Subject: Responses to reviewers' comments and suggestions

Manuscript Number: NHESS-2022-223

Title: Improvements to the detection and analysis of external surges in the North Sea

Dear reviewer,

We want to thank you again and like to express our appreciation for the constructive comments– we believe that the consideration of these comments has greatly enhanced the quality of the revised manuscript.

Please find enclosed in this letter detailed answers to reviewers' comments, as well as the corresponding actions performed to the original manuscript submission.

Sincerely,

Alexander Müller, Dr. Birgit Gerkensmeier, Benedikt Bratz, Clemens Krautwald, Dr.-Ing. Olaf Müller, Dr.-Ing. Nils Goseberg, Dr. Gabriele Gönnert

2 REVIEWER #2

- *"Does the paper address relevant scientific and/or technical questions within the scope of NHESS? Yes.*
- *Does the paper present new data and/or novel concepts, ideas, tools, methods or*
- *results? Yes.*
- *Are these up to international standards? Yes.*
- *Are the scientific methods and assumptions valid and outlined clearly? See comments below.*
- *Are the results sufficient to support the interpretations and the conclusions? Yes.*
- *Does the author reach substantial conclusions? Yes.*
- *Is the description of the data used, the methods used, the experiments and calculations made, and the results obtained sufficiently complete and accurate to allow their reproduction by fellow scientists (traceability of results)? See comments below.*
- *Does the title clearly and unambiguously reflect the contents of the paper? Yes.*
- *Does the abstract provide a concise, complete and unambiguous summary of the work done and the results obtained? Yes.*
- *Are the title and the abstract pertinent, and easy to understand to a wide and diversified audience? Yes.*
- *Are mathematical formulae, symbols, abbreviations and units correctly defined and used? If the formulae, symbols or abbreviations are numerous, are there tables or appendixes listing them? Yes, mostly.*
- *Is the size, quality and readability of each figure adequate to the type and quantity of data presented? Yes.*
- *Does the author give proper credit to previous and/or related work, and does he/she indicate clearly his/her own contribution? Yes.*
- *Are the number and quality of the references appropriate? Yes.*
- *Are the references accessible by fellow scientists? Yes.*
- *Is the overall presentation well structured, clear and easy to understand by a wide and general audience? See comments below.*
- *Is the length of the paper adequate, too long or too short? Adequate.*
- *Is there any part of the paper (title, abstract, main text, formulae, symbols, figures and their captions, tables, list of references, appendixes) that needs to be clarified, reduced, added, combined, or eliminated? The figure and table captions should be clarified and sometimes added for better understanding.*
- *Is the technical language precise and understandable by fellow scientists? See comments below.*
- *Is the English language of good quality, fluent, simple and easy to read and understand by a wide and diversified audience? Yes.*
- *Is the amount and quality of supplementary material (if any) appropriate? Yes.*

The manuscript " Improvements to the detection and analysis of external surges in the North Sea" by Müller et al. analyses external surges in the North Sea based on existing different datasets. As a result, the knowledge about external surges as well as the detection with an automated algorithm is improved. Also, a new data set is provided to account for past external surges.

External surges can have a significant impact on increasing storm surges. Therefore, they are extremely relevant for coastal protection concerns. In the past, there was always the problem that external surges could only be detected indirectly. Observational data, for example at the Heligoland gauge, or hydrodynamic numerical models were used for this purpose. However, there have always been limitations with regard to the significance and correct contribution of the physical processes, especially during the transition of the external surge to the shelf. Therefore, I was very much looking forward to reading the manuscript.

I liked the approach and the implementation very much. There is a common thread and the graphics are also appealing. A good idea was implemented with proven methodology. To summarize, the manuscript is mostly written in a clear and concise manner whose proposed research outcome has a strong applicative character. There is an undisputed relevance of the methodology and applications presented. However, as my comments below indicate, there is a certain need for discussion and clarification at some points, which can be resolved and clarified with a little more careful work. Therefore, I recommend publication not before a minor revision."

Answer to General Comment of Reviewer #2:

Thank you for the thorough and detailed review of the manuscript and many suggestions for improvement. Thank you as well for your interest in discussion of the results and suggestion to improve the analysis.

Answers to the specific and technical comments can be found below.

Specific comments

Comment 2.1:

"In the manuscript the background, the data and the methodology are described and finally presented and discussed. I have no fundamental reservations about the methodology and the general procedure, nor have I found any errors that are obvious to me. Overall, the paper makes a good impression. But especially the methodology needs a revision. It is difficult to fully follow the aspects and those used. This should be reworked so that a clear structure is recognizable. Also, the assumptions made and the different datasets to finally determine the external surges should be made clearer."

Answer to comment 2.1:

Thank you for your improvements of the structure of the article. We have reorganized the methodology section at multiple points, namely:

- The methodology to determine external surges is now divided into three parts to resemble the automated steps of the analysis.
- The methodology was extended with a short introduction into the three subsections.
- Assumption, especially on the wind setup on the British coast are so named as such.

Lines 182-188:

To isolate external surges from hydrographic records, all other major components have to be eliminated. The methodology was developed by Gönnert (2003), but was adapted for this study to make the time steps independent from tidal cycles and therefore enable a more refined analysis. While Gönnert (2003) used the average residual between high and low water, resulting in time steps of 6.25 h, here the time steps are shortened to 1 h, allowing more precise assertions of the duration, height, and timing of the events. Figure 2 gives an overview over the general steps in this process, consisting of three automated steps, each eliminating one contributor to water levels, and the consecutive manual quality check. Each step explained in further detail in the following three subsections.

Lines 193-205:

2.3.1 Calculation of the non-tidal residual

First, the astronomical tide (h_{astro}) is eliminated from the water level records to retrieve the non-tidal residual (Δh). The analysis and prediction have a long history on the North Sea coast and are still continuously improved (e.g. Amin 1982; Müller-Navarra 2013; Boesch und Müller-Navarra 2019). As continuous predicted tides are available from the respective authorities, no separate analysis in conduced in this study. By removing the predicted tide, the non-tidal residual remains, which includes all parts of the water level, that do not follow the harmonic forcing of the sun and the moon:

$$\Delta h_i = h_{obs,i} - h_{astro,i} \quad (3)$$

2.3.2 Wind setup hindcast

In step 2, the influence of meteorological factors on the non-tidal residual during external surges has to be discussed to derive the wind setup and the residual individually. In this study, wind setup describes the effect of wind and air pressure on the water level, but only including local forcing. The residual therefor describes the remaining components of the water level, for example remote meteorological forcing, lateral oscillations or the influence density variations.

