Determining extreme water return levels at un-gauged sites: a case study of the Schleswig-Holsteins coastline and Islands in north-west Germany

Arne Arns[1], Thomas Wahl[2], Ivan D. Haigh[3] and Jürgen Jensen[4]

Abstract

This paper describes a study aimed at estimating extreme water return levels for small marsh islands located offshore of the German Schleswig-Holstein coastline in the Wadden Sea. This is a challenging task as no water level records of sufficient length necessarily for undertaking extreme analysis exist for this region. To transfer water level information measured at gauged sites in neighbouring areas, to un-gauged sites in our study region, the concept of regionalization is adopted and adjusted from a riverine (where it has been previously applied) to a coastal setting. The regionalization 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. The bias-correction is transferred to the water levels predicted at every coastal and island model grid point in our study area. Hence, the simulated water levels are highly accordant with the observed water levels. Using a recommended procedure to conduct extreme value analyses from a companion study, water return levels suitable for coastal defence design conditions are estimated along the entire coastline of Schleswig-Holstein, including the offshore islands.

Key words: Extreme value statistics, storm surges, return periods, hydrodynamic modelling, North Sea, Germany

1. Introduction

Rising sea levels along with changes in storminess and their impact on the likelihood of coastal flooding around the world are among the most debated scientific issues of our time. This is mainly due to the fact, that the occurrence of these weather related events is essentially stochastic (Schneider et al., 2007). Nevertheless, the use of observational data has become a very important feature in decision-making and future planning. For instance, large parts of the German coastline are protected by coastal defence structures which have been built according to design levels determined from water level measurements.

In the German Bight, multi-decadal records of high and low waters exist at several sites, but high frequency datasets, measuring the whole water level curve, are only available since the late 1990s. However, for some regions (e.g. some islands in the German Wadden Sea) no or just a few water level measurements of any sort exist. 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 convey information about the likelihood of extreme water level events from gauged to surrounding un-gauged sites.

A workaround method is to use the concept of regionalization. This concept has been applied widely in hydrology, where catchment attributes and spatial proximity are used as a similarity measure. Based on this measure one can decide which information is to be transferred from the catchment to the site of interest. It is assumed that catchments with similar attributes behave in a similar manner in flood frequency response (Merz and Blöschl, 2005). However, this classical approach of regionalization is not applicable in coastal areas (at least not in the German Bight) as water level records show unique characteristics as they are strongly affected by local influences. A regionalization approach therefore needs to account for coastal

[1] Research Institute for Water and Environment, University of Siegen, Siegen, Germany, arne.arns@uni-siegen.de
[2] Institute of Advanced Studies – FoKoS (Research Group Civil Security), University of Siegen, Siegen, Germany., thomas.wahl@uni-siegen.de
[3] National Oceanography Centre, University of Southampton, Southampton, UK, I.D.Haigh@soton.ac.uk
[4] Research Institute for Water and Environment, University of Siegen, Siegen, Germany, juergen.jensen@uni-siegen.de

attributes. The aim of this paper is to develop a regionalization approach to determine extreme water level probabilities, especially for areas where few or no water level measurements exist. As a case study we do this for the coastline of Schleswig-Holstein in the German Wadden Sea. In this area, small islands with historical importance are located that could be threatened in the future with sea level rise, thus requiring accurate assessment of flood risk. The study area and the considered tide gauge data sets are described in Sects. 2 and 3, respectively. In Sect. 4, the regionalization method is introduced. Sect. 5 (and the related sub-sections) describes the hydrodynamic numerical model, its configuration and calibration, as well as the bias correction method and the validation. The results from the extreme value analysis are presented in Sect. 6 and paper closes with short summary provided in Sect. 7.

2. Study area

The Halligen are small low lying islands located off the coastline of Schleswig-Holstein in northern Germany (Figure 1). They are frequently exposed to storm surges and are inundated up to 50 times a year. Only 10 Halligen have survived to present day. It has been estimated that through the last centuries, around 100 Halligen have been destroyed by large storm surges (Quedens, 1992). Besides having a historic-cultural importance, the Halligen help to significantly reduce the storm surge impacts on the mainland of Schleswig-Holstein by providing a natural barrier of protection. It is thus of great interest, to preserve these small Islands.

The Halligen are inhabited by around 350 residents, but they have no dikes. In order to protect the inhabitants from regular inundation, houses have been built on dwelling mounds. Residents have learned to cope with extreme conditions, but it is expected that the Halligens' shapes and conditions will be gradually negatively affected, especially as a consequence of rising sea levels. Between the 13th and 20th century, approximately 50% of the surface area of the Halligen irretrievably vanished as a result of large storm surge (Quedens, 1992).

The Halligen are surrounded by the North Frisian Wadden Sea. With an area of approximately 9.000 km², this depositional coastline is one of the world’s largest intertidal wetlands. In 2009, the North Frisian Wadden Sea was added to UNESCO´s World Heritage List.

