMULATIONS AND SOURCES OF VARIABILITY

In order to investigate the contributions of external forcing and natural variability on ESL variations, as well as effect of the downscaling process on ESL variability, two additional experiments have been performed (see Table A.2).

First, to analyze the contribution of external forcing and natural variability, we have – additional to the downscaling shown in the main text (experiment *011*) – downscaled a second member of the parent global *Last-Millennium* simulations ('past1000r1', see Moreno-Chamarro et al., 2017), covering the period 1400-1850 (experiment *012*), with the first 100 years again used as spin-up. While the ESL statistics in terms of a quantile-quantile plot are comparable, the two downscalings show different temporal ESL variations, indicating that the externally forced variability is small compared to the natural variability.

Table A.2 | Downscaling simulations used in this study

run ID	parent GCM simulation (MPI-ESM)	time period
010	past1000r2 + historicalr4	1500-2000 (+100 year spin-up)
011	past1000r2 + historicalr4	1000-2000 (+100 year spin-up)
012	past1000r1	1500-1850 (+100 year spin-up)

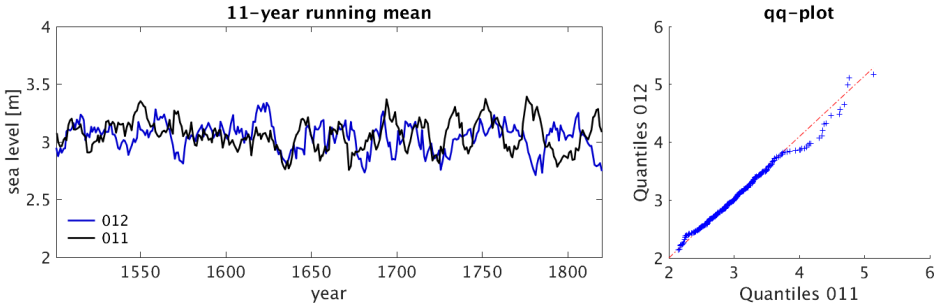


Figure A.4 | Comparison of ESL at Cuxhaven between runs *011* and *012* for the years 1500-1850. Left: Time series for 11-year running means. Right: Quantile-Quantile plot.

Second, to investigate the effect of the downscaling process on ESL variability, the global PMIP3 simulation 'past1000r2' (1000-1850; Moreno-Chamarro et al., 2017) and the subsequent 'historicalr4' (1850-2000) have been downscaled twice, once over one continuous 1000 year simulation (1000-2000AD.; experiment *011*; used for results shown in the main text), and once over

the second 500 years (1500-2000AD.; experiment *010*) with slightly different initial conditions. Both simulation are preceded by a 100 year long spin-up. An overview of the downscaling simulations is given in Table A.2.

Although there is a tendency towards higher extremes in 010, especially for return periods greater than 30 years, the main features of long-term variability and spectral characteristics are not affected by the downscaling (Fig. A.5). Nonetheless, the downscaling with REMO-MPIOM allows a more detailed simulation of large-scale dynamics (e.g. European blocking events).

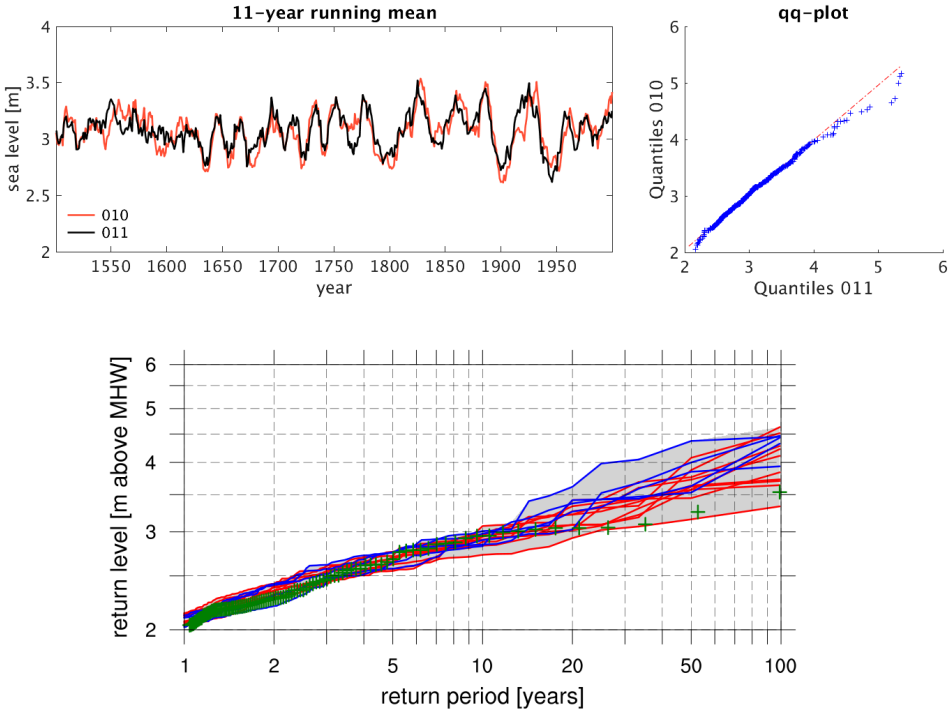


Figure A.5 | Comparison of ESL at Cuxhaven between runs *010* and *011* for the years 1500-2000. Left: Time series for 11-year running means. Right: Quantile-Quantile plot. Bottom: Return value plot of simulated sea level at Cuxhaven [m over MHW] (colored lines representing 100-year long segments of the full 1000 years) against observations from tide gauges (green crosses) for run *011* (red, 1000-2000) and for run *010* (blue, 1500-2000)

ESL VARIABILITY: SPATIAL COHERENCE

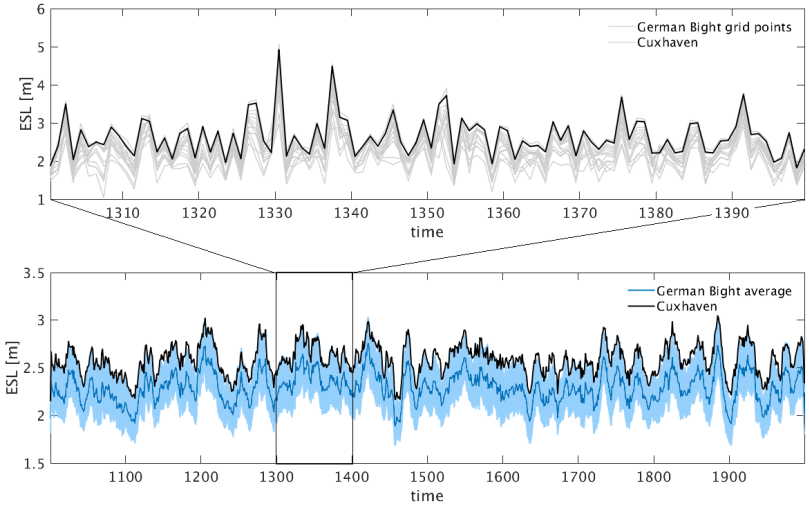


Figure A.6 | Comparison of ESL along the German Bight. Top: Annual maximum sea level at Cuxhaven in comparison to other gridpoints at the German Bight coast (grey). Bottom: 11-year running mean of the annual maximum sea level at Cuxhaven (black) and the spatially aggregated annual maxima along the German Bight coast (blue) ± one standard deviation (light blue).