The methodology to determine the wind setup varies between the tide gauges due to different local and regional bathymetry and time lag between wind and water level maxima (Proudman and Doodson, 1926; Dibbern and Müller-Navarra, 2009).

Lines 312-332:

2.3.3 Calculation of the residual

The meteorological influences from wind setup and air pressure are used in Eq. (7) to calculate the residual in Cuxhaven and Texel. For the wind setup, the tables, that were compiled in Sect. 3.3.2, are interpolated linearly.

$$\Delta h_R = \Delta h - \Delta h_W - \Delta h_{AP} \quad \textbf{(4)}$$

During timespans, where only tides and local meteorological forcing influence the water level, the residual should be approximately 0 cm, but many deviations remain, which may be caused by external surges, but can also be due to the simplified approach to hindcast the wind setup, natural variability, or other local phenomena. To find a collective term for these deviations, we call them surges, without specifying their origin. In step 3, the surges are filtered for external surges based on the following conditions proposed by Gönnert (2003):

- The non-tidal residual in Aberdeen is greater than 40 cm.
- Surge heights generally increase from Aberdeen to Immingham, caused by local winds, with a maximum decrease of 10 cm.
- The arrival at Immingham is at least 2 h later than at Aberdeen, but no more than 5 h later. The time of arrival is defined as the time when the non-tidal residual is over 10 cm in height.

For each external surge, the maximum height and arrival time at each tide gauge is determined and the non-tidal residual (for Aberdeen and Immingham) or the residual (for Texel and Cuxhaven) as well as the data quality flags are plotted for 6 h before the arrival of the external surge in Aberdeen to 42 h after the arrival.

Finally, a manual quality check is conducted on the plots of the external surges. A total of four surges were eliminated due to either of the following reasons:

- Missing tide gauge data during the beginning of the surge in Aberdeen or Immingham. The height or the offset of the surge may be influenced by these missing values.
- Long surge periods that can be attributed to wind setup in Aberdeen and Immingham

Lines 213-225:

Due to the different main wind direction, the coincidence of either two of the last three factors would require drastic and sudden changes in wind conditions.

For the Immingham tide gauge, Rossiter (1959) concluded that the following parameters influence the residuals the most:

- The residual at Aberdeen with a time lag of 5 h
- Northeast winds over the North Sea with a time lag of 6 h

The first accounts for local air pressure as well as external surges propagating southwards along the British coast. As the entrance of external surges is hindered by eastern winds, these two main factors are assumed to not coincide in general. For these two British tide gauges it can therefore be assumed that wind setup does not need to be accounted for during the detection of external surges, which was proven successful in previous studies (Gönnert, 2003). Further research on the meteorological conditions causing external surges is certainly needed, especially focusing on wind and air pressure patterns. This detailed analysis could also enable more detailed statements about the co-occurrence of the weather patterns, that cause non–tidal residuals in Aberdeen and Immingham. This analysis is, however, beyond the scope of this study but will be the focus of further work on external surges in the North Sea.

Comment 2.2:

"The same applies to the results chapter. I recommend starting with an explanation of the final data sets and then explaining the generated results. Furthermore, structural aspects have been noticed here. Parts of the results should be moved to the methodology, others to the discussion. Then the manuscript becomes a good technical article with a lot of practical usable content."

Answer to comment 2.2:

Thank you again for your help to improve the structure. The results chapter, especially the section on weather situations has been restructured based on the comments of Reviewer#1, which match your suggestions, that certain paragraphs are part of the methodology or discussion respectively. The data sets are now mentioned first in Section 2.5 as well.

Line 345-373:

2.5 Meteorology

External surges found in this study are mainly compared to the dataset of Gönnert (2003), that spans the years 1971 to 1995. To distinguish between the datasets, the dataset of Gönnert is hereafter called DataSet1 and the dataset derived from the automated approach is called DataSet1.

The occurrence of external surges in the North Sea basin is strongly coupled to storm systems in the North Atlantic. Although the process of the physical meteo–oceanic coupling is not within the scope of this work, it will be important to correlate the observations from the tide gauge data to general weather pattern. To briefly repeat meteorological conditions in a context of surge generation, the European weather situations during the beginning of the observed external surges are assembled first. The European weather situations were originally defined by Hess and Brezowski (1977) and determined by Werner and Gerstengarbe (2010) for the duration from 1881 to 2009. From 2010 onwards, the records from German Meteorological Service (2021) are used. The Agency for Roads, Bridges and Waters Hamburg (2012) found 61 of the 73 external surges of Dataset1 to occur during four weather situations (WZ: western situation cyclonic, WA: western situation anticyclonic, SWA: south western situation anticyclonic, BM: high pressure bridge Central Europe).

The characterisation of external surges with respect to low pressure cells can be summarized from Werner and Gerstengarbe (2010) as follows:

- WZ: Single disturbances with high–pressure cells in between travel from Ireland over the British Isles, North and Baltic Sea towards Eastern Europe. The driving low–pressure cells is located north of 60° N.
- WA: The central low–pressure cell is often located north of 65° N with single disturbances travelling from west of Scotland over Scandinavia towards the Baltic.
- SWA: A low–pressure system is mostly located over the middle of the North Atlantic and the western Norwegian Sea. Single disturbances travel to the northeast.
- BM: A high–pressure bridge between the Azores and Eastern Europe with an eastward directed frontal zone north of it and single disturbances travelling eastwards.

In the analysis of DataSet2, an additional weather situation is identified, that correlates with an increased number of external surges (NWZ, north western situation cyclonic) and has the following characteristics:

- NWZ: Extensive low–pressure area over Scotland, the Norwegian Sea and Scandinavia with single disturbances travelling over the British Isles towards eastern Central Europe.

