



Improving JULES Soil Moisture Estimates through 4D-En-Var Hybrid Assimilation of COSMOS-UK Soil Moisture Observations

Ramesh Visweshwaran^{1,2}, Elizabeth Cooper³, and Sarah L Dance^{1,2,4}

¹Department of Meteorology, University of Reading, Reading, UK.

²National Centre for Earth Observation (NCEO), Reading, UK.

³UK Centre for Ecology & Hydrology (UKCEH), Wallingford, UK.

⁴Department of Mathematics & Statistics, University of Reading, Reading, UK.

Correspondence: Ramesh Visweshwaran (v.ramesh@reading.ac.uk)

Abstract. Accurate soil moisture estimates are essential for effective management and operational planning in various applications, including flood and drought response. However, soil moisture values derived from land surface models often exhibit significant deviations from in-situ observations. Data assimilation combines model information with observations to enhance prediction accuracy. Previous studies have typically focused on either estimating the initial soil moisture state or optimizing Pedotransfer Function (PTF) constants, which link soil texture to the hydraulic properties of the land surface models. In contrast, in this study, we employ a novel approach by performing joint state-parameter assimilation for the JULES model. We optimized both the PTF constants and the initial soil moisture conditions simultaneously. Using Four-Dimensional Ensemble Variational hybrid data assimilation, we ingested field-scale soil moisture observations from the Cosmic-ray Soil Moisture Monitoring Network across 16 diverse sites in the UK. The results demonstrate that joint state-parameter assimilation significantly enhances the accuracy of soil moisture estimates, improving the average Kling Gupta Efficiency values from 0.33 to 0.72 across different soil characteristics. These findings indicate that our proposed joint state-parameter assimilation framework holds great potential for enhanced predictive accuracy in land surface models.

1 Introduction

Soil moisture plays a significant role in land surface processes due to its strong influence on various components of the water cycle. Soil moisture controls the partitioning of incoming rainfall into surface runoff, percolation, and subsequent subsurface runoff (Entekhabi et al., 2010). Thus, accurate monitoring of soil moisture is essential for proper water resource management, including irrigation, and responses to floods and droughts (Roy et al., 2021; Visweshwaran et al., 2022a; Hou and Pu, 2024; Sadhwani and Eldho, 2024; de Roos et al., 2024).

Soil moisture conditions can be monitored using in-situ instruments, remote sensing from both ground- and space-based platforms, or estimated by forcing land surface models (LSMs) and hydrological models with meteorological observations. Remotely sensed observations from sensors such as Soil Moisture and Ocean Salinity (SMOS; Kerr et al. (2001)), Soil Moisture Active Passive (SMAP; Entekhabi et al. (2010)), Advanced Scatterometer (ASCAT; Wagner et al. (2013)), Sentinel 1 (Torres et al., 2012), and Advanced Microwave Scanning Radiometer-2 (AMSRE; Njoku et al. (2003)) determine soil moisture



conditions across a wide area but provide information only at the surface level due to low penetration capacity (Albergel et al.,
25 2012). Conversely, in-situ observations from networks like SMOSMANIA (Calvet et al., 2007) and SMOSREX (de Rosnay
et al., 2006) provide accurate soil moisture conditions at both surface and subsurface levels. However, these observations are
limited to point locations at selected sites, offering less spatial coverage compared to satellite-based methods. In recent years,
new approaches have been developed to measure soil moisture at the field scale. The Cosmic-ray Soil Moisture Monitoring
Network (COSMOS-UK), provided by the UK Centre for Ecology and Hydrology (UKCEH), uses cosmic ray neutron sensors
30 (CNRS) to monitor soil moisture conditions. These sensors capture data over a horizontal footprint of approximately 25 to
30 hectares and measure to a vertical depth of up to 80 cm with high accuracy (Cooper et al., 2021b; Stanley, 2023). These
observations are available a 30-minute temporal scale, providing continuous, real-time monitoring of soil moisture variations.

Land surface models can estimate soil moisture conditions at various soil depths, including surface, subsurface, and deeper
soil layers, and across a range of spatial and temporal scales, typically down to a few meters spatially and at sub-daily to
35 weekly intervals temporally. However, these estimates are uncertain. These uncertainties can stem from errors in the forcing
variables, such as rainfall and temperature data, and the model's inherent structural limitations, including parameter inaccuracies
and improper initialization of the soil moisture conditions (Alvarado-Montero et al., 2017). Modern LSMs like the Joint UK
Land Environment Simulator (JULES) require empirical soil models or pedotransfer functions (PTFs) to translate more easily
measurable soil properties like soil texture into hydraulic parameters (hydraulic conductivity and soil matric suction) essential
40 for running the model. Errors in the specified values of PTF constants can spread through land surface models, impacting the
accuracy of model forecasts.

Combining land surface models and observations synergistically can provide better soil moisture estimates than either ap-
proach in isolation. Data assimilation facilitates this process by optimally combining information from land surface models and
observations measured from satellites or ground sensors, accounting for uncertainties in both the models and the observations
45 (Beven and Binley, 1992; Kumar et al., 2022; Xu et al., 2025). Recently, land data assimilation systems using soil moisture
observations have gained considerable attention (Pinnington et al., 2018; Mason et al., 2020; Kim et al., 2021; Tian et al., 2021;
Hung et al., 2022; Visweshwaran et al., 2022b; Heyvaert et al., 2024). While sequential assimilation methods (like Ensemble
filters) are widely used due to their relatively simple and cost-effective implementation (Cooper et al., 2019; Gong et al., 2024),
dynamically updating model states and parameters at frequent intervals may not be consistent with the stable (non-chaotic) na-
50 ture of land surface processes (Shi et al., 2015; De Lannoy and Reichle, 2016). On the other hand, variational techniques (like
4D-Var) can use a longer assimilation window timeframe and establish time-invariant state-parameter values over the duration
of this window (Smith et al., 2013; Yang et al., 2016; Pinnington et al., 2017). Due to the high dimensionality and complexity
of JULES, along with the frequent release of new versions, calculating the gradient of the dynamical model and the observation
operator has become both computationally expensive and inefficient. Hybrid methods, like 4D-En-Var, approximate the model
55 gradient and leverage the advantages of both variational and ensemble approaches, providing more accurate estimates over
long time windows without requiring the explicit calculation of the full model gradient.

Previous studies on parameter estimation in land surface models using soil moisture observations have primarily focused on
directly improving the model's soil parameters to enhance prediction accuracy (Lü et al., 2011; Zha et al., 2019; Seo et al.,



2021; Visweshwaran et al., 2024). We have attempted to update the underlying PTF constants to understand their impact on model forecasts, and only a few studies have explored this on the JULES model (Pinnington et al., 2021; Cooper et al., 2021b). Addressing PTF constants is crucial, as it preserves the relationships between soil physical parameters (such as matric suction and hydraulic conductivity) and reduces uncertainty propagation from poor calibration. For studies that include multiple sites, assimilating observations from these sites together can produce a common posterior PTF constant set, representative of various soil characteristics, which can then be applied in regions lacking in-situ observations.

Beyond parameters, the model's initial soil moisture condition also plays a significant role in model forecasts. While optimizing model parameters minimizes long-term bias (systematic error) in forecasts, state assimilation addresses stochastic errors (Nearing et al., 2018). The initial soil moisture conditions influence not only current land surface processes but also future conditions through its long-term memory effect (Walker and Houser, 2001). Many studies have therefore focused on updating the initial soil moisture state in land surface models (Ridler et al., 2014; Dumedah and Walker, 2014; Chen et al., 2015).

Accurate forecasting requires precise information on both the initial conditions and model parameters (Smith et al., 2013). This dual focus on both parameter and state updates is important, yet remains underexplored in the context of the JULES land surface model. Unlike previous work, we introduce a novel approach by implementing joint state-parameter estimation for the JULES model, simultaneously updating both the PTF constants and the initial soil moisture conditions. This joint assimilation framework allows for a more integrated approach to managing uncertainty and represents the first application of this method for the JULES model, aiming to improve predictive accuracy across various soil characteristics. To demonstrate the advantage of state-parameter assimilation over individual state-only or parameter-only assimilation, in this study, we performed soil moisture assimilation in three scenarios. 1) state-only assimilation to update the initial soil moisture condition, 2) parameter-only assimilation to update the PTF constants, and 3) joint state-parameter assimilation to update both the initial soil moisture condition and Cosby's PTF constants. COSMOS-UK soil moisture observations from 2017 across 16 sites in the UK were integrated into the community JULES land surface model using a 4D-En-Var hybrid assimilation approach. Observations from all 16 sites were used together to derive common PTF constants representative of the different soil types in the UK, and the model's initial conditions for each site.

Section 2 describes the details of the study area, COSMOS-UK observations, and the JULES and Cosby models. It then outlines the methodology used in the study and the metrics applied to evaluate the model's performance before and after assimilation. Section 3 discusses the performance of JULES during state estimation, parameter estimation, and joint state-parameter assimilation, highlighting how field-scale soil moisture assimilation improves JULES's ability to estimate and forecast soil moisture. Finally, Section 4 emphasizes the significance of joint state-parameter assimilation in enhancing JULES performance across diverse soil characteristics and discusses the broader implications of these findings for future applications.



90 2 Methods

2.1 Study Area and COSMOS-UK Observation

To assess the performance of the JULES model before and after assimilating field-scale soil moisture observations through the 4D-En-Var scheme, 16 sites across the UK were selected (Figure 1). The sites were chosen based on soil characteristics, the availability of COSMOS-UK soil moisture, and meteorological observations. The soil texture at nine locations (Cardington: 95 CARDT, Bickley Hall: BICKL, Crichton: CRICH, Waddesdon: WADDN, Hollin Hill: HOLLN, Easter Bush: EASTB, Rothamsted: ROTH, Hartwood Home: HARTW, Chobham Common: CHOBH) is categorized as “typical mineral soil,” where the JULES model is expected to perform well. Conversely, Sheepdrove: SHEEP, Gisburn Forest: GISBN, Moorhouse: MOORH, and Sourhope: SOURH have soils with "high organic content", while Chimney Meadows: CHIMN, Porton Down: PORTN, and Lullington Heath: LULLN feature “calcareous mineral soil”. The details about the soil characteristics were taken from 100 Antoniou et al. (2019).

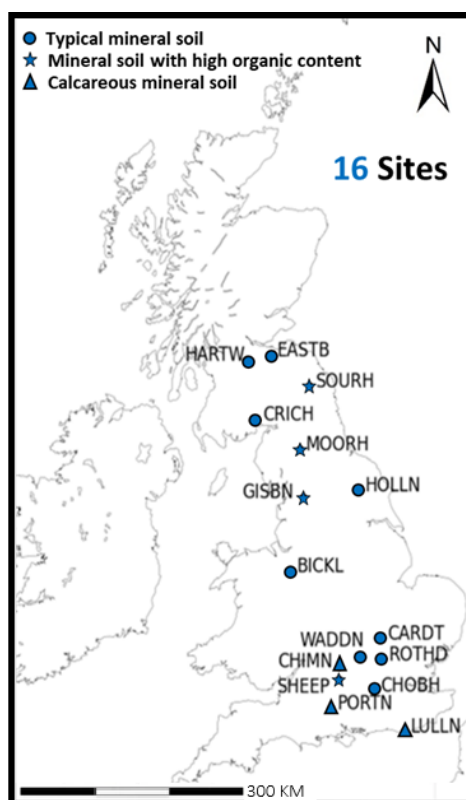


Figure 1. Locations of COSMOS-UK sites used for assimilation purposes.



The meteorological dataset required to run the JULES model was obtained from the COSMOS-UK sub-daily database (Cooper et al., 2021b; Stanley, 2023). These data were collected every 30 minutes at the same 16 locations as soil moisture observations were measured. Meteorological observations from 2016 to 2018 were used in this study. The year 2016 was designated as the spin-up period to initialize the JULES model, while 2017 served as the assimilation window for 4D-En-Var, and 2018 was used for forecast evaluation of JULES after assimilation.

For assimilation purposes, daily soil moisture observations were obtained from the Cosmic-ray Soil Moisture Observing System, which has ~50 sites across the United Kingdom. This system provides soil moisture time series observations across different sites, covering different soil types, land uses, and environmental conditions. The CNRS sensor at each site monitors epithermal cosmic-ray neutrons attenuated by water molecules in the soil profile, providing near-real-time soil moisture variations integrated over a variable sensing depth; this depth is included with the dataset at a daily timestep (Evans et al., 2016). CRNS provides information for an area up to 120,000 m^2 and an approximate depth of 10 to 80 cm around its peripheral area, depending on local site conditions (Antoniou et al., 2019). The soil moisture estimates from the sensors are calibrated for each site to allow for local conditions (Evans et al., 2016).

2.2 JULES Land Surface Model

The Joint UK Land Environment Simulator is a community land surface model that can operate both independently and as the land surface component within the Met Office Unified Model (Clark et al., 2011; Best et al., 2011). Meteorological data, including precipitation, wind speed, temperature, air pressure, and solar radiation (incoming shortwave and longwave radiation), are required inputs for running the JULES model. A detailed explanation of the energy and water flux components of JULES is provided in Best et al. (2011).

The JULES model assumes four vertical soil layers to represent soil moisture conditions, with default thicknesses of 0.1 m, 0.25 m, 0.65 m, and 2 m. The Richards equation defines the vertical movement of water flux between the layers,

$$\frac{d\theta_l}{dt} = w'_{l-1} - w'_l - E'_l - R_{bl}, \quad (1)$$

where θ and the subscript l represent the soil water content and soil layer, respectively, and w'_{l-1} and w'_l represent the downward water flux infiltrating from the layer above and out to the lower layer, respectively. The evapotranspiration uptake by the plant roots from each soil layer and the lateral movement of water are represented by E'_l and R_{bl} , respectively. In this study, R_{bl} is set to 0 since the assimilation is done separately on 16 sites, treating JULES as one-dimensional in vertical columns.

