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

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

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
25 (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
30 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
35 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 variabil-
40 ity 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
45 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,
50 even when land surface models are run at a high resolution (∼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 radius of ∼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
55 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 predic-
60 tions. These techniques can either be used to update the soil moisture state 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 to update the parameters of the model and 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 always takes into account the relative uncertainties given to both model and observed estimates to find a maximum-a-posteriori estimate. Previous
65 studies have used data assimilation to update the parameters of land surface models. 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.

70 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
75 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 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. After assimilation we found a 20% reduction in error for the JULES model using the new calibration of the Tóth et al. (2015)
80 pedotransfer functions when compared to the SMAP observations. As independent validation we also found an average 22% reduction in error and 16% increase in correlation over 11 COSMOS probe sites in the model domain.

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
85 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

90 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. 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

95 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 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.

2.2 Pedotransfer functions

The JULES model implements both the Brooks and Corey (1964) and the van Genuchten (1980) soil models. The JULES

100 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:

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

105 where θ_{res} is the residual soil moisture, θ_{sat} the saturated soil moisture, α and $(N - 1)$ parameters of the van Genuchten (1980) soil model, K_{sat} the saturated hydraulic conductivity, $\phi_a, \ldots, \phi_o$ 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 bulk density (ρ), fraction of clay and silt (f_{clay}, f_{silt}) and the cation exchange capacity (CEC) 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

110 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

Hydrology and
Earth System
Sciences
Discussions

115 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.

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.

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.

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

120   was designed 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 mal-
        functioned. 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 with a prescribed error of 0.05 m^3 m^{-3}. This error is a slight inflation
        of the SMAP baseline aim of 0.04 m^3 m^{-3}. However, other studies have found slightly higher values for the error in SMAP

125   observations (Li et al., 2018; Zhang et al., 2019). As part of the retrieval procedure SMAP relies on some ancillary informa-
        tion, 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 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

130   minimised. 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 1.
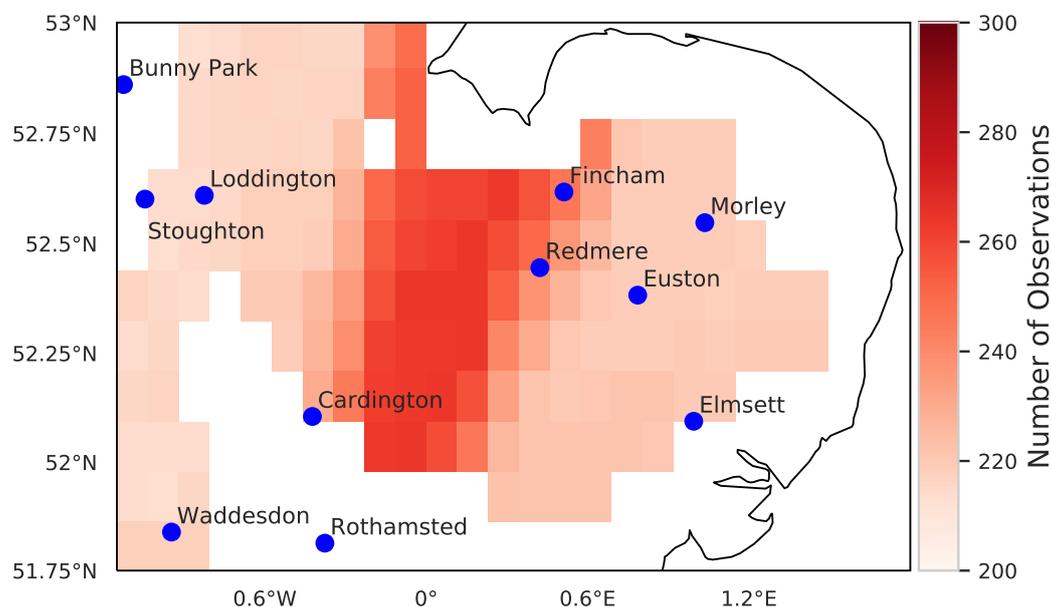

Figure 1. Location of COSMOS probes and number of SMAP observations assimilated in experiment period (2016). No colour corresponds
to no observations being assimilated in that location due to low quality retrieval or surface flag.

2.4   COSMOS–UK Observations

The COSMOS–UK network has been producing observations of soil moisture and other meteorological variables at an ex-
panding number of stations (currently 52) since 2013 (Stanley et al., 2019). For the area of interest in this paper we have 11

135   stations available to us with data for the relevant time period, see Figure 1. Some of these stations may not be representative of

JULES model estimates due to the current setup of JULES not considering some processes (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). The COSMOS–UK network report the representative depth of each soil moisture measurement. To make a fair comparison between the COSMOS–UK and JULES soil moisture estimates we have constructed a variable depth algorithm for JULES which takes

140 a weighted average of the different soil layers of the model given the relative depth of the COSMOS–UK observation.

