

A Geoprocessing Framework using Python and ESRI SDE

Howard Butler

**Center for Survey Statistics &
Methodology**

Iowa State University

Beginning

Middle

End

National Resources Inventory



- National longitudinal survey of multiple natural resources
- Estimates of status and change over time on nonfederal lands
 - Land cover/use changes
 - Soil erosion trends
 - Urbanization and loss of prime farmland
 - Wetland dynamics
- <http://www.nrcs.usda.gov/technical/NRI/>

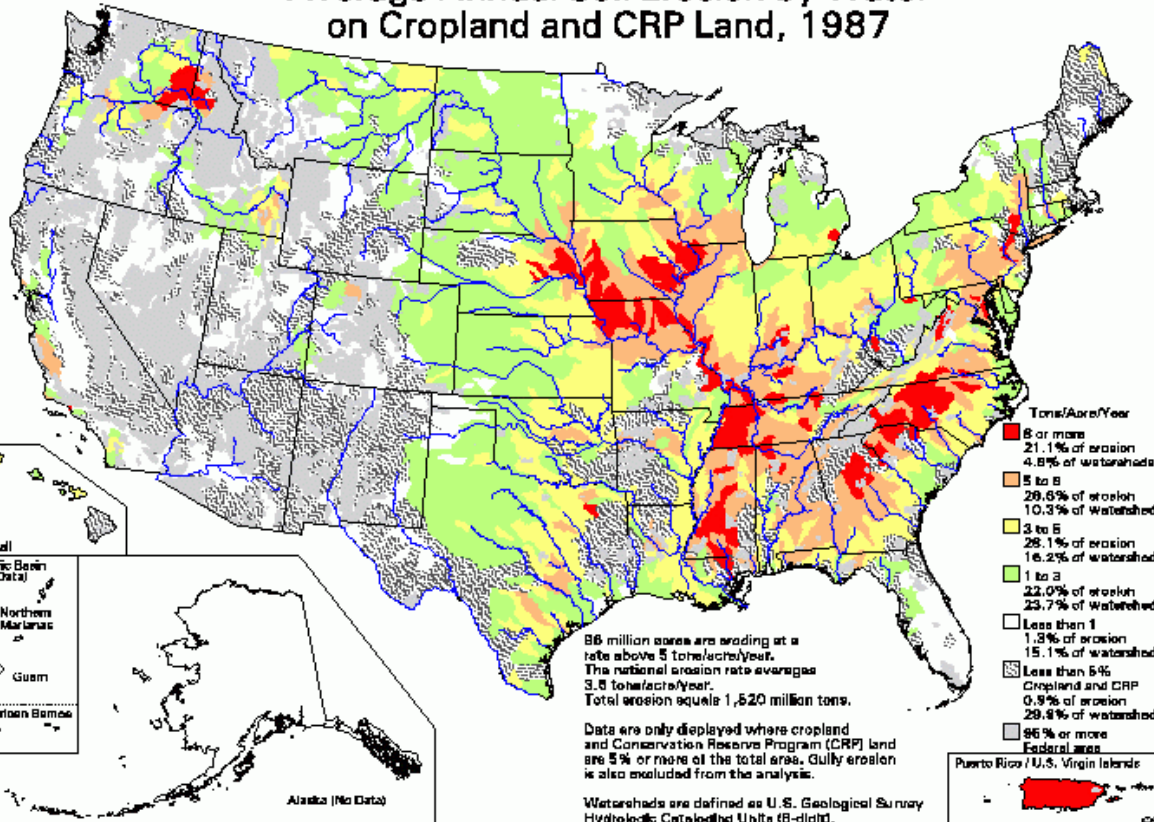
Beginning

Middle

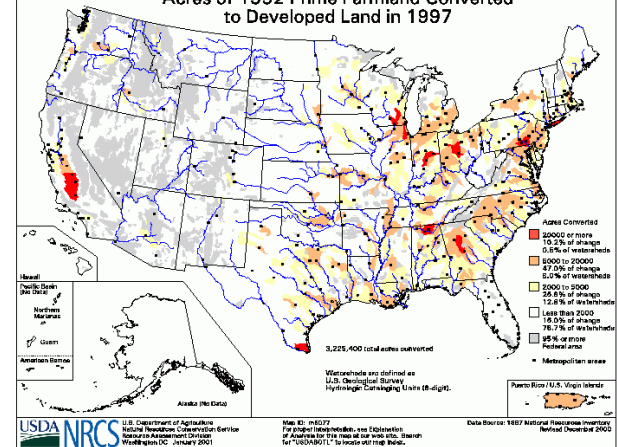
End

National Resources Inventory

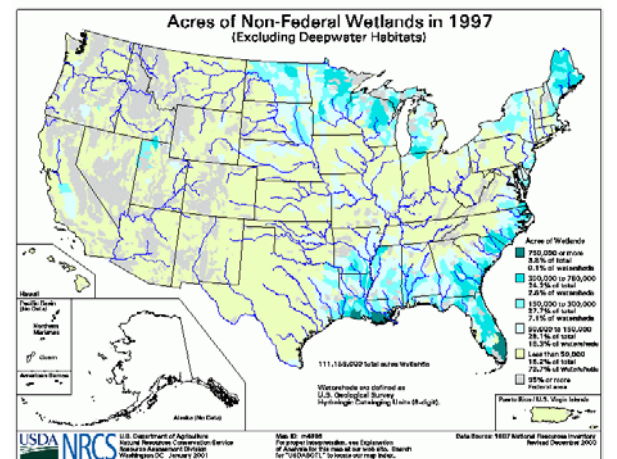
Average Annual Soil Erosion by Water on Cropland and CRP Land, 1987



Acres of 1992 Prime Farmland Converted to Developed Land in 1997



Acres of Non-Federal Wetlands in 1997 (Excluding Deepwater Habitats)



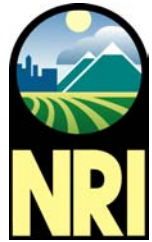
Tables

Table 8 - Changes in land cover/use between 1992 and 1997 (data per 1,000 acres)

Land cover/use in 1992	Land cover/use in 1997								1992 total
	Cropland	CRP land	Pastureland	Rangeland	Forest land	Other rural land	Developed land	Water areas & federal land	
