This study assimilates SMAP (9 km) surface soil moisture estimates to generate 1 km soil moisture, fluxes (sensible, latent), and temperature estimates for a region (Central, SE) of CONUS that's of critical interest (high density of in situ networks) for drought monitoring. Other datasets (ERA5 [reanalysis], GLASS [ML], GLEAM [model], SMAP [remotely sensed]) are compared with the available in situ records for soil moisture or other land-atmosphere variables as available. The 1 km SMAPDA product shows relatively good performance, and reasons are suggest further improvement for soil moisture modeling. This study helps advance downscaling/modeling soil moisture with land surface models (specifically Noah-MP). However, as the first reviewer has mentioned, further discussion is required to help frame the significance of this experiment. The limited data generated also needs to be considered (1 year, 2016). How does comparing a 1 km SMAP DA-modeled soil moisture product against other soil moisture products (modeled or otherwise) for this 1 year contribute to either the LSM, soil moisture/climate community, or other stakeholders?

We sincerely thank the reviewer for the thoughtful and constructive comments. The detailed feedback has been instrumental in improving the clarity, rigor, and overall quality of the manuscript. Our responses to each comment are provided below, directly following each bullet point.

Major comments:

1. Why was SMAP 9 km used to downscale to a 1 km resolution? SMAP-based 1 km soil moisture datasets exist (i.e., the NSIDC-SMAP 1 km product by Fang et al., 2022) that is available for this study period and area. Sentinel-1 based soil moisture would also be good to consider/compare as well (Brocca et al., 2024).

Reply: Thanks for the comment. We acknowledge the existing SMAP-based 1 km soil moisture datasets could potentially provide finer-scale soil moisture info and may be incorporated into the assimilation process. However, as mentioned in Fang et al. (2022), the spatial coverage and availability of the 1-km downscaled dataset is notably reduced compared to the 9-km dataset (e.g., Fig. 1 as shown below).

Regarding the Sentinel-1 based high-resolution soil moisture datasets, there are multiple demonstrations such as those documented in Brocca et al. (2024), Filippucci et al. (2021), Foucras et al. (2020), Gao et al. (2017), and even the one jointly use SMAP and Sentinel-1 (Meyer et al. (2022)). However, despite promising results, most of them were generated and tested in a much smaller domain. This is primarily due to the contrasts in sensor characteristics between these two satellites (shown in the table below):

Feature	SMAP	Sentinel-1
Platform	NASA	ESA (Copernicus)
Sensor Type	Radiometer (passive) + Radar (active – failed in 2015)	Synthetic Aperture Radar (SAR, active)
Frequency	L-band (1.41 GHz)	C-band (5.4 GHz)
Resolution	~36 km (radiometer), ~9 km (discontinued radar)	~1 km (can be resampled to ~100 m)
Revisit Time	~2-3 days	6–12 days (depends on latitude and orbit)

While Sentinel-1 has higher spatial resolution (~1 km) than SMAP, it has relatively lower radiometry sensitivity, much longer revisit time (approximately 3 to 4 times longer), and requires more complex preprocessing. This means Sentinel-1 is most likely less sensitive to subtle differences in soil moisture content and have limited ability to capture day-to day variability. In this case, SMAP is more suitable for regional-to-global scale applications, whereas Sentinel-1has more potential in local-scale monitoring.

Given the above, we did not attempt to assimilate any of the downscaled (higher spatial resolution) datasets. To acknowledge that those emerging data/products may be potentially used in the future, we've summarized discussions here and added into the introduction from L98 to L109.

