UNIVERSITÄT SIEGEN

Mitteilungen des Forschungsinstituts Wasser und Umwelt der Universität Siegen

Arne Arns

Regional to local assessment of extreme water levels

Methods and application to the northern part of the German North Sea coastline

Heft 7

2014

ISSN 1868-6613

Arne Arns

Regional to local assessment of extreme water levels

Methods and application to the northern part of the German North Sea coastline

Erscheinungsort: Siegen
Erscheinungsjahr: 2014
D 476

Mitteilungen des Forschungsinstituts Wasser und Umwelt der Universität Siegen
Heft 7 | 2014

Herausgeber:
Forschungsinstitut Wasser und Umwelt (fwu)
der Universität Siegen
Paul-Bonatz-Str. 9-11
57076 Siegen

Druck:
UniPrint, Universität Siegen

ISSN 1868-6613

Vorwort

Mit der ersten Promotion am Forschungsinstitut Wasser und Umwelt (fwu), die nach der Integration des Fachbereichs Bauingenieurwesen am fwu durchgeführt werden konnte, wurde eine eigene fwu-Schriftenreihe etabliert. Neben den Promotionen am fwu sollen in dieser Schriftenreihe die Ergebnisse von Institutsveranstaltungen, Konferenzen und Workshops sowie andere Forschungsergebnisse, die im Kontext des fwu erarbeitet werden, veröffentlicht werden. Bis dahin wurden die Forschungsergebnisse in verschiedenen internen und externen Schriftenreihen publiziert.

Eine Übersicht der bisher veröffentlichten Schriftenreihen kann der letzten Seite entnommen werden. In dem vorliegenden Heft 7 (2014) wird die Promotion von Arne Arns mit dem Titel „Regional to local assessment of extreme water levels - Methods and application to the northern part of the German North Sea coastline" in Papierform veröffentlicht; die digitale Veröffentlichung erfolgte am 14.08.2014 über die Universitätsbibliothek Siegen.

Inhaltlich befasst sich die Arbeit von Herrn Arns mit dem Auftreten extremer Sturmfluten. Diese natürlichen Ereignisse gehören zu den größten geophysikalischen Bedrohungen in Küstengebieten und haben in der Vergangenheit an der Deutschen Nordseeküste zu großen Schäden geführt. Der Abschätzung zukünftiger extremer Sturmfluten sowie der Bemessung von Küstenschutzbauwerken kommt daher eine besondere Bedeutung zu. Im Rahmen der Dissertation von Herrn Arns wird die Verwendung extemwertstatistischer Verfahren zur Ermittlung der Höhen und Häufigkeiten von Sturmflutwasserständen untersucht. Eine große Herausforderung im Küsteningenieurwesen ist die Entwicklung einer Methodik zur Regionalisierung extremer Wasserstände in unbepegelten Küstengebieten. In Verbindung mit von Herrn Arns entwickelten Empfehlungen wird eine Methodik zur Ermittlung der Höhen und Häufigkeiten von extremen Sturmfluten bzw. Wasserständen entlang der gesamten Küstenlinie vorgestellt. Ein Schwerpunkt der Forschungen war dabei die Sicherung der Halligen im nordfriesischen Wattenmeer.

Die von Herrn Arns durchgeführten Arbeiten wurden im Rahmen des vom Kuratorium für Forschung im Küsteningenieurwesen (KFKI) begleiteten Forschungsvorhabens „Entwicklung von nachhaltigen Küstenschutz- und Bewirtschaftungsstrategien für die Halligen unter Berücksichtigung des Klimawandels" (ZukunftHallig) durchgeführt. Das Projekt hatte eine Laufzeit von 3 Jahren und wurde vom Bundesministerium für Bildung und Forschung (BMBF) unter der Leitung des Projektträgers Jülich (PTJ) gefördert (Fördernummer: 03KIS093). Fachlich wurde das Vorhaben von der projektbegleitenden Gruppe des KFKI begleitet. Wir möchten uns hierfür herzlich bedanken.

Abschließend möchte ich mich für die Mitbetreuung der Promotion bei meinen Kollegen Prof. Dr.-Ing. Holger Schüttrumpf von der Rheinisch-Westfälischen Technischen Hochschule Aachen (RWTH Aachen), bei Herrn Prof. Dr. rer.-nat. Athanasios Vafeidis von der Christian-

Albrechts-Universität zu Kiel (CAU) und bei Herrn Prof. Dr.-Ing. Ulf Zander von der Universität Siegen herzlich bedanken.

Siegen im August 2014

Univ.-Prof. Dr.-Ing. Jürgen Jensen

Regional to local assessment of extreme water levels

Methods and application to the northern part of the German North Sea coastline

Vom Department für Bauingenieurwesen der Naturwissenschaftlich-Technischen Fakultät der Universität Siegen angenommene

Dissertation

zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)
von

Dipl.-Ing. Arne Arns

Referent:	**Univ.-Prof. Dr.-Ing. Jürgen Jensen** Universtät Siegen
Korreferent:	**Univ.-Prof. Dr.-Ing. Holger Schüttrumpf** Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen

Tag der Einreichung: 31.03.2014

Tag der mündlichen Prüfung: 10.07.2014

About the cover picture:

The cover picture shows the landing pier at Hallig Langeness. In general, Halligen are small low lying islands located off the coastline of Schleswig-Holstein. They have no dikes and as they are frequently exposed to extreme water levels, they are inundated up to 50 times a year. Nevertheless, they are inhabited by around 270 residents. In order to protect the inhabitants from regular inundation, houses have been built on artificial dwelling mounds (i.e. the mounds on the cover picture). Usually, such protection measures are constructed towards certain design levels which are calculated using observed water levels as input. In the Halligen area, there are no tide gauges available that provide sufficient information to reliably conduct traditional extreme value analyses. This thesis shows a methodology to provide return level estimates for such areas.

Contact: Arne Arns | e-Mail: arne.arns@gmail.com

Danksagung

Diese Dissertation entstand zwischen 2009 und 2014 während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Forschungsinstitut Wasser und Umwelt (fwu) der Universität Siegen. Im Folgenden möchte ich einer Fülle von Personen danken, die mich bei der Arbeit an dieser Dissertation unterstützt haben, sei es durch Anregungen, Ratschläge und Kritik, oder aber auch durch die notwendige Ablenkung von der Arbeit.

Mein größter Dank geht an meinen Doktorvater, Herrn Prof Dr.-Ing. Jürgen Jensen, der mir diese Arbeit überhaupt erst ermöglicht hat. Durch die richtige Balance aus Interesse, Anregungen, Kritik und gewahrter Zurückhaltung hinsichtlich Vorgaben hat er mir ein perfektes Umfeld für die Dissertation geschaffen. Vielen Dank hierfür! Danken möchte ich außerdem Herrn Prof. Dr.-Ing. Holger Schüttrumpf für die Übernahme des Korreferats sowie für hilfreiche Diskussionen zu unterschiedlichsten Fragestellungen, die im Rahmen unseres gemeinsamen Forschungsprojektes aufkamen. Dank gebührt darüber hinaus Herrn Prof. Dr.-Ing. Ulf Zander für den Vorsitz der Promotionskommission sowie Prof. Dr. rer. nat. Athanasios Vafeidis, der sich bereit erklärt hat, die Prüfungskommision zu vervollständigen.

Besonderer Dank gilt meinen Kollegen und Freunden vom fwu, die mir jederzeit mit Rat und Tat zur Seite standen. Im speziellen danken möchte ich: Jessica Schmidt, Sandra Sziburies, Vitalij Kelln, Sebastian Niehüser, Ugur Öztürk, Volker Spieß, Andre Stettner-Davis, Jörg Wieland. Ein besonderer Dank gilt Christoph Mudersbach, Jens Bender und Sönke Dangendorf, die mit hilfreichen Kommentaren und Korrekturen einen wertvollen Beitrag zu dieser Arbeit geleistet haben sowie Thomas Wahl, der diese Arbeit maßgeblich betreut hat. Eine besseres Arbeitsumfeld kann sich ein Doktorand kaum wünschen.

Große Teile dieser Dissertation sind im Rahmen des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Forschungsvorhabens ZukunftHallig entstanden, welches zwischen 2010 und 2014 unter der Leitung vom Projektträger Jülich (PTJ) von mir betreut wurde. Insbesondere die regelmäßigen Treffen und fachlichen Diskussionen mit unseren Projektpartnern aus Aachen, Göttingen und Husum sowie die Treffen mit der projektbegleitenden Gruppe des Kuratoriums für Forschung im Küsteningenieurwesen (KFKI) haben einen wertvollen Beitrag zur Lösung diverser Fragestellungen geliefert. Im Rahmen eines vom Deutschen Akademischen Austausch

Dienst (DAAD) geförderten Projektes im Programm „Group of Eight Australia-Germany Joint Research Cooperation Scheme“ konnten überdies wichtige Auslandserfahrungen gesammelt und nachhaltige Kontakte geknüpft werden. Part I dieser Dissertation ist maßgeblich im Rahmen dieser Kooperation entstanden. Für die Unterstützung möchte ich Prof. Charitha Pattiaratchi und Ivan Haigh danken.

Einige der im Folgenden dargestellten Ergebnisse entstanden unter Verwendung des Softwarepaketes MIKE21, welches mir im Rahmen der Dissertation von DHI-WASY zur Verfügung gestellt wurde. Vielen Dank dafür!

Ein besonderer Dank gilt meinen Eltern, meinen Schwiegereltern sowie allen nicht genannten Personen meiner Familie, deren moralische Unterstützung und Geduld während meiner gesamten Ausbildung von immenser Bedeutung waren.

Schlussendlich danken möchte ich meiner Frau Julia. Danke dafür, dass du mir immer wieder vor Augen führst, dass das Leben aus mehr als Verpflichtungen besteht und für all das, was man mit Worten nicht ausdrücken kann!

Siegen, im April 2014

Arne Arns

Abstract

This thesis investigates the use of extreme value statistics to estimate both the heights (i.e. return levels) and occurrence probabilities (i.e. return periods) of extreme water levels, which can cause considerable loss of life and millions of dollars of damage (Cunnane, 1987). Over the past five decades, several approaches for estimating extreme water levels have been developed. Currently, different methods are applied not only on transnational, but also on national scales, resulting in a heterogeneous level of protection. Applying different statistical methods can yield significantly different estimates of return water levels, but even the use of the same technique can produce large discrepancies, because there is subjective parameter choice at several steps in the model setup.

In this thesis, the main direct methods (i.e. the block maxima method and the peaks over threshold method) to estimate return levels and periods are compared, considering a wide range of strategies to create the extreme value datasets and a range of different model setups. The focus is on testing the influence of the main factors, which can significantly affect the estimates of extreme value statistics. Finally, to provide guidance for coastal engineers and operators, an objective approach for setting up the model is recommended. If this is applied routinely around a country, it will help overcome the problem of heterogeneous levels of protection resulting from different methods and varying model setups.

However, these recommendations can often not be considered for practical applications as the availability of water level data is a limitation in many regions. For example, for the North Frisian part of the German North Sea there are only a few water level records available and these are currently too short to apply traditional extreme value analysis methods. As tidal characteristics in the German Bight are highly influenced by shallow water effects and the shape of the coastline, they can differ significantly between stations (see e.g. Jensen and Müller-Navarra, 2008). It is thus difficult to directly convey information about the likelihood of extreme hydrologic events from gauged to surrounding un-gauged sites. To transfer water level information measured at gauged sites to un-gauged sites in the study region, the regional frequency analysis (RFA) concept (which has been previously applied to a riverine setting) is adopted and adjusted for application to a coastal setting. The proposed method is based on a numerical multi-decadal model hindcast of water levels for the whole of the North Sea. Predicted water levels from the hindcast are

bias-corrected using the information from the available tide gauge records. Hence, the simulated water levels agree well with the measured water levels at gauged sites. Combining the bias-corrected water levels and the recommendations that were made in the first part of this thesis provides a procedure to estimate return water levels suitable for coastal defence design conditions. The return levels are estimated continuously along the entire coastline of the study area, including the offshore islands. A similar methodology to that applied here could be used in other regions of the world.

One of the most discussed aspects in coastal engineering at the moment is concerned with the possible impact of sea level rise (SLR) and the associated changes in extreme water levels on coastal defense structures. The methodologies presented above can be used to calculate present day design levels for coastal defenses but do not account for SLR and potential nonlinear changes in the tidal characteristics, which in turn may affect the results from extreme value statistics. This is why the impact of SLR on extreme water levels is investigated using a numerical model that covers the entire North Sea and has its highest spatial resolution in the northern part of the German Bight. At most locations, the model run highlights that storm surge and return water levels are significantly different from the changes in MSL alone, a finding somewhat different from former studies in that area having major implications for the design of coastal defenses.

Furthermore, the analyses indicate that these increases in storm surge water levels are mainly caused by nonlinear changes in the tidal components which are spatially not coherent. The response of the tidal propagation to SLR is investigated based on the results from a tidal analysis of each individual event. These analyses point to changes in individual constituents, such as increases in the M_2 amplitude and decreases in the amplitudes of frictional and overtides accompanied by less tidal wave energy dissipation. Attributed effects are changes in phase lags of individual constituents leading to a different tidal modulation, thus additionally increasing tidal water levels.

Kurzfassung

In dieser Dissertation wird die Verwendung extemwertstatistischer Verfahren zur Abschätzung der Höhen und Häufigkeiten von Sturmflutwasserständen untersucht. In den vergangenen Dekaden wurden hierzu verschiedene Ansätze entwickelt. Bisweilen konnte sich auf nationaler wie auch auf internationaler Ebene jedoch kein allgemein gültiges Verfahren etablieren, weshalb die aktuell existierenden Schutzstandards nicht vergleichbar sind. Denn sowohl die Verwendung unterschiedlicher Modelle, als auch die Verwendung unterschiedlicher Einstellungen bei ein und demselben Modell kann zu großen Differenzen in den Ergebnissen extremwertstatistischer Auswertungen führen.

Im Rahmen der Dissertation werden die beiden primär verwendeten direkten Verfahren (d.h. die Block Maxima und die Peak Over Threshold Methode) zur Ermittlung der Höhen und Häufigkeiten von Sturmfluten unter Verwendung eines weiten Spektrums an Vorgehensweisen miteinander verglichen. Der Fokus liegt dabei auf der Ermittlung der Sensitivität der verwendeten Modelle gegenüber den bisweilen subjektiv zu wählenden Modelleinstellungen. Ausgehend von diesen Analysen werden Empfehlungen zur objektiven und vergleichbaren Verwendung extremwertstatistischer Modelle im Küsteningenieurwesen entwickelt. Werden diese Empfehlungen konsistent verwendet (auf nationaler sowie auf internationaler Ebene), kann hierdurch die Vergleichbarkeit der Schutzstandards an individuellen Küstenstandorten deutlich erhöht werden.

Für die Verwendung der Empfehlungen werden Wasserstandsinformationen benötigt, die eine ausreichend lange Periode abdecken. In vielen Gebieten sind diese Informationen jedoch limitiert. So existieren in großen Teilen der nordfriesischen Nordseeküste (einschließlich der Inseln und Halligen) insgesamt nur wenige Pegelstationen, deren Aufzeichnungen gegenwärtig nur wenige Jahre abdecken. Im Hinblick auf extremwertstatistische Analysen sind diese Informationen i.d.R. nicht ausreichend. Da die Wasserstände in der Deutschen Bucht durch nichtlineare Effekte (z.B. Flachwassereffekt) beeinflusst werden, weisen selbst nahegelegene Aufzeichnungen oft stark unterschiedliche Charakteristika auf (siehe z.B. Jensen and Müller-Navarra, 2008). Aus diesem Grund ist es nur bedingt möglich, die Höhen und Häufigkeiten an unbepegelten Standorten direkt aus den umliegenden bepegelten Standorten abzuleiten. In der Dissertation wird daher eine Methodik zur Ermittlung extremer Wasserstände in unbepegelten Küstengebieten entwickelt. Die Vorgehensweise orientiert sich zunächst am

Konzept der regionalen Frequenzanalyse (RFA), welche zuvor bereits im Bereich binnenhydrologischer Fragestellungen verwendet wurde. Aufbauend darauf wird eine neuartige Methodik entwickelt, welche auf numerisch simulierten Wasserständen der gesamten deutschen Nordseeküste basiert. Die simulierten Wasserstände werden mit Hilfe der Beobachtungsdaten korrigiert, so dass die simulierten und die beobachteten Wasserstände an den Pegelstationen vollständig übereinstimmen. In Verbindung mit den oben genannten Empfehlungen werden diese Wasserstandsinformationen zur Ermittlung der Höhen und Häufigkeiten von extremen Wasserständen entlang der gesamten Küstenlinie des Untersuchungsbereiches verwendet.

Mit Hilfe der zuvor genannten Methoden lassen sich Aussagen zur Sturmflutgefährdung unter gegenwärtigen Bedingungen treffen. Potentielle Änderungen in den Randbedingungen, wie etwa ein Anstieg des mittleren Meeresspiegels (MSL), werden dabei vernachlässigt. Jedoch können durch solche Änderungen Effekte induziert werden, die zu nichtlinearen Änderungen in den höheren Wasserständen führen. Prognosen zur zukünftigen Entwicklung von Sturmflutwasserständen unterliegen somit gewissen Unsicherheiten. Aus diesem Grund wird der Einfluss des Anstieges im MSL auf Extremwasserstände an einem numerischen Modell untersucht. Das Modell umfasst die gesamte Nordsee sowie Teile des Nordatlantiks, weist jedoch im Bereich der Deutschen Bucht die höchste Auflösung auf. Die Untersuchungen zeigen für die meisten Standorte, dass die Änderungen in den extremen Wasserständen in weiten Teilen des Untersuchungsgebietes signifikant höher sind als der Anstieg des MSL. Hierbei zeigt sich räumlich jedoch kein einheitliches Bild. Darüberhinaus zeigen die Untersuchungen, dass die erhöhten Sturmflutwasserstände maßgeblich in der astronomisch induzierten Komponente (d.h. der Reaktion des Wasserkörpers auf die Gezeitenkräfte) begründet sind. So konnte z.B. eine Erhöhung der Amplitude der dominanten M_2 Tide beobachtet werden, während in den Obertiden sowie den aus Reibung induzierten Tiden ein Amplitudenrückgang beobachtet wurde. Insbesondere für die Bemessung von Küstenschutzanlagen sind diese Ergebnisse von großer Bedeutung.

Content

List of Figures

List of Tables

List of abbreviations and symbols

Abbreviation	Full name
ACD	Admiralty chart datum
AD	Anno Domini
AMAX	Annual maxima
AR4	4th Assessment Report of the IPCC
AR5	5th Assessment Report of the IPCC
AU	Australia
Bor	Borkum station
BM	Block maxima
BP	Before present
Bus	Büsum station
CDF	Cumulative distribution functions
CIs	Confidence intervals
Cux	Cuxhaven station
Dag	Dagebüll station
Emd	Emden station
ENS	Entire North Sea configuration
EVA	Extreme value analysis
FRA	France
GB	Great Britain
GBi	German Bight configuration
GER	Germany
GEV	Generalized extreme value distribution
GIA	Glacial isostatic adjustment
GPD	Generalized Pareto distribution
Hel	Helgoland station
Hoe	Hörnum station
HR	High resolution values
Hus	Husum station
HW	Total observed high water level peaks
IDW	Inverse distance weighting
IID	Independent and identically distributed
IPCC	Intergovernmental Panel on Climate Change
IRPS	Index of return period stability
Lis	List station
LtA	Leuchtturm Alte Weser station
LW	total observed low water level peaks
MHW	Mean high water
MLE	Maximum likelihood estimation
MRD	Moment ratio diagram
MRLP	Mean residual life plot
MSE	Mean square error
MSL	Mean sea level
NHN	German reference datum '*Normalhöhennull'*
NL	Netherlands
Nor	Norderney station
NOR	Norway

NST	North Sea tracking model
PLP	Plotting position
POT	Peak over threshold
PV	Peak values
RCP	Representative concentration pathways
RFA	Regional frequency analysis
RMSE	Root mean squared error
Sch	Schlüttsiel station
SLR	Sea level rise
SSL	Standard storm surge length
SSP	Stability plot
TGZ	Tide gauge zero
TOR	Tide only run
TSR	Tide surge run
UK	United Kingdom
Wil	Wilhelmshaven station
Wit	Wittdün station
Wyk	Wyk station

Symbol	**Dimension**	**Description**
$2MK_5$	[°/h]	Fifth diurnal constituent (frequency = 0.20280)
$3MK_2$	[°/h]	Seventh diurnal constituent (frequency = 0.28331)
A	[m], [cm]	Amplitude
a	[-]	Plotting position parameter
C_D	[-]	Drag coefficient
c	[m/s]	Wave speed
c_h	[-]	Homogeneous combinations
c_p	[-]	Number of possible combination
D_i	[-]	Discordancy measure
d	[-]	Index of agreement
f	[N/m²]	Forces per mass unit
f_b	[N/m²]	bottom stress
f_s	[N/m²]	surface stress
g	[m/s²]	The earth's gravitational attraction
g_n	[°/h]	Phase lag
H	[-]	Heterogeneity measure
h	[m], [cm]	Water depth
K_1	[°/h]	Luni-solar declinational diurnal constituent
K_2	[°/h]	Luni-solar declinational semidiurnal constituent
k	[-]	Dimensionless factor, friction parameter
L	[m], [km]	Wavelength
M_2	[°/h]	Principal lunar semidiurnal constituent (frequency = 0.08051)
M_3	[°/h]	Lunar terdiurnal constituent (frequency = 0.12076)
M_4	[°/h]	First overtide of M_2 constituent (frequency = 0. 16102)
M_6	[°/h]	Second overtide of M_2 constituent (frequency = 0.24153)
M_8	[°/h]	Third overtide of M_2 constituent (frequency = 0.32204)
m	[-]	Number of stations to be drawn

μ	[-]	Location parameter
μ_i	[-]	Index flood
N	[-]	Number of sites
N_2	[°/h]	Larger Lunar elliptic semidiurnal constituent (frequency = 0.07899)
n	[-]	The number of elements/observations
n_m	[$m^{1/3}$/s]	Manning's n
O_1	[°/h]	Lunar declinational diurnal constituent (frequency = 0.03873)
P_1	[°/h]	Solar diurnal constituent (frequency = 0.04155)
P_E	1/a	Exceedance probability
P_U	1/a	Non-exceedance probability
Q_1	[°/h]	Larger lunar elliptic diurnal constituent (frequency = 0.03721)
Q(F)	[-]	Distribution
q(F)	[-]	Regional distribution
q_i	[1/a]	Probability plotting position
r	[val./yr]	r-largest values of each year
r^2	[-]	Coefficient of determination
ρ	[kg/m³]	Density
S	[m], [cm]	standard deviation
S_2	[°/h]	Principal solar semi diurnal constituent (frequency = 0.08333)
σ	[cm]	Scale parameter
σ_n	[°/h]	Angular velocity
$\tilde{\sigma}$	[cm]	Transformed scale parameter
T	[yr]	Return period
T_i	[d], [h]	difference times of the random variable of the extremal index
t_d	[d]	Declustering time
θ	[-]	Extremal index
τ_b	[N/m²]	Bottom friction
τ_s	[N/m²]	Wind stress
U	[m/s]	Wind in zonal direction
u	[m/s]	Total current
u_0	[m], [cm]	Threshold
V	[m/s]	Wind in meridional direction
x	[m], [cm]	Sample values
$\bar{x}$	[m], [cm]	Mean value
x_m	[m], [cm]	Modeled water levels
x_o	[m], [cm]	Observed water levels
xmax	[m], [cm]	The largest of all values in a sample
W	[m/s]	Wind speed
WL	[cm],[m]	Water level
z	[m], [cm]	Block maxima values
ξ	[-]	Shape parameter

This work considers SI units.

1 Introduction

Storm surges are among the most hazardous geophysical risks in coastal regions and are often associated with significant losses of life and property (von Storch, 2012). The North Sea, and the German coastline in particular, has a long history of severe storm surges. For example, a large storm occurred in 1164 when thousands of people lost their lives. This storm surge caused the first great damaging flood after the construction of dikes along the German coast (Petersen and Rohde, 1977). In January 1362, probably the greatest North Sea flood disaster in historical times occurred, where more than half the population of the marshland districts along today's federal state of Schleswig-Holstein drowned (Lamb, 1991). The number of reported fatalities ranges from 11,000 (e.g. Gram-Jensen, 1985) to 100,000 (Lamb, 1991). Another large storm occurred in November 1570, and it has been suggested that between 100,000 and 400,000 people were drowned in countries bordering the North Sea (Lamb, 1991). However, the death tolls given above are all based on chronicles and are thus highly uncertain. More recently, developments in coastal flood risk management in northern Europe accelerated following the 1953 floods which killed more than 2,000 people around the coastline of the southern North Sea (Gerritsen, 2005; Baxter, 2005; McRobie et al., 2005) and floods in the German Bight in 1962 when more than 300 people lost their lives (Bütow, 1963; von Storch and Woth, 2006).

Rising mean sea levels (MSL) will additionally increase the likelihood of coastal flooding around the world (Seneviratne et al., 2012), adversely impacting rapidly growing coastal communities. For instance, in 2005, 136 port cities had populations exceeding one million and thirteen of the twenty mega cities (populations > 8 million) in the world were port cities (Nicholls et al., 2008). Globally, it is estimated that more than 200 million people are already vulnerable to coastal flooding in these cities and other coastal settlements (Nicholls, 2011). According to recent projections, global MSL might rise by up to +2.0 m in the 21st century alone (see Nicholls et al., 2011 for an overview) marked by a

considerable regional variability (see e.g. Dangendorf et al., 2012; Wahl et al., 2013). As MSL rises, the risk of beach erosion and salt water intrusion into groundwater systems increases. It also directly affects extreme water level events by shifting the frequency distributions of storm surges to higher base levels (i.e. events of a given height occur more frequently) (Hunter, 2010). The Intergovernmental Panel on Climate Change's (IPCC) 4th Assessment Report (AR4) highlighted that *"societal impacts of sea level change primarily occur via the extreme levels rather than as a direct consequence of mean sea level change"* (Bindoff et al., 2007). In some coastal regions extreme water levels could further be amplified by changes in storminess, although there are still significant uncertainties regarding possible future changes in storm activity (Meehl et al., 2007). It is thus essential that the flood risk is accurately evaluated and defenses are upgraded where necessary (Coles and Tawn, 2005; Haigh et al., 2010a). This in turn requires a profound description of the complex natural processes which usually exhibit both, a spatially varying and random behavior.

Simplified, natural processes can be described with analytical approaches and models, helping towards a physical understanding of the underlying system. A model can be regarded as a simplified description of complex natural processes considering only those characteristics, which are important for the intended application (Haußer and Luchko, 2011). In hydrosciences, the utilization of models is multifarious covering e.g. analytical, numerical or statistical models. Such models are either used to obtain a specific answer to a specific problem (predictive) or to improve the understanding of natural processes (investigative). According to Blöschel and Sivapalan (1995), the development of both types of models traditionally follows a range of steps. These are:

- *The data collection*
- *The development of conceptual models describing the important characteristics*
- *The translation of the conceptual models into mathematical models*
- *The calibration of the mathematical models*
- *The validation of the models*

If the models are successfully validated, they can be used for several applications. If, however, the validation is not satisfying, one of the previous steps has to be repeated (Gutknecht, 1991).

Models in hydrosciences can consider different scales, covering e.g. individual molecules of water, the atmospheric circulation as well as the Earth's water balance (Dyck and Peschke, 1995). The term scale usually describes the characteristics with respect to both time and/or length (Blöschel and Sivapalan, 1995). A formal definition of the term *scale* is provided in Blöschel and Sivapalan (1995) who differentiate between the '*process scale*', the '*observation scale*' and the '*modeling scale*'. The '*process scale*' is defined as the scale that natural phenomena exhibit in space and time covering e.g. the duration, period (cycle) or spatial extent of a process. In the time domain, the '*process scale*' often exhibits preferred time-scales of a day or a year having a spectral gap in between. In the space domain there is no such clear evidence for preferred scales (Gupta and Waymire, 1993). The '*observation scale*', by contrast, is defined by the limitations of measurement techniques including the spatial and temporal extent of a data set or the spacing between stations. Ideally, a process under investigation should be observed on the scale it actually occurs. In reality, however, this is rarely the case as most hydrological processes occur on large scales but only samples recorded at one or a few stations are available.

The '*modeling scale*' is partly related to natural (hydrological) processes but also to the intended application of the model (Blöschel and Sivapalan, 1995). Fig. 1 shows the typical '*modeling scales*' (adapted from Dooge, 1982; 1986) with respect to both, time and space. In respect of the temporal scales, individual events and seasonalities may be considered as to occur at a relatively short and restricted period of time (e.g. within hours or at an annual cycle). The '*long term scale*' by contrast covers a broader period including decades to centuries. With respect to the spatial scales, the figure shows that the considered scales cannot be characterized by a single size but cover a wider range of values (e.g. a catchment can cover a few ha to a few km²) partly allowing for overlaps between the individual scales.

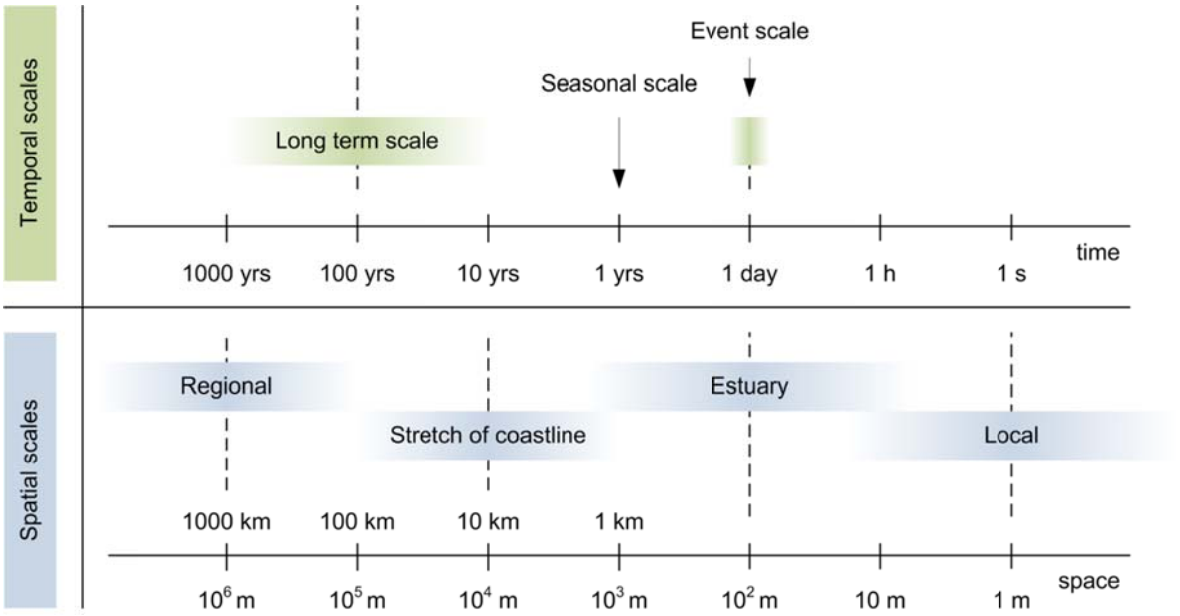


Fig. 1 Typical modeling scales adapted from Dooge (1982; 1986)

In hydrosciences, the modeling scale can be referred to characteristics and targets, but also to the applied methodologies (which is the most important aspect in this thesis) and all those can differ from one scale to another. However, in some cases there may also be differences within one individual scale. For instance, design levels for coastal defenses are usually defined using some form of statistical models (Dixon and Tawn, 1994). These models are mostly based on extreme value statistics, a special discipline in probability theory that deals with rare events, such as coastal floods (Coles, 2001). Using Fig. 2 as an example, the procedure of such models is usually as follows:

a) *Local water level records are used to derive an independent extreme value sample that is assumed to be representative of the predominating water level pattern.*
b) *Extreme value statistics are used to infer the characteristics of randomness that generated the data of the investigated process. This theoretical description enables to estimate the probabilities of events that are more extreme than any that have been observed (Coles, 2001). A predefined probability or return period (e.g. 200 years) can then be used to derive design levels for coastal defenses.*

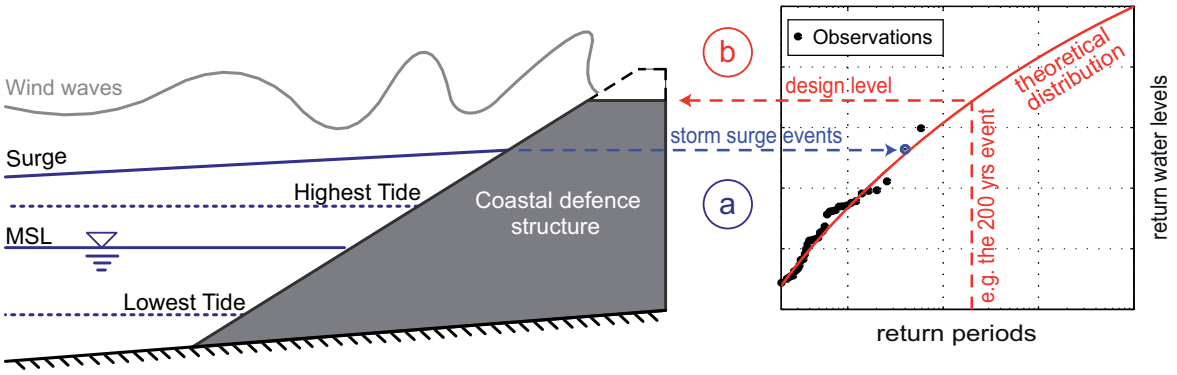


Fig. 2 An example on the application of extreme value statistics for the design of coastal defenses

Over the last five decades, several different extreme value analysis (EVA) methods for estimating the heights (i.e. return levels) and occurrence probabilities (i.e. return periods) of extreme still water levels have been developed (see Haigh et al., 2010b for an overview). There is, however, currently no universally accepted method available. Instead, different methods have been applied on transnational but also on national level (i.e. trans-regional and regional scale) resulting in a heterogeneous level of protection. Even the use of the same method can produce large discrepancies, because there is subjective choice at several steps in the model setup. In Germany, for instance, coastal protection is organized by government departments in federal states, who define design water levels using different methods. Applying different statistical methods can yield significantly different estimates of return water levels. As a result, it is difficult to assess the level of protection offered by defenses across the different federal states and equally difficult to compare this with neighboring defenses in the Netherlands and Denmark, where again different statistical techniques are used. To provide coastal protection of consistent standard (at least valid on regional scales), design levels need to be consistently calculated based on an objectively defined model setup.

However, an accurate assessment of return water levels using traditional extreme value analysis methods requires records of sufficient length (> 30 years; Haigh et al., 2010b), indicating one of the largest pitfalls of extreme value models, as the availability of measured water levels is limited in many regions. In the German Bight, multi-decadal records of high and low waters exist at several sites, but for some regions (e.g. at some small islands in the German Wadden Sea) no, or only very short and incomplete water level measurements exist. In practical applications it is often assumed that at-site (i.e. using local water level records from a tide gauge station) estimates can be transferred to un-

gauged surroundings. Nevertheless, water levels in the German Bight can differ significantly between stations as they are strongly influenced by shallow water effects and the complex topography of the coastline (see e.g. Jensen and Müller-Navarra 2008). Simply transferring information about the likelihood of extreme water level events from gauged to surrounding un-gauged sites is thus highly debatable and can cause erroneous return level estimates. Thus, more elaborate procedures to adequately transfer information from local to regional scales, such as the combination of numerical and statistical models, are required.

The usage of so called '*scaling'* methods helps (at least partly) to overcome this issue. Literally, the term '*scaling'* implies to increase or reduce the size of a feature (Blöschel and Sivapalan, 1995). With respect to hydrological applications, this refers to transferring information between scales which can further be differentiated into *up-* and *downscaling* (see Fig. 3). *Upscaling* indicates a transfer to larger scales as e.g. a *temporal upscaling* by using a 30 year record (i.e. a sample of individual events) to estimate a 100 year flood or e.g. a *spatial upscaling* by assuming the characteristics derived from one individual (local) tide gauge to be representative for an entire area (regional). *Downscaling* by contrast indicates a transfer to smaller scales (Gupta et al., 1986).

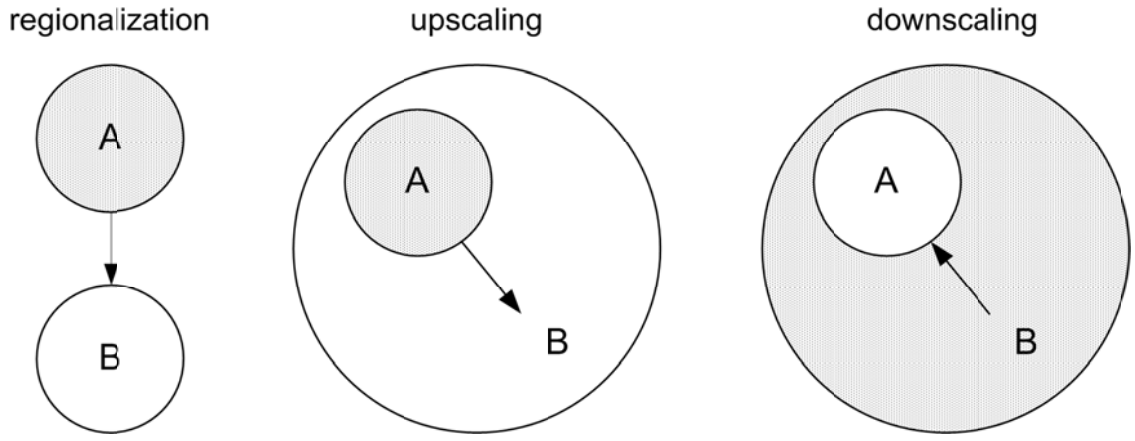


Fig. 3 Possible ways to transfer information between different scales and regions

The scaling issue also provides the basis for the concept of regionalization (Dyck and Peschke, 1995). In hydrosciences, the regionalization describes an approach that investigates larger areas or regions that have similar hydrological characteristics (see Fig. 3). However, instead of transferring information between scales, information is transferred from one location to another (Kleeberg, 1992). A general definition of regionalization is

given in Becker (1992), stating that *"regionalization is the regional transfer or area-wise generalization of a feature or a function […] or the parameter of that function"*. However, such scaling methods may also induce errors, as distributing information over space and time usually involves some sort of interpolation (Blöschel and Sivapalan, 1995), i.e. some effects are neglected while others may be amplified. The appropriateness of the applied method thus needs to be validated.

The return water level assessment is not only uncertain regarding the heterogeneous assessment procedures or the limited water level information but also with respect to possible future projections related to climate change. Recent analyses highlighted that global MSL rose by 2.0 mm/year from 1971 to 2010 (Church, 2013). As consequence from an increased ocean warming and the increased loss of mass from glaciers and ice sheets, future rates of sea level rise (SLR) are expected to very likely exceed those observed during 1971 to 2010 (IPPC 2013). Until recently, most coastal protection strategies assumed that changes in extreme water levels during the 21st century will be dominated by changes in MSL, and hence design water levels were raised to an amount equivalent to the projected SLR (Smith et al., 2010). These results are limited to the assumption of a similar long-term behavior between mean and extreme water levels. In the German Bight, however, Mudersbach et al. (2013) showed that trends in extreme high water levels differed significantly from those in MSL from the mid-1950s to approximately 1990, indicating the presence of nonlinear interactions between the different sea level components (i.e. MSL, tide, surge). This is contrary to most other locations around the world, where observed changes of extremes are equal to those of the MSL. In order to plan adequate adaptation strategies to cope with climate change challenges it is therefore essential that reliable projections of extreme water level changes become available. This in turn requires a profound understanding of the physical processes driving these changes, i.e. all relevant driving factors for regional and local changes in water level extremes need to be thoroughly investigated.

This brief summary shows that return levels need to be meticulous estimated to offer both an appropriate level of protection over the lifetime of the structure but also to avoid costly over design. This can only be achieved using and objective and consistent approach that accounts for local and regional effects at both present and future conditions. In view of the intention of this thesis, the '*regional'* and '*local'* scales are of particular importance, either focusing on an entire region (i.e. ideally including a number of

locations) or locally confined conditions (i.e. at individual stations/locations). In terms of temporal scales, this thesis mainly focuses on individual events (e.g. individual storm surges) that are used to make inferences about the likelihood of extreme events that are far beyond the observed period.

1.1 Research questions

The previous section showed that return water level estimates are often used to design coastal defense structures. However, the methodologies used for calculating return water levels are not consistent on transnational and in some cases even on national level (i.e. on regional scale). The first main objective (objective #1) of this thesis is thus to develop a methodology to obtain objective and stable results from extreme value analyses based on an automatically selected model setup which is spatially consistent on (at least) the regional scale. The associated research question of objective #1 is:

1) *How to estimate comparable, robust and consistent return water levels on regional scales?*

The overall intention of objective #1 is to provide guidance for coastal engineers, managers and planners who use these methods or the results produced by them.

Nevertheless, even if a universally accepted method has been established, the return level estimation can be challenging if there are only a few measured water level records available in a region, that are currently too short to apply traditional extreme value analysis methods. The second main objective (objective #2) of this work is thus to develop a methodology to estimate return water levels where no or just too short water level records exist. The research question of objective #2 is:

2) *How to estimate return water levels in un-gauged areas?*

The intention of the objectives #1 and #2 is to enable the calculation of present day return levels suitable for coastal defense design but both do not account for SLR and potential nonlinear changes in the tidal characteristics, which in turn may affect the results from extreme value statistics. One of the main challenges in coastal engineering is to estimate how SLR alters the design levels of coastal defenses. Until recently, most coastal protection strategies assumed that changes in extreme water levels during the 21st century will be dominated by changes in MSL. Nevertheless, a recent assessment by Mudersbach et al. (2013) showed discrepancies in trends of mean and extreme sea levels in the German

Bight indicating that the estimation of future design levels by raising extreme water levels by an amount equivalent to the projected SLR may underestimate the impact of SLR on return levels in some areas. This is why the third main objective (objective #3) is to investigate the effect of SLR on return levels in the German Bight. The research question of objective #3 is

3) How does sea level rise affect extreme water levels?

The thesis mainly addresses those three research questions but also investigates several other objectives which will be highlighted at the beginning of each part (see also the following section).

1.2 Structure of this work

This thesis pursues the three main objectives (#1, #2, #3), each of which is addressed separately in individual Parts (I, II, III). The structure of the work is highlighted in Fig. 4. The uppermost level of this figure shows a country-wise separation of scales, highlighting that the international and national sub-scales are unambiguously separated by borders. This assumption may hold for political issues but is often unrewarding for hydrological studies. This is why a further differentiation is made by subdividing the international and national scales into the trans-regional (e.g. global, the entire North Sea), regional (e.g. the German Bight, federal states) and local (e.g. individual stations) sub-scales (compare to Fig. 1). The dashed diagonal lines between those three sub-scales indicate that only blurred borders exist allowing for overlaps. As a result, models, studies or processes with scales that are border located cannot be assigned to one single scale but may be integrated in the assessment of both bordering scales. All three parts mainly focus on regional and local scales. However, individual applications such as the numerical model of Parts II and III partly also consider trans-regional scales as e.g. the entire North Sea.

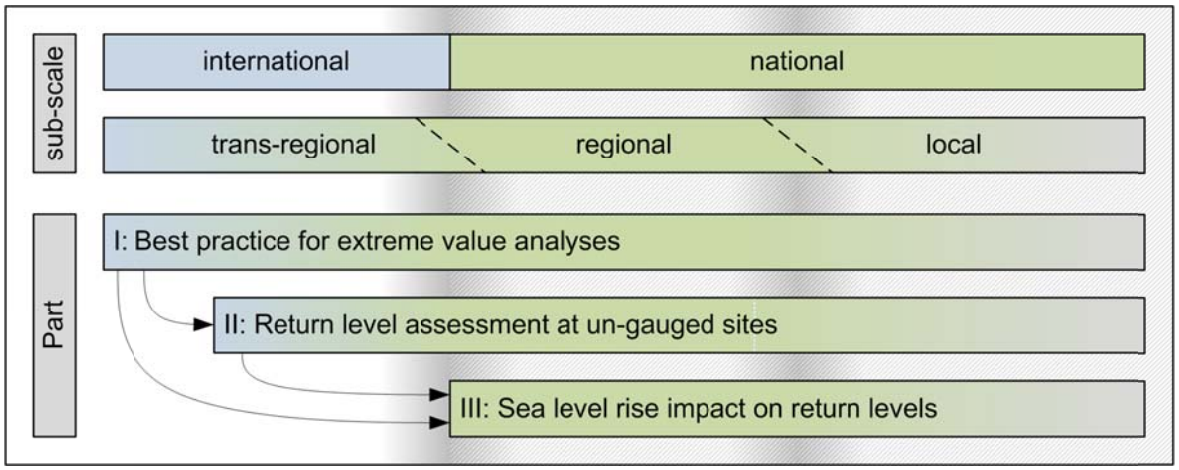


Fig. 4 Structure of the thesis

Part I investigates the applicability of commonly used approaches for estimating return water levels and provides guidance for coastal engineers, managers and planners who use extreme water level assessment methods or the results produced by them. The structure of Part I is as follows: Sect. 5 summarizes the various approaches and required operations for extreme value analyses described in literature. Results from analyzing the performance of different model set-ups, primarily based on the Cuxhaven record, are shown in Sect. 6. In Sect. 7 the transferability of the findings that are based on the Cuxhaven record is tested. The results are summarized and discussed in Sect. 8. In Sect. 9 the key findings and recommendations are given.

Part II describes an approach to assess return water levels at un-gauged sites. This part is based on the recommendations given at the end of Part I. Part II is organized as follows: Sect. 11 summarizes the operations required for performing the so called regional frequency analysis (RFA) method followed by an assessment of its applicability and limitations to water level data in the study area. Sect. 12 describes a hydrodynamic numerical model, its configuration and calibration, as well as a bias correction method that was undertaken and lastly the validation of this approach. The results from combining the numerical model output and extreme value analyses are presented in Sect. 13 followed by a summary and a discussion in Sect. 14 .Part II closes with the key findings in Sect. 15

Part III investigates the impact of SLR on extreme water levels in the German Bight. Part III is the last Part of the thesis and considers the recommendations of Part I as well as the numerical model of Part II. Part III is organized as follows. Sect. 16 and Sect. 17 provide a brief review of papers dealing with changes in extreme water levels and its

components. The sections are intended to justify the model assumptions and to review the current knowledge. In Sect. 18, a brief summary on physical background knowledge regarding tidal water levels and their response to different factors is given. The applied methodology is described in Section 19. The results are presented and discussed in Section 20, followed by the key findings in Section 22.

The thesis ends with general conclusions in Sect. 23 and a discussion of possible future research activities in Sect. 24.

2 Study Area

This thesis was conducted as part of the '*ZukunftHallig'* research project investigating the future development of the North Frisian Halligen. The Halligen are small low lying islands located off the coastline of Schleswig-Holstein (the most northern federal state) in Germany (see the blue shaded areas in Fig. 6d). The Halligen are surrounded by the North Frisian Wadden Sea. With an area of approximately 9.000 km², the Wadden Sea is one of the world's largest intertidal wetlands and in 2009 it was added to UNESCO´s World Heritage List.

Hardly any other landscape has experienced such major changes in the last few centuries as the North Frisian Wadden Sea. These changes were mainly caused by the last ice age (Quedens, 1992) which had its maximum around 20,000 years ago. During this cold period, large parts of the global water were bound in the continental ice sheets and this is why the MSL was more than 120 m below today's level (von Storch et al., 2009). In the subsequent phase, the still ongoing Holocene era which started around 11,000 BP, the temperature increased. At this time, large parts of the present-day Wadden Sea were dry land. From north to south, these land masses were partly separated by watercourses. In the eastern part, the North Sea and the land masses were separated by ramparts of sand and geest, created by the moraines of the last ice age (late Pleistocene). As a consequence from the increasing temperature, the ice sheets started to melt. This in turn caused sea level to rise and large areas of the formerly dry North Sea basin were flooded. However, the SLR was not uniform and showed temporary phases of stagnation and even decreases of up to 2 m (Quedens, 1992). During this phase, large marshlands developed along the geest areas of the eastern North Sea basin. These areas were hardly affected by Sea water resulting in a desalination of the groundwater (Quedens, 1992) enabling the development of extensive fen- and woodlands.

Around 2,000 years ago (the so called Dunkirk Regression), the SLR induced flooding of large parts of the fen- and woodlands caused most of the vegetation to die-off.

At the same time, a progressive deposition of marine sediments provided by the floods caused some areas (e.g. the small Halligen islands) to emerge. As these areas were very fertile, people started settling around 1,000 AD. The settlers transformed the natural into cultivated land and created drainage systems, dikes and sluices (Riecken 1985). Additionally, people started to dig for peat that developed from the died-off vegetation of the fenlands. The peat was used as a combustile and for salt extraction purposes. Especially the latter was a valuable asset at this time. All those interventions caused large parts of the underground to dry out and the already low lying areas subsided.

Over the centuries, the common impact of cultivation, subsidence, MSL changes and storm surges caused massive land losses in this region. Between the 13th and 20th century, approximately 50% of the surface area of the Halligen irretrievably vanished as a result of large storm surges. It has been estimated that around 100 Halligen have been destroyed over the last centuries; only 10 Halligen have survived to present day (Quedens, 1992). Comparing historical maps from the 13th century and the 17th century reveals that there have been enormous land losses in this 300 year period (see Fig. 5). Such land losses continued until the present but with a much smaller pace.

a) b)

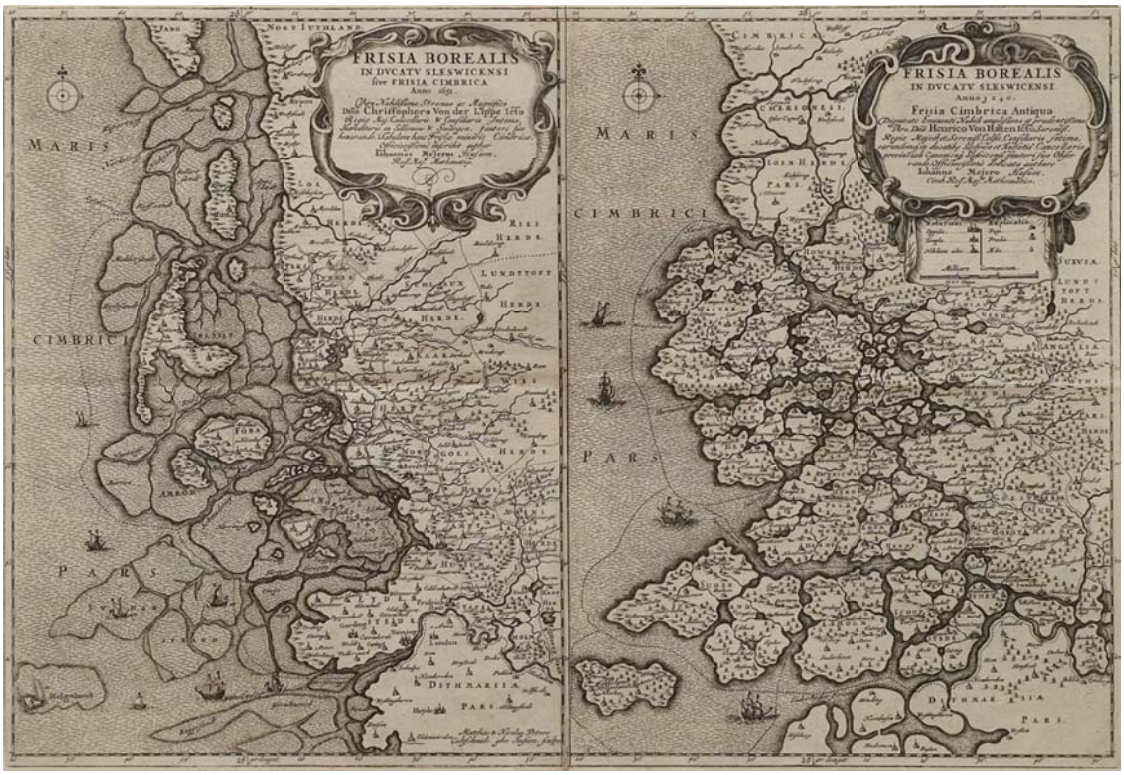


Fig. 5 The coastline of the federal state of Schleswig-Holstein in the years a) 1651 and b) 1241 (Danckwerth, 1963)

Today, the Halligen have no dikes and as they are frequently exposed to extreme water levels, they are inundated up to 50 times a year. Nevertheless, they are inhabited by around 270 residents. In order to protect the inhabitants from regular inundation, houses have been built on artificial dwelling mounds. Residents have learned to cope with extreme conditions, but it is expected that the Halligen will be gradually negatively affected, especially as a consequence of rising sea levels. Besides having a historic-cultural importance, the Halligen are believed to reduce the storm surge impact for the mainland coast of Schleswig-Holstein by providing a natural barrier of protection (although this effect has not yet been investigated in detail and quantified). It is thus of great importance to preserve these small Islands and it is the reason why this area was selected as a case study for this thesis.

In particular, Part II of this thesis shows a methodology to assess return levels in un-gauged areas. The methodology was originally developed to provide return level estimates for the Halligen, as there are no tide gauges available which provide sufficient information to conduct reliable extreme value analyses. In Part III, the impact of SLR on return levels is investigated focusing on the Halligen area. However, all parts of the thesis are also valid for the northern part of the German Bight and partly also for the entire German Bight (e.g. Part I).

3 Data sets

3.1 General remarks

Tide gauge records play a critical role in the assessment storm surges (Luther et al., 2007) as they provide information about the magnitudes and frequencies of extreme water levels. As each Part of this thesis (at least partially) focuses on different scales (see Sect. 1.2), a range of different tide gauge records is considered as elucidated in more detail in the subsequent paragraphs. All relevant information on the tide gauge data sets used in the individual parts is summarized in Tab. 1. In the "number of years" column, the indices (pv) and (hr) refer to the resolution of the datasets with the first one indicating peak value datasets (i.e. high and low waters) and the latter indicating high resolution datasets (i.e. at least hourly values). The tide gauges listed in Tab. 1 are shown in the subpanels a) to e) of Fig. 6. The areas emphasized with green rectangles are presented in more detail in the subsequent subpanels, e.g. the green rectangle of Fig. 6a is highlighted in Fig. 6b. All tide gauge stations that were considered in this thesis are displayed as circles. The larger circles, i.e. the tide gauge stations that are not enclosed by the green rectangles, contain information about the gauge numbers and in which part they were used.

Tide gauge records are generally subject to both anthropogenic (e.g. changes in the devices or locations of instruments, erroneous data processing) and natural (e.g. defects due to waves and/or salinity) influences (see e.g. Aguilar et al., 2003) which can bias records. Thus, all data sets were visually checked for common errors (such as isolated data spikes and timing errors; see Pugh (1987) p56–57 for a description of these) and suspicious records were excluded.

