This study investigated the effects of spatial heterogeneity in atmospheric forcing, land use and land cover, soil properties, and topography on the modeling of evapotranspiration, runoff, and their components (i.e., canopy evaporation, ground evaporation, transpiration, surface and subsurface runoff). The design of the numerical experiments is reasonable, the methods are innovative, and the results are insightful. With the above considerations, I suggest a publication with several clarifications.

Detailed comments
- L121--L124: Better put them in the Methodology section.
- L135, 2013a Li et al., 2013: typos?
- L136--L140: What is the purpose of these statements? They duplicate the discussion. On the other hand, more descriptions of ELM on the used parameterizations are desirable.
- L166--L167: How did you resample soil properties? How soil texture and organic matter are upscaled, and how soil color is downscaled?
- L199--L209: The calculation of Sobol's sensitivity index is still a bit confused to me. Assuming that X is the 30-year monthly value, SENSITIVITY(X) is the Sobol's sensitivity index of X, and 30-YEAR-SEASONAL-AVERAGE(X) is the seasonal average of X, did you calculate the index as SENSITIVITY(30-year-seasonal-average(X)) or 30-year-seasonal-average(SENSITIVITY(X))?
- Table 4: Did you first propose this approach?
- L260: It would be to compare the spatial variability with the temporal variability or the mean value of ET/P.
- L305--L328 and Figure 5: I did not get the objective of these contents. Since ATM and LULC are dominant, it is quite natural that their contributions to the spatial variability are complementary. Deletion seems fine and would make the paper concise.
- L392-L395: Do the interplay between the spatial variabilities in ATM, LULC, SOIL, and TOPO increase or decrease the overall variability?
- Section 3.5 and Figure 8: It is interesting to see the comparison between ERA5-Land and the 0.125-degree ELM simulations. Since the ERA5-Land atmospheric forcing is interpolated from the ~31-km ERA5 data with the consideration of elevation dependency (doi:10.5194/essd-13-4349-2021) rather than upscaled from the finer-scale observations,
the spatial variability from ERA5-Land would be smaller than that from ELM. In Figure 8, it would be better to show the difference rather than absolute difference to check this.