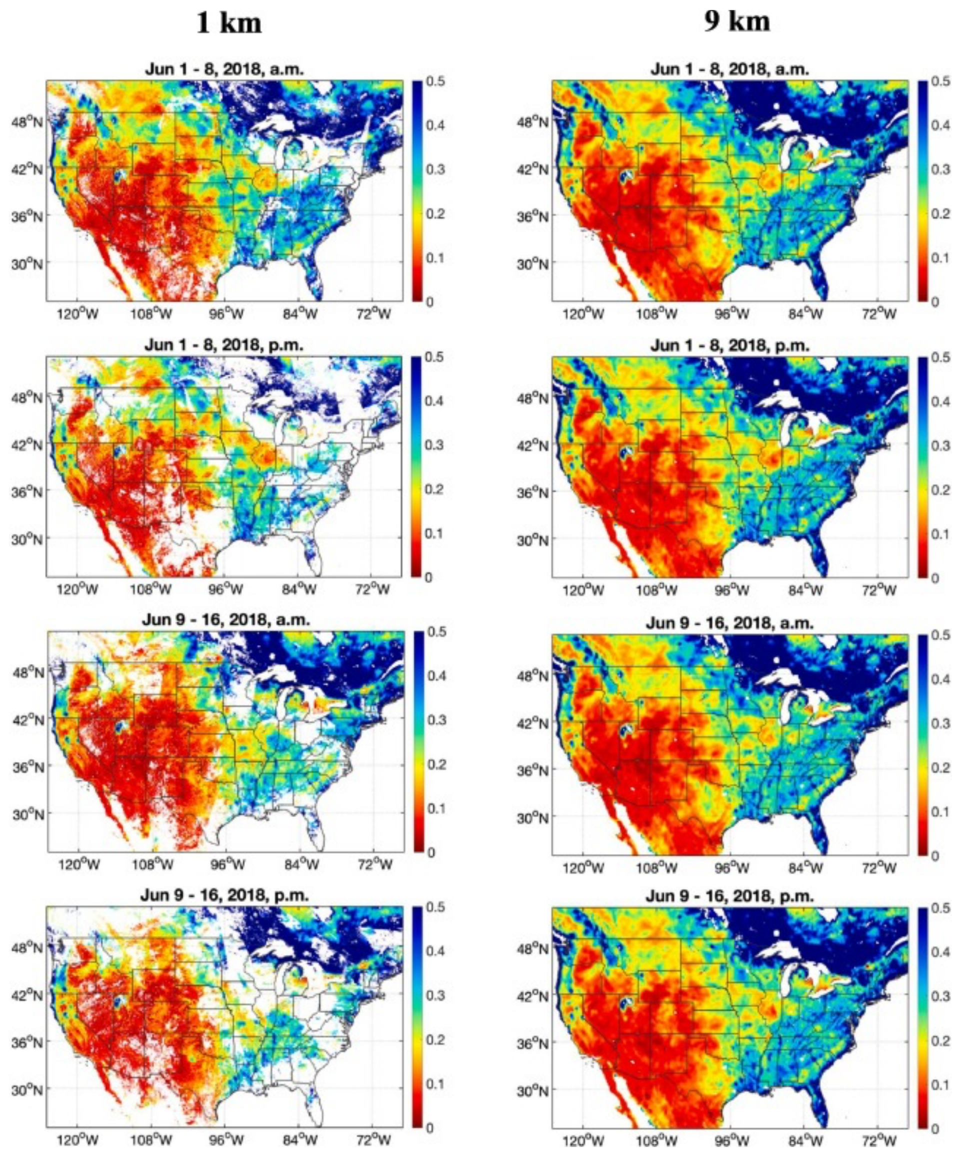

Fig. 1 Comparison of 1-km and 9-km SMAP soil moisture composite maps for the periods of Jun 1 – 8, 2018 and Jun 9 – 16, 2018. (adopted from Fig. 5 in Fang et al. (2022))

Reference:

Fang, B., Lakshmi, V., Cosh, M., Liu, P.-W., Bindlish, R., & Jackson, T. J (2022). A global 1-km downscaled SMAP soil moisture product based on thermal inertia theory. Vadose Zone Journal, 21, e20182. https://doi.org/10.1002/vzj2.20182

Lakshmi, V. and B. Fang. 2023. SMAP-Derived 1-km Downscaled Surface Soil Moisture Product, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/U8QZ2AXE5V7B.