Furthermore, the candidate-reference approach (Aguilar et al., 2003) was used to identify discrepancies between records. In the candidate-reference approach, the ratio between a candidate and a reference time series is calculated. The reference is typically a

Tab. 1 Tide gauge records considered in Part I, II and III

#	Site name	Country	Tidal regime	Location		Tide/ surge ratio	Years	Period considered	Completeness (%)	Part I	Part II			Part III
				Lon	Lat						Cal.	Cor.	Val.	
1	Aberdeen	UK	Semidiurnal	57.13300	-2.10000	-	1 (hr)	2006	100	-	√	-	-	-
2	Lowestoft	UK	Semidiurnal	52.47000	1.75000	-	1 (hr)	2006	100	-	√	-	-	-
3	Whitby	UK	Semidiurnal	54.48000	-0.61000	-	1 (hr)	2006	100	-	√	-	-	-
4	K 13a Platform	NL	Semidiurnal	53.21700	3.22000	-	1 (hr)	2006	100	-	√	-	-	-
5	Calais	FRA	Semidiurnal	50.95000	1.85000	-	1 (hr)	2006	89.6	-	√	-	-	-
6	Tregde	NOR	Semidiurnal	58.00000	7.56700	0.5	77 (hr)	1927 - 2003	73.5	√	-	-	-	-
7	Newlyn	GB	Semidiurnal	50.10000	-5.55000	8.2	94 (hr)	1915 - 2008	98.1	√	-	-	-	-
8	Fremantle	AU	Mixed	-32.05000	115.73333	1.1	113 (hr)	1897 - 2009	90.8	√	-	-	-	-
9	Fort Denison	AU	Semidiurnal	-33.85000	151.23333	3.5	91 (hr)	1914 - 2004	98.1	√	-	-	-	-
10	List	GER	Semidiurnal	55.01826	8.44140	1.4	74 (pv)	1936 - 2009	99.5	√	-	√	√	√
11	Hörnum	GER	Semidiurnal	54.75979	8.29700	1.7	1 (hr) 74 (pv)	2006 1936 - 2009	98.9 99.5	- √	√ -	- √	- √	- √
12	Wittdün	GER	Semidiurnal	54.63344	8.38493	-	40 (pv)	1970 - 2009	100	-	-	√	√	√
13	Wyk	GER	Semidiurnal	54.69333	8.57638	-	40 (pv)	1970 - 2009	100	-	-	√	√	√
14	Dagebüll	GER	Semidiurnal	54.73220	8.68801	2.0	74 (pv)	1935 - 2009	98.9	√	-	√	√	√
15	Schlüttsiel	GER	Semidiurnal	54.68270	8.75590	-	40 (pv)	1970 - 2009	100		-	√	√	√
16	Husum	GER	Semidiurnal	54.47394	9.02585	2.3	75 (pv)	1935 - 2009	99	√	-	√	√	√
17	Büsum	GER	Semidiurnal	54.12222	8.85916	2.5	74 (pv)	1935 - 2009	98.5	√	-	√	√	√
18	Helgoland	GER	Semidiurnal	54.17660	7.89522	-	40 (pv)	1970 - 2009	100	-	-	√	√	√
19	Pellworm Harbour	GER	Semidiurnal	54.53000	8.70000	-	40 (pv)	1970 - 2009	100	-	-	-	√	-
20	Cuxhaven	GER	Semidiurnal	53.86778	8.71750	2.3	91 (pv)	1918 - 2011	100	√	√	√	√	√
21	Wilhelmshaven	GER	Semidiurnal	53.51444	8.14500	2.6	74 (pv)	1935 - 2009	99	√	-	√	√	-
22	LT Alte Weser	GER	Semidiurnal	53.86333	8.12750	2.6	109 (pv)	1900 - 2009	97.2	√	-	√	√	-
23	Borkum	GER	Semidiurnal	53.55750	6.74778	2.6	74 (pv)	1935 - 2009	99	√	-	√	√	-
24	Norderney	GER	Semidiurnal	53.69806	7.15853	2.2	75 (pv)	1935 - 2009	100	√	√	√	√	-
25	Emden	GER	Semidiurnal	53.33670	7.18600	-	40 (pv)	1970 - 2009	100	-	-	√	√	-

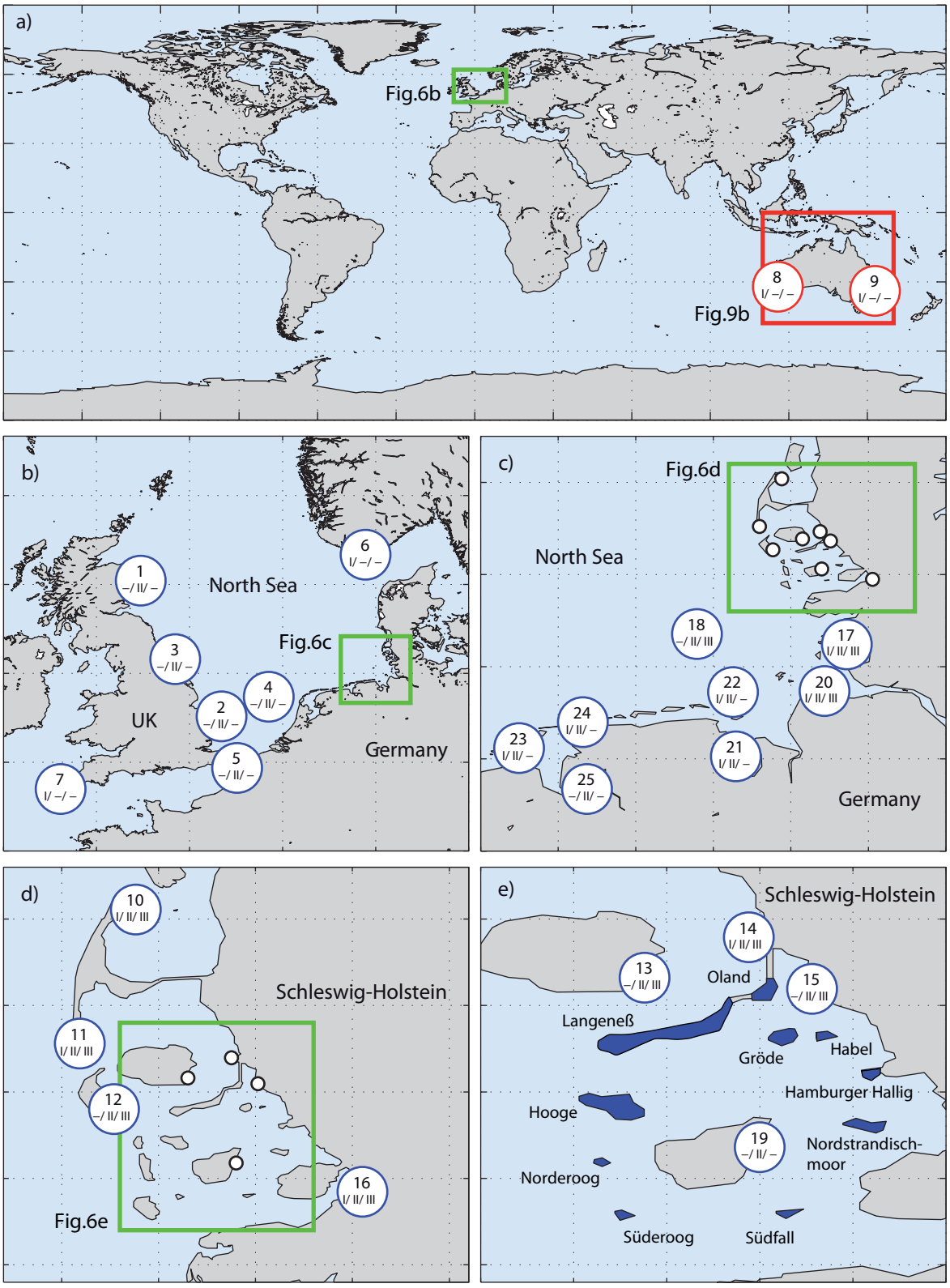


Fig. 6 Study areas that were considered in the thesis. The areas highlighted with the green rectangles are presented in more detail in the subsequent subpanels. The larger circles, i.e. the tide gauge stations that are not enclosed by the green rectangle, contain information about the gauge number and where it was used (Part I, II and/or III); the smaller dots are the remaining tide gauge stations

virtual station resulting from averaging nearby stations, whereas the candidate is the station that is intended to be checked for errors. The candidate and the reference series are compared by calculating their ratios, indicating discrepancies relative to each other. After a visual detection (an automatic detection has not yet been applied), erroneous periods were deleted manually. Fig. 7 exemplarily highlights the application of this methodology. In all subpanels, the reference was constructed using a range of different tide gauge records from the German Bight, each covering the period 1999 to 2009 (for illustration purposes only). As an example, the reference series is compared to three different candidates (i.e. the Hörnum, Husum and List records). The time scale under consideration for this example is 24 hours, i.e. each comparison is conducted for a timeframe of one day. In subpanels a) and c), there are no suspicious deviations visible. Subpanel b) by contrast shows a sudden drop of the ratio in 2009 indicating inhomogeneities in the Husum tide gauge record. In Fig. 8, this inhomogeneity is shown in detail, highlighting that the candidate time series is erroneous on July 29th, 2009. After detection the erroneous period was deleted.

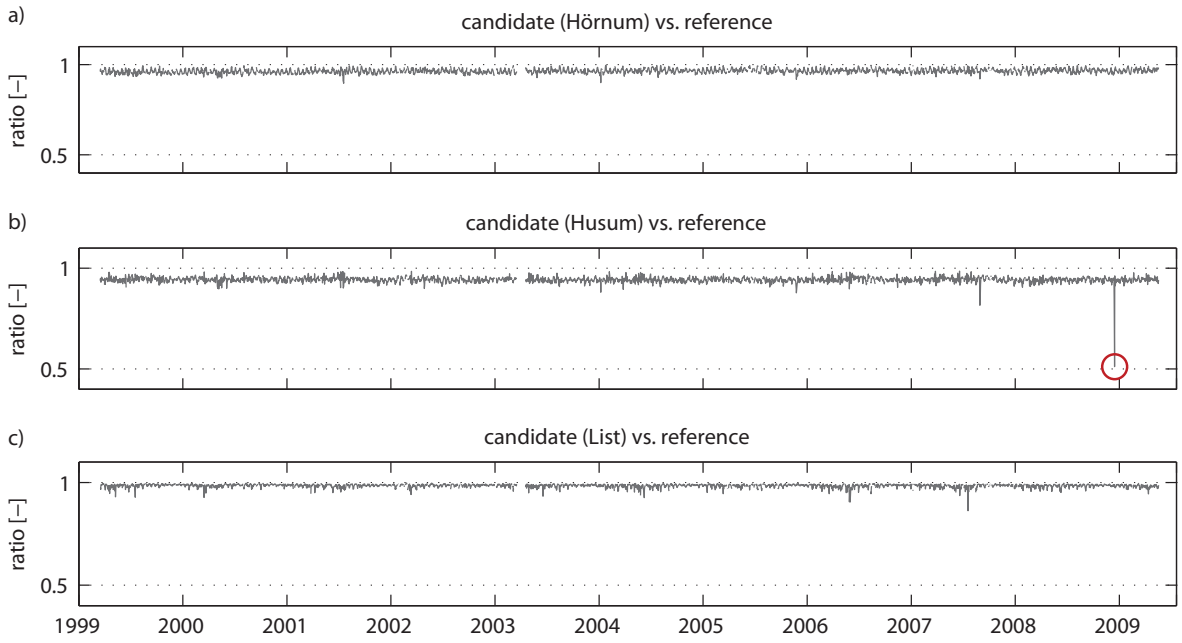


Fig. 7 Application of the candidate-reference method for detecting inhomogeneities in tide gauge data. The spike (highlighted as red circle) in subpanel b) indicates inhomogeneities in the candidate data set

3.2 Data used in Part I

Unless stated otherwise, the analyses and results described in Sect. 6 (Part I) are all based on the Cuxhaven record. The Cuxhaven tide gauge is located at the Elbe estuary and

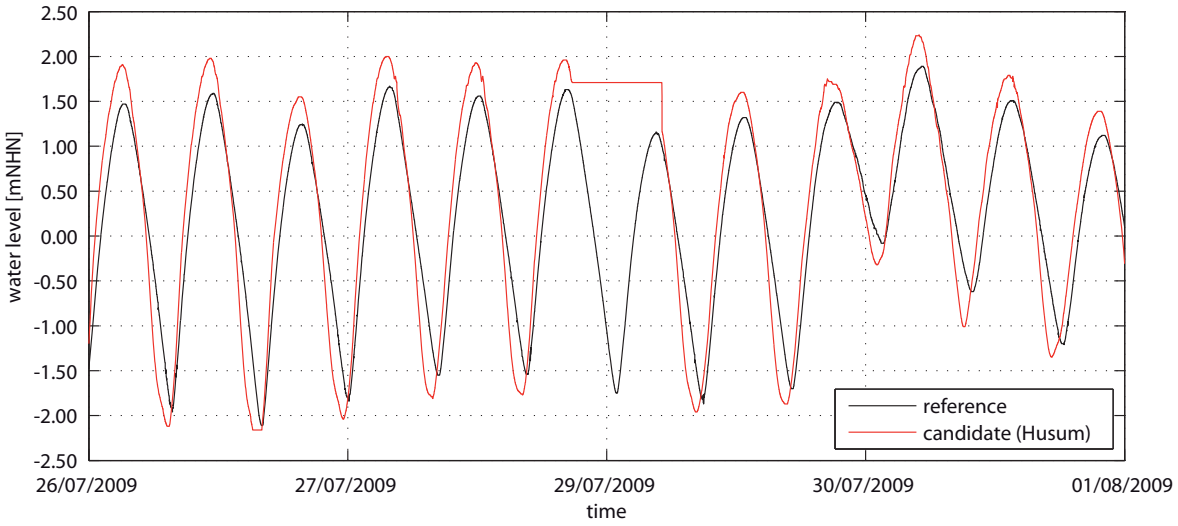


Fig. 8 Exemplarily depiction of the inhomogeneity that has been detected using the candidate-reference approach

provides high and low waters from 1843. High resolution data with at least hourly values are available since 1918 (Jensen, 1984; Jensen et al., 1992). In order to investigate the transferability of the results, water level records from 13 further tide gauges are also considered, these are: List, Hörnum, Dagebüll, Husum, Büsum, Alte Weser, Wilhelmshaven, Norderney, Borkum, Newlyn and Tregde, all located in northern Europe (Germany (GER), Great Britain (GB) and Norway (NOR)) while Fremantle and Fort Denison are located on the west- and east coast of Australia (AU). While Germany is the primary focus, the other international sites were selected in order to prove transferability of the proposed methods. Newlyn was chosen as it is one of the best documented and longest high frequency records in the world. Tregde provides a long time series of high frequency data and has a small tide/surge ratio (see the following paragraph) in comparison to the remaining sites. The Australian datasets represent two of the longest records from the southern Hemisphere. All German stations are referred to the German reference datum '*Normalhöhennull'* (NHN). Tregde, Freemantle and Fort Denison are referred to station Tide Gauge Zero (TGZ) which is linked to locally fixed benchmarks. Newlyn is referred to Admiralty Chart Datum (ACD) (http://www.bodc.ac.uk/).

To obtain criteria, at which of the tide gauges the analyzed methods of Part I are applicable, the ratio of tidal to non-tidal variation is used (see Dixon and Tawn, 1999). The tide/surge ratio was calculated as follows: a tidal analysis for each individual year was conducted with the Matlab t-tide package (Pawlowicz et al., 2002), considering a standard

set of 67 tidal constituents. The astronomical tide was then subtracted from the observed water levels in order to obtain the non-tidal residual. The tide/surge ratio was derived by dividing the 95th percentile of the tide by the 95th percentile of the surge. The ratios of tidal to non-tidal variation at each of the 14 tide gauges in the two study areas of Part I are shown in Fig. 9. The stations 6-7, 10-11, 14, 16-17 and 20-24 from Tab. 1 are all located in northern Europe and shown in Fig. 9a. The tide/surge ratio in this area is smallest in Tregde and slightly increasing along the German coastline from north to south. The largest tide/surge ratio is found in Newlyn, where it is more than 16 times higher than in Tregde. In Fig. 9b, the ratios of tidal to non-tidal variation of the two Australian tide gauge stations are shown, with one at the east and one at the west coast. The tide/surge ratio of these two stations differs by a factor of around 3 (see Tab. 1, Fig. 6 and Fig. 9b).

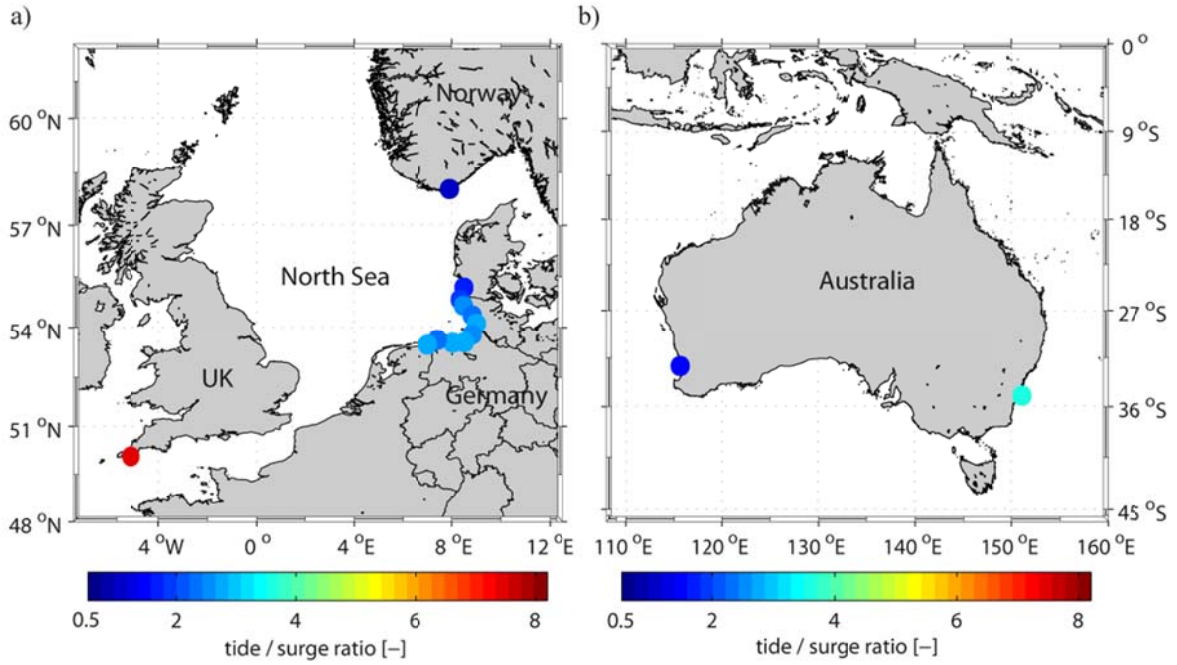


Fig. 9 The ratio of tidal to non-tidal variation at (a) 12 stations located in northern Europe and (b) at the two Australian tide gauge stations

Part I uses water level information that was available when the author started working of this thesis (i.e. the German gauges cover the early 1900s to 2009). To check the validity of the findings with recent (i.e. longer) water level records, the Cuxhaven data set was updated including water levels observed in the years 2010 to 2011 and the most relevant test (i.e. the stability of the distribution tested in Sect. 6.3) was repeated with this new data set.

3.3 Data used in Part II

Part II is based on a number of tide gauges along the coastlines of the United Kingdom (UK), the Netherlands (NL), France (FRA) and Germany (GER) (see Tab. 1) the locations of which are shown in Fig. 6. The hydrodynamic model that is setup in Part II requires a consistent vertical datum that is most useful for the water level response (Luther et al., 2007). This is why all water level records are consistently referred to the German vertical reference datum '*Normalhöhennull'* (NHN). To calibrate a numerical model, high resolution tide gauge data along the inner North Sea were used, covering the British East Coast, the English Channel, the Dutch coastline and the German Bight. The calibration was performed using a storm surge event that ocured on November 1st, 2006 and mainly affected the German coastline. For the bias-correction of the model output and the regional frequency analysis (RFA), high water levels for the period from 1970 to 2009 from all German Bight tide gauges except Pellworm Harbor were used; the water level record of Pellworm Harbor was used for validation purposes (see Sect.12.4).

3.4 Data used in Part III

Part III is primarily based on modeled water levels; observational data was only used to calibrate the model and this is already done in Part II.

PART I

HOW TO ESTIMATE COMPARABLE, ROBUST AND CONSISTENT RETURN WATER LEVELS ON REGIONAL SCALES?

4 Motivation

Over the past five decades, several approaches for estimating probabilities of extreme still water levels have been developed. Currently, different methods are applied not only on transnational, but also on national scales, resulting in a heterogeneous level of protection. In Germany, for instance, coastal protection is organized by government departments in federal states, each using different methods. Specifically, the German coastline has a total length of around 1,500 km with the two federal states Lower Saxony and Schleswig-Holstein directly bordering the North Sea. Two additional states, Hamburg and Bremen, are situated along tidal rivers (Elbe and Weser) that are strongly influenced by North Sea extreme water level events (see Fig. 10). All states have developed their own methods (although with some level of coordination) to derive design water levels (only Lower Saxony and Bremen use the same approach). In Lower Saxony and Bremen, a deterministic approach is used to calculate design water levels, i.e. the mean tidal high water level is superimposed with the largest observed storm surge, the difference between the largest spring tide and mean tidal high water, and a projected mean sea-level rise (NLWKN, 2007; see also Liese and Luck, 1978). By contrast, design water levels in Hamburg are based on an empirically derived design flood for Cuxhaven which is transferred from Cuxhaven to Hamburg using a two-dimensional hydrodynamic numerical model of the Elbe River (LSBG, 2012; see also Siefert, 1968; Siefert, 1998; Gönnert et al., 2013). Extreme value analyses are not part of the design procedure, but are applied afterwards in order to calculate the return period of the derived design water level. In the federal state of Schleswig-Holstein, the latest policy is to statistically derive design water levels associated with a 200-year return period using an extreme value analysis of the largest value per year superimposed with a projected mean sea-level rise (LKN, 2012; see also Wemelsfelder, 1939; Hundt, 1955; Führböter, 1976, Jensen, 1985). However, the choice of the model setup remains undefined. Hence, there is a considerable risk of

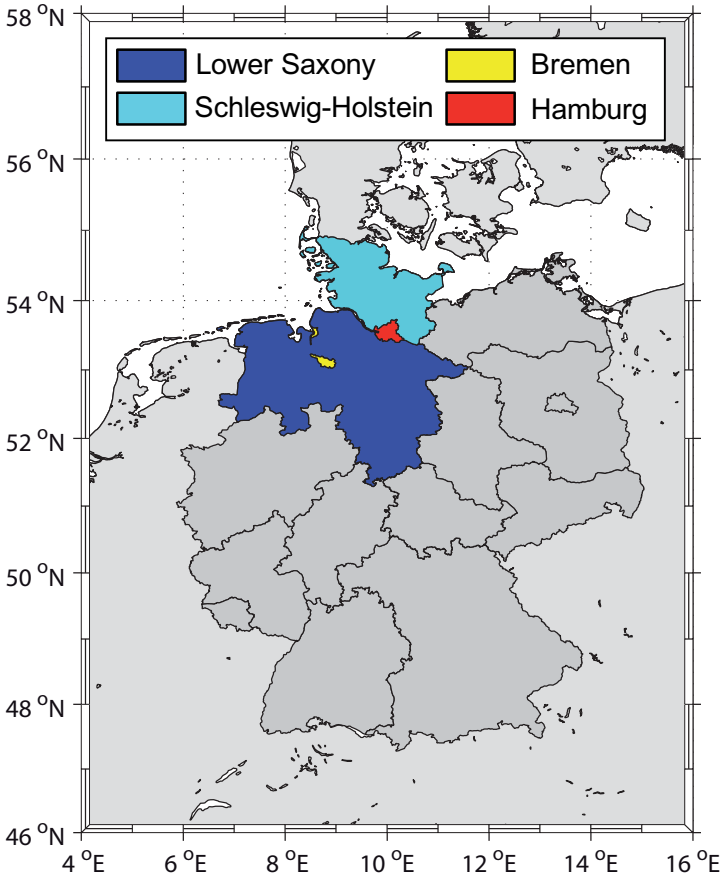


Fig. 10 The 16 federal states of Germany (depicted in dark grey). The four federal states being exposed to North Sea tides are shown in different colors according to the legend

subjectively influencing the return water level estimates. For a comprehensive review on the assessment of extreme water levels in the German Bight until 1985 can be found in Jensen (1985).

The application of different statistical methods can yield significantly different estimates of return water levels, and even the use of the same technique can produce large discrepancies, because there is subjective parameter choice at several steps in the model setup. This is why Part I focuses on a comparison of return level estimates using the two main direct methods (i.e. the block maxima method and the peaks over threshold method) considering a wide range of strategies to create extreme datasets and using a wide range of parameters in the model set up. Both of these methods have previously been applied to estimate return levels in Germany. The sensitivity of these direct methods to three important factors is tested, each of which can significantly influence the results of the statistical analyses. These three factors are: the detrending of the datasets; the sample that is created according to the chosen model; and the sensitivity of both distributions when

steadily reducing the dataset lengths. The final point is undertaken to examine the consistency of the considered direct methods for datasets covering different record lengths. Overall, Part I has three objectives:

(1) To briefly review the various steps involved in applying each method and describe the advantages and disadvantages of particular techniques involved;

(2) To test the sensitivity of the result from the extreme value analysis to the three factors mentioned above (i.e. detrending, sampling, and choice of distribution) and to develop an objective approach resulting in robust and stable return water level estimates that are applicable for design purposes; and

(3) To test the transferability of the defined approach, by applying this methodology to datasets from sites distributed along the northern European and Australian coastlines.

The overall aim of Part I is to provide guidance for coastal engineers, managers and planners who use these methods or the results produced by them. The challenge is in objectively obtaining stable results from extreme value analyses that are based on an automatically selected model setup and are spatially consistent on a national or even a transnational scale.

5 Theoretical background

Part I of the thesis focuses on the block maxima (BM) and the peaks over threshold (POT) methods. The main steps for estimating extreme water level probabilities, using these two methods, are summarized in Fig. 11. The first step involves detrending the input water level dataset to remove the influence of long-term trends such as mean sea-level rise.

The second step involves reducing the high frequency dataset to a sub-set of values to which the extreme value distribution is fitted. In the case of the BM method the data is reduced to r-largest values per year, or in the case of the POT method, to independent (i.e. from unique storm events) values above a given threshold. In steps three and four a choice is made regarding the parameter estimation technique that is applied and which extreme value distribution the sampled data will be fitted to. In step five, the appropriateness of the

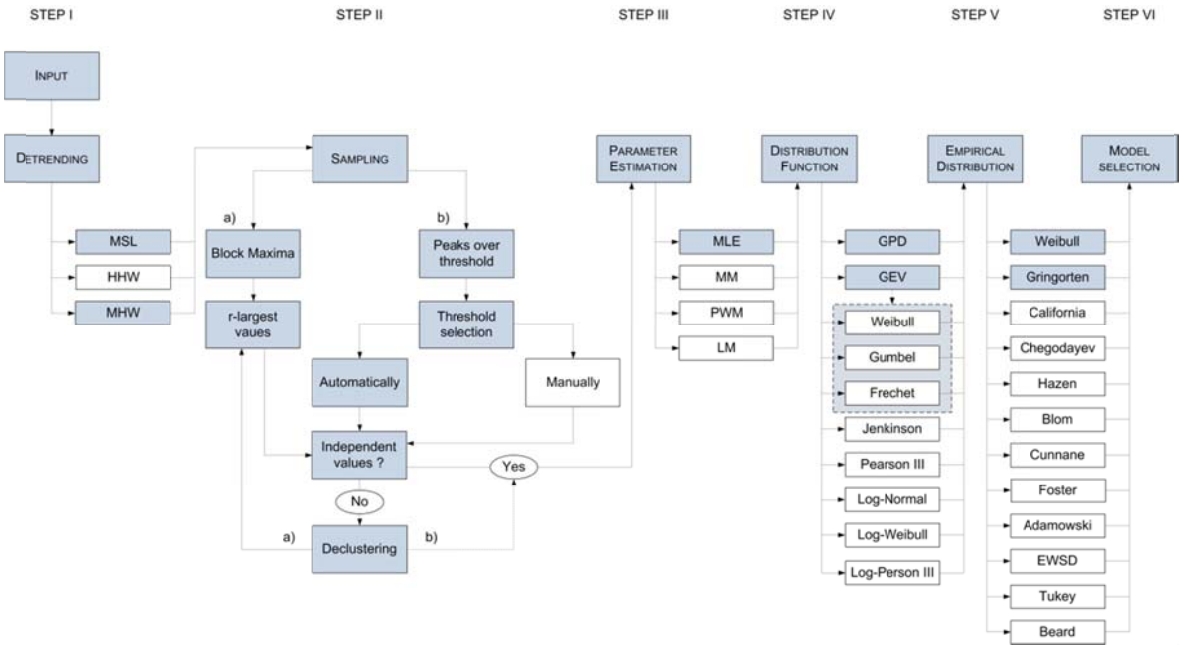


Fig. 11 Possible ways of conducting extreme value analyses. The procedures used in Part I are highlighted in blue color. As the GEV comprises the Weibull, Gumbel and Fréchet distribution these three are also shaded in blue

fit is tested against an empirical distribution (often referred to as plotting positions) (see e.g. Jensen, 1985). The five subsections below expand on each of these steps in turn and describe the main alternative techniques listed in the literature. It needs to be emphasized, that the list in Fig. 11 is not exhaustive, as there are many less common available methods not included. The thesis mainly focuses on the procedures highlighted in blue color.

5.1 Detrending

Detrending water level records prior to statistical analyses fulfils physical needs, as changes in water levels related to climate change can be compensated by 'adjusting' the data sets to present conditions. In probability theory on the other hand, the basic need for detrending datasets is founded on mathematical considerations as a fundamental assumption is that the dataset is independent and stationary (see e.g. Jensen, 1984; Rao and Hamed, 2000). An extreme value sample is considered to describe a random process, comprising independent and identically distributed (IID) random variables. However, individual values of a natural process often depend on recent past values (see Sect. 5.2 for a detailed discussion). Furthermore, the random behavior of these values often varies over time (e.g. due to seasonal effects), whereas stationarity assumes that the distribution of any subset of a sample remains the same (Coles, 2001). Water level time series can be assumed stationary if they are free of significant trends, shifts, or periodicity. This implies that the statistical parameters (e.g. mean and standard deviation) of the process do not change with time (Mudersbach, 2010). In real-world applications, many methods considered for time series analysis postulate some kind of stationarity. It is thus necessary to perform a test for stationarity, justifying the use of certain models (Wang et al., 2005).

Knowledge of the underlying physical system usually helps to identify trends in datasets. To date, numerous studies have been conducted in order to examine how extreme sea levels have changed and what were the driving factors. Woodworth and Blackman (2004) performed a global study and concluded that variations in extreme water levels are related to regional climate change and variability. They showed that secular changes and the inter-annual variability of extreme water levels were similar to those of the mean sea level (MSL) in most areas. These findings were consistent with the results of studies for the English Channel (Haigh et al., 2010a; Pirazzoli et al., 2006) and Australia (Church et al., 2004).

When dealing with trends, there are limiting issues. The identification of a trend in a time series remains subjective to some extent, as the trend cannot be clearly distinguished from variability and cyclic behavior in datasets. Mudersbach et al. (2013) showed that the variability of extreme water levels is much larger than the MSL variability in the German Bight. Similar findings were reported by Douglas (1991) and Haigh et al. (2010a) for other areas. Due to the variability in higher percentiles, trends of extreme sea levels are masked. MSL derived trends are thus more reliable than trends derived from extreme sea level datasets. At most stations around the world, it can be considered appropriate to use MSL trends to correct the datasets for the statistical analyses, as MSL trends represent an appropriate proxy for extreme water levels trends. In the case of the German Bight, Mudersbach et al. (2013) found similar results for the time period prior to the mid 1950s, concluding that changes in extreme sea levels were not significantly different from MSL changes. However, from the 1950s to the mid 1980s, estimated extreme high sea levels were found to have risen significantly faster than the MSL at all considered tide gauge sites. The authors concluded that these changes were primarily an effect of changes in the ocean tides. Therefore, detrending the datasets in this area by MSL only is not appropriate.

A common method for detrending is to use a simple linear regression (see e.g. AghaKouchak et al., 2013) which is applied to the complete dataset. In reality the temporal evolution of water levels is often far from linear, i.e. there are usually phases of accelerating or decelerating often associated with short and/or long-term periodicity. Hence, particularly in the physical meaning of detrending, the use of a linear correction can be a misleading assumption. An alternative approach consists in using a moving average of the chosen variable (i.e. MSL, mean high water, annual maxima etc.). This method accounts for different temporal changes across the whole dataset as it allows for correcting trends as well as periodicities on various timescales. The time scale of interest can be accounted for by defining an appropriate window size for the moving average.

Dixon and Tawn (1994) stated that water level datasets exhibit a non-stationarity resulting from seasonal changes of water levels with the majority of higher water levels occurring in the winter seasons. They showed that neglecting seasonality could result in a significant underestimation of the return water levels highlighting the necessity to account for both long-term changes and seasonal fluctuations when detrending the data sets. It is thus one of the main assumptions in fitting a distribution function to datasets, that the IID criteria are fulfilled. Particularly for the application of extreme value statistics, the

presence of trends or non-stationarities is undesirable. Ideally, any remaining trends or non-stationarity should be identified and removed to leave an approximately stationary process (Hawkes et al., 2008). An alternative approach is to incorporate non-stationarities in the calculation of return water levels as shown for example by Dixon and Tawn (1994), Méndez et al. (2007) and Mudersbach and Jensen (2010). Sect. 6.1 explores the effects of using different detrending approaches.

5.2 Sampling

The block maxima (BM) method is based on the assumption that the generalized extreme value distribution (GEV) is a good approximation to the distribution of the r-largest water level events within a certain time span (Dixon and Tawn, 1994). The choice of the model determines the way the sample is created. Many studies use the annual maximum (AMAX) value (i.e. r = 1 value/ year), of each year of the record (e.g. Acero et al., 2011). However, it is a wasteful method if further data on extremes are available (Coles, 2001). Further, the 2nd or even the 3rd largest values in a given year can be larger than the AMAX value in another year. Consequently, the AMAX method was extended by Smith and Weissman (1994) (see also Smith, 1986; Tawn, 1988) in order to include a fixed number of independent variables with r > 1 values/ year, the so called r-largest values of each year, into the sample. By applying the r-largest order statistics along the UK coastline, Dixon and Tawn (1994) showed that a choice of r = 8 values/ year appears to yield robust estimates. However, despite incorporating more of the observed extreme data in the estimation of extreme value statistics, even this method can be wasteful if one block contains more extremes than another (Coles, 2001). Especially in areas where water levels show a large inter-annual variability, the largest event within one year may hardly exceed the mean high water level of all years. Considering these events as extreme is misleading. Thus, the low efficiency of the BM method is its largest pitfall (i.e. large estimation uncertainties caused by small sample sizes).

The peak over threshold (POT) approach by contrast is much more efficient (if a not very high threshold is justified) as it considers all values exceeding a certain threshold. Hence, a POT derived sample comprises not only one or a fixed number of events per year. It rather allows for a more rational selection of events fulfilling the criteria of being “extreme” (Lang et al., 1999). In the POT approach, the aim is to develop robust estimates when the model distribution for the exceedances above a threshold is the generalized

Pareto distribution (GPD; Dupuis, 1998). By comparing AMAX estimates with POT estimates, Cunnane (1973) concluded that the POT approach produces a smaller sampling variance than the AMAX method if the POT derived dataset contains at least 1.65 extreme events per year. The key challenge consists in the determination of an optimal threshold as several important features of frequency modeling are very sensitive to the selected value. If the selected threshold is too low, it causes a bias because the model assumptions are invalid (i.e. values might not be independent or non-extreme data are included in the sample). If the threshold is too high, the variance is large because only few data points are included in the analyses. In extreme value analyses, where models are likely to be extrapolated beyond the observations, this may lead to large differences in the results.

Therefore, diagnostic procedures are needed for the threshold selection and there are in general two different methods that can be adopted. The first one is based on physical criteria whereas the second one is mathematically motivated (Lang et al., 1999). In the physically based approach, the threshold defines a water level that, if exceeded, results in inundation. This approach is often used in river engineering while for coastal waters mathematically based methods are usually preferred. Over the last decades, a number of these mathematically based selection methods were proposed, being either parametric or non-parametric. Rosbjerg et al. (1992) introduced a parametric procedure based on a factor k, the mean value $\bar{x}$ and the standard deviation S of the original dataset. The threshold u_0 is calculated as $u_0 = \bar{x} + k \cdot S$, and the authors recommended using a factor of k = 3.

Coles (2001) suggested using mean residual life plots (MRLP) for the threshold selection. As an example, Fig. 12a shows the MRLP applied to the Cuxhaven tide gauge record (see Tab. 1) which is calculated as

$$\left\{ \frac{1}{n_u} \sum_{i=1}^{n_u} \left(x_{(i)} - u_0 \right) : u_0 < x_{max} \right\} \quad , \qquad \text{Equation 1}$$

where x(i) consist of the n_u observations which exceed the threshold u_0; xmax is the largest of all values in the sample. The concept of this method is that if the GPD provides a good approximation to the threshold exceedances, the MRLP should approximately be linear (u_0 is considered at the starting point). Fig. 12a shows, that there are several possible segments in the MRLP where this condition is complied as e.g. at $u_0 = \{185, 265, 385\}$ cmNHN. However, which of those three thresholds is the 'correct' one? As an alternative, Coles

(2001) provides the use of the stability plot (SSP) which investigates the shape parameter dependence on the threshold u_0. The recommendation is to use the threshold value, where the shape parameter remains constantly as highlighted in Fig. 12b ($u_0 = 255$ cmNHN). Although widely used (see e.g. Morton and Bowers, 1996; Choulakian and Stephens, 2001), the example shows that both, the MRLP and the SSP are not simple to apply in practice as they rely on subjective judgment in interpreting the resulting graphs and thus cannot be easily converted into automatic selection algorithms.

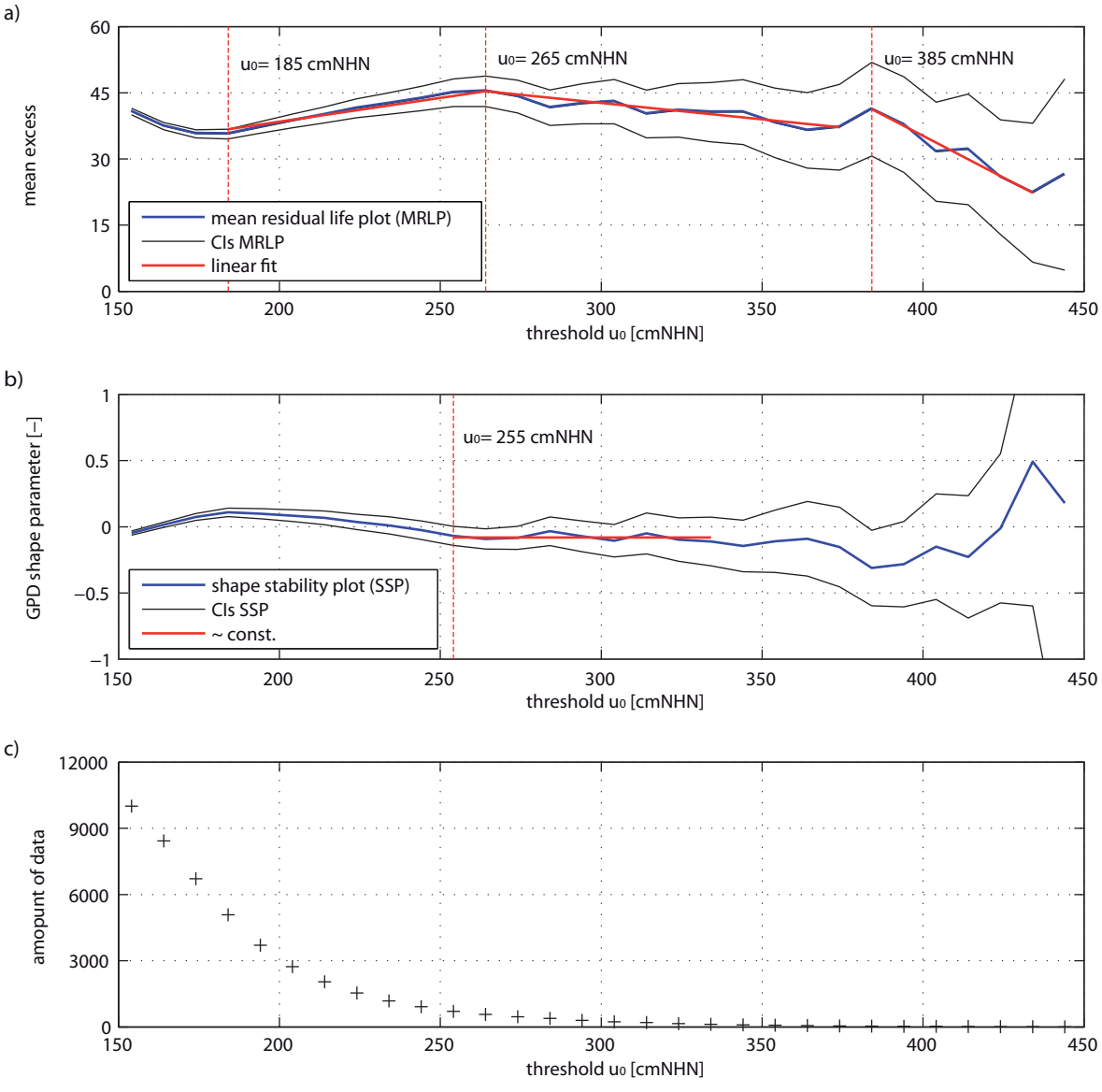


Fig. 12 a) Mean residual life plot (MRLP); b) shape stability plot (SSP); c) number of threshold excesses

Thompson et al. (2009) and Zhang and Ge (2009) considered hypothesis tests in combination with samples derived with different threshold values. Percentiles are often

used to derive threshold values, with the range of percentiles varying between the 97.5th (Environment Agency, 2011) and the 99.5th percentile (Grabemann and Weisse, 2008). This leads to the question: Which percentile leads to an appropriate threshold?

Furthermore, water level datasets can exhibit dependencies (so called clusters), which are mostly related to the same meteorological forcing. For practical applications, a limiting condition for extreme sea levels is often assumed where the events $x_i > u_0$ and $x_j > u_0$ are independent if u_0 is sufficiently high and the occurring times of i and j are far enough apart (Coles, 2001). For the BM method, independence of maxima can be achieved by selecting a large block size, while the theoretical assumptions are less critical in practice. Apart from the excesses of the selected threshold modeled by a GPD, the exceedance times are also modeled by a Poisson process.Therefore the POT approach unifies the asymptotic models, while it can include non-stationary phenomena in a more complete way. The independence assumption is however more critical compared to the BM model and declustering techniques have to be used. In literature, various methods are proposed for identifying the "correct" declustering time of extreme data samples. Zachary et al. (1998) used a standard storm surge length (SSL) between 24 and 72 h, while in the special case of northern latitude storms, Mathiesen et al. (1994) used a SSL of 120 to 168 h. As a result of analyzing an autocorrelation function, Soares and Scotto (2004) used a SSL of 480 h for the North Sea. For most environmental problems however, it is not realistic to assume equally sized storm clusters as the correlation structure of different annual time series can vary significantly. The consideration of a constant declustering time or SSL for the entire time series is thus highly debatable (Soukissian and Arapi, 2011).

Due to the variety of available methods, declustering is often influenced by subjective choices and the selection of the declustering parameters is largely arbitrary. To overcome this issue, an automatic and objective declustering scheme is needed. As the reciprocal of the mean cluster size, the extremal index is an important parameter that measures the degree of clustering of stationary extreme value datasets (Smith and Weissmann, 1994). The extremal index θ is defined as

$$\theta(u_0) = \frac{2\{\sum_{i=1}^{n-1}(T_i - 1)\}^2}{(n-1)\sum_{i=1}^{n-1}(T_i - 1)(T_i - 2)} \quad , \qquad \text{Equation 2}$$

where T_i are the difference times of the random variable, n is the sample size, and u_0 is the threshold value (Ferro and Segers, 2003). An alternative interpretation of the extremal

index is that 1/σ is the limiting mean cluster size (Coles, 2001). In Sect. 6.2, the influence of using different declustering times on the results of extreme value statistics is investigated.

5.3 Parameter estimation

One of the main objectives in statistical modeling is to use the sample information to make inferences on the distribution of the population. Assuming that the sample is an independent realization of the overall population, the sample can be used to estimate the unknown statistical parameters of the population. The parameter estimation method should meet the conditions of being robust and of showing a small variability against the sample size. Throughout this thesis, the model parameters are obtained using the common maximum likelihood estimation (MLE) approach (Davison and Smith, 1990; Hosking and Wallis, 1987; Smith, 1986). The MLE is a general and flexible method to estimate the unknown parameters of a distribution (Coles, 2001; Naveau et al., 2005). However, according to Katz et al. (2002), the performance of MLE can be extremely erratic for small samples ($n \leq 25$), especially when estimating extreme quantiles of the GEV distribution. Other common parameter estimation methods are the method of moments (Sachs, 1997) and the L-moments method (Hosking and Wallis, 1997). Although the results are sensitive to the chosen method, the influence is typically smaller compared to choosing different thresholds, detrending approaches or distributions (see e.g. Hosking and Wallis, 1997; Brabson and Palutikof, 2000).

5.4 Theoretical distribution

Currently the GEV and the GPD are the most commonly used distributions for extreme value analyses, and hence why this thesis only focuses on these two distributions. According to the Fisher-Tippett theorem (Fisher and Tippett, 1928), all limit distributions of IID partial maxima (or BM) series are GEV distributed (Neves and Fraga-Alves, 2008). The GEV is defined as

$$GEV = exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \quad , \qquad \text{Equation 3}$$

where μ is the location parameter, σ the scale parameter, ξ the shape parameter and the block maxima values are z. This formula combines the Gumbel, Fréchet and Weibull families into one single distribution (Coles, 2001).

The use of POT methods is linked to the GPD, considering all values exceeding a threshold u. This was proven by Balkema and de Haan (1974) and Pickands (1975) showing that the limiting distribution for the excesses over a sufficiently high threshold is the GPD (Neves and Fraga-Alves, 2008). The GPD encompasses a number of common extreme functions (Hawkes et al., 2008) and is defined as

$$GPD = 1 - \left[1 + \frac{\xi y}{\tilde{\sigma}}\right]^{-1/\xi} \quad , \qquad \text{Equation 4}$$

with

$$\tilde{\sigma} = \sigma + \xi(u - \mu) \quad , \qquad \text{Equation 5}$$

where parameters are the same as above (Coles, 2001).

The values of the GEV parameters are affected by the block size considered. In contrast, the parameters of the GPD distribution of threshold excesses are not, as the shape parameter ξ of the GPD is invariant to the block size, while the transformed scale parameter $\tilde{\sigma}$ is the sum of changes in μ and σ (Coles, 2001).

5.5 Empirical distribution

Probability plots are useful for visually examining the character of extreme datasets (Stedinger et al., 1992). By using plotting position formulae, the probability of exceedances or non-exceedances can be calculated for observed events. According to De (2000), probability plots are used to fit a certain probability distribution to a given dataset, to identify outliers and to visually assess the goodness of fit (see also Jensen, 1985). Nowadays, empirical distributions are primarily used for the latter two aspects, while fitting a distribution to datasets is usually achieved using analytical procedures. It is, however, still customary to supplement analytically obtained results with empirically derived plotting positions (De, 2000). Most plotting position formulae are special cases of the general formula (Hirsch et al., 1992):

$$qi = \frac{i - a}{n + 1 - 2a} \quad , \qquad \text{Equation 6}$$

where q_i is the probability plotting position for the i^{th}-largest event, a is the plotting position parameter and n the number of observations. The most used plotting position formulae are the Weibull formula with a = 0 and the Gringorten formula with a = 0.44, both of which are considered hereafter.

5.6 Return level assessment

From a mathematical perspective, the estimation of return periods that are far beyond the observed period is doubtful. According to Kleeberg and Schumann (2001), the return period estimation is limited to approximately two to three times the observed period (see also DVWK, 1991; DWA, 2012). This, however, has not been proven mathematically but can be regarded as an acknowledged rule of technology. Nevertheless, there are designing guidelines available demanding for return periods of up to 10.000 years (e.g. the design of reservoirs in Germany according to DIN 19700-12:2004-07, or coastal defense structures in the Netherlands that protect the lower parts of the country, see Vrijing et al., 2007). This is why this thesis considers return levels up to 10.000 years. In the main area under investigation, coastal defenses are planned to withstand the one in a 200 years event. This is why most assessments in this thesis focus on the 200 years event. Larger return periods are grey shaded to highlight the doubtfulness of those results.

6 Method set-up and results

This section comprises the results from investigating the impact of the main factors (i.e. detrending, sampling, and selection of the distribution function) which can influence the estimation of extreme still water levels, using the long Cuxhaven record. Unless noted otherwise, all analyses are performed using the same model set-up (except the factor whose influence is tested): the trend is corrected using a moving average trend correction with a window size of 1 year, hereafter referred to as Detrend C (see Section 4.1). BM samples are generated using the r-largest annual events, with r = 1 value/ year to r = 6 values/year. In the POT approach, the threshold is selected in order to match the number of events to the number of the r-largest events in the BM derived sample (i.e. 1 to 6 values/year). Independence of the BM sample and the threshold exceedances was achieved using a declustering time of t_d = 1.5 days.

6.1 Detrending

As mentioned in Sect. 5.1, observed maxima have risen faster than the MSL in the German Bight, and hence it is not appropriate to detrend the data in this region using only MSL derived trends. It is more appropriate to use trends in mean high water (MHW),which include changes in MSL, as well as other observed changes in tidal range and storminess (Mudersbach et al., 2013). The datasets are thus corrected using the high water peak values, which are derived from the original datasets; hence no external input other than the data itself is needed. The influence of the long-term trend correction is evaluated using a linear fit covering the entire dataset (Detrend A) and a 19-year moving average fit accounting for the nodal cycle of 18.6 years (Detrend B) (Haigh et al., 2010b). To yield a good approximation of the seasonality as well as the long-term trend of the data, a moving average trend correction with a window size of 1 year (Detrend C) is also tested. The trend adjusted datasets are found by subtracting the estimate of the trend from the original series. For the sensitivity study, r = 1 value/year and r = 6 values/year are used in the BM

approach. In the POT approach, the threshold is selected in order to match the number of events to the number of the r-largest events in the BM derived sample (i.e. 1 and 6 values/year).

As shown in Fig. 13a, discrepancies of up to a few decimeters occur depending on the method which is used for the trend correction. The differences between the three techniques are small when using BM with $r = 1$ value/year. Using $r = 6$ values/year, however, leads to considerable differences in the estimated return water levels, with the highest estimates occurring when the 1-year moving average (Detrend C) is applied. The results from using values of $r = 2$ to $r = 5$ values/year (not shown) are approximately linear. Smaller differences are found for the POT based return water levels. As shown in Fig. 13b, the POT approach is less susceptible to the trend correction applied. Again, the highest return water levels are estimated with the Detrend C approach. Compared to the other two approaches, the difference is up to two decimeters. Differences between the return water levels derived with the Detrend A and Detrend B approaches are small in comparison. The results highlighted here are presumably case dependent. Therefore it is not intended to draw a global conclusion on magnitudes of difference, but it is worth emphasizing that the detrending approach influences the results. This is why using the seasonal adjustment is recommended as it appears most accurate.

6.2 Sampling

6.2.1 Block maxima method

To evaluate the impact of the selected block maxima (BM) sample on return water level estimations, the r-largest values per year are sequentially increased. Around the UK, Dixon and Tawn (1994), recommend using $r = 8$ values/year. In the case of the German Bight, however, samples comprising $r \geq 2$ values/year can lead to a substantial overestimation of higher return levels (see also Sect. 6.3). To highlight this, r-values ranging from $r = 1$ to $r = 6$ values/year are considered as shown in Fig. 13c. With an increasing number of r values per year, considerably larger return water levels are obtained, especially for higher return periods. Comparing water level percentiles highlights that higher percentiles tend to have a larger variance than lower percentiles (Mudersbach et al., 2013). With an increasing r, improved estimates of the unknown parameters of the distribution can be achieved as the variance of the sample reduces, which in turn reduces

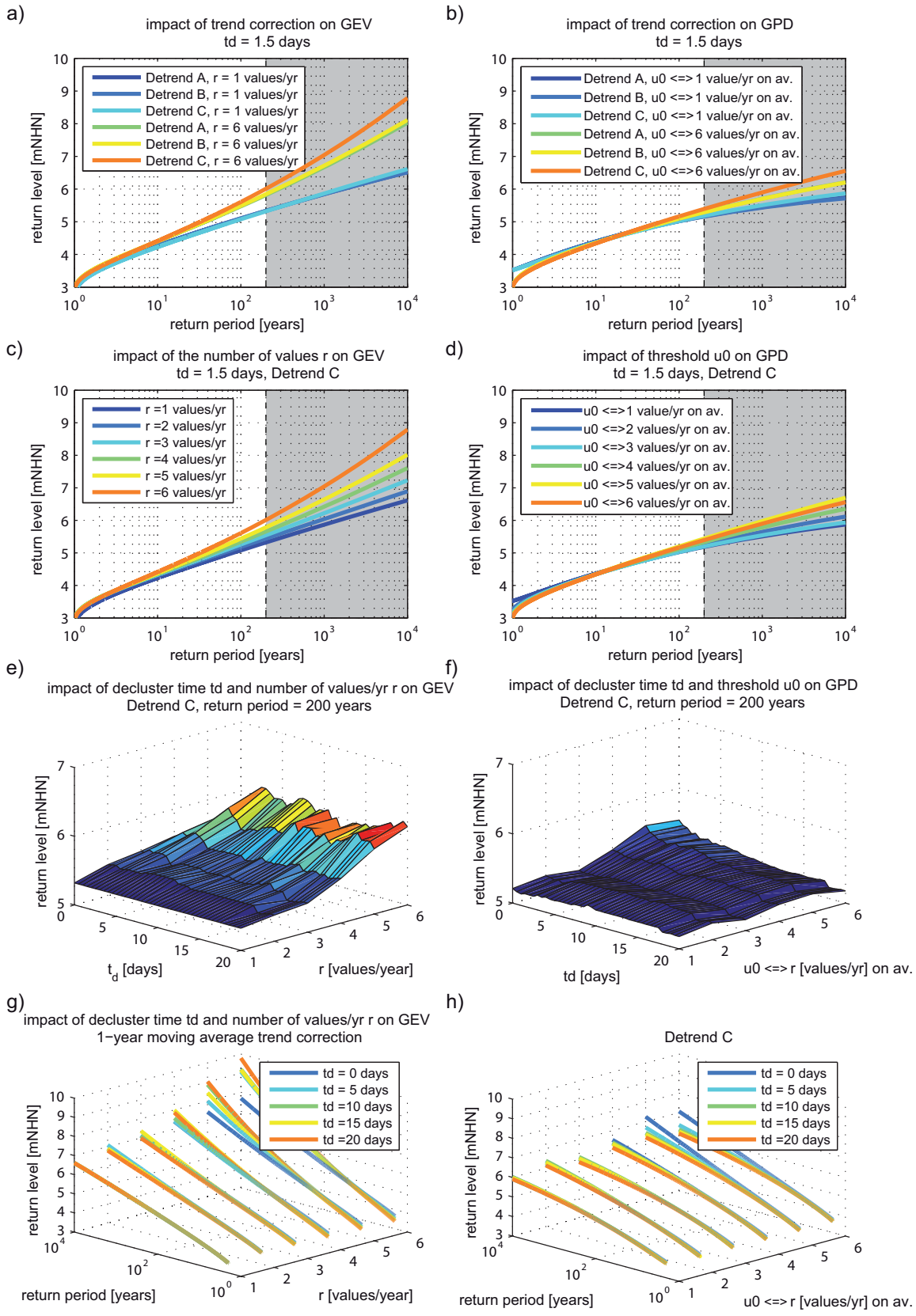


Fig. 13 Influencing factors of the GEV (left) and GPD (right). Note the different scaling of the z-axis (return level) for subplots e and f; Detrend A considers a linear trend correction, Detrend B a 19-year moving average trend correction and Detrend C a one year moving average trend correction; t_d denotes the declustering time in days; u_0 is the threshold level; r describes the number of values per block

confidence intervals (Dixon and Tawn, 1994). At the same time, however, individual events not well captured by the theoretical distribution (GEV) can be included and this can potentially increase the bias. According to Coles (2001), the choice of r is therefore a tradeoff between bias and variance. The individual return water level curves show gradually increasing shape parameters, resulting in a change of sign. From a physical point of view the question arises whether a shape parameter with $\xi > 0$ is eligible to describe the behavior of extreme events, as a progressive curve implies steady accelerating return water levels, reaching no upper limit.

