Reply to interactive comments on: Towards identification of critical rainfall thresholds for urban pluvial flooding prediction based on citizen flood observations, hess-2017-751

Reviewer #1

Comment 1.1: *Does the novelty/contribution lie in the use of crowdsourced data? If so, the paper is not written as if it is. If the novelty is indeed in the use of crowdsourced data, then the paper should focus on the crowdsourced data and make more of a discussion/examination of the use of the data to make a strong case of its associated difficulties/advantages. As it is, the introduction and lit review are only in a general sense with no focus on crowd sourcing. And only about 1/3 of the methods and results are based on crowdsourced data*

Response 1.1:

The main novelty of our work does not lie in the use of crowdsourced data, but rather in the formulation of a data-driven approach to predict urban pluvial flooding without the need to run hydrodynamic models. An important outcome of this approach is the derivation of critical rainfall thresholds, which can be used to predict flood occurrence and test the performance of urban stormwater systems. To clarify the novelty of our approach, we will highlight this more explicitly in Introduction (subsection 1.3).

Comment 1.2: *I am not sure if the historical complaint reports used can be called "crowdsourced". The reports are in the order of 10-100 per day (as inferred by Fig. 2). However, in today's context, crowdsourcing commonly refers to sourcing from a large pool of people using the Internet, smartphones, surveillance cameras etc. to obtain observations in the hundreds to thousands to millions. Thus, it may be misleading to describe the observations referred to in this paper as crowdsourced.*

Response 1.2:

Thanks for this comment. In fact, many definitions of crowdsourcing and related terms like citizen science and citizen observatories can be found in the literature, covering a wide range of data sources and collection methods. After reviewing the literature we agree that the term citizen observatories better fits the nature of our dataset than crowdsourcing, following definitions given in (Buytaert et al., 2014; Herman Assumpção et al., 2017). Hence, we will adopt this terminology instead.

Comment 1.3: *There could be a temporal resolution mismatch between the flood complaint reports used and a storm event. That is, in a storm event, it may not be possible to determine*

which reports coincide with the peak rainfall. Thus, there are some uncertainties in the "crowdsourced" observations that are nontrivial and may affect the validity of the methods/results.

Response 1.3:

We conducted our analyses at daily resolution, as citizen reports were captured at this temporal resolution; sub-daily resolution information was not available. Within the daily time window, we considered the maximum rainfall depth at three different temporal resolutions: 15 minutes, 60 minutes and 24 hours. This approach is similar to the one adopted, for instance, by (Spekkers et al., 2015). Given that flood response in highly impervious areas like the districts in Rotterdam is very short (less than an hour up to maximum a few hours), the gap between rainfall peak and flood observation is likely to be small. We will emphasize this particularity of urban flood response better in the Methods section (subsection 3.2).

Comment 1.4: *The paper found a strong correlation between surface imperviousness and the number of flooding reports and concluded that "there is some explanatory power behind degree of imperviousness as an urban pluvial flooding parameter". However, the correlation is likely a spurious correlation due to the correlation between population density and the number of flooding reports, and population density and imperviousness. Thus, this conclusion of the paper is not well-justified. The may be advisable for the authors to use other methods (e.g. multivariate linear regression) to exclude the effects of population density and arrive at a more justifiable conclusion.*

Response 1.4:

Thanks for this valuable comment, we agree that multivariate analysis will add value to the study. In the revised version, we will add results from a multivariate regression analysis. In other words, we will analyze the relationship between the number of reports and the imperviousness, rainfall intensity, and the population density. Furthermore, an analysis of variance (ANOVA) will be added to find the dominant factors explaining variability in the number of flood observations.

Comment 1.5: *Other minor points: 1. Page 4, Line 7, km^2 instead of km2 should be used. 2. Page 5, Line 10, I guess the area of the green roof park should be 40,000 m2 instead of 40.000 m2. 3. Page 6, Line 25, equation (1) should be appeared before equation (2).*

Response 1.5:

Thanks for pointing out this issue, we will correct this.

Comment 1.6: *The authors should also remove duplicated flooding reports in their temporal and spatial correlation analysis.*

Response 1.6:

In the Results section, we explained our method to identify and delete duplicated flooding reports (subsection 4.2, page 13). However, this explanation should have been placed in the Methods Section; this will be corrected in the new version of the manuscript.

Comment 1.7: *Instead of using the rainfall intensity at the center of the study area for the temporal correlation analysis, I would suggest the authors to use the aerial average rainfall intensity of the whole study area. This is to take into account the rainfall spatial variability, which could be rather high at the scale of the study area.*

Response 1.7:

We agree with this comment, which was also raised by Reviewers #2 and #3.

We will redo the analysis, using an aerial average rainfall intensity, instead of rainfall over the central radar pixel.

Comment 1.8: *What is the unit of dt, mm or mm/h? It seems to be mm/h according to the definition in equation (1), but is shown to be mm in Figures 2, 3, and 4. The authors should also provide more clarified explanations for the variables in equation (1).*

Response 1.8:

The coefficient d_t stands for the rainfall depth within a given time window and has unit mm. The time windows t that we adopted are 15 minutes, 60 minutes and 24 hours, respectively. So they are expressed as $d_{0.25}$, d_1 and d_{24} in terms of hours.

We will make this clearer in the revised version.

References

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T.C., Bastiaensen, J., De BiÃ¨vre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G., Zhumanova, M., 2014. Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. Front. Earth Sci. 2, 1–21. doi:10.3389/feart.2014.00026

Herman Assumpção, T., Popescu, I., Jonoski, A., Solomatine, D.P., 2017. Citizen observations contributing to flood modelling: opportunities and challenges. Hydrol. Earth Syst. Sci. Discuss. 1–26. doi:10.5194/hess-2017-456

Spekkers, M.H., Clemens, F.H.L.R., ten Veldhuis, J.A.E., 2015. On the occurrence of rainstorm damage based on home insurance and weather data. Nat. Hazards Earth Syst. Sci. 15, 261–272. doi:10.5194/nhess-15-261-2015