Quantifying the uncertainty in the prediction of soil properties from NIR and MIR soil-spectra using local and regional spectral libraries.

T Breure[1], A Milne[1], R Webster[1], J Hannam[2], S Haefele[1] , R Corstanje[2]

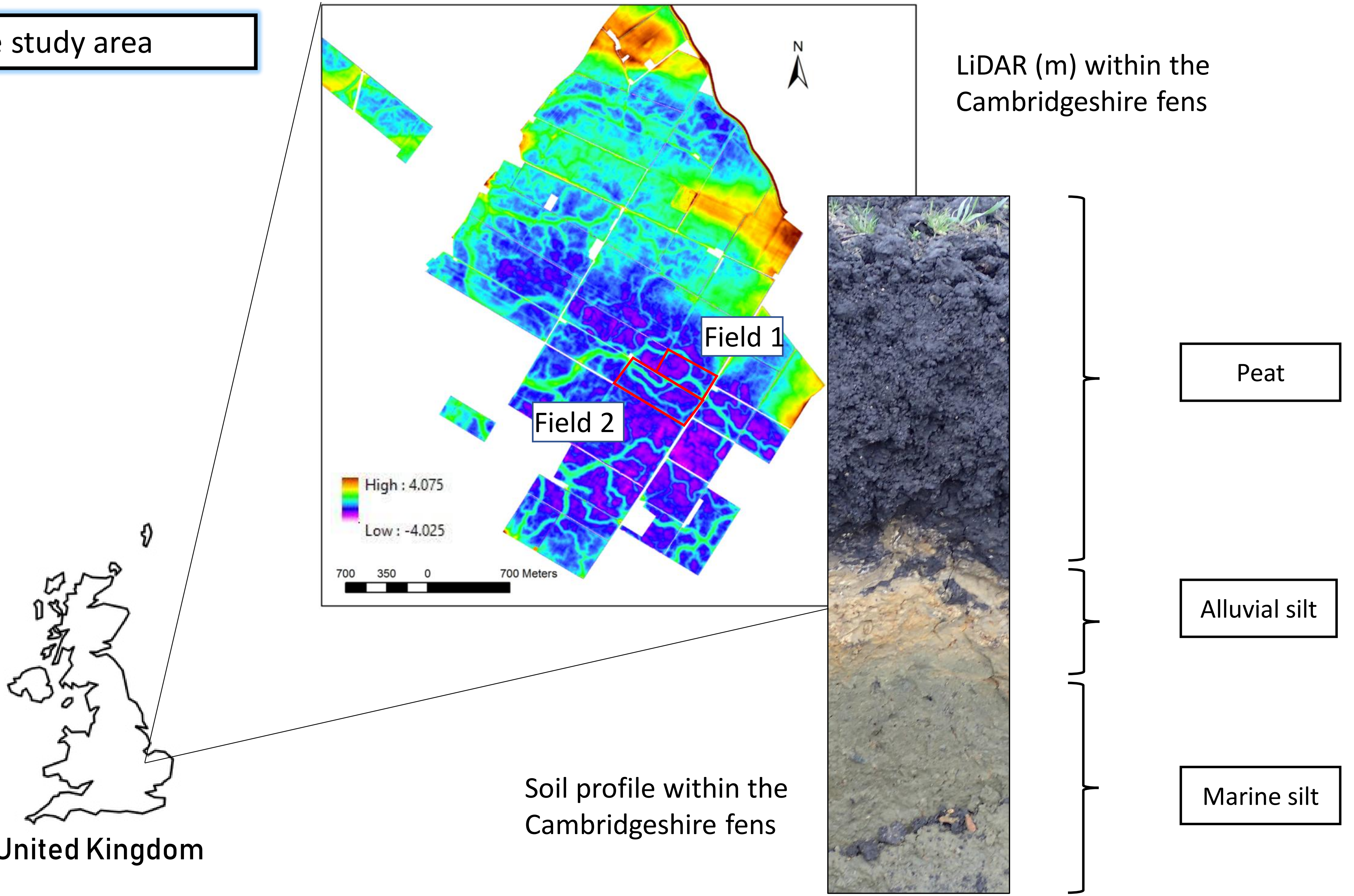
[1] Rothamsted Research, Harpenden, UK; [2]Cranfield University, Cranfield, UK

Objectives

IR spectroscopy to predict laboratory reference values of soil properties is a well established subject within the literature.
For this to be practically useful, however, the estimates of soil properties need to be sufficiently accurate which can be resource intensive.
One method to reduce this effort is by using an existing dataset of a larger spatial extent (regional library) that has paired laboratory measurements and IR spectra

Objective 1	Quantify the errors in prediction of soil properties from local and regional libraries
Objective 2	Quantify uncertainty of soil properties across case-study fields using predictions from local and regional libraries

