

Introduction

Soil moisture is a key driver of hydrological processes such as flooding during wet periods, and vegetation growth in dry periods, but is highly variable in time and space

Soil moisture can vary due to:

- landcover
- landscape position
- evaporation
- soil type
- rainfall input

Disentangling the different drivers and spatial relationships in soil moisture is important to delivering forecasts of water availability. There is a need for data science approaches which allow computationally efficient modelling (as opposed to numerical models) while remaining interpretable with quantified uncertainty (as opposed to, for example, neural networks).

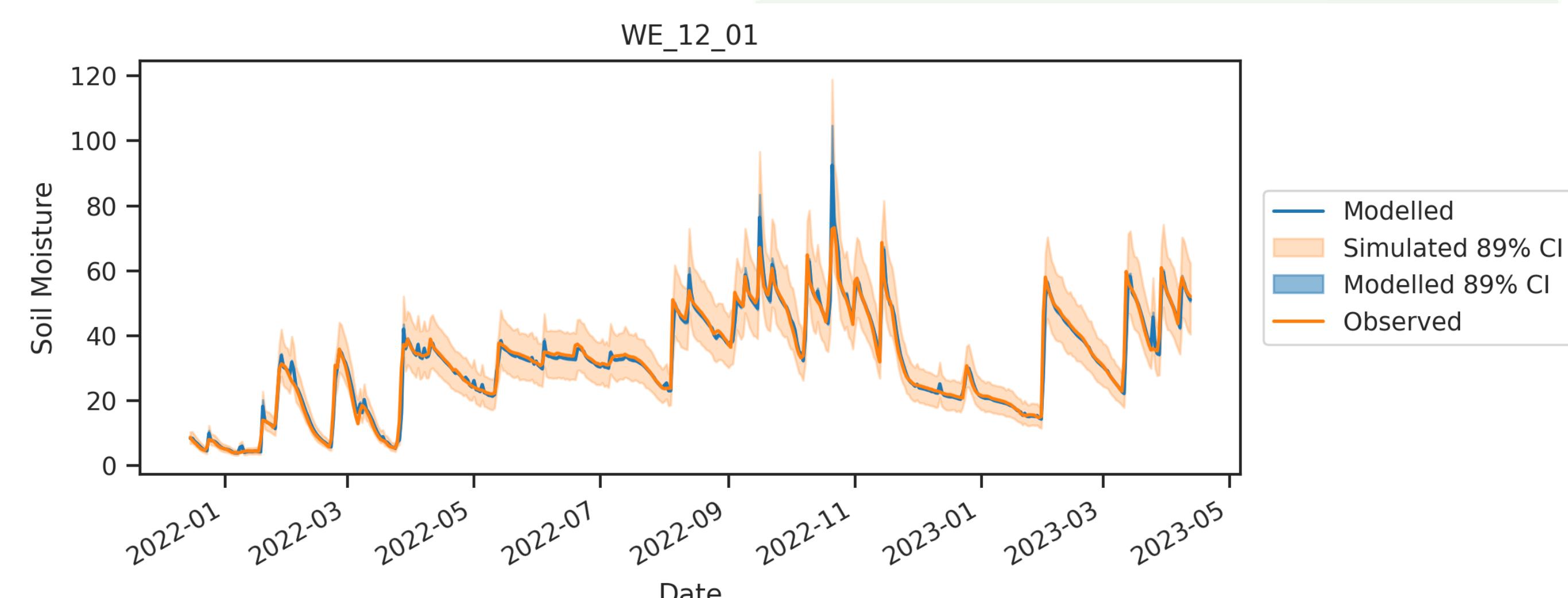


Figure 1: Model predictions of soil moisture at probe WE-12-01.

