Digital Soil Mapping in Practice: Essex County, VT
Jessica Philippe
3.22.18
What are we doing here?

Digital Soil Mapping in Practice

The Essex County Story

- The Raster Soil Survey (RSS) of Essex County, Vermont was posted to the NRCS Geospatial Data Gateway in 2014
- First officially published RSS in the nation
- Initial survey that resulted in both an RSS and SSURGO product
- Completed by actual field soil scientists in what is now the 12-STJ MLRA Soil Survey Office in St. Johnsbury, Vermont
- 430,000 acres, four soil scientists, about five years
  - After about 3 years of software development (in partnership with Dr. Xun Shi of Dartmouth College) and concept validation
Who am I?

A Soil Scientist
And unofficial GIS specialist

- Working for NRCS for 14 years
- B.S. in Natural Resources Planning and M.S. in Soil and Water Science
- I started my soils career just as ArcSIE was coming into existence; I’ve never had to knock on any doors or use a stereoscope

But I promise, I’ve dug a LOT of holes.
Digital Soil Mapping

- The generation of geographically referenced soil databases based on quantitative relationships between spatially explicit environmental data and measurements made in the field and laboratory (McBratney et al., 2003)

- The spatial prediction of soil classes or properties from point data using a statistical algorithm
Digital Soil Mapping

Conventional soil mapping
- “Where is the boundary between two soils?”
- Focus on the marginal areas

Digital soil mapping
- Central concept is well defined
- Variation expressed across the landscape
Digital Soil Mapping

Knowledge-based DSM

Expert Knowledge
Knowledge-based DSM

**Purpose:** extract the soil-landscape model that lives in an experienced soil scientist’s head and turn it into something consistent and reproducible.
Knowledge-based DSM

ArcSIE - main tool in our DSM Toolbox

- Terrain Derivatives
- Soil Inference
- Post Processing (Harden, Sliver Removal, Diversity)*
- Vectorization*

*To meet SSURGO standard

ArcSIE was designed for field soil scientists to implement knowledge-based raster soil mapping
Essex County, Vermont

430,000 acres in NE Vermont
MLRA 143
Initial Survey
Essex County, Vermont

Typical Landscape

- 90% wooded
- Mainly glaciated hills and mountains with narrow valleys
- Elevations 1000-3000 ft.
- Dominantly loamy Wisconsin age till
- Forestry and recreation are main land uses
Essex County, Vermont

Poor access

Very little existing field data
Essex County, Vermont

The easy stuff:

- Shallow soils on upper hillslope positions and mountains in background
- Deeper soils formed in dense till on footslopes in the foreground

Major landforms in Essex County
Essex County, Vermont

- Outwash, alluvial, and lacustrine soils were mapped traditionally (digitally)
- Organic soils were heads-up digitized
Digital Soil Mapping

Basic Steps

- Data processing
- Landscape stratification
- Modeling
  - Knowledge extraction
  - Environmental layer selection
  - Soil inference
- Field verification
- Post processing
- Correlation and publication

Iterative
Data Processing

LiDAR or otherwise

The time/resources needed to prepare for the DSM process should not be underestimated.
Data Processing

LiDAR processing for Essex included:

- Reprojecting
- Mosaicking small tiles
- Clipping DEM to work area(s)
- Filling NoDATA areas
- Spike shaving
- Smoothing
- Resampling
Landscape Stratification

Stratify the mapping area by parent material/landform

- Provides basis for all subsequent soil mapping
- Different models or even approaches are used within each parent material type
- Done manually in Essex; automated techniques may be possible
Landscape Stratification

**Peer Review**

- Since the parent material mapping was done manually, an extensive peer review process was implemented
  - Step 1. Each soil scientist delineates PM based on LiDAR signatures and other available data
  - Step 2. Group peer review, identifying areas that need field checking
  - Step 3. Each soil scientist goes out to the field in their respective area and edits PM based on what they find
  - Step 4. Soil scientists make edits to each other’s PM lines
  - Step 5. Group peer review and final edits
  - Step 6. Certification
Landscape Stratification

Field work is crucial

- Once relationships are well understood, more time can be spent in other areas
- Backs up decisions during peer review process
Elevation 200–600m is typical for soil A.