ESL VARIABILITY: RELATION TO BSL

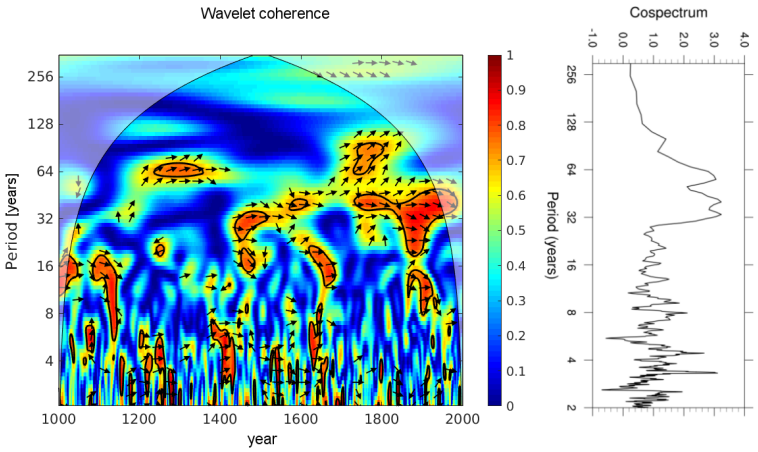


Figure A.7 | Wavelet coherence and cospectrum of winter median and annual maximum sea level. Arrows to the right (left) indicate a positive (negative) correlation and upward (downward) arrows indicate a lag (lead) of winter median sea level. Thick contours designate the 5% significance level against red noise, the cone of influence is shown as a lighter shade.

ESL VARIABILITY: RELATION TO EXTERNAL FORCING

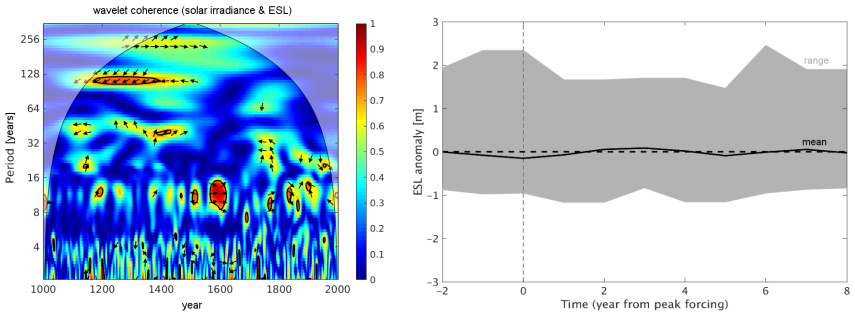


Figure A.8 | *Left:* Wavelet coherence between solar irradiance and ESL. Arrows to the right (left) indicate a positive (negative) correlation and upward (downward) arrows indicate a lag (lead) of solar irradiance. *Right:* Mean lagged ESL after major volcanic activity (perturbance $> 4\,\mathrm{W\,m^{-2}}$).

B

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

This appendix follows the Supplementary Material for the submitted article "Rising extreme sea levels in the German Bight under enhanced CO_2 levels – a regionalized large ensemble approach for the North Sea" (Lang & Mikolajewicz) and contains Figures B.1 - B.6, complementary to the ones shown in the main article.

INTERNAL ESL VARIABILITY

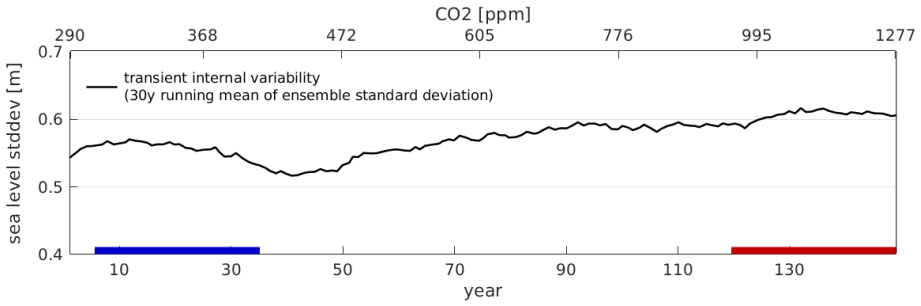


Figure B.1 | Transient internal variability of ESL at Cuxhaven in terms of the ensemble standard deviation. The two time slices used for the climate change analysis are marked by the blue and red bars.

ATMOSPHERIC DRIVERS: SUMMER RESPONSE

Additionally to the winter response shown in the main article we show here the corresponding summer response for 10m wind and large-scale storminess.

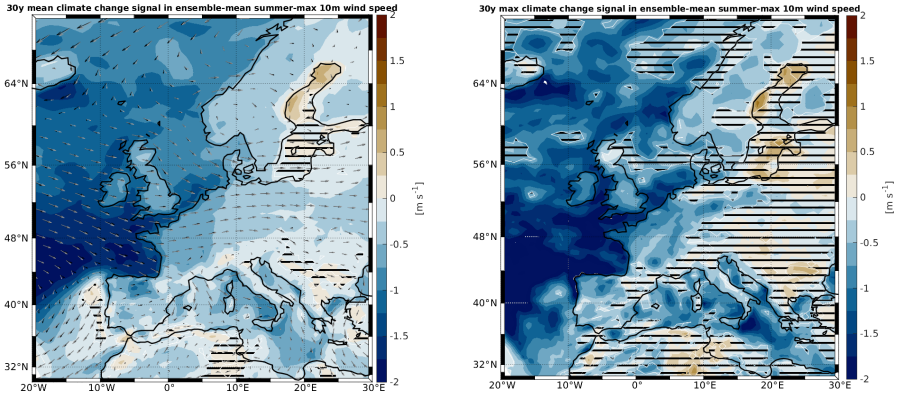


Figure B.2 | Summer change in 10 m wind of 1-year return period (left) and 30-year return period (right).

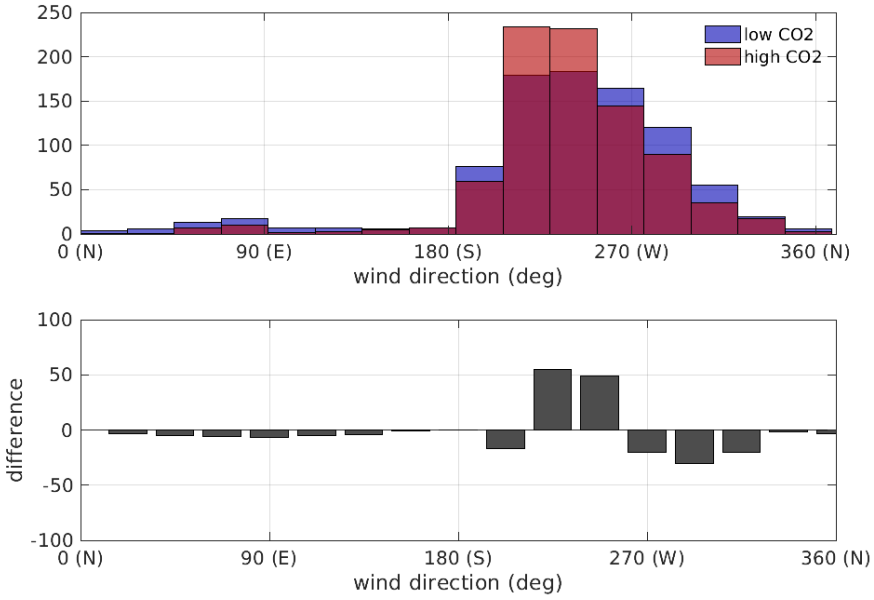


Figure B.3 | Summer change in 10 m wind direction of 1-year return periods. Areas not significant on the 95% confidence level are striped out.