The JULES model implements Darcy's law to update the soil moisture content at each layer based on hydraulic conductivity. The Brooks and Corey model (Brooks and Corey, 1964) is used to represent the relationship between soil water content (θ) and hydraulic conductivity (K_h),

$$\frac{\theta}{V_{\text{sat}}} = \left(\frac{\psi}{\text{sathh}} \right)^{-\frac{1}{b}}, \quad (2)$$

and

$$K_h = \text{satcon} \left(\frac{\theta}{V_{\text{sat}}} \right)^{2b+3}, \quad (3)$$



where b and ψ represent the exponent in the soil hydraulic characteristic, and matric suction, respectively and V_{sat} , s_{athh} ,
135 and satcon represent the volumetric soil moisture content, matric suction, and hydraulic conductivity at saturation condition,
respectively. Pedotransfer functions are generally used to determine the values of these hydraulic properties, which depend on
soil characteristics.

2.3 Cosby's Pedotransfer function

The volumetric soil moisture content can be estimated using hydraulic conductivity and soil matric suction, as proposed by
140 the Brooks and Corey method. Pedotransfer Functions are often employed to determine the values of these parameters by
establishing a relationship with easily measurable properties like soil texture and specific density (Marthews et al., 2014). In this
study, the Cosby Pedotransfer Function (Cosby et al., 1984) is employed, considering the input and output data requirements of
the soil hydraulic model. Furthermore, Cosby's PTF has been used in conjunction with the JULES land surface model across a
wide range of UK soils and has shown promising results (e.g. Cooper et al., 2021a). The mathematical explanation of Cosby's
145 PTF is summarized in the following equations. The soil-water retention exponent, b , is defined as,

$$b = K1 + K2f_{\text{clay}} - K3f_{\text{sand}}, \quad (4)$$

where f_{clay} , f_{silt} , and f_{sand} represent the fraction of clay, silt and sand, respectively, in percentage. The coefficients $K1$, $K2$, and
 $K3$ are Cosby constants. The volumetric water content at saturation, V_{sat} , is represented by,

$$V_{\text{sat}} = K4 - K5f_{\text{clay}} - K6f_{\text{sand}}, \quad (5)$$

150 where V_{sat} represents the maximum water-holding capacity of the soil. It depends on the soil texture via coefficients $K4$, $K5$,
and $K6$. The soil matric suction (in meters), s_{athh} , describing the force with which water is held in the soil at saturation is
represented by,

$$s_{\text{athh}} = 0.01 \times 10^{(K7 - K8f_{\text{clay}} - K9f_{\text{sand}})}. \quad (6)$$

Equation 6 shows a logarithmic relationship with the clay and sand fractions, and the parameters $K7$, $K8$, and $K9$ are Cosby
155 constants. The saturated hydraulic conductivity ($kg\ m^{-2}\ s^{-1}$), satcon , which describes how easily water moves through the
soil when fully saturated, is represented by,

$$\text{satcon} = 10^{(-K10 - K11f_{\text{clay}} + K12f_{\text{sand}})}. \quad (7)$$

The critical point volumetric water content (V_{crit}) corresponds to the water content (in $m^3\ m^{-3}$) at a specific matric suction
(usually -33 KPa, or field capacity) and the wilting point (V_{wilt}) is the volumetric water content (in $m^3\ m^{-3}$) at suction of -1500
160 KPa, beyond which plants can no longer extract water from the soil. The terms V_{crit} and V_{wilt} are represented by,

$$V_{\text{crit}} = V_{\text{sat}} \left(\frac{s_{\text{athh}}}{3.364} \right)^{\frac{1}{b}}, \quad (8)$$



and

$$V_{\text{wilt}} = V_{\text{sat}} \left(\frac{\text{sathh}}{152.9} \right)^{\frac{1}{b}}. \quad (9)$$

165 Lastly, the soil thermal capacity (h_{cap}) representing the ability of the soil to store heat and the soil thermal conductance (h_{con}) showing the measure of the soil's ability to transfer heat are represented by,

$$h_{\text{cap}} = (1 - V_{\text{sat}})(2.376 \times 10^6 f_{\text{clay}} + 2.133 \times 10^6 f_{\text{silt}}), \quad (10)$$

and

$$h_{\text{con}} = 0.025 V_{\text{sat}} \left(1.16^{f_{\text{clay}}(1-V_{\text{sat}})} \cdot 1.57^{f_{\text{sand}}(1-V_{\text{sat}})} \cdot 1.57^{f_{\text{silt}}(1-V_{\text{sat}})} \right). \quad (11)$$

170 The hydraulic parameters b , V_{sat} , sathh and satcon are primary input requirements to run the JULES model and these parameters are estimated using sand and clay fractions as shown in equations (4-7). The parameters h_{cap} , h_{con} , V_{crit} and V_{wilt} are derived parameters of V_{sat} . In this study, we use prior values of $K1$ to $K12$ as described by Cosby et al. (1984).

2.4 4D-En-Var Assimilation

In this study, we adopted the 4D-En-Var hybrid assimilation method to update JULES's initial soil moisture condition and/or Cosby PTF constants ($K1$ to $K12$). We consider the discrete non-linear dynamical model, defined as

$$175 \quad \mathbf{z}_t = \mathbf{f}_{t-1 \rightarrow t}(\mathbf{z}_{t-1}, \mathbf{p}_{t-1}), \quad (12)$$

where $\mathbf{z}_t \in \mathbb{R}^n$ represents the state vector at time t , and \mathbf{f} is the non-linear model operator that evolves the model state vector from time $t-1$ to t ($t = 1, 2, \dots, N$ timesteps). The model parameter vector at time $t-1$ is described by $\mathbf{p}_{t-1} \in \mathbb{R}^m$, and the parameters are assumed to be time-invariant, i.e., $\mathbf{p}_t = \mathbf{p}_{t-1}$. Initial conditions provide the starting values for \mathbf{z}_0 and are propagated forward through the model operator.

180 Let $\mathbf{y}_t \in \mathbb{R}^{k_t}$ represent the observation vector at time t , where k_t is the length of the observation vector. These observations are related to the model state through the observation operator $\mathbf{h}_t : \mathbb{R}^n \rightarrow \mathbb{R}^{k_t}$, which maps the state vector to the observation space:

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{z}_t) + \epsilon_t, \quad (13)$$

185 where $\epsilon_t \in \mathbb{R}^{k_t}$ is the observation error term, assumed to be unbiased, Gaussian, and uncorrelated in time, with a covariance matrix \mathbf{R}_t .

Pinnington et al. (2020) developed the Land Variational Ensemble Data Assimilation Framework (LAVENDAR) for parameter estimation in JULES, and we modified this algorithm to update both the state and/or parameters for the three scenarios conducted in this study. This optimization is achieved by estimating the control vector, \mathbf{x}_t , which encapsulates the variables targeted during assimilation, such as the initial soil moisture states, PTF constants, or both, depending on the assimilation 190 scenario.



1. In Scenario 1 (state-only assimilation: SC 1), we aim to estimate the initial soil moisture states, denoted by $z \in \mathbb{R}^n$, where $n = 64$ for the 4 soil layers at each of the 16 sites. The control vector for this scenario is represented as $x_t = [z_t]^T \in \mathbb{R}^n$.
2. In Scenario 2 (parameter-only assimilation: SC 2), we focus on estimating the constants of the PTF, denoted by $p \in \mathbb{R}^m$, where $m = 12$ for the 12 Cosby constants. Here, the control vector is represented as $x_t = [p_t]^T \in \mathbb{R}^m$.
- 195 3. In Scenario 3 (state-parameter assimilation: SC 3), we simultaneously estimate both the initial soil moisture states and the PTF constants. The augmented system control vector x_t is then determined by appending the parameters to the model state vector. It is represented as $x_t = [p_t \quad z_t]^T \in \mathbb{R}^{m+n}$.

The 4D-En-Var method is similar to 4D-Var, as it assimilates all observations from a predetermined time window and spatial domain together in a batch. In 4D-Var, we seek to minimize the cost function (J), which is the sum of the discrepancy
 200 between the model estimate and the prior term (x_b representing the prior vector at time = 0, ‘b’ for ‘background’ synonymous with ‘prior’), weighted by the background error covariance matrix (B), and the differences between all the model states and observations, weighted by the observation error covariance matrix (R). The cost function (J) is defined as

$$J(x_a) = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \frac{1}{2} \sum_{t=0}^N (h_t(x_t) - y_t)^T R^{-1} (h_t(x_t) - y_t), \quad (14)$$

where N represents the length of the assimilation window. The principal aim is to optimize a state and/or parameter vector (posterior vector: x_a , ‘a’ for ‘analysis’ synonymous with ‘posterior’) that minimizes J . The posterior vector is usually determined
 205 by applying a gradient-based descent algorithm iteratively to the cost function and its gradient. To avoid calculating tangent linear model, adjoint model and B^{-1} for JULES, 4D-En-Var applies a control variable transform to precondition the matrix B (Bannister, 2017). This approach allows B to be defined as the product of the perturbation matrix, x'_b , and its transpose given by

$$210 \quad B \approx x'_b x'^T_b, \quad (15)$$

where

$$x'_b = \frac{1}{\sqrt{N_e - 1}} \left(x_b^1 - \bar{x}_b, x_b^2 - \bar{x}_b, \dots, x_b^{m-1} - \bar{x}_b, x_b^{N_e} - \bar{x}_b \right). \quad (16)$$

The posterior vector x_a is then defined as

$$x_a = \bar{x}_b + x'_b w, \quad (17)$$

215 where w is a vector of length N_e and N_e is the ensemble size. The term x_b^i can come from a previous forecast (in that case \bar{x}_b is the mean of the N_e ensemble members) or sampled from a known distribution $\mathcal{N}(\bar{x}_b, B)$. Substituting equations 17 and 15 into 14, we get a new cost function and its gradient as a function of w given by

$$J(w) = \frac{1}{2} w^T w + \frac{1}{2} \sum_{t=0}^N (y'_{b,t} w + h_t(\bar{x}_b) - y_t)^T R^{-1} (y'_{b,t} w + h_t(\bar{x}_b) - y_t), \quad (18)$$



and

$$\nabla J(\mathbf{w}) = \mathbf{w} + \sum_{t=0}^N (\mathbf{y}'_{b,t})^T \mathbf{R}^{-1} (\mathbf{y}'_{b,t} \mathbf{w} + \hat{\mathbf{h}}_t(\bar{\mathbf{x}}_b) - \hat{\mathbf{y}}_t). \quad (19)$$

The term $\sum_{t=0}^N \mathbf{y}'_{b,t} \approx \sum_{t=0}^N \mathbf{h}_t(\mathbf{x}'_b)$ is the perturbation matrix in observation space, and it is computed by

$$\mathbf{y}'_{b,t} = \frac{1}{\sqrt{N_e - 1}} (\mathbf{h}_t(\mathbf{x}_b^1) - \mathbf{h}_t(\bar{\mathbf{x}}_b), \mathbf{h}_t(\mathbf{x}_b^2) - \mathbf{h}_t(\bar{\mathbf{x}}_b), \dots, \mathbf{h}_t(\mathbf{x}_b^{N_e-1}) - \mathbf{h}_t(\bar{\mathbf{x}}_b), \mathbf{h}_t(\mathbf{x}_b^{N_e}) - \mathbf{h}_t(\bar{\mathbf{x}}_b)). \quad (20)$$

This allows the computation of the posterior without explicitly determining the tangent linear model, adjoint model, and \mathbf{B}^{-1} . Minimizing the cost function (equation 18) using the gradient provides the posterior estimate of \mathbf{w} , which, when substituted into equation 17, yields the posterior estimate of the control vector \mathbf{x}_a .

2.5 Experiment Design

The implementation of COSMOS-UK soil moisture assimilation into the JULES model using the 4D-En-Var is done in three scenarios (as outlined in Section 2.4). In the first and second scenarios, the initial soil moisture state and Cosby's PTF constants are updated individually, respectively, to understand their individual impact on the JULES forecast. In the third scenario, joint state-parameter assimilation is performed to understand the combined impact. The steps involved in the implementation are outlined below and shown in Figure 2.

