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 #2

I read with interest your manuscript titled: "Identifying the non-exceedance probability of extreme storm surges as a component of natural-disaster management using tidal-gauge data from Typhoon Maemi in South Korea". The manuscript introduces a novel methodology for deriving non-exceedance probability diagrams of extreme surge storms. Clustered separated peaks-over-threshold simulation was developed, and various probability density function models were fitted to the empirical data for investigating the risk of storm surge height. Weibull probability density distribution was found to fit the empirical data. This manuscript introduces a novel simulation method for derivation of exceedance diagrams of storm surges that can contribute to many other natural hazards phenomena such as floods, forest fires, etc. The paper deserves minor revisions as follows:

1. The title of the manuscript is too long consider shorter title such as: "Nonexceedance probability of extreme storm surges using tidal-gauge".

- We appreciate this insightful comment, and as recommended, have modified the title of the manuscript to better reflect its content.

Estimation of the non-exceedance probability of extreme storm surges in South Korea using tidal-gauge data

2. The abstract does not reflect the novelty of the methodology. Consider revision.

- We are very grateful to the reviewer for providing this important advice. Our revisions in response to the above comments can be found in the revised Abstract. It 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.

3. Please add a research framework diagram, emphasize the core phases of the methodology, particularly: the Threshold selection iterative process and the clustering of storm surge data.

- We are grateful for this insightful suggestion. Our general approach, workflow, threshold selection interactive process, and clustering regarding estimating non-exceedance probability of extreme surges using tidal gauge data. These can be seen below.

