

The authors have done a very thorough job revising the manuscript, and it is much improved. I have only minor comments before I believe the manuscript will be ready for publication.

Major comments:

I find the phrase “subgrid-scale physics schemes” a little general. For a coupled model, this could refer to the representation of hydrology in ParFlow-CLM (or other hydrologic models), but in this study it refers to physics schemes within WRF. In the abstract (or even the title), could you clarify that the physics schemes here refer to the physics within WRF?

I also have one lingering comment about the poor agreement between streamflow simulations and observations, as well as the related temperature bias. The poor streamflow fit suggests that even if regional climate models like WRF outperform observations with respect to precipitation, a temperature bias could confound this with respect to their ability to act as accurate hydrologic model forcings. I think it would be valuable to mention as early as the abstract (and possibly expand in discussion) that all the WRF configurations have a temperature bias relative to PRISM that’s comparable to what we would expect due to climate change impacts. Maybe that will just tell readers that bias correction is still important, but it seems like information that should be available early.

Page 1

Line 21 – should be “uncertainty from synoptic-scale forcings...”?

Line 25 – “delayed earlier” – This is a contradiction in my view. Were flows delayed or earlier?

Page 4

Line 101 – “However, neither... do not...” Did you intend the double-negative here?

Page 18

Table 3 – The standout conclusion from this table, in my view, is that all models have pretty terrible R^2 of precipitation relative to PRISM. Is this due to stochasticity in the daily time series generation, or something else? If it’s event-scale stochasticity, would a different metric be more appropriate?

Figure 4 provides a time-averaged spatially explicit comparison between PRISM and modeled temperature; why don’t we have the same thing provided for precipitation?

Page 19

First line of page - Based on Table 3, NCAR-CFSR seems like a more attractive fit to PRISM than the BSU simulations. The temperature R^2 is slightly worse, but the precipitation is substantially better. Are there other factors that made the BSU simulations better?