Case study area

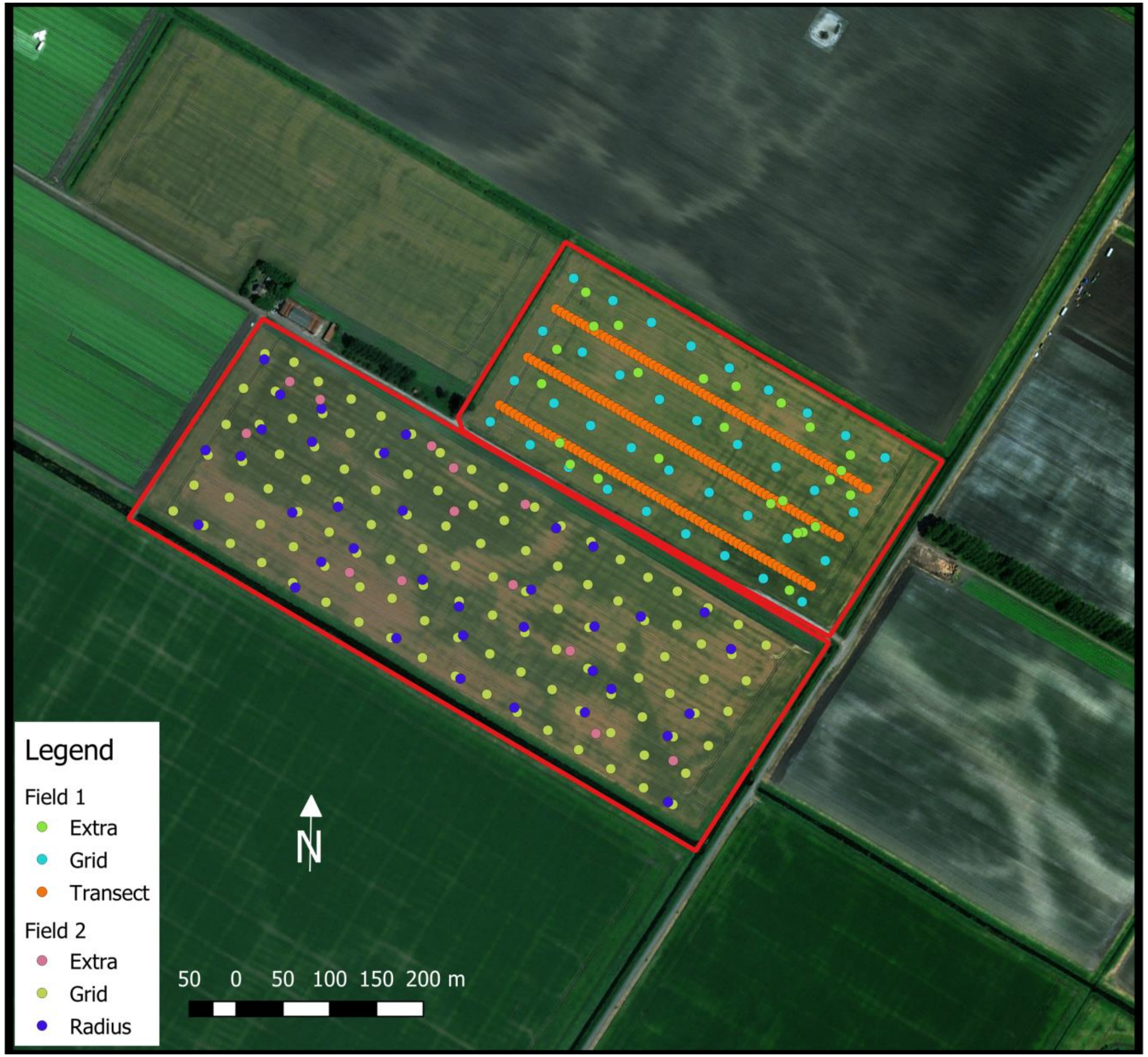


30 samples from each field have been analysed in the laboratory for:

