The hydrological performance of LID has been widely evaluated. The study focused on the basin runoff. It is interesting. And the article is well written. However, the utility and authenticity of conclusions are limited in view of this being largely a model-based simulation study wherein no observed data are used for model calibration with LID practices. The parameter values have a high impact on LID performance. Meanwhile, the model calibration (no LID practices) period is only 1 year without a warm period, which cannot represent the model performance in different climate conditions (e.g. dry and wet). A thorough analysis of the model results based on observed data is needed. In addition, the urban drainage system was not considered which has a great impact on the urban rainfall-runoff process.

1. The authors emphasized this study used a relatively simple semi-distributed model to reduce the model uncertainty caused by over-complex models. However, there is no comparison between the SUPERFLEX model and other models on model performance. Also, the NSE values of calibration and validation were 0.68 and 0.69, which are not high compared with some distributed physical models at some catchments. I suggest adding something to introduce what kind of phenomena this model can simulate and what are the limitations of its applicability. I also expect a short discussion on why this model is better than other models designed for similar purposes, e.g. SWMM.

2. The high peak flow is usually mainly contributed by surface runoff. In urban areas, due to the drainage system, the peak flow usually occurs quickly. In Figure 7, on 9th, November, the peak flow occurred quickly due to the heavy rainfall on the same day. But why the peak flow was higher on 11th, November while the rainfall was much smaller than the one on 9th, November? Is it due to parameter settings? Please explain it and show the observed values.

3. In Fig. 7, why the peak flow at the rural catchment was much lower than the one at the
urban catchment on 9-11, November, while it was opposite on 17th, November? Could you show the observed values here? Is it caused by the model parameter setting? e.g. original soil water.

4. The authors discuss the effects of urbanization, single LID, and mix LID using different rainfall events. Please show the simulation results with the same rainfall event.

5. Since the total period for calibration and validation is not long. Please compare the simulations and observations at the outlets of rural and urban catchments. Check the model performance at these two sites.

6. The study area is a large basin with 4544 km². There is only 1 precipitation station in the rural catchment. Please discuss its impact on this study and consider how to improve it.

7. Scenario A is 100% in current, Scenario B is 70% in current and Scenario C is 50% in current. The per capita living space for A, B, C is 0.9, 0.85, and 0. Is it an appropriate assumption?

8. Nowadays there is a tendency in writing Discussions making them longer with a greater number of comparisons with results reported in other publications. I’m not sure if I’m happy with such discussion, but please, consider extending this part.