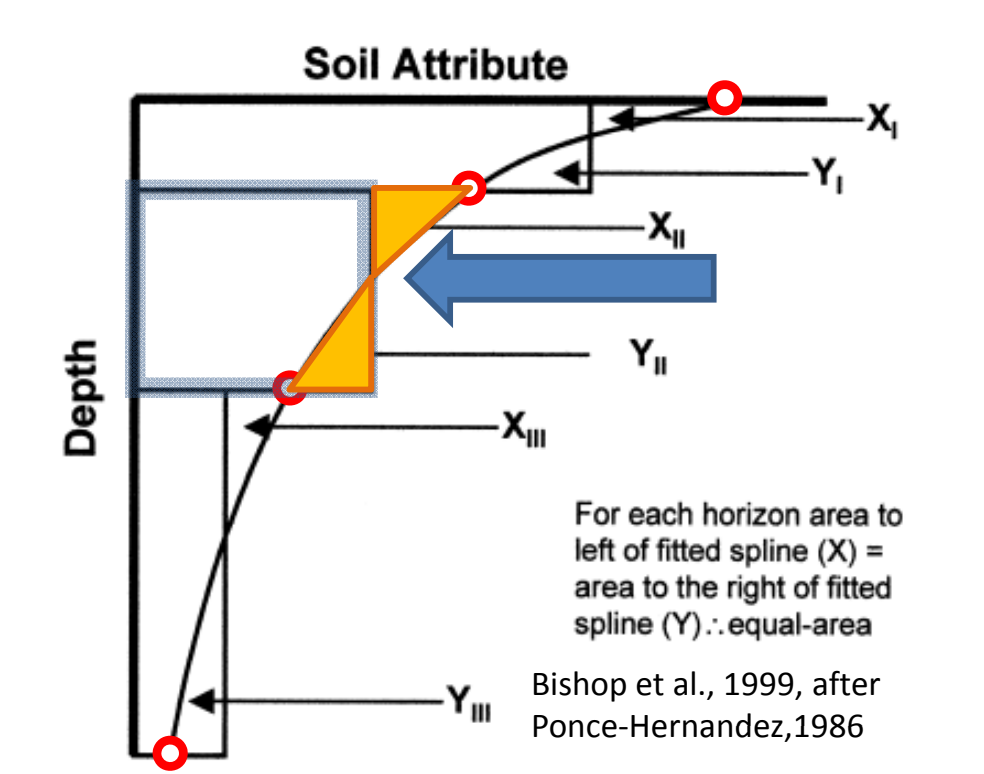
Gridded STATSGO2 property maps

Nathan Odgers, Zamir Libohova, James Thompson

Recap: equal-area splines

- Equal-area depth splines
- Consist of a series of local quadratic polynomials with "knots" or joins at horizon boundaries



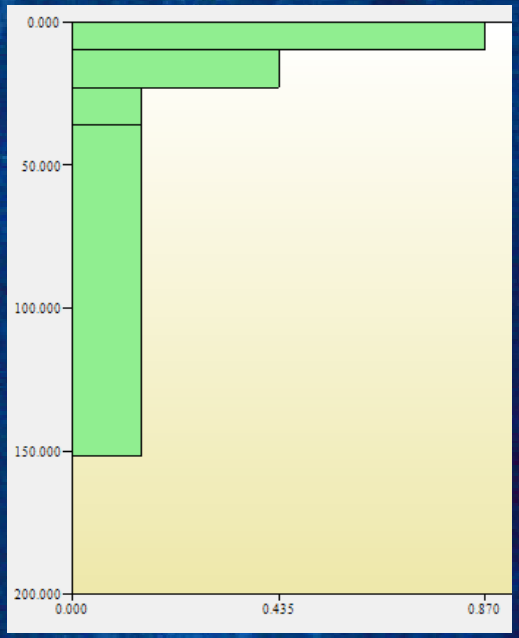
Bishop et al., 1999, after Ponce-Hernandez,1986

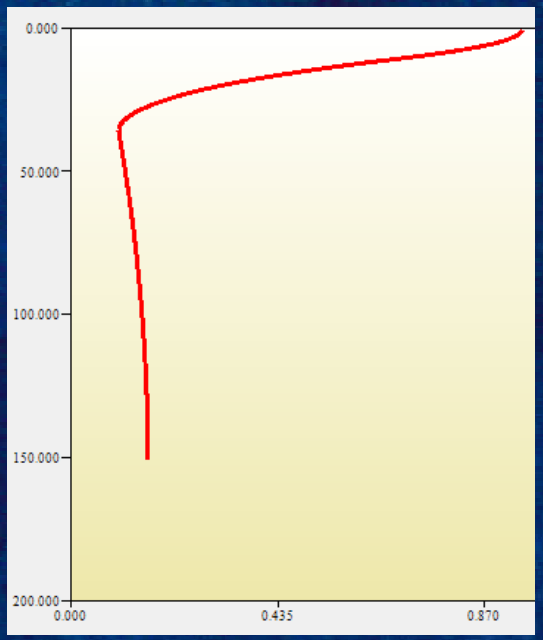
Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91: 27-45

Ponce-Hernandez, R., Marriott, F.H.C. and Beckett, P.H.T., 1986. An improved method for reconstructing a soil profile from analyses of a small number of samples. Journal of Soil Science, 37: 455-467

Recap: equal-area splines

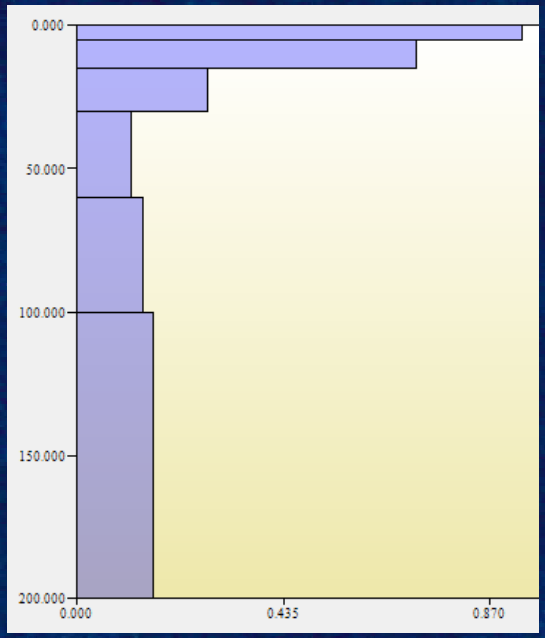
Soil profile with property data



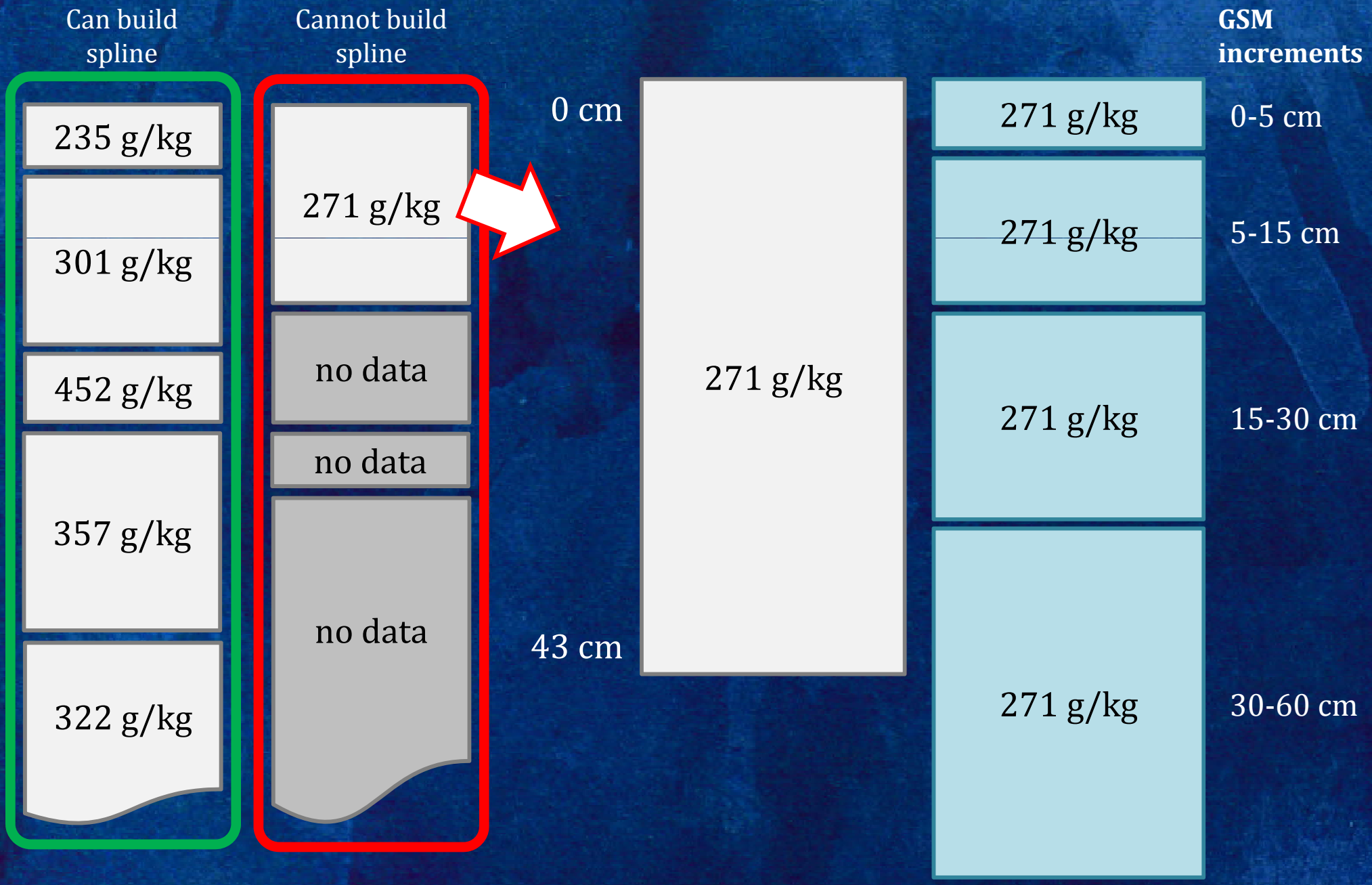


Continuous spline (1 cm resolution)

