Referee Comments #2

Comment	Authors response and changes in manuscript
[1] The first comment is purely formal. Authors state in lines 22-23 (pag 3) that they are going to call “run-up level” to the combined water elevation (mean water level and wave run-up contributions. This is misleading since it is not the unusual approach in the literature. It should be better to use something like “total water level” to avoid confusion with the standard wave-induced run-up.	After re-consideration of the terms used in the manuscript we agree that using “run-up level” to represent the combination of still water level and wave run-up might be misleading and cause confusion with the wave related run-up. Thus, we will replace “run-up level” with “total water level” throughout the manuscript as suggested.
[2] *Lines 13-14 (pag 3).* Coastal floods are also a consequence of storm-surges. Please rephrase the sentence.	We agree that coastal floods are also a consequence of storm surges and that the sentence is not properly formulated. As this sentence is not very relevant for our introduction (which is already quite long), we decided that the whole sentence will be removed from the manuscript.
[3] *Lines 6 (pag 4).* In general terms, the wave-induced component of the water level at the shoreline is the run-up and not the wave height (a different thing is that you approach the run-up with the wave height but this depends on how you calculate it).	This terminological mistake will be corrected to the text where the terms of equation (1) are explained i.e. “wave height” will be changed to “wave run-up”. See also our response to your comment [12.2].
[4] *Line 1 (pag 5).* Long-term mean sea level does not change from decade to decade. Mean sea level is continuously varying and “long-term” refers to the low-frequency component which, apparently, you consider to be associated to periods in the scale of decades (or longer).	We agree that the sentence was poorly formulated. However our purpose in the manuscript is to distinguish between the sea level variations taking place at short time scale (e.g. storm surges) and those that happen slowly within long time span (e.g. mean sea level change). The sentence will be reformulated in a following manner: “The long-term mean sea level on the Finnish coast, on decadal time scale, is affected by the global mean sea level, the post-glacial land uplift and the Baltic Sea water balance (Johansson et al., 2014).”
[5] *Section 3.1. Long-term sea level.* You are using long-term estimations of sea level at selected horizons based on a paper that is under review. If this component is important for your calculations, it can be difficult for some readers to trust on it without having access to the scientific work supporting used values.	The paper we are referring to has now been published. The reference list will be updated accordingly: Pellikka, H., Leijala, U., Johansson, M. M., Leinonen, K., Kahma, K. K., 2018. Future probabilities of coastal floods in Finland. Continental Shelf Research, 157, 32-42. DOI: 10.1016/j.csr.2018.02.006.
[6] *Section 3.2*. Please change the heading. Here you are not describing variability but just the existing data. They are simply water level measurements acquired by using tidal gauges. Use something similar	We agree that the heading was too complicated as the aim of this section is to just describe the tide gauge data. Thus we will change the heading of Section 3.2 simply to “Sea level data”.

