Land Surface Data Assimilation

R. Visweshwaran

PDRA, NCEO and University of Reading

Prof. Sarah Dance (NCEO, and UoR)

Dr. Elizabeth Cooper (UKCEH)

DARC-UoR: Data Assimilation Training





09/06/2025





National Centre for Earth Observation

Land Surface Processes



Figure: Different components of land surface processes. Bonan (2008)

Land surface processes describe the exchange of energy, water, and carbon between the land and the atmosphere.









Evolution of Land Surface Models (LSM)



Land Surface Data Assimilation (LSDA)



Optimizing Soil Moisture Assimilation Strategies for Improved Hydrological Predictions in JULES Model





09/06/2025





Part 1: JULES Land Surface Model

Joint UK Land Environment Simulator (JULES)

By Met Office Surface Exchange Scheme (MOSES)

Soil water content for each layer

$$\frac{dV_k}{dt} = W'_{k-1} - W'_k - E'_k - B'_k = 0$$

 \pmb{V}_k represent soil water content at layer k

 W_k^\prime represent diffusive fluxes determined through Darcy's law

 E'_k is SM extraction which follows exponential distribution

 $-\frac{R_k}{k}$ is subsurface lateral runoff (i.e. 1D Jules)





0.1m 0.25m 0.65m 2.0m Figure: Soil layer configuration





National Centre for Earth Observation

Part 1: JULES - Soil Hydraulics Module

Flow in the unsaturated zone: **Brooks and Corey**

Data Assimilation

esearch Centre

$$S_{e} = \left(\frac{S - S_{wo}}{1 - S_{wo}}\right) = \left(\frac{\Psi}{\Psi_{s}}\right)^{-1/b}$$

$$S = S_{wo} + (1 - S_{wo})\left(\frac{\Psi}{\Psi_{s}}\right)^{-1/b}$$

$$W = V_{wo} + (V_{sat} - V_{wo})\left(\frac{\Psi}{\Psi_{s}}\right)^{-1/b} \text{ and } \frac{K}{K_{s}} = \left(\frac{V}{V_{wo}}\right)^{2b+3}$$

$$0$$
Soil characteristic curve

 V_{wo} : Irreducible water content V_{sat} : Saturated water content Ψ : Capillary pressure S_e : Effective saturation

University of

Reading

 Ψ_s : Bubbling pressure (AEP) *b*: Fitting coefficient

K: Hydraulic conductivity

 $K_{\rm s}$:Hydraulic conductivity at saturation

0

V: Soil water content





1

7

Part 1: PedoTransfer Function (PTF)

Easily Measurable Variables



$$b = K_{1} + K_{2} f_{Clay} - K_{3} f_{Sand}$$

$$V_{sat} = K_{4} - K_{5} f_{Clay} - K_{6} f_{Sand}$$

$$\Psi_{s} = 0.01 * 10^{(K_{7} - K_{8} f_{Clay} - K_{9} f_{Sand}}$$

$$K_{s} = 10^{(-K_{10} - K_{11} f_{Clay} + K_{12} f_{Sand})}$$

$$V_{crit} = V_{sat} \left(\frac{\Psi_{s}}{3.364}\right)^{\frac{1}{b}}$$

$$V_{wilt} = V_{sat} \left(\frac{\Psi_{s}}{152.9}\right)^{\frac{1}{b}}$$

Difficult-to-Measure Variables

Cosby PTF

 $h_{cap} = (1 - V_{sat}) (2.376 * 10^6 f_{Clay} + 2.133 * 10^6 f_{Silt})$ $h_{con} = 0.025^{V_{sat}} \left(1.16^{f_{Clay}(1 - V_{sat})} * 1.57^{f_{sand}(1 - V_{sat})} * 1.57^{f_{silt}(1 - V_{sat})} \right)$

Cosby, B. J. et, al., (1984). Water Resources Research, https://doi.org/10.1029/WR020i006p00682,1984

