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 #3

The manuscript details an important and highly relevant topic: Development of probability of storm surge occurrences which seeks to develop a risk analysis for predicting natural hazards. The motivation and the importance of the study are clearly presented, and it has fruitful information and a review of the state-of-the art regarding empirical analysis for storm surges. The proposed minor revisions are followed:

1. The methodology paragraph is highly important in this paper. Please provide a general approach and workflow.

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

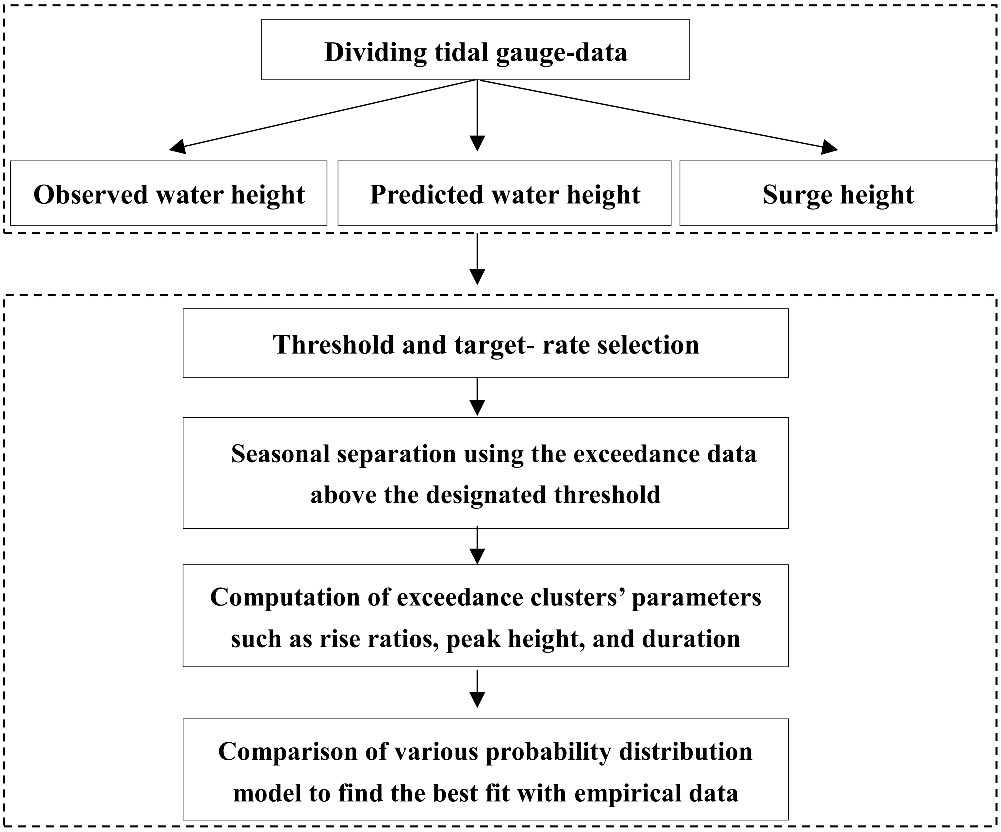

Figure 4. General approach and workflow

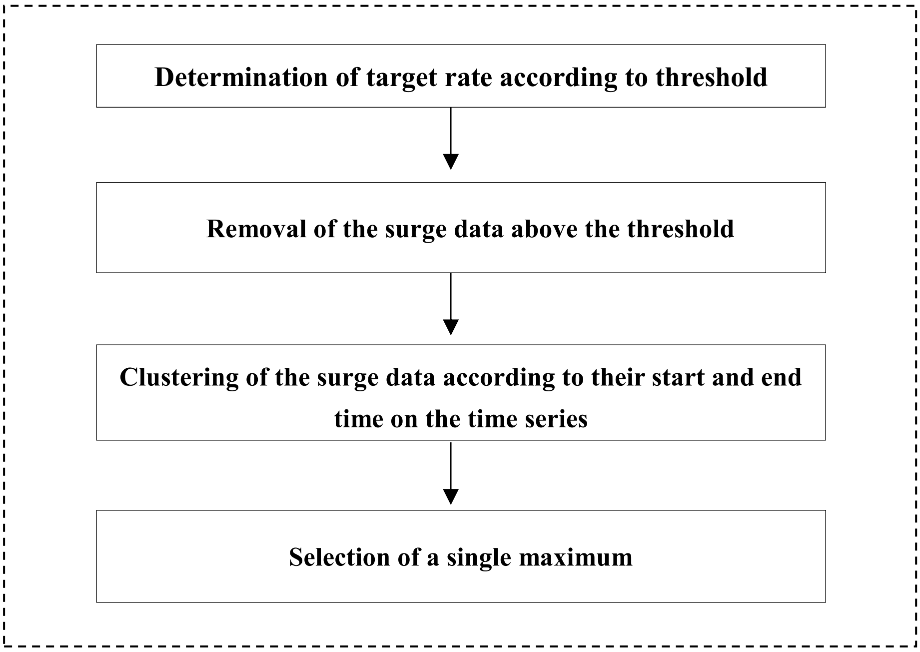

Figure 11. Clustering flowchart

2. Please improve the quality of the figures.

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

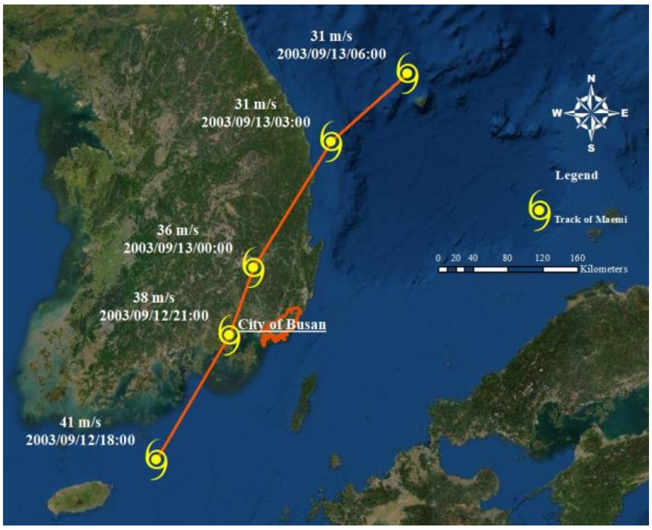

Figure 1. Track and wind speed of Maemi, 2003

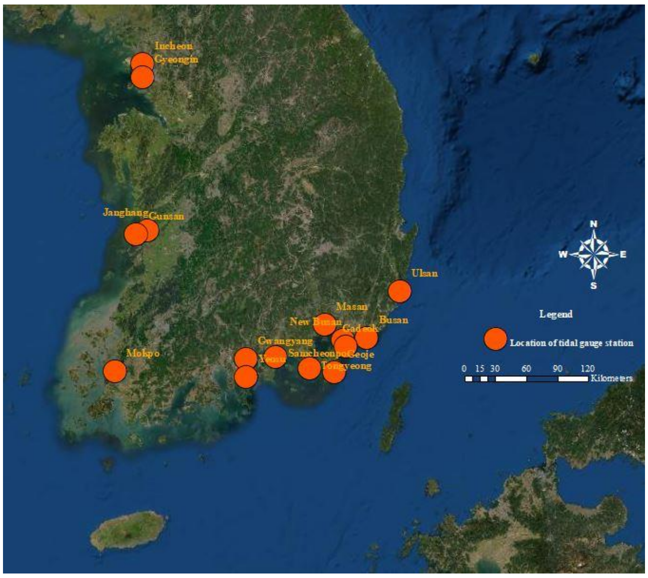

Figure 2. Locations of the 15 tidal-gauge stations on the western and southern coasts of South Korea as of 2003

3. Please emphasize the contribution of the research.

- We thank you for this comment. We are grateful to the reviewer for this helpful suggestion. The following passage has accordingly been added

Typhoons cause numerous fatalities and immense property damage, and their frequency has recently been increasing. Nevertheless, typhoon risk assessments are not yet sufficiently comprehensive to estimate either the damage levels from such events, or the probability of their occurrence. If they are to effectively plan for typhoons, governments and the insurance industry will need accurate estimates of both. Prompted by the high levels of damage inflicted by the high surge during South Korea's most severe typhoon, Maemi, this research has estimated the risk of storm surges through non-exceedance probability using MLE. Specifically, we estimated extreme storm surges' non-exceedance probability in accordance with their water levels, with such levels serving as references for non-exceedance probability above a certain threshold. We applied various methodologies to obtain more reliable thresholds, and a threshold-selection algorithm that utilised target rate and