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 involves running an ensemble of JULES models, with each model in the ensemble utilising a distinct soil ancillary data–set. Each ancillary file is created by sampling from the normal distribution defined by taking a

145 10% standard deviation around the prior mean in table 1 then using the unique set of sampled parameters within equations (1) applied to the HWSD for the domain. 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

150 observation operator which aggregates the JULES grid cells within each SMAP pixel. This will introduce an additional source of representativity error alongside the observational error of SMAP and the inherent model error within JULES. For this reason and due to the large number of observations assimilated in our years assimilation window (28698) we inflate the specified observational error by a factor of four. The method of inflation is common practice in meteorological data assimilation when assimilating streams of data with vast numbers of observations (Minamide and Zhang, 2017). It is largely used to account for

155 the fact that the representation of observational errors are often over simplified and do not account for any error correlations that may be present in time, space or between satellite bands (Stewart et al., 2008; Pinnington et al., 2016).

The LAVENDAR framework takes the JULES simulated soil moisture for each ensemble member (applying the SMAP spatially aggregated observation operator) and the SMAP observations and performs the method of Four-Dimensional Ensemble Variational (4DEnVar) data assimilation to find a new 'most likely' set of parameters for the pedotransfer functions given the

160 uncertainties in both the observations and prior parameter estimates. This new set of most likely parameters is then used in a method similar to that of the Iterative Ensemble Kalman Smoother (IEnKS) (Bocquet and Sakov, 2013) to find a posterior set of 50 parameter vectors, for more information on how this is calculated see the Appendix A. This subsequent posterior ensemble 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.

165 ## 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
170 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

175 In Figure 2 we show the difference between the mean soil moisture in 2016 for the prior and posterior JULES ensemble. The
grid cells that are darker blue correspond to the posterior ensemble being wetter after assimilation and grid cells that are darker
red correspond to the posterior being drier. 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. 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 to behave differently rather than
180 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. Figure 2 also allows us to see the high–resolution of the JULES model
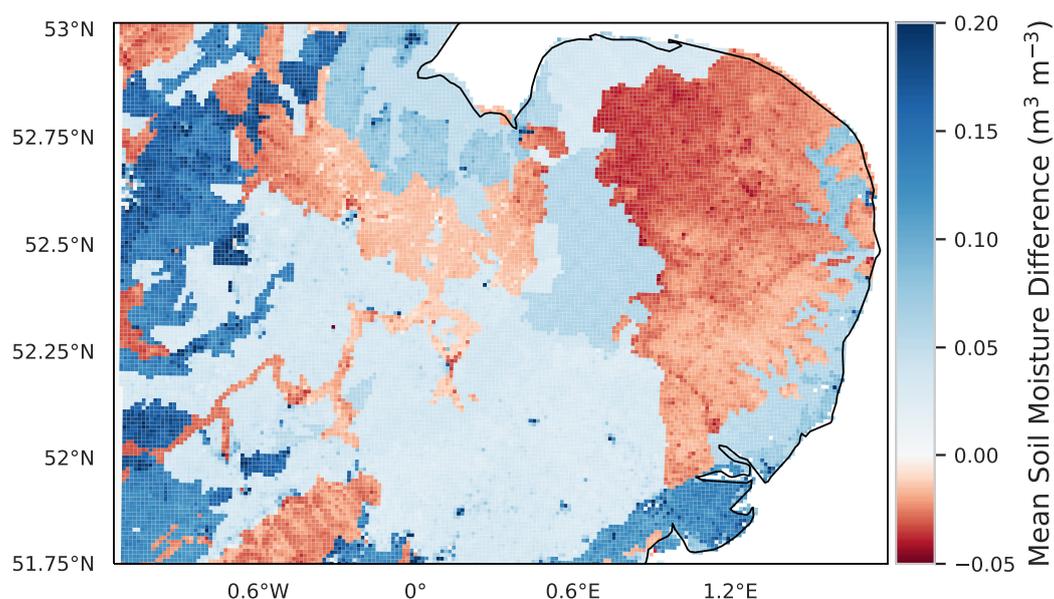when run with the CHESS data, for this domain we have over 30,000 individual model grid cells.

Figure 2. 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.

