The authors present a detailed analysis of various indicators typically used to assess the compound flood potential for 11 locations along the coast of China. It is clear after reading the manuscript that the authors have put a substantial amount of work in the execution of the methodology: selecting different thresholds to quantify the statistical dependence, looking at the influence of seasonality, sea-level rise and weather patterns for marginal or joint extremes of storm surge and precipitation. The main objectives of this paper do not appear scientifically novel to me but rather a thorough application of current methods. If the goal of the paper is to provide new insights on the compound flood potential in China, I would discuss this further in the discussion and highlight in the conclusion how your findings complement or contrast results from other local or global studies. Alternatively, another journal like NHESS could be more suitable to report such findings since I think that the fact that the paper focuses on the Asian coastline is a particularly relevant point for risk assessments. I listed below a few major and minor comments for the authors to consider.

Major comments:

- Throughout the paper, it seems that the authors interchangeably use the terms ‘flood’ and ‘event’ which is very confusing. In Figure 2, the authors clearly state what they define as a compound or non-compound event. However, these events (points on Figure 2) do not necessarily generate floods. Yet, this confusion is omnipresent throughout the manuscript, for example:
    - The title mentions “compound flood events” whereas the abstract mentions that “This paper investigates the potential compound effects”. Those statements have very different implications when interpreting the results and conclusion.
    - The authors mention three different definitions of ‘compound events’: the first one from Zscheischler et al. (2018), the second one from Wahl et al (2015) and the third one suggested by the authors in Figure 2. These three definitions are different so the authors should be clear about this and discuss the limitations of this selection in the discussion section. As correctly mentioned in the introduction, Zscheischler et al. (2018) could consider any combinations (also both non-extreme) to be a compound flood: selection has to be done based on impact (which is not known in this case). Wahl et al (2015) would consider any points in Zone 1, 2 or 3 to be a compound flood. I appreciate the fact that the authors are clear in the text and always mentioning when they refer to Zones 1/2/3 vs Case 1/2 but it becomes very confusing when interpreting results and conclusions in terms of compound and non-compound. The dependence and frequency analysis is done in terms of Case 1/2 but the weather maps and analysis of the typhoon dataset are done in terms of the Zones. In both sections, the terms ‘compound’ is used but I am not sure anymore what it really means as it refers to different areas in Figure 2. Clearly there is some value in this analysis but the discussion and conclusion have to be carefully rephrased to express the limitations of these definitions.
- The analysis based on the typhoon database is interesting but highly uncertain, especially when generalized with respect to compound/non-compound events. The authors

acknowledge that convective rainfall events are probably excluded from the typhoon dataset but no information on the damage from these events caused is added. Yet, conclusions about compound/non-compound flood events are made. As the authors state on line 316, we do not know whether those events lead to no damage or significant damage. This could lead to very different conclusions than the ones presented here. I find this analysis interesting but I would recommend acknowledging the fact that you focus only on typhoons for this analysis and instead show the influence of both drivers on damages when only considering typhoons, and not generalize it to compound/non-compound events.

- Did the authors consider comparing their results based on skew surge instead of storm surge? When performing a tidal analysis, small errors in the phase of the tide can lead to large storm surge peaks. This could have a large influence on your correlation. The authors mention on line 106 that the data has been checked for common errors but do not elaborate further.

Other comments:

- Convective rainfall is discussed is the discussion section (section 5) but is actually not mentioned when describing the weather patterns (section 4.4). I would recommend introducing this weather pattern earlier if you mention it for Hong Kong, this will help the reader understand all weather systems conducive to flooding.
- The authors mention on lines 303-305 that few regional assessments from hydrodynamic models have been conducted for compound flooding. Such analysis has been conducted at the global scale and it could be interesting to comment on this with respect to the patterns found in your study:
  Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H. C., & Ward, P. J. (2020). The effect of surge on riverine flood hazard and impact in deltas globally. Environmental Research Letters, 15(10), 104007.
- Some limitations are discussed in the conclusions (paragraph starting in line 350). I would move those points and elaborate them in a separate section or combine it with the discussion. Similarly, I would say that the analysis of the typhoon database in the discussion belongs more to the result section than discussion.
- Line 96: “where tropical cyclones impacts are more severe”. I am not sure why it is important to mention this here. Maybe make this clearer and/or add reference to support this because this is not clear to me when looking at Figure 1.
- Line 107: what do the authors mean by “earlier” here?
- Line 115: This is minor but it would be more logical to write sea *level* pressure for SLP instead of sea *surface* pressure
- Line 118: Maybe use the term “Defining” instead of “Selecting” as compound events are described in various ways in this paper.
- Line 137: Maybe change the word “appropriate”. Both annual maxima and POT can be used in this type of analysis as shown by previous literature.
- On Figure 4, I suggest changing the label of the colorbar to highlight that it is a difference. Otherwise, the negative values seem strange at first sight.

- Line 210: “To better understand the timing of events leading to joint dependence throughout the year”. This sentence is not clear to me. I would suggest rephrasing it.
- Line 239-240: “The summer monsoon brings continuous precipitation since June to August in southern China. Thus, the dependence is higher in the summer compared to the typhoon season”. Does this conclusion applies to all the gauges or only the last ones discussed (TG7 and TG10)? It would be useful to elaborate a bit more because I am not sure I understand this as currently phrased. Does the summer monsoon also generate storm surge? If the dependence is higher, this implies that storm surge and precipitation are more strongly correlated. If only the rainfall is higher but the storm surge is random, the correlation will be insignificant.
- Line 326: explain “gale” briefly?
- I would suggest labeling the gauges again on Figure 3b. This makes comparison of both panels a and b easier.
- I would strongly recommend carefully checking the manuscript for typos and other mistakes. Below are a few examples I found:
    - Line 89, 131, 132, 232: spaces missing
    - The description of the zones is sometimes flipped with what is shown on Figure 2. For example in the description of Figure 2, I think it should be “i.e. high precipitation and high storm surge, respectively”(line 134). This is also the case on line 162.
    - In Figure 2 and 8, the x-axis label should be “Precipitation”
    - Line 319: I think the word average is missing in (US$ 5 million per event)?
    - Line 334: remove ‘were’ or add ‘that’