Methodology

Objective: Demonstrate a Bayesian model for one-day ahead forecast framework that quantifies uncertainties and disentangles the relative importance of different drivers across treatment/control groups

The data are from a dense observation network at Llara farm in Narrabri (NSW), part of a landscape rehydration project. Soil moisture probes are located across two 40ha sites with some control and treatment areas in each. The treatment aims to reduce overland flow velocities and increase infiltration. The study uses 16 months of 10-min data aggregated to daily across 32 gauges at a common depth below the surface (200 mm).

A Hierarchical Linear Model (HLM) is fit using a Bayesian approach (built using the NumPyro PPL). Partial pooling provides a description of gauge (g) level parameters, grouped by site (s) and treatment (tr). Soil moisture at time t (SM_t) is modelled as a linear regression of:

- R : daily rainfall from a nearby SILO gauge (parameter β_R) with a power transformation (ξ) for a non-linear effect
- an trend term which accounts for drying (parameter β_{trend})
- an autoregressive term using the soil moisture from the previous timestep (SM_{t-1} with parameter β_{AR1})

$$SM_t = f(R, SM_{t-1}) + \epsilon \quad \epsilon \sim \mathcal{N}(0, \sigma)$$

$$SM_{t[g]} = \beta_{R[g]} \cdot R^\xi + \beta_{trend[g]} + \beta_{AR1[g]} \cdot SM_{t-1[g]}$$

then for the base model with no interaction on the rainfall effect

$$\beta_{[g]} = \mathcal{N}(\alpha_{[s,tr]}, \tau_{[s,tr]}) \quad \text{for } R, \text{ trend, AR1}$$

$$\alpha_{[s,tr]} \sim N(0, 2) \quad \text{for } R, \text{ trend, AR1}$$

or in a model with an interaction of rainfall and the antecedent soil moisture conditions (with matching specification of group level priors to above)

$$\beta_R = \beta_{r0} + \beta_{r1} SM_{t-1}$$

Results

Table 1: Model performance across the 32 gauges. Performance is measured using the Brier Skill Score (BSS) with a baseline of persistence from the previous day's observation (closer to 1 is better).

Rainfall effect	BSS (mean)	BSS > 0.2	BSS > 0.5	BSS > 0.7
β_R	0.31 (± 0.31)	24	10	0
$\beta_{r0} + SM_{t-1}\beta_{r1}$	0.58 (± 0.11)	32	25	4

The model adequately reproduces the observations with a mean Brier Skill Score (against a baseline of persistence) of 0.58 across the gauges (Table 1). The inclusion improvement of the model when including an interaction of SM_{t-1} in the rainfall effect (β_R) shows the importance of antecedent soil moisture conditions in modelling change.

