Dear Sir/Madam:

We sincerely thank the Referees for their detailed comments that greatly helped us to improve the manuscript. We have given careful consideration to all remarks and have expanded some sections of the paper. Also, we have carefully edited the entire text and removed several minor issues and typos.

The comments of Referees are represented using normal font below and our response and a description of corrections using italics.

Sincerely

Tarmo Soomere, Katri Pindsoo and Maris Eelsalu
03 November 2019

Anonymous Referee #1 (Received and published: 3 June 2019)

The study is aimed to investigate the alongshore variability of the empirical statistical distribution of maximum wave set-up occurrence in a morphologically complicated situation. The area in exam is embedded in the Gulf of Finland, in the eastern Baltic Sea. The selected shoreline has been divided in very small segments, each containing the coastal grid output of a triply nested climatological run of the WAve Model. The maximum set-up has been calculated algebraically from the properties of the wave field at the breaker line, the water depth and the orientation of the shoreline in each segment for each member of the climatology. At each segment, the frequency of occurrence is then plotted against the simulated maximum wave set-up and a quadratic-exponent (three-parameter) law is fitted to the data. In 3/4 of the segments the higher-order coefficient is found equal to zero at the 95% confidence level. In all the other cases, the leading quadratic coefficient is not null at the 95% confidence level, so a Wald (invert Gaussian) distribution is assumed. The method used for the evaluation of the wave set-up is fairly standard and consistent, the statistical analysis is rather on the qualitative side, but the results show a sort of internal coherence.

There are in my opinion several problems in the study that should be addressed in order to improve the quality of the work. Some of the problems, listed as points in the specific comments, would probably require some further analysis on the data, some work on the figures and a general review of the text. Many weaker points are listed in the technical corrections.

Thank you for highlighting the weak points and, in particular, for suggestions to improve the manuscript.

Specific comments

1. In my opinion the main problem is in the results: it was found that 3/4 of the coastal points have an exponential distribution, while all the others have a completely different distribution. But there is no clear indication in the study about the reason behind it. There are some educated hypothesis, but no direct link is provided which can relate the type of distribution with some physical quantity like the angle of approach to the coastline, the wave climate or the bathymetry. This means there is no way to generalise the results outside the area of interest. As a result the work may have a distinct ‘geographical’ interest but is affected by a lack of ‘physical’ significance. Further effort should be devoted to understand the reason.

We agree with this comment but still think that it is particularly important to identify and highlight the situations where the distribution of a certain phenomenon (wave set-up heights in our context) may follow a different distribution than expected, even when the reasons behind the emergence of such a distribution are not fully clear. To our knowledge, there exist no systematic studies of the distribution of set-up heights. In this context, the core result is that this height in the majority of coastal sections follows an exponential distribution. To our understanding, this is a novel result that expands our knowledge of nearshore processes. This is already interesting as the heights of single waves and the time series of significant wave heights usually follow a Rayleigh *or Weibull distribution. The applicability of an exponential distribution for the description of set-up heights seems to be a universal feature as it also becomes evident when we employed time series of offshore measured wave properties for such calculations. The emergence of another distribution in some coastal segments is, to our understanding, a far more interesting feature, for which we have no good explanation as yet. However, as it emerges on many occasions, we think it is important to at least mention its presence and the possible implications of its presence. We have reformulated some sections of the manuscript so that this line of thinking becomes clearer.*

2. The value of the leading coefficient in the quadratic expression in the exponent of the distribution is found remarkably close to zero in all cases, but it is not zero at 95% confidence level in a large fraction of cases. 95% is pretty high but it is not a matter of faith. My impression is that the results would change significantly if different levels were chosen, i.e. 98% or 90%. The dependence of the results (in terms of the number of cases having exponential or else distribution) on an arbitrary choice would show a weakness of the method, indicating a lack of robustness in the statistical analysis.
Yes, the main conclusion of the study is that with quite a high confidence, one can use an exponential distribution to approximate the distribution of set-up heights. The use of the 95% confidence level follows a common line of thinking that this threshold is a good indicator for saying that certain things have some property that is statistically significant. The use of 90% level of statistical significance would lead to an about twofold increase in the number of sections with nonzero coefficient at the quadratic term (in terms of statistical significance at this level) and the use of 98-99% level would render almost all but 6–7 segments following an exponential distribution of set-up heights. This variation is, however, an intrinsic feature of the application of this type of statistical analysis but would still lead to the conclusion that in a few locations the distribution of set-up heights seems to follow an inverse Gaussian distribution.

3. By the way, I would have expected as a first tentative analysis the standard extreme value approach, using a fitting of the empirical CDF by means of plotting position functions. It is less subjective than the method used in the present study and has the added value to introduce a return period, which would be welcome in this case. With 3 parameters at disposal and plotting over a log-log scale, it takes a lot to discard a Weibull distribution. If all cases could be described with a similar distribution then the observation at 1. would be irrelevant.
We agree that this analysis would be interesting and important; however, our goal here is not to look at extreme values. The main point of Section 3 about maximum set-up heights is to compare the results obtained using the simplified approach of Soomere et al. (2013), with those calculated using the approach in this paper. We intentionally explain the differences in terms of maximum set-up heights because the differences in mean set-up heights are much smaller.
The underlying theorems (e.g., Coles, S.: An introduction to statistical modelling of extreme values. Springer, 3rd printing, Springer, London, 208 pp., 2004) state that the distribution of extreme values (e.g., the block maxima) is independent of the particular distribution of the underlying time series (e.g., of set-up heights) under a quite general set of conditions. Therefore, establishing the (parameters of the relevant generalised) extreme value distribution provides no information about the underlying distributions (which we try to quantify). This aspect is explained in the Discussion section. However, we shall definitely perform this kind of exercise in the future.

4. Some of the panels in figures 6,7 show that the higher values of the set-up have the same probability. This is very odd, and led me thinking if there might be some problem with the independence of the data. It looks like the entire block of data would belong to the same storm. In the description of the methodology it should be described in detail how the problem of the serial correlation of the data has been taken care.

The reason for repeating values of probabilities is that the highest values of set-up heights that have the same (very low) probability have occurred only once or twice during the considered time period. As the underlying wind data set has a time step of 3 hours and usually extreme wave storms last no more than 5–6 hours in the Gulf of Finland, it is unlikely that any set of very large set-up values would belong to the same storm. As the entire simulation contains 103 498 single instants of wave properties, one occasion corresponds, theoretically, to 10^{-3}*%. As we exclude the cases with zero set-up heights (e.g., waves propagating offshore), the number of instants of wave properties varies between 40 000 and 70 000 for different coastal segments. Consequently, one occasion in such segments corresponds to 2–3x*10^{-3}*%.*

The level of serial correlation of single wave set-up heights is implicitly minimized by using an approach for rapid reconstruction of the wave climate. It is based on the sequence of wind properties once every 3 hours and contains a minimum amount of "memory" of wave fields. Moreover, in the context of our analysis, the possibility of serial correlation should have no impact on the results as we focus on probability density functions and do not carry out any analysis in which serial correlation may have a role (e.g., sequences of events, block maxima or similar).

5. The wind data gaps are a big problem if they are systematic in the upper percentiles. It should be taken care of in some way, and discussed in the conclusion.

We agree that gaps in wind data are always a problem in reconstructions of marine hydrodynamic fields. As our conclusions rely on the shape of the probability density function for relatively frequently occurring set-up heights (and we even exclude the most infrequent, equivalently the highest, set-up events from some parts of the analysis), it is likely that the impact of gaps in the recordings of the strongest winds have very little impact on our conclusions. We explain this aspect in the revised manuscript.

6. Some figures are very hard to read, in particular figure 4.

It is true that we tried to pack too much information into a small space but we did not find a better way to present this figure (now Fig. 5).

What is the rationale about the choice of the cases illustrated in figure 6,7?

Figure 6 (now 8) is meant to demonstrate how different the shapes of various distributions are, and where the basic difference between Gaussian, Weibull and Wald distributions is. The underlying set-up data are just for illustration (but still represent the most frequent case of exponentially distributed set-up heights).

The panels of Fig. 7 are presented for coastal segments with different orientation. As numbers of coastal segments are hardly visible in Fig. 2, we added a scheme with the location of these segments.

We reshaped and expanded the presentation of the information on Figs. 6, 7 (now 7, 8) at the beginning of Section 3.2 and edited it for clarity so that Figures 6 and 7 (now 7 and 8) are interchanged in the revised manuscript. Also, it is now made clear that the main outcome of the analysis is that generally an exponential distribution matches the empirical distributions of set-up heights.

7. There are some technical points which should be better explained, in particular the presence of 'data gaps in the distribution (the lowest set-up height that did not occur in 1981-2016)' at page 11.

We meant gaps in the calculated empirical distributions, not in the time series of wind or wave properties. It is natural that some specific set-up heights in such empirical distributions (in our case with a resolution of 1 cm) simply do not occur. This happens for very large values of set-up heights that are populated by a few events. We interpret the presence of such a gap as an indication that the number of occasions for the relevant set-up height (and for the higher values of set-up) is too small for use in the estimates of the shape of the entire distribution.

Is the statistical analysis in the range [0.01-0.4] really necessary?

Yes, because our aim is to understand the basic properties (such as the shape) of the probability density function of set-up heights. To avoid possible distortions by the presence of a few very large set-up heights (that may have a completely different distribution as discussed above), we perform in parallel the analysis of the shape of the distribution based only on well-defined values of this distribution. The possible difference in the results is implicitly demonstrated on Fig. 7 (Fig. 6 in the original manuscript) where the inclusion of large set-up values "bends" the approximation towards a concave upwards shape.

Higher than 40 cm set-up events do not occur in some coastal segments. If they occur, the number of such events is usually a few dozen, and only in a few segments exceeds 100. Therefore, such events form less than 0.1% of all set up events but their presence may considerably modify the shape of the approximation to the empirical distribution of set-up heights.

How the angle of incidence with the normal to the coastline was evaluated?
We employed the numerically evaluated wave properties from the standard WAM model. The angle of incidence of waves is evaluated based on the mean approach direction of waves at the centre of the model grid cell and a piecewise linear approximation of the shoreline. This is now described in more detail in the revised manuscript

How the phase and group velocities were estimated?
The phase and group speed at the wave model grid cell were estimated based on the standard expressions of linear theory using the wave number k evaluated from the dispersion relation for linear monochromatic waves $2\pi/T_p = \sqrt{gk \tanh kh_0}$ *with the period equal to the peak period* T_p *and the water depth* h_0 *equal to the model depth for the particular cell. The phase and group speed at the breaker line were estimated from the dispersion relation for long waves:* $c_{gb} = c_{fb} = \sqrt{gh_b}$ *. This is now described in the manuscript.*

8. The language should be improved.
We have asked a colleague-native speaker to edit the manuscript.

Technical corrections

There are many slightly inaccurate statements in the text which should be adjusted:
1 p.1 line 29: actual tides are not perfectly regular in many coastal areas (astronomical tides are).
Thank you; our remark was not necessary, and we deleted it and adjusted the sentence as follows: „A reasonable forecast of the joint impact of tides, …"

1 p.2 line 16: 'neither completely independent nor completely dependent' does not give a lot of information.
This formulation has been replaced by the following: „On most occasions severe coastal flooding occurs under the joint impact of several drivers"

1 p.2 line 19: significant wave heights. And it must be defined somewhere, because there is no definition in the manuscript. Add a symbol like Hm0 if spectral. What is H0? Use it in all the manuscript consistently.
Thank you; we have followed this recommendation and explained after the definition of significant wave height that the rest of the manuscript uses simply „wave height", having in mind significant wave height.

1 p.3 line 16: Normally instruments and model refers to statistical properties of wave
fields: significant wave height, peak period, mean period and mean direction.
Yes, we adjusted the expression accordingly

2.1 p.4 line 25: reference not found.
Thank you, it is added.

2.1 p.5 eqn (2): Here averaged eta is a function, it is customary to indicate the arguments in parentheses.
Yes, η is, technically, a function of four explicit arguments (h, h_b, theta_b, gamma_b) and one implicit argument (a spatial coordinate, say, x) via h(x). As its particular values are not used in the further derivation and η itself is eliminated from the final expression, it seems to us that showing explicitly these arguments would make the equation overly complicated to read, so we only added remark that it depends on the location x along the coastal profile. Also, as the overbar is really not necessary, we omitted it from the formulas.

2.1 p.5 lines 5-6: the meaning of ‘formal’ is unclear, the choice of ‘formal’ and ‘actual’ is not particularly fortunate.
Yes, indeed; we have reformulated these expressions.

In Figure 1: introduce the axis - it is not obvious the sign of h and eta, introduce d and d*.
Thank you. We have replaced the entire figure, inserted both axes, and made clear how the variables are defined. The signs of the variables are explained in the figure caption. The use of H for wave heights and h for depths (thickness of water layer) may be confusing indeed, so we have changed notations so that „d“ is showing the depth.

2.2 p.6 line 8: Wave directions? wind?
An extra „the“ distorted the point.

2.2 p.6 line 9-10: Suggestion: it is possible to analyse the set-up for different values of forcing and wave propagation geometry.
We plan to extend the analysis for the entire Baltic Sea shore with a somewhat lower resolution (perhaps 1-2 nautical miles) to determine how many of the results are locally confined.

2.2 p.6 line 13: Highest significant wave height is sufficient.
Thank you.

2.2 p.6 line 14-16: In meteorology it is customarily is to indicate wind coming from west as westerly, wind going toward west as westward. Eastern storm is unclear.
Thank you, it is now formulated in an unambiguous manner

2.2 p.6 line 28: It is not actually the model implementation and it does not increase the efficiency of the model. It is a simplified method of reproducing the wave climate avoiding to processing all the time series.
Thank you; this is exactly what we are doing, and now we say so.

2.2 p.7 line 6: That is an understatement. The wave simulation depends on wind, if the wind is not adequate the simulation is just noise.
Yes, of course, we deleted it.

2.2 p.7 line 6: ‘In particular..’ actually that is a completely different matter.
Deleted.

2.2 p.7 line 7: Someone might argue that wave directions and propagations in shallow waters and complex morphology depend more on bathymetry than on wind direction.
Yes, we added a short comment on this. In fact, we would like to have as accurate as possible information about both wave directions and heights offshore (at the selected wave model grid cells). We assume that the further propagation of the waves over ~500 m long stretch to the breaker line depends almost entirely on the local bathymetry and geometry, and calculate it as exactly as we can (and ignore wind input over this stretch).

2.2 p.7 line 19: This is a huge problem for the statistical analysis. In my opinion every other choice (interpolation, replacement with model data, looking for other sources of data) would be better than simply not considering the data corresponding to gaps. See point 5.
We agree that this is crucial in all applications where, e.g., the time series, maximum events or worst-case scenarios are addressed. We focus exclusively on the shape of empirical distributions of set-up height. We deliberately avoid consideration of the „tails“ of these distributions that are based on a relatively small number of time instants and could be affected by missing data. Moreover, we actually look at how the shape of these distributions changes along the study area. This shape is mostly governed by small and „intermediate“ values of wave and set-up heights that are all represented by a large number of time instants. In this sense it is likely that our results are quite insensitive with respect to a relatively large number of missing entries in the wind data set. This conjecture is confirmed by a comparison of modelled and measured wave statistics for 2012–2016 at the end of Section 2.2.

2.3 p.7 line 21: Suggestion: in water depth >4 m
Thank you, it makes the expression more transparent

2.3 p.7 line 25: See note 2.2 p.6 line 14-16
Thank you, it is now formulated in an unambiguous manner

2.3 p.7 line 28: Suggestion: oversimplified
Thank you; this is better indeed.

2.3 p.7 line 31: How could significant wave height be monochromatic? ‘As usual’ is not enough to justify the assumption.
We speak about „numerically evaluated wave field for each time instant [that is] is monochromatic“. We reshaped the formulation. Also, we make clear now that the approach & assumptions largely follow those made in (Soomere et al., 2013).

2.3 p.8 line 1: The mean wave direction provided by the model is not referred to the normal to the shoreline. This derivation must have been a successive operation which should be described appropriately.
Thank you for highlighting this gap. We have amended the description: „.The approach direction θ_0 *with respect to the onshore-directed normal to the shoreline is calculated from* θ *based on an approximation of the relevant (about 500 m long) coastal segment by a straight line that follows the average orientation of the shoreline in this segment.“*

2.3 p.8 line 7: If Hb is a water level and H0 is a significant wave height they are different quantities having the dimension of length.
Hb is the wave height at the breaking line. Water level was denoted using lowercase h, so we think the expressions are fine. Still, we have denoted water level by „d“ in the revised manuscript.

