Response to Anonymous Referees

Dear Editor,

Thank you very much for the constructive feedback from you and the reviewers. We have addressed all the reviewers' comments. Please find a point-by-point reply below.

Kind regards

Ting Zhang

Response to Referee #2

In this study, the authors thoroughly explored the factors driving compound flooding in an urban area by combining hydrodynamic modeling and multivariate statistics. Their research provided insights into the analysis of the relative contribution of rainfall and tidal levels on compound flooding and the roles of different flood scenarios. The study is fairly novel and findings are beneficial for flood risk management and applications. Further, the scope and findings of the study align with the scope of the journal "Hydrology and Earth System Sciences," and I recommend its publication but after some revisions. Below, the authors can find my suggestions and recommendations, which I believe would contribute to improving the manuscript and clarifying the methodology.

Major Comments:

1. Copulas and marginal distributions are fitted to the "tidal levels." Is this just the astronomical tide, or do these values include wind-driven surge (in the case of Typhoon/Cyclone conditions) and river discharge components as well? The authors mention "storm surge" as a critical factor for compound flooding several times (L346, L456, L564), but it's not clear how storm surge is taken into account when modeling "tidal levels" in this study. If the tidal range is higher and the storm surge component is relatively very small (or if only the astronomical tide is considered), fitting an extreme value model might be fundamentally incorrect since the astronomical tide is deterministic. Therefore, a clear explanation of the terminology is needed. Furthermore, are the tidal boundary conditions used at the sluice gate (as outflow boundary conditions?)? The authors mention a tidal gate, but the location is not shown in the figure.

Response: Thank you very much for your meticulous review and valuable comments. You raised the issue regarding the fitting of extreme values for "tidal levels," particularly how to account for storm surge factors. In this study, "tidal levels" is not solely based on astronomical tides but also includes the water level rise caused by wind-driven surge and river discharge. The study area experiences abundant rainfall and frequent typhoons, leading to significant tidal level increases due to wind-driven surge and discharge from heavy rainfall. For example, Typhoon Mangkhut in 2018 resulted in the highest tidal level recorded at the tide gauge in this study, reaching 3.28 m. Therefore, the tidal level is not deterministic. We will provide a clear explanation of the terminology in the revised manuscript.

Regarding the issue of sluice gate and tidal boundary conditions, as you correctly pointed out, the sluice gate uses tidal level as the outflow boundary condition. This setup ensures that the model can account for tidal influences. The tidal gate in this paper is the sluice gate. The location of the sluice gate is shown in Figure 1. We will unify the terminology of pictures and words.

2. When selecting the sample for joint probability analysis, the tidal level is chosen as the "highest tidal level of the corresponding day." This method implies that the calculated tidal design levels could occur within a 24-hour period from the rainfall event. In the numerical modeling process, a single event is used as a representative for the tidal levels process, but the timing of the rainfall peak and tidal peaks in generating design scenarios is not explained. Additionally, figure 8(a) shows the peak rainfall and peak tide almost coinciding, which is rare in reality (not the most probable 200-year event). The 3-hour rainfall duration shows the peak rainfall occurring approximately half an hour after the peak tide, whereas the 12-hour rainfall duration shows the peak rainfall occurring approximately 3 hours before the peak tide. How are the corresponding time lags between peak rainfall and peak tide determined for each scenario? Are they coming from a probabilistic basis? This timing significantly impacts the resultant flood volume and might be the main reason for the variations discussed around L430.

Response: Thank you for your detailed review of our research methodology and valuable suggestions. Our study selected the tidal level on May 22, 2020, as the typical tidal level process, based on the "Comprehensive Plan for Drainage (Rainwater) and Flood Prevention in Guangzhou (2022-2035)." The peak time of the typical tidal level is close to the peak time of the 24-hour design rainfall. For tide levels with rainfall durations of 1 h, 3 h, 6 h, and 12 h, this study designed the tidal levels with the same rising and ebb tide times, with the peak tide level occurring in the middle of the simulation period. The corresponding explanation will be supplemented in the revised manuscript. In future work, we will consider increasing the lag time and establishing the joint distribution of the three variables to design rainfall and tide level combinations. Additionally, we will discuss in detail about the impact of the lag time between peak rainfall and peak tidal level on flood volume in the revised manuscript and explain the reasons for the variations in L430.

