Editor

Your manuscript "Improving Soil Moisture Prediction of a High-Resolution Land Surface Model by Parameterising Pedotransfer Functions through Assimilation of SMAP Satellite Data" has been subjected now to review by three reviewers. All of them recommend major revision. Given the many and different comments by the reviewers, the manuscript is borderline to rejection. Please consider whether it is possible to handle all reviewer comments in due time. The main points to be handled are: (i) the introduction should be strengthened with references; (ii) the manuscript lacks details in many places, for example regarding the data assimilation setup, measurement operator, handling of the vertical scale mismatch (SMAP vs model); (iii) time series of estimated soil parameters should be presented, and an evaluation for filter inbreeding (underestimation of variance); (iv) the handling of bias should at least be discussed, and additional simulations may be necessary; (v) the results and discussion sections should be improved with more scientific interpretation and discussion.

In summary, I suggest major revision and additional review of the paper. Please consider whether such a revision is feasible in the available time framework.
In your answer to the main points and detailed comments, please indicate how comments have been handled exactly, indicating also whether text has been deleted and what the position of newly included text blocks is. I am looking forward to the new version of the paper.

Dear Professor Hendricks-Franssen,

We thank you for providing a decision on our manuscript and the clarification of points which need addressing. We have made substantial changes to the manuscript in line with the reviewers comments and have conducted additional model and experiment runs to provide additional output variables and address issues of the SMAP vs model depth mismatch. We have increased the level of detail throughout the manuscript. In relation to point (iii) we are unable to provide time-series of soil parameter values since we use a smoother with a single assimilation window. This means that we consider all the available observations over the spatial domain in a time window of a given length. In our case, we chose this length to be the entire experiment time frame. This can be done since we are searching for parameters which are static in time. In retrospect we think we had not adequately described the DA technique so have strengthened this as requested and included a diagram. For point (iii) we have instead included maps of how the soil parameters change and also a plot of time-series RMSE and ensemble spread. We have discussed the handling of bias and strengthened both the results and discussions section, including additional figures and reference to other scientific literature. Thank you again for considering our manuscript and please find our detailed responses and updates below along with a marked-up version of the new manuscript.

Kind Regards,
Ewan Pinnington

Reviewer 1 (R#1)

The study describes results of a data assimilation experiment, assimilating soil moisture data of the Soil Moisture Active Passive mission into the UK land surface model JULES. The assimilation updates states and parameters. Resulting soil moisture is compared to SMAP data and data of an independent network of cosmic ray neutron probes.
The title and general content of the manuscript are promising, while the manuscript itself exhibits lack of detail which would be required for following the study and reproducing the results. Below, my concerns, starting with the general ones, and followed by detailed comments.
We thank the reviewer for their comments which will undoubtedly help to strengthen this manuscript. We outline below our responses and updates to the paper.

1. Well known bias in the SMAP satellite product and impact on pedotransfer functions is not discussed (e.g. Reichle et al. https://doi.org/10.1029/2019MS001729 or Colliander et al 2017 https://doi.org/10.1016/j.rse.2017.01.021). This would be a key asset of the paper.
2. Which SMAP level data was used. It will help the reader in understanding the results. Please point this out in the introduction and methods sections. What are the implications?
This is a good point. As per the paper mentioned, if the Level-4 SMAP product was used that is biased high there could possibly be an impact on the retrieved pedotransfer function (PTF) parameters. This would likely exhibit itself in PTF parameters that would artificially increase the values of the saturated soil moisture and possibly decrease saturated conductivity given the underlying soil textural information. We have included discussion of this at line 387:
"For SMAP any bias contained in the observations could cause us to retrieve PTF parameters that result in erroneous soil hydraulic conductivity's and ultimately degrade the performance of other model components. It has been shown that the Level-3 9 km SMAP observations used here do not have a significant bias (Colliander et al., 2017) especially in temperate regions (Zhang et al., 2019). The fact that after assimilation of the SMAP data we not only reduce the RMSE of JULES compared to SMAP but also reduce the RMSE of JULES compared to independent COSMOS estimates also gives us confidence that the bias in the assimilated SMAP data is relatively low."

We used the L3 SMAP v3 9-km radiometer-radar combined product. We have included this in the introduction at line 84:
"[…]high quality SMAP data (here we use Level-3 SMAP soil moisture observations) and a high distribution of COSMOS probes [..]"

We have also included more information in the methods section on the Level-3 data and bias as requested at line 146:
"For the work in this paper we use the 9 km Level-3 soil moisture product (version 3) this product has a relatively low bias (Colliander et al., 2017, Zhang et al., 2019). However, it has been shown there is a wet bias present in the Level-4 SMAP product (Reichle et al., 2017). As part of the retrieval procedure SMAP relies on some ancillary information, one example of this is soil texture where the Harmonized World Soil Database (HWSD) (Fischer et al., 2008) is used to calculate the soil dielectric constant for use within the retrieval algorithm. The use of such ancillary data in the retrieval could introduce additional biases into the SMAP soil moisture estimates that are not consistent with estimates from the land surface model we are comparing to. However, as the HWSD is also used to create the JULES soil parameter ancillary files this effect should be minimised."

3. Discussion is not based on literature but merely on own postulations. A good guide is located here: https://www.biosciencewriters.com/How-to-Write-a-Strong-Discussionin-Scientific-Manuscripts.aspx
We agree the discussion could be strengthened and have endeavoured to do so by restructuring and including more literature. See also later related points.

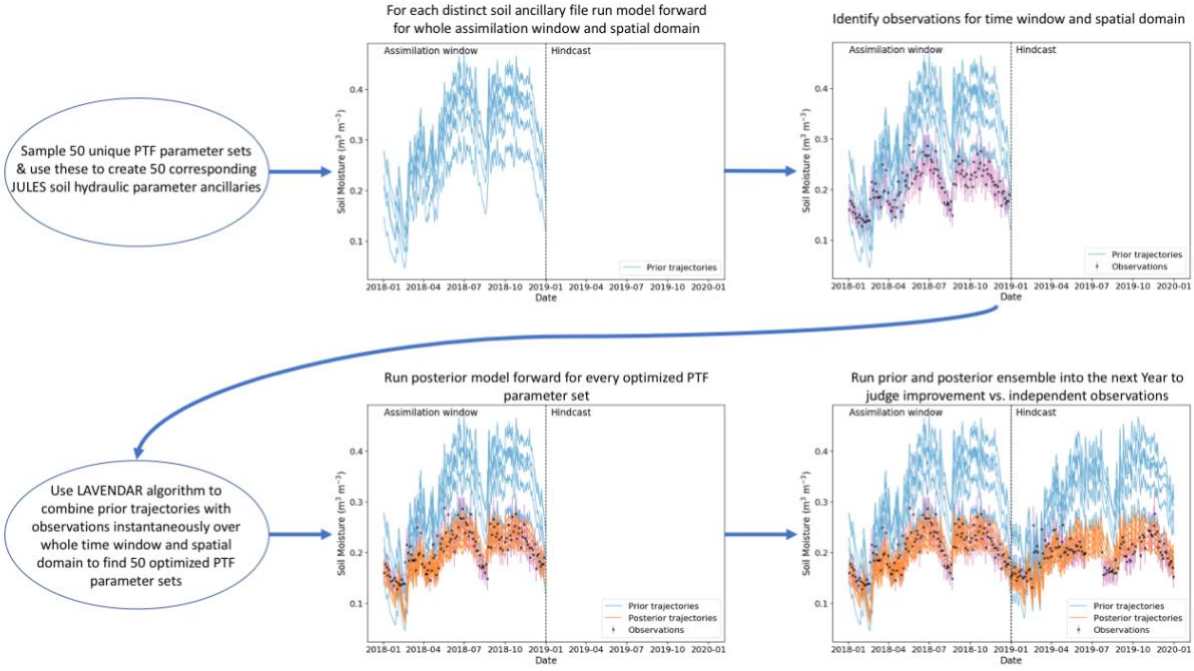
4. Please add conceptual details on how the 4DEnVar (an optimization method) is combined with EnKF (optimization) (see page 7 lines 159-164). I imagine this can be done by text or together with a figure. Also address why are both optimization methods combined at all?
5. Please add how is the state vector in Appendix A is composed in the present case (variables, parameters, IenKS posterior?) and which units do the variables in Appendix A have.
6. Please clarify, what are prior and posterior with respect to two data assimilation methods? How can posteriors be worse then priors considering that the results are optimized using the evaluation data? Please plot as well the data assimilation performance over time with regard to RMSE and parameter convergence as for example in Poterjoy et al. 2017 https://doi.org/10.1175/MWR-D-16-0298.1, Botto et al. 2018 https://doi.org/10.5194/hess-22-4251-2018 and Baatz et al. https://doi.org/10.5194/hess-21-2509-2017 .
7. Please add results after the 4DEnVar assimilation in order to demonstrate what an additional assimilation yields in terms of skill.
We have grouped together points 4-7 here as we believe these all stem from us not adequately describing the data assimilation technique used in the current manuscript. We have referenced a previous paper centred around the development of the technique and have not supplied enough information here for readers to properly understand what we have done.

4. 4DEnVar is not combined with the EnKF, 4DEnVar is a hybrid technique combining elements of both ensemble and variational data assimilation methods. On reflection the way we have described this in the manuscript is not clear and we believe this has caused a misunderstanding of the results. The method we have used is closer to that of the Iterative Ensemble Kalman Smoother (IEnKS). We have removed references to 4DEnVar to avoid confusion and have strengthened the description of the DA method at line 181. We agree that the use of a diagram will be beneficial to illustrate the technique (see diagram below).
" In order to estimate the identified pedotransfer function parameters we use the LAVENDAR data assimilation framework (Pinnington et al., 2020). This framework utilises a hybrid DA technique similar to that of the Iterative Ensemble KalmanSmoother (IEnKS) (Bocquet and Sakov, 2013). A smoother is different than a filter (e.g. the Ensemble Kalman Filter (Evensen,2003)) in that it uses batches of observations which are taken over a time window of given length and the whole spatial domain, as opposed to just in a time instant. These observations are combined with the model evolution over this window and a minimization process is performed to obtain initial conditions for the state/parameter values. It is possible to run a sequence of smoother steps for successive windows, but our study only uses one year long assimilation window as the parameters we are optimising do not vary in time.
Using a smoother instead of a filter has advantages (Lorenc and Rawlins, 2005) in that (a) more observations can be used to constrain the problem solution, and (b) information from the model evolution is implicitly used in the search process. However, using a smoother requires computing the Jacobian of the model, the so–called tangent linear model (TLM) and the related adjoint model (AM). The TLM/AM (Courtier et al., 1994). Computing and maintaining the TLM/AM is not a trivial task, and in fact we do not have this for JULES. The IEnKS solves this problem by replacing the role of the TLM/AM by 4–dimensional covariances, i.e. covariances defined over time and space. These covariances are computed as sample estimators of a given ensemble. The iterative nature of the method means that it finds the solution to the minimization problems using inner iterations rather than a single step (hence the variational nature), and this helps when the distributions of the variables/parameters of interest are not Gaussian. We provide details of the method in Appendix A.

Furthermore, to understand the variants of the ensemble Kalman Smoother and its position within the hybrid DA methods, the reader is referred to Evensen (2018).

We show a schematic of how this system works in Figure 3, […]"

We have also updated text in the introduction to make this more clear at line 71:
"Many previous studies optimising model soil parameters have taken a filtering DA approach (Moradkhani et al., 2005; Montzka et al., 2011;Han et al., 2014; Baatz et al., 2017; Botto et al., 2018) leading to the recovery of a time-series of parameter values as additional data is assimilated through time. In this study we use a smoother method, i.e. one that uses all observations in the spatial domain within a time window of a given length. Then, the static parameters are obtained by a single minimization process (which can contain iterative steps). Smoothers can be used in a sequence of 'analysis windows' (as it is done in operational numerical weather prediction), but in this study we only use one of these windows since the parameters we search for do not vary in time."



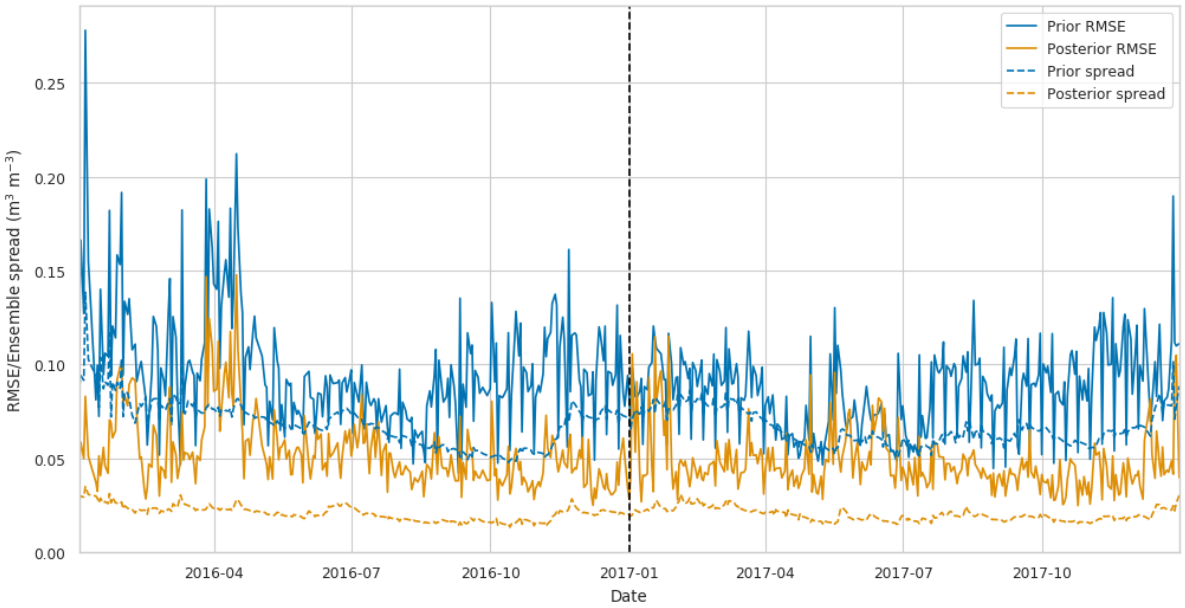
5. In Appendix A the state vector is the vector of 15 PTF parameters as defined in section 2.2 Table1 We have included this in the Appendix at line 460:
"In the case of this paper the variables and parameters correspond to the 15 PTF parameters in Table 1. […] In our experiments each x^i_b corresponds to a unique set of 15 PTF parameters and Ne = 50. […] where y are the observations for the whole time-window and spatial domain (here 2016 SMAP observations over the East of England, with units m3 m−3), H and h are the linearised and non-linear observation operator respectively (here the JULES model, which includes both a time integration and conversion into observation space to match the SMAP observations) and R is the observation error covariance matrix (here containing the error estimates for the assimilated SMAP observations)."

6. There is only a single assimilation step being used, we have aimed to make this clearer (see point 4 response). Using only one assimilation window is feasible because the parameters we are looking for are static in time. In Figures 4 to 7 the prior is just the mean and standard deviation of the 50 prior JULES ensemble members before DA and the posterior is the mean and standard deviation of the 50 posterior JULES ensemble members after DA. We have expanded the description of prior and posterior in the results section at line 240:
"The input to the data assimilation routine is an ensemble of 50 unique Tóth et al. (2015) PTF parameter sets drawn from a prior distribution (representing our best a priori guess to the true PTF parameters), the corresponding JULES runs (2016-2017) for each PTF parameter set and all the SMAP

observations for the year 2016 over the experiment domain. The output of the data assimilation is an ensemble of 50 optimised (posterior) PTF parameter sets, valid for the whole experiment domain and time, this allows us to calculate the posterior JULES soil ancillary files for each optimized parameter set and the corresponding posterior JULES model runs for 2016-2017."

We are optimizing 15 PTF parameters for the whole time window (<28000 observations) and the whole spatial domain (<30000 gridcells) in a single assimilation step by minimising a cost function. This is unlike sequential methods such as the EnKF or ETKF which step through time updating estimates at each step with available observations. We retrieve a single set of 15 PTF parameters valid over the whole domain and for the whole time period. This means that the optimisation may have to degrade the fit at certain locations to allow the 15 PTF parameters to improve the picture as a whole. This could be due to errors at these locations in driving data, the underlaying soil property map or indeed in the model structure (as is the case over urban areas in our results). We have included text to this effect at line 276:
"As we are minimising a cost function to find optimised values of PTF parameters valid for the whole spatial and temporal domain it is possible the optimisation may have to degrade the fit of the model estimates to the SMAP observations at certain locations in order to improve the picture as a whole. This could be due to errors at these locations in driving data, the underlying soil property map or indeed in the model structure."

As the DA method here is fundamentally different from the techniques in the papers mentioned, we are not able to reproduce the requested plots for parameter convergence as we retrieve just one set of parameters valid for the whole time window. However, we can plot the RMSE over time for both the prior and posterior ensemble members (see plot below). We have also aimed to increase the distinction between previous filtering DA methods and the variational smoother method we have used for this paper (see also response to point 4). We have included relevant text at line 296:
"In Figure 10 we show the RMSE averaged in space for the JULES model prior and posterior mean estimate, when compared to SMAP, alongside the JULES model prior and posterior ensemble spread. At all times the posterior JULES RMSE is lower than that of the prior, showing that the DA system has found a set of PTF parameters that improve the fit to the SMAP observations through time, this continues into the hindcast period (2017) when judged against observations that were not included in the DA cost function. We find slight peaks in the RMSE values throughout the time period corresponding to wetter conditions, this could be due to slight errors in the precipitation driving data used to force the model. It is optimal to have an ensemble spread that matches the magnitude of the ensemble mean RMSE and this relationship should hold given a large enough ensemble size (Houtekamer and Mitchell, 1998). We can see that this relationship holds for our prior estimates. However, after DA the posterior ensemble spread is slightly lower than that of the ensemble mean RMSE. This is perhaps unsurprising as we are conducting just a single assimilation step using all observations (over 28000) at once in space and time with a relatively small ensemble size (50). This can lead to some of the posterior parameter distributions becoming narrow, as with increasing observations we increase the confidence in our posterior, thus tightening the retrieved distributions and reducing the model ensemble spread. This result suggests that ensemble inflation (Anderson and Anderson, 1999) may be necessary if this ensemble was to be used in subsequent assimilation experiments."



7. We believe this is already shown and hopefully with the strengthened description of the DA algorithm this has become more clear (see response to point 4).

8. Please expand on why to add another 1% SWC error to SMAP (from 0.04 to 0.05 cm3/cm3, page 6 line 123) and multiply by four (20cm3/cm3 error?) for observation inflation, a rather seldomly used method. Inflation is rather used for covariance inflation during the run time of the data assimilation experiment (e.g. Jamal and Linker 2020, https://doi.org/10.1002/vzj2.20000 or Whitaker et al. 2011 DOI: 10.1175/MWR-D-11-00276.1). Please cite more studies were observation inflation is directly used and discuss why a bias aware data assimilation method was not used (e.g. Ridler et al. 2018 https://doi.org/10.2166/nh.2017.117)

Although the baseline aim for SMAP is 0.04 cm3/cm3 other studies have found higher values; 0.043 (Colliander et al 2017 https://doi.org/10.1016/j.rse.2017.01.021), 0.054 (Zhang et al. https://doi.org/10.1016/j.rse.2019.01.015), 0.057 (Li et al. https://doi.org/10.3390/rs10040535). We have included extra text at line 153:
"We prescribe an error of 0.05 m3 m−3 for SMAP observations in the assimilation algorithm. Although the SMAP baseline aim for error is 0.04 m3 m−3 other studies have found slightly higher values for the error in Level-3 SMAP observations (0.043 m3 m−3 (Colliander et al., 2017), 0.057 m3 m−3 (Li et al., 2018) and 0.054 m3 m−3 (Zhang et al., 2019)), we therefore chose a value between these studies."

Although observation error inflation is seldom used in sequential filtering data assimilation it is quite common place in variational methods and smoothers (such as the one in this paper), especially in numerical weather prediction (for example, Wang et al. http://dx.doi.org/10.1029/2019JD031029, Bormann et al. http://dx.doi.org/10.21957/gq8j2gjp7, Fowler et al. https://doi.org/10.1002/qj.3183, Hilton et al. https://www.ecmwf.int/node/15331). Observation error inflation is required due to the fact that all observations are used at once in the assimilation whereby we minimise a cost function containing a prior term and an observational term. The greater the number of observations in the observational cost function term, the higher the weight they have in the optimization. This can lead to the prior term being completely negated and hence the retrieval of unphysical parameters. Observation error inflation would not be required if the correct specification for the observation error correlations (in space and time) and model error was included. These, however, are hard to diagnose and it has been shown that in the absence of such information inflation is required for an optimal DA

system (Stewart et al. https://doi.org/10.1002/qj.2211). It has also been shown that for variational DA model errors can be included in the observational cost function term by inflating the diagonal variances, (Howes et al. https://doi.org/10.1002/qj.2996). We hope the improved description of the DA technique will also help here and the distinction between sequential and variational DA. Although we agree a bias aware data assimilation could be more optimal, the one proposed is in relation to a sequential technique (the ETKF) and we are using a variational method. We have added text around this at line 207:
"It has been shown that, for variational methods such as the one used in this paper, these additional sources of error (model error, representativity error, etc.) can be included in the observational term of the cost function by inflating the diagonal observation error variance (Howes et al., 2017). Although observation error inflation is rare in relation to sequential DA methods it is commonly used with variational methods and especially in numerical weather prediction (Hilton et al., 2009; Bormann et al., 2015; Minamide and Zhang, 2017; Fowler et al., 2018; Wang et al., 2019). Observation error inflation is required due to the fact that all observations are used at once in the assimilation whereby we minimise a cost function containing a prior term and an observational term. The greater the number of observations in the observational cost function term, the higher the weight they have in the optimization. This can lead to the prior term being completely negated and hence the retrieval of nonphysical parameters. Observation error inflation would not be required if the correct specification for the observation error correlations (in space and time), model error and representativity error were included. These, however, are hard to diagnose and it has been shown that in the absence of such information observation error inflation is required for an optimal DA system (Stewart et al., 2014). For this reason and due to the large number of observations assimilated in our one year assimilation window (28698) we inflate the specified observational error by a factor of four. If a filtering DA system were being used utilising a bias aware DA system such as that presented by Ridler et al. (2017) could help represent some of the additional sources of error discussed here."

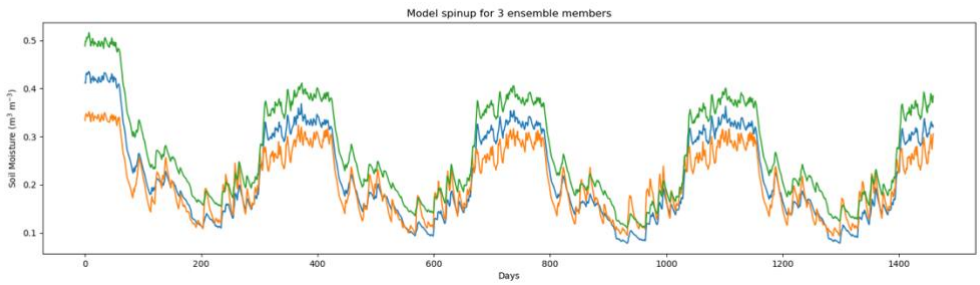
9. Please add legend to the graphs (Figure 6, 7 etc.).
Legends have been added to plots as requested.

10. Please discuss cross-correlation among the parameters of pedotransfer functions. From Equation 1 in the author's paper, it is clear that many parameters cross-correlate. Take for example Phi a and Phi c crosscorrelate strongly. What is the impact on saturated soil hydraulic conductivity?
We agree added discussion on this would be beneficial. We have added discussion at line 369:
"The correlated nature of the PTF parameters in equation (1) presents a potential source of equifinality (e.g. both ϕa and ϕc both act to increase the magnitude of θsat in the presence of clay soils), this means that we could achieve the same soil hydraulic conductivity with multiple realisations of PTF parameters at any individual grid cell. The effect of this is greatly reduced as we are performing the optimization over the whole domain and not on a grid cell by grid cell basis. In effect this means the unique soil properties at each of the 30614 model grid cells act as orthogonal constraints within the DA algorithm and reduce the issue of equifinailty for the optimized PTF parameters as the DA algorithm is having to fit the assimilated soil moisture observations for many different soil textures at once. It may also be possible to improve results further by including information on such correlations within our prior. Such estimates have been included in a variational DA framework for the carbon cycle and shown to improve posterior estimates (Pinnington et al., 2016)."