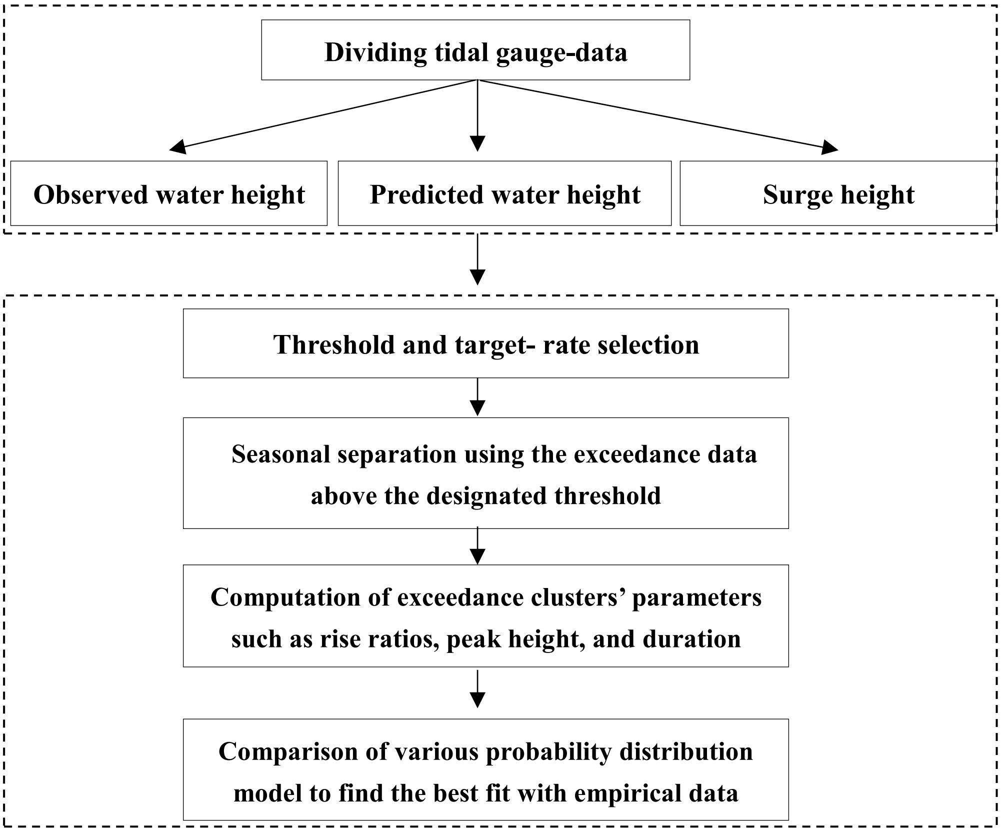

Figure 4. General approach and workflow

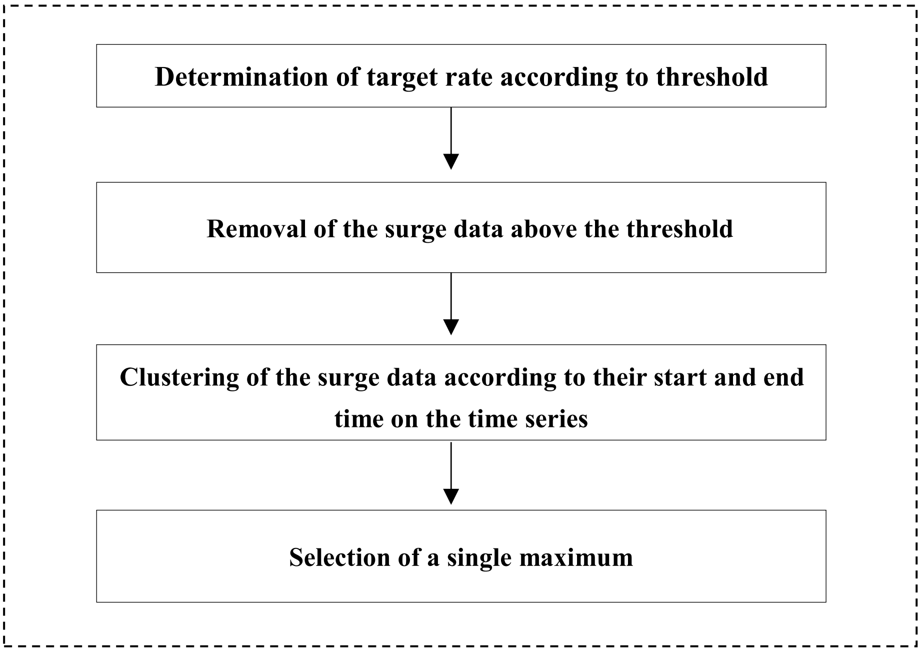

Figure 11. Clustering flowchart

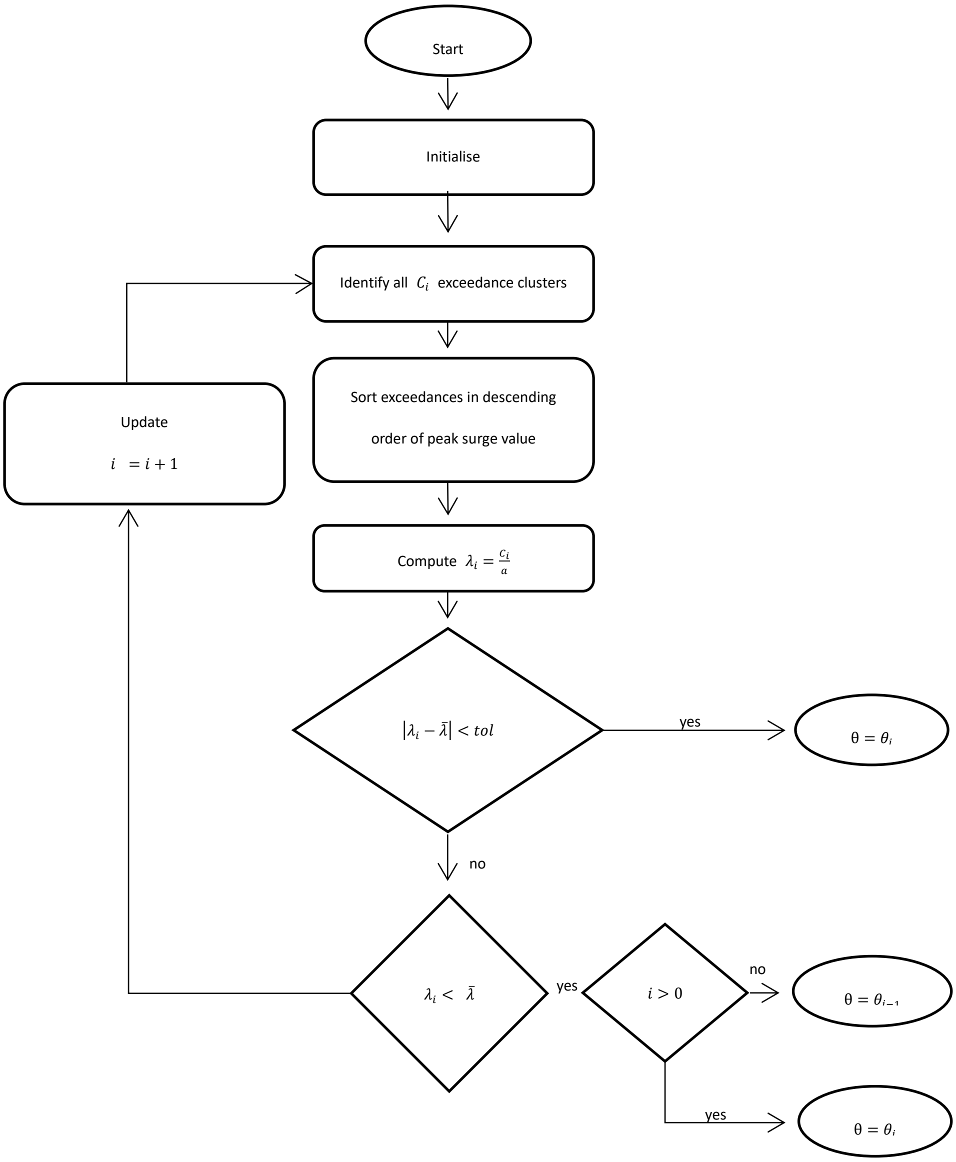

Figure 6. Threshold-selection flowchart

4. The statistics analysis presentation: it is suggested to add Analysis of Variance output data for all regression and probabilistic distribution goodness of fit as follows:
   1) Please provide Coefficient of determination and correlation coefficient, regression variance and Standard Error (SE) of the sea level fluctuations in Figure 4. Perhaps add a Table
   2) Please provide detailed statistical data of the different probabilistic distributions I Figures 17-19.
   3) Please present the parameters of the selected distribution (Weibull) and discuss this with reference to the literature.
   4) Please discuss the statistical significance of the model (P=Value, R2, C.I. P.I.).

- Thank you for your comment. Pursuant to this reviewer comment, the deeper description of statistical analysis was discussed. These can 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				

Based on our simulations, exceedances water height above the designated threshold were computed using MLE estimates. Table 11 presents the distribution parameters of the storm-surge parameters that were computed, each using a different probability model. These distribution parameters were based on the exceedance above the algorithmically designated threshold of 29.15cm, mentioned above.

Table 11. Probability-distribution parameters of the storm-surge parameters

	GPD		Beta	
Season	ξ	σ	α	β
Cold	0.02	0.34	3.12	3.45

Warm	0.51	0.33	2.87	1.89