2.3 p.8 line 7: Missing reference
Our apologies; it is added to the reference list.

2.3 p.8 line 8: It should be explained how the phase speed and the group velocity were evaluated.
Everything was done using the expressions from standard linear wave theory. The steps are now described in more detail.

2.3 p.8 line 14-15: How the assumptions might affect the results? to be discussed in the conclusion.
It is now discussed in the concluding section.

2.3 p.8 line 14-15: But it is used in the successive section, isn’t it?

As the assumption of ignoring waves that approach the shore at large angles was used in several earlier papers, we think it is sensible to discuss in the body of the paper how strongly and where it may affect the results in terms of maximum set-up heights.

Also, we removed the remark on p.8, lines 15-16 of the original manuscript about the isolation of the worst possible scenarios as this aspect was not addressed in this manuscript.

3.1 p.9 line 5-15: See previous observation. The 'simpler method' is used throughout
the manuscript: it should really be described better.

We replaced the „simpler method" by „S2013" and defined it as „the approach [used in] (Soomere et al., 2013)"

Fig.4 is very hard to read but it gives the impression that the results are very different.

The results are somewhat noisy indeed; however, this unfortunately often happens when (wave) maxima are compared. One point of Fig. 4 is to show that the approach used in earlier studies led to fairly large uncertainties (and possibly incorrect values for some coastal areas) and that the use of the entire range of approaching waves and more adequate approximations (here Eqs. 1–8) will make a clear difference for such areas. Another point is that in some 80% of the segments in question, the maximum values are almost insensitive with respect to the ignoring of waves that approach the shore at angles larger than 10-15 deg.

The text is rather confusing and seemingly incoherent. One is tempted to understand that the set-up is greater for greater angles of incidence than for normal waves.

We substantially reshaped the text so that is hopefully clearer and more consistent now. We also tried to make clear that if the highest set-up heights are created by waves with greater angle of incidence, then this feature reflects the absence of high waves that approach the shore at small angles.

It may be worth to observe that numerical statistical models like WAM are not able to deal with diffraction and reflection of waves.

Thank you, the original text mixed up reflection and refraction.

3.2 p.10 line 7-10: It is not clear how to verify the statement, the analysis is rather qualitative and the figures describe only a very small part of the set of 174 segments.

Suggestion: replication → simulation

Our original formulation was deceptive indeed and has been corrected; see also reply to the next observation.

3.2 p.10 line 18: The discussion seems to exclude the possibility that somewhere in the whole region considered there could be a Poisson process, which is contradicted by the results in p. 11 line 10-15.

Thank you; one of our main points is that for some ¾ of cases an exponential distribution (equivalently, a background Poisson process) serves as an acceptable fit. We added the sentences: „A feasible option for matching the empirical distribution in the majority of the coastal segments is an exponential distribution that reflects a background Poisson process. This distribution is represented by a straight line in log-linear coordinates used in Fig. 6, 7 and also describes, e.g., storm surges in the study area (Soomere et al., 2015)." to make that clear, and deleted some less relevant remarks and repetitions on lines 16-19 of the original manuscript.

3.2 p.10 line 24: Suggestion: approximation → fitting procedure

Thank you, we have changed the wording accordingly

3.2 p.10 line 28: 'Unexpected' does not explain the reason of the high values. On what basis the high values are assumed outliers? Looking at fig. 7, if the range of the setup is considered only in the range [0.01,0.4] maybe the distribution could have been 'forced' to be exponential. This part of the text is not sufficiently clear.

The text was indeed too compact. We have added an explanation as to what is meant and why we think it is reasonable to exclude such values from the fitting procedure: „Thirdly, a few locations have several outliers – a small number of remarkably high set-up events that obviously do not follow the general appearance of the empirical distribution of set-up heights for the particular location (Fig. 6b, d, f). Such events apparently reflect severe storms in which the wind pattern has been favourable for the development of very large waves whereas such waves approach a certain coastal segment under a small angle. The presence of similar outliers is characteristic, for example, for time series of sea level in Estonian waters (Suursaar and Sooäär, 2007) and is associated with situations when strong storms blow from a specific direction. Similarly to outliers in sea level time series, these events likely follow an extreme value distribution (Coles, 2004) and thus should be left out from the analysis of the shape of the distribution of set-up heights."

Anonymous Referee #2 (Received and published: 5 July 2019)

The authors discuss the often less analysed, especially in climate science, but important topic of wave set-up heights and occurrence of these heights along a 60 km long part of the Estonian coastline at the Gulf of Finland. Whereas the overall topic raised by the authors is of interest, the way it is presented make it hard for the reader to understand the overall purpose of this study. In the present form, I can not recommend to accept this manuscript. However, after addressing some major concerns by the authors I would suggest to accept the publication after major revisions.

Thank you for the recommendations towards improving the manuscript. We have carefully worked through and implemented all these recommendations in the revised manuscript.

General comments

The main criticism is that in the current form it is not clear what is the scientific question they wanted to tackle in the first place and what is the main conclusion at the end of this work. Is it to establish some kind of climatology of wave set-up height in that area? Or is it to improve the general understanding of properties of the wave set-up?

Thank you for this. We agree that the formulation of the core question (or hypothesis) of the study and the main conclusion were not perfect. The basic aim is to find out which type of the theoretical distribution describes the probability of occurrence of different wave set-up heights. As set-up may contribute up to 1/3 of the nearshore water level rise during strong storms, this question is of clear interest not only from the theoretical viewpoint but also for practical coastal management.

There is massive research into statistical properties of other drivers of high local water levels and reach of large water waves. The relevant distributions are very different: a Gaussian distribution for the water volume of the Baltic Sea, exponential for storm surges, quasi-Gaussian for water levels at the shores of the Baltic Sea, Weibull for different significant wave heights, and Weibull or Rayleigh for wave run-up. The process of wave set-up should be, to our understanding, no exception. Thus, it is just necessary to find out the shape of its typical probability density function. This is, in essence, the precondition for building an adequate climatology of set-up heights. We describe these arguments in more detail in the revised manuscript.

When we started the analysis of set-up heights, the expectation was that the result would be one of the listed distributions. Surprisingly, a Wald distribution popped up in a non-negligible manner. We have adjusted the Introduction, Discussion and Conclusions to make these aspects more clear.

Can one translate the findings to other coastal regions, at least can one assume some?

This question was also asked by Referee #1. It is possible to some extent. The applicability of an exponential distribution for the description of set-up heights seems to be a universal feature as it becomes evident also when we employed time series of offshore measured wave properties for such calculations. The presence of quasi-Gaussian of Weibull statistics for sheltered coastal areas is apparently a local feature. The emergence of another *distribution in some coastal segments is, to our understanding, a far more interesting feature, for which we have no decent explanation yet. However, as it emerges on many occasions, we think it is important to at least mention its presence and possible implications of its presence.*

What is the reason to find an approximation of the empirical distribution?
The knowledge of the shape and parameters of such a distribution is often crucial in various forecasts and management decisions.

Also, the authors should put more emphasis in the manuscript, as one of the most cited and may be important references Soomere et al. 2013 is missing in the reference list.
We apologise for this shortcoming that is of course improved now. Also, we have revised the entire manuscript for clarity and correct use of English.

Specific comments

If it is understood correctly they have used a single wind observation in favour of simulated wind information as forcing for the wave model. They argued that the simulated wind of atmospheric models are some times inaccurate and have some problems in representing the real wind fields. While this is certainly true, the analysis would benefit to present a short validation of the simulated wave field against some observation in the study area to show that a single forcing site is capable to represent the coastal wave climate reasonable.
We agree with this criticism. To show that the modelled data set reasonably represents the local wave climate, we included into the paper a comparison of buoy measurements in the vicinity of the study area with modelled wave data close to the buoy (end of Section 2.2 plus one new figure). The buoy data are available from 2012 and the comparison is performed for the time interval of 2012–August 2016. The basic properties of wave heights are represented reasonably. The bias is 0.05 m (which we consider as almost perfect, given the distance of the wind measurement site from the study area). The rms difference is 0.5 m. That is about twice as large as for the most recent direct simulations and on the level of this quantity for some other sub-basins of the Baltic Sea; however, it is reasonable for our purposes as we do not need the exact sequence of wave events. The scatter diagram (not shown in the manuscript) is almost perfectly symmetric. The appearance of the relevant empirical probability distributions of the occurrence of different wave height is similar for both data sets.

Which Equation is exactly used for estimating the wave set-up height. They stated Eq (8) is used, but they also introduce the effect of shoaling and refraction is this afterwards used or neglected.
The mathematics of wave set-up is a complicated matter for waves that approach the shore at large incident angles and we probably compressed this material too much. In fact we use a sequence of equations. We start from Eq. (9) on page 8 to evaluate the changes to the wave properties owing to refraction and shoaling on their way from the model grid point to the breaker line. This equation gives us properties of this wave field at the breaker line. Thereafter we employ Eq. (7) to find the set-up height. The derivation of this equation is provided because of intense discussion of the physics of breaking of obliquely approaching waves in Hsu et al. (2006) and Shi and Kirby (2008). We inserted a short version of this scheme into the text at the end of Section 3.2.

Also, two Equations have the same reference number (9).
We apologise for the presence of two equations (9). The second one on page 11 is not really used and is only presented to illustrate the functional form of the probability density function of a Wald distribution. It is renumbered in the revised manuscript.

Some of the figures should also be refined. Figure 2 shows the model domain and the coastal segments with coloured dots that represent the value of a parameter of an estimated approximation. At this point this is too much information for the reader, one does not know what parameter "*a*" stands for. For instance, a colour coding, or a

simple labelling of the number of the coastal segment would be of much more help later on, as in the other figures only the index number of the coastal section is displayed on the x-axis.
Thank you for this observation. We have removed this colour coding from Fig. 2 and created a new figure 8 to highlight the spatial distribution of the values of this coefficient.

It is not clear what is the difference between the left and right side of figure 5, besides more storms are involved on one side. What is the colour coding for?
The left panel of Fig. 5 shows the time when the all-time highest waves occurred in the study area in 1981–2016, whereas the right panel of this figure shows the timing of the all-time highest set-up events. This figure is meant to explain that high waves in the nearshore of an area with complicated geometry does not necessarily mean high wave set-up. The colour coding is necessary to separate storms in single years.

Anonymous Referee #3 (Received and published: 16 July 2019)
The authors present wave set-up calculations based on pre-computed wave model results for a shoreline with complex geometry located in the Gulf of Finland. The wave set-up distributions are based on long time series (1981-2016) and they are calculated for separate segments along the coast. The wave set-up is approximated by accounting also for varying incident angles, and the approximation is compared to a more simple approach. The main finding is that while the wave set-up distribution is well represented by a standard exponential distribution in most segments, in about 25% the distribution does not follow any standard distribution (Gaussian, Weibull, Rayleigh). At these locations the distribution is fairly well approximated by an inverse Gaussian distribution, meaning that the probabilities of the highest wave set-ups are higher with respect to the mean value compared to e.g. a Rayleigh distribution.
The paper is mostly well written and within the scope of Ocean Science. The results are also interesting. I can recommend that this paper is published in Ocean Science after major revisions. I have three main concerns. The first is that this paper is apparently a resubmission, but the results and conclusions have changed since the first paper, and I don't know what to make of this. The second concern is regarding the precalculated wave data. The third is how the distributions are presented and, to a degree, interpreted.
Many thanks for this. We hope to meet all the concerns in the revised manuscript.

<u>Main comment #1</u>: The manuscript seems to be a resubmission of this paper: https://www.earth-syst-dynam-discuss.net/esd-2016-76/. Still, the previous version of the paper presents different results, suggesting that the relevant distribution is inverse Gaussian at all locations instead of just 25%. The papers seem to be based on the same data, so I don't know what has changed. The first version of the paper claims:
"The distribution of wave set-up heights matches a Wald (inverse Gaussian) distribution along the entire study area. Even though different sections of the study area are open to different directions and host substantially different wave regimes, the leading term of the exponent in the associated inverse Gaussian distribution varies insignificantly along the study area and generally is close to –1."
The current version claims:
"Even though different segments of the study area host substantially different wave regimes, the leading term of this polynomial is usually small (between –0.005 and 0.005) and varies insignificantly along the study area. Consequently, the distribution of wave set-up heights substantially deviates from a Rayleigh or Weibull distribution (that usually reflect the distribution of different wave heights). In about ¾ of occasions it is fairly well approximated by a standard exponential distribution. In about 25% of coastal segments it matches a Wald (inverse Gaussian) distribution. "
The authors should make it clear why their results have changed from the previous paper.

The manuscript is a substantially revised version of our paper titled "Inverse Gaussian distribution of wave set-up heights along a shoreline with complicated geometry", originally submitted to Earth Systems Dynamics and available at https://www.earth-syst-dynam-discuss.net/esd-2016-76/.

In our cover letter we informed the editor of Ocean Science about the history of the manuscript.

The Editor of Earth System Dynamics, Dr Anna Rutgersson mentioned in her final comment that „the presentation of the methods and results was not very clear" but encouraged some rewriting and a resubmission (which we understood as a hint that the paper could be better placed elsewhere).

The rejection from Earth Systems Dynamics was, in some way, fortuitous. Namely, during the rewriting and additional check of the results we discovered a bug in the script for the calculation of set-up heights and for the subsequent evaluation of the parameters of their probability density function.

Removing this bug led to the conclusion that typically (in about 75% of coastal segments of the study area) set-up heights follow a classic exponential distribution but in about ¼ of the cases they follow a Wald distribution. In this sense the conclusions have changed significantly and it might be sensible to retract the early version from the website of Earth Systems Dynamics. We are willing to do so if it is acceptable according to the relevant policy of Earth Systems Dynamics. Alternatively, we can discuss this issue in the current paper, with our apologies to the readers of the discussion version in Earth Systems Dynamics. In particular, we have inserted a relevant footnote into the manuscript (at the beginning of Section 3.2), and are willing to follow the recommendations of the editors.

Main comment #2: It is not clear from the paper how the wave data has been obtained or how accurate it is. The main assumption is that the wave data in the nearshore areas reach a steady state quickly, which is why pre-computed maps can be used. I assume the maps are chosen based on the wind direction and the wind speed. Still, has the runs themselves been modelled using wind data from a numerical weather prediction model? If not, how have the Baltic Sea wide simulations been produced that provide the boundary data? Data from atmospheric models have been used by several studies in the Baltic Sea (including nearshore areas) and it is the industry standard. There is no reason to not produce the maps using proper wind data.

The details of the scheme for rapid reconstruction of wave properties are extensively described in Soomere (2005). This publication is open to everybody and we think that it is not appropriate to repeat the details in this manuscript. We understand that the scheme is not perfect. In particular, it assumes that the wind field is homogeneous all over the Baltic Sea. This assumption is the weakest aspect of the use of one-point wind data, either modelled or measured. Even though this assumption is usually not valid, the geometry of the Gulf of Finland and the location of the study area is such that for most of wind directions the waves that are generated in the Baltic Proper at latitudes to the south of the Gulf of Finland simply do not reach the study area. In this sense the results of calculations of the underlying maps are invariant with respect to the source of wind data.

Are you using WAM 4.5.1 or a later version? Cycle 4.5.1 had a bug in the depth-induced wave breaking source term that lead to unphysical high values at certain points. It has been fixed in cycle 4.5.4. This must be ruled out as an explanation for your surprising results.

The maps were calculated using an early version of WAM cycle 4. As wave heights are not very large and both the wave periods and the propagation distance over shallow areas with possible depth-induced breaking are fairly small (usually no more than 1-2 grid cells), it is unlikely that bugs of this kind have any identifiable impact on the results.

Using Kalbådagrund data to force nearshore areas will overestimate the wind speed and probably result in an inaccurate wind direction. The authors cite some of their previous work, but I think some more details are needed also in this paper.

We agree with all this; however, it is still the case that atmospheric models have difficulties with reproduction of strong wind directions for the Gulf of Finland. We provided more detailed arguments in the published response to the Referee and still think that the measured data more adequately represent the wind direction than the modelled data near the study area, at least during strong wind events that govern the wave heights larger than the mean wave height.

