Development of a Strategic Spatial Sampling Design for the State of Florida

D. Sarkhot, S. Grunwald, N.B. Comerford, W.G. Harris, and G.L. Bruland

Funding source:
USDA-NRI Soil Processes
USDA-CSREES#69579

Need for Landscape Scale Studies

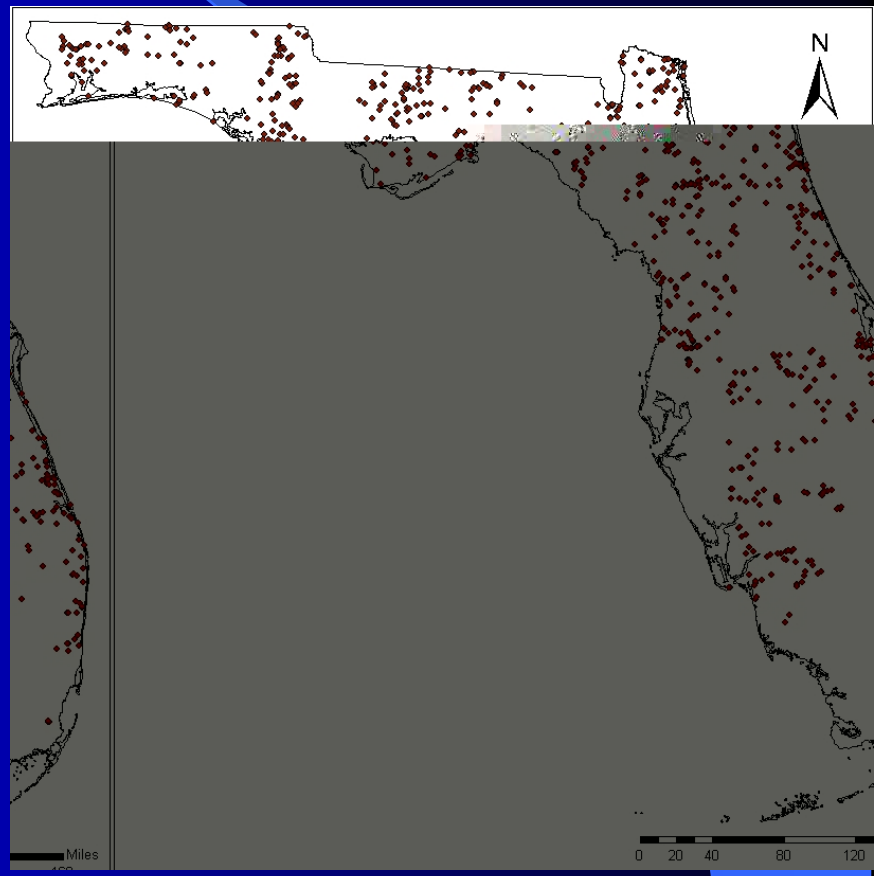
- Development of landscape scale estimates of soil properties
  - Describe the spatial variability across landscapes
  - Decision making
  - Exploiting economical opportunities such as carbon trading
- Development of maps for research, extension and educational purposes

Need for Appropriate Sampling Design

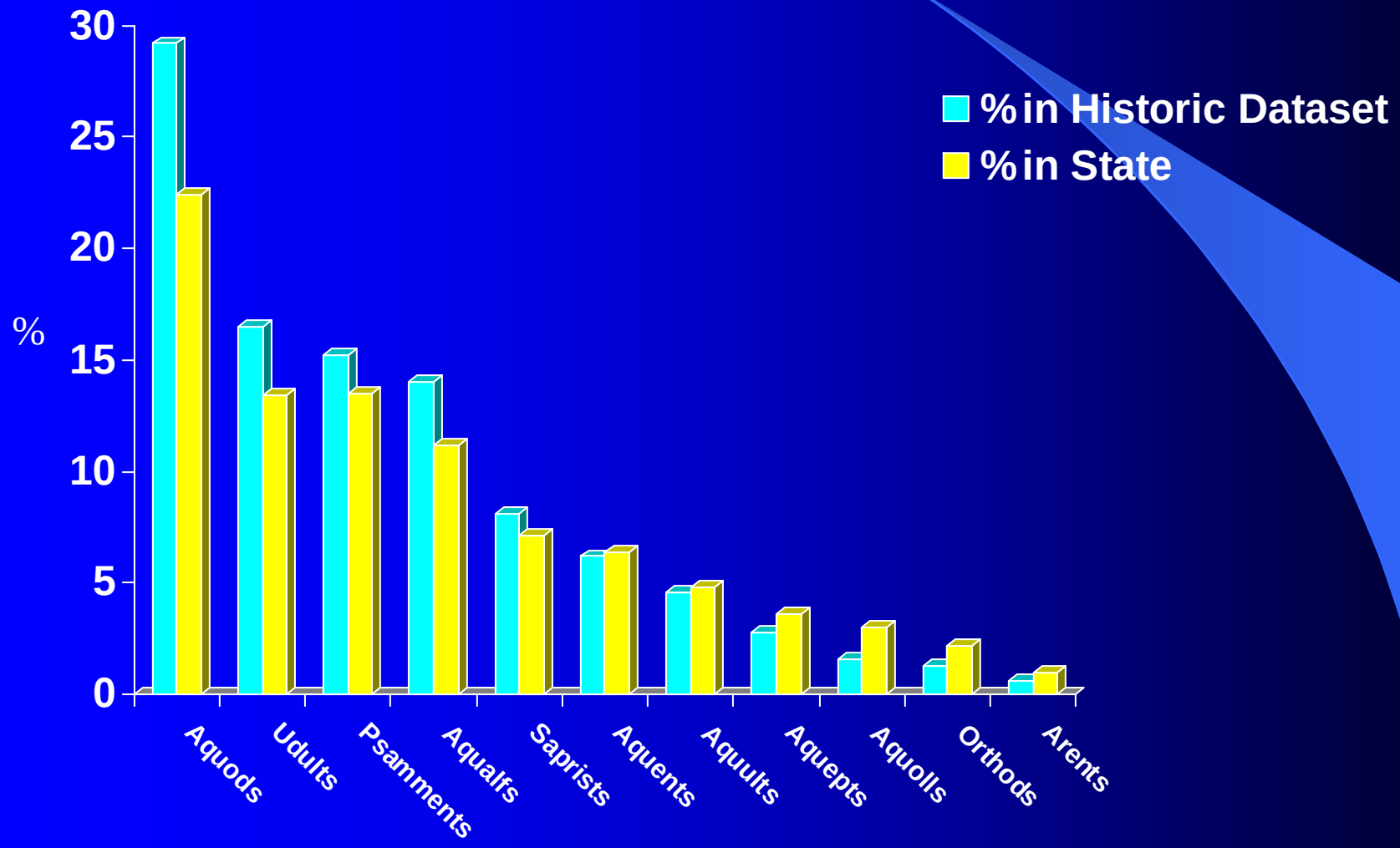
- To build on existing soil knowledge base e.g. NRCS Soil Data Mart
- To ensure adequate coverage of environmental landscape factors e.g. Land Use, Soil Type
- To ensure even geographical distribution

