

Using vis-NIR to predict soil organic carbon and clay at national scale: validation of geographically closest resampling strategy

Yi Peng, Maria Knadel, Mette B. Greve, Mogens H. Greve

E-mail address: yi.peng@agro.au.dk

Background

Predicting soil organic carbon at field scale using a national soil spectral library.

J. Near Infrared Spectrosc. 21, 213-222. Peng, et al 2013.

➤ 7 km soil monitoring grid

1. Full library & spiking library

2. Resampling strategy

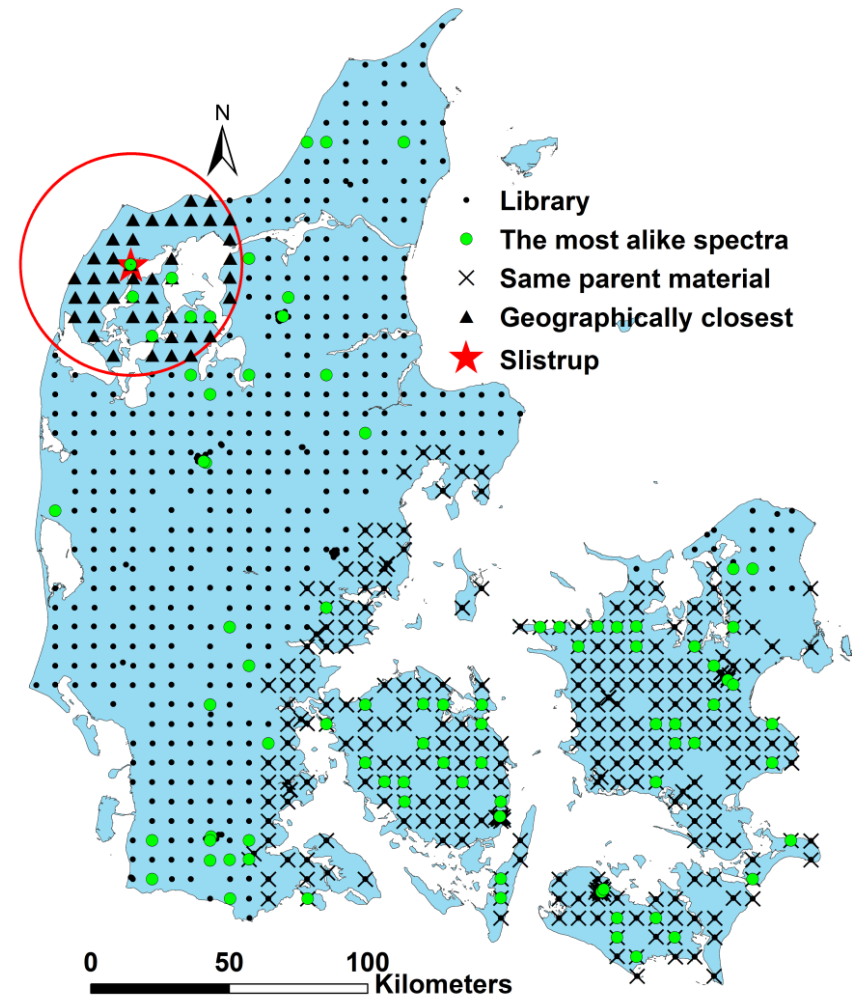
➤ Geographically closest (30km)

➤ Same parent material

➤ Mahalanobis distance

➤ Only one local field used of validation

➤ Distance?



Material and methods

Danish soil spectral library:

- 4650 spectra with SOC data (Range: 0.01-56.21%)
- 4512 spectral with clay values (Range: 0.1-72.28%)

Resampling strategy:

- Geographically closest 20, 30, 40, 50 points (soil profiles)

Chemometric modelling and validation:

