General response: We sincerely thank the reviewers for their valuable feedback, which we have used to improve the quality of our manuscript. The reviewer comments are provided below in **bold font**, and specific concerns have been numbered. Our responses are given in normal font, and changes/additions to the manuscript are given in blue text.

Responses to Referee #1

General comments

I regret to say that, in my opinion, the quality of this paper is not suitable for publication. Language is poor and technical terms are associated rather randomly resulting in nonsense sentences (e.g. L69, 74, 135, 143, 154, 197, just to mention a few).

Leaving aside the general confusion of the text (denoting lack of familiarity with theoretical concepts), the data analysis is the usual superficial application of bivariate copulas already reported in hundred of papers, iterating widespread mistakes and lacking whatever uncertainty assessment (I cannot understand why uncertainty assessment is fundamental to provide a decent univariate frequency analysis, but it suddenly disappears and is neglected when moving to multivariate frequency analysis, which is affected by the additional uncertainty related to the unknown dependence structure).

Some results that are interpreted as empirical findings are just theoretical constraints that do not provide any insight.

Justifying a paper with sentences like "*Presently, a few scholars have conducted studies on the interaction and joint distribution of storm surges and waves*" makes no sense: the use of joint distributions and copulas have been used in coastal/ocean engineering for at least 15 years worldwide to model not only surge and significant wave height but also other met-ocean variables; these methods have also been incorporated in national guidelines (e.g. UK and Netherlands, to mention a few) for several years, and are subject to ongoing improvement and update. So, please perform a decent preliminary literature review before running out-of-the-shelf computer packages.

Response: Your comments and suggestions are appreciated. We believe most of them are valuable to the improvement of this manuscript, though we are unable to agree with all of them. Kindly take your time and find our detailed response and revisions below.

Specific comments

1. **L17: "*the surge height shows an increasing trend closer to the coastline*" ?? -> the surge height shows higher values closer to the coastline.**

   Response: Thanks for the suggestion on language improvement.

   Line 17-18: Second, the surge height shows higher values as locations get closer to the coastline, and the significant wave height becomes greater further from the coastline.

2. **L18: "*when one variable is constant, the simultaneous, joint, and conditional risk probability tends to decrease as the other variable increases*." Marginal and joint distributions are monotonic functions! This behaviour is related to their mathematical properties and has nothing to do with the analysed variables.**

   Response: Sure it is. But even when we have a function as simple as y = ax + b, and suppose a is positive, we may still want to emphasize that y will increase as x increases. So, in many cases, it makes perfect sense, isn't it? We did make slight revisions to avoid misunderstanding, as follows.

   Line 18-21: Third, the marginal and bivariate cumulative distribution is monotone increasing functions. Correspondingly, the simultaneous, joint, and conditional probabilities decrease monotonically. Therefore, improving the protection standard for either variable can effectively reduce the bivariate probability in the engineering design.

3. **L58: "*the study of the joint probability distribution of tropical cyclone storm surges and waves is conducive to improving the accuracy and precision of joint hazard assessment*." Yes, this is true if we neglect the fact that joint distributions are affected by the same uncertainty of their marginals and the additional uncertainty of the dependence structure.**

   Response: Uncertainty is of great importance when we quantify their joint probability. But even with the existence of uncertainty or even greater uncertainty, which no one denies, joint probability analysis can still provide valuable insights. The text was revised as below, with uncertainty mentioned.

   Line 46-49: The interaction, concurrency, or chain of hazards makes the comprehensive hazard assessment of compound hazards with significant uncertainties. In addition, in the design process of sea dikes, breakwaters, and harbors, the surge height and significant wave height in different return periods are often separately considered (MWR, 2014; MOT, 2015, 2018), disregarding the correlation between storm surge and waves.

   Line 50-51: Therefore, the joint probability distribution of tropical cyclone storm surges and waves and their uncertainty analysis are crucial in the integrated assessment of joint hazard intensity (Xu et al., 2022).