Brocca, L., and Coauthors, 2024: Exploring the actual spatial resolution of 1 km satellite soil moisture products. Science of The Total Environment, 945, 174087, https://doi.org/10.1016/j.scitotenv.2024.174087.

Gao, Q., M. Zribi, M. J. Escorihuela, and N. Baghdadi, 2017: Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution. Sensors, 17, 1966, https://doi.org/10.3390/s17091966.

Meyer, R., W. Zhang, S. J. Kragh, M. Andreasen, K. H. Jensen, R. Fensholt, S. Stisen, and M. C. Looms, 2022: Exploring the combined use of SMAP and Sentinel-1 data for downscaling soil moisture beyond the 1 km scale. Hydrology and Earth System Sciences, 26, 3337–3357, https://doi.org/10.5194/hess-26-3337-2022.

Filippucci, P., L. Brocca, R. Quast, L. Ciabatta, C. Saltalippi, W. Wagner, and A. Tarpanelli, 2022: High-resolution (1-km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River basin as a case study. Hydrology and Earth System Sciences, 26, 2481–2497, https://doi.org/10.5194/hess-26-2481-2022.

Foucras, M., M. Zribi, C. Albergel, N. Baghdadi, J.-C. Calvet, and T. Pellarin, 2020: Estimating 500-m Resolution Soil Moisture Using Sentinel-1 and Optical Data Synergy. Water, 12, 866, https://doi.org/10.3390/w12030866.

2. Why is only SMAP AM used for the assimilation? The afternoon product is mentioned, but I did not see any further justification for excluding it. Also, note that the SMAP L3 product is not exactly daily due to missing coverage. Was any gap-filling used prior to the data assimilation?

Reply: We would like to clarify that the exclusion of SMAP PM overpass was only employed when analyzing the intercomparison among different soil moisture data/products. The assimilation process applies all valid overpasses, both AM and PM.

There was no gap-filling applied prior to the data assimilation. Based on the time label in the SMAP data, the assimilation was performed at those pixels with valid SMAP data coverage in an hourly frequency.

3. The resolution of each respective product appears to be an important point for explaining the pattern in performance (e.g., GLASS & SMAPDA vs. SPL3SMP_E/ERA5-Land/GLEAM). Why didn't the authors attempt a higher resolution grid (e.g., 500 m) for the forcing? Given the Rouf et al. (2021) paper investigated up to that resolution with Noah-MP, such a study may be more impactful than the current one presented. Their study also cite increasing accuracy with increased forcing resolution.

Reply: Thank you for pointing out the potential sensitivities due to the model resolutions and forcing data.

Rouf et al. (2021) used physically based downscaling method to obtain 500-m forcing data from 12.5 km resolution NLDAS-2 atmospheric variables (air temperature, pressure and humidity; longwave and shortwave radiation; wind speed and wind direction) and demonstrated positive impacts on soil moisture results. However, the downscaled forcing remains model-based and did not include fine-scale precipitation (remains in 12.5 km grid spacing). The coarse resolution precipitation could introduce large uncertainty in simulating soil states as precipitation amount at each grid point determines how much water may penetrate to soil layers. Thus, as opposed to the approach taken in Rouf et al. (2021), we replaced the precipitation forcing in the NLDAS-2 by Stage IV (grid spacing of 4 km), a radar-observation-constrained precipitation data valid for the CONUS region.