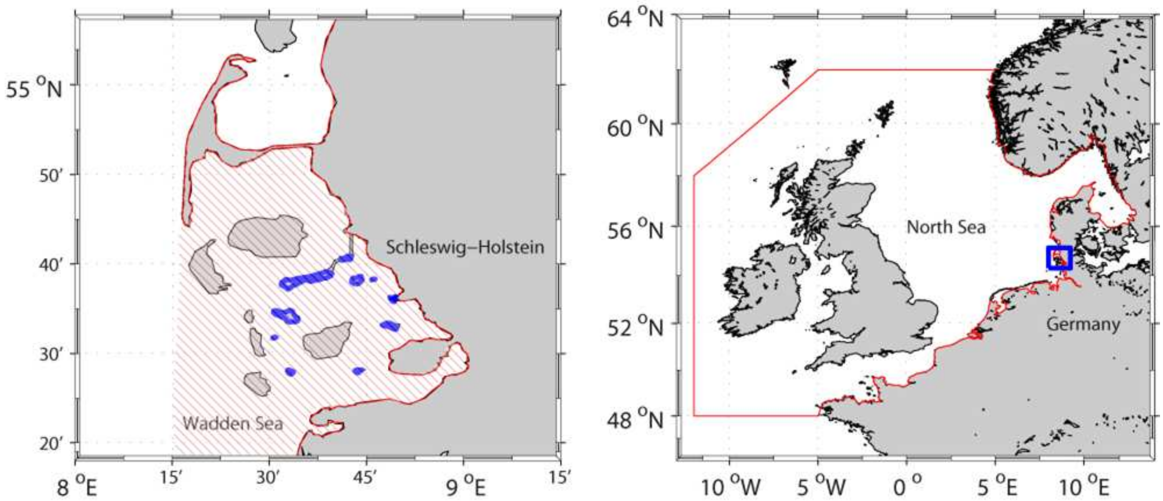

Figure 1. Study area with a) the main area of investigation, where the blue areas show the Halligen and the red shaded area the North Frisian part of the Wadden Sea, and b) the model boundaries used for the hydrodynamic numerical model.

3. Tide gauge data

Water level records from a number of tide gauges along the UK, Dutch, French and German coastlines were considered for the present study. All water level records were referenced relative to the German ordnance datum "Normalhöhennull" (NHN). The individual records were checked for common errors and suspicious data were deleted. To calibrate the 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 the storm surge event of November 1st, 2006. For the bias-correction of the model output, described below, tidal high water levels for the period from 1970 to 2009 from all German Bight tide gauges (except Pellworm Harbour) were used. Water level records of Pellworm Harbour were used for validation purposes (see Sect. 5.4).

Table 1. Tide gauges used to calibrate, correct and validate the model output. The (*) indicates, that tidal high and low waters are available; at all other stations, high resolution values (1-minute) were used. The marker (✓) indicates in which computational step the data was used; stations highlighted with (✓✓) are graphically shown in the paper.

Nr.	Tide gauge location	Country	Years	Availability [%]	Calibration	Correction	Validation
1	Aberdeen	GB	2006	100	✓✓	-	-
2	Lowestoft	GB	2006	100	✓✓	-	-
3	Whitby	GB	2006	100	✓✓	-	-
4	Texel	NL	2006	100	✓✓	-	-
5	Calais	FRA	2006	89.6	✓✓	-	-
6	Hörnum	GER	2006 1970 – 2009	98.9 100*	✓✓ -	- ✓✓	 ✓
7	Cuxhaven	GER	1970 – 2009	100*	✓✓	✓	✓
8	Norderney	GER	1970 – 2009	100*	✓✓	✓	✓
9	Helgoland	GER	1970 – 2009	100*	-	✓	✓
10	Wittdün	GER	1970 – 2009	100*	-	✓	✓
11	Wyk	GER	1970 – 2009	100*	-	✓	✓
12	Husum	GER	1970 – 2009	100*	-	✓	✓
13	Dagebüll	GER	1970 – 2009	100*	-	✓	✓✓
14	List	GER	1970 – 2009	100*	-	✓	✓
15	Büsum	GER	1970 – 2009	100*	-	✓	✓
16	Schlüttsiel	GER	1970 – 2009	100*	-	✓	✓
17	LT Alte Weser	GER	1970 – 2009	100*	-	✓	✓
18	Wilhelmshaven	GER	1970 – 2009	100*	-	✓	✓
19	Borkum FB	GER	1970 – 2009	100*	-	✓	✓
20	Emden	GER	1970 – 2009	100*	-	✓	✓
21	Pellworm Hafen	GER	1970 – 2009	100*	-	-	✓✓

4. Regionalization

The occurrence probability of a certain storm surge event is typically calculated using a limited sample of the entire population. An objective in statistical modelling 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, it can be used to estimate the unknown statistical parameters of the population (Arns et al., under review). The availability of sufficient data is thus one of the crucial aspects when performing extreme value analyses. Using small samples yields imperfect estimates and consequently unrealistic occurrence probabilities (Rao and Hamed, 2000).

As hydrological data is often limited in space and time, methods to indirectly derive occurrence probabilities have been developed; these are called regionalization methods. Such methods compensate for the lack of data at individual stations (Stedinger et al., 1993). The application essentially pursues two objectives. The first one is to enlarge the data basis in gauged areas. The causal extension of hydrological information can be achieved by merging individual samples into one single homogeneous group. When using n samples at the location x={1,2,3,…n}, having sizes Sx, yields a new sample consisting of $n_R=\sum Sx$ values; the resulting sample n_R is considered to be representative for a homogeneous region, which was identified by using homogeneity tests (see e.g. Dalrymple, 1960; Wiltshire, 1986; Hosking and Wallis, 1997). Using this kind of regionalization represents a substitution between space and time as different long records within an area are used to compensate shorter records (Rao and Hamed, 2000). Provided that the considered records are from the same distribution, the parameters can be estimated more robust. One of the most common regionalization methods is the so-called Flood-Index-Method proposed by Dalrymple (1960). Assuming regional homogeneity, the individual samples of different locations have a common distribution (Rao and Hamed, 2000); essential differences are only found in a scaling factor, called the Index-Flood (e.g. mean high water). Occurrence probabilities can be derived by multiplying the regional distribution ("growth curve") and the Index-Flood.