to heading of section 3.3 (e.g. tidal data, water level data).	
[7] *Lines 6-7 (Page 8).* Please remove the last sentence "The significant wave height is …". If you want to use a definition of Hs use a formal one (e.g. based on spectral moments).	We will replace the formal definition given on page 8, line 4 with the one using spectral moments. The "layman" definition will also be removed as redundant, as you suggested.
[8] *Lines 23-25 (pag 8)*. See comment [1].	See our response to your comment [1].
[9] *Lines 10-11 (pag 9).* When you explain which sea levels are used to obtain the probability distributions you mention that use sea levels with a given frequency (5 events/year or less in your case). This is equivalent to perform an extreme analysis in which you use a subset of your data composed by extreme events. Then, the usual way should be to select sea-level events by applying the POT method using a given threshold (which will result in a varying number of events per year that, in your case, is up to five events per year) and then fitting the obtained subset by a probability function (exponential in your case).	In the POT method two limits are usually set. One is the threshold (e.g. a certain sea level value), which will give us a certain amount of events per year (on average). The second limit determines the distance between two points. This second limit is set to remove events that are not independent, which enables the final data set to converge to a Pareto distribution. The second limit can be in the order of 24-72 hours, but for sea level data in the Baltic Sea the correlation might be significantly longer (in the order of months). This is because the slow changes in the total water volume in the Baltic Sea. If a proper POT method is applied, the resulting distribution converges to a generalized Pareto distribution (GPD) and will no longer simply be the tail of the original distribution. The main point is the following: in order to use our method, we ultimately need to revert back to the full distribution, since the statistical combination with the wave run-up will otherwise not be possible. If we have fitted a GPD to the data we got by applying the POT method and use that tail to extrapolate the original data, then we are extrapolating the original distribution with a fit that has been made to a different distribution (the GPD). This is obviously something we want to avoid. By fitting the exponential distribution to the tail, we are essentially using the POT method to the extent it is possible in our case. Using the POT method "to its fullest" would change the distribution, since the entire point with the method is to converge the subset of the original data to a GPD. Since we are not only interested in the extreme values, but need the full distribution to combine the sea level data with the wave run-up, the traditional use of the POT method is not a suitable tool for our purposes. See also our response to comment [1] from #1 Reviewer.

[10] *Section 4.3*. You determine an attenuation factor for both coastal locations to derive local wave time-series from 15-year long offshore measurements. This is equivalent to derive an empiric wave propagation model instead of using a numerical model. However, your coastal wave time series are just 31 days long in Jätkäsaari and 11 days long in Länsikari (section 3.3). Given the short-time duration of these records, it is necessary to have more details on this analysis to trust on reconstructed long-term wave time series at both coastal sites. For instance, it should be great to have Hs coastal-Hs offshore plots at both locations under different conditions (T, θ) to see the expected uncertainty in the reconstruction.	First, we want to stress that the attenuated time series we get with the transfer function is not a valid realisation of the wave height time series for the entire 15 year period. The main idea is, that while we can get the typical values (although with a slight positive bias because the measurements were made in the autumn) directly from the measurements, we cannot get the rare exceedances. We therefore determined a transfer function that was adjusted to accurately model the highest values in the measurement time series. When used on the longer open sea wave buoy measurements we can then get information about the nearshore wave height during the more extreme wave events that have happened outside of our short measurement period. Figure RC_C shows that the estimated distribution from the transfer function coincides with the observed values at the tail of the observed distribution. We acknowledge that this is not an optimal way, but it was a practical solution to extract as much information from the existing data set as possible. However, the method presented in the paper is in no way reliant on the method we used to determine the wave height distribution. We will add the information shown in Figure RC_C also to the manuscript, in order to show in more detail how the wave run-up distributions were formed.
[11] *Line 1 (pag 11).* This is a complicated ay to say that you use the maximum run-up, Ru_{max} instead of $Ru_{2\%}$.	Our purpose is to say that we use the maximum run-up and we acknowledge that our way of saying this could be more straightforward. We will rephrase the explanation as follows: ”The final step was to estimate the wave run-up, i.e. the maximum vertical elevation of the water in relation to the still water level. We defined the wave run-up using the highest single wave during an hour, since this will produce one well defined event when combined statistically with the water level data.” This sentence is followed by more detailed explanation of the selected method.
[12.1] *Line 3 (pag 11).* It should be great to include a typical coastal profile of the study sites (maybe after Fig 2) to see how steep they are, especially since you are using this characteristic to approach Ru by H.	We have added a picture of the shoreline at Jätkäsaari (Figure RC_D, from Björkqvist et al., 2017). Although this figure shows the wave damping chambers, the shoreline is similar at other locations