The detailed analysis of weather situations is presented in Sect. 3.2, while Sect. 3.3 analyses the influence of external surges on storm surges in the German Bight.

Comment 2.3:

"The last aspect is that it is not always clear to me whether the presented methodology really works in an automated manner. There is always talk of manual additions and adjustments. This should be made clearer again.

Further points can be found in the following paragraph and listing."

Answer to comment 2.3:

Thank you for this important aspect that we have clarified in the revision. Manual and automated steps are now differentiated clearer in the methodology and we added this to the discussion. With the restructuring of the methodology, the distinction between manual and automated steps should now be easier. We have also added comment on which parts are automated and mark automated and manual steps in Figure 2. In the abstract, the method is now described as semi-automated to give a more accurate description.

Lines 14-15:

This work describes an improved and semi-automated method to detect external surges in sea surface time histories.

Lines 186-191 and Figure 2:

Figure 2 gives an overview over the general steps in this process, consisting of three automated steps, each eliminating one contributor to water levels, and the consecutive manual quality check. Each step explained in further detail in the following three subsections.

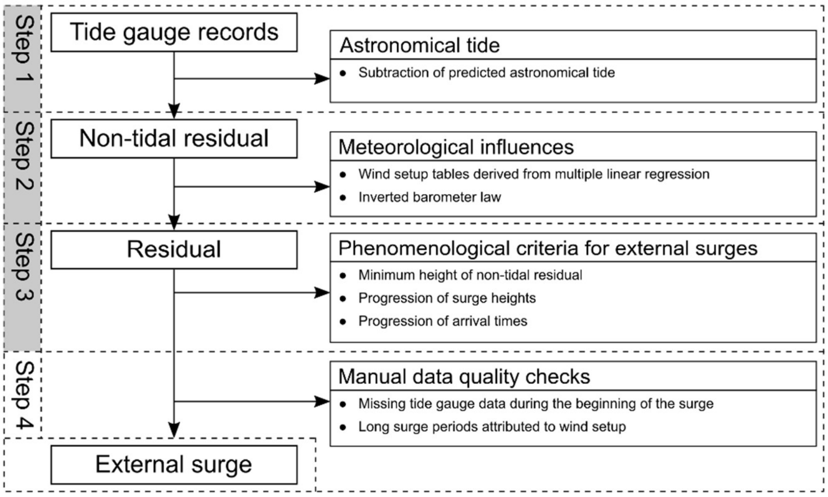

Figure 2: Schematic representation of the process to identify external surges. Grey backgrounds mark automated steps in the analysis.

Lines 322-323:

Finally, a manual quality check is conducted on the plots of the external surges. A total of four surges were eliminated due to one of two reasons: […]

Technical comments

Comment 2.4:

"General:

- *Please define clearly "surge", "storm surge", "surge residual", "external surge", "nontidal residual", "residual" etc. In between it is written very hard to follow.*

Answer to comment 2.4:

Thank you for this comment to make the manuscript more concise. We have added further explanation to several sections to define the phrases. The manuscript has additionally been checked for correct use of these.

Lines 38-40:

The most common effects related to wind setup are storm surges, caused by high winds over the North Sea, which represent the proportionately largest increases in extreme water levels.

Lines 47-48:

They were determined to originate from the response of the Northeast Atlantic to low pressure cells moving eastwards and to be amplified at the continental shelf.

Lines 197-196:

By removing the predicted tide, the non-tidal residual remains, which includes all parts of the water level, that do not follow the harmonic forcing of the sun and the moon:

Lines 201-203:

In this study, wind setup describes the effect of wind and air pressure on the water level, but only including local forcing. The residual therefor describes the remaining components of the water level, for example remote meteorological forcing, lateral oscillations or the influence density variations.

Lines 315-319:

During timespans, where only tides and local meteorological forcing influence the water level, the residual should be approximately 0 cm, but many surges remain, meaning any deviations, caused by the simplified approach, natural variability or phenomena like external surges. In step 3, these surges are filtered for external surges based on the following conditions proposed by Gönnert (2003):

Comment 2.5:

- *no blank character before "%"*

Answer to Comment 2.5:

Thanks for the improvements to the typesetting. We corrected all occurrences in the manuscript.

Comment 2.6:

- *uniform cross-references (e.g. Figure or Fig.)*

Answer to comment 2.6:

Thank you. We checked, that the cross-references are formatted according to the NHESS submission guidelines ("Figure" at the beginning of a sentence, abbreviation in running text).

Comment 2.7:

- *easterly/westerly winds*

Answer to comment 2.7:

Thank you for your correction. All directions have now been checked again and corrected.

Comment 2.8:

- *try to avoid "in overall good agreement", "very often", "highly unlikely" or "relatively small" without quantitative numbers or references*

Answer to comment 2.8:

Thank you for this encouragement of concise language. We have changed several occurrences of these phrases throughout the manuscript, others were already deleted during the rework of Section 3.3. See for example, the answer to comment 2.1.

Lines 56-57:

In contrast to the extensive knowledge available about tides, wind surge and RMSL rise, knowledge about the phenomenon of external surges is still ~~relatively~~ rare.

Lines 107-109:

However, available knowledge of external surges is ~~still relatively~~ small and translation of knowledge and impacts resulting from external surges into practical approaches, e.g. in terms of concepts for design level for coastal protection facilities, is still ~~very~~ rare.

Lines 167:

Missing or ~~highly~~ unlikely data like […]

Lines 307:

Hindcast and measurement are in ~~overall~~ good agreement,

Comment 2.9:

- *inverted barometer effect (IBE) or inverted barometer law (IBL)?*

Answer to comment 2.9:

Thank you for asking for clarification. Both names are used in literature to describe the relation. In the manuscript it is now uniformly called “inverted barometer law” (IBL).

Comment 2.10:

- *Describe figures and tables before they are shown and referred to them.”*

Answer to comment 2.10:

Thank you for this remark to improve the connection between text and figures. The figures and tables were now moved closer to their first mention in the text, but are always after the mention in the text. We however suspect possible changes during the typesetting, especially for the printed/ pdf version, and trust the judgement of the editor.