Property estimates at specific increments

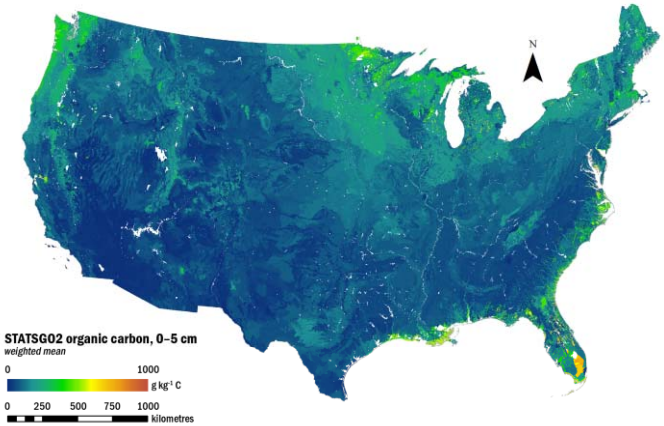


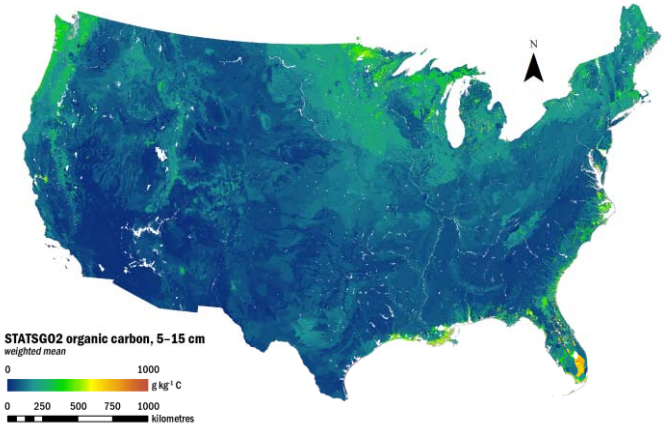
Recap: equal-area splines

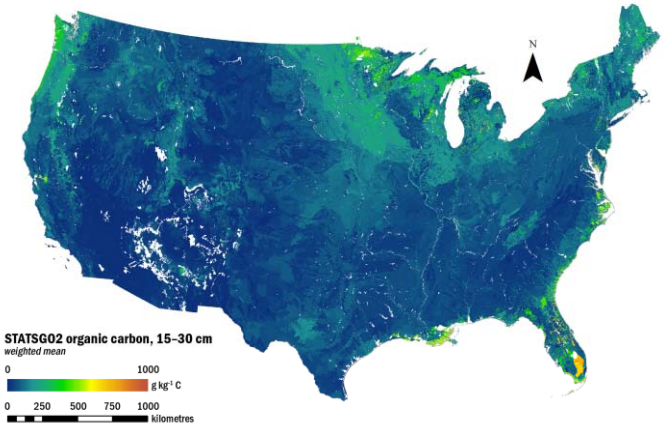


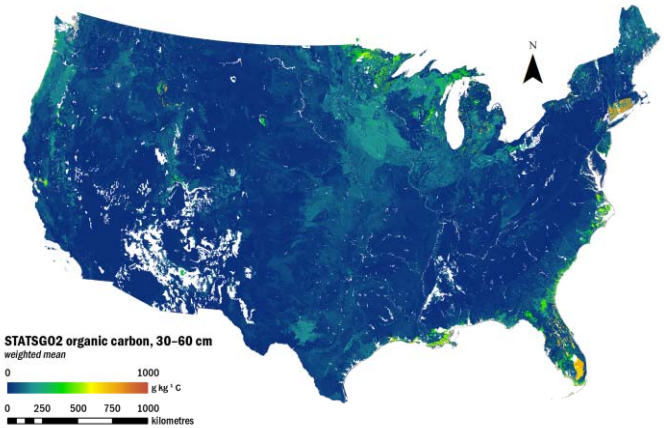
Property maps

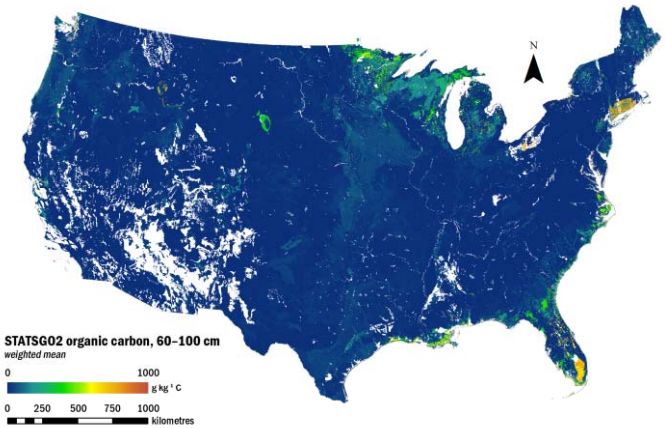
Organic carbon

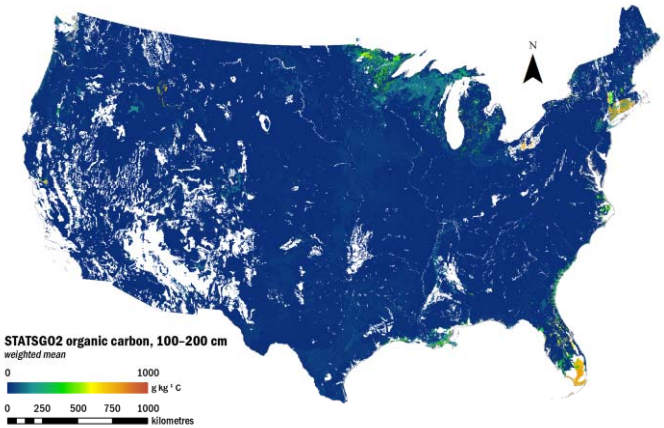




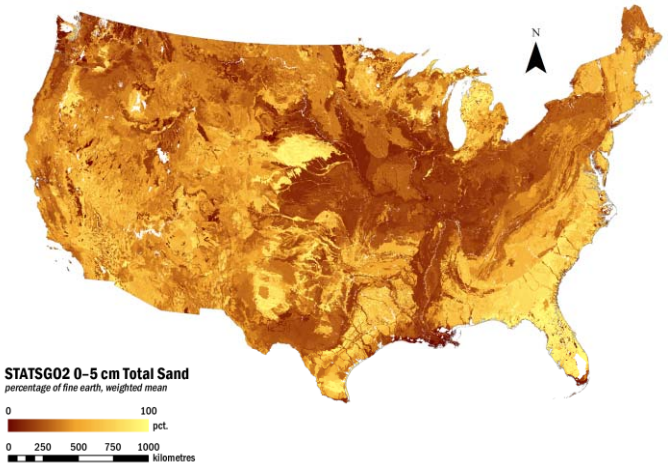


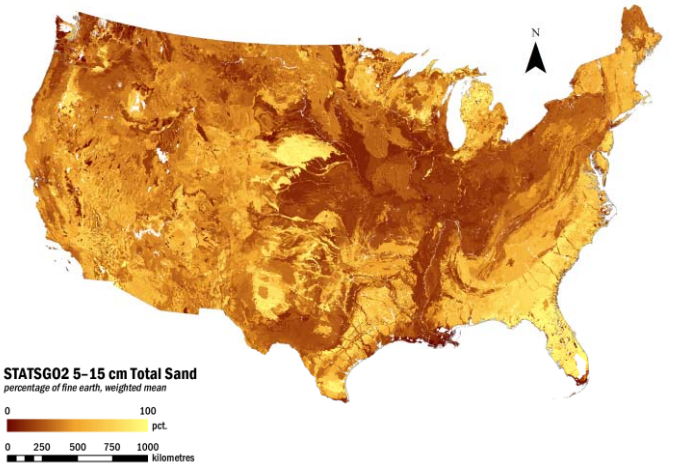


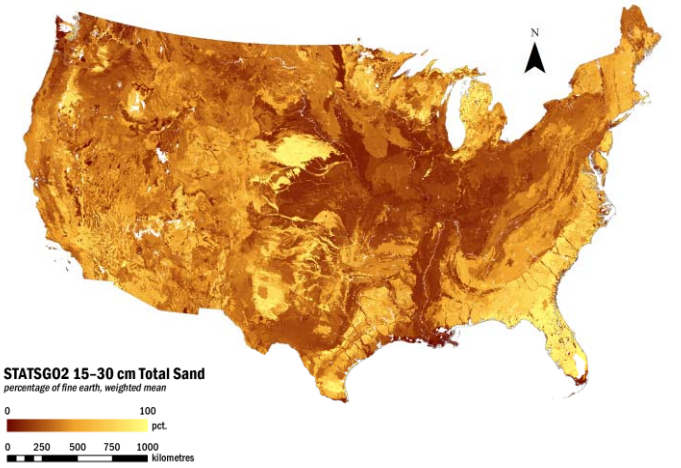


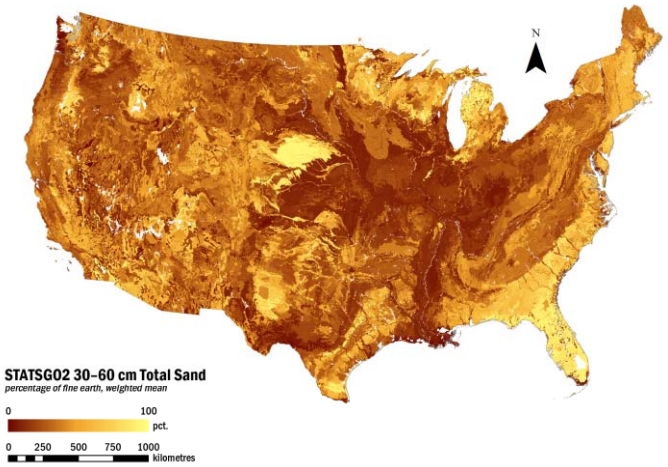


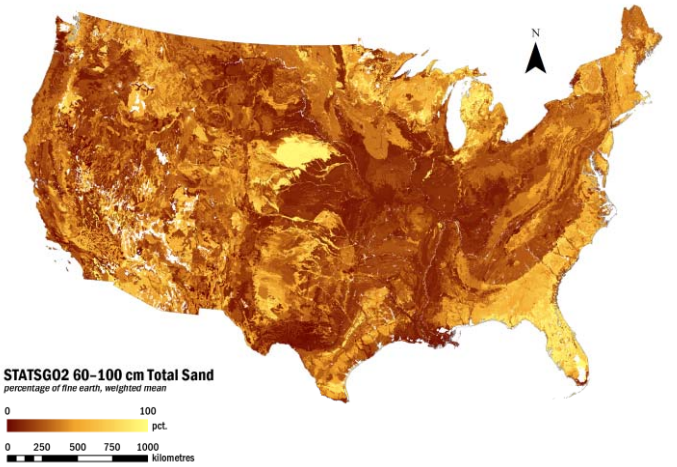
Sand

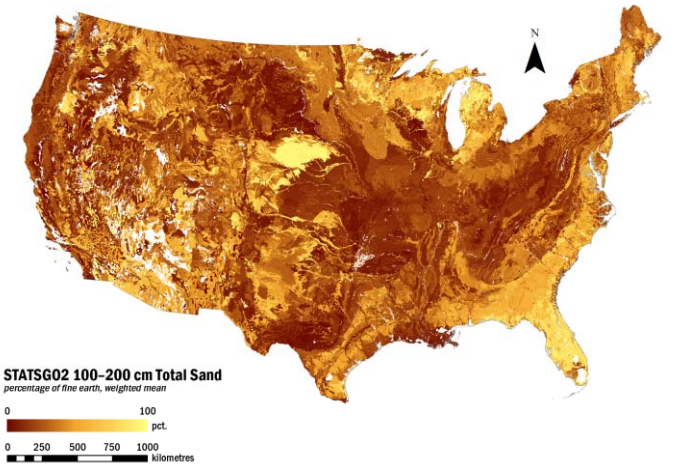