	that are not equipped with wave damping chambers (see Figure 1 in Björkqvist et al., 2017).
[12.2] The concept of *run-up height* needs to be defined to avoid misunderstandings. The run-up height is usually defined as the vertical distance between highest run-up level Ru and deepest run-down Rd. However, when we simply use wave run-up we refer to the vertical distance with respect to the mean water level. Please, clarify what you are using.	We agree that concepts need to be clearly and uniformly defined throughout the paper. In this study we define the "run-up" as the maximum vertical elevation of the water in relation to the still water level during a certain period. We will define this clearly in the manuscript and remove "run-up height" definition to avoid misunderstandings.
[13] *Line 5 (pag 11)*. The use of the relationship Hmax = 2 Hs need to be justified. The ratio Hs/Hmax can be quite variable depending on local conditions (see e.g. Oliveira et al. 2018, Ocean Engineering 153, 10-22). One possibility to select the value to be used is to obtain it from the wave data recorded at your offshore location.	We calculated this based on the wave data recorded at the nearshore location (not the offshore location, since the typical wave periods are much longer there). The mean zero-upcrossing period (Tz, calculated as Tm02 from the spectral moments) are around 3 seconds (3.2 s at Länsikari and 2.8 s at Jätkäsaari). This means about 1200-1300 waves during an hour, which results in Hmax being between 1.9Hs and 2Hs (Figure RC_A). See also our response to the fourth comment from Reviewer #1.
[14] The use of "full" reflection needs to be justified (or simply says that it is arbitrarily selected to be conservative). The study of Björkqvist et al (2017c) used to justify this selection was done in front of a Caisson breakwater. Since we do not know how the coast is (see comment [12.1]), it is difficult to see if the application of this reflection coefficient is appropriated for the site.	See our response to your comment [12.1].
[15] *Section 4.5.* Since you have 15 years of simultaneous data of water level and waves, why you did not convert these series into a single series of total water level (by simple summation) and then to obtain the probability distribution. This can give you a good estimation of the "real" joint probability distribution of water levels (for all components) under current conditions. This could be used to compare with the obtained one by combining individual probability functions.	This is an excellent point, which we would certainly have done if we had the data to do it. However, as addressed in point [10], the wave heights that are estimated do not produce a proper time series, but are used to complete the tail of the distribution based on the measurements (see Figure RC_C). Since the transfer function is constructed with the aim to get the highest tail (not to e.g. minimize the bias), it means that the lower wave heights are overestimated. This is acceptable, since they are not used to construct the distribution.
[16] *Section 5.* It is not clear which is the contribution of this analysis to overall results. If you are just using theoretical distributions, you do not need any data (?). However, for a real case (as it is yours) you should fit a probability distribution (Weibull in your	This is a relevant point. The purpose of the "sensitivity test" is to study how different wave height conditions (based on theoretical wave run-up distributions) affect the total water level when the still water level distribution is kept unchanged. We agree that the contribution of this section to the overall results is not clear, and calling it a "sensitivity

case) and retain the best fit (with the corresponding Weibull parameters). Of course, if you change your Weibull parameters your results will change. You want to include here a sensitivity analysis but, there is no sensitivity analysis (nor uncertainty) associated to your previous selections (Ru formula (H), relationship between Hs and Hmax, refraction model, etc…). If you want to do a formal sensitivity analysis, probably you should account for the different contributions through the entire assessment.	test" is misleading. In order to clarify our aim of the analysis we will remove Section 5 and instead reorganize the "Results" section to include: 1) the case study at Helsinki, 2) the study with these theoretical wave distributions, and 3) a comparison between the results of these.
[17] *Section 7.2. Lines 4-8 (pag 22).* See comment [15].	See our response to your comment [15].
[18] *Lines 9-20 (pag 22).* This is true but this is also less and less common. As it is written, it seems that this is the most used approach. At present, flood assessments for combined water level-wave contributions, usually consider full time series instead of monthly maxima.	We acknowledge that we generalized unnecessarily the use of block maxima. We will rephrase the sentence: "Using block maxima of sea level variations — such as the monthly maxima used by Pellikka et al. (2018) — in our analysis would implicitly restrict the study of the joint effect to cases where the still water level is high, thus excluding combinations of moderate still water level and high waves."
[19] *Lines 21 to 25 (pag 22).* More than the short-term variability in waves, probably, you must also consider the potential long-term variability in wave conditions for long time projections (see e.g. Méndez et al. 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model." *JGR Oceans* 111,C7).	Using the verified wave model data from Björkqvist et al. (2018) we calculated the mean significant wave height at the GoF wave buoy for the years 1965-2005 (the hindcast cannot resolve the nearshore conditions of Länsikari and Jätkäsaari). The results are shown in Figure RC_E for both ice-free statistics and ice-included (as Hs=0) statistics. In both statistics the trend is small, and not statistically significant according to a t-test. This is supported by Kudryavtseva and Soomere (2017). The authors used satellite altimetry data (1996-2015) and found no statistically significant trend in the Gulf of Finland. Of course, the absence of evidence is not evidence of absence, but using the current knowledge we have no means to predict the future changes of the significant wave height in the GoF. We have mentioned using long-term scenarios for wave conditions in the Discussion part of the manuscript, as a potential improvement on our method in future studies.