The validity of the modelling approach comes into question exactly because the authors present surprising results. Running a high-resolution model for 35 years is clearly a huge task. Creating the maps with a proper forcing is not. I think the Estonian Weather Service runs HIRLAM, which has been tried and tested in the GoF. A minimum requirement would be that the authors should identify the times that are responsible for creating the "surprising" high wave set-ups and validate the instances by running the "full model". This will satisfy that, to the degree we can trust wave models, the surprising results are real.
Thank you for this. We put considerable effort into retrieving locally produced high-resolution HIRLAM wind fields but this turned out to be a complicated problem. As the time foreseen for the revision approached rapidly, we chose another approach. To demonstrate the quality of the modelling approach for rapid reconstruction of local wave properties, we provide in the revised version a fairly large (under the circumstances) comparison of measured and reconstructed wave data at the end of Section 2.2. This is also done to address the comment by Referee #2.

The comparison is performed for the time interval of 2012–August 2016 for which the buoy data are available in the vicinity of the study area. The basic properties of wave heights are represented reasonably. The bias is 0.05 m (which we consider as almost perfect, given the distance of the wind measurement site from the study area and the simplicity of the calculation scheme). The rms difference is 0.5 m that is about twice as large as for the most recent direct simulations. A large part of this difference is because of mismatch of timing of modelled and measured wave events. This value is on the level of this quantity for some other sub-basins of the Baltic Sea (Björkqvist et al., 2018). We think that it is reasonable for our purposes as we do not need the exact sequence or timing of wave events. The scatter diagram (not shown in the manuscript) is almost perfectly symmetric. The appearance of the relevant empirical probability distributions of the occurrence of different wave heights is similar for both data sets.

Main comment #3: Usually the distributions are viewed and fitted by looking at the cumulative function, not the probability density function. The cumulative function is more stable, and I don't see any reason not to use it here.
We agree with this; however, the use of a cumulative density function on many occasions smooths out differences in the shapes of the empirical and theoretical distributions. This feature is particularly relevant in our case as for large arguments the Wald distribution becomes similar to a Gaussian distribution. Our intention is to demonstrate that a large number (or proportion) of the empirical distributions follow at least locally a clearly different law than the classic distributions. This difference is better noticeable in terms of probability density functions.

The comments regarding "gaps" in the distribution are also confusing. What is the physical meaning of these gaps? The "first gap" is controlled by the accuracy of the wave data (10 cm, 1 cm etc), so why is this a meaningful point? There is also no reason to expect that experimental/modelled data should be without these gaps, so the existence of one is not a defining point in the distribution.
We meant gaps in the calculated empirical distributions, not in the time series of wind or wave properties. It is natural that some specific set-up heights in such empirical distributions (in our case with a resolution of 1 cm) simply do not occur. This happens for very large values of set-up heights that are populated by a few events. We interpret the presence of such a gap as an indication that the number of occasions for the relevant set-up height

(and for the higher values of set-up) is too small for use in the estimates of the shape of the entire distribution. We explain this aspect more clearly in the revised manuscript.

I am also a bit worried about the highest points being correlated. Is this the case? Then one event that is overestimated might "destroy" the entire distribution. Are the highest values recorded during the time ice was present? This needs to be evaluated carefully.
We generally try not to include the highest values of wave set-up into estimates of the shape of the probability distributions for the same reasons. [Even though a comparison of measured and modelled data indicates that the ignoring of ice has not contributed to the „surprising" results.] This is why we look at the data points until the first „gap" in the distribution or limit the set-up heights to 0.4 m. Another reason for systematically ignoring the largest set-up heights is that they may represent another population of realisations with a completely different distribution (Coles, 2004). We plan to perform the analysis of extremes of set-up heights in the future using the technique of extreme value distributions.

If you have found that the distribution of the wave set-up is really completely different for two neighbouring grid points about 500 m apart, then also give an example of the detailed shoreline structure that leads to such a dramatic variation so that the reader can be convinced what is going on and find some physical basis for this change in behaviour. Providing the structure of the shoreline that leads to this inverse-Gaussian behaviour would make this study highly useful for other coastal areas.
The relevant comments are added to the description of the results. While some segments with non-Poisson features correspond to areas of rapid change in the orientation of the coastline, some other become evident along almost featureless shore segments. In general, a few segments having Gaussian or Weibull statistics are indeed exceptional whereas in the vicinity of segments having statistics that resemble an inverse Gaussian distribution the changes in the governing parameter „a" are fairly smooth.

It is unclear how the second order polynomial fitting is done. It seems like you are fitting a polynomial to the data in the log-linear plot (is the variable p the probability density?). Please write the formula for the exponential distribution and how the polynomial fits into that, and for what fit this generalized exponential function becomes the "normal" exponential function. This will make it so much easier to follow the discussion and agree with the conclusions.
This is a good recommendation. We have added the relevant remarks into Section 3.2, removed a typo in the description of the variables in the polynomial (p=probability was wrong; it is replaced by z=set-up height).

I will save minor comments for the revised manuscript.
We have carefully revised the entire manuscript and hope that the number of smaller issues is minimised.

Variability of distributions of wave set-up heights along a shoreline with complicated geometry

Tarmo Soomere[1,2], Katri Pindsoo[1], Maris Eelsalu[1]

[1]Laboratory of Wave Engineering, Department of Cybernetics, School of Science, Tallinn University of Technology, Akadeemia tee 21, Tallinn, 12618, Estonia
[2]Estonian Academy of Sciences, Kohtu 6, Tallinn, 10130, Estonia

Correspondence to: Katri Pindsoo (katri.pindsoo@ioc.ee)

Abstract. The phenomenon of wave set-up may substantially contribute to the formation of devastating coastal flooding in certain coastal areas. We study the appearance and properties of empirical probability density distributions of the occurrence of different set-up heights on an approximately 80 km section of coastline near Tallinn in the Gulf of Finland, the eastern Baltic Sea. The study area is often attacked by high waves propagating from various directions and the typical approach angle of high waves varies considerably along the shore. The distributions in question are approximated by an exponential distribution with a quadratic polynomial as the exponent. Even though different segments of the study area have substantially different wave regimes, the leading term of this polynomial is usually small (between –0.005 and 0.005) and varies insignificantly along the study area. Consequently, the distribution of wave set-up heights substantially deviates from a Rayleigh or Weibull distribution (that usually reflect the distribution of different wave heights). On about ¾ of the occasions it is fairly well approximated by a standard exponential distribution. In about 25% of the coastal segments it qualitatively matches a Wald (inverse Gaussian) distribution. This property signals that very high extreme set-up events may, in some locations, occur substantially more frequently than is expected from the probability of occurrence of severe seas.

1 Introduction

Global sea level rise in most of existing projections of climate change (Cazenave et al., 2014) is often associated with major consequences for the coastal zone (Hallegatte et al., 2013). The resulting economic damages to low-lying coastal areas (Darwin and Tol, 2001) may lead to a significant loss of worldwide welfare by the end of this century (Pycroft et al., 2016). Global sea level rise, however, contributes only a small fraction to the most devastating coastal floods. These events, in addition to being economically extremely damaging (Meyer et al., 2013), may also lead to massive loss of life and destruction of entire coastal communities (Dube et al., 2009).

A devastating flood is usually caused by the interplay of several drivers, with fundamentally different predictability, and physical, dynamic and statistical properties. A reasonable forecast of the joint impact of tides, low atmospheric pressure (inverted barometric effect), wind-driven surge and wave-induced effects requires a cluster of dedicated atmospheric, ocean circulation and wave models. The resulting high water levels may be additionally amplified by specific mechanisms and events such as tide–surge interactions (Batstone et al., 2013; Olbert et al., 2013), meteorologically driven long waves

Deleted: in
Deleted: about
Deleted: long
Deleted: largely
Deleted: is
Deleted: host
Deleted: In
Deleted: it could be
Deleted: The g
Deleted: to
Deleted: by almost 2%
Deleted: The g
Deleted: in
Deleted: ing
Deleted: additionally
Deleted: es
Deleted: lives
Deleted: er
Deleted: ifica
Deleted: ing
Deleted: al
Deleted: While tides are almost perfec
regular and caused by extra-terrestrial
drivers, a

(Pattiarachi and Wijeratne, 2014; Pellikka et al., 2014; Vilibic et al., 2014) or seiches (Vilibic, 2006; Kulikov and Medvedev, 2013). In addition, wave-driven effects at the waterline such as wave set-up and run-up (Stockdon et al., 2006) may greatly contribute to the damaging potential of extreme water levels. These phenomena are driven by momentum carried by waves and have different time scales and appearance. When a wave crest reaches the shore, the resulting temporary inland movement of the water, with a time scale comparable with the wave period, is termed run-up (see, e.g., Didenkulova, 2009 for an overview and references). In contrast, wave set-up is the increase in the mean water level due to the release of momentum of breaking waves.

Along with contemporary numerical simulations and a direct search for worst-case scenarios (e.g., Averkiev and Klevanny, 2010), the use of the probabilistic approach is another classic way to quantify the properties of extreme water levels and related risks. The relevant pool of literature contains substantial amounts of work on statistical parameters of water level variations (e.g., Serafin and Ruggiero, 2014; Fawcett and Walshaw, 2016), and extreme water levels and their return periods (e.g., Purvis et al., 2008; Haigh et al., 2010; Arns et al., 2013). Similar probabilistic analysis has been extensively applied to the average and extreme wave properties (e.g., Orimolade et al., 2016; Rueda et al., 2016), wave-driven effects at the waterline (Holland and Holman, 1993; Stockdon et al., 2006), and properties of meteotsunamis (Geist et al., 2014, Bechle et al., 2015). On most occasions severe coastal flooding occurs under the joint impact of several drivers. This feature generates the necessity to consider multivariate distributions of their properties. Most often, the simultaneous occurrence of storm surges and large waves is addressed (e.g., Hawkes et al., 2002; Wadey et al., 2015; Rueda et al., 2016b). A few studies also include an analysis of joint distributions of significant wave heights, periods and directions (Masina et al., 2015).

Typical probability distributions of different constituents of extreme water levels may be fundamentally different. The distribution of observed and numerically simulated water levels is usually close to Gaussian (Bortot et al., 2000; Johansson et al., 2001; Mel and Lionello, 2014; Soomere et al., 2015). The total water level in semi-sheltered seas with extensive subtidal or weekly-scale variability may contain two components. In the Baltic Sea, one of these (that reflects the water volume of the entire sea) has a classic quasi-Gaussian distribution whereas the other component (that reflects the local storm surge) has an exponential distribution and apparently mirrors a Poisson process (Soomere et al., 2015) similar to the non-tidal residual in the North Sea (Schmitt et al., 2018). The probabilities of occurrence of different single wave heights are at best approximated either by a Rayleigh (Longuet-Higgins, 1952), Weibull (Forristall, 1978) or Tayfun distribution (Socquet-Juglard et al., 2005). The probability distribution of run-up heights usually follows the relevant distribution for incident wave heights (Denissenko et al., 2011) but can be approximated by a Rayleigh distribution even if the approaching wave field does not represent a Gaussian process (Denissenko et al., 2013). The empirical probabilities of average or significant wave heights in various offshore conditions usually resemble either a Rayleigh or a Weibull distribution (Muraleedharan et al., 2007; Feng et al., 2014) while Pareto-type distributions are more suitable for the analysis of meteotsunami heights (Bechle et al., 2015).

In this paper we focus on the appearance and properties of empirical distributions of wave-driven local water level set-up. This process, called set-up in the following, is a classic phenomenon on open ocean coasts. It may often provide as much as 1/3 of the total water level rise during a storm (Dean and Bender, 2006) and significantly contribute to extreme sea level

Deleted: On top of that

Deleted: . They

Deleted: the

Deleted: In

Deleted: the drivers of

Deleted: are neither completely independent nor completely dependent o each other

Deleted: Importantly, t

Deleted: a

Deleted: one

Deleted: conditions

Deleted: ly

Deleted: a

Deleted: what follows

Deleted: at

events (Hoeke et al., 2013; Melet et al., 2016, 2018). The impact of this phenomenon *inter alia* contributes to the overall level of danger in the coastal zone because, for example, the baseline level of wave run-up includes the local elevation of water level owing to set-up. The physics of set-up has been known for half a century (Longuet-Higgins and Stewart, 1964). Adequate parameterizations of this phenomenon have been introduced more than a decade ago (Stockdon et al., 2006) and many models take it into account to a certain extent (SWAN, 2007; Roland et al., 2009; Alari and Kõuts, 2012; Moghimi et al., 2013).

The contribution from wave set-up still provides one of the largest challenges in the modelling of storm surges and flooding (Dukhovskoy and Morey, 2011; Melet et al., 2013). This reflects the intrinsically complicated nature of its formation. First of all, the set-up height strongly depends on the approach angle of waves at the breaker line. This angle is well-defined only if the coastline is almost straight, the nearshore is mostly homogeneous in the alongshore direction and the wave field is close to monochromatic (Larson et al., 2010; Viška and Soomere, 2013; Lopez-Ruiz et al., 2014; 2015). Generally, this angle is a complicated function of shoreline geometry, nearshore bathymetry, wave properties and instantaneous water level. Even if the basic statistical properties of wave fields (usually given in terms of significant wave height, mean or peak period, and mean propagation direction) are perfectly forecast or hindcast in a nearshore location, the evaluation of the further propagation of waves is a major challenge because, for example, refraction properties and the location of the breaking line change with the local water level.

Several studies have focused on the maxima of set-up heights over certain coastal areas (Soomere et al., 2013; O'Grady et al., 2015) or the maximum contribution from set-up to the local water level extremes (Pindsoo and Soomere, 2016). The problem of evaluation of maximum set-up heights has a relatively simple solution on comparatively straight open ocean coasts. The nearshore of such coasts is usually fairly homogeneous in the alongshore direction and the highest waves tend to approach the shore under relatively small angles. These features make it possible to use simplified schemes for the evaluation of the joint impact of refraction and shoaling in the nearshore (e.g., Larson et al., 2010). On many occasions it is acceptable to assume that waves propagate directly onshore (O'Grady et al., 2015) or to reduce the problem to an evaluation of the properties of the highest waves that approach the shore from a relatively narrow range of directions (Soomere et al., 2013). In areas with complicated geometry and especially in coastal segments where high waves may often approach at large angles it is necessary to take into account full refraction and shoaling in the nearshore (Viška and Soomere, 2013; Pindsoo and Soomere, 2015).

Even though high storm surges are often associated with severe seas, the formation of high set-up depends on many details of the storms and the impacted nearshore. It does not necessarily exhibit its maximum level in the coastal sections that are affected by the highest waves. The maximum storm surge and maximum set-up usually do not occur simultaneously (Pindsoo and Soomere, 2015). On the contrary, in coastal areas with complicated geometry each short segment may have its own 'perfect storm' that creates the all-highest sum of storm surge and set-up (Soomere et al., 2013). These observations call for further analysis of the properties of the set-up phenomenon.

Deleted: is

Deleted: this phenomenon

Deleted: feature

Deleted: the

Deleted: of this phenomenon

Deleted: actual

Deleted: a

Deleted: one

Deleted: (possibly increased

Deleted:)

Deleted: wave

Deleted: e.g.

Deleted: along

Deleted: . Moreover, accurate wave hindcast and forecast are still a challeng many water bodies because of insufficie quality of the wind information (Nikolki et al., 2014)

Deleted: In

Deleted: under

Deleted: to the shore normal

Deleted: In

As described above, research into the statistical properties of the main drivers of high local water levels and the reach of swash generated by large waves that attack the shore have revealed that the relevant distributions of the magnitude of these drivers are very different. They may include a Gaussian distribution for the water volume of the Baltic Sea (Soomere et al., 2015), an exponential distribution for storm surges (Schmitt et al., 2018), a quasi-Gaussian distribution for water levels at the shores of the Baltic Sea (Johansson et al., 2011), a Weibull distribution for different significant wave heights (Feng et al., 2014), and a Weibull or Rayleigh distribution for wave run-up heights (Denissenko et al., 2013). The knowledge of the shape and parameters of such distributions is often crucial in various forecasts and management decisions.

In this paper we address the basic features of statistical distributions of set-up heights along an approximately 80 km coastal section in the vicinity of Tallinn Bay in the Gulf of Finland, the Baltic Sea. The shoreline of the study area has a complicated geometry and contains segments with greatly different orientations. The goal is to identify the typical shapes of the distributions of the probability of occurrence of simulated wave set-up heights and to analyse the alongshore variability of these distributions.