4. **L69: "*However, the constructed joint distribution is still a probabilistic result, and further search for constraint relations is needed to provide a basis and guidance for disaster prevention and mitigation. Therefore, this paper quantitatively analyzes the*

occurrence probability of storm surge and wave combinations based on the fitting results of the copula function. **" Words have a meaning and should be used accordingly! First, you say that "constraint relations" are needed because joint distributions are still a probabilistic (which seems to be a minus, whatever that sentence means), and then joint distributions are used neglecting constraints. Those sentences contradict each other.**

Response: We reorganized the text as follows, expressing that, even with a joint probability distribution (a 3d surface or a 2d curve with the probability of one variate fixed) available, it is not enough for setting design criteria. And what needs to do furtherly is to find a specific combination of surge and wave probability, and obtain their respective SH and SWH (i.e., two scalar values).

Please also refer to comment # 22, which was supposed to explain the same thing.

Line 59-60: However, as the bivariate joint probability is a three-dimensional surface, the intercepted curve with the specified probability or one variable fixed is a curve rather than a specific scalar, which is not enough for setting the protection criteria.

5. **L74: "*In the design process of sea dikes, breakwaters, and harbors, the surge height and significant wave height in different return periods is separately considered… disregarding the correlation between storm surge and so that the calculated water level may be underestimated or overestimated.*" This is incorrect, underestimation/overestimation means that an estimate is smaller/larger than a true value. In real-world applications, (i) the true value is unknown, so you never know if an estimator overestimates/underestimates, and (ii) univariate probabilities cannot be compared with joint probabilities as they refer to different processes! The choice of the required probabilistic model depends on the specific problem/failure mechanism. Univariate analysis is perfectly fine if there is a unique target design variable, and it correctly estimates the required probability (these concepts are explained here https://link.springer.com/article/10.1007/s00477-014-0916-1).**

   Response: (1) For sure, it is not easy or even practically impossible to estimate the so-called TRUE value of the intensity for a given probability. But it is also fair and safe to say that, without considering their interaction, it is highly possible that the TRUE value will be underestimated or overestimated, since the physical interaction mechanism between surge and wave is there. (2) On the other hand, the value of univariate analysis and its application of course are there, which is not the focus of this manuscript. That said, we only want to emphasize the importance of joint probability in the text. But we did revise the sentence a little bit to avoid misunderstanding.

   Line 44-49: However, tropical cyclone disasters in coastal areas are caused by the combined effect of storm surges, waves, and other hazards, and the mechanism of disaster generation is very complex. The interaction, concurrency, or chain of hazards makes the comprehensive hazard assessment of compound hazards with significant uncertainties. In addition, in the design process of sea dikes, breakwaters, and harbors, the surge height and significant wave height in different return periods are often separately considered (MWR, 2014; MOT, 2015, 2018), disregarding the correlation between storm surge and waves.

6. **L80: "*optimize*" -> fit. As the confusion due to meaningless terminology already affects a significant part of the literature, please, use consistent technical terms without inventing a new vocabulary!**

   Response: Relax, but we did not mean to create a new term, and we did not create a new term. When we use the word "optimize," we mean to select the optimal marginal distribution function and proper copula function through the AIC, BIC, and goodness of fit of the K-S test for all nodes, not just a simple fitting process. In order to clarify that, the text was revised as follows.

   Line 61-66: The goal of this paper is to explore the joint probability characteristics of tropical cyclone storm surges and waves, and to apply the joint probability distribution surface to investigate the methods and steps for the design values of the protection criteria for storm surges and waves under combined scenarios. First, based on historical tropical cyclone surge heights and significant wave height, we fit the marginal distribution and copula function of nodes in the study area, and we use the maximum likelihood method to estimate the parameters. Then, the optimal functions are selected based on the Kolmogorov-Smirnov (K-S) test, AIC, and BIC for all nodes.