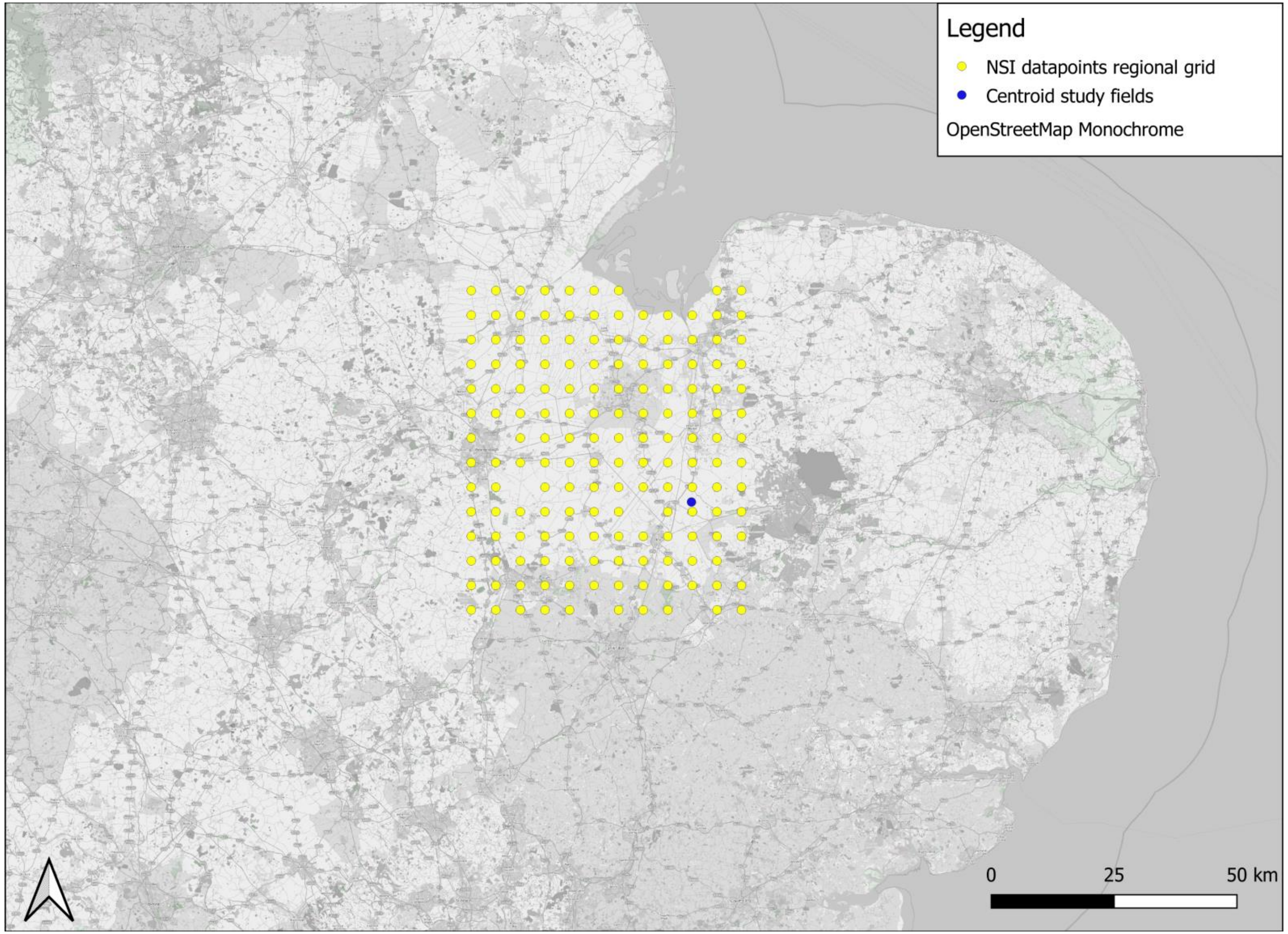
- Organic C (%)
- Clay (%)
- Potassium (%)

These laboratory reference values together with their paired spectral measurements make up the local library