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

Figure 3 shows the prior and posterior parameter distributions for the 15 optimized parameters of the Tóth et al. (2015) pedo-transfer 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.
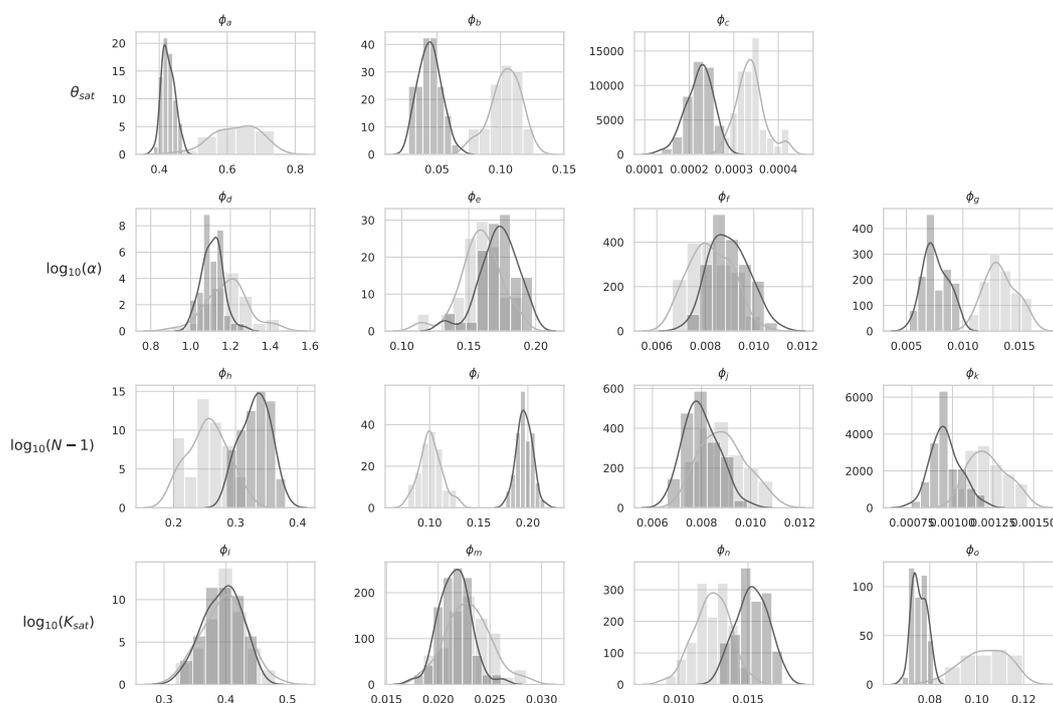
Figure 3. Distributions of prior and posterior pedotransfer function parameters grouped by the term in the equations (1) that they relate to (see row labels). Light grey: parameter distribution for the prior ensemble, dark grey: parameter distribution for the posterior ensemble.

Figure 4 shows the error reduction after performing data assimilation when comparing JULES spatially aggregated estimates to SMAP observations. 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 (see Figure 5). 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 Figure5 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 6 & 7. For Figure 6 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

200 observations that have not been used in the data assimilation framework. From Figure 6 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 7 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.
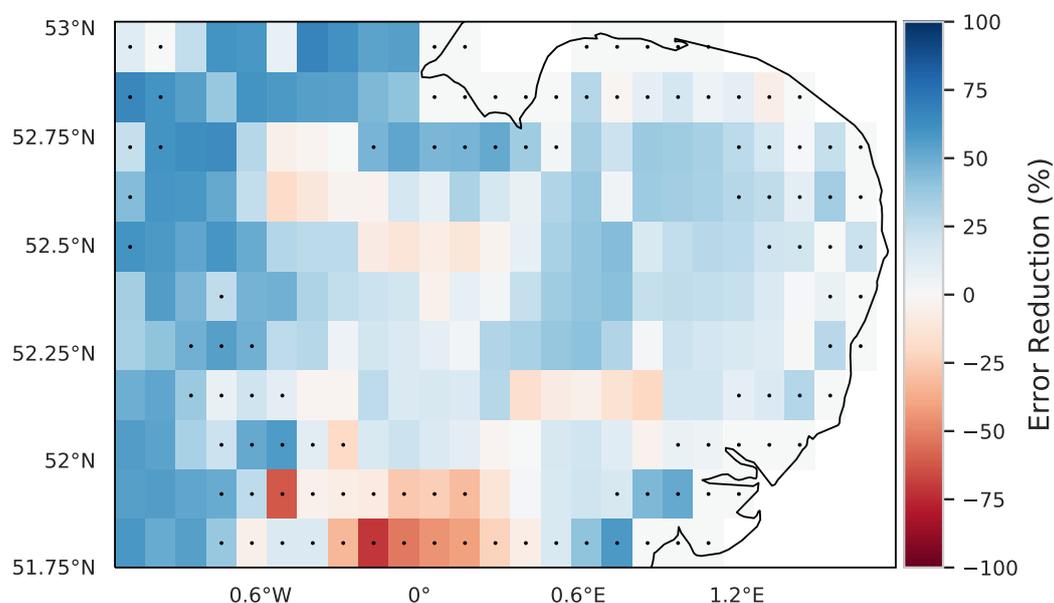
Figure 4. Map showing the difference between Root-Mean Squared Error (RMSE) when JULES spatially aggregated estimates are compared
to SMAP observations for the prior and posterior ensemble. Blue corresponds to reductions in RMSE for the posterior ensemble, red to an
increase. Grid-cells displaying stippling signify low quality SMAP pixels which have not been used in the assimilation procedure. Over the
whole domain we find an average reduction in RMSE of 20% after data assimilation.