7. **L81: "*function by the passing rate of the K-S test.*" When performing tests, the only meaningful "rate" is the rejection rate! (see here https://www.sciencedirect.com/science/article/pii/S0309170817305845 for an explanation)**

   Response: Thanks for your comment and suggestion. The term "passing rate" was misused in the original text, and it did not refer to the results of the K-S test, and we have changed the original statement in the revised manuscript. In the process of selecting the optimal marginal distribution function and copula function, the goodness-of-fit of each marginal distribution function and copula function is calculated by the K-S test for each node. If the p-value of the K-S test of a node is greater than 0.05, the hypothesis of "the sample obeys a certain theoretical distribution" is not rejected, which means that the node passes the K-S test. Then, the ratio of the number of nodes passing the K-S test to the number of all nodes is calculated, and based on this, the optimal function is determined by combining the AIC, BIC, and D-values of the K-S tests.

   Line 63-66: First, based on historical tropical cyclone surge heights and significant wave height, we fit the marginal distribution and copula function of nodes in the study area, and we use the maximum likelihood method to estimate the parameters. Then, the optimal functions are selected based on the Kolmogorov-Smirnov (K-S) test, AIC, and BIC for all nodes.

8. **L135: If the result of rewording concepts is this one, then it is preferable copying and pasting from some good paper/handbook. These sentences are just a set of randomly chosen words that make no sense whatsoever.**

   Response: Thanks for the suggestion on language.

Line 122-127: Sklar's theorem (Sklar, 1973) elucidates the role that copulas play in the relationship between multivariate distribution functions and their univariate margins and states that any multivariate joint distribution can be described by a univariate marginal distribution function and a couple describing the dependence structure between the variables (Nelsen, 2006). Let $F(x)$ and $G(y)$ be the marginal distributions of x and y, C is the copula, and $H(x, y) = C(F(x), G(y))$, where H is the bivariate joint distribution function of x and y (Serinaldi, 2015). Therefore, the copula function is widely utilized in multi-hazard joint probability analysis of natural disasters (Chen et al., 2019; Lee et al., 2013).

9. **L151: Archimedean, elliptical and quadratic are not families but classes of copulas, and they are not the only existing classes. Archimedean copulas can be multiparametric as well! Please, do not use models without knowing the underlying theory, as the result is just a collection of misconceptions, meaningless statements, and misinterpreted figures and tables.**

   Response: We believe this is a language problem, which has been revised by rewording. Joint probability analysis is not something like unsolved mathematical puzzles. And we believe avoiding subjective judgment with no consolidated evidence will improve the quality of the review.

   Line 143-146: Many families of copulas exist and mainly include the following: Meta-elliptical copulas (normal and t), Archimedean copulas (Clayton, Gumbel, Frank, and Ali-Mikhail-Haq), Extreme Value copulas (Gumbel, Husler-Reiss, Galambos, Tawn, and t-EV), and other families (Plackett and Farlie-Gumbel-Morgenstern). Among the various families of copulas, the Archimedean copula is more popular for hydrologic applications (Chen and Guo, 2019).

10. **L154: "*analyze the joint probabilities of the marginal distributions of two variables*" -> "the joint probabilities of two variables". Joint probabilities of the marginal distributions do not exist!**

    Response: Thanks for the comments. Revision is made as follows.

    Line 146-149: The commonly employed Archimedean copula functions include Gumbel, Clayton, and Frank (Table 1), which are selected to analyze the joint probabilities of two variables, the surge height and significant wave height of tropical cyclone, and then the maximum likelihood method is used to estimate the parameters of the copula function.

11. **L175: The conditional probabilities and the corresponding conditional distributions and return periods are not joint (multidimensional) but univariate (unidimensional).**

    Response: Thanks for your comment. But we do not think your comment is correct, especially in the context of this manuscript.

    We believe that the conditional probabilities and return periods can be bivariate distributions.