Where information is spatially limited, regionalization methods are used to generate information for un-gauged areas. The information of interest, e.g. the Index-Flood, can be obtained from catchment characteristics. Based on a similarity measure it is decided, which information is to be transferred to the site of interest. It is assumed that catchments with similar attributes behave similar in flood frequency response (Merz and Blöschl, 2005). Many approaches have been developed in the recent past but these were mainly designed for use in inland hydrology. As tidal characteristics in the German Bight are highly influenced by shallow water effects, they can differ significantly between stations. It is thus difficult to convey information about the likelihood of extreme hydrologic events from gauged to un-gauged sites. A regionalization approach therefore needs to account for coastal attributes, which are often locally confined and can be captured via numerical modelling.

5. Hydrodynamic numerical model

5.1 Model configuration

To generate continuous water levels for the entire German Bight, a hindcast for the 40-year period from 1970 to 2009 was performed with a 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 (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 southernmost 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, 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's Government-Owned Company for Coastal Protection, National Parks and Ocean Protection (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 chart 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. These were derived from a global tide model provided by MIKE21 (DHI), including the eight primary harmonic constituents (K1, O1, P1, Q1, M2, S2, N2 und K2, see e.g. Andersen, 1995). Additionally, the Mean Sea Level (MSL) was considered using an index-time series for the entire North Sea from Wahl et al. (in press); the time series was derived using data from 30 tide gauges located around the North Sea basin. As each year of the considered 40-yr hindcast was run separately, the MSL at the open boundaries was adjusted according to the annual MSL time series.

The surge component of the model 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 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° and a temporal resolution of 3 hours. The bed resistance was set to a constant Manning's number of n=0.022 [-] (corresponds to k_{st}=45 $m^{1/3}/s$) (see also section 5.2). The model was run for a two day warm up period and results were stored with at an interval of 10 minutes.

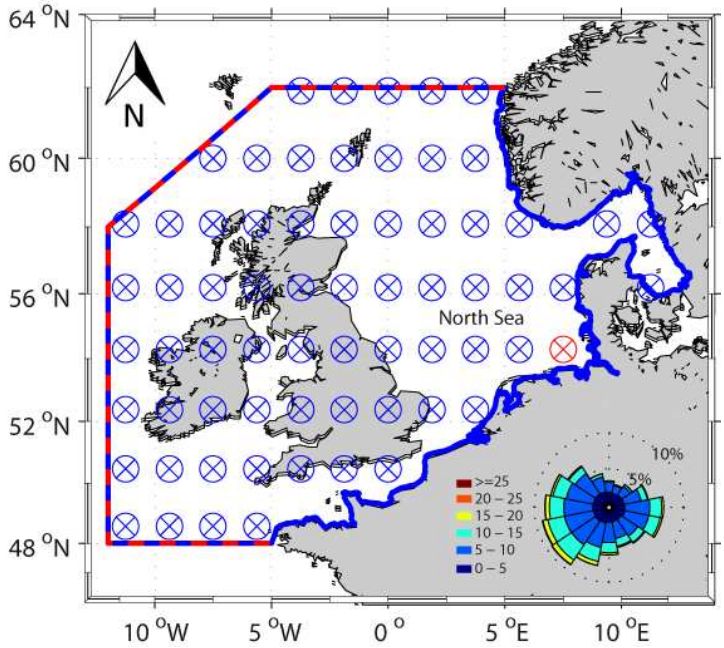

Figure 2. Grid points of wind- and pressure fields (blue circles): the red/blue dotted line represents the open boundary of the model; the wind rose located in the lower right corner shows the mean wind speeds in m/s at the location closest to the German Bight (red circle) between 1970 and 2008.

5.2 Model calibration

The model was calibrated using stepwise variations of the bed resistance, using Manning's n-values between $0.022 \leq n \leq 0.028$ [-]. This was to ensure only physically plausible Manning coefficients were considered (see e.g. Novikov and Bagtzoglou, 2006). The evaluation of the models behaviour 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 modelling to assess the model performance, each having assets and drawbacks. In this work we used the *Coefficient of determination (r^2)* and the *Index of agreement (d)*. The *Coefficient of determination* is defined as the squared value of the coefficient of correlation (Krause et al. 2005) between observed (Wo) and modelled (Wm) water levels and is calculated as follows:

$$r^2 = \left\{ \sum_{i=1}^{n} \left(Wm_i - \overline{Wm} \right) \left(Wo_i - \overline{Wo} \right) \div \left(\sqrt{\sum_{i=1}^{n} \left(Wo_i - \overline{Wo} \right)^2} \sqrt{\sum_{i=1}^{n} \left(Wm_i - \overline{Wm} \right)^2} \right) \right\}^2 \tag{1}$$

with $0 \le r^2 \le 1$. A value of $r^2=0$ [-] denotes that there is no correlation between Wo and Wm, whereas a value of $r^2=1$[-] indicates that Wo and Wm 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

$$d = 1 - \sum_{i=1}^{n} \left| Wo_i - Wm_i \right|^2 \div \sum_{i=1}^{n} \left(\left| Wm_i - \overline{Wo} \right| + \left| Wo_i - \overline{Wo} \right| \right)^2 \; . \tag{2}$$

The efficiency criteria resulting from the comparison between the observed and modelled water levels are summarized in Table 2. For calibration purposes, the storm surge of November 2006 was used. After removing the warm up period, the input time series cover a period of one week. For simplicity, constant Manning’s n-values were used spatially across the model domain. The results highlight, that the overall agreement was highest along the UK stations. Slightly higher differences occurred in the German Bight. These differences are most probably attributed to shallow water effects that occur in this region and which are possibly not captured properly by the model.