Cropland	362,606.3	2,062.6	9,210.5	1,555.5	1,937.1	1,722.2	2,902.1	318.9	382,315.2
CRP land	2,250.8	30,464.9	796.6	297.2	184.4	40.2	7.7	0.3	34,042.1
Pastureland	8,523.5	96.6	106,543.2	1,562.3	6,272.3	897.1	1,979.8	172.7	126,047.5
Rangeland	1,977.8	21.1	696.4	400,770.5	1,600.8	779.0	1,283.2	250.9	407,379.7
Forest land	759.9	22.9	1,887.7	1,170.0	395,273.0	950.2	4,771.1	372.2	405,207.0
Other rural land	707.7	27.9	753.5	474.0	1,447.4	46,744.4	448.5	22.4	50,625.8
Developed land	27.9	0.0	24.0	53.7	76.0	2.8	86,850.3	0.0	87,034.7
Water areas and federal land	144.0	0.0	80.0	94.0	164.2	5.6	9.0	450,980.9	451,477.7
1997 total	376,997.9	32,696.0	119,991.9	405,977.2	406,955.2	51,141.5	98,251.7	452,118.3	1,944,129.7

1992 land cover/use totals are listed in the right hand vertical column, titled "1992 total". 1997 land cover/use totals are listed in the bottom horizontal row, titled "1997 total". The number at the intersection of rows and columns with the same land cover/use designation represents acres that did not change from 1992 to 1997. Reading to the right or left of this number are the acres that were lost to another cover/use by 1997. Reading up or down from this number are the acres that were gained from another cover/use by 1997.