We also fit our empirical data to five different probability-distribution models – i.e., Fréchet, Gamma, GPD, Lognormal, and Weibull – as seen in Figures 19 and 20, using the case of storm-surge data. Calculation of mean squared error between the probability models and the empirical data revealed that the Gamma and Weibull distributions had the best fit to the data for both cold and warm seasons when MLE was used for estimating parameters of the probability model. These findings support previous ones by Bardsley (2019) regarding the Weibull distribution's appropriateness to extreme-value estimation. According to Bardsley, such a distribution could explain enough to enable extrapolation of the degree beyond the utilised data history, provided that the scale and shape parameter of the distribution are positive (meaning that the probability model has a good fit to the data). In the case of our own research, the shape and scale parameters were 1.87 and 5.21, respectively, indicating that the Weibull distribution model will likely have a good fit to large amounts of data beyond the dataset we used.

5. Please elaborate the literature review with reference to up-to-date publications, please refer to;
   1) with regards to exceedance diagrams

[1] Bermúdez, M., Cea, L., and Sopelana, J. (2019). "Quantifying the role of individual flood drivers and their correlations in flooding of coastal river reaches." Stoch Environ Res Risk Assess, 33(10), 1851-1861, https://doi.org/10.1007/s00477-019-01733-8/.

[2] Buchana, P., and McSharry, P. E. (2019). "Windstorm risk assessment for offshore wind farms in the North Sea." Wind Energy (Chichester, England), 22(9), 1219-1229, https://doi.org/10.1002/we.2351/.

[3] Catalano, A. J., Broccoli, A. J., Kapnick, S. B., and Janoski, T. P. (2019). "High-Impact Extratropical Cyclones along the Northeast Coast of the United States in a Long Coupled Climate Model Simulation." Journal of Climate, 32(7), 2131-2143, https://doi.org/10.1175/JCLI-D-18-0376.1/.

[4] Chen, Y., Li, J., Pan, S., Gan, M., Pan, Y., Xie, D., and Clee, S. (2019). "Joint probability analysis of extreme wave heights and surges along China's coasts." Ocean Engineering, 177 97-107, https://doi.org/10.1016/j.oceaneng.2018.12.010/.