To justify the use of Stage IV precipitation data, we examined the changes in soil moisture estimates due to the replacement in precipitation forcing data during the two intensive observation periods of HI-SCALE field campaign (Fast et al. 2018). The RMSEs were computed against several in-situ data sets collected by the Oklahoma Mesonet in a domain covering the ARM SGP site and adjacent area. The results in Fig. 2 and Fig. 3 demonstrate that soil moisture simulated by the OL simulation further improves after replacing the NLDAS-2 precipitation data in the forcing data with the Stage IV precipitation data. The comparison of instantaneous rain rate obtained from NLDAS-2 and Stage IV (ST4) precipitation at 00 UTC of August 30, 2016 (Fig. 4) also clearly shows that the Stage IV provides more heterogeneous precipitation distribution over the study domain.

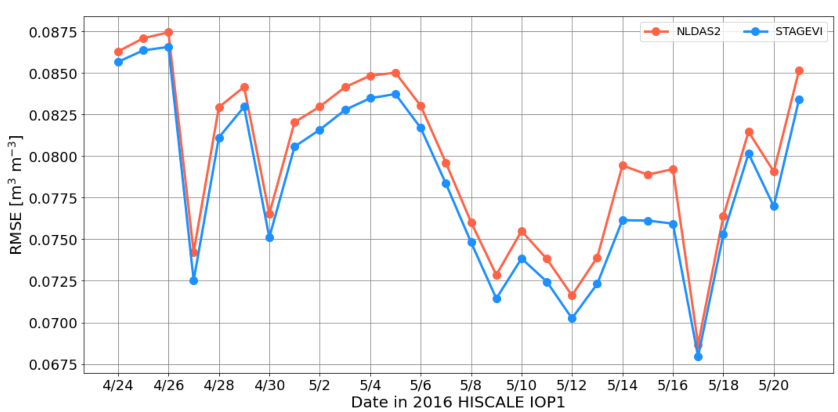

Fig. 2 Root-mean-square error (RMSE) of soil moisture computed simulated by the open loop simulations using NLDAS-2 and Stage IV precipitation during the HI-SCALE IOP1 (April 24 - May 20, 2016).

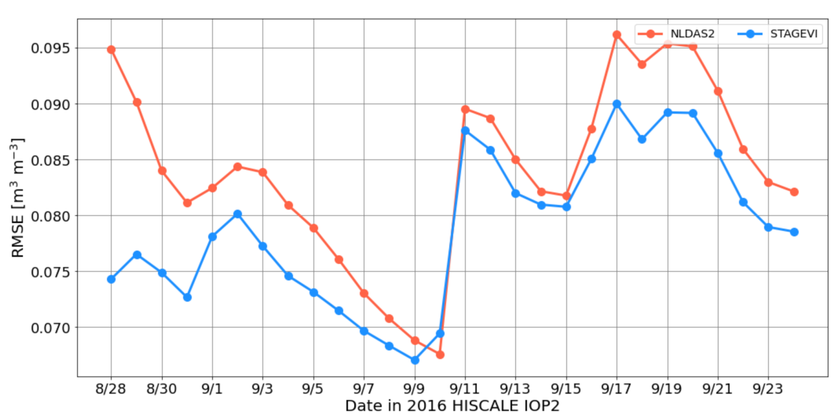

Fig. 3 Similar to Fig. 1 but for HI-SCLAE IO2 (Aug. 28 - Sep. 23, 2016).