205 ## 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 8 we show results at
the Cardington COSMOS sites, here we can see the posterior JULES estimate is a large improvement from the prior, although
210 some of the driest values are still not captured. Figure 9 shows results for Morley COSMOS site where both prior and posterior
JULES estimates perform similarly. There are also some sites where even after calibration we still do not capture the COSMOS
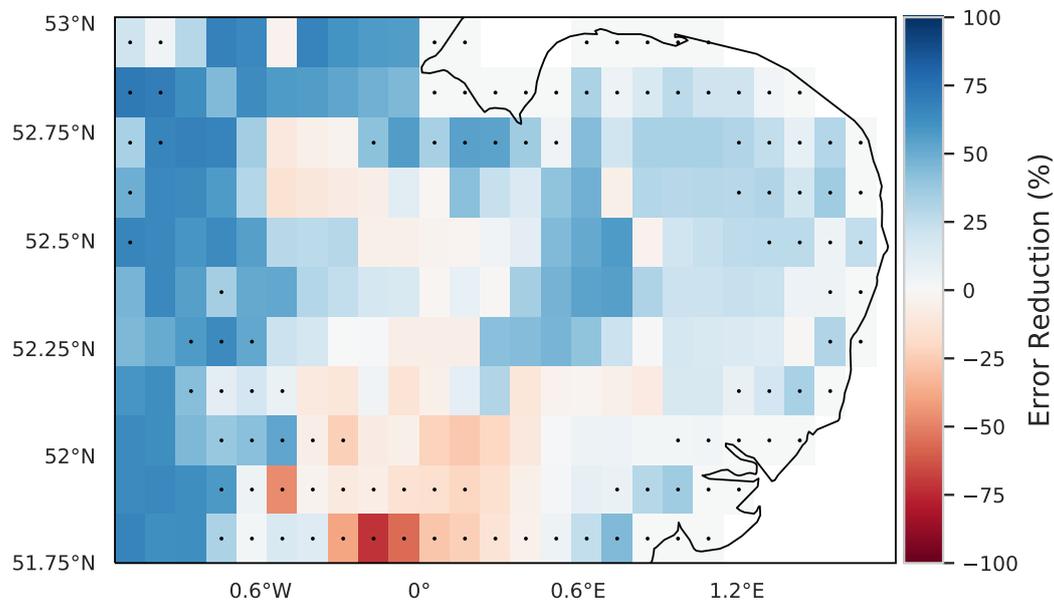
Figure 5. Map showing the difference between Root-Mean Squared Error (RMSE) when JULES spatially aggregated estimates are compared to SMAP observations for the prior and posterior ensemble in the hindcast period (2017), where no observations have been assimilated. Blue corresponds to reductions in RMSE for the posterior ensemble, red to an increase. Grid-cells displaying stippling signify low quality SMAP pixels which have not been used in the assimilation procedure. Over the whole domain we find an average reduction in RMSE of 21% after data assimilation for the hindcast period.

estimates, Stoughton in Figure 10 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. Figure 11 is an example where both prior and posterior perform equally poorly. In table 3 we show summary statistics for the 11 COSMOS sites, we see that when looking over all sites the

215 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.

4 Discussion

Typically pedotransfer functions have been calibrated using point samples of soil. We have shown that satellite data can provide additional information to find parameters that will allow for the improved prediction of soil moisture from land surface