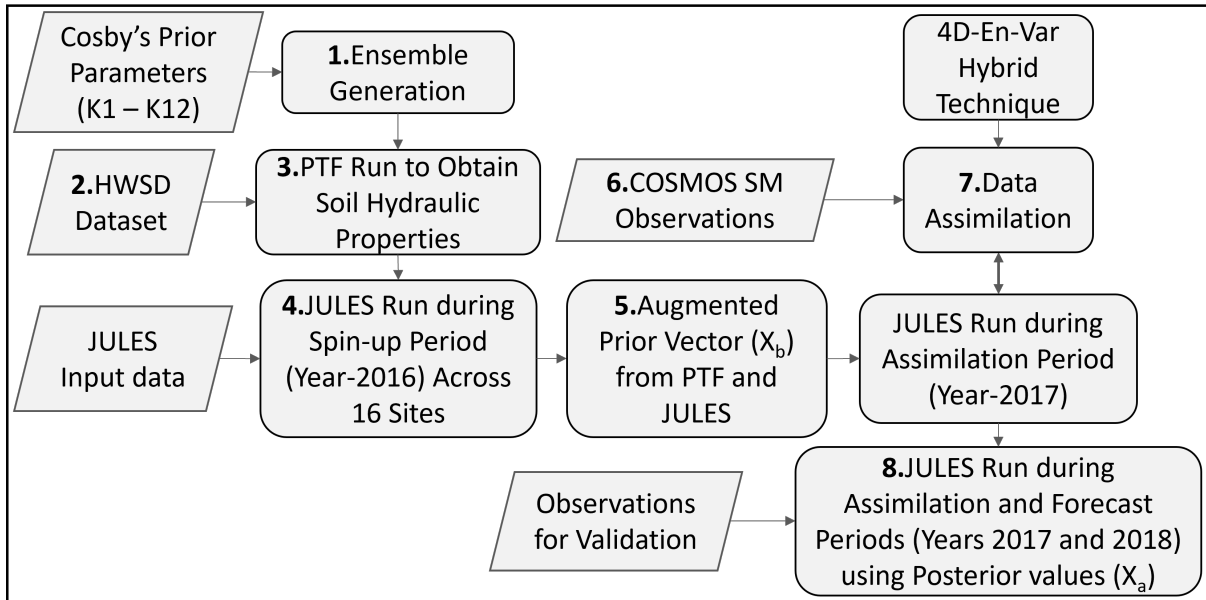


Figure 2. Schematic showing the application of 4D-En-Var assimilation scheme on JULES land surface model during state-parameter estimation (Scenario 3). The numbered boxes represent the steps adopted in this study.



1. **Ensemble generation:** A fifty-member ensemble (i.e. $N_e = 50$) was generated for each of the 12 Cosby PTF constants by sampling from a Gaussian distribution with mean values as outlined in Cosby et al. (1984). A 10 percent standard deviation around the mean was used, following the successful implementation of JULES on these sites by Cooper et al. (2021a).
235
2. **Harmonized World Soil Database (HWSD):** Soil texture information, comprising the percentage of sand, silt and clay, is obtained from the HWSD dataset (Fischer et al., 2008) for all 16 sites. The HWSD data, originally in a 30 arc-second spatial scale gridded raster format, was downscaled to a 1 km resolution, and the gridpoints passing through the COSMOS-UK sites were selected.
- 240 3. **Pedotransfer function:** Cosby's PTF is applied using the HWSD data and the 50 $K1 - K12$ sets to generate a 50-member ensemble of unique soil hydraulic parameters.
4. **Spin-up period:** The soil hydraulic parameters (50×12) obtained from the PTF are fed into the JULES model during the warm-up period extending from January to December 2016.
5. **Prior vector:**
 - 245 (a) **Scenario 1 (SC 1):** During state estimation, the JULES 50-member ensemble forecast at the end of the spin-up period, represented by four soil layers, forms the initial soil moisture condition for the start of the assimilation period. These 4 layers for each of the 16 sites are combined to represent the prior control vector ensemble (i.e., $\{\mathbf{x}_b^{(i)} \in \mathbb{R}^{64} : i = 1, 2, \dots, N_e\}$).
 - (b) **Scenario 2 (SC 2):** During parameter estimation, Cosby's prior parameter 50-member ensemble, which is common
250 to all 16 sites, forms the prior ensemble with a dimension of $\{\mathbf{x}_b^{(i)} \in \mathbb{R}^{12} : i = 1, 2, \dots, N_e\}$.
 - (c) **Scenario 3 (SC 3):** During state-parameter estimation, JULES 50-member ensemble forecast (containing 4 soil layers) from the spin-up period for each site and a common Cosby's prior parameter ensemble representing all the 16 sites are augmented to represent the prior ensemble (i.e. $\{\mathbf{x}_b^{(i)} \in \mathbb{R}^{76} : i = 1, 2, \dots, N_e\}$).
- 255 6. **Observations:** As outlined in Section 2.1, COSMOS-UK provides soil moisture observations at field scale around the CRNS instrument at a 30-minute frequency. We require the JULES soil moisture values corresponds to the same soil depth as the observations. Therefore, a weighted depth approach is employed, averaging the JULES estimates (soil moisture over 4 layers) to adjust to the level of COSMOS-UK soil moisture depth. It should be noted that hourly observations are averaged to a daily scale, with analysis indicating that the standard deviation of the hourly data around the daily mean is approximately 20% (Cooper et al., 2021a). Further, observation errors are assumed to be uncorrelated. Considering
260 these aspects, a standard deviation of 50% of the daily mean observation is assigned to \mathbf{R} for these sites. Further details on the weighted depth approach and the choice of \mathbf{R} can be found in Cooper et al. (2021a).
7. **Data assimilation:** The 4D-En-Var assimilation scheme is employed within the JULES model for the year 2017 (1-year time window) to obtain posterior values for the state (for each site) and/or parameters (common for all sites).



8. **Model forecast:** Steps 3 and 4 are repeated with posterior PTF values, and the JULES model is rerun for the assimilation
 265 period (year 2017) and for the forecast period (year 2018) with the posterior initial conditions and soil ancillary files.

2.6 Evaluation Metrics

To assess the model's performance during assimilation and forecast periods, two evaluation metrics are calculated: Root Mean
 Square Error (RMSE) and Kling-Gupta Efficiency (KGE). These metrics enable a comprehensive assessment of the model's
 performance by quantifying discrepancies between observed and modelled soil moisture values over a two-year period. The
 270 RMSE and KGE metrics are defined as follows,

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^N (\text{SM}_m^t - \text{SM}_o^t)^2}{N}}, \quad (21)$$

where N is the total number of observations. In this study, $N = 730$ for the two-year study period, and SM_m and SM_o are the
 modelled and observed soil moisture values, respectively. The term KGE is defined as

$$\text{KGE} = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}. \quad (22)$$

275 The KGE, as proposed by Gupta et al. (2009), provides an overall assessment of soil moisture distribution trends by incor-
 porating three components: correlation, variability, and bias. In this context, r represents the Pearson correlation coefficient,
 capturing the correlation between modelled and observed soil moisture time series. The term $\alpha = \sigma_m / \sigma_o$ reflects the model's
 ability to capture the variability in the observations, where σ_m and σ_o are the standard deviations of the modelled and observed
 data, respectively. Meanwhile, $\beta = \mu_m / \mu_o$ represents the model bias, where μ_m and μ_o denote the mean soil moisture values
 280 for the model and observations. Together, these components provide a robust evaluation of model performance.

For all components and the KGE, values closer to 1 indicate good model performance. In general, the KGE value ranges
 from $-\infty$ to 1, with 1 representing a perfect fit. A KGE value of 0 signifies model performance equal to the average value of
 the observations, and any value below -0.14 indicates poor model performance (Knoben et al., 2019).

3 Results

285 3.1 Scenario 1: Individual effect of state assimilation on JULES performance

During state assimilation, the initial conditions are updated for each site, and the posterior JULES estimates are compared
 against COSMOS-UK observations. The adjustments made to the initial conditions during assimilation shifted the posterior
 estimates closer to the observations during the initial phase, as illustrated in Figure 3 with examples from the HOLLN and
 SHEEP sites. These sites were chosen to represent two distinct soil types—typical mineral soil (HOLLN) and soil with high
 290 organic content (SHEEP) since JULES performance is likely to vary depending on soil characteristics. Although the initial
 adjustments improved the estimates, this effect decayed rapidly after 3 to 4 months. Notably, for most sites, including HOLLN
 and SHEEP, both the prior and posterior estimates showed no significant differences after 6 months. This can be attributed to



the dynamic nature of soil moisture, which diminishes the impact of updated initial conditions on JULES estimates over longer durations (seasonal to annual scale).

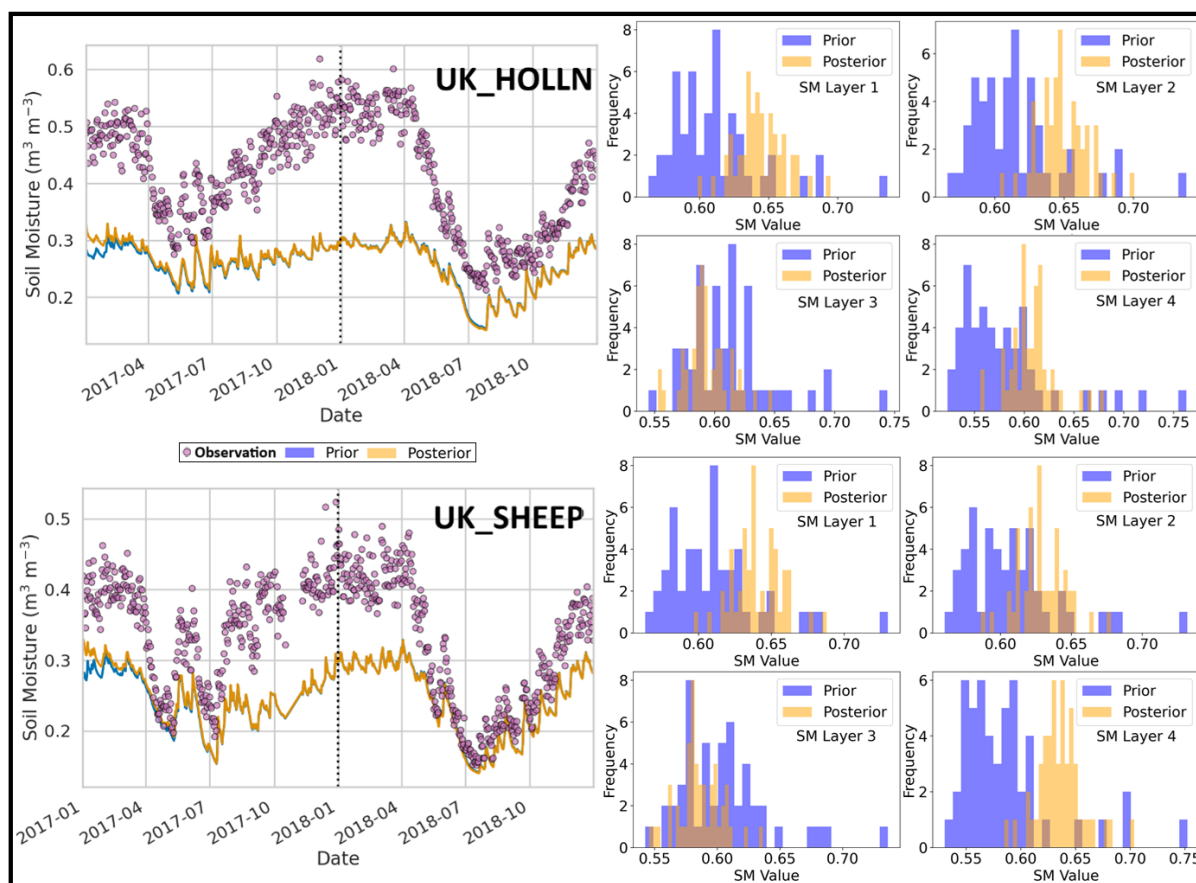


Figure 3. The figure shows the results of the state-only assimilation scenario (SC 1). The left panels display the modeled (ensemble mean) and observed soil moisture time series at the HOLLN and SHEEP sites during the assimilation (2017) and forecast (2018) periods. The right panels show the prior and posterior distributions of the initial condition state (January 1, 2017) across 50 ensemble members and four soil layers.

295 KGE metrics (Figure 4) also showed a similar outcome, with moderate improvements between the prior (Blue) and poste-
 300 rior (Brown) values, except at the GISBN site. Although Figure 4 includes results from other scenarios as well, this section
 specifically discusses scenario 1. The anomaly at the GISBN site might be due to the presence of a high amount of coniferous
 trees, which can complicate the calibration of the CRNS instrument and reduce the reliability of COSMOS-UK soil moisture
 observations. Across all sites, the posterior JULES model showed improvement in the α and β components (capturing the vari-
 ability and bias, respectively). However, there were no significant changes in the correlation component (r) of the model. This
 is due to the high correlation between prior values and observations before assimilation. Table 1 shows the RMSE between the



modelled soil moisture and COSMOS-UK observations. For the state-only scenario, RMSE values reduced at all sites, with an average reduction of 39.5% for all 16 sites.

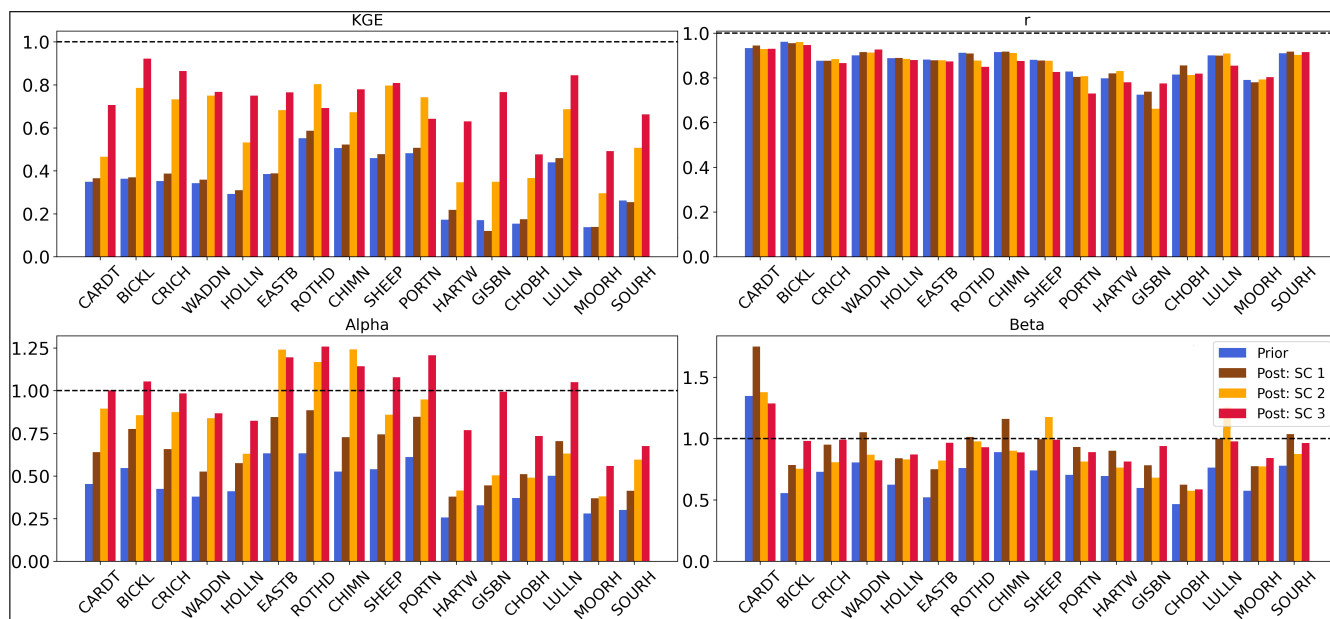


Figure 4. KGE metrics and its components showing the performance of JULES model across all the 16 sites in UK for three assimilation scenarios: state-only (SC 1), parameter-only (SC 2), and joint state-parameter (SC 3).