Florida Soil Characterization Database

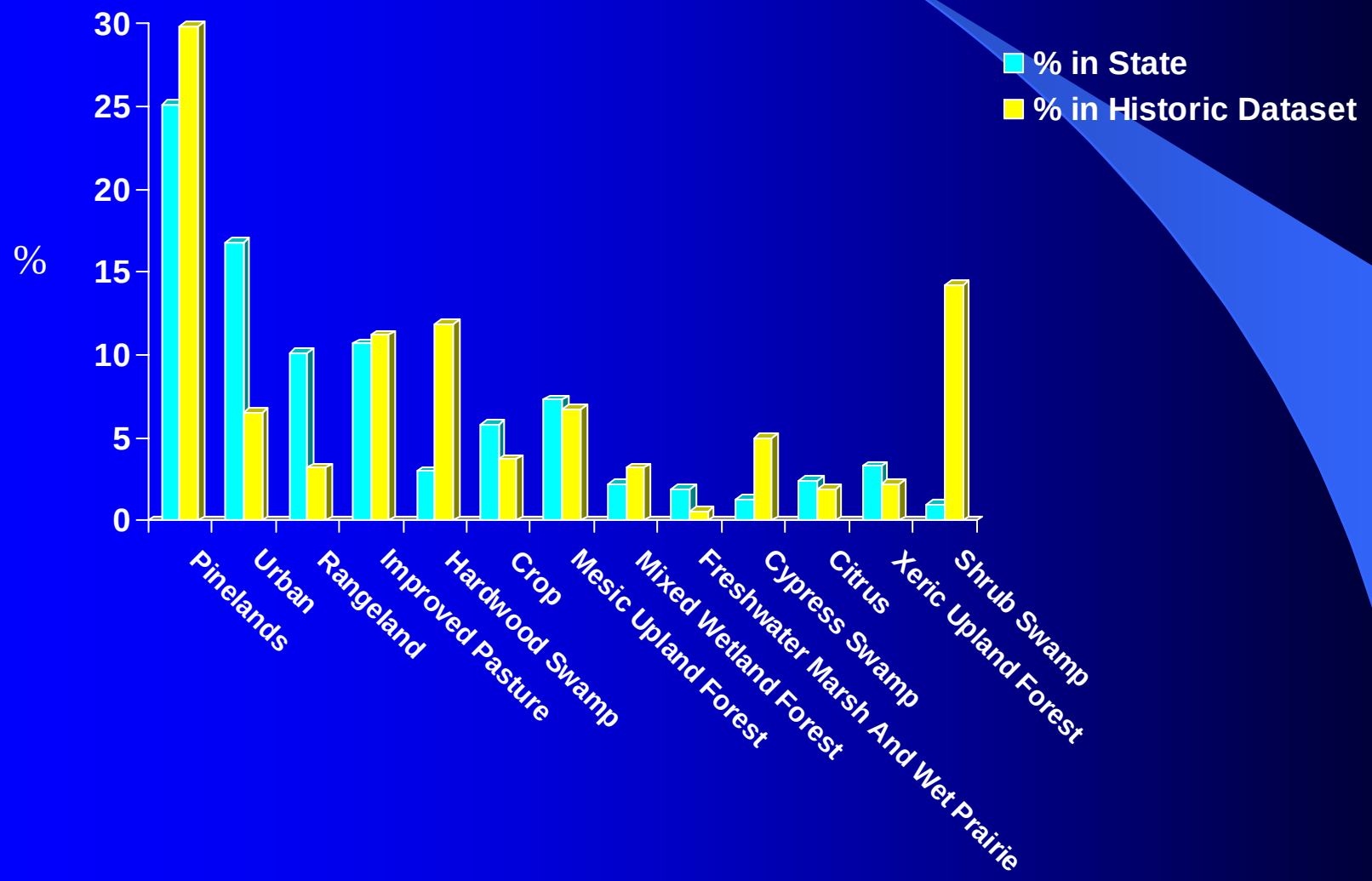
- Funded by “Accelerated Soil Survey Program”
- Database collected from 1965 to 1996
- Covering 58 counties
- 1288 profiles
- 8272 horizons
- 144 different physical and chemical soil properties (incl. soil C)



Distribution of Historic Samples across Suborders



Distribution of Historic Samples across Land uses



Role of GIS in Soil Survey

- Analyzing and integrating different data layers
- Useful for upscaling the measured soil properties

SCORPAN Model

$S_a = f(s, c, o, r, p, a, n)$

Where S_a = soil attribute to be predicted,

s = soil properties, c = climate, o = organisms, r = relief, p = parent material, a = age, n = space

McBratney et al., 2003

Objectives

- Measure and map soil carbon storage for the state of Florida
- Understand the effects of various environmental factors
- Compare the present soil C with historic measurements (~1990).
- Identify representative sampling locations that capture the underlying soil C variability
- To include 50% of the sites from the historic Florida Soil Characterization Database

Stratified Random Sampling

Advantages:

- **Captures the underlying variability of landscape factors**
- **Represents important subpopulations**
- **Efficient**
- **Strategic sampling across a landscape**
- **Greater accuracy of estimation due to lower sampling error**

Stratified Random Sampling

Disadvantages:

- **Necessary to identify “right” strata**
- **More knowledge of the population is needed.**
- **More complex to develop and analyze the results**
- **Depends on the accuracy of the data used for identifying the strata**