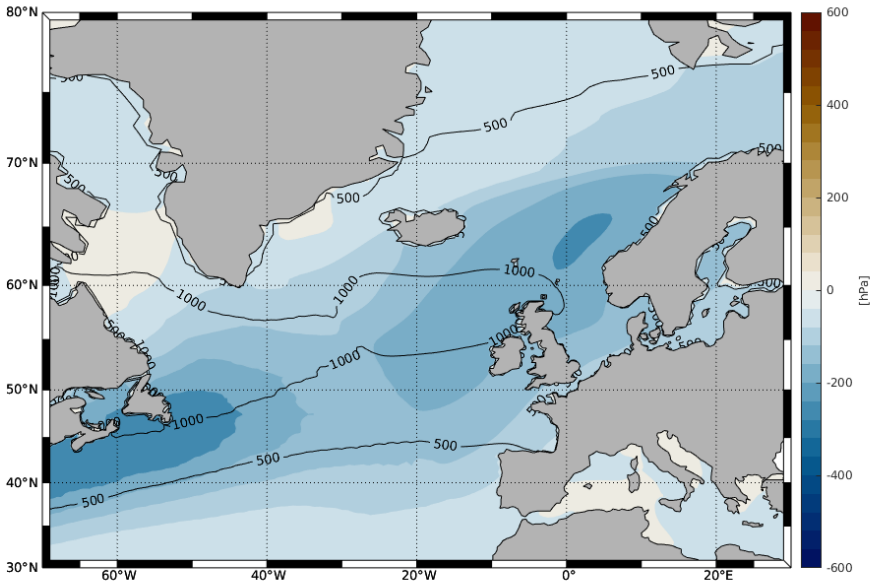


Figure B.4 | Summer change in storminess (based on 2-5-day Blackmon-filtered SLP).

LAGGED EOF ANALYSIS

Additionally to the 1st EOF shown in the article, we include here the 2nd and 3rd lagged EOF as well as statistics of their respective principal components.

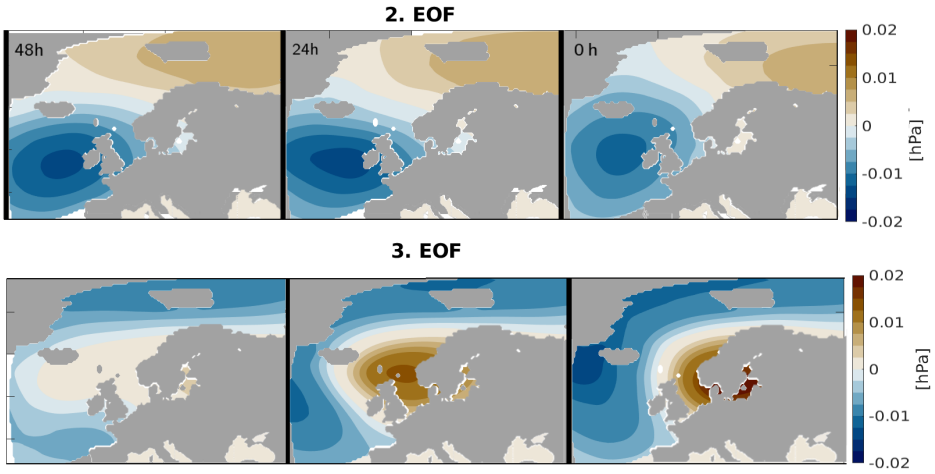


Figure B.5 | Second and third lagged EOF of SLP associated with ESL at Cuxhaven.

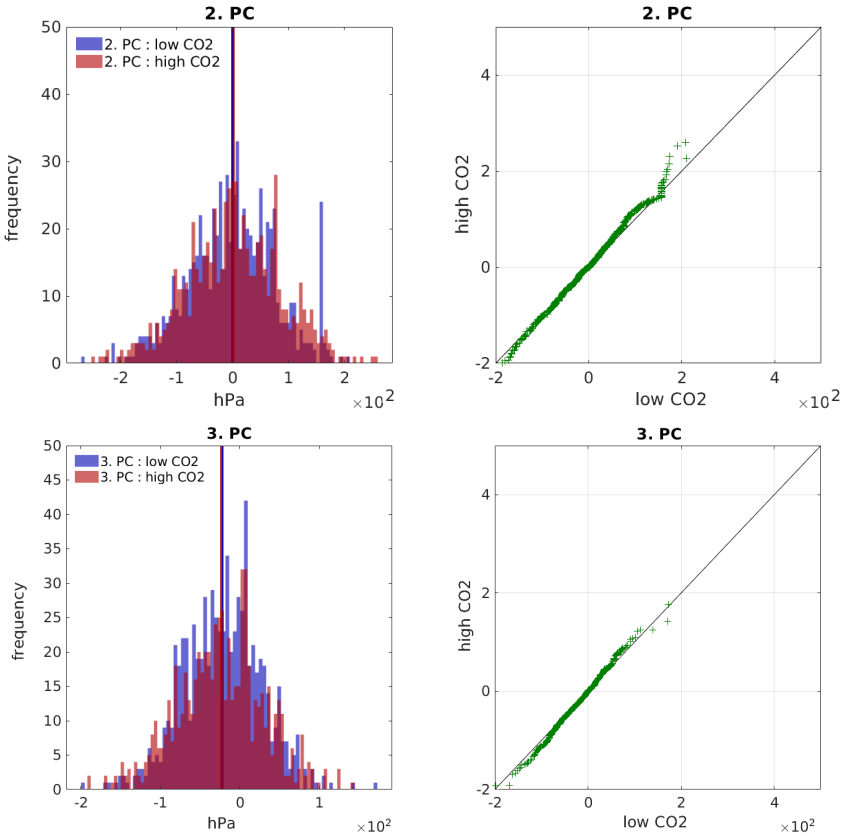


Figure B.6 | Statistics of second and third principal components.

BIBLIOGRAPHY

Araújo, I. B. & D. T. Pugh (2008). "Sea levels at Newlyn 1915–2005: analysis of trends for future flooding risks." In: *Journal of Coastal Research* 24.sp3, pp. 203–212.

Arends, F. (1974). *Physische Geschichte der Nordsee-Küste und deren Veränderungen durch Sturmfluten seit der Cymbrischen Fluth bis jetzt: erster und zweiter Band*. Schuster.

Arns, A, T Wahl, I. Haigh, J Jensen, & C Pattiaratchi (2013). "Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise." In: *Coastal Engineering* 81, pp. 51–66.

Arns, A., T. Wahl, S. Dangendorf, & J. Jensen (2015). "The impact of sea level rise on storm surge water levels in the northern part of the German Bight." In: *Coastal Engineering* 96, pp. 118–131.

Arns, A., S. Dangendorf, J. Jensen, S. Talke, J. Bender, & C. Pattiaratchi (2017). "Sea-level rise induced amplification of coastal protection design heights." In: *Scientific reports* 7, p. 40171.

Bamber, J. L., M. Oppenheimer, R. E. Kopp, W. P. Aspinall, & R. M. Cooke (2019). "Ice sheet contributions to future sea-level rise from structured expert judgment." In: *Proceedings of the National Academy of Sciences* 116.23, pp. 11195–11200.

Barriopedro, D., R. García-Herrera, P. Lionello, & C. Pino (2010). "A discussion of the links between solar variability and high-storm-surge events in Venice." In: *Journal of Geophysical Research: Atmospheres* 115.D13.

Beersma, J., K. Rider, G. Komen, E Kaas, & V. Kharin (1997). "An analysis of extra-tropical storms in the North Atlantic region as simulated in a control and 2× CO2 time-slice experiment with a high-resolution atmospheric model." In: *Tellus A* 49.3, pp. 347–361.

Bengtsson, L., K. I. Hodges, & E. Roeckner (2006). "Storm tracks and climate change." In: *Journal of Climate* 19.15, pp. 3518–3543.

Bengtsson, L., K. I. Hodges, & N. Keenlyside (2009). "Will extratropical storms intensify in a warmer climate?" In: *Journal of Climate* 22.9, pp. 2276–2301.

Beniston, M., D. B. Stephenson, O. B. Christensen, C. A. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylhä, B. Koffi, et al. (2007). "Future extreme events in European climate: an exploration of regional climate model projections." In: *Climatic change* 81.1, pp. 71–95.

Blackmon, M. L. (1976). "A climatological spectral study of the 500 mb geopotential height of the Northern Hemisphere." In: *Journal of the Atmospheric Sciences* 33.8, pp. 1607–1623.