Table 2: Efficiency criteria based on the best fit with n=0.033 [-].

Criteria	Hörnum	Cuxhaven	Norderney	Aberdeen	Lowestoft	Whitby	Texel	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

The results from using different n-values are shown in Figure 3. The black curves represent the observed water levels at a specific location; the red curves show the modelled 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.

5.3 Bias-correction

The calibration exercise allowed us to minimise the differences between the observed and the modelled water levels. However, with regards to extreme value analyses, small differences in the input water levels can produce large discrepancies in return water level estimates, particularly at large return periods. All water level observations are prone to natural as well as anthropogenic influences that cannot be captured by a numerical model. This causes a bias between the distributions derived from observed and modelled water levels. Thus the bias is attributed to input deficiencies e.g. resolution or scaling effects. The wind as an example has 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.

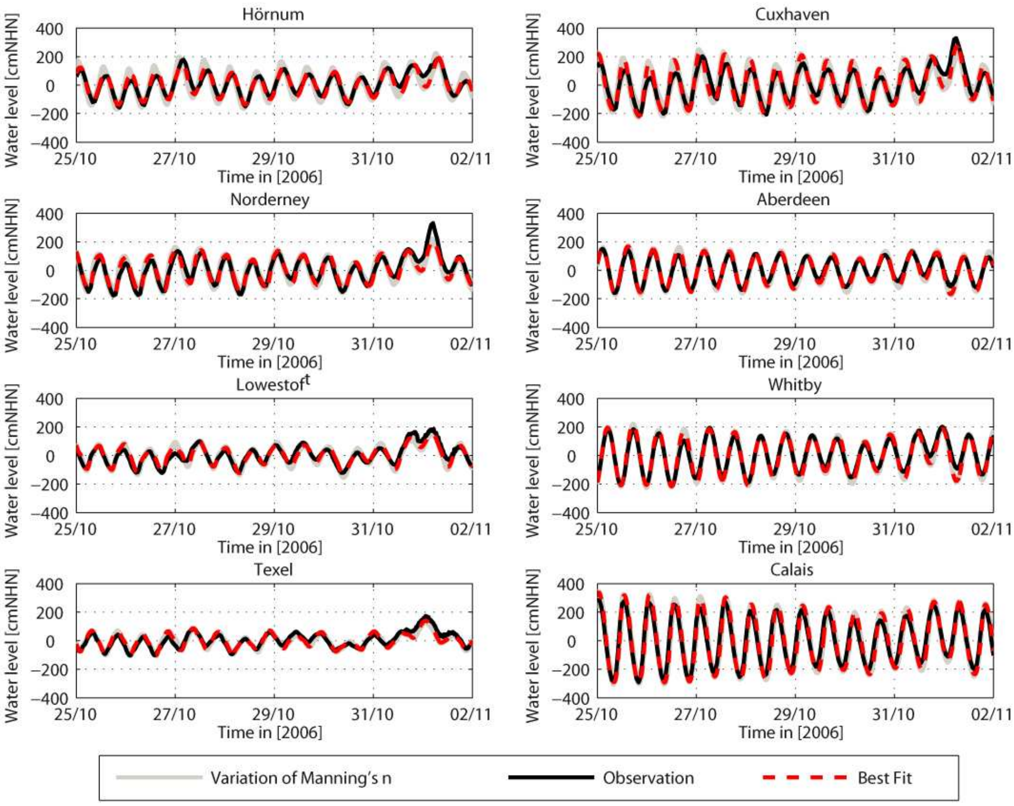

Figure 3. Model calibration at different location

For this reason, a bias correction method for the modelled water levels was developed; this correction almost completely eliminates the bias. The correction is performed in three steps. Firstly, tidal high water levels of observed and modelled water levels are computed and sorted in ascending order. Secondly, the differences between the distributions of observed $F_n(Wo_{,js})$ and modelled $F_n(Wm_{,js})$ tidal high waters at a certain location s and a specific year j are calculated as follows:

$$B_{c,js} = F_n\left(Wo_{js}\right) - F_n\left(Wm_{js}\right). \tag{3}$$

The bias-correction (B_C) is added to the distributions of the modelled $F_n(Wm_{,js})$ tidal high waters, so that the bias is eliminated and the resulting values correspond to the observational tidal high waters:

$$F_n\left(Wo_{js}\right) = Bc_{js} + F_n\left(Wm_{js}\right). \tag{4}$$

At each location, where observational data is available within the considered period, the bias can be eliminated using this procedure. However, as the model also generates water levels between the gauged sites, the bias-correction needs to be transferred to these locations. Thus, thirdly, the bias-correction is interpolated to the locations between the tide gauge sites. The interpolation is performed using the *Inverse Distance Weighted* (IDW) Method (e.g. Environment Agency, 2011).