The paper is organised as follows. Section 2 introduces the method of evaluation of the maximum set-up height for obliquely approaching waves. It also provides a short overview of the simplified method for rapid reconstruction of long-term wave climate, the forcing data for the underlying wave model and the procedure of evaluation of properties of breaking waves based on the output of the wave model. Section 3 presents an analysis of spatial variations in the appearance of the empirical distribution of wave set-up heights in the study area. The core result is an estimate of the typical shape of empirical probability distributions of different set-up heights along the coast. A part of these distributions substantially deviate from the listed distributions and exhibit an unexpectedly large proportion of high set-up events compared to the classic Gaussian, Rayleigh or Weibull statistics. Several implications of the results are discussed in Section 4.

2 Methods and data

2.1 Set-up height for obliquely incident waves

The classic concept of wave set-up (Longuet-Higgins and Stewart, 1962) relates the local increase in the water level with the release of the onshore component of radiation stress in the process of wave breaking. Based on this concept, it has been demonstrated that, in ideal conditions, the maximum set-up height $\eta_{\max}$ (with respect to the still water level) created by a train of monochromatic waves with a constant height propagating directly onshore along a planar impermeable beach is (McDougal and Hudspeth, 1983; Hsu et al., 2006)

$$\eta_{\max} = d_b \frac{40\gamma_b^2 - 3\gamma_b^4}{128} \approx \frac{5}{16}\gamma_b H_b , \qquad (1)$$

Deleted: intense

Deleted: in

Deleted: water

Deleted: one

Deleted: about

Deleted: long

Deleted: this test

Deleted: that are open to fairly

Deleted: directions

Deleted: wave model

Deleted: evaluation of wave time serie

Deleted: is

Deleted: of

Deleted: extreme

Deleted: along

Deleted: T

Deleted: similar

Deleted: of different wave conditions

Deleted: in comparison with

where $\gamma_b = H_b/d_b$ is the breaking index that is assumed to be constant all over the surf zone, d_b is the still water depth at the breaker line and H_b is the wave height at the breaker line (Fig. 1). Expression (1) is used in many engineering applications (Dean and Dalrymple, 1991) and studies into the properties of set-up of waves that approach the shore at a relatively small angle (see Soomere et al., 2013 and references therein).

If waves approach under a non-negligible angle θ with respect to the shore-normal, the situation is much more complicated. Shi and Kirby (2008) argue that the water level set-down at the breaker line is invariant with respect to the approach angle of waves. The average deviation η of the sea surface from the still water level within the surf zone of an impermeable planar beach is (Hsu et al., 2006; Shi and Kirby, 2008; the power of γ_b in the first term at the right-hand side of their expression being corrected):

$$\eta = \frac{\gamma_b^2 \sin^2 \theta_b}{2h_b\left(8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b\right)}\left(d^2-d_b^2\right) - \frac{3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}{8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}\left(d-d_b\right) - \frac{\gamma_b^2}{16} d_b . \qquad (2)$$

The last term at the right-hand side of Eq. (2) represents the water level set-down η_b at the breaker line and θ_b is the wave approach direction at breaking. Here $d = d(x)$ represents the water depth counted from the still water level at a particular distance x from the shoreline and η is a function of x. The maximum wave set-up $\eta_{\max}$ occurs somewhere inland where $d_{\max}$ is negative and the thickness of the water sheet $d^* = d + \eta_{\max} = 0$ and thus $\eta_{\max} = -d_{\max}$. For this location, Eq. (2) reduces to:

$$\eta_{\max} + d_{\max} = \frac{\gamma_b^2 \sin^2 \theta_b}{2h_b\left(8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b\right)}\left(d_{\max}^2-d_b^2\right) - \frac{3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}{8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}\left(d_{\max}-d_b\right) - \frac{\gamma_b^2}{16} d_b + d_{\max} = 0 . \qquad (3)$$

For shore-normal waves $\theta_b = 0$ and Eq. (3) reduces to a linear equation:

$$-\frac{3\gamma_b^2}{8+3\gamma_b^2}\left(d_{\max}-d_b\right) - \frac{\gamma_b^2}{16} d_b + d_{\max} = 0 . \qquad (4)$$

In this case the maximum set-up height $\eta_{\max}$ is defined by Eq. (1).

For obliquely approaching waves Eq. (3) is a quadratic equation with respect to $q = d_{\max}/d_b$:

$$\frac{\gamma_b^2 \sin^2 \theta_b}{2\left(8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b\right)}\left(q^2-1\right) - \frac{3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}{8+3\gamma_b^2-2\gamma_b^2 \sin^2 \theta_b}\left(q-1\right) - \frac{\gamma_b^2}{16} + q = 0 . \qquad (5)$$

This equation can be rewritten as

$$8q^2\gamma_b^2 \sin^2 \theta_b + 128q + \gamma_b^2\left(40-3\gamma_b^2\right) - \gamma_b^2 \sin^2 \theta_b\left(40-2\gamma_b^2\right) = 0 . \qquad (6)$$

Deleted: under

Deleted:

Deleted: *h*

Deleted: formal

Deleted: in the

Deleted: the formal water depth

Deleted: actual

Deleted: depth

Equation (6) has two negative solutions for physically reasonable values of γ_b. The physically relevant solution to Eq. (6) must be bounded and should be almost equal to $q \approx -5\gamma_b^2/16$ for very small approach angles $\theta_b \approx 0$. Therefore, the expression

$$q_1 = \frac{-32 + \sqrt{1024 - 2\gamma_b^4 \sin^2\theta_b \left[40 - 3\gamma_b^2 - \sin^2\theta_b \left(40 - 2\gamma_b^2\right)\right]}}{4\gamma_b^2 \sin^2\theta_b} \qquad (7)$$

provides the desired solution. Equation (7) deviates from expression (30) of Hsu et al. (2006) by reasons discussed by Shi and Kirby (2008). The maximum set-up height for obliquely approaching waves is thus

$$\eta_{\max} = -q_1 d_b = -\frac{q_1 H_b}{\gamma_b}. \qquad (8)$$

2.2 Wave time series in the nearshore of the study area

We evaluate the shape and parameters of the empirical probability distribution set-up heights along an approximately 80 km coastal segment of Tallinn Bay and Muuga Bay (Fig. 2). The study area is an example of a wave-dominated micro-tidal region. The shoreline is locally almost straight for scales up to a kilometre or two. Several relatively straight parts along the Suurupi Peninsula (grid points 1–10 in Fig. 2) and the area of Saviranna (grid points 137–143 in Fig. 2) are open to the north. However, at larger scales (from a few kilometres) the coast contains large peninsulas and bays deeply cut into the mainland. The shores of these landforms are open to different directions and have greatly different wave regimes (Soomere, 2005). As the formation of set-up crucially depends on the wave height and direction (or approach angle), this type of coastal landscape makes it possible to analyse the wave set-up distribution for coastal sections with radically different wave climates, and also the associated magnitudes of set-up (Soomere et al., 2013).

The fetch length in the Gulf of Finland is >200 km for westerly and easterly winds but <100 km for all other wind directions. The highest significant wave height (5.2 m) in the Gulf of Finland has been recorded twice in a location just a few tens of km to the north of the study area (Tuomi et al., 2011). The strong winds in this region blow predominantly from the south-west and north-north-west. Easterly storms are less frequent but may generate waves as high as those generated by westerly storms (Soomere et al., 2008). Strong storms with winds from the north-north-west may generate significant wave heights >4 m in the interior of Tallinn Bay (Soomere, 2005). The varying mutual orientation of high winds, propagation direction of waves and single shoreline segments makes it possible to identify potential alongshore variations in the distributions of set-up heights.

We employ time series of wave properties (significant wave height, wave period and propagation direction) reconstructed using the wave model WAM cycle 4 and one-point high-quality wind information from the vicinity of the study area. The wave model is implemented in a triple nested version with the resolution of the innermost grid about 470 m (Soomere, 2005). The study area is divided into 174 coastal segments with a length of about 500 m (Fig. 2). Each segment corresponds

Deleted: the

Deleted: statistical

Deleted: about

Deleted: long

Deleted: of this section of the norther
shore of Estonia in the Gulf of Finland,
north-eastern Baltic Sea,

Deleted: on

Deleted: host

Deleted: the

Deleted: in question

Deleted: severity of the

Deleted: n

Deleted: n

Deleted: all-time

Deleted: predominant

Deleted: directions

Deleted: are

Deleted: n

Deleted: as high

Deleted: as the

Deleted: n

Deleted: tern storms

to a nearshore wave model grid cell. Experience with this model in the Baltic Sea and Finnish archipelago indicates that it is important to adequately represent the wave growth in low wind and short fetch conditions (Tuomi et al., 2011; 2012). To meet this requirement, the model uses an increased frequency range of waves up to 2.08 Hz. Ignoring the presence of sea ice may lead to a certain overestimation of the overall wave energy in the region but does not significantly distort the shape of the probability distribution of different wave heights (Fig. 3). It is therefore likely that the shape of distributions of set-up heights and the variation in these distributions along the shoreline are also reconstructed adequately.

We employ a simplified method for rapid reconstruction of long-term wave climate. The computations are speeded up by replacing calculations of the time series of sea state by an analysis of precomputed maps of wave properties. This simplification avoids reconstruction of all the wave time series and relies on a favourable feature of the local wave regime, namely that wave fields rapidly become saturated and have relatively short memory in the study area (Soomere, 2005). Consequently, a reasonable reproduction of wave statistics is possible by the assumption that an instant wave field in Tallinn Bay is a function of a short time period of wind dynamics. This assumption justifies splitting the calculations of time series of wave properties into independent sections with duration of 3–12 hours. The details of the model set-up, bathymetry used, and the implementation and validation of the outcome have been repeatedly discussed in the literature (Soomere, 2005; Soomere et al., 2013).

The described approach makes it possible to circumvent one of the major issues of replication of Baltic Sea wave fields, namely, the frequent inconsistency of different modelled wind data sets (Nikolkina et al., 2014). Similar to, for example, wave-driven sediment transport, wave set-up is intrinsically sensitive with respect to the wave propagation direction. As the nearshore wave directions in areas with complex geometry and bathymetry may be greatly impacted by local features, it is crucial to properly reconstruct the offshore wave directions. This is only possible if the wave model has correct information about wind directions. This is an issue in the Gulf of Finland where atmospheric models often fail to reproduce wind directions (Keevallik and Soomere, 2010). To overcome this issue, we use wind data from an offshore location in the central part of this gulf. The wind recordings at Kalbådagrund (59°59′ N, 25°36′ E, a caisson lighthouse located on an offshore shoal) are known to represent marine wind properties exceptionally well (Soomere et al., 2008). Even though this site is located at a distance of about 60 km from the study area, it is expected to correctly record wind properties in the offshore that govern the generation of surface waves in the open sea.

Wind properties at Kalbådagrund were recorded starting from 1981 once every 3 hours for more than two decades, but since then they have been filed at a higher time resolution. To ensure that the forcing data is homogeneous, we downsampled the newer higher-resolution recordings by selecting the data entries once in 3 hours. The entire simulation interval 1981–2016 contained 103 498 wind measurement instants with a time step of 3 h. In about 9000 cases (less than 10% of the entire set) either wind speed or direction was missing. These time instants were excluded from the further analysis. As some of these instants involved quite strong winds, our analysis may underestimate the highest wave set-up events for some segments of the shore. However, as we are interested in the statistical properties of most frequently occurring set-up heights and in the alongshore variations of these properties, it is likely that omitting these data does not substantially impact the results.

Deleted: The e

Deleted: The i

Deleted: of

Deleted: apparently

Deleted: The model is implemented u

Deleted: scheme that is designed

Deleted: s

Deleted: time series

Deleted: Wave

Deleted: long-term

Deleted: .

Deleted: N

Deleted: ,

Deleted: section

Deleted: the

Deleted: the

Deleted: of

Deleted: a number of short

Deleted: used

Deleted: ,

Deleted: the

Deleted: ly

Deleted: e.g.,

Deleted: The quality of wave hindcast primarily depends on the adequacy of th wind information. In particular

Deleted: I

Deleted: therefore

Deleted: force

Deleted: with

Deleted: the top of

Deleted: impeccably

Deleted: are mostly responsible for

Deleted: in

Deleted: As this resolution of wind measurements was employed for more t two decades,

Deleted: t

Deleted: in

To roughly estimate the adequacy of the described method for speeding up the estimates of the properties of wave climate via rapid reconstruction of time series of approximate wave properties, its outcome is compared with the results of wave measurements made by Marine Systems Institute, Tallinn University of Technology, at Tallinnamadal (Fig. 2, 59°42.723′ N, 24°43.890′ E). As the results of buoy measurements are available starting from 2012, the comparison is performed for the time interval of 2012–August 2016. The measured wave properties are compared with the modelled properties at the closest grid point of the sea area represented at 470 m resolution at 59°41′ N, 24°45′ E. The buoy is located about 3 km from the border of this area and the distance between the buoy and the centre of the closest grid cell is 3.34 km. The comparison (Fig. 3) only includes the instants when both measured (green) and modelled (red) wave parameters were available. We use for comparison the significant wave height H (often denoted as H_S and below called wave height) defined as $H=\sqrt{m_0}$, where m_0 is the zero-order moment of the one-dimensional wave spectrum.

The basic properties of wave heights such as the maximum (measured 5.58 m, modelled 5.77 m), mean (measured 0.643 m, modelled 0.697 m) and median (measured 0.40 m, modelled 0.54 m) are represented reasonably. The bias of the model (about 0.05 m) is at the same level as the typical bias for modelled wave properties in the Baltic Sea in the most recent simulations (Björkqvist et al., 2018). As our study basically relies on statistical properties of wave fields (the probability of occurrence of seas with different significant wave height, period, and direction), the analysis below, strictly speaking, does not require an exact reconstruction of the sequence of wave events. In this context, the root mean square difference of the modelled and measured time series wave heights (0.5 m) is reasonable. This value is about twice as large as in (Björkqvist et al., 2018) for the Gulf of Finland (0.20–0.31 m) or northern Baltic Proper (0.26 m) and comparable to the level of this quantity for the Bothnian Sea (0.31–0.56 m, Björkqvist et al., 2018).

The appearance of the relevant empirical probability distributions of the occurrence of different wave heights is similar for both data sets (Fig. 3). The location and height of the peaks of these distributions (that represent the properties of most frequently occurring waves) have almost perfect match. The model overestimates to some extent the frequency of waves with heights of 0.8–1.5 m and underestimates the frequency of highest waves in the area (>2 m). The overall appearance of the distribution for modelled wave heights resembles a Weibull distribution whereas this distribution for measured wave heights higher than 0.4 m better matches an exponential distribution and exhibits much larger variability in the frequency of very high (>3 m) waves.

2.3 Nearshore refraction and shoaling

The nearshore grid cells selected for the analysis (Fig. 2) are located in water depth ≥4 m in order to avoid problems with reconstruction of wave heights under possible intense wave breaking in these cells in the strongest storms. Some of the cells are located in much deeper water, at a depth of 20–27 m. The nearshore of the study area contains various underwater features and bottom inhomogeneities. This means that shoaling and refraction may considerably impact the wave fields even along the relatively short paths (normally ≤1 km in our model setup) from the model grid cells to the breaker line. The

Deleted: deep water

Deleted: massive

Deleted: locations

Deleted: with

Deleted: water

Deleted: way

Deleted: until

Deleted: i

Deleted: ng

predominant wind directions in strong storms are from the south-west, north-north-west and west (Soomere et al., 2008). Consequently, high waves often approach some of the selected grid cells at large angles with respect to the shore-normal. Therefore, it is not acceptable to assume that the incidence angles are small. As a result, oversimplified approaches to replicate the changes in wave properties in the immediate nearshore (Lopez-Ruiz et al., 2014; 2015) and even advanced approximations of refraction and shoaling (Hansen and Larson, 2010) may fail.

