Data Assimilation with JULES

Ewan Pinnington

Tristan Quaife

Karina Williams

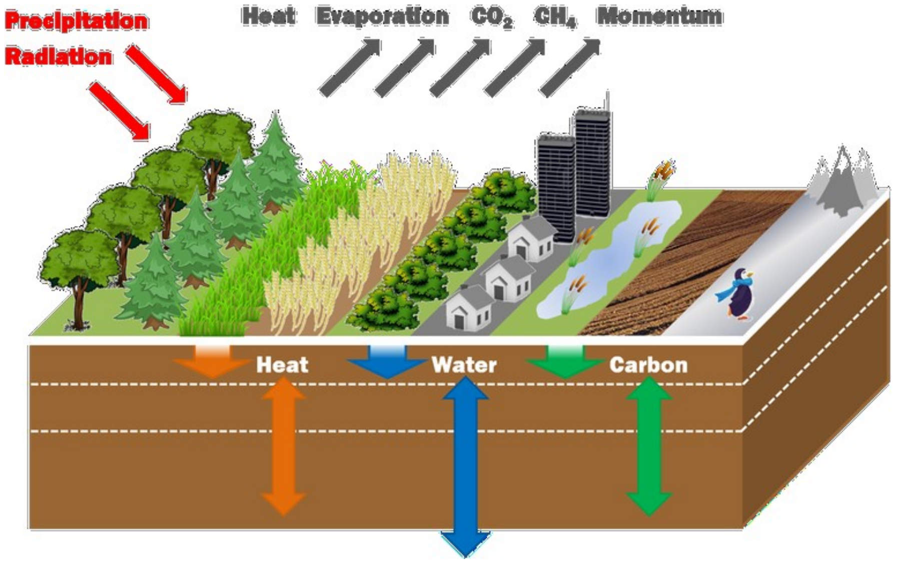
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

NERC SCIENCE OF THE ENVIRONMENT

Data Assimilation (DA) with JULES

- Data Assimilation (DA) techniques combine observations with process models in order to update parameter and state variables and improve model predictions.
- In order to conduct DA need to be able to easily run the model whilst varying the values of parameters and state variables.
- Lots of different techniques appropriate for different problems.



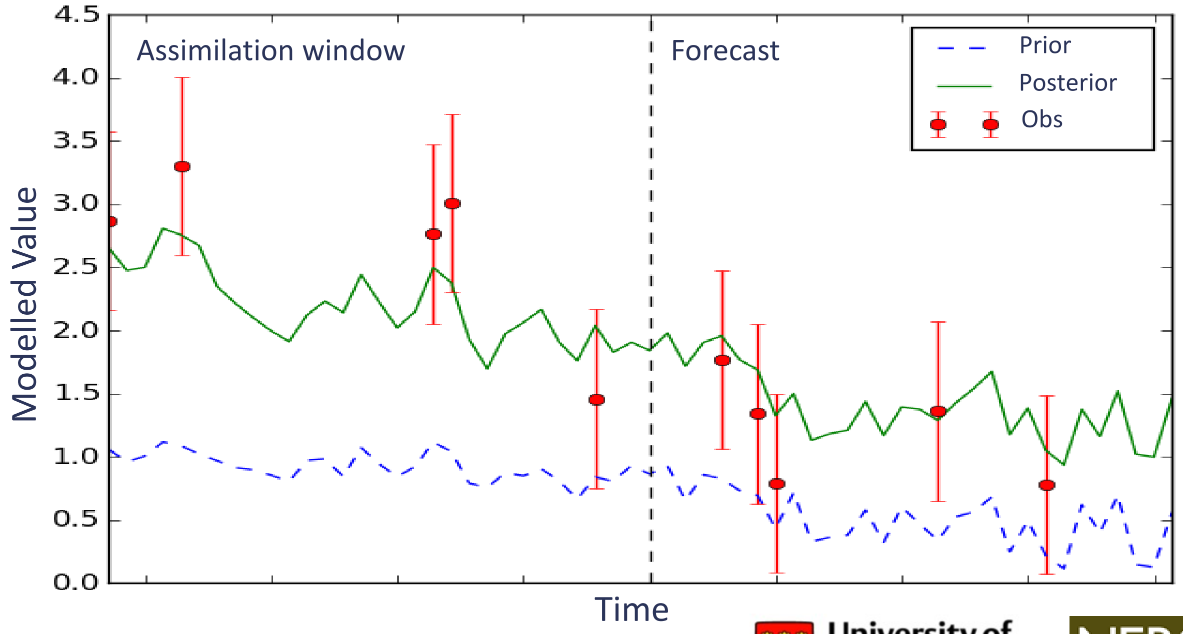
JULES model schematic

Four-Dimensional Variational (4DVar) DA

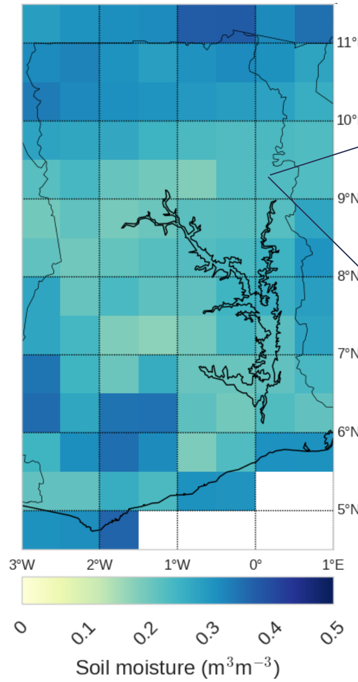
- Combine all sources of information to find best estimate to the state of a system.
- Do this by minimising a cost function.
- Typically requires the derivative of the model. This is an issue for JULES!

4DVar cost function:

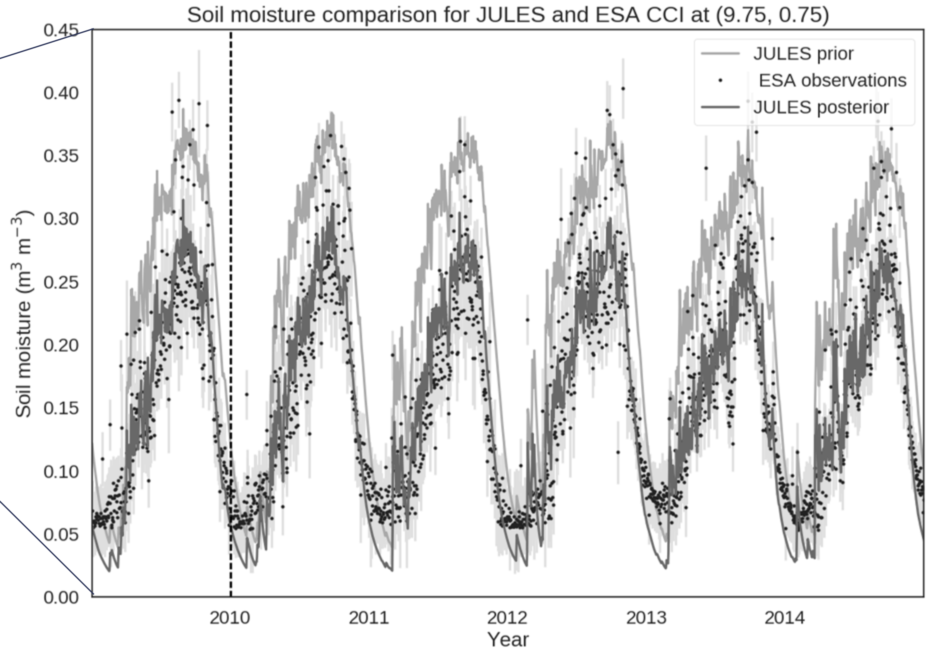
$$J(\mathbf{x}_0) = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b)}_{\textbf{Prior}} + \underbrace{\frac{1}{2}(\hat{\mathbf{h}}(\mathbf{x}_0) - \hat{\mathbf{y}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{h}}(\mathbf{x}_0) - \hat{\mathbf{y}})}_{\textbf{Observations}}$$



Improving soil moisture estimates for Ghana



JULES modelled soil moisture over Ghana



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, Hydrol. Earth Syst. Sci., 22, 2575-2588, https://doi.org/10.5194/hess-22-2575-2018, 2018.