Comment 2.11 & 2.12.:

“Line 18: surges instead of surge”

“Line 20: tend to occur”

Answer to comment 2.12:

Thank you, the errors have been corrected.

Comment 2.13:

“Line 29: abbreviation of regional mean sea level is MSL or RMSL? If it is changed then adjust throughout the document.”

Answer to comment 2.13:

Thank you for the question. To avoid confusion, the regional mean sea level is now called RMSL.

Comment 2.14:

“Line 33: Probably add some sources for studies dealing with the German coastline? E.g. Dangendorf?”

Answer to comment 2.14:

We agree, that MSL rise also has to be considered locally, especially in coastal protection measures. The aim of the section, however, is to name the relevant processes on a regional scale. A localized discussion of these processes would either add unnecessary length to the introduction or skew the introduction towards a certain area. In our eyes, the Germen Bight is covered adequately in the cited literature as they all include tide gauges on the German coast or in the case of Albrecht et al. analyse them exclusively.

Comment 2.15:

"Line 36: and interactions of the individual effects…"

Answer to comment 2.15:

Thank you for this comment. Nonlinear interactions are now mentioned explicitly as part of extreme sea levels.

Lines 37-38:

For the North Sea coast, extreme water levels are driven by various effects such as wind setup, tides, ~~and~~ intricate local effects, and their interactions (Mikhailova 2011).

Comment 2.16:

"Line 35-45: Some references are missing for the numbers given and the statements."

Answer to comment 2.16:

Thank you for this comment. Several sources have been added to this section.

Lines 36-49:

Besides long–term changes, short–term extreme sea levels in particular pose challenges for risk management and coastal protection in low lying coastal regions. For the North Sea coast, extreme water levels are driven by various effects such as wind setup, tides, ~~and~~ intricate local effects, and their interactions (Mikhailova 2011). The most common effects related to wind setup are storm surges ~~events~~, caused by high winds over the North Sea, which represent the proportionately largest increases in extreme water levels. The disastrous impacts of storm surges are well known for centuries along the North Sea coast and have been an inherent part of coastal protection strategies (Siefert 1991). Well–known impacts of climate change are accounted for in recent coastal protection strategies (Dronkers und Stojanovic 2016).

Besides tides and storm surges, external surges represent an additional phenomenon that generates increased regional extreme water levels in particular along the British, Dutch and German North Sea coast. This phenomenon was first discovered in the 1940s (Corkan, 1948) along the British coastline bordering the North Sea. Increases in water level time histories could not be well explained by local wind setup or tides and were observed to propagate counter–clockwise through the shelf sea similarly to the propagation of tides. They were determined to originate from the response of the Northeast Atlantic to low pressure cells moving eastwards and to be amplified at the continental shelf. Reaching heights of 1.3 m on the British coast and 1.2 m in the German Bight (Koopmann, 1962; Gönnert, 2003), […]

Comment 2.17:

"*Line 57: I'm confused about the start and end year of the data. Please specify.*"

Answer to comment 2.17:

Thank you for highlighting possible misunderstanding about the data of this study and data from previous studies. The section was changed to clarify the distinction.

Lines 61-64:

A secondary aim is to apply the novel automated approach to isolate external surges from water level records since 1995. Those can be combined with the dataset of Gönnert (2003), spanning from 1971 to 1995, and thus ~~to~~ generate a comprehensive dataset of external surges in the North Sea covering a period from 1971 to 2020.

Comment 2.18:

"*Line 59: chapter number wrong?*"

Answer to comment 2.18:

Thank you for also checking correct formatting. The chapter number has been corrected.

Comment 2.19:

"*Line 70: Which period?*"

Answer to comment 2.19:

Thank you for asking. The timespan that was investigated by Gönnert (2003) has been added to the text. For detailed information see Table 5.

Lines 76-78:

Still, besides the work of Gönnert (2003) who manually investigated external surges in the North Sea occurring between 1971 and 1995, no comprehensive datasets of external surges, identifying their occurrence within observational records are known to the authors.

Comments 2.20 and 2.21:

"Line 94: can or cannot?

Line 95: Is there a difference or not? Please clarify."

Answer to comments 2.20 and 2.21:

Thank you for this question. The periods on meteotsunamis and external surges differ by about an order of magnitude and their possible directions of travel. This information has been added to the text. Nonetheless, the knowledge about the interaction of air pressure and sea level as well as the wave transformation on the continental shelf, generated by research on meteotsunamis might be applicable to external surges as well. We now give an example of the differing characteristics on meteotsunamis and external surges to

Lines 100-103:

They have also occurred in the North Sea (Jong, 2004; Sibley et al., 2016) but can ~~clearly~~ be differentiated by their characteristics like period, amplitude and their effect on water levels. For example, the recorded meteotsunamis in the North Sea have periods in the order of one hour and can also travel northwards, while external surges have longer periods of 8 hours to several days and travel counter clockwise (Gönnert 2003).

Comment 2.22:

"Line 100: repetition ("still relatively/very rare")"

Answer to comment 2.22:

The sentence has been rephrased. This sentence was also part of the changes due to comment 2.4.

Comment 2.23:

"Line 104: chapter number wrong?"

Answer to comment 2.23:

Thank you. This chapter number has been corrected as well.

Comment 2.24:

"Line 106: expression ("Not only can")"

Answer to comment 2.24:

Thank you for the language advice. The sentence has been rephrased and moved to the first bullet point to make the section more compact.

Lines 115-117:

The methodology can not only be used to include additional locations in the detection of external surges, but it can also support local authorities in considering external surges for the protection of their communities.

Comment 2.25:

"*Line 123: "surges" is missing*"

Answer to comment 2.25:

Thank you, we have added the missing word.

Comment 2.26:

"*Line 123-128: Some references are missing.*"

Answer to comment 2.26:

Thank you, references about the topography of the North Sea and the Wadden Sea have been added. One number in the text has been adjusted to match the cited source.