The three steps to perform the bias-correction are shown in Figure 4. Figure 4a shows all grid-points (black dots) of the model, for which water level time series are available from the 40-year hindcast. Tide gauge locations are depicted with green circles (see the column "correction" in table 1). To clarify the methodology, the tide gauge of Hörnum was chosen as an example (red circle in Figure 4a). In Figure 4b, the distributions of observed (black line) and modelled (red line) high waters for Hörnum are shown. The bias-correction, i.e. difference between the two distributions according to equation (3), is depicted as blue line. Any value of the cumulative distribution between $0 \le F_n(x) \le 1$ yields a value to correct the modelled data. Using the example of $F_n(0.2)$ yields a correction of Δh=4.955 cm. In Figure 4c, the interpolation of the bias-correction to all grid points is shown. The corrected water levels are calculated by summing up the distribution of the modelled water levels and the interpolated bias-correction at each individual location. The figure indicates that bias-corrections for $F_n(0.2)$ are less in the northern parts (coastline of Schleswig-Holstein) than in the western parts (coastline of Lower Saxony) of Germany. These findings are similar for all the other $F_n(x)$, highlighting the good quality of the bathymetry used along the northern German coastline. Figure 4d shows a comparison of observed and modelled high waters before (red dots) and after (blue dots) the bias-correction is applied. The direct model output tends to underestimate higher high-waters whereas lower high waters are overestimated.

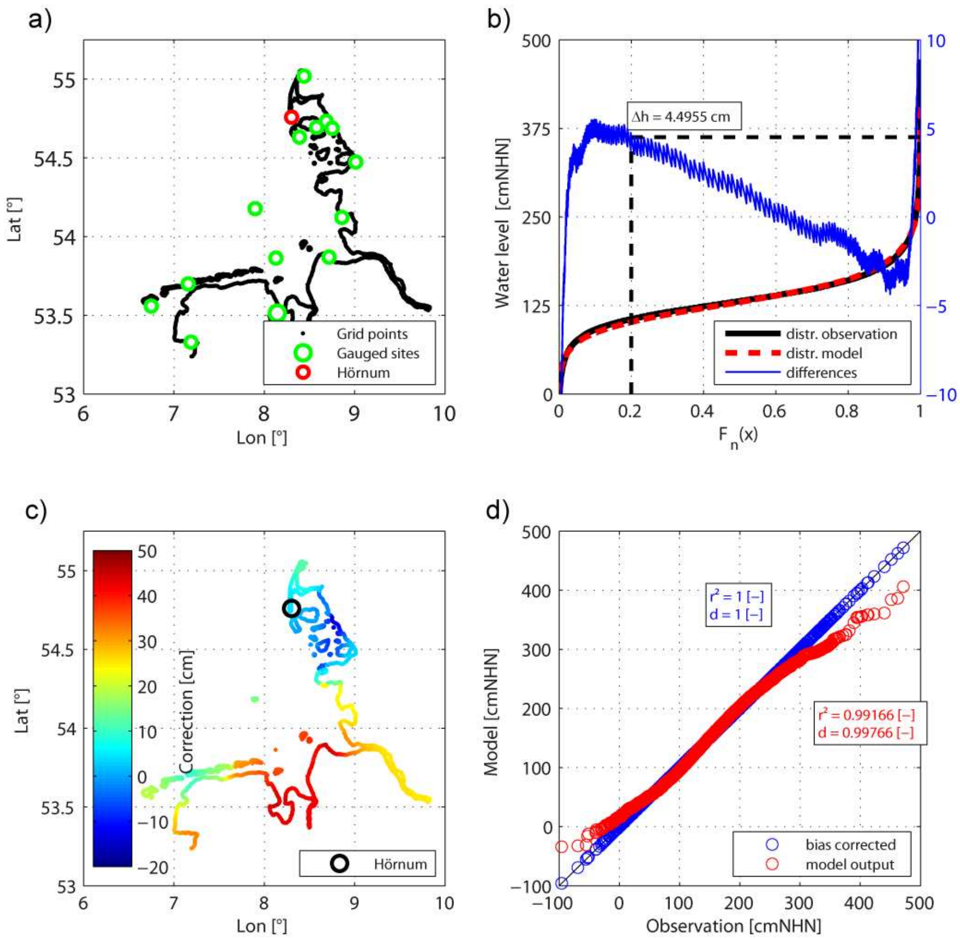

Figure 4. Example of performing the bias-correction

By analysing water levels derived with a hydrodynamic model, it is possible to consider local effects due to complex coastlines and bathymetries. The bias-correction helps to compensate model deficiencies, particularly resulting from the coarse input forcing data. Assuming that local disturbances are also acting on neighbouring grid points, the bias-correction is interpolated to nearby stations. This statement is validated in the next section.

5.4 Validation

In the previous section a methodology to correct the bias of numerically simulated water levels was presented using the tide gauge of Hörnum as an example. To validate the approach, the same methodology is applied to Pellworm Harbour. This 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 modelled water levels of Pellworm Harbour are adjusted using the bias-correction that has been interpolated from neighbouring stations. This is to test the general applicability of the methodology.