- Assimilated ESA CCI satellite observations of soil moisture to optimise soil parameters of JULES.
- Found a 20% reduction in unbiased-RMSE for 5-year hindcast.
- DA method was slow!
- In order to consider larger scales different approach to DA is required.

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

Ewan Pinnington | e.pinnington@reading.ac.uk

University of Reading

NERC SCIENCE OF THE ENVIRONMENT

Four-Dimensional Ensemble Variational (4DEnVar) DA

- Approximates 4D-Var Using an ensemble of model trajectories.
- Does not require the derivative of the model.
- Much faster than technique used in soil moisture work.
- Requires no code modification.
- Easily parallelisable.

$$\mathbf{X}'_b = \frac{1}{\sqrt{N-1}}(\mathbf{x}^{b,1} - \overline{\mathbf{x}}^b, \mathbf{x}^{b,2} - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}^{b,N} - \overline{\mathbf{x}}^b), \quad (1)$$

$$\mathbf{B} \approx \mathbf{X}'_b \mathbf{X}'^T_b \quad (2)$$

$$\mathbf{x}_0 = \mathbf{x}_b + \mathbf{X}'_b \mathbf{w}, \quad (3)$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}})^T \hat{\mathbf{R}}^{-1} (\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}}) \quad (4)$$

Ewan Pinnington | e.pinnington@reading.ac.uk

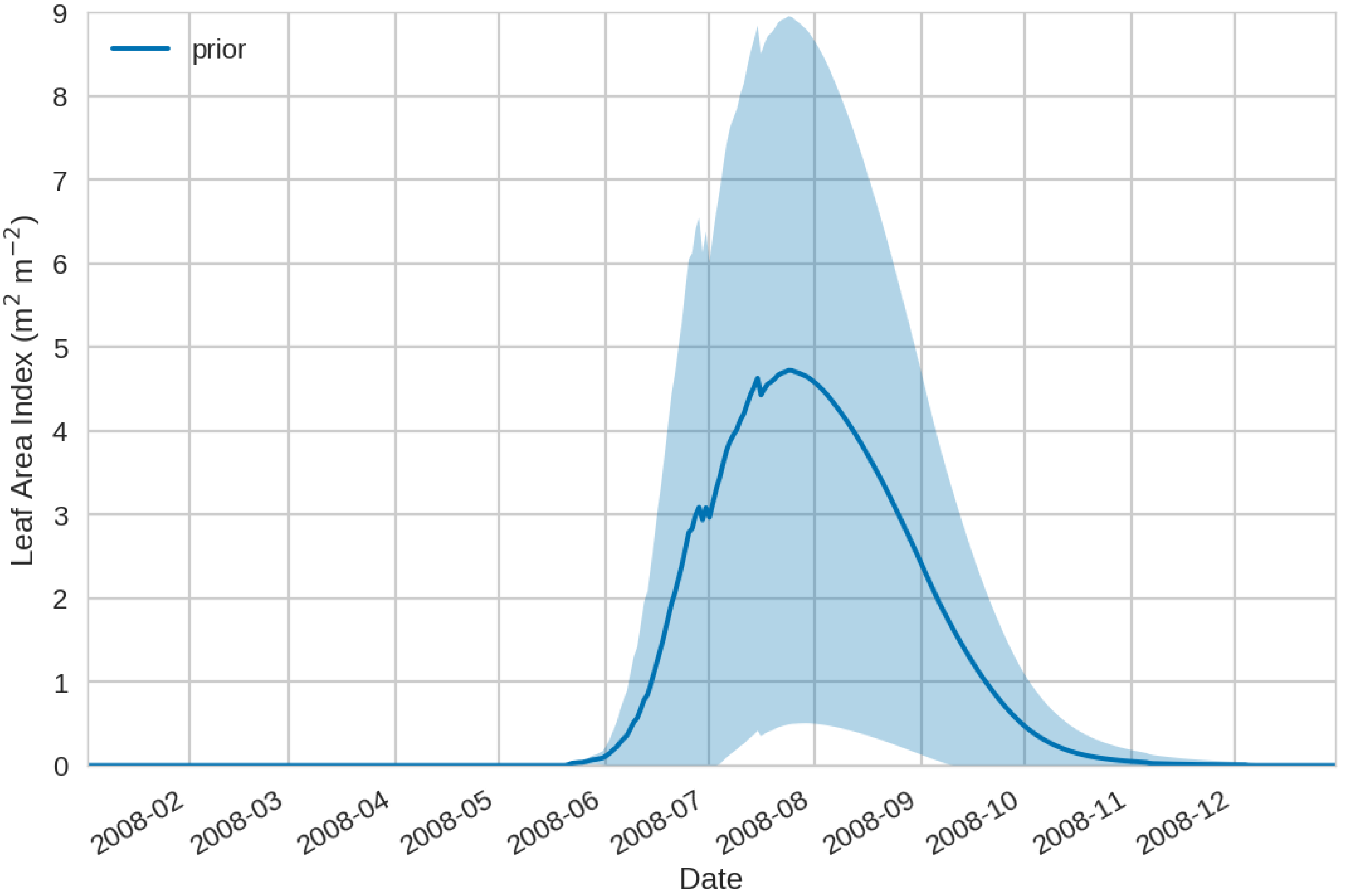
LaVEnDAR

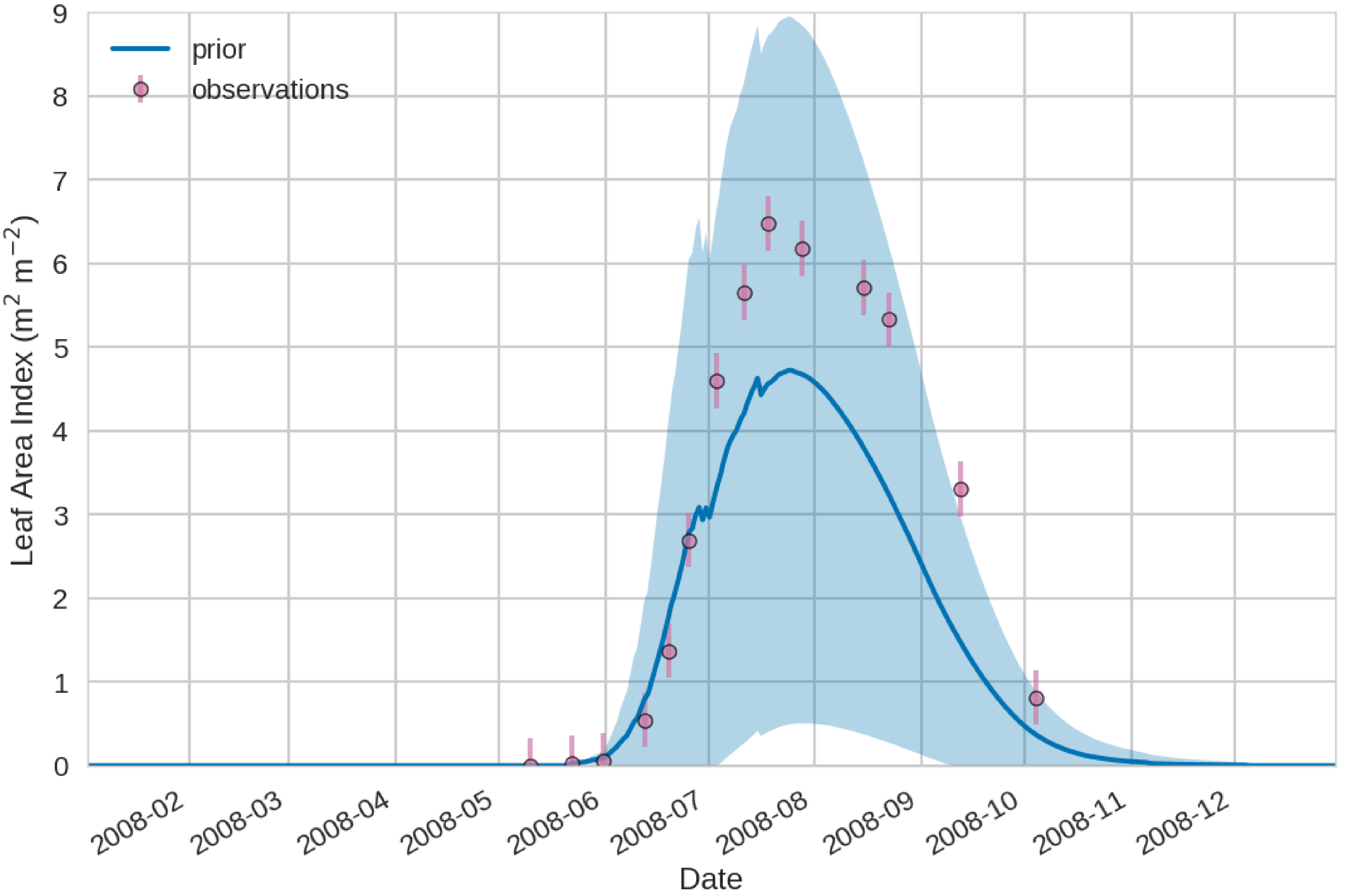
- The Land Ensemble Variational Data Assimilation fRamework (LaVEnDAR) implements 4DEnVar for land surface models.
  - https://github.com/pyearthsci/lavendar
- Written wrappers for JULES controlling functionality. Allows us to easily run ensemble of models with different parameters in parallel and perform DA.
- Have tested system for JULES-Crop at Mead maize flux site in Nebraska, USA.
  - Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., and Scoby, D.: The Land Variational Ensemble Data Assimilation fRamework: LaVEnDAR, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-60, in review, 2019.