N
STATSGO2 0–5 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres


N
STATSGO2 5–15 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

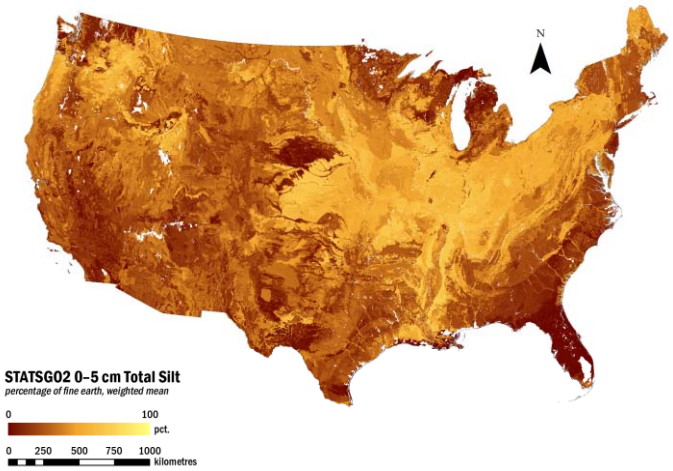

N
STATSGO2 15–30 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

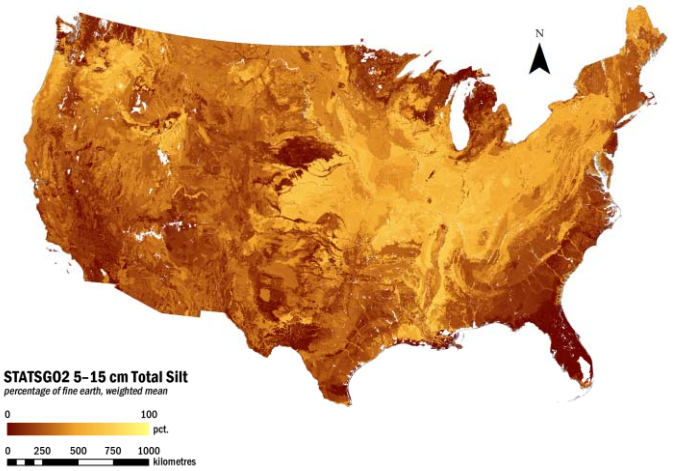

N
STATSGO2 30–60 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

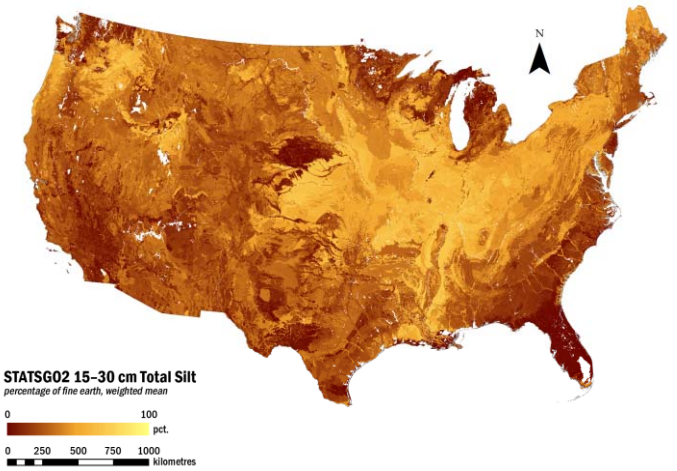

N
STATSGO2 60–100 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

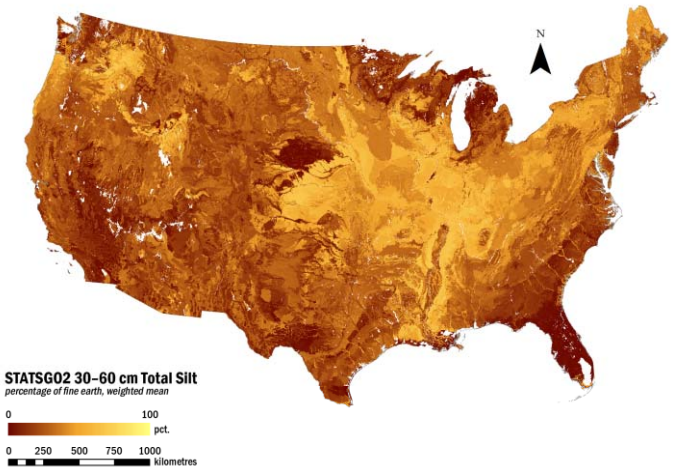

N
STATSGO2 100–200 cm Total Sand
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