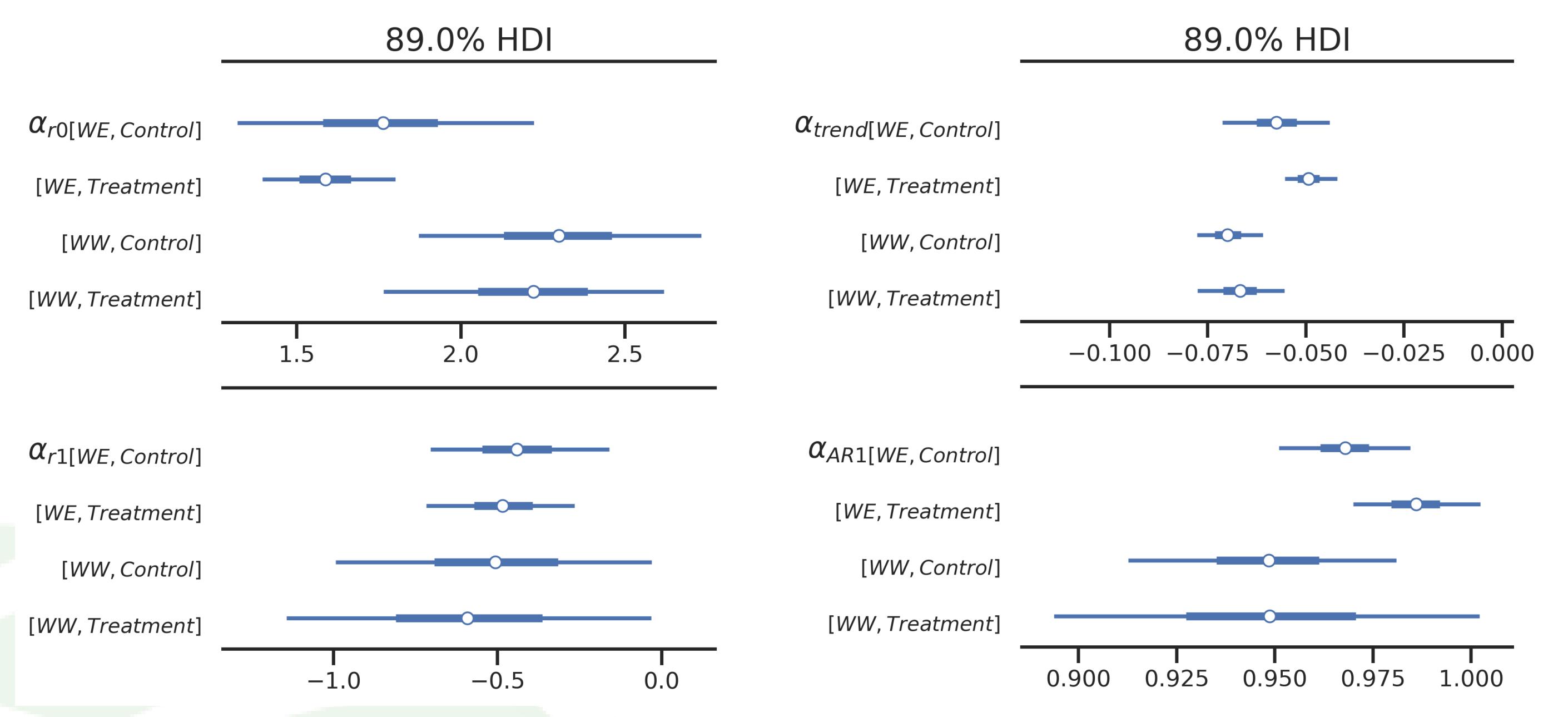


Figure 2a: Group level parameter means for combinations of site and control/treatment. Top: the rainfall effect intercept (α_{r0}). Bottom: the interaction of antecedent soil moisture with rainfall (α_{r1}).