Lines 132-138:

The North Sea is a shelf sea connected to the Northeast Atlantic through a 400 km wide opening to the Norwegian Sea and the narrower English Channel, which is 34 km wide at its narrowest point. The depth of the North Sea is on average 95 m but varies between less than 20 m at the Dogger Bank and 700 m in the Norwegian Trench. Also, the coastal morphology in this focus area is very diverse: fjords at the Scottish and Norwegian coasts, cliffs on the west coasts and marshlands at the southern and south–eastern coasts, which border the Netherlands, Germany, and Denmark (OSPAR Comission 2000). These marsh coasts also include the unique biotope of the Wadden Sea (Lotze et al. 2005; Reise et al. 2010).

Comment 2.27:

"*Line 133: surge action?*"

Answer to comment 2.27:

Thank you for this comment, the sentence was changed due to a comment of Review#1 and the term was omitted.

Lines 144-146:

The height of the external surge at the Aberdeen tide gauge was previously used as a proxy for the height during its entrance into the North Sea (Bruss et al., 2011; Ganske et al., 2018), while Immingham is close to the location where most external surges reach their maximum height.

Comment 2.28:

"*Line 140: more specific: tidal high and low water back to 1843*"

Answer to comment 2.28:

Thank you for the addition. This is now clarified in the text.

Comment 2.29:

"Line 141: this statement is risky. analyses of the tide gauge in Cuxhaven repeatedly show different behavior compared to other tide gauges in the German Bight. Furthermore, there is a lack of information from references."

Answer to comment 2.29:

We agree with this critique and thank you for bringing this to our attention. The statement has been omitted as suggested by Reviewer#1, as it does not provide essential information.

Lines 151-153:

The tide gauge in Cuxhaven has been regularly used as reference tide gauge in German coastal protection considerations as it provides a continuous data series (continuous record is available since 1918, tidal high and low waters since 1843)~~, which is nearly undisturbed (e.g. by man-made impacts),~~ and is influenced neither by barriers like the East Frisian Islands nor by the Elbe estuary fluxes.

Comment 2.30:

"Table 1: Can you reduce the size of the table. It seems very large for the relatively little bit of information. Most of the information is given from 151-157 anyway."

Answer to comment 2.30:

Thank you for this suggestion. We were able to combine columns and omitted the mostly empty last column, which notably reduced the overall size. The important information about gaps in the data of the Texel tide gauge is now included in the text.

Table 1:

Table 1: Location, offset to NAP, data sources and percentage of flagged data (flags from the data providers, missing and unreasonable values) for each measurement station.

Tide gauge (Lat., Lon.)	Offset tide gauge zero (TGZ) to NAP [m]	Tide gauge data and astronomical tide source	Tide gauge data time steps	Weather data source	Weather data time steps	Flagged data points [%]
Aberdeen (57° 08' 38.6" N, 02° 04' 48.5" W)	−2.45	NOC British Oceanographic Data Center (BODC) (2021)	15 min quality checked	-	-	2.0
Immingham (53° 37' 51.7" N, 00° 11' 9.7" W)	−4.10			-	-	2.2
Texel (53° 07' 14.2" N, 4° 43' 59.4" E)	0.00	Rijkwaterstaat (2021)	10 min quality checked	Koninklijk Nederlands Meteorologisch Instituut (KNMI) (2021) Station De Kooy	1 h	4.4

Cuxhaven (53° 52' 03.7" N, 8° 43' 02.7" E)	−5.03	1995–1997: Wasserstraßen- und Schifffahrtsverwaltung des Bundes (WSV) (2021a) 1998–2020: Wasserstraßen- und Schifffahrtsverwaltung des Bundes (WSV) (2021b) Astronomical tides: Federal Maritime and Hydrographic Agency (BSH) (2021)	1995–1997: 1 h 1998–2020: 1 min Astronomical tide: 10 min quality checked	DWD Climate Data Center (CDC) (2021) Stations *UFS Deutsche Bucht, UFS TW Ems,* Helgoland, Cuxhaven	1 h quality checked	2.4

Lines 169-170:

In the data of the tide gauge at Texel, two gaps, spanning more than a week, exist between 11 January and 22 November 2002 and 16 June and 2 July 2020.

Comment 2.31:

"Line 159-160: "Mean wind direction (averaged over the last 10 min before the timestamp)" This information is important. But I don't understand how you can average the last 10 minutes when the weather data has a resolution of 1 hour?"

Answer to comment 2.31:

Thank you for this question. We suppose a misunderstanding: the averaging was not conducted for this study but is part of the processing of data before it is published by the providers. As we agree that this is so far not pointed out in the manuscript, we have rephrased the paragraph.

Lines 176-178:

The weather data used to hindcast wind setup at the tide gauges at Texel and Cuxhaven is provided by the meteorological institutions as the following products:

- Mean wind direction (average of the last 10 min before the full hour)

Comment 2.32:

"*Figure 2: what does "surge heights" mean?"*

Answer to comment 2.32:

As their source is still unknow, surges in the residual can not be identified as external surges at this step. A definition of surges has been added in response to comment 2.4. (cf. line 313-315), which should also clarify the term surge height in Fig. 2.

Comment 2.33:

"*Line 168: "(2003)" is missing"*

Answer to comment 2.33:

Added.

Comment 2.34:

Line 171: “The methodology of this step varies” of which step?

Answer to comment 2.34:

Thank you for the question. During the revision of the section to answer to comment 2.4, it has been clarified, that this statement refers to Step 2 of the process.

Comment 2.35:

“Line 175: Did you calculate the astronomical tide? Which method is used? What about uncertainties?”

Answer to comment 2.35:

The astronomical tides were provided as well by the providers of the tide gauge data or other government agencies. Table 1 has been altered to clarify the sources of the tide data. For the British tide gauges, not the astronomical tide is provided in the raw data, but the water level and the non-tidal residual, so that the astronomical tide can be calculated. For further information on the methods and uncertainties, the authors refer to the providing agencies, as their discussion would be beyond the scope of this paper. The uncertainties are not always provided with the data, but are now mentioned in lines 158-162.

Comment 2.36:

“*Line 193: What is surge heights here? External surge or residual or non tidal residual?”*

Answer to comment 2.36:

Thank you for asking for clarification. The term surge is now defined in response to comment 2.4. As air and water temperature can also influence the calculated height of other types of surges (return surges, lateral oscillations, …), the more general term is used here.

