Interactive comment on "Using data assimilation to optimize pedotransfer functions using large-scale in-situ soil moisture observations" *by* Elizabeth Cooper et al.

Anonymous Referee #2

This paper clearly and neatly shows a study on optimizing constants in the underlying Cosby peotransfer functions used by JULES model via assimilating daily-averaged COSMOS-UK soil moisture data through LaVEnDAR data assimilation approach. With calibrated values for PTFs constants, the paper shows updated soil hydraulic parameters representing on field scale and comparison results to those on small (∼cm) scale. With 'vsat' updated being large and 'satcon' and 'sathh' being small, underestimations of soil moisture shown as prior are corrected and simulated soil moisture as posterior shows consistency to in situ measurements. The proposed method in this paper is an alternative attractive way to contribute to improving soil water flow and heat transport simulations by land surface models.

I have four major comments and few mirror comments on the manuscript. I would suggest the consideration of accepting this paper after the author addresses major comments.

Major comments

1. At line 157-162, "The daily soil moisture measurements we use are averaged from 30 minute soil moisture measurements. . . . . .uncertainty in the daily values is approximately 20%. We have inflated this here to 50% observation error. . . . . .in fact there will likely be intra-site correlations between observation errors due to site-specific instrument calibration."

Here "uncertainty in the daily values is approximately 20%", what does uncertainty mean? Is it the standard deviation of soil moisture at a daily scale or 20% is an estimate accounting for the conversion from neutron counts to soil moisture? Is inflated 50% observation error as a result of an optimized one, how? How can it be proved that inflated error accounts for intra-site correlations between observation errors due to site-specific instrument calibration?

2. In Fig. 3, posterior shows matching to in situ measurements except for the underestimation of soil moisture during the soil wetting period (around 2018-04 and 2018-11), why? Is it related to PTFs structure itself? Compared to Fig. 3, please in Fig. 4, it is better to give numbers such as the correlation coefficient and RMSE.

3. At line 148, it is mentioned that soil texture information for each site was taken from the Harmonised World Soil Database (HWSD) (Fischer et al., 2008). As soil texture information is a base for obtaining optimized constants for pedotransfer functions, how about the quality of HWSD compared to in situ measurements? Fig. 8 and Fig. 9 show almost the same values for topsoil and subsoil, soil profile in the site is homogenous or because of used HWSD product? How do the optimized constants for pedotransfer functions and associated soil moisture change with different soil texture inputs? Additionally, please if available, add (measured) soil constituents for each site in Table 3 for

readers' information.

4. At line 245, "The new distributions allow the model to access higher soil moisture values, potentially correcting for a deficiency in supporting datasets, parameter values or process representation in JULES", please clarify supporting datasets, do you mean the deficiency of soil properties dataset?

Mirror comments

1. In Table 1, the unit of satcon, Ks shall be kg m-2 s-1. Please check.

2. In Table 3, for the last cell, please complete the phrase "mineral (soil) with very high organic content". Please explain the difference between Grassland/heath and Grassland.

3. In Fig. 10, what does the blue line mean?

4. Please keep the citation consistent, for example, (Best et al. (2011),Brooks and Corey (1964)), (Cosby et al., 1984; Marthews et al., 2014). At line 168, Gupta et al. (2009); Knoben et al. (2019)

5. Please replace "in-situ" by "in situ", which follows the convention—Latin phrases should not be hyphenated (e.g. "in situ", not "in-situ").