Example for JULES-Crop

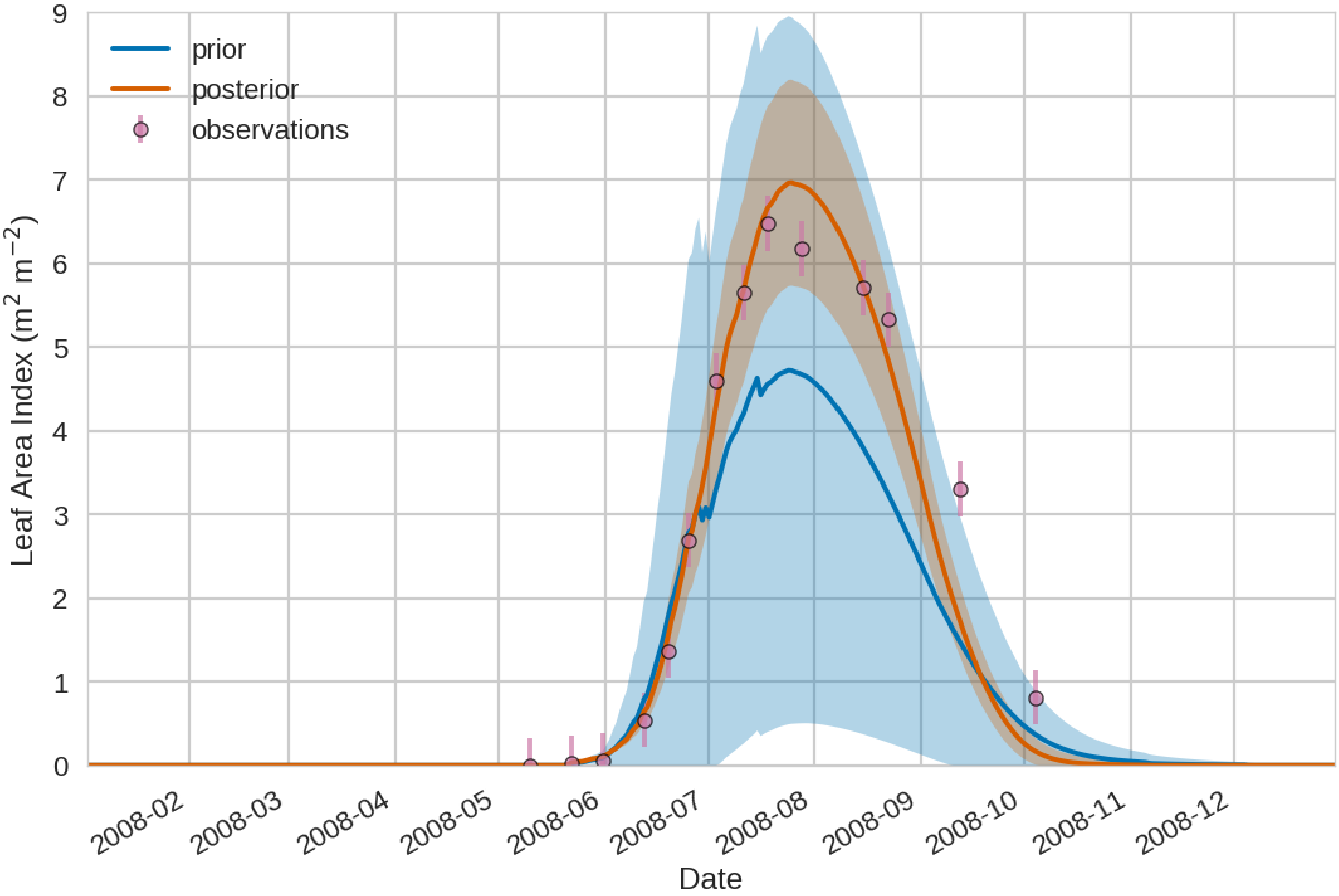
- Optimising 7 model parameters controlling crop behaviour, motivated by Williams et al. (2017).
  - Williams, K.et al.: Evaluation of JULES-crop performance against site observations of irrigated maize from Mead, Nebraska, Geosci. Model Dev., 10, 1291-1320, https://doi.org/10.5194/gmd-10-1291-2017, 2017.
- Run JULES 50 times in parallel with different values for the 7 parameters to produce a prior ensemble.
- Assimilating Leaf Area Index (LAI), flux tower Gross Primary Productivity (GPP) estimates and canopy height.

Parameter	Description
α	quantum efficiency of photosynthesis (mol CO_2 mol^{-1} PAR)
n_{eff}	nitrogen use efficiency (mol CO_2 m^{-2} s^{-1} kg C (kg N)$^{-1}$)
f_d	scale factor for dark respiration (-)
μ	allometric coefficient for calculation of senescence (-)
ν	allometric coefficient for calculation of senescence (-)
γ	coefficient for determining specific leaf area (-)
δ	coefficient for determining specific leaf area (-)


prior
Leaf Area Index (m^2 m^{-2})
0
1
2
3
4
5
6
7
8
9
2008-02
2008-03
2008-04
2008-05
2008-06
2008-07
2008-08
2008-09
2008-10
2008-11
2008-12
Date