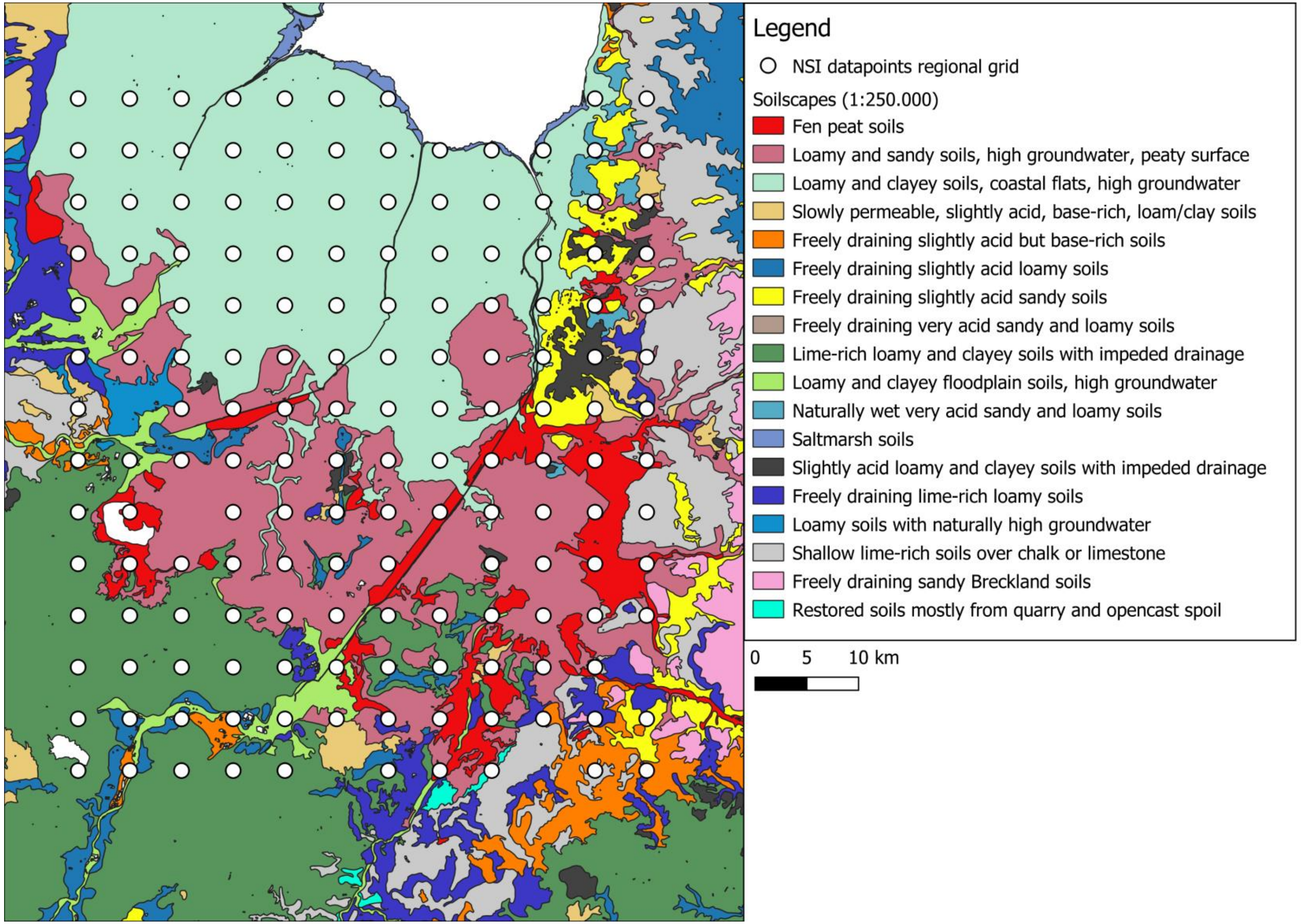
Sampling design



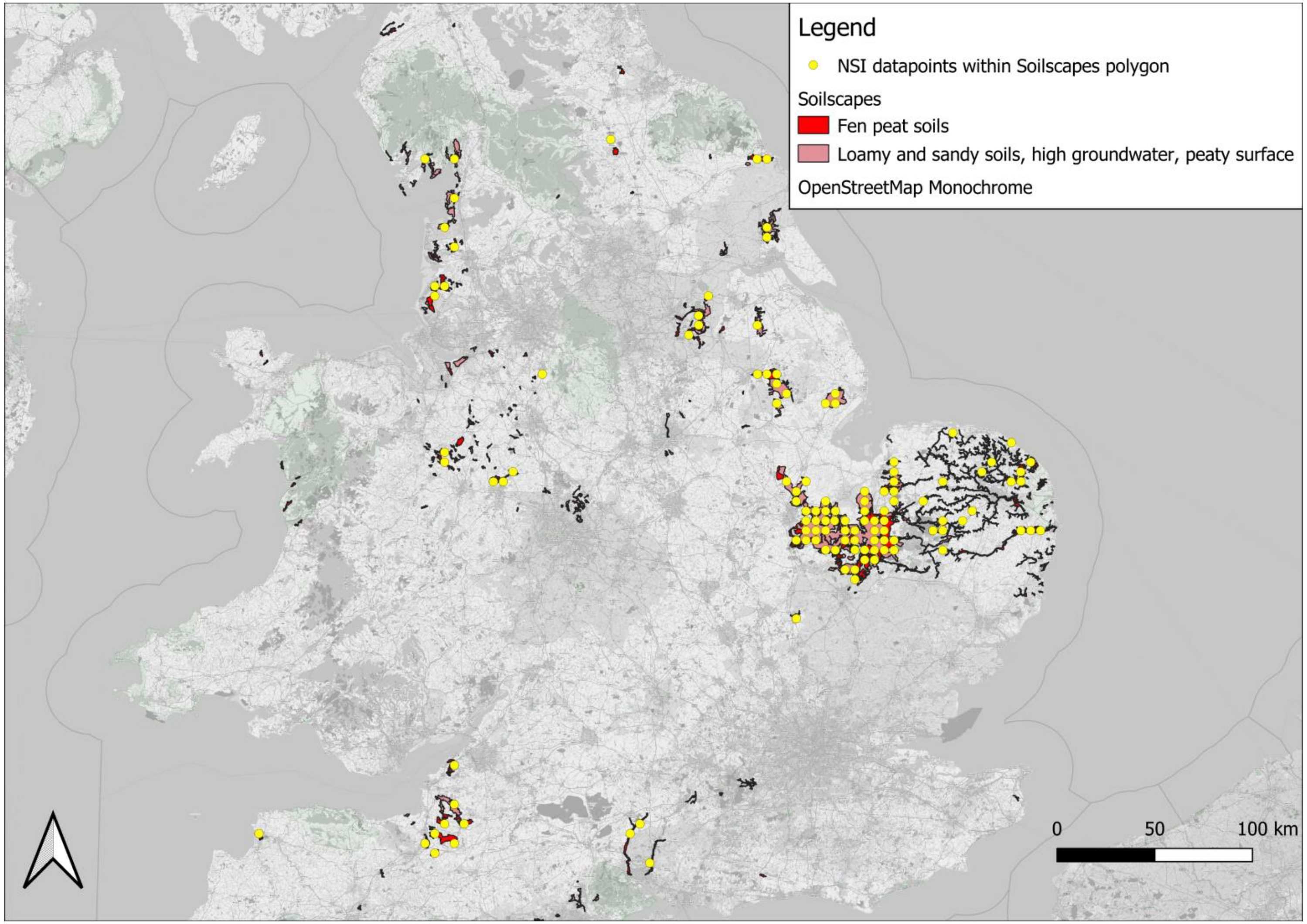
National soil inventory samples section by regional grid