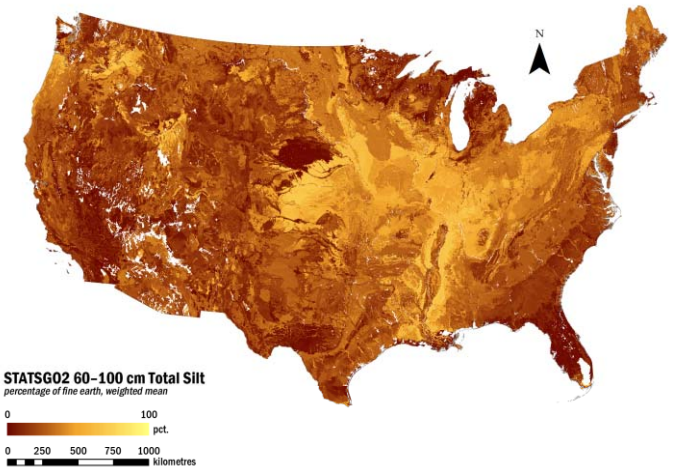
Silt

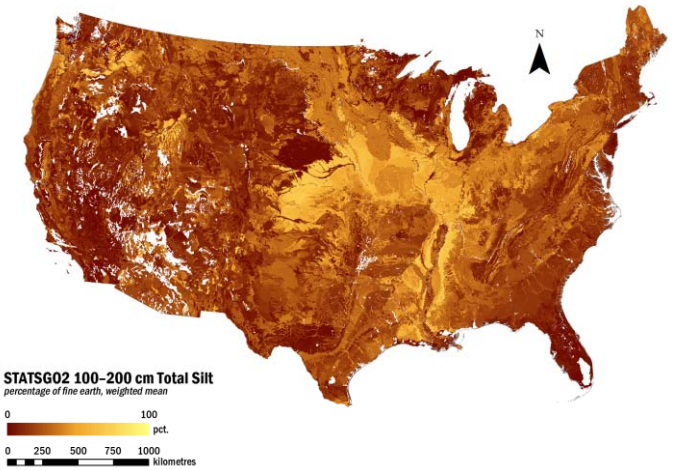

N
STATSGO2 0–5 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres


N
STATSGO2 5–15 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

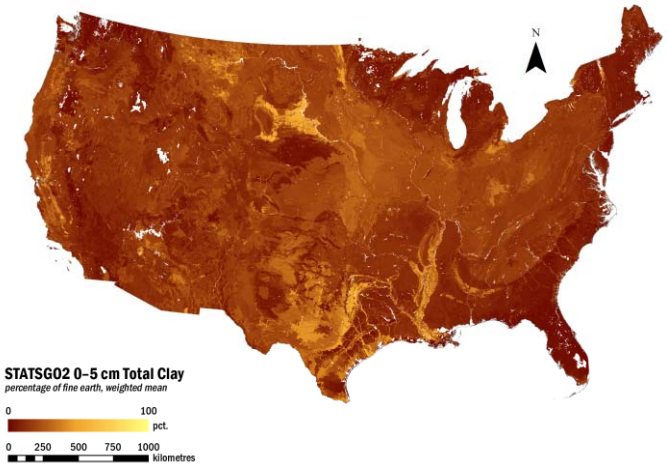

N
STATSGO2 15–30 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

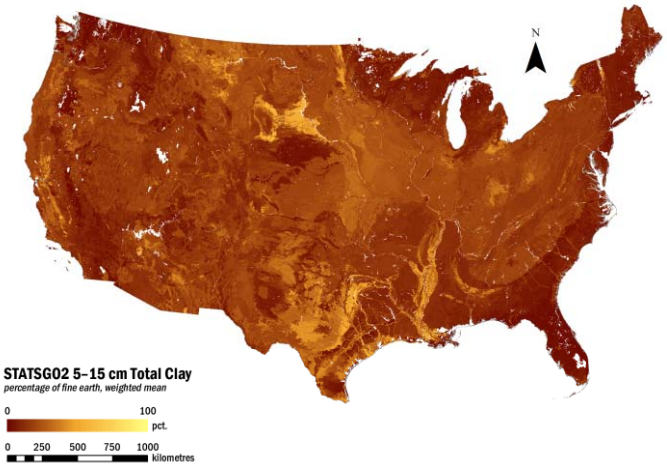

N
STATSGO2 30–60 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

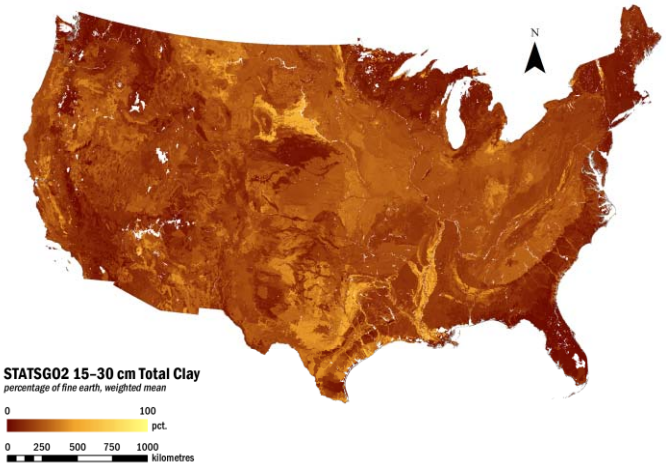

N
STATSGO2 60–100 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

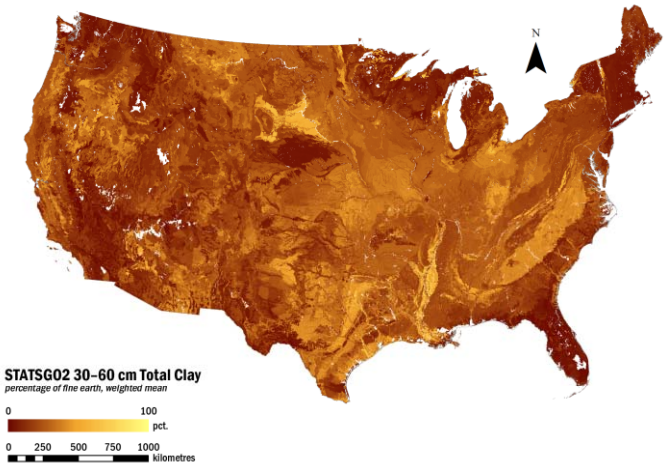

N
STATSGO2 100–200 cm Total Silt
percentage of fine earth, weighted mean
0
100
pct.
0 250 500 750 1000
kilometres

