Estimating soil moisture using a Bayesian data assimilation approach

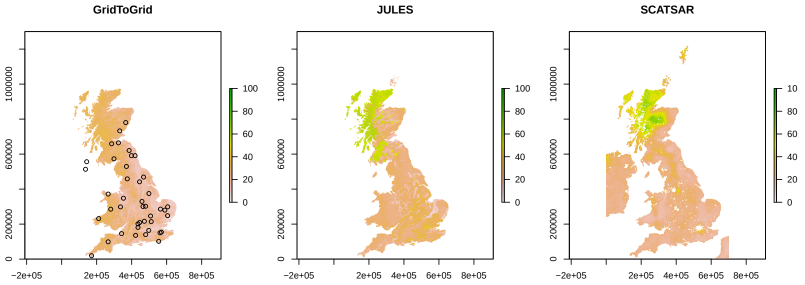
Peter Levy

2019-08-12

Background

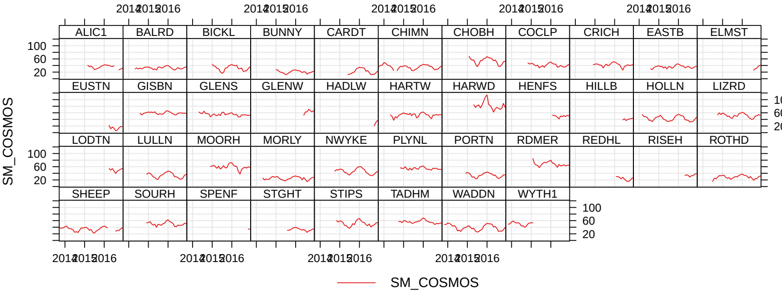
- Soil moisture affects many processes
  - plant growth, gas emissions, flood risk . . .
- Several different sources of data
  - direct measurements
  - hydrological models
  - satellite measurements
  - new Scatterometer - Synthetic Aperture Radar product (SCATSAT) gives daily 1-km soil moisture in near real-time
- How to combine data sources?
  - e.g. calibration of satellite estimates
  - e.g. upscaling of direct estimates
- How to propagate uncertainty?

Soil Moisture Estimates



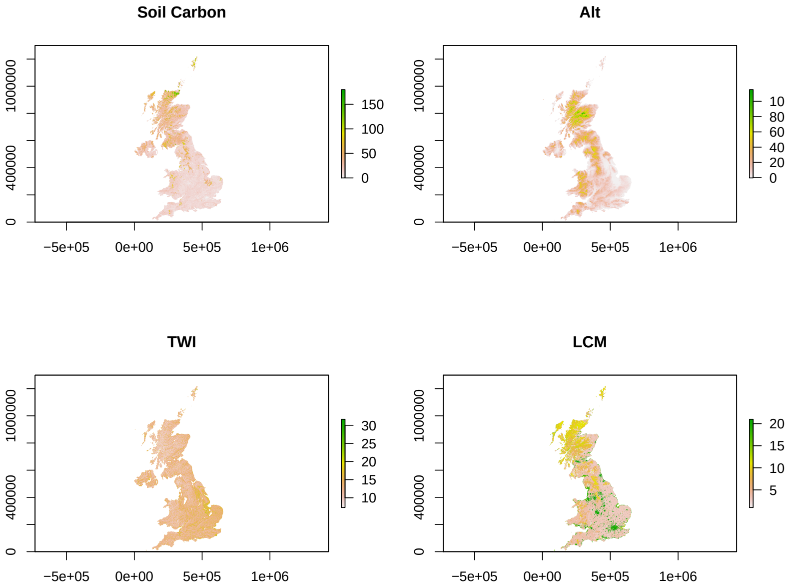
- GridToGrid hydrological model
- JULES surface energy balance model
- SCATSAR satellite observations

COSMOS



- calibrated ground-truth data
- 42 sites
- daily data for 2013-2015

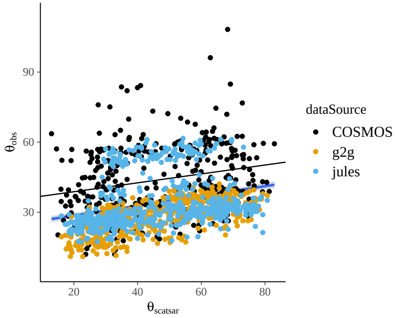
Other related variables

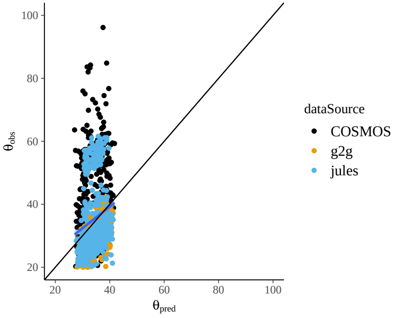


Approach

- base on SCATSAR satellite data
- calibrate on COSMOS data - closest to absolute truth
  - but only 42 sites
- incorporate information from Grid-To-Grid and JULES
  - detailed spatial-temporal patterns (1-km, daily)
- account for different numbers of observations (42 vs 200,000)
- Bayesian parameter estimation to account for uncertainty

