Dear Referee #2:

Thank you for the positive comments and constructive suggestions on this paper, which we fully taken into account in the revised version of the paper. In the supplement we address and reply to the questions below.

1. Page 3, lines 16-18, 'However, we find that few researchs use...': I do not think that this is right. Dynamic models have maybe not been used much to test the measures for mitigation by permeable pavement and green roofs, but such coupled models (1d pipe drainage network model and 2d surface flow model) exist and are used for urban flood management (just a few arbitrarily chosen examples: R. Loewe, C. Ulrich, N.Sto. Domingo, O. Mark, A. Deletic and K. Arnbjerg-Nielsen (2017): Assessment of urban pluvial flood risk and efficiency of adaptation options through simulations - A new generation of urban planning tools. Journal of Hydrology 550, 355-367. B. Russo, D. Sunyer, M. Velasco and S. Djordjevic (2015): Analysis of extreme flooding events through a calibrated 1d/2d coupled model: the case of Barcelona (Spain). Journal of Hydroinformatics 17(3), 473- 490. M. J. Burns, J. E. Schubert, T. D. Fletcher and B. F. Sanders (2015): Testing the impact of at-source stormwater management on urban flooding through a coupling of network and overland flow models. WIREs Water 2. 291-300)

Re: Our statement about coupled models is imprecise. Thank you for your kind reminder, and we will modify the expression and here is the revised version of the last paragraph of Introduction.

It is noteworthy that peak flow reduction, runoff reduction, and hydrograph delay are widely used indexes when evaluating the performance of LID practices (Ahiablame and Shakya, 2016; Qin et al., 2013; Zhang et al., 2016). However, these indexes are not very intuitive and how LID practices perform on urban inundation is more beneficial to local residents, such as providing guides for their travel behaviours.Indeed, some 1D-2D models have been applied for flood management such like ESTRY-TUFLOW (Fewtrell et al., 2011), InfoWorks ICM (Russo et al., 2015) and MIKE FLOOD (Loewe et al., 2017). However, most of these models are not free that limits their applications, therefore the open-source model (like SWMM) with LID module that can be coupied to simulate the urban inundation is needed in recently researches (Burns et al., 2015, Wu et al., 2017, Hu et al., 2017).

Therefore, the goal of this study is to demonstrate through a case study the effectiveness of LID practices to mitigate urban inundation in an urban watershed. The specific objectives were to (1) establish a 1D-2D hydrodynamic model coupled SWMM and IFMS Urban; and (2) evaluate the effectiveness of LID practices under different scenarios and hazard levels; and (3) explore the effiency of designed scenarios that related to the effectiveness of LID practices and the proportion of implementation areas. This study hopes to enrich the inundation mitigation research of LID on an urban watershed scale and provide some references to urban stormwater management and inundation mitigation for local government.

2. 2.2 Data, part 4: What were the criteria for removing nodes and pipelines? A reduction from 4502 to 597 pipelines and from 1175 to 653 nodes seems a bit more that deleting some redundant and incorrect data. How was it tested that the data were redundant?

Re: Indeed, the actural drainage networks are compulsory and substantial. Nontheless, SWMM cannot accurately simulate when the data is huge. Besides, after the data conversion process for applying into the SWMM, some overlaps and break points for the pipelines are generated, which makes lots of nodes and pipelines useless. Therefore, we have to simply the drainage data for model building and the criteria shown below:

a. Add nodes when the pipeline is too long;

b. Keep or add the corner nodeschanging diameter nodes, or large variation range of slope nodes;

c. Keep the parallel pipelines and nodes on both sides of the roads;

d. Delete the useless nodes and pipelines in this model.

3. 2.4 Coupling the SWMM/IFMS Urban models: As written above, I think that one does not learn much about the coupling. Also, Figure 4 does not help in this respect. One just learns that the models were coupled. But how were they coupled? Is inflow and outflow from and to manholes possible? What were the criteria for inflow and outflow? What was the spatial resolution of the geometry of a street? What timesteps were chosen for coupling? Either more discussion about the coupling is needed, which means that one also need to know more about the numerical schemes used for the two different models, or it does not make sense to have a section for this part.

Re: SWMM is a 1D rainfall-runoff model which use the given hydrology data and hydrodynamics to simulate the quantity and quality of rainfall-runoff. Notheless, when the node overflow occurs, SWMM can't simulate the spatial and temporal distributions of surface inundation, but the IFMS Urban can using 2D shallow water equations. However, the simulations of IFMS Urban must base on the simulated results of SWMM. So we coupled these two models to realize the simulate on the spatial and temporal distributions of surface inundation. What's more, the process of data conversion and model coupling are all accomplished in IFMS Urban, and it doesn't need other software programming or specialized knowledge, which is convenient for researchers and non-expert users. So we don't want to make it complicated or list algorithm and formula in this part. The spatial resolution of the geometry of a street is 15 m. The timestep of calculation is 10 s and the timestep of output is 200 s.