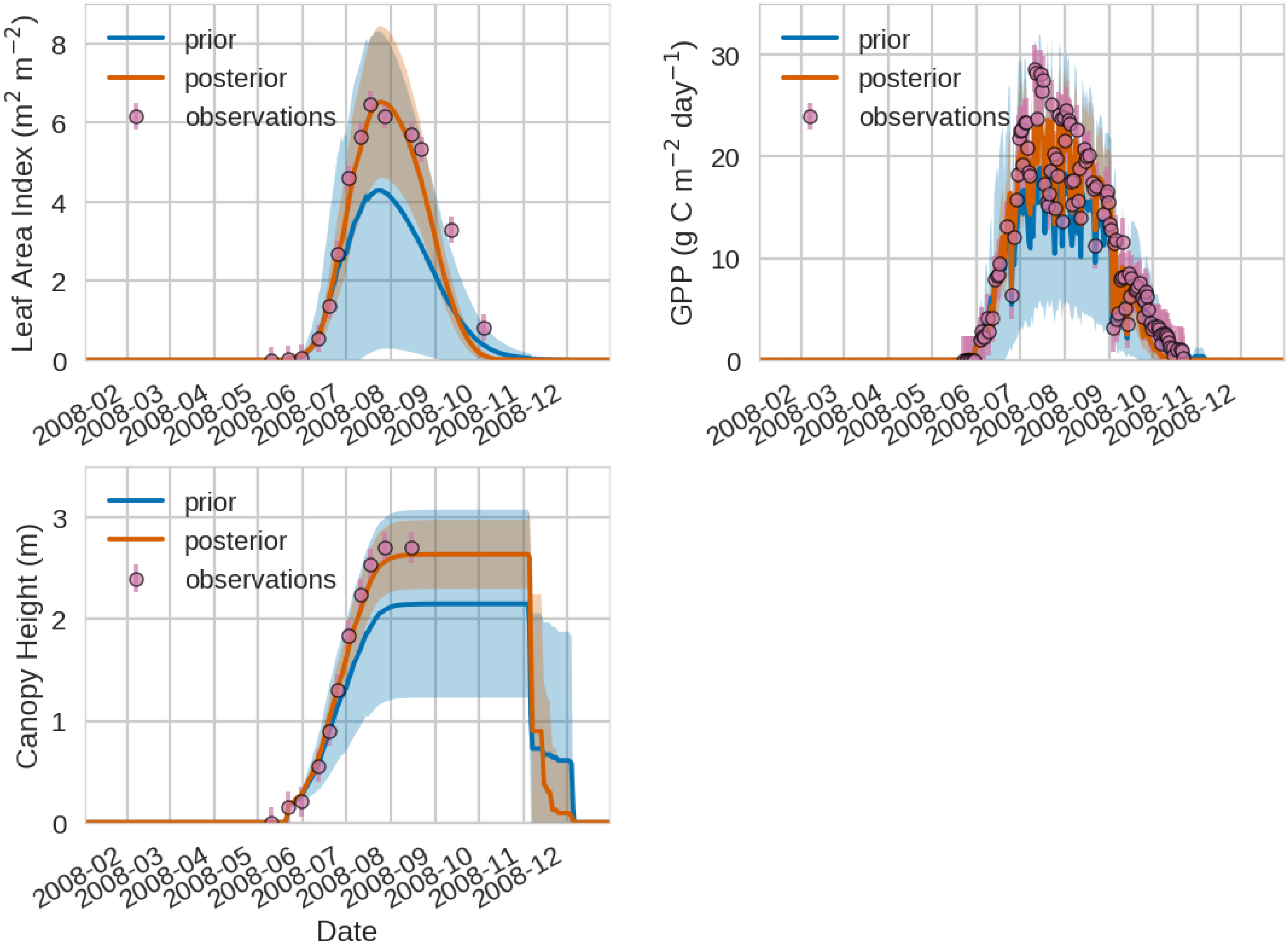


prior
observations
Leaf Area Index (m^2 m^{-2})
Date
0
1
2
3
4
5
6
7
8
9
2008-02
2008-03
2008-04
2008-05
2008-06
2008-07
2008-08
2008-09
2008-10
2008-11
2008-12

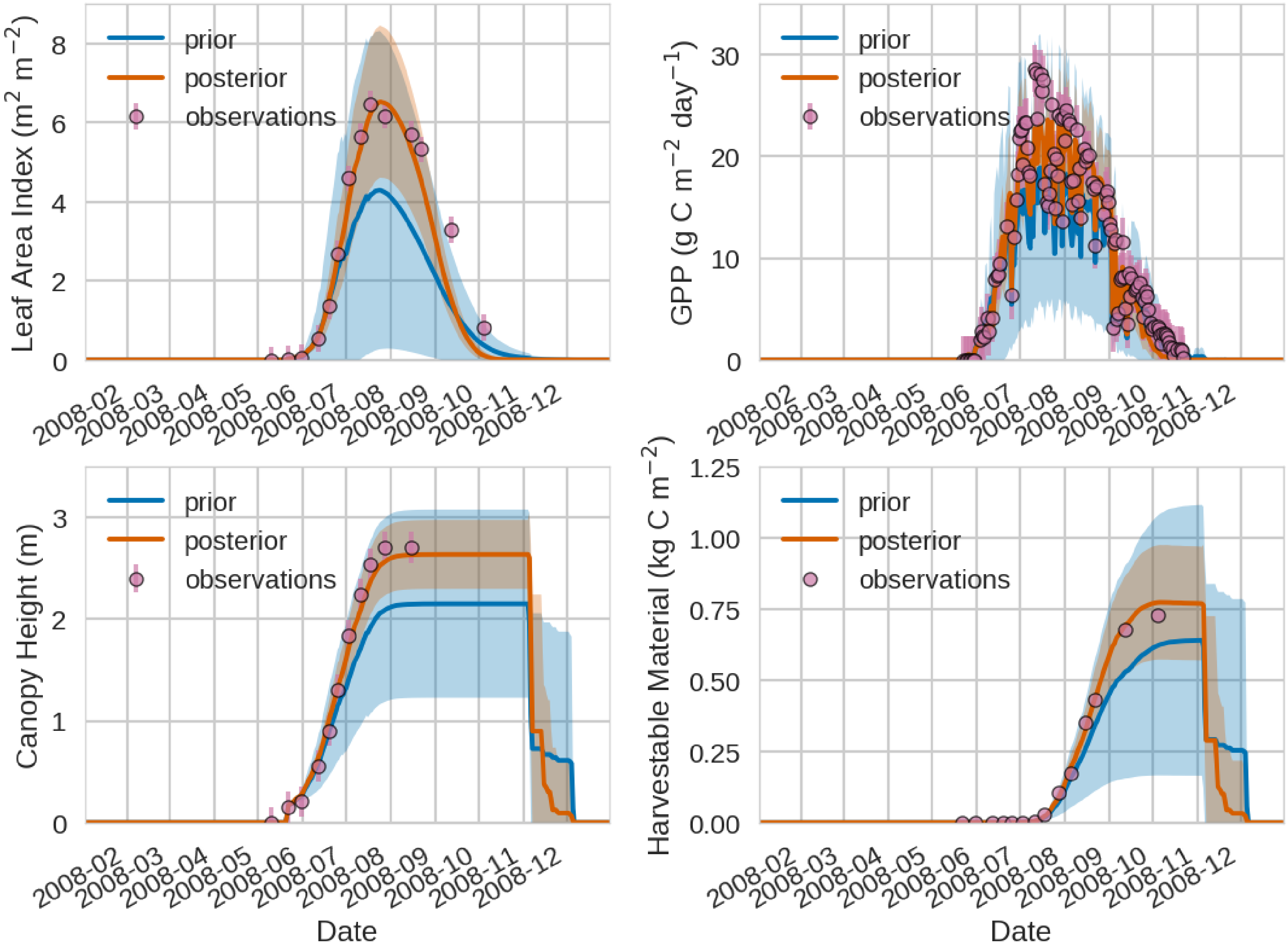


prior
posterior
observations
Leaf Area Index (m^2 m^{-2})
Date
0
1
2
3
4
5
6
7
8
9
2008-02
2008-03
2008-04
2008-05
2008-06
2008-07
2008-08
2008-09
2008-10
2008-11
2008-12