Model 1

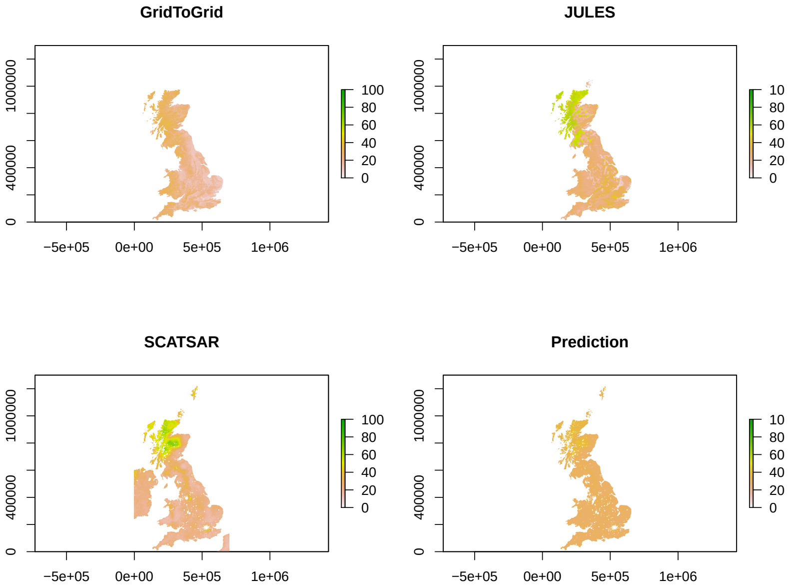




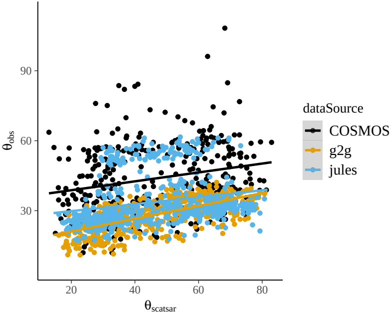
Simplest calibration approach - use all data in a linear model

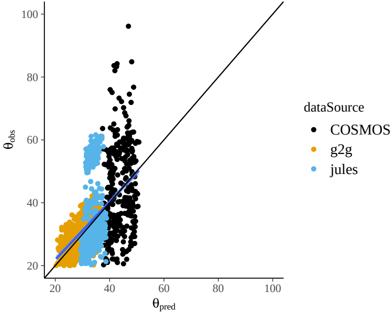
$$\theta_{obs} = \beta_0 + \beta_1 \theta_{scatsar}$$

Model 1


GridToGrid
JULES
SCATSAR
Prediction
0
400000
1000000
−5e+05
0e+00
5e+05
1e+06
100
80
60
40
20
0
10