11. Please expand on the JULES hydrologic water components (ET, ground water, surface water flow, overland flow, infiltration, snow). How exactly was the 4 year spin up done? Was it done in ensemble mode? How were parameters perturbed? Please provide groundwater and soil moisture development over time at four cosmic ray neutron probe locations during the spinup period to elucidate the reader about the spinup performance.

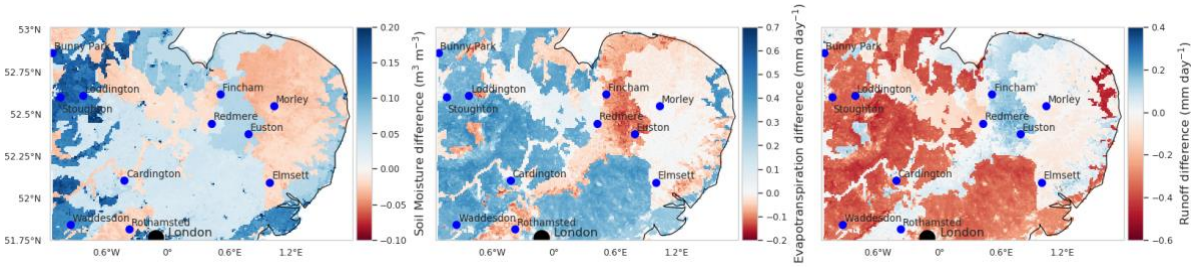
The spin up is done for each prior and posterior ensemble member, with the parameters either being sampled from the defined prior distribution or as outputs from the DA system in the case of the posterior. The model is run from an initial value (defined by the saturated soil moisture model parameter) over the same year of forcing data to reach an equilibrium soil moisture state for any given set of parameters. The plot below shows this for three distinct ensemble members which are all defined by unique sets of PTF parameters. We can see how these unique realisations of PTF parameters define unique soil moisture trajectories. The JULES model does not contain a groundwater component in the current configuration but we have added a plot of the spin-up for soil moisture to the supplementary material along with increased description of the spinup technique at line 111:

"It is necessary to find an appropriate initial state before running a land surface model such as JULES and it has been shown that without a suitable spin-up period forecast skill can be impacted (Maurer and Lettenmaier, 2004). We include a 4 year spin up period at the start of each JULES run to allow the soil moisture state to reach a point of equilibrium after parameter values are changed. For the JULES spin-up the model is run from an initial value (defined by the saturated soil moisture model parameter) over the same year of forcing data, here 2015, to reach an equilibrium soil moisture state for any given set of soil hydraulic parameters. We show this model spin-up for 3 unique soil parameter sets at the same location in Figure S4."



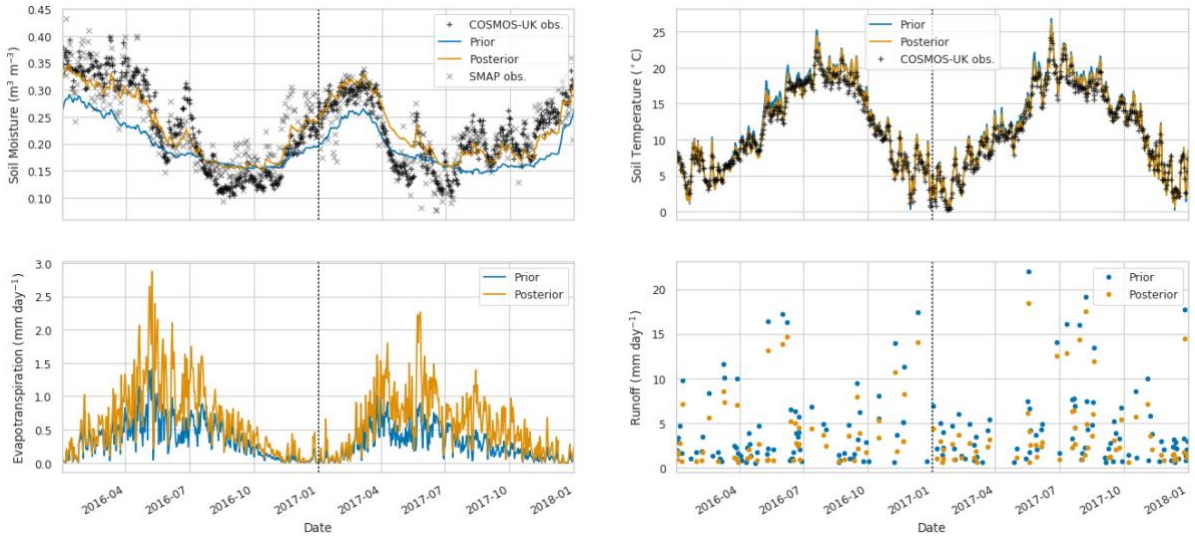
We have rerun the JULES experiments outputting additional water budget variables and have included maps of how these variables change before/after data assimilation along with relevant text at line 258:

"In Figure 6 we show the difference between mean water budget variable estimates (soil moisture, evapotranspiration and runoff) in 2016 for the prior and posterior JULES model ensemble. The grid cells that are darker blue correspond to the posterior ensemble estimate being higher after assimilation and grid cells that are darker red correspond to the posterior estimate being lower. We can see that after calibration of the pedotransfer function parameters the domain has not had a uniform increment to the value of mean soil moisture, evapotranspiration or run off. This is due to the fact that soil texture specific parameters have been optimised allowing the different distinct areas of soil type defined by the HWSD (see Figure 1) to behave differently rather than having a uniform correction across the modelled area. Across the whole domain we find an average increase of 0.03 m^3 m^{-3} in mean soil moisture estimates after data assimilation. We can see that in order to update PTF parameter values to find soil moisture estimates that more closely match the SMAP observations both evapotranspiration and run off model estimates have also been modified. In areas of sandy soils wetter soil moisture values have been achieved by a decrease in evapotranspiration offsetting a slight increase in runoff. In areas of high clay content wetter soil moisture values have been achieved by a larger decrease in run off compared to an increase in evapotranspiration. For silty soils we find a drier value of soil moisture for the posterior compared to the prior with a less prominent impact on evapotranspiration and run off. Figure 6 also allows us to see the high–resolution of the JULES model when run with the CHESS data, for this domain we have over 30,000 individual model grid cells."



For each of the COSMOS probe figures we have also included these additional water budget variable for the specific location, also including in-situ observations and model estimates to soil temperature in the top layer. We have added relevant text around this at line 314:

"In Figure 11 we show results at the Cardington COSMOS site, here we can see the posterior JULES estimate is a large improvement from the prior, although some of the driest values are still not captured. From Figure 11 we can also see that there is an increase in evapotranspiration and a decrease in runoff, this effect can also be seen from Figure 6. Figure 12 shows results for Morley COSMOS site where both prior and posterior JULES estimates perform similarly, we also have less of an update to evapotranspiration but a decrease in modelled runoff. There are also some sites where even after calibration we still do not capture the COSMOS estimates, Stoughton in Figure 13 is such an example where both prior and posterior estimates are too dry. However, here the posterior estimate is still much improved from the prior. We also find large increases in evapotranspiration and reductions in runoff for Stoughton. Figure 14 is an example where both prior and posterior perform equally poorly. The fact that the estimates and updates after DA are so different for Figures 11 - 14 despite all using the same PTF parameters highlights the effect that the underlying soil properties are having on soil hydraulic conductivity."



12. In this realm, a discussion of main characteristics, limitations and specifics of the study area with regard to SMAP data is essential to understand the manuscript. This would include addressing topography, land cover, other factors.

We have added this at line 159:

"The experiment area of the East of England is predominantly flat arable land which should allow for good quality SMAP retrievals, there are also coastal and urban areas where SMAP retrievals will be unreliable. This area is also prone to cloud cover which could cause gaps in the SMAP observational record."

13. Equation 1 – please list the units of the parameters in these physical equations.

These have been added.

14. Page 7 line 145 – why did the authors chose 10% standard deviation when it is well known that many van Genuchten parameters and soil hydraulic conductivity is logarithmic scale. What does 10% standard devation mean? Does it mean 0.63+/-0.063 for phi a and 0.0003 +/-0.00003 for phi c for example?

The reviewer is correct in their example of a 10% standard deviation, this is used to define a Gaussian distribution that 50 unique parameter sets are sampled from. It is true that van Genuchten and soil hydraulic conductivity parameters can be described by logarithmic distributions, but it is less clear what the best distributions are for the PTF parameters that are used to calculate the van Genuchten and soil hydraulic conductivity parameters. We therefore made a naïve assumption of a 10% standard deviation for our prior distribution and did not look further at this as we achieve good results when compared to in-situ COSMOS probe data. It is an important point that this is an area that could be investigated further in future studies. We have added text to this effect at line 200:

"Each ensemble members ancillary file is created by sampling from the normal distribution defined by mean xb and variance $(0.1 \times xb)^2$, where xb = (ϕa,ϕb,...,ϕo) with ϕa,...,ϕo taking the values given in table 1, then using each unique set of sampled parameters within equations (1) applied to the HWSD maps of soil properties (see Figure 1) for the whole domain. Although van Genuchten and hydraulic conductivity parameters can be described by logarithmic distributions it is less clear what distribution is best for the PTF parameters optimized here. We therefore made the naive assumption of a normal distribution in the first instance as this gave us good results."

15. Why did the authors not use a known weighting function for JULES soil moisture to compare with cosmic ray neutron sensors. Köhli et al. 2014 https://doi.org/10.1002/2015WR017169 Baatz et al. https://doi.org/10.5194/hess-21-2509-2017 or Shuttleworth et al. 2014 doi:10.5194/hess-17-3205-2013 provide already well tested methods. How does the author's method compare with these results?

Apologies we did not make this clear. The method of Baatz et al., 2014 is used by COSMOS-UK. The JULES operator was developed as part of the Hydro-JULES project (this paper also falls under this project) by colleagues at UKCEH (Cooper et al. https://doi.org/10.5194/hess-2020-359). We have strengthened the description of this at line 167:

"There are many studies translating the cosmic–ray neutron intensity measured at COSMOS probe sites to soil moisture (Baatz et al., 2014; Bogena et al., 2015; Köhli et al., 2015). There have also been efforts to relate modelled soil moisture to cosmic–ray neutron intensity, such as the COsmic-ray Soil Moisture Interaction Code (COSMIC) (Shuttleworth et al., 2013;Rosolem et al., 2014). The COSMOS–UK network use the N_0–method described by Baatz et al. (2014) to diagnose values for the soil moisture and then the method of Köhli et al. (2015) to calculate the representative depth for each COSMOS probe measurement. To make a fair comparison between the COSMOS–UK and JULES soil moisture estimates we have constructed a simple variable depth algorithm for JULES which takes a weighted average of the different soil layers of the model given the relative depth of the COSMOS–UK observation. This is defined as […] where θ_D is the JULES modelled soil moisture at the COSMOS-UK representative depth (D) and θ_{10}, θ_{25} and θ_{65} are the top, second and third layer soil moisture estimates from the JULES model."

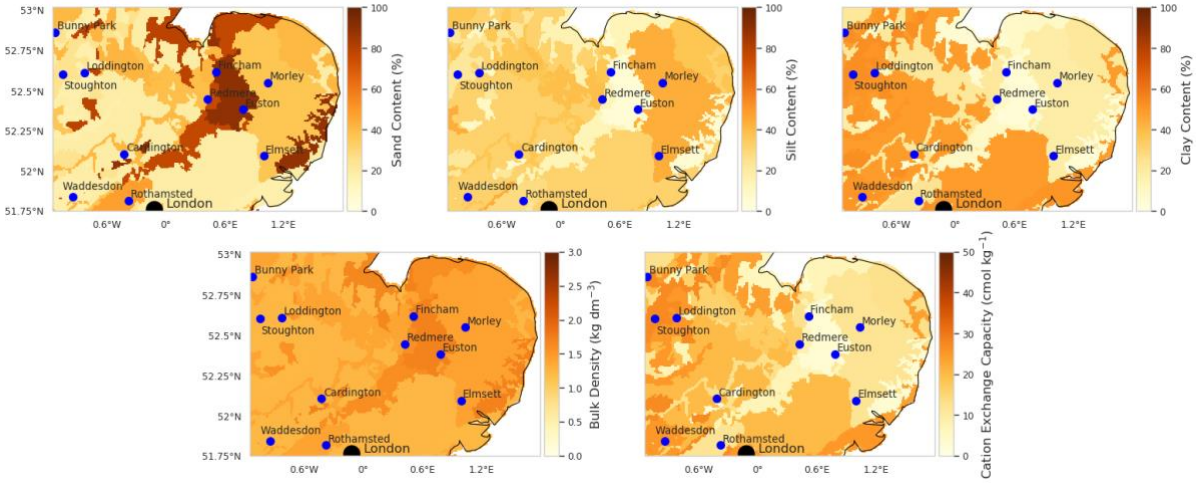
16. Aside, Desilets and Zreda, 2013 doi:10.1002/wrcr.20187, 2013 consider the diameter being 600 meter, not the radius.

Updated.

17. Figure 2: please add a map of soil textures. Please discuss the sharp light blue – dark red gradient at 0.9E. Is this an artifact from data assimilation?

The adding of soil texture maps is a useful suggestion and will help with interpretation of the results (we include these below also). We can see that the dark red gradient at 0.9E in Figure 2 is a result of a

distinct area of soil texture in the HWSD and how this is responding to the pedotransfer functions. Extra text has been added to the manuscript (see response to point 11).



18. Page 9 line 196 – adding London in all maps for the non UK citizens would be a great asset.
Noted, see above.

19. Page 10 line 206 – please define observation operator and outline the details on how this operator was developed, calibrated and validated. There are existing operators already (see point 15).
We agree we did not provide sufficient detail in the initial paper here, please see response to point 15 where we have now included these details.

20. Page 13: Please separate discussion and outlook clearly. The authors use repeatedly phrases on future work e.g. 'work is being undertaken' (line 238), 'we will' (line 241), 'is possible' (line 244) 'could be' (line 245) 'it may' (line 247) and so on. . . Also references to e.g. GRACE are missing.
We have restructured the discussion and moved such statements into the final 2 paragraphs.

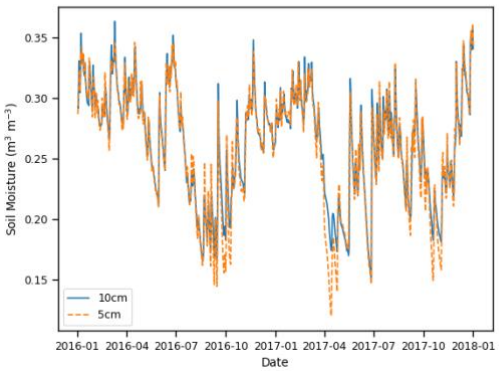
21. Also, a discussion on literature with previous published assimilation experiments on soil hydraulic parameters will be useful. Here, the paper can give a valuable contribution to existing literature. Especially considering the authors going the extra step to assimilate often cross correlated parameters of pedotransfer functions.
We have included more literature within the discussion and expanded comment on how optimising the PTF parameters differs from previous studies focusing on soil model parameters. Please see line 343 onwards.

22. Figure 11: Symbols with a center point are more precise and clearer than circles. Please use smaller dots, or even better symbols with a center point such as +,*, x and use different symbols for Cosmic Ray Calibration data and SMAP data points. Also please add SMAP soil moisture to the plots with cosmic ray neutron probe data, although these are not the equivalent depths as cosmic ray neutron probe soil moisture.
Plots have been updated in accordance with the Reviewers comments.

Reviewer 2 (R#2)

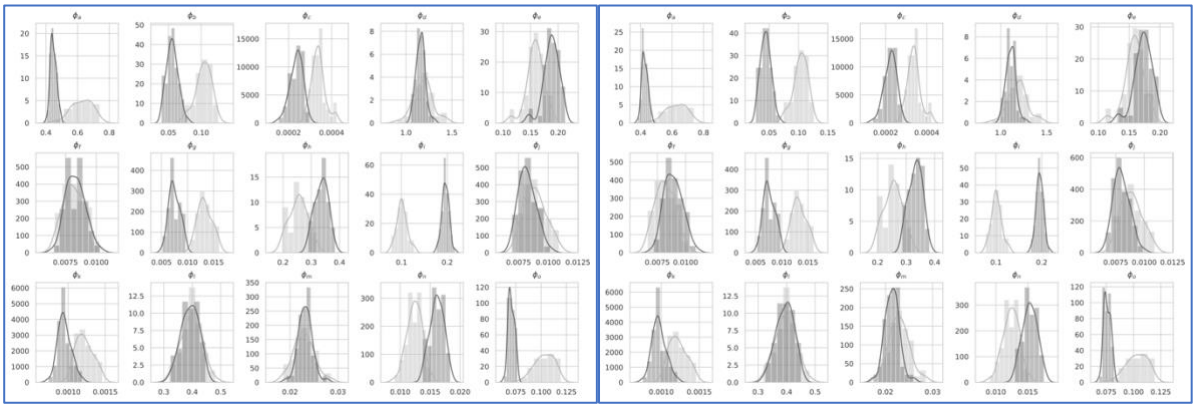
Accurate soil moisture simulation has always been a tough issue due to various sources of errors, including biased forcing, unrealistic model parameters, defect model structure and/or parameterizations. Focusing on uncertainties in pedotransfer functions, this study calibrates some of the key pedotransfer parameters through the assimilation of SMAP soil moisture product, and have obtained lower RMSD and higher correlation coefficients in posteriors. Independent evaluation against COSMOS observations also suggests promising results.
In general, this work presents a good example of utilizing satellite data to improve land surface models. The current layout and interpretation within the manuscript are mostly valid to me, except some remained concerns on the detailed DA implementations and soil moisture evaluations, as depicted below.
We thank the reviewer for their comments which will help us to improve this manuscript. We outline our responses and the changes we have made below.

1. My biggest concern is on the comparison of modelled soil moisture from a relatively 'thick' layer of 0-0.1 m with SMAP retrievals, which in most conditions corresponds to only a few centimeters of the topsoil (_2.5 cm, according to Zheng et al. 2019). Under some circumstances, soil moisture may vary a lot with depth. Is soil moisture mostly consistent and exhibits less vertical gradient within the 0-0.1m layer across the study domain? Otherwise the evaluation and the subsequent conclusions presented in this study maybe questioned. Please elaborate. Reference: Zheng, D., Li, X., Wang, X.,Wang, Z., Wen, J., van der Velde, R., Schwank, M., & Su, Z. (2019). Sampling depth of L-band radiometer measurements of soil moisture and freeze-thaw dynamics on the Tibetan Plateau. Remote Sensing of Environment, 226, 16-25
We agree the comparison between SMAP and the model top 10cm could present issues due to the representative depths. However, as stated in your comment the model soil moisture variability does not change a great deal across the top 10cm as shown in the below plot where we have run JULES with a 5cm soil depth. We made the choice to use 10cm as this is the default JULES top layer soil depth and we wanted the optimized soil parameter ancillary files to be useful to the wider JULES community. To ensure the effects of this choice were minimal on the results we have re-run the experiments using a 5cm top layer in JULES. We attach plots for the retrieved parameters in both cases and can see that the optimised distributions are very similar whether a 10cm or 5cm top layer is used. We have added text on this at line 104:
"The soil scheme is made up of 4 separate layers with depths of 0.1 m, 0.25 m, 0.65 m and 2 m respectively. We have chosen to keep JULES in its default soil-layer setup so that our optimised parameters are relevant to the wider JULES modelling community. This is despite the fact that SMAP satellite observations are typically sensitive to the top ~5cm of soil (Entekhabiet al., 2010), with some studies suggesting L-band radiometer measurements may only be sensitive to the top ~2.5cm (Zhenget al., 2019). This could introduce an additional source of error into our DA system. To ensure that the effect of this is not too great we show that there is only small difference in soil moisture between depths of 10cm and 5 cm in the JULES model in Figure S1. We have also re-run the entire data assimilation experiment with a 5 cm top soil layer in JULES and show that the recovered parameter distributions are similar to those recovered with a 10 cm top soil layer in Figure S2."



5cm experiment

10cm experiment



2. Looks typo in the third equation of Eq(1): should â´LEˇ _e f_clay be â´LEˇ _f f_clay ?
Corrected.

3. For the pedotransfer parameters shown in Table 1, are they independently calibrated grid by grid, or they share the same values across the whole domain?
These share the same values across the whole domain. We have clarifed this within the text at line 137, we have also strengthen description around the data assimilation technique (see Reviewer #1 point 4).
“The DA system used here optimises values for the parameters in table 1 across the whole domain rather than on a grid-by-grid basis. In this way the varied soil properties across the domain give us a form of orthogonal constraint within the assimilation and allow us to recover a single set of pedotransfer functions that are valid in space and time.”

4. L138-140: it is interesting to know to which depth each COSMOS monitors soil wetness. Together with results shown in section 3, it can help understand to what extend the innovation introduced into the surface layer can propagate into deep soils. That being said, I also expect the authors to spend a short paragraph to discuss this issue.
We have added this information at line 172:
“The COSMOS sites in our experiment domain have a representative depth of between 14 cm and 40 cm dependent on conditions when measurements are made.”
We have also added discussion on this in the results section at line 329:
“The COSMOS-UK observations we have used for independent validation of the results are representative of depths from 14 cm up to around 40 cm. The SMAP satellite observations, used within the assimilation algorithm to find a new set of pedotransfer functions for the experiment domain, are representative of soil moisture for the top 2.5 - 5 cm of soil. Therefore the fact that after assimilation we find such a distinct improvement at in-situ COSMOS probe locations indicates that

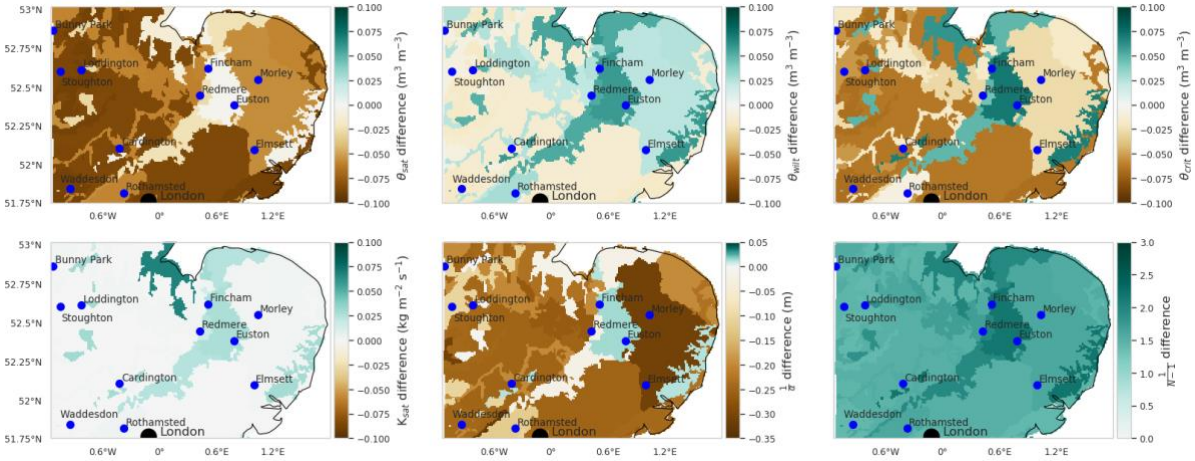
although the SMAP observations are only sensitive to shallow depths, by combining these with the JULES model we are also improving estimates at deeper levels."