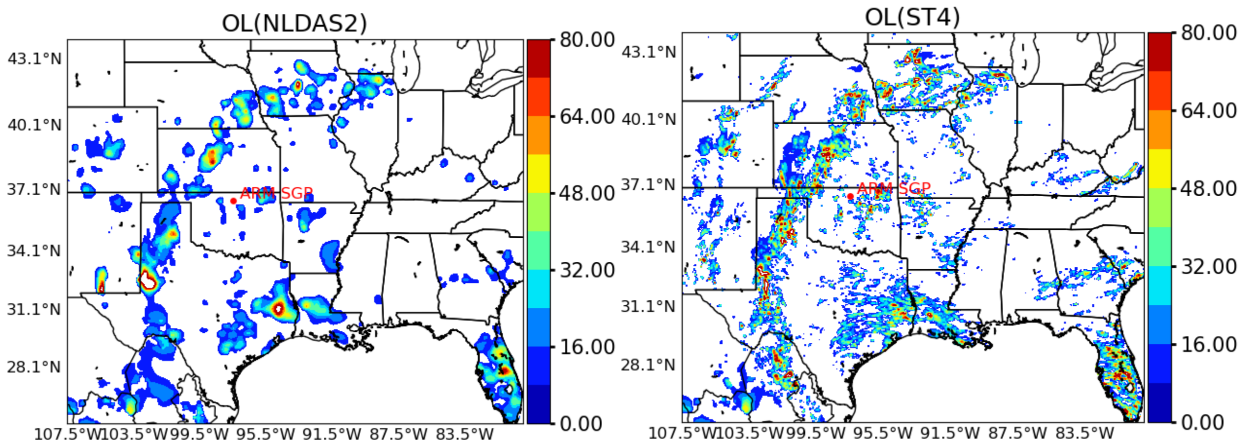

Fig. 4 Precipitation rate (unit: mm day^{-1}) maps extracted from NLDAS-2 (NLDAS2, left panel) and Stage IV (ST4, right panel) at 00 UTC on August 30, 2016.

Reference:

Fast, J. D., and Coauthors, 2018: Overview of the HI-SCALE Field Campaign: A New Perspective on Shallow Convective Clouds. Bull. Amer. Meteor. Soc., https://doi.org/10.1175/BAMS-D-18-0030.1.

4. Why was only the upper soil moisture compared? I understand SMAP is the limiting factor here, but I'd argue the benefit of a (DA) model is to investigate the variables (e.g., root zone soil moisture) not directly sensed by the satellite. Given the comparison with ERA5-Land, soil moisture at lower depths would be helpful for comparison/validation as well.

Reply: We acknowledge that evaluation/comparison in the context of root-zone soil moisture would potentially provide more insights into how model/assimilation acted to represent soil moisture in deeper soil layers. Nevertheless, as you noted, SMAP and GLASS SM provide soil moisture estimates only for the surface layer, and other products and in-situ observations use different depths or layer thicknesses to report soil moisture. In-situ observations likely report a point at a vertical level (e.g., 25 cm), while the interpretation of model-based products is rather unclear (point or layer-average?). The table here summarizes the depths/layers of soil column used in the models and in-situ measurements:

Dataset	Depths/layers (cm)	Notes
SMAPDA	10, 30, 60, 100	4-layer model output
ERA5-Land	0–7, 7–28, 28–100, 100–289	4-layer model output
USCRN	5, 10, 20, 50, 100	In-situ sensors
SCAN	5, 10, 20, 50, 100	In-situ; may vary by site
OKMet	5, 25, 60	In-situ; fewer layers

Therfore, unless different datasets report soil moisture at similar depths or report vertically integrated water amount (which seems complicated for in-situ measurements), we feel it is difficult to make a fair comparison and draw a firm conclusion for deeper soil layers. Having said that, other ongoing work compares soil moisture simulated in Large-Eddy simulations using SMAPDA as the soil initial condition to in-situ measurements at ARM SGP site to assess SMAPDA's realism in the deeper soil layers. Also, we discuss the potential of this further analysis in the Section 5 (Summary and discussion) from L638 to L639."

5. An open-loop simulation would be important for a base comparison. How do you justify not running OL simulations and comparing them against the DA outputs? Furthermore, more details on the modeling set-up could be provided. How much spin-up time was used? Although the purpose for analyzing data from 2016 is given (major drought year), either more work or reasoning needs to be done for limiting the study period to one year. E.g., how do you know your results are representative of the datasets themselves, and not just valid for times of drought?

Reply: Thank you for commenting on the absence of open-loop (OL) simulation in the analysis. To address the comment, we've extensively updated the result section with additional results obtained from OL simulation. Meanwhile, the Section 3.3 was also revised to include more descriptions about the model setup for the OL simulation.