Buchanan, M. K., M. Oppenheimer, & R. E. Kopp (2017). "Amplification of flood frequencies with local sea level rise and emerging flood regimes." In: *Environmental Research Letters* 12.6, p. 064009.

Buisman, J. (2006). "Duizend jaar weer, wind en water in de lage landen, deel 5: 1675-1750." In: *Uitgeverij Van Wijnen-Franeker, ISBN* 5.90, p. 5194.

Calafat, F., D. Chambers, & M. Tsimplis (2012). "Mechanisms of decadal sea level variability in the eastern North Atlantic and the Mediterranean Sea." In: *Journal of Geophysical Research: Oceans* 117.C9.

Camphuysen, C. J. & M. F. Leopold (1994). *Atlas of seabirds in the southern North Sea*. Tech. rep. NIOZ [etc.]

Carson, M., A. Köhl, D. Stammer, A. Slangen, C. Katsman, R. Van de Wal, J Church, & N White (2016). "Coastal sea level changes, observed and projected during the 20th and 21st century." In: *Climatic Change* 134.1-2, pp. 269–281.

Cazenave, A. & W. Llovel (2010). "Contemporary sea level rise." In: *Annual review of marine science* 2, pp. 145–173.

Chafik, L., J. E. Ø. Nilsen, & S. Dangendorf (2017). "Impact of North Atlantic teleconnection patterns on Northern European sea level." In: *Journal of Marine Science and Engineering* 5.3, p. 43.

Chang, E. K., S. Lee, & K. L. Swanson (2002). "Storm track dynamics." In: *Journal of climate* 15.16, pp. 2163–2183.

Chen, X., S. Dangendorf, N. Narayan, K. O'Driscoll, M. N. Tsimplis, J. Su, B. Mayer, & T. Pohlmann (2014). "On sea level change in the North Sea influenced by the North Atlantic Oscillation: local and remote steric effects." In: *Estuarine, Coastal and Shelf Science* 151, pp. 186–195.

Christensen, J. H., T. R. Carter, M. Rummukainen, & G. Amanatidis (2007). *Evaluating the performance and utility of regional climate models: the PRUDENCE project*.

Church, J., P. Clark, A Cazenave, J. Gregory, S Jevrejeva, A Levermann, M. Merrifield, G. Milne, R. Nerem, P. Nunn, et al. (2013). "Sea level change. Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change." In: *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, pp. 1137–1216.

Cid, A., M. Menéndez, S. Castanedo, A. J. Abascal, F. J. Méndez, & R. Medina (2016). "Long-term changes in the frequency, intensity and duration of extreme storm surge events in southern Europe." In: *Climate dynamics* 46.5-6, pp. 1503–1516.

Coles, S., J. Bawa, L. Trenner, & P. Dorazio (2001). *An introduction to statistical modeling of extreme values*. Vol. 208. Springer.

Crowley, T. J., G Zielinski, B Vinther, R Udisti, K Kreutz, J Cole-Dai, & E Castellano (2008). "Volcanism and the little ice age." In: *PAGES news* 16.2, pp. 22–23.

Dangendorf, S., T. Wahl, H. Hein, J. Jensen, S. Mai, & C. Mudersbach (2012). "Mean sea level variability and influence of the North Atlantic Oscilla-

tion on long-term trends in the German Bight." In: *Water* 4.1, pp. 170–195.

Dangendorf, S., C. Mudersbach, T. Wahl, & J. Jensen (2013a). "Characteristics of intra-, inter-annual and decadal sea-level variability and the role of meteorological forcing: the long record of Cuxhaven." In: *Ocean Dynamics* 63.2-3, pp. 209–224.

Dangendorf, S., C. Mudersbach, J. Jensen, G. Anette, & H. Heinrich (2013b). "Seasonal to decadal forcing of high water level percentiles in the German Bight throughout the last century." In: *Ocean Dynamics* 63.5, pp. 533–548.

Dangendorf, S., T. Wahl, E. Nilson, B. Klein, & J. Jensen (2014a). "A new atmospheric proxy for sea level variability in the southeastern North Sea: observations and future ensemble projections." In: *Climate dynamics* 43.1-2, pp. 447–467.

Dangendorf, S., F. M. Calafat, A. Arns, T. Wahl, I. D. Haigh, & J. Jensen (2014b). "Mean sea level variability in the North Sea: Processes and implications." In: *Journal of Geophysical Research: Oceans* 119.10.

Dangendorf, S., S. Müller-Navarra, J. Jensen, F. Schenk, T. Wahl, & R. Weisse (2014c). "North Sea storminess from a novel storm surge record since AD 1843." In: *Journal of Climate* 27.10, pp. 3582–3595.

Daniell, P. (1946). "Discussion on Symposium on autocorrelation in time series." In: *J. Roy. Stat. Soc.* 8, p. 88.

Davies, A. & R. Flather (1977). "Computation of the storm surge of 1 to 6 April 1973 using numerical models of the North West European continental shelf and the North Sea." In: *Deutsche Hydrografische Zeitschrift* 30.5, pp. 139–162.

De Winter, R., A Sterl, & B. Ruessink (2013). "Wind extremes in the North Sea Basin under climate change: An ensemble study of 12 CMIP5 GCMs." In: *Journal of Geophysical Research: Atmospheres* 118.4, pp. 1601–1612.

DeConto, R. M. & D. Pollard (2016). "Contribution of Antarctica to past and future sea-level rise." In: *Nature* 531.7596, pp. 591–597.

Debernard, J. B. & L. P. Røed (2008). "Future wind, wave and storm surge climate in the Northern Seas: a revisit." In: *Tellus A: Dynamic Meteorology and Oceanography* 60.3, pp. 427–438.

Donat, M., G. Leckebusch, S Wild, & U Ulbrich (2011). "Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations." In: *Natural Hazards and Earth System Sciences* 11.5, p. 1351.

Donat, M. G., G. C. Leckebusch, J. G. Pinto, & U. Ulbrich (2010a). "European storminess and associated circulation weather types: future changes deduced from a multi-model ensemble of GCM simulations." In: *Climate Research* 42.1, pp. 27–43.

— (2010b). "Examination of wind storms over Central Europe with respect to circulation weather types and NAO phases." In: *International Journal of Climatology* 30.9, pp. 1289–1300.

Elizalde, A, M Gröger, M Mathis, U MIkOLAJEw-ICZ, K BüLOw, S Hüttl-Kabus, B Klein, & A GANSkE (2014). "MPIOM-REMO A Coupled Regional Model for the North Sea." In: *KLIWAS-Schriftenreihe KLIWAS-58/2014. DOI* 10.

Ezer, T., I. D. Haigh, & P. L. Woodworth (2016). "Nonlinear sea-level trends and long-term variability on Western European Coasts." In: *Journal of Coastal Research* 32.4, pp. 744–755.

FEMA, H. K. R. (2005). "Reconstruction Guidance Using Hurricane Katrina Surge Inundation and Advisory Base Flood Elevations." In: ed. by F. E. M. Agency.

FHH (2019). "Prävention Hochwasserschutz in Hamburg heute. – Hrsg.: Freie und Hansestadt Hamburg." In: *Behörde für Umwelt und Energie Hamburg*.

Feng, J., H. von Storch, W. Jiang, & R. Weisse (2015). "Assessing changes in extreme sea levels along the coast of C hina." In: *Journal of Geophysical Research: Oceans* 120.12, pp. 8039–8051.

Feser, F., M Barcikowska, O Krueger, F. Schenk, R Weisse, & L Xia (2015). "Storminess over the North Atlantic and northwestern Europe—A review." In: *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 350–382.

Fischer-Bruns, I., H. von Storch, J. González-Rouco, & E. Zorita (2005). "Modelling the variability of midlatitude storm activity on decadal to century time scales." In: *Climate dynamics* 25.5, pp. 461–476.

Fischer, E. M., U. Beyerle, & R. Knutti (2013). "Robust spatially aggregated projections of climate extremes." In: *Nature Climate Change* 3.12, p. 1033.