5. L149-150: how is the observation operator like? Do you simply spatially average estimates from all the 1 km grids, and how do you project increments from the 9 km grid back to the 1 km grids? Please clarify. In addition, which variables are exactly included in the joint state-parameters?
Yes, we spatially average the 1km model estimates to the 9km SMAP grid. The variables included in the joint state-parameter vector are just the 15 pedotransfer function (PTF) parameters, with 50 realisations of these making up the ensemble. Each realisation will also uniquely define a model trajectory of soil moisture. Unlike sequential DA techniques we solve the problem for all observations over the whole domain at once by minimising a cost function. For this method it is not necessary to project any increments back to the 1km grid as the increments we find correspond to which parameter sets allow us to best fit the data given all relevant uncertainties. We have clarified this in the text at line 208:
"In order to compare the 1 km estimates of soil moisture from JULES to the 9 km SMAP estimates we create an observation operator which aggregates the JULES grid cells within each SMAP pixel by taking a spatial average of all JULES estimates which fall in the bounds of the SMAP grid cell. There is no need to project increments from the spatially averaged 9 km model estimates back to the 1 km model grid as the assimilation is only optimising the 15 PTF parameters (ϕa, ϕb, ..., ϕo) for the whole domain and the update to soil moisture will be implicit."

6. L153: ": : :by a factor a four: : :"–not sure how this is done, may need to provide more details on the implementation of inflation.
This is also noted by Reviewer 1 (comment number 8). We have included additional text on this at line 214:
"It has been shown that, for variational methods such as the one used in this paper, these additional sources of error (model error, representativity error, etc.) can be included in the observational term of the cost function by inflating the diagonal observation error variance (Howes et al., 2017). Although observation error inflation is rare in relation to sequential DA methods it is commonly used with variational methods and especially in numerical weather prediction (Hilton et al., 2009; Bormann et al., 2015; Minamide and Zhang, 2017; Fowler et al., 2018; Wang et al., 2019). Observation error inflation is required due to the fact that all observations are used at once in the assimilation whereby we minimise a cost function containing a prior term and an observational term. The greater the number of observations in the observational cost function term, the higher the weight they have in the optimization. This can lead to the prior term being completely negated and hence the retrieval of nonphysical parameters. Observation error inflation would not be required if the correct specification for the observation error correlations (in space and time), model error and representativity error were included. These, however, are hard to diagnose and it has been shown that in the absence of such information observation error inflation is required for an optimal DA system (Stewart et al., 2014). For this reason and due to the large number of observations assimilated in our one year assimilation window (28698) we inflate the specified observational error by a factor of four. If a filtering DA system were being used utilising a bias aware DA system such as that presented by Ridler et al. (2017) could help represent some of the additional sources of error discussed here."

7. Fig. 3: if possible, better to show prior and posterior distributions of some of the soil hydraulic parameters (e.g. _sat,Ksat) in Eq(1) as well, as they directly regulate soil water within the land model.
It will be difficult to show probability distributions of the hydraulic parameters as they vary across the domain dependent on the underlaying soil texture map. Instead we have included maps of the resultant soil hydraulic parameters and how these have changed after DA (see below). Text also added at line 251:

"This can be seen in Figure 5 where we show the updated PTF parameters effect on the mean estimate to the JULES model soil parameters when applied to the spatial maps of soil properties from the HWSD. We can see how different areas of distinct soil texture (see Figure 1) behave differently based on the PTF parameter updates after DA. For some parameters we see the majority of gridcell parameter values increase or decrease, θsat and 1/(N−1) respectively. Whereas for 1/α and θcrit we see an increase or decrease in grid cell parameter values dependent on the underlying soil properties (sandier soils lead to an increase, less sand more clay correspond to a decrease)."



8. L192: urban areas are known to have problems in both remote sensing and land surface modeled soil moisture. I would suggest excluding urban areas in all the plots in Figs.(2, 4-5). Meanwhile, the authors may want to show some of the COSMOS sites in these plots to help better interpret results in Figs. 8-11.

We agree including the COSMOS stations on the plots may help interpretation and will do so (see above). We have left urban areas just as a point of discussion on current limitations.

Reviewer 3 (R#3)

The paper explores the use of SMAP soil moisture products with the JULES land surface model with a data assimilation framework. The framework is applied in a region of the UK where soil properties from pedotransfer functions are constrained with data assimilation. The topic has potential and the paper started very well with its Introduction and Methods sections. However, I found the Results and Discussion very weakly presented, without in-depth analyses and implications. It is not clear what is the lessons learned and how it can benefit the wider community. In addition, these two sections read much more like a technical report. There are many additional tests that can be made to improve this study (I've made some suggestions). For that reason, I believe this paper manuscript requires considerable changes, hence I recommend major revisions before making my decision on its acceptance.

We thank the reviewer for their comments and have aimed to strengthen the paper in line with their specifications. Below we present our responses to comments and outline the changes we have made.

List of comments:

L63-64: Notice there are several approaches that constrain model parameters that do account for uncertainties, please refer to works by Keith Beven, Jim Freer, Jasper Vrugt, Grey Nearing, Hamid Moradkhani, Martyn Clark; to name a few.

We agree that it is beneficial to mention such studies and have included the following at line 63:
"Unlike traditional calibration procedures data assimilation and other associated Bayesian optimisation methods always take into account the relative uncertainties given to both model and observed estimates to find a maximum-a-posteriori estimate (Beven and Binley, 1992; Thiemann et al., 2001; Vrugt et al., Moradkhani et al., 2005; Nearing et al., 2010; Mizukami et al., 2017)."

L64-65: First, can the authors please point out the references for the 'Previous studies' mentioned in the sentence?

Apologies, these were missed here. We have added these on line 66:
"Previous studies have used data assimilation to update the soil parameters of land surface models (Rasmy et al., 2011; Sawada and Koike, 2014; Yang et al., 2016; Han et al.,2014)"

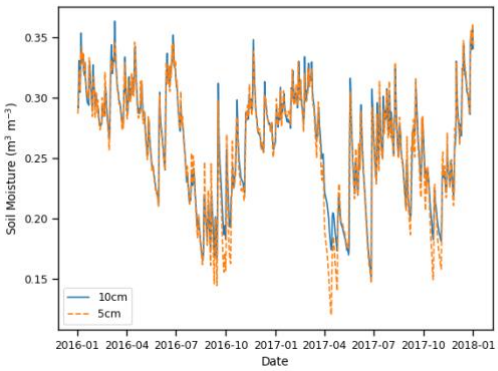
L65-66: Note that usually, the term data assimilation has been used in different ways by the atmospheric sciences and land surface modeling community in relation to the hydrological modeling community. 'Data assimilation' in general refers to using/fusing observed quantities to better constrain model components (i.e., parameter, states, etc...). Typically, the use of 'parameter estimation', 'state estimation', or 'dual parameter-state estimation' would be more clear. The reason I am mentioning this is because, although not technically a classic data assimilation application, the group by Luis Samaniego in UFZ Germany has explored similar approaches to this one using their mHM with their MPR framework. Additional work 'assimilating' both state and parameters include groups from Harrie-Jan Hendricks-Franssen, for example.

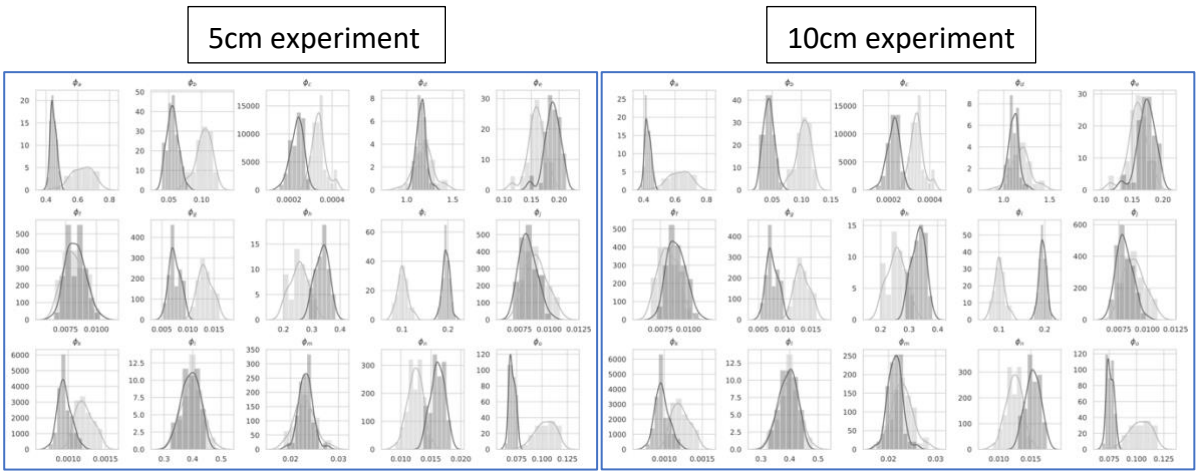
We agree. We are coming at this problem from a different background and so accept it may be helpful to update the wording here to make things more clear for the reader. We have updated the wording to include "state estimation" and "parameter estimation" line 60:
"These techniques can either be used for state estimation to update soil moisture values of the model in real-time as new observations are available […] or for model parameter estimation to find improved calibrations which better represent the observations […]."
We have also discussed work by Luis Samaniego in the discussion section at line 370.

L79-81: This seems to be related to Results, not sure why it is included at the end of the Introduction section.

We have removed the relevant text.

L94-95: The direct information obtained from SMAP is typically for the first few centimeters of soil; yet your JULES model is configured with a relatively thick initial soil layer and only 4 layers in general. Have the authors considered revising their soil layers in JULES? Have they done any simple sensitivity study to check how influential the choice of soil layer discretization is when assimilating SMAP data. If I recall correctly, CLM (which is similar to JULES) is run with a much finer soil layer discretization.
Reviewer #2 had a similar concern (see R#2 comment number 1). We chose to keep JULES in its default soil layer configuration so that the optimized soil ancillaries would be useful to the wider JULES community. We have run JULES with a 5cm top layer to confirm that the vertical variability at this depth in the model is not too great (see below). We have also re-run the entire experiment using this soil layer to confirm that we retrieve very similar parameter distributions from the DA procedure for both the 5cm and 10cm soil depth JULES models. Relevant text has been added at line 104: "The soil scheme is made up of 4 separate layers with depths of 0.1 m, 0.25 m, 0.65 m and 2 m respectively. We have chosen to keep JULES in its default soil-layer setup so that our optimised parameters are relevant to the wider JULES modelling community. This is despite the fact that SMAP satellite observations are typically sensitive to the top ~5cm of soil (Entekhabiet al., 2010), with some studies suggesting L-band radiometer measurements may only be sensitive to the top ~2.5cm (Zhenget al., 2019). This could introduce an additional source of error into our DA system. To ensure that the effect of this is not too great we show that there is only small difference in soil moisture between depths of 10 cm and 5 cm in the JULES model in Figure S1. We have also re-run the entire data assimilation experiment with a 5 cm top soil layer in JULES and show that the recovered parameter distributions are similar to those recovered with a 10 cm top soil layer in Figure S2."





L113-115 and Table 1: It is unclear to me how the prior is used. Don't you need an emsemble (i.e., range) for each prior factor shown in this Table? How is a single prior applied in this case?
For our prior we have 50 realisations of the parameters in Table 1. Each realisation is drawn from a normal distribution defined by the mean (shown in Table 1) and standard deviation (taken as 10% of

the mean). This gives us the light grey distributions shown in the plots above. We have expanded the text to this effect at line 200:
"Each ensemble members ancillary file is created by sampling from the normal distribution defined by mean xb and variance $(0.1 \times xb)^2$, where $xb = (\phi a, \phi b, \ldots, \phi o)$ with $\phi a, \ldots, \phi o$ taking the values given in table 1, then using each unique set of sampled parameters within equations (1) applied to the HWSD maps of soil properties (see Figure 1) for the whole domain. Although van Genuchten and hydraulic conductivity parameters can be described by logarithmic distributions it is less clear what distribution is best for the PTF parameters optimized here. We therefore made the naive assumption of a normal distribution in the first instance as this gave us good results."

L138-140: Can the authors be more specific about this? There are many studies that have used the COSMIC operator which is available (refer to works by Jim Shuttleworth, Rafael Rosolem, Harrie-Jan Hendricks-Franssen, as examples). Have the authors consider implementing this operator?
Section 2.6: Needs to be expanded as it is very vague and general.
Apologies, we should have added more here and also referenced the works noted above. We have included this additional information at line 167:
"There are many studies translating the cosmic–ray neutron intensity measured at COSMOS probe sites to soil moisture (Baatz et al., 2014; Bogena et al., 2015; Köhli et al., 2015). There have also been efforts to relate modelled soil moisture to cosmic–ray neutron intensity, such as the COsmic-ray Soil Moisture Interaction Code (COSMIC) (Shuttleworth et al., 2013;Rosolem et al., 2014). The COSMOS–UK network use the N_0–method described by Baatz et al. (2014) to diagnose values for the soil moisture and then the method of Köhli et al. (2015) to calculate the representative depth for each COSMOS probe measurement. To make a fair comparison between the COSMOS–UK and JULES soil moisture estimates we have constructed a simple variable depth algorithm for JULES which takes a weighted average of the different soil layers of the model given the relative depth of the COSMOS–UK observation. This is defined as […] where θ_D is the JULES modelled soil moisture at the COSMOS-UK representative depth (D) and θ_{10}, θ_{25} and θ_{65} are the top, second and third layer soil moisture estimates from the JULES model."

Figure 3: Typically, DA are justified as an operational tool for models (in the case of state estimation). This figure here shows the Bayesian optimization approach (prior –> likelihood –> posterior) which is fine. However, I'd be interested to see the time-series of the final soil parameters (produced with the updated pedotransfer function) to check for any inconsistencies in the way a particular parameter change from time to time. I'd expect soil properties to be fairly constant (relatively to the fluxes and states in the JULES model). Also, the authors should consider checking which of the PDFs shown in the figure are expected to be significantly different. One way to do this is for example by checking whether two samples come (or not) from the same probability distribution. This can be easily done with a two-sample Kolmogorov-Smirnov test.
This is indeed an interesting comment and Reviewer #1 had a similar query (point #6), it is important to make the following distinction. Data assimilation can be used to determine the value of state variables and parameters. In our case we are interested in determining the value of the 15 PTF parameters which are fixed in time. We use an implementation of the Iterative Ensemble Smoother. Being a smoother, this method performs data assimilation over a time window (labelled assimilation window). The method uses all observations over the spatial domain and during the time window (here the year 2016) in a single minimisation process. This minimisation is done via an iterative routine. Hence, a single set of estimated PTF parameters are obtained. A smoother usually requires the Jacobian of the evolution model, the so-called tangent linear model (TML) and adjoint model (AM) in the 4D-variational literature. We do not have such a model for JULES. The Ensemble nature of the method allows to replace the role of the TLM/AM by operations involving 4D sample covariances, i.e. covariances defined over space and time and computed from the ensemble.

It is also important to mention that we chose the length of the assimilation window to be the whole length of the experiment. This is of course, only a choice, but we justify it given that the parameters we are looking for do not vary in time, or if they do, this is in time-scales which are considerably longer than the duration of the experiment. In numerical weather prediction, for instance, the situation is different. In such a scenario one is more interested in the actual values of the time-evolving state variables. For instance, operational centres like Met Office and ECMWF use 12-hour windows to perform data assimilation (e.g with a 4D-variational technique), and then cycle this process for subsequent windows. We have included plots of how the soil parameters change after DA when produced with the prior or posterior PTF parameters (see below). We have also included additional description of the DA method at line 181:

" In order to estimate the identified pedotransfer function parameters we use the LAVENDAR data assimilation framework (Pinnington et al., 2020). This framework utilises a hybrid DA technique similar to that of the Iterative Ensemble KalmanSmoother (IEnKS) (Bocquet and Sakov, 2013). A smoother is different than a filter (e.g. the Ensemble Kalman Filter (Evensen,2003)) in that it uses batches of observations which are taken over a time window of given length and the whole spatial domain, as opposed to just in a time instant. These observations are combined with the model evolution over this window and a minimization process is performed to obtain initial conditions for the state/parameter values. It is possible to run a sequence of smoother steps for successive windows, but our study only uses one year long assimilation window as the parameters we are optimising do not vary in time.

Using a smoother instead of a filter has advantages (Lorenc and Rawlins, 2005) in that (a) more observations can be used to constrain the problem solution, and (b) information from the model evolution is implicitly used in the search process. However, using a smoother requires computing the Jacobian of the model, the so–called tangent linear model (TLM) and the related adjoint model (AM). The TLM/AM (Courtier et al., 1994). Computing and maintaining the TLM/AM is not a trivial task, and in fact we do not have this for JULES. The IEnKS solves this problem by replacing the role of the TLM/AM by 4–dimensional covariances, i.e. covariances defined over time and space. These covariances are computed as sample estimators of a given ensemble. The iterative nature of the method means that it finds the solution to the minimization problems using inner iterations rather than a single step (hence the variational nature), and this helps when the distributions of the variables/parameters of interest are not Gaussian. We provide details of the method in Appendix A. Furthermore, to understand the variants of the ensemble Kalman Smoother and its position within the hybrid DA methods, the reader is referred to Evensen (2018).

We show a schematic of how this system works in Figure 3, […]"

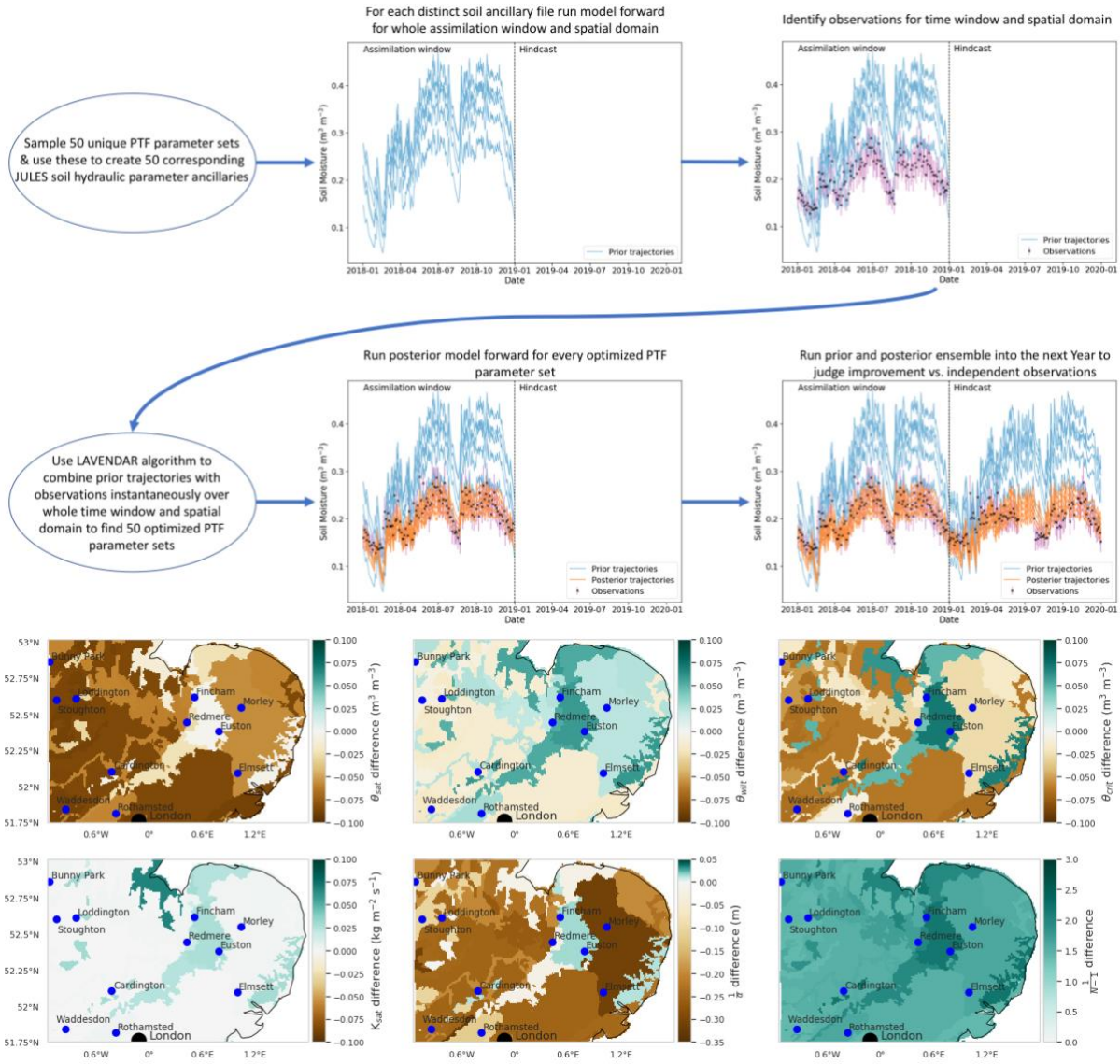


Figure 6: It is important to show how the prior and posterior spread compare with the actual RMSE calculated against the actual observation to check for consistencies with the DA setup. Without this analysis shown (for some points and maybe regionally), it is hard to diagnose the DA results. The goal is for the spread to have the same magnitude of the RMSE (not too large, nor too small)

We agree this is a useful check to make. We have added plots of the prior and posterior spread in the model predictions of soil moisture averaged in space to compare to the model RMSE averaged in space (see below, Reviewer #1 also had similar comments at point #6). We can see from this plot that for the prior we have the desired relationship with the ensemble spread being around the same magnitude as the prior RMSE. However, for the posterior we do find an ensemble spread with a slightly lower magnitude than the posterior RMSE. This is perhaps unsurprising as we are conducting just a single assimilation step but using all <28000 observations at once in space and time, so we may find that some of the posterior parameter distributions become too narrow, as with increasing observations we increase the confidence in our posterior, thus tightening the retrieved distributions. If we were to use our posterior optimized parameters in onward experiments we would require some form of ensemble inflation. We have added additional discussion on this in the results section at line 296:

"In Figure 10 we show the RMSE averaged in space for the JULES model prior and posterior mean estimate, when compared to SMAP, alongside the JULES model prior and posterior ensemble spread. At all times the posterior JULES RMSE is lower than that of the prior, showing that the DA system has

found a set of PTF parameters that improve the fit to the SMAP observations through time, this continues into the hindcast period (2017) when judged against observations that were not included in the DA cost function. We find slight peaks in the RMSE values throughout the time period corresponding to wetter conditions, this could be due to slight errors in the precipitation driving data used to force the model. It is optimal to have an ensemble spread that matches the magnitude of the ensemble mean RMSE and this relationship should hold given a large enough ensemble size (Houtekamer and Mitchell, 1998). We can see that this relationship holds for our prior estimates. However, after DA our posterior ensemble spread is slightly lower than that of the ensemble mean RMSE. This is perhaps unsurprising as we are conducting just a single assimilation step using all observations (over 28000) at once in space and time with a relatively small ensemble size (50). This can lead to some of the posterior parameter distributions becoming narrow, as with increasing observations we increase the confidence in our posterior, thus tightening the retrieved distributions and reducing the model ensemble spread. This result suggests that ensemble inflation (Anderson and Anderson, 1999) may be necessary if this ensemble was to be used in subsequent assimilation experiments."