Table 1. RMSE values between the JULES simulated soil moisture and COSMOS-UK observations.

Site	RMSE			
	Prior	Posterior state-only (SC 1)	Posterior parameter-only (SC 2)	Posterior state-parameter (SC 3)
CARDT	0.09	0.07	0.08	0.08
BICKL	0.14	0.08	0.03	0.03
CRICH	0.12	0.06	0.04	0.04
WADDN	0.10	0.06	0.07	0.07
HOLLN	0.17	0.10	0.06	0.06
EASTB	0.16	0.10	0.04	0.03
ROTHD	0.08	0.03	0.04	0.04
CHIMN	0.06	0.06	0.05	0.05
SHEEP	0.10	0.05	0.05	0.05
PORTN	0.11	0.06	0.07	0.06
HARTW	0.17	0.10	0.10	0.10
GISBN	0.23	0.17	0.09	0.07
CHOBH	0.28	0.21	0.20	0.20
LULLN	0.09	0.04	0.04	0.04
MOORH	0.25	0.17	0.13	0.11
SOURH	0.13	0.08	0.07	0.06

3.2 Scenario 2: Individual effect of parameter assimilation on JULES performance

305 During parameter-only assimilation, Cosby’s PTF constants were updated, leading to shifts in the posterior PTF constants compared to their respective prior values, as shown in Figure 5. This shift resulted in changes to the model’s hydraulic parameters, illustrated in Figure 6. For instance, the parameter V_{sat} increased due to higher values of K_4 and K_5 , suggesting an increase in the model’s representation of the soil’s water-holding capacity at full saturation. In contrast, satcon decreased due to changes in K_{10} to K_{12} values. This indicates that the model now simulates slower water movement through the soil when fully saturated, which may suggest less permeable soil conditions in the model. In addition, parameters sathh and b were reduced due to the decreased values of K_7 to K_9 and changes in K_1 to K_3 values, respectively. Finally, V_{crit} and V_{wilt} did not exhibit significant changes between prior and posterior distributions. Overall, the posterior distributions are narrower than the prior distributions, since the 4D-En-Var technique serves as an approximate minimum variance estimator.

310

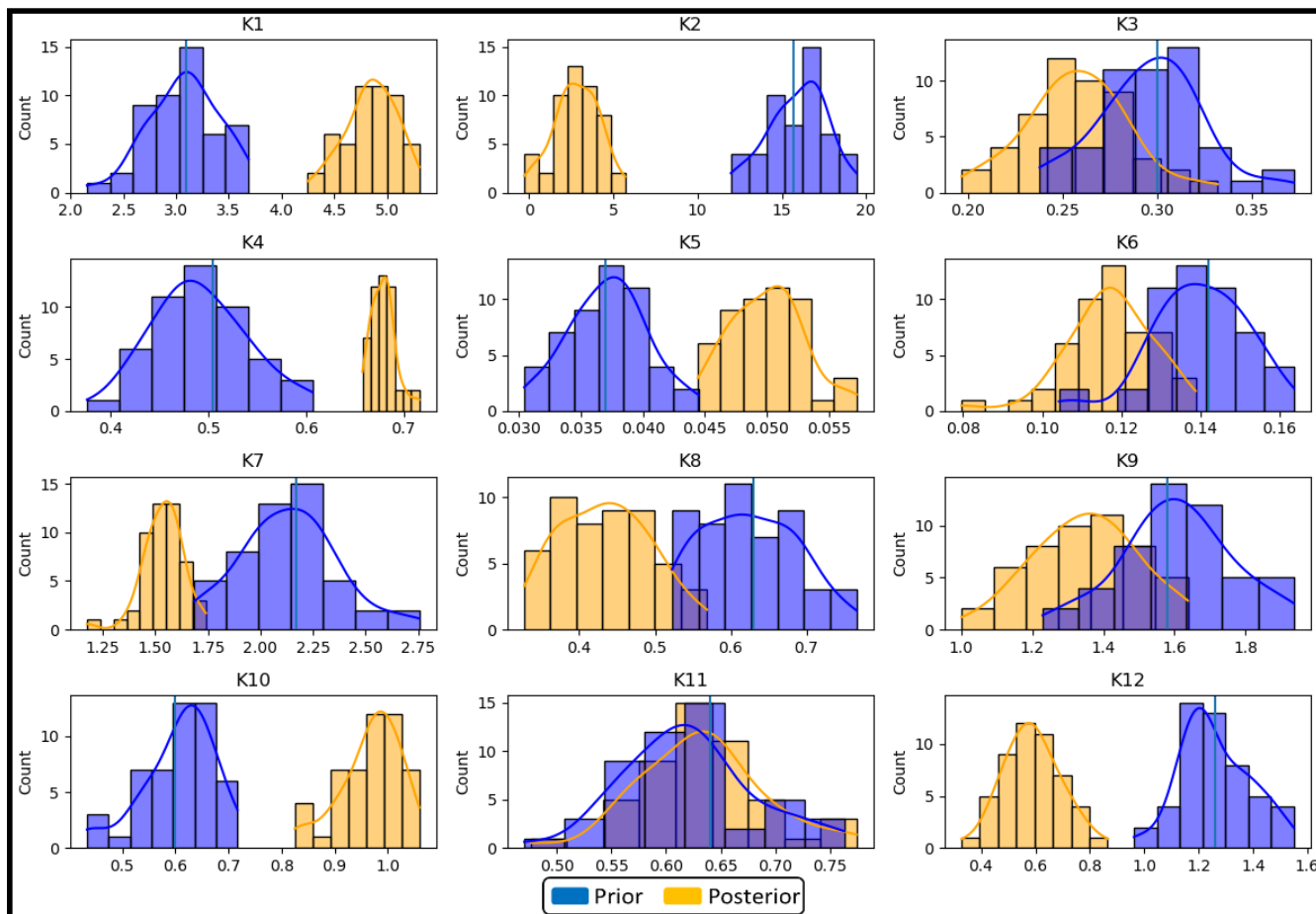


Figure 5. Histograms showing the prior and posterior distribution values of Cosby's PTF constants ($K1$ to $K12$) during parameter-only assimilation scenario (SC 2) across 16 sites. Blue (prior) and orange (posterior) lines represent kernel density estimates. Bin widths are individually adjusted to reflect the range and spread of values for each constant.

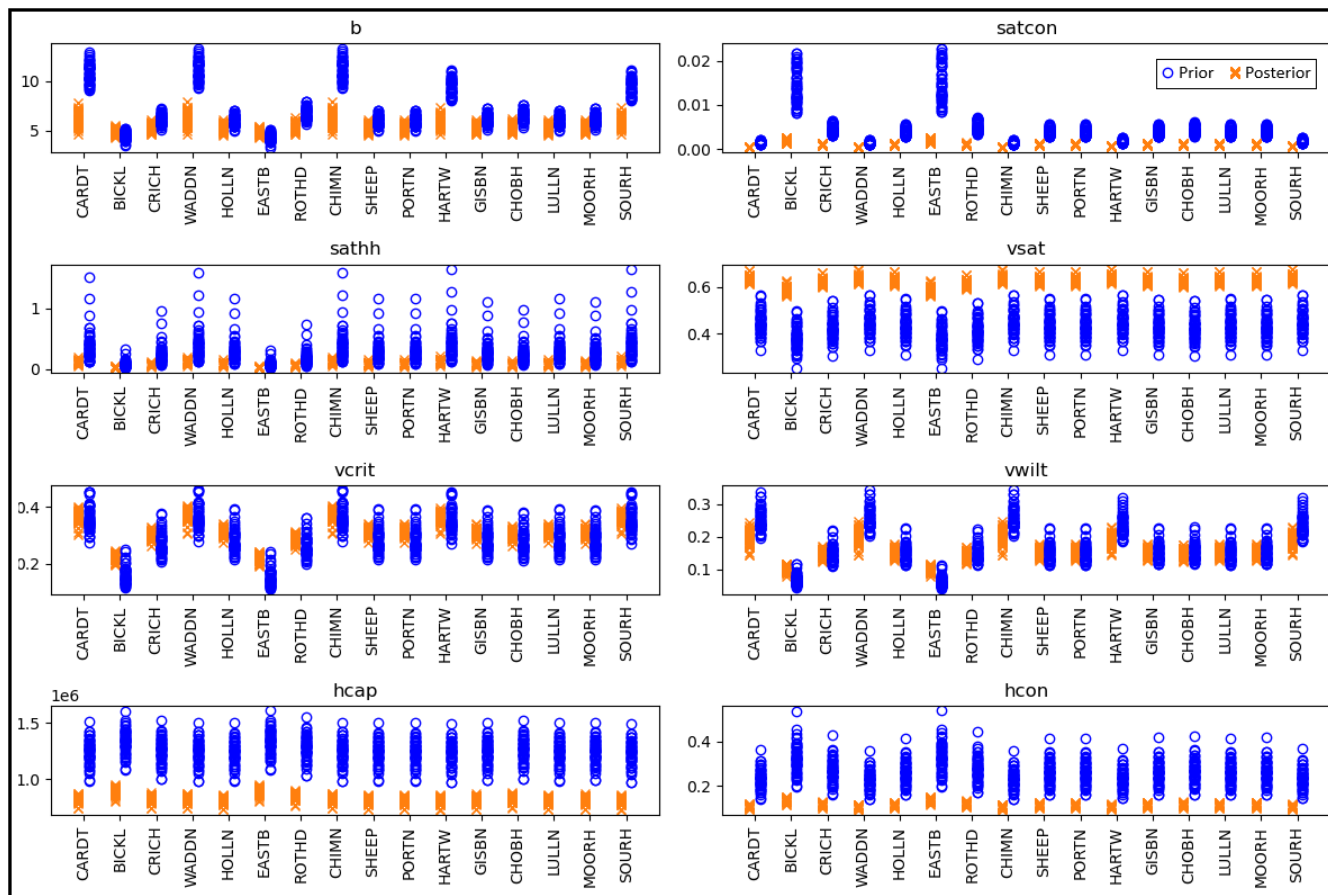


Figure 6. Ensemble prior and posterior values of different soil hydraulic parameters at 16 sites across the UK during parameter-only assimilation scenario (SC 2).

Figure 7 shows the ensemble mean JULES and COSMOS-UK soil moisture time series for the HOLLN and SHEEP sites, which have different soil characteristics. From the figure, it is evident that the posterior soil moisture time series is more consistent with the COSMOS-UK observations than the prior values. Visually, the model estimates exhibit better correlation and lower bias with the COSMOS-UK observations. KGE metrics and their components also showed a similar outcome, indicating improved performance for the posterior estimates (Figure 4). The reduction in bias was a major contributor to the improvement in KGE metrics. This is due to changes in the V_{sat} , b , and $satcon$ values, which allowed more water to be stored in the modelled soil, thereby reducing the underestimation in JULES soil moisture estimates. When we analyzed the differences in RMSE values (Table 1) between the prior and parameter-only scenarios, the posterior RMSE values showed decreased values, with an average reduction of 47.4% across all 16 sites.