220 models at varying spatial scales. There are obviously uncertainties which have not been taken into consideration within the DA procedure to find optimized pedotransfer function parameters. There will be errors in the underlying soil texture map from the HWSD which will imapct results, although these will not be as large for European countries as they will be in other locations with sparser soil texture information. There will also be inherent bias and errors in both the observations and model. For
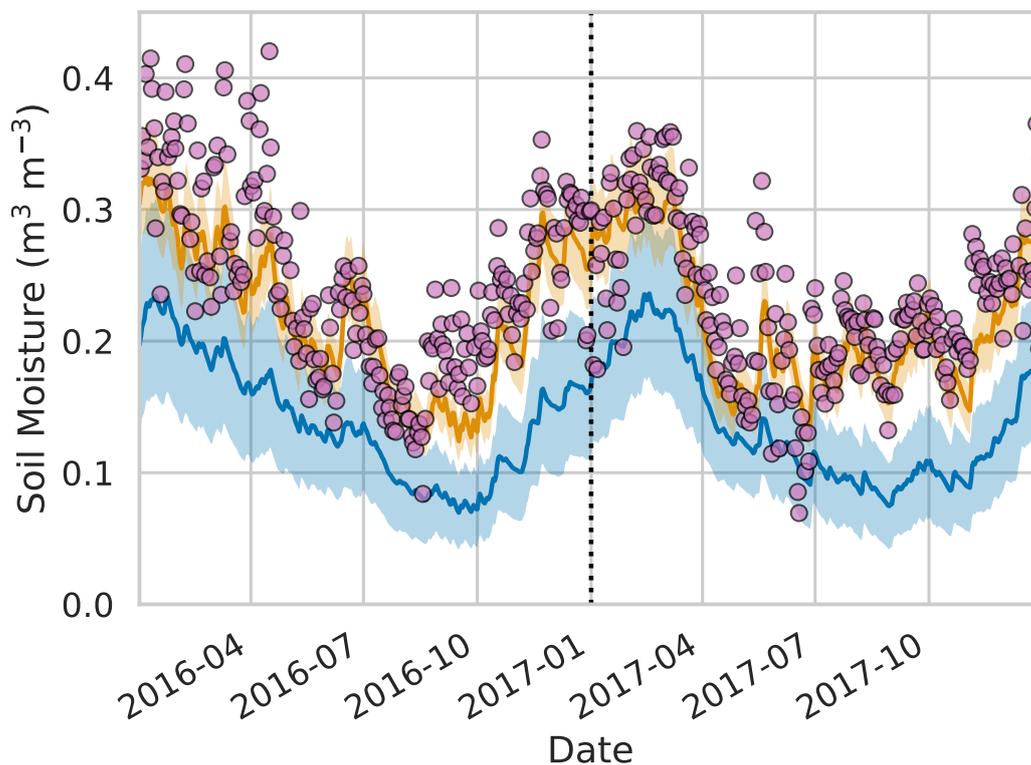
Figure 6. Time–series of soil moisture for 52.96° N 0.40° W. Pink dots: SMAP observations, blue line and shading: prior JULES mean and ensemble spread, orange line and shading: posterior ensemble mean and spread. Black dotted line represents the end of the assimilation window and start of the hindcast period.

the SMAP observations we have dealt with this by inflating the observation uncertainty within the DA algorithm. However,
225   specifying the errors arising from structural uncertainties and missing processes within the JULES model is more difficult. We can see these errors manifesting themselves in our comparisons to COSMOS–UK observations in Figures 8 to 11. Figure 8 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 9 and 10 respectively) are positioned on arable land with typical mineral soils and while we
230   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 11), 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
235   estimates of soil moisture from microwave satellite sensors over peatland are problematic, so it is understandable that we are
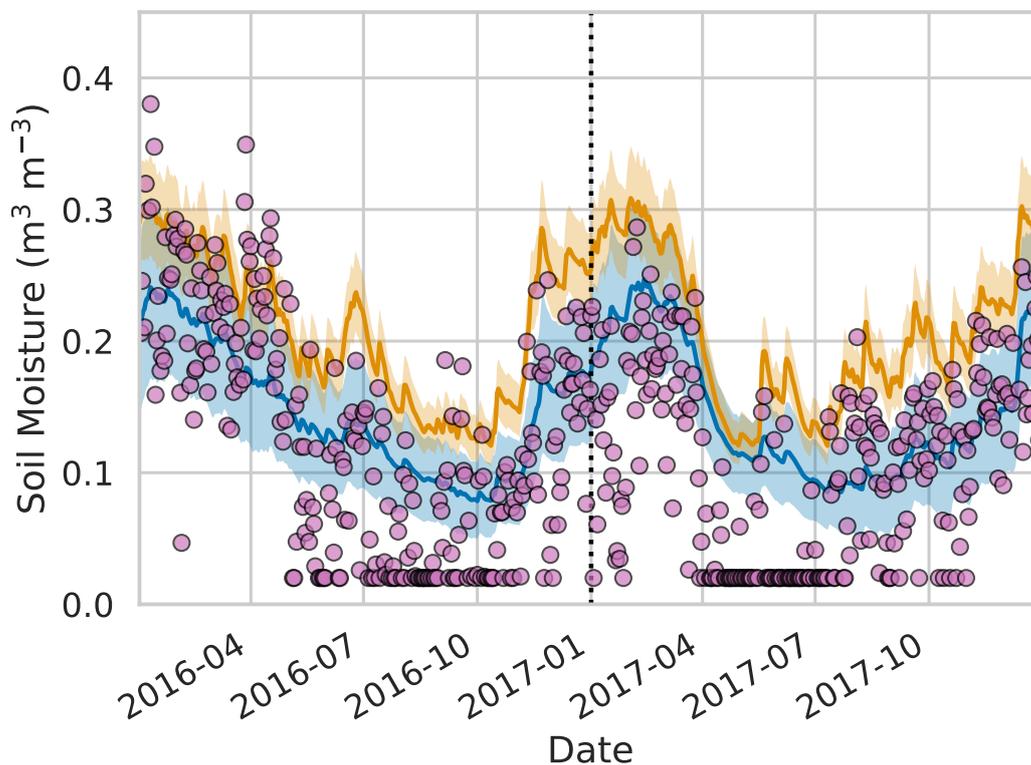
Figure 7. Time–series of soil moisture for $51.81°$ N $0.17°$ W. Pink dots: SMAP observations, blue line and shading: prior JULES mean and ensemble spread, orange line and shading: posterior ensemble mean and spread. Black dotted line represents the end of the assimilation window and start of the hindcast period.