Pedotransfer function



Food and Agricultus Organization of the United Nations



8





Part 2: Ensemble Variational (En-Var) Hybrid Assimilation



To avoid computing the AM and reduce the computational time of estimating B^{-1} , a preconditioned matrix is used using **control variable transform**

$$B \approx X_b' X_b'^T \quad \text{and} \quad x_0 = \overline{x_b} + X_b' w$$
$$J(w) = \frac{1}{2} w^T w + \sum_{t=0}^N \left(h_t \left(m_t(\overline{x_b}) \right) + H_t M_t X_b' w - y_t \right)^T R_t^{-1} \left(h_t \left(m_t(\overline{x_b}) \right) + H_t M_t X_b' w - y_t \right)$$
$$\nabla J(w) = w + \sum_{t=0}^N (H_t M_t X_b')^T R_t^{-1} \left(h_t \left(m_t(\overline{x_b}) \right) + H_t M_t X_b' w - y_t \right)$$

Perturbation matrix in Observation space

$$\boldsymbol{H}_{t}\boldsymbol{M}_{t}\boldsymbol{X}_{b}^{\prime} \approx \boldsymbol{Y}_{b}^{\prime} = \frac{1}{\sqrt{m-1}} \begin{pmatrix} h_{t}\left(\boldsymbol{m}_{t}\left(\boldsymbol{X}_{b}^{1}\right)\right) - h_{t}\left(\boldsymbol{m}_{t}(\overline{\boldsymbol{x}_{b}})\right), \ h_{t}\left(\boldsymbol{m}_{t}\left(\boldsymbol{X}_{b}^{2}\right)\right) - h_{t}\left(\boldsymbol{m}_{t}(\overline{\boldsymbol{x}_{b}})\right), \ \dots, \\ \dots, \ h_{t}\left(\boldsymbol{m}_{t}(\boldsymbol{X}_{b}^{m})\right) - h_{t}\left(\boldsymbol{m}_{t}(\overline{\boldsymbol{x}_{b}})\right) \end{pmatrix}$$



4D Var: Cost

function

Ensemble model run



En-Var

DA





Part 3: COSMOS-UK Observations



Temporal resolution: Daily Spatial resolution: Field scale across UK Duration: Three years 2016 (Warm-up period) 2017 (DA run) and 2018 (Forecast period)







Figure: Typical COSMOS-UK observation site (<u>https://cosmos.ceh.ac.uk/</u>)





National Centre for Earth Observation

Site Locations







Assimilation Scenarios

Three Scenarios

1) State estimation of initial soil moisture condition

2) Parameter estimation of PTF constants

3) Dual state-parameter estimation

 X_{h}^{l} = Augmented

state-parameter

vector

<u>SC 3</u>











Methodology

Dual state-parameter estimation



PTF parameter: Prior-Posterior Distributions



KGE Metrics



Figure: 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).

1. r (Correlation): JULES is already showing very good trend leaving less scope to improve them further after DA

2. α (Ratio of variability :- $\frac{\sigma_{model}}{\sigma_{obs}}$): This component influenced the KGE values maximum 3. β (Ratio of Bias :- $\frac{\mu_{model}}{\mu_{obs}}$): Increased V_sat and decreased K_s values after assimilation retained more water in the soil column reducing the bias difference.





15

Hydro-JU



Figure: 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)

Observation Prior Posterior



National Centre for Earth Observation

Scenario comparison



Hydraulic parameter changes are more pronounced in SC2, leading to greater variability in JULES outputs. Incorporating initial conditions moderated this effect, enabling JULES to align more closely with COSMOS observations.







Limitations



Future directions

Ultimate Goal: Produce a KM scale soil moisture product for UK using JULES land surface model

 Potentially to be used by UKCEH and Met-office for their hydrological outlook portal



Fig: UKCEH's Hydrological Outlook Map

Step 1: Perform state-parameter assimilation on JULES for 16 COSMOS-UK sites across UK (<u>https://doi.org/10.5194/egusphere-2024-3980</u>) **Step 2:** Identify optimal assimilation window length and perform cyclic update of JULES initial soil moisture condition (Ongoing) **Step 3:** Multivariate assimilation of soil moisture and streamflow observations within state-parameter framework for gridded JULES model (Future)











Thank you







Part 1: JULES Parameter Estimation

Geosci. Model Dev., 13, 55–69, 2020 https://doi.org/10.5194/gmd-13-55-2020 © Author(s) 2020. This work is distributed under the Creative Commons Attribution 4.0 License.



