Reviewer 3

General Comments:

Authors present an approach which combines soil moisture predictions from the JULES land surface model with in-situ field scale observational data measured by cosmic ray neutron sensors of 16 sites. Cosby et al. (1984) pedotransfer functions were used to compute soil hydraulic parameters for the JULES model. The manuscript shows that JULES model performs better in the prediction of soil moisture if the constants of the pedotransfer functions are calibrated based on field-scale soil moisture observations. This way soil physics parameters of the JULES are not directly optimized. The manuscript presents a new approach to improve performance of JULES model in soil moisture prediction. It is a high quality research, has interesting results and is well structured. Only one aspect could be explicitly clarified, if soil textural information was derived from a course resolution raster dataset in the presented analysis. If that is the case, it would be important to discuss how uncertainty of soil textural data influences the performance of the prior JULES run.

The authors thank the reviewer for their useful comments. We respond to the reviewer's comments here; we have numbered the specific comments and regrouped them in some cases. Our responses are given in blue text and planned changes to the manuscript are given in green.

Specific Comments

1. Title, L126, L252 and L262: In most of the text COSMOS-UK observations was mentioned as field-scale observations, except in the title, L126, L252 and L262, where large-scale is written. It might be better to call it field-scale. Please revise entire text to be consistent in using filed-scale and large-scale. L126: In the above text COSMOS-UK observations was mentioned as field-scale observations, here "large-scale" is written. It might be better to call it field-scale. Please revise it.

We will revise the use of 'large scale' to 'field scale' throughout the paper.

2. L84-90: Reference of equations 8-11 is not clear, could you please clarify it or add the reference?

We will clarify the sources of these equations with extra text after equation 11:

Equations (8) and (9) are rearrangements of equation (2) at fixed values of matric suction corresponding to the wilting and critical points. Equation (10) is a linear combination of the assumed heat capacities of sand, silt and clay, weighted by their relative fractions, and equation (11) is as given in Dharssi et al (2009).

3. L80: In the original Cosby et al. (1984) paper (Table 4 on page 686), the multiple linear regression of the "Absolute value of the soil matric suction at saturation" uses silt% and sand%, but the equation 6 of the manuscript includes clay% and sand%. Please recheck the equation or add further reference if a modified version of Cosby et al. (1984) pedotransfer functions are used.

   Table 2: The constants needs a further check, compared to Table 4 of Cosby et al. (1984), because of the following. It is not clear: - why k2 and k3 are multiplied by 100; - why k4 is divided by 100 and in the same time the original values of k5 and k6 are kept; for predicting

volumetric water content in m3/m3: also k5 and k6 has to be divided by 100 or do you consider sand and clay content as g/g (not weight %); - why k7, k8, k9 constants differ from the original constants, please note that in the original PTF silt% and sand% are the predictors as mentioned above, please clarify in the text why the constants differ from that of Cosby et al. (1984); - why k11 and k12 are multiplied by 100, do you consider sand and clay content as g/g (not weight %)? If you find after the check that constants of Cosby PTF is are those are built in the JULES model it might be helpful to check those also in the model code. Please add the units and fraction limits of clay, silt and sand content in line 91.

The differences in the values of constants between Cosby et al (1984) and here are in part due to conversions of units (from e.g. inches per hour to kg m^{-2} s^{-1} for Ksat). As the reviewer notes, we have used clay and sand as predictors in all equations, using the fact that $f_{sand} + f_{silt} + f_{clay} = 1$; this also changes the value of some of the constants. The ptf constant values we have given here match those in table 1 of Marthews et al (2014) with a small exception. While Marthews et al (2014) express clay, sand and silt fractions as percentages, we use fractions (i.e. in Marthews et al, $f_{sand} + f_{silt} + f_{clay} = 100$). This means that the multipliers given for f_{sand} and f_{clay} are 100 times larger in our version. We will add a reference to Marthews et al (2014):

The values of the constants given here match those in Marthews et al (2014) (with soil fraction multipliers adjusted for fraction, rather than percentage, of soil by weight).

To clarify, the PTF is not built into the JULES code; users are required to provide values for the soil physics parameters, but can calculate these via any choice of PTF (or other method).

The units for f_{sand} are fraction by weight, i.e. dimensionless, and we will add this to the text at line 91.

4. L95-97: Please list meteorological data required by JULES to derive soil moisture prediction.

We will add this information.

The required input variables are: air pressure, air temperature, humidity, downward fluxes of shortwave and longwave radiation, precipitation and wind speed.