As elevation deviates from this range, the soil’s similarity to type A gradually decreases.
Modeling

Knowledge Represented as a Rule

Remember – for the raster mapping process we’re trying to define the central concept of the target classes.
Modeling

Essex County lodgement till

central concepts

Cabot (poorly drained)
Colonel (somewhat poorly drained)
Dixfield (moderately well drained)
Modeling

Traditional soil components vs. modeled classes from ArcSIE

In the model, the soil formative environment of certain soil series are targeted as the central concept, but the actual soil class that ends up being mapped generally spans more than one series concept (born out through field verification) and with proper documentation better represents the soil continuum.
Modeling

Knowledge Extraction

Document everything the soil scientist knows about the local soil formative environment

<table>
<thead>
<tr>
<th>Target soil</th>
<th>Major landform</th>
<th>Hillslope position</th>
<th>Elevation</th>
<th>Slope gradient</th>
<th>Slope aspect</th>
<th>Vegetation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabot</td>
<td>Mountain, hill, till plain, swamp</td>
<td>Most common on footslopes and toeslopes. In drainageways on backslopes. In depressions on flats and broad ridges.</td>
<td>Mostly below 750m. Never above 900m.</td>
<td>Mostly 0 to 15 percent.</td>
<td>Mostly above 35 percent.</td>
<td>Not important.</td>
</tr>
</tbody>
</table>

**Additional notes:** Cabot soils occur in all till areas. They grade to Histic Humaquepts and Histosols in very wet locations, and Wilmington and Colonel soils (Spodosols) in drier locations.
Environmental layer selection

Select digital layers that best represent the documented soil-landscape model. In the case of Essex County, slope (30m neighborhood) and wetness index were critical.

* Both generated from LiDAR data using ArcSIE
Modeling

Environmental layer selection

Many possible layers were evaluated but not used for various reasons.

- Landforms (overly simplistic and inconsistent in Essex County landscape)
- 25 cell neighborhood relative position (did not consistently add value to the model)
Actual wetness index rule for Cabot soil: pixels with a wetness index value of 6.3 and above are assigned a full membership in the Cabot class. At a wetness index value of 4.8, membership in the Cabot class is 0.5.
Output from ArcSIE is in the form of a fuzzy raster map. Fuzzy membership values of each pixel represent the similarity of that pixel location to the typical soil formative environment defined by the curves.

Fuzzy inference results for Cabot. Darker colors represent higher membership.
Fuzzy results from each soil class are “hardened” together so each pixel in the survey is assigned to the soil class with the highest fuzzy membership.
Modeling

Field Verification
By the end of the survey, 9041 georeferenced points were collected, over 400 of which were full or almost-full pedon descriptions.
Modeling

Results

• Slope phase (from 60m slope neighborhood, reclassified to standard slope breaks) is integrated with hardened results.

• In the legend, the first digit represents soil class from the model; the second digit is slope phase.
Raster Product

**Raster class concept validation**

- 187 well-documented field samples were used to validate raster class concepts
- Ranges AND RVs come from this aggregated field data
Raster Product

NASIS Population

- A raster soil component is spatially represented by a single pixel; a raster soil map unit comprises the entire population of each raster soil component.
- Component tabular data is correlated to each pixel value.
- The data mapunit contains only one component, populated based on the aggregated field data.
Raster Product

Delivery

- Distribution format is File Geodatabase modeled after gSSURGO

Raster dataset

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>VT009_raster_subset</td>
<td>File Geodatabase Raster Dataset</td>
</tr>
</tbody>
</table>

SSURGO export tables

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>chashto</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chconsistency</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chdesignsuffix</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chfrags</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chhorizon</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chpores</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chservlet</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chructgrip</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chtext</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chtexture</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chtexturegrip</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chtexturenod</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>chuuniform</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocanopycover</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocropyld</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocropfeatures</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocrocoext</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocrocoext</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocrocoext</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocrocoext</td>
<td>File Geodatabase Table</td>
</tr>
<tr>
<td>ccocrocoext</td>
<td>File Geodatabase Table</td>
</tr>
</tbody>
</table>

Pre-built relationship classes
Interpretation component
“hydric rating” = “Yes”
Raster Product

Raster Soil Survey

- RSS is available for only a subset of the Essex County Soil Survey Area (lodgment till areas)
- Raster components were used to create SSURGO for the same area
- Multiple mapping approaches were combined to form SSURGO for the entire county
  - About 85% DSM, 15% traditional (but digital)
To eventually meet SSURGO standards, some of the raster classes get combined to form logical map units.
SSURGO Product

Post Processing

- “Sliver removal” helps achieve a target minimum size delineation (5 acres in Essex County)
SSURGO Product

SSURGO Processing

- A significant amount of post-processing is required to go from raster to vector
- After vectorizing, it took two expert soil scientists six months (so one FTE) to manually edit the vectors to meet cartographic expectations
- Must be done at the very end of the process (so there are no draft vector maps until the entire project is complete)

...So, yes, vectorizing is possible. But the Essex County crew recommends spending the needed resources elsewhere.
Essex County RSS

To find it:

- Head to the NRCS GeoSpatial Data Gateway (https://gdg.sc.egov.usda.gov/), search for Essex County, Vermont, and scroll down to Soils data.

There's also a Fact Sheet and FAQs located on the NRCS Digital Soil Mapping website, located here