Comment on hess-2021-501

Anonymous Referee #1

___

Referee comment on "The effects of spatial and temporal resolution of gridded meteorological forcing on watershed hydrological responses" by Pin Shuai et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-501-RC1, 2021

___

The paper explores the impacts of publicly available gridded meteorological forcing (GMFs) on the simulated hydrology using an integrated hydrologic model. The choices of the GMFs are excellent as they are widely used in the hydrology community of the continental US. The selected site also allows analyzing different and complex settings (different topography, geology, and land cover) and is characterized by a heterogeneous pattern of forcing. The model is very suitable for this kind of study. Overall, the manuscript is well written and the objectives as well as the methodology are adequate and clearly explained.

However, some issues that need to be addressed:

- Although the authors acknowledge the mismatch between the coarse resolution of the NLDAS and the size of their watershed, they didn't elaborate more on the impacts of this match. The watershed size is half the NLDAS grid and the precipitation is homogeneous, so technically NLDAS isn't adequate for this particular watershed, and analyzing the impacts of temporal resolution using NLDAS isn't accurate. The authors could analyze the impacts of the temporal resolution by using PRISM data instead of NLDAS. PRISM data could be converted from daily to hourly using NLDAS patterns for example.
- The authors compare both the impacts of spatial and temporal resolution, but they did so using different forcing. Hence, it is difficult to highlight which one has the highest impact on the simulated hydrology. By analyzing the impacts of the temporal resolution using PRISM, the authors could compare the impacts of spatial and temporal resolution and highlight the most important resolution for this particular watershed.
- What are the physical explanations for the differences between NLDAS, PRISM, and DAYMET? Are these differences due to the physics, or the types of data they used?
- It is difficult to compare data with different spatial and temporal resolutions in addition to the differences in the methods used to generate these datasets. How can we differentiate errors due to the spatial and temporal resolution of the GMFs from the methods used to generate them?

- Because these datasets provide different variables, it could be great to know which of these forcing variables (e.g., precipitation or temperature) drive the observed impacts. Is a high-resolution precipitation more important that a high-resolution temperature or solar radiation?
- The authors used the DAYMET solar radiation when running the model with the PRISM dataset, as their results have shown by merging these datasets some simulated variables such as ET were found to be equal. What are the impacts of solar radiation? Is a high-resolution solar radiation needed to simulate the system? The energy balance (ET and SWE) plays an important role in this watershed, and because most of the key forcing variables were the same, so were the simulated hydrologic variables.
- The authors should isolate the impacts of the different variables. They could perform a simulation with different precipitation (PRISM, NLDAS, DAYMET) and keep the other variables identical, then the same temperature and precipitation and the other variables remaining identical. This will allow to better understand how the uncertainties of these variables drive the observed differences. Performing the comparisons step by step will allow providing insights into how each components acts and the aggregated effects.
- State from the beginning that there are no ground measurements of groundwater levels and soil moisture. This is important to better understand the goal of the study.
- In the absence of a reasonable number of ground measurements, the authors could compare the random errors embedded in each of the datasets. The triple collocation analysis allows computing the random errors associated with each dataset without knowing the truth (ground measurements). The authors could employ the framework and better analyze the differences between the GMFs first and the simulated hydrology. Otherwise, it is hard to know which dataset is accurate or contains fewer biases.
- Are the findings watershed-dependent? How can we apply these findings to other watersheds?

Other comments

- WRF-Hydro is not an integrated hydrologic model
- Paragraph 100: add the source of the annual precipitation values
- Does the water table depth (in areas located outside the watershed because there are no measurements in the watershed) go beyond 28m depth? what are the boundary conditions at the lower limits and the other limits?
- What do you mean by high snow year and low snow year, please provide numbers
- How did you test the pinup. Did you compare the storages? How did you find that the spin-up period is accurate?
- Even if the model is not calibrated it should be able to reproduce the measured streamflow?
- Figure 1: Add the location of the watershed in the US or at least the western US. What is NHDPlus, add the definition and the source of the dataset
- Figure 2: missing legend of soil and geology map
- Figure 4: put the temporal variations figures on the same graph for visual comparisons
- Figure 6: smoothen ET for a better comparison
- Figure 7: Plot the differences
- Figure 6: ET is misleading. Compare daily and hourly ET similarly with the same units.

- Figure 8: plot the temporal variations figures together
- Figure 10: plot the differences
- Figure 12: Is the groundwater table equal to 3000m?
- Figure 13: plot the temporal variations figures together
- Figure 15: Are the daily NLDAS ET is twice higher than the hourly NLDAS ET? Why the differences are less pronounced for the 8d?
- What explains the low snow associated with hourly NLDAS? Is the temperature higher with daily NLDAS?
- Paragraph 360 is cut in my version "Additionally?"