    The concept of the conditional probability of a multidimensional random variable is the probability of $(X > x)$ occurring under the condition that $(Y > y)$, which is calculated by the formula $P_| = P\big((X > x)\big|(Y > y)\big) = \frac{P(X>x,Y>y)}{P(Y>y)}$. Therefore, it is a conditional distribution of the two-dimensional variable (X, Y). In addition, a preliminary introduction to the concepts of multivariate probabilities and return periods is provided in the paper (https://link.springer.com/article/10.1007/s00477-014-0916-1), which also states that conditional probabilities and conditional return periods are multivariate.

12. **L197: "*The marginal and joint probabilities of storm surge and wave scenarios cannot be directly employed as reference values for engineering protection standards.*" Return periods are just indicators derived from probabilities: if the former can be used, the latter can be used as well.**

    Response: According to your suggestion, we have deleted "marginal" from the original manuscript and only proposed that the joint probability cannot be used as a reference value.

    This is mainly because the intensity values of storm surges and waves under different univariate return periods can be used as a reference for marine engineering protection. However, the bivariate joint probability/return period is a three-dimensional surface. The intercepted curve under the specified occurrence probability or return period is a curve, so it is difficult to be used directly as a reference value for engineering protection criteria.

    Line 191-193: The joint probabilities of storm surge and wave scenarios is hard to be directly employed as reference values for engineering protection standards, since the bivariate joint probability/return period is a three-dimensional surface, and the intercepted curve under the specified occurrence probability or return period is a curve, not a scalar.

13. **Sect. 3.4.2: this awkward method is quite useless because AND and OR return level curves are complementary, and their surfaces are monotonic and diagonally symmetric for the considered Archimedean copulas. This means that the point estimate (x, y) resulting from the (unnecessary) non-linear optimization is just the intersection of the main diagonal of the copula and the OR level curve with RP = k. No optimization is required, just some familiarity with the theoretical properties of the models one intends to use.**

    Response: Thanks for your comments. But we have a different opinion from you, and no revision has been made.

    We believe it is essential to use a non-linear optimization method to estimate optimal design values for storm surge and wave combinations. This is because although $RP_\cap$ and $RP_\cup$ are complementary and the three-dimensional surface is monotonic, they are not perfectly diagonally symmetric. Therefore it is not the intersection of the main diagonal and $RP_\cup = K$ that corresponds to the largest value of $RP_\cap$.

14. **L214: "*determine whether it passes the 95% significance level.*" "Passing the 95% significance level" makes not sense. A test can only reject or not reject the reference/null hypothesis at a given significance level. The significance level is not 95% but 5%.**

Response: Thanks for your comments. We have now changed the wording to "*the p-value of K-S test is used to determine whether the hypothesis that the sample obeys a certain theoretical distribution is rejected*" in the revised manuscript.

Line 208-211: Next, the time series of the bivariate annual maximum value for all nodes are fitted with marginal functions by five functions, including Gumbel, Weibull, gamma, exponential, and generalized extreme value distribution (GEV), and the p-value of K-S test is used to determine whether the hypothesis that the sample obeys a certain theoretical distribution is rejected.

15. **Table 3 and corresponding discussion: Using the rate of no rejection to select the best model makes no sense for several reasons:**

    1) "No rejection" does not mean acceptance, and if two or more models are not rejected for a given data set, the only possible conclusion is that the information/data is not enough to discriminate among the models.

    2) The comparison includes distributions with different number of parameters (1-parameter Exponential, 2-parameter Gamma, Gumbel and Weibull (assuming that the Weibull does not include a location parameter), and the 3-parameter GEV). So, GEV has the obvious advantage of being more flexible due to higher parameterization.

    3) Performing a test on 1665 nodes (i.e., 1665 times) is a multiple testing exercise. Under independence (which is not valid in this case), if the null hypothesis is valid, the number of no rejections has an expected value equal to 1582 (95%), and a binomial distribution, thus meaning that the 95% confidence interval of the number of no rejections is (1564, 1599). Therefore, Exponential and Gumbel cannot be discarded for the variable SH.