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. 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. It

240   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).

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

245   compare 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
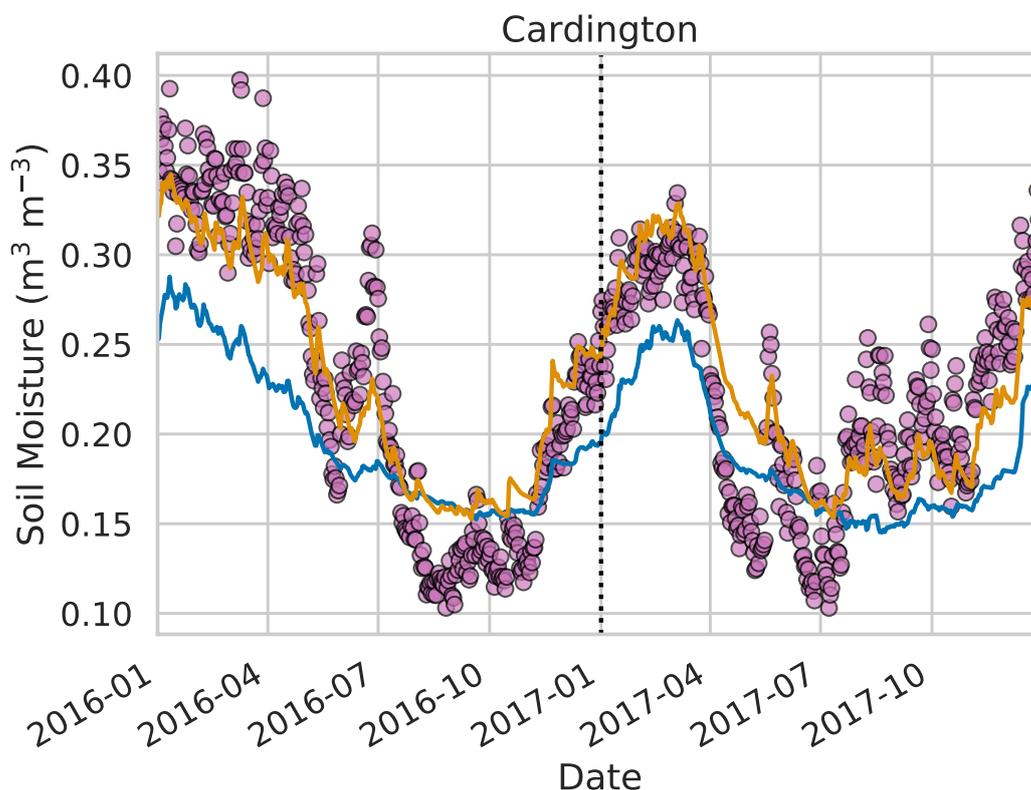
13

Figure 8. Times–series of soil moisture at Cardington COSMOS site. Pink dots: COSMOS soil moisture observations, blue line: prior JULES estimate, orange line: posterior JULES estimate.

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. The

250   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 the initial application of this technique we have focused on a specific region at a high resolution. Here we have utilised 256

255   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

260   land grid cells globally. Using our fairly modest experimental setup and assuming a linear scaling repeating at the global scale
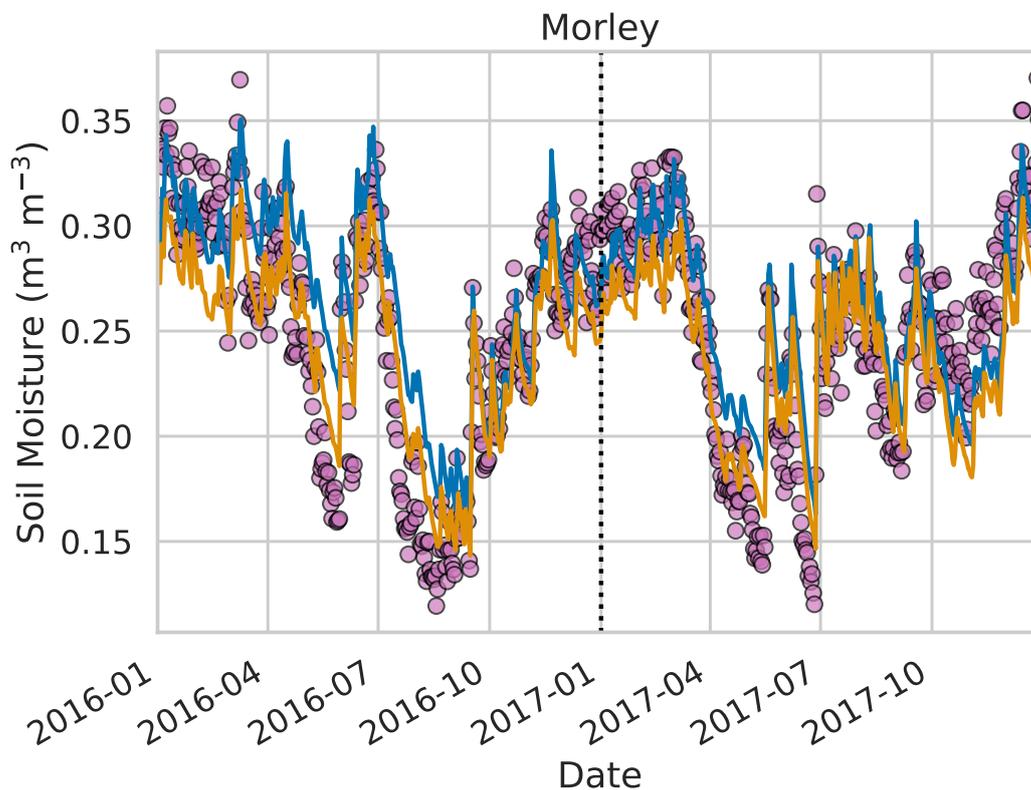
Figure 9. Times–series of soil moisture at Morley COSMOS site. Pink dots: COSMOS soil moisture observations, blue line: prior JULES estimate, orange line: posterior JULES estimate.