[5] Davies, G., Callaghan, D. P., Gravois, U., Jiang, W., Hanslow, D., Nichol, S., and Baldock, T. (2017). "Improved treatment of non-stationary conditions and uncertainties in probabilistic models of storm wave climate." Coastal Engineering (Amsterdam), 127 1-19, https://doi.org/10.1016/j.coastaleng.2017.06.005/.

[6] Fawcett, L. and Walshaw, D. (2016). "Sea-surge and wind speed extremes: optimal estimation strategies for planners and engineers." Stoch Environ Res Risk Assess, 30(2), 463-480, https://doi.org/10.1007/s00477-015-1132-3/.

[7] Hisamatsu, R., Tabeta, S., Kim, S., and Mizuno, K. (2020). "Storm surge risk assessment for the insurance system: A case study in Tokyo Bay, Japan." Ocean &

Coastal Management, 189, https://doi.org/10.1016/j.ocecoaman.2020.105147/.
[8] Ke, Q., Jonkman, S. N., van Gelder, P. H. A. J. M, and Bricker, J. D. (2018). "Frequency Analysis of Storm-Surge-Induced Flooding for the Huangpu River in Shanghai, China." Journal of Marine Science and Engineering, 6(2), https://doi.org/10.3390/jmse6020070/.
[9] McInnes, K., Hoeke, R., Walsh, K., O'Grady, J., and Hubbert, G. (2016). "Application of a synthetic cyclone method for assessment of tropical cyclone storm tides in Samoa." Nat Hazards, 80(1), 425-444, https://doi.org/10.1007/s11069-015-1975-4/.
[10] Silva-González, F., Heredia-Zavoni, E., and Inda-Sarmiento, G. (2017). "Square Error Method for threshold estimation in extreme value analysis of wave heights." Ocean Engineering, 137 138-150, https://doi.org/10.1016/j.oceaneng.2017.03.028/.
[11] Wahl, T., Mudersbach, C., and Jensen, J. (2015). "Statistical Assessment of Storm Surge Scenarios Within Integrated Risk Analyses." Coastal Engineering Journal, 57(1), https://doi.org/10.1142/s0578563415400033/.

2) with regards to storm surge risk assessment;

[1] Zhu, Y., Xie, K., Ozbay, K., Zuo, F., and Yang, H. (2017). "Data-driven spatial modeling for quantifying networkwide resilience in the aftermath of hurricanes Irene and Sandy." Transp.Res.Rec., 2604(1), 9-18.
[2] Yum, S., Kim, J. H., and Wei, H. (2020). "Development of vulnerability curves of buildings to windstorms using." Journal of Building Engineering, Article in Press.
[3] Ke, Q., Jonkman, S. N., van Gelder, P. H. A. J. M, and Bricker, J. D. (2018). "Frequency Analysis of Storm-Surge-Induced Flooding for the Huangpu River in Shanghai, China." Journal of Marine Science and Engineering, 6(2), https://doi.org/10.3390/jmse6020070/.

- We are very grateful to the reviewer for providing these valuable references. As requested, we have added reviews of the studies recommended above by the reviewer. These can also be seen below.

Ke et al. (2018) studied these new frequencies of storm-induced flooding, with the aim of formulating new safety guidelines for flood-defence systems in Shanghai, China. They proposed a methodology for estimating new flooding frequencies, which involved analysing annual water-level data obtained from water-gauge stations along a river near Shanghai. The authors reported that a generalised extreme value (GEV) probability-distribution model was the best fit to the empirical data, and this led them to advocate changes in the recommended height of the city's flood wall. However, Ke at al. only considered annual maximum water levels when analysing flooding frequencies, which could have led to inaccurate estimation of the exceedance probability of extreme natural hazards such as mega-typhoons, which may bring unexpectedly or even unprecedentedly high water levels. In such circumstances, the protection of human society calls for highly accurate forecasting systems, especially as inaccurate estimation of the risk probability of these hazards can lead to the construction of facilities in inappropriate locations, thus wasting time and money as well as endangering life. Moreover, the combined effect of sea-level rises and tropical storms is potentially even more catastrophic than either of these hazards by itself.