    However, met-ocean variables come from a model and refer to grid points of a connected area, they are surely strongly correlated in space. This means that the binomial distribution strongly underestimates the actual variability of the number of no rejections because of information redundancy. Therefore, other distributions for both SH and SWH might not be discarded. This is not surprising as these distributions are fitted to just 60/65 annual maxima, i.e. a very small sample size.

    4) "No rejection" rate for Weibull and SWH is reasonably wrong because (i) WH/SWH (in deep water) has been historically modelled by Rayleigh distribution, which a particular case of Weibull, (ii) Weibull was shown to be a very good model in shallow water for WH (see https://doi.org/10.1016/j.coastaleng.2022.104130), and (iii) Weibull and 3-paremeter Weibull are closely related to GEV, and Weibull is also a pre-asymptotic distribution in EVT.

    In summary the selection strategy seems to work just because it neglects the foregoing theoretical aspects.

    Response: Thanks for pointing out the problems in the marginal distribution selection process. We changed/refined the selection process as follows:

(1) The p-value of the K-S test is used to determine whether each node rejects the hypothesis that the sample obeys a certain theoretical distribution (Table 3).

(2) The optimal function is screened by three metrics, AIC, BIC, and the D-value of the K-S test. The smaller the AIC, BIC, and D-value of the K-S test, the better the fit, thus determining the optimal marginal function for each node.

(3) The percentage of each optimal function to the number of all nodes was calculated (Table 3), and the function with the highest rate was selected as the optimal marginal function for the storm surges and waves.

Through the above steps, GEV was finally determined as the marginal function for storm surges and waves:

(1) GEV belongs to 3 parameters, and the parameter fitting is relatively flexible.

(2) The no-rejection numbers of the fitted GEV functions for storm surge and waves are 1665 and 1657, respectively, both of which are in the 95% confidence interval.

(4) Among all functions, the GEV function has the highest percentage of preferences, 30% and 27%, respectively.

Line 135-140: In this paper, the maximum likelihood method is used to estimate the fit parameters, based on which the optimal marginal functions for storm surges and waves are screened by the following steps: firstly, the p-value of the K-S test is used to determine whether each node rejects the hypothesis that the samples obey a certain functional distribution; secondly, the optimal function for each node is screened by the three metrics, AIC, BIC and D-value of the K-S test. The smaller the AIC, BIC, and D-value of the K-S test, the better the fit, thus determining the optimal marginal function for each node.

Line 211-212: Then, we counted the frequency of each function passing the K-S test and its percentage as well as the frequency of the optimal function and its percentage (Table 3).

Line 217-221: Based on the statistical results, it was found that for fitting the SH, the K-S test of the GEV function had the highest no-rejection rate of 100%, and the corresponding preference ratio was 30.04%, so GEV was set as the optimal marginal function in this study. For SWH fitting, the number of nodes with no rejection in the K-S test of the GEV function was 1657, accounting for 99.52% of the total number of nodes, and the corresponding percentage of preferences was also higher than that of other fitting functions.

Table 1 Frequency and percentage of five functions passing the K-S test and the optimal function for all nodes of SH and SWH

Marginal function	Surge height				Significant wave height			
	Frequency of K-S test passed	Percentage of K-S test passed (%)	Frequency of optimal function	Percentage of optimal function (%)	Frequency of K-S test passed	Percentage of K-S test passed (%)	Frequency of optimal function	Percentage of optimal function (%)
Gamma	1508	90.57	183	10.99	1464	87.93	159	9.55
Exponential	1567	94.11	216	12.97	1076	64.62	95	5.71

Gumbel (right)	1615	97.00	350	21.02	1629	97.84	149	8.95
Weibull (max)	1469	88.23	416	24.98	300	18.02	494	29.66
GEV	**1665**	**100.00**	**500**	**30.04**	**1657**	**99.52**	**768**	**46.13**

16. Figure3: What about complementing these figures with confidence/prediction intervals?

Response: Thanks for your suggestion. We add 95% confidence intervals to Figure 3a and Figure 3c in the revised manuscript.