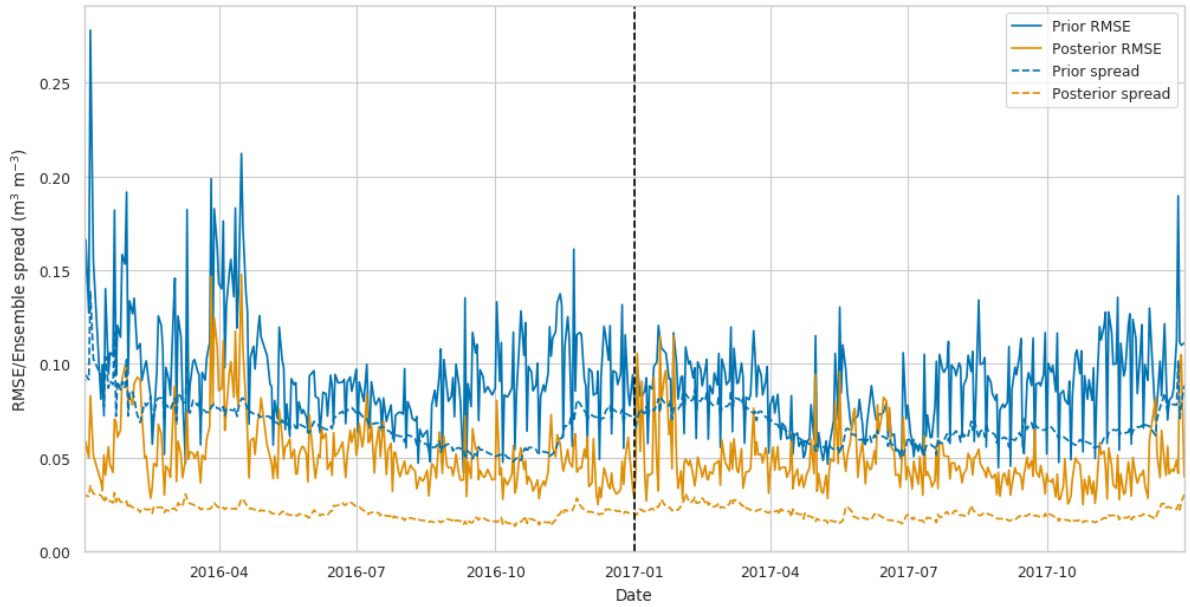


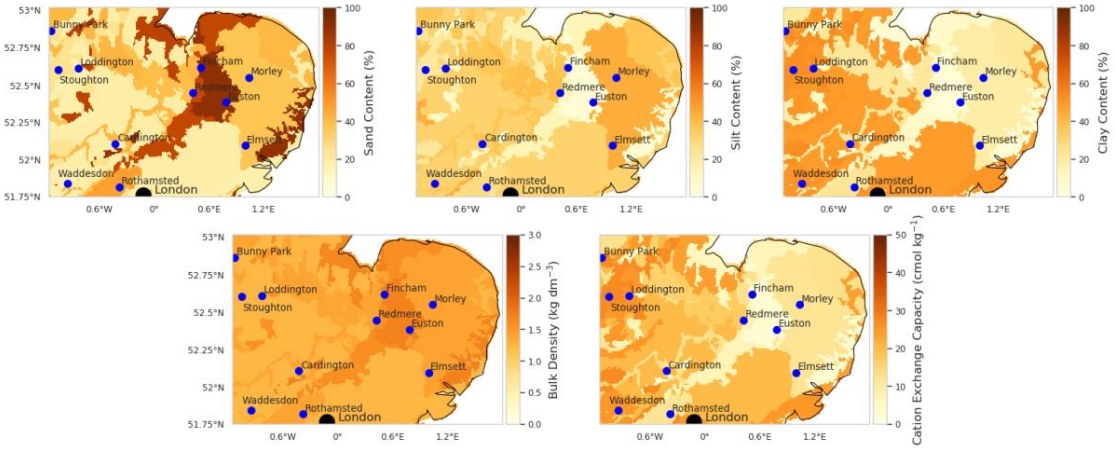
Figure 4: It is not clear to me how RMSE is calculated in percentage. Maybe I missed something. Can the authors made this clear in the captions.

For the spatial plots of error reduction we have just calculated the percentage change between the prior JULES prediction RMSE (compared to SMAP) and the posterior JULES prediction RMSE, both averaged in time. We have defined this as an equation in the manuscript at line 274.

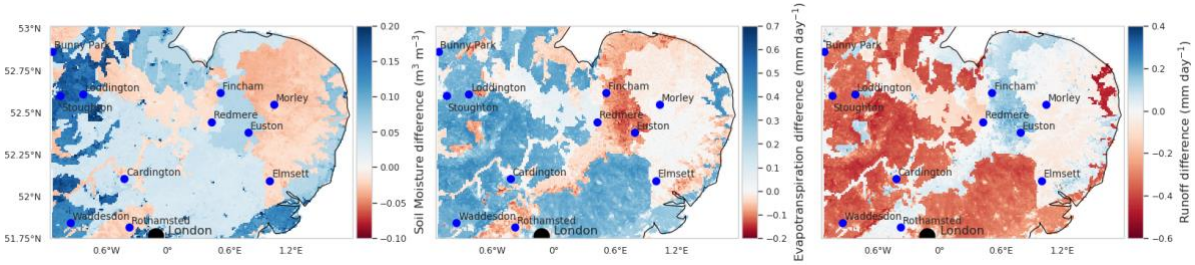
Results section: I found the results section to be presented in a very weak way. It seems to be rushed with the same regional map shown only for different metrics. The section is written almost like a technical report just going from figure to figure with very little in-depth analysis. How does the soil moisture in the region change from time to time (the metrics are only aggregated for the period)? Are the soil properties and consequently soil moisture profiles realistic? What are the impacts on other components of the model? Does 'improving' soil moisture improves other fluxes in JULES? My understanding is that COSMOS-UK also has flux data that can be used (H, LE, G???). The simple exercise of assimilating soil moisture to constrain parameters and/or states and evaluate the impact on soil moisture only does not seem to be particularly novel in my opinion (the DA frameowrk and the use of COSMOS-UK do, but should be explored further). This item is a major issue I have with the current manuscript.

We agree that the results section could be expanded to add more detail on the impact of the DA to other model components. We have had similar comments from other reviewers. We have added plots of soil texture and how the resultant soil parameters change after DA. Also we agree looking at the performance of the models through time would be useful and have included a plot of RMSE through time over the modelled domain (see plot above). Unfortunately flux observations are not routinely available from COSMOS-UK, we do have access to soil temperature observations which we have included to judge the performance of another model component. We have also included other water budget variables (ET, Run off) for the domain and at each of the current 4 COSMOS sites shown so that we can judge the impact of the DA on these variables (please see below for the included plots). We believe that the ability to calibrate pedotransfer functions at a large scale using a considerable amount of satellite data (<28000 observations) in an innovative data assimilation system does present novelty. Especially when it is shown that from this very large scale we are able to improve independent in-situ estimates from the model. However, we do agree that including extra variables will strengthen the paper and have included those stated above along with a strengthening of the analysis text in the Results section (see line 240 for start of results section).

Soil properties:



Impact on water budget variables:



COSMOS-UK site example (Cardington):

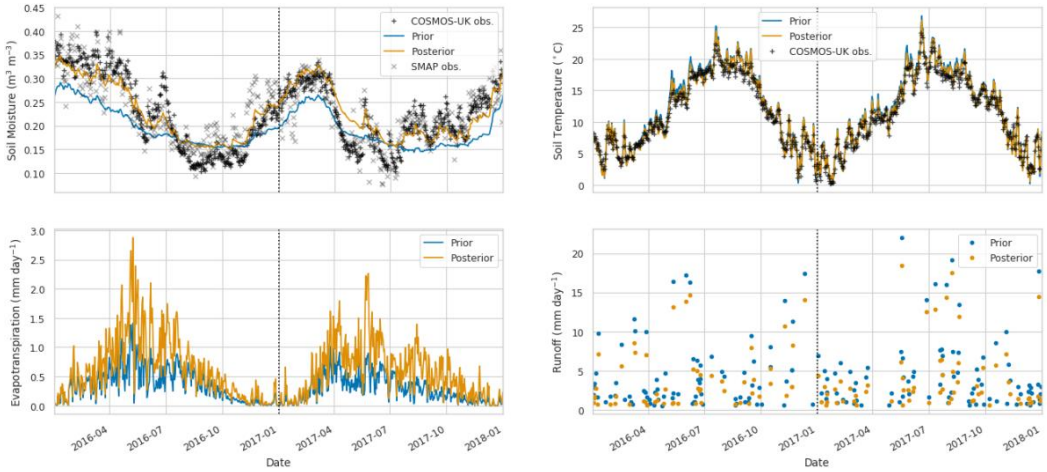


Figure 10: There seems to be some systematic biases in the model that suggests non-optimal DA setup (DA requires errors to be around a zero mean). How much that impacts the results? Are there other sites with similar issues (can you expand the discussion)? Have you tried some initial pre-calibration prior to running the DA to reduce/remove the biases?

We have added text addressing this at line 334:

"The large errors in our prior JULES estimates for the COSMOS sites in Figure 13 and 14 could point towards some systematic bias within the model. However, it is important to note that the COSMOS-UK observations are independent of the data assimilation. For the assimilated SMAP observations it may be optimal to have errors centred around zero but for the independent in-situ validation data there will be many competing errors that may make this impossible. There will be errors in the forcing meteorology (here we are using CHESS 1 km forcing data and not observed in-situ meteorology), errors in the model grid and its representativity to the in-situ location, structural model errors (we currently have no ground water model in JULES and some in-situ sites may be more ground water dominated), errors in the vegetation fractions, and many more. At the larger SMAP scale many of these effects will be minimised when looking at the 9 km spatial scale that is more representative of modelled estimates."

Discussion section: I also found the discussion a bit weak. Very little is further discussed and explored. Sometimes the discussion is mainly focused on aspects that can be done in the future. I'd suggest the authors to define 2-5 clear objectives –> questions –> hypotheses that can be presented in more detail in the Results section, and discussed more in-depth in this current section.

We agree the discussion section could be improved and had similar comments from R#1. We have expanded the discussion section and included more literature (see line 343 onwards for discussion section).

At the end of the introduction we have also included the 2 main objectives of this study, line 86:

"The objectives of our study are as follow:

- –To examine the ability of 9 km Level-3 SMAP data to update pedotransfer parameters in a 1 km land surface model.
- –To assess the resulting prediction of modelled soil moisture against (a) SMAP data from a different time period and (b) independent in-situ data from the COSMOS-UK network.

Improving Soil Moisture Prediction of a High–Resolution Land Surface Model by Parameterising Pedotransfer Functions through Assimilation of SMAP Satellite Data

Ewan Pinnington[1], Javier Amezcua[1], Elizabeth Cooper[2], Simon Dadson[2,3], Rich Ellis[2], Jian Peng[3], Emma Robinson[2], and Tristan Quaife[1]

[1]National Center for Earth Observation, Department of Meteorology, University of Reading, Reading, UK
[2]UK Centre for Ecology and Hydrology, Wallingford, UK
[3]University of Oxford, Oxford, UK

Correspondence: Ewan Pinnington (e.pinnington@reading.ac.uk)

Abstract. Pedotransfer functions are used to relate gridded databases of soil texture information to the soil hydraulic and thermal parameters of land surface models. The parameters within these pedotransfer functions are uncertain and calibrated through analyses of point soil samples. How these calibrations relate to the soil parameters at the spatial scale of modern land surface models is unclear, because gridded databases of soil texture represent an area average. We present a novel approach for calibrating such pedotransfer functions to improve land surface model soil moisture prediction by using observations from the Soil Moisture Active Passive (SMAP) satellite mission within a data assimilation framework. Unlike traditional calibration procedures data assimilation always takes into account the relative uncertainties given to both model and observed estimates to find a maximum likelihood estimate. After performing the calibration procedure we find improved estimates of soil moisture for the JULES land surface model (run at a 1 km resolution) when compared to estimates from a cosmic-ray soil moisture monitoring network (COSMOS–UK). The spatial resolution of these COSMOS probes is much more representative of the 1 km model grid than traditional point based soil moisture sensors. For 11 cosmic–ray neutron soil moisture probes located across the modelled domain we find an average 22% reduction in root-mean squared error, a 16% reduction in unbiased root-mean squared error and a 16% increase in correlation after using data assimilation techniques to retrieve new pedotransfer function parameters.

1 Introduction

Land surface models are important tools for translating meteorological forecasts and reanalyses into real-world impacts by providing schemes for how energy, water and other matter will interact with the Earth's surface, outputting relevant diagnostics and variables and understanding the role of variability in the terrestrial hydrological cycle in the Earth system. As the spatial resolution of available meteorological information has become increasingly fine (Clark et al., 2016) it is necessary to ensure

land surface models can utilise this information at its native resolution in order to provide outputs that are as accurate as possible for local populations. In this paper our focus is on soil moisture which plays an essential role in agriculture (Asfaw et al., 2018), weather and climate prediction (Hauser et al., 2017) and land surface energy partitioning (Beljaars et al., 1996; Bateni and Entekhabi, 2012). The modelling of soil moisture is highly sensitive to driving precipitation and model parameterisations (Pitman et al., 1999). Typically, models of soil moisture will determine parameters based on spatial datasets of soil texture information using pedotransfer functions such as those defined by Cosby et al. (1984) for the Brooks and Corey (1964) soil model. The majority of pedotransfer relationships are calibrated for point samples of soil for a specific geographic location (Cosby et al., 1984; Wösten et al., 1999; Schaap et al., 2004; Tóth et al., 2015). Selecting the appropriate set of pedotransfer functions for the modelled area will allow for more representative results. It is unclear how these calibrations of pedotransfer functions and their resulting soil model parameters relate to the varying spatial scales of modern land surface models, and indeed the use of additional streams of information from remote sensing and in-situ observations is seen as increasingly important for calibration and validation (Van Looy et al., 2017). Pedotransfer functions can be continuous or discrete (setting predefined model parameters for different ranges of soil texture). Discrete examples of pedotransfer functions can be found in Wösten et al. (1999) for the van Genuchten (1980) soil model. Continuous versions of these functions may be preferential as they provide greater heterogeneity for resulting soil model parameter maps which may be more realistic. Tóth et al. (2015) provide more recent examples of continuous pedotransfer functions for the van Genuchten (1980) model. For this paper continuous functions will also allow us to seek updated parameter values that improve the prediction of a land surface model at a given spatial scale and properly account for uncertainty in both the soils information and resulting model predictions.

There now exists a large amount of information from different satellite missions relating to the spatial and temporal variability of soil moisture. These can be based on either active (*e.g.* The Advanced Scatterometer (ASCAT) (Wagner et al., 2013)) or passive (*e.g.* The Soil Moisture Ocean Salinity (SMOS) mission (Kerr et al., 2001)) observing instruments with good results found when combining both (*e.g.* the Soil Moisture Active Passive (SMAP) mission (Entekhabi et al., 2010)). The NASA SMAP mission was originally designed with both an active and passive sensor on board, soon after launch in January 2015 the active sensor malfunctioned. Sentinel 1 is now used as the active component in the SMAP soil moisture retrieval. Recent validation studies have shown SMAP to perform well in comparison with other satellite estimates (Montzka et al., 2017; Chen et al., 2018; Peng et al., 2020). These remotely sensed products are available at scales comparable to current land surface models from 50 km down to 9 km. Traditional in-situ observations of soil moisture are made at a single point using a variety of different methods (Walker et al., 2004). These in-situ measurements provide accurate estimates to the true state of the amount of water contained within the soil. However, the scale of such measurements can be unrepresentative to the scales of the model, even when land surface models are run at a high resolution ($\sim$1 km). The recent developments of cosmic–ray neutron sensing soil moisture probes (Zreda et al., 2008) somewhat alleviates this issue. Cosmic–ray neutron probe observations have a variable spatial footprint dependent on atmospheric air density (130 m - 240 m (Köhli et al., 2015) with some studies quoting a diameter of $\sim$600 m Desilets and Zreda (2013)) that is much more representative of land surface model estimates than that of traditional soil moisture probes. There are now good networks of cosmic–ray probes within several countries (Zreda et al., 2012). This is true in the UK where the COsmic–ray Soil Moisture Observing System United Kingdom (COSMOS-UK) network (Evans

et al., 2016) has been established by The UK Center for Ecology and Hydrology (UKCEH) and returning observations since 2013 (Stanley et al., 2019). These observations can act as valuable validation data of both satellite and land surface model soil moisture estimates (Duygu and Akyürek, 2019).

Data assimilation provides methods for combining new observations with land surface models in order to improve predictions. These techniques can either be used for state-estimation to update soil moisture values of the model in real-time as new observations are available (Liu et al., 2011; Draper et al., 2012; De Lannoy and Reichle, 2016; Kolassa et al., 2017) or for model parameter-estimation to find improved calibrations which better represent the observations (Rasmy et al., 2011; Sawada and Koike, 2014; Yang et al., 2016; Pinnington et al., 2018). Unlike traditional calibration procedures data assimilation and other associated Bayesian optimisation methods always take into account the relative uncertainties given to both model and observed estimates to find a maximum-a-posteriori estimate (Beven and Binley, 1992; Thiemann et al., 2001; Vrugt et al., 2003; Moradkhani et al., 2005; Nearing et al., 2010; Mizukami et al., 2017) Previous studies have used data assimilation to update the soil parameters of land surface models (Rasmy et al., 2011; Sawada and Koike, 2014; Yang et al., 2016; Han et al., 2014) . However, we are unaware of any studies using data assimilation to update the parameters of pedotransfer functions to improve land surface model predictions. Updating the parameters of these pedotransfer functions by combining them with observations from satellites addresses a key uncertainty within their calibration with respect to land surface models, adding additional information about spatial heterogeneity and the larger scales of both satellite and land surface model estimates. Many previous studies optimising model soil parameters have taken a filtering DA approach (Moradkhani et al., 2005; Montzka et al., 2011; Han et al., 2014; Baatz et al., 2017; Botto et al., 2018) leading to the recovery of a time-series of parameter values as additional data is assimilated through time. In this study we use a smoother method, i.e. one that uses all observations in the spatial domain within a time window of a given length. Then, the static parameters are obtained by a single minimization process (which can contain iterative steps). Smoothers can be used in a sequence of 'analysis windows' (as it is done in operational numerical weather prediction), but in this study we only use one of these windows since the parameters we search for do not vary in time.

We have used the Land Variational Ensemble Data Assimilation Framework (LAVENDAR) (Pinnington et al., 2020) to combine soil moisture estimates from the NASA SMAP mission with the Joint UK Land Environment Simulator (JULES) model run at a high-resolution (1 km) and update the parameters of the Tóth et al. (2015) pedotransfer functions for the van Genuchten (1980) soil model. In our experiments we assimilated 2016 SMAP data and then ran a hindcast for the year 2017. The experiments were conducted over a sub-domain of the UK due to considerations of computational expense. We selected the region of East Anglia due to it being equally susceptible to flooding and drought and therefore displaying a good dynamic range of soil moisture values. This region also had a good availability of high quality SMAP data (here we use Level-3 SMAP soil moisture observations) and a high distribution of COSMOS probes to allow for thorough validation of any results. While reducing the spatial domain in our experiments eased the computational load we were still modelling over 30,000 grid points due to the high-resolution of the JULES model. The objectives of our study are as follow:

- To examine the ability of 9 km SMAP data to update pedotransfer parameters in a 1 km land surface model.

– To asses the resulting prediction of modelled soil moisture against (a) SMAP data from a different time period and (b) independent in-situ data from the COSMOS-UK network.

2 Method

2.1 JULES land surface model

The Joint UK Land Environment Simulator (JULES) is a community developed process based land surface model and forms the land surface component in the next generation UK Earth System Model (UKESM). A description of the energy and water fluxes is given in Best et al. (2011), with carbon fluxes and vegetation dynamics described in Clark et al. (2011). We drive the JULES model with the Climate Hydrology and Ecology research Support System meteorology (CHESS) dataset (Robinson et al., 2017) which is a 1 km daily dataset of meteorological variables, an example implementation of JULES with the CHESS-met dataset can be found in Martínez-de la Torre et al. (2019). In our experiments we have used JULES version 5.3, the code and model settings are available through the MetOffice JULES repository (https://code.metoffice.gov.uk/trac/jules), with Rose suite number u-bq357. This model setup is based on the Rose suite u-au394 used to create the CHESS-land dataset (Martinez-de la Torre et al., 2018). The JULES model utilises the Harmonized World Soil Database (HWSD) (Fischer et al., 2008) as the underlying soil texture map for the creation of its soil parameter ancillaries using a pedotransfer function, see Figure 1. The HWSD has been gap-filled in urban areas where no information is available as we ran JULES without urban tiles switched on. The soil scheme is made up of 4 separate layers with depths of 0.1 m, 0.25 m, 0.65 m and 2 m respectively. We have chosen to keep JULES in its default soil-layer setup so that our optimised parameters are relevant to the wider JULES modelling community. This is despite the fact that SMAP satellite observations are typically sensitive to the top $\sim$ 5 cm of soil (Entekhabi et al., 2010) , with some studies suggsting L-band radiometer measurements may only be sensitive to the top $\sim$ 2.5cm (Zheng et al., 2019) . This could introduce an additional source of error into our DA system. To ensure that the effect of this is not too great we show that there is only a small difference in soil moisture between depths of 10cm and 5 cm in the JULES model in Figure S1. We have also re-run the entire data assimilation experiment with a 5 cm top soil layer in JULES and show that the recovered parameter distributions are similar to those recovered with a 10 cm top soil layer in Figure S2 and S3. It is necessary to find an appropriate initial state before running a land surface model such as JULES and it has been shown that without a suitable spin-up period forecast skill can be impacted (Maurer and Lettenmaier, 2004) . We include a 4 year spin up period at the start of each JULES run to allow the soil moisture state to reach a point of equilibrium after parameter values are changed. For the JULES spin-up the model is run from an initial value (defined by the saturated soil moisture model parameter) over the same year of forcing data, here 2015, to reach an equilibrium soil moisture state for any given set of soil hydraulic parameters. We show this model spin–up for 3 unique soil parameter sets at the same location in Figure S4.

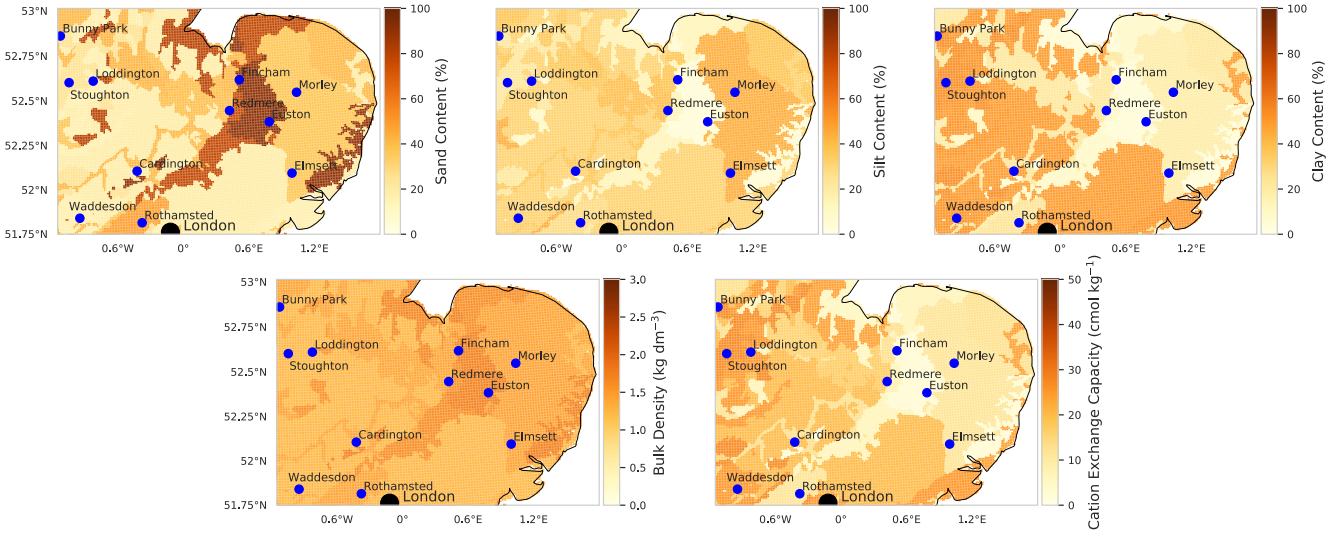


Figure 1. Maps of soil properties from the Harmonized World Soil Database (HWSD) (Fischer et al., 2008) used in the creation of the JULES soil parameter ancillaries with the Tóth et al. (2015) pedotransfer functions. Blue dots show locations of COSMOS-UK probes, black dot shows location of London, UK.