An example of creating a sample of the r-largest values with r = 3 values/year is shown in Fig. 14a, with the resulting sample depicted as blue circles. For illustration purposes, only the time span between 1935 and 1945 is shown. The figure highlights that the r-largest sample does not only comprise extreme values. Between 1937 and 1938, as an example, three relatively small events are selected to be included in the sample. The impact of including these non-extreme values in the sample is shown in the resulting return period plot in Fig. 14b. The calculation of the return water levels is based on the entire data set covering 1918 to 2009. In this case, the GEV is fitted to a sample giving more weight to lower water levels leading to a mismatch in higher water levels. In Cuxhaven, the use of r-largest order statistics with r > 1 value/year can thus lead to a significant overestimation of return water levels.

6.2.2 POT method

To evaluate the impact of the selected POT sample on return water level estimations, the threshold value u_0 is selected in order to consistently match the number of events in the BM derived samples. The samples derived this way are tested for applicability with the GPD using a χ^2 hypothesis test that compares the sample with a reference probability distribution (here the GPD). The influence of the threshold u_0 on the return water level estimates is shown in Fig. 13d. When considering the defined range of thresholds, it is apparent that increasing thresholds lead to decreasing return water level estimates. Generally, an increasing threshold means that the average annual number of values reduces, which is similar to a decreasing number of r-values considered in the BM approach. The discrepancy between the estimated return water levels using different thresholds is nevertheless far smaller than using varying numbers of r in the BM approach.

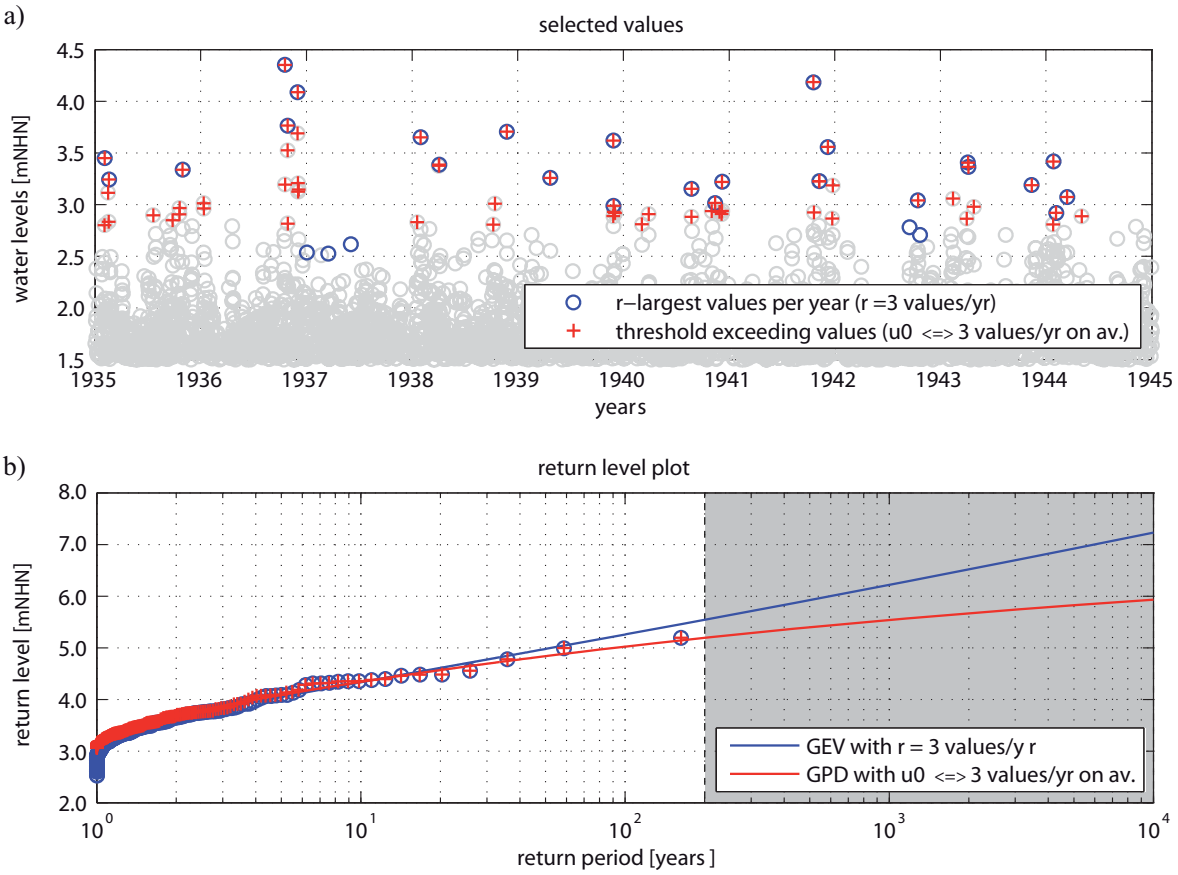


Fig. 14 (a) An example of creating a sample, using either the BM or the POT method, (b) the resulting return level plot for the GEV and GPD

Similar to the BM approach, the uncertainties of the POT method are highly depended on the number of values included in the sample. A low threshold causes a bias in the asymptotic distribution's tail whereas higher thresholds generate fewer excesses and the model is fitted to a data set with a larger variance (Coles, 2001).

The sample is created using the POT approach (red crosses in Fig. 14b) with a threshold leading to the average annual number of values being equal to the annual number used in the r-largest approach in Fig. 14 (i.e. r = 3 values/year). Comparing the two samples derived with the BM approach and the POT approach clearly shows that there are some years in the POT based sample, where no values are taken into account. In the BM derived sample, however, these events are considered as “extreme” events.

The threshold selected above is based on practical considerations. However, it needs to be outlined, that it is important to use objective and stable threshold selection techniques that do not rely on subjective choices. To analyze the performance of different threshold selection techniques, the robustness or stability of the results is tested. The

stability of a feature is defined as the agreement between its results when applied to randomly selected subsamples of the input data (Kuncheva, 2007). The stability is investigated using samples that are steadily reduced by one successive year until the sample reaches a lower limit of 10 years. The resulting return water level estimates of the GPD using different thresholds are then compared to the GEV using the entire Cuxhaven record (i.e. from1918 to 2009). Following this approach it is assumed that the GEV, when fitted to the entire record, yields reliable return water level estimates, hereafter referred to as the 'reference truth'.

For the threshold selection in the POT approach, the following methods are tested: the parametric procedure introduced by Rosbjerg et al. (1992); the non-parametric normality based hypothesis testing approach presented by Thompson et al. (2009); the KS-statistics proposed by Zhang and Ge (2009); and percentile based approaches (e.g. Environment Agency, 2011; Grabemann andWeisse, 2008). Except for the percentile based approach, at almost all considered stations the listed methods were rejected, since the obtained threshold values resulted in considerably less than the 1.65 extreme events per year criterion (see Sect. 5.2) recommended by Cunnane (1973). As the aim is to find a stable and time invariant method, only approaches which produce a smaller sampling variance than the AMAX method are considered.

Another approach is tested in which empirical distributions (or plotting positions) are compared with theoretical distributions, both derived from samples with different thresholds. Samples were generated using exceedances above a range of thresholds that are determined using a χ^2 hypothesis test, detecting all threshold values which create a Pareto distributed sample. Each sample is used to calculate the empirical distributions, using the common Weibull- and Gringorten formulae, as well as the theoretical distribution (GPD). Assuming that the plotting positions represent the "true" exceedance (or non-exceedance) probabilities for a given water level, the root mean square error (RMSE) between plotting positions and theoretical distributions is calculated and used as a measure of consistency between both. The threshold level leading to the lowest RMSE is selected. This approach is referred to as the Plotting Position (PLP) based approach.

The stability of return water levels, based on different threshold selection methods at the tide gauge of Cuxhaven, is shown in Fig. 15. According to the latest policy of the federal state of Schleswig-Holstein, design water levels have to be estimated for a 200-year return period (see Sect. 4). This is why all return water levels are calculated for a return

period of 200 years. The findings presented here are, however, also valid for other return periods. The threshold selection methods considered are the percentile based approaches, with percentiles ranging from the 97.5th (referred to all annual high tides, this is 17.6 values/year on average) to the 99.7th percentile (~2.1 values/year on average), and the PLP based approaches. The grey shaded area represents the 95% confidence bounds of the "reference truth" (i.e. the 200 year return water level derived with the annual maxima method and the GEV for the entire data set). Until 1976, most curves show a relatively stable behavior. The confidence bounds of the reference truth are exceeded before 1976 for some of the percentiles derived with the PLP approach. In 1976, all curves exhibit a strong decrease, resulting in a distinctively smaller return water level and then vary diversely after this. The figure illustrates the importance of the 1976 storm surge event for the estimation of return water levels along the German coastline. This is reasonable because in large parts of the German Bight, the 1976 event resulted in the highest water levels on record to date

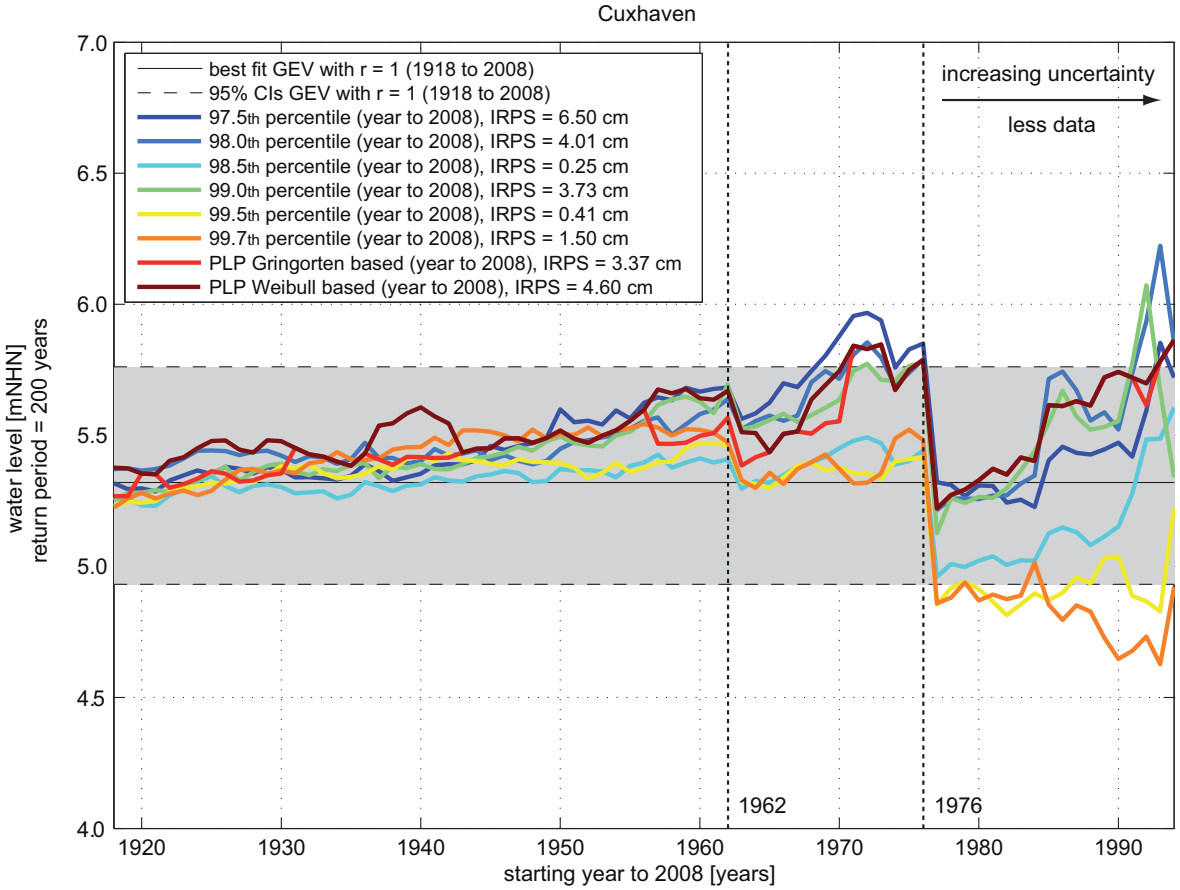


Fig. 15 Stability of GPD estimates using different threshold selection methods at Cuxhaven station. The grey shaded area shows the 95% confidence bounds of the reference truth

(Jensen and Müller-Navarra, 2008; see Fig. 56 in the B Appendix for a detailed illustration of the occurrence of the highest observed high waters along the entire German Bight between 1970 and 2009). Also the impact of the 1962 event is visible but depending on the threshold level, the impact is less than from the 1976 event.

To objectively compare the performance of the investigated threshold selection techniques, the Index of Return Period Stability (IRPS) is introduced in this thesis. This index is defined as

$$IRPS = s^2 \cdot \max. \Delta h \cdot \bar{h} \quad , \qquad \text{Equation 7}$$

saying that the IRPS is the product of the sample variance s^2 of the estimated return water level curve, its maximum distance to the reference truth max. Δh and the mean distance to the reference truth $\bar{h}$ (see also Fig. 16).

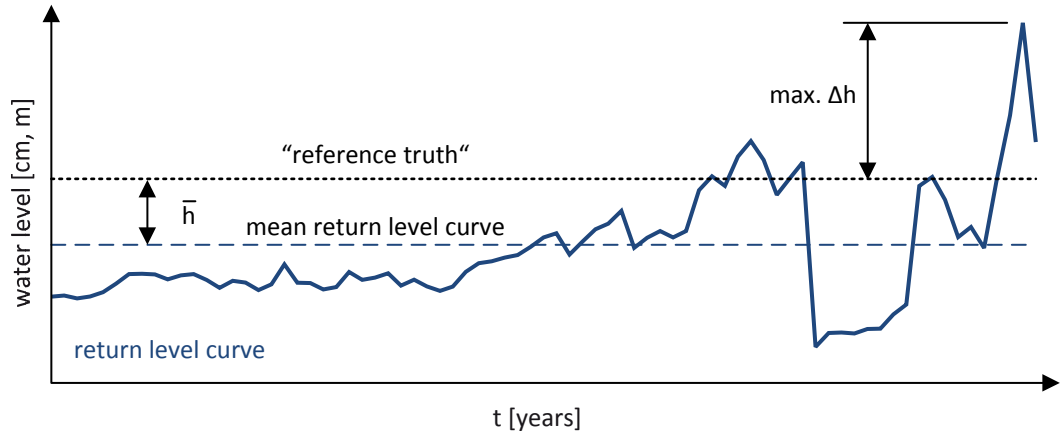


Fig. 16 Illustration of the calculation of the Index of Return Period Stability (IRPS)

In Fig. 17a, the time span considered is 1918 to 1976, as estimations after 1976 are assumed to not be reliable for the reasons mentioned above. In the case of Cuxhaven, the return water level curve based on the 98.5th percentile leads to the smallest IRPS; other percentiles as for example the 99.5th and the 99.7th percentiles also lead to relatively small IRPS values and are thus suitable to be chosen as threshold values. As many tide gauges are subjected to local influences, the transferability of these findings is tested for an additional nine water level records from tide gauges along the German Bight as well as three international tide gauge records (see Sect. 3.2).

The IRPS values for all stations (Fig. 17a) highlight that lower percentiles lead to higher IRPS values whereas higher thresholds, especially the 99.7th percentile, result in relatively small IRPS values. In all cases, the IRPS value using the 99.7th percentile is at least among the two lowest, except for Büsum where it causes the 3rd lowest IRPS value. An average IRPS value, calculated as the mean value across all stations, is shown in Fig. 17b. The average value using the 99.7th percentile has the lowest IRPS value of all considered thresholds. The second lowest IRPS value is achieved by using the 99.5th percentile which, however, leads to much higher IRPS values.

6.2.3 Declustering

To investigate the influence of declustering on the estimation of return water levels, different declustering times (see Sect. 5.2) ranging from $t_d = 0$ to $t_d = 20$ days are applied to the samples derived with the BM and the POT methods. Fig. 13e and f shows the results for the 200-year event. Fig. 13g and h displays the entire return period curves (covering return periods up to 10,000 years). The impact of the declustering time on the 200 year level using different r-values is shown in Fig. 13e. For r = 1 value/year, the return water

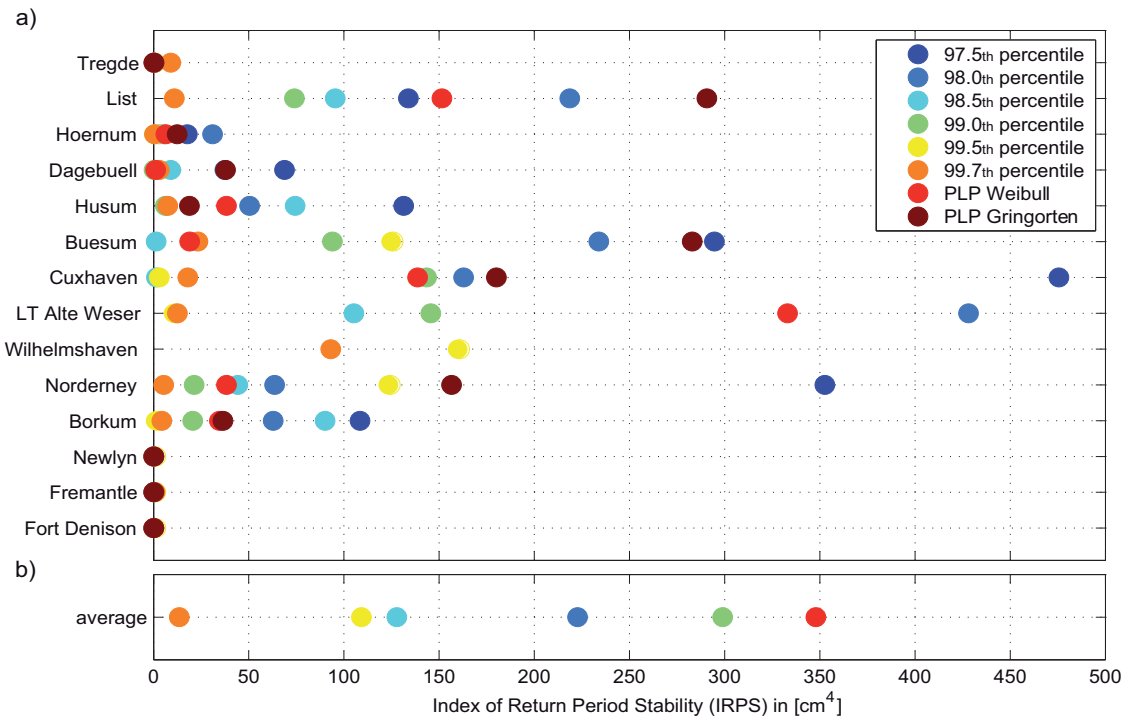


Fig. 17 (a) IRPS of all considered stations depicted as colored points for different threshold values (at Cuxhaven, the value of the 98.5th percentile is covered by the value of the 99.5th percentile and therefore not visible), (b) mean IRPS values of all stations

levels are hardly affected by the declustering time. With increasing r-values, however, both the overall height of the return water levels and the influence of the declustering time increase, resulting in a larger variability with no distinct pattern. To examine the influence of the declustering time on a wider range of return periods, Fig. 13g shows the impact for return periods ranging from one to 10,000 years. Again, the calculation of return periods using r = 1 value/year is not (or negligibly) affected by the declustering time, whereas especially for higher values of r, return water levels strongly depend on the declustering time t_d.

Results from equivalent analyses for the POT approach are shown in Fig. 13f and h. Postulating the same sample size as in the BM approach, return water level estimates in the POT approach show less dependency on the threshold values (see Fig. 13f). They show, however, a stronger dependency on the declustering time, leading to significant overestimation if the declustering time is short (t_d < 1 days). Especially using a short time span t_d for declustering and a relatively low threshold u_0 leads to a significant overestimation of the return water levels. If the threshold u_0 and the declustering time t_d exceed a certain value (here u_0 equals a water level resulting in a sample of 4 values/year on average, and t_d is > 1 days), the resulting return water levels show almost no variability. The impact of the declustering time t_d (see Sect. 5.2) and that from using different threshold values u_0 for return periods from one to 10,000 years are shown in Fig. 13h and indicate (similar to the BM approach) that the return water levels depend on the declustering time considered. Nevertheless, the range of possible outcomes for one and the same return period using different declustering times is by far smaller than in the BM approach.

The sensitivity of both methods to the declustering time is determined by the dependency between adjacent events in the samples. With high thresholds or low r-values, the events within a sample are rare and obviously not connected. For low thresholds or higher r-values, the likelihood of two neighboring values to be dependent is much higher, which might lead to a sample that violates the IID criteria (see Sect. 5.1).

6.3 Distribution

As pointed out by Hawkes et al. (2008), the choice of a proper distribution function should be guided not only by a goodness-of-fit test but also by the robustness of the fit. The performance of the GEV and the GPD is therefore tested by focusing on the

robustness and stability of the particular distribution. The stability of both methods is evaluated using a water level timeseries that is steadily reduced by one successive year. The last year included is 2009 with the starting year steadily increasing from 1918 to 1998, until the sample reaches a lower limit of 10 years in length. In all cases, the return water level estimates with a return period of 200 years and the associated confidence intervals (CIs) are calculated and plotted against the considered starting time. To analyze the stability of the BM method, r-largest values ranging from r = 1 to r = 6 values/year are used to create the samples (see the blue curves in Fig. 18a–f). The stability of the POT

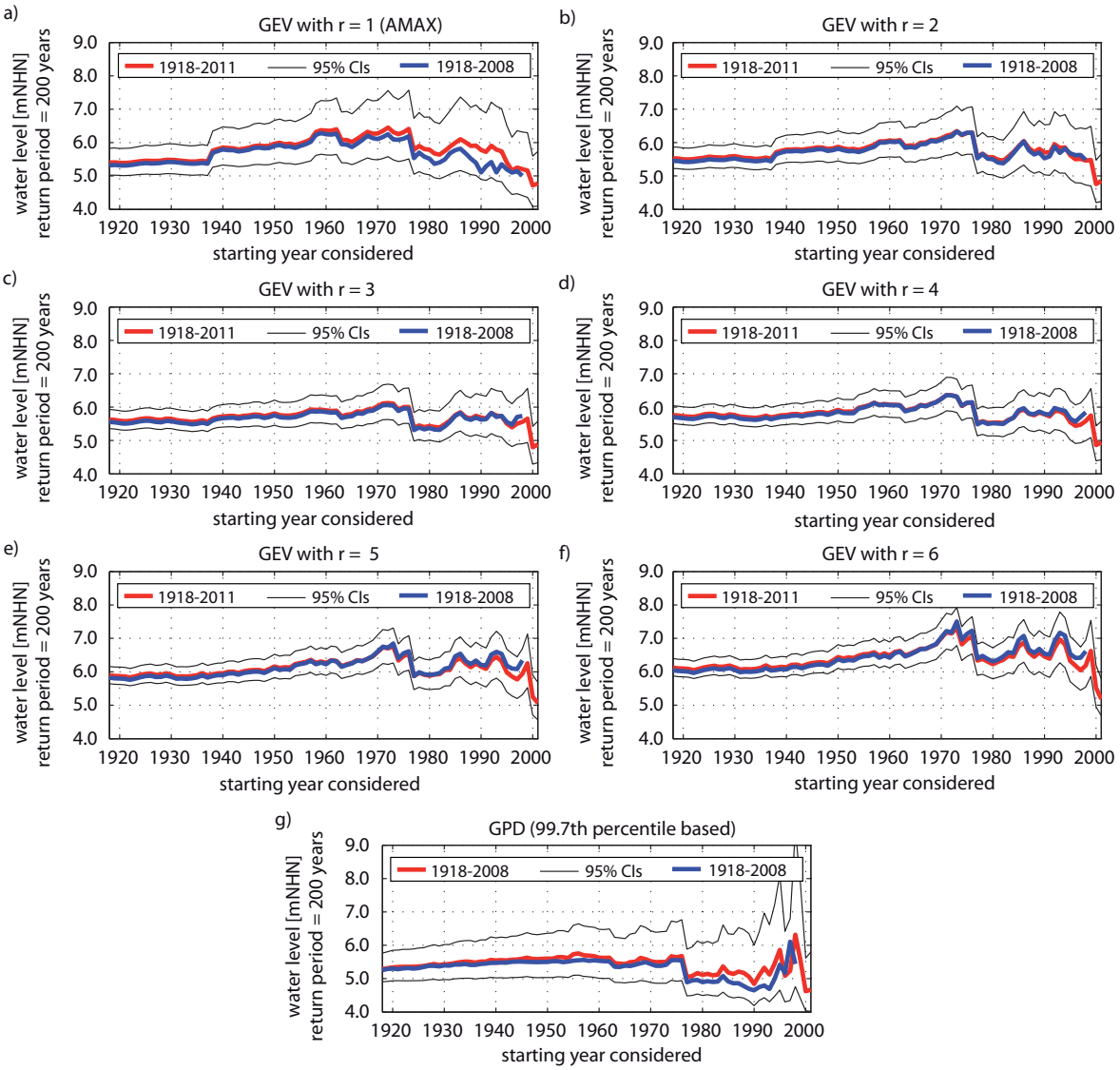


Fig. 18 (a) -(f) Stability of the GEV using r = 1to r = 6 values/year, (g) stability of the GPD using a threshold of u0 = 99.7th percentile. The blue curve shows the period from 1918 to 2009; the red curve shows the updated record covering the period from 1918 to 2011

method is analyzed by applying the 99.7th percentile based threshold, which was identified to be most appropriate for the tide gauges considered here (see Sect. 6.2.2).

As already mentioned above, the GEV with r = 1 value/year is stable when a long record is used. This behavior changes from 1938 onwards, when the GEV derived return water level estimates begin to stagger, with large discrepancies of up to 0.9 m in the resulting return water levels. A similar behavior is observed for the GEV with r = 2 to r = 6 values/year (see the blue curves in Fig. 18b to f) and for all stations considered along the German coastline (results not shown). To obtain reliable and stable return water level estimates for the German Bight using the GEV, the use of datasets which start in 1937 or earlier is recommended. In this thesis, the GEV derived return water levels for the period from 1918 to 2009 are considered as "reference truth".

Comparing the CIs of the r-largest cases highlights that the CIs narrow with an increasing number of values included. The mean value of the estimated return water levels also increases with the number of r-values, leading to the largest absolute estimates when using $r = 6$ values/year. The use of the GEV with $r > 1$ value/year thus seems to overestimate return water levels. The stability of the GPD indicates that, in contrast to all cases of the GEV, the GPD leads to very stable return water level estimates until the starting year of the considered fraction of the time series is in 1977 (see the blue curve in Fig. 18g). Using a sample that does not include 1976s values creates unstable results leading to lower return water level estimates. With the starting year in 1997 or later, return water levels increase again.

To check the validity of the above findings with recent water level records, the Cuxhaven data set was updated by including the water levels of the years 2010 to 2011. The updated record was used to create the samples using exactly the same approach as described above. In Fig. 18, the return level estimates of the 200-year event based on the updated water level record are shown as red curves. The results highlight that the water level update does hardly affect the return water levels of all approaches, if the 1976 storm surge event is included. Neglecting this event causes discrepancies in the $r = 1$ largest/values and the POT approach highlighting how important it is to include the 1976 event in a return level assessment. On the other hand, it also highlights that the period 1976 (or earlier) to 2009 is currently long enough to obtain reliable return level estimates. This, however, needs to be checked periodically or after the occurrence of intense storm surges.

7 Transferability

In the last section, it was shown that the GPD leads to more stable return water level estimates than the GEV for the Cuxhaven record. To validate this hypothesis for other stations, a further 13 tide gauge datasets (see Tab. 1) are analyzed. As described in Sect. 6.2.1, the use of the GEV with $r > 1$ value/year tends to overestimate return water levels when applied to the Cuxhaven tide gauge record. In order to investigate the transferability of the different methods, only $r = 1$ value/year for the GEV are considered. In the POT approach, the 99.7^{th} percentile is used for an automated threshold selection. All return water levels are calculated for a return period of 200 years.

Results from applying the GEV and the GPD are shown in Fig. 19, for all 14 sites, using the model set-up specified in Sect. 6. At all German tide gauge sites, the findings are consistent with the Cuxhaven site, where the GPD performs much more stable than the GEV. Only the return water level estimates at Norderney and Borkum (Fig. 19i and j) show a slight tendency to decrease after the dataset is reduced to 1960 or later, with a magnitude at a maximum of 0.1 m. This magnitude is, however, far smaller than the one resulting from the GEV, which is for these two cases of the order of up to 1.0 m. As with the Cuxhaven record, all other German datasets show good agreement between the GEV and GPD derived return water level estimates up until 1938. Afterwards, the GEV derived return water level estimates begin to fluctuate, causing large discrepancies between the GEV and the GPD. The findings using the Cuxhaven dataset can thus be confirmed for nine other tide gauge sites in the German Bight.

The results for Newlyn (model set-up is the same as before) yield very stable results for both GEV and GPD with only small fluctuations and negligible differences between the two models (Fig. 19k). Dixon and Tawn (1999) stated that direct methods tend to underestimate return water levels at sites where the non-tidal variation is small compared to the variation in astronomical tidal levels. The ratio of tidal to non-tidal variability at

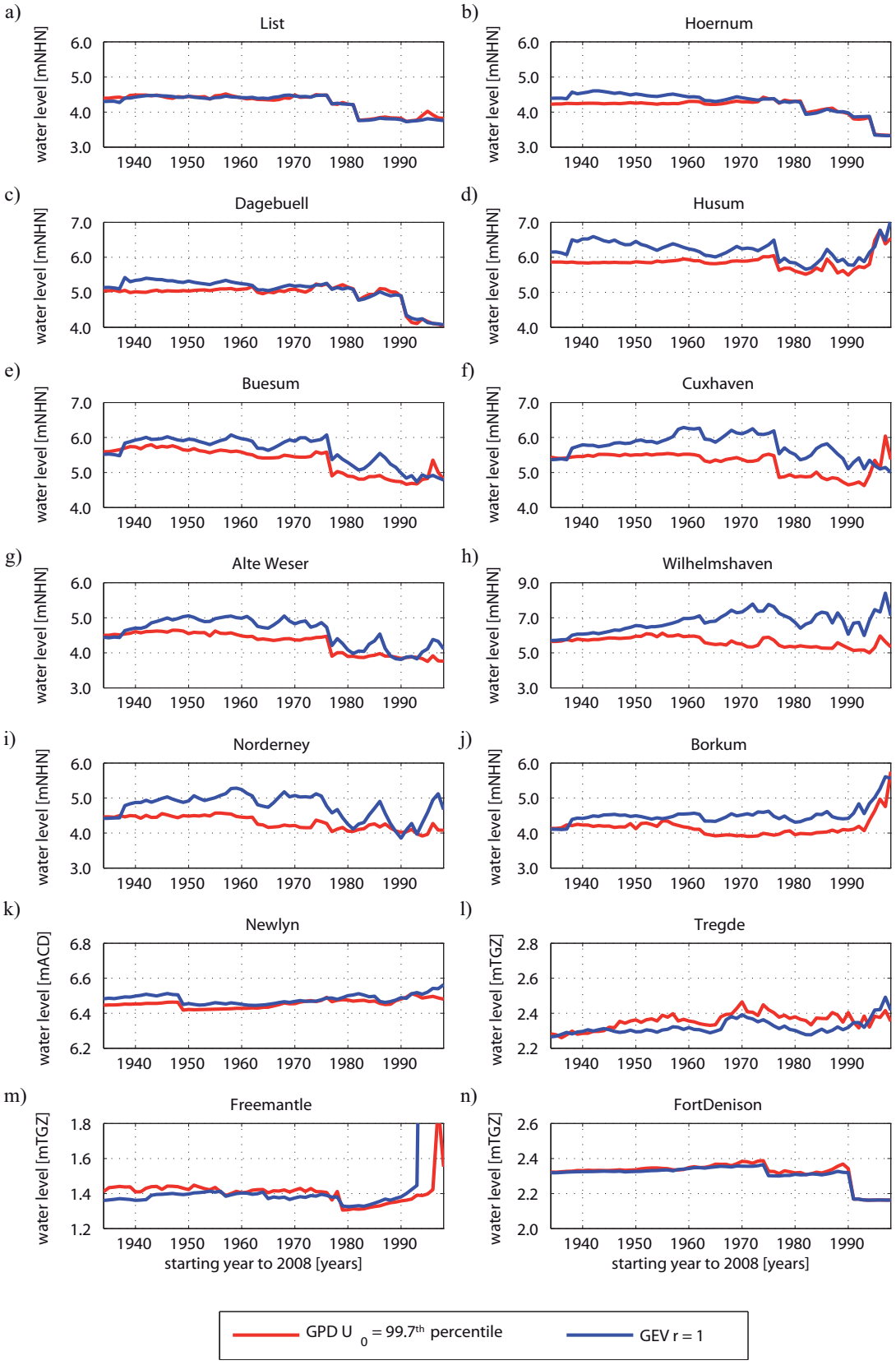


Fig. 19 Results of the GEV with r = 1 value/year and the GPD with a threshold at the 99.7th percentile of 14 tide gauge records. All return water levels are estimated using steadily reducing datasets. Note the different scaling on the y-axes

Newlyn is 8.2 [-] (i.e. the highest ratio considered in this study). In terms of the eligibility of the direct methods it is thus highly debatable if the Newlyn dataset can be used in combination with the GEV and the GPD, as extreme value theory assumes a random process and the Newlyn data set mainly consists of deterministic (tidal) parts.

The application of the GEV and the GPD to the Tregde dataset, which is dominated by non-tidal components resulting in a tide/surge ratio of 0.5 [-] (i.e. the smallest ratio considered in this study), indicates that both GEV and the GPD derived return water level estimates are stable for each time step (Fig. 19l). Only the last few years show slightly higher variability. Deviations between the GEV and the GPD results are negligible.

For the Australian sites, Fremantle and Fort Denison with tide/surge ratios of 1.1 [-] and 3.5 [-], respectively, there are no significant differences between the results of the GEV and the GPD (Fig. 19m and n). Both methods show stable return water level estimates with negligible differences. For the Fort Denison dataset, however, both methods show a sudden drop in the return water levels when the sample begins in 1990 or later. This may be a result of the shorter remaining time span considering that the Fort Denison dataset ends in 2004 (see Tab. 1). In Fremantle by contrast, there is a sudden increase at the same time, leading too much higher return water levels. This is most probably a result of the small sample size considered, as the combination of the threshold selection and the declustering time of $t_d = 1.5$ days results in a sample that includes less than 0.9 values/year on average. Both examples show that using too short datasets yields large uncertainties in return water level estimates. In contrast, results also highlighted that using 30-years of data could be as accurate as the results from using 100-years of data (or more), provided that the model set-up is appropriately chosen.

In order to objectively assess the stability and eligibility of the evaluated methods, the IRPS is used again. Fig. 20 shows the IRPS values for the GEV with r = 1 value/year (denoted as red dots) and for the GPD with $u_0 = 99.7^{th}$ percentile (denoted as blue dots) for each of the considered sites. For all German sites, the GPD yields a much smaller IRPS than the GEV. This is also valid for Newlyn and Fremantle but with negligible differences (note the log scaled abscissa). At Fort Denison and Tregde, the GEV leads to an IRPS that is slightly below the one for the GPD, but these differences are small. The averaged values across all tide gauge sites (lower part of Fig. 20) confirm the results for most of the individual stations, with the GEV having a higher IRPS than the GPD. The large difference of the average IRPS between the two methods is partly caused by the results for

Wilhelmshaven, as the application of the GEV causes much larger fluctuations at this station than the GPD ($GEV_{IRPS} = 7.36 \times 10^7$, $GPD_{IRPS} = 1.74 \times 10^5$). Excluding the results of Wilhelmshaven reduces the average IRPS of the BM method from $GEV_{IRPS} = 6.20 \times 10^6$ to $GEV_{IRPS} = 1.01 \times 10^6$ and the average IRPS of the POT method from $GPD_{IRPS} = 2.67 \times 10^4$ to $GPD_{IRPS} = 1.54 \times 10^4$, still showing that the GPD leads to more stable results than the GEV.

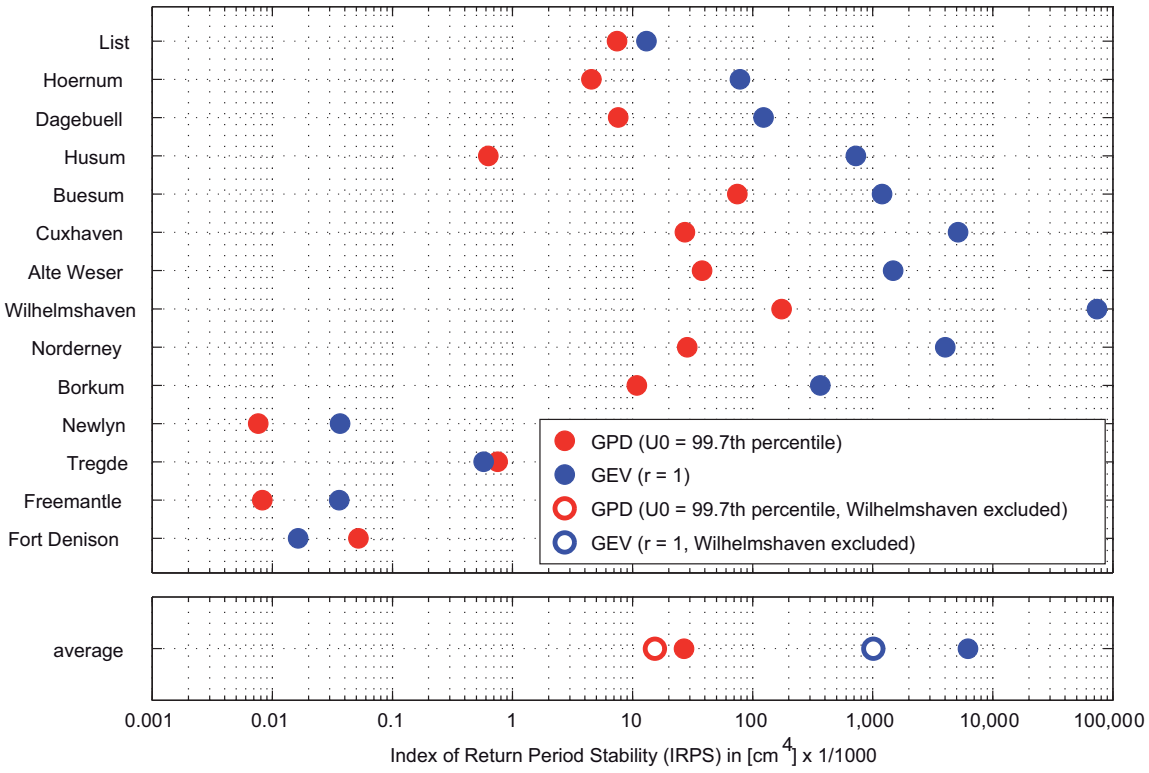


Fig. 20 IRPS of all considered tide gauges records using the entire time period available at the individual stations (see Tab. 1). The calculations are based on the curves of Fig. 19, using the results of the GEV with r = 1 value/year as “reference truth”. Note the log scaled abscissa

8 Summary and discussion

One of the main objectives of Part I was to determine how sensitive each of the direct methods was to three factors, which can significantly influence the results. First the influence of using different detrending methods in order to yield a stationary time series for the extreme value statistics were analyzed. Therefore, three different detrending approaches were used. By comparing the BM method and the POT method, it was found that the BM method is more sensitive to the trend correction than the POT method. Following Hawkes et al. (2008), the use of a trend derived from a 1-year moving average for the correction in order to create a stationary dataset is recommended, as this approach is the only one accounting for the influence of the considerable seasonal variability in sea level time series. For large parts of the world, it is valid to use MSL derived trends to correct the data before doing the statistical analyses. In the German Bight, however, observed maxima have risen faster than the MSL over the last 60 years. Therefore it is recommended using trends derived from high water peaks for this area (and other areas where MSL changes differed from MHW changes in the past). For the other stations considered here, it did not make any difference whether to use MHW or MSL derived trends for the correction.

Second, the influence of using different techniques to generate a sample of extreme values was investigated. To investigate the performance of the BM method, a range of different samples with r = 1 to 6 values/year were created. The results from using different r-values were noticeably different, showing the sensitivity of the statistical assessment to the extreme sample used. In comparison, the POT method using a range of thresholds that lead to a sample matching in size the BM derived samples was investigated. The discrepancy among the estimated return water levels using appropriate thresholds was far smaller than using the different BM samples.

In terms of threshold selection, for the POT approach, the main purpose of the study consisted in outlining the importance of using objective and stable threshold

modeling techniques that do not rely on subjective choice. As there is no comprehensive guideline available detailing how to select an appropriate threshold, a broad range of threshold selection methods were analyzed. The analyses showed that the use of the 99.7th percentile leads to the most stable return water level estimates along the German Bight, where the tide/surge ratio ranges from 1.4 [-], at the most northern point, to 2.6 [-] at the most western point. This was confirmed by transferring these analyses to four international tide gauge datasets, with tide surge ratios ranging from 0.5 [-] in Tregde (NOR) to 8.2 [-] in Newlyn (GB). However, at these four international tide gauges, the differences between using the BM or POT approach were much smaller.

Often, extreme sea levels do not appear randomly dispersed in time, but exhibit clusters, with a higher density of sampling points in some periods. For estimating return water levels, it is therefore important to decluster extreme value samples. Here it was shown that the selection of the declustering time distinctively influences the results. However, the objective selection of independence criteria is a very complex problem as the decision whether two events are independent or not is often subjective. To date, there are no physically based criteria available to calculate declustering times objectively. However, to obtain consistent and reproducible results, it is suggested to use the reciprocal of the extremal index as it is objective and well-recognized. Applying the extremal index to all examples resulted in a mean declustering time of $t_d \approx 1.5$ days for each individual station. This is why a fixed value of $t_d = 1.5$ days is considered for all analyses except the analysis of the effect of the declustering itself.

The third factor under examination was the sensitivity and eligibility of the GEV and the GPD. The aim was to establish an automated model set-up that did require any steps involving subjective choices. To assess the stability of the results from the GEV and the GPD, the sensitivity of both distributions when steadily reducing the datasets lengths was investigated. When considering samples starting in 1937 or earlier, the GEV led to stable and reliable return water level estimates along the German Bight. Using samples covering a shorter time span led to unstable results. The results from the GPD, in contrast, were very stable. Using a 99.7th percentile derived threshold for analyzing the Cuxhaven record yielded negligible differences considering any of the starting years between 1918 and 1976. In the German Bight, this is up to 40-years less than what is recommended for the BM approach (i.e. at least data from 1937 onwards). The international stations were less susceptible to the chosen distribution, with both distributions showing relatively stable

results. For all stations considered in Part I, it was shown that when using the GPD, around 30-years of data (in the German Bight since 1976 or earlier) can be as valuable as 100-years if the model is properly set-up.

9 Key findings of Part I

The overall aim of Part I was to provide guidance for estimating extreme still water levels (with little subjectivity) addressing coastal engineers, managers and planners. Based on the sensitivity tests, the recommended approach is as follows (and is summarized in Fig. 21).

- *Use a high water peak time series starting in 1976 or earlier as input data.*
- *Create a stationary dataset using a 1-year moving average trend correction of the high water peaks.*
- *Calculate a threshold using the 99.7th percentile of the high water peaks.*
- *Use the extremal index for declustering.*
- *Fit the GPD to the extreme value sample.*

These recommendations are valid for the German Bight as well as for the international stations considered in the study (the first point is relevant to just the German Bight), but would benefit from being adapted (e.g. finding the required time series length) and verified further for other locations around the world (see also Arns et al., 2013a).

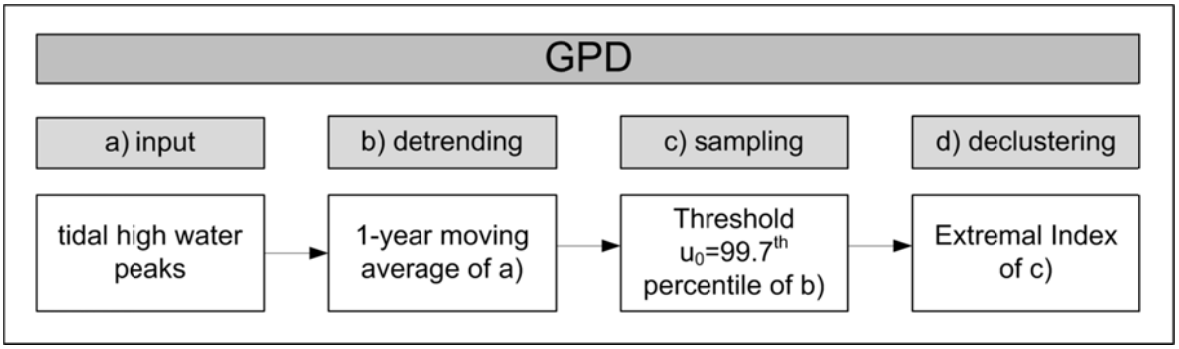


Fig. 21 Recommended procedure for calculating return water levels from time series of observed total water levels

PART II

HOW TO ESTIMATE RETURN WATER LEVELS IN UN-GAUGED AREAS?

10 Motivation

The coastline of the federal state of Schleswig-Holstein located in the northern part of the German Bight (see Fig. 6d), is protected by coastal defense structures which have been built to withstand extreme still water levels with an exceedance probability of $P_E = 0.0005$ (i.e. a 200- year event). Accurately calculating the associated return water levels using traditional extreme value analysis methods requires records of sufficient length (> 30 years; Haigh et al., 2010b), indicating the largest pitfall of this approach, as the availability of measured water levels is limited in many regions. In the German Bight, multi-decadal records of high and low waters exist at several sites, but for some regions (e.g. at some small islands in the German Wadden Sea) no or only very short and incomplete water level measurements exist. As water levels in the German Bight are strongly influenced by shallow water effects and the complex topography of the coastline, they can differ significantly between stations (see e.g. Jensen and Müller-Navarra, 2008). In such cases it is difficult to convey information about the likelihood of extreme water level events from gauged (local scale) to surrounding un-gauged areas (local to regional scale).

One way of working around this problem is to use the regional flood frequency analysis (RFA) approach. This technique is based on regional homogeneity, assuming that samples from individual sites exhibit similar spatial and temporal statistical characteristics over a larger area and can be described by a common regional distribution (Rao and Hamed, 2000). Quantiles can then be estimated by transferring this distribution to locations that are within the assigned region. The RFA approach has widely been applied in hydrology, where river catchment attributes and spatial proximity are used as a measure to decide which information can be appropriately transferred from the catchment to a particular site of interest. This concept is based on the assumption that catchments with similar attributes behave in a similar manner in terms of flood frequency response (Merz and Blöschl, 2005).

An alternative approach is to apply extreme value analyses methods to water level data sets derived from hydrodynamic numerical model simulations. Similar to at-site (i.e. the direct analysis of tide gauge records from a particular site) analyses, return water levels are calculated for each individual grid point along the coastline and islands within the model domain. Such a methodology was recently applied successfully to the coastlines of the UK. On behalf of the Environment Agency (EA), Dixon and Tawn (1994, 1995, 1997) provided a single coherent estimate of extreme still water level probabilities at high resolution all around the UK coastline using their Spatially Revised Joint Probability Method which was based on both tide gauge data and a multi-decadal predicted water level hindcast. A major update of that study has recently been completed (Batstone et al., 2013; Environment Agency, 2011), which improved the basic statistical assumptions (resulting in the Skew Surge Joint Probability Method) and used longer tide gauge records that are now available. A similar study has recently been completed for Australia that provided a consistent estimate of the probabilities of extreme water levels at high resolution all around the Australian coastline (see Haigh et al., 2013a, 2013b) and is freely available for coastal engineers, managers and planners via a web-based tool (www.sealevelrise.info). This shows that estimates of extreme water level probabilities are starting to be calculated systematically at high resolution all around the coastline of countries (e.g. the UK and Australia).

One of the above mentioned approaches could be used to overcome the pitfall of traditional extreme value analysis methods requiring a certain period of input data. Overall, Part II of this thesis has the following two objectives:

(1) To test the applicability of the RFA at a coastal setting.

(2) To develop an alternative approach to determine return periods of extreme still water levels for areas where few and short, or no, water level measurements exist.

Both of these approaches (RFA, numerical model based) are tested for applicability. As a case study, this is conducted for the coastline of Schleswig-Holstein located in the northern part of the German Wadden Sea (see Sect. 2).

11 Regionalization

11.1 Principle of the method

The estimation of heights and occurrence probabilities of extreme events such as floods or storm surges is typically conducted using a limited sample of the considered variable, inferring the distribution of the entire population. By assuming that the sample is an independent realization of the overall population, it can be used to estimate the unknown statistical parameters of the population (see e.g. Coles, 2001). However, in practice, hydrological data is often limited in both space and time yielding imperfect parameter estimates and as a result unrealistic occurrence probabilities (Rao and Hamed, 2000). The availability of sufficient data is thus one of the crucial limitations when performing statistical analyses. This is why RFA methods have been developed. They enable the indirect estimation of occurrence probabilities. Such methods, first applied in hydrology, are based on the assumption that river catchments with similar attributes behave similar in flood frequency response. They thus compensate for the lack of data at individual stations (Stedinger et al., 1993) by transferring hydrological information from gauged to related un-gauged sites.

The regionalization of statistically derived design floods was pioneered by Dalrymple (1960), who merged data from different stations of a region into a unified probability model. The concept is based on the assumption that essential differences between distributions of individual sites within a homogeneous region (in a statistical sense) are only found in a scaling factor, called the '*index-flood*' μ_i (e.g. mean high water, but also any other parameter may be used). This is why the first step in RFA is to identify a homogeneous region (e.g. the region described by the black rectangle in Fig. 22a) for which the flood frequency can be approximately described by a single (regional) distribution that is representative for all sites (N) located in the region (Hosking and Wallis, 1993; Rao and Hamed, 2000; see the black curve in Fig. 22b). The local

distributions $Q_i(F)$ at individual sites i = 1,…,N may then be calculated by multiplying the regional distribution q(F), with 0 < F < 1, and the index-flood μ_i according to

$$Q_i(F) = \mu_i q(F) \quad . \qquad \text{Equation 8}$$

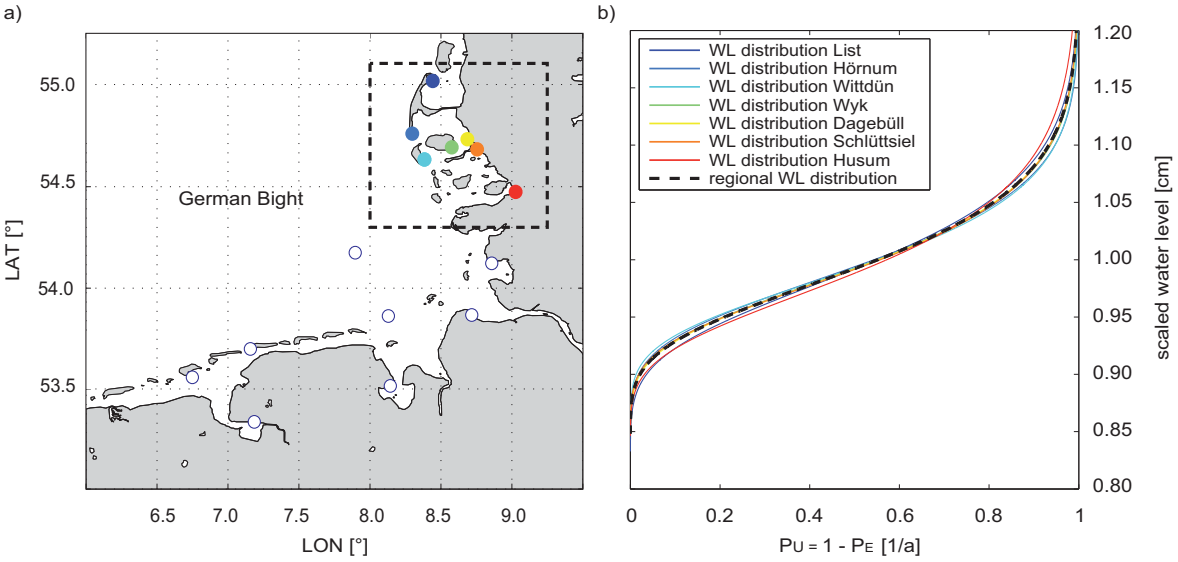


Fig. 22 Principle of the RFA method with a) the locations of the individual stations (colored dots in the black rectangle) and b) the water level (WL) distributions at those stations scaled by the average value of each data set

The RFA essentially pursues two objectives. The first objective is to enlarge the data basis in gauged areas in order to enhance the precision of flood estimates in the study area. Provided that the considered records are from the same distribution, samples from the joint use of measured at-site data using a number of stations can yield more robust parameter estimates. Using this kind of regionalization represents a substitution between space and time as different long records within a region are used to compensate shorter records (Rao and Hamed, 2000). With respect to practical applications, this concept does not necessarily need to define boundaries between regions but rather includes sites that are similar to the site where information is to be transferred to. In a mathematical sense, extreme value samples are considered to describe a random process, comprising independent and identically distributed (IID) random variables. However, water level datasets can exhibit dependencies (so called clusters), which are mostly related to the same meteorological forcing. This has to be taken into account when performing RFA of coastal data sets. The second objective is to generate information for un-gauged sites. Where information is spatially limited (i.e. little or no data is available in a specific area),

regionalization methods can be used to infer hydrologic information from one site (or region) to another. A similarity measure is used to decide which information is to be transferred to the site of interest (Merz and Blöschl, 2005).

The RFA has most often been applied to riverine areas (see e.g. Wiltshire, 1985) and there have only been few studies that adjusted this concept to a coastal setting. Van Gelder and Nykov (1998) tested the RFA along the North Sea coast of the Netherlands. They concluded that sites along the Dutch coastline do not form a homogeneous region and recommended including sites from surrounding countries instead. Mai et al. (2006) investigated the use of RFA for the prediction of extreme waves along the Dutch and extreme storm surges along the Vietnamese coastlines. Aiming at unbiased extreme estimates, they used the concept of RFA to extend samples of individual stations by using the entire information within a homogeneous region. They concluded that RFA yields a more accurate prediction of extreme quantiles but does not account for dependencies between sites. Bardet et al. (2011) conducted a RFA for surges derived from 21 tide gauge records along the French Atlantic and English Channel coastlines. By merging all available data sets they constructed a regional sample covering an effective duration of 601 years. They found the RFA to lead to more reliable results for some sites than at-site analyses (i.e. more robust parameter estimates) but also point to the limitations of these findings linked to the dependency between the individual events within the sample arising from the same forcing (e.g. storm). A study similar in methodology and study area but using only 18 tide gauge records was by Bernadara et al. (2011). They concluded that the RFA is generally applicable to surge data sets and may help to overcome the drawbacks in at-site analyses (i.e. reduced uncertainties). However, they also point to the need of validating these findings and suggest using numerical model simulations for that purpose. For the German Bight, there is no published study available dealing with RFA and its application to coastal water level records. Hence, the most common RFA approach is adapted and its application to the German Bight is investigated.

