Interactive comment on "A coupled atmospheric-hydrologic modeling system with variable grid sizes for rainfall-runoff simulation in semi-humid and semi-arid watersheds: How does the coupling scale affects the results?" *by* Jiyang Tian et al.

Anonymous Referee #1

Received and published: 9 March 2020

The paper presents a study to explore the impact of coupling an atmospheric model with a hydrological model (lumped and distributed) using different model grid sizes for the simulation of river flows. The results showed that for precipitation events with uneven spatiotemporal distribution, a higher resolution model can lead to better flow simulations, whereas for events with uniform spatial distribution, the coupling scale has less impact on flow simulation and in this case a lumped model performs as good as a distributed model. The paper is interesting and well written with some minor spelling

mistakes, but there are a number of questions that the authors need to address before the paper is recommended for publication.

1- The paper is essentially divided in two parts: a) A coupled atmospheric-hydrologic system with variable grid sizes for rainfall-runoff simulation and b) the development of a new distributed hydrological model. The distributed model helps to answer the research question posed in the first part (How coupling affect the results?). However, the authors are actually presenting a new distributed hydrological model (why?) rather than using a distributed model widely accepted by the hydrological community. why do we need a new model? In order to answer your research question, you need to demonstrate that your hydrological model works well for a number of storms and that the model has been properly calibrated. I think this area is a bit weak in the paper and needs further results. For instance, the authors calibrate the model with 7 storms, but they do not give any indication on how the calibration was achieved, what type of storms are used and so on. Usually to calibrate hydrological models, continuous rainfall-runoff time series are needed rather than individual events in order to account for initial conditions in the model such as soil moisture, catchment wetness, etc. In addition, the authors quoted a model calibration efficiency (NSE) of 0.686, but three out of four events used in the validation showed an efficiency higher than 0.75. Normally the performance of the model in the validation phase is worse than in the calibration, but this is not the case in your analysis. why? For the storms used for validation, how do you account for the initial catchment conditions? It is obvious that if the model starts completely dry, the results will be affected by this even if the model is calibrated properly. please expand on this. In addition, it is unclear from all the equations used to describe the model, which are model parameters. maybe you can summarise all the model parameters in a table and include their range of values. what about model parameter calibration uncertainty? What ranges of model parameters did you use and why? It is well known that different parameter sets can produce a similar model performance (equifinality). You need to look at parameter uncertainty and maybe produce an ensemble of hydrographs rather than a deterministic one. what about the uncertainty in the observations (e.g. rain

gauges, flow stations, etc)? You did not mention any of this in the paper. 2- The second point is about the use of different WRF microphysics parameterisations (MP). Table 4 shows that different MPs are used to simulate each storm event. How do you isolate the impact of the WRF MPs in your results? How do you justify that the difference in the results is due to the different grid resolutions rather than the fact that different MPs are used to simulate each event? What is the performance of WRF simulating those storms?

Other comments

The abstract talks about the "Hebei model", but this model is not known in the literature and you have not introduced this model yet. In page 3: "a higher error rate". Do you mean "larger errors"? the use of "error rate" might be confusing. "variation pattern" - again unclear what you mean here. "Hebei model" Is this model published? If so, you need to include a proper reference. If not, then you should describe the model in the methodology and do not use "Hebei model" until this has been described. "1x1km ... 9x9km" Are these spatial scales within the WRF model domains or are you talking about the spatial scales of the hydrological model? Section 2. The description of the catchments should be placed before the description of the events otherwise how do you know which catchment outlet are you talking about here? Is Cv a standard metric to characterise the spatial and temporal distribution of precipitation processes? If so, you need to include a reference. I believe to characterise the spatial distribution of precipitation is better to use semi variograms, correlograms or by looking at the spatial correlation of the precipitation field. Likewise, with the temporal correlation. Could you please expand and justify why a simple metric like this was used? Cv here is highly dependent on the number of rain gauge stations available. Is the WRF rainfall field used to compute Cv? table 2 - I think you need to explain what values of Cv correspond to a highly variable event in space and time. how did you come up with the critical values of Cv (0.4 and 1 for spatial/temporal distribution respectively)? you need to justify these values. Section 2.2 can you provide more information of these catchments (e.g. catch-

C3

ment descriptors)? e.g. apart from catchment area, mean annual rainfall, catchment slope, mean flow, predominant land use and soil types, type of geology, percentage of urban area, etc. you can summarise all this info in a table. this is important to understand the catchment response to precipitation. Section 3.1.1. "in consistence" consistent? The grid resolution of the coarser domain is 9km, but the NCEP analysis is about 100km. How does the WRF model handles this discrepancy with the initial and boundary conditions between the outer domain and the analysis? What's the temporal resolution of the analysis? Section 3.1.2. Need to define all acronyms for the physical parameterisations. Section 3.2.2 "... has widely been applied in Hebei Province ..." add references. Summarise all model parameters in a table with the range of values for calibration. How do you account for different soil types, land use, etc? These will have important implications in terms of runoff production. Do you use ET (evapotranspiration) or only E (evaporation)? do you have forest in any of the basins? Expand on the calibration and validation of the hydrological model, including both, lumped and distributed models. fig 11. Unclear if these results are for catchment A or B. Table 5. how do you isolate the impact of the different WRF microphysics parameterisations in your results?