2.2 Pedotransfer functions

The JULES model implements both the Brooks and Corey (1964) and the van Genuchten (1980) soil models. The JULES implementation of these models can be found in Clark et al. (2011). In this paper we have used the van Genuchten (1980) soil model and have selected a set of pedotransfer functions from Tóth et al. (2015). The Tóth et al. (2015) pedotransfer functions have been calibrated across a large range of European soils and should be representative of the study area. The mathematical formulation of these pedotransfer functions is:

$$
\begin{aligned}
\theta_{res} &= \begin{cases} 0.041 & f_{sand} \geqslant 2 \\ 0.179 & f_{sand} < 2 \end{cases} \\
\theta_{sat} &= \phi_a - \phi_b \rho^2 + \phi_c f_{clay} + \kappa_a \mathrm{pH}^2 \\
\log_{10}(\alpha) &= -\phi_d - \phi_e \rho^2 - \phi_f f_{clay} - \phi_g f_{silt} + \frac{\kappa_b}{(C_{organic}+1)} + \kappa_c \mathrm{pH}^2 + \kappa_d \mathrm{topsoil} \\
\log_{10}(N-1) &= -\phi_h - \phi_i \rho^2 - \phi_j f_{clay} - \phi_k f_{silt} + \frac{\kappa_e}{(C_{organic}+1)} \\
\log_{10}(K_{sat}) &= \phi_l - \phi_m f_{clay} - \phi_n f_{silt} - \phi_o \mathrm{CEC} + \kappa_f \mathrm{pH}^2 + \kappa_g \mathrm{topsoil},
\end{aligned}
\tag{1}
$$

where θ_{res} is the residual soil moisture ($\mathrm{m^3\ m^{-3}}$), θ_{sat} the saturated soil moisture ($\mathrm{m^3\ m^{-3}}$), α and $(N-1)$ parameters of the van Genuchten (1980) soil model (-), K_{sat} the saturated hydraulic conductivity ($\mathrm{kg\ m^{-2}\ s^{-1}}$), $\phi_a, \ldots, \phi_o$ are model parameters to be optimised (values given in table 1) and $\kappa_a, \ldots, \kappa_g$ are static model parameters (values given in table 2). We optimise

the parameters controlling the impact of the bulk density ρ (g cm^{-3}), fraction of clay and silt (f_{clay}, f_{silt}) (%) and the cation exchange capacity (CEC) (meq 100g^{-1}) as these terms have a first order impact on the outputted van Genuchten (1980) soil parameters. The organic carbon content ($C_{organic}$) (%), soil pH value and topsoil flag have a less pronounced effect on the van Genuchten (1980) soil parameters. We treat the top two soil layers of JULES as topsoil (topsoil $= 1$) and the bottom two as subsoil (topsoil $= 0$). From equations (1) we can see that defining a soil as topsoil will act to increase the saturated hydraulic conductivity and the value of α, which will both allow water to flow more freely through the soil. The prior values for the parameters $(\phi_a, \ldots, \phi_o)$ are shown in table 1. We used the values given by Tóth et al. (2015) for the prior except for ϕ_o for which we found better results (experiments not shown) when the magnitude of this parameter was increased. To create the JULES soil parameter ancillary files these pedotransfer functions are applied to soil texture information from the HWSD (Fischer et al., 2008) at a 1 km resolution. The DA system used here optimises values for the parameters in table 1 across the whole domain rather than on a grid-by-grid basis. In this way the varied soil properties across the domain give us a form of orthogonal constraint within the assimilation and allow us to recover a single set of pedotransfer functions that are valid in space and time.

Parameter	Prior value
ϕ_a	0.63052
ϕ_b	0.10262
ϕ_c	0.0003335
ϕ_d	1.16518
ϕ_e	0.16063
ϕ_f	0.008372
ϕ_g	0.01300
ϕ_h	0.25929
ϕ_i	0.10590
ϕ_j	0.009004
ϕ_k	0.001223
ϕ_l	0.40220
ϕ_m	0.02329
ϕ_n	0.01265
ϕ_o	0.10380

Table 1. Prior values for parameters of the Tóth et al. (2015) pedotransfer functions used in experiments.

2.3 SMAP Observations

The NASA Soil Moisture Active Passive (SMAP) satellite mission provides estimates of soil moisture every 2-3 days (Entekhabi et al., 2010). The mission is an orbiting observatory with a passive radiometer and an active radar instrument. SMAP was de-

Parameter	Value
κ_a	0.0002904
κ_b	0.40515
κ_c	0.002166
κ_d	0.08233
κ_e	0.2568
κ_f	0.26122
κ_g	0.44565

Table 2. Static parameter values for the Tóth et al. (2015) pedotransfer functions used in experiments.

signed to deliver a 36 km spatial resolution estimate of soil moisture from the passive instrument alongside a 9 km estimate from a retrieval using both the passive and active sensors. After its launch in January 2015 the radar instrument malfunctioned. Subsequently ESA's Sentinel 1 mission has been used as a replacement for the active sensor. For the work in this paper we use the 9 km Level-3 soil moisture product (version 3) this product has a relatively low bias (Colliander et al., 2017; Zhang et al., 2019) . However, it has been shown there is a wet bias present in the Level-4 SMAP product (Reichle et al., 2017) . As part of the retrieval procedure SMAP relies on some ancillary information, one example of this is soil texture where the Harmonized World Soil Database (HWSD) (Fischer et al., 2008) is used to calculate the soil dielectric constant for use within the retrieval algorithm. The use of such ancillary data in the retrieval could introduce additional biases into the SMAP soil moisture estimates that are not consistent with estimates from the land surface model we are comparing to. However, as the HWSD is also used to create the JULES soil parameter ancillary files this effect should be minimised. We prescribe an error of 0.05 m^3 m^{-3} for SMAP observations in the assimilation algorithm. Although the SMAP baseline aim for error is 0.04 m^3 m^{-3} other studies have found slightly higher values for the error in Level-3 SMAP observations (0.043 m^3 m^{-3} (Colliander et al., 2017) , 0.057 m^3 m^{-3} (Li et al., 2018) and 0.054 m^3 m^{-3} (Zhang et al., 2019)), we therefore chose a value between these studies. We have only used SMAP observations corresponding to best retrieval quality flag and surface flag in experiments. The effect that removing poor quality observations has on the total number of observations assimilated can be seen in Figure 2. The experiment area of the East of England is predominantly flat arable land which should allow for good quality SMAP retrievals, there are also coastal and urban areas where SMAP retrievals will be unreliable. This area is also prone to cloud cover which could cause gaps in the SMAP observational record.

2.4 COSMOS–UK Observations

The COSMOS–UK network has been producing observations of soil moisture and other meteorological variables at an expanding number of stations (currently 52) since 2013 (Stanley et al., 2019). For the area of interest in this paper we have 11 stations available to us with data for the relevant time period, see Figure 2. Some of these stations may not be representative of JULES model estimates due to the current setup of JULES not considering some processes

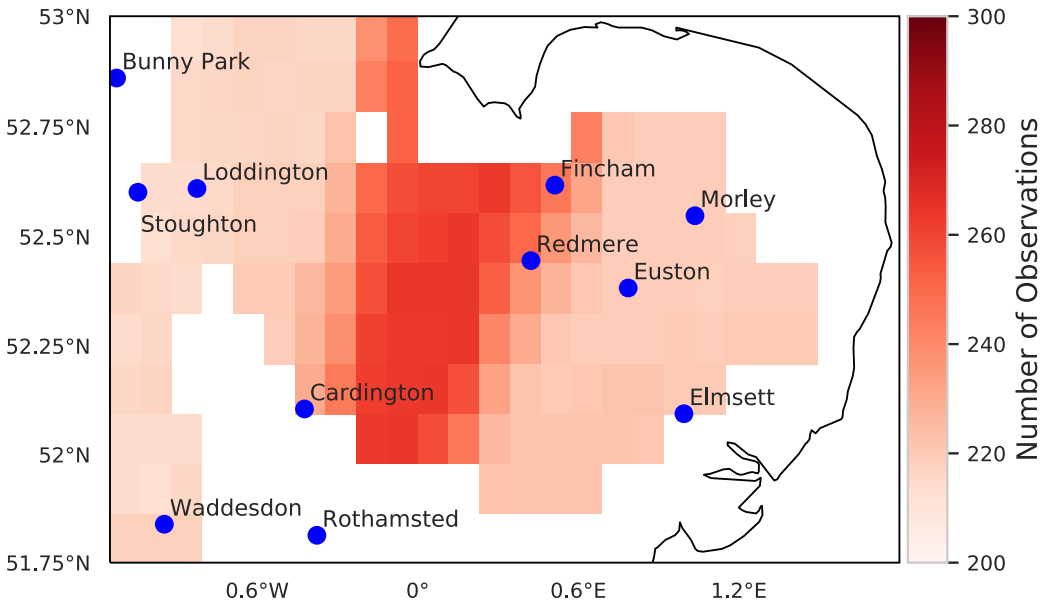


Figure 2. Location of COSMOS probes and number of SMAP observations assimilated in experiment period (2016). No colour corresponds to no observations being assimilated in that location due to low quality retrieval or surface flag.

(ground water, organic soils, urban tiles). Cosmic–ray sensing soil moisture probes have a variable depth as well as horizontal sensitivity (Zreda et al., 2008). There are many studies translating the cosmic–ray neutron intensity measured at COSMOS probe sites to soil moisture (Baatz et al., 2014; Bogena et al., 2015; Köhli et al., 2015) . There have also been efforts to relate modelled soil moisture to cosmic–ray neutron intensity, such as the COsmic-ray Soil Moisture Interaction Code (COSMIC) (Shuttleworth et al., 2013; Rosolem et al., 2014) . The COSMOS–UK network use the N_0–method described by Baatz et al. (2014) to diagnose values for the soil moisture and then the method of Köhli et al. (2015) to calculate the representative depth for each COSMOS probe measurement. The COSMOS sites in our experiment domain have a representative depth of between 14 cm and 40 cm dependent on conditions when measurements are made. To make a fair comparison between the COSMOS–UK and JULES soil moisture estimates we have constructed a simple variable depth algorithm for JULES which takes a weighted average of the different soil layers of the model given the relative depth of the COSMOS–UK observation. This is defined as

$$\theta_D = \begin{cases} \theta_{10}, & \text{if } D \leq 10 \text{ cm} \\ \frac{10}{D}\theta_{10} + \frac{(D-10)}{D\theta_{25}}, & \text{if } 10 \text{ cm} < D \leq 25 \text{ cm} \\ \frac{10}{D}\theta_{10} + \frac{25}{D\theta_{25} + \frac{(D-35)}{D}\theta_{65}}, & \text{if } 35 \text{ cm} < D \leq 65 \text{ cm} \end{cases} , \quad (2)$$

where θ_D is the JULES modelled soil moisture at the COSMOS-UK representative depth (D) and θ_{10},θ_{25} and θ_{65} are the top, second and third layer soil moisture estimates from the JULES model.

2.5 Data Assimilation Framework

In order to estimate the identified pedotransfer function parameters we use the LAVENDAR data assimilation framework (Pinnington et al., 2020). This framework utilises a hybrid DA technique similar to that of the Iterative Ensemble Kalman Smoother (IEnKS) (Bocquet and Sakov, 2013) . A smoother is different than a filter (e.g. the Ensemble Kalman Filter (Evensen, 2003)) in that it uses batches of observations which are taken over a time window of given length and the whole spatial domain, as opposed to just in a time instant. These observations are combined with the model evolution over this window and a minimization process is performed to obtain initial conditions for the state/parameter values. It is possible to run a sequence of smoother steps for successive windows, but our study only uses one year long assimilation window as the parameters we are optimising do not vary in time.

Using a smoother instead of a filter has advantages (Lorenc and Rawlins, 2005) in that (a) more observations can be used to constrain the problem solution, and (b) information from the model evolution is implicitly used in the search process. However, using a smoother requires computing the Jacobian of the model, the so–called tangent linear model (TLM) and the related adjoint model (AM). The TLM/AM (Courtier et al., 1994) . Computing and maintaining the TLM/AM is not a trivial task, and in fact we do not have this for JULES. The IEnKS solves this problem by replacing the role of the TLM/AM by 4–dimensional covariances, i.e. covariances defined over time and space. These covariances are computed as sample estimators of a given ensemble. The iterative nature of the method means that it finds the solution to the minimization problems using inner iterations rather than a single step (hence the variational nature), and this helps when the distributions of the variables/parameters of interest are not Gaussian. We provide details of the method in Appendix A. Furthermore, to understand the variants of the ensemble Kalman Smoother and its position within the hybrid DA methods, the reader is referred to Evensen (2018) .

We show a schematic of how this system works in Figure 3, this involves running an ensemble of JULES models, with each model in the ensemble utilising a distinct soil ancillary data–set. Each ensemble members ancillary file is created by sampling from the normal distribution defined by mean $\mathbf{x}_b$ and variance $(0.1 \times \mathbf{x}_b)^2$, where $\mathbf{x}_b = (\phi_a, \phi_b, \ldots, \phi_o)$ with $\phi_a, \ldots, \phi_o$ taking the values given in table 1, then using each unique set of sampled parameters within equations (1) applied to the HWSD maps of soil properties (see Figure 1) for the whole domain. Although van Genuchten and hydraulic conductivity parameters can be described by logarithmic distributions it is less clear what distribution is best for the PTF parameters optimized here. We therefore made the naive assumption of a normal distribution in the first instance as this gave us good results. In this type of experiment the number of ensemble members will control the quality of the results, with a larger ensemble more likely to identify the optimum parameters. However, running a land surface model at a 1 km spatial resolution over the specified domain is computationally expensive, we therefore use an ensemble size of 50 in our experiments. In order to compare the 1 km estimates of soil moisture from JULES to the 9 km SMAP estimates we create an observation operator which aggregates the JULES grid cells within each SMAP pixel by taking a spatial average of all JULES estimates which fall in the bounds of the SMAP grid cell. There is no need to project increments from the spatially averaged 9 km model estimates back to the 1 km model grid as the assimilation is only optimising the 15 PTF parameters $(\phi_a, \phi_b, \ldots, \phi_o)$ for the whole domain and the update to soil moisture will be implicit. The aggregated spatial observation operator will introduce

an additional source of representativity error alongside the observational error of SMAP and the inherent model error within JULES. It has been shown that, for variational methods such as the one used in this paper, these additional sources of error (model error, representativity error, etc.) can be included in the observational term of the cost function by inflating the diagonal observation error variance (Howes et al., 2017) . Although observation error inflation is rare in relation to filtering DA methods it is commonly used with variational methods and smoothers, especially in numerical weather prediction (Hilton et al., 2009; Bormann et al., 2015; Minamide and Zhang, 2017; Fowler et al., 2018; Wang et al., 2019) . Observation error inflation is required due to the fact that all observations are used at once in the assimilation whereby we minimise a cost function containing a prior term and an observational term. The greater the number of observations in the observational cost function term, the higher the weight they have in the optimization. This can lead to the prior term being completely negated and hence the retrieval of nonphysical parameters. Observation error inflation would not be required if the correct specification for the observation error correlations (in space and time), model error and representativity error were included. These, however, are hard to diagnose and it has been shown that in the absence of such information observation error inflation is required for an optimal DA system (Stewart et al., 2014) . For this reason and due to the large number of observations assimilated in our one year assimilation window (28698) we inflate the specified observational error by a factor of four. If a filtering DA system were being used utilising a bias aware method such as that presented by Ridler et al. (2017) could help represent some of the additional sources of error discussed here.

2.6 Experiment Formulation

We conducted our pedotransfer function parameter estimation for the year of 2016 using all SMAP observations in this period. We also ran the prior and posterior JULES ensembles into 2017 so that we could judge the results against independent SMAP observations in a 'hindcast' experiment. Allowing us to judge if any skill added by the assimilation persisted into the future. For the 2016-2017 period we then used the available COSMOS probe observations for validation comparing both prior and posterior JULES soil moisture estimates to these observations. Using the COSMOS–UK observations in this way gives us a better understanding of whether information added by the assimilation of SMAP observations can help to improve model estimates at in–situ scales.

3 Results

3.1 Assimilation Output

The input to the data assimilation routine is an ensemble of 50 unique Tóth et al. (2015) PTF parameter sets drawn from a prior distribution (representing our best a priori guess to the true PTF parameters), the corresponding JULES runs (2016-2017) for each PTF parameter set and all the SMAP observations for the year 2016 over the experiment domain. The output of the data assimilation is an ensemble of 50 optimised (posterior) PTF parameter sets, valid for the whole experiment domain and time, this allows us to calculate the posterior JULES soil ancillary files for each optimized parameter set and the corresponding

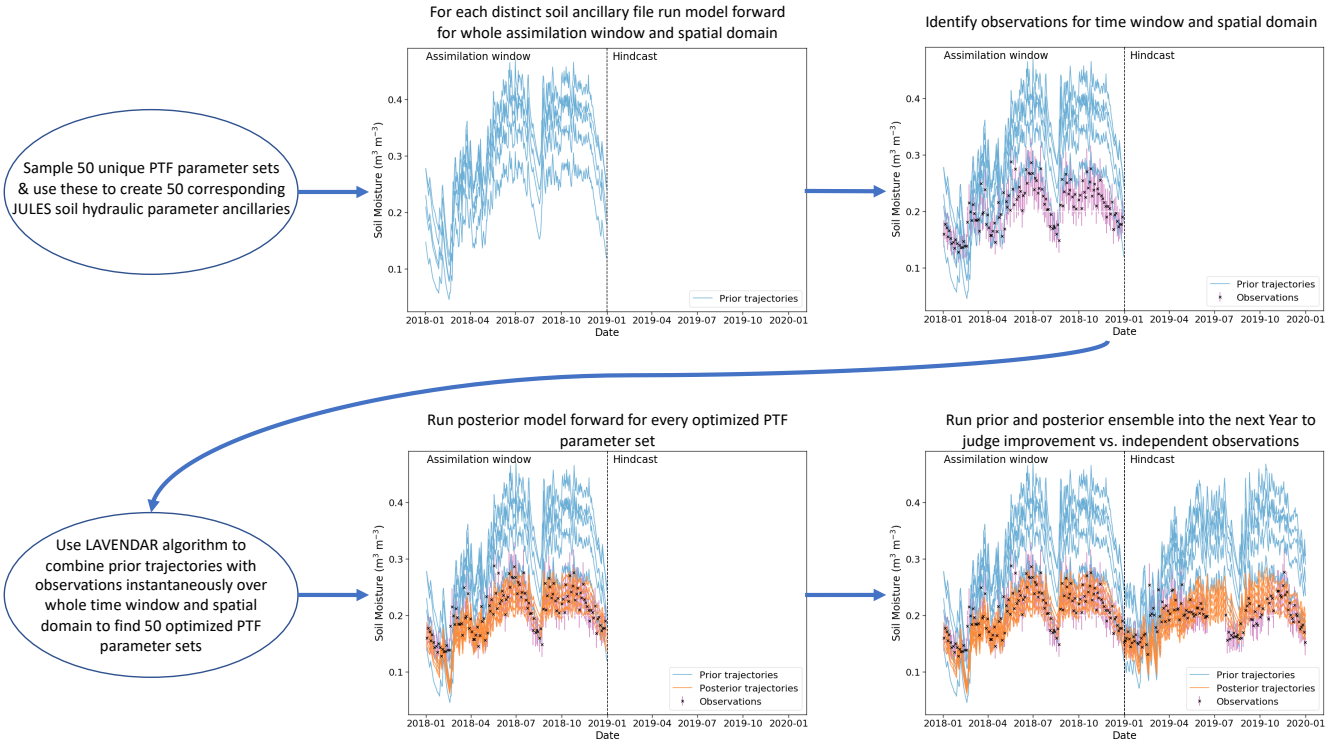


Figure 3. Schematic of the LAVENDAR data assimilation framework, showing the workflow for the experiment.

posterior JULES model runs for 2016-2017. Figure 4 shows the prior and posterior parameter distributions for the 15 optimized parameters of the Tóth et al. (2015) pedotransfer functions. Prior distributions for the 50 JULES ensemble members are shown in light grey with posterior distributions shown as dark grey. We can see that while the DA procedure made large updates to some parameters compared to their prior values others have not changed, with their mean appearing to be in a very similar place. One of the parameters with a strong change is ϕ_a which is decreased compared to the prior, this parameter controls the absolute magnitude of the saturated soil moisture (θ_{sat}). Decreasing it will reduce the absolute saturated soil moisture and allow the soil texture information to have more impact on the diagnosed van Genuchten (1980) model parameter. This can be seen in Figure 5 where we show the updated PTF parameters effect on the mean estimate to the JULES model soil parameters when applied to the spatial maps of soil properties from the HWSD. We can see how different areas of distinct soil texture (see Figure 1) behave differently based on the PTF parameter updates after DA. For some parameters we see the majority of grid cell parameter values increase or decrease, θ_{sat} and $\frac{1}{N-1}$ respectively. Whereas for $\frac{1}{\alpha}$ and θ_{crit} we see an increase or decrease in grid cell parameter values dependent on the underlying soil properties (sandier soils lead to an increase, less sand more clay correspond to a decrease).

In Figure 6 we show the difference between mean water budget variable estimates (soil moisture, evapotranspiration and run off) in 2016 for the prior and posterior JULES model ensemble. The grid cells that are darker blue correspond to the posterior

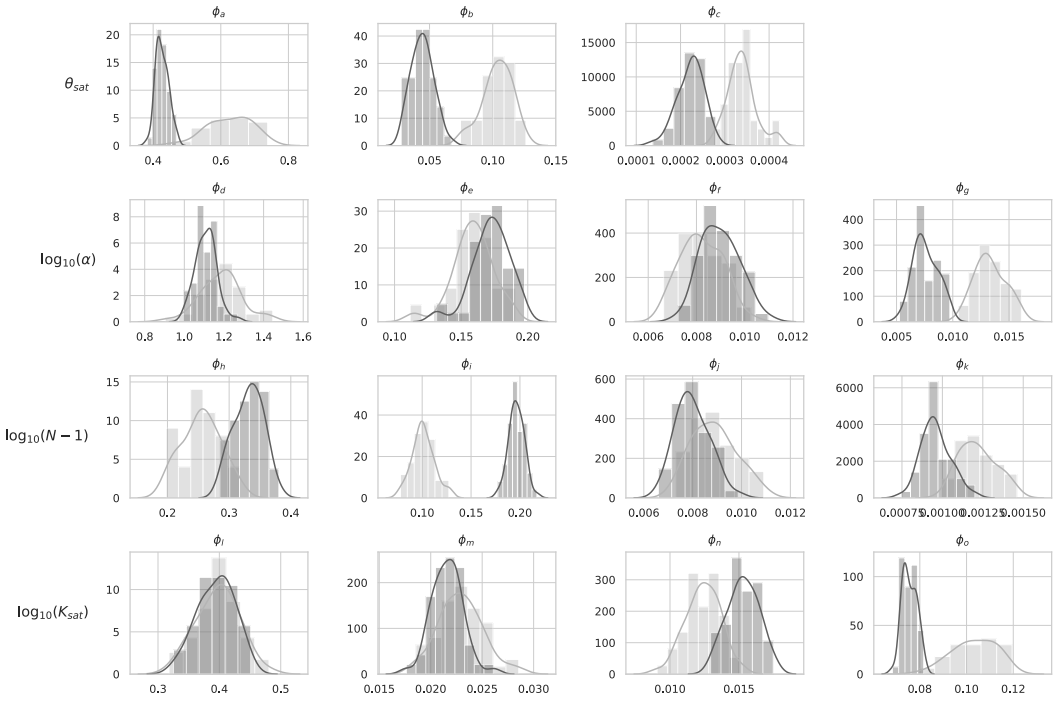


Figure 4. Distributions of prior and posterior pedotransfer function parameters grouped by the term in the equations (1) that they relate to (see row labels). Light grey: parameter distribution for the prior ensemble, dark grey: parameter distribution for the posterior ensemble.

ensemble estimate being higher after assimilation and grid cells that are darker red correspond to the posterior estimate being lower. We can see that after calibration of the pedotransfer function parameters the domain has not had a uniform increment to the value of mean soil moisture, evapotranspiration or run off. This is due to the fact that soil texture specific parameters have been optimised allowing the different distinct areas of soil type defined by the HWSD (see Figure 1) to behave differently rather than having a uniform correction across the modelled area. Across the whole domain we find an average increase of 0.03 m^3 m^{-3} in mean soil moisture estimates after data assimilation. We can see that in order to update PTF parameter values to find soil moisture estimates that more closely match the SMAP observations both evapotranspiration and run off model estimates have also been modified. In areas of sandy soils wetter soil moisture values have been achieved by a decrease in evapotranspiration offsetting a slight increase in runoff. In areas of high clay content wetter soil moisture values have been achieved by a larger decrease in run off compared to an increase in evapotranspiration. For silty soils we find a drier value of soil moisture for the posterior compared to the prior with a less prominent impact on evapotranspiration and run off. Figure 6 also allows us to see the high–resolution of the JULES model when run with the CHESS data, for this domain we have over 30,000 individual model grid cells.

