Manuscript number: NHESS-2020-379

My co-authors and I would like to express our gratitude to the reviewers for their constructive feedback and suggestions for strengthening our research. The changes we have made to the attached file in response to such feedback and suggestions have been highlighted in blue to facilitate their identification. I would also like to offer my apologies for the length of time it took us to prepare this response.

Referee #4

1. Abstract: Since this is a scientific paper, it is recommended to summarize the research methodology and conclusions (major outcomes) in more detail.

- We are very grateful to the reviewer for this important advice. Our revisions in response to the above comments can be found in the revised Abstract, which can also be seen below.

Global warming, one of the most serious aspects of climate change, can be expected to cause rising sea levels. These, in turn, have been linked to unprecedentedly large typhoons that can cause flooding of low-lying land, coastal invasion, seawater flows into rivers and groundwater, rising river levels, and aberrant tides. To prevent typhoon-related loss of life and property damage, it is crucial to accurately estimate storm-surge risk. This study therefore develops a statistical model for estimating such surges' probability, based on surge data pertaining to Typhoon Maemi, which struck South Korea in 2003. Specifically, estimation of non-exceedance probability models of the typhoon-related storm surge was achieved via clustered separated peaks-over-threshold simulation, while various distribution models were fitted to the empirical data for investigating the risk of storm surges reaching particular heights. To explore the non-exceedance probability of extreme storm surges caused by typhoons, a threshold algorithm with clustering methodology was applied. To enhance the accuracy of such non-exceedance probability, the surge data was separated into three different components: predicted water level, observed water level, and surge. Sea-level data from when Typhoon Maemi struck was collected from a tidal gauge station in the City of Busan, which is vulnerable to typhoon-related disasters due to its geographical characteristics. Fréchet, Gamma, log-normal, Generalised Pareto, and Weibull distributions were fitted to the empirical surge data, and the researchers compared each one's performance at explaining the non-exceedance probability. This established that Weibull distribution was better than any of the other distributions for modeling Typhoon Maemi's peak total water level. Although this research was limited to one city in the Korean Peninsula and one extreme weather event, its approach could be used to reliably estimate non-exceedance probabilities in other regions where tidal gauge data are available. In practical terms, the findings of this study, and future ones adopting its methodology, will provide a useful reference for designers of coastal infrastructure.

2. Line 43: I can't find this paper (Hwang and Deodatis 2013) in your reference. Check please.

- We are grateful to the reviewer for pointing out that this was missing. We have now added it to the References list.

3. Line 71: I recommend that the names of Sea of Japan and East Sea should be written as "East/Japan Sea" together.

- We are grateful to the reviewer for this helpful suggestion, which has been adopted.

4. Line 85: In your Table 2 & 3, Fill the last column in line1 -> Total Sum & Average in line2-3 -> Incidence(sum) & Incidence(average)

- We thank the reviewer for this comment. The manuscript has been revised accordingly, as shown below.

Table 2. Incidence of typhoons and typhoon landfall in South Korea, 1952-2019, by month

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
Typhoons, *n*	29	15	15	45	67	115	245	351	322	238	152	73	1,678
Typhoons, avg.	0.54	0.28	0.46	0.83	1.24	2.13	4.54	6.52	5.96	4.41	2.81	1.35	31.07
Incidence (sum)	0	0	0	0	1	18	65	70	45	5	0	0	206
Incidence (average)	0.0	0.0	0.0	0.0	0.02	0.33	1.2	1.3	0.87	0.09	0.0	0.0	3.81

Table 3. Incidence of typhoons and typhoon landfall in South Korea, 2010-19, by month

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
Typhoons, *n*	4	3	4	5	12	18	33	43	56	34	16	7	235
Typhoons, avg.	0.4	0.3	0.4	0.5	1.2	1.8	3.3	4.3	5.6	3.4	1.6	0.7	23.5
Incidence (sum)	0	0	0	0	0	0	3	11	7	5	2	0	28
Incidence (average)	0	0	0	0	0	0	0.3	1.1	0.7	0.5	0.2	0	2.8

5. Line 94-96: South Korea operated 17 tidal gauge stations. -> South Korea operated 46 tidal gauge stations. And (n=5), (n=10), (n=2) ~. Check please.