3. 41 and 42 are used to evaluate the spatial interactions between drainage units (DUs). The equations only rely on the masses and distances of two drainage units. To clarify the conclusions around L445-L450, it is essential to understand what the

spatial interaction force represents. For example, the spatial interaction force between DU22 and DU6 (which are far away from each other) is 0.75. Does this relatively higher value suggest a higher level of interconnectedness or influence on each other, than even adjacent drainage units? Please explain. Also, it is not clear how the spatial interaction of the drainage units relates to the main objectives of the study.

Response: Thank you very much for your meticulous review of the spatial interaction between drainage units in our study. The primary aim of this research is to analyze the sources of compound flooding. The study on the spatial interaction of drainage units seeks to provide a scientific basis for identifying key areas for flood management. The spatial interaction between drainage units is based on the urban gravity model, which posits that the intensity of interaction between cities is proportional to their size (mass) and inversely proportional to their distance. In this study, the spatial interaction between drainage units is analogous to the interaction between cities. Spatial interaction force represents the strength of the connection between drainage units.

Theoretically, the spatial interaction force between DU22 and DU6 is relatively high, indicating a higher level of interconnectedness compared to adjacent drainage units. Upon careful examination of the calculation process, we found that the drainage pipeline length of DU22 is significantly greater than that of other drainage units on the left bank, resulting in a higher spatial interaction force between DU22 and other drainage units, even with more distant to DU6. This is related to the simplification of the model assumptions. In this study, the mass indicators only included flood volume and drainage pipeline length, which may lead to an incomplete interpretation of mass. Additionally, the weights of mass and distance are equal, which could introduce uncertainty. Future work will involve conducting sensitivity analyses to select more mass indicators and set indicator weights to reduce uncertainty in the results. We will supplement the corresponding explanations and analyses in the revised manuscript.

4. It is necessary to provide additional details on the types of data used in the study. For instance, specify whether rainfall data is gridded or gauge data, and if gauge data is used, indicate the number of gauges involved. What are the locations of the tidal data and rainfall used? If so, do you assume a uniformly distributed rainfall over the entire catchment?

Response: Thank you very much for your meticulous review and attention to the data types. The rainfall and tide level data used in this study are gauge data. We obtained rainfall data from 3 gauges and tidal data from 1 gauge. We assumed that the distribution of rainfall is uniform across the entire catchment area, so we used the average value from the 3 gauges. This assumption is based on the representativeness of the gauges and the limitations in obtaining long-sequence data. Future work will consider using data from more gauges or high-resolution gridded data to improve the accuracy and reliability of the model. We will include detailed descriptions of the data in the revised manuscript and indicate the specific locations of these gauges in the figures.

5. The paper lacks a proper discussion about the main assumptions and limitations of the process. For instance, using only 16 years of data to estimate a 200-year design event introduces significant uncertainty in modeling the dependence structure and the tail distributions. Since the study proposes a "universal" method, it is crucial for readers to understand these assumptions and uncertainties, especially when the method is to be applied to other study sites.

Response: Thank you very much for your suggestion. We will add a section in the revised manuscript to discuss in detail the assumptions and limitations of this study. This will be very helpful in enhancing the rigor and applicability of the "universal" method presented in this paper.

Minor Comments

1. L11 “currently” repeats.

Response: Thank you for pointing out the repetition error in the text. We will remove the redundant "currently" in the revised manuscript.