SCORPAN Model

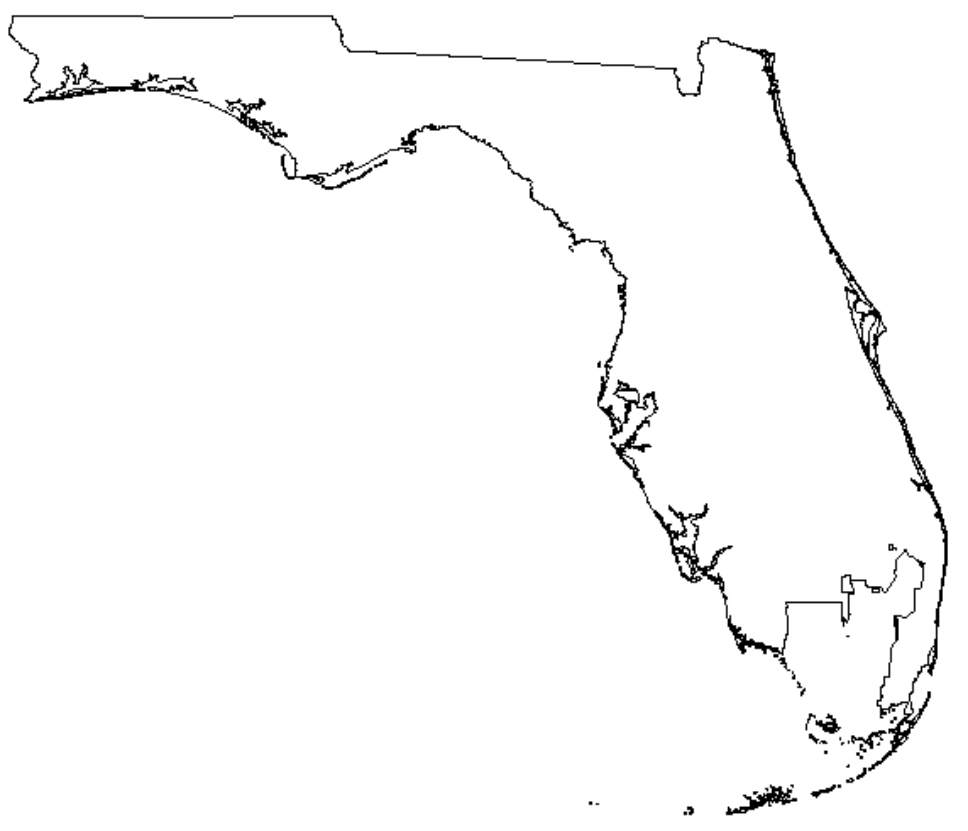
- **S = Soil properties** → Suborder
- C = Climate
- **O = Organisms** → Land Use
- R = Relief
- P = Parent material
- A = Age
- N = Space

Parton, 1987; Schulze and Freibauer, 2005

Delineation of the Study Area

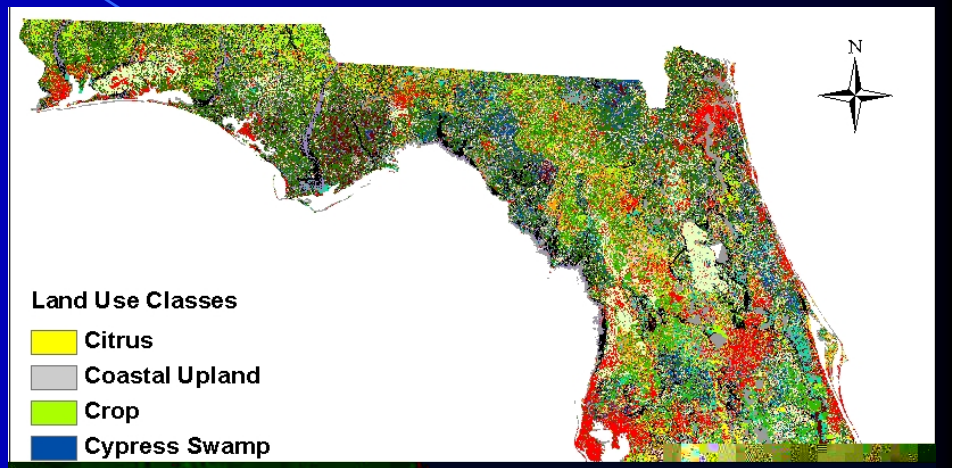
Areas excluded

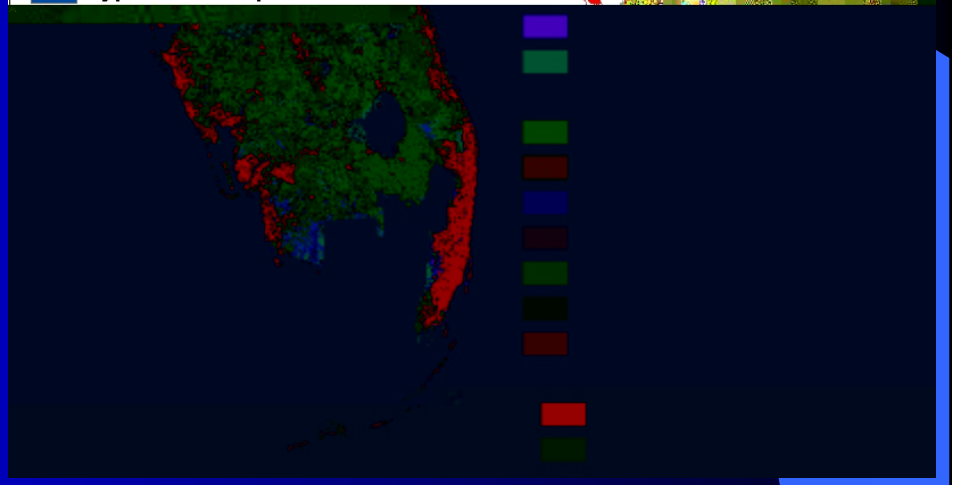
- The Greater Everglades area
- Specialty layers
  - Army sites
  - Interstate and state highways
  - Railways
  - Airports
  - Cemeteries



Land Use 2003

- Florida Vegetation and Land Cover Data (2003)
- Resolution: 30m
- Treatment (GIS):
  - Reclassification
  - Delineation
  - Area calculation



