

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2020-371-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2020-371

Wayne Angevine (Referee)

Referee comment on "Surface representation impacts on turbulent heat fluxes in the Weather Research and Forecasting (WRF) model (v.4.1.3)" by Carlos Román-Cascón et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-371-RC1>, 2021

The manuscript describes tests of several aspects of land surface representation in WRF with respect to heat flux from the surface. It is generally well-written. The problem is very important. The results should be useful to the many readers applying WRF to their problems. The presentation could be improved by removing some sections and clarifying others. I offer a number of general comments, but overall it is a good paper, and addressing the comments should not require doing more simulations.

General comments:

1. I have concerns about the use of different land surface models with the same input data. Each LSM has its own "climate", and the input data (assuming it all comes from FNL, as implied by table 2) comes from a different model with a different climate. This is strongly demonstrated by Angevine et al. (2014) in ACP for the same project. Have the model outputs been checked to be sure that there are not strong spinup effects in the soil? I particularly suspect that this is responsible from some of the behavior of the RUC LSM, which is known to have a different soil moisture baseline from Noah.
2. The evaluation of PBL height (sections 2.1.3 and 4) is too incomplete to be useful. In such a small area, it is not possible to learn anything important by looking at individual columns, there is too much interaction between columns at any reasonable wind speed. The authors say as much at the end of section 4. The paper would be strengthened by removing these sections.
3. Throughout the paper, RMSE is used as the major metric. There are two problems with this. First, RMSE includes both bias and random error. It is much more useful to treat the two separately (bias and standard deviation). Second, it is not clear how the RMSE is calculated. Please state clearly exactly what time series are being compared, including

time and locations (pixels or categories).

4. Noah-MP is an extremely complex LSM with many configuration options, intended for use in ensemble systems. It is not clear that using the default options is correct. It might be necessary to consult the scheme developers, or to drop this option from the comparison.

5. The urban land class has large biases. This is a known problem with the LSMs in WRF when run without an urban parameterization. Basically the LSM treats the urban class as a slab of concrete with no moisture availability. Since urban parameterization is out of the scope of the paper, and the AAF comparison is problematic for this category, I recommend ignoring the urban class except for a brief mention.

6. In the mosaic approach, I would have expected that soil properties depend on the tile class, not just radiative properties. Is this not the case? This arises in line 510 discussing RUC, but we need to know what is varied in Noah also.

7. In the final paragraph of the conclusions, a number of uncertainties in the flux observations are mentioned. It would have been good to address these more formally throughout the paper, but I would not recommend going back to do that now. The paragraph also urges even more comprehensive deployments in the future. Given the size and scope of BLLAST, I think it is unlikely that we will ever see a better-instrumented area. What is needed is better methods of coping with the inevitable limitations of the observations. We will definitely never see comprehensive instrumentation at global scale, so we need to think better about how to do a good job of modeling places we can't measure.