Comment 2.37, 2.38 & 2.39:

“*Line 195: zonal/meridional wind”*

“*Line 198: tables”*

“*Line 206: with instead of at”*

Answer to comment 2.37:

Thank you for your suggestion to improve the language of the manuscript. We gladly incorporated the terms.

Line 228-229:

[…] like ~~x and y components of the~~ quadratic and cubic wind speeds of the zonal and meridional winds, […]

Line 230:

These tables group the measurements by two criteria to account for non-linear interactions between tide and wind setup:

Line 239:

The time lag between wind measurements and water level deviation is given by Müller-Navarra and Giese (1999) with 3 h.

Comment 2.40:

"*Line 210: reference?*"

Answer to comment 2.40:

The value was assessed by also applying the MLR to the tide gauge Cuxhaven. The RMSE of this hindcast was slightly higher than the RMSE of the hindcast using the tables of the BSH, but their hindcast was similar. We therefore chose to use the hindcast of the BSH to obtain more precise results for the residual. Further we the regression factor for the static air pressure to compare it to the factor of the IBL, which showed good agreement. The section was slightly rephrased to highlight this as proof of our assumption instead of giving literature references.

Lines 239-245:

The time lag between wind measurements and water level deviation is given by Müller-Navarra und Giese (1999) with 3 h. For the influence of static air pressure, the *inverted barometer law* (IBL) is used. It states that a decrease in air pressure by 1 hPa correlates linear to an increase in water level by 1 cm (Koopmann 1962), even though this correlation is simplified and can vary depending on local factors (Ponte und Gaspar 1999; Müller-Navarra und Giese 1999; Olbert und Hartnett 2010)~~, it is found to be in good agreement with the conditions at Cuxhaven.~~ A simplified analysis of wind setup at Cuxhaven, using only wind and air pressure measurements, determined linear factors between 0.9 and 1.25 cm hPa^{-1}, depending on tidal phase and main wind direction, showing that the proposed value of 1 cm hPA^{-1} is a good first approximation of the relationship.

Comment 2.41:

"Line 212: in line 190ff was mentioned that "The influence of other factors like water and air temperature or wind setup should be taken into consideration during the interpretation of surge heights, nonetheless."?"

Answer to comment 2.41:

Thank you for highlighting the need for more detailed explanation of the methodology. Wind and air pressure are the two main contributors to wind setup in Cuxhaven. To simplify the analysis, we have focused on these but are aware, that the calculated height of the residual loses accuracy. Therefore, heights have to be interpreted more carefully. A short discussion has been added.

Lines 247-249:

This certainly reduces the accuracy of the hindcast and has to be discussed during analysis of the height of the external surges, but previous studies have produced satisfactory results based on wind and air pressure data (Dibbern and Müller–Navarra, 2009; Jensen et al. 2013).

Comment 2.42:

"*Table 2: MLR is not defined yet"*

Answer to comment 2.42:

Thank you for bringing this to our attention. The abbreviation is now introduced earlier and hopefully also better describes the method of Müller–Navarra and Giese (1999).

Lines 227-228:

For the tide gauge at Cuxhaven, Müller–Navarra and Giese (1999) developed an empirical forecast model using multiple linear regression (MLR) of various meteorological input factors […]

Comment 2.43:

"*Line 220: Function 0 is a constant term…"*

Answer to comment 2.43:

Thank for this correction. The number in the manuscript was changed.

Comment 2.44:

"*Table 3: Did I understand that correctly? The methodology of Müller-Navarra and Giese (1999) was used to build AND validate a MLR for Texel. For Cuxhaven only validation was done?"*

Answer to comment 2.44:

The 1995 to 2004 data was used to obtain the regression factors. This was done for Texel as well as Cuxhaven. The 2005 to 2020 data was used to independently verify the regression model. This was done for the tables of the BSH as well to

1. Obtain a base value for the RMSE to compare the results of the regression to
2. Verify the quality of the hindcast using the tables from the BSH

The building and verification of the regression model for Cuxhaven was omitted in the text, because it was finally not used in the determination of the external surges.

Comment 2.45:

"*Figure 3: RMSE instead of RMS-error"*

Answer to comment 2.45:

Thank you, the figure has been corrected.

Comment 2.46:

"Line 247: Which significance level? "0.8 and 0.92" is not listed in Table 4. Please use similar digits."

Answer to comment 2.46:

We agree, that this information is missing. The confidence level has been added to line 289 as well as the description of Tab.4. The precision of the numbers in the text has been adjusted to match the precision of Tab. 4.

Comment 2.47:

"Line 254: and instead of und"

Answer to comment 2.47:

Corrected.

Comment 2.48:

"Figure 4: You do not mention the overestimation of the hindcast especially in b). Did you check model criteria for the MLR? Probably the MLR does not capture all relevant processes?"

Answer to comment 2.48:

Thank you for this question. The performance of the hindcast is addressed in lines 307-311 and possible improvements are discussed in lines 573-587. The few cases were the hindcast overestimates the wind setup are now also addressed. Implementing the discussion points would require additional work, that is unfortunately beyond the scope of this paper. This includes checking for additional relevant processes and the inclusion of known, but less relevant influences. Additionally, nonlinear interaction should be considered in future approaches.

Lines 307-311:

Hindcast and measurement are in ~~overall~~ good agreement, but show the tendency of hindcasts underestimating wind setup, which is more pronounced at Texel tide gauge, with the exception of three hindcasts of high wind setup during low water phases at the Cuxhaven tide gauge. This is probably due to the reduced set of predictors of wind setup. The method was tested for the British tide gauges as well but did not produce sufficient results, probably due to the wrong set of key predictors. Approaches to improve the wind setup for all locations are discussed in Sect. 5.

Comment 2.49:

"*Line 286: I miss these analyses at the point, since a new and improved approach is to be developed? Please explain briefly why this is not done here.*"

Answer to comment 2.49:

The criteria for external surges were adapted from Gönnert (2003). Deviating from the established knowledge about external surges would require extensive research about the propagation of external surges, for example the development of a hydrodynamic model and higher resolution. While we have great interest to conduct such research in the future, it is unfortunately beyond the scope of this paper. At the same time, different sets of criteria would skew the comparison between the existing dataset of Gönnert (2003) and the new dataset. If a longer offset between Aberdeen and Immingham is found to be physically possible, a reanalysis of the 1971-1995 data would be required as well.