National Resources Inventory



- > 300,000 area segments
 - 160 acres each
 - 70,000 samples / year
 - Low altitude photography

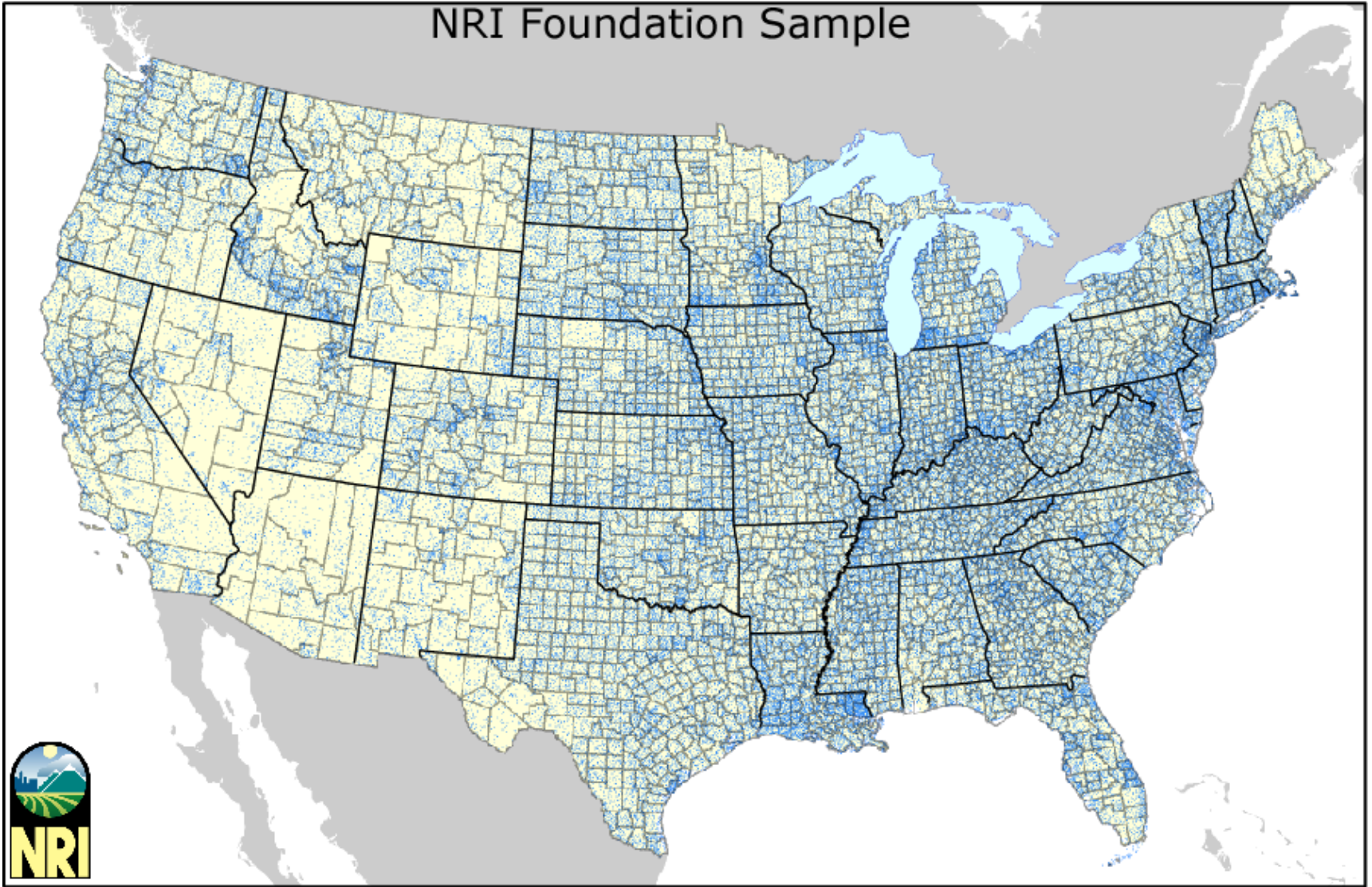
Beginning

Middle

End

National Resources Inventory

NRI Foundation Sample



National Resources Inventory



- 1/3 digital -- 2/3 analog
- > 18 gb of support data
 - Federal land
 - Water
 - Sample location layers
 - Support imagery