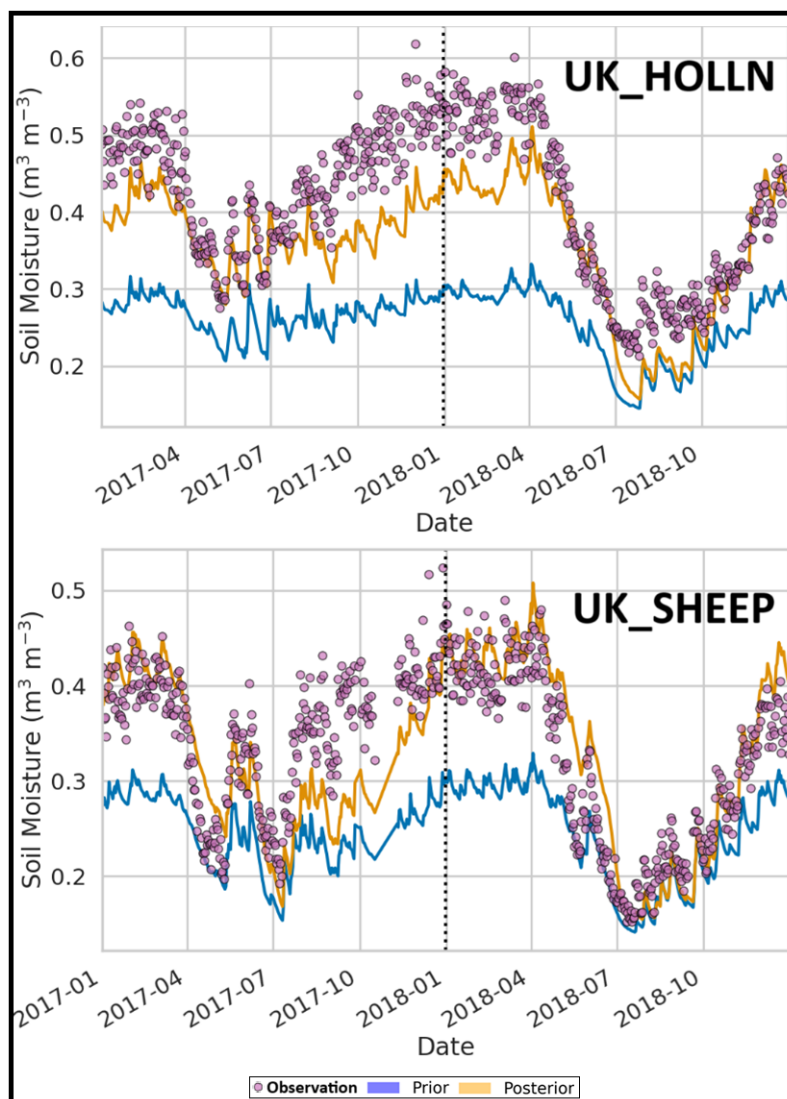


Figure 7. Modelled (ensemble mean) and observed soil moisture time series at HOLLN and SHEEP sites during assimilation (2017) and forecast periods (2018) for parameter-only assimilation scenario (SC 2).

3.3 Scenario 3: Combined effect of state-parameter assimilation on JULES performance

During joint state-parameter assimilation, we optimized both Cosby’s constants (Figure 8) and JULES’s initial soil moisture condition (Figure 10). Figure 8 shows that the posterior parameter distribution (orange) deviated from its respective prior distribution (blue) after assimilation, though the deviations are less pronounced compared to the parameter-only scenario. This suggests that updating the initial soil moisture conditions in this scenario has partially alleviated the adjustments required in the PTF constants. Due to this, all eight soil hydraulic properties in the Brooks and Corey model also showed deviations

325



from their respective prior values as shown in Figure 9. In general, the ensemble members span a narrower range in the
 330 posterior distribution than in the prior distribution, as seen in the parameter-only assimilation scenario, suggesting a reduction
 in variability of the hydraulic parameters after assimilation. The values of satcon were reduced due to the decreased values of
 K_{10} and K_{12} . Conversely, V_{sat} showed an increased value due to higher K_4 and K_5 values after assimilation. However, both
 V_{sat} and satcon exhibited smaller changes compared to the parameter-only scenario. Finally, b and sathh did not show much
 335 change in distribution after assimilation. These results indicate that the posterior JULES can store more water in the soil due
 to the combined effect of higher retention capacity and slower movement of water through the soil.

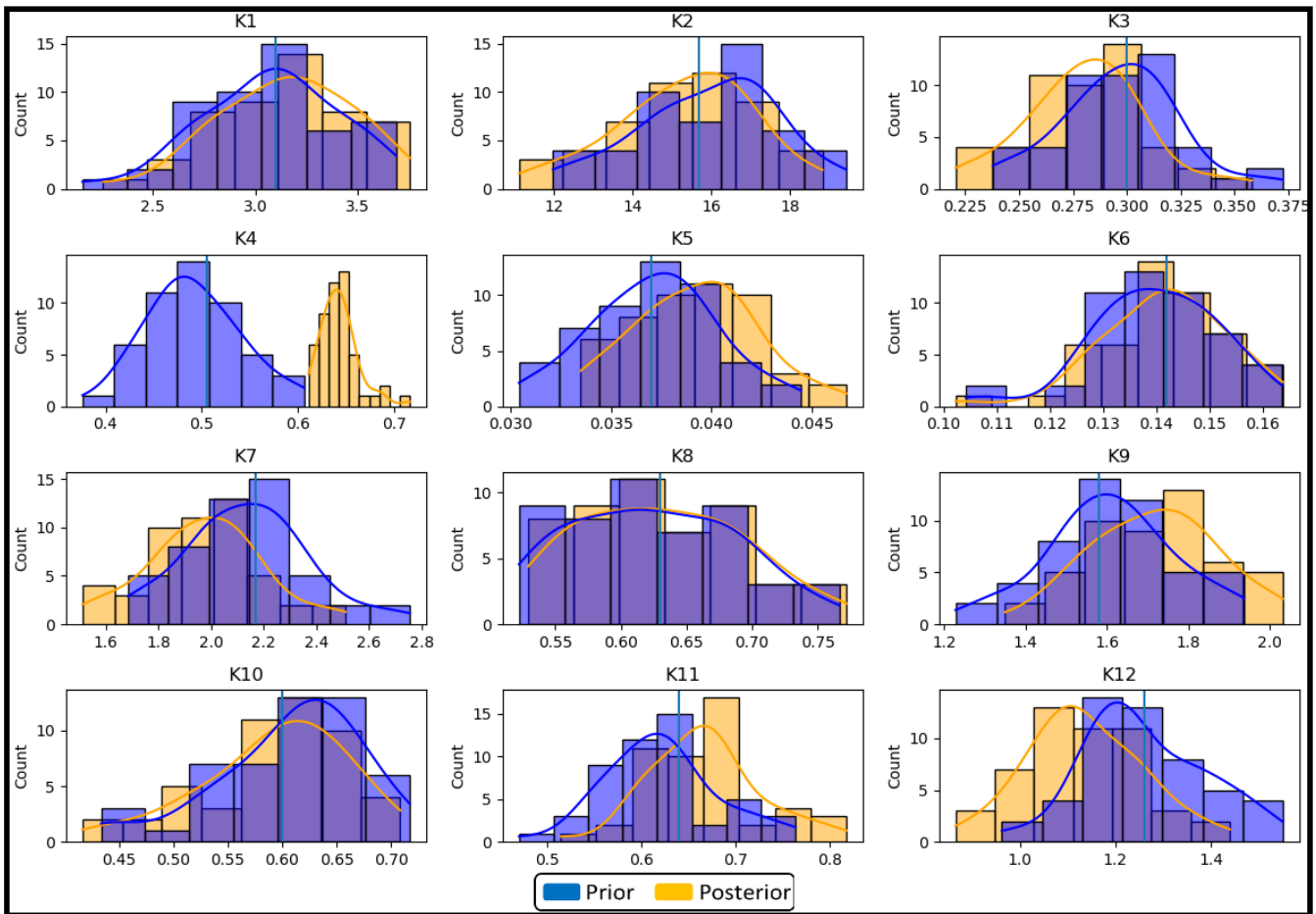


Figure 8. Histograms showing the prior and posterior distribution values of Cosby’s PTF constants (K_1 to K_{12}) during state-parameter assimilation (SC 3) across 16 sites. Blue (prior) and orange (posterior) lines represent kernel density estimates. Bin widths are individually adjusted to reflect the range and spread of values for each constant.

Similarly, Figure 10 shows the prior and posterior distribution of initial soil moisture conditions for two locations, GISBN and CARDT. These sites were selected to illustrate the assimilation’s impact on different soil types— GISBN with high organic



content and CARDT with mineral-rich soil. From the figure, we can see that the initial soil moisture condition has increased in GISBN and decreased in CARDT after assimilation. This is because, on January 1, 2017, the model's prior estimates were lower compared to the observations for the GISBN location and higher for the CARDT location. Accordingly, the initial condition was updated to be closer to the observation values. The prior and posterior distributions for all 16 locations are shown in Figures A1- A4 of Appendix A.

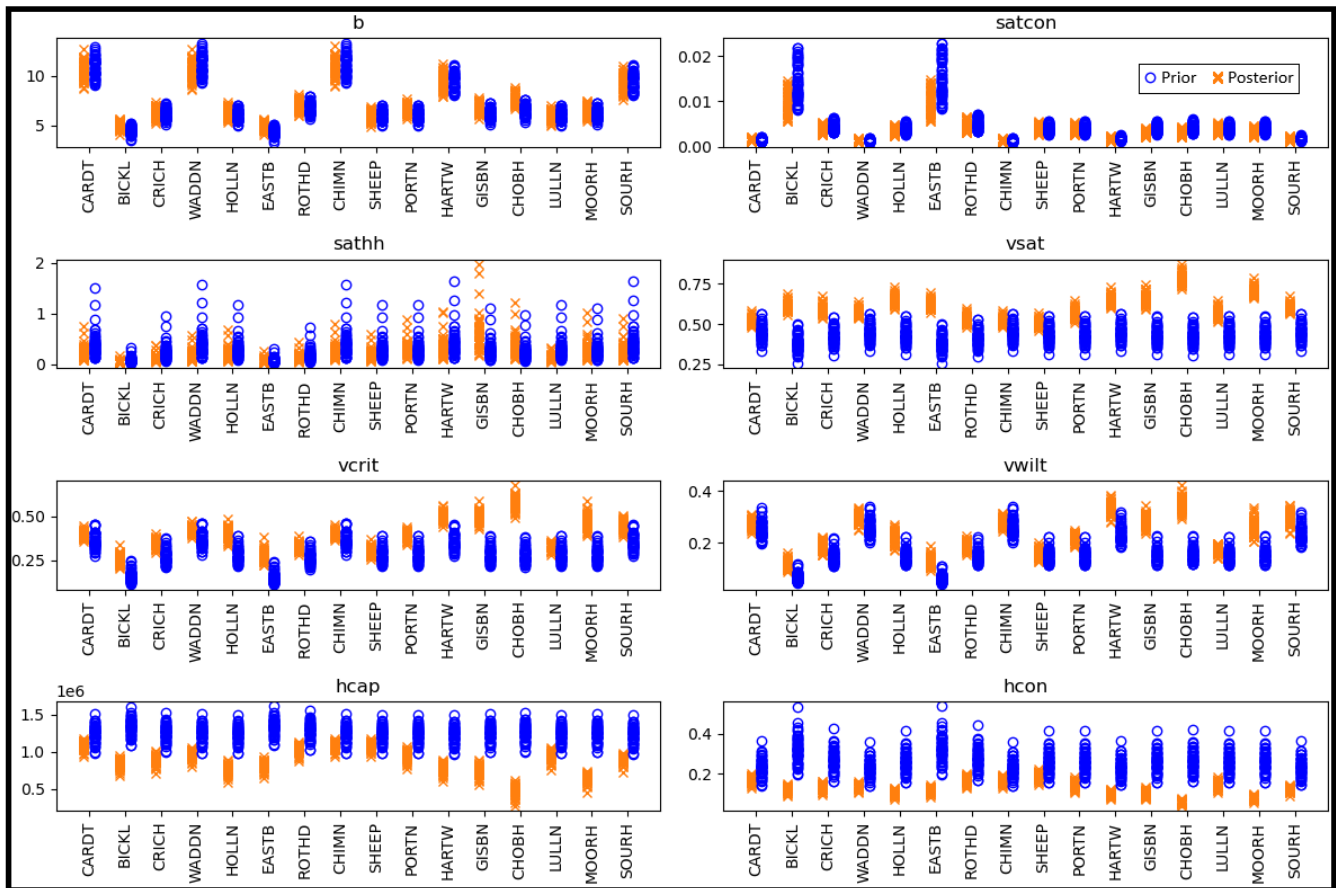


Figure 9. Ensemble prior and posterior values of different soil hydraulic parameters at 16 sites across the UK during state-parameter assimilation scenario (SC 3).

Figure 11 shows the ensemble mean soil moisture time series for two locations with different soil characteristics. Table 1 and Figure 11 show that the posterior soil moisture estimates are more consistent with the COSMOS-UK observations. RMSE reduced by an average of 50% across all sites, with the maximum reduction observed at BICKL (78.5%) followed by GISBN (69.5%). Overall, these results suggest that the bias in the model is significantly reduced, the posterior JULES estimates showed a good correlation with COSMOS-UK observations, and the variability in the observations is captured more accurately.

