Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2019-272-AC1, 2019

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Urban pluvial flood risk assessment – data resolution and spatial scale when developing screening approaches on the micro scale" *by* Roland Löwe and Karsten Arnbjerg-Nielsen

Roland Löwe and Karsten Arnbjerg-Nielsen

rolo@env.dtu.dk

Received and published: 6 November 2019

1 Author's summary

We wish to thank the reviewer for taking the time for a thorough review the manuscript and for providing constructive comments. In brief, we agree with the issues raised by the reviewer and will adress them as detailed in our reply to each comment below.

This involves the following major changes:

- *New discussion section "5.5 Generalization and application" which includes a stepwise workflow for deriving suitable scales in a new case study and discusses the limitations linked to topography and urban layout raised by the reviewer*
- *Streamline terminology and symbols related to the different data resolutions in the workflow figure (Fig.3) and throughout the text.*
- *Include scatterplots showing the effect of data transformations in flood damage regression in the supporting material*
- *Improve explanations in the manuscript as suggested by the reviewer in the detailed comments.*

Both reviewers point out language issues, so we suggest that will have the manuscript proofread by a language editing service before final submission.

2 Reviewer Summary

Review of manuscript "Urban pluvial flood risk assessment - data resolution and spatial scale when developing screening approaches on the micro scale" by Roland Löwe and Karsten Arnbjerg-Nielsen submitted to NHESS The authors present a study analyzing the impact of aggregation scale of high resolution DEM, imperviousness and building data on urban pluvial flood risk assessments. The study intends to quantify these impacts and to identify the optimal scale for data aggregation to be used in "flood screening", i.e. for low computational flood hazard and risk assessment considering different flood adaptation scenarios and urban developments. The authors thus deal with a topic that has been of a long standing concern in flood risk research and add an at least useful, but potentially also important contribution to the question of optimal scales to be used in flood risk assessments, here with a particular focus on urban pluvial floods.

This is an appropriate summary of our study.

3 Major comments

The study is generally well designed and presented, the data analysis solid and the conclusion are supported by the results. Overall I don't have any major objections to the presented work, but I suggest to enhance the discussion of the implications of the findings for urban pluvial flood risk assessments in more detail, as well as the generalization/transferability of the results. This would enhance the manuscript and increase the potential impact of the work.

In section 5.4 about the limitations of the work the authors state that the regression models likely have to be newly fitted for different topography and urban structures, but that they expect that the identified optimal scales are generic. Unfortunately the authors did not provide any reason why they expect that the optimal scales are generic, i.e. transferable to any other urban flood risk study. This needs to be provided. I actually would challenge this statement. From my experience and understanding of the problem, I would argue that the urban texture/layout also controls the optimal scale for risk assessment. In the context of this work it should control at least the optimal scale of the imperviousness regression.

Think of cities with wide roads and sidewalks designed for car traffic (e.g. American suburbs) vs. old towns with narrow streets and sidewalks and/or steep topography (e.g. old European cities with medieval city centers). It can be reasoned that at least the optimal resolution for the imperviousness regression is likely different for these urban structures. If the authors argue against this, proper arguments should be given. Otherwise the limitations of the study results in terms of transferability needs to be

extended.

We agree with the reviewer – optimal scales must depend on the density of urban developments, which can vary between cities. We suggest elaborating on these issues in the discussion section and include them in a new section "5.5 Generalization and application", which will also address the reviewers next comment.

Furthermore, the manuscript would profit if the authors provide recommendation/blueprints of how the presented optimal scales and regressions can be used in other urban flood risk studies/assessments and assessment of flood management/mitigation/urban development plans. What would be the procedure to follow? What are the minimal data and model requirements? This is currently a bit blurry and not well defined. A more detailed illustration of the use of the results/findings would surely increase the uptake of the study in research as well as in practice.

The new section "5.5 Generalization and application" will include a stepwise workflow towards creating a screening setup for flood risk in a new case study.

4 Detailed comments

Besides these general concerns, I have some specific small comments listed below.