National soil inventory samples selected by regional grid overlaid on Soilscapes map



National soil inventory sample selection by stratification



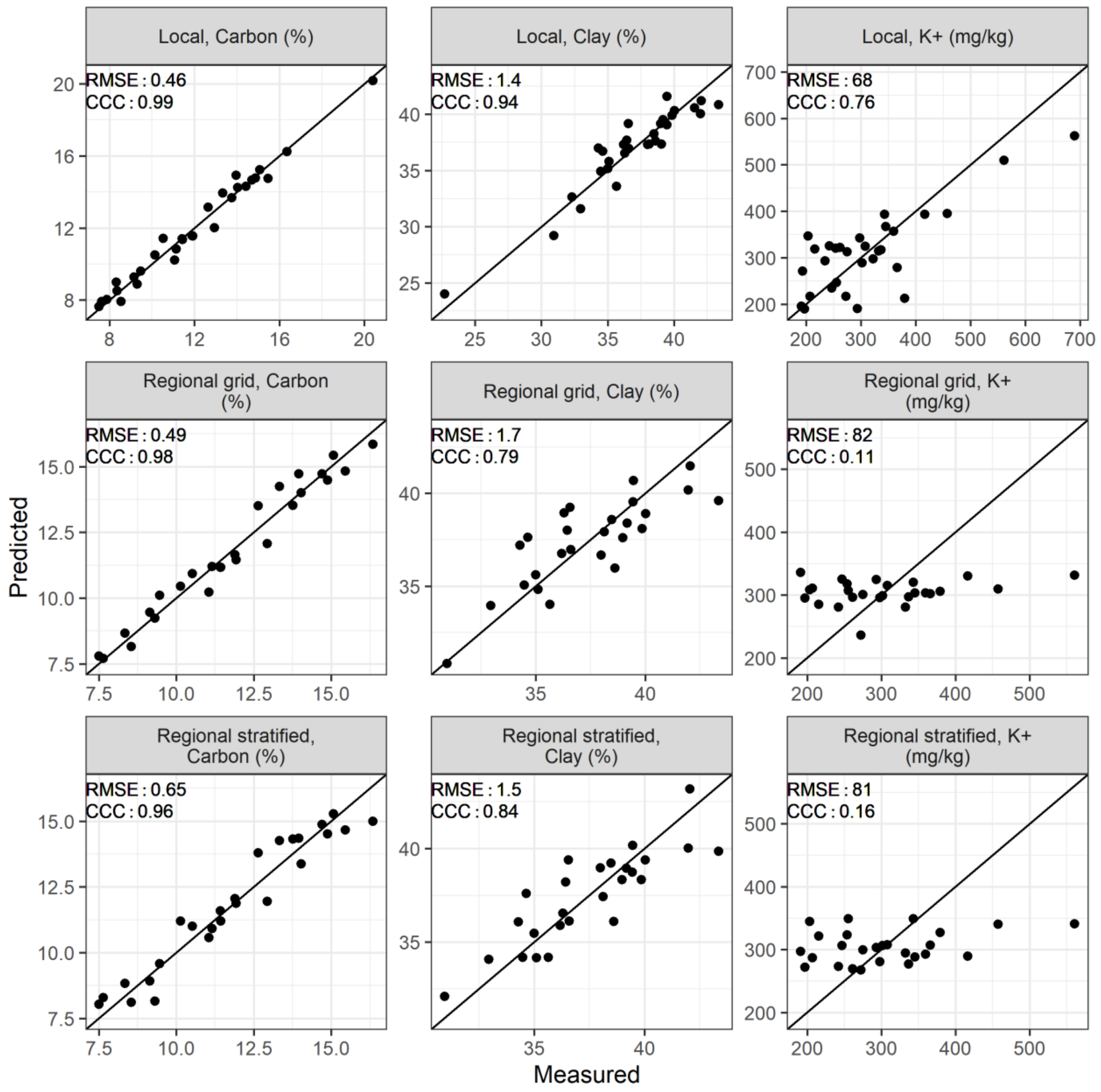
Methodology for deriving the prediction model