Figure

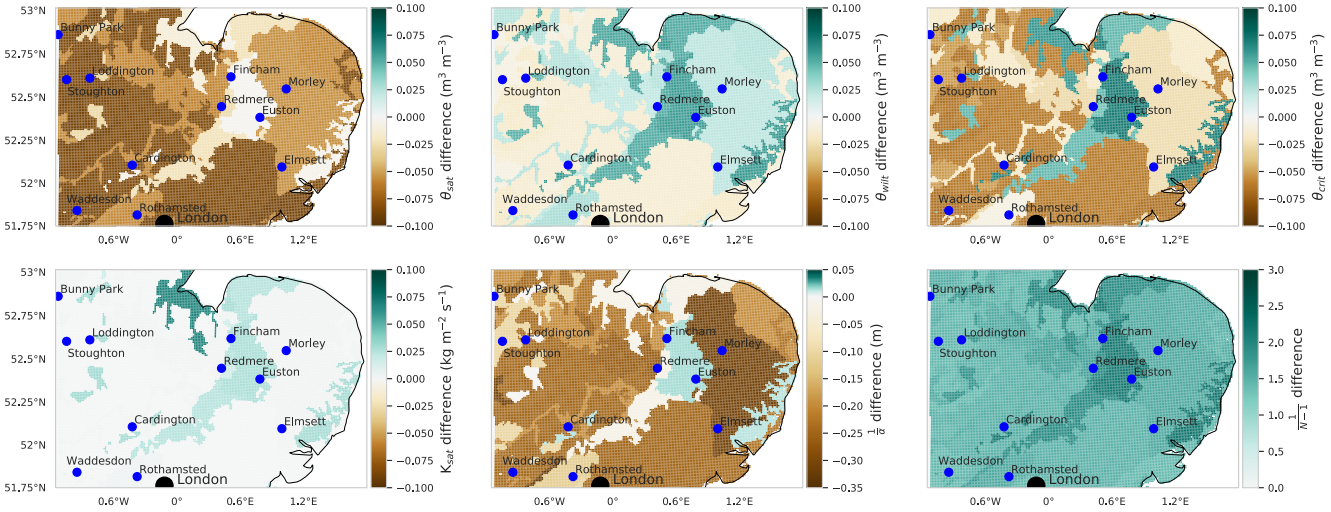


Figure 5. Maps showing the difference between the prior and posterior mean JULES model soil parameters, created by applying the prior/posterior PTF's to the HWSD maps of soil properties. Brown corresponds to a decrease in the soil parameter after data assimilation, green to an increase.

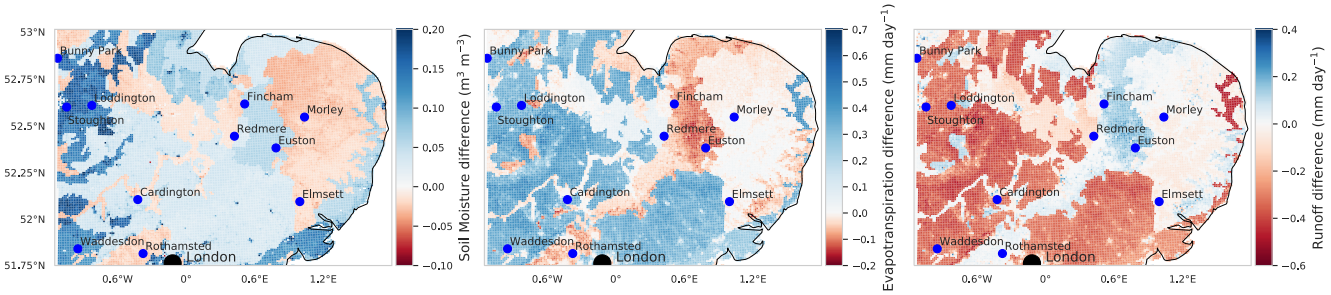


Figure 6. Map showing the difference between yearly mean soil moisture for the prior and posterior ensemble of JULES model runs in 2016. Blue corresponds to the posterior ensemble estimate being wetter, red corresponds to the posterior being drier.

7 shows the error reduction after performing data assimilation when comparing JULES spatially aggregated estimates to SMAP observations. This is computed as $100 \times \frac{(RMSE_{prior} - RMSE_{post})}{RMSE_{prior}}$, where $RMSE_{prior}$ is the JULES prior ensemble mean RMSE when compared to 2016 SMAP observations and $RMSE_{post}$ is the JULES posterior ensemble mean RMSE when compared to 2016 SMAP observations. As we are minimising a cost function to find optimised values of PTF parameters valid for the whole spatial and temporal domain it is possible the optimisation may have to degrade the fit of the model estimates to the SMAP observations at certain locations in order to improve the picture as a whole. This could be due to errors at these locations in driving data, the underlying soil property map or indeed in the model structure. For the majority of the domain we find a reduction in error after assimilation, with a mean error reduction of 20% in 2016 and 21% in 2017. The exception to this

being the area corresponding to the city of London. There are two reasons for this, firstly we have not assimilated SMAP soil moisture estimates over this area due to the surface flag corresponding to poor quality observations (poor quality SMAP grid cells are shown in Figure 7 with stippling). Secondly the setup of JULES we have used in our experiments does not have the urban tile turned on, instead we have had to gap-fill the HWSD over London with the surrounding grid cells soil type. This means that soil moisture estimates for this location will not be realistic. To visualise what the time-series of results looks like we plot SMAP observations and JULES model predictions for different pixels in Figure 8 & 9. From these figures we can see the distinct seasonal dynamics of soil moisture in this region, with the highest moisture being in the winter months and a distinct dry down from April into the summer months. This seasonal cycle is seen for both the JULES model and SMAP obsereved estimates. For Figure 8 we can clearly see the improvement in the posterior JULES ensemble estimate when compared to the prior. This improvement continues into the 2017 hindcast period when judged against observations that have not been used in the cost function of the data assimilation framework. We can see that although the dynamics in 2017 are distinct from those used for calibration in 2016 we still match the SMAP estimates for dry-down and re-wetting of the soil in this period. From Figure 8 we can also see the spread in our model estimates with the JULES ensemble standard deviation displayed as shading. This spread is decreased from the prior to posterior estimates. In Figure 9 we plot the results for a SMAP pixel over London where the posterior error increases compared to the prior. However, we can see that the SMAP observations do not appear reliable here with many observations hitting the lower bound of soil moisture in the SMAP retrieval. In Figure 10 we show the RMSE averaged in space for the JULES model prior and posterior mean estimate, when compared to SMAP, alongside the JULES model prior and posterior ensemble spread. At all times the posterior JULES RMSE is lower than that of the prior, showing that the DA system has found a set of PTF parameters that improve the fit to the SMAP observations through time, this continues into the hindcast period (2017) when judged against observations that were not included in the DA cost function. We find slight peaks in the RMSE values throughout the time period corresponding to wetter conditions, this could be due to slight errors in the precipitation driving data used to force the model. It is optimal to have an ensemble spread that matches the magnitude of the ensemble mean RMSE and this relationship should hold given a large enough ensemble size (Houtekamer and Mitchell, 1998) . We can see that this relationship holds for our prior estimates. However, after DA the posterior ensemble spread is slightly lower than that of the ensemble mean RMSE. This is perhaps unsurprising as we are conducting just a single assimilation step using all observations (over 28000) at once in space and time with a relatively small ensemble size (50). This can lead to some of the posterior parameter distributions becoming narrow, as with increasing observations we increase the confidence in our posterior, thus tightening the retrieved distributions and reducing the model ensemble spread. This result suggests that ensemble inflation (Anderson and Anderson, 1999) may be necessary if this ensemble was to be used in subsequent assimilation experiments.

3.2 Comparison to COSMOS–UK

After performing the data assimilation procedure we use the observation operator described in section 2.4 to compare the prior and posterior JULES 4–layer soil moisture estimates to the 11 COSMOS probes located in our experiment domain. For each COSMOS site we select the nearest JULES grid–cell to the given site longitude and latitude. In Figure 11 we show results at the Cardington COSMOS site, here we can see the posterior JULES estimate is a large improvement from the prior, although some of

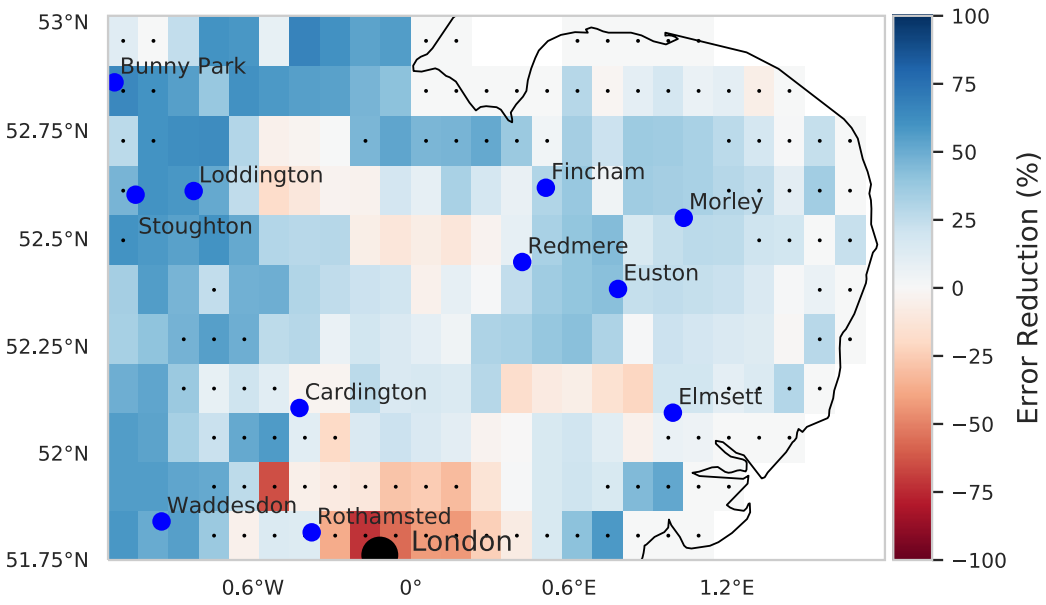


Figure 7. Map showing the difference between Root-Mean Squared Error (RMSE) when JULES spatially aggregated estimates are compared to SMAP observations for 2016 and 21% for 2017.

the driest values are still not captured. From Figure 11 we can also see there is an increase in evapotranspiration and a decrease in runoff, this effect can also be seen from Figure 6. Figure 12 shows results for Morley COSMOS site where both prior and posterior JULES estimates perform similarly, we also have less of an update to evapotranspiration but a decrease in modelled runoff. There are also some sites where even after calibration we still do not capture the COSMOS estimates, Stoughton in Figure 13 is such an example where both prior and posterior estimates are too dry. However, here the posterior estimate is still much improved from the prior. We also find large increases in evapotranspiration and reductions in runoff for Stoughton. Figure 14 is an example where both prior and posterior perform equally poorly. The fact that the estimates and updates after DA are so different for Figures 11 - 14 despite all using the same PTF parameters highlights the effect that the underlying soil properties are having on soil hydraulic conductivity. At all sites the JULES model predicts top layer soil temperature well when both prior and posterior estimates are compared to in-situ observations. In table 3 we show summary statistics for soil moisture at the 11 COSMOS sites, we see that when looking over all sites the posterior estimate yields a 16% increase in correlation, 16% reduction in unbiased Root-Mean-Squared Error (ubRMSE) and a 22% reduction in Root-Mean-Squared Error (RMSE) when compared to the prior.

The COSMOS-UK observations we have used for independent validation of the results are respresentative of depths from 14 cm up to around 40 cm. The SMAP satellite observations, used within the assimilation algorithm to find a new set of pedotransfer functions for the experiment domain, are representative of soil moisture for the top 2.5 - 5 cm of soil. Therefore the fact that after assimilation we find such a distinct improvement at in-situ COSMOS probe locations indicates that although

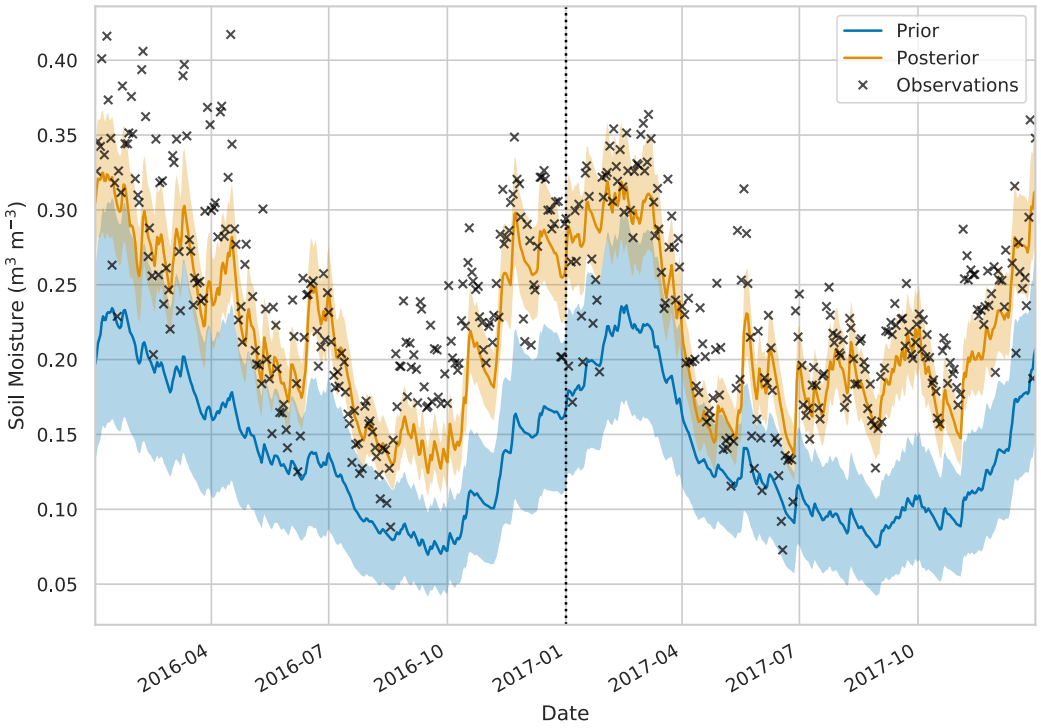


Figure 8. Time–series of soil moisture for 52.96° N 0.40° W. Black crosses: SMAP observations, blue line and shading: prior JULES mean and ensemble spread, orange line and shading: posterior ensemble mean and spread. Black dotted line represents the end of the assimilation window and start of the hindcast period.

the SMAP observations are only sensitive to shallow depths, by combining these with the JULES model we are also improving estimates at deeper levels. The large errors in our prior JULES estimates for the COSMOS sites in Figure 13 and 14 could point towards some systematic bias within the model. However, it is important to note that the COSMOS-UK observations are independent of the data assimilation. For the assimilated SMAP observations it may be optimal to have errors centred around zero but for the independent in-situ validation data there will be many competing errors that may make this impossible. There will be errors in the forcing meteorology (here we are using CHESS 1km forcing data and not observed in-situ meteorology), errors in the model grid and its representativity to the in-situ location, structural model errors (we currently have no ground water model in JULES and some in-situ sites may be more ground water dominated), errors in the vegetation fractions, and many more. At the larger SMAP scale many of these effects will be minimised when looking at the 9 km spatial scale that is more representative of modelled estimates.

4 Discussion

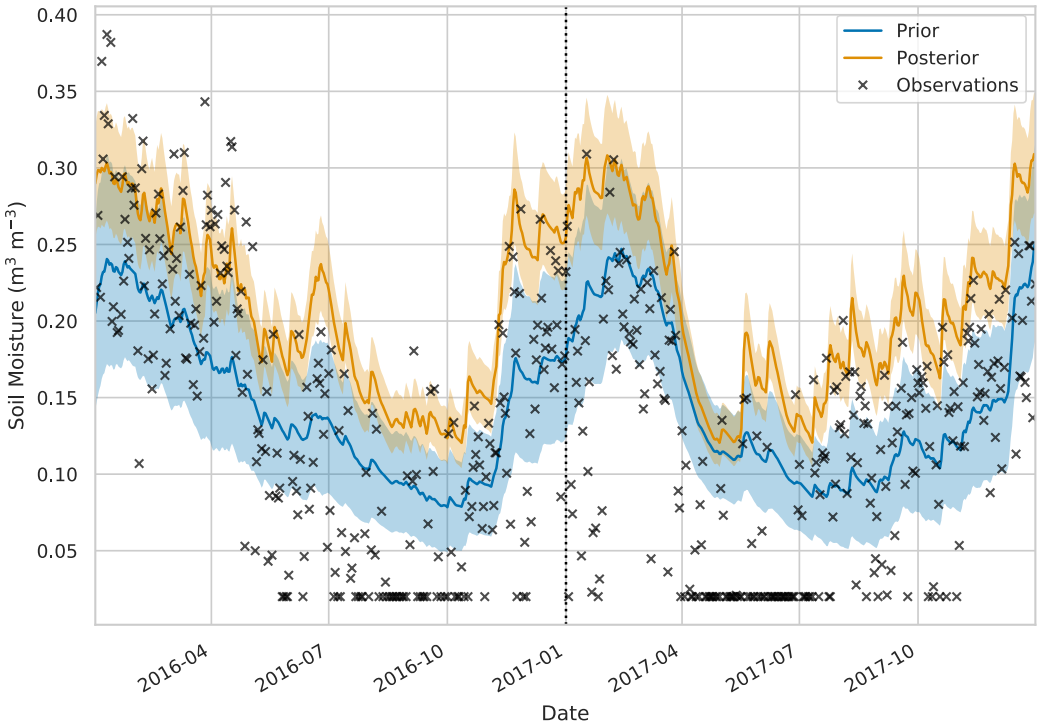


Figure 9. Time–series of soil moisture for 51.81° N 0.17° W. Black crosses: SMAP observations, blue line and shading: prior JULES mean and ensemble spread, orange line and shading: posterior ensemble mean and spread. Black dotted line represents the end of the assimilation window and start of the hindcast period.

This study aimed to determine the suitability of satellite observations to optimise pedotransfer functions and improve soil moisture estimates for a land surface model. Currently pedotranfer functions are calibrated through analyses of point soil samples and it is unclear how these calibrations and their resultant soil model parameters relate the the varying spatial resolutions of modern land surface models. Adding additional information from satellite estimates into the callibration of pedotransfer functions should address a key uncertainty with respect to the larger scales of land surface model estimates.

We used the LAVENDAR hybrid data assimilation framework (Pinnington et al., 2020) to optimise the parameters of the Tóth et al. (2015) pedotransfer functions by combining them with SMAP Level-3 9 km satellite observations and the JULES land surface model run at a 1 km resolution. This framework outputs a single set of PTF parameters valid in space and time by utilising all data at once through the minimisation of a cost function. The optimized pedotranfer functions found after DA were shown to improve model estimates of soil moisture when compared to SMAP data from a different time period (21% reduction in RMSE) and independent in-situ observations from the COSMOS-UK network (16% increase in correlation, 16% reduction in ubRMSE and 22% reduction in RMSE over 11 sites). This demonstrates that satellite observations can be used to update pedotransfer functions and improve estimates of soil moisture for land surface models. Previous studies

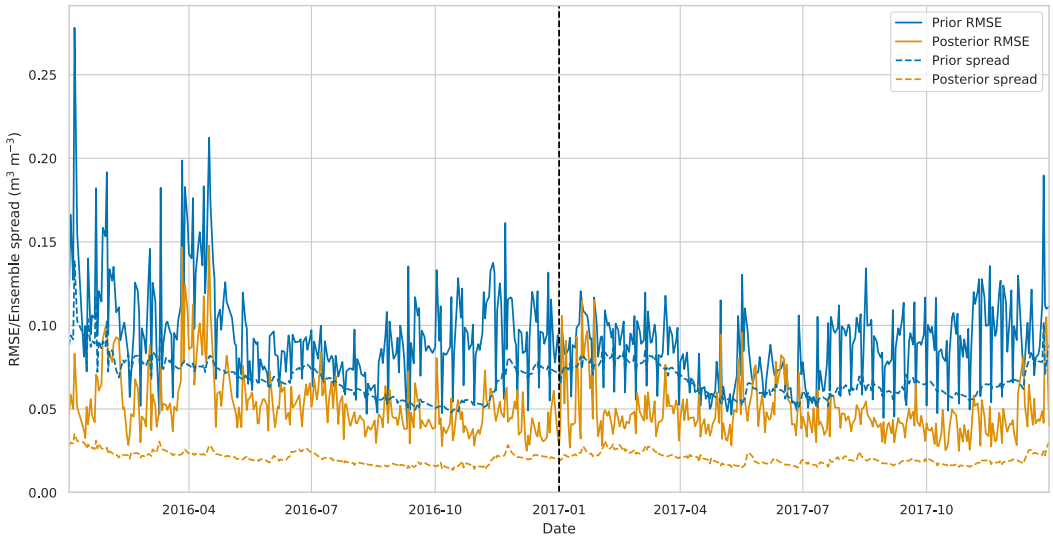


Figure 10. Spatially averaged RMSE and ensemble spread for JULES prior and posterior model estimate. Blue solid line: prior JULES RMSE, blue dashed line: prior JULES ensemble spread, orange solid line: posterior JULES rmse, orange dashed line: posterior JULES ensemble spread. Black dotted line represents the end of the assimilation window and start of the hindcast period.

have shown that satellite observations can be used to improve model estimates of soil moisture by directly updating soil model parameters on a grid by grid basis. Han et al. (2014) used observations from the Soil Moisture Ocean Salinity (SMOS) mission (Kerr et al., 2001) to update parameters of the Community Land Model (CLM) in a Local Ensemble Transform Kalman Filter (LETKF) and improved model estimates. Yang et al. (2016) used a variational method to combine observations from the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) (Kawanishi et al., 2003) with a land surface model to improve estimates over the Tibetan and Mongolian Plateau. (Nearing et al., 2010) used calibration techniques to update NOAH land surface model parameters using synthetic aperture radar imagery at a site in Arizona, USA. Our results show similar improvements are achieved by updating PTF parameters with SMAP satellite data. We also demonstrate that information from such satellite observations which are representative of a larger spatial area (9 km) and shallow soil depth (2.5 - 5 cm) allow us to improve 1 km model estimates at independent COSMOS probe sites. The COSMOS probes are representative of a much smaller spatial scale (∼300 m) and a deeper soil layer (14 - 40 cm), meaning that by combining SMAP observations with he JULES model we are able to find PTF's that better represent finer spatial scales and deeper soil moisture's.