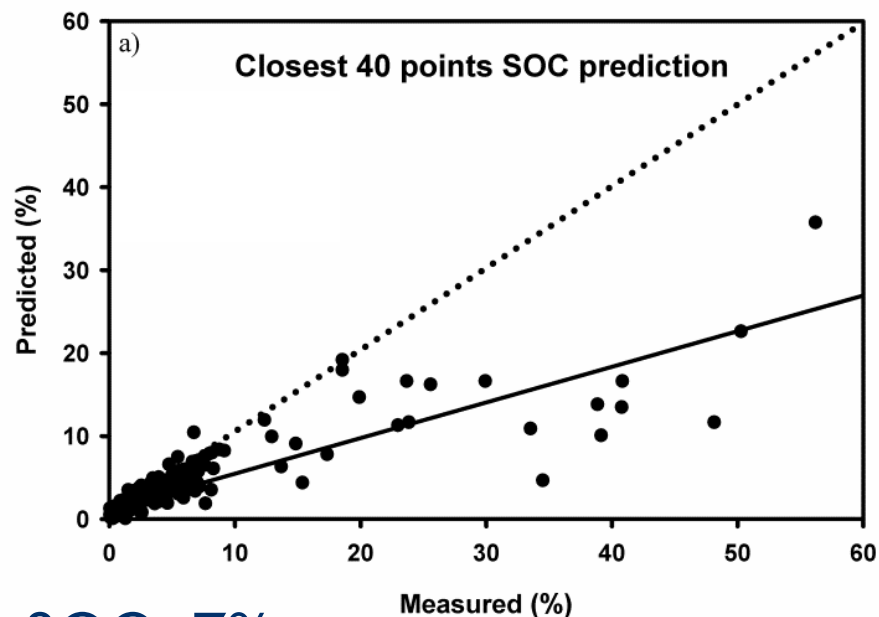
- Partial least square regression (PLSR)
- Standard deviation of differences (SDD) value of **0.17%** (SOC range: 0.4 to 3 %) (Peng et al., 2013)
- Standard deviation of reproducibility value of **1.3%** (clay range: 2 to 23%) (Sørensen and Dalsgaard, 2005)
- We accept $RMSE \leq 2 \times \text{lab measurement error}$

Results

SOC predictions

(Closest 40 point)

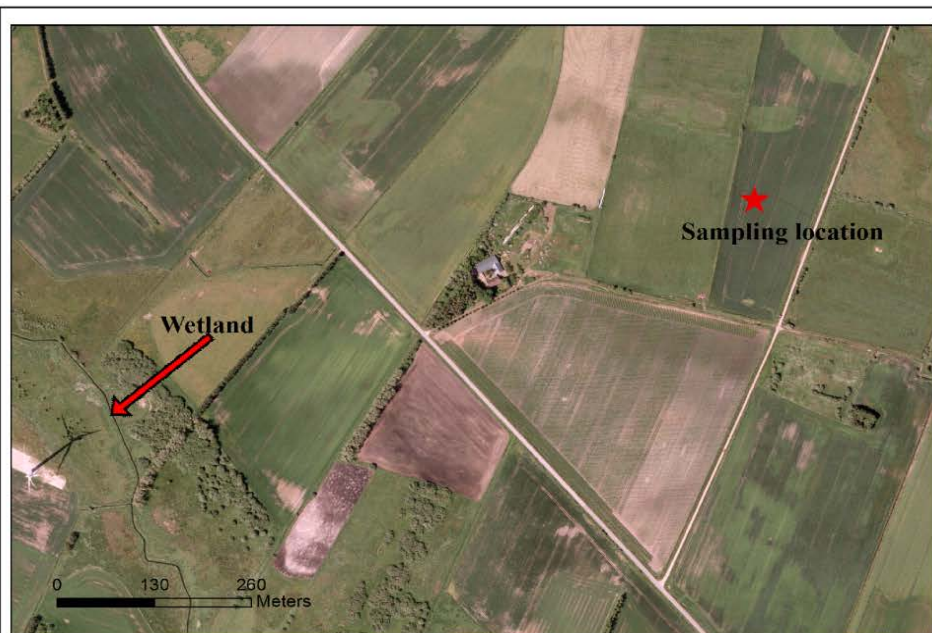
$R^2=0.76$; $RMSE=4.01$; $RPIQ=0.35$



$SOC < 7\%$

$R^2=0.82$; $RMSE=0.54$; $RPIQ=2.16$

$2 \times \text{lab error} = 0.34\%$



★ Sampling location

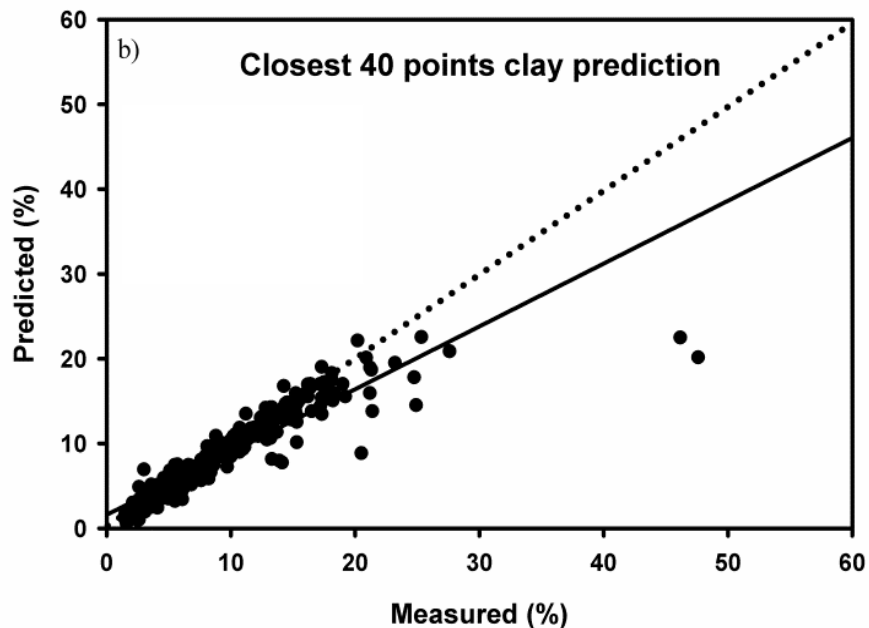
Measured SOC = 25.6%
Predicted SOC = 16.2%