11.2 Identification of homogeneous regions

The identification of homogeneous regions is usually based on some sort of similarity measure. From its origin, the region under consideration is usually a catchment, postulating a relationship between catchment attributes and hydrological processes. This is why similarity was traditionally founded on spatial proximity as climate and catchment

attributes (e.g. size, geology, climate conditions) are likely to only vary smoothly in space (Merz and Blöschl, 2005). However, hydrologic variables may also reveal small scale variability whereas catchments that are far apart may still be similar (see e.g. Pilgrim, 1983). In RFA, different similarity measures can be used including multiple regressions of flood quantiles, moments or catchment attributes as well as the pooling of catchments or stations into homogeneous groups (Merz and Blöschl, 2005). However, with respect to coastal waters it is difficult to define enclosed catchments appropriate for the use in regional flood frequency analyses as distinct boundaries do not exist. This is why this thesis focuses on approaches that are based on pooling individual stations into homogeneous groups, where the stations have a similar distribution (Hosking and Wallis, 1993).

In the literature, several analytical approaches for testing regional homogeneity have been proposed, for example: the heterogeneity measure H (Hosking and Wallis, 1993; 1997); the discordancy measure $D_{i,}$ (see e.g. Hosking and Wallis, 1993; Rao and Hamed, 2000); or the Wiltshire method (Wiltshire, 1986; for a review and discussion of different methods, see e.g. Viglione et al., 2006; Castellarin et al., 2008). This study uses an assessment of L-moment dispersion as it was suggested by Hosking and Wallis (1993). This method compares standardized L-moments of individual stations identifying those distinctly discordant with the entire group by plotting each stations L-Cv (scale) vs. L-skewness (see Hosking, 1990) and constructing concentric ellipses (one and two times the standard deviation), with discordant stations being those outside the outer ellipse (Hosking and Wallis, 1997). The same method has already been used by several other studies that applied RFA to coastal data sets (e.g. van Gelder and Nykov, 1998; Mai et al., 2006; Bardet et al., 2011).

This part of the thesis considers a total number of 15 stations along the German North Sea coastline, 9 are located in the federal state of Schleswig-Holstein (SH) and the remainder in the federal state of Lower-Saxony (LS) (see Tab. 2). In a first attempt, all 15 stations were considered and it was assumed that they form one single homogeneous region (referred to as Attempt A). This region, however, was statistically heterogeneous causing large discrepancies (in terms of the root mean squared error (RMSE) and maximum differences in return levels referred to as Δ max.) when compared to at-site analyses (see Tab. 2). It was therefore decided to assign the stations to two different regions with each region consisting of at least 4 stations. Following these assumptions, all

possible combinations were investigated, each of which contains n_i elements of the given set differing from one another by at least one element. The number of possible combinations c_p can then be calculated as:

$$c_p = \frac{n!}{(n-m)! * m!} \quad , \qquad \text{Equation 9}$$

where n is the total number of stations and m the number of stations to be drawn without repetitions. In this study, this amounted to $c_p = 15.808$ possible combinations and nearly half of those ($c_h = 7.678$) were found to be homogeneous. The homogeneity assessment is exemplarily highlighted in Fig. 23, showing the 1σ and 2σ (referred to multiples of the standard deviation) error ellipses around the L-moment dispersion in Regions I and II. In this particular case, Region I is constructed using all nine stations within the federal state of Schleswig-Holstein whereas Region II consists of all six stations located in the federal state of Lower-Saxony (referred to as Attempt B; see Tab. 2).

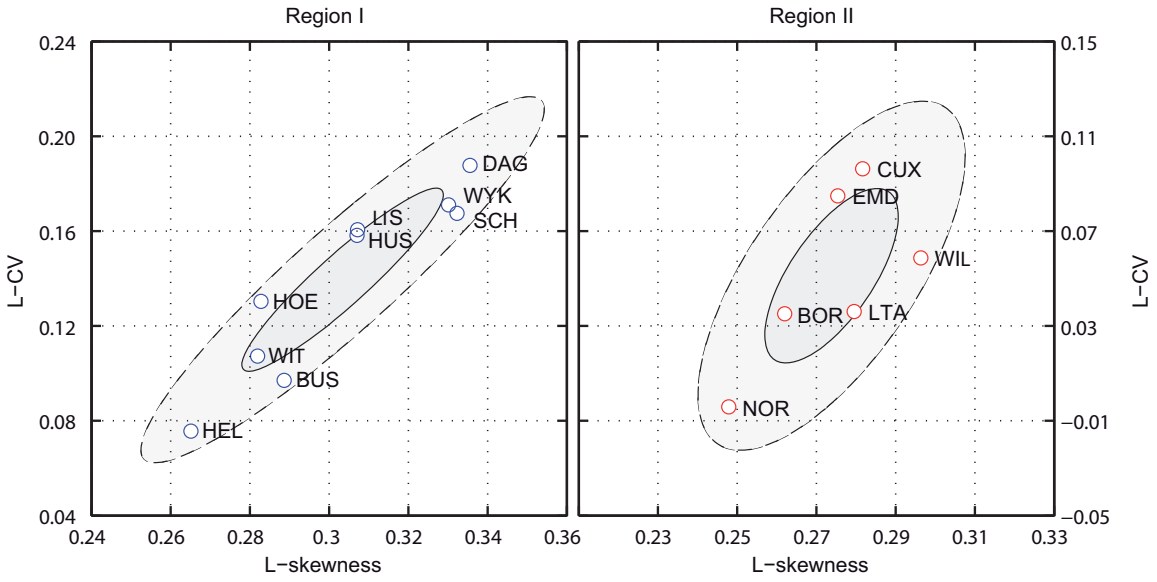


Fig. 23 Error ellipse of Region I and II following Attempt B (see Tab. 2)

11.3 Regional distribution

In RFA, the individual samples within each region are assumed to have a common distribution and essential differences are only found in a scaling factor. The selection of an appropriate regional distribution is conducted using L-moment ratio diagrams (MRDs)

Tab. 2 Homogeneity assessment of different regions

Location		Heterogeneous			Homogeneous								
Station	State	A	RMSE [cm]	Δ max. [cm]	B	RMSE [cm]	Δ max. [cm]	C	RMSE [cm]	Δ max. [cm]	D	RMSE [cm]	Δ max. [cm]
Lis	SH	R. I	9.22	60.12	R. I	9.11	3.03	R. I	8.93	46.19	R. I	8.75	28.51
Hoe	SH		9.83	18.88		9.84	5.60	R. II	9.93	53.99	R. II	9.84	47.69
Wit	SH		0.51	16.63		10.57	6.36	R. I	10.72	17.96		10.41	32.74
Wyk	SH		11.04	42.33		11.00	24.01		10.99	27.40	R. I	11.05	17.08
Dag	SH		10.73	38.05		10.70	19.38		10.72	22.84		10.81	17.23
Sch	SH		11.81	42.24		11.77	23.34		11.71	26.84		11.73	17.37
Hus	SH		13.76	67.80		13.97	57.06		13.39	51.59		13.18	31.01
Bus	SH		11.53	89.23		11.73	82.79		11.24	73.53		11.04	53.61
Hel	SH		10.29	39.45		8.87	40.80	R. II	7.71	11.56	R. II	7.83	12.91
Wil	LS		10.98	21.09	R. II	12.27	61.76	R. I	11.63	20.50		12.56	62.17
Cux	LS		10.33	67.33		11.12	43.92		10.65	17.50	R. I	10.66	16.82
Emd	LS		11.24	17.79		12.95	95.16		11.68	46.32		11.40	27.56
LtA	LS		10.25	36.36		10.09	24.65		10.47	54.38	R. II	9.60	15.03
Bor	LS		10.79	38.62		7.89	16.15	R. II	7.16	14.52		7.16	11.88
Nor	LS		9.73	45.22		9.50	25.27		9.08	14.49		9.14	15.92
∑	---	---	162.04	641.14	---	161.38	589.28	---	156.01	499.61	---	**155.16**	**407.53**
Ø	---	---	10.80	42.74	---	10.76	39.29	---	10.40	33.31	---	**10.34**	**27.17**

where L-moment ratios of individual stations as well as their regional average are plotted against given distributions. MRDs are increasingly used in literature (Peel et al., 2001), providing a visual indication which distribution may be appropriate to describe the regional sample.

In Fig. 24, the MRD of Attempt B (i.e. the example used in the previous section) is shown. Stations (circles) assigned to both Region I (blue) and II (red) as well as their regional average (crosses) cluster around the generalized Pareto distribution (GPD). This is why the GPD is assumed to give the best regional fit to the data. Occurrence probabilities can be derived by multiplying the regional distribution (here the GPD) and the index-flood according to Equation 8. In this step, a large degree of uncertainty may be introduced as the index flood from individual stations may have a large variability reflecting the hydrologic diversity within a region (Bocchiola et al., 2003).

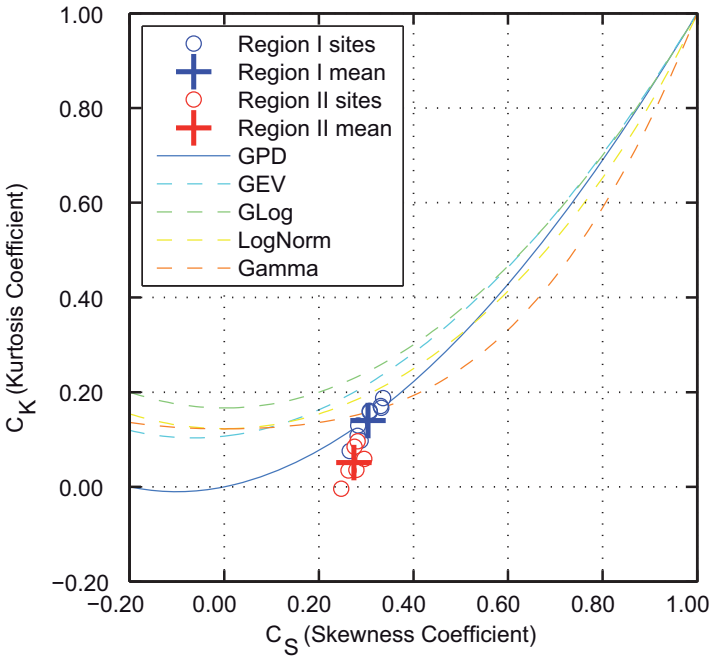


Fig. 24 Moment ratio diagram comparing Region I stations (blue dots) and mean (blue cross) and Region II stations (red dots) and mean (red cross) with different distributions as given in the legend

11.4 Choice of appropriate regions

Sect. 11.2 highlighted that the '*pooling groups*' approach may be useful to identify homogeneous regions. To transfer hydrologic information from gauged to un-gauged sites, a similarity measure is needed that helps to identify which un-gauged sites may be assigned to which homogeneous region. Weiss et al. (2013) present a methodology to identify homogeneous regions for RFA intended to be used with extreme skew surges. Their approach is based on identifying typical storm footprints that are appropriate to describe local storm surge characteristics. In the study area of Part II, however, extreme water levels cannot be characterized by meteorological forcing alone. Instead there is a complex interaction between different forces, such as from astronomical and meteorological conditions as well as their response to the extensive tidal flats in the Wadden Sea (Jensen and Müller-Navarra, 2008). Consequently, there is no published similarity measure available (at least to the authors knowledge) that can be used to objectively assign individual stations or un-gauged locations into one of the two regions of the case study.

Here, the objective of RFA is to infer return water levels for un-gauged sites. This requires that the regional distribution appropriately describes the water level distribution at

each individual station that is considered. To objectively assess the performance of each of the $c_h = 7.678$ homogeneous combinations, the differences in return levels at all 15 stations using RFA according to Equation 8 and at-site analysis were calculated. As a selection of these, Tab. 2 presents three cases resulting in statistically homogeneous combinations (i.e. Region I & Region II), where regions are constructed by assigning the individual stations either to federal states (Attempt B), to locations (i.e. island or shoreline, Attempt C) or using the combinations causing the smallest errors (i.e. RMSE between the regional distribution from the RFA and from analyzing the observations at the 15 study sites covering return periods up to 10.000 years averaging all 15 stations; Attempt D).

Tab. 2 shows that Attempt D has the smallest deviations (sum and on average) of all three homogeneous examples compared to the at-site analyses. The allocation of stations, however, shows no distinct pattern but seems to be arbitrary. Attempt C shows slightly larger errors and stations tend to be separated either located on islands (Region II) or the mainland (Region I) but showing exceptions as e.g. LT Alte Weser (LtA). The largest errors are found for Attempt B. This attempt, however, has the advantage, that individual stations may unequivocally be assigned to a region, i.e. the two federal states. It may be argued that a federal state may not be a suitable criterion to characterize hydrologic responses. However, some studies reported differences in atmospheric forcing affecting mean (Dangendorf et al., 2013a; Wahl et al., 2013) and extreme water levels (Dangendorf et al., 2013b) along the German Bight where tide gauges located in the south-western part are more exposed to north-westerly winds while tide gauges along the north-eastern coast are mainly affected by south-westerly winds. Therefore, it was decided to continue with Attempt B hereafter. Fig. 25a and Fig. 25b show differences in exceedance probabilities $Q_i(F)$ with $1 \leq T \leq 10.000$ years at the 15 sites from RFA and at-site analyses.

This comparison shows differences between return levels to deviate up to ~83 cm for Region I and up to ~95 cm for Region II. The largest discrepancies in both regions are found for higher return periods and decrease for smaller return periods. With respect to practical applications, this has major implications. In Schleswig-Holstein for instance, coastal defenses are built according to the 200-year design level. For this particular return period, the figures show deviations of up to 60 cm, where positive deviations indicate that the use of RFA underestimates the at-site results, i.e. the RFA significantly underestimates the required level of protection. From these findings it is concluded that it is difficult to convey information about the likelihood of extreme water levels from gauged to un-gauged

sites in the German Bight using the RFA. Instead, regionalization approaches that account for local storm surge characteristics are required, such as using multi-decadal model hindcasts, which are explored in the next section.

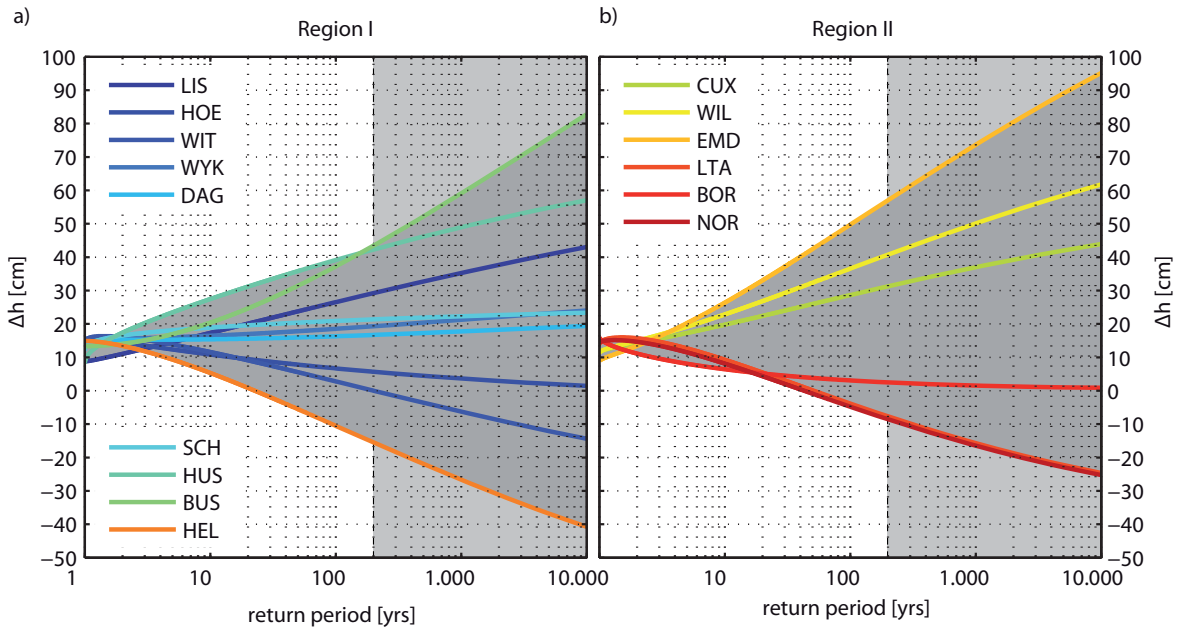


Fig. 25 Differences in return water levels from comparing regionalized water levels and at site analysis of Attempt B in a) Region I and b) Region II

12 Alternative regionalization approach

12.1 Model configuration

To generate continuous water levels for the entire German Bight, a 40-year hindcast for the period from 1970 to 2009 was performed with a process-based hydrodynamic numerical model. A two-dimensional, depth-averaged barotropic tide-surge model of the entire North Sea has been configured using the Danish Hydraulic Institute's (DHI) Mike21 FM (flexible mesh) model suite. The software is based on the numerical solution of the incompressible Reynolds averaged Navier-Stokes equations; the spatial discretization is achieved using a flexible mesh. The model was configured within a coastline provided by the National Oceanic and Atmospheric Administration (NOAA) with a resolution of 1:250.000 km (http://www.ngdc.noaa.gov/mgg_coastline/). The resolution of the coastline was resampled to 30 km along the open boundaries, increasing to 10 km in the northern- and southern-most parts of the European mainland coastline. In between these locations (Scandinavia, the Netherlands, Belgium, France), the resolution was successively resampled until reaching a maximum resolution of 1 km in the German Bight.

The bathymetric data (see Fig. 26), interpolated onto the model grid, was obtained from a range of different sources. In the northern part of the German Bight, high resolution (~ 15 m) survey maps of the Wadden area provided by the Schleswig-Holstein Agency for Coastal Defense, National parks and Marine Conservation (LKN-SH) were used. In this particular area, the Halligen are located. To account for influences on currents resulting from these small islands, a Digital Elevation Model (DEM) covering all of the ten existing Halligen was integrated into the model. The DEM was also provided by the LKN-SH. In the remaining parts of the German Bight, a bathymetric dataset with a resolution of 1 nautical mile provided by the Federal Maritime and Hydrographic Agency (BSH) was interpolated onto the grid. Apart from the German Bight, the General Bathymetric Chart of the Oceans (GEBCO) data provided by the British Oceanographic Data Centre (BODC)

with global coverage and a resolution of 0.5° was used. All datasets were corrected to the German reference datum NHN.

At the open boundaries, the model was driven by astronomical tidal levels (see Fig. 26). These were derived from a global tide model provided by MIKE21 (DHI), including the eight primary harmonic constituents (K_1, O_1, P_1, Q_1, M_2, S_2, N_2 und K_2, see e.g. Andersen, 1995). In order to capture the effect of external surges, the locations of the open boundaries were chosen to be far enough away of the shelf break. Additionally, the Mean Sea Level (MSL) was considered using an index-time series for the entire North Sea from Wahl et al. (2013); the time series was derived using data from 30 tide gauges located around the North Sea basin. As each year of the considered 40- year hindcast was run separately, the MSL at the open boundaries was adjusted according to the annual average MSL values from the time series.

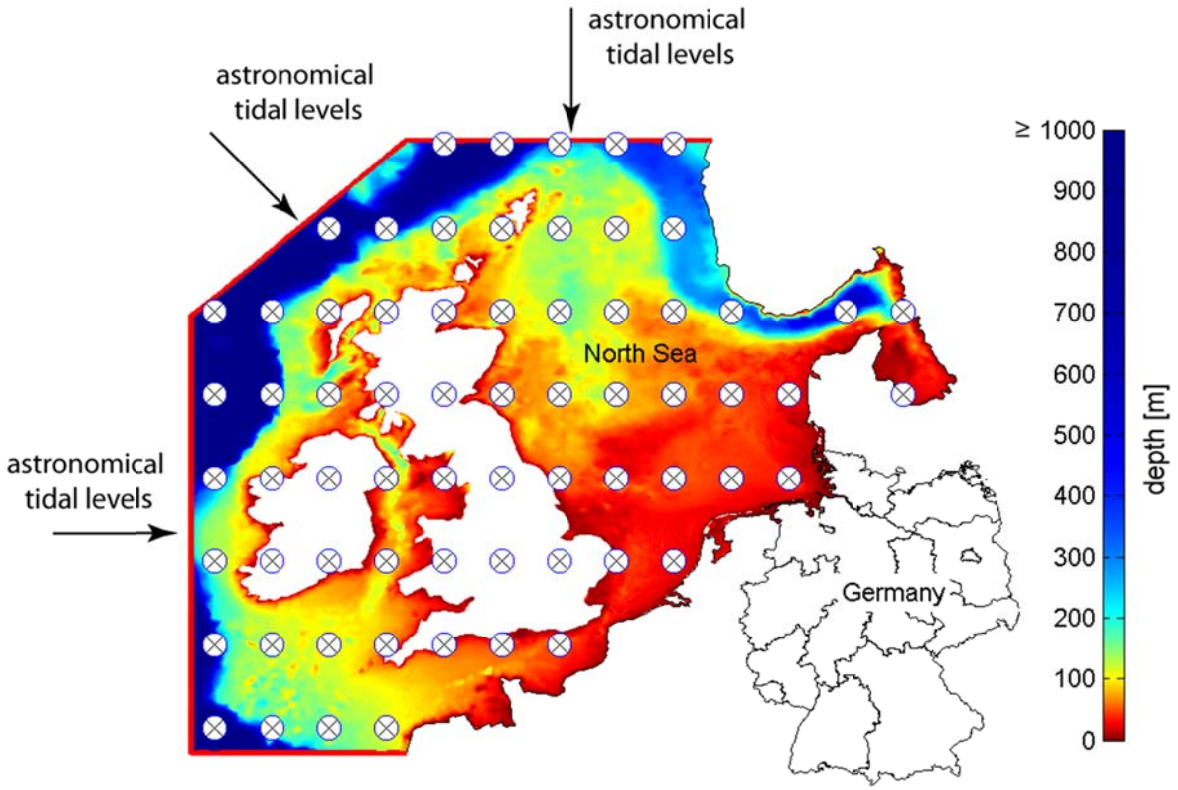


Fig. 26 Grid points of wind- and pressure fields (blue circles): the red line represents the open boundary of the model; the colorbar shows the model depths [m]

The surge component was generated by forcing the model with mean sea level pressure fields and U and V components of 10 m wind fields provided by the Cooperative Institute for Research in Environmental Sciences (CIRES) 20th Century Reanalysis V2

Project (Compo et al., 2011) of the Earth System Research Laboratory, US National Oceanic & Atmospheric Administration (NOAA). These fields are available with a spatial resolution of 2° (see Fig. 26) and a temporal resolution of 6 hours (3 hours in the forecast). The bed resistance was set to a constant Manning's number of $n_m = 0.022$ [-] (corresponds to $k_{st} = 45\ m^{1/3}/s$) (see also Sect.12.2). The model was run for a two day warm up period (a test using longer warm up periods did not show any changes) and results were stored at an interval of 10 minutes for every model grid point.

12.2 Model calibration

The model was calibrated using stepwise variations of the bed resistance, considering Manning's n-values in the range of $0.020 \leq n_m \leq 0.028$ [-]. For simplicity, constant Manning's n-values were used spatially across the entire model domain. The evaluation of the models behavior and performance was conducted by comparing simulated and observed water levels. As shown in Krause et al. (2005), a large number of efficiency criteria are available in hydrologic modeling. Here, the *Coefficient of determination (r^2),* the *Index of agreement (d)* and the *Root Mean Squared Error (RMSE)* were used. The *Coefficient of determination* is defined as the squared value of the coefficient of correlation (Krause et al., 2005) between observed (x_o) and modeled (x_m) water levels and is calculated as follows:

$$r^2 = \left\{ \frac{\sum_{i=1}^{n}(x_{mi} - \overline{x_m})(x_{oi} - \overline{x_o})}{\sqrt{\sum_{i=1}^{n}(x_{oi} - \overline{x_o})^2}\ \sqrt{\sum_{i=1}^{n}(x_{mi} - \overline{x_m})^2}} \right\}^2 \quad , \qquad \text{Equation 10}$$

with $0 \leq r^2 \leq 1$. A value of $r^2 = 0$ [-] denotes that there is no correlation between observed and modeled water levels whereas a value of $r^2 = 1$[-] indicates that observed and modeled water levels are identical. The Index of agreement, proposed by Willmot (1981), is the ratio of the *mean square error* (MSE) and the potential error (Krause et al., 2005). It is defined as

$$d = \frac{1 - \sum_{i=1}^{n}|x_{oi} - x_{mi}|^2}{\sum_{i=1}^{n}(|x_{mi} - \overline{x_o}| + |x_{oi} - \overline{x_o}|)^2} \quad . \qquad \text{Equation 11}$$

Additionally, the *Root Mean Squared Error (RMSE)* was calculated as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{oi} - x_{mi})^2} \quad , \qquad \text{Equation 12}$$

where n denotes the number of values considered and x_o and x_m are defined as above. The efficiency criteria resulting from the comparison between the observed and modeled water levels are summarized in Tab. 3. The results highlight, that the overall agreement was highest along the UK stations. Slightly higher differences occurred in the German Bight

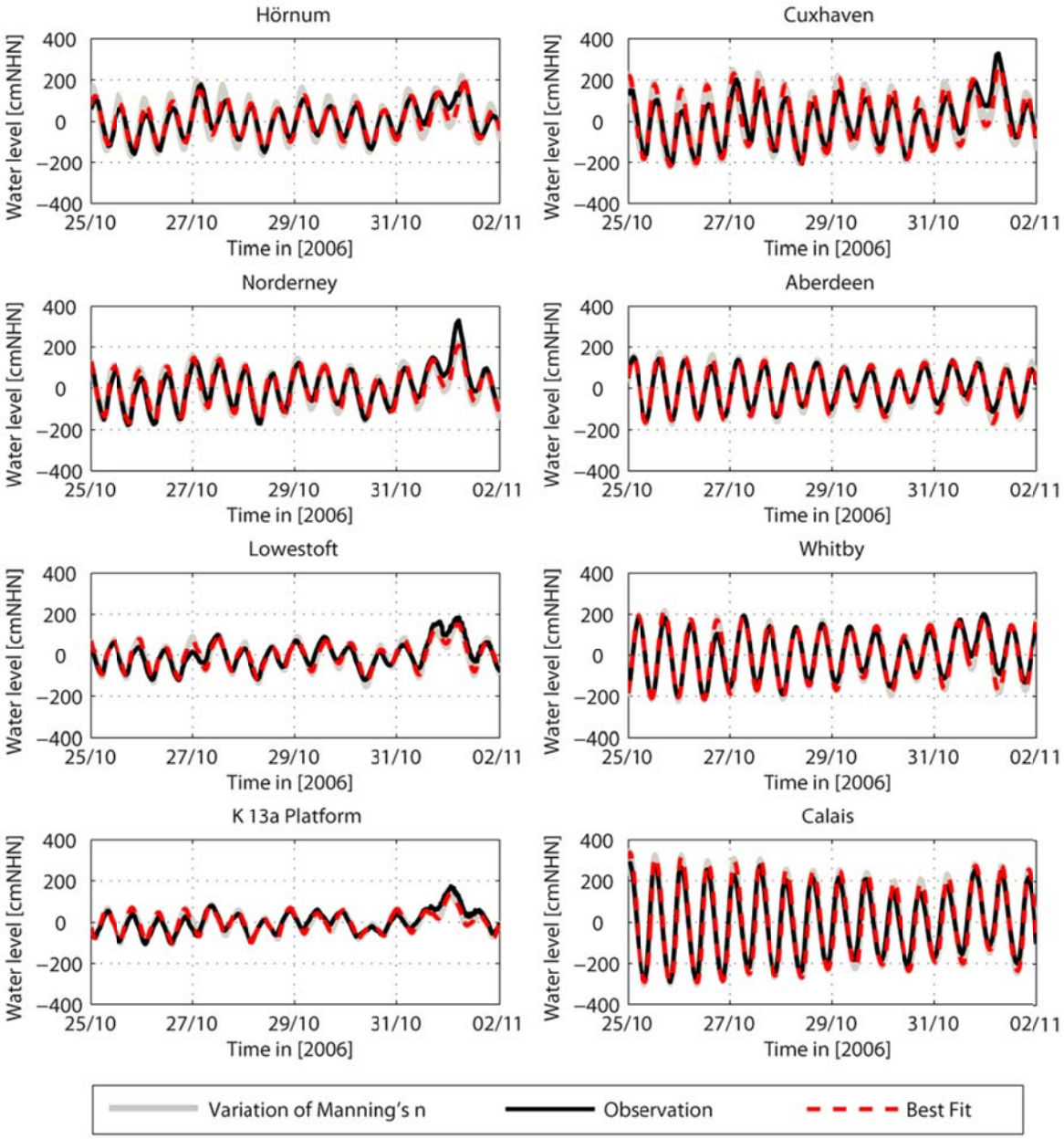


Fig. 27 Model calibration at different location

and are most probably attributed to using one representative bed resistance instead of defining regions of different resistances as well as from shallow water effects that occur in this region and which are possibly not captured properly by the model. The results from using different n-values are shown in Fig. 27 (see also C Appendix).

The black curves represent the observed water levels at the locations given in the calibration (cal.) column of Tab. 1; the red curves show the modeled water levels when using the best fit according to the efficiency criteria; grey shaded curves show the results from the remaining n-values, which did not yield the best fit.

Tab. 3 Efficiency criteria based on the best fit with Manning's n = 0.022 [-]

Criteria	Hörnum	Cuxhaven	Norderney	Aberdeen	Lowestoft	Whitby	K 13a P.	Calais
r^2 [-]	0.91	0.88	0.89	0.97	0.86	0.95	0.85	0.94
d [-]	0.98	0.96	0.97	0.99	0.96	0.99	0.96	0.98
RMSE [cm]	16.64	31.08	21.92	13.26	17.25	19.76	14.61	33.20

12.3 Bias-correction

The calibration exercise allowed to minimize the differences between the observed and the modeled water levels (bias) at individual stations but there are still some differences present. The possible sources of such differences are multifarious including the parameterization that is conducted in the model set-up, allowing for a range of different strategies. Furthermore, all water level observations are prone to natural and anthropogenic influences that cannot entirely be captured by a numerical model. For instance, the wind fields that were used have a temporal resolution of 3 hours and a spatial resolution of 2°; for simulating storm surges, this might be too coarse in order to capture all local meteorological effects. The bias can also be attributed to input deficiencies e.g. resolution or scaling effects. With regards to extreme value analyses, this bias can produce large discrepancies in return water level estimates, particularly at higher return periods.

Thus, the modeled water levels are corrected prior to performing the extreme value analysis. The bias correction can be assumed as a function to transfer the modeled variable into a corrected variable (Piani et al., 2010). This function is created by describing the differences between a pair of variables (e.g. observed and modeled water levels at a tide gauge station) with a parametric or a non-parametric fit (Mudelsee et al., 2010). In this

thesis, a bias correction method to derive reliable water level data covering the entire German Bight and the period 1970 to 2009 is developed, i.e. a period where many tide gauge records exists (see Tab. 1). Hence, a non-parametric transfer function for each individual year of the 40-year hindcast is used.

The bias correction is based on three computational steps. Firstly, high water levels of observed x_o and modeled water levels x_m are computed and sorted in ascending order. Secondly, the differences (bias) between the cumulative distribution functions (CDF) of observed $Q(x_{o,js})$ and modeled $Q(x_{m,js})$ high waters at tide gauge station s and for year j are calculated as follows:

$$B_{c,js} = Q(x_{o,js}) - Q(x_{m,js}) \qquad \text{Equation 13}$$

The differences ($B_{c,js}$) are added to the distributions of the modeled high waters $Q(x_{m,js})$ in order to eliminate the bias at each individual station; the resulting values correspond to the high waters derived from tide gauge records:

$$Q(x_{o,js}) = B_{c,js} + Q(x_{m,js}) \qquad \text{Equation 14}$$

This procedure can be used to eliminate the bias at each gauged station and for each period where observational data is available. However, as the model also generates water levels between the gauged sites, the bias-correction needs to be transferred to these locations. In the third stage, the bias-correction is interpolated from all 15 tide gauge stations envisaged for correction purposes (see correction (cor.) column in Tab. 1) to the locations between the gauged sites. The interpolation is performed for each year individually using the *Inverse Distance Weighted* (IDW) Method (e.g. McMillan et al., 2011).

The three steps to perform the bias-correction are shown in Fig. 28. Fig. 28a shows all grid-points (black dots) of the model along the coast, for which water level time series are available from the 40-year hindcast. Tide gauge locations are highlighted as green circles (see the correction (cor.) column in Tab. 1). To illustrate the methodology, the tide gauge of Hörnum was chosen as an example (red circle in Fig. 28a).

In Fig. 28b, the distributions of observed (black line) and modeled (red line) high waters for Hörnum are shown. The bias, i.e. difference between the two distributions according to Equation 13, is shown as blue line. Any bias having a probability between 0

and 1 yields a value to correct the modeled data. For instance, the correction for *Q(x = 0.2)* amounts to Δh = 4.95 cm. Fig. 28c shows the interpolation of the bias-correction for *Q(x = 0.2)* to all grid points. The corrected water levels are calculated by summing the distribution of the modeled water levels and the interpolated bias-correction at each individual location. Fig. 28c indicates that the bias-corrections for *Q(x = 0.2)* are less in the northern parts (coastline of Schleswig-Holstein) than in the western parts (coastline of Lower Saxony) of Germany (see Fig. 55 in the B Appendix). These findings are similar for all the other *Q(x)*, highlighting the good quality of the bathymetry used along the northern German coastline and its impact on simulated water levels (see Fig. 55 for detailed

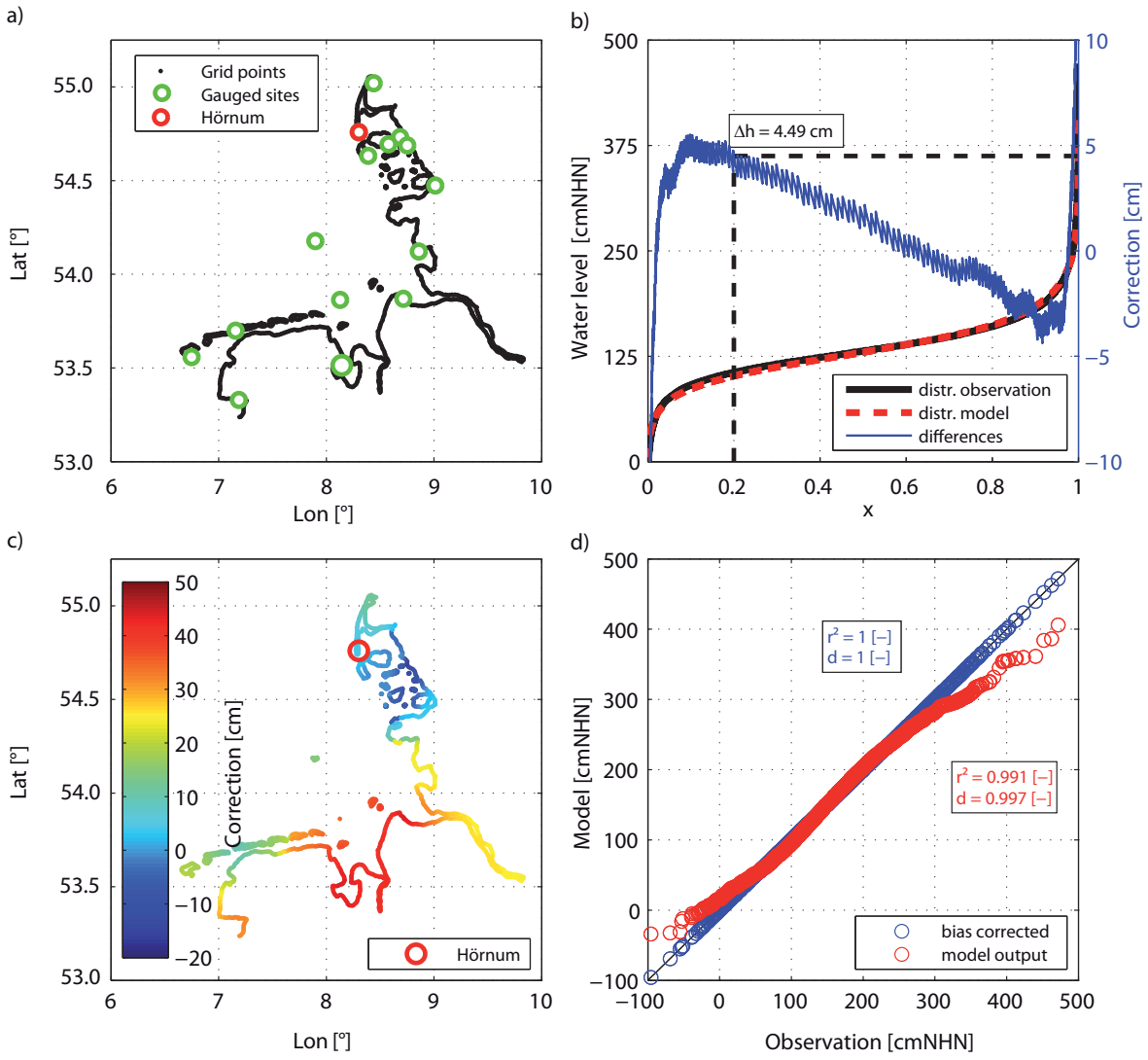


Fig. 28 Example of performing the bias-correction with: a) showing all grid-points (black) and tide gauges (green) of the model along the coast; b) the distributions of observed (black) and modeled (red) high waters for Hörnum tide gauge; c) the bias-corrections for Q(x = 0.2); d) a comparison of observed and modeled high waters before (red) and after (blue) the bias-correction

illustration of the bias-correction along the entire German Bight). Fig. 28d shows a comparison of observed and modeled high waters before (red dots) and after (blue dots) the bias-correction. The un-corrected model output tends to underestimate higher high-waters whereas lower high waters are overestimated.

12.4 Validation

To verify the validity of the approach described above at un-gauged sites, the same methodology is applied to Pellworm Harbor (see Fig. 6e). This tide gauge station has not been used to correct the water levels along the German North Sea coastline. Instead it has been removed from the pool of tide gauge records considered for correction purposes, so that the modeled water levels of Pellworm Harbor are adjusted using the bias-correction that has been interpolated from neighboring stations.

Fig. 29a shows the regression of observed and modeled tidal high water levels at Pellworm Harbor (red dots). As before in Hörnum, the largest differences are found in the highest and lowest high water levels ($r^2 = 0.99$ [-], $d = 0.98$ [-]). Applying the bias-correction (blue dots) eliminates most of the deviations, but does not lead to complete equality ($r^2 = 0.999$ [-], $d = 0.999$ [-]). The remaining differences between the distributions of observed and bias-corrected water levels are shown in Fig. 29b, indicating larger differences for lower high waters. The same behavior is evident from Fig. 29c, where lower percentiles show a tendency to have larger deviations than higher percentiles.

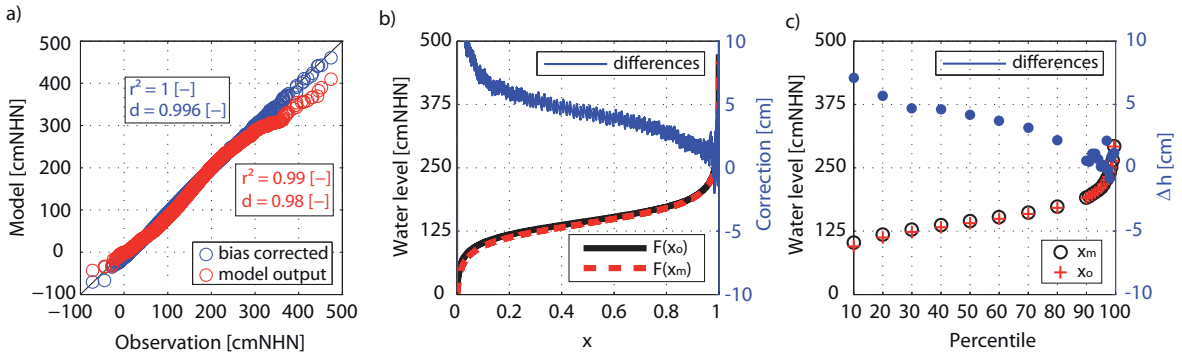


Fig. 29 Validation of bias-corrected water levels at Pellworm Harbor

A compilation of all efficiency criteria applied to all 16 validation sites (see validation column (val.) in Tab. 1) is shown in Fig. 30. The red dots depict the comparison of observed and modeled water levels at individual stations; the blue dots show the

comparisons of observed vs. modeled and bias-corrected water levels. As expected, the bias-correction increases the *coefficient of determination r²* at all stations, reaching values of $r^2 \approx 1$ [-] (Fig. 30a). Fig. 30b shows a similar effect for the *index of agreement d*. At all stations, the *index of agreement* is improved to $d \approx 1$ [-]; the improvement at Wittdün, Wyk and Dagebüll is small, as the *index of agreement* was already high at these stations before the bias-correction was applied. In summary, the validation shows that especially higher high water levels (i.e. storm surge water levels) derived from numerical model simulations are very well represented when the bias-correction is applied.

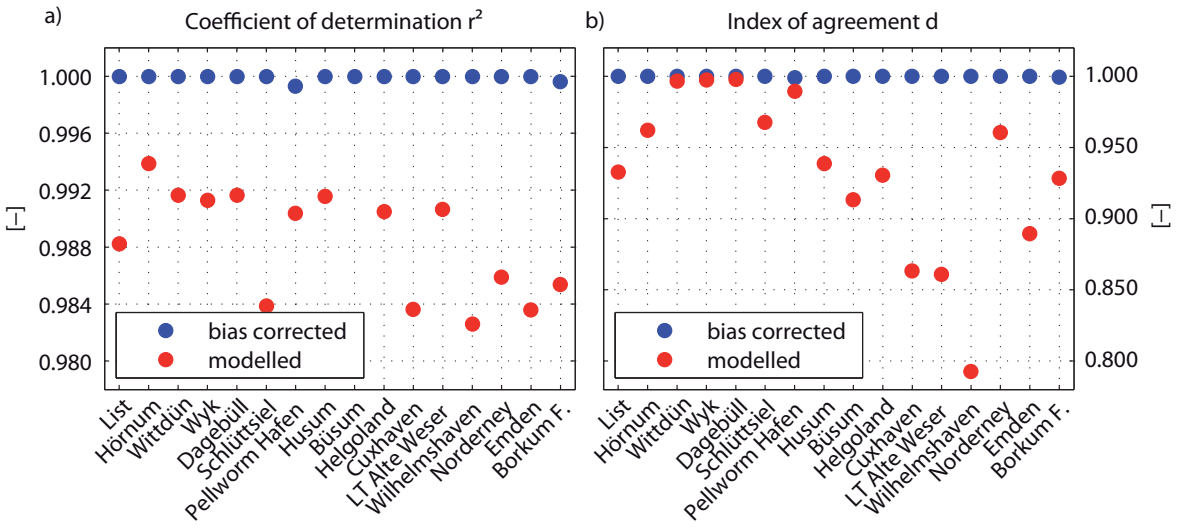


Fig. 30 Compilation of efficiency criteria applied to 16 stations

13 Extreme value analysis

Following the bias correction stage, extreme value analyses were conducted for the whole North Sea coastline of Schleswig-Holstein (north-eastern German Bight). For this stretch of coastline, the model provides water level time series at about 900 coastal grid points that are located approximately every kilometer (i.e. the mean distance). All return water levels are estimated using the approach recommended in Sect. 9.

Return water levels for Hörnum are shown in Fig. 31a. The return periods T and associated return water levels are calculated using both tide gauge records and water level time series derived from the model hindcast after applying the bias-correction. As expected (Hörnum was considered for the correction), there are no differences in the estimates from the observed (blue line) and modeled (red line) water levels. Fig. 31b shows the results for Pellworm Harbor (not considered for the correction). In this case, slight differences in the return water level estimates are found. However, up to return periods of approx. T = 400 years, the differences are below $\Delta h \leq 2$ cm reaching $\Delta h \leq 5$ cm for a return period of

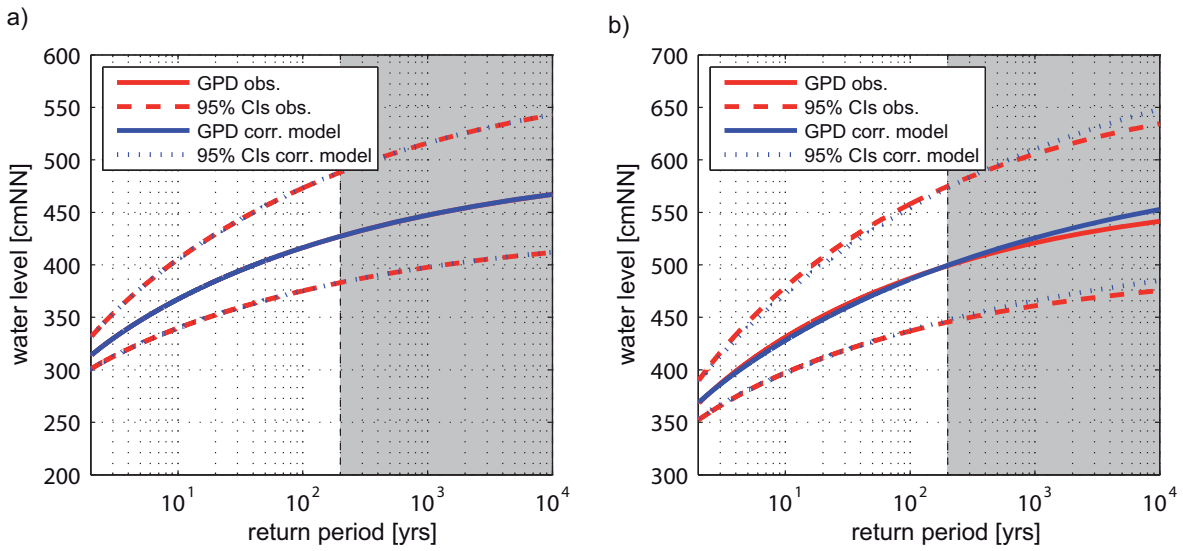


Fig. 31 Return water levels for a) Hörnum and b) Pellworm Harbor

T = 1,000 years. The maximum of Δh = 11 cm is found for a return period of T = 10,000years. However, the input period only covers 40 years, and therefore the extrapolation to 1,000 years or even 10,000 years is debatable. The deviations referred to estimates based on observational data are therefore considered acceptable. This is why it is concluded that the above presented bias-correction is suitable to be used with modeled water levels in the German Bight, which are envisaged to serve as input for extreme value analyses. Fig. 32a schematically shows regionalized water levels with a return period of 200 years for the entire coastline of Schleswig-Holstein (see D Appendix for a range of maps showing the return levels T={10, 20, 50, 100, 200, 500 and 1,000} years for different stretches of the coastline drawn to scale). Return water levels in the southern parts of Schleswig-Holstein are higher than in the northern parts, most likely as a result of shallow water effects and meteorological (i.e. wind pile up in enclosed bays) forcing. Fig. 32b exemplarily shows regionalized return water levels for the Hallig Hooge, highlighting the benefit of the regionalization approach proposed here. There are no tide gauge measurements available in this area that could be used to calculate return water levels. The regionalization, however, enabled to reliably derive return water levels for this un-gauged

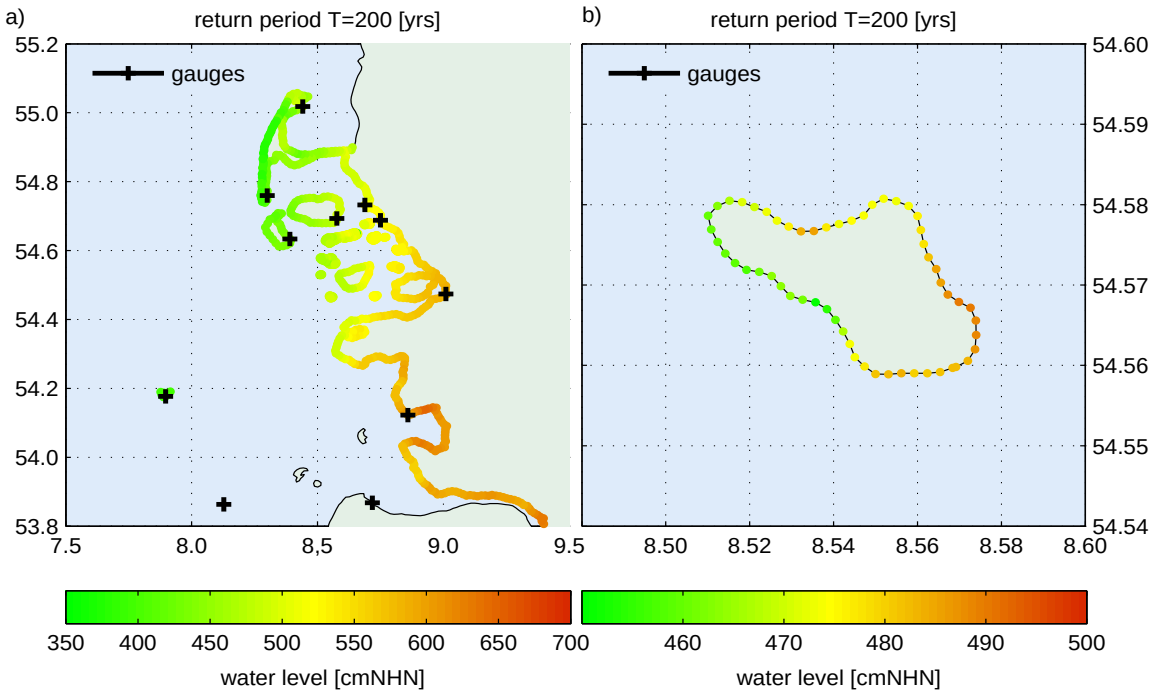


Fig. 32 Schematical illustration of the a) regionalized return water levels along the coastline of Schleswig-Holstein; b) the regionalized return water levels at the Hallig Hooge

region. This information can be used as a basis for the design of protection measures and is also useful for risk analyses in un-gauged regions like the Halligen.

Fig. 32 shows that return water levels are not similar along the edge of Hallig Hooge but vary up to ~40 cm with larger return levels in the eastern part and slightly lower return levels in the western part of the Hallig. These findings are surprising as the return levels are smallest at the side which is exposed to the open sea. However, similar characteristics can be observed in the water level sample. Fig. 33a shows the occurrence of the highest to lowest water levels along the edge of Hallig Hooge considering all storm surges between 1970 and 2009. The figure indicates that water levels the south-eastern part where generally higher that in the remaining parts, showing water level differences of up to ~15 cm (see Fig. 33b). These discrepancies are mainly caused by the pile up of water in the German Bight but may also be related to shallow water and reflection effects.

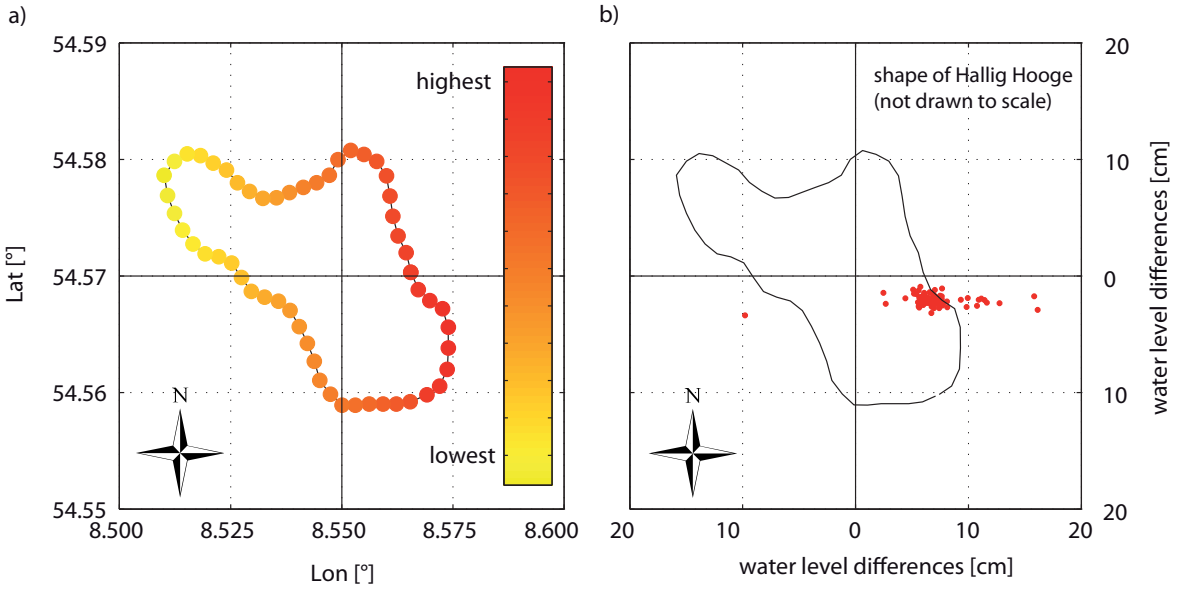


Fig. 33 Water levels differences along the edge of Hallig Hooge with a) the occurrence of the highest to lowest water levels and b) the water level differences

14 Summary and discussion

Part II of the thesis investigates two different methods to estimate return water levels at sites where only little or even no measured water level data is available. The first method under investigation was the regional frequency analysis (RFA), which is based on the concept of regional homogeneity, stating that sites with similar statistical characteristics behave similar in flood frequency response. The use of the index flood approach is investigated where all sites within a homogeneous region are assumed to have a common distribution and differ only by a site dependent scaling factor (index flood). The assessment showed that this very simple approach can principally be used to derive return water levels at gauged sites but causes large deviations compared to at-site analyses. This was mainly caused by differences in shape parameters between samples of individual stations that are combined within a region. To yield better results, a more sophisticated approach thus needs to account for both, the scaling factor as well as the shape of the samples of individual stations.

Furthermore, in RFA it is required to assign individual stations to a set of regions. This is usually accomplished by using a similarity measure which refers to similar characteristics that significantly influence the variable under consideration. With respect to coastal environments this is a challenging task because there are many factors that can substantially impact water levels, such as wind direction, wind speed, tides, topography and non-linear effects. Another drawback in applying RFA to coastal water levels is that the regional quantile needs to be transferred to un-gauged sites (if unambiguously assigned to a region) but there are no scaling factors (i.e. index flood) available that can be used to calculate the quantiles at un-gauged stations. As a workaround, numerically generated water level information may be used. However, the local characteristics in the German Bight cause differences in water levels between neighbouring stations and artificially generated data thus needs to account for a wide range of different conditions. This can be achieved by a multi decadal hindcast.

A second method was tested that combines the output from a model hindcast and extreme value analyses. A comparable method has recently been applied along different stretches of coastlines around the globe (as e.g. in Australia) but these approaches are focusing on the correction of the parameters that are derived from the extreme value sample that was created using the numerical model output. These approaches were adopted and modified to satisfy the characteristics along the entire coastline of Schleswig-Holstein in northern Germany. The benefit of the presented approach is that the simulated water levels are adjusted to the observations and can thus directly be used to derive extreme value samples but also for the detection of trends.