- Pursuant to this reviewer comment, the following supplementary sentences were added.

When Typhoon Maemi struck the Korean Peninsula in 2003, South Korea was operating 17 tidal-gauge stations, of which eight had been collecting data for 30 years or more. They were located on the western (n=5), southern (n=10), and eastern coasts (n=2).

6. Line 99: These expressions are already expressed in Figure 1 and are redundant, so please remove them.
- As recommended, the redundant expressions were removed.

7. Line 105: Authors did not to describe about QC processes (i.e. interpolation, outliers etc) of tidal gauge data, cause tidal stations have lots of errors in raw data. Rising rate of MSL can be varied considerably according to the OC processes.
- We are grateful to the reviewer for pointing this out. Accordingly, a deeper description of statistical analysis for QC processes has been added, and can also be seen below.

KHOA makes hourly observations of water height at the Busan tidal-gauge station, and the annual means presented in this paper have been calculated from that hourly data. As can be seen in Figure 3, plotting MSL for each year confirms that short-term water-level variation merely masks the long-term trend of sea-level increase. Therefore, on the assumption that MSL variation was a function of time, a linear regression was performed, with the resulting coefficient of slope indicating the rate of increase (Yoon and Kim, 2012). The data utilised to estimate MSL for the tidal gauge station in Busan was provided by KHOA, which performed quality control on it before releasing it to us. Additionally, however, a normality test was performed, and the results (as shown in Table 7) indicated that the hourly sea-level data followed a normal distribution, at a significance >0.05. The Kolmogorov-Smirnov normality test was adopted as being well-suited to datasets containing more than 30 items.

Table 7. Kolmogorov-Smirnov normality test of sea-level fluctuation data from Busan tidal-gauge station

	Statistic	df	Significance	Pearson correlation
Sea level fluctuation	0.084	473352	0.200	0.96

As can be seen in Figure 3, the average rate of increase in MSL at Busan's tidal-gauge station from 1962 to 2019 was 2.4mm per year, yielding a difference of 16.31cm between the end of that period and the beginning. This finding is broadly in line with Yoon and Kim's (2012), that the rate of MSL increase around the Korean Peninsula as a whole between 1960 and 2010 was about 2.9mm/year. Also, linear regression analysis of the sea-level fluctuation data for 1965-2019 was utilised to discern the MSL trend. The significance level of 0.000 (<0.05) obtained via analysis of variance

(ANOVA; Table 8) indicates that the regression model of sea-level fluctuations was significant. Also, its correlation coefficient (0.96) indicated a strong positive relationship between sea-level rise and recentness. The coefficient of determination (R2) was utilised to describe how well the model explained the collected data. The closer R2 is to 1, the better the model can predict the linear trend; and here, it was 0.74, as shown in Table 9. This means that the linear-regression model explained 74% of the sea-level variation. While this result suggests that the linear-regression analysis for sea-level fluctuation at the tidal gauge station in Busan is reliable, however, such results may not be generalisable because variation in the data could have been due to several factors, including geological variation and modification of gauge points.

Table 8. Linear-regression coefficients, sea-level fluctuations at Busan tidal-gauge station

	Non-standardised Coefficients		Standardised Coefficients	t	Significance Probability (P-value)
	B	Standard Error	Beta		
(Constant)	-422.23	35.022	0.887	-12.06	0.00
Sea level fluctuations	0.246	0.018		13.97	0.00

Table 9. Summary of analysis of variance results, sea-level fluctuations at Busan tidal-gauge station

Model	Sum of Squares	df	Mean squares	F	Sig.	Adjusted R2
Regression	830446354.04	41	20254789.12	32109.38	.000b	0.74
Residual	298566787.86	473310	630.81			
Total	1129013141.90	473351				

8. Line 106: Highest recorded water levels -> Recorded highest water levels

- We are grateful to the reviewer for pointing out this anomaly. The usual word order, “highest recorded”, is now used consistently.