The term "flood screening" should be explained/defined in the introduction. The authors expect the reader to be familiar with the term, but this cannot be assumed. Moreover, the term is not widely used (to my knowledge), and thus different readers are likely to associate different meanings to the term.

The term will be defined in the introduction. We also noticed that the term is used in varying ways in the manuscript. We suggest using "Flood screening setup" to refer to the overall setup for fast flood risk assessment (i.e., the combination of fast urban development simulation, simulation of flood hazard and damage calculation), while the *simulation of flood hazard should be referred as "fast flood simulation".*

I found it occasionally difficult to follow the different aggregation scales used in the different analysis (Δx_{fit}, Δx_{pred}). Additionally different terms are used in the manuscript, e.g. Δx_{fit} as fitting resolution or data resolution. This should be harmonized. Additionally it would be beneficial to clearly separate these terms in order to easy the understanding of the work done in the different sections, although I also don't have a precise suggestion how this can be achieved. One way could be a clear definition at the start of the method section, e.g. in a table:

Symbol	Description as used in text	Explanation used in analysis xy
Δx_{fit}	Data resolution	....
Δx_{pred}	Prediction resolution	....

The description should then be used constantly throughout the text.

There are 3 resolutions to distinguish:

- *Δx_{fit} – being the data resolution used when training the regression models (varied from 25 to 2000m - both, in imperviousness and damage regression)*
- *Δx_{pred} – being the data resolution at which predictions are generated from the regression models (varied between and 25 and 2000m for imperviousness regression, and kept fixed at 2000m for damage regression)*
- *Δx_b – being the resolution of the building data used for predicting imperviousness as input to the 2D flood simulations, as well as to compute the flooding building area as input to the damage regression models*

We prefer clarifying the usage of different data resolutions in Figure 3 over inserting a new table, because we would expect that the reader would try to understand the

dataflow from this figure. The figure will make explicit reference to different data resolutions used in the different parts of the analysis. In addition, we suggest revisiting the text and inserting explicit references to Δx_{fit}, Δx_{pred} and Δx_b when discussing resolutions.

The regression results are compared to a benchmark simulation based on highly detailed input data. This is totally valid, but ideally a quality statement of the benchmark should be provided. If there is no quality assessment of the benchmark possible (because of lacking data/observations), then there should be at least a statement that benchmark is not validated and could thus also be (far) off reality. Of course this does not touch the validity of the results, because the benchmark could likely be tuned to be close to reality as possible if validation data is available.

We will include a corresponding statement in the Methods section. As reasoned by the reviewer, our aim was to generate flood map which is realistic rather than to reproduce observed conditions.

In Figure 3 and associated text it is stated that only 8 aggregation levels (resolutions) for imperviousness (simulated flooded areas) are used for the regression of the damage functions, but there is no reason given for the reduction. I assume that this is because of reduction of possible resolution combinations without compromising the overall results, but it needs to be stated. *Indeed, we have performed flood simulations for a limited set of resolutions, because additional simulations require substantial manual effort, provide limited insight and make it difficult to present results in an understandable manner. The statement suggested by the reviewer will be included in the figure caption for Fig. 3.*

In section 4.1 it is stated that the optimal solution derived from Figure 4 is in the order of 400m, because the curves in Figure 4F have a local minimum at about 400m for prediction resolutions of 500m – 2000m. However, the standard deviation of RMSE for a prediction resolution of 250m has no minimum, but is always below the standard deviation of RMSE of the higher prediction resolution for all fitting resolutions. Therefor I cannot really follow the conclusion that 400m is the optimal fitting resolution for estimating the impervious area. This should be explained better. Moreover, the caption of figure 4 should state that it deals with the regression functions of the imperviousness. This is currently missing, thus impairing the understandability of the figure without reading the associated text section.

We will clarify the figure caption and provide an explanation for the artefact at 250m prediction resolution in the main text. A detailed explanation is provided below.