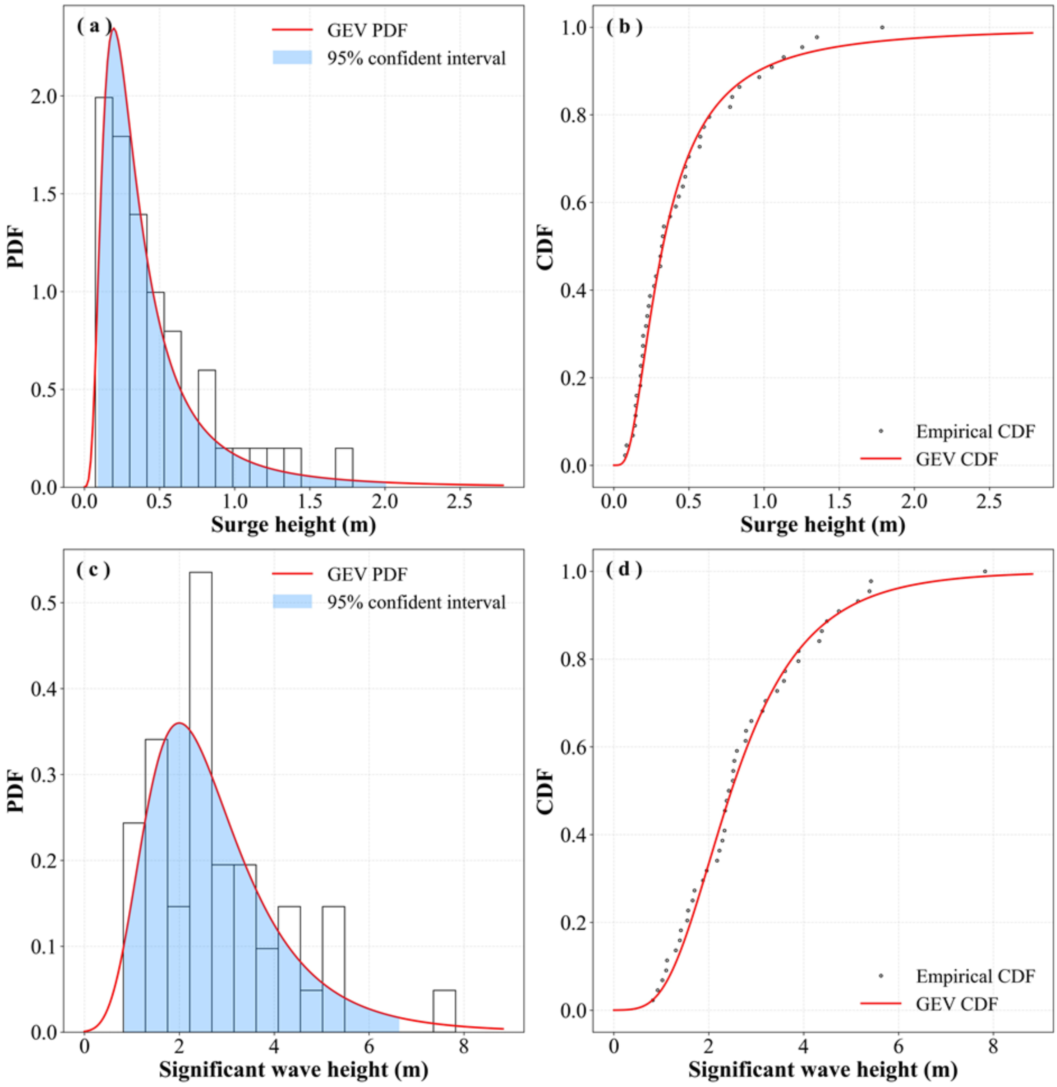

Figure 3: Fitting results of the PDF and CDF of the SH and SWH based on the GEV function (using node (110.5142° E, 20.2768° N) as an example)

17. L249: "*In general, the SWH increases with an increasing return period ...*" In general? Distributions are monotonic functions: higher quantiles always correspond to higher return period for whatever (continuous) random variable… always, not "in general"!