Hydrol. Earth Syst. Sci., 25, 1617–1641, 2021 https://doi.org/10.5194/hess-25-1617-2021 © Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Hydrol. Earth Syst. Sci., 25, 2445–2458, 2021 https://doi.org/10.5194/hess-25-2445-2021 © Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Ewan Pinnington¹, Javier Amezcua¹, Elizabeth Cooper², Simon Dadson²⁻³, R Emma Robinson², Ross Morrison², Simon Osborne⁴, and Tristan Quaife¹ ¹National Centre for Earth Observation, Department of Meteorology, University ²UK Centre for Earth Observation, Department of Meteorology, University ³School of Geography and the Environment, University of Oxford, Oxford, UK ³School of Geography and the Environment, University of Oxford, Oxford, UK ⁴Met Office Field Site, Cardington Airfield, Shortstown, Bedford, UK ⁵Department of Remote Sensing, Helmholtz Centre for Environmental Research – Permoserstraße 15, 04318 Leipzig, Germany ¹UK Centre for Ecology and Hydrology, Wallingford, UK

⁶Remote Sensing Centre for Earth System Research, Leipzig University, 04103 Le

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 pa-

rameters 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 grid-

ded 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 pre-

diction 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 uncertain-

ties given to both model and observed estimates to find a

maximum likelihood estimate. After performing the calibra-

tion procedure, we find improved estimates of soil moisture

and heat flux for the Joint UK Land Environment Simulator

(JULES) land surface model (run at a 1 km resolution) when

compared to estimates from a cosmic-ray soil moisture mon-

itoring network (COSMOS-UK) and three flux tower sites.

The spatial resolution of the 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

Received: 17 June 2020 – Discussion started: 15 July 2020 Revised: 24 February 2021 – Accepted: 25 February 2021 – Published: 31 March

Elizabeth Cooper¹, Eleanor Blyth¹, Hollie Cooper¹, Rich Ellis¹, Ewan Pinnington³, and Simon J. Dadson^{1,2} ¹UK Centre for Ecology and Hydrology, Wallingford, UK ²School of Geography and the Environment, South Parks Road, Oxford, OX1 3QY, UK ³National Centre for Earth Observation, Department of Meteorology, University of Reading, Reading, UK

March : Correspondence: Elizabeth Cooper (elicoo@ceh.ac.uk) Received: 8 July 2020 – Discussion started: 6 August 2020

moisture probes 1 Revised: 13 March 2021 – Accepted: 21 March 2021 – Published: 10 May 2021 an average 22 % 1

reduction in unbi

crease in correlat to retrieve new pt els are important in hydrological, ecological, and meteorological applications. In recent years, the availability of widearea soil moisture measurements has increased, but few stud-

1 Introduction ies have combined model-based soil moisture predictions

with in situ observations beyond the point scale. Here we Land surface mo teorological foreces show that we can markedly improve soil moisture estimates by providing sch from the Joint UK Land Environment Simulator (JULES) ter will interact v land surface model using field-scale observations and data diagnostics and v assimilation techniques. Rather than directly updating soil ability in the terr moisture estimates towards observed values, we optimize tem. As the spatia formation has be constants in the underlying pedotransfer functions, which reit is necessary to late soil texture to JULES soil physics parameters. In this information at its way, we generate a single set of newly calibrated pedotransthat are as accura fer functions based on observations from a number of UK paper, our focus i sites with different soil textures. We demonstrate that calirole in agricultur prediction (Haust brating a pedotransfer function in this way improves the soil titioning (Beljaar moisture predictions of a land surface model at 16 UK sites, leading to the potential for better flood, drought, and climate

projections.



(PTFs) to relate readily available or easy-to-measure soil characteristics such as soil texture to the soil hydraulics parameters required by the model (e.g. Van Looy et al., 2017)

Hydrology and

Earth System

Sciences

There are a number of different types of pedotransfer function, as noted in Van Looy et al. (2017) and Hodnett and Tomasella (2002), with different inputs and outputs depending partly on the representation of soil physics processes of the chosen land surface model. In "class" approaches, soil types are clustered into groups, and hydraulic model parameters are then obtained from a lookup table (Wösten et al., 1999); this results in discrete soil hydraulics parameter sets. Alternatively, continuous pedotransfer functions take soil characteristic information from each sample of interest and apply the function to produce continuous soil hydraulics parameter sets (e.g. Cosby et al., 1984; Hodnett and Tomasella, 2002; Schaap et al., 2001).