	Méndez et al. (2006) used a POT-method where the coefficients of the GDP-distribution where allowed to vary in time. The time varying parameters can capture some of the seasonal variability that is lost if the POT-method is used on the entire data set with only one set of parameters. However, since the method implemented in this paper uses the full distributions, all seasonal variations are already present in the data, and no special methods are required to account for them.

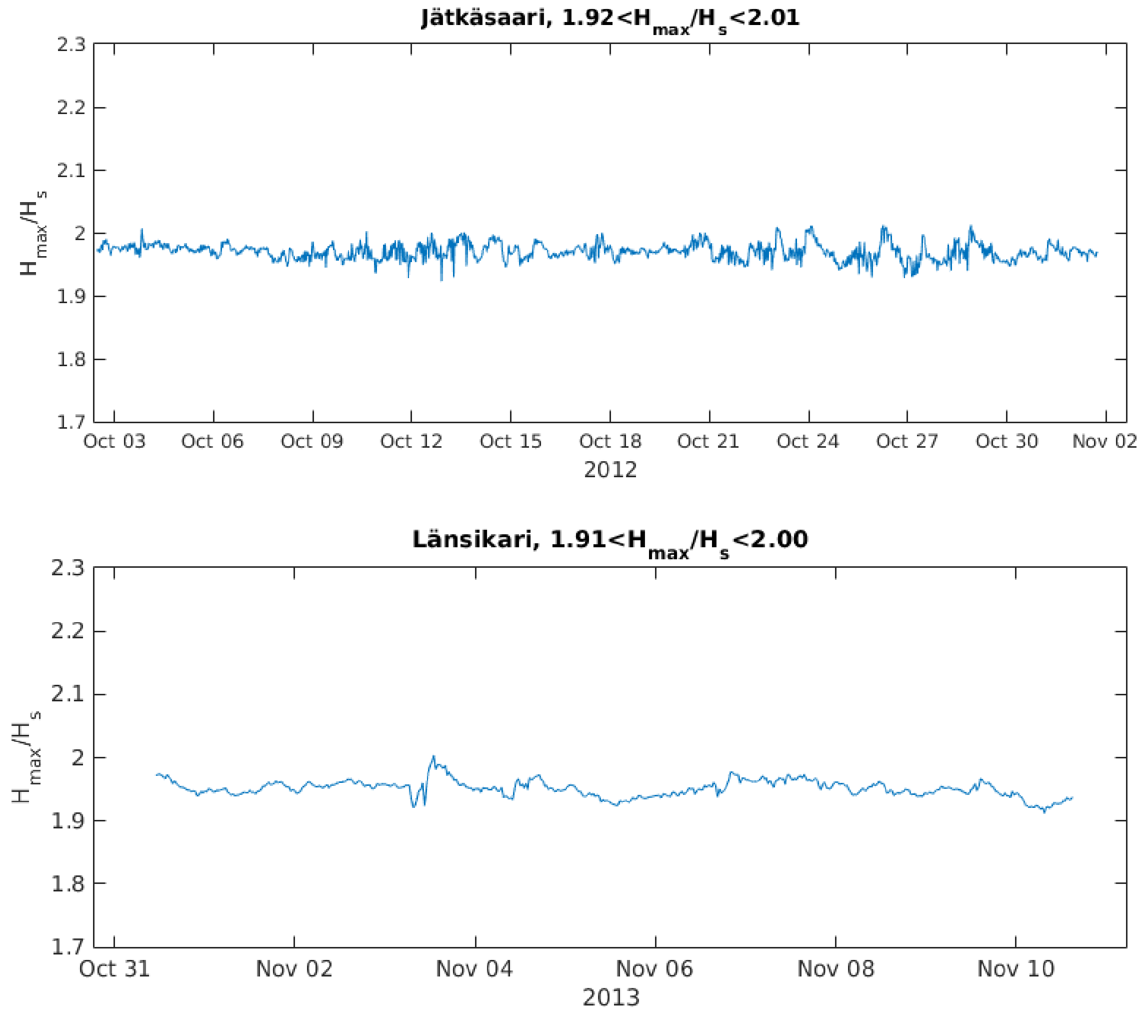

Figure RC_A. The ratio between the highest single wave and the significant wave height estimated from the Rayleigh distribution at Jätkäsaari and Länsikari.

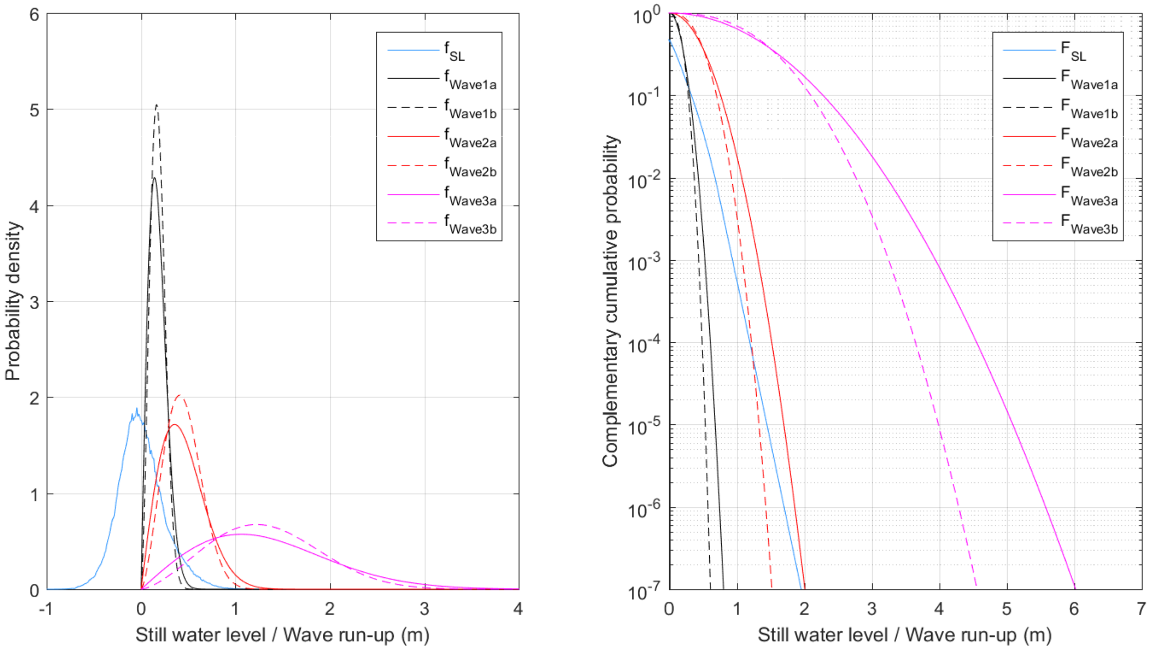

Figure RC_B. Pdfs (on the left) and ccdfs (on the right) for the still water level and the six theoretical wave run-up distributions.

