Improved spectral estimation of multiple soil properties by stratification on ancillary and spectral data

Harm Bartholomeus*, Clifton Sabajo, Annemarie van Groenestijn, Lammert Kooistra

Harm.Bartholomeus@wur.nl

Present state of soil spectroscopy

- Good prediction accuracies are achieved with chemometric techniques for multiple soil properties
  - Local calibrations → global calibrations
- Large datasets are being created
  - Large ranges in soil properties
  - Large variations in "background" reflectance
- "New" data can not directly be added if
  - Other setups/protocols are used
  - There are differences in mineralogy

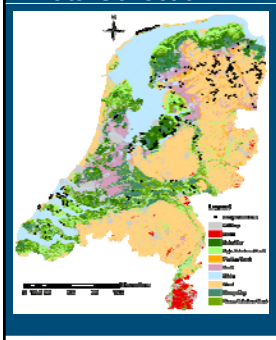
Research question

Can stratification improve model robustness and estimates of multiple soil properties using large spectral databases?

Stratification based on

- auxillary data
  - Soil texture map
- spectral properties
  - K-means clustering on spectral data (PCA transformed)
- pre-classification
  - Two-step approach
  - Basically: you could stratify based on e.g. SOM easily, but to which class should it be assigned?

Data Collection



422 soil samples from natural ecosystems

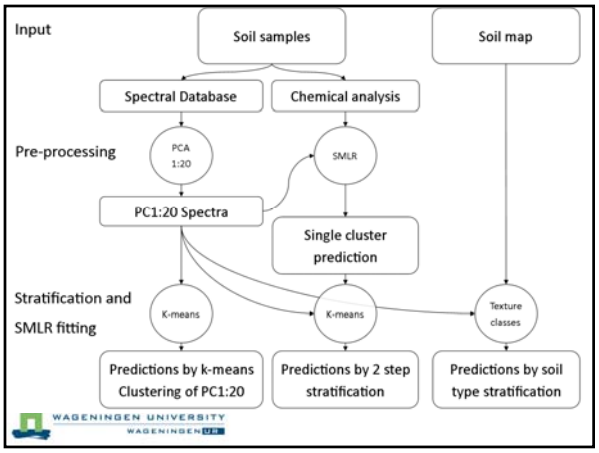
- Wide range soil types
- Wide range soil property concentrations

Laboratory analysis

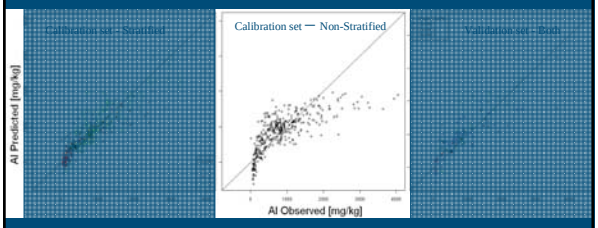
- pH, **SOM,** Nitrogen (total and soluble)
- NH4, NO3, P-total, PO4
- Mg, Ca, K, Al

Samples measured in laboratory setup with ASD Fieldspec Pro FR & Contact probe (protocol Soil Spectroscopy Group)

Source: Wamelink, Slim et al., 2009

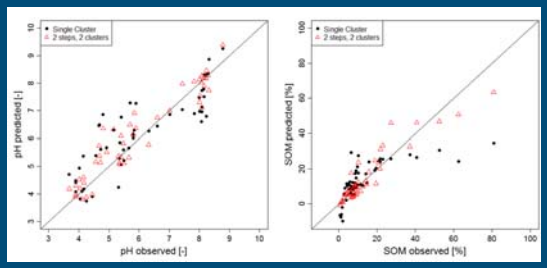


Results



Calibration set: N=371
Validation set: N=51

Results



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Results - Summarized

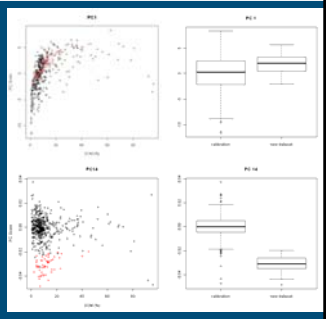
Table 1: Ratio of Performance to Deviation (RPD) values for stratified models by means of different stratification procedures and cluster sizes. The green cells indicate an improved performance, compared to the prediction by means of the non-stratified reference models.

Attribute	-	Soil Type	2-Step	2-Step	PC1-20	PC1-20
# clusters	1	3	2	3	2	3
NT	1.52	1.45	1.69	**1.82**	1.72	1.72
PT	2.22	1.92	2.50	2.51	**2.55**	2.10
K	1.29	**1.36**	1.13	1.06	1.01	0.91
MG	1.33	**1.45**	1.35	1.31	1.31	1.32
N-NH4	1.35	1.28	1.31	**1.46**	1.30	1.39
N-NO3_NO2	1.16	1.13	1.17	**1.18**	1.16	1.10
NTS	1.38	1.46	1.41	1.43	**1.47**	1.34
AL	1.67	1.65	2.19	2.26	2.20	**2.28**
CA	1.76	1.92	**1.99**	1.89	1.89	1.65
SOM	1.62	1.93	**2.75**	1.71	1.79	1.96
pH	2.46	2.72	**3.11**	2.81	3.00	2.95

WAGENINGEN UNIVERSITY
WAGENINGEN UR

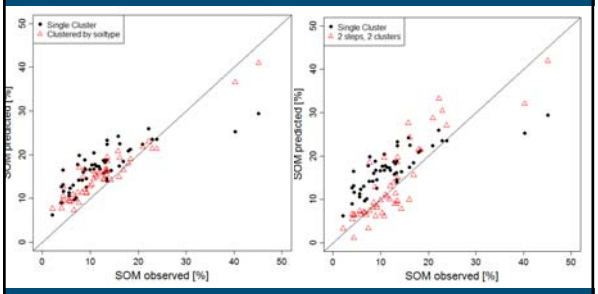
Estimate within field variation of local dataset

- Single field
- 51 Samples
- 2 soil types
- SOM
- Selection of PC's because of different setup
  - Illumination geometry
  - Observation geometry
  - Sample preparation



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Results



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Conclusions

- ✓ Stratification of VNIRS models can improve the estimate of multiple soil properties compared to the use of a single model.
- ✓ In general, a two-step approach yields the most consistent results, although there is no method which is always better than the others.
- ✓ With appropriate stratification methods, the potential of large spectral databases for VNIR soil spectroscopy and proximal soil sensing can better be used.

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Thank you for your attention

WAGENINGEN UNIVERSITY
WAGENINGEN UR