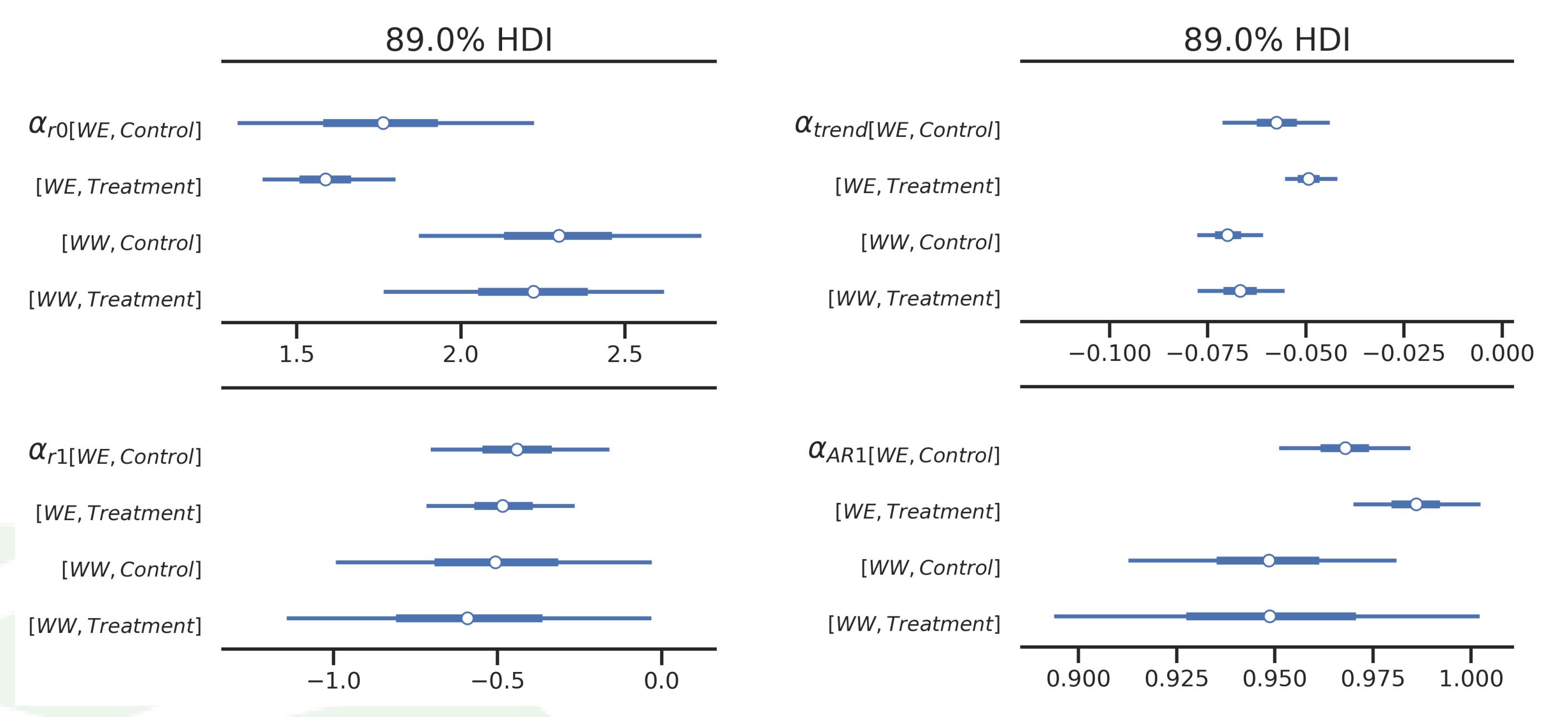


Figure 2b: Group level parameter means for combinations of site and control/treatment. Top: the overall trend parameter or intercept term (α_{trend}). Bottom: the autoregressive term applied to SM_{t-1} (α_{AR1}).

The treatment sites show slightly reduced responsiveness to rainfall (more apparent at WE), though there is much greater variation in the rainfall effect between site WE and WW (Figure 2a). Treatment sites tend to show slower drying rates (again more apparent at WE), but the variability between sites is still much greater between sites than treatments (Figure 2b).

Furthermore, the interaction between rainfall and past soil moisture is mostly small or negative suggesting that higher previous soil moisture leads to less change in soil moisture with rainfall. In other words drier soils infiltrate more water. The lack of difference observed between the treatments could simply be related to the fact that the treatments were only established 16 months before.