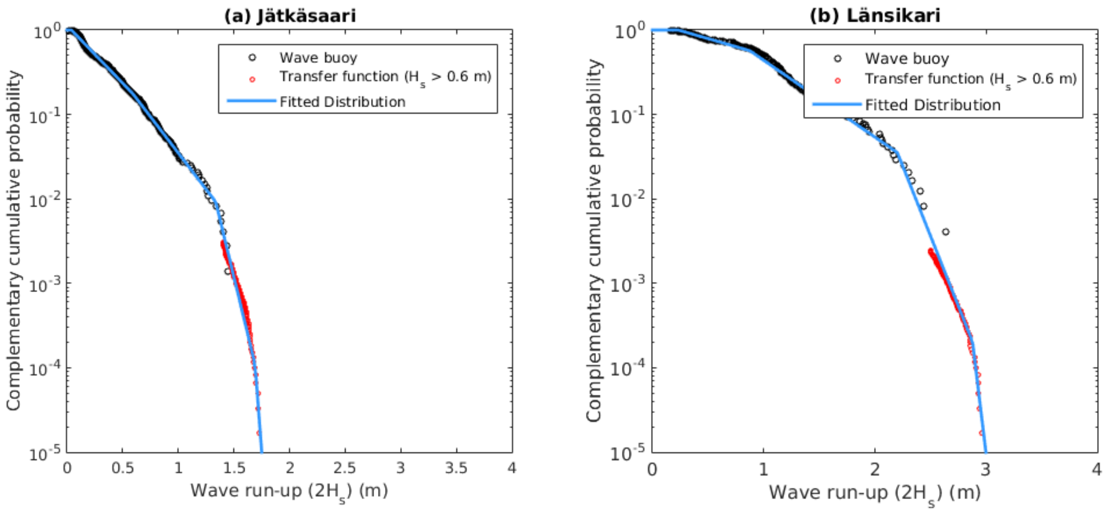

Figure RC_C. Wave run-up distributions for the two locations in the Helsinki archipelago: Jätkäsaari and Länsikari.