The correlated nature of the PTF parameters in equation (1) presents a potential source of equifinality (e.g. both ϕ_a and ϕ_c both act to increase the magnitude of θ_{sat} in the presence of clay soils), this means that we could achieve the same soil hydraulic conductivity with multiple realisations of PTF parameters at any individual grid cell. The effect of this is greatly reduced as we are performing the optimization over the whole domain and not on a grid cell by grid cell basis. In effect this

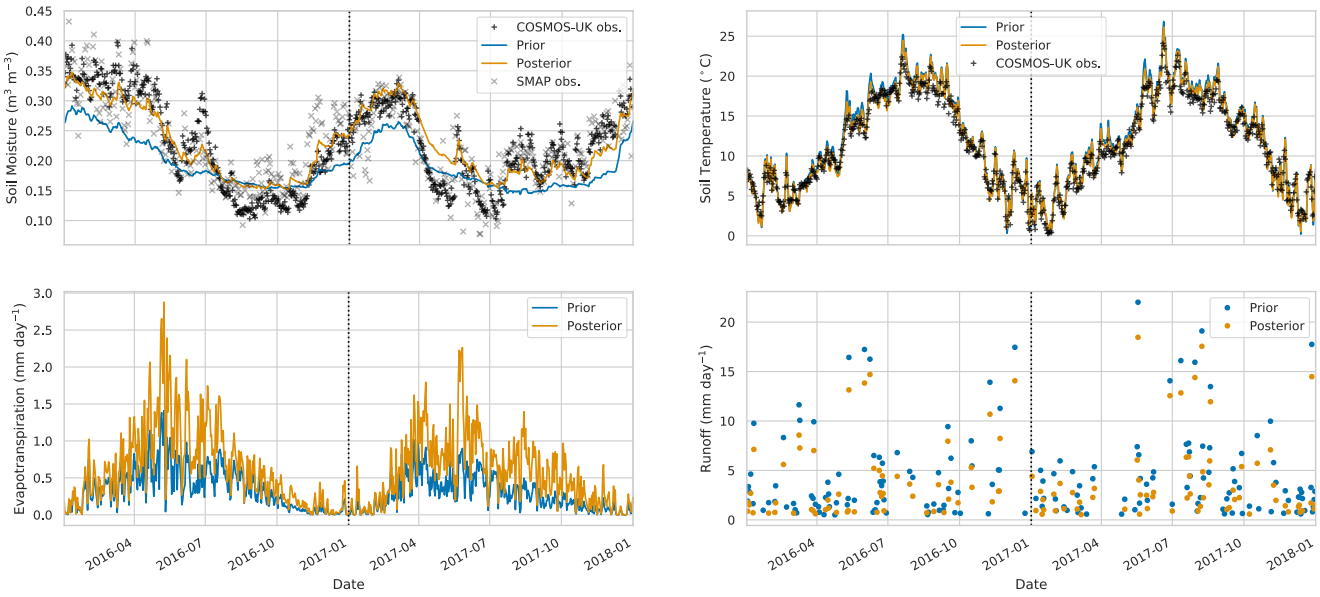


Figure 11. Times–series of water budget variables and soil temperature at Cardington COSMOS site. Black plus signs: COSMOS-UK observations, grey crosses: SMAP observations for closest 9 km pixel, blue line: prior JULES estimate for closest 1 km grid cell, orange line: posterior JULES estimate for closest 1 km grid cell.

means the unique soil properties at each of the 30614 model grid cells act as orthogonal constraints within the DA algorithm and reduce the issue of equifinailty for the optimized PTF parameters as the DA algorithm is having to fit the assimilated soil moisture observations for many different soil textures at once. It may also be possible to improve results further by including information on such correlations within our prior. Such estimates have been included in a variational DA framework for the carbon cycle and shown to improve posterior estimates (Pinnington et al., 2016) . Previous studies have noted the issue of equifinality when optimising soil model parameters on a grid by grid basis (Beven, 2001) . Samaniego et al. (2010) proposed the multiscale parameter regionalization method to alleviate this issue by performing a spatial uniforming function and linking parameters at coarser scales to those at finer resolutions. Our technique also allows for a vastly reduced parameter space by moving from updating gridded soil model parameters to instead optimizing a single set of pedotransfer function parameters valid in space and time. This could also lead to issues as we are not considering uncertainty in the underlying soil property database (Fischer et al., 2008) , which could have contain errors (Tifafi et al., 2018) . It may be appropriate when performing such a technique at a larger scale that the optimization is split up into different calibration zones as it has been shown that pedotransfer functions in certain regions can have a different form (e.g. tropical soils (Marthews et al., 2014)).

Within the DA procedure used to optimise the PTF parameters there are uncertainties that have not been explicitly prescribed. There will be inherent bias and errors in both the observations and model. For SMAP any bias contained in the observations could cause us to retrieve PTF parameters that result in erroneous soil hydraulic conductivity's and ultimately degrade the

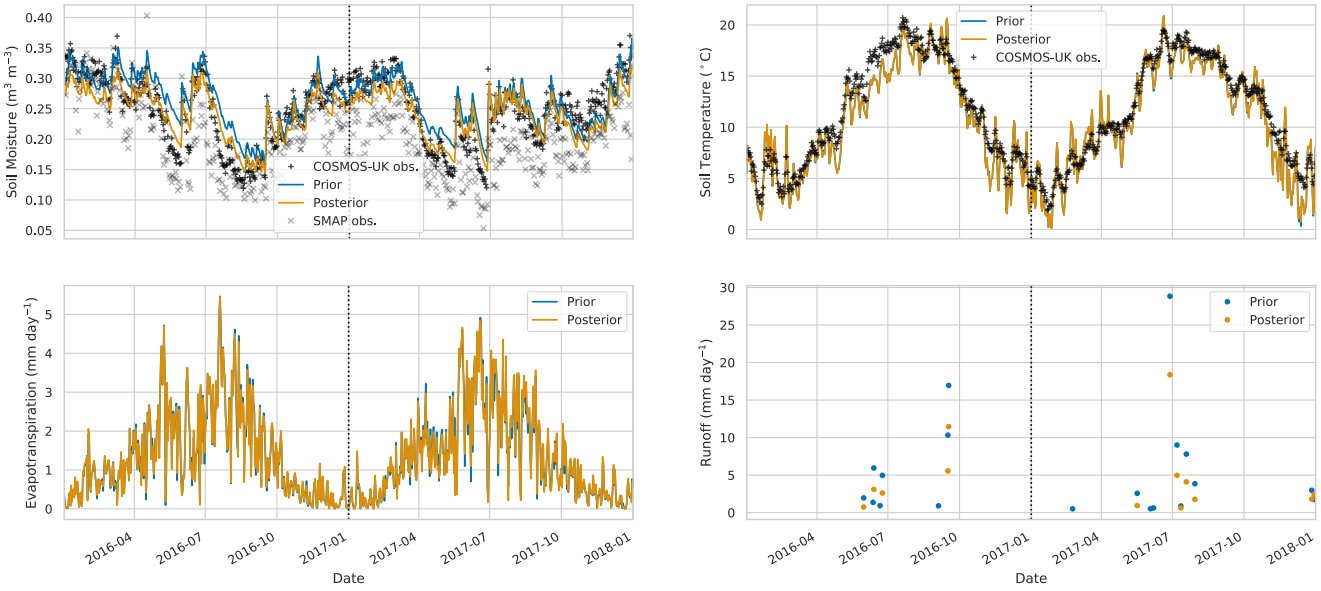


Figure 12. Times–series of water budget variables and soil temperature at Morley COSMOS site. Black plus signs: COSMOS-UK observations, grey crosses: SMAP observations for closest 9 km pixel, blue line: prior JULES estimate for closest 1 km grid cell, orange line: posterior JULES estimate for closest 1 km grid cell.

performance of other model components. It has been shown that the Level-3 9 km SMAP observations used here do not have a significant bias (Colliander et al., 2017) especially in temperate regions (Zhang et al., 2019) . The fact that after assimilation of the SMAP data we not only reduce the RMSE of JULES compared to SMAP but also reduce the RMSE of JULES compared to independent COSMOS estimates also gives us confidence that the bias in the assimilated SMAP data is relatively low. We have dealt with the many errors contained within our DA cost function by inflating the observation uncertainty within the observation error covariance matrix, as described in section 2.5. However, specifying the errors arising from structural uncertainties and missing processes within the JULES model is difficult. We can see these errors manifesting themselves in our comparisons to COSMOS–UK observations in Figures 11 to 14. Figure 11 displays results for the Cardington cosmic–ray probe, this site is a level well–managed grassland with a typical mineral soil and is therefore well modelled by JULES which has the ability to represent the processes of such a site. Both the Morley and Stoughton sensors (Figure 12 and 13 respectively) are positioned on arable land with typical mineral soils and while we model Morley well we struggle to match the magnitude of the Stoughton observations. It is possible that different management practices at the respective sites are impacting on the ability of JULES to predict the observations. In this paper we have not run JULES with its in–built crop model turned on, so that the model will struggle to represent heavily managed crops that behave distinctly from a grassland. The site at which both prior and posterior perform worst is Redmere (Figure 14), this cosmic-ray probe is again on arable land but with a soil type of peat. In its current configuration JULES does not model organic soils and estimates of soil moisture from microwave satellite sensors over

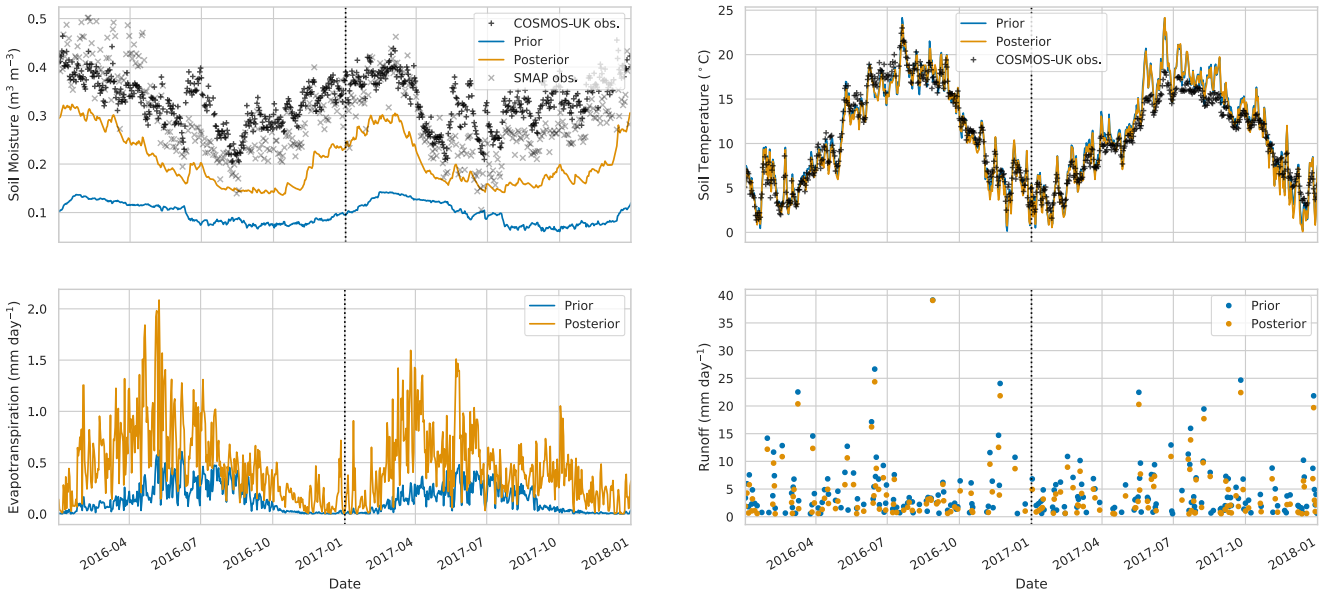


Figure 13. Times–series of water budget variables and soil temperature at Stoughton COSMOS site. Black plus signs: COSMOS-UK observations, grey crosses: SMAP observations for closest 9 km pixel, blue line: prior JULES estimate for closest 1 km grid cell, orange line: posterior JULES estimate for closest 1 km grid cell.

peatland are problematic (Zhang et al., 2019) , so it is understandable that we are unable to match the much wetter conditions observed at this site. The accuracy of JULES posterior estimates is also contingent on the assimilated SMAP observations, so if SMAP estimates have large errors compared to cosmic-ray probe observations JULES will be unable to improve from its prior predictions.

In the initial application of this technique we have focused on a specific region at a high resolution. Here we have utilised 256 processors to run the JULES model ensemble, with each JULES run utilising message parsing interfaces to disaggregate the spatial domain of the model and split the computational load across multiple processors. In this set up it has taken approximately 1.5 days to complete 100 JULES model runs, with each model being for 30614 grid cells and over 6 years (2016 to 2017, with a 4 year spin-up). In order to find a set of pedotransfer function parameters valid at the global scale, using the technique presented here, we would need to decrease the spatial resolution. Working at the scale of 0.5 degrees we would have approximately 67000 land grid cells globally. Using our fairly modest experimental setup and assuming a linear scaling repeating at the global scale would still only take a little over 3 days. However, it may be beneficial to focus on regional efforts to ensure the optimised pedotransfer functions best reflect the behaviour of local soils. The global domain could then be decomposed into sub-regions with specific parameters being found for each distinct region.

Both SMAP and COSMOS–UK observations represent a valuable resource for validation and improvement of land surface models and could be further utilised still. It is possible that our formation of a spatially aggregated observation operator to compare

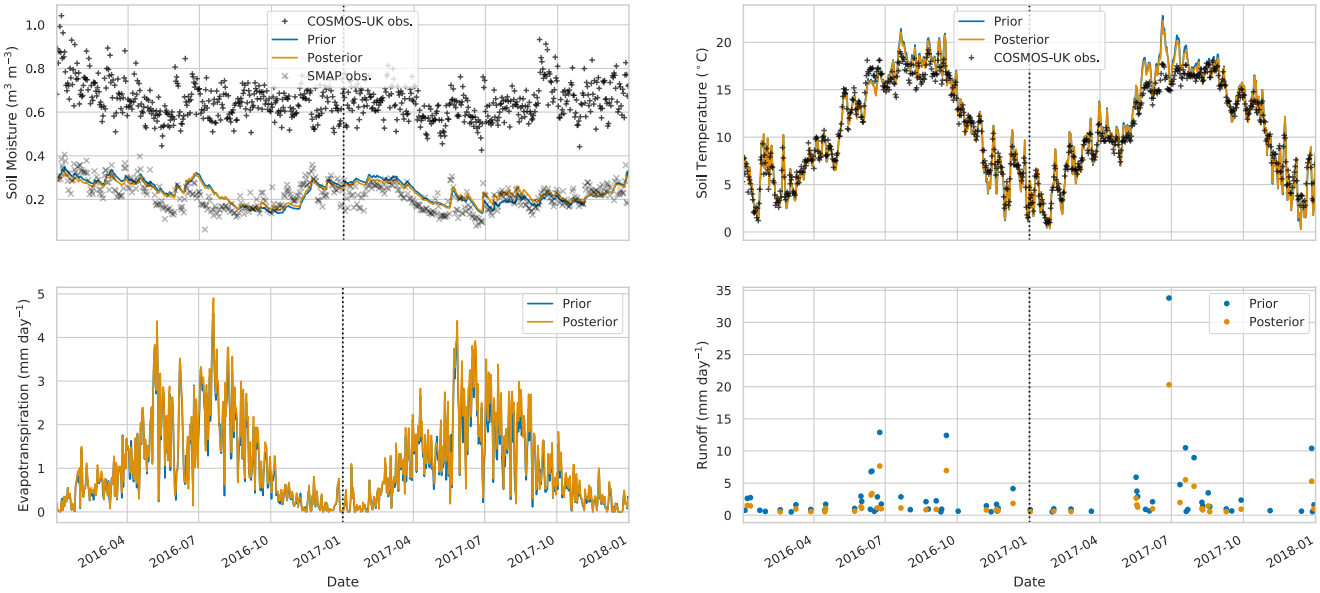


Figure 14. Times–series of water budget variables and soil temperature at Redmere COSMOS site. Black plus signs: COSMOS-UK observations, grey crosses: SMAP observations for closest 9 km pixel, blue line: prior JULES estimate for closest 1 km grid cell, orange line: posterior JULES estimate for closest 1 km grid cell.

SMAP 9 km estimates to JULES 1 km estimates could be improved upon and that more signal may be coming from the centre of the satellite pixel, so that we could weight these JULES model pixels more highly within the observation operator. In future work it may also be beneficial to build towards a full radiative transfer scheme on top of JULES to assimilate the raw brightness temperature observations from the SMAP satellite to increase the representativty between the observations and the model and reduce sources of bias that may be introduced by the use of ancillary data in the soil moisture retrieval. Other studies utilising different land surface models have have shown this works well (Han et al., 2014; Yang et al., 2016; Lievens et al., 2017) The COSMOS–UK observations could also be used within the data assimilation algorithm, rather than just acting as validation, to capture information on another spatial scale. Much work would be needed here to process and organise site level driving data and understand the different characteristics of each site before combining these observations with the JULES land surface model.

In this paper we have focused on the optimisation of pedotransfer function parameters to improve estimates of water balance from land surface models. In other regions across the globe where underlying soil texture maps are highly uncertain it may be necessary to also consider optimising estimates of soil properties per-grid cell, given satellite and in-situ observations (Pinnington et al., 2018). This could further increase the skill of estimates in problematic areas. There is also the opportunity to incorporate other streams of observations into the data assimilation procedure. For example the use of stream flow data could give us a powerful integrated constraint on land surface model estimates of water balance and run-off (Abbaszadeh et al., 2020) .

	Correlation		ubRMSE		RMSE	
Site	Prior	Posterior	Prior	Posterior	Prior	Posterior
Bunny Park	0.86	0.89	0.02	0.02	0.07	0.04
Cardington	0.85	0.91	0.05	0.03	0.06	0.03
Elmsett	0.81	0.82	0.04	0.04	0.16	0.17
Euston	0.90	0.92	0.04	0.04	0.05	0.04
Fincham	0.83	0.85	0.02	0.02	0.19	0.13
Loddington	0.45	0.79	0.06	0.04	0.39	0.31
Morley	0.86	0.89	0.03	0.03	0.03	0.03
Redmere	0.33	0.35	0.08	0.08	0.43	0.43
Rothamsted	0.85	0.89	0.03	0.03	0.05	0.07
Stoughton	0.30	0.76	0.05	0.04	0.24	0.13
Waddesdon	0.63	0.87	0.07	0.05	0.27	0.19
All Sites	0.70	0.81	0.045	0.038	0.18	0.14

Table 3. Summary statistics for comparison of JULES-CHESS soil moisture estimates to COSMOS probe observations over the experiment period. Over all sites we find a 16% increase in correlation, 16% reduction in ubRMSE and 22% reduction in RMSE after performing the calibration using LAVENDAR.

Flux tower observations of latent and sensible heat could also provide useful constraints on assimilation outputs. Within the Hydro–JULES project work is being undertaken to improve the representation of hydrological processes at different scales, especially lateral soil water flow and groundwater. The development of the new JULES groundwater component will allow for the use of observations from the Gravity Recovery and Climate Experiment (GRACE) satellites (Tapley et al., 2004) which have the ability to monitor changes in the Earth's underground water storage. It will be informative to re–run this parameter estimation experiment again as new processes are added to the model to understand the effect on the retrieved pedotransfer function parameters. We will then be able to see where we might be over–fitting these parameters to account for current structural deficiencies within the model (such as the current lack of a groundwater model).

5 Conclusions

We have presented novel methods for calibrating pedotranfer functions used to create the soil parameter ancillaries of a land surface model by using satellite data from the NASA SMAP mission. After the retrieval of an optimized parameter set, using new hybrid data assimilation techniques, we find an average 20% reduction in error for JULES model estimates of soil moisture when compared to SMAP satellite estimates. There are still areas which remain problematic such as working over urban locations and peatlands. These will require additional modelling efforts and new model components. The resultant posterior pedotransfer functions also improve the prediction of soil moisture for the JULES land surface model when compared to independent in-situ

estimates from the COSMOS–UK network. At 11 COSMOS–UK research sites distributed across the experiment domain we find an average 16% increase in correlation, 16% reduction in ubRMSE and a 22% reduction in RMSE for the posterior pedotransfer functions compared to the prior.

Code availability. The code used in experiments is available from the MetOffice JULES repository (https://code.metoffice.gov.uk/trac/jules) under Rose suite number u-bq357. The LAVENDAR data assimilation first release is available here: https://github.com/pyearthsci/lavendar (last access: 20 February 2019).

Appendix A: Computing the posterior ensemble

In this appendix we summarise the process to get the analysis (or posterior) ensemble of extended state variables (variables and parameters) In the case of this paper the variables and parameters correspond to the 15 PTF parameters in Table 1. The following steps are a recapitulation and continuation of the equations in Pinnington et al. (2018).

Let us start with a background ensemble of N_e joint state-parameter vectors:

$$\mathbf{X}_b = \left[\mathbf{x}_b^1, \mathbf{x}_b^2, \ldots, \mathbf{x}_b^{N_e}\right]. \tag{A1}$$

In our experiments each $\mathbf{x}_b^i$ corresponds to a unique set of 15 PTF parameters ($\mathbf{x}_b^i = (\phi_a^i, \phi_b^i, \ldots, \phi_o^i)$) and $N_e = 50$. We can define the sample background (or prior) mean as:

$$\bar{\mathbf{x}}_b = \frac{1}{N_e}\sum_{n=1}^{N_e}\mathbf{x}_b^n \tag{A2}$$

and sample background perturbation matrix as

$$\mathbf{X}_b' = \frac{1}{\sqrt{N_e - 1}}\left[\mathbf{x}_b^1 - \bar{\mathbf{x}}_b, \mathbf{x}_b^2 - \bar{\mathbf{x}}_b, \ldots, \mathbf{x}_b^{N_e} - \bar{\mathbf{x}}_b\right]. \tag{A3}$$

The ensemble background error covariance matrix defined by

$$\mathbf{P}_b = \mathbf{X}_b'\mathbf{X}_b'^T. \tag{A4}$$

To reduce the difficulty in finding the ensemble analysis mean, we use an incremental and pre-conditioned algorithm. Incremental means that we express the analysis mean which is a perturbation from the background mean, i.e.:

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b + \delta\mathbf{x}. \tag{A5}$$

The pre-conditioned part means that the departure $\delta\mathbf{x}$ can be written by a control variable pre-multiplied by a conditioning matrix. In particular we choose the departure vector to be written as a linear combination of the background ensemble of perturbations, i.e.

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b + \mathbf{X}_b'\mathbf{w}_a, \tag{A6}$$

where $\mathbf{w}_a$ is a vector of weights, which becomes the object we are solving for in the estimation process. This formulation has been used in several formulations, starting with Bishop et al (2001) and Wang et al (2004). We do not use localisation in this work, but in the presence of localisation it would be applied in the manner of Hunt et al (2007). This vector of weights is the minimiser of a cost function which can be written in ensemble space as:

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\bar{\mathbf{x}}_b) - \hat{\mathbf{y}})^T\hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\bar{\mathbf{x}}_b) - \hat{\mathbf{y}}) \tag{A7}$$

with gradient

$$\nabla J(\mathbf{w}) = \mathbf{w} + (\hat{\mathbf{H}}\mathbf{X}'_b)^T\hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}}), \tag{A8}$$

where $\hat{\mathbf{y}}$ are the observations for the whole time-window and spatial domain (here 2016 SMAP observations over the East of England, with units $m^3\ m^{-3}$), $\hat{\mathbf{H}}$ and $\hat{\mathbf{h}}$ are the linearised and non-linear observation operator respectively (here the JULES model, which includes both a time integration and conversion into observation space to match the SMAP observations) and $\hat{\mathbf{R}}$ is the observation error covariance matrix (here containing the error estimates for the assimilated SMAP observations).

In practice we do not compute the linearised version of JULES. Instead one can define statistics in the observation space in the following manner. The background ensemble of N_e joint state-parameter vectors in observation space is obtained by applying the observation operator to each ensemble member:

$$\mathbf{Y}_b = \left[\mathbf{y}_b^1 = \hat{\mathbf{h}}\left(\mathbf{x}_b^1\right), \mathbf{y}_b^2 = \hat{\mathbf{h}}\left(\mathbf{x}_b^2\right), \ldots, \mathbf{y}_b^{N_e} = \hat{\mathbf{h}}\left(\mathbf{x}_b^{N_e}\right)\right] \tag{A9}$$

The sample background mean in observation space is:

$$\bar{\mathbf{y}}_b = \frac{1}{N_e}\sum_{n=1}^{N_e}\mathbf{y}_b^n \tag{A10}$$

and the sample background perturbation matrix in observation space is:

$$\mathbf{Y}'_b = \frac{1}{\sqrt{N_e - 1}}\left[\mathbf{y}_b^1 - \bar{\mathbf{y}}_b, \mathbf{y}_b^2 - \bar{\mathbf{y}}_b, \ldots, \mathbf{y}_b^{N_e} - \bar{\mathbf{y}}_b\right] \tag{A11}$$

Using these considerations, (A7) and (A8) become (approximately):

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\mathbf{Y}'_b\mathbf{w} + \bar{\mathbf{y}}_b - \hat{\mathbf{y}})^T\hat{\mathbf{R}}^{-1}(\mathbf{Y}'_b\mathbf{w} + \bar{\mathbf{y}}_b - \hat{\mathbf{y}}) \tag{A12}$$

and

$$\nabla J(\mathbf{w}) = \mathbf{w} + (\mathbf{Y}'_b)^T\hat{\mathbf{R}}^{-1}(\mathbf{Y}'_b\mathbf{w} + \bar{\mathbf{y}}^b - \hat{\mathbf{y}}). \tag{A13}$$

Computing the minimum of the cost function (A12) using gradient (A13) yields the maximum-a-posteriori estimate $\mathbf{w}_a$ which inserting into equation (A6) gives us the maximum-a-posteriori estimate to the parameter and/or state variables $\mathbf{x}_a$. The analysis error covariance matrix ($\mathbf{P}_a$) is given by (Evensen, 2003):

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\hat{\mathbf{H}})\mathbf{P}_b \implies \mathbf{X}'_a\mathbf{X}'^T_a = (\mathbf{I} - \mathbf{K}\hat{\mathbf{H}})\mathbf{X}'_b\mathbf{X}'^T_b \tag{A14}$$

where $\mathbf{K}$ is the Kalman gain matrix and

$$(\mathbf{I}-\mathbf{K}\hat{\mathbf{H}}) = (\mathbf{I}+\hat{\mathbf{H}}\mathbf{X}_b{}'^{T}\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{X}_b)'^{-1} \approx (\mathbf{I}+\mathbf{Y}_b{}'^{T}\hat{\mathbf{R}}^{-1}\mathbf{Y}'_b)^{-1}. \tag{A15}$$

Then

$$\mathbf{X}'_a\mathbf{X}'^{T}_a = \mathbf{X}'_b(\mathbf{I}-\mathbf{K}\hat{\mathbf{H}})\mathbf{X}'^{T}_b \implies \mathbf{X}'_a = \mathbf{X}'_b(\mathbf{I}+\mathbf{Y}_b{}'^{T}\hat{\mathbf{R}}^{-1}\mathbf{Y}'_b)^{-\frac{1}{2}} \tag{A16}$$

i.e. the analysis ensemble of perturbations can be obtained by a right multiplication of the background ensemble of perturbations times a matrix of weights defined as:

$$\mathbf{W}_a = (\mathbf{I}+\mathbf{Y}_b{}'^{T}\hat{\mathbf{R}}^{-1}\mathbf{Y}'_b)^{-\frac{1}{2}}. \tag{A17}$$

In our case the matrix square root is computed via Cholesky decomposition. Finally the posterior ensemble of N_e parameter/state vectors ($\mathbf{X}_a$) is constructed as

$$\mathbf{X}_a = \left[\mathbf{x}^a+\mathbf{X}'_{a,1}, \mathbf{x}^a+\mathbf{X}'_{a,2}, \ldots, \mathbf{x}^a+\mathbf{X}'_{a,N_e}\right]. \tag{A18}$$

This posterior parameter ensemble and corresponding set of JULES runs can then be used to provide uncertainty estimates on our posterior model predictions and can also be used in future calibration studies or as an ensemble forecast for state estimation.