Using insurance data from when Typhoon Maemi made landfall on the Korean Peninsula, Yum et al. (2021) presented vulnerability functions linked to typhoon-induced high wind speeds. Specifically, the authors used insurance data to calculate separate damage ratios for residential, commercial, and industrial buildings, and four damage states adopted from an insurance company and a government agency to construct vulnerability curves. Mean squared error and maximum likelihood estimation (MLE) were used to ascertain which curves most reliably explained the exceedance probability of the damage linked to particular wind speeds. Making novel use of a binomial method based on MLE, which is usually used to determine the extent of earthquake damage, the same study found that such an approach explained the extent of the damage caused by high winds in the Korean Peninsula more reliably than other existing methods such as theoretical probability method.

Zhu et al. (2017) explored recovery plans pertaining to two New York City disasters, Hurricanes Irene and Sandy, using data-driven city-wide spatial modelling. They used resilience quantification and logistic modelling to delineate neighbourhood tabulation areas, which were smaller units than other researchers had previously used, and which thus enabled the collection of more highly detailed data. They also introduced the concept of loss of resilience to reveal patterns of recovery from these two hurricanes, again based on their smaller spatial units. Moran's I was utilised to confirm that loss of resilience was strongly correlated not only with spatial characteristics, but also with socioeconomic ones, and factors like the location of transport systems. However, given the particularity of such factors, Zhu et al.'s results might not be generalisable beyond New York City; and they made no attempt to predict future extreme events' severity or frequency.

The sharp differences in the results of the past studies cited above are due to wide variations in both the data they used and their assumptions. The present study therefore applies all of the methods used in previous studies of Hurricane Sandy's return period to estimate that of Typhoon Maemi, and in the process, establishes a new model.

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.

6. Please see further comments and typo-edit in the attached.

- We would again like to thank Reviewer #1 for the above insightful and constructive comments on our manuscript. We hope that all of them have now been addressed, but of course, are happy to make further changes if needed.

7. Some of the Figures need Legend and improved resolution.

- We are grateful for these constructive comments. The original Figures 1 and 2 have been revised accordingly.

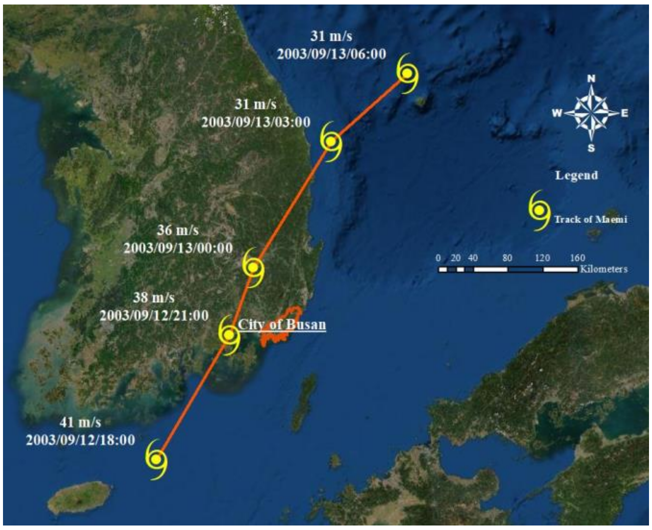

Figure 1. Track and wind speed of Maemi, 2003

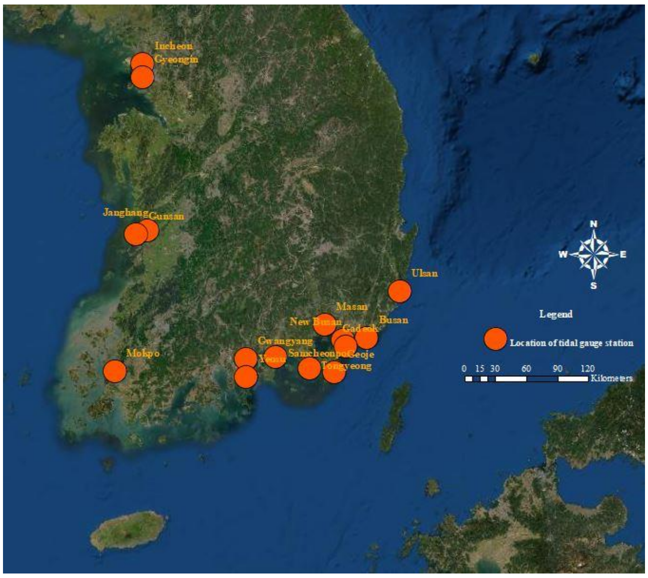

Figure 2. Locations of the 15 tidal-gauge stations on the western and southern coasts of South Korea as of 2003

8. Some references in the text are miss from the bibliographic list.

- Thank you for your comment. The missing references have been added to the References list.

-