9. Line 110: Table 5 is presented but not cited in the text.

Line 115: In Table 5, New Busan -> New Busan Port

- Pursuant to this comment, Table 5 is now mentioned in the main body of the manuscript, as follows.

The same approach was applied to the data from the 10 tidal-gauge stations on the south coast, as shown in Table 5 and Table 6, below.

10. Line 129: Figure 4 and 5 -> Figure 3 and 4
- As recommended, the figure numbers have been corrected.

11. Line 130: I think it would be appropriate to delete Figure 3 because it is not important in the context described in the paper.
- As recommended, this figure has been removed.

12. Line 132: In Figure 3 & 4, Sea level (cm) -> Mean sea level (cm)

    Line 136: Figure 4. Sea-level fluctuations around the mean, Busan -> Figure 4. Mean Sea-level fluctuations in Busan

- Pursuant to this reviewer comment, the wording of these captions has been standardized to “Mean Sea-level fluctuations”.

13. Line 139: "Looking at the sea-level history in Figure 3, it is clear that the data trend between 1956 and 1961 is anomalous. As this may have been due to quality-control issues with the observations from that period, it has been excluded from this study, and only data from 1962 to 2019 have been used, as shown in Figure 4." -> If you accept the recommendation, please delete this context.
- We are grateful for these constructive comments. As recommended, the 1956-61 data have been removed from analysis, and the text modified accordingly.

14. Line 154: In Table 7, New Busan -> New Busan Port
- We thank for your comment. The term “Port of New Busan” is now used consistently throughout.

15. Line 155: As can be seen from Table 7
- We thank the reviewer for pointing this out. The table number has now been corrected.

16. Line 185: I can’t find this paper (Lee et al. 2008) in your reference. Check please.
- We are grateful to the reviewer for pointing this out. The missing reference has now been added to the References list.

17. Line 201: description on Previous research results

    Line 231: Introduction on return period of Hurricane using GEV, MLE, CSPS

- We are grateful for this insightful suggestion. As requested, we have added reviews of the relevant studies, as shown below.

Bermúdez et al. (2019) studied flood drivers in coastal and riverine areas as part of their approach to quantifying flood hazards, using 2D shallow-water models to compute the correlation between extreme events and flood drivers. They also adopted ordinary least-squares regression analysis to construct a 10,000-year time series, and computed water levels' exceedance probabilities for comparison. However, the possibility of river discharges, sea-wave trends, and tidal fluctuations were not considered in their study.

The wrecking of windfarms by extreme windstorms is of considerable concern in the North Sea region, which is home to 38 such farms belonging to five different countries. According to Buchana and McSharry's (2019) Monte Carlo simulation-based risk-management study, the total asset value of these windfarms is €35 billion. It used a log-logistic damage function and Weibull probability distribution to assess the risks posed to windfarms in that region by extreme strong winds, and exceedance probability to predict the extent of financial loss from such damage, in terms of solvency capital requirement (SCR). The same study also simulated the results of various climate-change scenarios, and the results confirmed that higher wind speed and higher storm frequency were correlated with rises in SCR: a finding that could be expected to help emergency planners, investors, and insurers reduce their asset losses.

According to Catalano et al.'s (2019) study of high-impact extratropical cyclones (ETCs) on the north-eastern coast of the Unites States, limited data caused by these storms' rarity made it difficult to predict the damage they would cause, or analyse their frequency. To overcome this, they utilised 1,505 years' worth of simulations derived from a long-coupled model, GFDL FLOR, to estimate these extreme events' exceedance probabilities, and compared the results against those of short-term time-series estimation. This revealed not only that the former was more useful for statistical analysis of ETCs' key characteristics – which they defined as maximum wind speed, lowest pressure, and surge height – but also that the use of a short time-series risked biasing estimates of ETCs' return levels upwards (i.e., underestimating their actual frequency). While these results regarding return levels and time-series were valuable, however, Catalano et al. did not distinguish between the cold season and the warm season of each year, which could also have led to biased results.