Flather, R. (1979). "Recent results from a storm surge prediction scheme for the North Sea." In: *Elsevier Oceanography Series*. Vol. 25. Elsevier, pp. 385–409.

— (1987). "Estimates of extreme conditions of tide and surge using a numerical model of the north-west European continental shelf." In: *Estuarine, Coastal and Shelf Science* 24.1, pp. 69–93.

Flather, R. & J. Smith (1998). "First estimates of changes in extreme storm surge elevations due to the doubling of CO sub (2)." In: *Global Atmosphere and Ocean System* 6.2, pp. 193–208.

Frankcombe, L. & H. Dijkstra (2009). "Coherent multidecadal variability in North Atlantic sea level." In: *Geophysical Research Letters* 36.15.

Gagen, M. H., E. Zorita, D. McCarroll, M. Zahn, G. H. Young, & I. Robertson (2016). "North Atlantic summer storm tracks over Europe dominated by internal variability over the past millennium." In: *Nature Geoscience* 9.8, p. 630.

Ganske, A., B. Tinz, G. Rosenhagen, & H. Heinrich (2016). "Interannual and multidecadal changes of wind speed and directions over the North Sea from climate model results." In: *Meteorologische Zeitschrift* 25, pp. 463–478.

Ganske, A., N. Fery, L. Gaslikova, I. Grabemann, R. Weisse, & B. Tinz (2018). "Identification of extreme storm surges with high-impact poten-

tial along the German North Sea coastline." In: *Ocean Dynamics* 68.10, pp. 1371–1382.

Gaslikova, L., I. Grabemann, & N. Groll (2013). "Changes in North Sea storm surge conditions for four transient future climate realizations." In: *Natural hazards* 66.3, pp. 1501–1518.

Gerber, M., A. Ganske, S. Müller-Navarra, & G. Rosenhagen (2016). "Categorisation of Meteorological Conditions for Storm Tide Episodes in the German Bight." In: *Meteorologische Zeitschrift*, pp. 447–462.

Giorgetta, M., E Roeckner, T Mauritsen, B Stevens, J Bader, T Crueger, M Esch, S Rast, L Kornblueh, H Schmidt, et al. (2012). "The atmospheric general circulation model ECHAM6." In: *Model description, Max Planck Inst. for Meteorol., Hamburg, Germany*.

Gómez-Navarro, J. & E. Zorita (2013). "Atmospheric annular modes in simulations over the past millennium: No long-term response to external forcing." In: *Geophysical Research Letters* 40.12, pp. 3232–3236.

Gönnert, G. (2004). "Maximum storm surge curve due to global warming for the European North Sea region during the 20th–21st century." In: *Natural hazards* 32.2, pp. 211–218.

Gönnert, G. & K. Sossidi (2011). "A new approach to calculate extreme storm surges." In: *Irrigation and drainage* 60, pp. 91–98.

Gönnert, G., B. Gerkensmeier, K. Sossidi, S. Thumm, J. Jensen, & T. Wahl (2012). "Katalog der Extremsturmfluten Zwischenbericht Teilprojekt." In:

Gregory, J. M., T Andrews, & P Good (2015). "The inconstancy of the transient climate response parameter under increasing CO2." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373.2054, p. 20140417.

Gregory, J. M., S. M. Griffies, C. W. Hughes, J. A. Lowe, J. A. Church, I. Fukimori, N. Gomez, R. E. Kopp, F. Landerer, G. Le Cozannet, et al. (2019). "Concepts and terminology for sea level: mean, variability and change, both local and global." In: *Surveys in Geophysics* 40.6, pp. 1251–1289.

Hadler, H, A Vött, J Newig, K Emde, C Finkler, P Fischer, & T Willershäuser (2018). "Geoarchaeological evidence of marshland destruction in the area of Rungholt, present-day Wadden Sea around Hallig Südfall (North Frisia, Germany), by the Grote Mandrenke in 1362 AD." In: *Quaternary International* 473, pp. 37–54.

Hagemann, S. & L. Dümenil (1999). "Application of a global discharge model to atmospheric model simulations in the BALTEX region." In: *Hydrology Research* 30.3, pp. 209–230.

Hanna, E., F. J. Navarro, F. Pattyn, C. M. Domingues, X. Fettweis, E. R. Ivins, R. J. Nicholls, C. Ritz, B. Smith, S. Tulaczyk, et al. (2013). "Ice-sheet mass balance and climate change." In: *Nature* 498.7452, pp. 51–59.

Heaps, N. (1967). "Storm surges." In: *Oceanography and Marine Biology: An Annual Review*.

Heimreich M. A. (Editor: Falck, N. (1819). *Nordfriesische Chronik*. Tondern.

Heyen, H., E. Zorita, & H. von Storch (1996). "Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea." In: *Tellus A* 48.2, pp. 312–323.

Hinkel, J., C. Jaeger, R. J. Nicholls, J. Lowe, O. Renn, & S. Peijun (2015). "Sea-level rise scenarios and coastal risk management." In: *Nature Climate Change* 5.3, p. 188.

Horsburgh, K. & C Wilson (2007). "Tide-surge interaction and its role in the distribution of surge residuals in the North Sea." In: *Journal of Geophysical Research: Oceans* 112.C8.

Howard, T., J. Lowe, & K. Horsburgh (2010). "Interpreting century-scale changes in southern North Sea storm surge climate derived from coupled model simulations." In: *Journal of Climate* 23.23, pp. 6234–6247.

Hu, A. & C. Deser (2013). "Uncertainty in future regional sea level rise due to internal climate variability." In: *Geophysical Research Letters* 40.11, pp. 2768–2772.

Hurrell, J (1995). "NAO index data provided by the climate analysis section, NCAR, Boulder, USA." In: *Online: http://www. cgd. ucar. edu/cas/jhurrell/indices. html*.

IPCC (2019). *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Poertner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegria, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]* Tech. rep. In press. IPCC.

Jacob, D. & R. Podzun (1997). "Sensitivity studies with the regional climate model REMO." In: *Meteorology and Atmospheric Physics* 63.1-2, pp. 119–129.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, et al. (2014). "EURO-CORDEX: new high-resolution climate change projections for European impact research." In: *Regional Environmental Change* 14.2, pp. 563–578.

Jakubowski-Tiessen, M. (1992). *Sturmflut 1717: die Bewältigung einer Naturkatastrophe in der frühen Neuzeit*. Vol. 24. Oldenbourg Verlag.

Jensen, J. & C. Mudersbach (2007). "Zeitliche Anderungen in den Wasserstandszeitreihen an den Deutschen Kusten." In: *Berichte zur deutschen Landeskunde* 81.2, p. 99.

Jensen, J. & S. H. Müller-Navarra (2008). "Storm surges on the German Coast." In: *Die Küste, 74 ICCE* 74, pp. 92–124.

Jensen, J., T. Frank, & T. Wahl (2011). "Analyse von hochaufgelösten Tidewasserständen und Ermittlung des MSL an der deutschen Nordseeküste (AMSeL)." In: *Die Küste, 78* 78, pp. 59–163.

Jungclaus, J., N. Fischer, H. Haak, K Lohmann, J Marotzke, D Matei, U Mikolajewicz, D Notz, & J. Storch (2013). "Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model." In: *Journal of Advances in Modeling Earth Systems* 5.2, pp. 422–446.

Jungclaus, J. H., K. Lohmann, & D. Zanchettin (2014). "Enhanced 20th century heat transfer to the Arctic simulated in context of climate variations over last millennium." In: *Climate of the Past* 10, pp. 2201–2213.

Kaniewski, D., N. Marriner, C. Morhange, S. Faivre, T. Otto, & E. Van Campo (2016). "Solar pacing of storm surges, coastal flooding and agricultural losses in the Central Mediterranean." In: *Scientific reports* 6, p. 25197.