In Figure 5a, the regression of observed and modelled tidal high water levels at Pellworm Harbour is shown (red dots). As before in Hörnum, the largest differences are found in the highest and lowest tidal high waters (r^2=0.99 [-], d=0.98 [-]). Applying the bias-correction (blue dots) eliminates most of the deviations, however, not yielding equality (r^2=0.999 [-], d=0,999 [-]). The remaining differences between the distributions of observed and bias-corrected water levels are shown in Figure 5b, indicating that the remaining differences are larger for lower high waters. The same behavior is evident from in Figure 5c, where lower percentiles show a tendency to have larger deviations than higher percentiles.

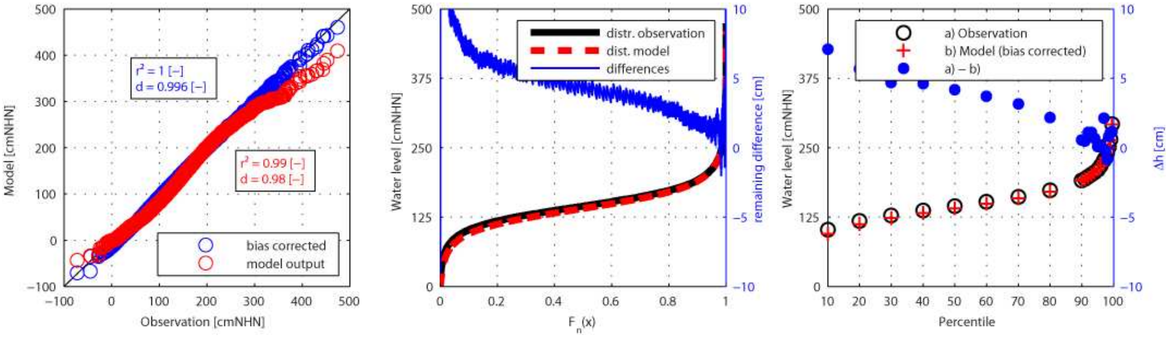

Figure 5. Validation of bias-corrected water levels at Pellworm Harbour

A compilation of all efficiency criteria applied to all 16 validation sites (see table 1) is shown in figure 6. The red dots depict the comparisons of observed and modelled water levels; the blue dots show the comparisons of observed and bias-corrected modelled water levels. Figure 6a highlights that the bias-correction increases the *Coefficient of determination* r^2 at all stations, reaching values of $r^2 \approx 1$ [-]. Figure 6 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 at these stations the index of agreement was already high before the bias-correction was applied. The validation showed 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.

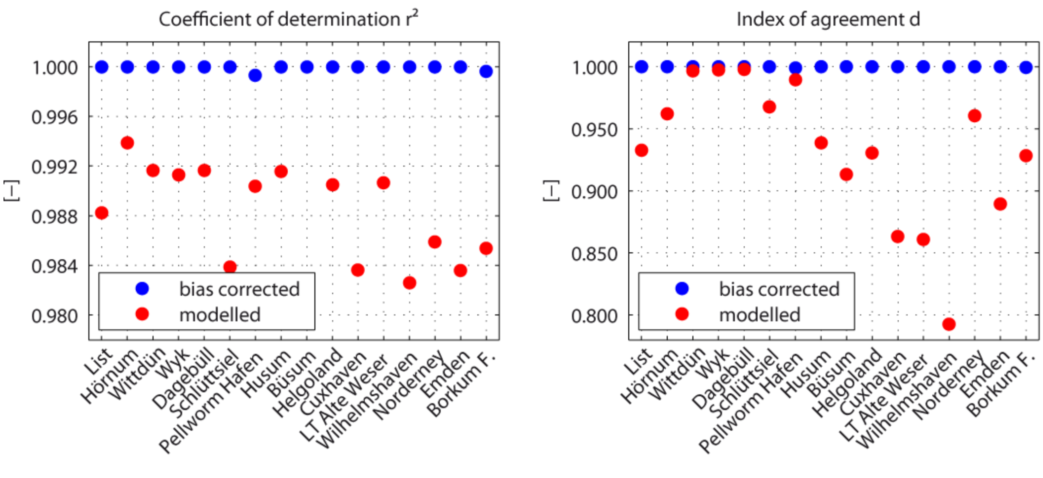

Figure 6. Compilation of efficiency criteria applied to 16 stations

6. **Extreme value analysis**

In this section, extreme value analyses are conducted for the North Sea coastline of Schleswig-Holstein (north-eastern German Bight). The analyses are based on the water level time series derived from the model hindcast after applying the bias-correction as discussed above. Along the coastline of Schleswig-Holstein, water level time series are available for about 900 coastal grid points, located approximately every 1 km (i.e. the mean distance) along the coast.

Over the past half a century, several different extreme value analysis methods for estimating probabilities of extreme still water levels have been developed (see Haigh et al., 2010 for an overview). The available methods are based on different approaches, each having strengths and weaknesses. In the "classical" approach, observed total water level records are directly used to derive return water levels; these methods are usually referred to as direct methods. In coastal hydrology, two of these direct methods are most often applied, namely the block maxima (BM) method and the peaks over threshold (POT) method. Arns et al. (under review) investigated the applicability of these two methods to water level records in the German Bight. It was shown, that the POT method generally yields better results if the model set-up is carefully chosen. Arns et al. (under review) provided recommendations of how to conduct consistent and reliable extreme value analyses of water levels in the German Bight, with minimal subjectivity. Here, we adopt the recommended approach, which is based on the on the assumption that the sample (i.e. all values above a threshold) is characterized by the generalized Pareto distribution (GPD). The GPD is defined as