2. In the abstract, the authors mention the Kendall return period for the combined event of rainfall and tidal levels, greater than the “Or” return period and less than the “And” return period. The Kendall return period is typically expected to fall between the "Or" and "And" scenarios for combined events. It's not a significant finding of the study to be mentioned in the abstract.

Response: Thank you for your meticulous review of the abstract content. Based on your suggestion, we will remove the description of the Kendall return period from the abstract.

3. L32, and L36, could you provide references?

Response: Thank you very much for your review comments. We have supplemented the relevant references as per your suggestions to enhance the accuracy and reliability of the paper. In response to your request, we have added the following references:

Dorrington, J., Wenta, M., Grazzini, F., Magnusson, L., Vitart, F., and Grams, C.: Precursors and pathways: Dynamically informed extreme event forecasting demonstrated on the historic Emilia-Romagna 2023 flood, EGUsphere, 2024, 1-27, https://doi.org/10.5194/egusphere-2024-415, 2024.

Jiao, Z., Zhang, Z., and Wu, L.: SAR-based dynamic information retrieving of the Beijing-Tianjin-Hebei flood-inundation happened in July 2023, North China, Geomatics, Natural Hazards and Risk, 15, 2366361, https://doi.org/10.1080/19475705.2024.2366361, 2024.

Marengo, J. A., Cunha, A. P., Seluchi, M. E., Camarinha, P. I., Dolif, G., Sperling, V. B., Alcântara, E. H., Ramos, A. M., Andrade, M. M., and Stabile, R. A.: Heavy rains and hydrogeological disasters on February 18th–19th, 2023, in the city of São Sebastião, São Paulo, Brazil: from meteorological causes to early warnings, Nat. Hazards, 120, 7997-8024, https://doi.org/10.1007/s11069-024-06558-5, 2024.

4. Figure 2: what represents the purple arrow from correlation analysis to “effects of social factors on flood”?

Response: Thank you very much for your meticulous review. We indeed recognize the error in the direction of the arrow and apologize for it. The original intention was to express that the analysis of "effects of social factors on flood" is based on the combination of rainfall and tide level events. We have made the necessary corrections to Figure 2 to ensure the arrows point correctly.

5. L61, in underdeveloped areas, fluvial flooding may indeed be more prevalent due to the natural terrain and lack of infrastructure to mitigate such events. However, it is not accurate to say that "only" fluvial flooding exists in these areas. Could you explain or provide references?

Response: Thank you for your correction. The statement in L61 is indeed inaccurate. In underdeveloped areas, fluvial flooding may be more common. However, it is inappropriate to simplify the flooding issues in these regions as "only fluvial flooding." These areas may also encounter different types of flooding events caused by factors such as tide. We will re-examine this paragraph and provide a more accurate description in the revised manuscript.

6. Provide references for the “Pilgrim & Cordery rainfall model”, “Maximum Possible Weighting Function” and “Pilgrim & Cordery rainfall model”.

Response: Thank you very much for your suggestions regarding the references. We have supplemented the corresponding references for the "Pilgrim & Cordery rainfall model" and the "Maximum Possible Weighting Function."

For the "Pilgrim & Cordery rainfall model," we will cite the following references:Pilgrim, D. H. and Cordery, I.: Rainfall temporal patterns for design floods, Journal of the Hydraulics Division, 101, 81-95, https://doi.org/10.1061/JYCEAJ.0004197, 1975.

For the "Maximum Possible Weighting Function," we will cite the following reference:Gibson, S., Wills, A., and Ninness, B.: Maximum-likelihood parameter estimation of bilinear systems, IEEE Transactions on Automatic Control, 50, 1581-1596, https://doi.org/10.1109/TAC.2005.856664, 2005.

7. Figure 6 seems incorrect. The theoretical distributions (GEV, Gamma) should be smooth unless you don't change the distribution parameters.