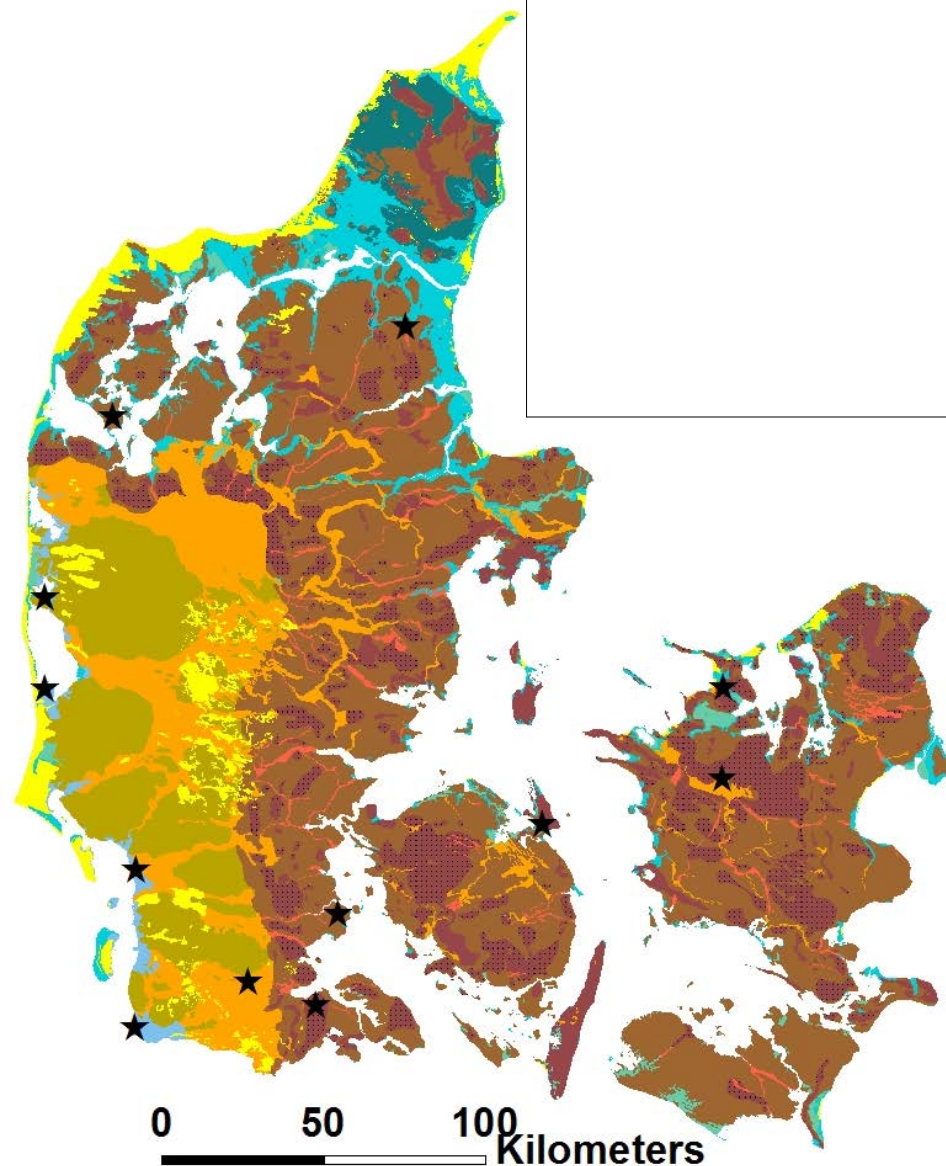
Clay predictions

Closest 40 points:

$R^2=0.86$; $RMSE=2.36$; $RPIQ=2.88$



$2 \times \text{lab error} = 2.6\%$



Conclusions

- 40 geographical closest sampling points
- SOC: land use/ landcover
- Clay: soil parent material and landscape