1. Partial least squares regression to predict soil properties from NIR and MIR spectra
2. Granger-Ramanathan model averaging using OLS regression

We quantify 3 sources of uncertainty:

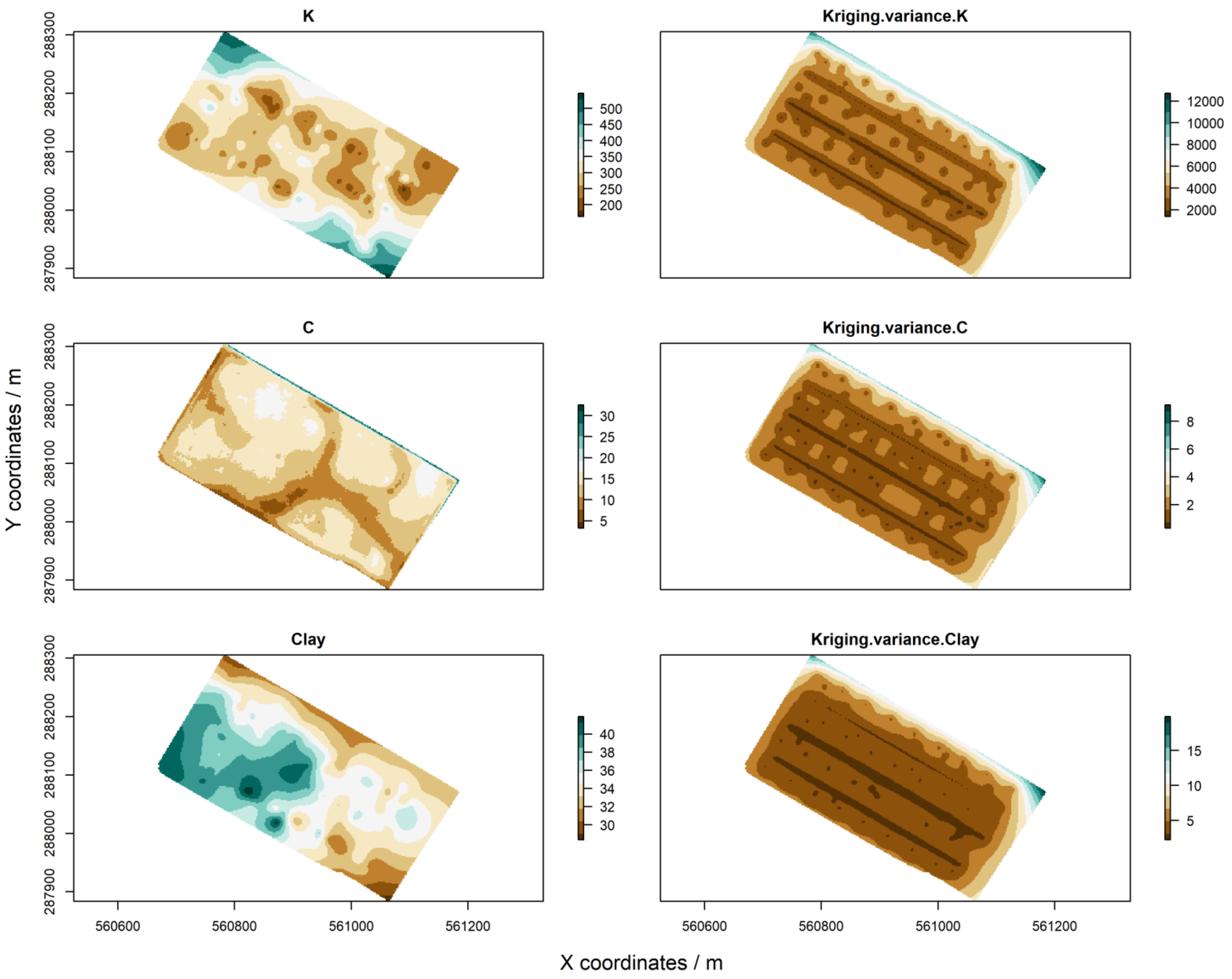
- Laboratory reference measurement error
- PLS regression prediction error
- Spatial prediction error