$$\mathrm{GPD} = 1 - \left[1 + \xi y \div \tilde{\sigma}\right]^{1 \div \xi}, \text{ with } \tilde{\sigma} = \sigma + \xi(u - \mu), \qquad (5)$$

where the location is depicted as μ, the scale as σ, the shape as ξ, and the threshold as u (Coles, 2001). The POT sample is created using a predefined threshold. The threshold selection is often subjective and this can potentially lead to different outcomes, especially when comparing the results from many sites along a coastline. In analysing different threshold selection criteria, Arns et al. (under review) showed that the 99.7^{th} percentile leads to stable and consistent results in the German Bight. Furthermore, it was shown that the storm surge of 1976 has to be included in the statistical analyses; this event was the highest one ever recorded in large parts of the German Bight. This is why the model hindcast covers the period from 1970 to 2009, including the years prior to 1976 to test the stability of the results from the extreme value analyses. Parameter estimation is conducted using the well-known *maximum likelihood method* (see. e.g. (Smith, 1986; Hosking and Wallis, 1987). The declustering of the sample is achieved by using the *extremal index* (see. e.g. Smith and Weissmann, 1994; Ferro & Segers, 2003; Coles, 2001).

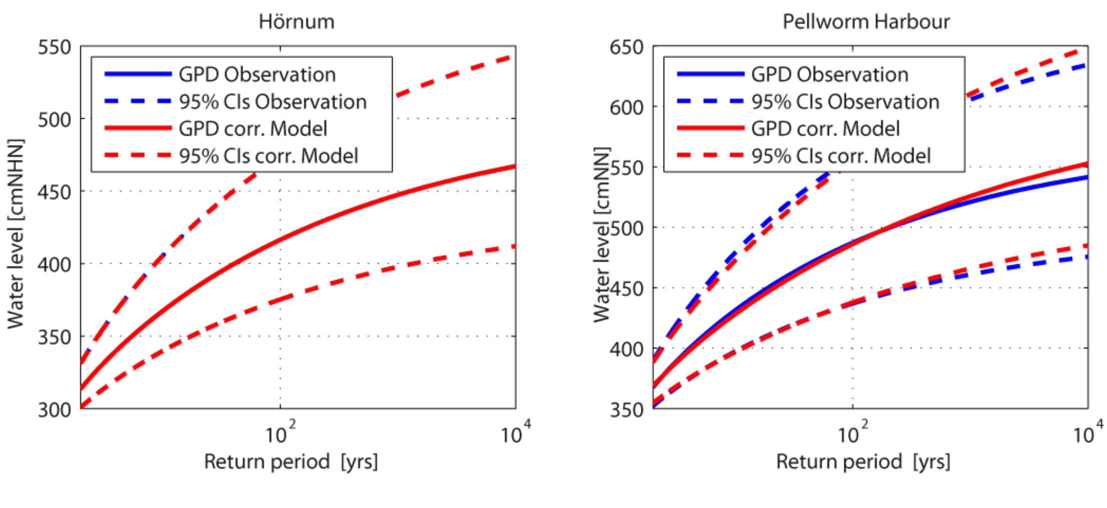

Figure 7. Return water levels for a) Hörnum and b) Pellworm Harbour

Return periods T and associated return water levels are calculated using both measured water level records and water level time series derived from the model hindcast after applying the bias-correction. In Figure 7a, return water levels for Hörnum are shown.

As expected (Hörnum was considered for the correction), there are no differences in the estimates from the observed (blue line) and modelled (red line) water levels. Figure 7b shows the results for Pellworm Harbour (not considered for the correction). In this case we find slight differences in the return water level estimates. However, up to return periods of approx. T=400 yrs, the differences are below $\Delta h \leq 2$ cm reaching $\Delta h \leq 5$ cm for a return period of T=1.000 yrs. The maximum of Δh=11 cm occurs in T=10.000 yrs. The input period only covers 40 yrs, so that extrapolation to 10.000 yrs or even 1.000 yrs is highly debatable. The deviations referred to estimates based on observational data are therefore considered acceptable. The bias-correction is suitable for the modelled water levels in the German Bight, which are envisaged to serve as input for extreme value analyses.

In Figure 8a, regionalized water levels with a return period of T=200 yrs are shown for the entire coastline of Schleswig-Holstein. Water levels in the southern parts of Schleswig-Holstein are higher than in the northern parts, most likely a result of shallow water effects and meteorological forcing. Figure 8b exemplarily shows regionalized return water levels for the Hallig Nordstrandischmoor, 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. Using the regionalization enables return water levels to be derived reliably for this un-gauged region. The information obtained here is used as a basis for the design of protection measures and are also useful risk analyses in un-gauged regions like the Halligen.

