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 #2

Anonymous Referee #2

Comment 2.1

Setting and boundary conditions of the study are very case specific (flat catchment, specific drainage infrastructure and operational regime (OP), rare data availability (OP data, flood reports)). Hence, transfer of findings to other cases - without having carried out similar analyses for other systems - would be, at least questionable. This clearly lowers the scientific significance of the work (reproducibility). Despite the fact that various interesting aspects of general relevance are discussed (use of citizen-reported flood incidents as ground truth data for urban flooding, change point analysis), identified (non-)correlations as well as rainfall thresholds – which relate to the key research questions in the paper - are exclusively valid for the particular case Rotterdam. I am wondering why this clearly limiting aspect has not been discussed in the manuscript. I strongly encourage the authors to address this aspect adequately, e.g. by clearly labeling results and findings as case specific (title, abstract, conclusions) and discussing the relevance of findings for other systems.

Response 2.1

Thank you very much for reviewing the paper and making helpful comments.

One of the aims of our study was to investigate the potential of citizen science data to identify flood-prone areas and thresholds that trigger flood occurrence. In addition, the thresholds can be used in a flood-early-warning context. We agree that the threshold values that we derived are specific to this case study. However, conclusions with respect to the use of this type of datasets and applicability of the threshold method are generic.

We will clarify this distinction more explicitly in the introduction, discussion and conclusions.

Comment 2.2

The title prominently suggests research on a currently popular topic: the collection and evaluation of crowdsourced data to extract meaningful information. The main text then reveals that 'crowdsourced data' are here understood as structured recordings on flood incidents reported from the public (!?) which the researchers "obtained" from an existing database (!?) – cf. page 3, line 39 -41. It is not entirely clear to me if the term "crowdsourced" refers to the fact that different people, i.e. the crowd reported the incidents or that the recordings are received from various different sources but are then formalized and archived. In a way this issue is somewhat peculiar since in previous publications the same data set had been named "citizen-reported flood incidents"... Irrespective of the fact if data used in this study can be referred to as "crowdsourced data" the novelty aspect (in the current version of the manuscript) is marginal. Hence I suggest reconsidering the title formulation or a thorough revision of the paper shifting the focus to the use of crowdsourced flood reports as such, e.g. discussing the quality of this source of information.

Response 2.2

We agreed. A similar point was raised by Reviewer #1.

We will replace the term “crowdsourced data” by the more general term “citizen observatories” and refer to (Buytaert et al. 2014; Paul et al. 2018), who summarized definitions of various citizen science-associated terms.

Comment 2.3

Using rainfall threshold values (which do not account for the spatial structure, moving patterns of storms) to warn for area-specific flooding incidents is a bit far-fetched. From my point of view, the presented results are not convincing enough to allow lumping changing factors (such as downstream drainage system behavior, operational regime, spatial rainfall variability) which eventually influence the degree of flooding and the location at which flooding occurs into a single rainfall threshold value that predicts flooding in a particular urban subcatchment. I suggest using a physically-based dual drainage model to systematically partition the influence of these aspects and to so put conclusions on a more solid basis.

Response 2.3

We agree that the spatial resolution used in our analysis to derive rainfall thresholds introduces uncertainty, as can be seen from the explanatory power that can be achieved. A higher resolution analysis for threshold identification was not possible due to limitations in the number of available data (see also results of spatial analysis in section 4.2). With the growing availability of citizen science data, more detailed analyses will be possible in the future. We will mention this point more explicitly in the discussion section.

Using a physically-based, high resolution model is unlikely to provide additional insights, because, as ten Veldhuis et al. (2010) have shown, models account for only a small part of flood-generating mechanisms in urban drainage systems. Additionally, the reliability of even very detailed physically-based models depends on availability of observational data to verify parameter assumptions. Thus far, no models have been shown to have satisfactory performance for the complex drainage system in Rotterdam.

An important motivation for our study was precisely to investigate the relationships that can be derived based on a data-driven analysis. Our investigation relies directly on citizen science data and reveals relationships between flooding and rainfall depth, an approach that is more directly accessible in many cases than complex models.