Kauker, F. & H Langenberg (2000). "Two models for the climate change related development of sea levels in the North Sea a comparison." In: *Climate Research* 15.1, pp. 61–67.

Kauker, F. & H. von Storch (2000). "Regionalization of climate model results for the North Sea." In: *externer Bericht der GKSS, GKSS 2000/28*.

Kempe, M. (2006). "'Mind the next flood!'Memories of natural disasters in Northern Germany from the sixteenth century to the present." In: *The Medieval History Journal* 10.1-2, pp. 327–354.

Knudsen, M. F., B. H. Jacobsen, M.-S. Seidenkrantz, & J. Olsen (2014). "Evidence for external forcing of the Atlantic Multidecadal Oscillation since termination of the Little Ice Age." In: *Nature Communications* 5, p. 3323.

Kolker, A. S. & S. Hameed (2007). "Meteorologically driven trends in sea level rise." In: *Geophysical Research Letters* 34.23.

Kopp, R. E., C. C. Hay, C. M. Little, & J. X. Mitrovica (2015). "Geographic variability of sea-level change." In: *Current Climate Change Reports* 1.3, pp. 192–204.

Kruhl, H. (1978). "Sturmflutwetterlagen." In: *Promet* 8.4, pp. 6–8.

Lamb, H. & K. Frydendahl (1991). *Historic Storms of the North Sea, British Isles and Northwest Europe*. Cambridge University Press.

Lang, A. & U. Mikolajewicz (2019). "The long-term variability of extreme sea levels in the German Bight." In: *Ocean Science* 15, pp. 651–668.

Langenberg, H, A Pfizenmayer, H von Storch, & J Sündermann (1999). "Storm-related sea level variations along the North Sea coast: natural variability and anthropogenic change." In: *Continental shelf research* 19.6, pp. 821–842.

Lehmann, J., D. Coumou, K. Frieler, A. V. Eliseev, & A. Levermann (2014). "Future changes in extratropical storm tracks and baroclinicity under climate change." In: *Environmental Research Letters* 9.8, p. 084002.

Lowe, J. & J. Gregory (2005). "The effects of climate change on storm surges around the United Kingdom." In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 363.1831, pp. 1313–1328.

Lowe, J., J. M. Gregory, & R. Flather (2001). "Changes in the occurrence of storm surges around the United Kingdom under a future climate scenario using a dynamic storm surge model driven by the Hadley Centre climate models." In: *Climate dynamics* 18.3-4, pp. 179–188.

Lowe, J. A., P. L. Woodworth, T. Knutson, R. E. McDonald, K. L. McInnes, K. Woth, V Swail, N. Bernier, S Gulev, K. Horsburgh, et al. (2010). "Past and future changes in extreme sea levels and waves." In:

MELUR (2012). *Generalplan Küstenschutz des Landes Schleswig-Holstein*. Tech. rep. Fortschreibung. Fortschreibung.

MFLR (2012). "(Ministerium für ländliche Räume des Landes Schleswig-Holstein): Generalplan Küstenschutz des Landes Schleswig-Holstein." In:

Maher, N., S. Milinski, L. Suarez-Gutierrez, M. Botzet, L. Kornblueh, Y. Takano, J. Kröger, R. Ghosh, C. Hedemann, C. Li, et al. (2019). "The Max Planck Institute grand ensemble-enabling the exploration of climate system variability." In: *Journal of Advances in Modeling Earth Systems* 11, pp. 2050–2069.

Marcos, M. & P. L. Woodworth (2017). "Spatiotemporal changes in extreme sea levels along the coasts of the North Atlantic and the Gulf of Mexico." In: *Journal of Geophysical Research: Oceans* 122.9, pp. 7031–7048.

Marcos, M., M. N. Tsimplis, & A. G. Shaw (2009). "Sea level extremes in southern Europe." In: *Journal of Geophysical Research: Oceans* 114.C1.

Marcos, M., F. M. Calafat, Á. Berihuete, & S. Dangendorf (2015). "Long-term variations in global sea level extremes." In: *Journal of Geophysical Research: Oceans* 120.12, pp. 8115–8134.

Marsland, S. J., N. Bindoff, G. Williams, & W. Budd (2004). "Modeling water mass formation in the Mertz Glacier Polynya and Adélie Depression, east Antarctica." In: *Journal of Geophysical Research: Oceans* 109.C11.

Martínez-Asensio, A., M. N. Tsimplis, & F. M. Calafat (2016). "Decadal variability of European sea level extremes in relation to the solar activity." In: *Geophysical Research Letters* 43.22.

Mathis, M., A. Elizalde, & U. Mikolajewicz (2018). "Which complexity of regional climate system models is essential for downscaling anthropogenic climate change in the Northwest European Shelf?" In: *Climate Dynamics* 50.7-8, pp. 2637–2659.

Mawdsley, R. J. & I. D. Haigh (2016). "Spatial and temporal variability and long-term trends in skew surges globally." In: *Frontiers in Marine Science* 3, p. 29.

Méndez, F. J., M. Menéndez, A. Luceño, & I. J. Losada (2006). "Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model." In: *Journal of Geophysical Research: Oceans* 111.C7.

— (2007). "Analyzing monthly extreme sea levels with a time-dependent GEV model." In: *Journal of Atmospheric and Oceanic Technology* 24.5, pp. 894–911.

Menéndez, M. & P. L. Woodworth (2010). "Changes in extreme high water levels based on a quasi-global tide-gauge data set." In: *Journal of Geophysical Research: Oceans* 115.C10.

Mikolajewicz, U., D. V. Sein, D. Jacob, T. Königk, R. Podzun, & T. Semmler (2005). "Simulating Arctic sea ice variability with a coupled regional atmosphere-ocean-sea ice model." In: *Meteorologische Zeitschrift* 14.6, pp. 793–800.

Miller, G. H., Á. Geirsdóttir, Y. Zhong, D. J. Larsen, B. L. Otto-Bliesner, M. M. Holland, D. A. Bailey, K. A. Refsnider, S. J. Lehman, J. R. Southon, et al. (2012). "Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks." In: *Geophysical Research Letters* 39.2.

Miller, L. & B. C. Douglas (2004). "Mass and volume contributions to twentieth-century global sea level rise." In: *Nature* 428.6981, pp. 406–409.

Mölter, T., D. Schindler, A. T. Albrecht, & U. Kohnle (2016). "Review on the projections of future storminess over the North Atlantic European region." In: *Atmosphere* 7.4, p. 60.

Moreno-Chamarro, E., D. Zanchettin, K. Lohmann, & J. H. Jungclaus (2017a). "An abrupt weakening of the subpolar gyre as trigger of Little Ice Age-type episodes." In: *Climate Dynamics* 48.3-4, pp. 727–744.

Moreno-Chamarro, E., D. Zanchettin, K. Lohmann, J. Luterbacher, & J. H. Jungclaus (2017b). "Winter amplification of the European Little Ice Age cooling by the subpolar gyre." In: *Scientific Reports* 7.1, p. 9981.

Mudersbach, C. & J Jensen (2010). "Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline." In: *Journal of Flood Risk Management* 3.1, pp. 52–62.

Mudersbach, C., T. Wahl, I. D. Haigh, & J. Jensen (2013). "Trends in high sea levels of German North Sea gauges compared to regional mean sea level changes." In: *Continental Shelf Research* 65, pp. 111–120.

Müller-Navarra, S., W Lange, S Dick, & K. Soetje (2003). "Über die Verfahren der Wasserstands-und Sturmflutvorhersage." In: *promet* 29.1–4, pp. 117–124.

MunichRe (2012). *Press Dossier*.

Neumann, B., A. T. Vafeidis, J. Zimmermann, & R. J. Nicholls (2015). "Future coastal population growth and exposure to sea-level rise and coastal flooding-a global assessment." In: *PloS one* 10.3, e0118571.

Niemeyer, H. (1983). "Über den Seegang an einer inselgeschützten Wattküste." In: *on the wave climate at an island-sheltered coast. BMFT-Forschungsbericht MF* 203.

