Responses to comments by referee #2

We would like to thank the referee for the careful and thorough reading of this manuscript and the thoughtful comments and constructive suggestions, which help improve this manuscript's quality. Our response follows:

Reviewer #1:

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.*

Response: We appreciate this feedback. In the discussion, the limitation of the model is mentioned as the heterogeneity within the rural and urban areas is not represented. Besides, an additional Section 2.3 SUPERFLEX model framework was created for a more detailed introduction about SUPERFLEX and its application. As mentioned in Section 2.3, it is admitted that some complex distributed hydrological model may bring better NSE values involving more parameters and more complex model structure, but it will suffer from the drawbacks including high data-requirement for calibration, equifinality, and model uncertainty (Beven, 1993; Savenije, 2001; Hrachowitz et al., 2014). This research selected a relatively simple semi-distributed model to describe the different hydrological conditions in urban and rural areas and at the same time avoid the problem caused by complex distributed models.

Secondly, this study involves LID practices that seem suitable for the urban stormwater model, similar to SWMM with available LID modules. But different hydrological responses of urban and rural areas are also concerned. Especially, the different time lags of urban and rural peaks are closely related to the research topic. It makes the description of hydrological attenuation processes important. Therefore, an adapted version of SUPERFLEX with LID modules is used to reflect the hydrological conditions of the whole catchment, especially the difference between urban and rural areas.

Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in water resources*, *16*(1), pp.41-51.

Savenije, H.H., 2001. Equifinality, a blessing in disguise?. *Hydrological processes*, *15*(14), pp.2835-2838.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H.H.G. and Gascuel-Odoux, C., 2014. Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water resources research*, *50*(9), pp.7445-7469.

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.*

   Response: We thank you for this question. The different performances of peak flow on 9^{th} and 11^{th} September are caused by the antecedent rains. Although the intensity of the first precipitation event is stronger than the second one, the initial water storage in reservoirs is small on 9^{th}, due to the lack of antecedent rains. For the second precipitation event in the rainy season, the antecedent rain, that is the first precipitation event on 9^{th}, fills up the reservoirs and makes the reservoirs more sensitive to the precipitation, which leads to a larger peak in the end. This mechanism is similar to the rural runoff generation for the last two of three successive basin peaks in Figure $6a_2$ (Figure 6 was re-drawn in the latest version), instead of the first peak, as mentioned in the manuscript (line 430).

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 hereï¼Ÿ Is it caused by the model parameter setting? e.g. original soil water.*

   Response: The answer to this question is similar to the answer to question 2. It is influenced by the antecedent rains. Because the 9^{th}, September is close to the end of the dry season, rural reservoirs are quite empty. However, the dense precipitation events in the rainy season fill the rural reservoirs gradually, making the rural areas more sensitive to precipitation events. Therefore, the rural areas start to generate large peaks from the 17^{th} onward. Besides, Figure 5 of model performance was re-drawn in the latest version. It compares some peak events between the observation and simulation, including the peak events in September 2018.

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.*

   Response: Thank you for this suggestion. As mentioned in the last answer, Figure 5 of model performance was re-drawn, as it adds the comparison of some peak events between the observation and simulation.

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.*

   Response: We thank you for this suggestion. However, because the urban and rural sub-catchments are not exactly the urban and rural areas of the whole catchment, the observed runoff data at the outlets of urban and rural sub-catchments could not be used directly to calibrate the semi-distributed model directly. But in order to take advantage of this information, the runoff data of urban and rural sub-catchments were used in two parts of this research.
   Firstly, as mentioned in Section 3.2.1 Semi-distributed model setup, the hydrological model starts from two simple lumped pre-models, one for a rural and one for an urban sub-catchment, respectively. In this process, the data collected from two sub-catchments (rural and urban) are used to calibrate the two lumped pre-models, respectively. Secondly, during the selection of parameter set for the semi-distributed model, the runoff characteristics of

simulated urban and rural sub-flows are also compared with the observed runoff time series of two sub-catchments.

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.*

Response: We appreciate this insightful question. A map of the Thiessen polygon for the precipitation data processing is provided below. As it shows, four precipitation stations are located in the rural areas, and two are in the sub-urban areas. But more dense precipitation stations are located in the eastern than the western part of the basin. This unbalanced station distribution may cause the inaccuracy of precipitation estimation. An improvement of this precipitation dataset could be using the gridded precipitation products retrieved from radar and satellite or derived from numerical models. These sorts of the gridded dataset are suitable for those distributed models but also suffer from the uncertainty of retrieval algorithms, instrument design, and physical understanding and parameterization (Yong et al., 2015; Tan et al., 2016; Duan et al., 2016; Tang et al., 2020; Sui et al., 2020).
In this research, a robust semi-distributed model is used, which assumes a uniform precipitation pattern among the study catchment. It will balance the possible underestimation and overestimation in the whole catchment, and also more suitable for our semi-distributed model.
In general, considering the research objective and model used in this study, using the Thiessen polygons method to calculate the areal precipitation based on the precipitation gauges seems appropriate.