Response: We agree that "in general" does not make sense here, and the sentence is deleted.

18. **Fig. 6: This figure (along with Fig. 7) is uninformative. The shape of the surfaces of generic joint PDFs and CDFs is well known. What matters is some diagnostic plot showing the goodness of fit (with uncertainty!)**

    Response: We have removed the two example figures, Fig. 6 and Fig. 7, from the manuscript. In addition, we calculated the goodness of fit for each grid, filtered the optimal copula function based on this, and did not show the diagnostic plots separately anymore.

19. **L270: "*When the intensity values of the two disaster-causing factors are equivalent,* $RP_{\cap}$ *is greater than* $RP_{\cup}$*, which indicates that* $P_{\cap}$ *is smaller than* $P_{\cup}$." This is not a result but the effect of theoretical constraints (see here https://link.springer.com/article/10.1007/s00477-014-0916-1)**

    Response: Thanks for your comments. According to Eq. 2 - Eq. 5, when the intensity values of the two hazards are equivalent, $RP_{\cap} > RP_{\cup}$ and $P_{\cap} < P_{\cup}$ is evident. So, it is removed in the revised manuscript.

    Line 260-261: Based on the optimal marginal distribution function and copula function, we calculate $RP_{\cap}$, $RP_{\cup}$, and $RP_{|}$ of SHs and SWHs.

20. **Figs 8-10: These figures just report what is expected according to the monotonic nature of bivariate distributions.**

    Response: Thanks for your comments, but we believe that Fig. 8 - Fig. 10 can provide valuable info to readers. Even theoretically, the distributions are monotonic. So no revision is made.

    On the one hand, they show the spatial distribution of simultaneous, joint, and conditional probabilities for different combinations of return periods for storm surges and waves, showing the high and low probability of occurrence in each region and thus highlighting regional differences.

    On the other hand, they also express the characteristic monotonically increasing (decreasing) change in the simultaneous, joint, and conditional probability as the return period of one variable is held constant and as the return period of the other variable increases, laying the groundwork for later research on reducing the bivariate probability by increasing the univariate return period protection criteria.

21. **L315: "$P_{\&}$" was not defined in the text. I guess it refers to the discretised classes in Sect. 3.3.3 and Fig. 11.**

    Response: Thanks for your reminder. The $P_{\&}$ in this text refers to the bivariate probability of the discrete hazard class, which is calculated as Eq. 8 in Sect. 3.3.3, and we have added $P_{\&}$ to Eq. 8 in the revised manuscript. In addition, $P_{\cup}$ in L288 of the original manuscript is incorrect and should be changed to $P_{\&}$.

    Line 178-179: We calculate the bivariate probabilities for discretized hazard level combination scenarios based on the marginal and copula functions of the storm surge and wave. The calculation formula is as follows:

$$P_{\&} = P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$$
$$= P(X \leq x_2, Y \leq y_2) - P(X \leq x_2, Y \leq y_1) - P(X \leq x_1, Y \leq y_2) + P(X \leq x_1, Y \leq y_1) = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1) \quad (8)$$

Line 288-290: We calculated two-dimensional $P_{\&}$ based on Eq. 8 for all nodes with discretized combinations of SH and SWH for a total of 25 scenarios and interpolate them into 1 km raster data using the cubic spline interpolation method (Figure 9).

22. **L409-410: These statements are not related to the preceding text; they are just generic sentences.**

Response: Thanks for the comments. The sentence was supposed to explain that only with joint probability available is not enough for design criteria determination. A method to determine the specific values of SH and SWH is needed. Please refer to comment #4. So, we reorganized the text as follows.

Line 394-397: Since the joint probability distribution of bivariate is a 3-dimensional surface. In order to obtain the specific scaler values of the two hazards as design criteria, in this study, the method for estimating the minimum return periods of SH and SWH was implemented, given their estimated joint probability distribution as a constraint.