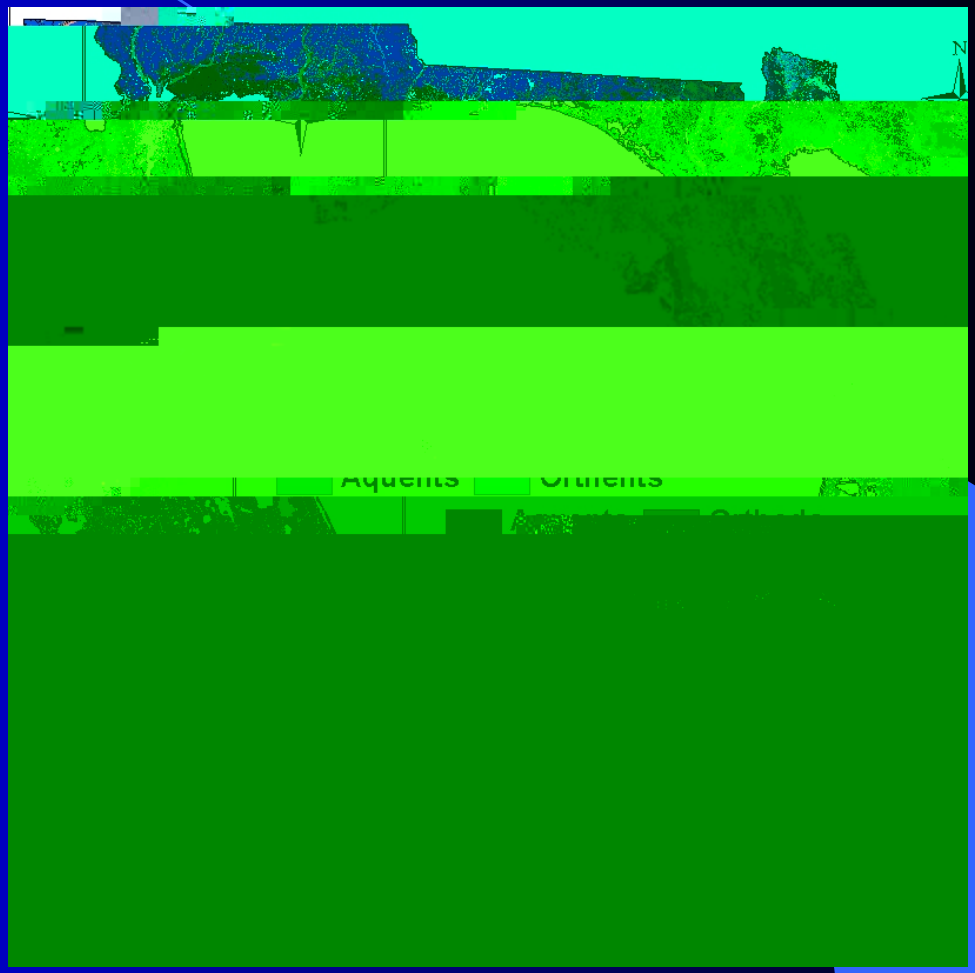


Data Source: Florida Fish and Wildlife Conservation Commission, Tallahassee, FL

Land Use Classes	Area (km^2)	%
Pinelands	**25,998**	**19.3**
Urban	16,073	11.9
Rangeland	12,516	9.3
Improved Pasture	11,992	8.9
Hardwood Swamp	7,993	5.9
Crop	7,734	5.7
Mesic Upland Forest	7,512	5.6
Mixed Wetland Forest	6,087	4.5
Freshwater Marsh and Wet Prairie	5,730	4.2
Cypress Swamp	5,330	4.0
Xeric Upland Forest	4,435	3.3
Shrub Swamp	3,955	2.9
Citrus	3,851	2.9
Other Agriculture	879	0.7
Exotic Plants	178	0.1
Coastal Upland	62	0.0
Total	134,830	100

Soil Suborders

- Suborder data from NRCS Soil Data Mart (2006)
- Scale 1:24,000
- Treatment (GIS)
  - Extraction of county data and merging
  - Delineation
  - Area calculation