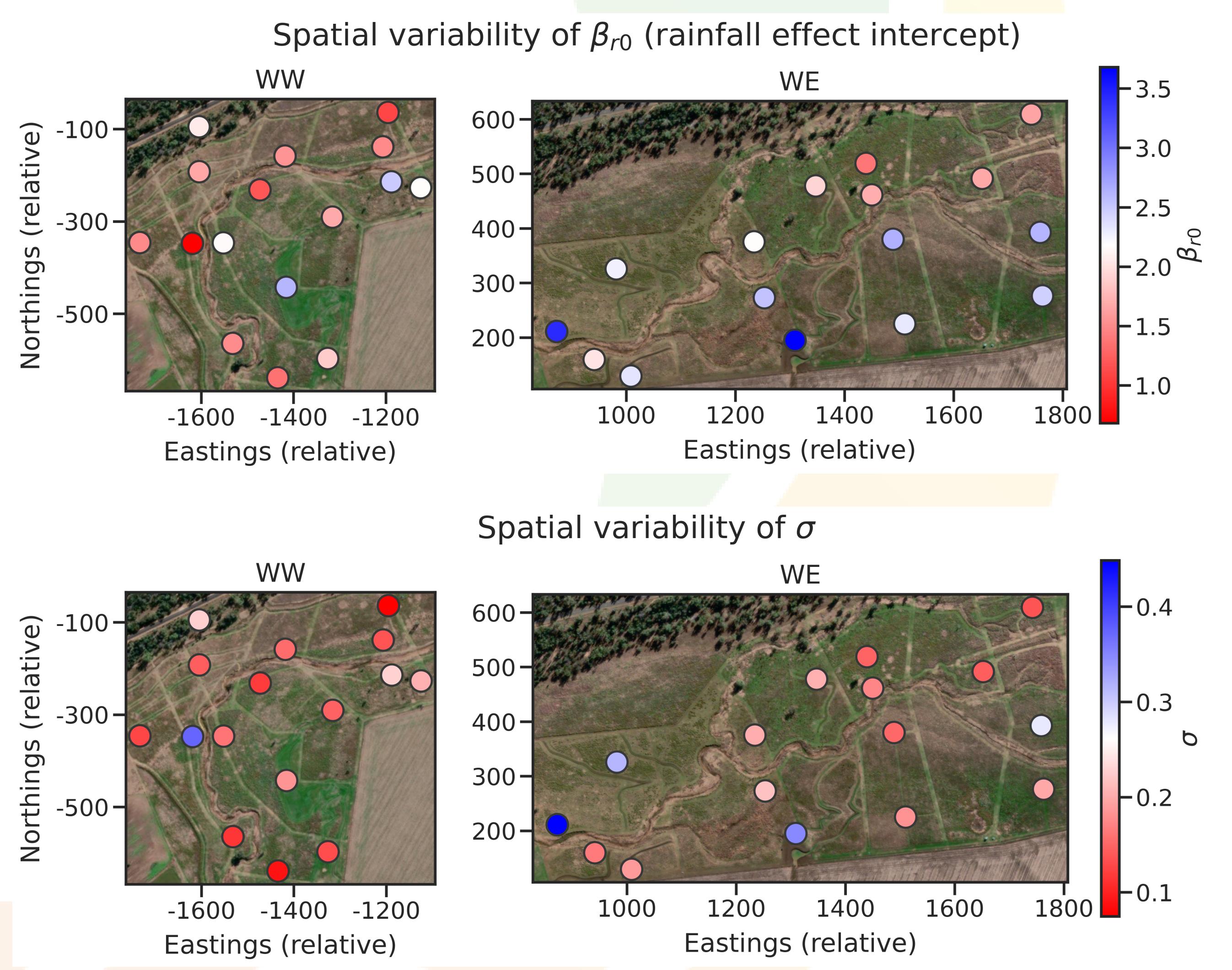


Figure 3: Top: Spatial variability in the strength of the parameter representing rainfall effect across the two sites (WE and WW). Bottom: Spatial variability in the variance of the residuals across the two sites (WE and WW).

Discussion

Figure 3 shows the spatial variability in the rainfall effect intercepts and the variance of the residuals. These broadly pick up the variation in soil, with coarser sandier soils on the northern side (and in WW) and deeper clay soils on the southern side in WE. Further work must be done to fully account for the spatial variability in the model, including for the residuals which were also found to be heavy-tailed.

The relevance of the framework is that it is interpretable, quantifies modelling uncertainty, can be easily extended to include more variables, and serves as a model for other environmental monitoring networks. Future work will include quantifying comparing with generalised additive models to capture non-linear effects and incorporating more spatial information (potentially using Gaussian Processes).

Code

github.com/dare-centre/llara-soil-moisture

Funding

Funding for the Landscape rehydration project is through a generous donation.

We thank the Australian Government for supporting this research through the Australian Research Council's Industrial Transformation Training Centre in Data Analytics for Resources and Environments (DARE) (project IC190100031).