Model 2

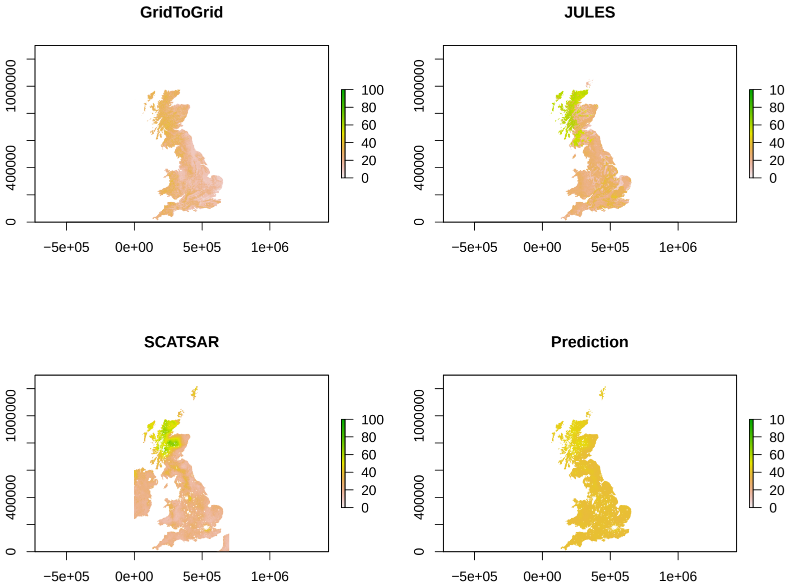




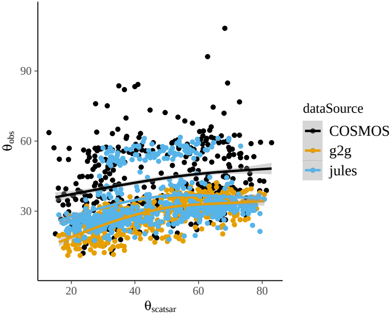
Add model discrepancy terms: slope and intercept for each data source j

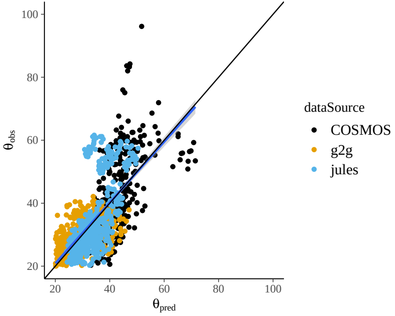
$$\theta_{obs,j} = \beta_0 + \beta_1 \theta_{scatsar} + \beta_{0j} + \beta_{1j} \theta_{scatsar}$$

Model 2


GridToGrid
JULES
SCATSAR
Prediction
1000000
400000
0
−5e+05
0e+00
5e+05
1e+06
100
80
60
40
20
0
10
80
60
40
20
0

Model 3

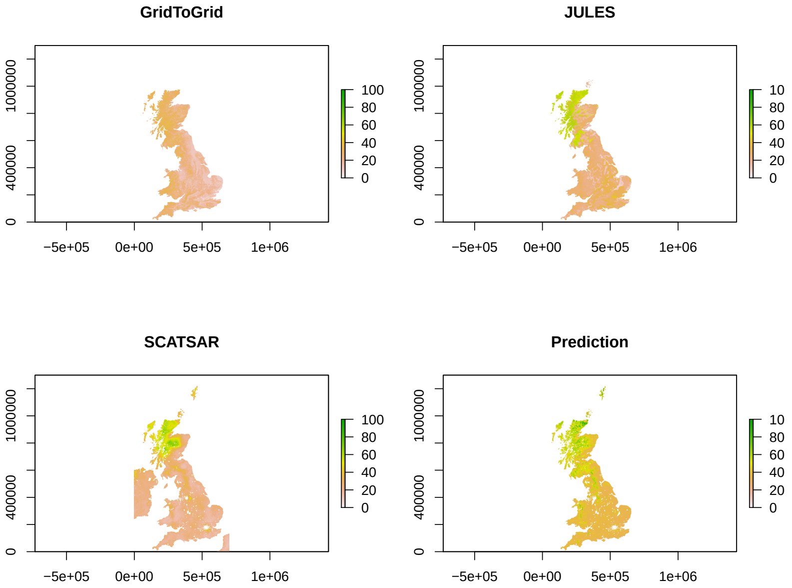




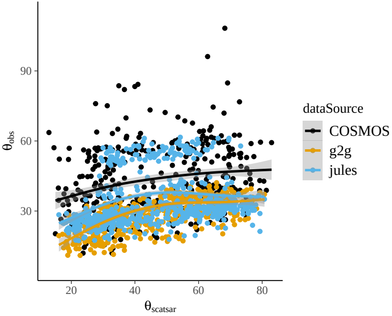
Add slopes/intercepts for soil carbon, altitude, topographic wetness index

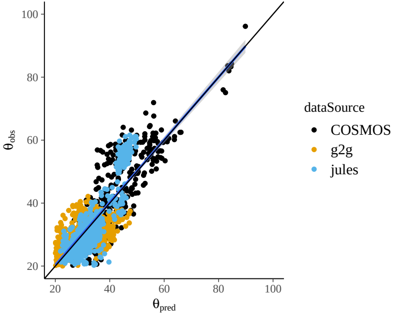
$$\theta_{obs,j} = \beta_0 + \beta_1 \theta_{scatsar} + \beta_{0j} + \beta_{1j} \theta_{scatsar} + \beta_2 C_{soil} + \beta_3 z_{alt} + \beta_4 I_{TWI}$$