Detailed explanation:

The standard deviation of the estimated regression model coefficients decreases when smaller data resolutions Δx_{fit} are considered during model fitting, i.e., we obtain more stable parameter estimates (not shown). The mean parameter estimates approach 1 for very fine data resolutions (not shown), i.e., the regression models only capture the roof area as impervious area. A strong negative bias is thus introduced in the regression predictions of impervious areas.

When considering large enough prediction resolutions (Δx_{pred}), where the pixels containing the buildings also include all the associated impervious areas, this bias leads to strong variability of the RMSE values computed during cross validation, despite smaller variability of the parameter estimates. The variability is driven by different areas being sampled for validation (e.g., more or fewer industrial areas). The bias disappears when coarser data resolutions Δx_{fit} are considered (leading also to the increase in COD values in Fig. 4D), however, at the expense of fewer data points being available, leading to instability in the parameter estimates and again an increasing variability of the RMSE values computed during cross validation. The data resolution where $\sigma(RMSE)$ is minimal (Δx_{fit} around 400m) indicates the optimal tradeoff, where regression predictions become unbiased, and the data are aggregated only to the necessary level. It is also the resolution where COD values in Fig. 4D reach their maximum.

For smaller prediction resolutions ($\Delta x_{pred} = 250m$), we observed an artefact where the biased regression predictions for small data resolutions Δx_{fit} do not lead to an increase in $\sigma(RMSE)$. In this case, substantial portions of the impervious areas are located in pixels where building areas are 0. The impervious area predicted by the regression models for these pixels is thus always 0 and does not depend on the regression coefficients. The absolute values of RMSE increase due to the bias. However, the variability of RMSE values (Fig. 4F) becomes determined by how much the predictions of imperviousness close to the buildings vary during cross validation. This variability decreases as the coefficients approach a constant value of 1.

In equation (1) the a_i needs to be explained in the text below. For better understanding the meaning of the equation should be explained in one sentence. The statement “we considered the following relationship” has only a vague relation to the text leaving room for speculation/confusion.

Will be adressed

Page 8, line 156: extend the sentence to “Buildings were not explicitly included in the DEM for flow calculation in this case.”

Will be adressed

Section 3.4.1 (page 10, line 215ff): Please provide argument/reasoning for the square root transformation used in equation (6). It is currently unclear why this transformation was performed. Ideally provide a figure in the supplement to justify/explain this transformation. Furthermore the coefficients b_{xi} in equation (6) need to be explained in the text below the equation.

The coefficients will be explained in the paper.

Scatterplots showing the relationship between flooded building area and flood damages (with/without data transformation) and a brief explanation will be included in the supporting material. The scatterplots are also attached in the end of this reply. We *have, in fact, experimented with a number of power and logarithmic transformations. The squareroot transformation turned out to be robust and can handle 0 values. The latter point is a problem particularly with the logarithmic transformation, which amplifies the impact of outliers and where regression predictions of flood damages for pixels with a flooded building area of zero are not guaranteed to be zero.*

Page 10, lines 232-234: to improve understandability, clearly state the difference between baseline flood map and the flood maps based on aggregated building data (buildings in the DEM blocking flows and not) again.

We will include a brief explanation of the baseline flood map in the text.

I would feel more comfortable to use the term "coefficient of determination COD" instead of NSE throughout the manuscript. Both have identical meanings, with NSE being adopted in the hydrological modelling community and typically used to compare simulated and observed (discharge) time series, which is clearly not the case in this study. COD is more widely and generically used. However, this is a suggestion, the authors are free to decide.

We will change "NSE" to "COD" in the text and the result figures.

Page 11, equation (7): explain subscript "CV 2000". I assume that this refers to "cross validation over the 2000m x 2000m sub-areas", but it needs to be explained.

Will be adressed

Page 17, line 368: what is meant here? "while" does not seem appropriate. Maybe ". . ., because coarse representations of imperviousness had little effect on the flow dynamics."

Will be rephrased to "Coarse representations of imperviousness and the resulting change in rainfall-runoff behaviour had little effect in comparison."

Occasionally the English reads a bit awkward/complicated, which is not of major con-

cern for me, but a grammar check by a native English speaker might improve the manuscript further.