To date, pedotransfer functions have been derived by fitting to results from field or laboratory experiments on pointor small-scale soil samples (centimetre to metre), despite the fact that land surface models are generally applied at larger (field to kilometre) scales. The recent development of novel in situ techniques for measurine soil moisture over



National Centre for Earth Observation

The Land Variational Ensemble Data Assimilation Framework: LAVENDAR v1.0.0

Ewan Pinnington¹, Tristan Quaife^{1,2}, Amos Lawless^{1,2}, Karina Williams³, Tim Arkebauer⁴, and Dave Scoby⁴

¹National Centre for Earth Observation, Department of Meteorology, University of Reading, Reading, UK ²School of Mathematical, Physical and Computational Sciences, University Of Reading, Reading, UK ³Met Office Hadley Centre, Exeter, UK

⁴Department of Agronomy and Horticulture, University of Nebraska–Lincoln, Lincoln, Nebraska, USA

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

Received: 28 February 2019 – Discussion started: 11 April 2019 Revised: 9 July 2019 – Accepted: 19 November 2019 – Published: 7 January 2020

Abstract. The Land Variational Ensemble Data Assimilation Framework (LAVENDAR) implements the method of fourdimensional ensemble variational (4D-En-Var) data assimilation (DA) for land surface models. Four-dimensional ensemble variational data assimilation negates the often costly calculation of a model adjoint required by traditional variational techniques (such as 4D-Var) for optimizing parameters or state variables over a time window of observations. In this paper we present the first application of LAVENDAR, implementing the framework with the Joint UK Land Environment Simulator (JULES) land surface model. We show that the system can recover seven parameters controlling crop behaviour in a set of twin experiments. We run the same experiments at the Mead continuous maize FLUXNET site in Nebraska, USA, to show the technique working with real data. We find that the system accurately captures observations of leaf area index, canopy height and gross primary productivity after assimilation and improves posterior estimates of the amount of harvestable material from the maize crop by 74 %. LAVENDAR requires no modification to the model that it is being used with and is hence able to keep up to date with model releases more easily than other DA methods.

pacts on human life. Most land surface models will converge to a steady state; their state vector tends toward an equilibrium defined by forcing variables (i.e., the meteorology experienced by the model) and the model parameters. This is quite unlike fluid dynamics models used for the atmosphere and oceans, which exhibit chaotic behaviour; a small change in their initial state can lead to large deviations in the state vector evolution with time. Consequently, for some land surface applications parameter estimation can have greater utility than state estimation (Luo et al., 2015). This article deals primarily with the problem of parameter estimation in land surface models, although the technique we introduce could easily be used for state estimation problems too.

Data assimilation (DA) combines models and data such that resulting estimates are an optimal combination of both, taking into account all available information about respective uncertainties. DA techniques are typically derived from a Bayesian standpoint and have been largely developed to service the needs of atmospheric and ocean modelling, especially where there is a need to provide near-real-time forecasts. Typically the focus of such activities is on estimating the optimal model state as the fundamental laws underlying fluid dynamics are well understood and many of the model

University of Reading



Effect of DA on Non-typical Soil Sites



Mineral soil with very high organic content

**	University of			
Y	Reading			

Grassland

Sourhope (SOURH)

Place	KGE-Prior	KGE- Post.	% increase	Rank
CARDT	0.35	0.71	102.85	10
BICKL	0.36	0.92	155.55	6
CHRICH	0.35	0.86	145.71	8
WADDN	0.34	0.77	126.47	9
HOLLN	0.29	0.75	158.62	5
EASTB	0.38	0.76	100	11
ROTHD	0.55	0.69	25.45	16
CHIMN	0.51	0.78	52.94	14
SHEEP	0.46	0.81	76.08	13
PORTN	0.48	0.64	33.33	15
HARTW	0.17	0.63	270.58	2
GISBN	0.17	0.77	352.94	1
СНОВН	0.15	0.48	220	4
LULLN	0.44	0.84	90.90	12
MOORH	0.14	0.49	250	3
SOURH	0.26	0.66	153.84	7

Correlation and Error Co-variance plots