It is shown that water levels derived from a hydrodynamic model can be used to calculate reliable return water levels. Regions with no or only few tide gauge stations can especially benefit from this approach. However, a precondition is to adequately correct the bias that is generated with the numerical simulations. The bias-correction is performed first at each individual station where water level observations exist. Then the correction is transferred to the neighboring sites points using a *Inverse Distance Weighting* interpolation method. As a result, regionalized return water levels at un-gauged sites are obtained, that account for locally confined coastal attributes. The assessment shows that return water levels that are estimated using the approach presented in this part of the thesis are highly consistent with the return water levels from at-site analyses. This information can thus be used for planning purposes and risk analyses.

15 Key findings of Part II

The overall aim of Part II was to develop an approach to assess current return water levels in un-gauged areas. This can be accomplished by following the procedure that is described above and can be summarized as follows:

- *Generate a numerical model hindcast covering the entire coastline and an adequate period as recommended in Sect. 9.*
- *Calculate a non-parametric bias-correction (transfer function) by comparing modeled and observed high water levels at existing tide gauge stations.*
- *Transfer the bias-correction to the entire coastline by using a spatial interpolation method such as the inverse distance weighting (IDW).*
- *Adjust the modeled high water levels to the observed high waters by combining the bias-correction and the modeled water levels.*
- *Estimate return water levels based on modeled and bias-corrected water levels.*

The procedure to bias-correct simulated water levels is also given in Fig. 53 (see A Appendix). This methodology is valid for the German Bight but can likely be applied for other locations around the world but needs to be verified before (see Arns et al., 2013b; see also Arns et al., under review a).

PART III

HOW DOES SEA LEVEL RISE AFFECT EXTREME WATER LEVELS?

16 Motivation

16.1 General

Extreme water level assessments usually include some form of statistical analyses based on the extreme value theory, which requires stationary data sets (see e.g. Jensen, 1984; Rao and Hamed, 2000). For many sites, however, extreme water level samples appear to have non-stationary features such as trends or cycles (Dixon and Tawn, 1994). This is why e.g. Coles (2001), Méndez et al. (2007), Menéndez and Woodworth (2010) and Mudersbach and Jensen (2010) introduced non-stationary approaches to include such temporal changes and fluctuations into the extreme value models. Hunter (2010) used a simpler approach for estimating future changes in the exceedance probabilities of extreme events. By assuming that changes in extreme water levels during the 21st century will be dominated by changes in MSL, he combined observations of present-day sea level extremes with sea level rise (SLR) projections. Until recently, most coastal protection strategies adapted this methodology by raising design water levels according to the projected SLR (Smith et al., 2010). Hence, the results are based on the assumption that mean and extreme water levels will rise by exactly the same amount. However, due to nonlinear interactions between the different sea level components (i.e. MSL, tide, surge) this may either under- or overestimate the impact of SLR on extreme water levels in some areas. This is why a profound understanding of the physical processes driving these changes is required, i.e. all relevant driving factors for regional and local changes in water level extremes need to be thoroughly investigated

16.2 Observed changes in storm surge water levels

A number of investigations dealing with changes in storm surge water levels based on observational data have been published and covered entire coastline stretches (e.g. the German Bight: Jensen, 1985; Mudersbach et al., 2013) or individual stations (e.g.

Cuxhaven tide gauge: von Storch and Reinhardt, 1997; updated by Weisse and von Storch, 2009). First attempts addressing the question whether extremes have changed at rates different to MSL at a global scale were by Woodworth and Blackman (2004). Their results were later updated and extended by Menéndez and Woodworth (2010). Using a quasi-global sea level data set, they found that extreme sea levels have increased at most locations around the world with the main conclusion that much of this increase was due to changes in MSL. Similar findings were reported by numerous authors for specific sites around Europe, e.g. Haigh et al. (2010b) for the English Channel, Marcos et al. (2009) and Tsimplis and Shaw (2010) for southern Europe, and Letetrel et al., 2010 for Marseille. A comprehensive review on this subject is provided by Woodworth et al. (2011). They conclude that changes in extremes were mostly driven by changes in MSL but also highlight that there were exceptions to this rule. Some authors have identified areas where extreme water levels appear to have changed at rates faster than those observed in MSL. For the Rhone Delta in southern France, Ullmann et al. (2007) showed maximum sea levels to have increased twice as fast as MSL during the 20th century. They attributed these differences to changes in the wind fields. At Oostende in Belgium, Ullmann and Monbaliu (2010) found that changes in the atmospheric circulation over the North Atlantic have triggered an increase in the wintertime 99th percentiles of water levels of +3.0 mm/year and of surges of approximately +1.0 mm/year from 1925 to 2000.

In the German Bight, Mudersbach et al. (2013) showed that trends in extreme high water levels differed significantly from those in MSL from the mid-1950s to approximately 1990. At six tide gauges, they performed trend analyses of different high percentile time series after being reduced to MSL (see Fig. 34). The residuals showed significant trends between +1.0 to+3.7 mm/year. They argued that this was a result of changes in the amplitudes of some of the main tidal constituents. Additionally, Dangendorf et al. (2013c) demonstrated that the observed changes in high sea levels in the German Bight were seasonally uneven distributed, with the highest rates of change during winter. This was coherent with simultaneous changes in local and large-scale wind fields. A final quantification of the contribution of both factors (i.e. tides and surges/wind) to the observed differences between changes in MSL and extreme water levels is still missing.

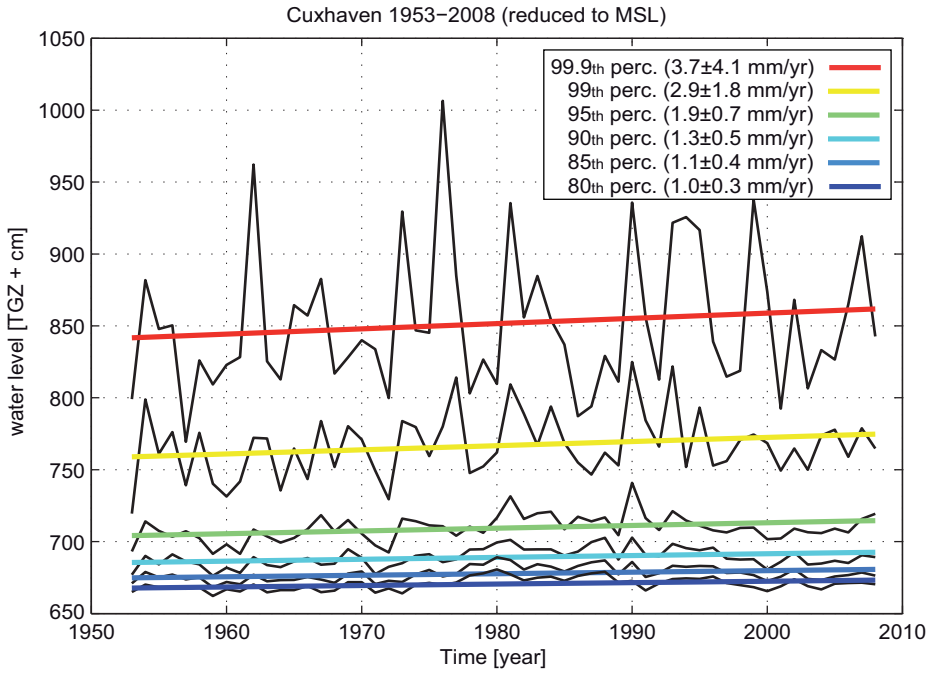


Fig. 34 Linear trends with standard errors of the 99.9th, 99th, 95th, 90th, 85th, and 80th percentiles of tidal high waters exemplarily shown at gauge Cuxhaven from 1953 to 2008 reduced to the corresponding MSL (adapted from Mudersbach et a., 2013)

16.3 Investigations on possible future changes in extreme water levels

Extreme water levels arise as a combination of the MSL, astronomical tides (hereafter referred to as tide), the dynamic response of the sea surface to atmospheric forces (hereafter referred to as surge), and the nonlinear interactions between them. Long term changes, as e.g. from climate change, in any of those components (see Sect. 17 for a literature review on changes in each of these components) may substantially alter the risk associated with extreme water levels (Weisse et al., 2011). Quantifying the individual contribution of each component to extreme water levels, however, is difficult using observational data. This is why numerous studies focus on model based investigations, considering one or more components to be changed.

Various studies reported responses of extreme water levels different to the assumed SLR for different locations around the globe. Flather and Khandker (1993) investigated the effect of SLR on extremes in the northern Bay of Bengal by comparing model runs of present day conditions (often referred to as control run) and +2.0 m SLR. Based on one single event, they showed that maximum surge heights were reduced by about +0.2 to +0.3 m whereas total water levels showed spatially different patterns with increases up to

+0.25 m as well as decreases by up to -0.5 m different from SLR. This indicates the nonlinear impact water depth changes may have on storm surges. For southeast Louisiana, Smith et al. (2010) investigated the impact of +0.5 and +1.0 m SLR on hurricane surge. They used six hypothetic hurricanes potentially causing 100-year return water levels. In areas where maximum surges occurred, they found that surges increase linearly with SLR and concluded this being a result of an unchanged interaction between the bottom and the surge propagation. In areas of moderate or lower surges (2.0 to 3.0 m) by contrast they found surges to increase by as much as triple the considered SLR. They argued that this could be caused by shallow water effects, e.g. an increase in the speed of propagation.

In the German Bight, first model based investigations on changes in extreme water levels from SLR were conducted by Stengel und Zielke (1994). Based on four historical storm surge events they compared control (observed period) and +1.0 m SLR scenario runs and found that changes in extreme water levels differed significantly from the considered SLR at most of the investigated sites. However, this nonlinear increase was not uniform throughout all investigated events and they attributed the different rates of changes to different storm types that cause extreme events in the German Bight. To investigate nonlinearities between SLR and extreme water levels in the North Sea, Kauker and Langenberg (2000) compared present-day storm surge simulations of a 20-year hindcast with a second hindcast considering +0.1 m of SLR. They detected mean water level changes by roughly the same amount that sea level was raised but could not detect any deviations between mean and extreme water levels. For the North Sea they concluded that storm surge heights do not increase faster than sea levels in general. A more recent study for the German Bight was conducted by Bruss et al. (2010). They investigated the effects of up to +1.0 m SLR on water levels by simulating the entire year of 1999, i.e. they considered only a few extreme events. Their results indicated changes in total water levels of up to ±0.25 m relative to the SLR, showing no constant increase of water levels in the German Bight.

Some studies have also investigated the impact of changes in atmospheric forcing on extreme water levels. Assuming a linear relationship between MSL and extreme storm surges, Woth et al. (2006) compared two storm surge model hindcasts of the entire North Sea each covering a 30-year period. The control runs covered the period from 1961 to 1990, the scenario runs (the output from different climate models was considered) were conducted for 2071 to 2100. Overall, they showed large changes along the continental

coast whereas changes along the UK coastline were insignificant. In the German Bight up to Denmark, changes in both, intensity and duration of extreme water levels became more important, with the 99.5th storm surge percentiles showing significant increases in all scenarios ranging from +0.2 to +0.3 m. This corresponds to a rise of around 20% in surge heights. Comparing North Sea extremes from present-day and possible future climate conditions in combination with changes in MSL, Lowe et al. (2001) found statistically significant changes induced by future meteorological forcing but could not detect significant indirect changes from SLR. Similar findings were reported for the UK (Lowe and Gregory, 2005) and Dutch coastlines (Sterl et al., 2009).

Tab. 4 Investigations on changes in water levels from SLR that were conducted by different authors

Authors	SLR	Sample	Total water levels	Remarks
Flather and Khandker (1993)	+2.0 m	1 event	-0.5m to +0.25 m	Spatially inconsistent
Stengel and Zielke (1994)	+1.0 m	4 events	Significantly different from SLR	Spatially inconsistent
Kauker and Langenberg (2000)	+0.1 m	20 yrs	Negligible differences	Valid for most coastal points
Bruss et al. (2010)	+0.5 to +1.0 m	1 yr (few extremes)	±0.1 and ±0.25 m relative to SLR	No constant increase
Smith et al. (2013)	+0.5 to +1.0 m	6 events	Up to three times the considered SLR	Spatially very inconsistent

16.4 Objectives of this study

This brief summary shows that there are only a few studies available which addressed the impact of SLR on extreme water levels in the German Bight. All concluded that changes in the extremes are mainly caused by MSL, a finding which could possibly depend on the model set up that has been used (i.e. only a few storm surge events were selected for the investigation). Due to these limitations, general conclusions can hardly be drawn on how SLR may alter extreme water levels in this area. Some previous studies also attempted to assess future changes in extreme water levels due to both changes in the meteorological forcing and SLR. Given that the uncertainties accompanied with possible changes in meteorological forcing are still very large (see Sect. 17.2 for a review of recent publications), the thesis focuses on the impact of SLR alone. Thus, the overall aim is to examine the impact of SLR on extreme water levels and the associated exceedance probabilities derived from extreme value statistics in the German Bight. Based on the

recommendations given Part I, all storm surge events where the water level exceeded the 99.7^{th} percentile between 1970 and 2009 were considered. This results in a sufficiently large sample of $n = 65$ events that is used to derive reliable return water levels. By following the overall objective, this part also assesses changes in

(1) the peak high water levels (hereafter referred to as high water levels),

(2) the high water occurrence times,

(3) the high water level distributions,

(4) the tidal constituents

(5) and identifies the spatial distribution of the observed changes.

17 Changes in potential driving factors

17.1 Tidal changes

One of the main parameters to describe the prevailing tidal characteristics is the tidal range, i.e. the difference between successive high and low waters. Based on tide gauge data, large increases in tidal range have been identified along the German (Jensen et al., 1992; Jensen and Mudersbach, 2007) and Dutch (Hollebrandse, 2005) coastlines. These findings have major implications, as increases in tidal range contribute to coastal erosion, but it may also alter coastal circulation patterns affecting nutrient supply and primary production (Jay, 2009). Changes in tidal range were also found in a model-based study by Flather and Williams (2000). By comparing a control run and a +0.5 m SLR scenario run, they found increased tidal ranges in the German Bight, the Skagerrak and the Kattegat and attributed these to reduced tidal dissipation from water depth increases.

Changes are not only apparent in tidal ranges but also in the basic astronomical movements of water levels represented by individual tidal constituents. For the Gulf of Maine, Ray (2006) showed that the M_2 constituent increased almost linearly with SLR throughout most of the 20th century, followed by a sudden drop in the early 1980s. These observations were explained with an enhanced resonance in the Gulf of Maine as a consequence of SLR. From analyzing 34 long records located in the Eastern Pacific Ocean, Jay (2009) found an increase of the mean total tidal amplitudes (0.59 mm/year) at most locations, but less than the present global SLR (1.7 mm/year). He argues that the observed tidal evolution is the result of a shift in the locations of amphidromic points. Analyzing twentieth century tide gauge records, Ray (2009) noticed that the amplitude of the semidiurnal constituent S_2 had decreased along the eastern coast of North America and at the mid-ocean site Bermuda. The observed rates of decrease were unusually (~10% per century) but also not consistent among the stations, i.e. nearly half of the stations showed increasing amplitudes while others showed a strong decrease. He argued that these changes

could be explained by variations in radiational forcing. Using a quasi-global data set of tide gage information, Woodworth (2010) found only little evidence for extensive regional changes in the main tidal constituents in Europe and the Far East, but he also pointed to changes in smaller regions as e.g. the German Bight.

Model-based investigations on changes in tidal constituents induced by SLR were performed by Flather and Kandker (1993). They found that +2.0 m SLR could lead to an increase in the M_2 amplitude of around 10 cm in the north eastern part of the Bay of Bengal and a decrease of around 15 cm in the north western parts of the bay. They concluded that tidal ranges may increase by up to +0.5 m in areas with large tidal ranges. In the German Bight, Stengel und Zielke (1994) investigated changes in tidal dynamics resulting from +1.0 m SLR. At some locations they found tidal ranges to increase by a factor more than 30% above that of SLR with largest increases occurring in the Elbe estuary. Plüß (2004) also addressed the impact of SLR on tidal constituents in the German Bight. Considering a period of ~14 days and a maximum SLR of +1.0 m he showed the M_2 constituent to increase by up to +0.05 m in the adjacent estuaries. For the M_4 constituent he could not detect any significant changes in most areas but slightly reduced amplitudes nearshore; similar findings were presented for the M_6 constituent, but with a smaller magnitude.

17.2 Changes in atmospheric forcing

Storm surges are the response of water levels to local and large-scale meteorological forcing. In the German Bight, they are usually generated by strong North Sea winds (> 25m/s) with directions prevailing between north and west (Jensen and Müller-Navarra, 2008). These migratory atmospheric disturbances tend to propagate along regionally confined storm tracks (Weisse et al., 2011). In coastal regions, they are the major geophysical risk often associated with significant losses of life and property (von Storch and Woth, 2008). Storms can be classified as tropical and extra-tropical storms, with the latter being common along the North Sea coastline. Attributes of winds can cause extreme impacts (Seneviratne et al., 2012), with persistent mid-latitude winds causing elevated water levels (e.g. McInnes et al., 2009) and long term changes in prevailing wind directions having the potential to impact wave climate and coastal stability (Pirazzoli and Tomasin, 2003). A number of recent publications have analyzed trends in mean and extreme wind conditions in the North Sea region. Using geostrophic winds in the southern

North Sea from 1876 to 1989 (Schmidt and von Storch, 1993) and pressure records from two Swedish stations from 1780 and 1823 to 2002 (Bärring and von Storch, 2004), considerable inter-annual and decadal variability was noticed but no evidence for significant long-term trends was found. Using 13 Dutch records of near-surface winds, Smits et al. (2005) found extreme wind speeds to have declined by up to 10% per decade. Contradictory, they found extreme wind speeds to have increased by up to 20% per decade using reanalysis data covering the same period and region. They attribute these inconsistencies to inhomogeneities in the reanalysis data although overestimations from station data cannot be excluded. There are also other studies reporting opposite trends between station and reanalysis data for some areas (see Seneviratne et al., 2012 and references therein). Using 20th century reanalysis data (Compo et al., 2011), Donat et al. (2011) detected significant long-term trends in the occurrence of annual maximum wind storms over Europe suggesting the increases could (at least partly) be a response to enhanced greenhouse gas emissions during past decades. Krüger et al. (2013) concluded that this upward trend was mainly an artifact of less data assimilated and larger inconsistencies before 1950, a finding controversially discussed in Wang et al. (2013). However, Dangendorf et al. (under review) confirmed the findings of Krüger et al. (2013) (even they also pointed to a better reanalysis quality back to ~1910) with an independent storm surge record at Cuxhaven and further demonstrated that the surges do not show any evidence for a significant long-term trend back to 1843.

Fewer studies report shifts in the North Atlantic storm track (see Weisse et al., 2011 and references therein), with decreased storm frequencies and nearly constant storm intensities in mid-latitudes, superimposed by inter-annual and decadal variability during the second half of the 20th century. For the same period, higher latitudes north of 60° N show increased storm frequencies and intensities. Based on reanalysis data, Siegismund and Schrum (2001) detected a shift of strong south westerly winds from the late autumn into early spring, a finding consistent with the occurrence times of the seasonal MSL peaks detected by Dangendorf et al. (2012).

With respect to future projections, an assessment of possible changes in the North Sea climate (WASA project) points towards a moderate increase of North Sea winds (WASA Group, 1998). From analyzing the outputs of the newest generation of atmosphere ocean global coupled climate models (AOGCMs; CMIP5) for the southern North Sea region, de Winter et al. (2013) found a possible shift towards more westerly winds. These

changes were, however, within the range of previously observed variations and can thus not unequivocally be attributed to climate change (Sterl et al., 2009).

17.3 Mea sea level changes

Sea level rise is one of the most important aspects of climate change (Woodworth et al., 2011; Wahl et al., 2013) and there is concern about the impact this could have on growing coastal communities (Nicholls and Cazenave, 2010). There have been numerous studies on the observed global and regional (North Sea) sea level changes (e.g. Church and White, 2006; Woodworth et al., 2009; Wöppelmann et al., 2008 and 2009; Haigh et al., 2009; Wahl et al., 2010 and 2011 and 2013), pointing to a considerable regional variability in the rates of sea level change (e.g. Church et al., 2004 and 2008; Dangendorf et al., 2012 and 2013a). These differences are mainly caused by regional meteorological, oceanographic and gravitational effects, vertical land movements as well as anthropogenic interventions (see e.g. Wahl et al., 2013 for a review for the North Sea). From a coastal management and planning perspective, regional or local relative sea level changes are thus most important (Nicholls et al., 2011) and it seems reasonable to assume that future changes in sea level will also exhibit a strong spatial variability (Wahl et al., 2011 and 2013). Projections of possible future sea levels have already been assessed on global (see

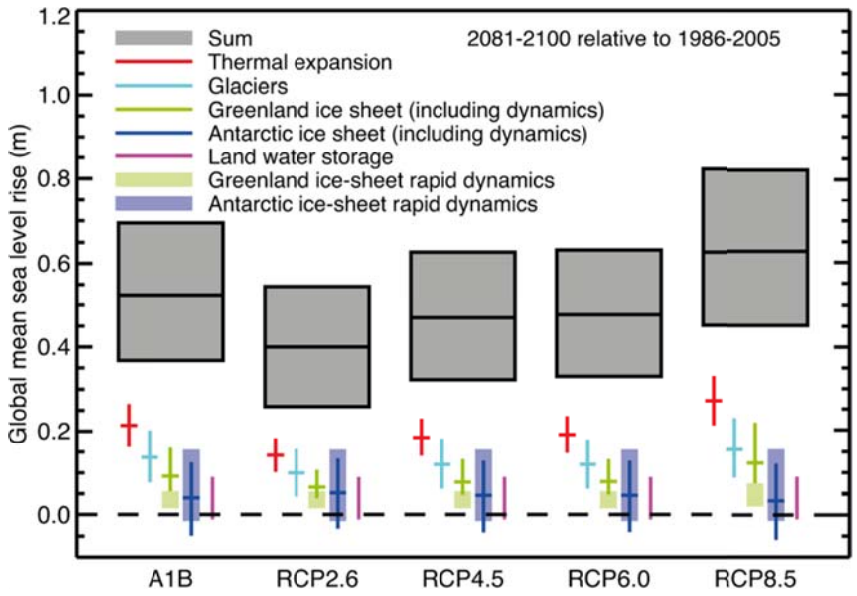


Fig. 35 Projections from process-based models with likely ranges and median values for global mean sea level rise and its contributions in 2081-2100 relative to 1986-2005 for the four RCP scenarios and scenario SRES A1B used in the AR4 (Church et al., 2013)

Meehl et al., 2007 for a review; Rahmstorf, 2007; Vermeer and Rahmstorf, 2009; Grinsted et al.,2010; Jevrejeva et al., 2010) or regional scales as e.g. in the northeast Atlantic Ocean (Katsman et al., 2008), along the UK (see e.g. Lowe et al., 2009), Dutch (Katsman et al., 2011) and Norwegian coastlines (Simpson et al., 2012; Nilsen et al., 2012). The 5th assessment report (AR5) of the IPCC (Church et al., 2013) suggests a model and scenario dependent range of 0.26 to 0.82 m in global MSL until 2081-2100 relative to 1986-2005 (see Fig. 35), a range that will further vary considerably on regional scales.

18 Analytical assessment

Tide gauges measure the combined and complex response of a water body to different forcing factors. Following Flather and Williams (2000), extreme water levels can be considered as the sum of MSL, tide, surge, and the (non-linear) interactions between them. In deep water, non-linear effects can be considered negligible but become increasingly important in shallow water where they are locally generated (Pugh, 1987). This separation into three main components is a simplification as long-period tides (e.g. seasonal cycles) as well as the MSL usually include contributions from meteorological forcing which cannot be entirely distinguished. In the North Sea, for example, the mean amplitude of the largest constituent M_2 has a seasonal modulation of between 1% and 2%. Part of this modulation is caused by astronomical effects, whereas the remaining part is thought to be a result of tide-surge interaction (Pugh and Vassie, 1976) or seasonally varying stratification (Müller et al., under review).

Simplified, each of these components can be described with analytical approaches, helping towards a physical understanding of the involved processes. Analytical approaches, however, cannot provide a full description of the ocean's response to irregular boundary conditions (e.g. variable water depths, specific meteorological patterns, bottom stresses, or sea level changes). A more elaborate way to model complicated responses of coastal water levels to a variety of boundary conditions is thus to use a numerical tide-surge model (Pugh, 1987). Here, both approaches are used to examine the impact of SLR on extreme water levels. First, assuming a SLR of 0.5 m, changes in tides and surges are investigated analytically to identify the major drivers impacting extreme water levels. Non-linear interactions are then assessed using a numerical model. Based on the shallow water equations, presumable impacts of SLR on extreme water levels are identified to verify the numerical model outputs.

18.1 Processes involved

Tidal water level oscillations are the local response of water masses to the gravitational forces of the Moon and the Sun acting on individual water particles (Kamphuis, 2000). In the deep ocean, the tidal spectrum is closely linked to the tide generating potential (Le Provost, 1991). As a tidal wave propagates from the deep ocean towards the shallower shelf, it is deformed by different effects, yielding higher harmonic tides known as shallow water tides (van Rijn, 2010). The shelf tides of the North Sea are dominated by semidiurnal (~twice a day) oscillations with a period of about 12 hours and 25 min. The largest force is generated by the Moon (M_2-constituent), reaching its maximum value once in 14.78 days, when the Moon is nearest to the Earth causing spring tides. Due to friction effects, spring tides do not occur when the Sun and the Moon are in line, but usually two or three days later (van Rijn, 2010). This phenomenon is known as the '*age of the tide*' (see e.g. Proudman, 1941).

Both, the motion of storm surges and the motion of tides belong to the class of long gravity waves (Flather and Kandhaker, 2000). Neglecting bottom friction, Airy (1842) and Lamb (1932) showed that the speed of a tidal wave c, with amplitude A small compared to the total water depth h (consisting of the water depth d and the wave amplitude η; see Fig. 36), and a water depth h small compared to the wavelength L, can be approximated by

$$c \approx \sqrt{g * h} \quad . \qquad \text{Equation 15}$$

This highlights that the tidal wave speed c only depends on the water depth h; g represents the gravitational acceleration. Thus it can be concluded that SLR increases c and thereby alters the "age of the tide" at particular places.

Observed total water level peaks (consisting of tide and surge) are known as high waters (HW) and low waters (LW). The differences between high and low waters are defined as tidal ranges. The latter can reach values of up to more than 3.0 m in the German Bight (see e.g. Wahl et al., 2011). Considering Equation 15, the tidal wave speed is proportional to the water depth, causing different wave speeds at high and low waters, with $c_{HW} > c_{LW}$. In consequence, the tide is deformed and differs more or less from the sinusoidal form. This asymmetric effect can be described by additional higher harmonics of the basic tides (van Rijn, 2010), where the evolution of these higher harmonics causes a

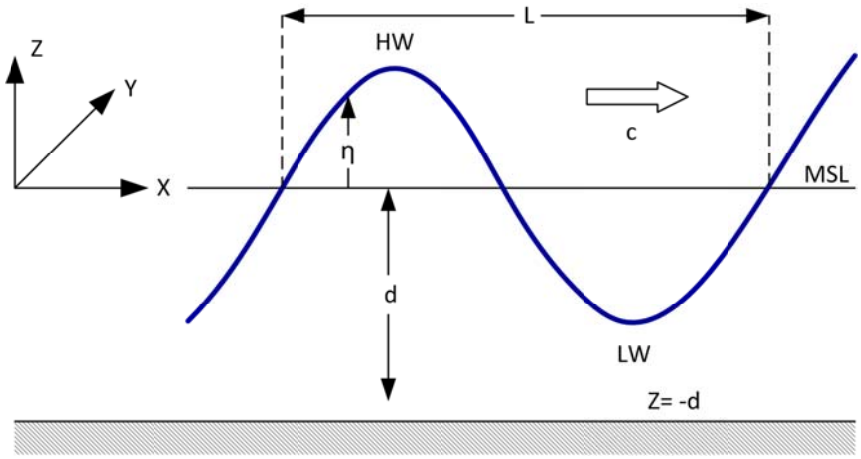


Fig. 36 Nomenclature used in the upcoming sections

reduction of the amplitudes of the basic tides. According to Equation 15, SLR would reduce the discrepancy between c_{HW} and c_{LW}, with less energy being transferred from basic tides to higher harmonics.

The deformation is furthermore amplified by shoaling, i.e. an increase in amplitudes of tidal waves entering shallower water (see e.g. van Rijn, 2010). This effect is described using the conservation of energy principle, a principle law of physics:

$$E(t) = \frac{1}{2} * \rho * g * A(t)^2 * \sqrt{g * h(t)} = const. \quad , \qquad \text{Equation 16}$$

with the water density ρ, the earth's gravitational attraction g, the water depth h, the tidal amplitude A, and considering the velocity component according to Equation 15. With ρ = g = const., Equation 16 reduces to

$$A(t)^2 * \sqrt{h(t)} = const. \quad , \qquad \text{Equation 17}$$

characterizing the ratio of amplitude A to water depth h. As an example, the amplitude of a wave travelling from depth h1 = 20 m to depth h2 = 10 m (h1/h2 = 2) is increased by a factor of f = 1.19 (example based on Pugh, 1981); considering a SLR of 0.5 m (h1/h2 = 1.95) slightly reduces this factor to f = 1.18. Thus, the amplitude of a tidal wave increases with decreasing depth, but the relative increase is slightly reduced when the ratio between the incident and amplified wave decreases (as e.g. due to SLR).

Both the astronomical tide and the surge are affected by frictional forces. North Sea tides are driven by co-oscillation with Northern Atlantic tides with an energy flux from

deep ocean (Atlantic) on to the shelf (North Sea). In shallow areas, bottom friction becomes increasingly important.

From numerical investigations of energy fluxes onto the northwest European continental shelf, Flather (1976) showed substantial energy dissipation occurring in this area. In the North Sea, most energy is dissipated in the shallower southern areas and the German Bight, having the potential to systematically change tidal patterns (e.g. reduction of tidal wave heights; Pugh, 1981). Bottom friction τ_b can be expressed as

$$\tau_b = k\rho_w u^2 \quad , \qquad \text{Equation 18}$$

with a friction parameter k, the water density ρ and the total current u (Flather and Kandhaker, 2000). From Equation 18, bottom friction is related to the square of the current speed. For instance, assuming an approximately constant water depth (h = const.), the tidal wave height reduces exponentially when moving further (see Fig. 37).

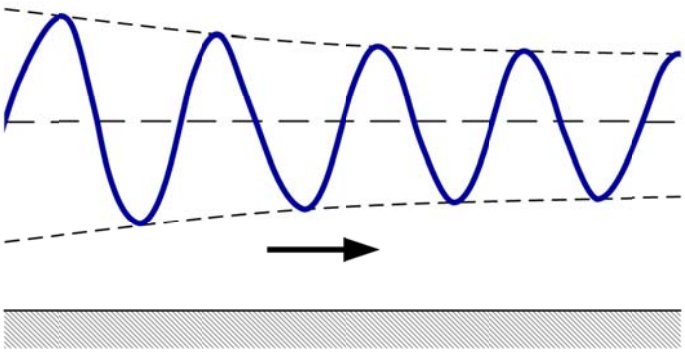


Fig. 37 Schematical illustration of the tidal wave deformation due to bottom friction

Extreme storm surges are a result of atmospheric pressure acting vertically on the sea surface and wind stress piling-up the water. The drag (or stress) on the sea surface due to the wind is defined as

$$\tau_S = C_D\rho_a W^2 \quad , \qquad \text{Equation 19}$$

and depends on a drag coefficient C_D, the wind speed W and the air density ρ_a (Pugh, 1981). The drag coefficient depends on wind speed and different empirical approaches exist to characterize the dependency. Wu (1982), for example, proposes a drag coefficient that reads as

$$C_D = \frac{0.8 + 0.0065 * W}{10^3} \quad , \qquad \text{Equation 20}$$

for any wind speed W between 0 and 50 ms^{-1}.

The motion of both storm surges and tides are described by the same dynamical equations (Flather and Kandhaker, 2000). As the horizontal sizes of tides and storm surges usually exceed the ocean depth significantly, the linear shallow water theory can be applied to estimate gravitational waves excited by external forcing (e.g. Nosov and Skachko, 2001).

The shallow water equations are a set of partial differential equations describing depth averaged flows under a free surface (Vater and Klein, 2009). The main model equations are based on 2D continuity and momentum equations. The continuity equation (which states a conservation of volume) can be derived from the conservation of mass equation in the Langrangian framework (which means that the observer follows an individual particle through space and time)

$$\frac{Dm}{Dt} = 0 \quad . \qquad \text{Equation 21}$$

In order to derive the continuity equation in the eulerian framework (i.e. the observer focuses on specific locations through which the fluid passes), the coordinate system used in Equation 21 is transferred and the Boussinesq approximation is applied. The latter is necessary to transfer the mass conservation equation into a volume conservation equation by eliminating the vertical coordinate. Hence, the continuity equation reads as

$$\frac{\partial h}{\partial t} + \frac{\partial uh}{\partial x} + \frac{\partial vh}{\partial y} = 0 \quad , \qquad \text{Equation 22}$$

describing the relation of surface elevation changes to net fluxes of water in or out of an control volume.

The momentum equation is based on Newton's 2nd Law,

$$\frac{D\vec{u}}{Dt} = \sum \vec{f} \quad , \qquad \text{Equation 23}$$

implying that acceleration (left side) equals the sum of all forces per mass unit $\vec{f}$ (right side) acting on each individual fluid particle. Referred to a control volume using Cartesian coordinates this reads as

$$\frac{D\vec{u}}{Dt} = \frac{\partial \vec{u}}{\partial t} + u\frac{\partial \vec{u}}{\partial x} + v\frac{\partial \vec{u}}{\partial y} \quad , \qquad \text{Equation 24}$$

Considering all forces, except coriolis forcing (which is simplified not considered here but needs to be taken into account in computer models) and applying the Reynolds-Stress-approximation yields a linear subset of equations of the form

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -g\frac{\partial z_s}{\partial x} - \frac{1}{h}\frac{\vec{\tau}_{b_x}}{\rho_w} + \frac{1}{h}\frac{\vec{\tau}_{s_x}}{\rho_a} + f_{x,tide} \quad , \qquad \text{Equation 25}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -g\frac{\partial z_s}{\partial y} - \frac{1}{h}\frac{\vec{\tau}_{b_y}}{\rho_w} + \frac{1}{h}\frac{\vec{\tau}_{s_y}}{\rho_a} + f_{y,tide} \quad , \qquad \text{Equation 26}$$

that are entered by nonlinearities through the quadratic term of bottom friction (see Equation 18; Le Provost, 1991) and the second and third term of Equation 25 and Equation 26 . The use of Equation 22, Equation 25 and Equation 26 enables to completely describe depth averaged flows. In Equation 25 and Equation 26, the bottom friction τ_b and wind stress τ_s are described by quadratic laws (Equation 18 and Equation 19). Forces per mass unit f from bottom stress f_b and surface stress f_s acting on each element are defined as

$$\vec{f_b} = \frac{1}{h}\frac{\vec{\tau}_b}{\rho_w} \quad , \qquad \text{Equation 27}$$

and

$$\vec{f_S} = \frac{1}{h}\frac{\vec{\tau}_S}{\rho_a} \quad , \qquad \text{Equation 28}$$

considering the reciprocal of the water depth h. These two equations indicate that increases in h yield reduced influences of f_b and f_s on total water levels.

Fig. 38 exemplarily highlights the relative impact of +0.5 m SLR on forces f induced by bottom friction and wind stress vs. the water depth. The abscissa shows the water depth related to the actual state, schematically illustrating how bottom friction and wind stress are acting in consequence of SLR at a fixed location. The figure indicates that

SLR reduces the impact on total water levels of both bottom friction and wind stress. The relative influence of SLR on wind stress forces is, however, larger than the accompanied influence on bottom friction.

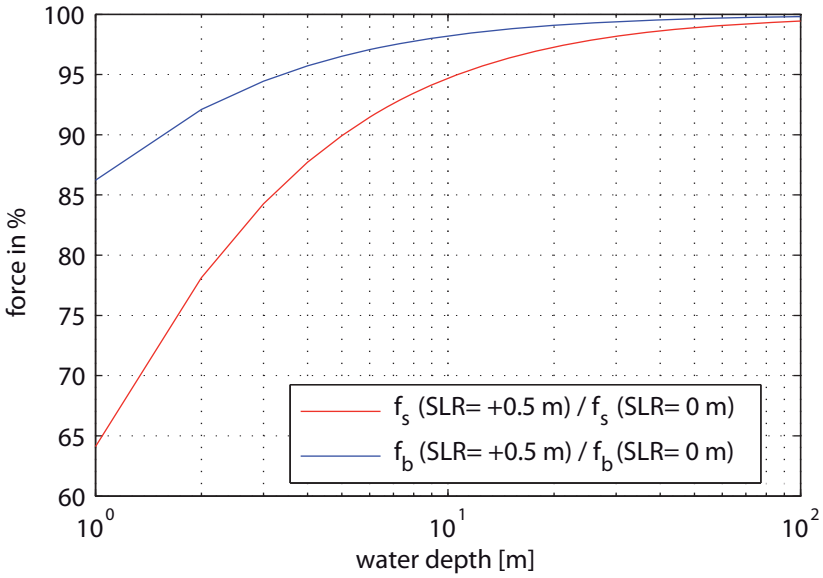


Fig. 38 Relative impact of a +0.5 m SLR on bottom friction (blue) and wind stress (red)

Both forces are impacting coastal water levels. On the one hand, the wind blowing parallel to a body of water imparts motion to the surface water towards (or away from) the coast causing increases (or decreases) in total water levels. From Equation 28 and Fig. 38 it is expected, that SLR causes a relative reduction in the wind setup. On the other hand, bottom friction causes water level decreases due to dissipation. In deeper water, the relative impact of bottom friction on water levels reduces. Thus, comparing the influence of increases in water depth (without considering 2nd order effects, such as changes in temperature, salinity, etc.) on both bottom friction and wind stress, the total water level is expected to change less than the SLR.

19 Methodology

19.1 Numerical model

The main purpose of Part III is to estimate the impact of SLR (e.g. as projected by 2100) on extreme water levels in the northern part of the German Bight. This is investigated using a numerical tide-surge model of the North Sea as described in Sect. 12.

19.2 Model specifications

The model specifications used for the present study are summarized in Tab. 5. The model was adjusted to two different configurations, either focusing on the northern German Bight (GBi) or the entire North Sea (ENS). The German Bight (GBi) configuration is intended to highlight the influence of SLR on extreme water levels in the shallow Wadden Sea areas. Results of the GBi configuration were outputted in equidistant spacing of 1 km along the coastline of Schleswig-Holstein (northern German Bight). The ENS configuration is envisaged to track model regions, where water level differences between the SLR and control runs (both described below) are largest. Results of the ENS model were outputted at ~1000 points with an equidistant spacing of 0.5 degrees, covering the entire model domain.

Tab. 5 Model configuration and setup used in Part III

configuration		German Bight (GBi)		entire North Sea (ENS)	
setup		Tide-Surge Run (TSR)	Tide Only Run (TOR)	Tide-Surge Run (TSR)	Tide Only Run (TOR)
run	control (0)	65 events	65 events	3 events	3 events
	SLR scenario (+)	65 events	65 events	3 events	3 events

In both configurations, the model was setup either using tidal forcing only, hereafter referred to as Tide Only Run (TOR), or combined tidal and atmospheric forcing, hereafter referred to as Tide-Surge Run (TSR). In the GBi configuration, both setups are used to calculate n = 65 extreme events which occurred between 1970 and 2009; this is referred to as the control run. The extreme events were identified, using the 99.7^{th} percentile threshold exceedances (see Sect. 9) of the Hörnum tide gauge record. In the ENS configuration, both setups were used to calculate the three largest events, which occurred on 01/1976, 11/1981, and 12/1999, and for the mean of the three events.

Additionally, a SLR scenario run was conducted to examine how this might affect extreme sea level events in the future. Regional MSL projections have recently been published in the AR5 (Church et al., 2013), but the model resolution is still relatively coarse for marginal seas such as the North Sea. To account for changes in MSL the global projections given in the AR5 (Church et al., 2013) are used, reporting that SLR will very likely exceed the observed rates during 1971 and 2010 due to increased ocean warming and increased loss of mass from glaciers and ice sheets. Based on climate projections in combination with process-based models they state that global MSL rise for 2081–2100 relative to 1986–2005 will likely be in the range of 0.26 to 0.82 m including uncertainties. This range covers four different Representative Concentration Pathways (RCPs) allowing for possible future climates, each of which is considered possible depending on how much greenhouse gas is emitted in the upcoming decades (see Fig. 35). For the SLR scenario the average of all four RCPs is used with z = 0.5 m and it is assumed that this is the global MSL rise by 2100. This is just an assumption which is used to simply estimate an appropriate SLR as input for the scenario run and it is not claimed that this estimate is elaborate. Here, an alternative could also be to use an arbitrary SLR.

Vertical land movements in the German Bight are considered as derived from the glacial isostatic adjustment (GIA) model of Peltier (2004) which were downloaded from the website of the Permanent Service for Mean Sea Level. In the study region, GIA amounts to ~0.44 mm/year on average (closest point to the study region: Lon. 8; Lat. 54.4). Assuming that vertical trends describe ongoing (at least until 2100) long-term processes, SLR projection and GIA influence can be summed up to a relative mean seal level (RMSL) rise scenario of +0.54 m. This projection is assumed to be valid for the study region; additional local effects from meteorological forcing should be captured by the model.

In the introduction it was highlighted that changes in atmospheric circulation and storminess are controversially discussed (see e.g. Weisse and von Storch, 2009 and references therein). In the light of these competing results, the SLR scenario used here assumes that wind conditions (speed and directions) do not change. Instead, all model runs using the TSR setup (including the SLR scenario run) are conducted using the same meteorological forcing from 1970 to 2009. However, changes in storminess may additionally increase (or decrease) storm surge water levels in the German Bight (Woth et al., 2006).

The SLR scenario run considers all boundary conditions to remain the same but assuming the MSL to have increased by an additional +0.54 m. This increase is added to the observed MSL between 1970 and 2009. Hence, the effects of SLR on storm surge water levels can directly be compared. In Fig. 39a, relative water depth changes in the model domain due to the SLR of +0.54 m compared to the 2009 MSL conditions are shown (simplified, the MSL index time series for the German Bight provided by Wahl et al. (2013) is used for the entire North Sea). The colorbar indicates depth changes on a percentage basis ranging from 0 to 20 %. Major changes occur along the southern North Sea, the Irish Sea, the English Channel and the south-eastern UK coast.

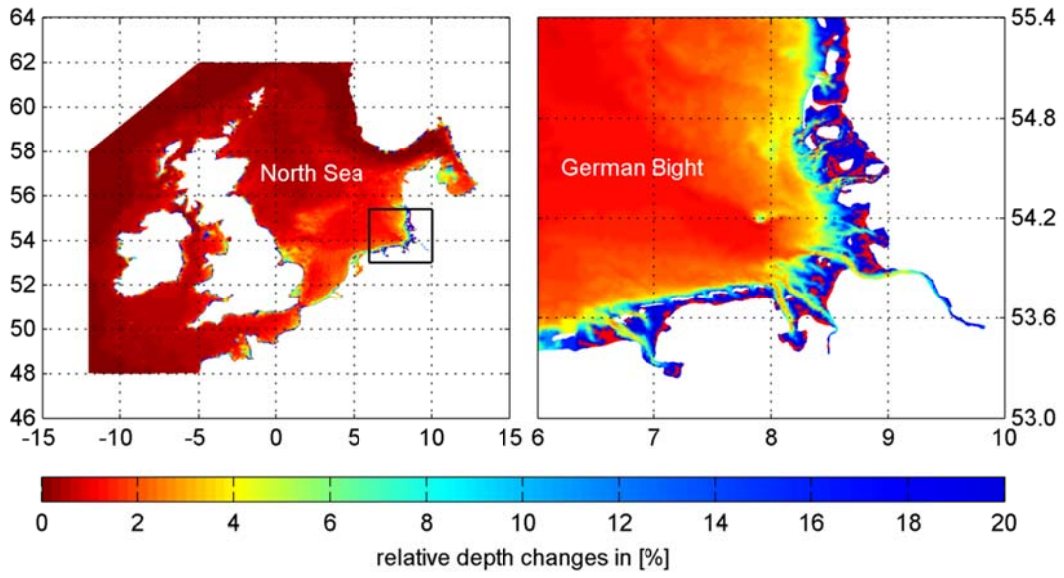


Fig. 39 Relative changes in water depth due to a SLR of 0.54 m for a) the entire model domain and b) the German Bight

Largest changes are found in the German Bight highlighted with a black rectangle. A detailed illustration of depth changes in this area is given in Fig. 39b. Some nearshore

tidal inlet areas show depth changes only slightly different from zero. Most nearshore areas, however, show remarkable relative depth changes of up to 20 %. In addition, SLR causes low lying areas formerly not inundated to be flooded, increasing the number of wet cells in the model domain.

19.3 Tidal analysis

Tide-generating forces may be expressed as series of harmonic constituents. At its open boundaries, the model used in this part is forced by eight primary tidal constituents (see Sect. 12) accounting for most of the tidal energy in the diurnal and semidiurnal frequency bands (Foreman et al., 1995). Due to nonlinearities arising from shallow water conditions, the model will additionally generate overtides (having multiple periods of the fundamental constituents) and compound tides (as linear combinations of multiple constituents), each of which are characterized by the amplitude and angular speed. With regards to the governing equations, the nonlinearities enter through the quadratic term of bottom friction and the advective terms. These nonlinear terms transfer momentum and energy from one frequency to another (Parker, 1991). Using a second order approximation of interactions, Le Provost (1976) showed the main origin of different nonlinear constituents, highlighting that shallow water effects are responsible for over- or compound tide generation whereas friction causes higher frequency and semidiurnal odd harmonics. Additionally, higher order interactions can significantly contribute to these constituents (Le Provost, 1991). Tidal water level oscillations can be described by a summation of N harmonic terms

$$tide = \sum_{n=1}^{N} A_n * cos(\sigma_n - g_n) \quad , \qquad \text{Equation 29}$$

with amplitudes A_n, angular speeds σ_n and phase lags g_n behind Equilibrium Tide at Greenwich (see e.g. Godin, 1972; Foreman, 1977; Pugh, 1987). This equation describes the different constituents resulting from linear and nonlinear processes. To analyze the impact of SLR on tidal response, a tidal analysis for each individual event was conducted using the Matlab t-tide package (Pawlowicz et al., 2002). Only constituents separated by at least a complete period from their neighboring constituents (Rayleigh criterion) and having a period of at least twice the sampling interval (Nyquist criterion, see e.g. Pugh, 1987) as given in Tab. 6 were considered.

The agreement of tidal constituents at Wittdün station derived from observational and modeled data for the control run is exemplarily shown in Fig. 40. Basic tides K_1 and M_2 and the higher harmonics (M_3, M_4, M_8) derived from modeled data are in line with the ones derived from observational data. M_6 and frictional tides $3MK_7$ and $2MK_5$ (red crosses in Fig. 40) are slightly underestimated (note the log-scaled axes) by the model, but this does not affect the overall conclusions of the present part of the thesis.

Tab. 6 Tidal constituents considered for analyses

Tide	K1	M2	M3	M4	M6	M8	2MK5	3MK2
Nonlinear*	no	no	no	yes	yes	yes	yes	yes

* Nonlinearities enter through the quadratic term of friction, spatial advection and mass conservation (Le Provost, 1991).

To investigate the influence of the spring-neap cycle on extreme water levels and their components, the Matlab t-tide package (Pawlowicz et al., 2002) was used for prediction, considering the angular speeds of M_2 and S_2. The cycle was calculated for the Cuxhaven water level record as it provides hourly values since 1918. The station was assumed to be representative for spring-neap cycles of all other stations along the German Bight.

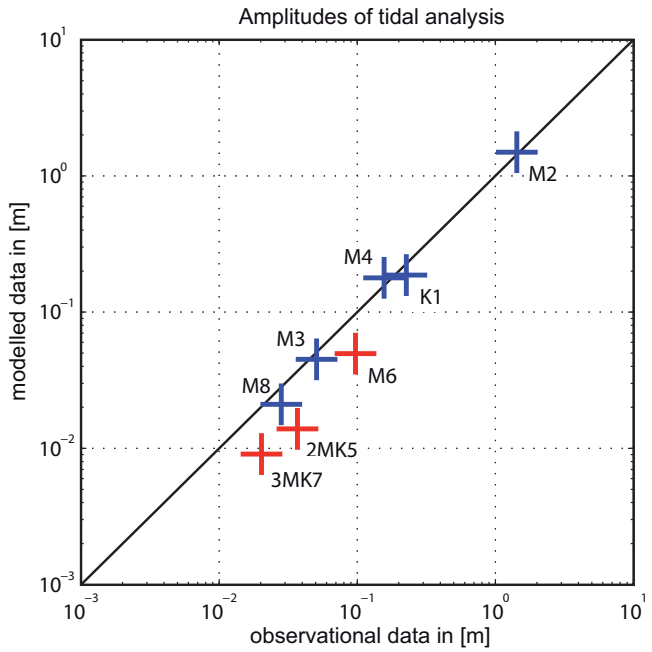


Fig. 40 Precision of tidal constituents derived from observational and modeled waterlevel records for periods where simultaneous records exist

19.4 Extreme water level assessment

To consistently compare return water level estimates from the control run and SLR scenario run, the recommendations for estimating extreme still water levels given in Sect. 9 are used. Differences in return levels are calculated as given in Fig. 41, highlighting that the differences are related to the theoretical distribution.

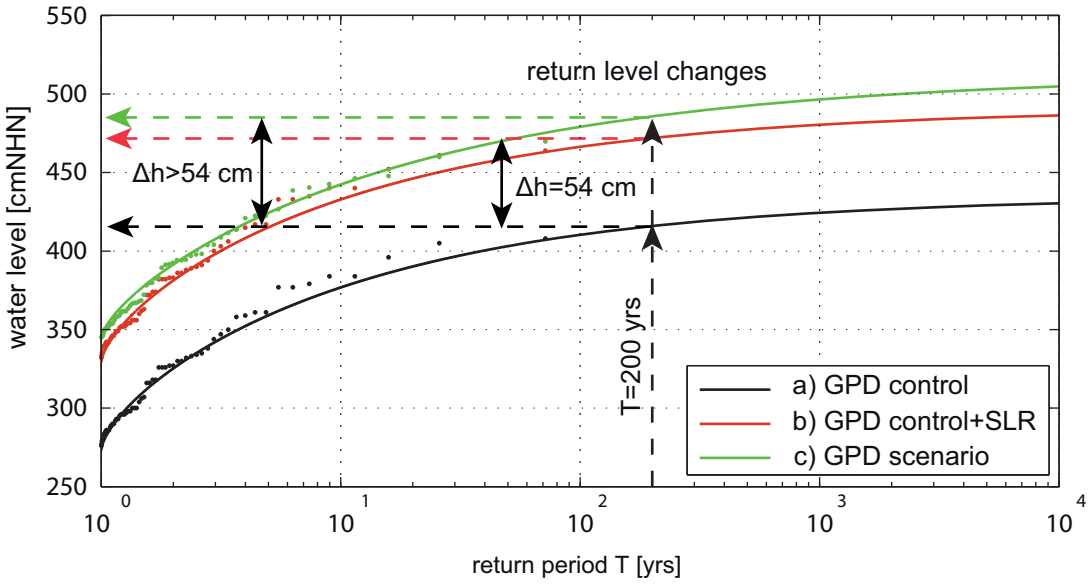


Fig. 41 Exemple of calculating the return level differences using a) observed water levels, b) considering the effect of SLR when using the MSL-Offset method and c) the numerical model simulations

20 Results

20.1 Changes in high water levels due to SLR

Changes in high water levels due to SLR are investigated using the German Bight (GBi) configuration with both the Tide-Surge Run (TSR) and the Tide Only Run (TOR) setup. In Fig. 42a and Fig. 42c, water level residuals are calculated as

$$h\,residuals = \sum_{i=1}^{65} hw_{i,(+)} - hw_{i,(0)} - SLR \quad , \qquad \text{Equation 30}$$

with the SLR scenario run high water levels $hw_{i,(+)}$, the control run high water levels $hw_{i,(0)}$ and a SLR of +0.54 m. The mean of all 65 events is shown. The colorbar encodes the height (h) residuals in cm for locations along the northern German Bight, where blue indicates reduced high water levels and red shows increased high water levels (i.e. blue denotes that not the full amount of SLR propagates into extreme water levels, whereas red denotes that changes in extreme water levels are larger than SLR alone); areas of residuals ~0 are highlighted in white. Insignificant residuals (not significantly different from zero based on 95% confidence intervals (CIs), i.e. CIs do not intersect with zero) are shown in black. Fig. 42a shows h residuals of the TSR to be generally above 0 cm with most locations showing significant positive changes in addition to SLR. Large h residuals are mostly found in the region bounded by latitudes 54.4 and 54.9, where the water is very shallow (Wadden Sea).

From Equation 28 it was expected, that water depth increases cause less surge. If there are no changes in the tides and/or no additional nonlinear influences occur, total storm surge water levels will increase less than SLR. The sensitivity study above, however, highlights increases in total storm surge water levels, which exceed the considered SLR. It is thus concluded that all h residuals can be attributed to tidal and nonlinear effects. To analyze how SLR alters tides in the study area, the TOR was conducted with the same

model configuration as before. In Fig. 42c, residuals of the TOR with color coding as above are shown. As before, all residuals are positive and the areas with largest residuals are also similar compared to the TSR. The residuals magnitude, however, is larger. In the eastern part of the region bounded by latitudes 54.4 and 54.9, the residuals are up to three times larger as compared to the TSR. This indicates that water level residuals in the study area are caused by nonlinear changes in the tidal component. In the TSR this is partly compensated by surge reduction due to increases in water depth. The spatial distribution of h residuals in TSR (Fig. 42a) is thus more uniform than in TOR (Fig. 42c); in the latter, largest h residuals gather between latitudes 54.5 and 54.7. The h residuals highlighted here are probably a result of reduced damping and deformation effects altering the tidal characteristic.