We suggest having the manuscript checked by a language editing service before final submission.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-272, 2019.

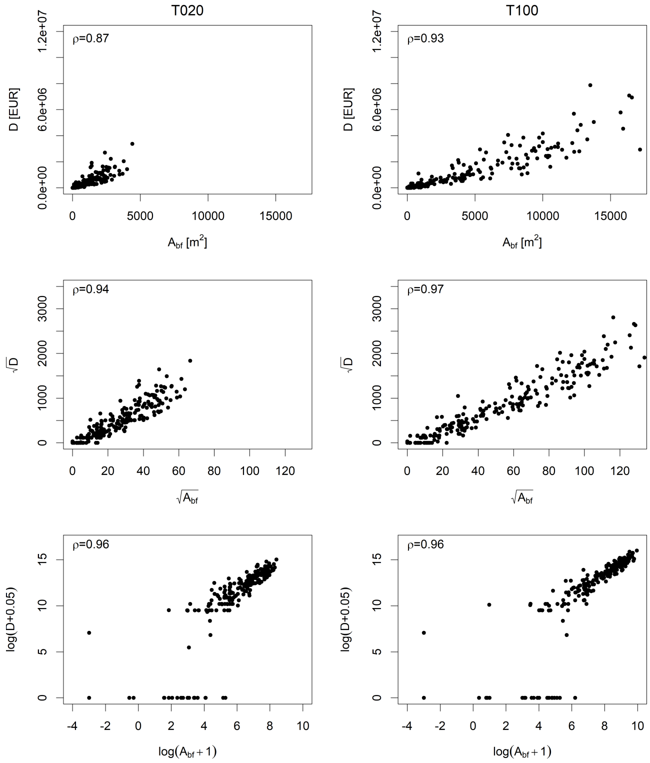

Fig. 1. Scatterplots of flood damages (Olsen et al., 2015) versus total flooded building area (Delta x_b=200m). Columns: T=20 (left) and 100 years (right). Rows: no transform, sqrt-sqrt, log-log transform

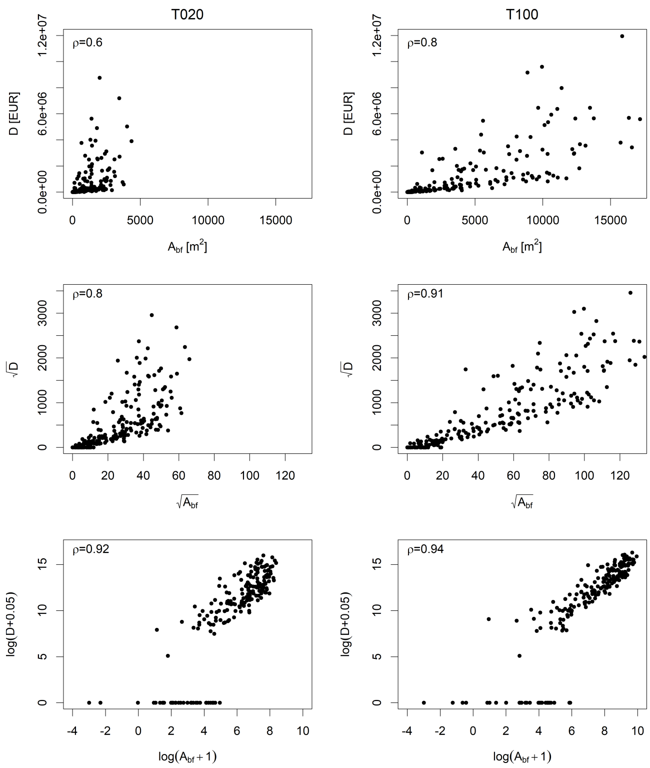

Fig. 2. Scatterplots of flood damages (Beckers et al, 2013) versus total flooded building area (Delta x_b=200m). Columns: T=20 (left) and 100 years (right). Rows: no transform, sqrt-sqrt, log-log transform