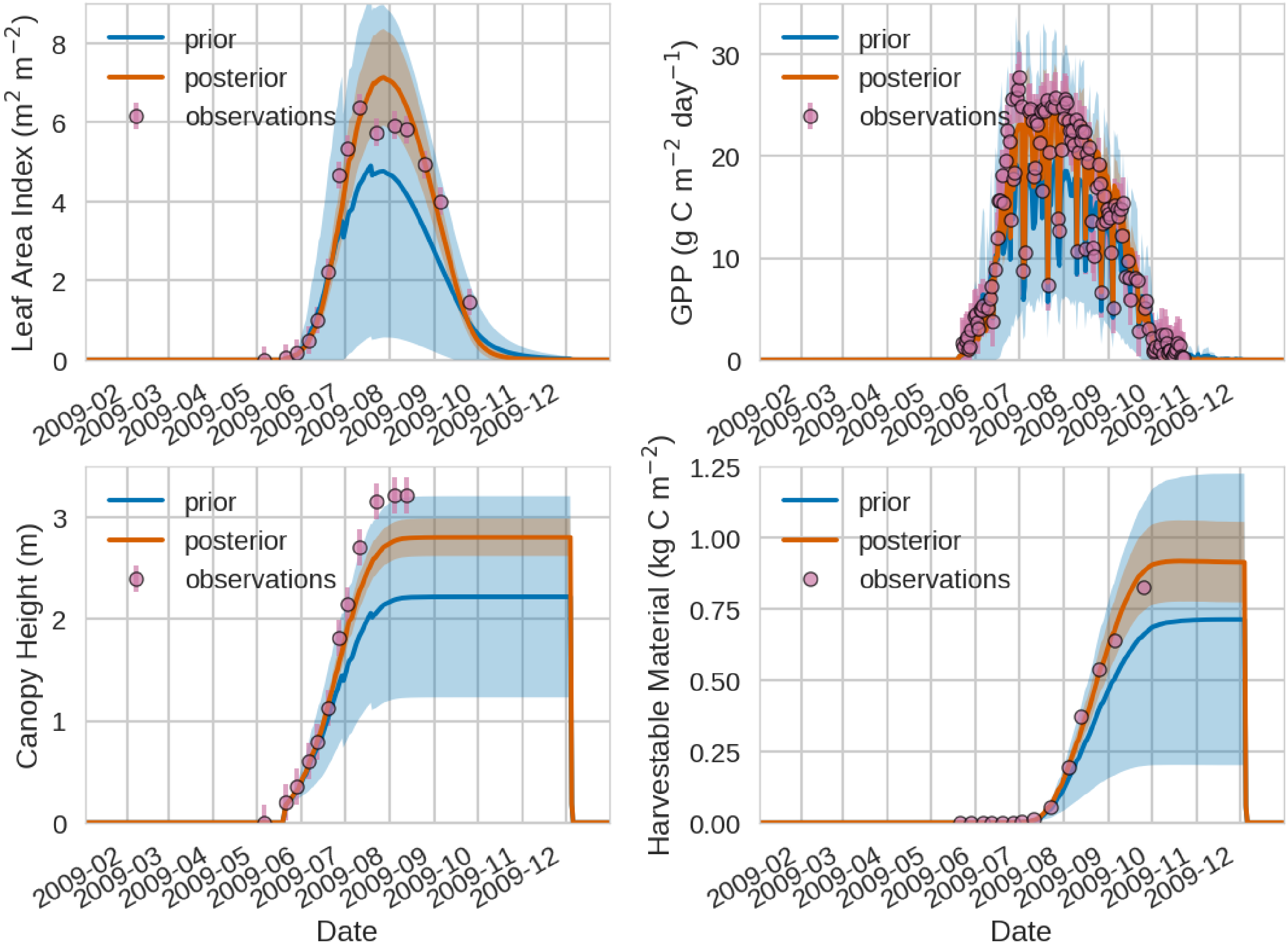
Assimilation results



Assimilation results

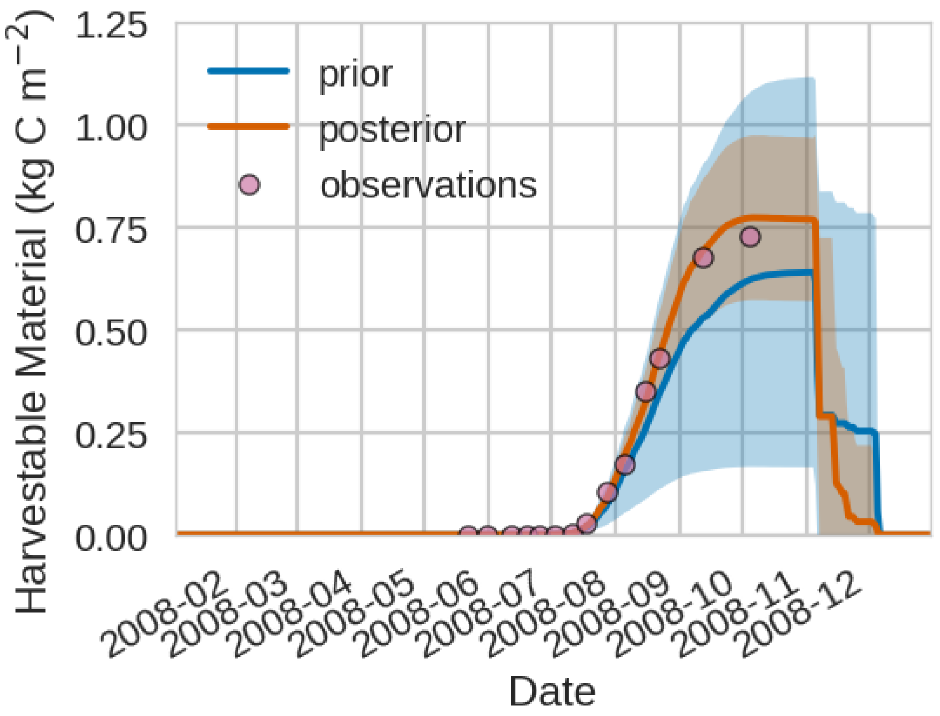


Hindcast for 2009



Assimilation Summary

- After assimilation the model prediction RMSE for 3 target variables (LAI, GPP and canopy height) is reduced by an average of 59%.
- As independent validation reduce the RMSE in prediction of harvestable material by 74%. Also improve model skill in hindcast for subsequent year.
- 4DEnVar performs well and negates the need for the computation of the model derivative which can be costly.