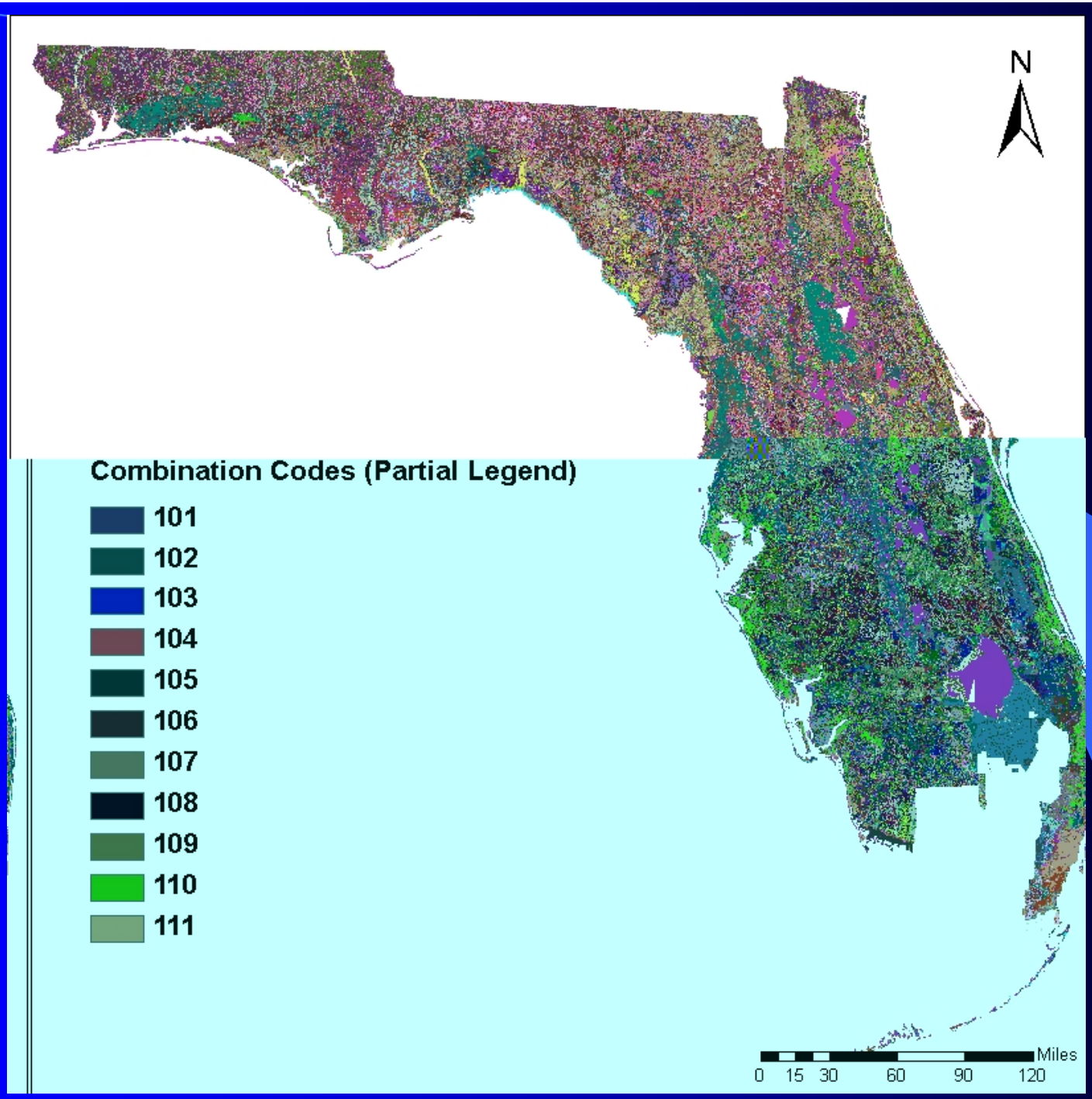


Data Source: Natural Resources Conservation Service - Soil Data Mart

Suborder	Area(km^2)	%
Aquods	**30,157**	**22.41**
Psamments	18,104	13.45
Udults	18,007	13.38
Aqualfs	15,105	11.22
Saprists	9,555	7.10
Aquents	8,624	6.41
Aquults	6,440	4.79
Aquepts	4,886	3.63
Aquolls	4,099	3.05
Orthods	3,001	2.23
Udalfs	2,035	1.51
Arents	1,329	0.99
Hemists	1,046	0.78
Orthents	888	0.66
Humods	391	0.29
Udepts	376	0.28
Fluvents	143	0.11
Water	10,244	7.61
Total	13,4577	100

Combination of Suborder and Land Use Strata

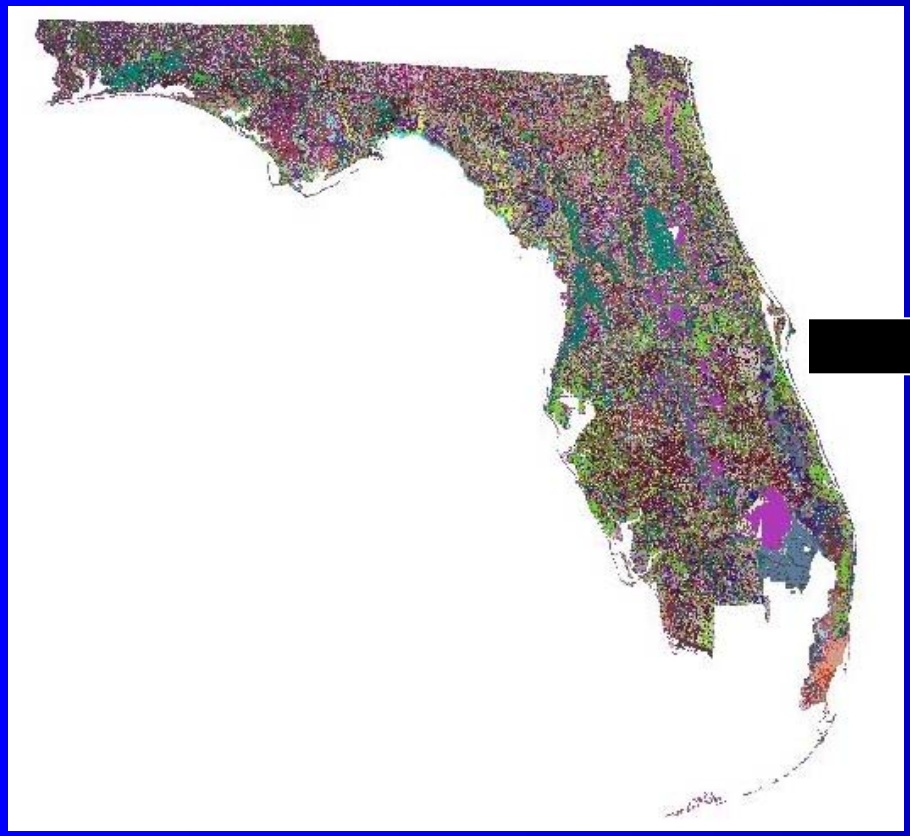
- **Assigning numbers to different land uses and suborders**
  - **Values from 100 to 1700 were assigned to the land use classes and**
  - **Values from 1 to 24 were assigned to the suborders.**
- **Addition of the two layers in the Spatial analyst (ArcGIS)**
- **E.g. 1700 + 1 = 1701, where the first two digits indicate the land use and the last two digits denote the suborder**

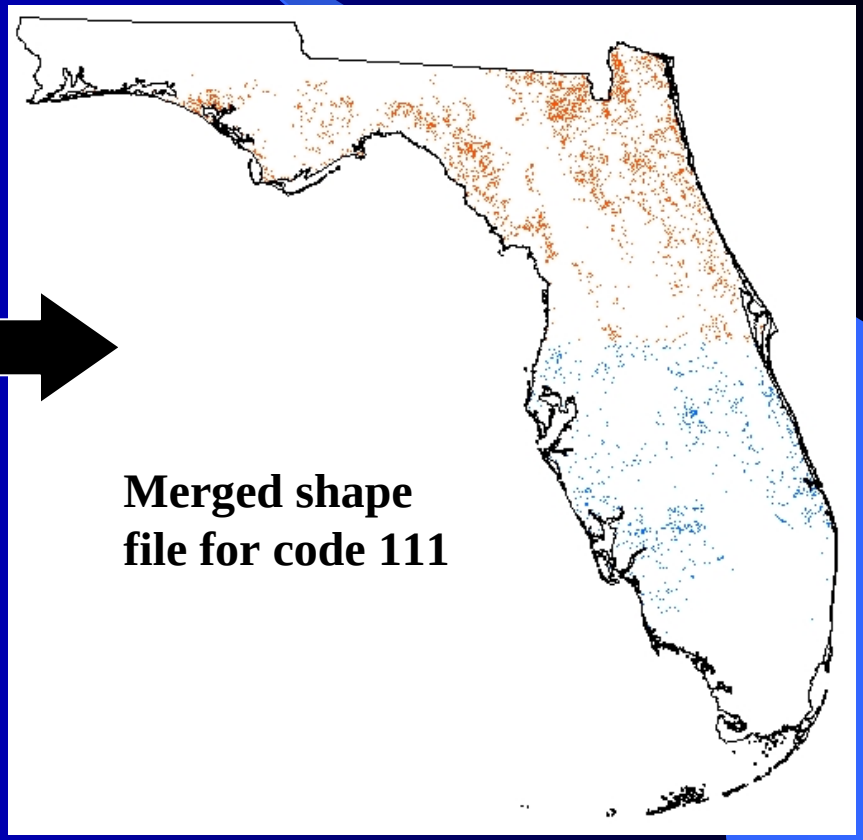

N
Combination Codes (Partial Legend)
101
102
103
104
105
106
107
108
109
110
111
Miles
0 15 30 60 90 120