In this revision, the 5-year (2012 to 2016) outputs from OL simulation are also leveraged in the computation of climatological references in the discussion of the 2016 drought in southeastern U.S. More details please refer to the paragraph between L541 and L549.

Minor comments:

6. The number of stations should be explicitly described, and/or mention the OK (the state) in Figure 1 when describing the ARM SGP sites. The white circles were a little difficult to identify at first. The number of stations for validation would also help contextualize Figure 4.

Reply: The number of observational sites for Oklahoma Mesonet and ARM SGP are added for clarification. The sentence in L206 is revised accordingly as: "The OKMet (120 sites) and ARM SGP (6 sites) observations are adopted…"

7. Line 45 - is (O (10 km)) a typo?

Reply: No, it represents it is in the order of 10 kilometers.

8. Line 320, Figure 4 - GLASS is stated to have more off-diagonal terms, but the concentration and shape of off-diagonal terms in the SMAPDA plot make that statement appear moot. The total RMSE for GLASS is also slightly lower than the SMAPDA product.

Reply: Good catch. We decide to drop this statement.

9. Line 482 - Sites 6, 8, 14 also show high bias during some months but are more sandy soils. What do you attribute the cause to be? Could it be a question of sampling? (e.g., less clay sampled available, whereas more sandy time series exist)

Reply: Although the sites 6, 8, and 14 also show relatively larger differences between models and observations compared to other sites with sandy soil, the observed soil moisture does not exceed 0.5

mm mm^{-3} at those sites. Whereas maximum soil moisture in the three clay-soil sites can be even larger than 0.6 mm mm^{-3}.

To understand if this conclusion is statistically robust, we examined the results for all the sites (USCRN and SCAN) located in the study domain. Fig. 5 shows the RMSE of SMAPDA soil moisture computed in a function of soil types. Although the sample size varies among soil types, the soil moisture error is generally larger at sites with clay soil than other soil types. An extended analysis focusing on its dependency on landcover types (Fig. 6) further shows the differences in error statistics in association with landcover types are less distinct. This suggests that our hypothetical conclusion is likely true despite other factors could potentially contribute to the errors. The description of extended analysis is included in the revised manuscript from L569 to L571.

