Comment on nhess-2022-48

Anonymous Referee #1

Referee comment on "Real-time urban rainstorm and waterlogging disaster detection by Weibo users" by Haoran Zhu et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2022-48-RC1, 2022

In this paper, the authors try to use NLP to real-time detect the rainstorm and waterlogging issue from the social media information. This study is interesting and essential in the big data era. However, the structure of the manuscript need improve since some paragraphs in the results should be in the methodology section. Besides, I highly recommend authors to conduct reading proof for the overall manuscript, especially for the tense issue. Also, there are many sentences in the introduction section without any reference.

Technical Question:

- In the data collection section, authors state that they extracted Weibo text based on the keywords within certain period. I am wondering if the texts post/repost from other place (different cities or countries) have been screened out as well. If not, how will this affect the results?
- In the word segmentation line 128, this study mentioned 'we prefer the second method' which is that every word contributes to sentiment value. Could you give any proof or reference that the second method is more preferrable? Also, if considering every word's sentiment value, will it affect the computational time and how much?
- In the data classifying and effect evaluation section, this paper stated b is 1, which means that the precision and recall ratio are equally important. Could you give the explanation why this value was used? To me, it seems that the recall is more important since we don't want many false negative result (Supposed to be waterlogging issue but was tested no issue).
- In the result section, authors mentioned the IDW (inverse distance weighted method) and local polynomial interpolation. Could you provide the detail information in the methodology section (no info in methodology at all)? The parameters are also

important in the interpolation method in ArcGIS, like cell size, power etc. Also, at line 259, 269 microblogs were selected out with location information. Is this 269 the total points in the ArcGIS analysis? Also, could you give the area of these 269 points to see the point density? The interpolation method might not be accurate enough if you have large area but only 269 points.

- In table 3, I am just wondering why the precision is so different between 'within case A' and 'Migratory validation'? An explanation should be added for this issue. At line 228, 'but it could be further improved with a larger training set', how do you make this conclusion for the validation set? How large the dataset will be optimal for this condition?
- In the methodology, it seems this study only use the Case A or part of Case A as training set. Have you tried combine the Case A and Case B, then select 80% of both as training test? It could be interesting to see how it affect the accuracy of the model.

Specific Question:

- Line 22, 'heavy rainstorm swept Beijing', give the precipitation information of the storm, how much is considering as heavy? 10-, 50-, 100-year storm?
- Line 23, please provide the reference or source of the death and economic loss.
- Line 26, please provide the reference for 'many studies…'
- Line 28, please provide the reference for 'many factors…'
- Line 29, please provide the reference for SWMM, the SWMM was developed by US EPA (environmental protection agency). They have their own official SWMM reference, it would be better to have it.
- Line 40, please provide the reference for 'As of March 2021,…'.
- Line 45, please provide the reference for 'as there is usually only one strong earthquake.'
- Line 46, please summarize the result of Nair et al. 2017
- Line 52, what does transform mean, like repost/share?
- Line 60, this paragraph is not necessary to me.
- Between line 175 and line 287, please move all the methodology part to methodology section and restructure both, like how you deal with data, how you generate the plot, how to divide the train test and training set, etc.
- Line 227, 'higher accuracy' did you mean 'higher precision'?
- Line 229, '10 words' did you mean 10 Chinese characters or 10 words after word segmentation?
- For figure 4,5, and 6, please check the standard ArcGIS plot. These figures missed the boundary box, legend information, north arrow and scale.