Comment: 2.4

The spatial correlation analysis between interval-specific rainfall depths and reported surface flooding observations leaves me a bit puzzled. Beside the fact that identified correlations are at the very low end in terms of occurrence and significance, the following points are at least debatable: i) As this particular part of the analysis is based on only six events I am asking how representativeness is ensured. I do not fully understand why the number of reports must be greater than 40 to allow an event to be included in the analysis. No justification is given on what the selection of this threshold is based on, nor a sensitivity analysis is conducted to show how results alter in case more events (with less reports) are selected. In any case it should be questioned to what extend six events provide enough input for a spatial correlation analysis to come to a meaningful conclusion. ii) Disentangling dependencies: the weak but still existing correlation between population and report density suggests an inherent dependence between the two variables: the more people live in an area, the more reports can be submitted. Hence it should be discussed to what these variables can be independent at all! Depending on the way how flood reports are submitted, normalizing over population density may be a first step to research this aspect. Technically different, but similar with regard to the dependency aspect, the statement ". . . imperviousness was confirmed to be an appropriate parameter to predict urban pluvial flooding" (p. 17, line 2-3) is somewhat trivial, i.e. misleading since i) yes, sealed surfaces produce higher surface runoff and ii) solely considered, the degree of imperviousness does allow a prediction of flooding potential – it must be considered in context with other factors. A multivariate analysis approach is recommended. iii) Considering drainage network capacity constraints: the spatial analysis somehow ignores the fact that urban flooding can substantially be influenced by hydraulic behavior of the actual drainage network (e.g. hydraulic capacity constraints further downstream in the network may lead to manhole overflows). In other words: here flooding is expected to occur right where the rain cell is present, suggesting that the main cause for surface flooding is the pure amount of rainfall at the spot maybe combined with a limited capacity of street inlets. It remains an open issue to what extend the found spatial correlation is influenced/biased through this aspect (still, it is outlined in the outlook for further research – p. 17, line 37) and if this could be a reason indeed for the decreasing correlation when increasing the spatial resolution of data (finer grid). Moreover, this aspect is very likely to become more relevant when researching systems with higher terrain elevation variability, i.e. elevation difference in the catchment.

Response 2.4
We agree with your comments.

In the new manuscript, we will implement the following analyses as you suggested: a multivariate 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 implemented afterwards to find the dominant factors without the need to implement normalization.

In regard to the number of reports, it is a compromise made to ensure there is at least one report per pixel. A selection of less than 40 reports leads to an event with no reports at some given pixels.

In fact, our study area is flat and the sewer (sub)districts that we consider are relatively small-sized, most of which are in a scale of five squared kilometers with multiple outflow points around the district borders. Therefore, we regard the rainfall as the main cause in the spatial analysis. We will provide more proof for this statement in Section 4.2.

Comment 2.5
The paper's layout is corrupt at many points, it seems that the manuscript had been submitted in

a rush.

Response 2.5
We realized after submission that there was problem in converting doc file into pdf file. We will pay close attention to this issue for submission next time.

Comment 2.6
Scatter plots, especially Fig. 2, 3, 5, are difficult to read. In particular outliers are difficult to spot.

Response 2.6
We agree with the fact that there is a problem of how we deal with and plot the outliers in an arbitrary way.
In the new version of the manuscript, we will implement the robust regression to eliminate the effect of outliers. Particularly, robust regression works by assigning a weight to each data point. Weighting is done automatically and iteratively using a process called *iteratively reweighted least squares*. In the first iteration, each point is assigned equal weight and model coefficients are estimated using ordinary least squares. At subsequent iterations, weights are re-computed so that points farther from model predictions in the previous iteration are given lower weight. Model coefficients are then recomputed using weighted least squares. The process continues until the values of the coefficient estimates converge within a specified tolerance.
Finally, we will compare the results of robust regression with those of linear regression (Section 4.1).

Comment 2.7
The discussion of particular events in the text referring to Fig. 4, 5 (p. 10, line 10, 22) is useless unless it is indicated in the graph and has a particular meaning.

Response 2.7
Thanks for this comment.
We will make an explanation in a better way. The effect of outliers on the regression performance will be included in the new analysis of the robust regression as described in Response 2.6.

Comment 2.8
The treatment of outliers in the OP data is occasionally fuzzy (p. 9, line 5ff; p. 10, line 10-11) and sometimes arbitrary (p. 15, line 4-6).

Response 2.8
Agreed.
We will replace this by a robust regression analysis (see the previous Response 2.6) and meanwhile we will compare the robust regression with the conventional linear regression in the paper.

Comment 2.9
The authors should revise the analysis to avoid the impression that leaving out particular data

was done to let results look a bit better.

Response 2.9

Agreed.

Applying a robust regression to the dataset will ensure objective identification of outliers.

References

Buytaert W, Zulkafli Z, Grainger S, et al (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

Paul JD, Buytaert W, Allen S, et al (2018) Citizen science for hydrological risk reduction and resilience building. Wiley Interdiscip Rev Water 5:e1262. doi: 10.1002/wat2.1262