Otterå, O. H., M. Bentsen, H. Drange, & L. Suo (2010). "External forcing as a metronome for Atlantic multidecadal variability." In: *Nature Geoscience* 3.10, p. 688.

Park, J., J. Obeysekera, M. Irizarry-Ortiz, J. Barnes, & W. Park-Said (2010). "Climate links and variability of extreme sea-level events at Key West, Pensacola, and Mayport, Florida." In: *Journal of waterway, port, coastal, and ocean engineering* 136.6, pp. 350–356.

Pattyn, F., C. Ritz, E. Hanna, X. Asay-Davis, R. DeConto, G. Durand, L. Favier, X. Fettweis, H. Goelzer, N. R. Golledge, et al. (2018). "The Greenland and Antarctic ice sheets under 1.5 C global warming." In: *Nature climate change* 8.12, pp. 1053–1061.

Pawlowicz, R., B. Beardsley, & S. Lentz (2002). "Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE." In: *Computers & Geosciences* 28.8, pp. 929–937.

Peeck, H., R Proctor, & C Brockmann (1983). "Operational storm surge models for the North Sea." In: *Continental Shelf Research* 2.4, pp. 317–329.

Peltier, W. (2001). "Global glacial isostatic adjustment and modern instrumental records of relative sea level history." In: *International Geophysics*. Vol. 75. Elsevier, pp. 65–95.

Perrette, M, F Landerer, R Riva, K Frieler, & M Meinshausen (2013). "A scaling approach to project regional sea level rise and its uncertainties." In:

Pinto, J. G., U Ulbrich, G. Leckebusch, T Spangehl, M Reyers, & S Zacharias (2007). "Changes in storm track and cyclone activity in three SRES ensemble experiments with the ECHAM5/MPI-OM1 GCM." In: *Climate Dynamics* 29.2-3, pp. 195–210.

Plüß, A. (2004). "Nichtlineare Wechselwirkung der Tide auf Änderungen des Meeresspiegels im Übergangsbereich Küste/Ästuar am Beispiel der Elbe." In: *Proc Klimaänderung Küstenschutz*, pp. 129–138.

Proctor, R. & R. Flather (1989). "Storm surge prediction in the Bristol Channel—the floods of 13 December 1981." In: *Continental Shelf Research* 9.10, pp. 889–918.

Pugh, D. (1987). *Tides, surges and mean sea-level: a handbook for engineers and scientists, 472 pp*.

Pugh, D. & J. Vassie (1978). "Extreme sea levels from tide and surge probability." In: *Coastal Engineering 1978*, pp. 911–930.

Rietbroek, R., S.-E. Brunnabend, J. Kusche, & J. Schröter (2012). "Resolving sea level contributions by identifying fingerprints in time-variable gravity and altimetry." In: *Journal of Geodynamics* 59, pp. 72–81.

Rockel, B. & K. Woth (2007). "Extremes of near-surface wind speed over Europe and their future changes as estimated from an ensemble of RCM simulations." In: *Climatic Change* 81.1, pp. 267–280.

Roeckner, E., G Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, et al. (2003). "The atmospheric general circulation model ECHAM 5. PART I: Model description." In:

Schirmer, M. (2018). "WARNSIGNAL KLIMA: Extremereignisse." In: ed. by L. et al. Chap. Küstenschutz bis und nach 2100 in Deutschland und den Niederlanden.

Schmidt, G., J. Jungclaus, C. Ammann, E Bard, P Braconnot, T. Crowley, G Delaygue, F. Joos, N. Krivova, R. Muscheler, et al. (2012). "Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1. 1)." In: *Geoscientific Model Development* 5, pp. 185–191.

Schmitz, H., D Habicht, & H Volkert (1988). "Barotropic numerical experiments on external surge generation at the edge of the north-western European shelf." In: *Gerlands Beiträge zur Geophysik* 97.5, pp. 422–437.

Schrum, C., J. Lowe, H. M. Meier, I. Grabemann, J. Holt, M. Mathis, T. Pohlmann, M. D. Skogen, A. Sterl, & S. Wakelin (2016). "Projected change—North sea." In: *North Sea region climate change assessment*. Springer, Cham, pp. 175–217.

Seierstad, I., D. Stephenson, & N. Kvamstø (2007). "How useful are teleconnection patterns for explaining variability in extratropical storminess?" In: *Tellus A: Dynamic Meteorology and Oceanography* 59.2, pp. 170–181.

Sein, D. V., U. Mikolajewicz, M. Gröger, I. Fast, W. Cabos, J. G. Pinto, S. Hagemann, T. Semmler, A. Izquierdo, & D. Jacob (2015). "Regionally coupled atmosphere-ocean-sea ice-marine biogeochemistry model ROM: 1. Description and validation." In: *Journal of Advances in Modeling Earth Systems* 7.1, pp. 268–304.

Sidorenko, D., T. Rackow, T. Jung, T. Semmler, D. Barbi, S. Danilov, K. Dethloff, W. Dorn, K. Fieg, H. F. Gößling, et al. (2015). "Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate." In: *Climate Dynamics* 44.3-4, pp. 757–780.

Slangen, A., M. Carson, C. Katsman, R. Van de Wal, A. Köhl, L. Vermeersen, & D. Stammer (2014). "Projecting twenty-first century regional sea-level changes." In: *Climatic Change* 124.1-2, pp. 317–332.

Slangen, A., R. S. Van de Wal, T. J. Reerink, R. C. De Winter, J. R. Hunter, P. L. Woodworth, & T. Edwards (2017). "The impact of uncertainties in ice sheet dynamics on sea-level allowances at tide gauge locations." In: *Journal of Marine Science and Engineering* 5.2, p. 21.

Small, C. & R. J. Nicholls (2003). "A global analysis of human settlement in coastal zones." In: *Journal of coastal research*, pp. 584–599.

Steffen, H. & P. Wu (2011). "Glacial isostatic adjustment in Fennoscandia—a review of data and modeling." In: *Journal of Geodynamics* 52.3-4, pp. 169–204.

Sterl, A., H. v. d. Brink, H. d. Vries, R. Haarsma, & E. v. Meijgaard (2009). "An ensemble study of extreme storm surge related water levels in the North Sea in a changing climate." In: *Ocean Science* 5.3, pp. 369–378.

Sturges, W. & B. C. Douglas (2011). "Wind effects on estimates of sea level rise." In: *Journal of Geophysical Research: Oceans* 116.C6.

Swingedouw, D., L Terray, C Cassou, A Voldoire, D Salas-Mélia, & J Servonnat (2011). "Natural forcing of climate during the last millennium: fingerprint of solar variability." In: *Climate Dynamics* 36.7-8, pp. 1349–1364.

Talke, S. A. & D. A. Jay (2020). "Changing tides: the role of natural and anthropogenic factors." In: *Annual review of marine science* 12, pp. 121–151.

Taylor, K. E., R. J. Stouffer, & G. A. Meehl (2012). "An overview of CMIP5 and the experiment design." In: *Bulletin of the American Meteorological Society* 93.4, pp. 485–498.

Thomas, M., J. Sündermann, & E. Maier-Reimer (2001). "Consideration of ocean tides in an OGCM and impacts on subseasonal to decadal polar motion excitation." In: *Geophysical Research Letters* 28.12, pp. 2457–2460.

Tsimplis, M. & P. Woodworth (1994). "The global distribution of the seasonal sea level cycle calculated from coastal tide gauge data." In: *Journal of Geophysical Research: Oceans* 99.C8, pp. 16031–16039.

Turki, I., N. Massei, & B. Laignel (2019). "Linking sea level dynamic and exceptional events to large-scale atmospheric circulation variability: A case of the Seine Bay, France." In: *Oceanologia* 61.3, pp. 321–330.

Valcke, S (2013). "The OASIS3 coupler: a European climate modelling community software." In: *Geoscientific Model Development* 6.2, pp. 373–388.

