Response to Referee #1:

We would like to thank the reviewer for the time of reading this manuscript and giving suggestions and inputs. Here is our response to the reviewer. The reviewer's comments are in blue ink, and our response follows in black ink. When applicable, the changes made in the manuscript are inserted in italic.

The article entitled "Track-dependency of tropical cyclone risk in South Korea" by Nam et al. addresses the issue of including cyclone tracks as the bridging factor between exposure and actual impacts of tropical cyclones. The article introduces the main problematic in the first section explaining the distinction between "active" hazards (e.g. rainfall and wind surges) and "potential" hazards (tropical cyclone intensity). Then, an extensive presentation of the datasets, tools and methods that the authors used is followed by a discussion of the results after applying a tree decision method that the authors propose for evaluating cyclones hazard. The method introduced in this paper seems to provide meaningful results when applied to the Korean peninsula and I would recommend the article to be published in NHESS.

This is a good way of summarizing our main idea of track information in risk assessment in the paper, and your positive comment is much appreciated!

My only major concern goes on the presentation and organization of the manuscript. In several parts, I found that the text is repetitive on the arguments and methods and thus less attractive to the reader. Furthermore, my opinion is that the results section is too long and difficult to follow. It should be divided to subsections in order to ease the reader with better articulation of the main findings. Finally, although understandable language could be improved.

It is true that we have endeavored to describe the details in our damage dataset and methods. We have revised the entire manuscript to be more concise and straightforward. Especially method section is rearranged and some parts are simplified. Also, following your suggestion, we divided the result section into three subsections: 3.1 TC hazards and risk of different track types, 3.2 Geographical impacts on TC risk distribution, and 3.3 Decision tree analysis results.

I would also recommend to the authors to include a discussion on the potential application of their method to different regions. It seems that the number of track clusters is detrimental for the complexity of the decision tree. Is the applicability of their method jeopardized if e.g. a broader region with complex geographical and exposure issues is taken into consideration?

First, to answer your question, it is true that for different regions, the number of track clusters can be larger and it can add complexity to the decision tree. However, there are various ways to incorporate track information into decision trees, because decision trees

can deal with both of numerical and categorical data. There are various ways to incorporate track information into decision trees. We adopted clustering method and made it categorical, but for other regions, track information could be divided into two variables such as the approaching angle and distance from the coast.

We have already included discussion of other methods of risk analysis for the potential application to different regions in the discussion section. However, we did not include how to apply the same method of combining fuzzy c-means clustering method (FCM) and decision tree model, for it is very detailed. Instead, we elaborated the reason why we adopted FCM in more details, so that the authors can have better ideas of dealing with track information in their research, if they want. The revised paragraph in the method section is as below:

"The 85 selected influential TCs are then grouped according to their track patterns using the fuzzy c-means clustering method (FCM). We clustered the track patterns, considering only the part of the tracks in the domain of 28°N–40°N and 120°E–138°E (grey boxes in Fig. 2) so that we could divide tracks focusing on the paths near South Korea, whose national TC risk distribution was examined with respect to these clustered track patterns. Best-track data have 6 hourly longitude and latitude location information of the TCs, and there can be other ways to preprocess track information with the purpose of categorizing their pattern in the area of interest. For example, one can group them with a certain longitude criterion (e.g. east versus west from 128E) or the approaching angle criterion (Hall and Sobel, 2013). We chose to use the FCM, for it is widely used for objectively dividing widespread data with amorphous boundaries. Some previous studies have shown this method to be effective for grouping TC track patterns (e.g., Kim et al. 2011)."

Finally, I would recommend a demonstration of their method application for the ensemble members of a tropical cyclone forecast. What is the hazard variability when the spread of the forecasted tracks is relatively large?

We introduced decision tree method primarily to assess the relationship of tropical cyclone damage with risk elements. In order to apply these results for prediction, we recommend using random (decision) forest method, which is, in a nutshell, a collection of decision trees. In fact, we are currently studying how to forecast risk with the forecasted tracks and the following hazard forecasts with random forests model. We included one paragraph at the end of the discussion section in our manuscript as below.

"Finally, if one wants to apply the decision tree methods in other risk analysis, note that decision trees have advantages to other data mining methods that it's easy to interpret, generating straightforward visualizations, but it is prone to overfitting and errors due to bias and variance. It's because decision tree determines an optimal choice at each node. Choosing the best answer at each step does not guarantee the global optimum. We used decision tree method to diagnose the relationship of risk elements in this paper. . If the model makes a different choice at a given step, the final result can be very

different, especially when dataset is small. For the current paper, to prevent these errors, we verified our results with pruning and cross-validation. However, for forecasting not understanding, stable and accurate result is more important than transparent and intuitive results. Random forests trains the model with different sample sets of the data (same mechanism to cross-validation in a way). It incorporates random repetition, which makes interpretation and visualization complicated, but random forests give more robust results that a single decision tree."