Using soil spectral libraries in support of proximal soil sensing

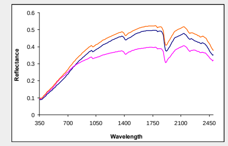
Yufeng Ge[1], Cristine L.S. Morgan[2], J. Alex Thomasson[1]

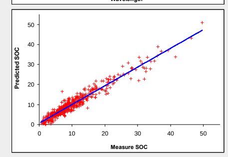
[1]Biological & Agricultural Engineering, Texas A&M University, TX, USA
[2]Soil & Crop Sciences, Texas A&M University, TX, USA

Workshop on Proximal Soil Sensing
May 15-18, 2011 | Montreal, CA

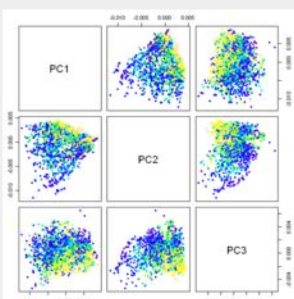
Optical Sensors for Proximal Sensing

- Advantages
  - By one measurement several soil attributes can be known
  - Sample presentation to the sensor can be simpler
  - Measurement along the soil profile
- A major disadvantage
  - Calibrations to relate optical measurements to target soil properties are required





Distribution of SOC in Principal Component Space



SOC (g kg^{-1})

- 0.1 ~ 1.0
- 1.0 ~ 1.6
- 1.7 ~ 2.4
- 2.5 ~ 3.6
- 3.7 ~ 5.4
- 5.5 ~ 8.2
- 8.3 ~ 14.2
- 14.2 ~ 88.5

Questions

1. How useful a spectral library would be for predicting local samples using optical sensors?

2. How much improvement can be achieved?
   - By selecting a subset of samples from the library
   - By using local boost samples

Objective

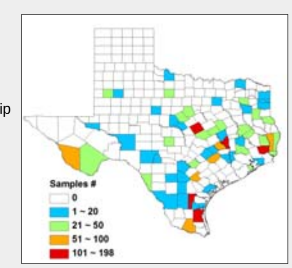
Use the Texas Soil Spectral Library to predict soil samples from local fields

1. Test two subset selection techniques
   a. Reduce redundancy (Kennard-Stone algorithm)
   b. Select similar samples (Spectral angle mapper)

2. Test the usefulness of local boosting samples

Texas Soil Spectral Library (***TSSL***)

- Soil sample spectra obtained with ASD AgriSpec 350-2500 nm
- ~ 3000 samples (~2400 have countyship in their records)
- ~130 foreign samples from Turkey, Honduras, and Jamaica



Methods

- Use the first derivative spectra and focus on soil organic carbon
- Model calibration: PLSR, 25 segments cross validation, the first local minimum for $RMSE_{cv}$
- 6 local fields (40-100 ha) 50 soil samples each field
  - 25 samples used for validation
  - 25 were reserved for local boosting
- Calibration Models
  - Full Texas Spectral Library (TSSL)
  - Kennard-Stone algorithm
  - Kennard-Stone algorithm + local Boosting
  - Spectral Angle Mapper
  - Spectral Angle Mapper + local Boosting

Kennard-Stone-algorithm

1. Find the 2 samples in the population with the highest distance between them.
2. For all candidate samples, find the smallest distances between each candidate and one of the already chosen samples.
3. Find the maximum of the smallest distances and the associated candidate is chosen.
4. Repeat procedure until a predefined size is reached.

--**result** keep 30% of library is redundant

Spectral angle mapper

Assume that the reflectance spectra of two soil samples are ***X*** ($x_1, x_2, \ldots, x_n$) and ***Y*** ($y_1, y_2, \ldots, y_n$), the spectral angle mapper measures the angle **θ** between them:

$$\theta(X,Y) = \arccos\left(\frac{<X,Y>}{|X||Y|}\right)$$

Where

$<X,Y> = x_1y_1 + x_2y_2 + \ldots + x_ny_n$ is the inner product of X and Y

$|X| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$ is the modulus of X

Results: *RPD Values*

Field	Library only(RMSD)	K-S	K-S+ Boost	SAM	SAM+ Boost
McLennan	1.6 (3.9)	1.6	1.6	1.3	1.5
Erath1	1.4 (9.2)	1.5	1.5	2.1	2.0
Erath2	1.6 (5.0)	1.6	1.7	1.5	1.7
Erath3	1.5 (5.7)	1.5	1.6	1.4	1.7
Comanche1	1.6 (6.9)	1.6	1.6	1.9	1.9
Comanche2	2.0 (4.7)	2.0	2.0	1.7	1.6

K-S = Kennard-Stone algorithm (70% retention)
SAM = Spectral Angle Mapper

Results: *Bias g kg^{-1}*

Field	Library only	K-S	K-S+ Boost	SAM	SAM+ Boost
McLennan	1.2	1.5	2.1	0.5	0.6
Erath1	-2.9	-2.5	-2.8	0.0	-2.0
Erath2	0.1	1.2	0.7	0.6	0.4
Erath3	-1.8	0.5	-0.5	-1.0	0.1
Comanche1	1.2	0.9	0.6	1.2	1.3
Comanche2	-0.7	0.0	-0.4	0.2	-0.6

K-S = Kennard-Stone algorithm (70% retention)
SAM = Spectral Angle Mapper

Conclusions

- Calibration with whole library calibration:
  - Satisfactory results; RMSE 3.9-9.2 g kg^{-1}
- Calibration with Kennard-Stone algorithm
  - Maintained or improved RMSD and RPD
  - Boosting maintained or improved over K-S alone
- Calibration with Spectral Angle Mapper algorithm
  - Larger gains and losses compared to K-S
  - 2:2:2 improved:same:worse
  - Boosting 3:2:1 improved:same:worsened