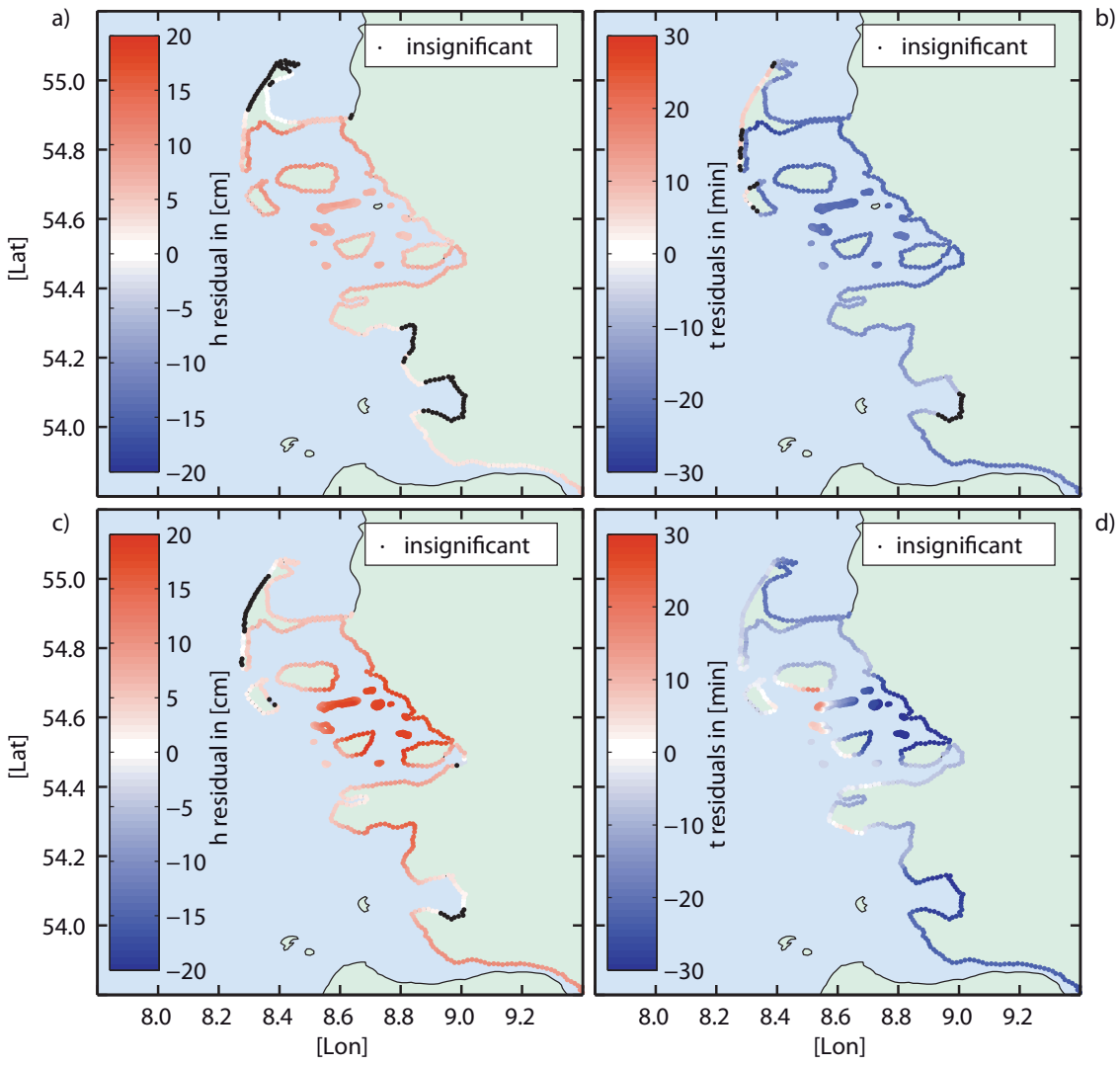


Fig. 42 a) h residuals of the TSR; b) t residuals of the TSR; c) h residuals of the TOR; d) t residuals of the TOR; the black dots show locations where changes were found to be insignificant

20.2 Changes in high water occurrence times due to SLR

Changes in high water occurrence times due to SLR are investigated using the GBi configuration with both setups (TSR and TOR). Occurrence time residuals are calculated as

$$t\ residuals = \sum_{i=1}^{65} t_{i,(+)} - t_{i,(0)} \quad , \qquad \text{Equation 31}$$

with the occurrence times of scenario run high water levels $t_{i,(+)}$ and the occurrence times of control run high water levels $t_{,(0)}$. Fig. 42b and Fig. 42d show time (t) residuals using TSR and TOR with color coding as above, except that units are changed to minutes. The t residuals show a pattern similar to that of the h residuals. In the TSR (Fig. 42b), t residuals are generally negative. In accordance with Equation 15, this is a result of increases in the tidal wave speed due to larger water depth. The same is found from the TOR (Fig. 42d), with larger t residuals in the eastern part (latitudes 54.4 to 54.8).

20.3 Spatial appearance of changes

To identify regions in the entire model domain where h and t residuals are large, the North Sea Tracking (NST) model configuration was used. Simulations were performed in two blocks, using either the TOR or the TSR setup. Using the TOR setup, h residuals of three events as well as the mean of all three events according to Equation 30 were calculated. These residuals are shown in Fig. 43, where the colorbar indicates h residuals between -6 and 6 cm and black dots highlight h residuals with a magnitude of $|h| > 2.5$ cm. The figure shows that the feedback of h residuals is spatially different, where the largest positive h residuals with $h > 2.5$ cm considering all three events occurred in the German Bight. Largest negative h residuals with $h < -2.5$ cm in all three events occurred on the western Scottish coast and in the western part of the English Channel.

In the remaining parts of the North Sea, minor changes ($|h| < 2.5$ cm) are found. A comparison of Fig. 43 and Fig. 39a indicates that the largest residuals occur in areas where relative depths changes are large. Apart from the German Bight, the bathymetry used for the model study is relatively coarse; this might cause inaccuracies in some regions. Similar results are found from the TSR setup (not shown here), whereas the magnitudes of changes are slightly lower. This is a result of the lowering effect of increasing water depths on

surges. It is thus concluded that major changes in tides are caused by locally varying depth changes altering shallow water effects. As a consequence, SLR causes an amplification of the tidal component in the German Bight, which in turn results in (nonlinear) increases of extreme water levels that are higher than the MSL rise alone.

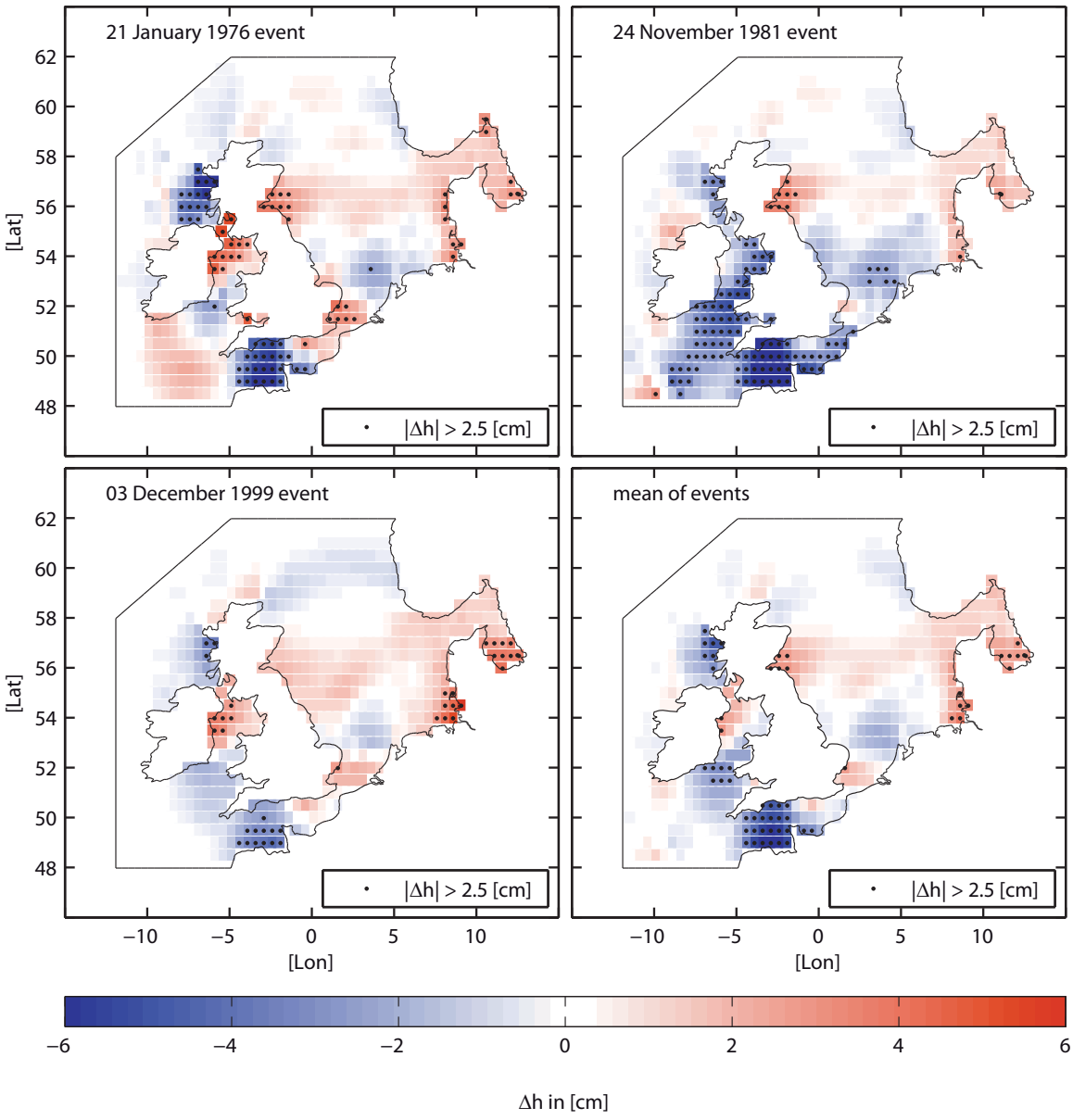


Fig. 43 h residuals of three major events in the considered period. The lower right subpanel shows the mean of h residuals from all three events

20.4 Changes in high water level distributions

The analyses above indicate that the mean tidal amplitudes are altered by rising water depths. However, for a reliable assessment of tidal changes it is of particular interest

not only to describe the mean values but also the corresponding spread. This is why frequency distributions of h residuals from the TOR setup (see Equation 30) at randomly distributed locations along the coastline of Schleswig-Holstein were calculated. In the centered panel of Fig. 44, mean h residuals resulting from the SLR scenario run vs. the control run along the Schleswig-Holstein coastline are shown. Blue crosses point the

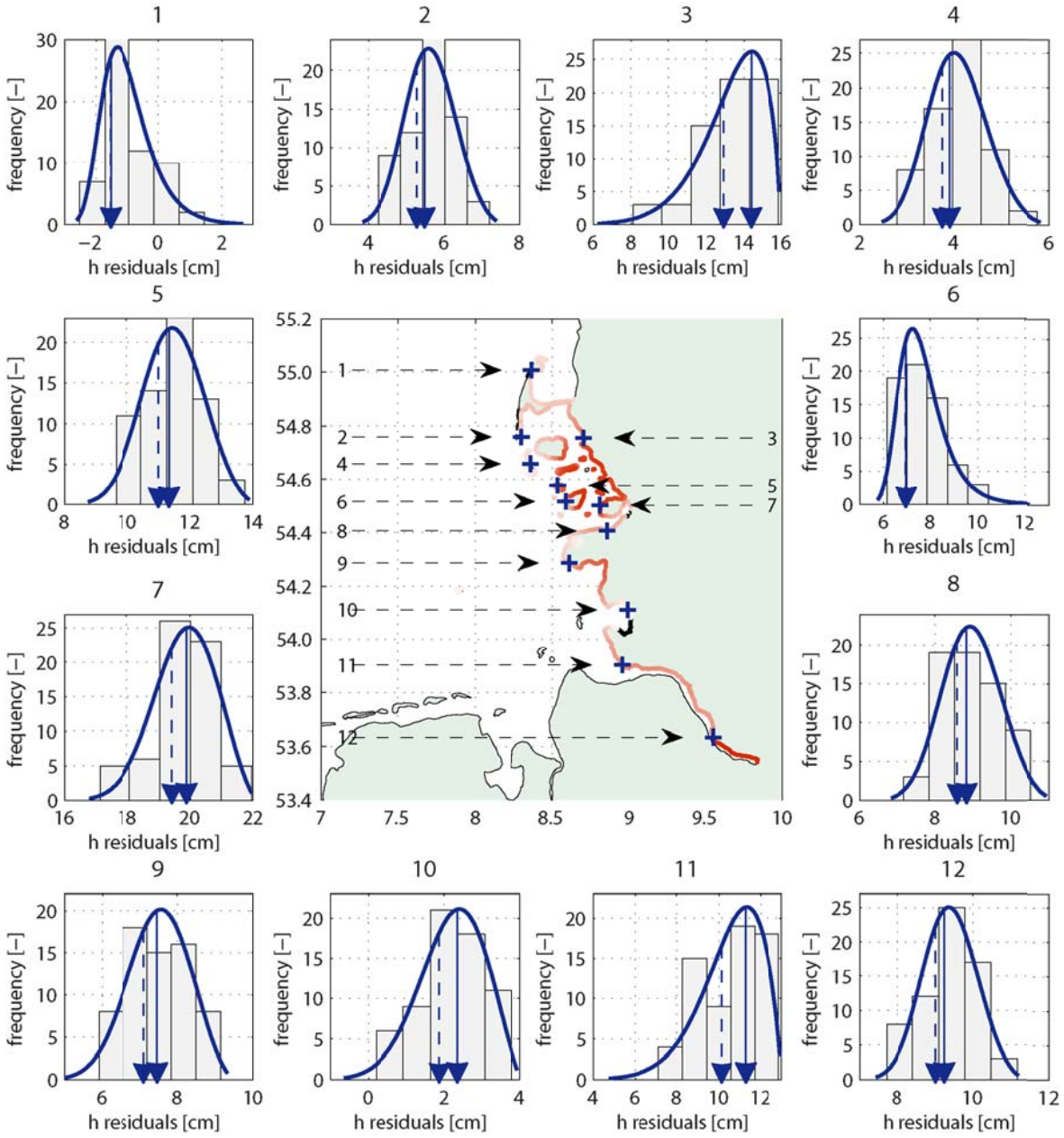


Fig. 44 Centered panel: Mean h residuals resulting from the SLR scenario run vs. the control run along the Schleswig-Holstein coastline and locations used in the outlying subpanels; outer subpanels: Distribution of differences from the scenario and the control run using TOR setup, with mean (solid lined arrow) and mode (dashed lined arrow)

locations shown in the outlying subpanels. All subpanels of Fig. 44 show the mean (solid lined arrow) and the modal value (i.e. the value that appears most often in a dataset; shown as dashed lined arrow) based on a parametric distribution function. Comparing all distributions shown here highlights the spatially varying impact of SLR on tidal high waters, with less impact on tidal high waters in the north-western part of the study area (e.g. locations 1, 2, 4) and larger impact on tidal high waters in the shallow areas located in the central study area (e.g. locations 3, 5, 7). Using a Kolmogorov-Smirnov test (KS-test) exhibits h residuals at locations 2, 4, 10 and 12 to be normally distributed, with higher location parameters in the SLR scenario run than in the control run. At the remainder, the h residuals follow a Generalized Extreme Value (GEV) distribution indicating systematic changes in tidal high water distributions; except location 1, all h residual distributions show increases in location parameters.

To investigate the origin of changes in distributions in more detail the focus is on location 3 (Dagebüll station). In Fig. 45a, the control run water levels are plotted against SLR scenario run water levels. Both datasets were derived using the TOR setup and were normalized to the mean value of the control run. The figure shows that the water levels from the SLR scenario run at location 3 are on average ~15.3 cm (additional to SLR) higher than those of the control run. A comparison of both datasets shows that the SLR scenario run water levels increase overproportionally by a factor of 1.08 compared to the control run water levels. This indicates that the differences between SLR scenario and control run water levels are higher when astronomically induced water levels are large and the other way round. Fig. 45c schematically shows the distributions of the two samples (from control and SLR scenario run); both samples are described by a GEV distribution with negative shape. This is reasonable as most the n = 65 events occurred close to spring tide whereas only few events occurred apart from spring tide, giving more weight to the higher water levels of the distribution. A comparison of both distributions shows that the considered SLR does hardly affect the shape (k) of the water level distribution but slightly increases the scale (σ). The largest effect of the SLR on the water level distribution is observed in the location parameter (μ) which increases by ~15 cm.

Fig. 45b is the same as Fig. 45a, but using TSR instead of TOR data. In this case, the water levels from the SLR scenario run at location 3 are on average ~11.6 cm (additional to SLR) higher than those of the control run also showing a linear dependency. In this case, however, the SLR scenario run water levels show a slight decrease by a factor

of 0.98 compared to the control run water levels. Fig. 45d again shows the distribution of the two samples. This time, both samples are described by a GEV distribution with positive shape highlighting the general behavior of extreme water levels which tend to occur less frequent. The shape and scale parameters are hardly affected by SLR but the location parameter increases by ~9.3 cm.

A comparison of Fig. 45c and Fig. 45d indicates that the largest increases in high water levels from SLR can be observed in the astronomically induced water levels. Meteorological forcing partly compensates this increase. However, the resulting total water level is still larger than could be expected from SLR alone. In this case, the decrease of surge heights due to deeper water is smaller than the water level increases due to alterations in tidal response, i.e. in frictional and shallow water effects.

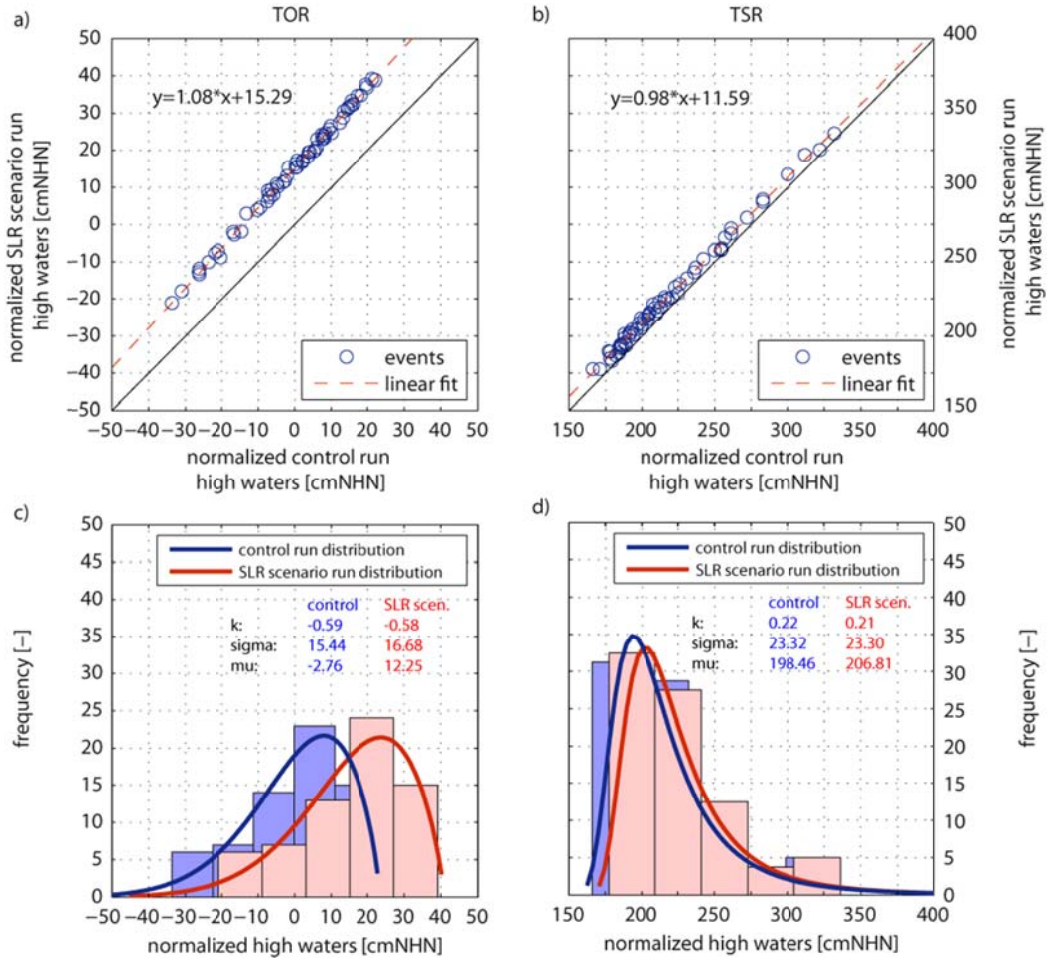


Fig. 45 a) Regression of the control and SLR scenario run tidal high waters of the TOR; b) Regression of the control and SLR scenario run high waters of the TSR; c) Distribution of control and SLR scenario run tidal high waters of the TOR; d) Distribution of control and SLR scenario run high waters of the TSR

20.5 Changes in tidal constituents

In Sect. 18 it has been shown, that changes in water depth impact the hydrodynamic characteristics and it is expected that SLR alters tidal amplitudes and phase lags (see Equation 29). In Fig. 46, tidal constituents from a tidal analysis (Pawlowicz, 2002) of all 65 SLR scenario run and control run events from the GBi configuration with TSR setup at Wittdün station (i.e. location 4 in Fig. 44) are shown. The subpanels a) to h) show all eight constituents that fulfilled the Rayleigh criterion. In total, the amplitudes of nonlinear constituents (all except K_1, M_2 and M_3; see Tab. 6) decrease when SLR is added, indicating that less energy is transferred from the largest tidal constituent M_2 to the nonlinear constituents. As a result, the amplitude of the M_2 constituent increases by a factor of 1.05, whereas the amplitudes of K_1 and M_3 remain nearly unaffected. With increasing water depths, a comparison of compound tide M_4 from control vs. M_4 from SLR scenario run shows a decrease by a factor of 0.7, most probably a result of reduced deformation effects (asymmetric effect, see e.g. Parker, 1991). A similar behavior is observed for the M_8 constituent, but the magnitude of change is smaller. Additionally, changes in bottom friction seem to decrease amplitudes of compound tides $2MK_5$ and $3MK_7$; only the M_6 overtide increases. The results for the frictional tides are less reliable as the amplitudes are small in comparison to the variability of changes (especially for $2MK_5$ and $3MK_7$). In subpanel i), all constituents are summed up according to Equation 29. This figure clearly shows that the ‘total amplitude’ of the tide increases with SLR.

The analyses described above were used to identify contributions of the nonlinear processes (friction, advection and continuity) altering the amplitudes of tidal constituents as a consequence of increasing water depths. From Sect. 18 it follows that larger water depths also cause increasing wave speeds, which in turn alter the tidal phase lag behind the equilibrium tide at Greenwich. However, these alterations are non-uniformly distributed along the entire frequency band of tidal records and may differ from one constituent to another. In tidal synthesis according to Equation 29, where a superposition of signals from different constituents is conducted, this may lead to non-linear changes in tidal water levels. In Fig. 47a, changes in the tidal phase lags at Wittdün station based on the same data as above are shown. The figure shows the mean phase lag changes of the eight tidal constituents as circles and the associated uncertainties as blue error bars; significant changes are highlighted as white-filled circles. The phase lags of all main constituents and

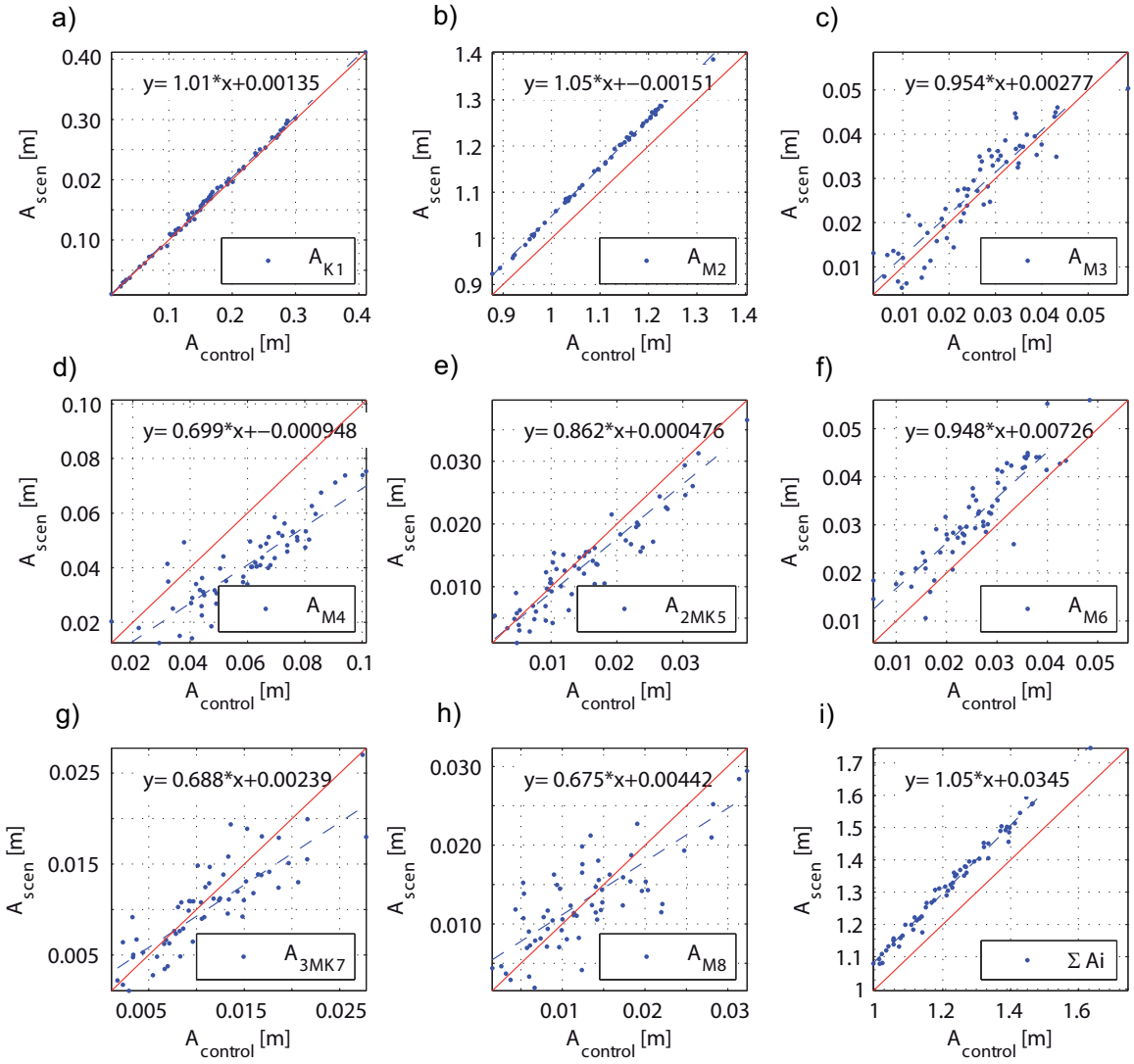


Fig. 46 Comparison of amplitudes from control vs. SLR scenario run. The subpanels a) to h) show individual constituents; subpanel i) shows the superposition of all constituents considered

overtides are reduced significantly, but the magnitude differs between constituents. The impact of changes in smaller amplitude constituents, as e.g. the compound tides considered here, is less and in this case not statistically significant. A comparison of the tidal synthesis based on the constituents derived from the control run (black curve) and the SLR scenario run considering changes in phases (blue curve) and changes in both phases and amplitudes (red curve) is shown in Fig. 47b. The tidal amplitudes and high waters at this particular station increase as a consequence of changes in the phase lags of individual constituents. When changes in both phase lags and amplitudes of the individual constituents are considered, the increase in high waters is even larger, but changes in the phase lags clearly dominate.

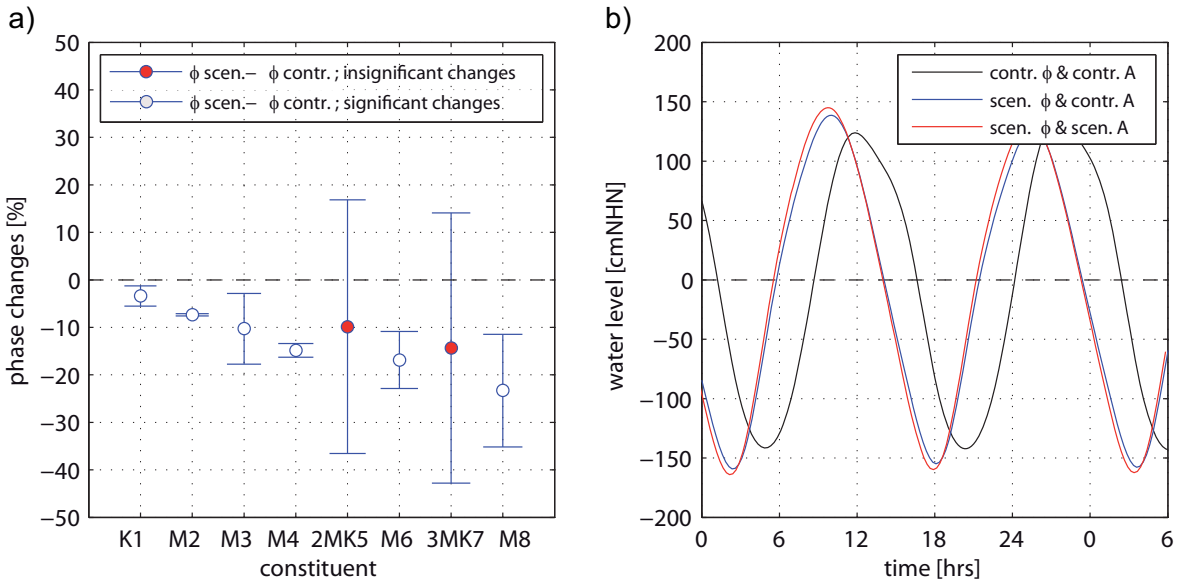


Fig. 47 a) Changes in phases and associated uncertainties. Significant changes are highlighted with white circles; b) Tidal synthesis according to Equation 29. The black curve shows the tidal signal using phases (Φ) and amplitudes (A) of the control run. The blue curve shows the tidal signal considering changes in phases obtained from the SLR scenario run; the red curve additionally considers changes in amplitudes from the SLR scenario run

20.6 Impact on EVA

From the discussion of the theoretical background and results presented above, it is obvious that changes in the water depth can alter the tidal propagation and high waters in the German Bight. This will also have an effect on return water levels derived from extreme value analysis and being of great relevance for design purposes. The return water levels shown in Fig. 48a for Wittdün station are based on the control run water levels (set (a); black), a linear superposition of SLR onto the control run water levels (set (b); blue) and the SLR scenario run water levels (set (c); red). The linear superposition assumes that SLR can be added linearly to derive future return water levels (i.e. the widely used MSL offset method, e.g. Hunter, 2010). Plotting positions (PLPs) are based on Gringorten's formula (see Sect. 5.5).

Differences in return water level estimates and PLPs based on data sets b) and c) are shown in Fig. 48b. The figure highlights that changes in return water level estimates due to SLR are nonlinear at Wittdün station. Smaller changes in the order of ~10 cm are found for exceedance probabilities between $P_E = 0.5$ [1/a] and $P_E = 0.1$ [1/a] and larger changes in the order of ~15 cm for exceedance probabilities between $P_E = 0.001$ [1/a] and $P_E = 0.005$ [1/a]. This nonlinear behavior can be explained with the differences visible in

the PLPs, where lower water levels which occur more frequently have – on average – slightly larger differences, whereas higher and less frequent water levels show slightly lower differences. The discussion and results presented above suggest that this is due to reduced surge generation with increasing water depth (see Sect. 18). As a consequence, the discrepancies between data sets (b) and (c) mainly cause changes in the distribution's (here GPD) shape parameter yielding a nonlinear feedback on return water level estimates.

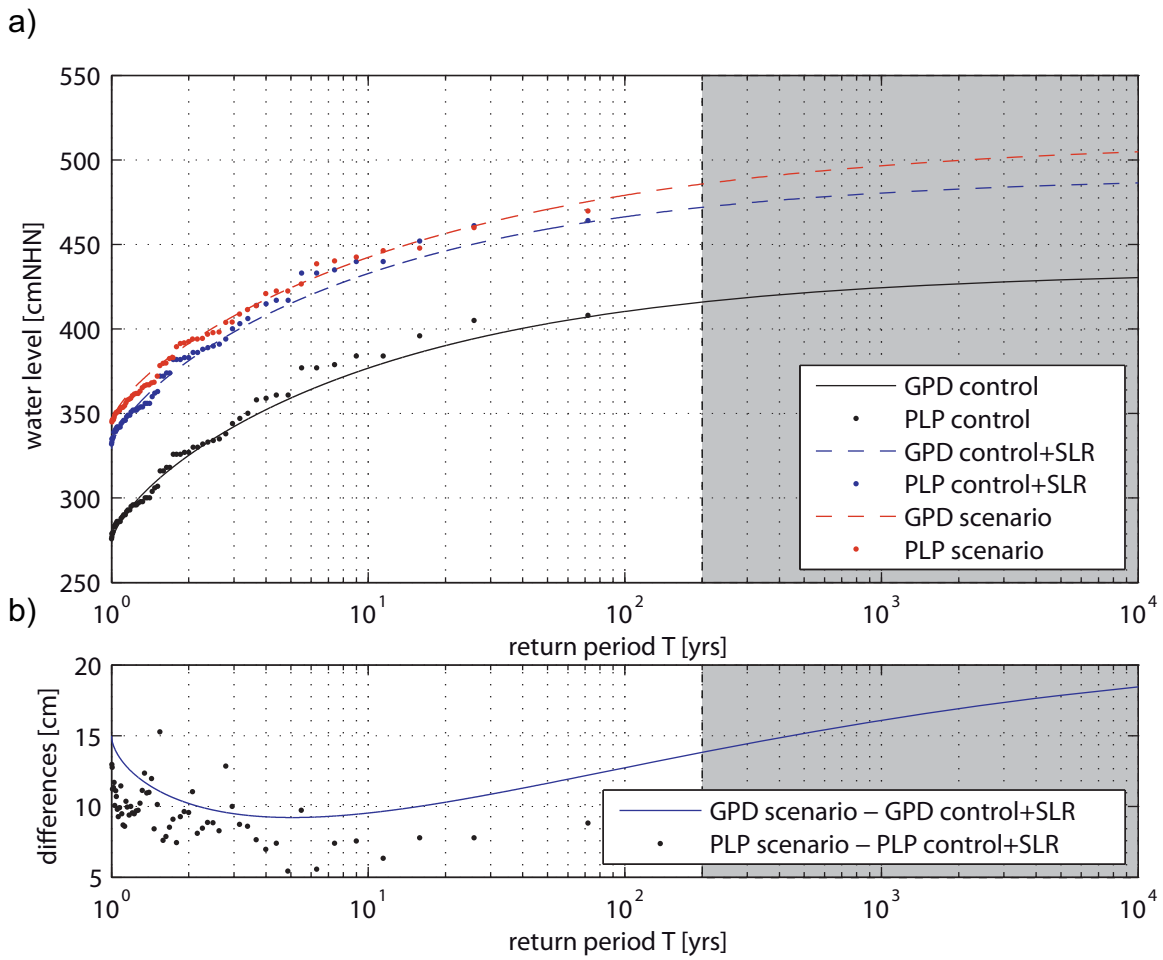


Fig. 48 (a) plotting positions (Gringorten) and best fit of the GPD using control (black), control +0.54 m (black) and SLR scenario (red) water levels. (b) Differences between SLR scenario and control +0.54 m GPD (blue curve) and plotting positions (black dots) at Wittdün station

This effect is not locally restricted but can be found along the entire coastline of Schleswig-Holstein, when return water levels with specific exceedance probabilities and derived from the control and SLR runs are compared. The magnitude of the nonlinear feedback differs within the investigation area. Fig. 49 shows the differences in exceedance probabilities P_E = {0.1; 0.02; 0.01; 0.005} at Schleswig-Holstein's coastline using data sets

(b) and (c). Hence, positive values indicate that SLR leads to changes in storm surge return water levels that are larger than the increase in MSL alone.

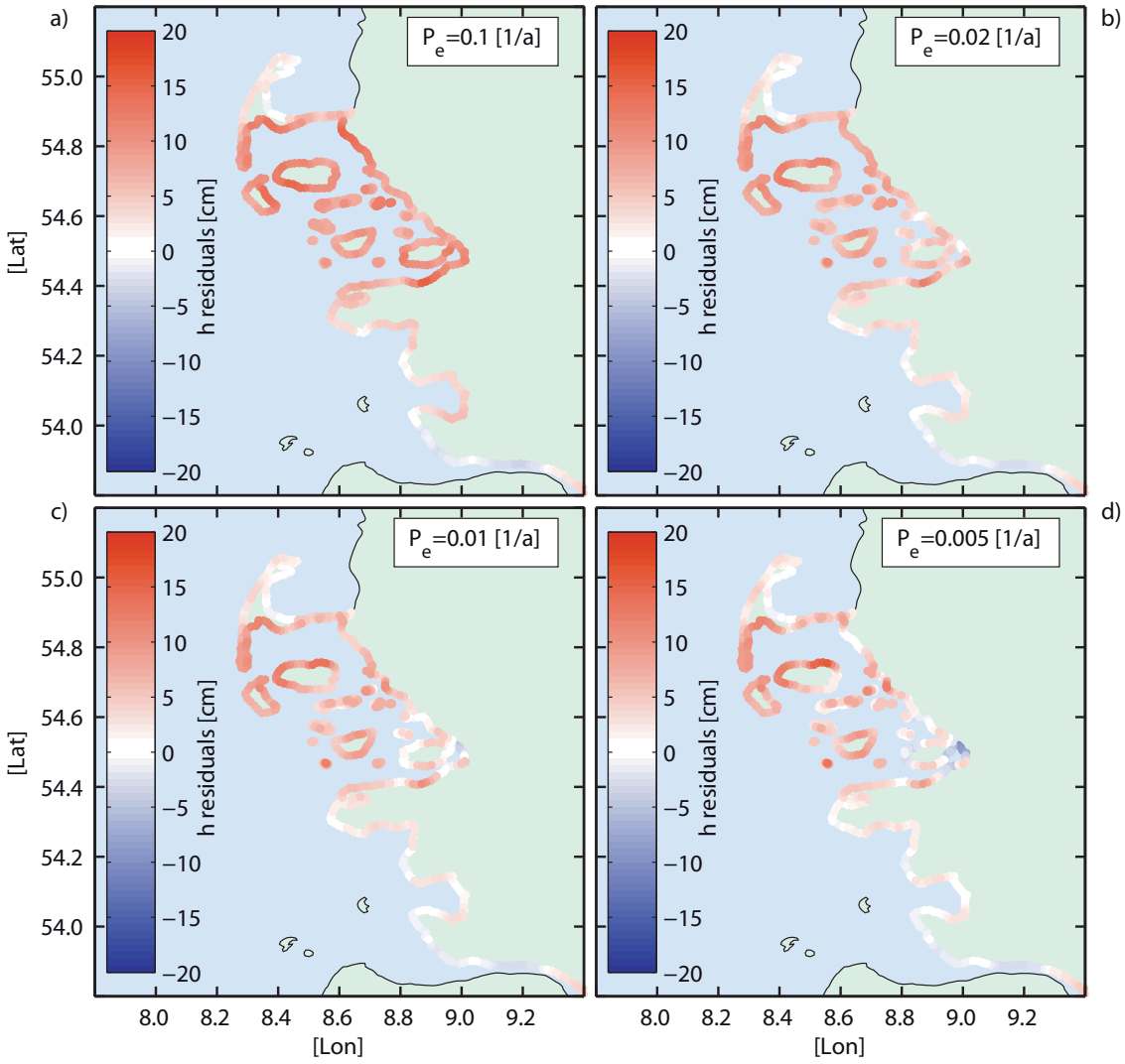


Fig. 49 Impact of SLR of +0.54 cm on return water levels in the northern German Bight; differences between control and SLR scenario runs are shown. The individual subpanels show the results for exceedance probabilities ranging from Pe = 0.1 to Pe = 0.005 [1/a]

Overall, largest nonlinear feedback on return water levels is found for higher exceedance probabilities, smaller changes for lower exceedance probabilities. However, return water levels behave different across the region. In the northern part of Föhr Island (Lon: 8.4 to 8.6; Lat: 54.6 to 54.8), return water levels show a strong increase for lower exceedance probabilities, e.g. the considered SLR is exceeded by ~17 cm for $P_E = 0.005$ [1/a]. Most of the Halligen show nearly constant increases in return water levels, exceeding the considered SLR for all exceedance probabilities by ~10 cm, with only a few

exceptions. Similar results are found for most parts of Amrum Island (Lon: 8.2 to 8.4; Lat: 54.6 to 54.7), the south-eastern part of Sylt Island (Lon: 8.2 to 8.4; Lat: 54.7 to 54.8), and parts of the mainland coastline, with return water levels exceeding the considered SLR by ~10 to 12 cm.

21 Summary and discussion

The impact of SLR on extreme water levels in the German Bight is investigated based on 65 extreme events that occurred between 1970 and 2009. The study uses a numerical tide-surge model covering the entire North Sea, whereas the bathymetric resolution is highest in the German Bight. The effect of SLR on extreme water levels is assessed by running the model for present-day conditions (no SLR) and for a SLR scenario of +0.54 m. At most locations, the model points to significant positive changes in extreme water levels relative to the MSL rise. The largest nonlinear increases in the order of +0.15 m occur in the shallow areas of the Wadden Sea. Two additional model runs were conducted (with and without SLR) where atmospheric forcing was neglected, i.e. only the effect of SLR on astronomical tidal water levels under constant atmospheric conditions were investigated. The results show that tidal high water changes from SLR are up to three times larger compared to the model run with meteorological forcing included. This indicates that water level residuals in the study area are mainly caused by nonlinear changes in the tidal components. Taking atmospheric forcing into account, by contrast, partly compensates tidal high water increases by surge reduction due to increases in the water depth. It is also shown that high water levels are shifted towards an earlier occurrence, and this is also mainly a result from water depth increases causing reduced shallow water effects and friction.

To track model regions where major water level changes occur, SLR induced changes in high waters of three extreme events covering the entire model domain are analyzed. The model shows a spatially different feedback; the largest positive residuals are mostly among those areas where relative depths changes are large (e.g. in the German Bight and in the most eastern part of the English Channel). These findings are coherent in the tide-surge and tide only runs, whereas the latter shows larger magnitudes. This is consistent with the theory that high water changes from SLR are strongly related to changes in shallow water effects. To explore the impact of SLR on the tidal response, a

tidal analysis was conducted for each individual event. The analyses point to changes in individual constituents, such as increases in the M_2 amplitude and decreases in frictional and overtides accompanied by less energy dissipation. Attributed effects are changes in phase lags of individual constituents leading to changes in the tidal modulation, which in turn results in an increase of tidal water levels.

The main purpose of Part III is to estimate the impact of SLR on the results from calculating return water levels using EVA in the northern part of the German Bight. The analyses highlight that this impact is nonlinear (with respect to exceedance probabilities) and spatially not coherent. In some locations, the increase in return water levels is nearly constant for all exceedance probabilities with values exceeding the considered SLR by 7-10 cm. Related to the large confidence bounds usually accompanied with extreme water level estimates, these changes are not significant. Designing coastal defense structures is usually based on the best fit of distribution functions. Hence, following those results, increases in design water levels are expected to be above the rate of SLR.

22 Key findings of Part III

The overall aim of Part III was to examine the impact of SLR on extreme water levels and the associated exceedance probabilities derived from extreme value statistics in the German Bight. By following the overall objective, Part III was also to assess changes in the high water levels, high water occurrence times, high water level distributions, tidal constituents, and the spatial distribution of the observed changes. The main outcomes can be summarized as follows:

- *Changes in high water levels are found to be significantly larger than the MSL rise.*
- *Changes in high waters are found to be largest in areas where the SLR scenario caused the largest relative water depths changes.*
- *High waters occur earlier as a consequence of SLR.*
- *SLR is found to alter phases and amplitudes of tidal constituents.*
- *In some locations, increases in return water levels are found to exceed the considered SLR by up to 15 cm (e.g. in the 10 years event).*
- *Changes in return water levels are spatially not coherent.*

These findings are valid for Schleswig-Holstein, a federal state that is located in the northern part of German Bight (see also Arns et al., under review b).

23 Overall summary and conclusions

The intention of this thesis was to investigate extreme water levels on regional and local scales, mainly focusing on the following three research questions:

1) *How to estimate comparable, robust and consistent return water levels on regional scales?*
2) *How to estimate return water levels in un-gauged areas?*
3) *How does sea level rise affect return water levels?*

Those three questions were investigated in three individual parts building up on each other. In the first part it was shown that the return water level assessment is inhomogeneous on the trans-regional to regional scale. In Germany for instance, the individual federal states use different methods to asses return water levels. Even if extreme value analyses are used, the procedures are subjectively setup as there is no objective guideline available. This is why the commonly used direct return level estimations methods were investigated and compared. The term '*direct'* implies the use of total observed water levels. Dixon and Tawn (1999) showed that the use of the BM method (see Sect. 5.2) can lead to a substantial underestimation of return water levels when tidal variations are large relative to surge variations. Allamano et al. (2011) presented similar results for the POT method (see Sect. 5.2). To overcome these issues, indirect methods have been introduced. The use of "indirect" methods is based on the idea of modeling the astronomical tidal and non-tidal components separately and inferring extreme sea levels as a combination of both (for more details see e.g. Pugh and Vassie, 1979, 1980; Middleton and Thompson, 1986; Walden et al., 1982; Tawn and Vassie, 1989; Tawn, 1992; Dixon and Tawn, 1994; Haigh et al., 2010b; Environment Agency, 2011). For some locations, however, the use of indirect methods is not yet feasible due to a lack of high resolution data. Haigh et al. (2010b) concluded that for the use of indirect methods, at least 20 years of records are required to derive reliable return water level estimates. In the German Bight, as an example, long datasets of high and low waters exist, whereas high resolution datasets

are available only since the late 1990s (Wahl et al., 2011). Using the Cuxhaven tide gauge dataset, Mudersbach et al. (2013) furthermore showed that from the mid 1950s changes in the ocean tides occurred. In the German Bight, the use of indirect methods is thus linked to unknown processes that can cause large uncertainties in tidal predictions and consequently in extreme water level probabilities derived with indirect methods. In the future it is intended to undertake a similar study using the main indirect methods.

Part I closes with recommendations for an objective use of direct estimation methods. These recommendations may help to overcome the inhomogeneous return level assessment in Germany. As the recommendations are primarily valid for Germany (i.e. regional scale) they should be verified for further locations around the globe. In addition, they are valid for the time period under investigation and need to be verified from time to time, e.g. after the occurrence of large storm surges or periodically.

In Part II it was shown that the data availability may be a limitation for extreme value analyses in some regions. This is why a new approach was introduced helping to transfer information from local to regional scales. The described approach was based on a numerical model hindcast, whereas the model output was corrected by a non-parametric transfer function (bias-correction). The model of Part II was calibrated against a set of tide gauge records covering the entire North Sea. Simplified, the model was calibrated globally, i.e. the entire model uses the same bed roughness. On the one hand, the accuracy of the raw model output might be improved by defining areas of different bed roughnesses; this may reduce the need for correcting the model output. On the other hand, this probably exacerbates the calibration exercise. The transfer function was created from observed and modeled water level information, i.e. it was only available at sites where tide gauge records existed. To obtain corrected water levels at un-gauged sites, the transfer function was interpolated to the surrounding areas. Following McMillan et al. (2011), the interpolation was performed with the inverse distance weighting (IDW) method without testing the performance of the IDW against other methods. To obtain the best possible results, the use of alternative interpolation techniques needs to be investigated in future.

In the first two parts of the thesis return levels were assessed under current (including the past) conditions. Part III by contrast investigated the impact of a 0.54 m SLR on future extreme water levels on regional to local scale. The study showed that future extreme water levels could be significantly larger than expected from SLR alone. These differences are mainly caused by changes in shallow water and frictional effects,

altering the tidal component of the total water levels. In Part III, a two-dimensional model is used, implying that frictional effects are only captured at the bottom. Effects of stratification and estuarine flow are not considered, i.e. interactions between tides and river discharges as well as internal tide generation are not represented (Foreman et al., 1995), but would possibly cause differences in the results. However, the differences are expected to be rather small and seasonal dependent (in the German Bight thermal stratification is only possible during the summer months) and would not affect the overall conclusions.

In Part III, changes in atmospheric forcing and in the bathymetric charts were neglected. This was important to obtain fundamental knowledge about the related processes and its drivers. The study is based on stationary bathymetric information, i.e. the seabed morphology does not change with time. In reality, however, the bathymetry is subject to a range of different forces such as currents that have the potential to change the seabed morphology (e.g. the deepening of channels, increases in Wadden areas). Recent investigations dealing with changes in the local bathymetries in the German Bight between 1982 and 2006 can be found in Kösters and Winter (2014). A more general (theorectical) statistical approach also dealing with morphdynamics the German Bight can e.g. be found in Siefert (1987). Such bathymetric changes may impact the findings of Part III as they can potentially cause both an amplification but also a reduction of the observed changes in extreme water levels (see e.g. Siefert and Lassen, 1987; Stengel and Zielke; 1994; Ferk, 1995).

The results of Part III may also help to solve an open issue that has been matter of debate throughout the last decades in Germany. Analyses of observational data show that tidal ranges in the German Bight increased (see e.g. Jensen et al., 1992; Jensen and Mudersbach, 2007) but up to date there is no published explanation available. The results presented here suggest that changes in tidal range may partly be attributed to SLR. In this study a SLR of +0.54 m was considered, causing increases in high waters of up to +0.15 m. Observational data, however, reveals higher increases in tidal ranges during periods with less SLR. Using Dagebüll station as an example, the actual SLR between 1970 and 2009 was only of the order of ~0.12 m accompanied by a tidal range increase of ~0.16 m. The latter was driven by both changes in tidal high waters and tidal low waters. Changes in tidal low waters were not considered in the present study but will be addressed in future investigations as this may have impacts, for example, on cooling systems of nuclear power stations in tidal rivers and for shipping. Considering the magnitudes of observed changes

indicates that phenomena other than SLR also contributed to the increases in local tidal ranges found in the German Bight. This may include morphologic changes from natural (e.g. erosion) and anthropogenic (e.g. dredging) impacts.

A combination of individual parts of this thesis can be used to objectively and reliably estimate regional to local return levels for current and future SLR conditions. These methodologies enable to estimate return levels for an entire coastline helping to obtain water level information in un-gauged areas. The results can be used for the design of coastal defenses of for risk analyses.

24 Recommendations for further research

Some parts of this PhD thesis originated as follow up of the two previous PhD theses at the fwu by Dr. Christoph Mudersbach and Dr. Thomas Wahl. Both authors examined different aspects regarding the application of extreme value statistics in coastal engineering and suggest a range of future activities that are needed for a better understanding and more reliable assessment of extremes. Among others, they highlight the need for to objectively assess the required time series length for conducting return level estimates, a regional assessment regarding possible changes in future extremes due to possible changes in the boundary conditions and to spatially extend return level assessments to provide information for ungauged areas.

The three main research questions that were posed in the beginning of this thesis as well as the related issues mostly accomplished those issues. Nevertheless, the work of the last years also suggests different possible future research activities which are discussed in the following paragraphs.

- Indirect assessment

  The return levels assessment in Part I was exclusively based on direct approaches as these are currently applied in Germany. These approaches are easy to handle but usually require long records of high waters. In the German Bight, long datasets of high and low waters exist. These records are needed to create extreme value samples and distribution functions are used to describe the stochastic behavior of the sample. However, water levels in the German Bight are a result of both, deterministic (astronomically, tide) and stochastic (atmospherically, surge) forcing. Allamano et al. (2011) showed that return water levels may substantially be underestimated when tidal variations are large relative to surge variations. The application of direct methods may thus introduce estimation errors. For upcoming

assessments it is thus important to test the application of indirect methods for estimating return water levels in the German Bight.

Indirect methods are usually based on a tide-surge separation and this is why they require high resolution records. A recent assessment by Haigh et al. (2010b) showed that a period of at least 20 years is needed to obtain reliable estimates with indirect methods. In the German Bight, high resolution datasets are available only since the late 1990s (Wahl et al., 2011). To perform a reliable indirect assessment, there are currently two options. These are:

c) *To wait until the high resolution records are long enough.*
d) *The use of an alternative approach to separate water level records into its components; the approach should not be based on long high resolution records.*

- Baltic Sea assessment

The entire work of this thesis needs to be validated and adjusted to the German (or entire) Baltic Sea. Specifically, the recommendations of Part I are probably not valid for the Baltic Sea as the largest observed storm surge in that area occurred in 1872. If this event is not taken into account, return levels are probably underestimated. Such an assessment would be the basis for the determination of return levels at gauged and, in a second step, for un-gauged sites as conducted in Part II of this thesis. SLR induced changes in extreme water levels from alterations in the tidal component as found in Part III are not expected for the Baltic Sea as tidal related water level changes are small. This, however, needs to be investigated in detail.

- North Sea assessment

To provide coastal protection of consistent standard, the return level assessment based on the recommended approach of Part I using regionalized data as proposed in Part II needs to be conducted for the entire North Sea or, in a first step, the entire German Bight including the coastline of Lower Saxony. This, however, requires that the bathymetric information currently used in the numerical model is updated using the best available bathymetric information in all parts of the model.

- Multivariate assessment

The assessments of Part I, II and III only focus on water levels (univariate data). In the future, all relevant loading parameters (e.g. water levels and wind waves) should jointly be investigated. In the German Bight, a similar study including waves and water levels has recently been conducted by Wahl (2012) but focusing on individual points rather than on an entire coastline. This study needs to be extended to the entire coastline providing information about the joint occurrence of water levels and waves. Especially the occurence of the latter is heavily dependent on local characteristics and would benefit from being provided on high spatial and temporal resolution.

The multivariate assessment has widely been used in scientific studies but is often difficult to apply in practical applications. Especially with respect to the design of coastal defenses it is difficult to interpret as it offers a range of design heights instead of one single value (see e.g. Grähler et al., 2013). An alternative might be to use a coupled (multi-stage) approach as e.g. conducted by Bender et al. (2014). In a first step, the joint occurrence probabilities from different input variables are assessed. In the next step, the consequences arising from all possible combinations having the same occurrence probability are considered. For identifying the event that has the largest consequences or impact (e.g. the largest inundation), the authors use a hydrodynamic numerical model using the results of the statistical model as input.

The coupling of statistical and dynamical models is, however, time consuming. As an alternative empirical instead of numerical models can be used for some applications. First attempts have been conducted by the author of this thesis within the research project “ZukunftHallig” combining water levels and possible wave impacts. Similar to the approach of Bender et al. (under review), the joint occurrence probabilities are assessed in a multivariate framework using numerically generated water levels and significant wave heights as input. The consequences, however, are assessed using a simple 2% wave runup formula given by Wassing (see e.g. van der Meer, and Stam, 1992), which is intended to describe the wave runup that is only exceeded by 2% of all occurences (EurOtop, 2007). The empirical assessment is conducted for all possible combinations having a certain return period (or probability of occurrence). The event causing the largest wave

runup can then be considered for design purposes (see the red cross in Fig. 50). However, to ensure that this approach can be considered for practical applications, further improvement (e.g. to include a more elaborate wave runup formula as given by the EuroOtop manual (EurOtop, 2007)) and verification is needed.

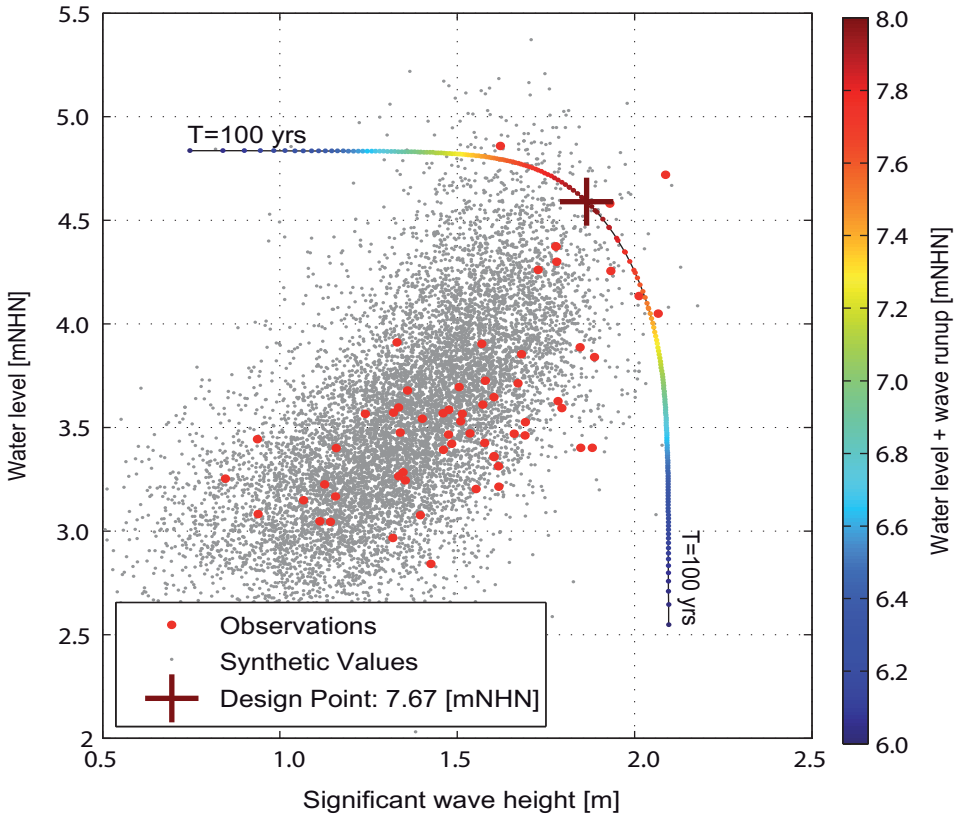


Fig. 50 Bivariate design of coastal defenses using the joint occurrence of water levels and significant wave heights as input

- Declustering

One of the fundamental assumptions in extreme value analysis is that the extreme value sample fulfills the IID criterion. For different reasons, water level data may sample preferentially, i.e. for some time periods higher densities of sample points occur than for other periods. An essential step in extreme value analyses is thus to detect and to decluster such samples. In Part I, the declustering was achieved by using the extremal index. A future strategy could be to develop a physically based approach using for example atmospheric patterns.