A joint-probability methodology was used to analyse extreme water heights and surges on China's coast by Chen et al. (2019). They obtained the sea-level data from nine gauge stations, and utilised 35 years' worth of simulation data with Gumbell distribution and Gumbell-Hougaard copula. The three major sampling methods proposed in the study were structural-response, wave-dominated, and surge-dominated sampling. The first was utilised to assess structures' performance in response to waves and surges. Joint-probability analysis revealed that such performance was correlated with extreme weather events in the target region, and that such correlation became closer when wave motion was stronger. Also, based on their finding that joint exceedance probability tended to overestimate return periods for certain water levels, Chen et al. recommended that offshore defence-facility designers use joint-probability density to estimate return levels of extreme wave heights. Yet, while their study provided a useful methodology, particularly with regard to sampling methods and probability modelling of return periods and structural performance, they only looked at China's coast, and therefore their findings are unlikely to be generalizable to the Korean Peninsula.

Davies et al. (2017) proposed a framework for probability modelling of coastal storm surges, especially during non-stationary extreme storms, and tested it using the El Niño-Southern Oscillation (ENSO) on the east coast of Australia. Importantly, they applied their framework to ENSO and seasonality separately. This is because, while ENSO affects storm-wave direction, mean sea level, and storm frequency, seasonality is mostly related to storm-surge height, storm-surge duration, and total water height. This separation has the advantage of allowing all storm variables of non-stationary events to be modelled, regardless of their marginal distribution. Specifically, Davies et al. applied non-parametric distribution to storm-wave direction and steepness, and parametric distribution to duration and surge using mixture-generalised extreme value probability modelling, which they argued was more useful than standard ones such as Generalized Pareto Distribution (GPD). This, they said, was because the statistical threshold in an extreme mixture model can be integrated into the analysis, whereas a GPD model should be given an unbiased threshold: if it is low, too many normal data may be included. Accordingly, they utilised bootstrapping for the confidence interval to show the uncertainty of the non-stationary aspects of the extreme events. Also, they added a Bayesian method to provide wider confidence intervals with less bias. Their findings are mainly beneficial to overcoming the challenges of GPD threshold selection; however, robust testing of their approach will require that it be applied to a wider range of abnormal climate phenomena.

Similar research was conducted by Fawcett and Walshaw (2016), who developed a methodology for estimating the return levels of extreme events such as sea surges and high winds of particular speeds, with the wider aim of informing practical applications such as design codes for coastal structures. They reported that two of the most popular existing methods for doing so, block maxima (BM) and POT, both have shortcomings, and concluded that a Bayesian approach would be more accurate. Specifically, they argued that BM and POT methods tend to waste valuable data, and that considering all exceedance via accurate estimation of the extremal index (reflecting uncertainty's natural behaviour) could compensate for this disadvantage. They further proposed the seasonal variations should be taken into consideration with the all exceedance data, where possible.

In response to Japanese government interest in unexpected flooding caused by extreme storm surges during typhoons and other high-wind events, Hisamatsu et al. (2020) simulated typhoons as a means of predicting the cost of the damage they would cause in Tokyo Bay, which is very vulnerable to such events due to its geographic and socio-economic characteristics. Using stochastic approaches, they modelled future typhoons over a 10,000-year period, and calculated flooding using a numerical surge model based on the probability of historical typhoons. These flooding calculations, in turn, were utilised to create a storm-surge inundation map, representing exceedance probabilities derived from stochastic hazard calculations pertaining to 1,000 typhoons. Next, the completed map was overlaid on government-provided values of Tokyo Bay's buildings and other infrastructural elements, to assess the spatial extent and distribution of the likely damage. The results showed that Chiba and Kanagawa would be the most damaged areas, and suffer financial losses of ¥158.4 billion and ¥91.5 billion, respectively, with an exceedance probability of 0.005 (as commonly used to estimate damage in the insurance industry). However, the real-estate values they used were two decades out of date at the time their study was conducted, meaning that further validation of their approach will be needed.