For this reason we calculate the joint impact of shoaling and refraction of approaching waves in the framework of linear wave theory. Following (Soomere et al., 2013), we assume that the numerically evaluated wave field for each time instant is monochromatic. The wave height is characterised by the numerically simulated (significant) wave height H, peak period T_p and mean propagation direction θ (clockwise with respect to the direction to the North). These properties are evaluated at the centre of each selected grid cell. The significant wave height at this location is denoted as H_0. The approach direction θ_0 at this location with respect to the onshore-directed normal to the shoreline is calculated from θ based on an approximation of the relevant (about 500 m long) coastal segment by a straight line that follows the average orientation of the shoreline in this segment. Similarly, it is assumed that the nearshore seabed from the centre of each grid cell to the waterline is a plane with isobaths strictly parallel to this straight line. Finally, we assume that breaking waves are long waves. Then the wave height H_b at the breaking line can be found as the smaller real solution of the following algebraic equation of 6th order (Viška and Soomere, 2013; Soomere et al., 2013):

$$\frac{H_b^5 g}{H_0^4 \gamma_b}\left(1-\frac{gH_b}{\gamma_b}\frac{\sin^2\theta_0}{c_{f0}^2}\right)=c_{g0}^2\left(1-\sin^2\theta_0\right). \qquad (9)$$

Here c_g is the group speed, c_f is the phase speed and the subscripts "0" and "b" indicate the relevant value at the centre of the particular wave model grid cell and at the breaker line, respectively. The phase and group speed at the wave model grid cell are estimated based on the standard expressions of linear theory using the wave number k evaluated from the dispersion relation for linear monochromatic waves $2\pi/T_p=\sqrt{gk\tanh kd_0}$ with the period equal to the peak period T_p and the water depth d_0 equal to the model depth for the particular grid cell. We assume that waves at the breaker line (where the water depth is d_b) are long waves. The set of assumptions is completed with the common notion that the breaking index is $\gamma_b=H_b/d_b=0.8$ (Dean and Dalrymple, 1991). Part of the introduced assumptions, such as a plane seabed and dry coast without any vegetation, monochromatic wave fields and a constant value of the breaking index for wind-seas as well as the ignoring of the wave period (or steepness) in the calculations are not fully realistic. The potential impact of these approximations is discussed in Section 4.

Deleted: storm

Deleted: e

Deleted: a

Deleted: ea

Deleted: under

Deleted:

Deleted: the

Deleted: using the technique develope in

Deleted: Viška and Soomere, 2013;

Deleted: . As usual, it is

Deleted: d

Deleted: and

Deleted: until

Deleted: e

Deleted: hore

Deleted: the

Deleted: and may lead to certain overestimation of the set-up heights. As are first of all interested in the basic features of the distributions of different up heights, we expect that the use of this of assumptions makes it possible to highlight the features of question in a m contrast manner than they would appear real wave fields. In other words, we mak an attempt to isolate the worst possible scenarios for wave set-up for generally realistic wave conditions.

The procedure of evaluation of set-up heights is thus as follows. We start from the numerically simulated significant wave height H_0, peak period T_p, and mean propagation direction θ with respect to the direction to the North. Next we calculate the phase (c_{fb}) and group speed (c_{gb}) of such waves at the model grid cell and find the propagation direction with respect to the shore normal θ_0. Equation (9) is employed subsequently to evaluate the changes to the wave height owing to refraction and shoaling on its way from the model grid point to the breaker line. The phase and group speed at the breaker line are estimated from the dispersion relation for long waves: $c_{gb} = c_{fb} = \sqrt{gd_b}$. The wave approach direction θ_b at the breaker line with respect to the shorenormal is calculated from Snell's law $\sin\theta / c_f = const$. Thereafter we employ Eqs. (7) and (8) to find the set-up height for the particular time instant.

Several earlier studies of extreme set-up heights (Soomere et al., 2013; Pindsoo and Soomere, 2015) followed this procedure but took into account only waves that propagated under high angles (not larger than ±15°) with respect to the shore-normal) and ignored the correction expressed in Eqs. (7, 8) for waves that approached under a nonzero angle. This approach is denoted S2013 below. It is adequate on the open ocean coasts where waves usually approach the shore under relatively small angles but it may fail in semi-sheltered basins with short fetch.

3 Results

3.1 Maximum set-up heights

The phenomenon of wave set-up is only significant if large waves propagate towards the shore. This is usually the case on open ocean coasts where swells almost always create set-up. The situation may be different in sheltered sea areas with complicated geometry where intense swells may be infrequent and the majority of the wind wave energy may propagate in an offshore direction, this being common in the study area. The wind regime of the study area is a superposition of four wind systems (Soomere et al., 2008). The most frequent wind direction is from south-west (that is, from the mainland to the sea). The proportion of wave fields that propagate onshore is 40–70% along the entire study shoreline (Fig. 4). The statistical properties of set-up heights discussed below thus represent 40 000–70 000 examples of wave fields in each coastal segment. The only exception is grid cell 107 (Figs. 2, 4) between Viimsi Peninsula and the island of Aegna that is sheltered for almost all wind directions.

We start from a comparison of maximum set-up heights evaluated using the above-described approach and the method employed in S2013. The two sets of estimates differ insignificantly (by less than 0.1 m) in about 80% of the coastal segments (Fig. 5). The alongshore variations in the maxima of set-up heights evaluated from Eq. (8) are considerably smaller than those estimated using the approach of S2013. The largest examples of set-up heights reach 1 m and the majority of maximum set-up heights for single coastal sections is 0.6–0.8 m in both sets of estimates.

Deleted: -

Deleted: have

Deleted: taken

Deleted: almost directly onshore

Deleted: This assumption

Deleted: apparently often

Deleted: valid

Deleted: . However,

Deleted: and thus is abandoned in thi
paper

Deleted: profound

Deleted: a certain

Deleted: often

Deleted: from the nearshore towards t
open sea

Deleted: typical

Deleted: . The depicted features are typical for

Deleted: area

Deleted: the one

Deleted: a simpler method (Soomere e
al., 2013) that

Deleted: took into account wave field
that propagated almost directly onshore (±15° with respect to the shore normal)

Deleted: results of the

Deleted: methods

Deleted: e

Deleted: variations of values

Deleted: a simpler method

The largest differences between the two sets become evident in segments that are sheltered from predominant storm directions, most notably in deeply cut bays. Estimates based on Eqs. (7, 8) are often remarkably (by up to 50%) higher in these sections than those derived using S2013. This feature signals that the highest waves approach the shore at a relatively large angle in such sections. This property shows the importance of the generation of remarkable set-up heights by obliquely approaching high waves. Therefore, ignoring waves that approach under large angles may substantially underestimate the maximum set-up height in some coastal segments.

In other words, the impact of refraction often overrides the effect of geometric blocking of waves by changing the orientation of the coastline. Refraction thus often redirects wave energy so that even beaches that are seemingly well sheltered geometrically may at times receive remarkable amounts of wave energy (*cf* Caliskan and Valle-Levinson, 2008). The differences in the maxima of set-up heights evaluated using the two approaches for such coastal sections are often 0.2–0.3 m and reach up to 0.5 m. Such a strong impact of refraction is thought to be responsible for a local increase in wave heights in the Baltic Sea (Soomere, 2003) and also in extreme ocean conditions (Babanin et al., 2011). The processes that are not resolved by phase-averaged wave models such as reflection and diffraction may add even more wave energy to seemingly sheltered coastal segments.

On the contrary, S2013 overestimates the maximum set-up height in a few locations at headlands that are fully open to the Gulf of Finland (Fig. 5). A likely reason for such a sporadic overestimation is the sensitivity of the formation of set-up with respect to the approach angle of waves. The magnitude of set-up rapidly decreases with an increase in the approach angle. This decrease is ignored in S2013. As a result, the height of set-up created by waves that approached at angles of 10–15 degrees with respect to shore-normal was overestimated. This feature also demonstrates the importance of the correct evaluation of refraction and shoaling.

Some differences between the results presented in this paper and those described in Soomere et al., (2013) and Pindsoo and Soomere, (2015) stem from the different time intervals used in the calculations. Simulations for 1981–2012 indicate that the maximum set-up heights in coastal areas open to the east were mostly from the 1980s (Soomere et al., 2013) even though the maximum wave heights occurred starting from the mid-1990s. This feature may be related to a change in the directional structure of strong windss with easterly storms being relatively weak for about two decades. There is increasing evidence, however, that this process has reversed and strong easterly storms have returned to the area. Evidence of this change is the event with significant wave height 5.2 m that was recorded in the Gulf of Finland during an extreme easterly storm on 29–30 November 2012 (Pettersson et al., 2013). Pindsoo and Soomere (2015) observed that many new all-time highest set-up events apparently occurred in coastal segments open to the east since 2012. This process has led to the generation of maximum simulated waves at a number of locations on the eastern Viimsi Peninsula near Leppneeme (grid cells 115–117 in Fig. 2) in 2010 (Fig. 6). These aspects will be addressed in more detail elsewhere.

Deleted: first of all

Deleted: ; often…in these sections of

Deleted: under

Deleted: This difference

Deleted: stresses

Deleted: potential …f the generation o

Deleted: the

Deleted: simpler method

Deleted: of

Deleted: s…the maximum set-up heig

Deleted: here …he effect of geometric

Deleted: In

Deleted: the simpler method

Deleted: that

Deleted: is sensitive …ith respect to th

Deleted: under …t angles of 10–15

Deleted: of

Deleted: latter set of

Deleted: from

Deleted: presented

Deleted: (…oomere et al., (2013) and

Deleted: direction

Deleted: so that

Deleted: n

Deleted: were

Deleted: became …eaker

Deleted: However, t

Deleted: n

Deleted: An

Deleted: The first

Deleted: e

Deleted: that …he event with all-time

Deleted: in …t a number of locations

3.2 Frequency of occurrence of set-up heights

The shape of the empirical distributions of the occurrence of set-up heights varies extensively in the study area (Fig. 7). This shape matches an exponential distribution in the majority (about 75%) of the model coastal segments[1]. Such a distribution is represented by a straight line in the semilogarithmic (log-linear) coordinates used in Figs. 7, 8. It apparently reflects a background Poisson process that also describesstorm surges in the study area (Soomere et al., 2015). This type of distribution appears for coastal segments that are open to the common strong wave directions. For example, grid point 23 (Fig. 7a) is open to the northwest and north, grid point 96 (Fig. 7c) is located near the western shore of Viimsi Peninsula and is open to the west and northwest, and grid point 145 (Fig. 7f) is widely open to the directions from the northwest to the northeast.

The distribution is convex upwards at a few locations that are sheltered from most of predominant approach directions of strong waves, including north-north-west. This shape is evident in the most sheltered location of eastern Kopli Bay (grid point 43, Fig. 7b) and to a lesser extent in coastal sections sheltered by the island of Aegna (grid point 106, Fig. 6d). The relevant empirical distributions of set-up heights can be reasonably approximated by a two-parameter Weibull or Gaussian distribution both of which have a convex upwards shape in log-linear coordinates.

A subset (about ¼) of the presented distributions exhibit a different, clearly concave upwards shape in log-linear coordinates. This feature is evident in coastal sections that are sheltered from a few predominant wave directions. Strong winds blow in this region usually from three directions: west, north-north-west, and east (Soomere et al., 2008). The segments that exhibit a concave upwards shape of the distribution (e.g., grid point 129, Fig. 6e) are mostly sheltered against waves that approach from the west or north-north-west but are open to the east. However, segments that are widely open to all strong wave directions (e.g., grid point 1, Fig. 7) may also exhibit a concave upwards shape of the empirical distribution of set-up heights. This concave upwards appearance clearly differs from the shape of the usual distributions of the magnitude of wave phenomena (Fig. 8) such as the classic (Rayleigh) distribution of single wave heights (Longuet-Higgins, 1952), the Tayfun distribution of the heights of largest waves, the Weibull family of distributions for the occurrence of various wave conditions, or the Rayleigh distribution for run-up of (narrow-banded) Gaussian wave fields (Didenkulova et al., 2008). Therefore, none of the above-mentioned distributions can be used for the universal approximation of the probabilities of different set-up heights.

To further explore the shape of the distributions of set-up heights and their possible variations along the shoreline we assume that these distributions belong to the family of general exponential distributions. The overall appearance of empirical distributions in log-linear coordinates (Fig. 7) suggests that their shape can be, as a first approximation, matched with a

Deleted: extensively

Deleted: are clearly non-Gaussian

Deleted: many locations of

Deleted: 6,

Deleted: , e.g.,

Deleted: Even though the replication o
wave propagation directions by the wave
model and the impact of refraction may
suffer from insufficient resolution of bot
the wind information and wave model,

Deleted: similar

Deleted: for the entire study area. The
appearance of these distributions

Deleted: semilogarithmic (

Deleted:)

Deleted: is often clearly concave upwards

Deleted: also

Deleted: all

Deleted: se

Deleted: The frequently occurring
concave upwards shape signals that even
the exponential distribution (that describ
e.g., storm surges in the study area,
Soomere et al., 2015) is not suitable for
their description. Therefore the backgrou
process of formation of high set-up even
is not a Poisson one (that would lead to a
basically linear shape of the distribution
question in log-linear coordinates).

---

[1] An early discussion version of this paper (available at https://www.earth-syst-dynam-discuss.net/esd-2016-76) contained a bug in the script for the calculation of set-up heights and for the subsequent evaluation of the parameters of their probability density function. This bug led to an erroneous conclusion about the frequency of occurrence of various kinds of distributions of set-up heights in different coastal sections as well as to severe overestimation of the frequency of matches of these distributions with an inverse Gaussian distribution.

quadratic polynomial $az^2 + bz + c$, where z is the set-up height. In other words, the empirical probability density $P(z)$ is approximated by the following function:

$$P(z) = \exp(az^2 + bz + c) \qquad (10)$$

The fitting is performed in log-linear coordinates used in Figs. 7, 8. In the case $a = 0$, distribution (10) reduces to a classic exponential distribution $\sim A\exp(bz)$. The values $a < 0$ correspond to convex upwards distributions that eventually can be approximated by a Weibull or normal distribution whereas $a > 0$ indicates that the distribution is locally concave upwards.

Such a fitting procedure is not straightforward for several reasons. Firstly, the number of nonzero points of the distributions in Fig. 7 is highly variable along the study area similar to the variation in the typical magnitude of the set-up. Secondly, the relevant empirical distributions have gaps for some value(s) of the set-up height. A natural reason for this feature is that we are looking at very low probabilities (down to 0.001%, that is, a few occasions) of occurrence of relatively high set-up events in 1981–2016. Thirdly, a few locations have several outliers. There are remarkably high set-up events that do not follow the general appearance of the empirical distribution of set-up heights for the particular location (Fig. 7b, d, f). Such events apparently reflect severe storms in which the wind pattern was favourable for the development of very large waves that approached a certain coastal segment at a small angle. The presence of similar outliers is characteristic, for example, for time series of sea level in Estonian waters (Suursaar and Sooäär, 2007) and is associated with situations when strong storms blow from a specific direction. Similar to outliers in sea level time series, these events likely follow an extreme value distribution (Coles, 2004) and thus should be left out of the analysis of the shape of the distribution of set-up heights.

To estimate the impact of these aspects on the results, we performed three versions of the fitting procedure. Firstly, we used all data points in the relevant distributions starting from the height of 0.01 m to evaluate the coefficients a, b and c. Secondly, we used for the same purpose only set-up heights from 0.01 m to 0.4 m (Fig. 7). This approach was not applicable in some locations where set-up heights did not reach 0.4 m. Thirdly, we evaluated these coefficients starting from the height of 0.01 m to the first gap in the empirical distribution (the lowest set-up height that did not occur in 1981–2016). Doing so made it possible to check whether the shape of the distribution is governed by the majority of events or if it is dominated by the presence of a few very large set-up heights (Fig. 8).

The particular values of the coefficients a, b and c depend to some extent on the chosen version (Fig. 7). The shape of the approximate distribution is invariant with respect to the particular choice. All distributions also match reasonably well data points corresponding to the largest set-up heights. The differences between the resulting theoretical distributions are mostly insignificant. The relevant estimates are located almost in the middle of the 95% confidence intervals of each other (Fig. 7). The coefficients a at the leading term of the approximating polynomial (Fig. 9, 10) are mostly very small, in the range of (–0.005, 0.005). Their 95% confidence intervals normally include the zero value. This feature indicates that on most occasions the parameter a can be set to zero and the distribution of set-up heights can be reasonably approximated with an exponential

Deleted: p … is the probability of a

Deleted: approximation…is, however,

Deleted: because of

Deleted: depending on…imilar to the

Deleted: host

Deleted: unexpectedly

Deleted: l…min

Deleted: performed

Deleted: analysis…on this approximat

Deleted: in three versionsways… Firs

Deleted: the level of ….4 m. However

Deleted: up

Deleted: roughly evaluate the effect o

Deleted: approximation …ersion (Fig

Deleted: of

Deleted: exponential

Deleted: generally …easonably well a

Deleted: differed to some extent from each other but the differences w…e…e

Deleted: whereas t

Deleted: in

Deleted: the hypothesis that …he

distribution at a 95% level of statistical significance. On such occasions, the entire process can be adequately approximated by a Poisson process and the parameter *b* characterises the vulnerability of the particular coastal segment with respect to the set-up phenomenon similarly to the analysis of storm-driven high water levels (Soomere et al., 2015).