5. L110-112: Please consider that CHIMN, PORTN, HARTW, LULLN are mineral soils too based on Table 3, therefore the sentence starting with “The Cosby pedotransfer function . . .” needs to be revised.

We will revise this sentence to read:

The Cosby pedotransfer function was designed to work for mineral soils, and the CRNS calibration is most reliable at sites with minimal vegetation. We therefore consider that the first seven sites listed in table 3 are those at which the JULES model can be expected to provide a good match to observations via our chosen PTF; soil types and land cover at the remaining sites mean that JULES may not be able to represent the observed soil moisture time series as accurately.

6. Table 3: Instead of the basic soil description it would be more informative to provide soil taxonomical information, i.e. name of soil suborders (USDA, Soil taxonomy) or reference soil

groups with principal qualifiers (WRB, 2014). If soil taxonomical information cannot be added, soil texture, organic carbon content and bulk density of topsoil and subsoil could be shown, if that is available for the COSMOS-UK sites.

L120: Are measured soil chemical and physical properties available for the COSMOSUK sites.

Unfortunately we do not have access to any further soil texture, chemical and physical properties, or taxonomical information for the soils at COSMOS-UK sites.

7. L119-120: sentence starting with “We have used . . .” is repetition of the first part of the sentence starting with “In this paper . . .” in line 95-96.

We feel that it is useful to remind the reader of this at this point in the paper.

8. L124: The reference for LaVEnDAR is given, but it might be helpful for the readers if a very short description of the data assimilation technique would be given in the text.

We will add the following short description of the algorithm:

LaVEnDAR optimises k1 to k12 here by minimising a cost function with two terms. The first term is a measure of the difference between the observed and modelled soil moisture, and the second term is a measure of the difference between prior and posterior values of k1 to k12.

9. L133: Please add the meaning of “75m” or delete it if it is not important.

We will clarify this by adding the following text

'The observed depth changes with soil moisture and with distance from the CRNS instrument; here we have used the reported observation depth at 75m from the CRNS. For each day, we calculate a depth-adjusted JULES soil moisture estimate, SM_{depth}, depending on the 75m observation depth value, D86, provided for that day, such that..'

10. L148: Is not measured soil texture available at the COSMOS sites? Uncertainty of texture taken from the Harmonised World Soil Database (HWSD) can be high, because its resolution is 30 arc-second. If texture is derived from a course resolution dataset the lower performance of prior JULES run can come from the uncertainty of clay, silt and sand content. It would be interesting to analyse the performance of prior JULES run at a site where measured soil texture can be used in the Cosby pedotransfer functions. If there is no measured soil texture data, better resolution national soil texture maps or 250 m resolution SoilGrids could provide more accurate soil textural information than HWSD does. Please consider to rerun analysis based on a more accurate soil texture dataset or explain why HWSD was used. It would be good to highlight importance of using measured soil texture if that is available.

Reviewers 1 and 2 made similar comments. We used a global soil texture dataset here because we wanted to make sure our method would work when local measurements are not available, and in fact we do not have soil texture data for the COSMOS sites, only the broad descriptions given in the COSMOS-UK user guide (v2). We feel that rerunning the experiments using an alternative soil texture database would lead to an interesting comparison with the work here, but is out of the scope of this paper, which aims to demonstrate a new method for calibrating ptf constants. We plan to add text to explain our choice to use the HWSD:

We assume that the soil texture values from the HWSD are correct; they are not changed during the data assimilation process. We used a global soil dataset rather than locally available soil texture observations to ensure that our method has the potential for extension to areas without local measurements. Other open source global soil texture products are also available (e.g. SoilGrids Hengl et al (2017)). We acknowledge that there may be discrepancies between the HWSD and local measurements (e.g. Zhao at al (2018)), but our choice to use the HWSD here follows recent successful integration of soil texture data from the HWSD with JULES in studies such as Martinez de la Torre (2019), Ritchie et al (2019) and Ehsan Bhuiyan et al (2019)

11. L158-161: Does it mean that higher observation error was used when results of soil moisture predictions was assessed than the error computed based in the measured data? The reasoning of it is not clear, could you please describe it? Sorry if I miss something.

The inflation of observation error is for use in the LaVEnDAR algorithm and is a reasonably common technique in data assimilation. We will clarify this with the following text

The daily soil moisture measurements we use are averaged from hourly soil moisture measurements; analysis of the data shows that the standard deviation of the hourly data around the hourly mean is approximately 20%. We have inflated this here to 50% observation error; we note that similar observation error covariance inflation techniques have been used in e.g. assimilation of satellite observations in numerical weather prediction (Fowler (2018), Hilton(2009)). The reason for inflating the observation error is essentially because we found that smaller observation error values impacted negatively on the posterior soil moisture results. We suggest that inflation of the observation error is necessary here to compensate for otherwise neglected sources of error (e.g. the error in converting neutron counts to soul moisture) and for the assumption of uncorrelated observation error; in fact there will likely be intra-site correlations between observation errors due to site-specific instrument calibration.