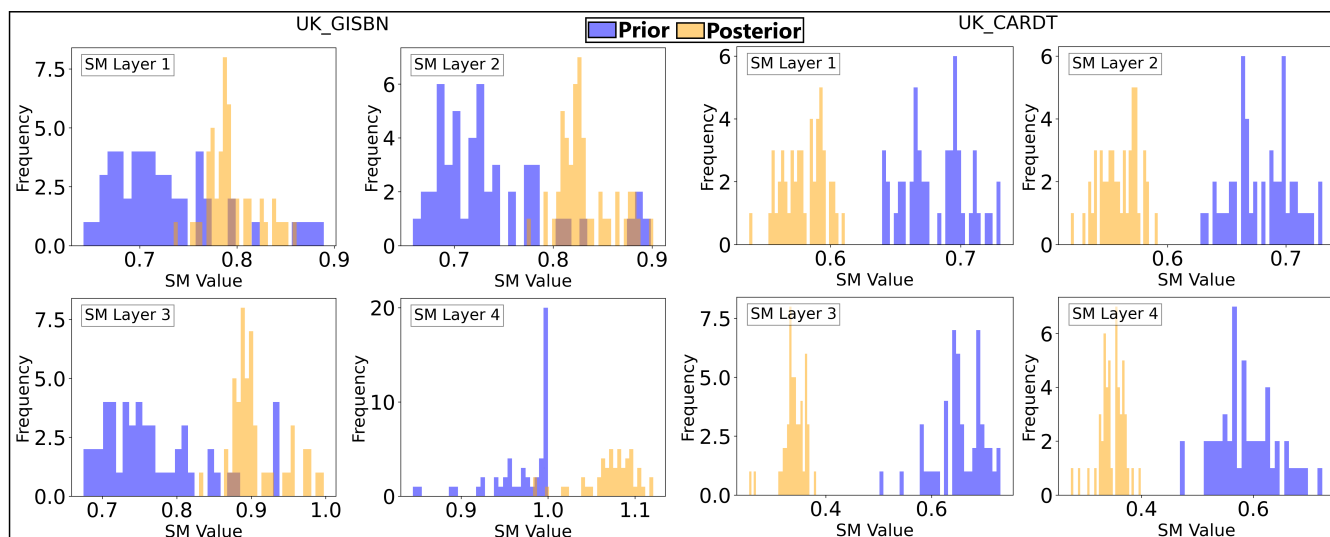


Figure 10. Prior (blue) and posterior (orange) distributions of JULES's initial soil moisture condition during state-parameter assimilation (SC 3) for two sites (GISBN and CARDT) across 50 ensemble members and four soil layers. Bin widths vary across subplots to allow comparison of prior and posterior value ranges for each soil layer.

KGE values showed significant improvement after assimilation, with the maximum increase observed at the GISBN site ($\approx 350\%$) and a minimum increase of $\approx 25\%$ at the CHIMN site (Figure 4). Breaking down the components of KGE, the combined effect of the α and β components contributed substantially to the increase in KGE. The r component was already high in the prior JULES estimates (averaging more than ≈ 0.8), leaving little scope for assimilation to improve the correlation with COSMOS-UK observations.

Out of the 16 locations, prior soil moisture estimates at 15 sites were negatively biased, as indicated by the β component (values less than 1, represented by a horizontal line). After assimilation, the hydraulic parameter satcon decreased, reducing the movement of water within the soil profile, while the parameter V_{sat} increased, improving the water holding capacity of the soil. This combined effect increased the water content in the soil column, reducing the underestimation of the prior JULES run. This improvement also enabled the model to better capture the variability in COSMOS-UK observations, leading to an improved α component.

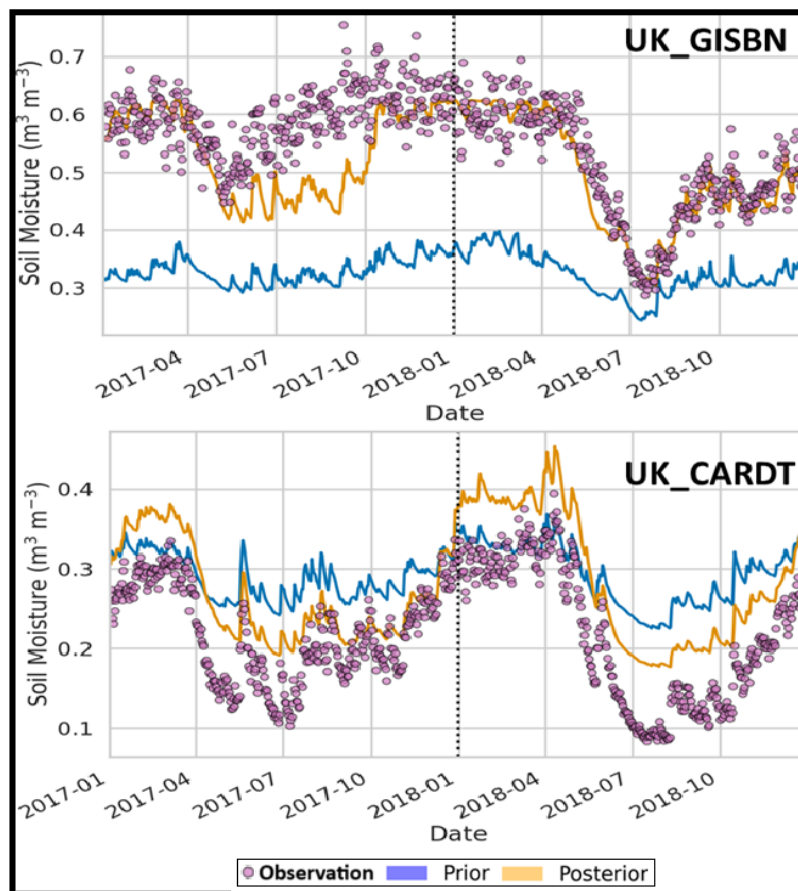


Figure 11. Modelled (ensemble mean) and observed soil moisture at GISBN and CARDT sites during assimilation (2017) and forecast periods (2018) for state-parameter assimilation scenario (SC 3).

4 Conclusions

360 In this work, we assessed the impact of assimilating field-scale soil moisture observations on the JULES community land surface model. We used daily COSMOS-UK observations for the year 2017 across 16 sites in the UK, employing the 4D-En-Var hybrid assimilation technique to update the state and/or parameter prior vector. While past studies focused on updating the initial condition of JULES (Seo et al., 2021) or by updating the underlying PTF functions (Pinnington et al., 2021; Cooper et al., 2021a), we performed state-parameter assimilation by simultaneously optimizing both the constants of the Pedotransfer function and JULES initial soil moisture condition.

365

To demonstrate the advantage of joint state-parameter assimilation over individual state-only or parameter-only assimilation, we designed three assimilation scenarios. Scenarios 1 and 2 determined the impact of optimizing initial soil moisture conditions and Cosby's PTF constants, respectively. In scenario 3, we assessed the impact of jointly optimizing both Cosby's PTF



constants and initial condition by performing joint state-parameter assimilation.

370 In general, results from all the experiments showed that it is possible to improve the performance of JULES by assimilating
field-scale soil moisture observations. Comparing the results of Scenario 3 with scenarios 1 and 2 demonstrated that state-
parameter assimilation can constrain the JULES model to perform much better. During Scenario 3, the average KGE value
across 16 sites increased from 0.33 to 0.72 after assimilation, representing an increase of 143%. In contrast, individual state-
only and parameter-only assimilation results showed only moderate improvements of 7.6% (KGE: 0.33 to 0.35) and 100%
375 (KGE: 0.33 to 0.66), respectively.

Soil characteristics have a strong influence on JULES performance. In general, the JULES model performs well in "typical
mineral soil" or "calcareous soil" but tends to show more errors in areas with high organic content. This may be due to the incor-
rect initialization of PTF constants and initial conditions during the warm-up period. Prior model results (Figure 4) demonstrate
this pattern, with an average KGE of 0.35 for "typical mineral soil" sites and 0.23 for "high organic content" sites. However,
380 assimilation results (scenario 3) showed maximum improvement at sites where JULES initially performed poorly. For instance,
if we rank the model improvement across the 16 sites, three of the top four improvements were seen at sites with high organic
content. GISBN, CHOBH, and MOORH showed improvements of approximately 350%, 250%, and 220%, respectively, in
KGE values. The post-assimilation model with updated PTF constants and initial conditions performed well across various soil
types within the UK, suggesting that this framework has strong potential for broader application within similar regions.

385 During state-only assimilation, the posterior JULES model performed marginally better than the prior run, with a mean im-
provement of 7.6% in KGE and 39.5% in RMSE. However, the effect of the updated initial soil moisture condition decayed
quickly in less than six months. For a few sites, the prior and posterior soil moisture time series coincided within three months
from the point of assimilation. This limitation can be addressed by assimilating more frequently using a smaller time window.
Having a longer time window of one year for parameter estimation and multiple shorter time windows of three months for state
390 assimilation could potentially yield better forecasts.

In conclusion, this study presents the first attempt at joint state-parameter assimilation on the JULES model, leading to signif-
icant improvements in soil moisture estimates. By optimizing both the initial conditions and PTF constants, we demonstrated
clear enhancements in model performance. These improvements have important implications for better flood forecasting and
disaster management, as more accurate soil moisture predictions can enhance preparedness and response strategies. This work
395 provides a valuable basis for further advancements in land surface data assimilation across different regions and soil types.

Code and data availability. JULES source code and the assimilation code used in this study are available from the MetOffice JULES repos-
itory (<https://code.metoffice.gov.uk/trac/jules>, last access: 30 August 2024) under Rose suite number u-dc794. Registration is required.
COSMOS-UK data are deposited annually in the NERC Environmental Information Data Centre (EIDC) ([https://doi.org/10.5285/b5c190e4-
e35d-40ea-8f8e-598da03a1185](https://doi.org/10.5285/b5c190e4-e35d-40ea-8f8e-598da03a1185), Stanley et al., 2021).



400 *Author contributions.* VR, EC, and SD devised the experiments. VR designed the Rose suite used here and ran the experiments. EC prepared the COSMOS-UK data. SD supervised the work and VR prepared the manuscript with inputs from all the co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was jointly supported by the Natural Environment Research Council's (NERC) HydroJULES-NEXT project (grant number: NE/X019063/1) and the Climate Change in the Arctic and North Atlantic Region and Impacts on the UK (CANARI) project
405 (grant number: NE/W004984/1) as a part of the National capability funding programmes. The authors gratefully acknowledge the provision by UKCEH of hydrometeorological and soil data collected by the COSMOS-UK project. COSMOS-UK is funded by the NERC (award no. NE/R016429/1) as part of the UK-SCAPE programme.



Appendix A: Updated JULES initial soil moisture condition during state-parameter scenario

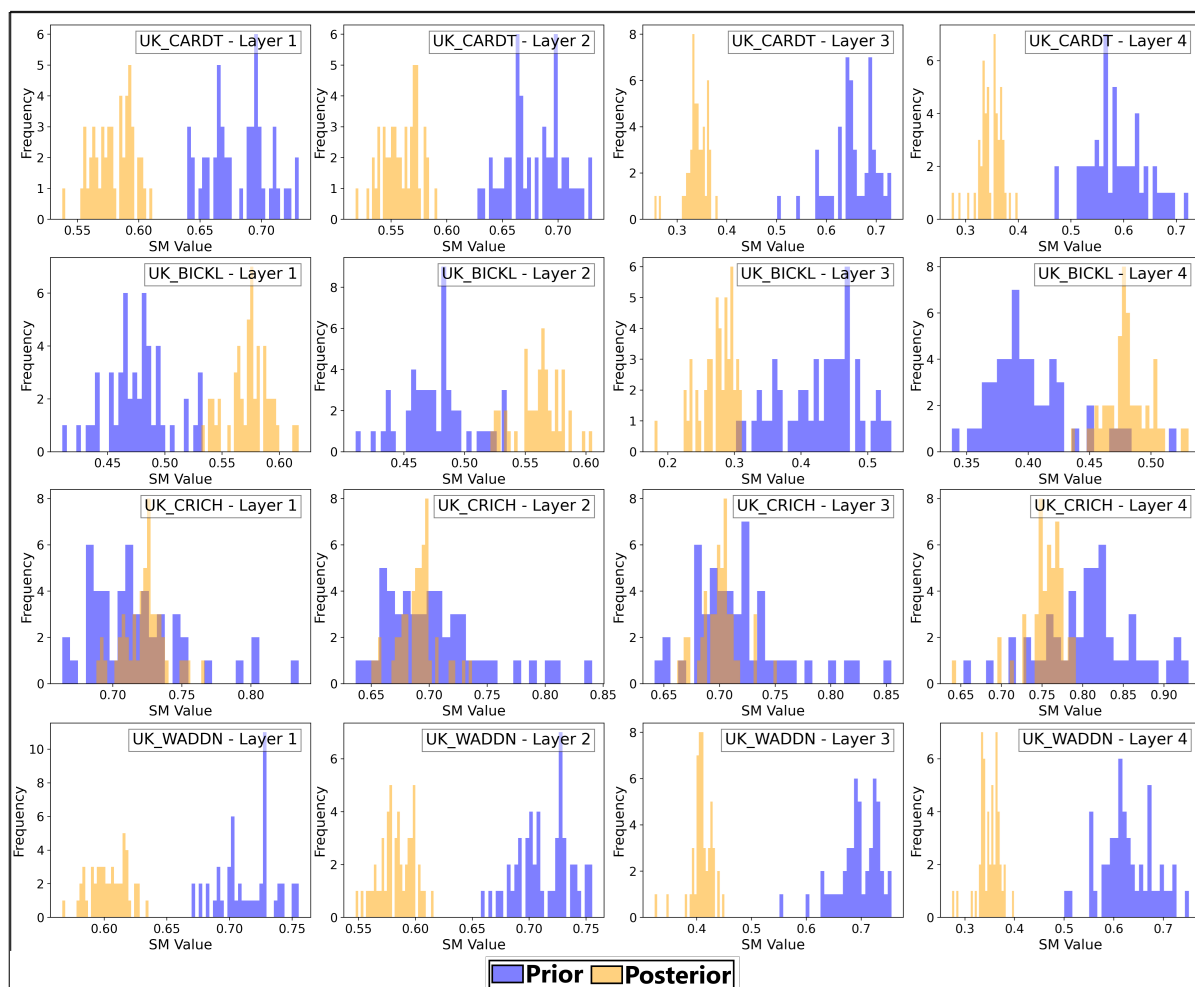


Figure A1. Prior and posterior distribution values of JULES’s initial soil moisture condition at CARDT, BICKL, CRICH, and WADDN COSMOS sites during state-parameter assimilation.