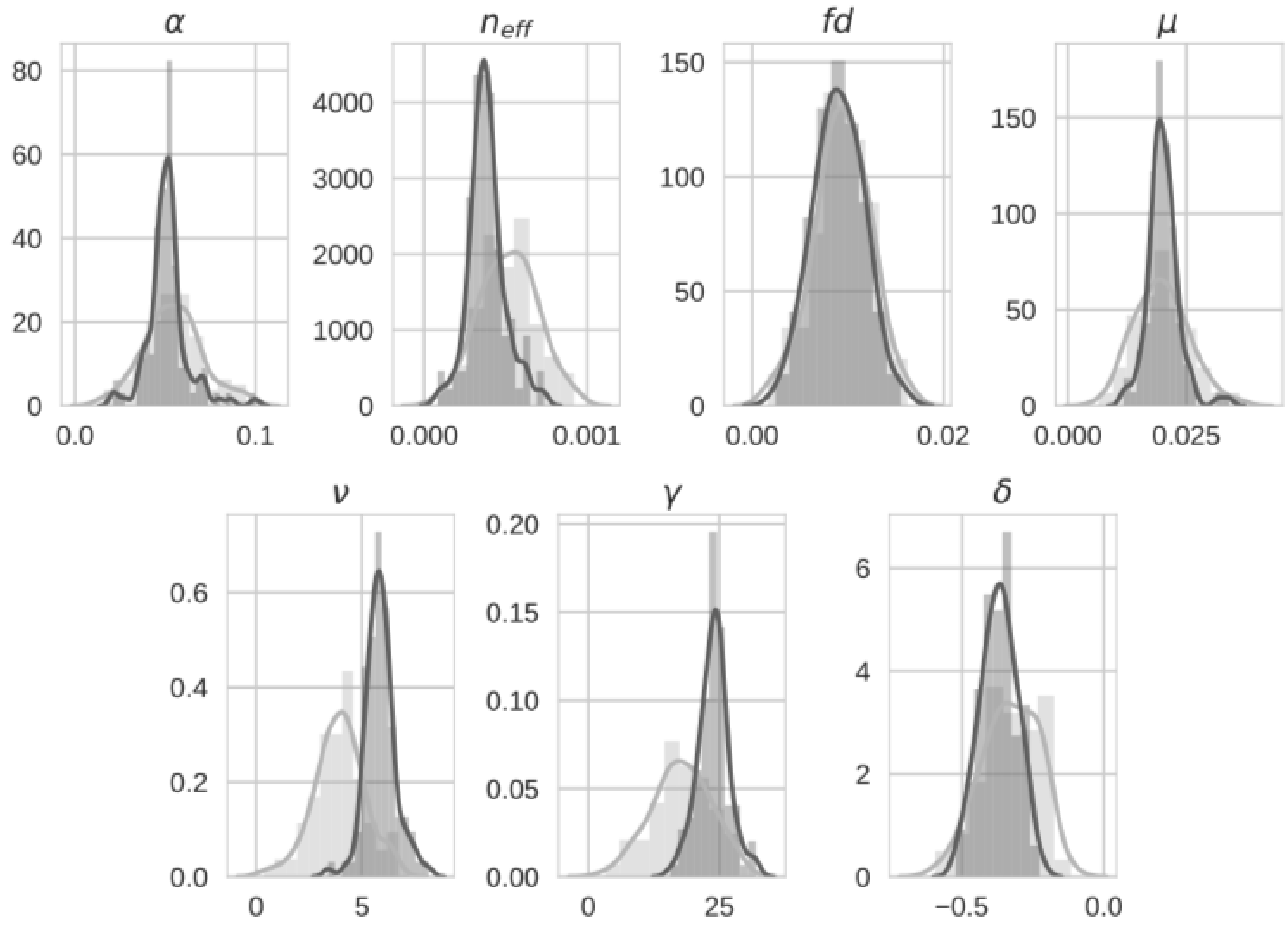


Conclusions

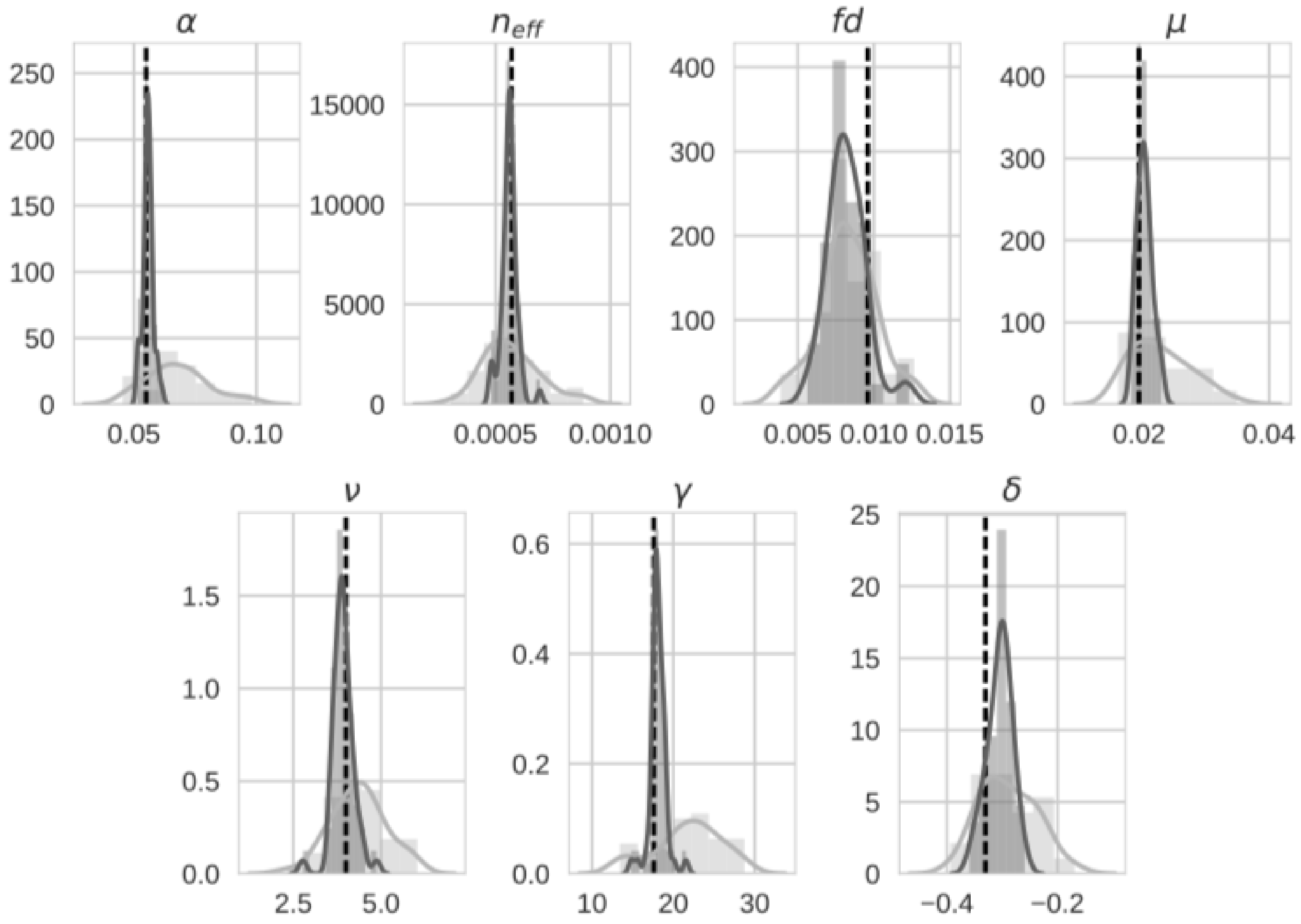
- 4DEnVar DA appears to be a good option for parameter and state estimation for JULES.
- We have implemented the LaVEnDAR technique with Python wrappers, but are looking into building this into Cylc workflow.
- Next steps:
  - Working on Hydro-JULES project assimilating SMAP data to improve soil moisture prediction for the UK.
  - Further developing the technique over Africa as part of NCEO Official Development Assistance project.