A few outliers of the parameter *a* in relatively sheltered coastal segments were negative and reached values down to –0.08 (Fig. 9). These values correspond to distributions with convex upwards shape in semilogarithmic coordinates and are thus qualitatively similar to the family of Gaussian or Weibull distributions.

Importantly, in about a quarter of the coastal segments in the study area, the parameter *a* is positive and its 95% confidence intervals do not include the zero value. In other words, in these locations the leading term *a* of the quadratic polynomial az^2+bz+c is positive at a 95% significance level. This feature corresponds to the concave-up appearance of the relevant distributions of set-up heights. This means that very large set-up events are systematically much higher and/or occur much more frequently than one could expect from the classic Gaussian, Weibull or Poisson-type statistics. The described features indicate that the empirical distribution of set-up heights can be, at least locally, approximated using an inverse Gaussian (Wald) distribution with a probability density function (Folks and Chhikara, 1978)

$$P=\sqrt{\frac{\lambda}{2\pi z^3}}\exp\left[-\frac{\lambda(z-\mu)^2}{2\mu^2 z}\right]. \qquad (11)$$

For a certain set of parameters λ (the shape parameter) and μ (the mean), a part of the graph of this function has a concave upward shape in semilogarithmic coordinates (Fig. 8) and thus fairly well approximates the empirical distributions of wave set-up at the relevant locations. For large values of *z* this function behaves similarly to the probability density function of a Gaussian distribution.

All coefficients of the quadratic approximation of the exponent vary insignificantly along the study area. This is remarkable because the shape of the relevant Weibull distribution for different wave conditions varies considerably in the study area (Soomere, 2005). The variations in the leading coefficient *a* are uncorrelated with the values of maximum set-up heights in the study area. It is thus likely that the locations where an inverse Gaussian distribution governs the properties of set-up heights appear because of a specific match of the directional structure of winds and the orientation of the coastline. This feature also signals that the basic features of the distribution are only weakly (if at all) connected with the properties of local wave climate. This conjecture is supported by comparatively small variations in the values of other parameters in the polynomial approximation (Fig. 9b,c). The values of *c* are all positive and mostly in the range of 2.5–4 (Fig. 9). The values of parameter *b* are, as expected, almost everywhere negative, concentrated around –0.2, typically varying between –0.1 and –0.4. A few locations with positive values of this parameter correspond to large negative values of *a*.

Deleted: In this case

Deleted: exclusively

Deleted: the

Deleted: of

Deleted: t

Deleted: is

Deleted: appearance first of all

Deleted: or

Deleted: 9

Deleted: in

Deleted: At

Deleted: largely

Deleted: along

Deleted: of

Deleted: along

Deleted: presence of

Deleted: is connected

Deleted: with

Deleted: in question

Deleted: to

Deleted: and

Deleted:

4 Discussion and conclusions

The analysis reveals that numerical estimates of maxima of wave set-up heights are relatively sensitive with respect to how the impact of radiation stress and the transformation of wave properties in the nearshore are evaluated. The magnitude of the related effects substantially depends on the bathymetry. Refraction can easily override the purely geometric effects of shoreline orientation changes and redirect substantial levels of wave energy into seemingly sheltered shore sections. This means that high-resolution information about wind (including wind directions) and bathymetry, together with advanced methods for the evaluation of propagation and impact of radiation stress in the nearshore in operational and hindcast models of coastal flooding, are required.

The core message from the analysis is that the empirical probability distribution of different set-up heights can usually be fairly well approximated by a standard exponential distribution $\exp(-\lambda z)$. When the exponent function of the general exponential distribution is approximated using a quadratic function, the coefficient of its leading term does not differ from zero at a 95% significance level for more than ¾ of the coastal segments of the study area. As the study area contains a variety of sections open to different directions and with radically different wave properties, it is likely that the qualitative shape of the distribution only weakly depends on the properties of the local wave climate.

Another important message is that the basic shape of this distribution function is concave upwards in a log-linear plot for a substantial number of coastal segments. The local shape of the relevant empirical distributions of wave set-up heights can be adequately approximated with a family of inverse Gaussian (Wald) distributions. Even though the absolute values of the coefficients of the leading term of such a quadratic approximation are relatively small, approximations with other classic distributions, such as Rayleigh or Weibull (Fig. 8), are evidently inappropriate. As the coefficient of the linear term of this quadratic approximation is relatively small (Fig. 9b), the use of a Lévy distribution might also be appropriate.

This result is intriguing because sensible approximations with inverse Gaussian (Wald) distributions are scarce in descriptions of geophysical phenomena. Perhaps the most well-known example of the use of a Wald distribution is to describe the time a Brownian motion (with positive drift) takes to reach a fixed positive level. Other examples include statistical properties of soil phosphorus (Manunta et al., 2002), long-distance seed dispersal by wind (Katul et al., 2005), and some models of failure (Park and Padgett, 2005).

The appearance of the distribution of modelled wave heights in the offshore (Fig. 3) is convex upwards in the range of relatively frequent wave heights of 0.5–1.7 m. It would thus be natural to expect that this property also becomes evident in setup heights. The distribution of measured wave heights largely follows an exponential distribution for 1.2–3.2 m high waves and is only slightly convex upwards for 0.5–1.5 m high waves. This difference in the distributions for modelled and measured wave heights suggests that the presence of convex upwards distributions of set-up heights in nature may be even more pronounced than is demonstrated in Figs. 7 and 9. This difference also signals that the several approximations described above are not responsible for the presence of convex upwards distributions of set-up heights. For example, a

Deleted: performed

Deleted: first of all

Deleted: of

Deleted: appearance of

Deleted: The impact of r

Deleted: feature calls for the necessity using

Deleted: .

Deleted: of

Deleted: be

Deleted: at

Deleted: of 95%

Deleted: this

Deleted: in question

Deleted: in

Deleted: at

Deleted: ones

Deleted: at

Deleted: of

Deleted: or

Deleted: and we have established no arguments in favour of this distribution wave set-up heights.

natural conjecture from Fig. 3 is that ignoring ice cover and the use of discontinuous wind data have at least partially supported the convex upwards shape of the distribution of wave heights.

Some of the introduced assumptions such as the ideal plane and rigid seabed, the presence of a dry coast without any vegetation, and ignoring the wave period and coastal slope (and thus wave reflection) in the calculations are not fully realistic. They all generally lead to an overestimation of set-up heights (Dean and Bender, 2006). As they impact set-up heights in the same manner, independently of wave properties, it is likely that they mainly stretch the resulting distributions of set-up heights towards larger values but do not modify their basic shapes. It might be expected that the impact of other simplifications such as the assumption of monochromatic wave fields, using a constant value of the breaking index and employing long wave approximation for breaking waves, generally emphasizes the role of approach directions. Therefore, it is likely that the set of assumptions used makes the established features more noticeable than they would be for real wave fields.

Finally, we note that the presented results do not require any modification of the classic estimates of extreme values of set-up heights and their return periods based on, for example, the block maximum method. Namely, the limiting distributions of independent block maxima follow either a Gumbel, Weibull or Frechet distribution notwithstanding the distribution of the underlying values (Coles, 2004). This general theorem is obviously also valid for any time series that follows an inverse Gaussian distribution. A subtle implication from the qualitative match of statistics of set-up heights with an inverse Gaussian distribution is that set-up events with heights close to extreme heights may be much more frequent than their estimates based on classic Gaussian or Weibull statistics and also clearly more frequent than similar estimates for Poisson processes. This increase in the probability of large wave set-up events is balanced by a similar decrease in the relative number of events with an average magnitude compared to normally distributed events. The described features basically indicate that the frequency and role of close to extreme set-up events (and their contribution to damages and economic losses) may be underestimated based on observations of similar events of average height. In particular, severe set-up events may occur substantially more frequently that could be expected from the probability of the occurrence of severe seas.

Deleted: on

Deleted: also

Deleted: as well.

Deleted: delicate

Deleted: ones

Deleted: -

Deleted: -

Deleted: associated

Deleted: it

Code availability: From the authors on request (Matlab scripts).

Data availability: Kalbådagrund wind data are available from the Finnish Meteorological Institute.

Sample availability: not applicable.

Authors team: Tarmo Soomere, Katri Pindsoo, and Maris Eelsalu

Author contribution. T. Soomere designed the study, derived the equations and approximations used in the paper, produced Figure 6, compiled the introduction and discussion, checked the consistency of the results, and polished the text. K. Pindsoo developed the scripts, ran the simulations, produced most of graphics, and drafted the body parts of the manuscript. M. Eelsalu recalculated the graphics, performed a comparison of the properties of the reconstructed waves with measured waves near the study area, evaluated the distributions of set-up heights for time series of wave properties measured in the northern Baltic Proper, drafted the relevant parts of the manuscript and produced updated maps.

Deleted: entire

Competing interests: The authors declare that they have no conflict of interest.

Disclaimer: All authors have approved the submitted version.

Acknowledgements

The research was financed by the institutional support of the Estonian Ministry of Education and Research (IUT33-3), the project "Sebastian Checkpoints – Lot 3 Baltic" of the call MARE/2014/09 and by the Estonian Research Council (ETAg) to two networks: the ERA-NET Rus+ network EXOSYSTEM and FLAG-ERA network FuturICT2.0. The authors are greatly thankful to the Finnish Meteorological Institute for making the Kalbådagrund wind data public, the Marine Systems Institute, Tallinn University of Technology, for providing wave data at Tallinnamadal, and to three anonymous referees for valuable suggestions towards improvement of the manuscript.

References

Alari, V. and Kõuts, T.: Simulating wave–surge interaction in a non-tidal bay during cyclone Gudrun in January 2005, in: Proceedings of the IEEE/OES Baltic 2012 International Symposium "Ocean: Past, Present and Future. Climate Change Research, Ocean Observation & Advanced Technologies for Regional Sustainability," May 8–11, Klaipėda, Lithuania. IEEE Conference Publications, doi: 10.1109/BALTIC.2012.6249185, 2012.

Arns, A., Wahl, T., Haigh, I. D., Jensen, J., and Pattiaratchi, C.: Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise, Coast. Eng., 81, 51–66, doi: 10.1016/j.coastaleng.2013.07.003, 2013.

Averkiev, A. S. and Klevanny, K. A.: A case study of the impact of cyclonic trajectories on sea-level extremes in the Gulf of Finland, Cont. Shelf Res., 30, 707–714, 2010.

Babanin, A. V., Hsu, T.-W., Roland, A., Ou, S.-H., Doong, D.-J., and Kao, C. C.: Spectral wave modelling of Typhoon Krosa, Nat. Hazards Earth Syst. Sci., 11, 501–511, doi: 10.5194/nhess-11-501-2011, 2011.

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., and Hunt, T.: A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines, Ocean Eng., 71, 28–39, 2013.

Bechle, A. J., Kristovich, D. A. R., and Wu, C. H.: Meteotsunami occurrences and causes in Lake Michigan, J. Geophys. Res-Oceans, 120, 8422–8438, doi: 10.1002/2015JC011317, 2015.

Björkqvist, J.-V., Lukas, I., Alari, V., van Vledder, G. Ph., Hulst, S., Pettersson, H., Behrens, A., Männik, A.: Comparing a 41-year model hindcast with decades of wave measurements from the Baltic Sea. Ocean Eng., 152, 57–71, doi: 10.1016/j.oceaneng.2018.01.048, 2018.

Bortot, P., Coles, S., and Tawn, J.: The multivariate Gaussian tail model: an application to oceanographic data. J. Roy. Stat. Soc. A-Appl. Stat. 49, 31–49, doi: 10.1111/1467-9876.00177, 2000.

Caliskan, H., and Valle-Levinson, A.: Wind-wave transformations in an elongated bay. Cont. Shelf Res., 28, 1702–1710, doi: 10.1016/j.csr.2008.03.009, 2008.

Cazenave, A., Dieng B., Meyssignac, B., von Schuckmann, K., Decharme, B., and Berthier, E.: The rate of sea-level rise, Nat. Clim. Change, 4, 358–361, doi:10.1038/nclimate2159, 2014.

Coles, S.: An introduction to statistical modeling of extreme values. Springer, 3rd printing, Springer, London, 208 pp., 2004.

Darwin, R. F., and Tol, R. S. J.: Estimates of the economic effects of sea level rise, Environ. Resource Economics, 19, 113–129, doi: 0.1023/A:1011136417375, 2001.

Dean R.G., and Dalrymple R.A: Water Wave Mechanics for Engineers and Scientists. World Scientific, 1991.

Dean, R. G., and Bender, C. J.: Static wave set-up with emphasis on damping effects by vegetation and bottom friction, Coast. Eng., 53, 149–165, 2006.

Denissenko, P., Didenkulova, I., Pelinovsky, E., and Pearson, J.: Influence of the nonlinearity on statistical characteristics of long wave runup, Nonlin. Proc. Geophys., 18(6), 967–975, doi: 10.5194/npg-18-967-2011, 2011.

Denissenko, P., Didenkulova, I., Rodin, A., Listak, M., and Pelinovsky, E.: Experimental statistics of long wave runup on a plane beach, J. Coast. Res., Special Issue 65, 195–200, doi: 10.2112/SI65-034.1, 2013.

Didenkulova, I.: New trends in the analytical theory of long sea wave runup. In: Applied Wave Mathematics (Quak, E. and Soomere, T., eds.), Springer, 265–296, doi: 10.1007/978-3-642-00585-5_14, 2009.

Didenkulova, I., Pelinovsky, E., and Sergeeva, A.: Statistical characteristics of long waves nearshore. Coast. Eng. 58, 94–102, doi: 10.1016/j.coastaleng.2010.08.005, 2008.

Dube, S. K., Jain, I., Rao, A. D., and Murty, T. S.: Storm surge modelling for the Bay of Bengal and Arabian Sea, Nat. Hazards, 51, 3–27, doi: 10.1007/s11069-009-9397-9, 2009.

Dukhovskoy, D.S., and Morey, S. L.: Simulation of the Hurricane Dennis storm surge and considerations for vertical resolution, Nat. Hazards, 58, 511–540, doi: 10.1007/s11069-010-9684-5, 2011.

Fawcett, L., and Walshaw, D.: Sea-surge and wind speed extremes: optimal estimation strategies for planners and engineers, Stoch Environ Res Risk Assess, 30:463–480, DOI 10.1007/s00477-015-1132-3, 2016.

Feng, X., Tsimplis, M. N., Quartly, G. D., and Yelland, M. J.: Wave height analysis from 10 years of observations in the Norwegian Sea. Cont Shelf Res 72, 47-56, doi: 10.1016/j.csr.2013.10.013, 2013.

Forristall, G. Z.: Statistical distribution of wave heights in a storm, J. Geophys. Res-Oceans, 83(NC5), 2353–2358, 1978.

Folks, J. L. and Chhikara, R. S.: The inverse Gaussian distribution and its statistical application–A review. J. R. Stat. Soc. Ser. B-Methodol., 40, 263–289, doi: 10.2307/2984691, 1978.

Geist, E. L., ten Brink, U. S., and Gove, M.: A framework for the probabilistic analysis of meteotsunamis, Nat. Hazards, 74, 123–142, doi: 10.1007/s11069-014-1294-1, 2014.

Hallegatte, S., Green, C., Nicholls, R. J., and Corfee-Morlot, J.: Future flood losses in major coastal cities, Nat. Clim. Change, 3, 802–806, 2013.

Hawkes, P. J., Gouldby, B. R., Tawn, J. A., and Owen, M. W.: The joint probability of waves and water levels in coastal engineering design, J. Hydraulic Res., 40, 241–251, 2002.

Haigh, I. D., Nicholls, R., and Wells, N.: A comparison of the main methods for estimating probabilities of extreme still water levels, Coast. Eng., 57, 838–849, 2010.