- Impact of morphodynamics on extreme water levels

The return water level assessment of Part II and III could be influenced by bathymetric changes. The main outcomes of the recently finished '*AufMod*' research project are dynamic bathymetries covering the entire German North Sea. The bathymetries consider annual changes between 1982 and 2006 and could be used instead of the stationary information. This enables to estimate the combined effect of bathymetric and MSL changes on extreme water levels and their assessment in the German Bight. To obtain more 'realistic' projections of future high- and return water levels, other changes (such as atmospheric, baroclinic, temperature) need to be taken into account.

References

Acero, F.J., García, J.A., Gallego, M.C., 2011. Peaks-over-threshold study of trends in extreme rainfall over the Iberian Peninsula. Journal of Climate 24, 1089–1105.

Acreman, M.C., Sinclair, C.D., 1986. Classification of drainage basins according to their physical characteristics: an application for flood frequency analysis in Scotland. Journal of Hydrology, Vol. 86, p. 365-380

AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., Sorooshian, S., 2013. Extremes in achanging climate. Detection, Analysis and Technology. Springer.

Aguilar, E., Auer, I., Brunet, M., Peterson, T.,Wieringa, J., 2003. Guidelines on climate metadata and homogenization. World meteorological organization. Secretariat of the World Meteorological Organization.

Airy, G.B., 1841. Tides and waves. Encyclopaedia Metropolitana (1817–1845), Mixed Sciences, Vol. 3.

Allamano, P., Laio, F., Claps, P., 2011. Effects of disregarding seasonality on the distribution of hydrological extremes. Hydrology and Earth System Sciences 15, 3207–3215.

Andersen, O. B., 1995. Global ocean tides from ERS 1 and TOPEX/POSEIDON altimetry, Journal of Geophysical Research, Volume 100, p. 25249-25259. DOI: 10.1029/95JC01389

Arns, A., Wahl, T., Haigh, I.D., Jensen, J., Pattiaratchi, C., 2013a. Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise, Coastal Engineering, Volume 81, p. 51-66, http://dx.doi.org/10.1016/j.coastaleng.2013.07.003.

Arns, A., Wahl, T., Dangendorf, S., Mudersbach, C., Jensen, J., 2013b. Ermittlung regionalisierter Extremwasserstände für die Schleswig-Holsteinische Nordseeküste. Hydrologie und Wasserbewirtschaftung, HW57, 2013, H. 6.

Arns, A., Wahl, T., Haigh, I.D., Jensen, J., under review a. Determining return water levels at un-gauged coastal sites: a case study for northern Germany, Ocean Dynamics, under review.

Arns, A., Wahl, T., Dangendorf, S., Jensen, J., under review b. The impact of sea level rise on extreme water levels in the northern part of the German Bight, submitted to: Coastal Engineering.

Balkema, A.A. and de Haan, L. (1974). Residual life time at great age, Ann. Probab., 2, 792–804.

Bardet, L., Duluc, C.M., Rebour, V., L`Her, J., 2011. Regional frequency analysis of extreme storm surges along the French coast. Nat. Hazards Earth Syst. Sci., 11, p. 1627-1639

Batstone, C., Lawless, M., Horsburgh, K., Blackman, D., Tawn, J., 2009. Calculating extreme sea level probabilities around complex coastlines. A best practice approach. Proceedings of the Irish National Hydrology Conference 2009.

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., Hunt, T., 2013. A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines. Ocean Engineering, Vol. 71, p. 28-39, http://dx.doi.org/10.1016/j.oceaneng.2013.02.003.

Bärring, L., von Storch, H., 2004. Scandinavian storminess since about 180. Geophysical Res. Letters, vol. 31, doi:10.1029/2004GL02044.

Baxter, P.J., 2005. The east coast Big Flood, 31 January–1 February 1953: a summary of the human disaster. Philos Trans R Soc A Math Phys Eng Sci 363, p. 1293–1312.

Becker, A., 1992. Methodische Aspekte der Regionalisierung. In. Regionalisierung in der Hydrologie. Deutsche Forschungsgemeinschaft. Mitt. XI der Senatskommision für Wasserforschung. VCH Verlagsgesellschaft, Weinheim.

Bender, J, Wahl, T., Müller, A., Jensen, J., under review. On Deriving Design Water Levels at River Confluences Using Copula Functions and Hydrodynamic Flow Models, submitted to: Water Res. Research.

Bernadara, P., Andreewsky, M., Benoit, M., 2011. Application of regional frequency analysis to the estimation of extreme storm surges. Journal of Geophysical Research, Vol. 116.

Bindoff, N.L. et al., 2007. Observations: oceanic climate change and sea level. In: Solomon, S. et al. (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, pp 385–432

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modelling: a review. Hydrological processes, vol. 9, pp. 251-190.

Bocchiola, D., De Michele, C., Rosso, R., 2003. Review of recent advances in index flood estimation. Hydrology and Earth System Sciences, Vol. 7, p. 283-296.

Brabson, B.B., Palutikof, J.P., 2000. Tests of the generalized Pareto distribution for predicting extreme wind speeds. Journal of Applied Meteorology 39, 1627–1640.

Bruss, G., Gönnert, G., Mayerle, R., 2010. Extreme scenarios at the German North Sea coast: a numerical model study. Proceedings of the 32nd International Conference on Coastal Engineering (ICCE) 2010.

Bütow, H., 1963. Die große Flut in Hamburg: Eine Chronik der Katastrophe vom Februar 1962. Verlag: Hansestadt, German.

Cartwright D. E., Tayler, R. J., 1971. New computations of the tide-generating potential. Geophysical Journal of the Royal Astronomical Society, 23, 45-74.

Castellarin, A., Burn, D.H., Brath, A., 2008. Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem? Journal of Hydrology, Vol. 160, p. 67-76

Choulakian, V., Stephens, M.A., 2001. Goodness-of-fit tests for the generalized Pareto distribution. Technometrics 43, 478–484.

Church, J.A., White, N.J., Coleman, R., Lambeck, K., Mitrovica, J.X., 2004. Estimates of the regional distribution of sea level rise over the 1950–2000 period. Journal of Climate 17, 2609–2625, http://dx.doi.org/10.1175/1520-0442(2004)017 < 2609:EOTRDO > 2.0.CO;2

Church, J.A., White, N.J., 2006. A 20th century acceleration in global sea-level rise. Geophysical Research Letters 33, L01602. http://dx.doi.org/10.1029/2005GL024826.

Church, J.A., White, N.J., Aarup, T.,Wilson, S.W., Woodworth, P.L., Domingues, C.M., Hunter, J.R., Lambeck, K., 2008. Understanding global sea levels: past, present and future. Sustainability Science 3, 9–22. http://dx.doi.org/10.1007/s11625-008-0042-4.

Church, J.A., Clark, P.U., Cazenave, A., Gregory, J.M., Jevrejeva, S., Levermann, A., Merrifield, M.A., Milne, G.A., Nerem, R.S., Nunn, P.D., Payne, A.J., Pfeffer, W.T., Stammer, D., Unnikrishnan, A.S., 2013. Sea Level Change. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Verlag, London.

Coles, S.G., Tawn, J.A., 2005. Bayesian modelling extreme surges on the UK east coast. Phil.Trans. Roy. Soc. A: Mathematical, Physical and Engineering Sciences. 363, 1387-1406,

Compo G.B., Whitaker, J.S., Sardeshmukh, P.D. et al. (2011). The twentieth century reanalysis project. Q J Roy Meteor Soc 137:1–28, DOI: 10.1002/qj.776

Cunnane, C., 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. Journal of Hydrology 18, 257–271.

Cunnane, C., 1987. Review of Statistical Models for Flood Frequency Estimation. Springer Netherlands.

Dalrymple, T., 1960. Flood-frequency analyses. U.S. Geological Survey Water Supply Paper 1543-A, 11.51

Danckwerth, K., 1963: Die Landkarten von Johannes Meier, Husum, aus der neuen Landesbeschreibung der zwei Herzogtümer Schleswig und Holstein. Hamburg.

Dangendorf, S., Wahl, T., Hein, H., Jensen, J., Mai, S., Mudersbach, C., 2012. Mean sea level variability and influence of the North Atlantic Oscillation on long-term trends in the German Bight. Water 4(1):170–195. doi:10.3390/w4010170

Dangendorf, S., Mudersbach, C., Jensen, J., Ganske, A., Heinrich, H., 2013a. Seasonal to decadal forcing of high water level percentiles in the German Bight throughout the last century. Ocean Dynamics, Vol. 63, Issue 5, pp. 533-548

Dangendorf, S., Wahl, T., Nilson, E., Klein, B., Jensen, J., 2013b. A new athmosperic proxy for sea level variability in the southeastern North Sea: observations and future ensemble projections. Climate Dynamics.

Dangendorf, S., Mudersbach, C., Wahl, T., Jensen, J., 2013c. Characteristics of intra-, inter-annual and decadal variability and the role of meteorological forcing: the long record of Cuxhaven. Ocean Dyn., Vol. 63(2–3):209–224

Dangendorf, S., Müller-Navarra, S., Jensen, J., Schenk, F., Wahl, T., Weisse, R., under review. North Sea storminess from a novel storm surge record since AD 1843, Journal of Climate.

Davison, A.C., Smith, R.L., 1990. Models for Exceedances over High Thresholds. 52, 393–442.

De,M., 2000. A new unbiased plotting position formula for Gumbel distribution. Stochastic Environmental Research and Risk Assessment 14, 1436–3240.

De Winter, R.C., Sterl, A., Ruessink, B.G., 2013. Wind extremes in the North Sea Basin under climate change: an ensemble study of 12 CMIP5 GCMs. J Geophys Res 118:1–12.

Dixon, M.J., Tawn, J.A., 1994. Extreme sea-levels at the UK A-class sites: site-by-site analyses. Proudman Oceanographic Laboratory Internal Document No. 65.

Dixon, M.J., Tawn, J.A., 1995. A semi-parametric model for multivariate extreme values. Statistics and Computing 5 (3), 215–252.

Dixon, M.J., Tawn, J.A., 1997. Estimation of extreme sea conditions. Spatial analyses for theUK coast. Final report. Ministry of Agriculture, Fisheries and Food.

Dixon, M.J., Tawn, J.A., 1999. The effect of non-stationarity on extreme sea-level estimation. Journal of the Royal Statistical Society: Series C: Applied Statistics 48, 135–151.

Dooge, J.C.I., 1982. Parameterization of hydrologic processes. in Eagleson, P.S. (Ed.). Land Surjace Processes in Atmospheric General Circulation Models. Cambridge University Press, London. pp. 243-288.

Dooge, J.C.I., 1986. Looking for hydrologic laws. Wat. Resour. Res., 22, 46s-58s.

Douglas, B.C., 1991. Global sea level rise. Journal of Geophysical Research 96, 6981–6992.

Dupuis, D.J., 1998. Exceedances over high thresholds: a guide to threshold selection. Extremes 1, 251–261.

Deutscher Verband für Wasserwirtschaft und Kulturbau e.V. (DVWK), 1999: Statistische Analsyse von Hochwasserabflüssen, DVWK-Merkblatt 251.

Deutsche Vereinigung für Wasserwirtschaft und Abfall e.V. (DWA), 2012: Ermittlung von Hochwasserwahrscheinlichkeiten, Merkblatt DWA-M 552.

DIN 19700-12:2004-07: Stauanlagen – Teil 12: Hochwasserrückhaltebecken (engl. Dam plants – Part 12: Flood retarding basins).

Dyck, S., Peschke, G., 1995. Grundlagen der Hydrologie. Verlag für Bauwesen, Berlin.

EurOtop, 2007. Wave Overtopping of Sea Defences and Related Structures: Assessment Manual. Die Küste, Heft 73.

Environment Agency, 2011. Coastal flood boundary conditions for UK mainland and islands. Project: SC060064/TR2: Design sea-levels. Environment Agency of England and Wales.

Ferk, U., 1995. Folgen eines beschleunigten Meeresspiegelanstiegs für die Wattgebiete der niedersächsischen Nordseeküste. Die Küste, Heft 57.

Ferro, Ch.A.T., Segers, J., 2003. Inference for clusters of extreme values. Journal of the RoyalStatistical Society: Series B 65, 545–556.

Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distribution of the largest and smallest member of a sample, Proc. Camb. Phil. Soc., 24, pp. 180-190.

Flather, R.A., 1976. A tidal model of the north-west European continental shelf. Memoires de la Society Royal des Sciences de Liege, 6 series, 10: 141-164.

Flather, R.A., Khandker, H., 1993. The storm surge problem and possible effects of sea level changes on coastal flooding in the Bay of Bengal. Climate and sea level change: observations, projections and implications. Warrick, R.A., Barrow, E.M., Wigley, T.M.L., Eds.Cambridge: Cambridge University Press.

Flather, R.A., Williams, J.A., 2000. Climate change effects on storm surges: methodologies and results. Climate scenarios for water-related and coastal impacts, (3), 66-72.

Foreman, M.G.G., 1977. Manual for tidal heights analysis and prediction. Canadian Pacific Marine Science Report No. 77-10, 10 pp.

Foreman, M.G.G., Walters, R.A., Henry, R.F, Keller, C.P., Dolling, A.G, 1995. A tidal model for eastern Juan de Fuca Strait and the southern Strait of Georgia JO. Journal of Geophysical Research: Oceans, Vol. 100, 721-740, DOI: 10.1029/94JC02721

Führböter, A., 1976. Über zeitliche Veränderungen der Wahrscheinlichkeiten von Extremsturmfluten an der Deutschen Nordseeküste. Mitteilungen des Leichtweiß-Instituts der TU Braunschweig, Heft 51.

Gerritsen, H., 2005. What happened in 1953? The Big Flood in the Netherlands in retrospect. Philos Trans R Soc A Math Phys Eng Sci 363, p. 1271–1291.

Godin, G., 1972. The Analysis of Tides. Liverpool University Press, 264 pp.

Gönnert, G., Müller, O., Gerkensmeier, B., 2013. Development of extreme storm surge events - results of the XtremRisK project. In: Klijn, F., Schwecckendiek, T. (Eds.),Comprehensive Flood Risk Management. Research for policy and practice.

Grabemann, I., Weisse, R., 2008. Climate change impact on extreme wave conditions in the North Sea: an ensemble study. Ocean Dynamics 58, 199–212.

Gram-Jensen, I.B., 1985. Sea floods. Danish Meteorological Institute, Climatological Papers No. 13. Copenhagen, 76 pp.

Grähler, B., van den Berg, M. J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De Baets, B., Verhoest, N.E.C., 2013. Multivariate retun periods in hydrology: a criticla and practical review focusing on synthetic design hydrograph estimation. Hydrol. Earth Syt. Sci., 17.

Grinsted, A., Moore, J.C., Jevrejeva, S., 2009. Reconstructing sea level from paleo and projected temperatures 200 to 2100 AD. Climate Dynamics. http://dx.doi.org/10.1007/s00382-008-0507-2 (published online first).

Gutknecht, D., 1991. Computer aided modeling for operational forecasting systems. Ann. Geeophysical Suppl. Vol. 9, C480.

Haigh, I.D., 2009. Extreme sea levels in the English Channel 1900 to 2006. (Ph.D. Thesis) University of Southampton, United Kingdom (198 pp.).

Haigh, I., Nicholls, R., Wells, N., 2010a. Assessing changes in extreme sea levels: application to the English Channel, 1900–2006. Continental Shelf Research 30, 1042–1055.

Haigh, I.D., Nicholls, R.,Wells, N., 2010b. A comparison of the main methods for estimating probabilities of extreme still water levels. Coastal Engineering 57, 838–849.

Haigh, I.D., Wijeratne, E.M.S., MacPherson, L.R., Pattiaratchi, C.B., George, S., 2013a. Estimating present day extreme total water level exceedance probabilities around the coastline of Australia: tides, extra-tropical storm surges and mean sea level. Climate Dynamics (in press-a). http://dx.doi.org/10.1007/s00382-012-1652-1.

Haigh, I.D., MacPherson, L.R., Mason, M.S., Wijeratne, E.M.S., Pattiaratchi, C.B., Crompton, R.P., George, S., 2013b. Estimating present day extreme total water level exceedance probabilities around the coastline of Australia: tropical cyclone induced storm surges. Climate Dynamics (in press-b). http://dx.doi.org/10.1007/s00382-012-1653-0.

Haußer, F., Luchko, Y., 2011. Mathematische Modellierung mit MATLAB. Eine praxisorientierte Einführung. Spektrum akademischer Verlag, Springer.

Hawkes, P.J., Gonzalez-Marco, D., Sánchez-Arcilla, A., Prinos, P., 2008. Best practice for the estimation of extremes: a review. Journal of Hydraulic Research 46, 324–332.

Hirsch, R.M., Helsel, D.R., Cohn, T.A., Gilroy, E.J., 1992. Statistical analysis of hydrological data. In: Maidment, D.R. (Ed.), Handbook of Hydrology.

Hollebrandse, F.A.P., 2005. Temporal development of the tidal range in the southern North Sea. Master thesis, TU Delft, Faculty of civil engineering and geosciences.

Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29, 339–349.

Hosking, J.R.M., 1990. L-moments analysis and estimation of distributions using linear combination of order statistics. Journal of the Royal Statistical Society, Vol. 52, p. 105-124

Hosking, J.R.M., Wallis, J.R., 1993. Some statistics useful in regional frequency analysis. Water Resources Research, Vol. 29, Issue 2, p. 271-281

Hosking, J.R.M.,Wallis, J.R., 1997. Regional Frequency Analysis. Cambridge University Press, Cambridge.

Hundt, C., 1955. Maßgebende Sturmfluthöhe für das Deichbestick der Schleswig-Holsteinischen Westküste. Die Küste, Jahrgang 3, Doppelheft 12.

Hunter, J., 2010. Estimating sea-level extremes under conditions of uncertain sea-level rise. Climatic Change, vol. 99, pp. 331-350.

IPCC, 2013. Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jansz, W., Blaeu, J., 2000. The Folio Atlases. Utrecht University, HES Punlishers BV.

Jay, D.A., 2009. Evolution of tidal amplitudes in the eastern Pacific Ocean. Geoph. Res. Letters, Vol. 36.

Jensen, J., 1984. Änderungen der mittleren Tidewasserstände an der Nordseeküste. Mitteilungen Leichtweiß-Institut der TU Braunschweig, Heft 83.

Jensen, J., 1985. Über instationäre Entwicklungen der Wasserstände an der deutschen Nordseeküste. Mitteilungen Leichtweiß-Institut der TU Braunschweig, Heft 88.

Jensen, J., Mügge, H.-E., Schönfeld, W., 1992. Analyse der Wasserstandsentwicklung und Tidedynamik in der Deutschen Bucht. Die Küste, Heft 53.

Jensen, J., Mudersbach, C., 2007. Zeitliche Änderungen in den Wasserstandszeitreihen an den Deutschen Küsten. In: Glaser R, Schenk W, Vogt J, Wießner R, Zepp H (eds) Berichte zur deutschen Landeskunde 81:99–112. Leipzig, Germany (in German).

Jensen, J.,Müller-Navarra, S., 2008. Storm surges on the German Coast. Die Küste, Heft 74, pp. 92–125.

Jevrejeva, S., Moore, J.C., Grinsted, A., 2010. How will sea level respond to changes in natural and anthropogenic forcings by 2100? Geophysical Research Letters, Vol. 37, http://dx.doi.org/10.1029/2010GL042947

Kamphuis, J.W., 2000. Introduction to Coastal Engineering and Management. World Scientific Publishing.

Katsman, C.A., Hazeleger, W., Drijfhout, S.S., van Oldenborgh, G.J., Burgers, G.J.H., Change, Climatic, 2008. Climate scenarios of sea level rise for the northeast Atlantic Ocean: a study including the effects of ocean dynamics and gravity changes induced by ice melt. Climatic Change 91 (3–4), 351–374.

Katsman, C.A., Sterl, A., Beersma, J.J., van den Brink, H.W., Hazeleger, W., et al., 2011. Exploring high-end scenarios for local sea level rise to develop flood protection strategies for a low-lying delta — the Netherlands as an example. Climatic Change. http://dx.doi.org/10.1007/s10584-011-0037-5.

Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Advances in Water Resources 25, 1287–1304.

Kauker, F., Langenberg, H., 2000. Two models for the climate change related development of sea levels in the North Sea - A comparison. Clim. Res., Vol. 15, pp. 61-67.

Kleeberg, H.B., 1992. Regionalisierung in der Hydrologie. DFG-Mitt. XI. VCH Verl. ges., Weinheim. 444 pp.

Kleeberg, H.B., Schumann, A.H., 2001: Zur Ableitung von Bemessungsabflüssen geringer Überschreitungswahrscheinlichkeiten. Wasserwirtschaft 91 (12), 609.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. in Geosc., Vol. 5, pp. 89-97.

Krueger, O., F. Schenk, F. Feser, R. Weisse, 2013a. Inconsistencies between long-term trends in storminess derived from the 20CR reanalysis and observations. J Clim, Vol. 26, pp. 868-874.

Kuncheva, L.I., 2007. A stability index for feature selection. Proceedings of the 25th IASTED International Multi-Conference: artificial intelligence and applications, pp. 390–395.

Lamb, H., 1932. Hydrodynamics. 6th edn. Cambridge University Press, 738 pp.

Lamb, H., 1991. Historic storms of the North Sea, British Isles and Northwest Europe. Cambridge University Press, Cambridge

Lang, M., Ouarda, T.B.M.J., Bobée, B., 1999. Towards operational guidelines for over threshold modelling. Journal of Hydrology 225, 103–117.

Le Provost, C., 1976. Theoretical analysis of the tidal wave's spectrum in shallow water areas. Memoires de la Society Royale des Sciences de Liege, Vol. 6, pp. 97.-111.

Le Provost, C., 1991. Generation of overtides and compound tides (review). Tidal Hydrodynamics, John Wiley, New York, pp. 269-296.

Letetrel, C., Marcos, M., Míguez, B.M., Woppelmann G., 2010. Sea level extremes in Marseille (NW Mediterranean) during 1885–2008, Cont. Shelf Res., Vol. 30, pp. 1267–1274.

Liese, R., Luck, G., 1978. Verfahren zum Nachweis von Veränderungen der Tidehochwasserstände in der Deutschen Bucht. Deutsche Gewässerkundliche Mitteilungen, Jahrgang. 22, Heft 5.

LKN (Landesbetrieb für Küstenschutz,Nationalpark und Meeresschutz Schleswig-Holstein), 2012. Generalplan Küstenschutz der Landes Schleswig-Holstein, Fortschreibung 2012. Entwurf für die Anhörung der Verbände.

Losada, I.J., Reguero, B.G., Méndez, F.J., Castanedo, S., Abascal, A.J., Minguez, R., 2013. Long-term changes in sea-level components in Latin America and the Caribbean. Global and Planetray Change, Vol. 104, p. 34-50

Lowe, J. A., Gregory, J. M., Flather, R. A., 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. Climate Dynamics, Vol. 18, pp. 179-188.

Lowe, J.A., Gregory, J.M., 2005. The effects of climate change on storm surges around the United Kingdom. Philosophical Transactions of the Royal Society A, 363, pp. 1313-1328.

Lowe, J. A. et al., 2009. UK Climate Projections Science Report: Marine and Coastal Projections. Report provided by Met. Office Hadley Centre, Exeter, UK.

LSBG (Landesbetriebe Straßen, Brücken und Gewässer), 2012. Ermittlung des Sturmflutbemessungswasserstandes für den öffentlichen Hochwasserschutz in Hamburg. Bericht Nr. 12/2012.

Luther, M.E., Merz, C.R., Scudder, J., Baig, S.R., Pralgo, J.L.T., Thompson, D., Gill, S., Hovis, G., 2007. Water Level Observations for Storm Surge. Marine Technology Society Journal, Vol. 41, Number 1, pp. 35-43.

Mai, V.C., van Gelder, P., Vrijling, J.K., 2007. Statistical methods to estimate extreme quantile values of the sea data. In: Proceedings of the Fifth International Symposium on Evironmental Hydraulics (ISEH-V), Tempe, Arizona, USA.

Marcos, M., Tsimplis, M.N., Shaw, A.G.P., 2009. Sea level extremes in southern Europe. J Geophys Res 114:C01007. doi:10.1029/2008JC004912

Mathiesen,M., Goda, Y., Hawkes, P.J., Mansard, E.,Martín,M.J., Peltier, E., Thompson, E.F., Van Vledder, G., 1994. Recommended practice for extreme wave analysis. 32 (6), 803–814.

McInnes, K.L., Macadam, I., Hubbert, G.D., O'Grady, J.G., 2009. A Modelling Approach for Estimating the Frequency of Sea Level Extremes and the Impact of Climate Change in Southeast Australia. Natural Hazards, Vol. 51, pp. 115-137.

McMillan, A., Batstone, C., Worth, D., Tawn, J., Horsburgh, K., Lawless, M., 2011. Coastal flood boundary conditions for UK mainland and islands. Project: SC060064/TR2: Design sea-levels. Environment Agency of England and Wales.

McRobie, A., Spencer, T., Gerritsen, H., 2005. The Big Flood: North Sea storm surge. Philos Trans R Soc A Math Phys Eng Sci 363, p. 1263–1270.

Meehl, G.A., et al., 2007. Global climate projections. In: Solomon, S., et al. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Méndez, F.J., Menéndez, M., Luceño, A., Losada, I.J., 2007. Analyzing monthly extreme sea levels with a time-dependent GEV model. Journal of Atmospheric and Oceanic Technology 24, 894–911.

Menéndez, M., Woodworth, P.L., 2010. Changes in extreme high water levels based on a quasi-global tide-gage dataset. J Geophys Res Oceans 115:C10011. doi:10.1029/2009JC005997

Merz, R., Blöschl, G., 2005. Flood frequency regionalisation – spatial proximity vs. catchment attributes. Journal of Hydrology, Volume 302, Issues 1–4, Pages 283-306

Middleton, Thompson, 1986. Return periods of extreme sea levels from short records. Journal of Geophysical Research 91, 11,707.

Morton, I.D., Bowers, J., 1996. Extreme value analysis in a multivariate offshore environment. Applied Ocean research 18, 303–317.

Mudelsee, M., Chirila, D., Deutschländer, T., et al., 2010. Climate Model Bias Correction und die Deutsche Anpassungsstrategie. Mitteilungen Deutsche Meteorologische Gesellschaft,
p. 2-7 (in German)

Mudersbach, C., Jensen, J., 2010. Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German north sea coastline. Journal Flood Risk Management 3 (1), 52–62. http://dx.doi.org/10.1111/j.1753-318X.2009.01054.x.

Mudersbach, C., Wahl, T., Haigh, I.D., Jensen, J., 2013. Trends in extreme high sea levels along the German north sea coastline compared to regional mean sea level changes. Continental Shelf Research (in press). ISSN 0278–4343. http://dx.doi.org/10.1016/j.csr.2013.06.016.

Müller, M., Cherniawsky, J.Y., Foreman, M.G.G., von Storch, J.S., under review. Seasonal variation of the M2 tide. Submitted to Ocean Dynamics.

Naveau, P.,Nogaj,M., Ammann, C., Yiou, P., Cooley, D., Jomelli, V., 2005. Statistical methods for the analysis of climate extremes. Comptes Rendus Geosciences 1013–1022.

Neves, C., Fraga-Alves, M.I., 2008. Testing extreme value conditions - an overvbiew and recent approaches. Revstat – Statistical Journal, Vol. 6, # 1, pp.83–100.

Nicholls, R.J.,Wong, P.P., Burkett, V.R., Codignotto, J.O., Hay, J.E., McLean, R.F., Ragoonaden, S., Woodroffe, C.D., 2007. Coastal systems and low-lying areas. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (Eds.), Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, pp. 315–356.

Nicholls, R.J.,Hanson, S., Herweijer, C., Patmore,N.,Hallegatte, S., Corfee-Morlot, J., Château, J.,Muir-Wood, R., 2008. Ranking port cities with high exposure and vulnerability to climate extremes: exposure estimates. OECD Environment Working Papers, No. 1. OECD Publishing. http://dx.doi.org/10.1787/011766488208.

Nicholls, R.J., Cazenave, A., 2010. Sea-level rise and its impact on coastal zones. Science 328(5985):1517–1520. doi:10.1126/science.1185782

Nicholls, R.J., 2011. Planning for the impacts of sea level rise. Oceanography 24 (2), 144–157. http://dx.doi.org/10.5670/oceanog.2011.34.

Nicholls, R.J., Hanson, S.E., Lowe, J.A.,Warrick, R.A., Lu, X., Long, A.J., Carter, T.R., 2011. Constructing sea-level scenarios for impact and adaptation assessment of coastal area: a guidance document. Supporting Material, Intergovernmental Panel on Climate Change Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA). (47 pp. Available online at http://www.ipcc-data.org/docs/Sea_Level_Scenario_Guidance_Oct2011.pdf).

Niedersächsischer Landesbetrieb fürWasserwirtschaft, Küsten- und Naturschutz (NLWKN), 2007. Generalplan Küstenschutz Niedersachsen/Bremen–Festland.

Nilsen, J.E.Ø., Drange, H., Richter, K., Jansen, E., Nesje, A., 2012. Changes in the past, present, and future sea level on the coast of Norway. NERSC Special Report 89, Bergen, Norway. 48 pp.

Nosov, M.A., Skachko, S.N., 2001. Nonlinear tsunami generation mechanism. Natural Hazards and Earth System Sciences, p. 251–253.

Parker, B.B., 1991. The relative importance of the various nonlinear mechanisms in a wide range of tidal interactions. In: B.B. Parker (ed.), Tidal Hydrodynamics, John Wiley and Sons, Inc. New York, pp. 237-268.

Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Comput Geosci 28(8):929–937. doi:10.1016/S0098-3004(02) 00013-4

Peel, M., Wang, Q.J., Vogel, R.M., McMahon, T.A., 2001. The utility of L-moment ratio diagrams for selecting a regional probability distribution. Hydrological Sciences Journal, Vol. 46, p.147-155.

Peltier, W.R., 2004. Global Glacial Isostasy and the Surface of the Ice-Age Earth: The ICE-5G(VM2) model and GRACE. Ann. Rev. Earth. Planet. Sci. 2004. 32,111-149.

Petersen, M., Rohde, H., 1977. Sturmflut: Die großen Fluten an den Küsten Schleswig-Holsteins und in der Elbe. Neumünster (Karl Wachholtz), 148 pp.

Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol., Vol. 99, p. 187-192

Pickands, J. (1975). Statistical inference using extreme order statistics, Ann. Statist., 3, 119-131.

Pilgrim, D.H., 1983. Some problems in transferring hydrological relationships between small and large drainage basins and between regions. Journal of Hydrology, Vol. 65, p. 49-72

Pirazzoli, P. A., Tomasin, A., 2003. Recent Near-Surface Wind Changes in the Central Mediterranean and Adriatic Areas. Int. Journal of Climatology, Vol. 23, pp. 963-973.

Pirazzoli, P.A., Costa, S., Dornbusch, U., Tomasin, A., 2006. Recent evolution of surge-related events and assessment of coastal flooding risk on the eastern coasts of the English Channel. Ocean Dynamics 56, 498-512.

Plüß, A., 2004. Küstenschutz in Hamburg - Nichtlineare Wechselwirkung der Tide auf Änderungen des Meeresspiegels im Übergangsbereich Küste/Ästuar am Beispiel der Elbe. In: Proc Klimaänderung Küstenschutz, pp. 129-138.

Proudman, J.F.R.S., 1941. The Effect of Coastal Friction on the Tides. Geophysical Journal International. Volume 5.

Pugh, D.T., Vassie, J.M., 1976. Tide and surge propagation off-shore in the Dowsing region of the North Sea. Deutsche Hydrographische Zeitschrift, 29, 163-213

Pugh, D. T., 1981. Tidal amphidrome movement and energy dissipation in the Irish Sea. Geophysical Journal of the Royal Astronomical Society, 67, 515-27.

Pugh, D.J., 1987. Tides, Surges and Mean Sea-Level. A Handbook for Engineers and Scientists. Wiley, Chichester, p. 472.

Pugh, D., Vassie, 1979. Extreme sea levels from tide and surge probability. Proceedings of the sixteenth Coastal Engineering Conference, 1, pp. 911–930.

Pugh, D., Vassie, 1980. Applications of the joint probability method for extreme sea level computations. Proceedings of the Institution of Civil Engineers, 69, pp. 959–975.

Quedens, G., 1992. Die Halligen. Breklumer Verlag, Breklum (in German)

Rahmstorf, S., 2007. A semi-empirical approach to projecting future sea-level rise.Science 315 (5810), 368-370. http://dx.doi.org/10.1126/science.1135456.

Rao, A.R., Hamed, K.H., 2000. Flood frequency analysis. CRC Press, New York.

Ray, R.D., 2006. Secular changes of the M2 tide in the Gulf of Maine. Cont. Shelf Res., Vol. 26, pp. 422-427.

Ray, R.D., 2009. Secular changes in the solar semidiurnal tide of the western North Atlantic Ocean. Geophysical Research Letters, Vol. 36.

Riecken, G., 1985. Die Halligen in Wandel. Husum-Druck- und Verl.-Ges.

Rosbjerg, D., Madsen, H., Rasmussen, P.F., 1992. Prediction in partial duration serieswith generalized Pareto-distributed exceedances. Water Resources Research 28, 3001–3010.

Sachs, L., 1997. Angewandte Statistik. Anwendung Statistischer Methoden. Springer Verlag, Berlin, Heidelberg, New York.

Schmidt, H., von Storch, H., 1993. German Bight storm analysed. Nature 365:791. doi:10.1038/365791a0

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., Zhang, X., 2012. Changes in climate extremes and their impacts on the natural physical environment. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Siefert, W., 1968. Sturmflutvorhersage für den Tidebereich der Elbe aus dem Verlauf der Windstaukurve in Cuxhaven. Mitteilungen des Franzius Institut der TU Hannover, Heft 30.

Siefert, W., 1987. Umsatz- und Bilanz-Analysen für das Küstenvorfeld der Deutschen Bucht. Grundlagen und erste Auswertungen. Die Küste, Heft 45.

Siefert, W., Lassen, H., 1987. Zum säkularen Verhalten der mittleren Watthöhen an ausgewählten Beispielen. Die Küste, Heft 45.

Siefert, W., 1998. Bemessungswasserstände 2085 A entlang der Elbe - Ergebnisse einer Überprüfung durch die Länderarbeitsgruppe nach 10 Jahren (1995/1996). Die Küste, Heft 60.

Siegismund, F., Schrum, C., 2001. Decadal changes in the wind forcing over the North Sea. Clim Res 18:39-45. doi:10.3354/cr018039

Simpson, M., Breili, K., Kierulf, H.P., Lysaker, D., Ouassou, M., Haug, E., 2012. Estimates of future sea-level changes for norway. Technical Report of the Norwegian Mapping Authority.

Smith, R.L., 1986. Extreme value theory based on the r largest annual events. Journal of Hydrology 86, 27-43.

Smith, R.L.,Weissman, I., 1994. Estimating the extremal index. Journal of the Royal StatisticalSociety: Series B: Methodological 56, 515–528.

Smith, J.M., Cialone, M.A., Wamsley, T.V., McAlpin, T.O., 2010. Potential impact of sea level rise on coastal surges in southeast Louisiana. Ocean Engineering, Vol. 37, pp. 37-47.

Smits, A., Klein-Tank, A.M.G., Können, G.P., 2005. Trends in storminess over the Netherlands 1962–2002. Int. Journal of Climatology, Vol. 25, pp. 1331-1344.

Soares, G., Scotto, M.G., 2004. Application of the r largest-order statistics for long-term predictions of significant wave height. Coastal Engineering 51, 387–394.

Sönnichsen, U., Moseberg, J., 1994. Wenn die Deiche brechen. Husum Druck- und Verlagsgesellschaft.

Soukissian, T.H., Arapi, P.M., 2011. The effect of declustering in the r-largest maxima model for the estimation of Hs-design values S. The Open Ocean Engineering Journal 4, 34-43.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency Analysis of Extreme Events. In: Handbook of Hydrology, ed. D.R. Maidment, McGraw-Hill, New-York, NY, pp. 18.1-18.66

Stengel, T., Zielke, W.,1994. Der Einfluß eines Meeresspiegelanstiegs auf Gezeiten und Sturmfluten in der Deutschen Bucht. Die Küste, H. 56.

Sterl, A., van den Brink, H., de Vries, H., Haarsma, R., van Meijgaard, E., 2009. An enesemble study of extreme storm surge related water levels in the North Sea in a changing climate, Ocean Science 5, p. 369-378.

Tawn, J.A., 1988. Bivariate extreme value theory: models and estimation. Biometrika 75, 397-415.

Tawn, J.A., 1992. Estimating probabilities of extreme sea-levels. Journal of the Royal Statistical Society: Series C: Applied Statistics 41, 77-93.

Tawn, J.A., Vassie, J.M., 1989. Extreme sea levels: the joint probabilities method revisited and revised. Proceedings of the Institution of Civil Engineers 87, 42-442.

Thompson, K.R., Bernier, N.B., Chan, P., 2009. Extreme sea levels, coastal flooding and climate change with a focus on Atlantic Canada. Natural Hazards 51, 139-150.

Tsimplis, M.N., Shaw, A.G.P., 2010. Seasonal sea level extremes in the Mediterranean Sea and at the Atlantic European coasts. Nat Hazards Earth Syst Sci 10:1457-1475. doi:10.5194/nhess-10-1457-2010.

Ullmann, A., Pirazzoli, P.A., Tomasin, A., 2007. Sea surges in Camargue (French Mediterranean coast): evolution and statistical analysis on the 20th century. Continental Shelf Research, Vol. 27, pp. 922-934.

Ullmann, A., Monbaliu, J., 2010. Changes in atmospheric circulation over the North Atlantic and sea surge variations along the Belgian coast during the 20th century. Int. J. Climatol., Vol. 30, pp. 558-568.

van der Meer, J.W., Stam, C.M., 1992. Wave runup on smooth and rock slopes of coastal structures. Journal of Waterway Port Coastal and Ocean Engineering, Vol. 118 (5).

van Gelder, P., Nykov, M.N., 1998. Regional frequency analysis of extreme water levels along the dutch coast using L-moments: A preliminary study. Proceedings of the 8th international probabilistic workshop, Szcecin, Poland.

Van Rijn, L.C., 2010. Tidal phenomena in the Scheldt Estuary. Report, Deltares.

Vater, S., Klein, R., 2009. Stability of a Cartesian grid projection method for zero Froude number shallow water flows. Num. Math., Vol. 113, p. 123-161

Vermeer, M., Rahmstorf, S., 2009. Global sea level linked to global temperature. http://dx.doi.org/10.1073/pnas.0907765106 .

Viglione, A., Laio, F., Claps, P., 2006. A comparison of homogeneity tests for regional frequency analysis. Water Resources Research, Vol. 43, Issue 3.

von Storch, H., Reichardt, H., 1997. A scenario of storm surge statistics for the German Bight at the expected time of doubled atmospheric carbon dioxide concentration, J. Climate 10, p. 2653-2662

von Storch, H., Woth, K., 2006. Storm surges - the case of Hamburg, Germany. ESSP OSC panel session on ‘‘GEC, natural disasters, and their implications for human security in coastal urban areas’’. Available online: http://www.safecoast.nl/editor/databank/File/hamburg-storms.pdf. Accessed 26 Jan 2012

von Storch, H., Woth, K., 2008. Storm surges, perspectives and options. Sustain Sci 3(1):33–43. doi:10.1007/s11625-008-0044-2

von Storch, H., Doerffner, J., Meinke, I., 2009. Die deutsche Nordseeküste und der Klimawandel. Hamburger Symposium Geographie, Band 1, Hamburg 2009, pp. 9-22.

von Storch, H., 2012. Storm Surges: Phenomena, Forecasting and Scenarios of Change. Proceeding of the International Union of Theoretical and Applied Mechanics (IUTAM).

Vrijling, J.K., Kanning, W., Kok, M., Jonkman, S.N., 2007: Designing robust coastal structures. In: Coastal structures 2007, Proceedings of the 5th coastal structures international conference, Venice, Italy.

Wahl, T., Jensen, J., Frank, T., 2010. On analysing sea level rise in the German Bight since 1844. Nat Hazards Earth Syst Sci 10:171–179. doi:10.5194/nhess-10-171-2010

Wahl, T., Jensen, J., Frank, T., Haigh, I.D., 2011. Improved estimates of mean sea level changes in the German Bight over the last 166 years. Ocean Dyn 61(5):701–715. doi:10.1007/s10236-011-0383-x

Wahl. T., 2012. Statistical Methods to assess the hydrodynamic boundary conditions for risk based design approaches in coastal engineering - Methods and application to the German North Sea coastline. Mitteilungen des Forschungsinstituts Wasser und Umwelt der Universität Siegen. Heft 4.

Wahl, T., Haigh, I.D., Woodworth, P.L., Albrecht, F., Dillingh, D., Jensen, J., Nicholls, R.J., Weisse, R., Wöppelmann, G., 2013. Observed mean sea level changes around the North Sea coastline from 1800 to present, Earth-Science Reviews, Vol.124, pp. 51-67, http://dx.doi.org/10.1016/j.earscirev.2013.05.003.

Walden, A.T., Prescott, P., Webber, N.B., 1982. An alternative approach to the joint probability method for extreme high sea level computations. Coastal Engineering 6, 71–82.

Wang, W., Van Gelder, P.H.A.J.M., Vrijling, J.K., 2005. Trend and stationarity analysis for stream flow processes of rivers in western Europe in the 20th century. IWA International Conference on Water Economics, Statistics, and Finance, pp. 451–461.

Wang, X.L., Feng, Y., Compo, G.P., Zwiers, F.W., Allan, R. J., Swail, V.R., Sardeshmukh, P. D., 2013. Is the storminess in the twentieth century reanalysis really inconsistent with observations? - A reply to the comment by Krüger et al. (2013), Geophys. Res. Abstr., Vol. 15.

WASA Group, 1998. Changing waves and storms in the Northeast Atlantic? Bull Am Meteorol Soc 79:741–760. doi:10.1175/1520-0477(1998)079 < 0741:CWASIT > 2.0.CO;2

Wemelsfelder, P. J., 1939. Wetmatigheden in het optreden van stormvloeden. De Ingenieur, No. 9.

Weiss, J., Bernadara, P., Benoit, M., 2013. A method to identify and form homogeneous regions for regional frequency analysis of extreme skew storm surges. Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards (EVAN2013), Siegen, Germany

Weisse, R., von Storch, H., 2009. Marine Climate and Climate Change: Storms, Wind Waves and Storm Surges. Springer-Verlag, Berlin Heidelberg New York.

Weisse, R., von Storch, H., Niemeyer, H.D., Knaack, H., 2011. Changing North Sea storm surge climate: An increasing hazard? Ocean & Coastal Management, Vol. 68, pp. 58-68.

Wiltshire, S.E., 1985. Grouping basins for regional flood frequency analysis. Hydrological Sciences Journal, Vol. 30, Issue 1, doi: 10.1080/02626668509490976, pp. 151-159

Wiltshire, S.E., 1986. Identification of Homogeneous Regions for Flood Frequency Analysis. Journal of Hydraulics, Vol. 84, pp. 287-307

Willmot, C.J., 1981. On the validation of models. Physical Geography, Vol. 2, p. 184–194

Woodworth, P.L., Blackman, D.L., 2004. Evidence for systematic changes in extreme high waters since the mid-1970s. Journal of Climate 17, 1190–1197.

Woodworth, P.L., Teferle, F.N., Bingley, R.M., Shennan, I., Williams, S.D.P., 2009. Trends in UK mean sea level revisited. Geophysical Journal International 176 (22), 19–30.

Woodworth, P.L., 2010: A survey of recent changes in the main components of the ocean tide. Continental Shelf Research, vol. 30, pp. 1680–1691.

Woodworth, P.L., Menéndez, M., Gehrels, W.R., 2011. Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels. Surv Geophys 32:603–618

Wöppelmann, G., Pouvreau, N., Coulomb, A., Simon, B., Woodworth, P.L., 2008. Tide gage datum continuity at Brest since 1711: France's longest sea-level record. Geophysical Research Letters 35, L22605. http://dx.doi.org/10.1029/2008GL035783.

Wöppelmann, G., Letetrel, C., Santamaria, A., Bouin, M.-N., Collilieux, X., Altamimi, Z., Williams, S.D.P., Miguez, B.M., 2009. Rates of sea-level change over the past century in a geocentric reference frame. Geophysical Research Letters 36, L12607. http://dx.doi.org/10.1029/2009GL038720.

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. Ocean Dyn., Vol. 56, pp. 3-15.

Wu, Jin, 1988: Wind-Stress Coefficients at Light Winds. J. Atmos. Oceanic Technol., 5, 885–888. doi:

Zachary, S., Feld, G.,Ward, G.,Wolfram, J., 1998. Multivariate extrapolation in the offshore environment. Applied Ocean research 20 (5), 273–295.

Zhang, X., Ge, W., 2009. A new method to choose the threshold in the POT model. Proceedings to the 1st International Conference on Information Science and, Engineering (ICISE2009).

A Appendix

Most analyses conducted within this thesis were performed and automated in MATLAB, a standard technical computing environment provided by MathWorks. In the following, the flow charts of the most important analyses and approaches are briefly introduced. The compilation is not supposed to be a user manual but should provide an overview of how the analyses have been conducted.

Regional extreme water level assessment

The flow chart in Fig. 51 shows how the regional extreme value assessment was conducted in this thesis. As input, numerically derived water levels (as given in the numerical model flow chart) were used and bias-corrected. The bias correction stage is shown in a seperate flow chart.

Numerical model

The flow chart in Fig. 52 shows how the model was setup, highlighting the different forcing and the spatial data that were considered.

Bias correction

The bias correction in this thesis mainly consists of a transfer function, which helps to transfer the modeled into the observed variable (see Fig. 53). Hereafter, the bias corrected water levels are similar to those observed. A detailed description of the bias correction stage can be found in Sect. 12.3 (Part II).

Return level assessment

Return levels are assessed either using modeled and bias-corrected or observed water levels (see Fig. 54). The procedure was developed in Part I of this thesis.

