General comments

The paper evaluates the performance of coupled atmospheric-hydrological modeling systems in order to improve the simulation of flood events. Precisely, it compares 1) the performance of coupling two hydrological systems (fully distributed WRF-Hydro and a simi-distributed Hydrological Engineering Center-Hydrological Modeling System (HECHMS)) with WRF model for four rainfall events in two mountainous areas in China, and 2) the lumped HEC-HMS using the observed gauge precipitation with the simi-distributed HEC-HMS using a 1 x 1 km rainfall output from the WRF model. They concluded that the accuracy of the simulation of the rainfall and the model's complexity influenced the performance of the coupling atmospheric and hydrological models. The manuscript offers an additional contribution to understanding the advantages and disadvantages in the selection of the hydrological model. However, the paper must be more organized in some sections. Thus, I recommend a minor revision of the manuscript before it can be accepted for publication in NHESS.

Specific comments

1. The main objectives and the motivation of this paper must be more clearly explained in the introduction. A reconstruction of the introduction is needed.
2. The authors must add a paragraph at the end of the introduction that explains the structure of this paper.
3. The authors must write information in Section 2 about the origin of the gauged data, the total number of stations at the two catchments, the hydrological data, and the initial boundary conditions for WRF.
4. A small paragraph about the atmospheric circulation and the structure / scale of the selected events will be helpful.
5. The authors must add a table with the WRF physics schemes used and the Land surface model. Are there previous sensitivity studies that used this options in similar simulated events? Please justify the selection.
6. More details about FNL data are needed.
7. The authors must elaborate on the calibration process of WRF-Hydro. Which method did you use, and why? What data did they use for the calibration?
8. A discussion that would put their results in a comparative context with other studies is missing.
9. Discuss about the potential of exploiting the presented results under operational forecasting applications, as those presented, for example, by Giannaros et al. (2021) and Varlas et al. (2024).

   https://www.mdpi.com/2571-9394/3/2/26

   https://www.mdpi.com/2073-4433/15/1/120

10. A paragraph about the limitations of this study and if there are uncertainties that impact its results must be added.

Technical corrections

1. Line 64. Which NWP model?
2. Line 64-65. Please rephrase

3. Line 67. Please elaborate on the results.
4. Line 70. Which are the studies that consider the difference?
5. Line 75. Use capital letters on “valiya veettil et al”.
6. Line 85. Please be more clear about the role of examining the lumped HEC85 HMS model with observations.Line 89. Please give reference to Figure 1.
7. Line 95. Please elaborate on the results of coupling WRF and WRF-Hydro.
8. Line 122. Please correct “2219km2” and “1760km2”.
9. Line 128. Which is the strom season?
10. Line 120. It must be “2.1”?
11. Line 141. Is there a reference for the coefficient of variance?
12. Line 164. Reference?
13. Line 167. A title on the x-axis is missing in Figure 2.
14. Line 173. Which version of WRF is used?
15. Line 187: A figure with the WPS domain configuration will be helpful.
16. Line 197. The information about WRF-Hydro must be reduced.
17. Line 242. The authors must explain how they compute the catchments' routing grids.
18. Line 254. There is no need for figure 3.
19. Line 255. The information about HEC-HMS model must be reduced.
20. Line 303. References?
21. Line 325. Reference?
22. Line 481. This section looks like a summary. The authors must highlight the strong points of the study.