Hoeke, R. K., McInnes, K. L., Kruger, J. C., McNaught, R. J., Hunter, J. R., and Smithers, S. G.: Widespread inundation of Pacific islands triggered by distant-source wind-waves, Glob. Planet Change, 108, 128–138, 2013.

Holland, K. T. and Holman, R. A.: The statistical distribution of swash maxima on natural beaches, J. Geophys. Res.–Oceans, 98(C6), 10271–10278, doi: 10.1029/93JC00035, 1993.

Hsu, T.-W., Hsu, J. R.-C., Weng, W.-K., Wang, S.-K., and Ou, S.-H.: Wave setup and setdown generated by obliquely incident waves, Coast. Eng., 53, 865–877, 2006.

Johansson, M., Boman, H., Kahma, K., and Launiainen, J.: Trends in sea level variability in the Baltic Sea, Boreal Environ. Res. 6, 159–179, 2001.

Katul, G. G., Porporato, A., Nathan, R., Siqueira, M., Soons, M. B., Poggi, D., Horn, H. S., and Levin, S. A.: Mechanistic analytical models for long-distance seed dispersal by wind, American Naturalist, 166, 368–381, doi: 10.1086/432589, 2005.

Keevallik, S., and Soomere, T.: Towards quantifying variations in wind parameters across the Gulf of Finland, Estonian J. Earth Sci., 59, 288–297, 2010.

Kulikov, E. A., and Medvedev, I. P.: Variability of the Baltic Sea level and floods in the Gulf of Finland, Oceanology, 53, 145–151, doi: 10.1134/S0001437013020094, 2013.

Larson, M., Hoan, L. X., and Hanson, H.: Direct formula to compute wave height and angle at incipient breaking, J. Waterw. Port C-ASCE., 136, 119–122, doi: 10.1061/ASCEWW.1943-5460.0000030, 2010.

Longuet-Higgins, M. S.: On the statistical distribution of the heights of sea waves, J. Mar. Res., 11, 1245–1266, 1952.

Longuet-Higgins, M. S., and Stewart, R. W.: Radiation stress and mass transport in gravity waves with application to ”surf-beats”, J. Fluid Mech., 8, 565–583, 1962.

Longuet-Higgins, M. S., and Stewart, R. W.: Radiation stresses in water waves: a physical discussion with applications, Deep-Sea Res., 11, 529–562, 1964.

Lopez-Ruiz, A., Ortega-Sanchez, M., Baquerizo, A., and Losada, M. A.: A note on alongshore sediment transport on weakly curvilinear coasts and its implications, Coast. Eng., 88, 143–153, doi: 10.1016/j.coastaleng.2014.03.001, 2014.

Lopez-Ruiz, A., Solari, S., Ortega-Sanchez, M., and Losada, M.: A simple approximation for wave refraction – Application to the assessment of the nearshore wave directionality, Ocean Modell., 96, 324–333, doi: 10.1016/j.ocemod.2015.09.007, 2015.

Manunta, P., Feng, Y., Goddard, T., Anderson, A. M., and Cannon, K.: Analysis of soil test phosphorus to assess the risk of P transport in a watershed. Comm. Soil Sci. Plant Analysis, 33, 3481–3492, doi: 10.1081/CSS-12001452, 2002.

McDougal, W. G., Hudspeth, R. T.: Wave setup/setdown and longshore current on non-planar beaches, Coast. Eng., 7, 103–117, 1983.

Masina, M., Lamberti, A., and Archetti, R.: Coastal flooding: A copula based approach for estimating the joint probability of water levels and waves, Coast. Eng., 97, 37–52, doi: 10.1016/j.coastaleng.2014.12.010, 2015.

Mel, R., and Lionello, P.: Verification of an ensemble prediction system for storm surge forecast in the Adriatic Sea, Ocean Dyn., 64, 1803–2814. doi: 10.1007/s10236-014-0782-x, 2014.

Melet, A., Almar, R., and Meyssignac, B.: What dominates sea level at the coast: a case study for the Gulf of Guinea, Ocean Dyn., 66, 623–636, doi: 10.1007/s10236-016-0942-2, 2016.

Melet, A., Meyssignac, B., Almar, R., and Le Cozannet, G.: 2018. Under-estimated wave contribution to coastal sea-level rise. Nature Clim. Change, 8, 234–239, doi: 10.1038/s41558-018-0088-y, 2018.

Meyer, V., Becker, N., Markantonis, V.,; Schwarze, R., van den Bergh, J. C. J. M., Bouwer, L. M., Bubeck, P., Ciavola,P., Genovese, E., Green, C., Hallegatte, S., Kreibich, H., Lequeux, Q., Logar, I., Papyrakis, E., Pfurtscheller, C., Poussin, J., Przyluski, V., Thieken, A. H., and Viavattene, C. Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps, Nat. Hazards Earth Syst. Sci., 13, 1351–1373, doi: 10.5194/nhess-13-1351-2013, 2013.

Moghimi, S., Klingbeil, K., Gräwe, U., and Burchard, H.: A direct comparison of a depth-dependent Radiation stress formulation and a Vortex force formulation within a three-dimensional coastal ocean model, Ocean Modell., 70, 132–144, http://dx.doi.org/10.1016/j.ocemod.2012.10.002, 2013.

Muraleedharan, G., Rao, A. D., Kurup, P. G., Nair, N. U., and Sinha, M.: Modified Weibull distribution for maximum and significant wave height simulation and prediction, Coast. Eng. 54, 630–638, doi: 10.1016/j.coastaleng.2007.05.001, 2007.

Nikolkina, I., Soomere, T., and Räämet, A.: Multidecadal ensemble hindcast of wave fields in the Baltic Sea. In: The 6th IEEE/OES Baltic Symposium Measuring and Modeling of Multi-Scale Interactions in the Marine Environment, May 26–29, Tallinn Estonia. IEEE Conference Publications, 9 pp., doi: 10.1109/BALTIC.2014.6887854, 2014.

O'Grady, J. G., McInnes, K. L., and Hoeke, R. K.: Forecasting maximum wave setup hazards around Australia. Coasts & Ports Conference 2015, 15–18 September 2015, Pullman Hotel, Auckland, 636–641, 2015.

Olbert, A. I., Nash, S., Cunnane, C., and Hartnett, M.: Tide-surge interactions and their effects on total sea levels in Irish coastal waters, Ocean Dyn., 63, 599–614, 2013.

Orimolade, A. P., Haver, S., and Gudmestad, O. T.: Estimation of extreme significant wave heights and the associated uncertainties: A case study using NORA10 hindcast data for the Barents Sea, Marine Structures 49. 1e17, http://dx.doi.org/10.1016/j.marstruc.2016.05.004, 2016.

Park, C. and Padgett, W. J.: Accelerated degradation models for failure based on geometric Brownian motion and gamma processes, Lifetime Data Anal., 11, 511–527, doi: 10.1007/s10985-005-5237-8, 2005.

Pattiaratchi, C., and Wijeratne, E. M. S.: Observations of meteorological tsunamis along the south-west Australian coast, Nat. Hazards, 74, 281–303, 2014.

Pellikka, H., Rauhala, J., Kahma, K.K., Stipa, T., Boman, H., and Kangas, A.: Recent observations of meteotsunamis on the Finnish coast, Nat. Hazards, 74, 197–215, doi: 10.1007/s11069-014-1150-3, 2014.

Pettersson, H., Lindow, H., and Brüning, T.: Wave climate in the Baltic Sea 2012. HELCOM Baltic Sea Environment Fact Sheets 2012. http://helcom.fi/baltic-sea-trends/environment-fact-sheets/hydrography/wave-climate-in-the-baltic-sea/, 2013.

Pindsoo, K., and Soomere, T.: Contribution of wave set-up into the total water level in the Tallinn area, Proc. Estonian Acad. Sci., 64(3S), 338–348, doi: 10.3176/proc.2015.3S.03, 2015.

Purvis, M. J., Bates, P. D., and Hayes, C. M.: A probabilistic methodology to estimate future coastal flood risk due to sea level rise, Coast. Eng. 55, 1062–1073, doi: 10.1016/j.coastaleng.2008.04.008, 2008.

Pycroft, J., Abrell, J., and Ciscar, J. C.: The global impacts of extreme sea-level rise: A comprehensive economic assessment, Environ. Resource Economics, 64, 225–253, doi: 10.1007/s10640-014-9866-9, 2016.

Roland, A., Cucco, A., Ferrarin, C., Hsu, T.-W., Liau, J.-M., Ou, S.-H., Umgiesser, G., and Zanke, U.: On the development and verification of a 2-D coupled wave-current model on unstructured meshes, J. Mar. Syst., 78, S244–S254, 2009.

Rueda, A., Camus, P., Mendez, F. J., Tomas, A., and Luceno, A.: An extreme value model for maximum wave heights based on weather types, J. Geophys. Res-Oceans, 121, 1262–1273, doi: 10.1002/2015JC01095, 2016a.

Rueda, A., Camus, P., Tomas, A., Vitousek, S., and Mendez, F. J.: A multivariate extreme wave and storm surge climate emulator based on weather patterns. Ocean Modell., 104, 242–251, doi: 10.1016/j.ocemod.2016.06.008, 2016b.

Sallenger Jr, A. H., List, J. H., Gelfenbaum, G., Stumpf, R. P., and Hansen, M.: Large wave at Daytona Beach, Florida, explained as a squall-line surge, J. Coast. Res., 11, 1383–1388, 1995.

Schmitt, F. G., Crapoulet, A., Hequette, A., and Huang, Y.: Nonlinear dynamics of the sea level time series in the eastern English Channel, Nat. Hazards, 91(1), 267–285, doi: 10.1007/s11069-017-3125-7, 2018.

Serafin, K. A., and Ruggiero, P.: Simulating extreme total water levels using a time-dependent, extreme value approach, J. Geophys. Res-Oceans, 119, 6305–6329, doi: 10.1002/2014JC010093, 2014.

Socquet-Juglard, H., Dysthe, K., Trulsen, K., Krogstad, H. E., and Liu, J. D.: Probability distributions of surface gravity waves during spectral changes, J. Fluid Mech., 542, 195–216, doi: 10.1017/S0022112005006312, 2005.

Soomere, T.: Anisotropy of wind and wave regimes in the Baltic Proper, J. Sea Res., 49, 305–316, 2003.

Soomere, T.: Wind wave statistics in Tallinn Bay, Boreal Environ. Res., 10, 103–118, 2005.

Soomere, T., Myrberg, K., Leppäranta, M., and Nekrasov, A.: The progress in knowledge of physical oceanography of the Gulf of Finland: a review for 1997–2007, Oceanologia, 50(3), 287–362, 2008.

Soomere, T., Pindsoo, K., Bishop, S. R., Käärd, A, and Valdmann, A.: Mapping wave set-up near a complex geometric urban coastline. Nat. Hazards Earth Syst. Sci., 13, 3049–3061, doi: 10.5194/nhess-13-3049-2013, 2013.

Soomere, T., Eelsalu, M., Kurkin, A., and Rybin, A.: Separation of the Baltic Sea water level into daily and multi-weekly components, Cont. Shelf Res., 103, 23–32, doi:10.1016/j.csr.2015.04.018, 2015.

Stockdon, H. F., Holman, R. A., Howd, P. A., and Sallenger, A. H.: Empirical parameterization of setup, swash, and runup, Coast. Eng., 53, 573–588, 2006.

Suursaar, Ü., and Sooäär, J.: Decadal variations in mean and extreme sea level values along the Estonian coast of the Baltic Sea. Tellus A 59(2), 249–260, doi: 10.1111/j.1600-0870.2006.00220.x, 2007.

[SWAN]: SWAN Technical Documentation. Delft University of Technology. The Netherlands, 2007.

Tuomi, L., Kahma, K. K., and Pettersson, H.: Wave hindcast statistics in the seasonally ice-covered Baltic Sea, Boreal Environ. Res., 16, 451–472, 2011.

Tuomi, L., Kahma, K. K., and Fortelius, C.: Modelling fetch-limited wave growth from an irregular shoreline, J. Mar. Syst., 105, 96–105, 2012.

Vilibic, I.: The role of the fundamental seiche in the Adriatic coastal floods, Cont. Shelf Res., 26, 206–216, doi: 10.1016/j.csr.2005.11.001, 2006

Vilibic, I., Monserrat, S., and Rabinovich, A. B.: Meteorological tsunamis on the US East Coast and in other regions of the World Ocean, Nat. Hazards, 74, 1–9, doi: 10.1007/s11069-014-1350-x, 2014.

Viška, M., and Soomere, T.: Simulated and observed reversals of wave-driven alongshore sediment transport at the eastern Baltic Sea coast, Baltica, 26 (2), 145–156, doi: 10.5200/baltica.2013.26.15, 2013.

Wadey, M. P., Brown, J. M., Haigh, I. D., Dolphin, T., and Wisse, P.: Assessment and comparison of extreme sea levels and waves during the 2013/14 storm season in two UK coastal regions, Nat. Hazards Earth Syst. Sci., 15, 2209–2225, doi: 10.5194/nhess-15-2209-2015, 2015.

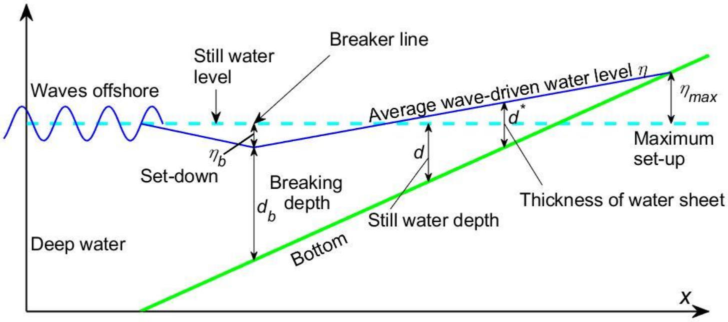


Figure 1: Scheme of the cross-section of a coastal area with wave set-up. The sign of η is positive if the average wave-driven water level exceeds the still water level and negative in the opposite case. The sign of d is positive in the area covered by still water and negative in the otherwise dry section of the coast. The quantity d^* is non-negative.

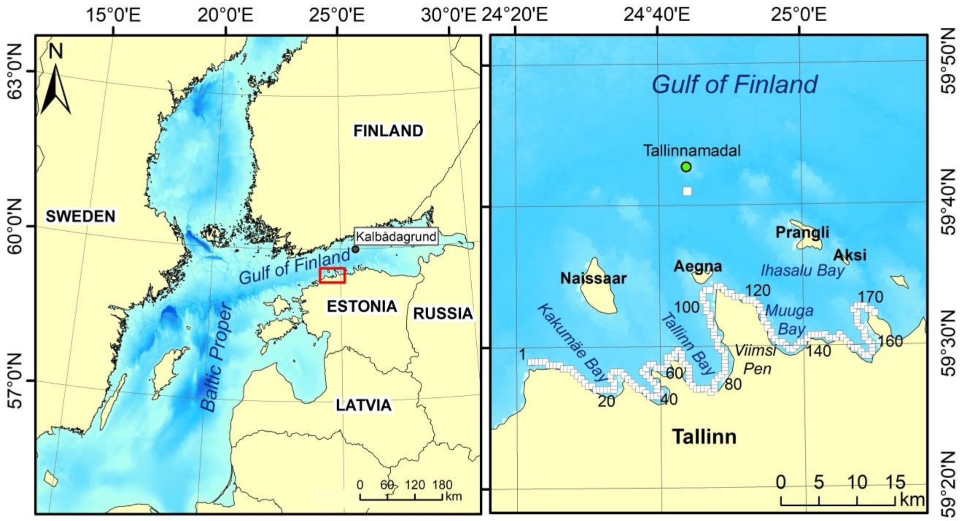
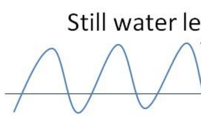


Figure 2: Study area (red box in the left panel) in the vicinity of Tallinn Bay. Small circles along the shoreline in the right panel indicate the nearshore grid cells of the wave model WAM with a resolution of about 470 m. The grid cells are numbered consecutively from the west to the east. The green circle shows the location of the wave buoy at Tallinnamadal and the white square to the south of it the closest grid point of the wave model used for comparison of modelled and measured wave data.