Author contributions. EP designed the data assimilation system and conducted all experiments with input from all co-authors. JA contributed to the underlying mathematical framework. EC wrote the algorithm to compare JULES soil moisture to cosmic-ray probe measurements. ER processed the HWSD and built the initial system for relating soil textural information to the parameters of JULES used in these experiments. EP prepared the manuscript with input from all co-authors.

Competing interests. No competing interests present.

Acknowledgements. This work was funded by the UK Natural Environment Research Council's Hydro–JULES project (NE/S017380/1). TQ and JA contribution was funded via the UK National Centre for Earth Observation (NCEO) at University of Reading (NCEO grant number: nceo020004). The authors gratefully acknowledge the provision by UKCEH of hydrometeorological and soil data collected by the COSMOS-UK project.

References

Abbaszadeh, P., Gavahi, K., and Moradkhani, H.: Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting, Advances in Water Resources, 145, 103 721, https://doi.org/https://doi.org/10.1016/j.advwatres.2020.103721, 2020.

Anderson, J. L. and Anderson, S. L.: A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts, Monthly Weather Review, 127, 2741–2758, https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2, 1999.

Asfaw, D., Black, E., Brown, M., Nicklin, K. J., Otu-Larbi, F., Pinnington, E., Challinor, A., Maidment, R., and Quaife, T.: TAMSAT-ALERT v1: a new framework for agricultural decision support, Geoscientific Model Development, 11, 2353–2371, https://doi.org/10.5194/gmd-11-2353-2018, 2018.

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, determination of soil moisture: Measurements and theoretical approaches, 2014.

Baatz, R., Hendricks Franssen, H.-J., Han, X., Hoar, T., Bogena, H. R., and Vereecken, H.: Evaluation of a cosmic-ray neutron sensor network for improved land surface model prediction, Hydrology and Earth System Sciences, 21, 2509–2530, https://doi.org/10.5194/hess-21-2509-2017, https://hess.copernicus.org/articles/21/2509/2017/, 2017.

Bateni, S. M. and Entekhabi, D.: Relative efficiency of land surface energy balance components, Water Resources Research, 48, https://doi.org/10.1029/2011WR011357, 2012.

Beljaars, A. C. M., Viterbo, P., Miller, M. J., and Betts, A. K.: The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies, Monthly Weather Review, 124, 362–383, https://doi.org/10.1175/1520-0493(1996)124<0362:TAROTU>2.0.CO;2, 1996.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geoscientific Model Development, 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, 2011.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrology and Earth System Sciences, 5, 1–12, https://doi.org/10.5194/hess-5-1-2001, https://hess.copernicus.org/articles/5/1/2001/, 2001.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, https://doi.org/10.1002/hyp.3360060305, 1992.

Bocquet, M. and Sakov, P.: Joint state and parameter estimation with an iterative ensemble Kalman smoother, Nonlinear Processes in Geophysics, 20, 803–818, https://doi.org/10.5194/npg-20-803-2013, 2013.

Bogena, H. R., Huisman, J. A., Güntner, A., Hübner, C., Kusche, J., Jonard, F., Vey, S., and Vereecken, H.: Emerging methods for noninvasive sensing of soil moisture dynamics from field to catchment scale: a review, WIREs Water, 2, 635–647, https://doi.org/10.1002/wat2.1097, 2015.

Bormann, N., Bonavita, M., Dragani, R., Eresmaa, R., Matricardi, M., and McNally, A.: Enhancing the impact of IASI observations through an updated observation error covariance matrix, https://doi.org/10.21957/gq8j2gjp7, https://www.ecmwf.int/node/8292, 2015.

Botto, A., Belluco, E., and Camporese, M.: Multi-source data assimilation for physically based hydrological modeling of an experimental hillslope, Hydrology and Earth System Sciences, 22, 4251–4266, https://doi.org/10.5194/hess-22-4251-2018, https://hess.copernicus.org/articles/22/4251/2018/, 2018.

Brooks, R. H. and Corey, A. T.: Hydraulic properties of porous media and their relation to drainage design, Transactions of the ASAE, 7, 26–0028, 1964.

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., and Aida, K.: Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation, Remote Sensing of Environment, 214, 1 – 13, https://doi.org/https://doi.org/10.1016/j.rse.2018.05.008, 2018.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, Geoscientific Model Development, 4, 701–722, https://doi.org/10.5194/gmd-4-701-2011, 2011.

Clark, P., Roberts, N., Lean, H., Ballard, S. P., and Charlton-Perez, C.: Convection-permitting models: a step-change in rainfall forecasting, Meteorological Applications, 23, 165–181, https://doi.org/10.1002/met.1538, 2016.

Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., Cosh, M., Dunbar, R., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernàndez, J., Gonzàlez-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J., Wu, X., Monerris, A., O'Neill, P., Entekhabi, D., Njoku, E., and Yueh, S.: Validation of SMAP surface soil moisture products with core validation sites, Remote Sensing of Environment, 191, 215 – 231, https://doi.org/https://doi.org/10.1016/j.rse.2017.01.021, http://www.sciencedirect.com/science/article/pii/S0034425717300329, 2017.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils, Water Resources Research, 20, 682–690, https://doi.org/10.1029/WR020i006p00682, 1984.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Quarterly Journal of the Royal Meteorological Society, 120, 1367–1387, https://doi.org/10.1002/qj.49712051912, 1994.

De Lannoy, G. J. M. and Reichle, R. H.: Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model, Hydrology and Earth System Sciences, 20, 4895–4911, https://doi.org/10.5194/hess-20-4895-2016, 2016.

Desilets, D. and Zreda, M.: Footprint diameter for a cosmic-ray soil moisture probe: Theory and Monte Carlo simulations, Water Resources Research, 49, 3566–3575, https://doi.org/10.1002/wrcr.20187, 2013.

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, Geophysical Research Letters, 39, https://doi.org/10.1029/2011GL050655, 2012.

Duygu, M. B. and Akyürek, Z.: Using Cosmic-Ray Neutron Probes in Validating Satellite Soil Moisture Products and Land Surface Models, Water, 11, 1362, https://doi.org/10.3390/w11071362, http://dx.doi.org/10.3390/w11071362, 2019.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The Soil Moisture Active Passive (SMAP) Mission, Proceedings of the IEEE, 98, 704–716, 2010.

Evans, J. G., Ward, H. C., Blake, J. R., Hewitt, E. J., Morrison, R., Fry, M., Ball, L. A., Doughty, L. C., Libre, J. W., Hitt, O. E., Rylett, D., Ellis, R. J., Warwick, A. C., Brooks, M., Parkes, M. A., Wright, G. M. H., Singer, A. C., Boorman, D. B., and Jenkins, A.: Soil water content

in southern England derived from a cosmic-ray soil moisture observing system – COSMOS-UK, Hydrological Processes, 30, 4987–4999, https://doi.org/10.1002/hyp.10929, 2016.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Evensen, G.: Analysis of iterative ensemble smoothers for solving inverse problems, Computational Geosciences, 22, 885–908, https://doi.org/10.1007/s10596-018-9731-y, https://doi.org/10.1007/s10596-018-9731-y, 2018.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H., Verelst, L., and Wiberg, D.: Global agro-ecological zones assessment for agriculture (GAEZ 2008), IIASA, Laxenburg, Austria and FAO, Rome, Italy, 10, http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/, [Accessed 2020-04-28], 2008.

Fowler, A. M., Dance, S. L., and Waller, J. A.: On the interaction of observation and prior error correlations in data assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 48–62, https://doi.org/10.1002/qj.3183, 2018.

Han, X., Franssen, H.-J. H., Montzka, C., and Vereecken, H.: Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations, Water Resources Research, 50, 6081–6105, https://doi.org/10.1002/2013WR014586, 2014.

Hauser, M., Orth, R., and Seneviratne, S. I.: Investigating soil moisture–climate interactions with prescribed soil moisture experiments: an assessment with the Community Earth System Model (version 1.2), Geoscientific Model Development, 10, 1665–1677, https://doi.org/10.5194/gmd-10-1665-2017, 2017.

Hilton, F., Collard, A., Guidard, V., Randriamampianina, R., and Schwaerz, M.: Assimilation of IASI radiances at European NWP centres, https://www.ecmwf.int/node/15331, 2009.

Houtekamer, P. L. and Mitchell, H. L.: Data Assimilation Using an Ensemble Kalman Filter Technique, Monthly Weather Review, 126, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2, 1998.

Howes, K. E., Fowler, A. M., and Lawless, A. S.: Accounting for model error in strong-constraint 4D-Var data assimilation, Quarterly Journal of the Royal Meteorological Society, 143, 1227–1240, https://doi.org/10.1002/qj.2996, 2017.

Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., Shibata, A., Miura, M., Inahata, H., and Spencer, R. W.: The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies, IEEE Transactions on Geoscience and Remote Sensing, 41, 184–194, https://doi.org/10.1109/TGRS.2002.808331, 2003.

Kerr, Y. H., Waldteufel, P., Wigneron, J. ., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, IEEE Transactions on Geoscience and Remote Sensing, 39, 1729–1735, 2001.

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., and Zacharias, S.: Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons, Water Resources Research, 51, 5772–5790, https://doi.org/10.1002/2015WR017169, 2015.

Kolassa, J., Reichle, R., and Draper, C.: Merging active and passive microwave observations in soil moisture data assimilation, Remote Sensing of Environment, 191, 117–130, 2017.

Li, C., Lu, H., Yang, K., Han, M., Wright, J., Chen, Y., Yu, L., Xu, S., Huang, X., and Gong, W.: The Evaluation of SMAP Enhanced Soil Moisture Products Using High-Resolution Model Simulations and In-Situ Observations on the Tibetan Plateau, Remote Sensing, 10, 535, https://doi.org/10.3390/rs10040535, 2018.

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker, J. P., and Wagner, W.: Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates, Geophysical Research Letters, 44, 6145–6153, https://doi.org/10.1002/2017GL073904, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL073904, 2017.

Liu, Q., Reichle, R. H., Bindlish, R., Cosh, M. H., Crow, W. T., de Jeu, R., De Lannoy, G. J. M., Huffman, G. J., and Jackson, T. J.: The Contributions of Precipitation and Soil Moisture Observations to the Skill of Soil Moisture Estimates in a Land Data Assimilation System, Journal of Hydrometeorology, 12, 750–765, https://doi.org/10.1175/JHM-D-10-05000.1, 2011.

Lorenc, A. C. and Rawlins, F.: Why does 4D-Var beat 3D-Var?, Quarterly Journal of the Royal Meteorological Society, 131, 3247–3257, https://doi.org/10.1256/qj.05.85, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.85, 2005.

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, Geoscientific Model Development, 7, 711–723, https://doi.org/10.5194/gmd-7-711-2014, https://gmd.copernicus.org/articles/7/711/2014/, 2014.

Martinez-de la Torre, A., Blyth, E., and Robinson, E.: Water, carbon and energy fluxes simulation for Great Britain using the JULES Land Surface Model and the Climate Hydrology and Ecology research Support System meteorology dataset (1961-2015) [CHESS-land], https://doi.org/10.5285/c76096d6-45d4-4a69-a310-4c67f8dcf096, https://doi.org/10.5285/c76096d6-45d4-4a69-a310-4c67f8dcf096, 2018.

Martínez-de la Torre, A., Blyth, E. M., and Weedon, G. P.: Using observed river flow data to improve the hydrological functioning of the JULES land surface model (vn4.3) used for regional coupled modelling in Great Britain (UKC2), Geoscientific Model Development, 12, 765–784, https://doi.org/10.5194/gmd-12-765-2019, https://www.geosci-model-dev.net/12/765/2019/, 2019.

Maurer, E. and Lettenmaier, D.: Potential effects of long-lead hydrologic predictability on Missouri River main-stem reservoirs, JOURNAL OF CLIMATE, 17, 174–186, https://doi.org/10.1175/1520-0442(2004)017<0174:PEOLHP>2.0.CO;2, 2004.

Minamide, M. and Zhang, F.: Adaptive Observation Error Inflation for Assimilating All-Sky Satellite Radiance, Monthly Weather Review, 145, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1, 2017.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020–8040, https://doi.org/10.1002/2017WR020401, 2017.

Montzka, C., Moradkhani, H., WeihermÃŒller, L., Franssen, H.-J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, Journal of Hydrology, 399, 410 – 421, https://doi.org/https://doi.org/10.1016/j.jhydrol.2011.01.020, http://www.sciencedirect.com/science/article/pii/S002216941100045X, 2011.

Montzka, C., Bogena, H. R., Zreda, M., Monerris, A., Morrison, R., Muddu, S., and Vereecken, H.: Validation of Spaceborne and Modelled Surface Soil Moisture Products with Cosmic-Ray Neutron Probes, Remote Sensing, 9, https://doi.org/10.3390/rs9020103, 2017.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state–parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, 28, 135 – 147, https://doi.org/https://doi.org/10.1016/j.advwatres.2004.09.002, 2005.

Nearing, G. S., Moran, M. S., Thorp, K. R., Collins, C. D. H., and Slack, D. C.: Likelihood parameter estimation for calibrating a soil moisture model using radar bakscatter, Remote Sensing of Environment, 114, 2564 – 2574, https://doi.org/https://doi.org/10.1016/j.rse.2010.05.031, 2010.

Peng, J., Pinnington, E., Robinson, E., Evans, J., Quaife, T., Harris, P., Blyth, E., and S., D.: High resolution soil moisture estimation and evaluation from Earth observation, In preparation, 2020.

Pinnington, E., Quaife, T., and Black, E.: Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the JULES land surface model, Hydrology and Earth System Sciences, 22, 2575–2588, https://doi.org/10.5194/hess-22-2575-2018, 2018.

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., and Scoby, D.: The Land Variational Ensemble Data Assimilation Framework: LAVENDAR v1.0.0, Geoscientific Model Development, 13, 55–69, https://doi.org/10.5194/gmd-13-55-2020, 2020.

Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I., Nichols, N. K., Wilkinson, M., and Quaife, T. L.: Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using Four-dimensional Variational data assimilation, Agricultural and Forest Meteorology, 228-229, 299 – 314, https://doi.org/10.1016/j.agrformet.2016.07.006, 2016.

Pitman, A. J., Henderson-Sellers, A., Desborough, C. E., Yang, Z. L., Abramopoulos, F., Boone, A., Dickinson, R. E., Gedney, N., Koster, R., Kowalczyk, E., Lettenmaier, D., Liang, X., Mahfouf, J. F., Noilhan, J., Polcher, J., Qu, W., Robock, A., Rosenzweig, C., Schlosser, C. A., Shmakin, A. B., Smith, J., Suarez, M., Verseghy, D., Wetzel, P., Wood, E., and Xue, Y.: Key results and implications from phase 1(c) of the Project for Intercomparison of Land-surface Parametrization Schemes, Climate Dynamics, 15, 673–684, https://doi.org/10.1007/s003820050309, 1999.

Rasmy, M., Koike, T., Boussetta, S., Lu, H., and Li, X.: Development of a Satellite Land Data Assimilation System Coupled With a Mesoscale Model in the Tibetan Plateau, IEEE Transactions on Geoscience and Remote Sensing, 49, 2847–2862, 2011.

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., Crow, W., Jackson, T. J., Jones, L. A., Kimball, J. S., Koster, R. D., Mahanama, S. P., Smith, E. B., Berg, A., Bircher, S., Bosch, D., Caldwell, T. G., Cosh, M., Gonzàlez-Zamora, A., Holifield Collins, C. D., Jensen, K. H., Livingston, S., Lopez-Baeza, E., Martínez-Fernàndez, J., McNairn, H., Moghaddam, M., Pacheco, A., Pellarin, T., Prueger, J., Rowlandson, T., Seyfried, M., Starks, P., Su, Z., Thibeault, M., van der Velde, R., Walker, J., Wu, X., and Zeng, Y.: Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements, Journal of Hydrometeorology, 18, 2621–2645, https://doi.org/10.1175/JHM-D-17-0063.1, https://doi.org/10.1175/JHM-D-17-0063.1, 2017.

Ridler, M.-E., Zhang, D., Madsen, H., Kidmose, J., Refsgaard, J. C., and Jensen, K. H.: Bias-aware data assimilation in integrated hydrological modelling, Hydrology Research, 49, 989–1004, https://doi.org/10.2166/nh.2017.117, 2017.

Robinson, E., Blyth, E., Clark, D., Comyn-Platt, E., Finch, J., and Rudd, A.: Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2, https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900, https://doi.org/10.5285/b745e7b1-626c-4ccc-ac27-56582e77b900, 2017.

Rosolem, R., Hoar, T., Arellano, A., Anderson, J. L., Shuttleworth, W. J., Zeng, X., and Franz, T. E.: Translating aboveground cosmic-ray neutron intensity to high-frequency soil moisture profiles at sub-kilometer scale, Hydrology and Earth System Sciences, 18, 4363–4379, https://doi.org/10.5194/hess-18-4363-2014, 2014.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, https://doi.org/10.1029/2008WR007327, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008WR007327, 2010.

Sawada, Y. and Koike, T.: Simultaneous estimation of both hydrological and ecological parameters in an ecohydrological model by assimilating microwave signal, Journal of Geophysical Research: Atmospheres, 119, 8839–8857, https://doi.org/10.1002/2014JD021536, 2014JD021536, 2014.

Schaap, M. G., Nemes, A., and van Genuchten, M. T.: Comparison of Models for Indirect Estimation of Water Retention and Available Water in Surface Soils, Vadose Zone Journal, 3, 1455–1463, https://doi.org/10.2136/vzj2004.1455, 2004.

Shuttleworth, J., Rosolem, R., Zreda, M., and Franz, T.: The COsmic-ray Soil Moisture Interaction Code (COSMIC) for use in data assimilation, Hydrology and Earth System Sciences, 17, 3205–3217, https://doi.org/10.5194/hess-17-3205-2013, 2013.

Stanley, S., Antoniou, V., Ball, L., Bennett, E., Blake, J., Boorman, D., Brooks, M., Clarke, M., Cooper, H., Cowan, N., Evans, J., Farrand, P., Fry, M., Hitt, O., Jenkins, A., Kral, F., Lord, W., Morrison, R., Nash, G., Rylett, D., Scarlett, P., Swain, O., Thornton, J., Trill, E., Warwick, A., and Winterbourn, J.: Daily and sub-daily hydrometeorological and soil data (2013-2017) [COSMOS-UK], https://doi.org/10.5285/a6012796-291c-4fd6-a7ef-6f6ed0a6cfa5, 2019.

Stewart, L. M., Dance, S. L., Nichols, N. K., Eyre, J. R., and Cameron, J.: Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system†, Quarterly Journal of the Royal Meteorological Society, 140, 1236–1244, https://doi.org/10.1002/qj.2211, 2014.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE Measurements of Mass Variability in the Earth System, Science, 305, 503–505, https://doi.org/10.1126/science.1099192, https://science.sciencemag.org/content/305/5683/503, 2004.

Thiemann, M., Trosset, M., Gupta, H., and Sorooshian, S.: Bayesian recursive parameter estimation for hydrologic models, Water Resources Research, 37, 2521–2535, https://doi.org/10.1029/2000WR900405, 2001.

Tifafi, M., Guenet, B., and Hatté, C.: Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France, Global Biogeochemical Cycles, 32, 42–56, https://doi.org/10.1002/2017GB005678, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GB005678, 2018.

Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., and Tóth, G.: New generation of hydraulic pedotransfer functions for Europe, European Journal of Soil Science, 66, 226–238, https://doi.org/10.1111/ejss.12192, 2015.

van Genuchten, M. T.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, Soil Science Society of America Journal, 44, 892–898, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y. A., Padarian, J., Schaap, M. G., Tóth, B., Verhoef, A., Vanderborght, J., van der Ploeg, M. J., Weihermüller, L., Zacharias, S., Zhang, Y., and Vereecken, H.: Pedotransfer Functions in Earth System Science: Challenges and Perspectives, Reviews of Geophysics, 55, 1199–1256, https://doi.org/10.1002/2017RG000581, 2017.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research, 39, https://doi.org/10.1029/2002WR001642, 2003.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., and Rubel, F.: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications, Meteorologische Zeitschrift, 22, 5–33, https://doi.org/doi:10.1127/0941-2948/2013/0399, 2013.

Walker, J. P., Willgoose, G. R., and Kalma, J. D.: In situ measurement of soil moisture: a comparison of techniques, Journal of Hydrology, 293, 85 – 99, https://doi.org/https://doi.org/10.1016/j.jhydrol.2004.01.008, 2004.

Wang, P., Li, J., Li, Z., Lim, A. H. N., Li, J., and Goldberg, M. D.: Impacts of Observation Errors on Hurricane Forecasts When Assimilating Hyperspectral Infrared Sounder Radiances in Partially Cloudy Skies, Journal of Geophysical Research: Atmospheres, 124, 10 802–10 813, https://doi.org/10.1029/2019JD031029, 2019.

Wösten, J., Lilly, A., Nemes, A., and Le Bas, C.: Development and use of a database of hydraulic properties of European soils, Geoderma, 90, 169 – 185, https://doi.org/https://doi.org/10.1016/S0016-7061(98)00132-3, 1999.

Yang, K., Zhu, L., Chen, Y., Zhao, L., Qin, J., Lu, H., Tang, W., Han, M., Ding, B., and Fang, N.: Land surface model calibration through microwave data assimilation for improving soil moisture simulations, Journal of Hydrology, 533, 266 – 276, https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.12.018, 2016.

Zhang, R., Kim, S., and Sharma, A.: A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes, Remote Sensing of Environment, 223, 82 – 94, https://doi.org/https://doi.org/10.1016/j.rse.2019.01.015, 2019.

Zheng, D., Li, X., Wang, X., Wang, Z., Wen, J., van der Velde, R., Schwank, M., and Su, Z.: Sampling depth of L-band radiometer measurements of soil moisture and freeze-thaw dynamics on the Tibetan Plateau, Remote Sensing of Environment, 226, 16 – 25, https://doi.org/https://doi.org/10.1016/j.rse.2019.03.029, 2019.

Zreda, M., Desilets, D., Ferré, T. P. A., and Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL035655, 2008.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R.: COSMOS: the COsmic-ray Soil Moisture Observing System, Hydrology and Earth System Sciences, 16, 4079–4099, https://doi.org/10.5194/hess-16-4079-2012, 2012.