Thank you for your valuable comments. We will proofread the manuscript again especially the tense issue and the references. And the reply for all the questions are as below.

Technical Question Part:

1. While extracting the Weibo texts, the keyboard we used was "Beijing rainstorm", which already restricted the places.

2. Microblogs have a 140-character limit and many of them usually contain only one sentence, which leads to less corpus. As many dictionaries of sentiment words, degree adverbs and negative words only contain hundreds of words, most microblogs will be judged as neutral ones with a zero sentiment value. The social media such as Weibo has provided us a lot of open corpus and many studies based on Word2vec have been carried on calculate each word' sentiment value, for example, the paper *A Study on Sentiment Computing and Classification of Sina Weibo with Word2vec* in 2014 IEEE International Congress on Big Data. These works expanded traditional sentiment words' dictionaries and made it more suitable for dealing with short texts such as microblogs.
   Most of the computational time spent on the Chinese word segmentation process. After getting the dictionaries of Chinese words' sentiment value, all the procedures are matching the key-value in the dictionary and add the value together, which cost little computational time. So, considering every word's sentiment value can hardly affect the total computational time.

3. In the paper *Real-time identification of urban rainstorm waterlogging disasters based on Weibo big data* on Natural Hazards 2018, the author stated b was 1, and we utilized it. On the other hand, both positive and false negative results are potential waterlogging points. If we get a lot of potential waterlogging points, we could plan the patrol route to reduce the time and human costs. However, if we miss some positive results with a lower precision, we will never consider some potential points.

4. The parameters will be added into the paper. These 269 microblogs are all the total points in the ArcGIS analysis as only they provide a location information which could be expressed by in terms of latitude and longitude. The area is 1,050 square kilometers and the point density is almost 0.25 in each square kilometer. After 2017, Weibo has forbidden personal account getting the microblogs' location information, so the number of points is small. This part is an attempt to illustrate that we can find out the potential waterlogging points through the Weibo big data, and the accuracy is discussed in the following. Maybe it is not accurate enough, but the result can show us the level of waterlogging to some extent.

5. In the social media, the expressions are flexible and changeable. In different issues, people will use different expressions to describe the same targets. For example, dragon kings, boats, rivers, etc. can be all used to describe a rainstorm, while in case A, many microblogs use dragon kings which are hardly mentioned in case B. So, the texts from the same sources are more similar and this leads to a higher accuracy. We checked some misclassified texts in case B, and mentioned that many errors

were caused by unique expressions. For example, a text says "the weather is good for fish-farming". "Fish-farming" is a new expression for bad weather in case B so the model don not understand it. If we have a larger training set including this expression, it may be classified correctly. Therefore, we suppose that a larger training set will improve this result. As for the optimal data set size, we hope it can contain all the expressions.

6. When we read the paper *Real-time identification of urban rainstorm waterlogging disasters based on Weibo big data* on Natural Hazards 2018, we observe that the author only extract microblog texts from one disaster. So, in this paper, one of the target is to verify whether the model learned from a disaster can be used in another or not. So we designed the 'Migratory validation', in which we use the model from case A to analyze the texts in case B. If we combine the Case A and B together, it seems that they combined to form one new disaster. We will train and test the texts in the same source again. We will try out this idea in future studies.

Specific Question part:

1. From the urban area's maximum precipitation, it is a 100-year storm, while from the total area's maximum precipitation, it is a 500-year storm. As we mainly discuss urban rainstorm and waterlogging disaster in this paper, we consider it as a 100-year storm.

Before modification	After modification
On July 21, 2012, a heavy rainstorm swept Beijing	On July 21, 2012, a 100-year rainstorm swept Beijing

2. The source comes from the Website of China Mountain Flood Disaster Prevention and Control：

   http://www.qgshzh.com/show/db15c68a-2f3c-429d-afcd-27e279f95e40
   We have added this reference in the paper.

3. This sentence is a conclusion of the next two paragraphs. The second paragraph of introduction discusses the studies of the hydrologic model and some simplified models in these four aspects, while the third paragraph shows some studies based on social network big data. To avoid ambiguity, we modify the sentence.

Before modification	After modification
Therefore, many studies have been carried on urban rainstorm and waterlogging's risk assessment, preventing, forecasting and early warning.	Therefore, we should focus on urban rainstorm and waterlogging's risk assessment, preventing, forecasting and early warning.

4. We have added the reference for this statement.
   Zhang Q, Wu Z, Zhang H, et al. Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China[J]. Journal

of Environmental Management, 2020, 271: 110951

5. We have added the reference.

   Rossman L A. Storm water management model user's manual, version 5.0[M]. Cincinnati: National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, 2010.

6. It comes from the *Microblog User Development Report 2020,* and we have added the reference.

   https://data.weibo.com/report/reportDetail?id=456

7. The statement is not precise, and we remove this sentence.

8. We have added a summary for Nair's research as they found that Random Forest performed the best in the research.

9. We have changed the word 'transform' into the word 'repost'.

10. We have removed this paragraph to simplify the paper.

11. We will reshape the paper's structure after receiving all the comments. Your suggestions are very valuable to us.

12. 'Higher accuracy' is an incorrect expression, and we have modified it into 'higher precision'.

13. '10 words' means 10 Chinese characters and we have illustrated it in the paper.

14. We will replace them with the standard ArcGIS plot.