Lines 341-344:

This would, however require extensive research about the propagation of external surges, which is beyond the scope of this paper. Using the dataset with a longer allowed offset between Aberdeen and Immingham would, however, invalidate the comparison of data collected by Gönnert (2003) and this study.

Comment 2.50:

"Figure 5: Can you further highlight in the figure what you describe in text?"

Answer to comment 2.50:

Thank you for the suggestion. Fig. 5b) has been altered to highlight to succession of the external surges and the description of Fig 5b) has been adapted.

Figure 5:

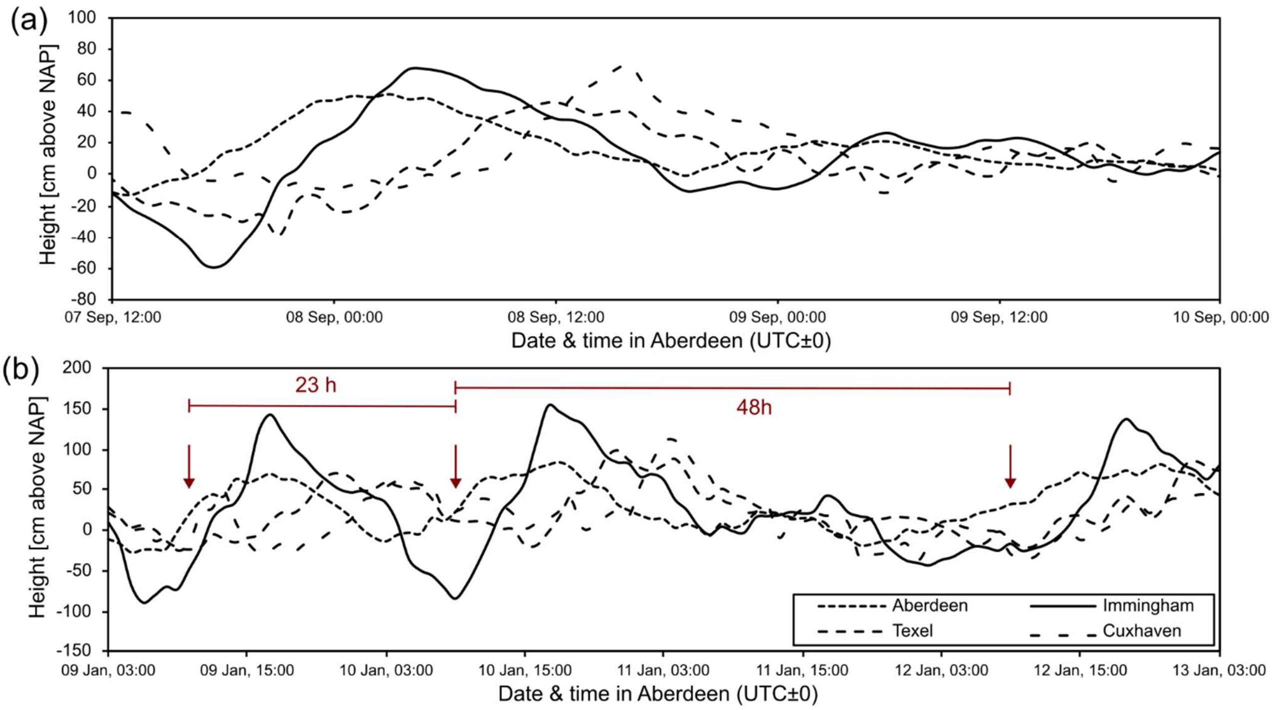

Figure 1: Height of the non–tidal residual (at the gauges in Aberdeen & Immingham) and the residual (at Texel & Cuxhaven) for examples of an external surge on 7 and 8 September 2000 (a) and a serial external surge on 9–13 January 2015 (b). The arrows in (b) mark the beginning of the external surges in Aberdeen (residual >20 cm) and the timing between the serial external surges is given.

Comment 2.51:

"*Line 299: and 50 cm higher at Cuxhaven*"

Answer to comment 2.51:

Thank you for this comment, the word has been added.

Comment 2.52:

"Line 305: Did you check this manually when these are not detected automatically?"

Answer to comment 2.52:

We agree that this distinction was not described clearly. The phenomenon was discovered while the external surges were quality checked for flawed data. This is now highlighted in the text and a is also a part of the discussion.

Lines 394-398:

Some of those external surges were also detected by the automated detection but grouped with other external surges, others do not match the required height or offset between Aberdeen and Immingham but were identified during the manual review of the external surges. Since they occurred shortly before or after a detected external surge and preceding external surges may alter the height and timing of the external surge, they are also included in the dataset. The grouping into external surge events has therefore been conducted manually.

Lines 560-561:

As of now, the automated detection cannot accurately detect all external surge in a series of external surges and differentiate between each external surge.

Comments 2.53, 2.54 & 2.55:

"Line 347: winds"

"Line 361: moving average"

"Line 386: surge = external surge?"

Answer to comments 2.53, 2.54 & 2.55:

Thank you for the corrections and your improvement of the language. The words have been corrected or clarified. Figure 8 and its description were also updated.

Comment 2.56:

"Line 421: significance is always a statistical expression. Did you perform a significance here?"

Answer to comment 2.56:

Thank you for the question. We did not conduct any test. We changed the expressions "significance" and "variance", as we agree, that they convey the impression of statistical treatment.

Line 501-503:

Two other weather situations also show an increase in relevance: BM and SWA. Meanwhile, the share of external surges caused during the weather situations WZ and WA slightly lost relevance. Still, these variations lie within the natural variability of the occurrence of external surges.

Comment 2.57:

"Line 434ff: This is methodology or should be defined earlier."

Answer to comment 2.57:

Thank you for the suggestion to improve the structure of the paper. The section has been moved during reworks according to comments 2.2 and comments 1.17 and 1.18 of Reviewer#1. The section 2.5 was added to the methodology, describing the relevant weather situations.