4. Page 5, lines 18-20: This sentence is unclear. Also: What is innovative about the coupling?

Re: SWMM is a 1D hydrodynamics model which can simulate the quantity and quality of rainfall-runoff but it can't simulate the urban inundation, while the IFMS Urban is a 2D model which can simunate the urban inundation but it must base on the results of SWMM. Through coupling, we build a 1D-2D hydrodynamic model that can simulate the spatial and temporal distributions of surface inundation. Based on this coupled model, we can evaluate the effectiveness of LID from inundation depth, inundation area and inundation time. And this coupled model both takes in the advantages of SWMM and IFMS Urban (open-source, free, great compatibility with ArcGIS and 2D inundation simulation), which is convenient for researchers and non-expert users.

5. Page 5, line 26: Why was a geostatistical method (Kriging) used for interpolation? I do not see the connection to geostatistics for a digital elevation model in a city.

Re: We need DEM when building the 2D model. However, the accuracy of DEM production from Geospatial Data Cloud can not meet our demand (for example, 6 m, 13 m). However, the high accuracy DEM is confidential and difficult to obtain in China. Alternatively, we find the ground elevation of nodes in pipe network data has a higher accuracy (for example, 6.588 m, 13.483 m), and the nodes on the roads are relatively dense. So we use a geostatistical method (Kriging) to get a high accuracy DEM of the roads with the elevation data of nodes on the roads.

6. Page6, top: Please explain why green roofs should not be possible in a dense construction land.

Re: We have not explained the details for this part and thank you for your kind reminder. In fact, there are some special attributes for buildings on the dense construction land in our research area. Through the detailed urban planning and field investigations of our research area, we found the 80% of the residential lands are urban villages, desnsely constructed on construction lands. The structures and shapes of roofs for urban villages are diversity which makes it difficult to build green roofs on them. Thereforce, we temporarily didn't set green roof in the dense construction land in this study.

7. Modeling part (Section 2): Please explain how green roofs and permeable pavements are realized in the model. I assume that a storage for a roof area is assigned (or an existing one is increased) and that there is a soil compartment which gets a connection to the paved area if the pavement is permeable. As this is the key process that is here investigated, I think it is necessary to outline these things (and it is not enough to refer to the manuals of the models).

Re: The simulation designs and parameter setting for PP and GR are listed in Table 1 of our paper, which are strictly desinged according to the manual of SWMM and some highly cited studies of LIDs (Ahiablame and Shakya, 2016; Chui et al., 2016; Kong et al., 2017; Qin et al., 2013).

8. Page 7, top: Please explain why the classification in hazard levels is made. What can be learned from the classification? It is written that the changes of inundation level are different for the different classes. But what does one make out of this fact? More discussion about consequences would be useful.

Re: Through the classification in hazard levels, we can explore the effectiveness of LID practices in different hazard levels, especial in the High level. Through the analysis in section 3.2 and 3.3, we can find that in the High levels, the inundation depth has been decreased (***depth reduction rates*** are from 22% to 40

%) and most inundation areas are downgraded from High levels to Medium or Low levels (***area reduction rates*** are from 71% to 90 %), but most inundation areas heavn't been eliminated and the ***depth reduction rate*** is lower than other levels (lower 38-40% than Low level). This indicates that LID practices can only ease the inundation depth and downgrade the inundation hazard level and can't thoroughly resolve the inundation problem in High level. And some other methods of stormwater management should be used together to deal with severe waterlogging at High level areas.

9. Page 7, line 4: Please name scenarios 1 to 4

Re: Amended as requested.

10. Figure 6: What is meant by percentage GR and PP? Both with the same percentage?

Re: The proportion means the percentage of the total available implementation areas of LID. Here the ***percentage GR and PP*** means the proportion of Scenario 1 to Scenario 4 (from 25% to 100%) in Figure 6.

11. Page 7, lines 14-18: I do not see where this conclusion comes from. Is this concluded from the numbers in Table 4? What is here meant by performance? Reduction of maximum inundation? This paragraph needs clarification.

Re: We did not put the data in the part that the ***depth reduction rates*** of 100% PP are 67%, 38% and 23% at Low, Medium and High levels, and the ***depth reduction rates*** of 100% GR are 61%, 31% and 21% at three hazard levels. Here the performance means the average depth reduction rate. We will reorganize this paragraph.

12. Page 7, lines 28-31: Again it is not clear where these numbers come from. I do not find it in the Figures. In Figure 6, the single 100 percent cases are not shown.