PLSR predictions of soil variables - Field 1



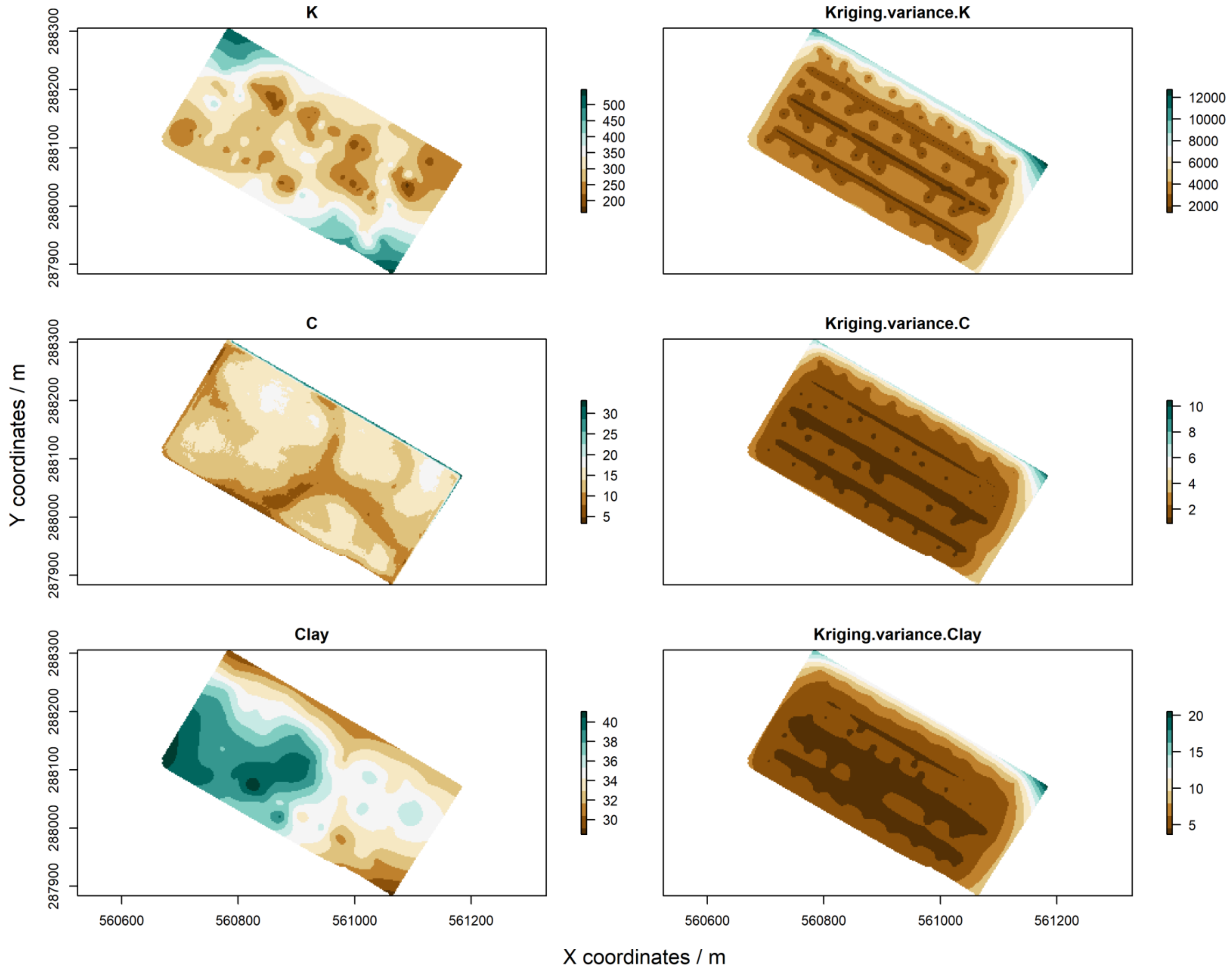
Local library kriging predictions accounting for laboratory reference measurement error

We used residual maximum likelihood estimation (REML E-BLUP) to estimate the variogram by fitting a linear mixed model with the trend accounted for as fixed effects