Model 3


GridToGrid
JULES
SCATSAR
Prediction
1000000
400000
0
−5e+05
0e+00
5e+05
1e+06
100
80
60
40
20
0
10
80
60
40
20
0

Model 4

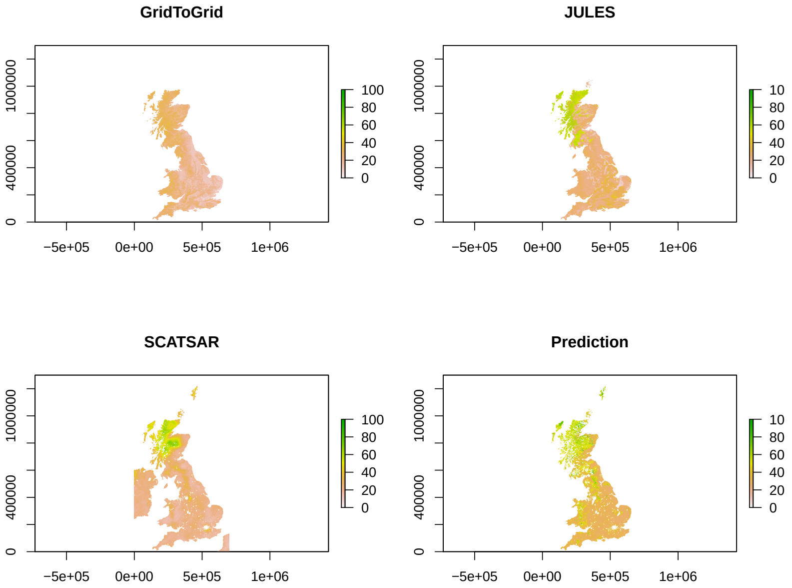




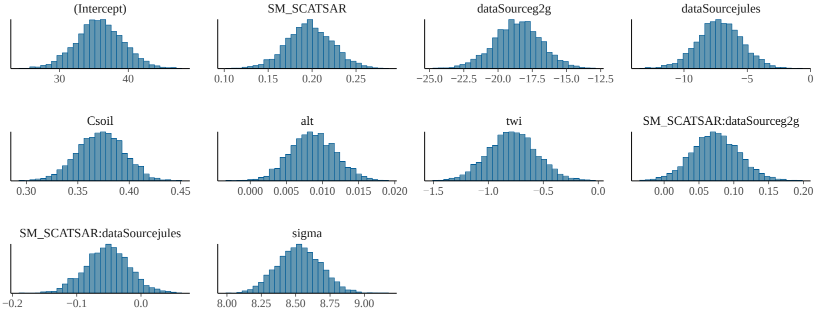
Add slopes/intercepts b for each Land Cover Map class k

$$
\begin{aligned}
\theta_{obs,j,k} = {} & \beta_0 + \beta_1 \theta_{scatsar} \\
& + \beta_{0j} + \beta_{1j} \theta_{scatsar} \\
& + \beta_2 C_{soil} + \beta_3 z_{alt} + \beta_4 I_{TWI} \\
& + b_{0k} + b_{1k} \theta_{scatsar}
\end{aligned}
$$

Model 4

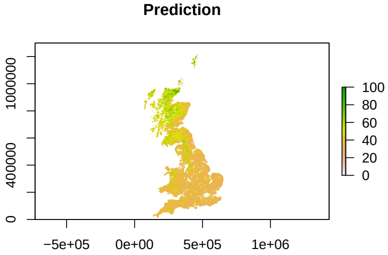

GridToGrid
JULES
SCATSAR
Prediction
0
400000
1000000
−5e+05
0e+00
5e+05
1e+06
100
80
60
40
20
0
10
80
60
40
20
0

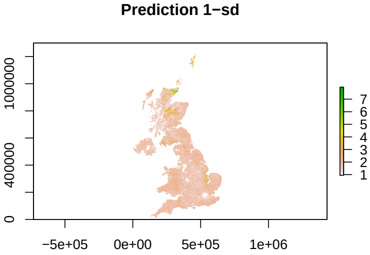
Uncertainty: Posterior distribution of parameters



Estimate model parameters using Hamiltonian MCMC in rstanarm.

Uncertainty in posterior predictions





Conclusions

Method allows us to:

- combine four different data sources in a probabilitically coherent way
- include uncertainty
- map posterior predictions and uncertainties

Further issues:

- representivity of COSMOS sites
  - cf. 1-km square
  - LCM classes
- computation time with daily and 1-km data
  - how much spatial/time averaging appropriate?