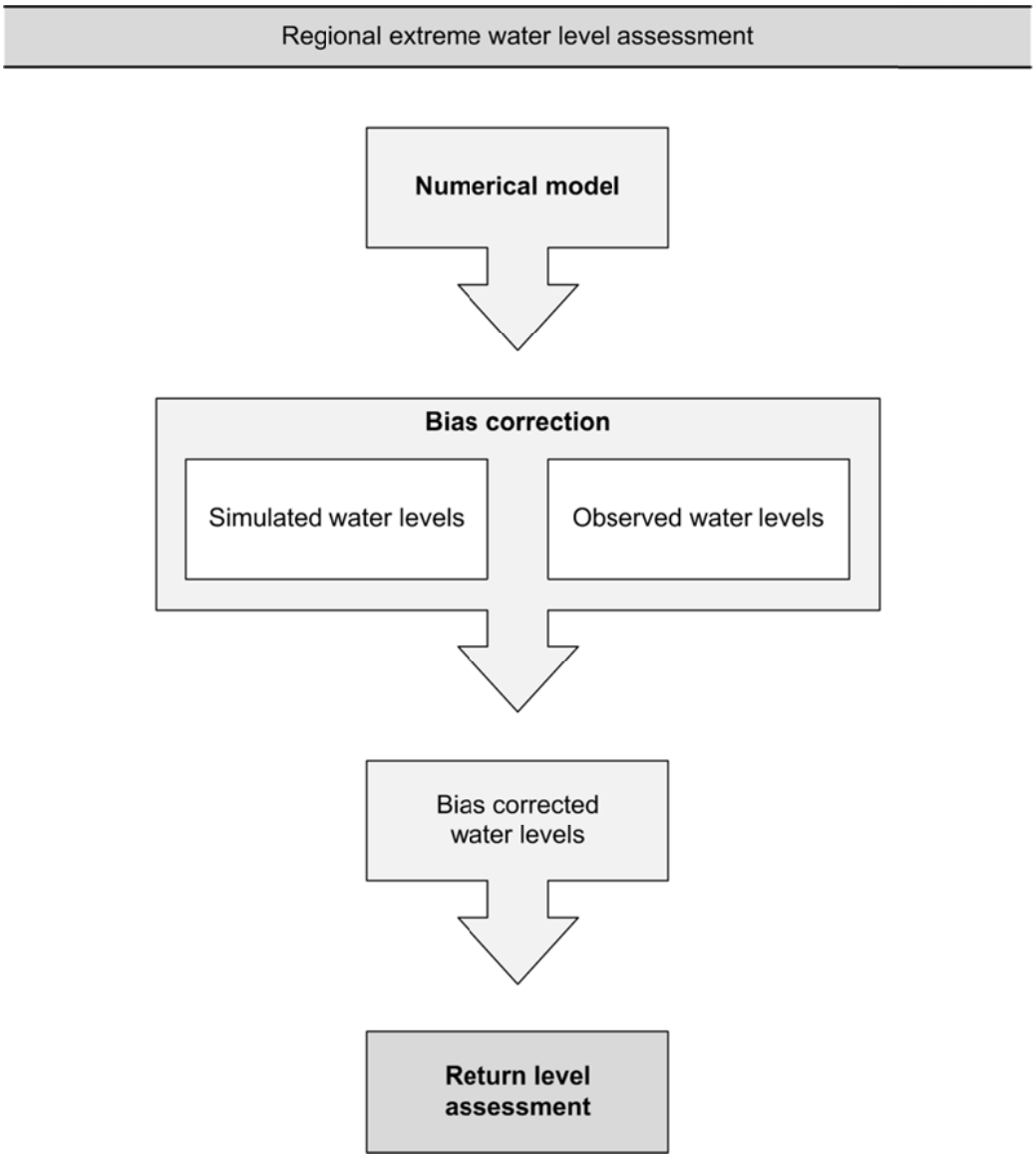


Fig. 51 Flow chart highlighting how regional extreme water levels were assessed

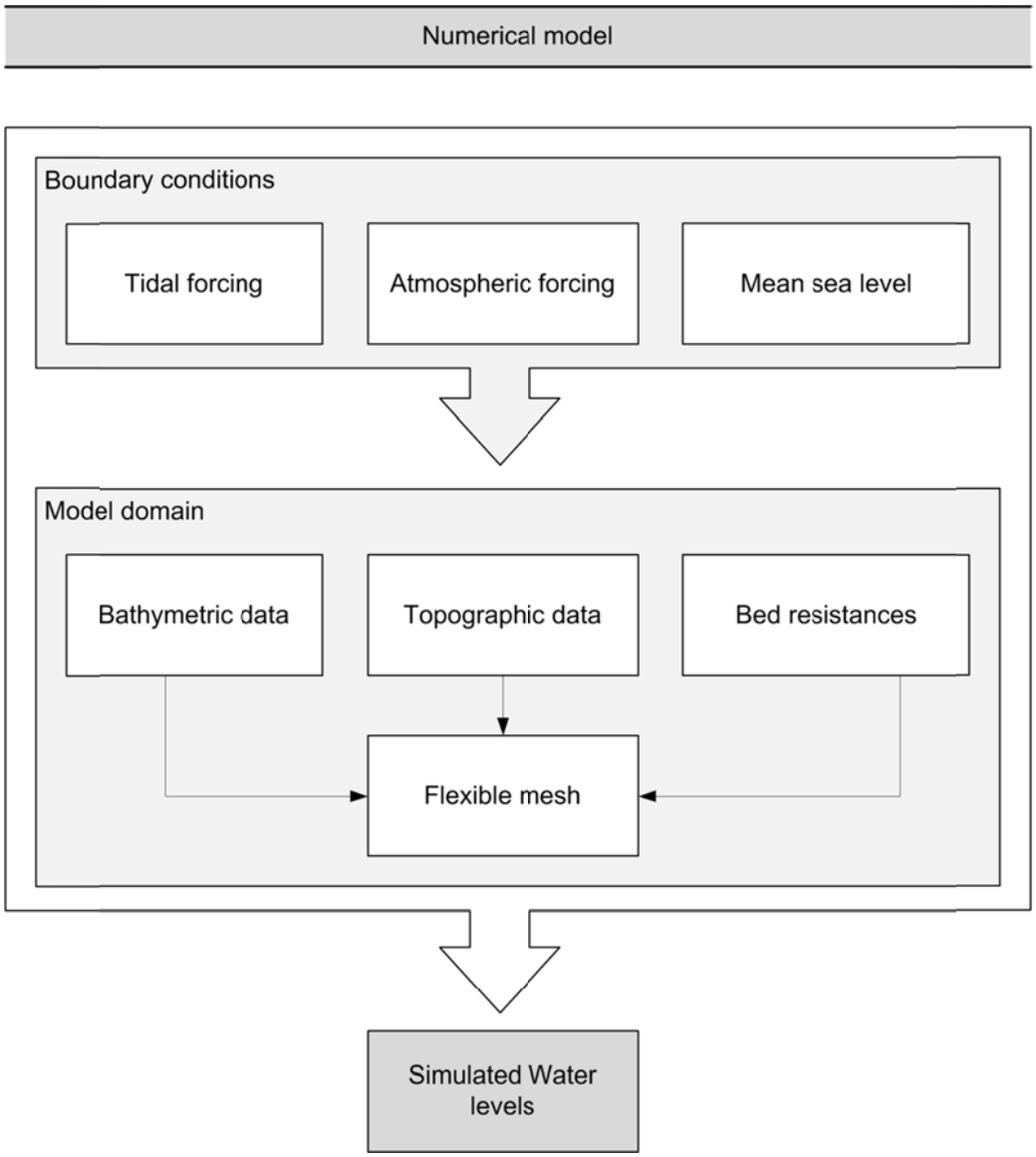


Fig. 52 Flow chart describing the setup of the numerical model that was used in Part II and Part III of this thesis

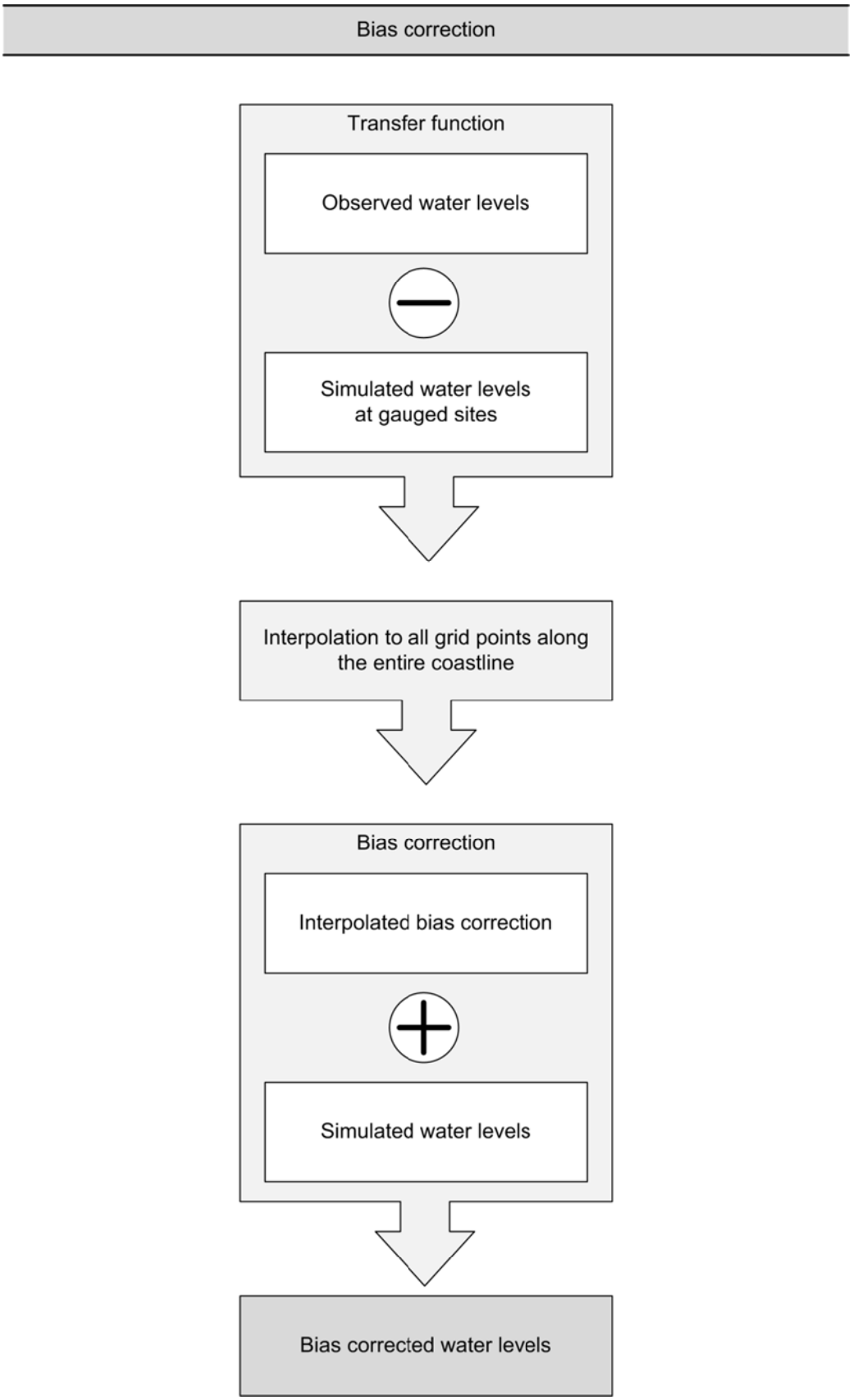


Fig. 53 Flow chart describing the bias-correction that was developed in Part II of this thesis

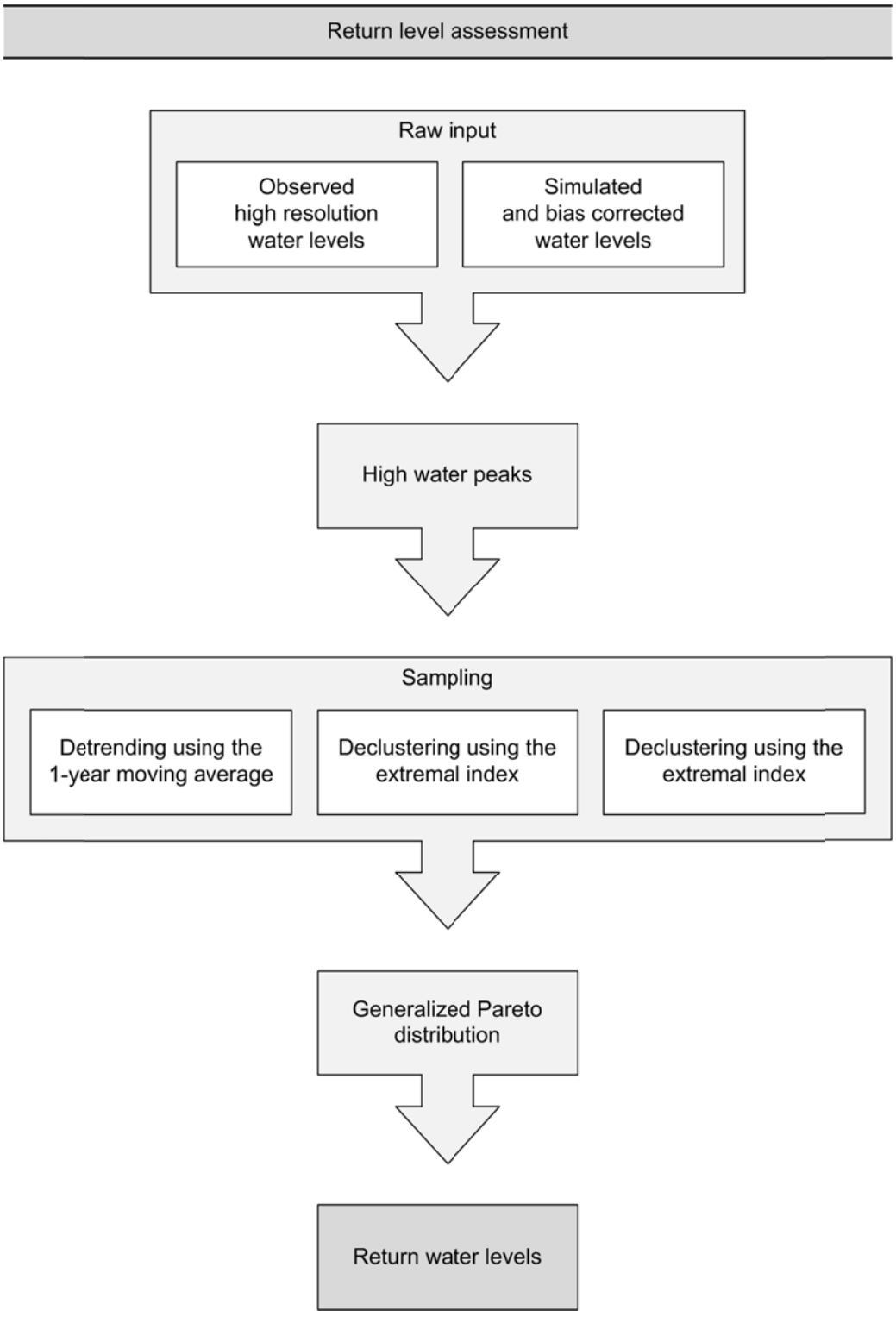


Fig. 54 Flow chart showing the return level assessment (see also Part I of this thesis)

B Appendix

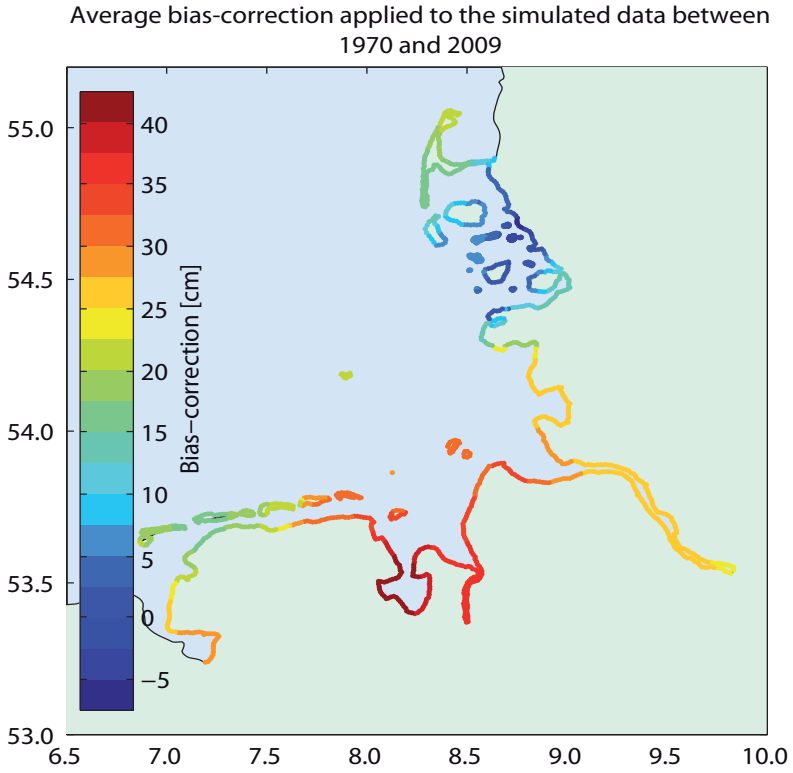


Fig. 55 Average bias-correction applied to the simulated data between 1970 and 2009

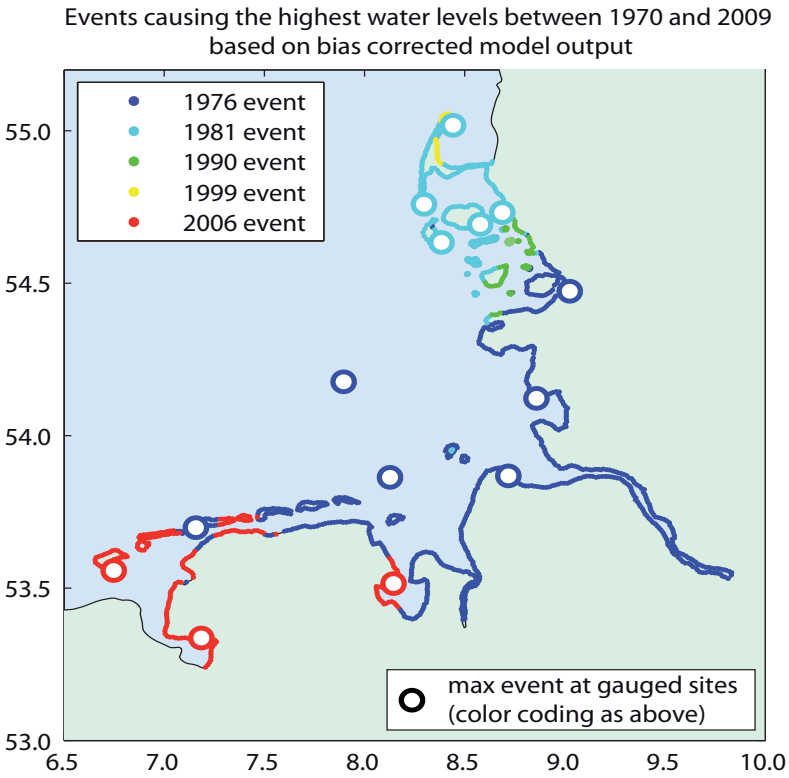


Fig. 56 Events causing the largest water levels between 1970 and 2009 based on the bias-corrected model output

C Appendix

The following pages show the results of the calibration exercise that was used to adjust the numerical model of Part II and Part III to natural consitions. The figures show differences between observed and simulated water levels and peaks at the stations Hoernum, Cuxhaven, Norderney, Aberdeen, Lowestoft, Whitby, Texel and, Calais.

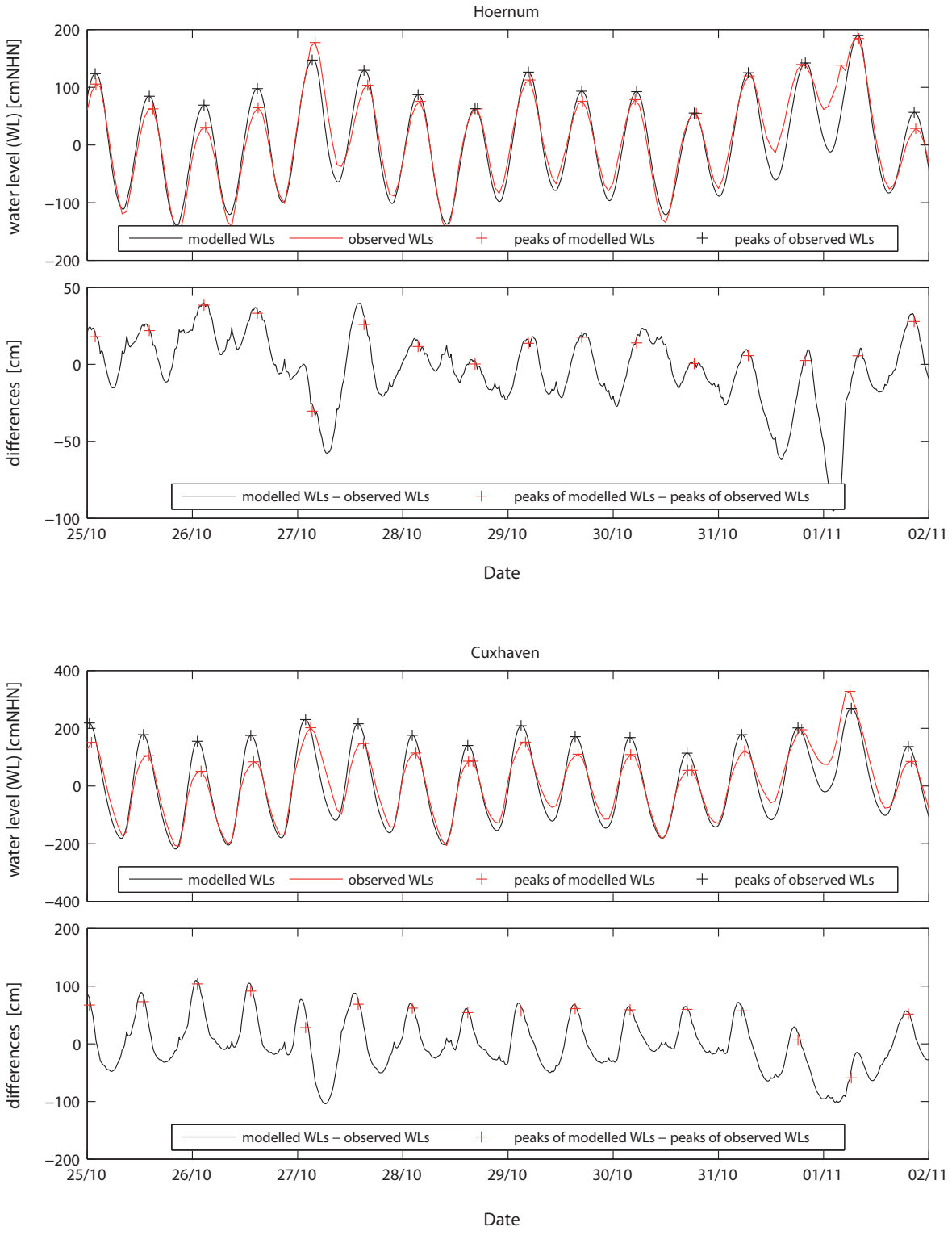


Fig. 57 Calibration at Hoernum (upper figure) and Cuxhaven (lower figure)

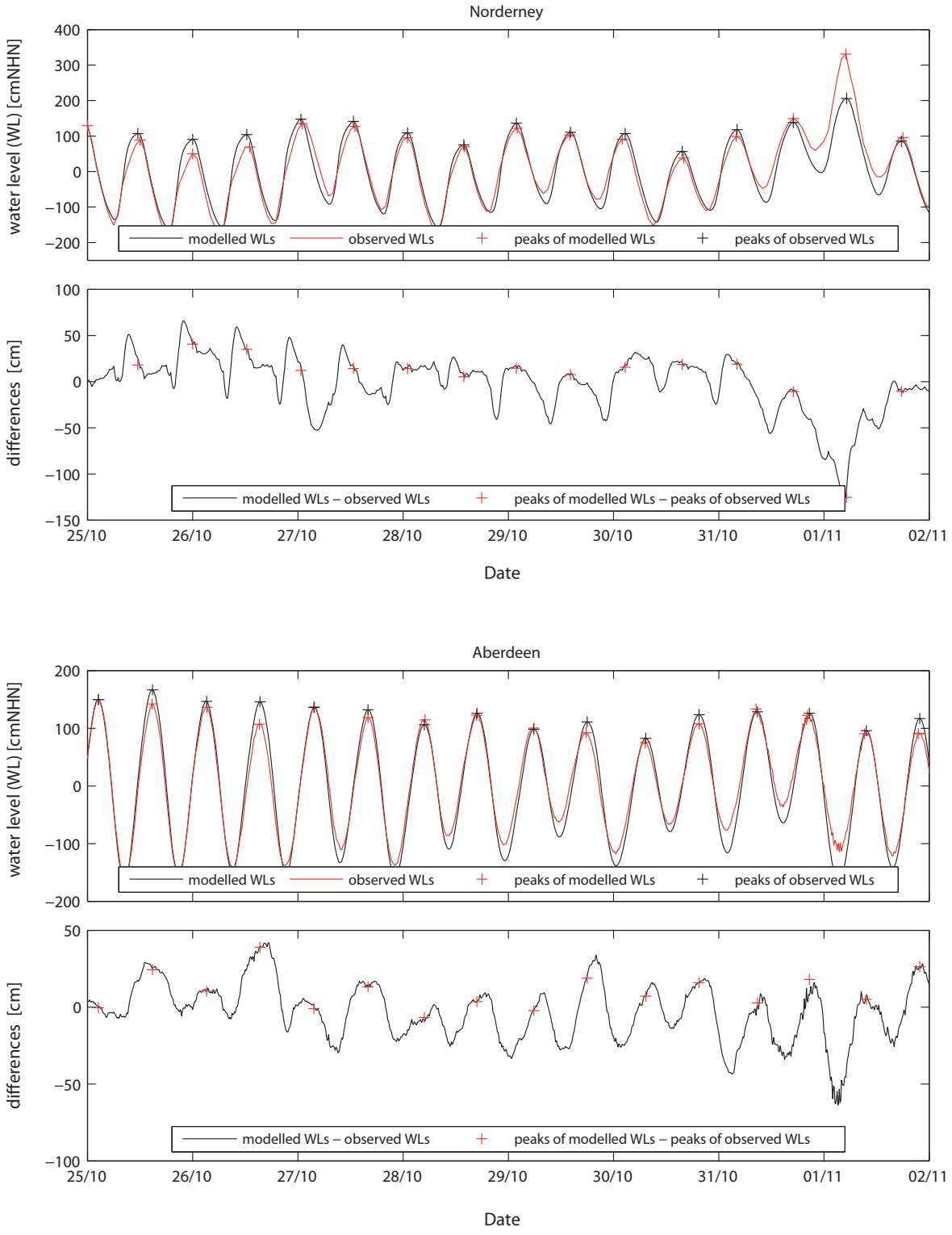


Fig. 58 Calibration at Norderney (upper figure) and Aberdeen (lower figure)

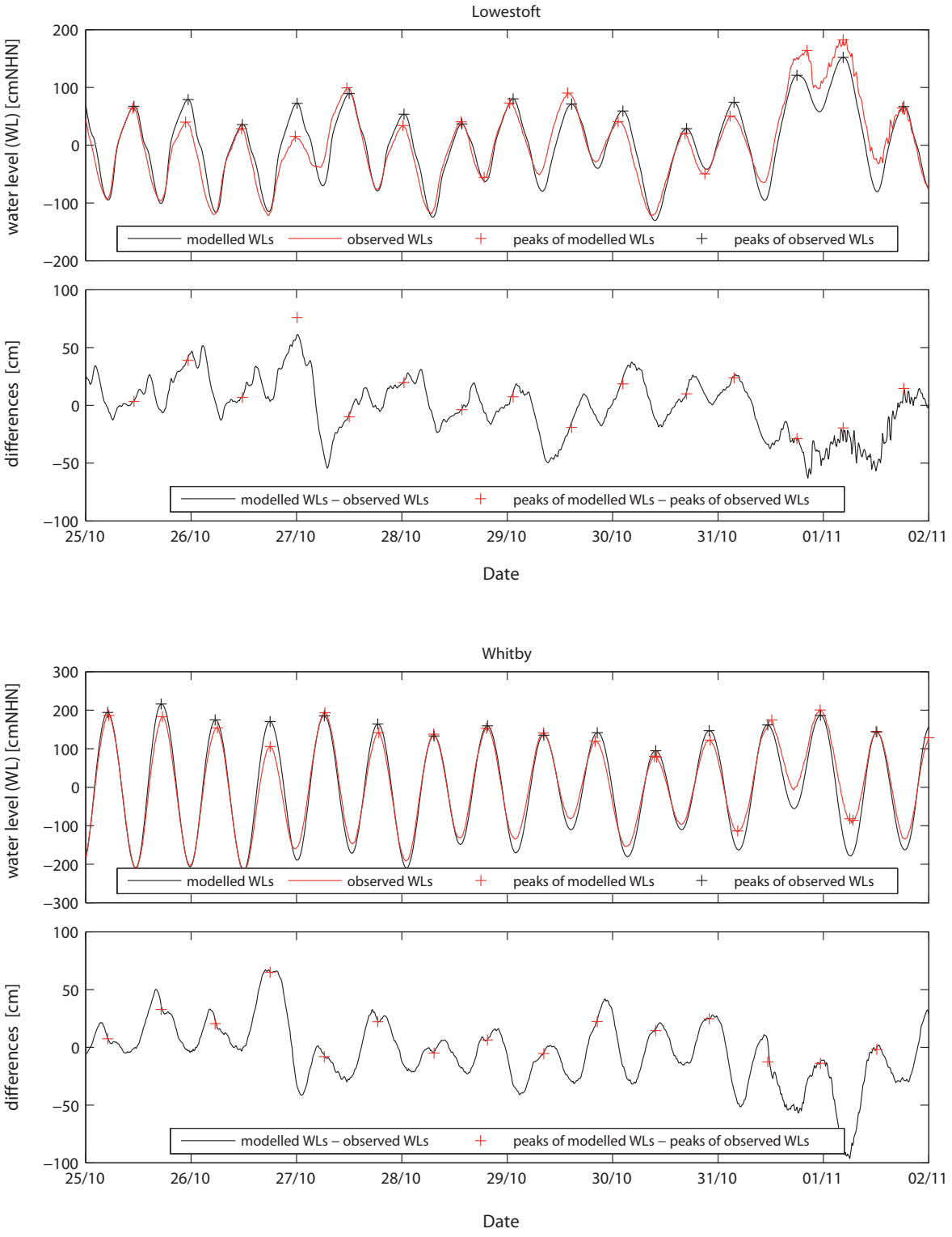


Fig. 59 Calibration at Lowestoft (upper figure) and Whitby (lower figure)

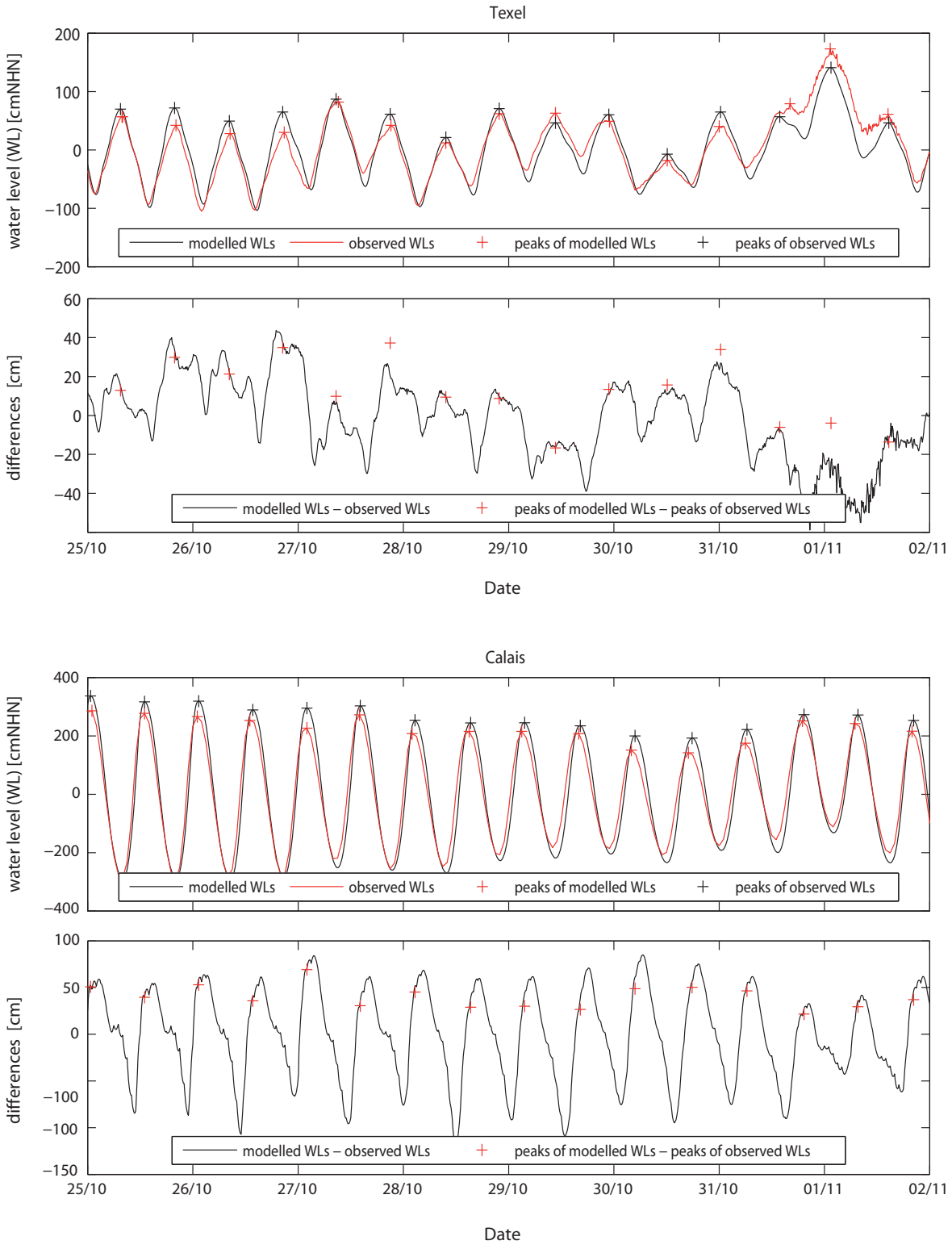
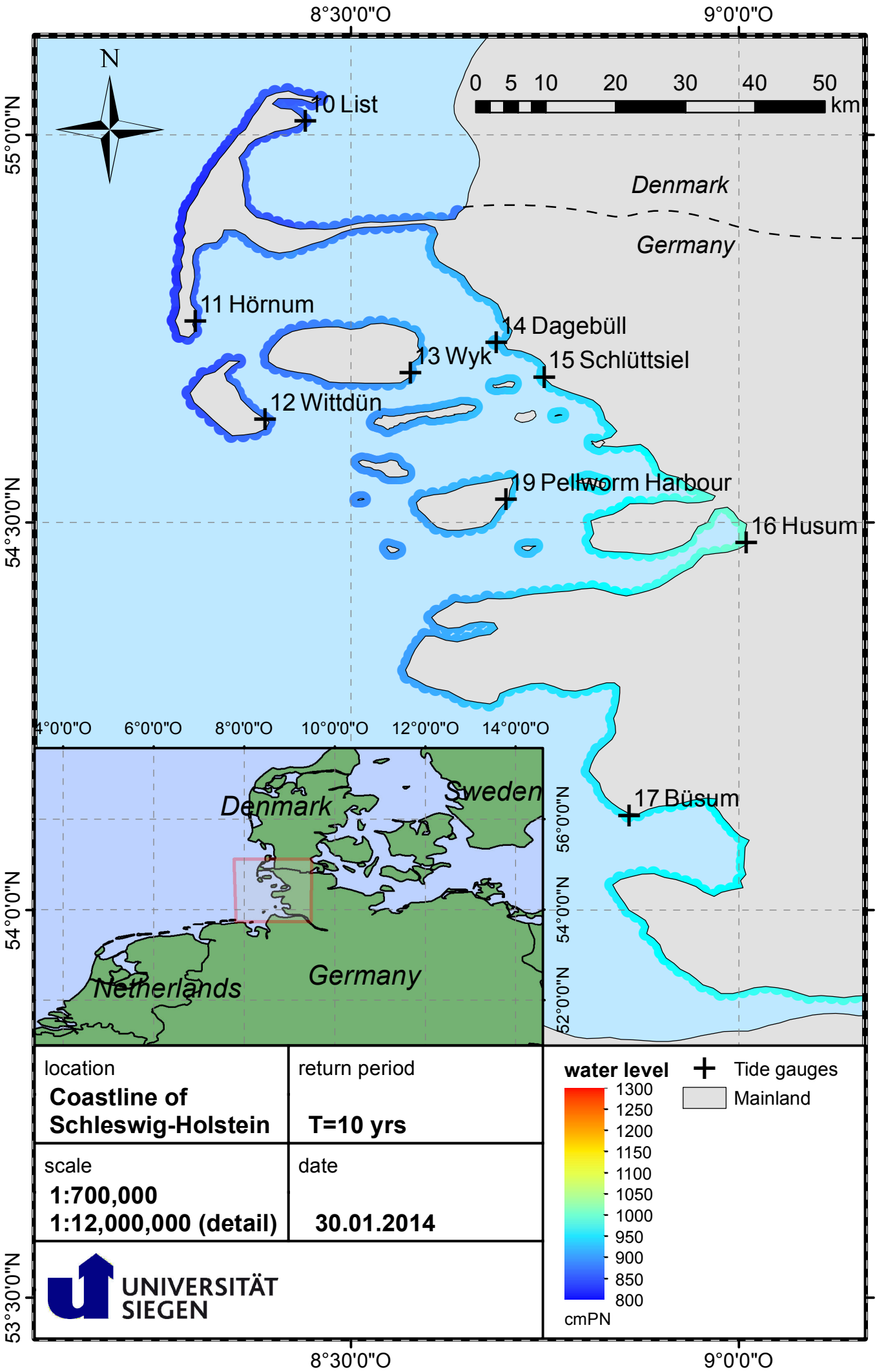
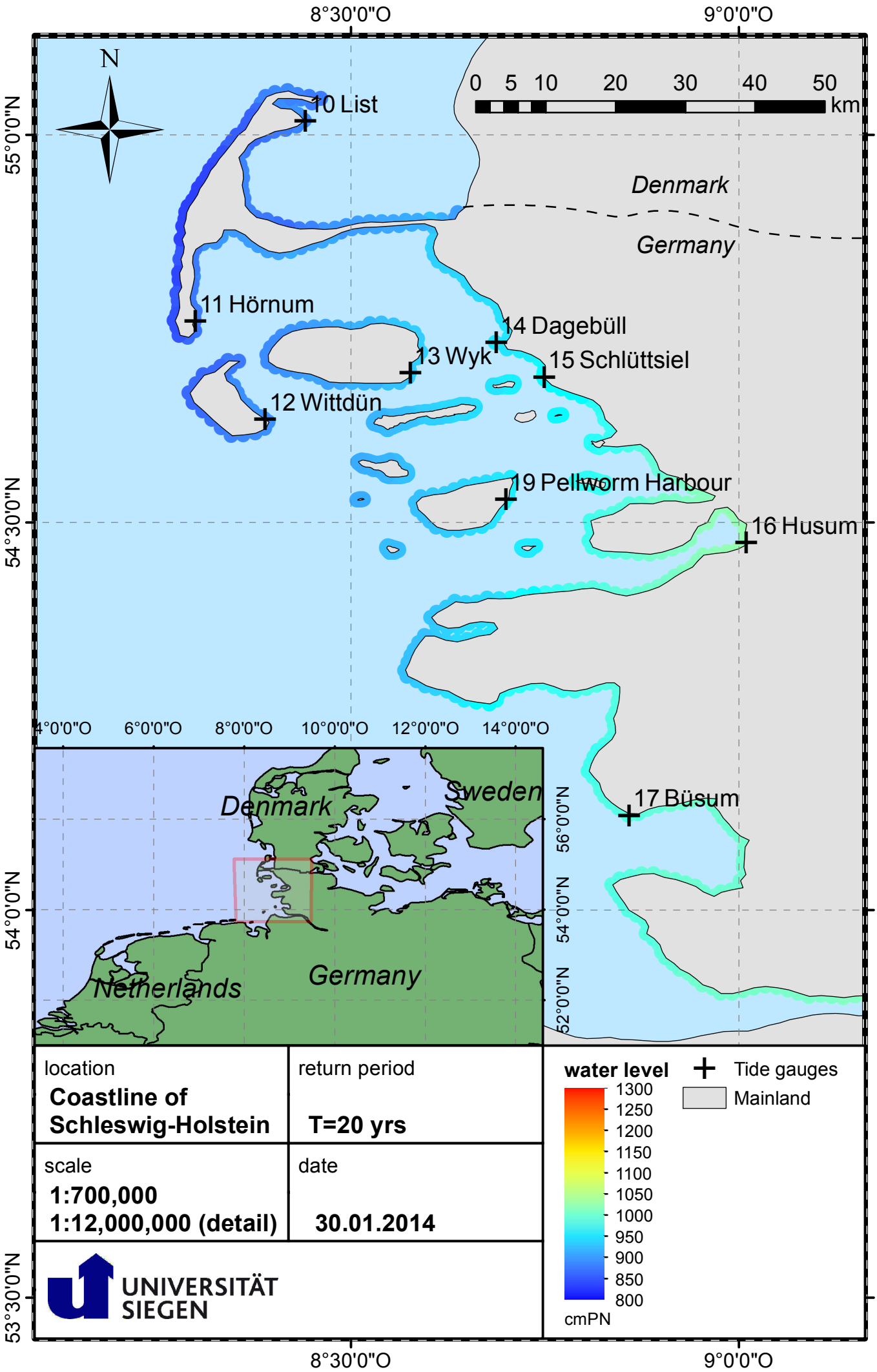


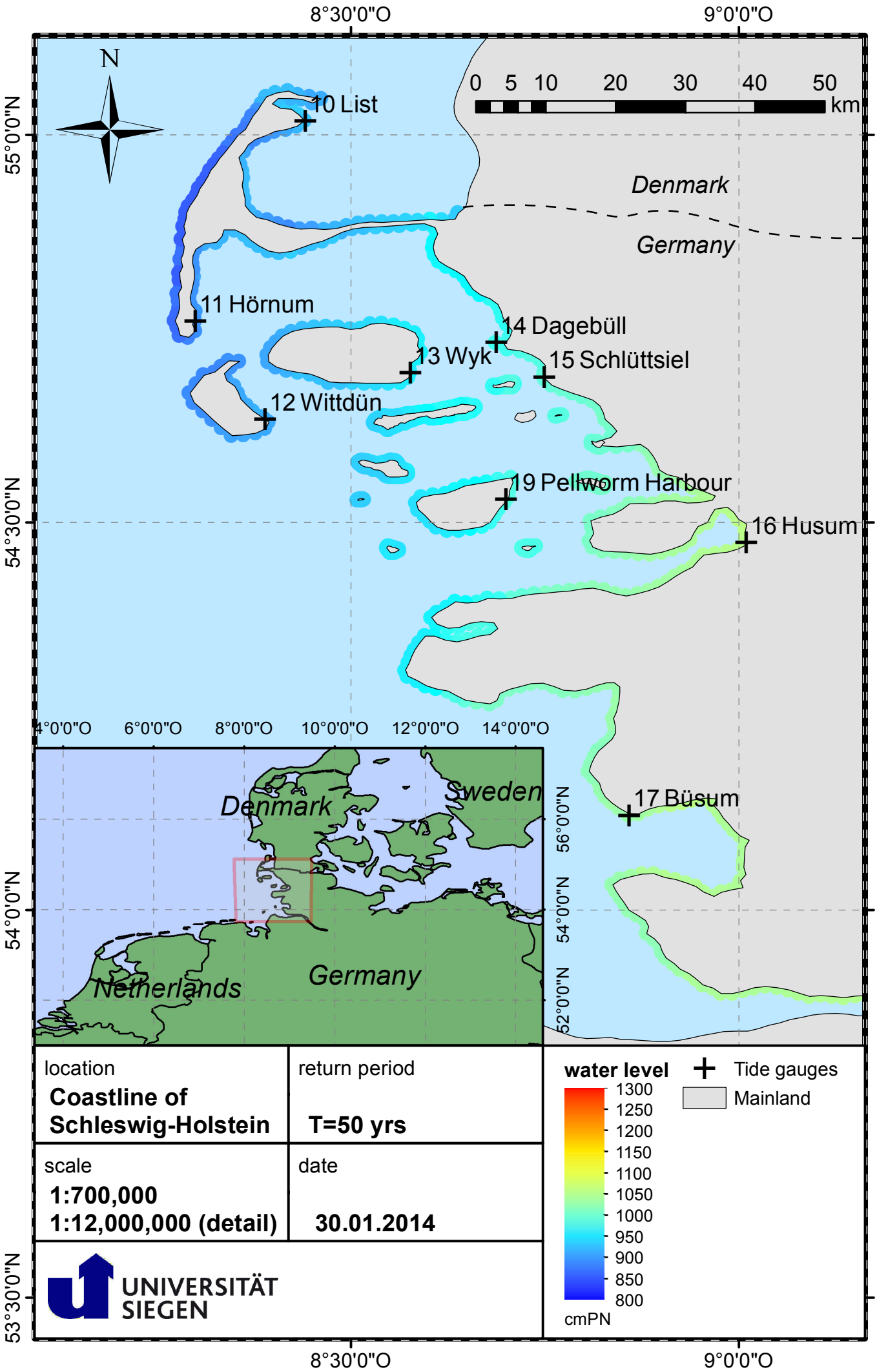
Fig. 60 Calibration at Texel (upper figure) and Calais (lower figure)

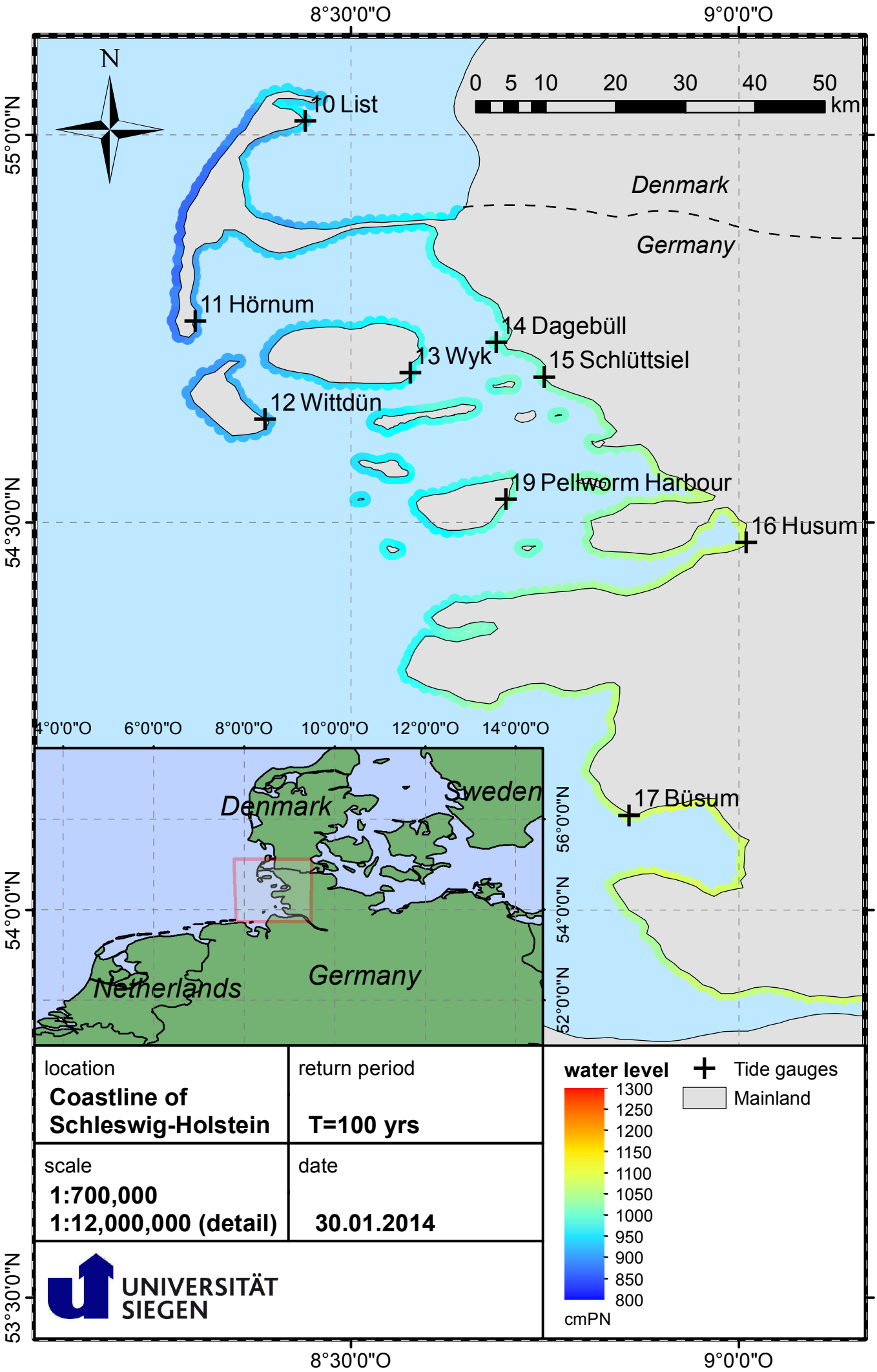
D Appendix

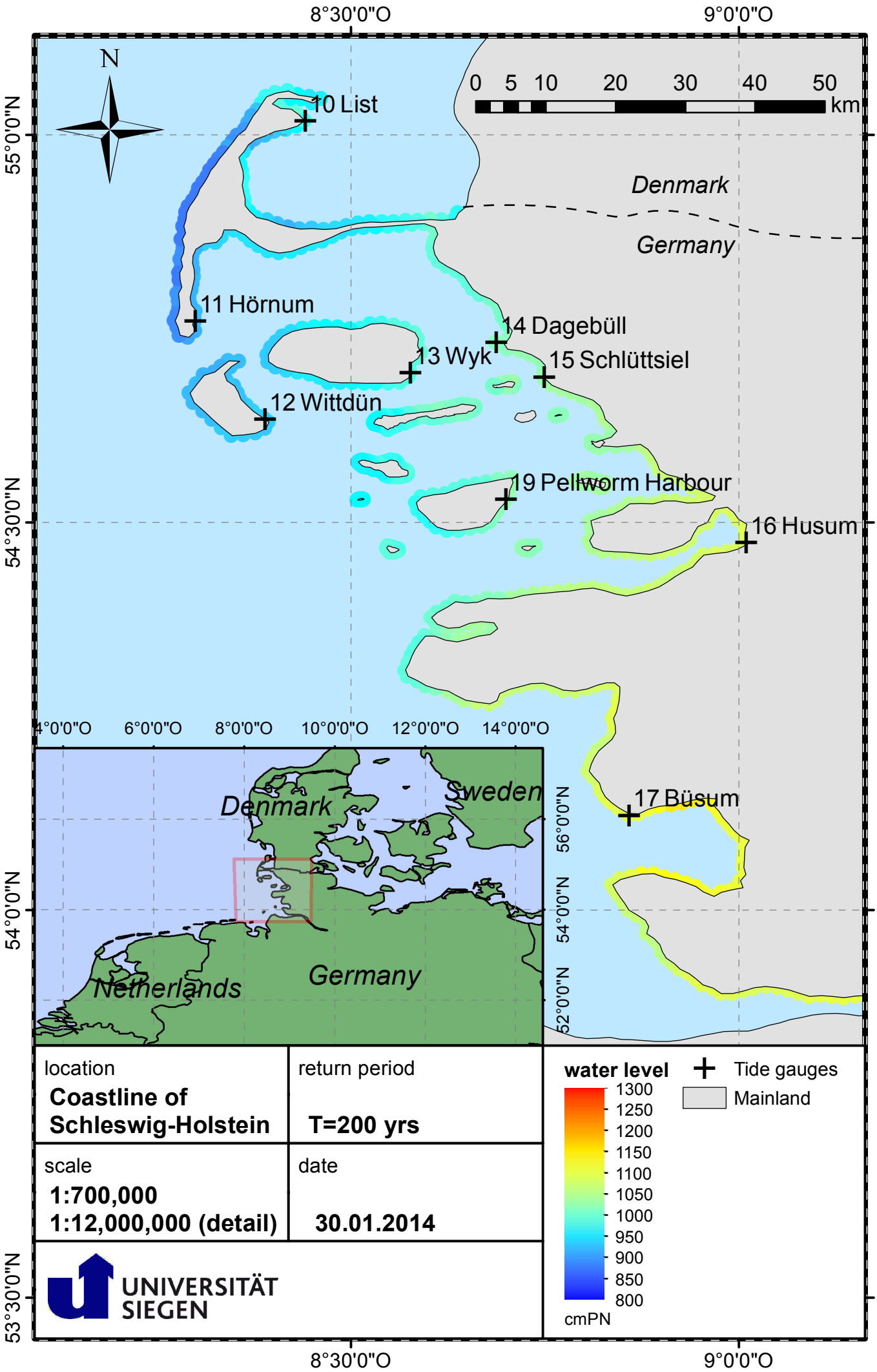
The following pages show a range of maps displaying the water levels of the return levels T={10, 20, 50, 100, 200, 500, 1000} years for the entire coastline of Schleswig-Holstein. The maps are drawn to scale.

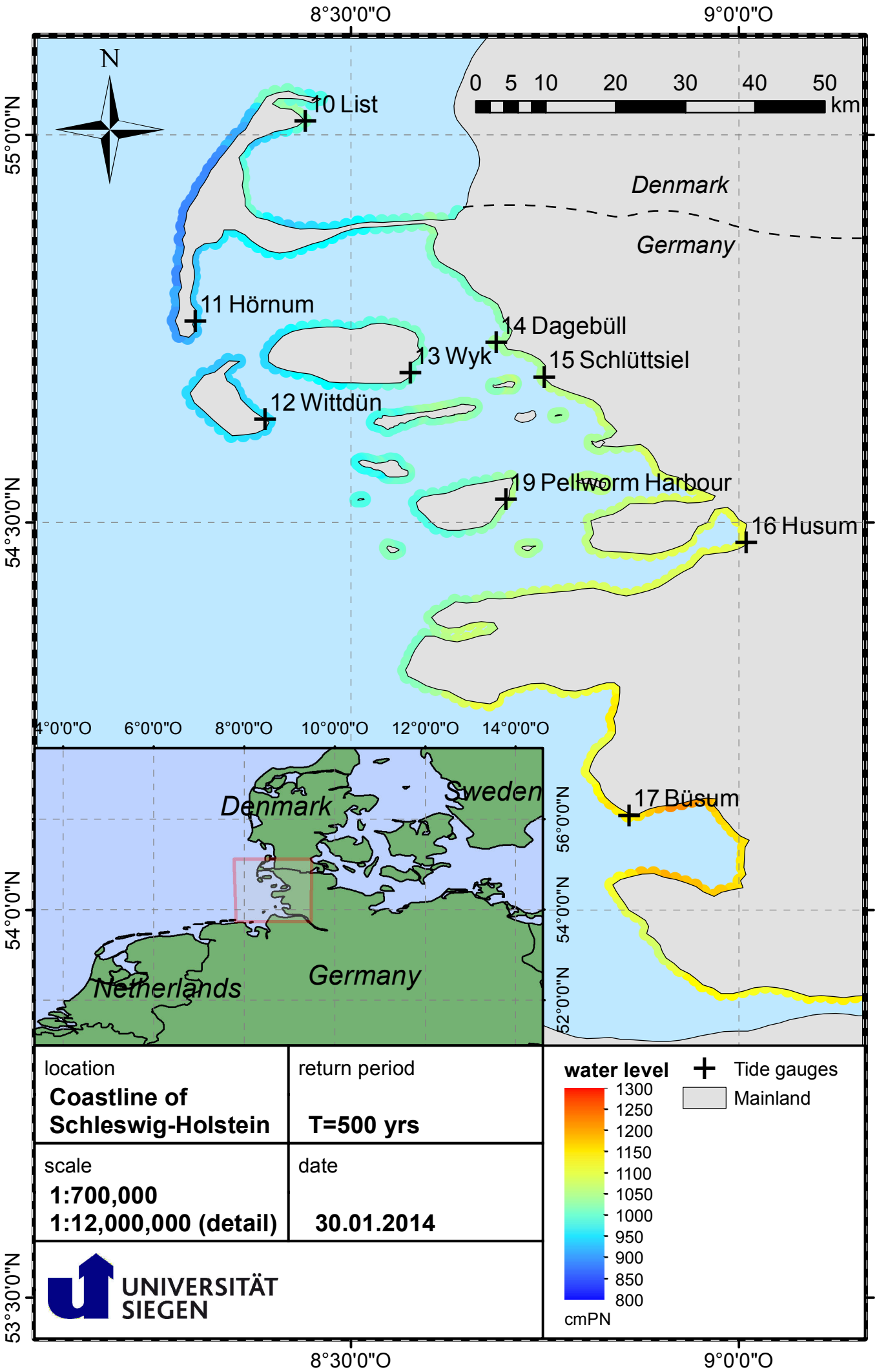

8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=10 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland

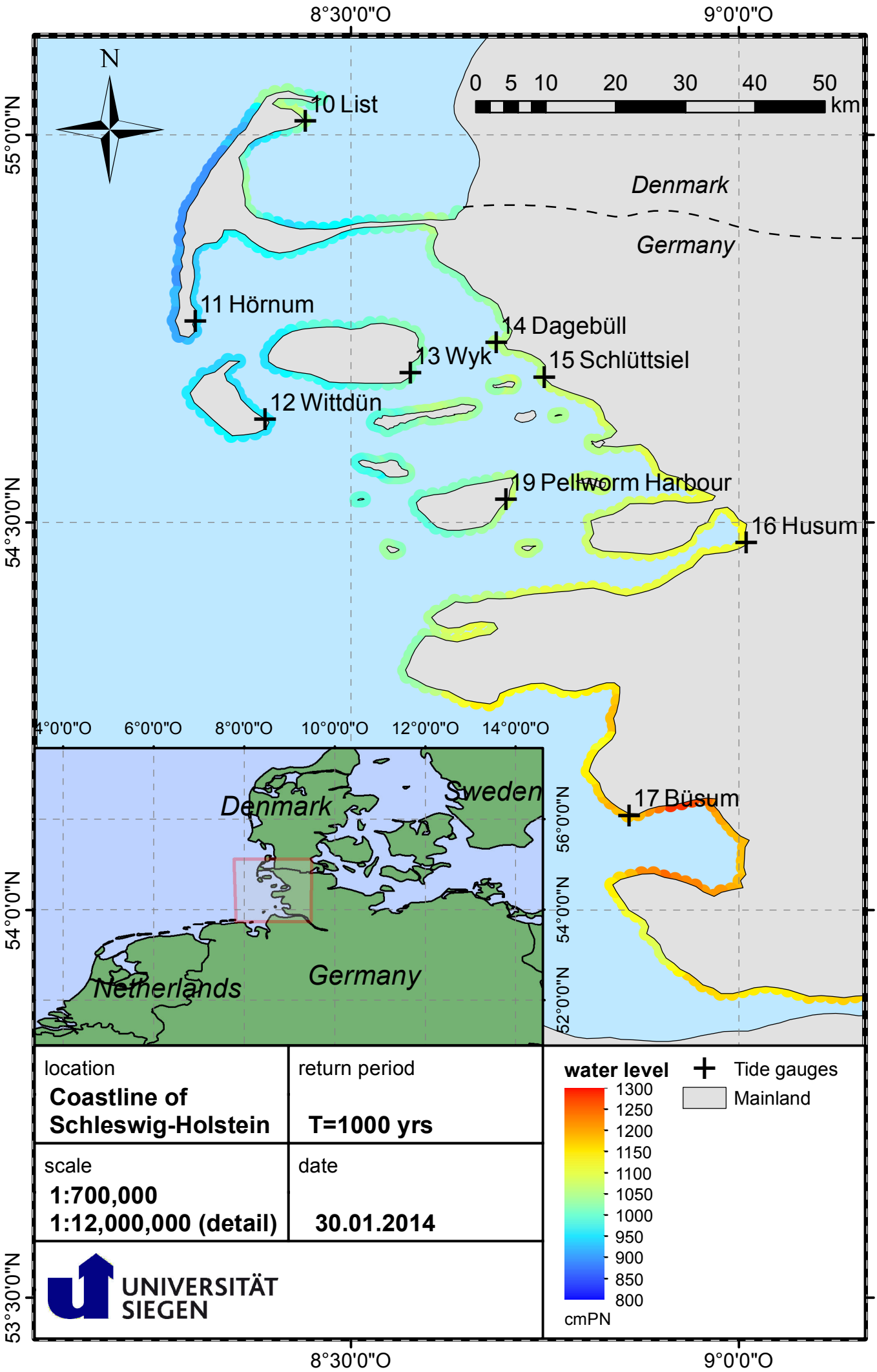

8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Denmark
Sweden
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=20 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland


8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=50 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland


8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=100 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland


8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=200 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland


8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=500 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland


8°30'0"O
9°0'0"O
55°0'0"N
54°30'0"N
54°0'0"N
53°30'0"N
N
0 5 10 20 30 40 50
km
10 List
11 Hörnum
12 Wittdün
13 Wyk
14 Dagebüll
15 Schlüttsiel
19 Pellworm Harbour
16 Husum
17 Büsum
Denmark
Germany
4°0'0"O
6°0'0"O
8°0'0"O
10°0'0"O
12°0'0"O
14°0'0"O
56°0'0"N
54°0'0"N
52°0'0"N
Sweden
Denmark
Germany
Netherlands
location
Coastline of Schleswig-Holstein
return period
T=1000 yrs
scale
1:700,000
1:12,000,000 (detail)
date
30.01.2014
UNIVERSITÄT SIEGEN
water level
1300
1250
1200
1150
1100
1050
1000
950
900
850
800
cmPN
Tide gauges
Mainland

Previously published in the fwu-series (Mitteilungen des Forschungsinstitus Wasser und Umwelt der Universität Siegen)

Issue No.	Title	Author(s)	Year
1	Untersuchungen zur Ermittlung von hydrologischen Bemessungsgrößen mit Verfahren der instationären Extremwertstatiskik - Methoden und Anwendungen auf Pegelwasserstände an der Deutschen Nord- und Ostseeküste	Mudersbach, Christoph	2010
2	CoastDoc 2010 - Beiträge zum 1. Doktorandenseminar CoastDoc, Universität Siegen	Jensen, Jürgen (Editor)	2011
3	Expertenseminar Watershed Management and Rural Sanitation	Bormann, Helge (Editor)	2012
4	Statistical methods to assess the hydrodynamic boundary conditions for risk based design approaches in coastal engineering - Methods and application to the German North Sea coastline	Wahl, Thomas	2012
5	Towards sustainable water quality Management	Bormann, Helge und Althoff, Ingrid	2013
6	Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards (EVAN 2013)	Jensen, Jürgen (Editor)	2014
7	Regional to local assessment of extreme water levels Methods and application to the northern part of the German North Sea coastline	Arne Arns	2014

Herausgeber:
Forschungsinstitut Wasser und
Umwelt (fwu) der Universität Siegen
Paul-Bonatz-Straße 9-11
57076 Siegen