Beginning

Middle

End

History

- ESRI Arc/Info coverage format
- Migrating toward ESRI SDE
- Oracle database for tabular data
- Changes/edits tracked with paper

Beginning

Middle

End

Impetus

- NRI is a panel (time series) survey
 - Tracks spatial data changes through time
- Federal land and water layers must be updated annually
- File system-based solution is messy and inflexible for time-series geospatial data

Beginning

Middle

End

Impetus

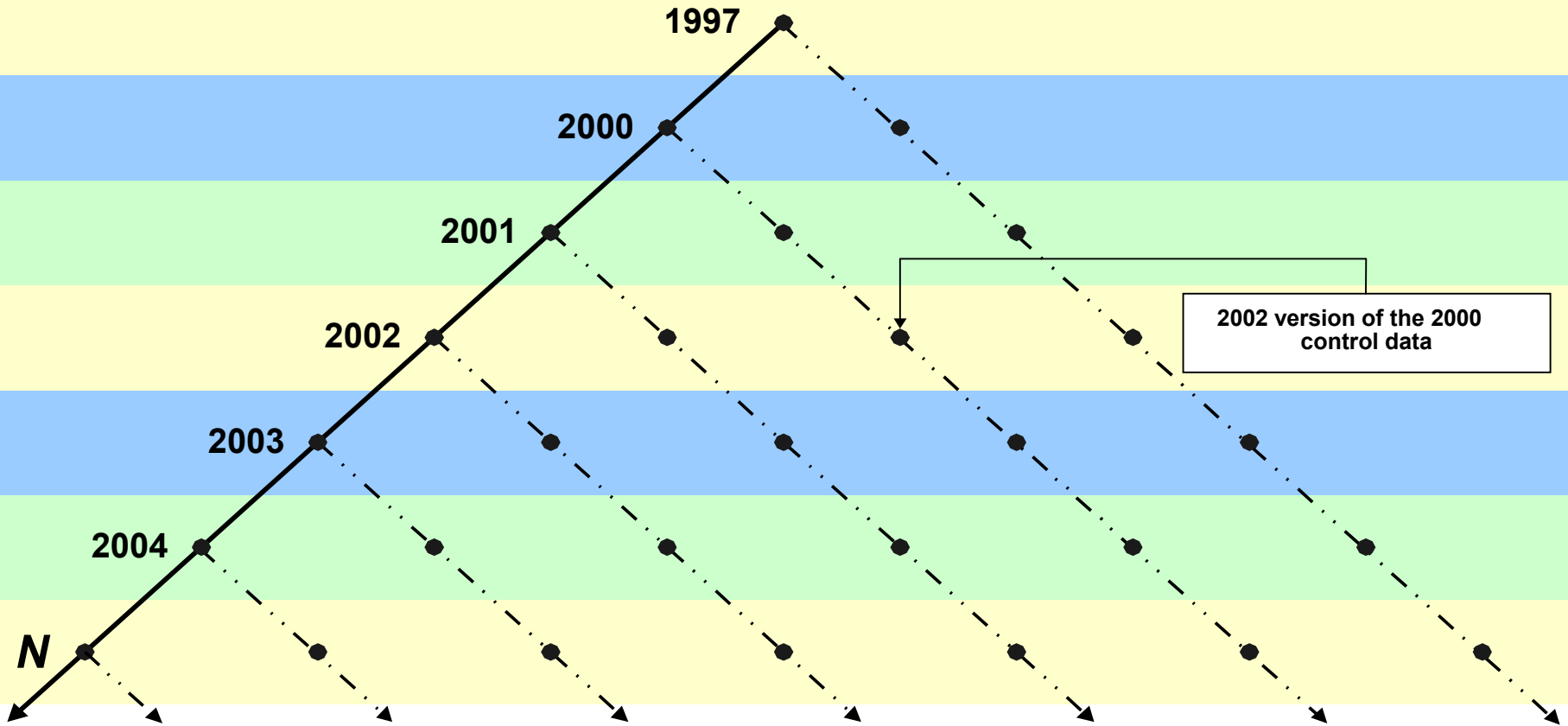
- Must be able to interoperate with existing Oracle data
- Campus licensing agreement makes ESRI software a cost-effective solution