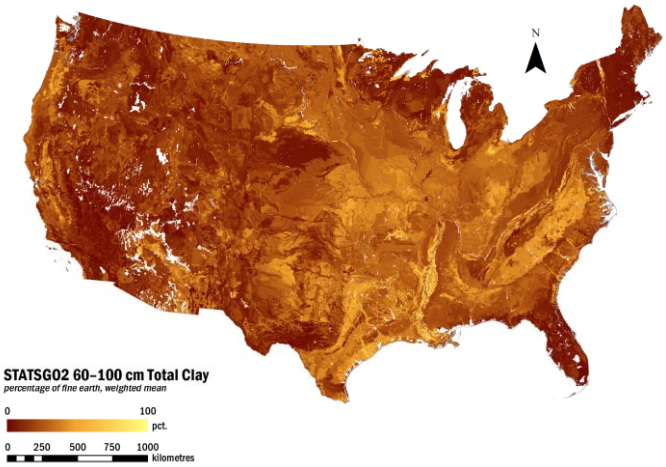
Clay

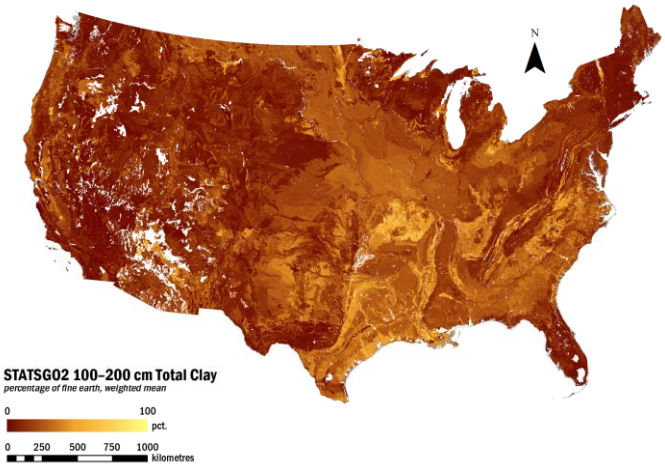

N
STATSGO2 0–5 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres


N
STATSGO2 5–15 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres

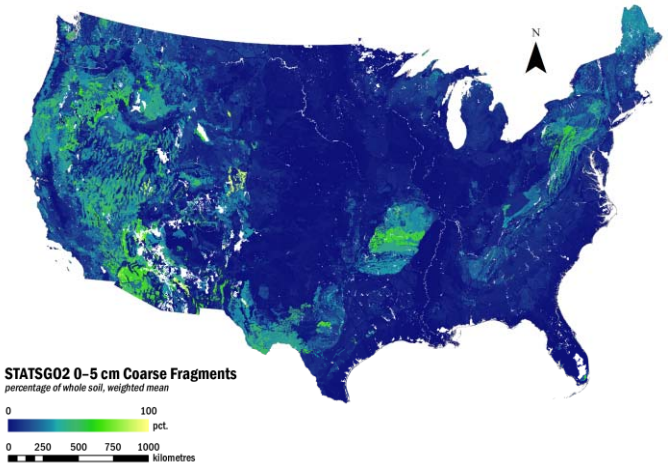

N
STATSGO2 15–30 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres

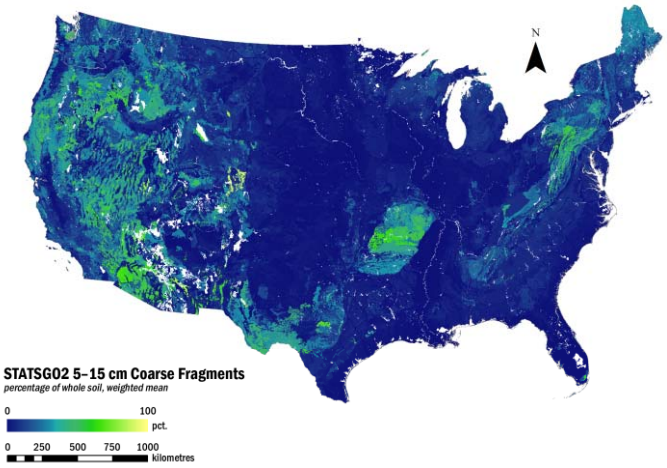

N
STATSGO2 30–60 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres

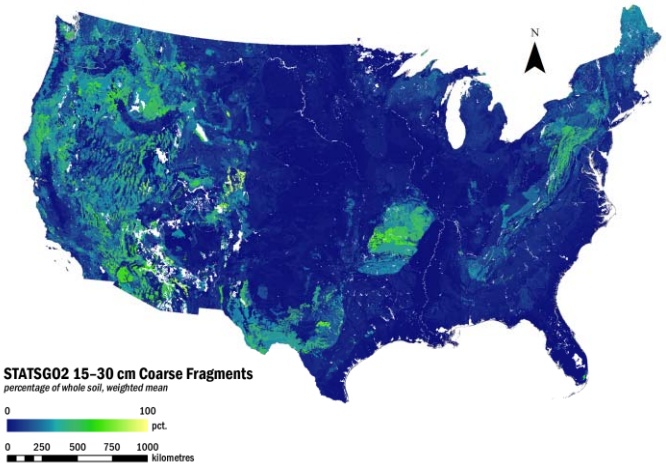

N
STATSGO2 60–100 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres

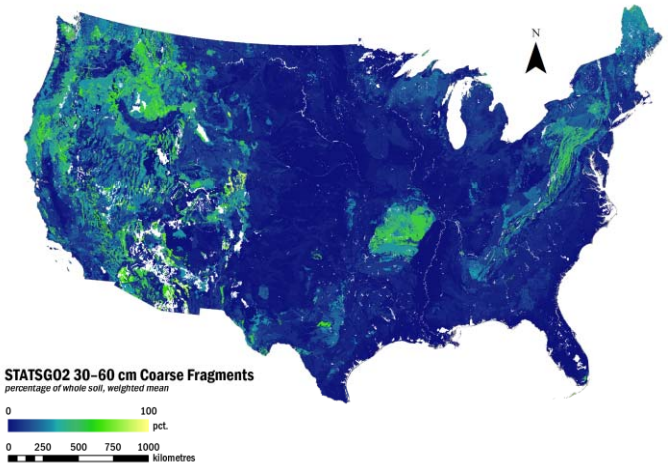

N
STATSGO2 100–200 cm Total Clay
percentage of fine earth, weighted mean
0
100
pct.
0
250
500
750
1000
kilometres

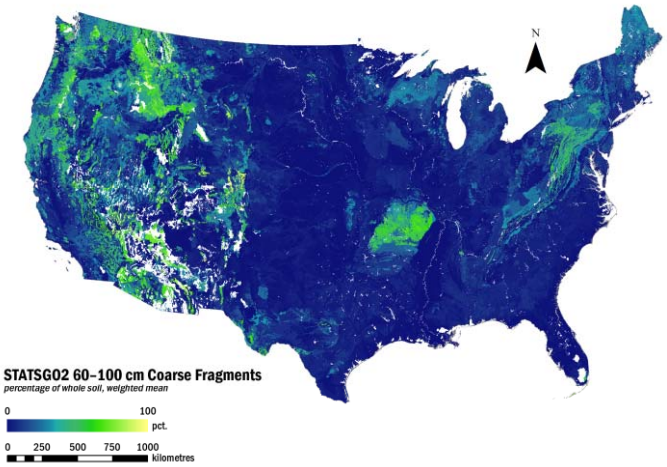
Coarse fragments

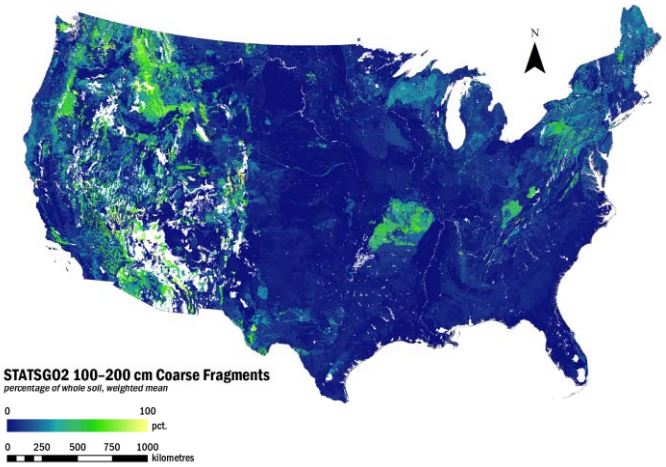












Uncertainty estimation

Background

- Soil property estimation:

$$\hat{S} = f(s, c, o, r, p, a, n) + \varepsilon$$

- Some degree of error is present in all our estimations
- Quantitative techniques enable us to estimate it
- Estimates of uncertainty are a required product for GlobalSoilMap.net

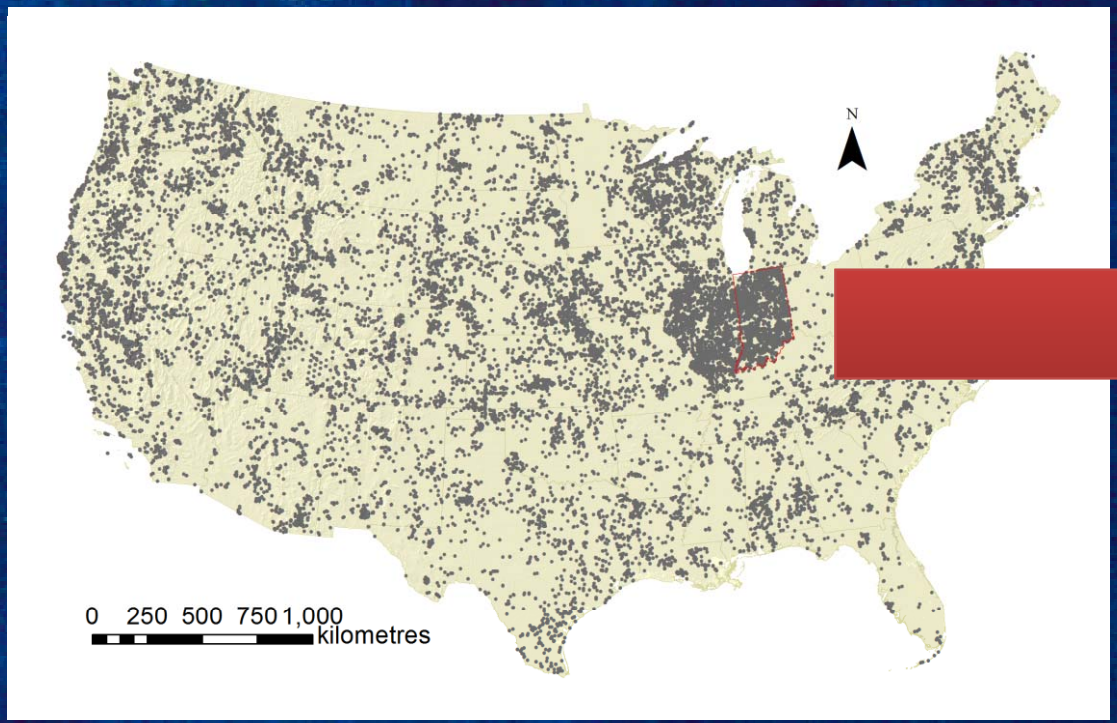
ε = true value − predicted value
s=soil, c=climate, o=biological factors, r=relief (topography), p=parent material, a=age (time), n=spatial position

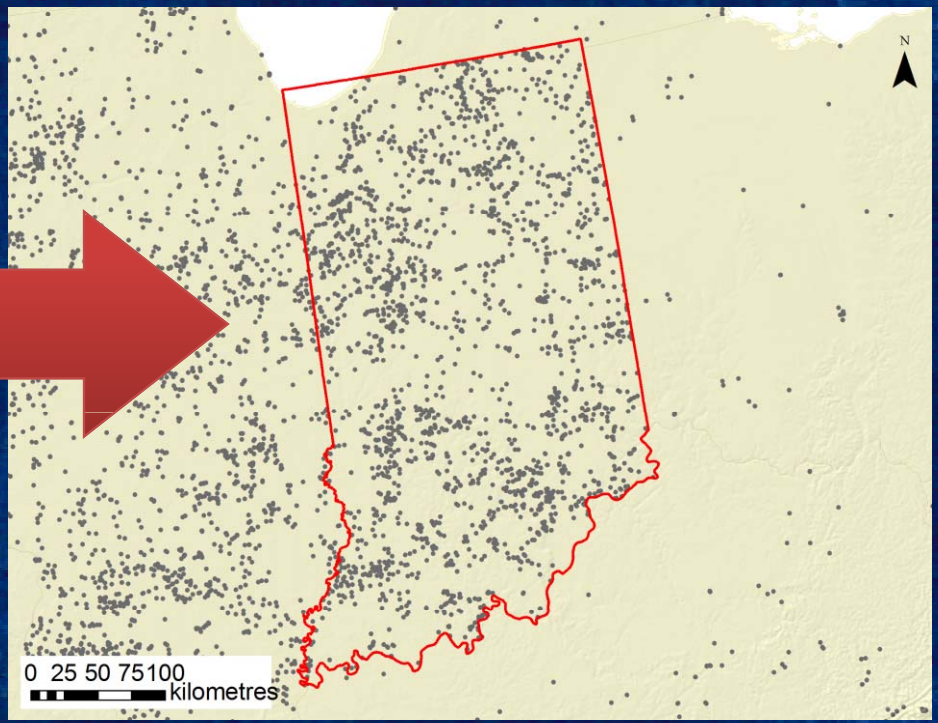
Method

1. Generate weighted means maps (~1:24,000 SSURGO)
2. Assemble georeferenced pedon (profile) data
3. Calculate absolute error between pedon and weighted mean for its map unit
4. Use regression tree to calibrate prediction model
5. Predict absolute error onto map grid

Georeferenced pedons

- Use Indiana as a case study
- Assembled 1711 georeferenced pedons with clay data
- Ran splines on pedons