Deleted:

Deleted: hosting

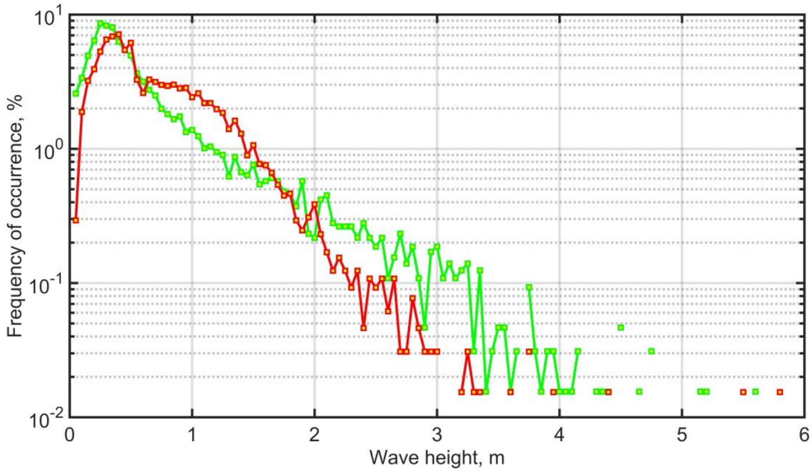


Figure 3: Empirical probability distributions of measured wave heights at Tallinnamadal (green) and modelled wave heights in the closest grid cell (red) in 2012–2016. The missing lines between the data points indicate gaps in the sets of frequencies.

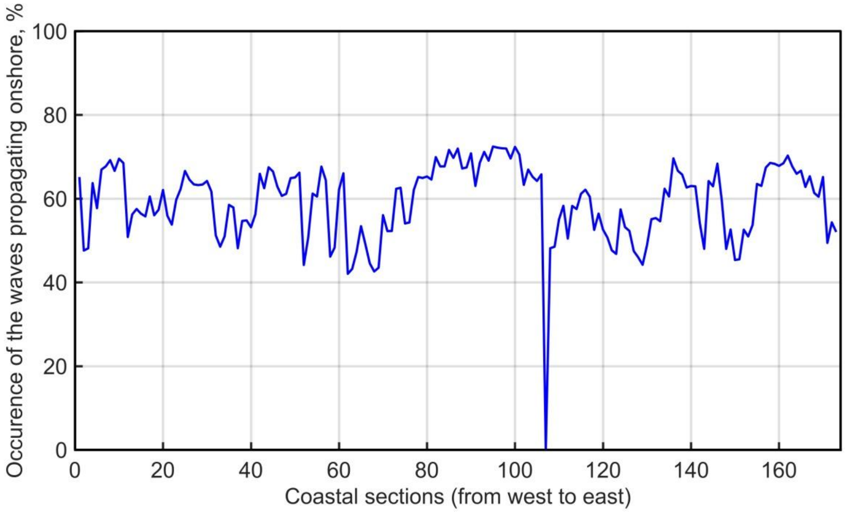


Figure 4: The percentage of occurrence of waves that propagate onshore and produce elevated wave set-up events in the study area.

Deleted: 3

Deleted: along

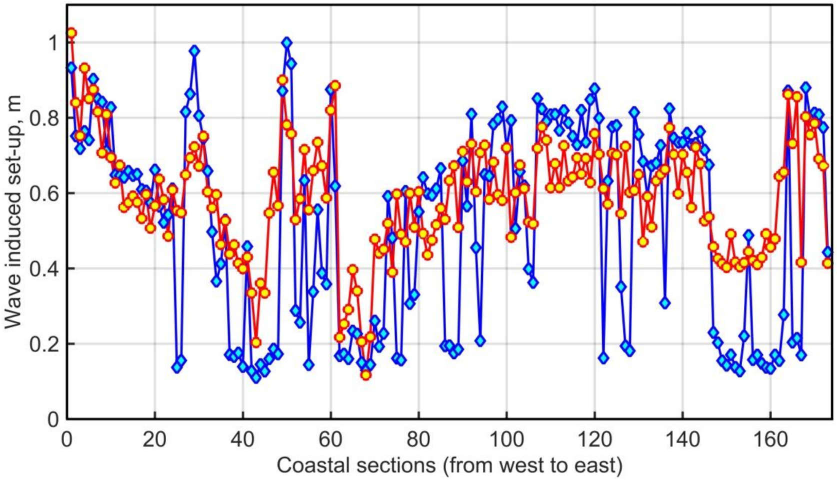


Figure 5: Maximum set-up heights evaluated using all onshore-propagating wave fields and Eqs. (1–8) (red circles) and similar heights evaluated using only those waves that approach the shore at an angle less than ±15° from shore-normal (blue diamonds).

Deleted: 4

Deleted: the onshore

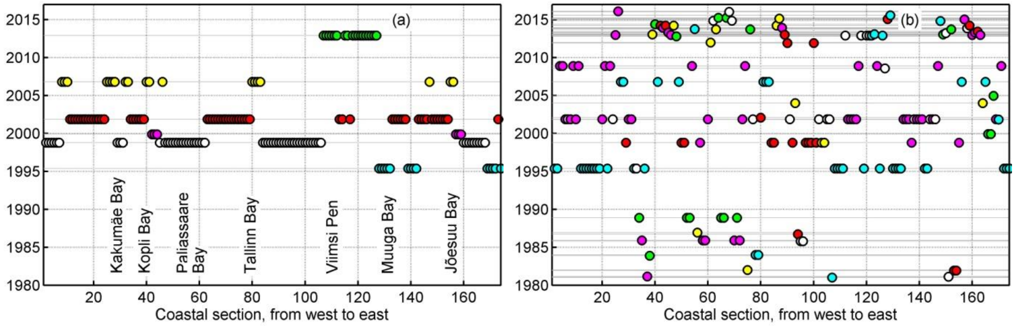


Figure 6: (a) Six storms that caused the highest waves in different coastal sections of the study area in January 1981–May 2016. Notice the cluster of green circles along the eastern coast of Viimsi Peninsula in an autumn storm of 2013. (b) 58 storms that caused the highest wave set-up in these sections in January 1981–May 2016. The set-up heights are evaluated similarly to the procedure in (Pindsoo and Soomere, 2015) using only waves approaching at an angle ±15° with respect to shore-normal. Colours vary cyclically in time and correspond to different storms in single years.

Deleted: 5

Deleted: the

Deleted:

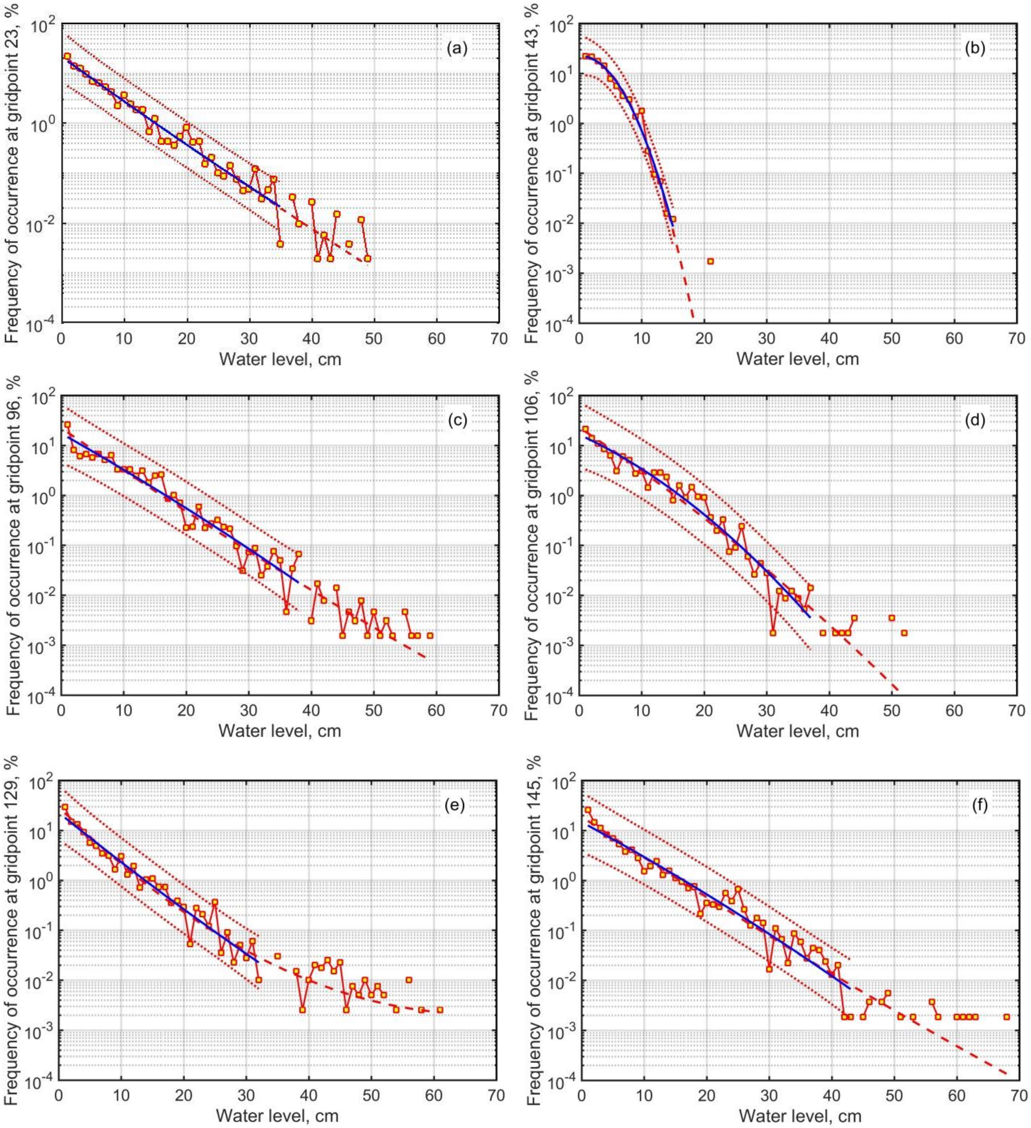


Figure 3: Simulated distributions of various set-up heights (red squares) at various locations in the Tallinn Bay area open to different directions. Blue line: interpolation with a quadratic function from the set-up height of 0.01 m to the first gap in the empirical distribution; red dotted lines: its 95% confidence intervals; red dashed line: similar interpolation using all data points. The interpolating lines evaluated using only the data points from 0.01 m to 0.4 m are fully masked by blue lines.

Deleted: of

Deleted: until

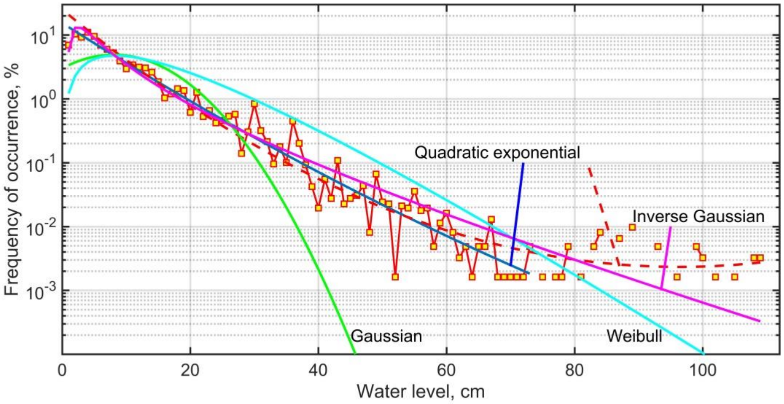
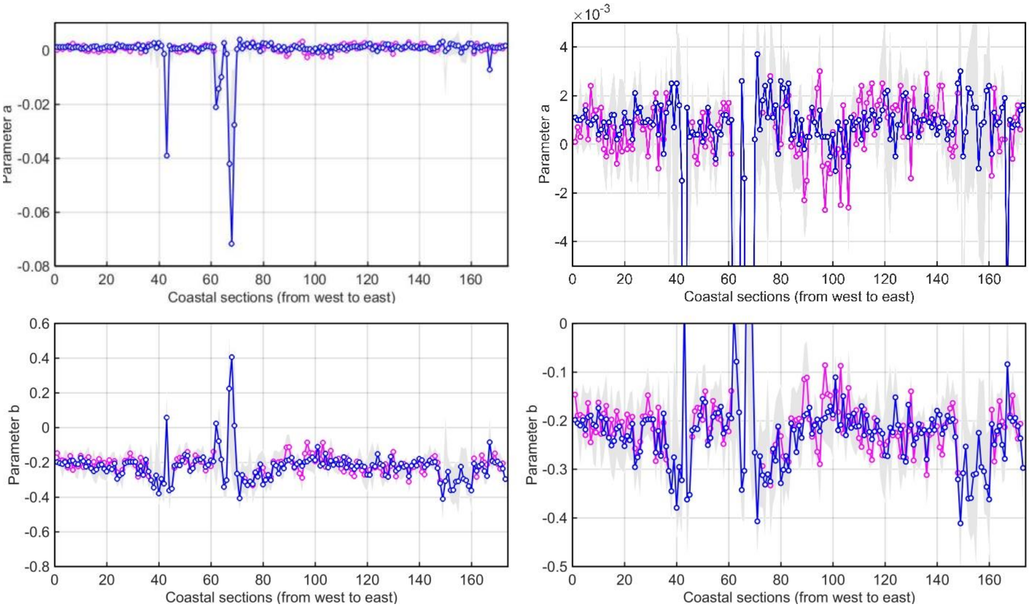


Figure 8: Simulated distributions of various set-up heights (red squares) in the easternmost coastal segment of the study area (grid point 1) and the relevant Gaussian, Weibull and inverse Gaussian (Wald) distributions evaluated using the method of moments. Blue line: interpolation of the empirical distribution in semilogarithmic coordinates with a quadratic function (equivalently, the formal local exponential distribution with a general quadratic exponent) from the set-up height of 0.01 m to the first gap in the empirical distribution; red dashed line: similar interpolation using all data points.

Deleted: 6

Deleted: normal

Deleted: until



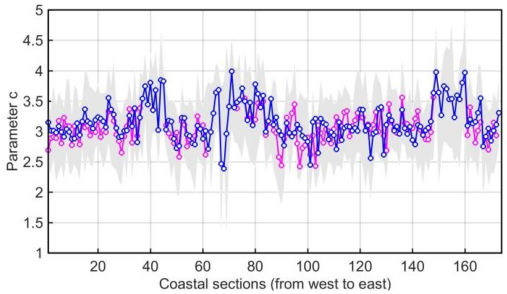


Figure 9: Alongshore variation of the coefficients a, b, c of the quadratic approximation az^2+bz+c of the exponent of empirical distributions of set-up heights in the Tallinn Bay area. For the parameters a and b, more detailed alongshore variation is presented in graphs with vertically stretched scales. Blue line: the respective parameter calculated for the range of set-up heights from 0.01 m to the first gap in the empirical distribution; the grey area marks the 95% confidence interval of this value, the pink line describes the values of the relevant parameter for the range of set-up heights from 0.01 m to 0.04 m.

Deleted: 8

Deleted: ,

Deleted: also

Deleted: until

Deleted: until

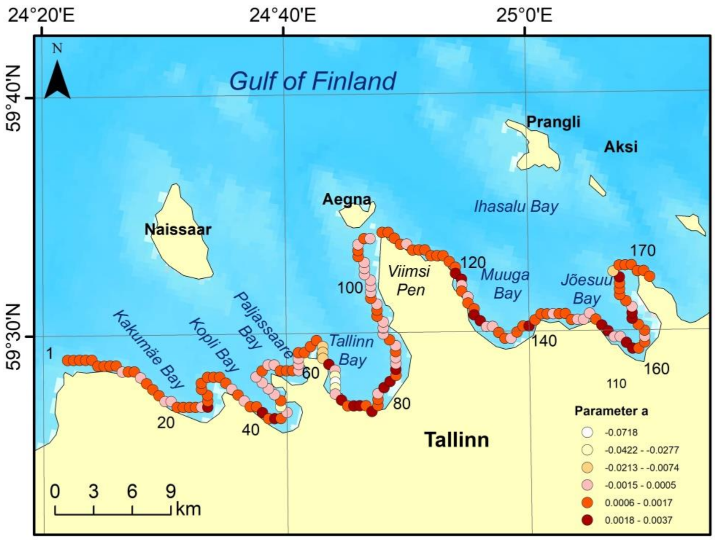


Figure 10: Alongshore variation in the coefficient of the leading term (colour code) in the approximation of the exponent of empirical distribution of set-up heights at single locations. The grid cells are numbered consecutively from the west to the east.