Land use + suborder code	Land_Use	Suborder	Area (sq.km.)	%
111	**Pinelands**	**Aquods**	**7267**	**7.8**
101	Pinelands	Udults	5967	6.4
511	Rangeland	Aquods	4746	5.1
411	Improved Pasture	Aquods	4451	4.7
1611	Urban	Aquods	3653	3.9
1606	Urban	Psamments	3270	3.5
706	Xeric Upland Forest	Psamments	2800	3
106	Pinelands	Psamments	2765	2.9
201	Crop	Udults	2532	2.7
104	Pinelands	Aquults	2402	2.6
114	Pinelands	Aqualfs	2451	2.6
801	Mesic Upland Forest	Udults	2243	2.4
202	Crop	Saprists	1857	2

Land use + suborder code	Land_Use	Suborder	No of samples from historic dataset	No of samples from reconnaissance dataset	Total
111	**Pinelands**	**Aquods**	**51**	**27**	**78**
101	Pinelands	Udults	42	22	64
511	Rangeland	Aquods	24	27	51
411	Improved Pasture	Aquods	24	23	47
1611	Urban	Aquods	25	14	39
1606	Urban	Psamments	23	12	35
706	Xeric Upland Forest	Psamments	20	10	30
106	Pinelands	Psamments	19	10	29
201	Crop	Udults	18	9	27
104	Pinelands	Aquults	17	9	26
114	Pinelands	Aqualfs	17	9	26
801	Mesic Upland Forest	Udults	16	8	24
202	Crop	Saprists	8	12	20

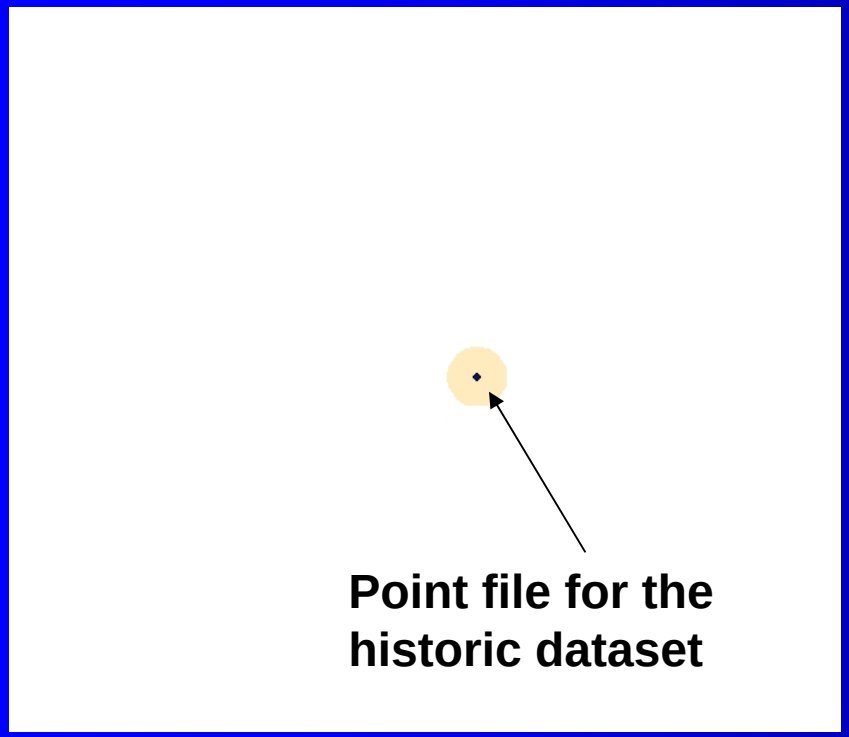
Separating the Land Use + Suborder Combinations