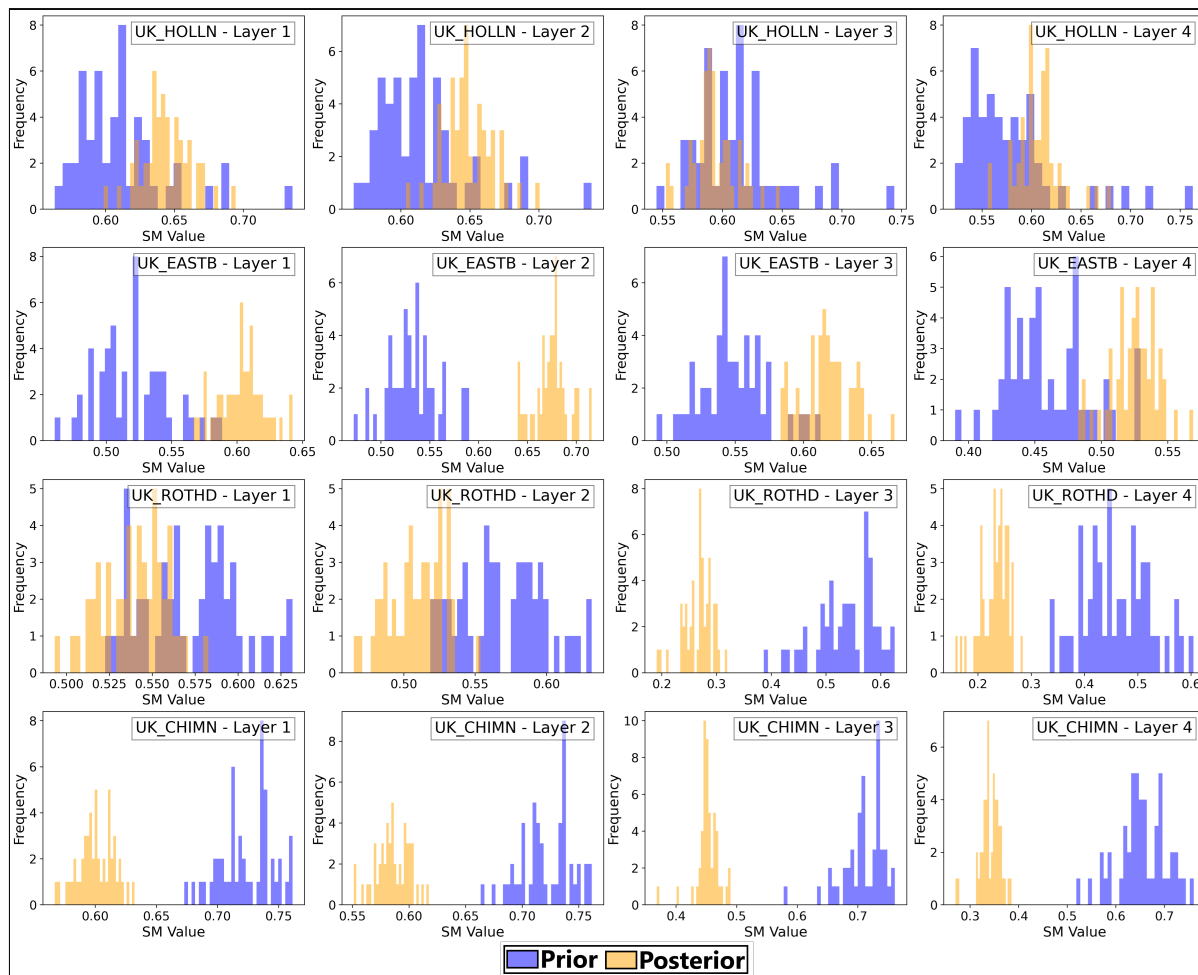


Figure A2. Prior and posterior distribution values of JULES’s initial soil moisture condition at HOLLN, EASTB, ROTH, and CHIMN COSMOS sites during state-parameter assimilation.

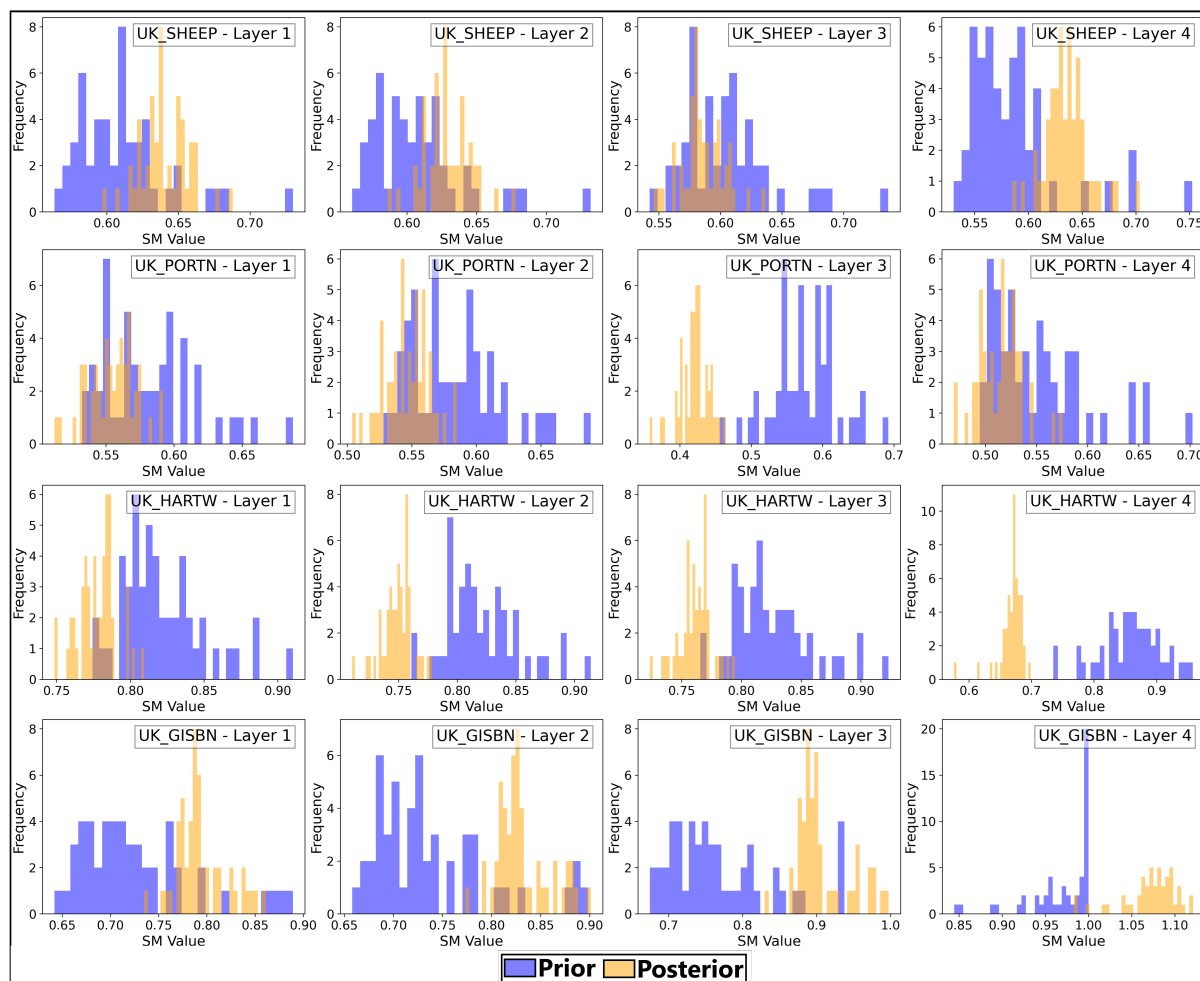


Figure A3. Prior and posterior distribution values of JULES's initial soil moisture condition at SHEEP, PORTN, HARTW, and GISBN COSMOS sites during state-parameter assimilation.

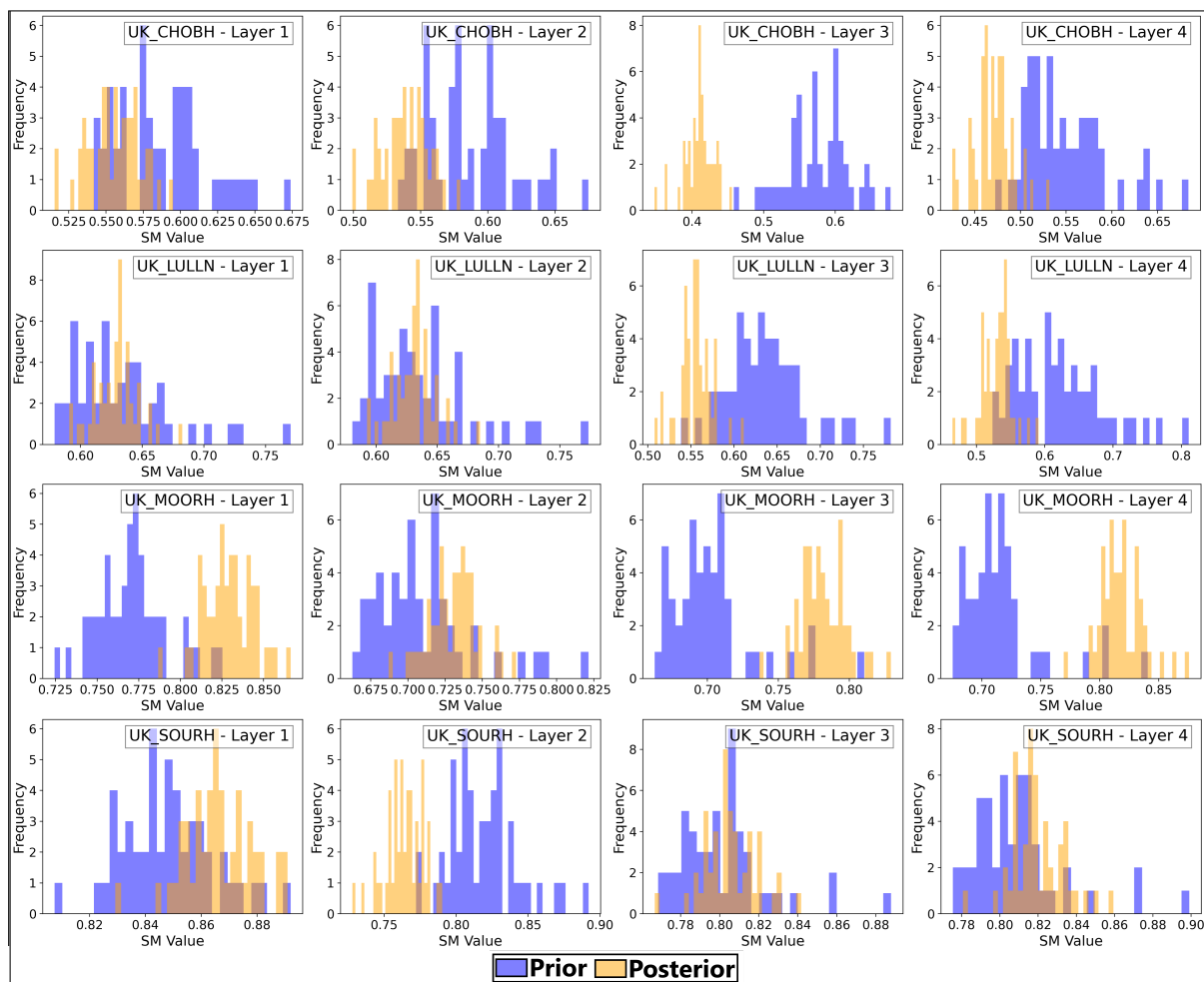


Figure A4. Prior and posterior distribution values of JULES's initial soil moisture condition at CHOBH, LULLN, MOORH, and SOURH COSMOS sites during state-parameter assimilation.