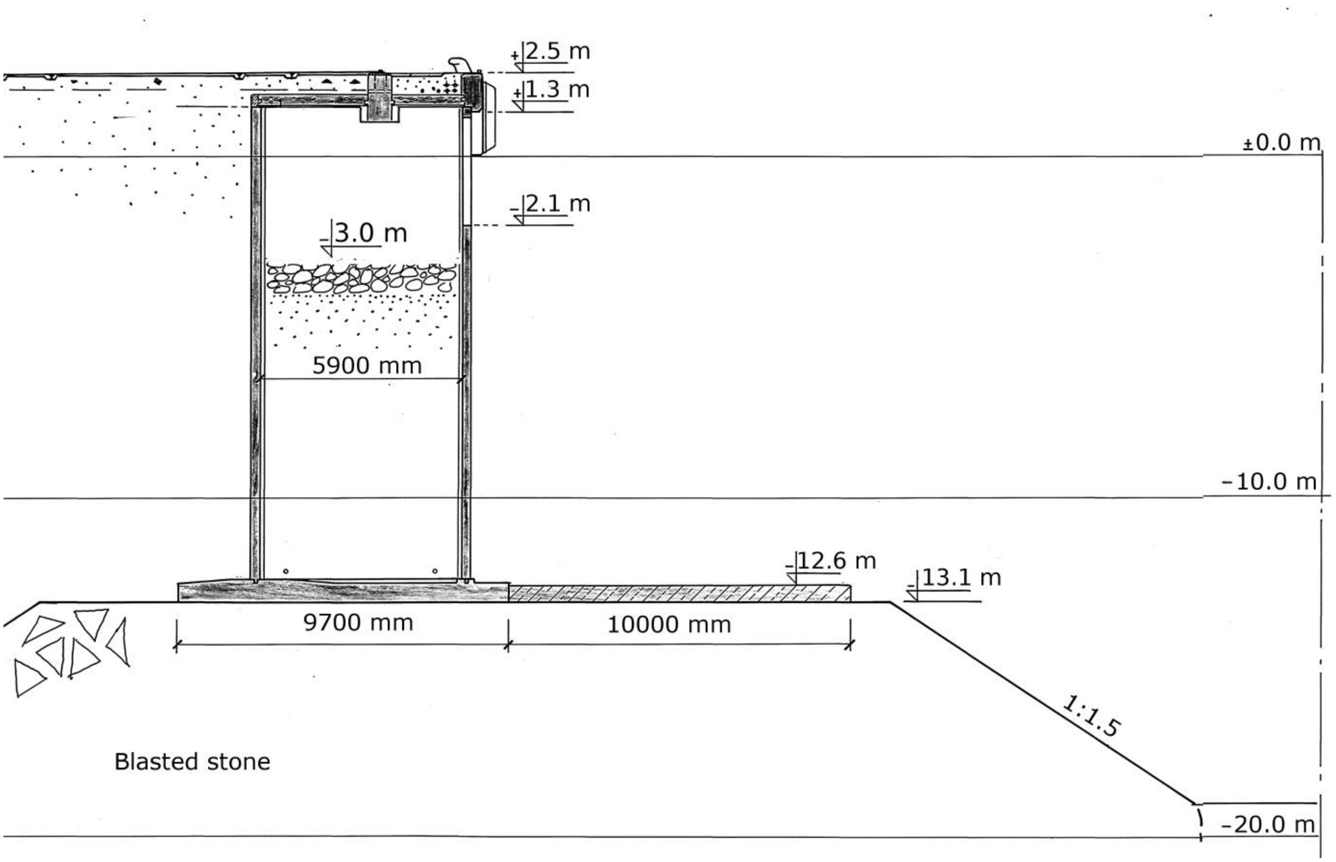

Figure RC_D. The shoreline at Jätkäsaari (from Björkqvist et al., 2017). Other parts of the shoreline are of similar shape (vertical walls), but are not equipped with wave damping chambers.