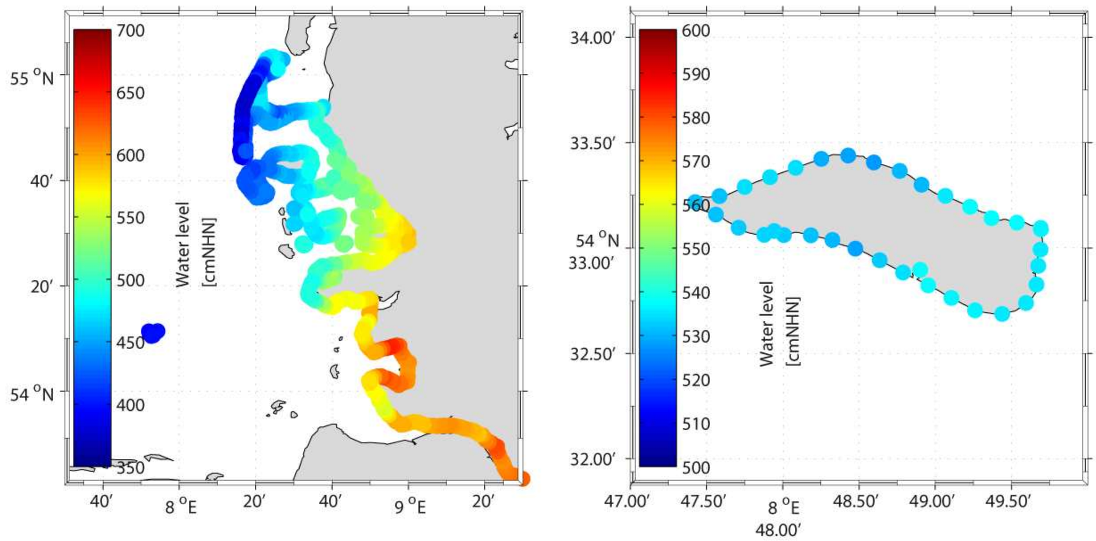

Figure 8. a) Regionalized return water levels along the coastline of Schleswig-Holstein; b) regionalized return water levels at the Hallig Nordstrandischmoor

7. Summary

In this paper we described a method to derive information about the likelihood of extreme water levels at sites with no or little water level data, and applied this along the entire coastline of Schleswig-Holstein in northern Germany. We showed that water levels derived from a hydrodynamic model can be used to calculate reliable extreme water return levels. Especially regions with no or only few tide gauge stations can benefit from this methodology. 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 neighbouring grid points using a *Inverse Distance Weighting* interpolation method. As a result, regionalized extreme water return levels at un-gauged sites are obtained, that account for locally confined coastal attributes. This information can be used for planning purposes and risk analyses.

Acknowledgements

All analyses presented here were performed within the “ZukunftHallig” research project, a German CoastalEngineering Research Council (KFKI) project, funded by the German Federal Ministry of Education and Research BMBF through the project management of Projektträger Jülich PTJ under the grant number 03KIS093.

References

Andersen, O. B. (1995), Global ocean tides from ERS 1 and TOPEX/POSEIDON altimetry, J. Geophys. Res

Arns, A., Wahl, T., Haigh, I.D., Jensen, J., Pattiaratchi, C., (under review). Extreme sea level statistics: direct methods and recommendations. submitted to: Coastal Engineering.

Coles, S. (2001): An introduction to Statistical Modeling of Extreme Values. Springer Verlag, London.

Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, B.E. Gleason, R.S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R.I. Crouthamel, A.N. Grant, P.Y. Groisman, P.D. Jones, M. Kruk, A.C. Kruger, G.J. Marshall, M. Maugeri, H.Y. Mok, Ø. Nordli, T.F. Ross, R.M. Trigo, X.L. Wang, S.D. Woodruff, and S.J. Worley, 2011: The Twentieth Century Reanalysis Project. Quarterly J. Roy. Meteorol. Soc., 137, 1-28

Dalrymple, T. (1960): Flood-frequency analyses, U.S. Geol. Surv. Water Supply Pap., 1543-A.

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

Ferro, C. A.T., Segers, J. (2003): Inference for Clusters of Extreme Values. Journal of the Royal Statistical Society. Series B (Statistical Methodology) , Vol. 65, No. 2 (2003), pp. 545-556

Haigh, I.D., Nicholls, R., Wells, N. (2010): A comparison of the main methods for estimating probabilities of extreme still water levels, Coastal Engineering, Volume 57, Issue 9, September 2010, Pages 838-849

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., Wallis, J.R. (1997): Regional frequency analysis, Cambridge University Press, Cambridge.

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

Krause, P., Boyle, D.P., Bäse, F. (2005): Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5, Pages 89–97.

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

Novikov, A., Bagtzoglou, A. (2006): Hydrodynamic Model of the Lower Hudson River Estuarine System and its Application for Water Quality Management, Water Resources Management, P 257-276

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

Rao, A.R. and Hamed, K.H. (2000). Flood frequency analysis. CRC Press, Boca Raton.

Schneider, S.H., S. Semenov, A. Patwardhan, I. Burton, C.H.D. Magadza, M. Oppenheimer, A.B. Pittock, A. Rahman, J.B. Smith, A. Suarez and F. Yamin, 2007: Assessing key vulnerabilities and the risk from climate change. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds., Cambridge University Press, Cambridge, UK, 779-810.

Smith, R.L. (1986): Extreme Value Theory Based On The r Largest Annual Events. Journal of Hydrology.

Smith, R.L., Weissman, I., 1994. Estimating the Extremal Index. Journal of the Royal Statistical Society. Series B (Methodological). 56, 515-528.

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.

Wahl T., Haigh I.D., Woodworth P.L., Albrecht F., Dillingh D., Jensen J., Nicholls R.J., Weisse R., Wöppelmann G. (accepted): Observed mean sea level changes around the North 3 Sea coastline from 1800 to present. Submitted to: Earth Science Reviews.

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

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