Re: We did not put the data in the part that the ***area reduction rates*** of 100% PP are 37%, 65% and 67% at Low, Medium and High levels, and the ***area reduction rates*** of 100% GR are 32%, 56% and 67% at three hazard levels. We will add the data in Figure 6.

13. Page 8, line 11: This needs explanation. Why is it difficult to mitigate? Is the reason the topography? I think that such a statement needs to be more specific.

Re: The topographical attributes, such as concaves and potholes, are easy to lead to some places got inundation on the road surfaces. If these places are not or not enough drainage pipes to drainage the rainwater, it is difficult for them to mitigate the influences of urban inundation even there are LIDs. Because of these long-time inundation time areas, the average inundation time increases 0.1 h after the implementation of LID practices (***question 15***).

14. Page 8, lines 12-13: How can one see in these figures that the infrastructure is not perfect? And what infrastructure is here meant and how does it influence the inundation? Also: How can one see from these figures that the LID practices are not perfect? In which sense are they not perfect?

Re: Here we want to explain why some places are difficult to mitigate (***question 13***). These sentences are not rigorous and we will modify them in the revised manuscript.

15. Page 8, lines 13-15: I could not follow the reasoning. Why does the mitigation of short-time inundation areas lead to an increase in the average inundation time with LID measures? Is here something meant along the lines: If a storage due to green roofs helps to keep water back, leading to less inundation depth, the storage will at the same time lead to a longer inundation time (it holds the water back, but releases it eventually)? I am just guessing and I think this needs a better explanation.

Re: Indeed, this is because the statistical number of urban inundation areas are not the same before and after the implementation of LID practices. Here we want to explain why the average inundation time increases 0.1 h after the implementation of LID practices. Because of the implementation of LID practices, the inundation time has been decreased in all hazard levels. However, for the Low level some short-time inundation areas previously affected by surface runoff are freed from urban flooding after the construction of the LID projects, which makes the total number of inundation areas decreases after the implementation of LID. More important, the most freed areas are short-time inundation areas. Although LID practices make existing urban inundation areas' inundation time shorten, the statistical data suggest that the average of the lasting inundation areas' inundation duration is a little longer than that before LID practices. It is also suggests the great effectiveness of LID practices at Low level. We will modify the sentences in line 11-15 and make them clearer to understand.

16. Section 4.1, Comparison of permeable pavement and green roofs: What is the reasoning of the different effects? This should be explained based on the mitigation mechanisms. The last sentence sounds a bit strong. I do not think that one test case can use as a proof, if no general reasoning is given for the different performances.

Re: The available implementation area of PP and GR is 5.95 km^2 and 8.92 km^2, respectively. Although the implementation area of PP is smaller than GR, the effectiveness of PP on urban immunation mitigation is greater than GR in this study (***question 11, 12***). Except the differences of LID parameters, the reason of the different effects might be that PP is built both on low ad high construction lands, while GR is only built on low density construction lands. Indeed, the effectiveness of PP for urban inundation mitigation were different from studies (Qin et al., 2013, Ahiablame and Shakya, 2016, Zhang et al., 2016, Hu et al., 2017), and PP can not always perform better because that the effectiveness is depended on the parameters, implementation area, spatial pattern, rainfall intensity, rainfall frequency and other factors in different regions. Here we want to give a reference for local government that PP might be a good choice for local areas because of the great effectiveness and the large potential for reconstruction in the built-up region (PP could be gradually applied in roads and parking lots, while GR is hard to implement in density construction lands, especially in the urban villages).

17. Page 8, line 29: I would be a bit more careful with the word ’comprehensively’. The paper shows one case study. I do not think that this is a comprehensive exploration of inundation mitigation in an urban watershed.

Re: We will delete the word.

18. Page 9, lines 10-14: As before, I do not see the point about infrastructure. How is poor infrastructure reflected in the model? If not at all: How can one draw any conclusions about this point from a modeling study that does not capture this effect? If yes: What exactly is meant by poor infrastructure and how is this realized in the model?

Re: The scentences in lines 10-14 are not rigorous. Indeed, we find that the efficiency decreases as the proportion of LID implementation increases from Scenario 1 to Scenario 4 and the efficiency of 25 % PP + 25 % GR is higher than other scenarios in this study. This indicates that the greater proportion of LID implementation might not lead to the higher efficiency, and we should not only consider the effectiveness but also the cost of LID practices in the construction of “Sponge City”.

19. Page 22-23: Maybe this sentence is only not formulated well. But I do not see how from this study one could see anything about landscape patterns (’we find that the...’ sounds as if it is a conclusion from this study). The landscape patterns are not discussed, so one cannot conclude about this point. For this reason, I can also not see how ’this provides a new perspective’. Or is here simply meant that this point should be studied in the future? In this case the sentences need to be reformulated.

