Review of the paper "Multiscale Modeling for Coastal Cities: Addressing Climate Change Impacts on Flood Events at Urban-Scale" submitted by Michele Bendoni

This study introduces a comprehensive modeling framework to assess the impacts of climate change on coastal and riverine flooding in European coastal cities. By integrating atmospheric data from the ALADIN63 regional climate model under the EURO-CORDEX project, researchers simulate future flood scenarios under RCP4.5 and RCP8.5 climate pathways. The modeling chain employs a wave model, storm surge model, river discharge model based on rainfall-runoff processes, and HEC-RAS for hydrodynamic simulations to assess flood extents and depths. The findings underscore the importance of high-resolution, integrated modeling approaches to accurately project and manage future flood risks in coastal urban areas.

I would like to compliment the authors for a well-written manuscript. The paper introduces a framework for assessing coastal and inland flooding at the urban scale in the European region under climate change. I recommend publication of the manuscript after properly addressing the following comments.

Main comments:

1. The manuscript does not provide any information on the validation of the numerical models used in the framework, specifically the wave model, storm surge model, river discharge model, or the hydrodynamic model. While the paper presents a generalized workflow, the quantitative conclusions (e.g., flood volumes, durations, extents) rely heavily on the reliability of the model outputs. Without presenting model validation or calibration results, it is difficult to assess the robustness of the findings. This is particularly critical for extreme hydrologic/hydrodynamic modeling, where prior studies have shown that numerical models often fail to estimate peak magnitudes accurately unless properly calibrated. I strongly recommend including a validation section for each model, either in the main text or supplementary materials, so that readers can better understand the limitations and reliability of the framework and derived results.

2. To provide boundary conditions for the hydrodynamic model, the time series of Q and total water levels are obtained by averaging superimposed 49-hour time series around the annual maxima peaks. The time series surrounding the annual maxima may exhibit multiple peaks and varied shapes across a wide range, particularly for Q, depending on the upstream catchment and rainfall characteristics. Therefore, averaging at each timestep and treating it as the representative hydrograph may overlook temporal dependencies and could be unrealistic. A similar issue may arise for storm surge and/or total water levels (see Quinn et al., 2014; Santamaria-Aguilar et al., 2017). To support the approach used in this study, a figure showing all individual and averaged hydrographs (possibly in the supplementary material) or appropriate citations of relevant studies that have previously applied this method would strengthen the justification.

3. The purpose of comparing the historical and evaluation simulations is unclear. These runs differ in only 7 years of annual maxima, with the remaining 26 years overlapping. It's mentioned that this analysis is done to "test the ability of the model to reproduce observable extreme events." I would like to see some discussion on how this is achieved in the author's reply in the open discussion. Additionally, while the authors suggest that this comparison of quantile plots helps to estimate

the degree of over- or under-estimation in projections (e.g., L 624-L625), the interpretation of these results and their contribution to the main analysis are not well explained.

4.

Other Comments:

1. Figures 6 & 7: The term “envelope of the water depth” is unclear. Do the flood maps represent the maximum depth reached at each grid cell throughout the simulation period, or are they snapshots at a specific moment in time? Please clarify this in the text or figure captions.

2. The authors claim that the uncertainty, which was not accounted for due to the absence of a multi-model ensemble, was recovered in the statistical analysis by calculating confidence intervals through the bootstrap method. However, these represent fundamentally different types of uncertainty and should not be considered as “recovered” from one another.

3. Line 87: Typo: the word “and” is misspelled.

4. Line 343: The sentence implies that the shape, timing, and magnitude of streamflow are determined solely by upstream hydrological processes. However, downstream conditions such as tides and storm surge can also play a significant role. Please revise the sentence to reflect this more accurately.

5. Lines 614–616: The phrase “associated with the highest waves” is unclear whether it refers to wave heights or wavelengths. Please specify.

6. Lines 654–657: It is unclear why the analysis in this section was conducted only for one location instead of all three sites. The phrase “number of events per year higher than specific values” would benefit from clarification; please indicate what those specific threshold values are.

Quinn, N., Lewis, M., Wadey, M. P., and Haigh, I. D.: Assessing the temporal variability in extreme storm-tide time series for coastal flood risk assessment, J Geophys Res Oceans, 119, 4983–4998, https://doi.org/10.1002/2014JC010197, 2014.

Santamaria-Aguilar, S., Schuerch, M., Vafeidis, A. T., and Carretero, S. C.: Long-term trends and variability of water levels and tides in Buenos Aires and Mar del Plata, Argentina, Front. Mar. Sci., 4, 1–15, https://doi.org/10.3389/fmars.2017.00380, 2017.