Another effort to estimate return periods was made by McInnes et al. (2016), who created a stochastic dataset on all cyclones that occurred near Samoa from 1969 to 2009. That dataset was utilized to model storm tides using an analytic cyclone model and a hydrodynamic model, which also took account of prevailing climate phenomena such as La Niña and El Niño when estimating return periods. The authors found that tropical cyclones' tracks could be affected by La Niña and El Niño, and more specifically, that the frequency of cyclones and storm tides during El Niño was consistent across all seasons, whereas La Niña conditions make their frequency considerably lower in La Niña season. Additionally, McInnes et al. proposed that sea-level rises had a more significant influence on storm tides than on future tropical cyclones did, based on their finding that future cyclones' frequency would be reduced as the intensity of future cyclones increased. Lastly, they found that the likelihood of a storm tide exceeding a 1% annual exceedance probability (i.e., a one-in-a-hundred year tide) was 6% along the entire coastline of Samoa. However, other effects such as sea level fluctuations and meteorological factors were not included in their calculations.

Silva-González et al. (2017) studied threshold estimation for analysis of extreme wave heights in the Gulf of Mexico, and argued that appropriate thresholds for this purpose should consider exceedances. They applied the Hill estimator method, an automated threshold-selection method, and the square-error method for threshold estimation in hydrological, coastal engineering, and financial scenarios with very limited data, and found that the square-error method had the most advantages, because it did not consider any prior parameters that could affect thresholds. The authors went on to propose improvements to that method, i.e., the addition of differences between quantiles of the observed samples and median quantiles from GPD-aided simulation. When GPD was utilised to estimate observed samples, it effectively prevented convergence problems with the maximum-likelihood method when only small amounts of data were available. The key advantage of Silva-González et al.'s approach is that the choice of a threshold can be made without reliance on any subjective criteria. Additionally, no particular choice of marginal probability distribution is required to estimate a threshold. However, to be of practical value, their method will need to incorporate more meteorological factors.

Lastly, Wahl et al.'s (2015) study of the exceedance probabilities of a large number of synthetic and a small number of actual storm-surge scenarios utilized four steps: parameterising the observed data; fitting different distribution models to the time series; Monte Carlo simulation; and recreating synthetic storm-surge scenarios. Specifically, projected 40cm and 80cm sea-level rises were used as the basis for investigating the effects of climate change on flooding in northern Germany. Realistic joint-exceedance probabilities were used for all parameters with copula models; and the exceedance probabilities of storm surges were obtained from the bivariate exceedance probability method with two parameters, i.e., the highest total water level with the tidal fluctuations and intensity. Wahl et al.'s findings indicated that extremely high water levels would cause substantial damage over a short time period, whereas relatively small storm surges could inflict similar levels of damage but over a much longer period. However, like various other studies cited above, Wahl et al.'s did not take account of seasonal variation.

18. Line 268: I can't find this paper (Pickands, 1975 & Scarrott and Macdonald 2012) in your reference. Check please.

- We are grateful to the reviewer for noticing this. The missing references have been added to the References list.

19. Line 333: Description on the surge height caused by typhoon Maemi.

- As recommended, we have added more description, which can be seen below.

The Korean Peninsula is bounded by three distinct sea-systems, generally known in English as the Yellow Sea, the Korea Strait, and the East Sea / Sea of Japan. This characteristic has often led to severe damage to its coastal regions. According to the Korea Ocean Observing and Forecasting System (KOOFS), Typhoon Maemi in September 2003 had a maximum wind speed of 54 metres per second (m/s), and these strong gusts caused an unexpected storm surge. This event caused US$3.5 billion in property damage, as shown in Table 1. All three of the highest peaks ever recorded by South Korea's tidal-gauge stations also occurred in that month.

Table 1. Largest typhoons to have struck the Korean Peninsula

Name	Date	Amount of Damage (US$)	Max. Wind Speed (10 min. avg., m/s)

Rusa	30 Aug.-1 Sep. 2002	4.3 billion (1st)	41
Maemi	12-13 Sep. 2003	3.5 billion (2nd)	54
Bolaven	25-30 Aug. 2012	0.9 billion (3rd)	53

The most typhoon-heavy month in South Korea is August, followed by July and September, with two-thirds of all typhoons occurring in July and August. Tables 2 and 3, below, present statistics about typhoons in South Korea over periods of 68 years and 10 years ending in 2019, respectively; and Figure 1 shows the track of Typhoon Maemi from 4-16 September 2003. As can be seen from Figure 1, Typhoon Maemi passed into Busan from the southeast, causing direct damage upon landfall, after which its maximum 10-minute sustained wind speed was 54 m/s. Typhoon Maemi prompted the insurance industry, the South Korean government, and many academic researchers to recognise the importance of advance planning and preparations for such storms, as well as for other types of natural disasters.