Comment 2.58:

"Line 464: how is the mean high water calculated? How many years? It is essential due to the rising base water level (MSL)"

Answer to comment 2.58:

Thank you for the question. Clarification has been added to the text. Since a moving average is used, MSL rise is accounted for in this definition.

Lines 527-530:

For this analysis, storm surges are defined as events reaching peaks of at least +1.5 m above mean high water (moving average of the previous 5 years) in Cuxhaven or a wind set up of at least 2 m is reached independent to the tidal phase […]

Comment 2.59:

"Line 488: In the BMBF-project easy-GSH the tide gauge of Heligoland was used to determine the influence of external surges. Is it possible to compare the results with yours?"

Answer to comment 2.59:

Thank you for this interesting comment. To the authors, two main approaches are known to account for external surges in hydrodynamic models of the North Sea: inclusion of the Northeast Atlantic: through a model-cascade including the North Atlantic, and residuals at specific tide gauges as proxies (for example Wick, Aberdeen or in the case of easy-GSH Heligoland), which are included as boundary conditions. Comparing these assumptions with measurement data of external surges is certainly interesting, as it would benefit storm surge modelling as well as sea level forecasts and warning systems. We aim to include this comparison in future research about external surges. However, a direct comparison of model input data and the external surge dataset includes several challenges, for example differentiating between external surges and other phenomena in input data, that cause residuals. Still, the Heligoland tide gauge is one of the most important in the aim to include additional tide gauges in the analysis, as it can provide more information on the effects of external surges on offshore tide gauges.

Comment 2.60:

"Line 494: Please clarify in the methodology which part of the detection is fully automated and which is manually complemented."

Answer to comment 2.60:

Thank you. Please see the answers to comments 2.3 and 2.52 for the additions to the manuscript:

3 Added references

The following references were added during the revisions of the manuscript:

Amin, M. (1982): On analysis and prediction of tides on the west coast of Great Britain. In: *Geophysical Journal International* 68 (1), S. 57–78. DOI: 10.1111/j.1365-246X.1982.tb06962.x.

Boesch, Andreas; Müller-Navarra, Sylvin (2019): Reassessment of long-period constituents for tidal predictions along the German North Sea coast and its tidally influenced rivers. In: *Ocean Sci.* 15 (5), S. 1363–1379. DOI: 10.5194/os-15-1363-2019.

Bruss, Gerd; Gönnert, Babriele; Mayerle, Roberto (2011): Extreme scenarios at the Germen north sea coast - a numerical model study. In: *Int. Conf. Coastal. Eng.* 1 (32), S. 26. DOI: 10.9753/icce.v32.currents.26.

Dronkers, Job; Stojanovic, Tim (2016): Socio-economic Impacts—Coastal Management and Governance. In: Markus Quante und Franciscus Colijn (Hg.): North Sea Region Climate Change Assessment. Cham: Springer International Publishing (Regional Climate Studies), S. 475–488.

Jänicke, Leon; Ebener, Andra; Dangendorf, Sönke; Arns, Arne; Schindelegger, Michael; Niehüser, Sebastian et al. (2021): Assessment of Tidal Range Changes in the North Sea From 1958 to 2014. In: *Journal of geophysical research. Oceans* 126 (1), e2020JC016456. DOI: 10.1029/2020JC016456.

Jensen, Jürgen; Mudersbach, Christoph; Dangendorf, Sönke (2013): Untersuchungen zum Einfluss der Astronomie und des lokalen Windes auf sich verändernde Extremwasserstände in der Deutschen Bucht.

Lotze, Heike K.; Reise, Karsten; Worm, Boris; van Beusekom, Justus; Busch, Mette; Ehlers, Anneli et al. (2005): Human transformations of the Wadden Sea ecosystem through time: a synthesis. In: *Helgol Mar Res* 59 (1), S. 84–95. DOI: 10.1007/s10152-004-0209-z.

Mehra, Prakash; Prabhudesai, R. G.; Joseph, Antony; Kumar, Vijay; Aga, Yogesh; Luis, Ryan et al. (2009): A one year comparison of radar and pressure tide gauge at Goa, west coast of India. In: 2009 International Symposium on Ocean Electronics (SYMPOL 2009). 2009 International Symposium on Ocean Electronics (SYMPOL 2009). Cochin, India, 2009: IEEE, S. 173–183.

Mikhailova, M. V. (2011): Interaction of tides and storm surges at the Elbe River mouth. In: *Water Resour* 38 (3), S. 284–297. DOI: 10.1134/S0097807811030079.

Müller-Navarra, Sylvin (2013): On Tidal Predictions by Mean of Harmonic Representation of Inequalities. Hamburg (Reports of the Federal Maritime and Hydrographic Agency (BSH), 50).

Müller-Navarra, Sylvin; Bork, Ingrid (2011): DEVELOPMENT OF AN OPERATIONAL ELBE TIDAL ESTUARY MODEL. In: *Int. Conf. Coastal. Eng.* 1 (32), S. 48. DOI: 10.9753/icce.v32.management.48.

OSPAR Comission 2000: Quality Status Report 2000. Region II - Greater North Sea. Hg. v. OSPAR Comission. London.

Reise, Karsten; Baptist, Martin; Burbridge, Peter; Dankers, Norbert; Fischer, Ludwig; Flemming, Burghard et al. (2010): The Wadden Sea - A universally outstanding tidal wetland. Wadden Sea Ecosystem No. 29. Commen Wadden Sea Secretariat. Wilhelmshaven.

Serafin, Katherine A.; Ruggiero, Peter; Barnard, Patrick L.; Stockdon, Hilary F. (2019): The influence of shelf bathymetry and beach topography on extreme total water levels: Linking large-scale changes of the wave climate to local coastal hazards. In: *Coastal Engineering* 150, S. 1–17. DOI: 10.1016/j.coastaleng.2019.03.012.

Siefert, Winfried (1991): Ober Eintrittswahrscheinlichkeiten von Windstau, Oberwasser und örtlichem Wind in einem Tidefluß am Beispiel der Elbe. In: *Die Küste* (52), S. 171–190.