Beginning

Middle

End

Version tracking of survey data



Why not PostGIS?

- PostGIS currently limited to PostgreSQL
- Inertia of existing data in ESRI formats is high
- Campus agreement
- No support for versions or edit states

Beginning

Middle

End

Why Python?

- Functionality required by NRI only exists in the C API
- C API is complex, cryptic, and obscures the task at hand

Beginning

Middle

End

Why Python? (cont)

- Our GIS team has limited C experience
 - dynamic/scripting languages
 - AML, Avenue, VB
- Easy transition to Python from these languages

Beginning

Middle

End

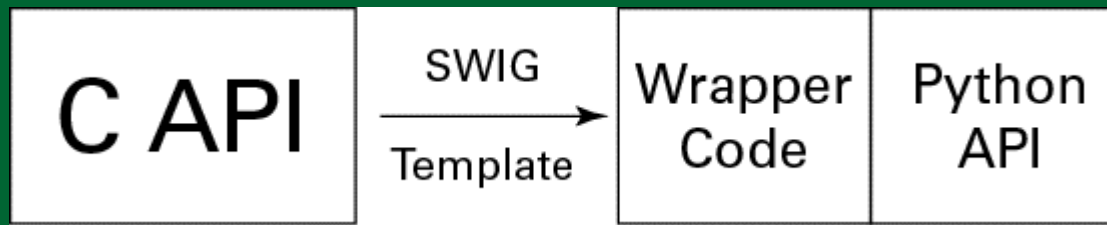
Advantages of Python

- Elementary data structures included
 - Lists, sets, strings, and hashes
 - Debugged and optimized by many folks
- Interpreted
 - Speeds up the wash-rinse-repeat cycle
 - Run-time type checking (but not compile time)
- Rapid development
 - Wield more power per line of code

Beginning

Middle

End



- SWIG to the rescue
- ~ 3k lines of template -> 13k lines of wrapper
- A Python SDE API built upon the wrapper code

Beginning

Middle

End

Why not just the wrapper?

- ESRI changes APIs frequently – changes at the C level don't have to be reflected at the Python level
- Allows for coherent exception handling and error reporting
- Allows for add-ons like Python DB-API 2.0 and pyshapelib support to be developed in Python

Beginning

Middle

End

What it currently supports

- Most of the shape handling functions
- Versioning and states
- Projection operations
- DB query operations

Beginning

Middle

End

Platforms

- Python 2.2 and greater
- Windows 2k/XP
- Solaris 9
- Linux 2.4?

Beginning

Middle

End

Sticky parts

- Not thread safe
- Leaks memory like the Titanic
- Assumes GIS programming experience
- Docs are very thin
- Still in flux – not quite beta but a bit better than alpha

Beginning

Middle

End

Future developments

- Would like clean integration with pyshapelib
- Python DB API 2.0 support
- Thread safety and memory issues cleanup

Beginning

Middle

End

Conclusion

- Cost of learning the C API is only incurred once
- Clearer code that addresses the problem instead of how to store the data, how to open a connection, how to loop through an array.....
- High performance but still easily programmed interface

Beginning

Middle

End

Where to download

hobu.stat.iastate.edu/pysde

Beginning

Middle

End