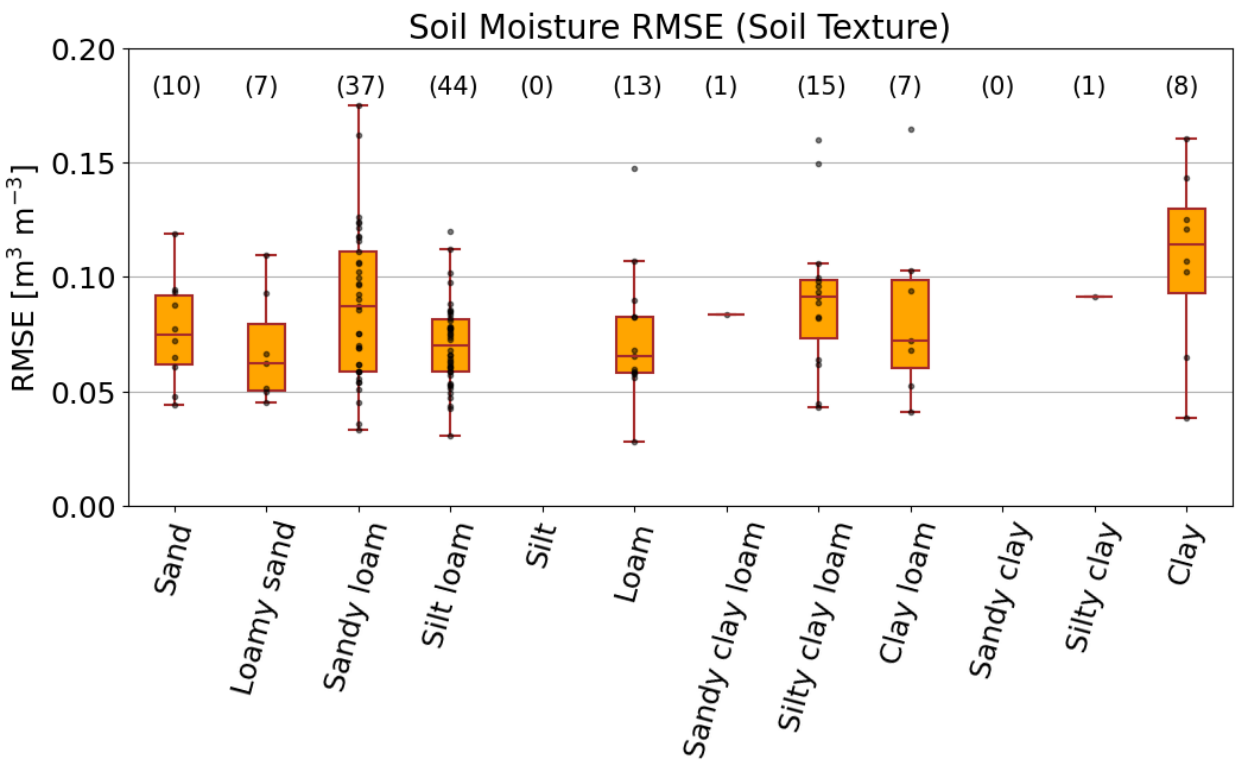

Fig. 5 SMAPDA soil moisture RMSE computed for all SCAN and USCRN sites as a function of soil texture. The number of sampling sites for each soil texture type is given in the parenthesis.