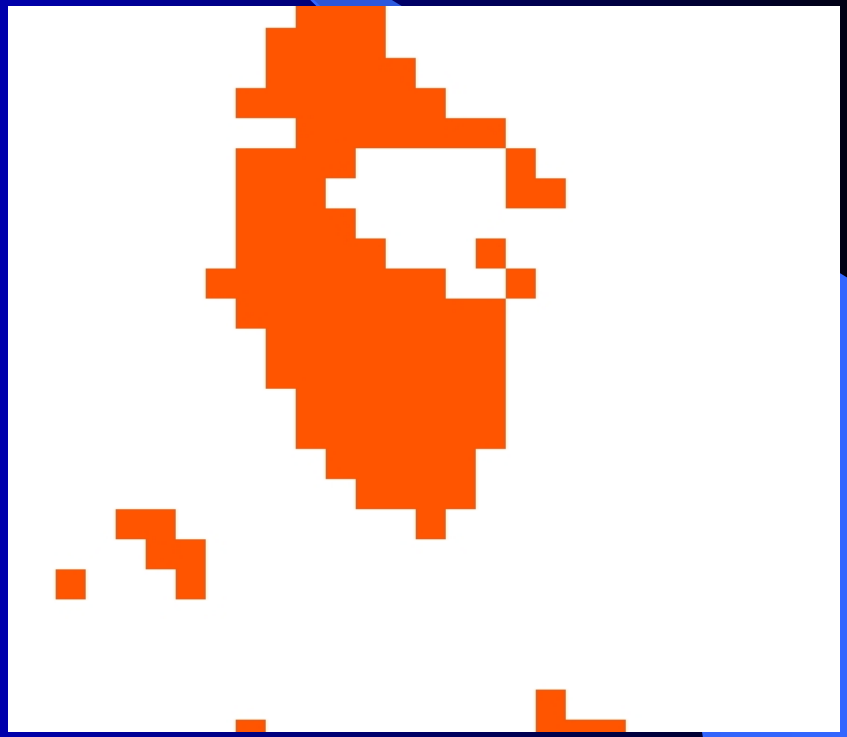


Excluding Historic Dataset

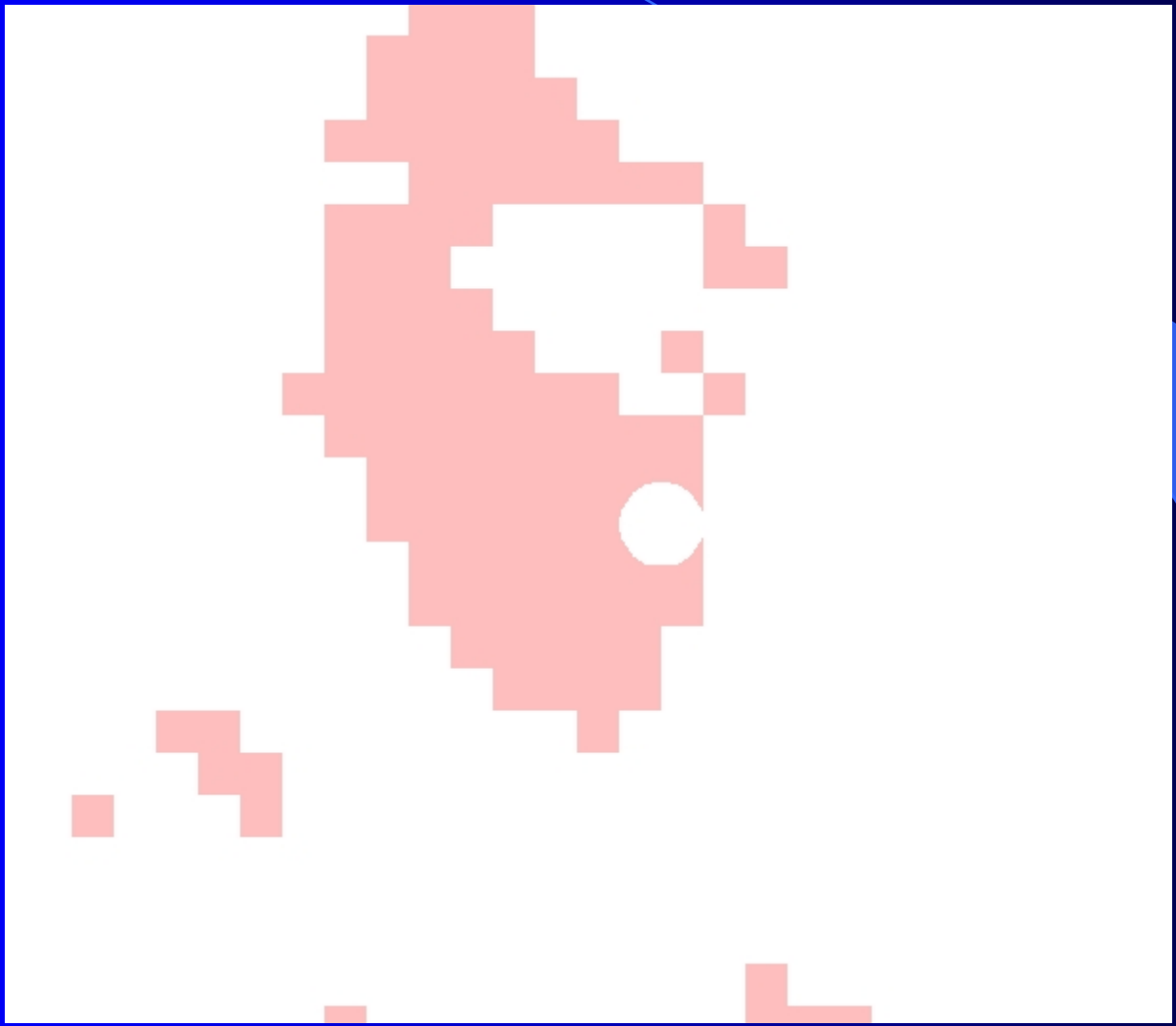
Shape file with 30m buffer around the historic site points