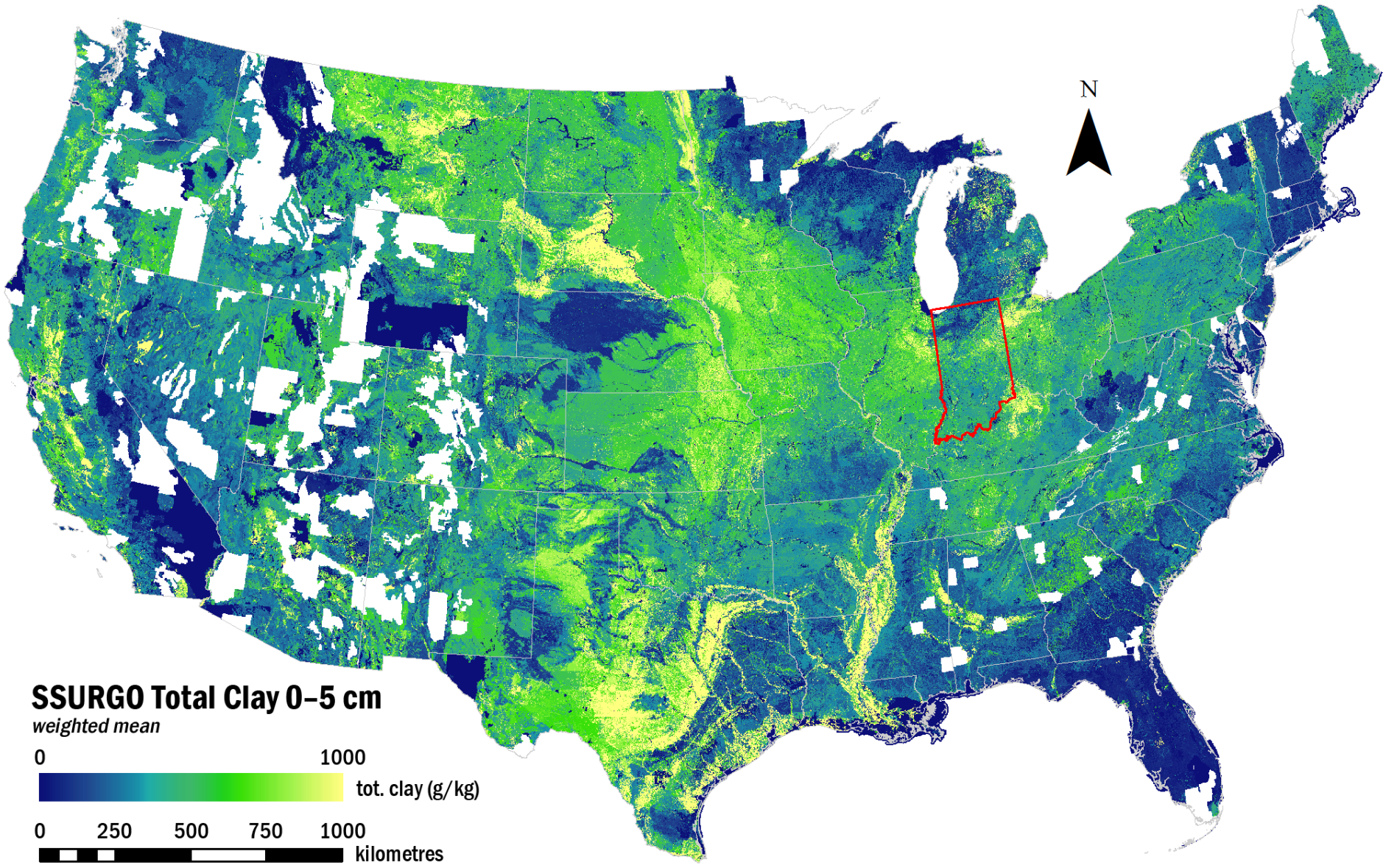


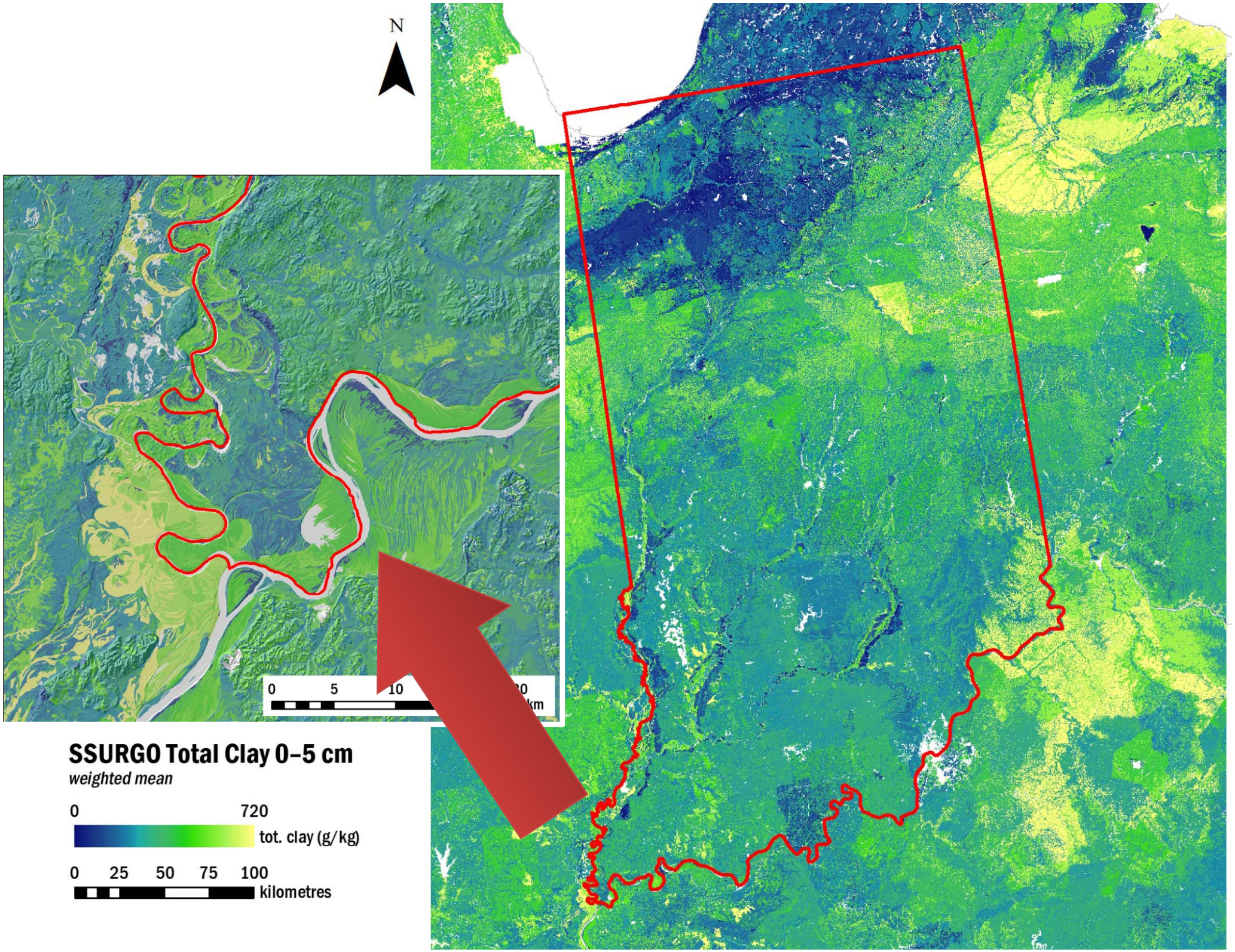
Calculate error

- Calculate the absolute error between each pedon's clay and the weighted mean for its map unit:

$$\varepsilon_{abs} = \sqrt{(weighted\ mean\ - pedon\ value)^2}$$

- Do this for each pedon, at each standard depth


N
SSURGO Total Clay 0–5 cm
weighted mean
0
1000
tot. clay (g/kg)
0
250
500
750
1000
kilometres


N
0
5
10
SSURGO Total Clay 0–5 cm
weighted mean
0
720
tot. clay (g/kg)
0
25
50
75
100
kilometres

Cubist

- Cubist builds regression trees
  - If-then rules with linear regression at the leaves

- Objective: build regression tree to estimate absolute error of total clay content
- Randomly selected 66% of the pedons (1129 samples) to build the tree
  - Remainder of pedons (582) for validation

Cubist covariates

- Terrain attributes
  - Elevation, aspect, convergence, slope, wetness, curvatures
  - Couldn't use accumulative attributes e.g. flow accumulation
- Land use
  - 2001 National Land Cover Database
- Geology
  - Age, bedrock type
- Spatial position (easting, northing)
- Soil survey area identifier, Major Land Resource Area id
- **Did not use:**
  - Soil map unit identifier (1:250,000 STATSGO2, 1:~24,000 SSURGO)

Ruleset

Rule 1: [686 cases, mean 40.2537041, range 0 to 309.0304, est err 27.7321072]
If survey_area in 9, 28, 30, 31, 32, 33, 34, 35, 37, 38, 39, 40, 41, 43, 49, 50, 51, 54, 58, 60, 61, 62, 64, 66, 67, 68, 70, 74, 76, 82, 85, 87, 89, 91, 92, 140, 142, 144, 145, 149, 150, 151, 152, 154, 155, 204, 218, 227, 228, 230, 232
then abs. error = 7.8850519 + 1.4 wetness + 63 slope

Rule 2: [406 cases, mean 64.4082947, range 0.3635917 to 439.403, est err 42.1156197]
If survey_area in 25, 26, 27, 29, 36, 42, 47, 53, 55, 59, 63, 65, 69, 72, 75, 78, 79, 83, 84, 86, 90, 141, 153
then abs. error = 50.0122528 + 187 slope - 2.2 aspect

Rule 3: [23 cases, mean 131.3504639, range 7.142974 to 476.0556, est err 93.3202209]
If elevation <= 264.973 & survey_area in 44, 57, 71, 77, 80, 81, 88, 143, 207, 231
then abs. error = 92.4117301