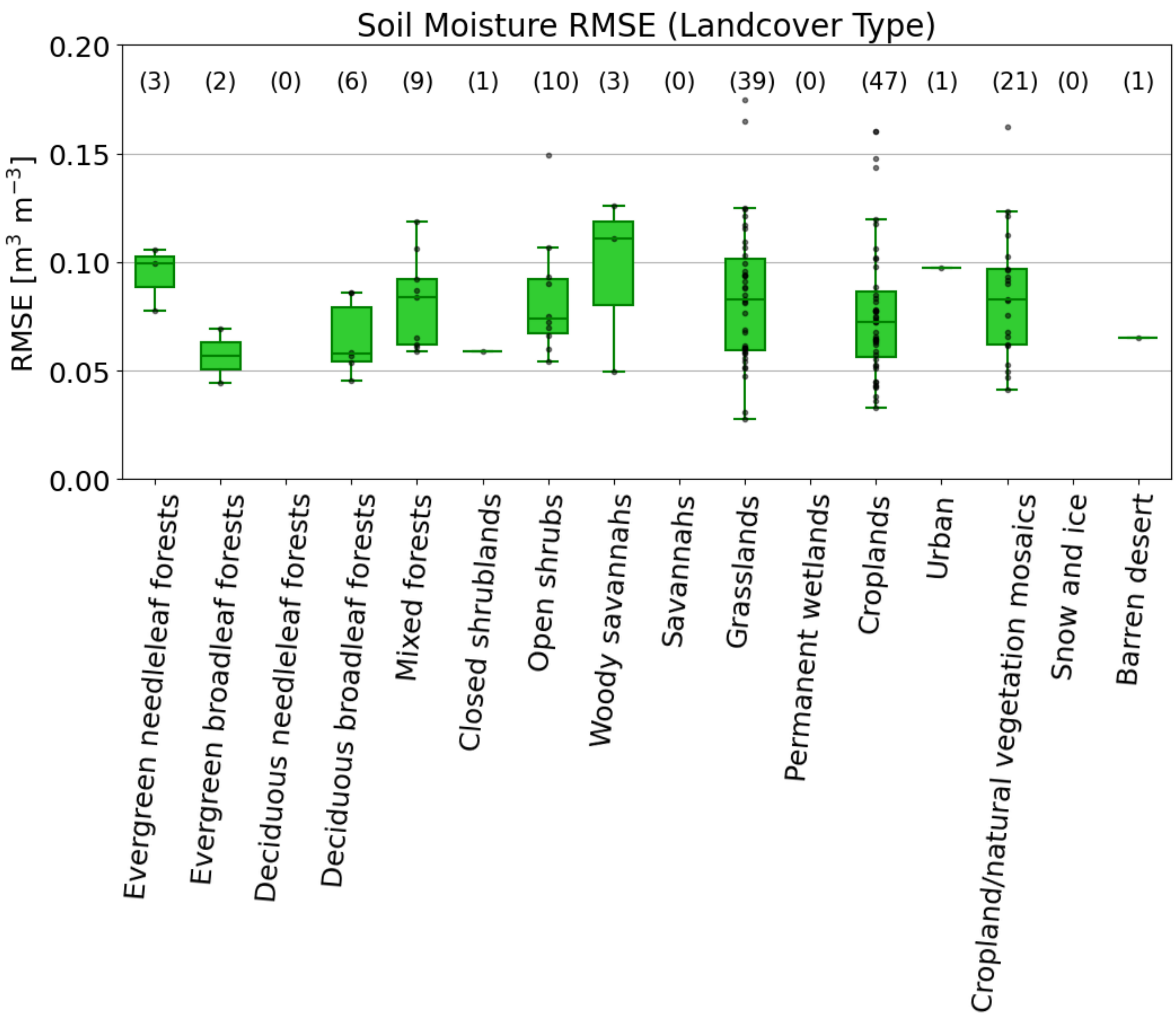

Fig. 6 Similar to Fig. 5 but for landcover types.

10. Line 444 - The soil moisture bias doesn't appear to be that great compared to the in-situ data. And LHF is only marginally larger for SMAPDA than the observed during the summer. The differences in the summer SHFs also show that they are not within error/standard deviations. Are there other causes you could attribute this discrepancy to?

Reply: According to the equation of SHF, the magnitude of SHF is primarily driven by the temperature gradient between the ground surface and the air:

$$SHF = \rho C_p \frac{T_s - T_a}{r_a}$$

Where T_s is the surface temperature (K), T_a represents air temperature (K). ρ is air density (J kg^{-1} K^{-1}), whereas C_p stands for specific heat capacity of air (kg m^{-3}). r_a denotes aerodynamic resistance (s m^{-1}), which is affected by wind speed, surface roughness, atmospheric stability. Note that the SHF may be changed drastically when the temperature gradient shifts direction or magnitude. As also noted by Xie et al. (2012, 2014), uncertainties in meteorological forcing (NLDAS-2) in terms of air temperature, wind, and stability can potentially modulate the difference in SHF between SMAPDA and local observations during the summer months, when local land-atmosphere interactions become more important than in other seasons. For example, the NLDAS-2 forcing is based on the North American Regional Reanalysis on the 32-km grid, and downscaled to 1/8° grid, or ~ 12 km grid spacing; this grid spacing is not enough to capture the local-scale variability of near-surface meteorological variables happening at point-scale observations.

Another uncertainty is from the measurement, for which we only used the flux measurement from the eddy covariance (EC) systems. It was noted that there can be large differences between the fluxes from EC and those estimated by the Bowen-Ratio energy balance approach. In this study, we focused on eddy covariance measurement that is widely used in other national/international network (e.g., FLUXNET). Lastly, Noah-MP provides an extensive range of model configurations, e.g., bottom boundary condition for the soil layers (with or without ground water), soil hydrology, bare-ground evaporation resistance, canopy radiative transfer, etc., which will carry the impact of soil moisture assimilation differently to the surface flux."