References

- 410 Albergel, C., de Rosnay, P., Balsamo, G., Isaksen, L., and Muñoz-Sabater, J.: Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations, *Journal of Hydrometeorology*, 13, 1442 – 1460, <https://doi.org/10.1175/JHM-D-11-0107.1>, 2012.
- Alvarado-Montero, R., Schwanenber, D., Krahe, P., Helmke, P., and Klein, B.: Multi-parametric variational data assimilation for hydrological forecasting, *Advances in Water Resources*, 110, 182–192, <https://doi.org/https://doi.org/10.1016/j.advwatres.2017.09.026>, 2017.
- Antoniou, V., Askquith-Ellis, A., Bagnoli, S., Ball, L., Bennett, E., Blake, J., Boorman, D., Brooks, M., Clarke, M., Cooper, H., Cowan, N., Cumming, A., Doughty, L., Evans, J., Farrand, P., Fry, M., Hewitt, N., Hitt, O., Jenkins, A., Kral, F., Libre, J., Lord, W., Roberts, C., Morrison, R., Parkes, M., Nash, G., Newcomb, J., Rylett, D., Scarlett, P., Singer, A., Stanley, S., Swain, O., Thornton, J., Trill, E., Vincent, H., Ward, H., Warwick, A., Winterbourn, B., and Wright, G.: COSMOS-UK user guide: users' guide to sites, instruments and available data (version 2.10), <https://doi.org/http://nora.nerc.ac.uk/id/eprint/524801/>, 2019.
- Bannister, R. N.: A review of operational methods of variational and ensemble-variational data assimilation, *Quarterly Journal of the Royal Meteorological Society*, 143, 607–633, <https://doi.org/https://doi.org/10.1002/qj.2982>, 2017.
- 420 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.
- 425 Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, 6, 279–298, <https://doi.org/https://doi.org/10.1002/hyp.3360060305>, 1992.
- Brooks, R. H. and Corey, A. T.: Hydraulic properties of porous media, *Hydrological Papers*, 3, colorado State Univ., Fort Collins, 1964.
- Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., and Pignatelli, B.: In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, in: 2007 IEEE International Geoscience and Remote Sensing Symposium, pp. 1196–1199, <https://doi.org/10.1109/IGARSS.2007.4423019>, 2007.
- 430 Chen, W., Huang, C., Shen, H., and Li, X.: Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation, *Advances in Water Resources*, 86, 425–438, <https://doi.org/https://doi.org/10.1016/j.advwatres.2015.08.003>, 2015.
- 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.
- 435 Cooper, E., Blyth, E., Cooper, H., Ellis, R., Pinnington, E., and Dadson, S. J.: Using data assimilation to optimize pedotransfer functions using field-scale in situ soil moisture observations, *Hydrology and Earth System Sciences*, 25, 2445–2458, <https://doi.org/10.5194/hess-25-2445-2021>, 2021a.
- Cooper, E. S., Dance, S. L., García-Pintado, J., Nichols, N. K., and Smith, P. J.: Observation operators for assimilation of satellite observations in fluvial inundation forecasting, *Hydrology and Earth System Sciences*, 23, 2541–2559, <https://doi.org/10.5194/hess-23-2541-2019>, 2019.
- 440 Cooper, H. M., Bennett, E., Blake, J., Blyth, E., Boorman, D., Cooper, E., Evans, J., Fry, M., Jenkins, A., Morrison, R., Rylett, D., Stanley, S., Szczykulska, M., Trill, E., Antoniou, V., Askquith-Ellis, A., Ball, L., Brooks, M., Clarke, M. A., Cowan, N., Cumming, A., Farrand, P., Hitt, O., Lord, W., Scarlett, P., Swain, O., Thornton, J., Warwick, A., and Winterbourn, B.: COSMOS-UK: national soil moisture and



- 445 hydrometeorology data for environmental science research, *Earth System Science Data*, 13, 1737–1757, <https://doi.org/10.5194/essd-13-1737-2021>, 2021b.
- 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/https://doi.org/10.1029/WR020i006p00682>, 1984.
- 450 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.
- de Roos, S., Bechtold, M., Busschaert, L., Lievens, H., and De Lannoy, G. J. M.: Assimilation of Sentinel-1 Backscatter to Update AquaCrop Estimates of Soil Moisture and Crop Biomass, *Journal of Geophysical Research: Biogeosciences*, 129, e2024JG008231, <https://doi.org/https://doi.org/10.1029/2024JG008231>, e2024JG008231 2024JG008231, 2024.
- 455 de Rosnay, P., Calvet, J., Kerr, Y. H., Wigneron, J.-P., Lemaître, F., Escorihuela, M.-J., Muñoz Sabater, J., Saleh Contell, K., Barrié, J., Bouhours, G., Coret, L., Cherel, G., Dedieu, G., Durbe, R., Froissard, F., Hoedjes, J., Kruszewski, A., Lavenu, F., Suquia, D., and Waldteufel, P.: SMOSREX : A Long Term Field Campaign Experiment for Soil Moisture and Land Surface Processes Remote Sensing, *Remote Sensing of Environment*, 102, 377–389, <https://doi.org/10.1016/j.rse.2006.02.021>, 2006.
- Dumedah, G. and Walker, J. P.: Intercomparison of the JULES and CABLE land surface models through as-
460 simulation of remotely sensed soil moisture in southeast Australia, *Advances in Water Resources*, 74, 231–244, <https://doi.org/https://doi.org/10.1016/j.advwatres.2014.09.011>, 2014.
- 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*,
465 98, 704–716, <https://doi.org/10.1109/JPROC.2010.2043918>, 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/https://doi.org/10.1002/hyp.10929>, 2016.
- 470 Fischer, G., F. N., S. P., Van Velthuisen, H., Verelst, L., and Wiberg, D.: harmonized-world-soil-database-v12, Global agro-ecological zones assessment for agriculture (GAEZ 2008), IIASA, Laxenburg, Austria and FAO, Rome, Italy, 10, <https://doi.org/http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/>, 2008.
- Gong, J., Liu, X., Yao, C., Li, Z., Weerts, A., Li, Q., Bastola, S., Huang, Y., and Xu, J.: State updating in the Xin’anjiang Model: Joint assimilating streamflow and multi-source soil moisture data via Asynchronous Ensemble Kalman Filter with enhanced Error Models,
475 *Hydrology and Earth System Sciences Discussions*, 2024, 1–38, <https://doi.org/10.5194/hess-2024-211>, 2024.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, 377, 80–91, <https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.08.003>, 2009.
- Heyvaert, Z., Scherrer, S., Dorigo, W., Bechtold, M., and De Lannoy, G.: Joint assimilation of satellite-based surface
480 soil moisture and vegetation conditions into the Noah-MP land surface model, *Science of Remote Sensing*, 9, 100129, <https://doi.org/https://doi.org/10.1016/j.srs.2024.100129>, 2024.



- Hou, Z. and Pu, Z.: Assessing CYGNSS Satellite Soil Moisture Data for Drought Monitoring with Multiple Datasets and Indicators, *Remote Sensing*, 16, <https://doi.org/10.3390/rs16010116>, 2024.
- Hung, C. P., Schalge, B., Baroni, G., Vereecken, H., and Hendricks Franssen, H.-J.: Assimilation of Groundwater Level and Soil Moisture Data in an Integrated Land Surface-Subsurface Model for Southwestern Germany, *Water Resources Research*, 58, e2021WR031549, <https://doi.org/https://doi.org/10.1029/2021WR031549>, e2021WR031549 2021WR031549, 2022.
- 485 Kerr, Y., Waldteufel, P., Wigneron, J.-P., 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, <https://doi.org/10.1109/36.942551>, 2001.
- 490 Kim, H., Lakshmi, V., Kwon, Y., and Kumar, S. V.: First attempt of global-scale assimilation of subdaily scale soil moisture estimates from CYGNSS and SMAP into a land surface model, *Environmental Research Letters*, 16, 074041, <https://doi.org/10.1088/1748-9326/ac0ddf>, 2021.
- Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, *Hydrology and Earth System Sciences*, 23, 4323–4331, <https://doi.org/10.5194/hess-23-4323-2019>, 2019.
- 495 Kumar, S., Kolassa, J., Reichle, R., Crow, W., de Lannoy, G., de Rosnay, P., MacBean, N., Giroto, M., Fox, A., Quaife, T., Draper, C., Forman, B., Balsamo, G., Steele-Dunne, S., Albergel, C., Bonan, B., Calvet, J.-C., Dong, J., Liddy, H., and Ruston, B.: An Agenda for Land Data Assimilation Priorities: Realizing the Promise of Terrestrial Water, Energy, and Vegetation Observations From Space, *Journal of Advances in Modeling Earth Systems*, 14, e2022MS003259, <https://doi.org/https://doi.org/10.1029/2022MS003259>, e2022MS003259 2022MS003259, 2022.
- 500 Lü, H., Yu, Z., Zhu, Y., Drake, S., Hao, Z., and Sudicky, E. A.: Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation, *Advances in Water Resources*, 34, 395–406, <https://doi.org/https://doi.org/10.1016/j.advwatres.2010.12.005>, 2011.
- 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>, 2014.
- 505 Mason, D., Garcia-Pintado, J., Cloke, H., Dance, S., and Munoz-Sabater, J.: Assimilating high resolution remotely sensed soil moisture into a distributed hydrologic model to improve runoff prediction, <https://doi.org/10.21957/5isuz4a91>, 2020.
- Nearing, G., Yatheendradas, S., Crow, W., Zhan, X., Liu, J., and Chen, F.: The Efficiency of Data Assimilation, *Water Resources Research*, 54, 6374–6392, <https://doi.org/https://doi.org/10.1029/2017WR020991>, 2018.
- 510 Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, S. T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, *IEEE Trans. Geosci. Remote. Sens.*, 41, 215–229, <https://api.semanticscholar.org/CorpusID:10317485>, 2003.
- 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.
- 515 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., Amezcua, J., Cooper, E., Dadson, S., Ellis, R., Peng, J., Robinson, E., Morrison, R., Osborne, S., and Quaife, T.: Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of SMAP satellite data, *Hydrology and Earth System Sciences*, 25, 1617–1641, <https://doi.org/10.5194/hess-25-1617-2021>, 2021.



- 520 Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I. L., Nichols, N. K., Wilkinson, M., and Quaife, T. L.: Understanding the effect of disturbance from selective felling on the carbon dynamics of a managed woodland by combining observations with model predictions, *Journal of Geophysical Research: Biogeosciences*, 122, 886–902, <https://doi.org/10.1002/2017JG003760>, 2017.
- Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., and Fensholt, R.: Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark, *Water Resources Research*, 50, 8962–8981, <https://doi.org/10.1002/2014WR015392>, 2014.
- 525 Roy, A., Narvekar, P., Murtugudde, R., Shinde, V., and Ghosh, S.: Short and Medium Range Irrigation Scheduling Using Stochastic Simulation-Optimization Framework With Farm-Scale Ecohydrological Model and Weather Forecasts, *Water Resources Research*, 57, e2020WR029004, <https://doi.org/10.1029/2020WR029004>, e2020WR029004 2020WR029004, 2021.
- Sadhvani, K. and Eldho, T.: Assessing the Effect of Future Climate Change on Drought Characteristics and Propagation from Meteorological to Hydrological Droughts—A Comparison of Three Indices, *Water Resources Management*, 38, 441–462, <https://doi.org/10.1007/s11269-023-03679-7>, 2024.
- 530 Seo, E., Lee, M.-I., and Reichle, R. H.: Assimilation of SMAP and ASCAT soil moisture retrievals into the JULES land surface model using the Local Ensemble Transform Kalman Filter, *Remote Sensing of Environment*, 253, 112 222, <https://doi.org/10.1016/j.rse.2020.112222>, 2021.
- 535 Shi, Y., Davis, K. J., Zhang, F., Duffy, C. J., and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: A multivariate real-data experiment, *Advances in Water Resources*, 83, 421–427, <https://doi.org/10.1016/j.advwatres.2015.06.009>, 2015.
- Smith, P. J., Thornhill, G. D., Dance, S. L., Lawless, A. S., Mason, D. C., and Nichols, N. K.: Data assimilation for state and parameter estimation: application to morphodynamic modelling, *Quarterly Journal of the Royal Meteorological Society*, 139, 314–327, <https://doi.org/10.1002/qj.1944>, 2013.
- 540 Stanley, S.;Antoniou, V.-E. A. L. E. J. D. M. M. H. N. A. J. P. M. D.-C. H. T. D. W. R. G. F. D. P. R. O. M. S. J. E. H. A. J.: Daily and sub-daily hydrometeorological and soil data (2013–2022) [COSMOS-UK], <https://doi.org/10.5285/5060cc27-0b5b-471b-86eb-71f96da0c80f>, 2023.
- Tian, S., Renzullo, L. J., Pipunic, R. C., Lerat, J., Sharples, W., and Donnelly, C.: Satellite soil moisture data assimilation for improved operational continental water balance prediction, *Hydrology and Earth System Sciences*, 25, 4567–4584, <https://doi.org/10.5194/hess-25-4567-2021>, 2021.
- 545 Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L’Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., and Rostan, F.: GMES Sentinel-1 mission, *Remote Sensing of Environment*, 120, 9–24, <https://doi.org/10.1016/j.rse.2011.05.028>, the Sentinel Missions - New Opportunities for Science, 2012.
- 550 Visweshwaran, R., Ramsankaran, R., Eldho, T. I., and Jha, M. K.: Hydrological Impact Assessment of Future Climate Change on a Complex River Basin of Western Ghats, India, *Water*, 14, <https://doi.org/10.3390/w14213571>, 2022a.
- Visweshwaran, R., Ramsankaran, R., Eldho, T. I., and Lakshmiarahan, S.: Sensitivity-Based Soil Moisture Assimilation for Improved Streamflow Forecast Using a Novel Forward Sensitivity Method (FSM) Approach, *Water Resources Research*, 58, e2021WR031092, <https://doi.org/10.1029/2021WR031092>, e2021WR031092 2021WR031092, 2022b.
- 555 Visweshwaran, R., Ramsankaran, R., and Eldho, T.: Improving modelled streamflow using time-varying multivariate assimilation of remotely sensed soil moisture and in-situ streamflow observations, *Advances in Water Resources*, 186, 104676, <https://doi.org/10.1016/j.advwatres.2024.104676>, 2024.



- 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., and Steinnocher, K.: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications, *Meteorologische Zeitschrift*, 22, 5–33, <https://doi.org/10.1127/09412948/2013/0399>, 2013.
- Walker, J. P. and Houser, P. R.: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *Journal of Geophysical Research: Atmospheres*, 106, 11 761–11 774, <https://doi.org/https://doi.org/10.1029/2001JD900149>, 2001.
- 565 Xu, Y., Calvet, J.-C., and Bonan, B.: The joint assimilation of satellite observed LAI and soil moisture for the global root zone soil moisture production and its impact on land surface and ecosystem variables, *Agricultural and Forest Meteorology*, 360, 110 299, <https://doi.org/https://doi.org/10.1016/j.agrformet.2024.110299>, 2025.
- 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.
- 570 Zha, Y., Zhu, P., Zhang, Q., Mao, W., and Shi, L.: Investigation of Data Assimilation Methods for Soil Parameter Estimation with Different Types of Data, *Vadose Zone Journal*, 18, 190 013, <https://doi.org/https://doi.org/10.2136/vzj2019.01.0013>, 2019.