Ewan Pinnington | e.pinnington@reading.ac.uk

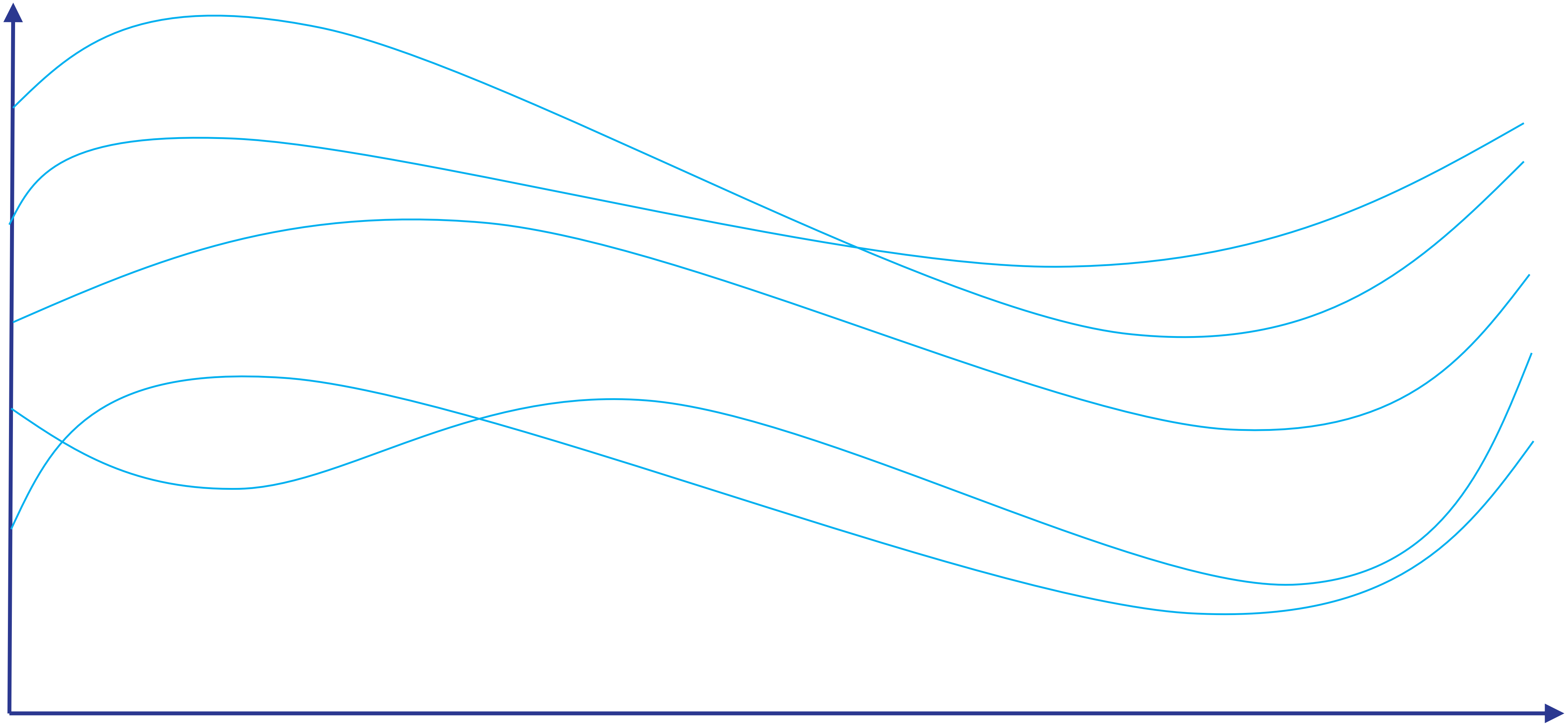
Extras – parameter distributions



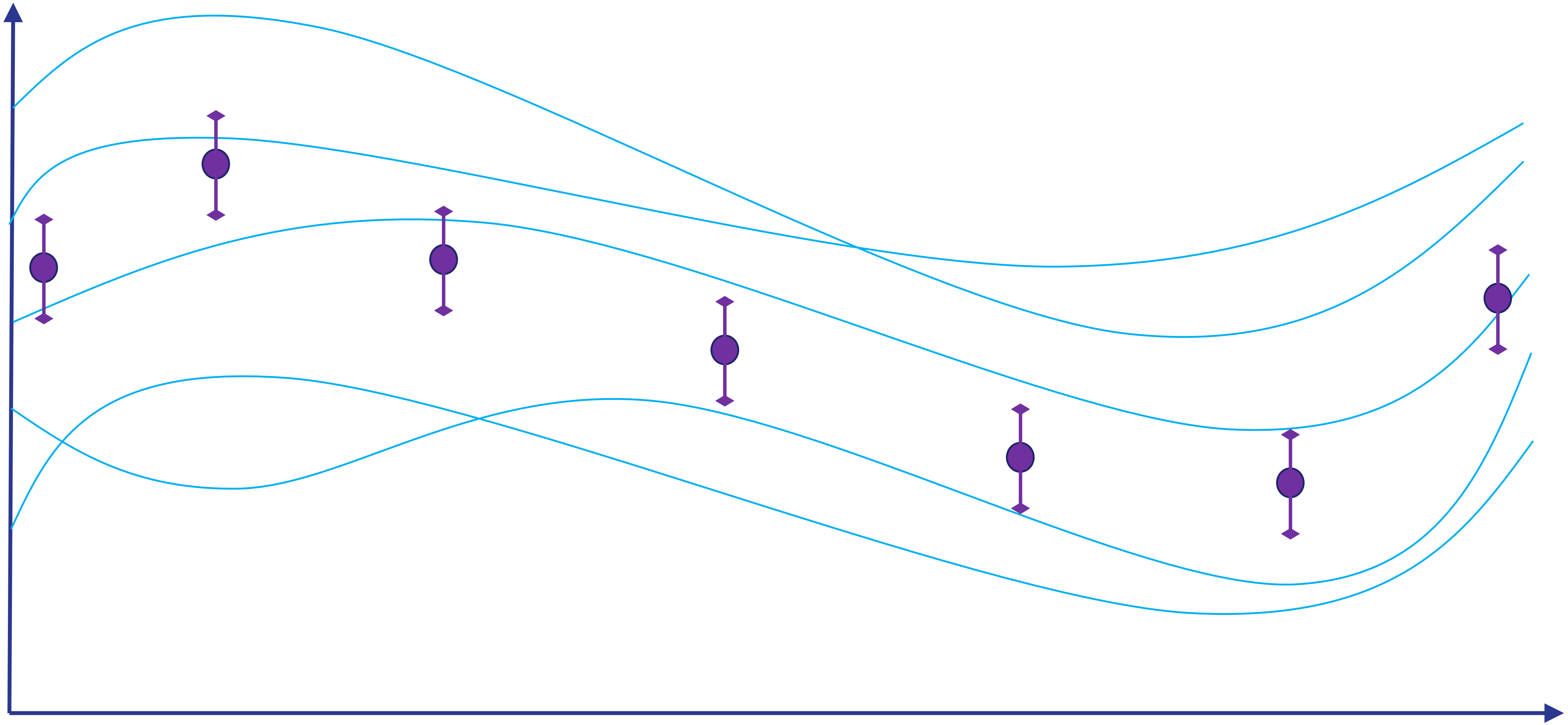
Extras – twin experiment



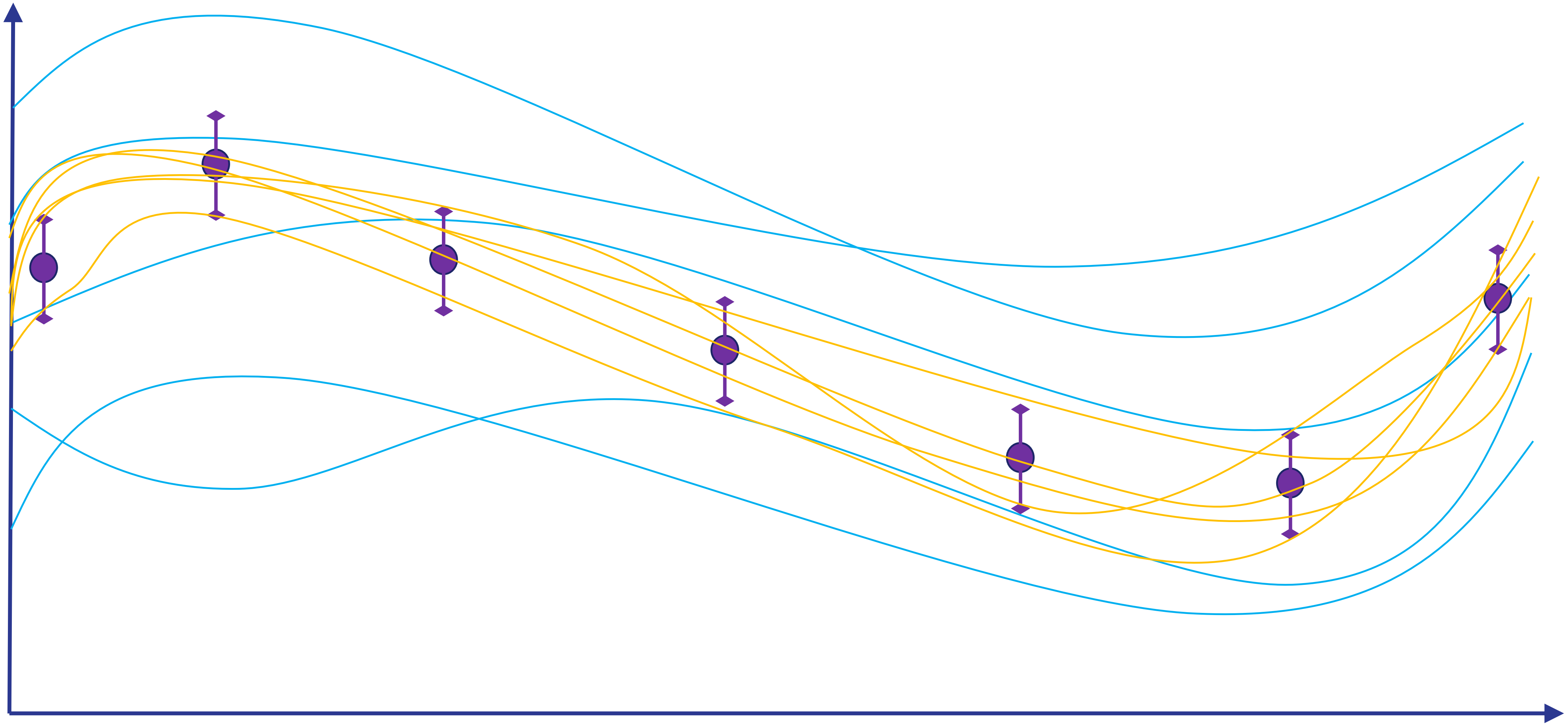
LaVEnDAR



LaVEnDAR



LaVEnDAR



Four-Dimensional Variational (4DVar) DA

4DVar cost function: $$J(\mathbf{x}_0) = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b)}_{\textbf{Prior}} + \underbrace{\frac{1}{2}(\hat{\mathbf{h}}(\mathbf{x}_0) - \hat{\mathbf{y}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{h}}(\mathbf{x}_0) - \hat{\mathbf{y}})}_{\textbf{Observations}}$$

Jacobian of cost function: $$\nabla J(\mathbf{x}_0) = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) - \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{h}}(\mathbf{x}_0) - \hat{\mathbf{y}})$$

where

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \ \hat{\mathbf{h}}(\mathbf{x}_0) = \begin{pmatrix} \mathbf{h}_0(\mathbf{x}_0) \\ \mathbf{h}_1(\mathbf{m}_{0\to 1}(\mathbf{x}_0)) \\ \vdots \\ \mathbf{h}_N(\mathbf{m}_{0\to N}(\mathbf{x}_0)) \end{pmatrix}, \ \hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_{0,0} & \mathbf{R}_{0,1} & \dots & \mathbf{R}_{0,N} \\ \mathbf{R}_{1,0} & \mathbf{R}_{1,1} & \dots & \mathbf{R}_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{N,0} & \mathbf{R}_{N,1} & \dots & \mathbf{R}_{N,N} \end{pmatrix} \text{ and } \ \hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_0 \\ \vdots \\ \mathbf{H}_N\mathbf{M}_{N,0} \end{pmatrix}$$

Improving soil moisture estimates for Ghana

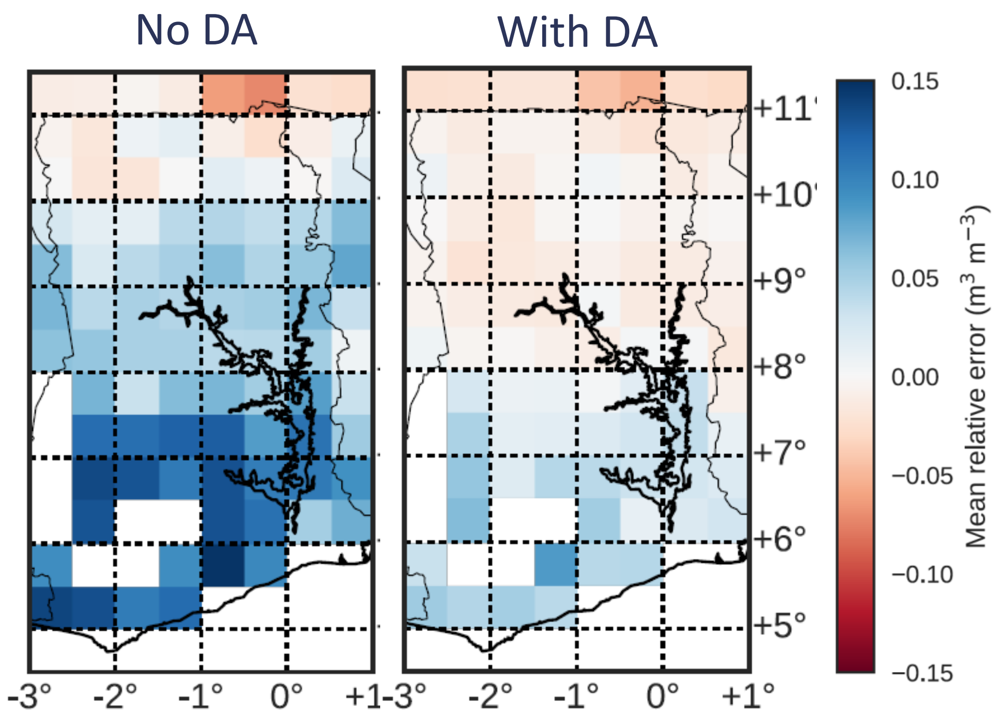
- Significant impact on model skill after DA. Assimilation not able to remove all biases associated with model driving data, especially onset of rains. Improved precipitation data required.