Because observed sea level usually differs from predicted sea level, Figure 5 depicts the former (as calculated through harmonic analysis) in blue. Predicted sea levels are shown in green, and surge height in red. As the figure indicates, the highest overall water level coincided with the highest surge during Typhoon Maemi, i.e., at 21:00 on 12 September 2003. Given a total water height of 211cm, the surge height was calculated as 73.35cm. The unexpectedly large height of the surge induced by Typhoon Maemi caused US$3.5 billion in property damage and many causalities in Busan, as mentioned in Table 1.

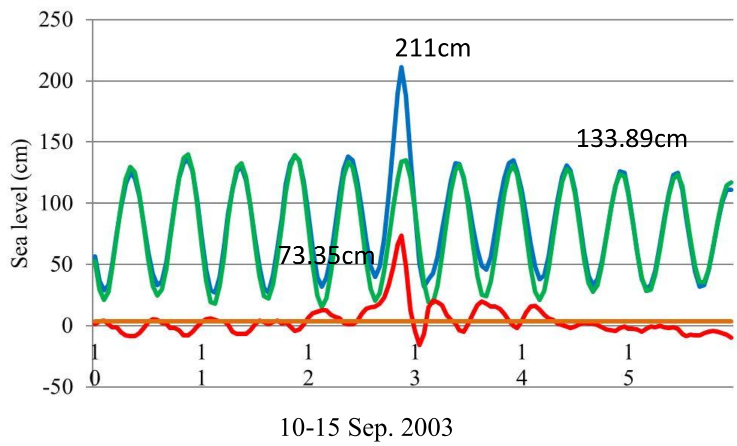

Figure 5. Observed (green), predicted (blue), and residual (red) water levels at Busan during Typhoon Maemi

20. Line 339: I can’t find where is the Figure 7 in the text.

- The original Figure 7 (“Threshold-selection flowchart”) has now been renumbered as Figure 6, both in its caption and in references to it in the main text.

21. Although the return period is importantly mentioned at the beginning, it is not presented in the results, and that only mentioning that further research is needed may lower the quality and justification of the paper, so additional supplementation is required.

- We are grateful to the reviewer for pointing this out. The following passage has accordingly been added:

Although the present research investigated various non-exceedance probability distributions of typhoon-driven storm surges, it only used a single extreme event in a specified region. As such, its findings may not be applicable to other regions, each of which has its own unique weather conditions, geographic features, and tidal characteristics. Future research should therefore include tidal and environmental data from a range of different regions and various extreme events to test the present study’s findings. Also, various natural-hazards indicators and environmental factors such as wind speed, pressure, rainfall, landslides, distance to waterways, and so forth may be useful variables in estimating the exceedance probabilities of typhoons and other natural hazards, and thus be beneficial to risk assessment and mitigation. Also, it should be borne in mind that much of the tidal-gauge data that this study utilised was from the fairly distant past. Thus, in similar future studies, efforts should be made to ensure that such data are reliable, especially in light of climate-change-driven patterns in sea-level behaviour.

Return periods based on various non-exceedance probability models should also be considered in future research, insofar as elaborated return-period estimation can be utilised to improve disaster-relief and emergency-planning efforts. Our comparison of various probability models to find the best fitting distribution models could be adapted to the simulation of time series of the past typhoons, and the collected simulated storm-surge time series then used to estimate typhoons’ return periods using bootstrapping of the exceedance data. Potentially, this would provide more exact return periods with confidence intervals. Lastly, future work on return periods should take account of trends in sea-level change, driven by climate change, which already pose a non-negligible risk to coastal buildings and other infrastructure. Advanced statistical methods such as Monte Carlo simulation, as well as deep-learning techniques, could be applied to make typhoon return-period estimates even more accurate.