Re: Thanks for pointing out the expression problem that these results are from Kim and Park (2016) and Giacomoni and Joseph (2017), and we will modify it in the revised manuscript.

20. Conclusions: I think it should be mentioned that the findings in this study apply to the one test case considered. It is not clear if the results are more general and could be transferred to other sites. In particular: Numbers can certainly not be transferred.

Re: This study is a simulation-based research on a local basis. Although the results cannot be transferable to other places directly, the analytical methods, including the coupling model, cost-effectiveness analysis during the sponge city construction can be transferable. We will list the main conclusions below:

1. The coupling model with SWMM and IFMS Urban can be applied to evaluate the effectiveness of LID for urban inundation risk mitigation and can be transferred to other sites.

2. The effectiveness of PP for urban inundation mitigation performs better than that of GR in this research. This conclusion might be different in other regions but it gives a reference for policy-maker on a local basis.

3. LID practices can only ease the inundation depth and downgrade the inundation hazard level but can't thoroughly resolve the inundation problem in High level. Therefore, some other methods of stormwater management should be used together to deal with severe waterlogging at High level areas.

4. The greater proportion of LID implementation might not lead to the higher efficiency, and we should not only consider the effectiveness but also the cost of LID practices in the construction of "Sponge City".

References

Ahiablame, L., and Shakya, R.: Modeling flood reduction effects of low impact development at a watershed scale, J Environ Manage, 171, 81-91, 10.1016/j.jenvman.2016.01.036, 2016.

Burns, M. J., Schubert, J. E., Fletcher, T. D., and Sanders, B. F.: Testing the impact of at-source stormwater management on urban flooding through a coupling of network and overland flow models, Wiley Interdiscip. Rev.-Water, 2, 291-300, 10.1002/wat2.1078, 2015.

Chui, T. F. M., Liu, X., and Zhan, W.: Assessing cost-effectiveness of specific LID practice designs in response to large storm events, Journal of Hydrology, 533, 353-364, 10.1016/j.jhydrol.2015.12.011, 2016.

Fewtrell, T. J., Neal, J. C., Bates, P. D., and Harrison, P. J.: Geometric and structural river channel complexity and the prediction of urban inundation, Hydrological Processes, 25, 3173-3186, 10.1002/hyp.8035, 2011.

Giacomoni, M. H., and Joseph, J.: Multi-Objective Evolutionary Optimization and Monte Carlo Simulation for Placement of Low Impact Development in the Catchment Scale, J. Water Resour. Plan. Manage.-ASCE, 143, 15, 10.1061/(asce)wr.1943-5452.0000812, 2017.

Hu, M., Sayama, T., Zhang, X., Tanaka, K., Takara, K., and Yang, H.: Evaluation of low impact development approach for mitigating flood inundation at a watershed scale in China, J Environ Manage, 193, 430-438, 10.1016/j.jenvman.2017.02.020, 2017.

Kim, H. W., and Park, Y.: Urban green infrastructure and local flooding: The impact of landscape patterns on peak runoff in four Texas MSAs, Applied Geography, 77, 72-81, 10.1016/j.apgeog.2016.10.008, 2016.

Kong, F. H., Ban, Y. L., Yin, H. W., James, P., and Dronova, I.: Modeling stormwater management at the city district level in response to changes in land use and low impact development, Environ. Modell. Softw., 95, 132-142, 10.1016/j.envsoft.2017.06.021, 2017.

Lowe, R., Urich, C., Domingo, N. S., Mark, O., Deletic, A., and Arnbjerg-Nielsen, K.: Assessment of urban pluvial flood risk and efficiency of adaptation options through simulations - A new generation of urban planning tools, Journal of Hydrology, 550, 355-367, 10.1016/j.jhydrol.2017.05.009, 2017.

Qin, H. P., Li, Z. X., and Fu, G.: The effects of low impact development on urban flooding under different rainfall characteristics, J Environ Manage, 129, 577-585, 10.1016/j.jenvman.2013.08.026, 2013.

Russo, B., Sunyer, D., Velasco, M., and Djordjevic, S.: Analysis of extreme flooding events through a calibrated 1D/2D coupled model: the case of Barcelona (Spain), J. Hydroinform., 17, 473-491, 10.2166/hydro.2014.063, 2015.

Wu, X., Wang, Z., Guo, S., Liao, W., Zeng, Z., and Chen, X.: Scenario-based projections of future urban inundation within a coupled hydrodynamic model framework: A case study in Dongguan City, China, Journal of Hydrology, 547, 428-442, 10.1016/j.jhydrol.2017.02.020, 2017.

Zhang, X., Guo, X., and Hu, M.: Hydrological effect of typical low impact development approaches in a residential district, Natural Hazards, 80, 389-400, 10.1007/s11069-015-1974-5, 2016.