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.

265   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 as a powerful integrated constraint on land surface model estimates of water balance and run-off. The development of the
270   new JULES groundwater component will allow for the use of observations from the Gravity Recovery and Climate Experiment (GRACE) satellites which have the ability to monitor changes in the Earth's underground water storage.
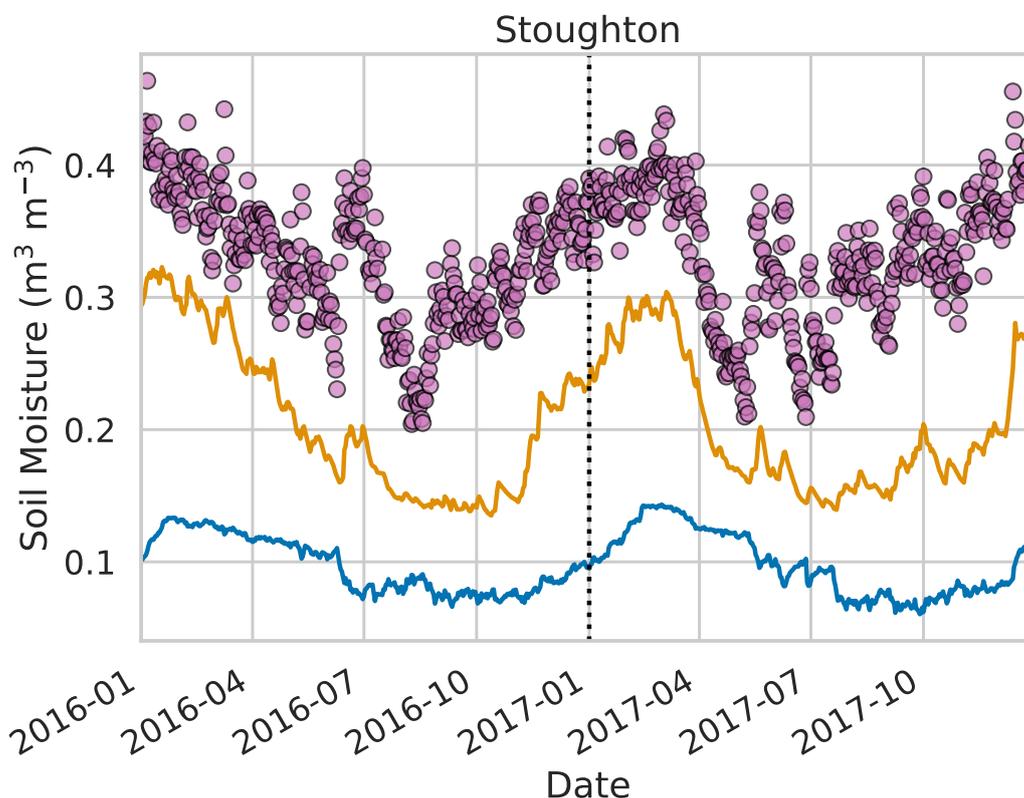
15

Figure 10. Times–series of soil moisture at Stoughton COSMOS site. Pink dots: COSMOS soil moisture observations, blue line: prior JULES estimate, orange line: posterior JULES estimate.