Vousdoukas, M. I., E. Voukouvalas, A. Annunziato, A. Giardino, & L. Feyen (2016). "Projections of extreme storm surge levels along Europe." In: *Climate Dynamics* 47.9-10, pp. 3171–3190.

Vousdoukas, M. I., L. Mentaschi, E. Voukouvalas, M. Verlaan, & L. Feyen (2017). "Extreme sea levels on the rise along Europe's coasts." In: *Earth's Future* 5.3, pp. 304–323.

WASA-group (1998). "Changing waves and storms in the Northeast Atlantic?" In: *Bulletin of the American Meteorological Society* 79.5, pp. 741–760.

Wahl, T. & D. P. Chambers (2015). "Evidence for multidecadal variability in US extreme sea level records." In: *Journal of Geophysical Research: Oceans* 120.3, pp. 1527–1544.

— (2016). "Climate controls multidecadal variability in US extreme sea level records." In: *Journal of Geophysical Research: Oceans* 121.2, pp. 1274–1290.

Wahl, T., J. Jensen, T. Frank, & I. D. Haigh (2011). "Improved estimates of mean sea level changes in the German Bight over the last 166 years." In: *Ocean Dynamics* 61.5, pp. 701–715.

Wahl, T., I. D. Haigh, R. J. Nicholls, A. Arns, S. Dangendorf, J. Hinkel, & A. B. Slangen (2017). "Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis." In: *Nature communications* 8, p. 16075.

Wakelin, S., P. Woodworth, R. Flather, & J. Williams (2003). "Sea-level dependence on the NAO over the NW European Continental Shelf." In: *Geophysical Research Letters* 30.7.

Wang, Y.-M., J. Lean, & N. Sheeley Jr (2005). "Modeling the Sun's magnetic field and irradiance since 1713." In: *The Astrophysical Journal* 625.1, p. 522.

Weare, B. C. & J. S. Nasstrom (1982). "Examples of extended empirical orthogonal function analyses." In: *Monthly Weather Review* 110.6, pp. 481–485.

Weisheimer, A., D. Decremer, D. MacLeod, C. O'Reilly, T. N. Stockdale, S. Johnson, & T. N. Palmer (2019). "How confident are predictability estimates of the winter North Atlantic Oscillation?" In: *Quarterly Journal of the Royal Meteorological Society* 145, pp. 140–159.

Weisse, R. & A. Plüß (2006). "Storm-related sea level variations along the North Sea coast as simulated by a high-resolution model 1958–2002." In: *Ocean Dynamics* 56.1, pp. 16–25.

Weisse, R., H. von Storch, U. Callies, A. Chrastansky, F. Feser, I. Grabemann, H. Guenther, A. Plüß, T. Stoye, J. Tellkamp, et al. (2009). "Regional meteorological–marine Reanalyses and climate change projections: Results for Northern Europe and potential for coastal and offshore applications." In: *Bulletin of the American Meteorological Society* 90.6, pp. 849–860.

Weisse, R., H. von Storch, H. D. Niemeyer, & H. Knaack (2012). "Changing North Sea storm surge climate: An increasing hazard?" In: *Ocean & Coastal Management* 68, pp. 58–68.

Weisse, R., D. Bellafiore, M. Menéndez, F. Méndez, R. J. Nicholls, G. Umgiesser, & P. Willems (2014). "Changing extreme sea levels along European coasts." In: *Coastal engineering* 87, pp. 4–14.

Woodworth, P., R. Flather, J. Williams, S. Wakelin, & S Jevrejeva (2007). "The dependence of UK extreme sea levels and storm surges on the North Atlantic Oscillation." In: *Continental Shelf Research* 27.7, pp. 935–946.

Woodworth, P. L. & D. L. Blackman (2004). "Evidence for systematic changes in extreme high waters since the mid-1970s." In: *Journal of Climate* 17.6, pp. 1190–1197.

Woodworth, P. L., M. Menéndez, & W. R. Gehrels (2011). "Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels." In: *Surveys in geophysics* 32.4-5, pp. 603–618.

Woth, K. (2005). "Regionalization of global climate change scenarios: An ensemble study of possible changes in the North Sea storm surge statistics." PhD thesis. Universität Hamburg Hamburg.

Woth, K., R. Weisse, & H. von Storch (2006). "Climate change and North Sea storm surge extremes: an ensemble study of storm surge extremes expected in a changed climate projected by four different regional climate models." In: *Ocean Dynamics* 56.1, pp. 3–15.

Yin, J. H. (2005). "A consistent poleward shift of the storm tracks in simulations of 21st century climate." In: *Geophysical Research Letters* 32.18.

Zanchettin, D., C. Timmreck, O. Bothe, S. J. Lorenz, G. Hegerl, H.-F. Graf, J. Luterbacher, & J. H. Jungclaus (2013). "Delayed winter warming: A robust decadal response to strong tropical volcanic eruptions?" In: *Geophysical Research Letters* 40.1, pp. 204–209.

de Vries, H., M. Breton, T. de Mulder, Y. Krestenitis, R. Proctor, K. Ruddick, J. C. Salomon, A. Voorrips, et al. (1995). "A comparison of 2D storm surge models applied to three shallow European seas." In: *Environmental Software* 10.1, pp. 23–42.

von Storch, H. & H. Reichardt (1997). "A scenario of storm surge statistics for the German Bight at the expected time of doubled atmospheric carbon dioxide concentration." In: *Journal of Climate* 10.10, pp. 2653–2662.

von Storch, H. & K. Woth (2008). "Storm surges: perspectives and options." In: *Sustainability Science* 3.1, pp. 33–43.

ACKNOWLEDGEMENTS

Writing this dissertation was only possible thanks to the support of the many people who helped me along this journey.

First and foremost, I thank Uwe Mikolajewicz for supervising my work over the last three years, for his guidance and for the trust in giving me freedom in my research. I thank Gerhard Schmiedl and Hartmut Grassl for the fruitful discussions, useful input and friendly atmosphere and support within our panel meetings, and Dorothea Bunzel for the good collaboration within our very interdisciplinary project and for the support at our various project meetings. Thanks also to the IMPRS office, namely Antje, Connie and Michaela, for making my life as a doctoral student so easy.

Many thanks to my ocean physics group colleagues for the good working atmosphere. A special thank here goes to Moritz Mathis, for helping me getting started at the MPI, introducing me to the model code, assisting me in setting up the simulations, and not least for his valuable input for my research at various stages of my PhD. Here, a big thank you also goes to Veit Lüshow, Sebastian Milinski and Johann Jungclaus for their valuable feedback on parts of this thesis or the therein included articles.

In addition, I want to thank my fellow PhD colleagues who became friends, especially Anke and Lydi who shared my path from the very first same day we started here at MPI. And of course a big shout-out to the entire MPI-Football team for the great team spirit and atmosphere on and off the pitch during those three seasons among Hamburg's football elite.

Finally, I want to thank my family, for your constant support and encouragement. And of course a big thank you to you, Fine – not only for proof-reading and improving my manuscripts, but for simply being there and always believing in me.

Thank you!

EIDESSTATTLICHE VERSICHERUNG

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

I hereby declare an oath that I have written the present dissertation on my own and have not used other than the acknowledged resources and aids. All passages taken literally or analogously from other publications are identified as such. I further declare that this thesis has not been submitted to any other German or foreign examination board and that the submitted written version corresponds to that on the electronic repository.

Hamburg, 01. April 2020

Andreas Lang

Hinweis / Reference

Die gesamten Veröffentlichungen in der Publikationsreihe des MPI-M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199

sind über die Internetseiten des Max-Planck-Instituts für Meteorologie erhältlich:
http://www.mpimet.mpg.de/wissenschaft/publikationen.html

All the publications in the series of the MPI -M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199

are available on the website of the Max Planck Institute for Meteorology:
http://www.mpimet.mpg.de/wissenschaft/publikationen.html