number of clusters to more accurately predict the height threshold. Additionally, we separated storm surges into cold-season and warm-season ones, as this allowed more reliable estimations, given their different frequencies in these seasons. Three parameters – exceedance, rise ratio, and duration – were separated from the storm surges and compared to ascertain their relationship. This established that exceedance and duration have a quite strong linear relationship. In previous research, total water level was utilised to estimate the possibility of future occurrences, but such an approach could lead to inaccurate results, for the reasons mentioned in the Literature Review section, above. Accordingly, in this study, we subcategorised total water levels into predicted, observed, and surge levels. Once that had been done, surge level was found to be the main factor influencing damage to coastal infrastructure, and thus, only it was applied to our estimates of non-exceedance probability.

Based on a quantitative risk assessment for extreme storm surges in a city on the Korean Peninsula that was severely damaged by Typhoon Maemi due to its geographical characteristics, this study has proposed a risk-management approach to such natural hazards based on the non-exceedance probabilities of extreme storm surges. Various probability-distribution models were tested within this framework to explore clustering and threshold-selection methods, and Weibull distribution was found to have the best fit to our empirical data. Our results suggest that the use of various probability models, clustering, and separation of tidal-gauge data as described above could all benefit the accuracy of natural-hazard return prediction. The present study's findings also confirm non-exceedance probability as a useful, geographically sensitive tool for government agencies, insurance companies, and construction companies conducting risk assessments, setting insurance prices, preparing safety guidelines, and setting policies aimed at reducing typhoon-related damage and financial losses.

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.

4. Please add more relevant literature review with up-to- date.

- We thank you for this comment. As recommended by the reviewer, 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.