Response: Thank you very much for pointing out the issue in Figure 6, and we apologize for the oversight. Upon careful examination, we found that the legends for the empirical distribution curve and the theoretical distribution curve were reversed, causing confusion. We have corrected Figure 6 to ensure that all legends accurately reflect their corresponding distribution types. The theoretical distribution curve is now smooth and will be presented in the revised manuscript.

8. What types of information does Figure 7 (b) provide? What are the insights we can gain into joint distribution probabilities? The color scale is also not given.

Response: Thank you very much for your detailed review of Figure 7(b). We apologize for not providing complete information. In Figure 7(b), we can see the joint probability of two variables for any combination. The graph shows that a significant portion of the joint probability is below 0.5, suggesting that the joint probability of rainfall and tide level has a sharp peak and heavy tail characteristic. This indicates that a majority of the data is concentrated within a specific range, while the rest of the data is spread out across a wide range of intervals. We will provide a detailed analysis of Figure 7(b) in the revised manuscript. Additionally, you correctly pointed out the absence of a color scale. The color scale is a crucial element for understanding the data distribution in the figure. We have modified Figure 7(b) to include the color scale, which will be presented in the revised manuscript.

9. Although the peak tidal levels could be negative (depending on the datum used), the gamma distribution is selected to fit the peak tidal levels for some rainfall durations. The gamma distribution is lower bounded by zero thus making it impossible to sample negative realizations from the fitted distribution (if necessary).

Response: Your opinion is indeed correct. The lower bound of the gamma distribution is zero, making it unsuitable for fitting data that may include negative values. In our study, the peak tidal levels are all positive, based on the reference datum we selected. Therefore, using the gamma distribution in this context is appropriate. We will include an explanation of this point in the revised manuscript to help readers better understand the conditions under which the gamma distribution is used.

10. The extreme sample selection is not well explained. Does it consider annual maxima? To ensure that events are independent, it is important to make sure that the outcome of one event does not affect the outcome of another. This can be achieved by using random sampling techniques and ensuring that each event is selected independently of the others. Additionally, it is important to have a sufficiently large sample size in order to make accurate inferences about the population. The sample

size should be determined based on the desired level of confidence and the variability of the data.

Response: Thank you very much for your attention and suggestions regarding the sample selection method. In this study, we considered the annual maximum values. First, we calculated the maximum rainfall for each year, forming set D, and used the minimum value in set D as the sample threshold. Then, from 16 years of historical data, we selected events greater than or equal to this threshold as preliminary rainfall samples. Following the principle of the minimum inter-event time of 6 hours between two samples, if the minimum inter-event time was less than 6 hours, we selected the larger of the two samples. Furthermore, if multiple events within a single day exceeded the threshold, we chose only the largest one to generate the final rainfall samples. For each selected rainfall sample P, we identified the highest tide level value Z on the same day as the tide sample. In this study, the rainfall thresholds for 1-hour, 3-hour, 6-hour, 12-hour, and 24-hour durations were 36 mm, 56 mm, 58 mm, 66 mm, and 78 mm, respectively. The final number of samples for each duration was 48, 39, 49, 49, and 52, respectively. A more detailed explanation will be provided in the methodology section of the revised manuscript.

11. Figure 10: The study area is in the legend (red). But cannot be seen in the figure.

Response: Thank you very much for pointing out the issue in Figure 10, and we apologize for the oversight. Upon careful examination, we found that the color annotation for the study area was indeed incorrect. The red marker you mentioned was not correctly displayed in the figure, which may have caused difficulties in understanding. We have corrected Figure 10 to ensure that the color of the study area in the legend matches the color displayed in the figure.

12. Figure 12: Do these different colors represent any information? If not use a single color.

Response: Thank you for your meticulous review of the color usage in Figure 12. Indeed, the different colors used in Figure 12 did not represent specific information, which may have caused unnecessary confusion. Based on your suggestion, we have modified these colors to a single color to make the graphical representation more concise and clear.