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 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 areas 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.
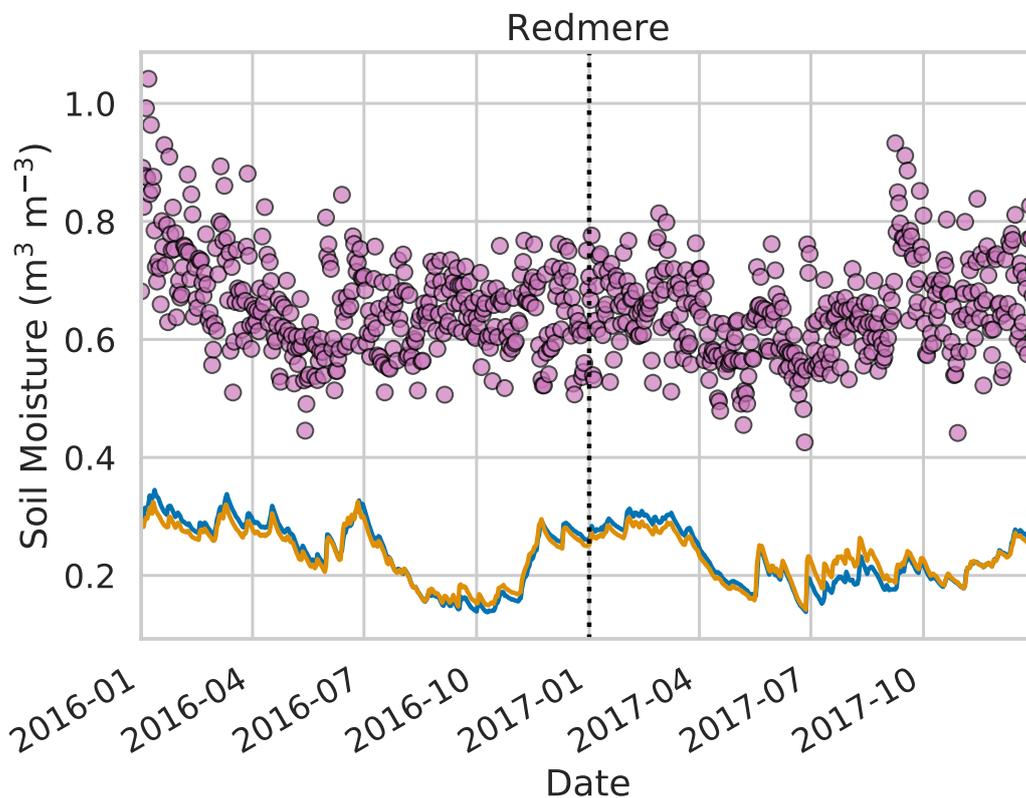
Figure 11. Times–series of soil moisture at Redmere COSMOS site. Pink dots: COSMOS soil moisture observations, blue line: prior JULES estimate, orange line: posterior JULES estimate.

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). 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:

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

Site	Correlation		ubRMSE		RMSE	
	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.

290   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

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

     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
300   matrix. In particular we choose the departure vector to be written as a linear combination of the background ensemble of perturbations, i.e.

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

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

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

305   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

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

310   where $\hat{\mathbf{H}}$ 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), $\hat{\mathbf{y}}$ are the observations (here SMAP) and $\hat{\mathbf{R}}$ is the observation error covariance matrix.

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

315   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:

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

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

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

325   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

$$330 \quad (\mathbf{I} - \mathbf{K}\hat{\mathbf{H}}) = (\mathbf{I} + \hat{\mathbf{H}}\mathbf{X}_b'{}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{X}_b')^{-1} \approx (\mathbf{I} + \mathbf{Y}_b'{}^T\hat{\mathbf{R}}^{-1}\mathbf{Y}_b')^{-1}. \tag{A15}$$

Then

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

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:

$$335 \quad \mathbf{W}_a = (\mathbf{I} + \mathbf{Y}_b'{}^T\hat{\mathbf{R}}^{-1}\mathbf{Y}_b')^{-\frac{1}{2}}. \tag{A17}$$

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

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

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

References

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-
350   11-2353-2018, 2018.

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-
355   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., Ged-
ney, 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.

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

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
365   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.

370   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.

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.

De Lannoy, G. J. M. and Reichle, R. H.: Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model,
375   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.

380   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,
385   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.

390   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.

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.

395   Hauser, M., Orth, R., and Seneviratne, S. I.: Investigating soil moisture–climate interactions with prescribed soil moisture experi-
ments: 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.

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.

400   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
405   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.

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.

410   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 meteorol-
ogy 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
415   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.

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.

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

Hydrology and
Earth System
Sciences
Discussions

Open Access

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.

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.

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.

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., and Nichols, N. K.: Correlated observation errors in data assimilation, International Journal for Numerical Methods in Fluids, 56, 1521–1527, https://doi.org/10.1002/fld.1636, 2008.

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.

460 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,
465 293, 85 – 99, https://doi.org/https://doi.org/10.1016/j.jhydrol.2004.01.008, 2004.

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,
470 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.

Zreda, M., Desilets, D., Ferré, T. P. A., and Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using
475 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.