**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.

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.

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.

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.

- I found it occasionally difficult to follow the different aggregation scales used in the different analysis ($\Delta x_{fit}$, $\Delta x_{pred}$). Additionally different terms are used in the manuscript, e.g. $\Delta x_{fit}$ 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 |
|---|---|---|
| $\Delta x_{fit}$ | Data resolution | …. |
| $\Delta x_{pred}$ | Prediction resolution | …. |

The description should then be used constantly throughout the text.

- 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.

- 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.

- 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.

- 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.

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

- 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.

- 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.

- 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.

- 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.

- 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."

- Occasionally the English reads a bit awkward/complicated, which is not of major concern for me, but a grammar check by a native English speaker might improve the manuscript further.