  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, Hydrol. Earth Syst. Sci., 22, 2575-2588, https://doi.org/10.5194/hess-22-2575-2018, 2018.

- Working with TAMSAT group at Reading. Ultimately build work into system predicting meteorological risk to agriculture.

  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, Geosci. Model Dev., 11, 2353-2371, https://doi.org/10.5194/gmd-11-2353-2018, 2018.

- In order to consider larger scales different approach to DA is required.



Mean relative error in volumetric soil moisture predicted by JULES in a hindcast experiment for 5 years for Ghana. Without Data Assimilation (left panel) and assimilating ESA CCI data (right panel)

Ewan Pinnington | e.pinnington@reading.ac.uk

Four-Dimensional Ensemble Variational DA

- Approximates 4D-Var Using an ensemble of model trajectories.
- Does not require the derivative of the model.
- Much faster than technique used in soil moisture work.
- Requires no code modification.
- Easily parallelisable.

$$\mathbf{X}'_b = \frac{1}{\sqrt{N-1}}(\mathbf{x}^{b,1} - \overline{\mathbf{x}}^b, \mathbf{x}^{b,2} - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}^{b,N} - \overline{\mathbf{x}}^b), \quad (1)$$

$$\mathbf{B} \approx \mathbf{X}'_b \mathbf{X}'^T_b \quad (2)$$

$$\mathbf{x}_0 = \mathbf{x}_b + \mathbf{X}'_b \mathbf{w}, \quad (3)$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}}) \quad (4)$$

$$\nabla J(\mathbf{w}) = \mathbf{w} - (\hat{\mathbf{H}}\mathbf{X}'_b)^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{X}'_b\mathbf{w} + \hat{\mathbf{h}}(\mathbf{x}^b) - \hat{\mathbf{y}}), \quad (5)$$

$$\hat{\mathbf{H}}\mathbf{X}'_b \approx \frac{1}{\sqrt{N-1}}(\hat{\mathbf{h}}(\mathbf{x}^{b,1}) - \hat{\mathbf{h}}(\overline{\mathbf{x}}^b), \hat{\mathbf{h}}(\mathbf{x}^{b,2}) - \hat{\mathbf{h}}(\overline{\mathbf{x}}^b), \ldots, \hat{\mathbf{h}}(\mathbf{x}^{b,N}) - \hat{\mathbf{h}}(\overline{\mathbf{x}}^b)) \quad (6)$$