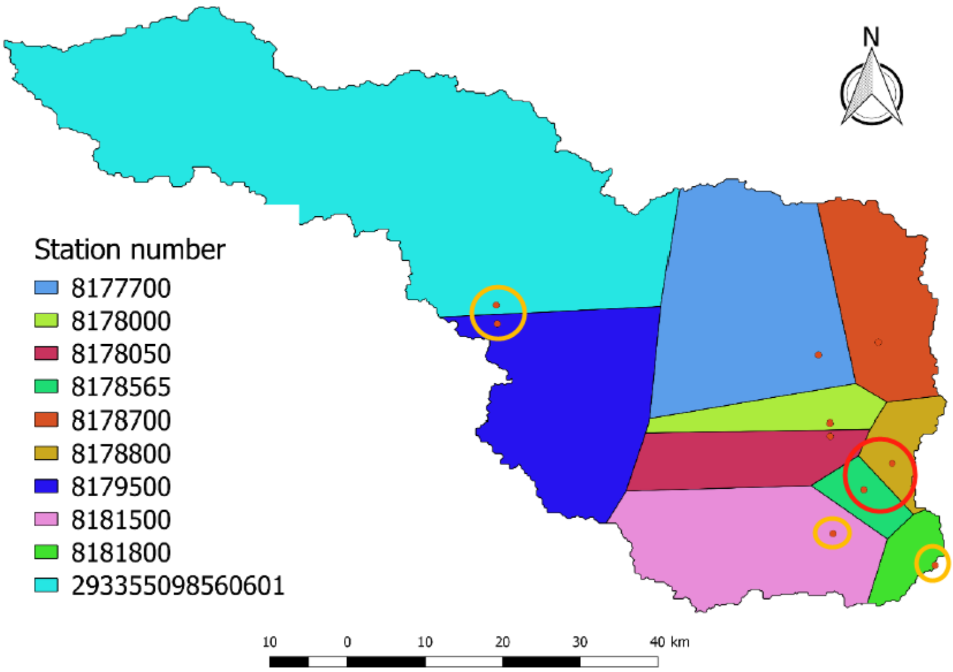

Figure 47. Thiessen polygons of available precipitation sites

Yong, B., Liu, D., Gourley, J.J., Tian, Y., Huffman, G.J., Ren, L. and Hong, Y., 2015. Global view of real-time TRMM multisatellite precipitation analysis: Implications for its successor global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *96*(2), pp.283-296.

Tan, J., Petersen, W. A., Tokay, A., 2016. A novel approach to identify sources of errors in IMERG for GPM ground validation. *Journal of Hydrometeorology*, *17*(9), 2477-2491.

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., Disse, M., 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Science of the Total Environment*, *573*, 1536-1553.

Tang, G., Clark, M. P., Papalexiou, S. M., Newman, A. J., Wood, A. W., Brunet, D., Whitfield, P. H., 2020. EMDNA: Ensemble Meteorological Dataset for North America. *Earth System Science Data Discussions*, 1-41.

Sui, X., Li, Z., Ma, Z., Xu, J., Zhu, S., Liu, H., 2020. Ground Validation and Error Sources Identification for GPM IMERG Product over the Southeast Coastal Regions of China. *Remote Sensing*, *12*(24), 4154.

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?*

Response: The first assumption is about the extent of urban infill development. Based on the urban development information provided by “City of San Antonio: Comprehensive Plan” (2010), it is known that the local government plans to terminate the unconstrained sprawl of the city but adopt the infill strategy. Due to the uncertainty of future urban development, three different extents of urban infill development are assumed, and a relatively broad range of values is given for the three scenarios from 100% to 50%.

For the second assumption about the compression of living space in the future, it is more difficult to predict, controlled by the economic policy and urban condition together. The accurate forecast of this number is difficult, and therefore the parameter uncertainty is admitted in section 5.2 Limitations. In addition, this study focuses on the hydrological influences of LID implementation rather than providing a future prediction for San Antonio City. Some degree of uncertainty on the social assumption could be acceptable in this hydrological discussion.

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.*

Response: Response: Thank you for this valuable advice. Some comparisons about the performance of four LID practices between literature and this research results were added in Section 4.3.1-4.3.4. Besides, another Section 5.3 Comparative analysis was created in the Section 5 Discussion to compare some arguments in this research refer to former studies.

Yours sincerely,

Xinxin Sui and Frans van de Ven