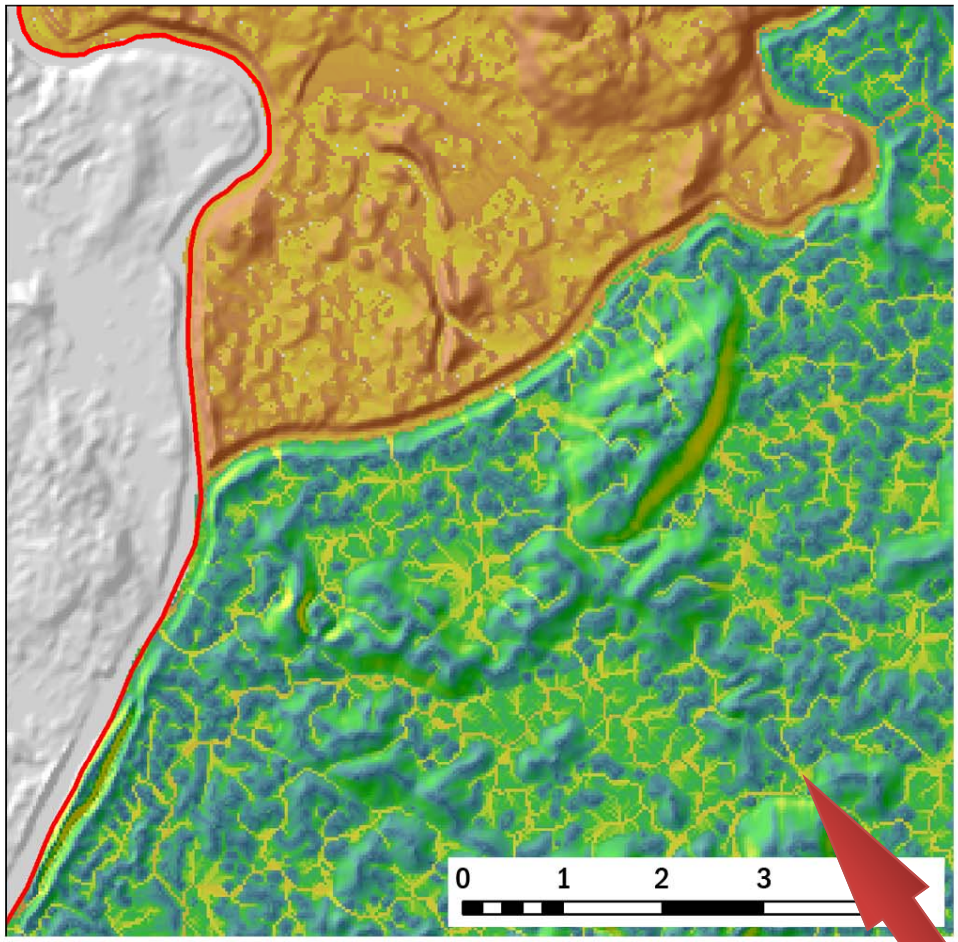
Rule 4: [14 cases, mean 197.5741272, range 2.626086 to 648.9125, est err 153.2537079]
If elevation > 264.973 & survey_area in 44, 57, 71, 77, 80, 81, 88, 143, 207, 231
then abs. error = -4233.7678031 + 15.97 elevation

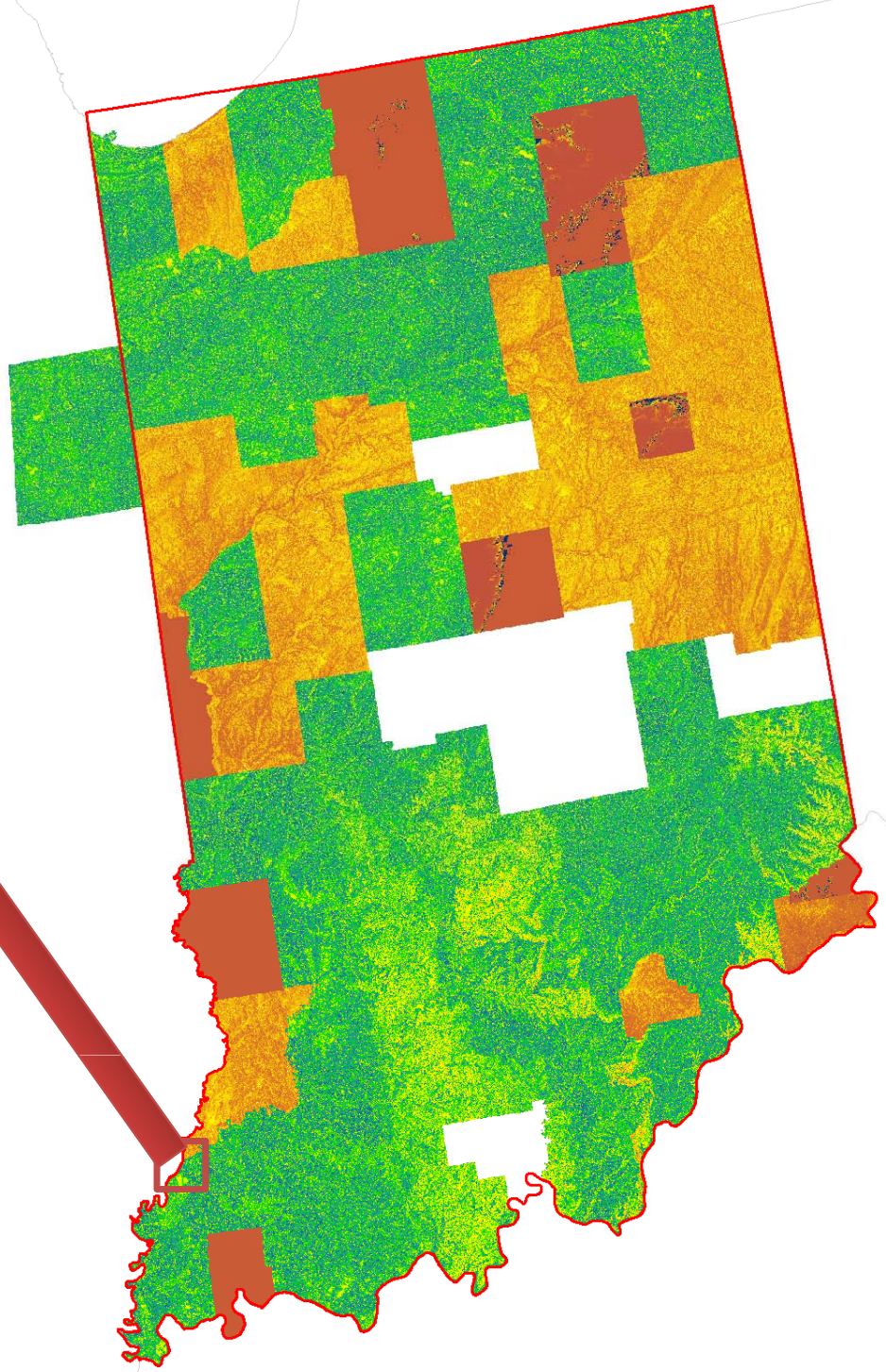
Evaluation on training data (1129 pedons):

Average |error|	34.8
Relative |error|	0.86
Correlation coefficient	0.48

Evaluation on test data (582 pedons):

Average |error|	40.8
Relative |error|	0.92
Correlation coefficient	0.49





SSURGO Total Clay 0–5 cm RMSE

0 1000 RMSE (g/kg tot. clay)

0 25 50 75 100 kilometres

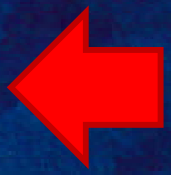
Ruleset

- This time, dropped soil survey area

Rule 1: [1129 cases, mean 53.4561577, range 0 to 838.7448, est err 42.2797699]
abs. error = 7.8850519 + 1.4 wetness + 63 slope

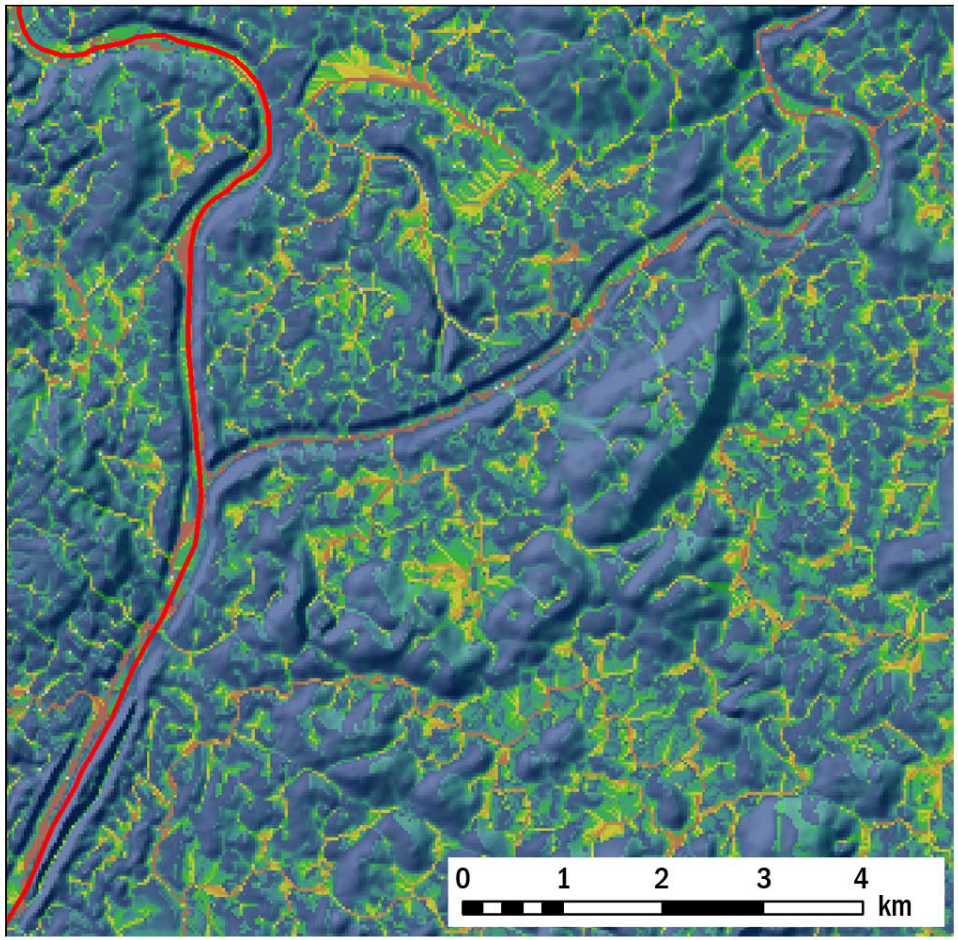
Evaluation on training data (1129 pedons):