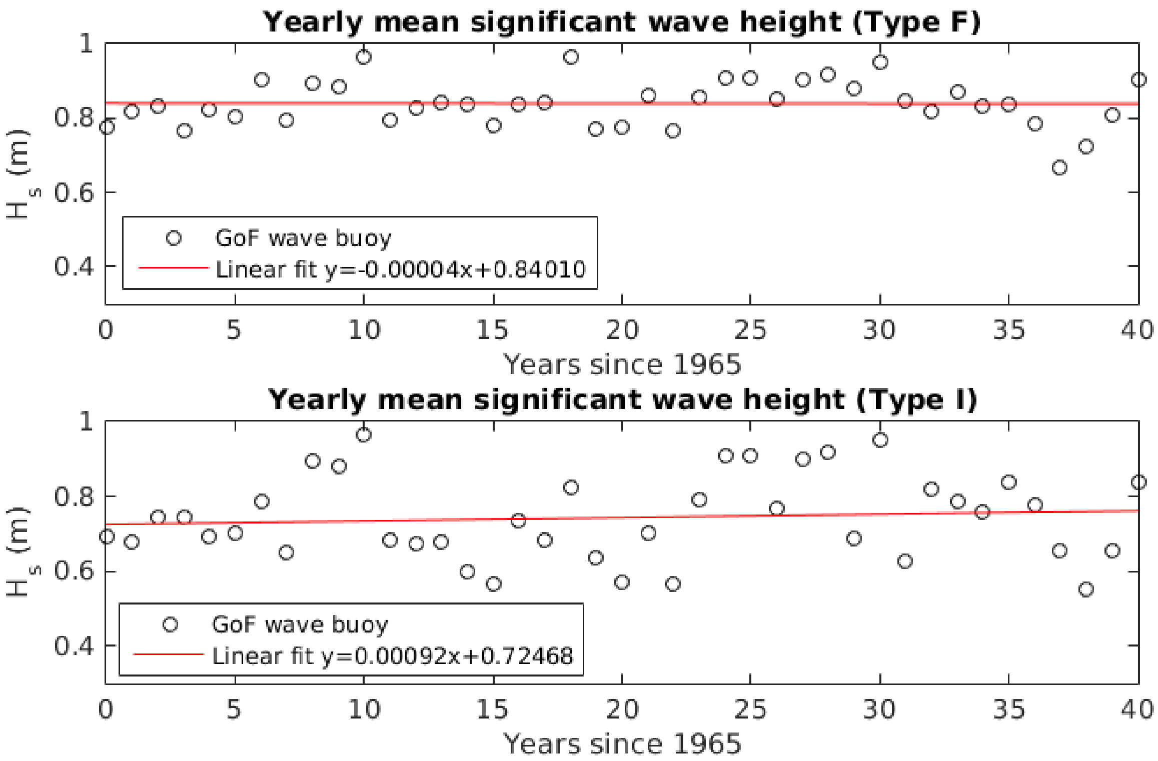

Figure RC_E. The yearly significant wave height at the Gulf of Finland wave buoy taken from the wave hindcast of Björkqvist et al. (2018). Trends were calculated for both the ice-free statistics and the ice-included statistics. Neither was statistically significant.

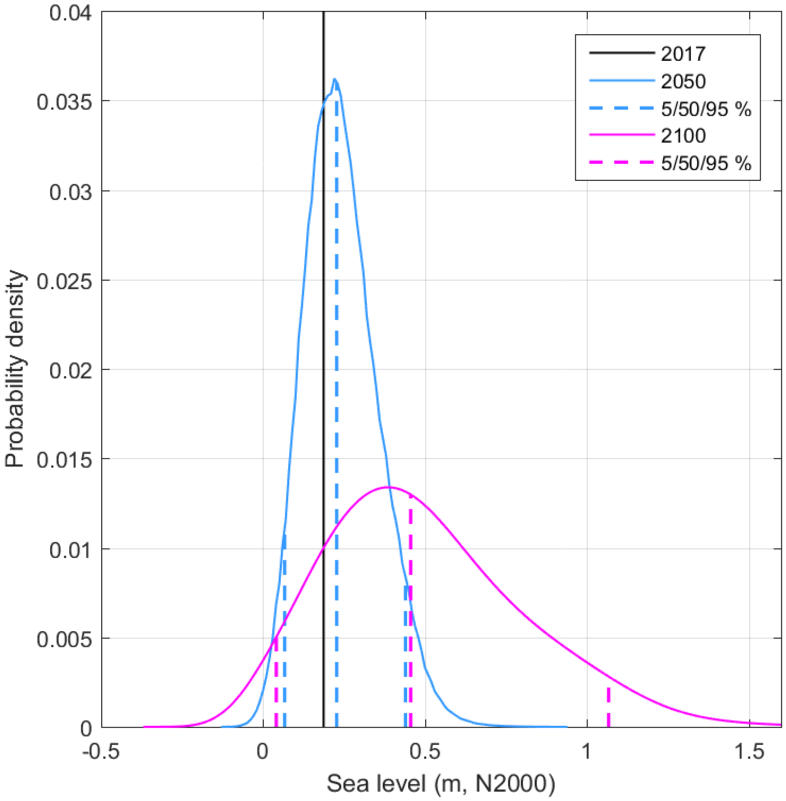

Figure RC_F. Probability density functions of future mean sea level at the Helsinki tide gauge for years 2050 and 2100 and the long-term mean sea level estimate of 0.19 m for year 2017. The 5th, 50th and 95th percentiles are shown for 2050 and 2100. The data in the Figure is from the results of Pellikka et al. (2018).