13. It's not clear how this observation is made. Despite differences in rainfall peak timing, rainfall, and tide peak timings leading to variations in calculated flooding volumes, the impact degree index is not affected by these differences.

Response: Thank you very much for your attention to our research results. Regarding the observation you mentioned, "Despite differences in rainfall peak timing, rainfall, and tide peak timings leading to variations in calculated flooding volumes, the impact degree index is not affected by these differences," we recognize that there was a writing error in the original text, and we apologize for this. We will reorganize and refine this section in the revised manuscript.

14. It's not clear how this conclusion is made: “This study also highlights that under the same recurrence interval, rainfall events with larger peak timings are more destructive than those with earlier peak timings”? Have you run many events for the same return period by only changing the peak timing?

Response: Thank you for your meticulous review. Regarding your mention: "This study also highlights that under the same recurrence interval, rainfall events with later peak timings are more destructive than those with earlier peak timings," and its previous sentence "Similar conclusions have been drawn by Cheng et al., stating that floods with shorter time lags and later peak timings intensify the impact of flooding" are all derived from the work of Cheng et al. (Cheng, T., Xu, Z., Yang, H., Hong, S., and Leitao Joao, P.: Analysis of Effect of Rainfall Patterns on Urban Flood Process by Coupled Hydrological and Hydrodynamic Modeling, Journal of Hydrologic Engineering, 25, 04019061, 10.1061/(ASCE)HE.1943-5584.0001867, 2020.).We will rewrite these sentences in the revised manuscript to clearly indicate that they are conclusions from other researchers and to cite the original source correctly.

15. The authors suggest a set of equations (from eq. 30 to eq. 40), but poorly explained. Consider elaborating more about the governing process of equations. Additionally, it's mentioned that “calculate the range of rainfall and tidal level design values for different durations from 2-yr RP to 200-yr RP using the following formula” (L286,), but the formula doesn’t give the range of rainfall, and since it divides the range by 200yr Rainfall value. So, ΔX_t will be a dimensionless parameter that is related to the amount of variability (range) of rainfall. Same for Tides. Same in L304, ΔVx_t and ΔVy_t do not quantify the variation of flooding volume. Accordingly, check Dx_t and Dy_t.

Response: Thank you for your detailed review of the equations and related calculation methods in our paper. We acknowledge that our explanation and description in this section were indeed insufficient, and we apologize for this. We will rephrase the control process of the equation set in the revised manuscript to help readers better understand the physical significance and application value of these equations.

Regarding your mention of L286, "calculate the range of rainfall and tidal level design values for different durations from 2-yr RP to 200-yr RP using the following formula." The original intention was to express "calculating the variation range of rainfall and tide levels from 2-year to 200-year rainfall and tide levels," where the rainfall and tide levels are derived from the design values in Section 3.3. Our incorrect expression misled your understanding, and we apologize for this. We will correct this statement and re-examine ΔVx_t, ΔVy_t, Dx_t, and Dy_t to ensure that all parameters and variables are clearly and correctly defined.

16. The paragraph from L543 to 558 is more about discussing the results. Consider moving it into the discussion section.

Response: Thank you very much for your suggestion, which is very helpful for improving the logical flow and structure of our paper. We will adjust the placement of this section to ensure that the overall structure of the paper is clearer and more coherent.

17. The paper is lengthy. Consider moving some of the extra results to the supplementary materials. For example, Tables 1 and 2. Commonly used equations, such as those for correlation and distributions, can be omitted from the text. Instead, it is sufficient to cite the relevant references for these well-known equations.

Response: Thank you very much for your suggestion, which is extremely beneficial for optimizing the length and organization of our paper. We will make adjustments and refinements to the manuscript as per your advice.

18.Check the way of citing publications in the manuscript according to the HESS guideline (e.g., L40, L42, L50).

Response: Thank you for pointing out the issue with the citation format. We will thoroughly check and revise the references to ensure they strictly adhere to the HESS guidelines.