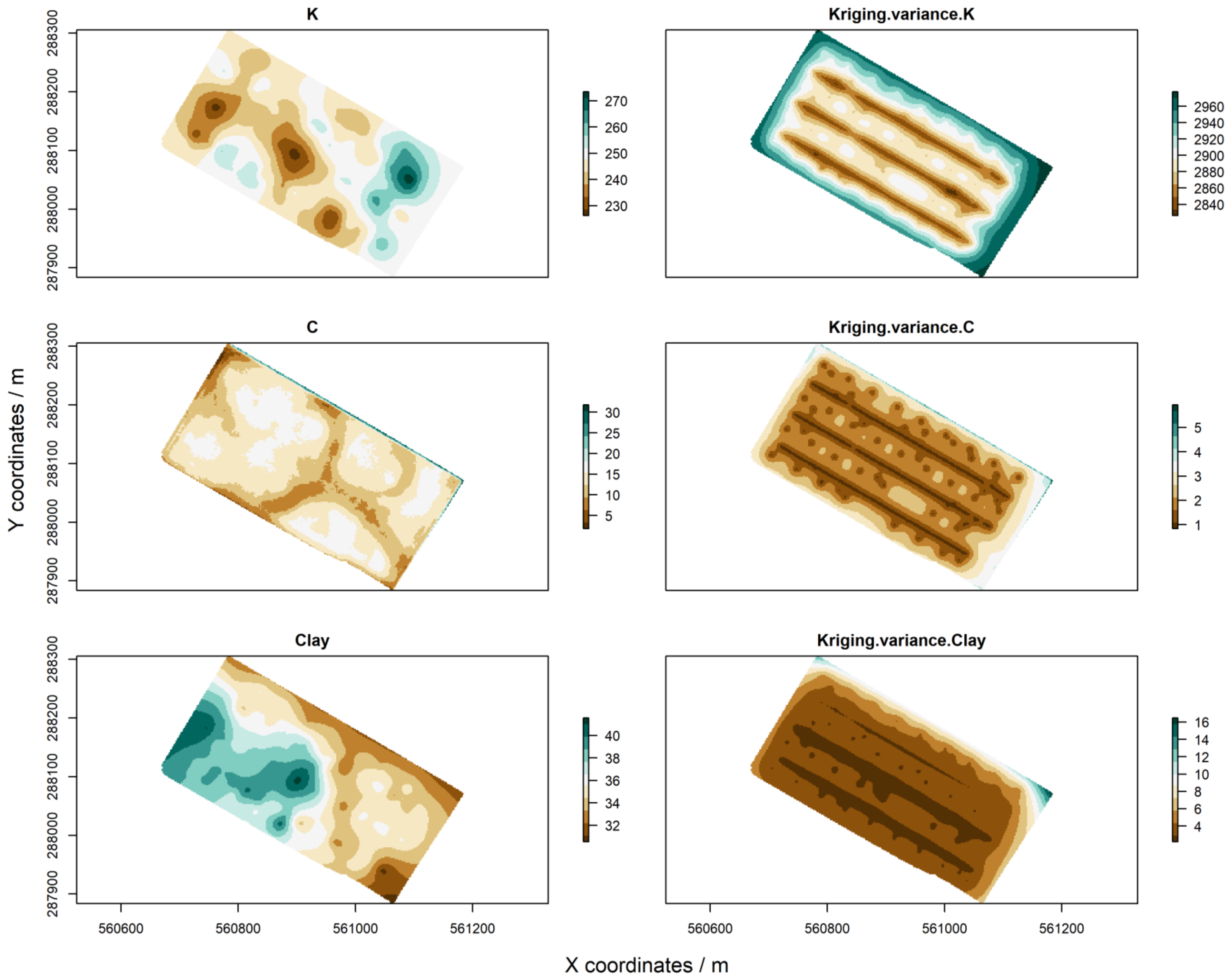


In case the nugget variance was fitted at a lower value than the known error, we fixed the nugget variance to the known error and re-estimated the model parameters.

Local library kriging predictions accounting for laboratory reference measurement error and PLS prediction error



Regional library kriging predictions accounting for PLS prediction error



Conclusions

- Kriging variance of Potassium did not change once accounted for the PLS regression prediction error (indicating short-range unresolved spatial variation)

- Kriging variance of organic C and Clay increased by 10% and 16%, respectively, once accounted for the PLS regression prediction error

- Kriging variance from the regional libraries (where we spiked the regional dataset with a small subset of local samples) was lower compared to the local dataset.

- However, these predictions do not take into account the laboratory reference measurement error from the regional data.

- Cross validation should determine whether low kriging variance is offset by higher errors in prediction accuracy