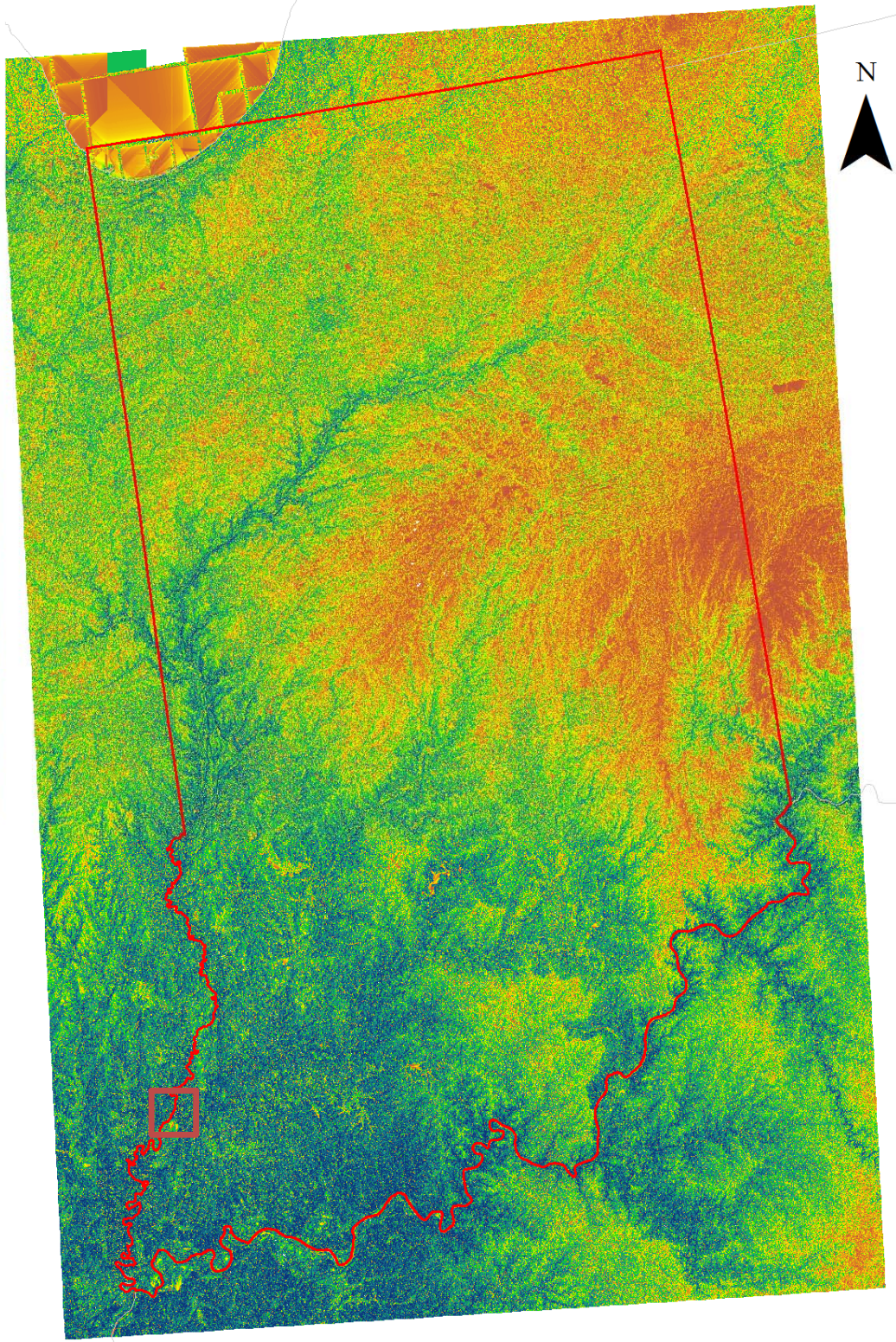
Average |error|	42.0
Relative |error|	0.99
Correlation coefficient	0.09



Evaluation on test data (582 pedons):

Average |error|	41.4
Relative |error|	1.00
Correlation coefficient	0.04





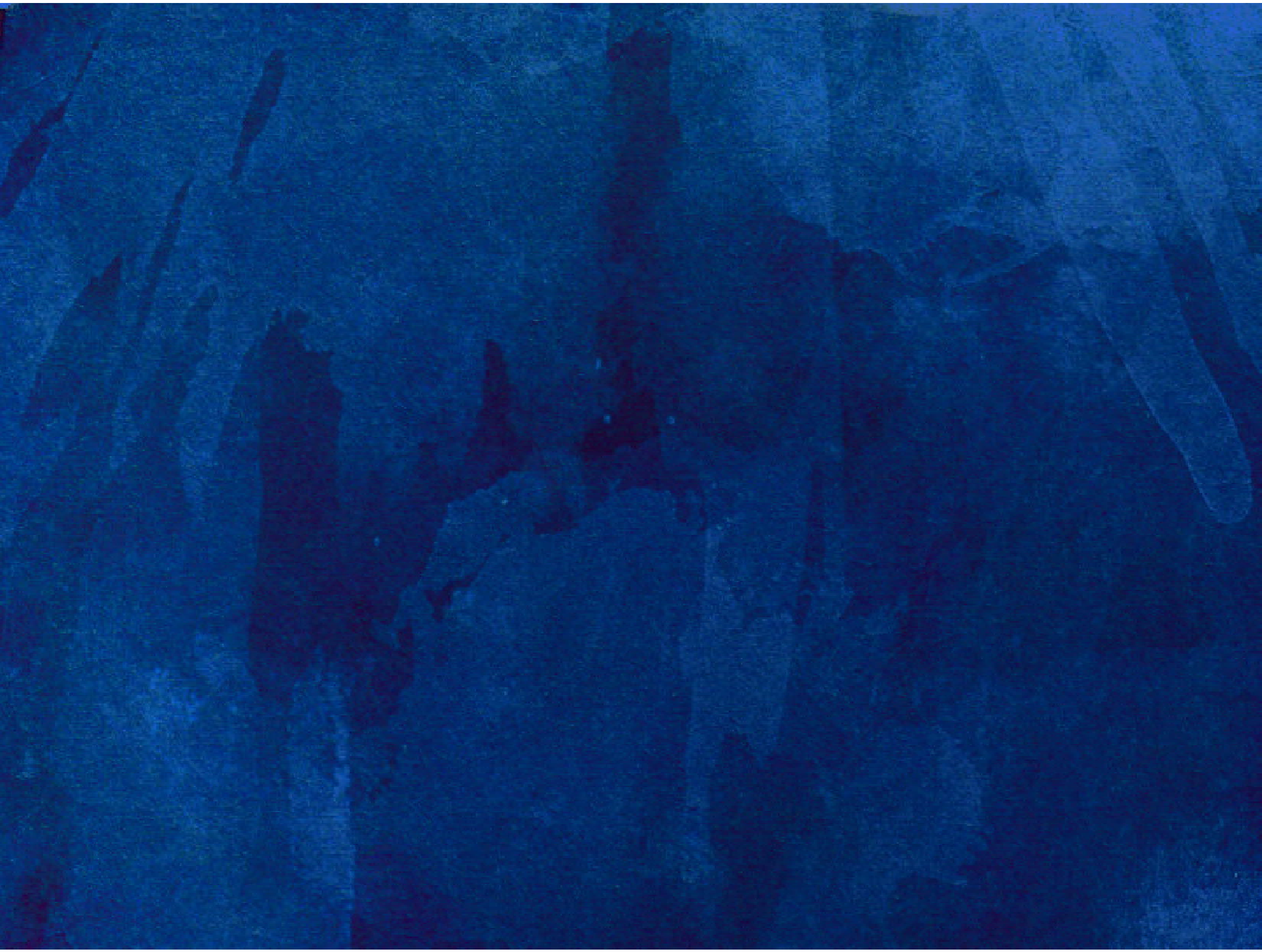
SSURGO Total Clay 0–5 cm RMSE

35 85
RMSE (g/kg tot. clay)

0 25 50 75 100
kilometres

Summary

- Polygon maps are often only exhaustive source of data for mapmaking
  - i.e. not enough point data to make maps using statistical/geostatistical approaches
  - Statistical approaches include their own measurement of uncertainty

Future work

- Compare absolute error approach with kriging standard error
- Idea is that if there is close correspondence then both are describing similar concepts
  - We can use one in place of the other