12. Figure 2. Maybe the following could be added: - Data assimilation (LaVenDAR), - 16 sets of field-scale obs,

We will update the schematic to include these suggestions

13. L185: Please add which software was used to compute the metrics and prepare plots.

We will add this information at line 185

‘We used python 3.7.1 to calculate metrics and prepare plots.’

14. L197: Please add under Materials and methods section which method was used to analyse if difference was significant.

We used ‘significant’ in a non-mathematical sense here. We will replace ‘significant’ with ‘marked’.

15. L205-206: It would be informative to roughly add the soil organic content of MOORH site, if measured value is available that would be the best. Could you please add reference to the CRNS regarding soil organic carbon content and texture that can be reliably measured?

Unfortunately we do not have any further reliable information about the soil organic content at MOORH.

16. L208-209: It could be mentioned that it is a disadvantage that CRNS measurement considers water held on the canopy to be soil moisture. Is there any solution for correcting the COSMOS soil moisture values if that happens?

Soil moisture measurements are calibrated at each COSMOS-UK site, and this aims to correct for water stored on vegetation. However, vegetation makes the calibration less reliable for a number of reasons. We will add text to make this clearer

'.. which is likely due to the fact that there are a large number of trees at this site. This means that the presence of aboveground biomass may make the site-specific calibration less reliable than at other sites (Baatz et al. (2014)). The high organic carbon content of the soil at Gisburn Forest likely also contributes to this as our chosen PTF is designed to work best with mineral soils. Interception is another processes which potentially complicates the calibration at sites with vegetation, although the authors of Bogena et al (2013) report that water intercepted by the canopy constitutes a negligible amount of the water detected in the CRNS footprint, even in coniferous forests. '

17. Figure 8. Please add soil depth that you consider topsoil.

The depths are 0 - 35cm and 35cm – 300cm for topsoil and subsoil layers respectively. We will add assumed depth information to the captions of figures 8 and 9.

18. L240: Do you think the profile-scale measurements could be successfully used in the presented data assimilation method?

An alternative approach would have been to use point scale measurements in our experiments. However, point sensors only measure the soil moisture in a very small area and are therefore not representative of the soil moisture on the scales that JULES is typically used. We see from point sensors at COSMOS-UK sites that sensors quite close to each other can measure quite different soil moisture values due to their different very localised conditions. We chose to use field-scale measurements here in order to average out the local variations in observed soil moisture and to better match the scales over which JULES is typically used.

19. L274: The code is available only for those who are registered for a Met Office account, it might be mentioned.

We will add text to clarify this.

Technical Corrections:

L91: . . . where fclay, fsilt and fsand are fractions of clay, silt and sand in the soil . . .

L143: Do you mean: “the value given in table 2”? Please revise it.

L193: . . . high soil organic carbon content . . .

L229: . . . 12 PTF . . .

Thank you for spotting these errors, which we will correct.

References:

Baatz, R., Bogena, H., Hendricks Franssen, H.-J., Huisman, J., Qu, W., Montzka, C., and Vereecken, H.: Calibration of a catchment scale cosmic-ray probe network: A comparison of three parameterization methods, Journal of Hydrology, 516, 231 – 244,https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.02.026 , 2014

Bogena, H. R., Huisman, J. A., Baatz, R., Hendricks Franssen, H.-J., and Vereecken, H.: Accuracy of the cosmic-ray soil water content probein humid forest ecosystems: The worst case scenario, Water Resources Research, 49, 5778–5791, https://doi.org/10.1002/wrcr.20463,https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20463, 2013

Dharssi, I., Vidale, P.L.and Verhoef, A., Macpherson, B., Jones, C., and Best, M.: New soil physical properties implemented in the Uni-fied Model at PS18, Met Office Technical Report Series, https://library.metoffice.gov.uk/Portal/Default/en-GB/RecordView/Index/251847, 2009. [Accessed 14th October 2020]

Marthews, T. R., Quesada, C. A., Galbraith, D. R., Malhi, Y., Mullins, C. E., Hodnett, M. G., and Dharssi, I.: High-resolution hydraulic parameter maps for surface soils in tropical South America, Geosci. Model Dev., 7, 711–723, https://doi.org/doi.org/10.5194/gmd-7-711-2014, 2014.