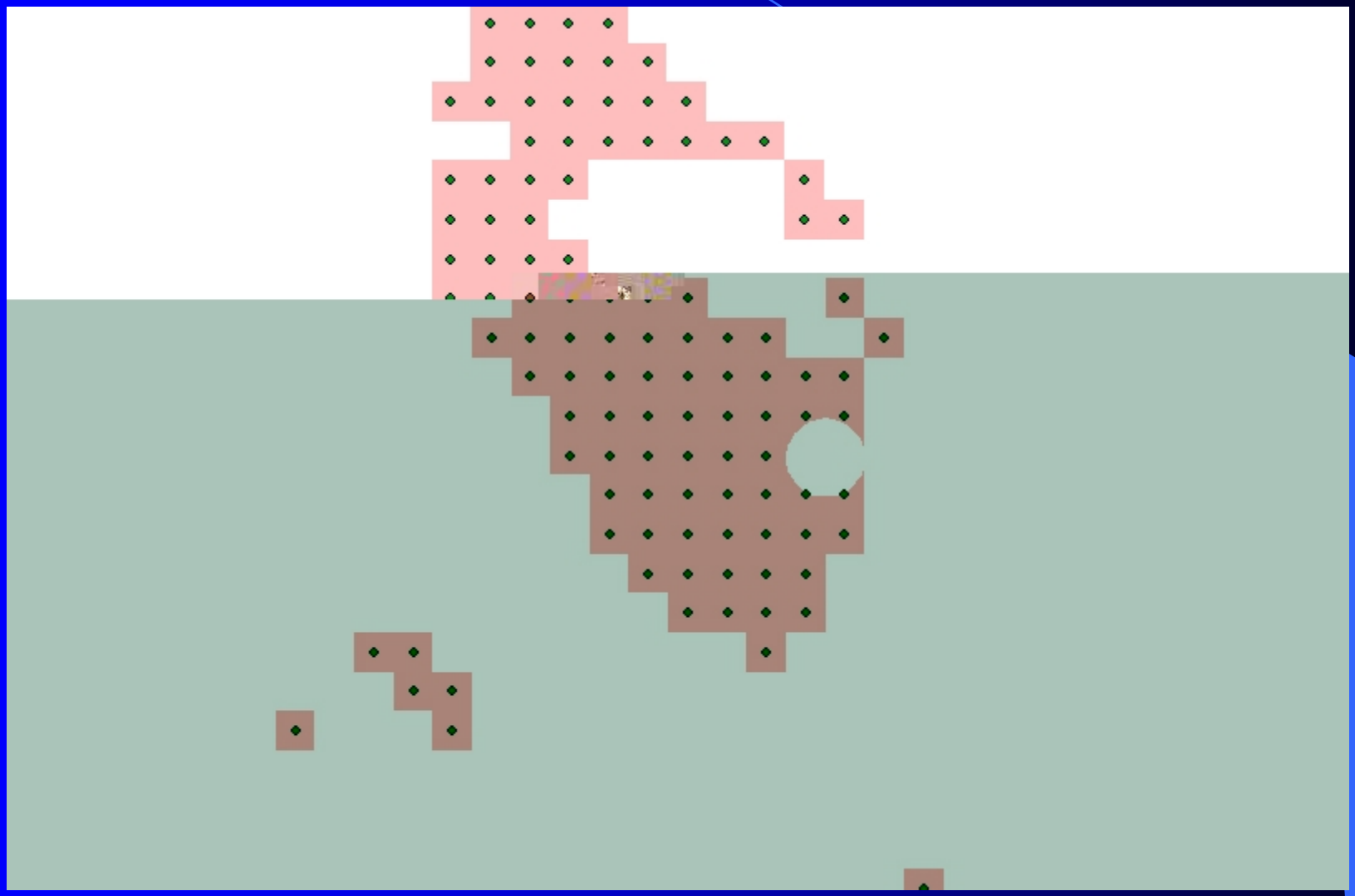
Merged shape file for the code 111



Dataset Without Historic Data



Constructing Point Files



Random Selection Using SPSS

Importing data for each code

↓

Data

↓

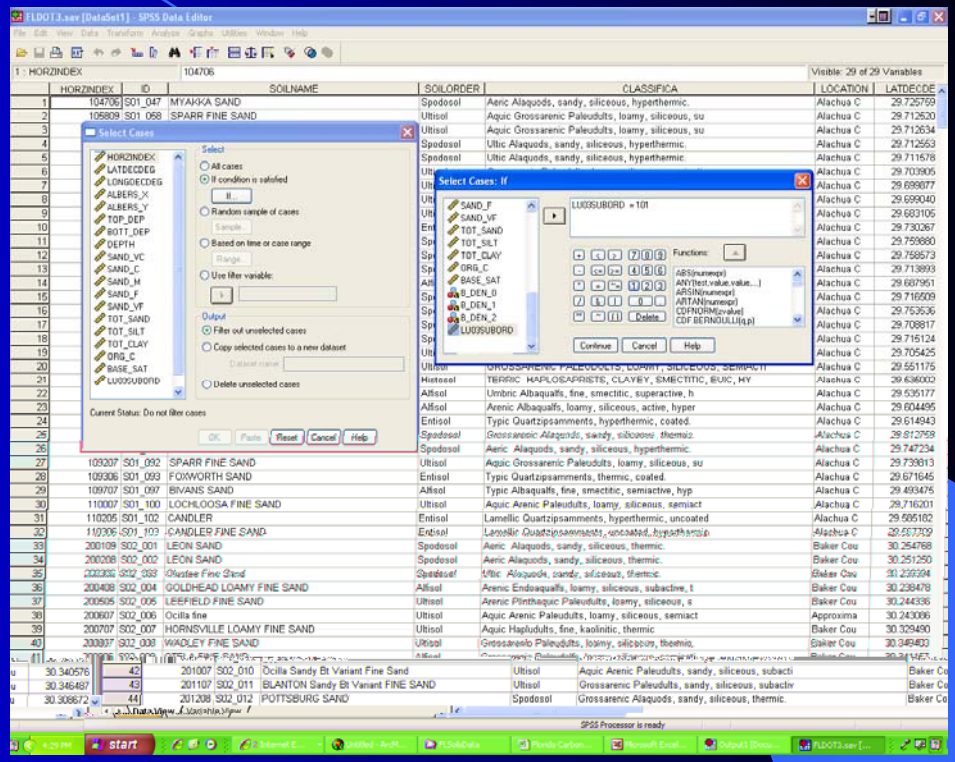
Select Cases

↓

Random sample of cases

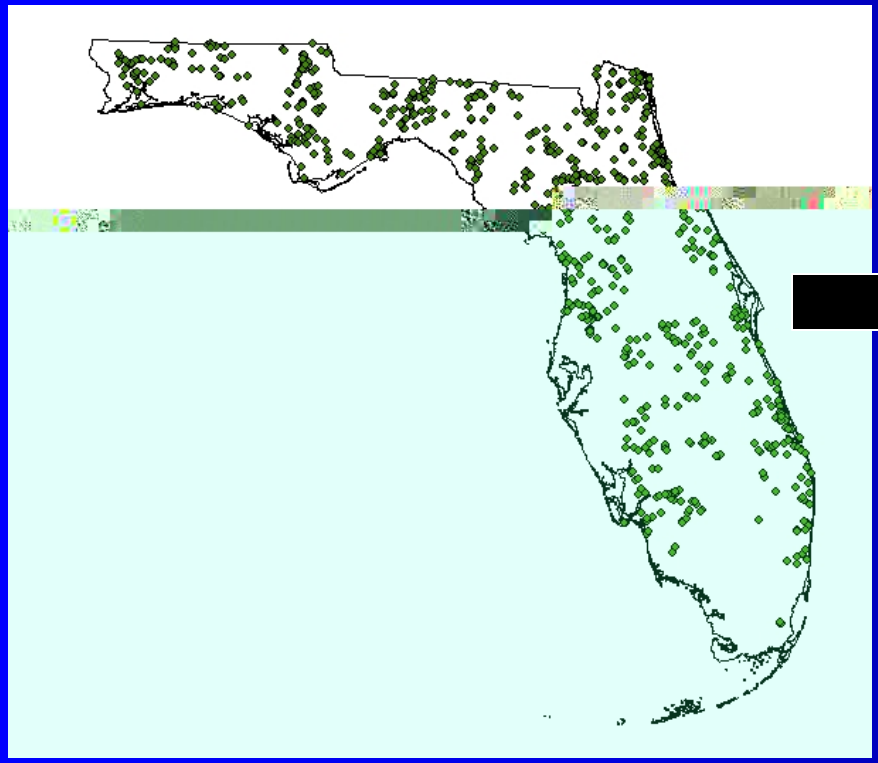
↓

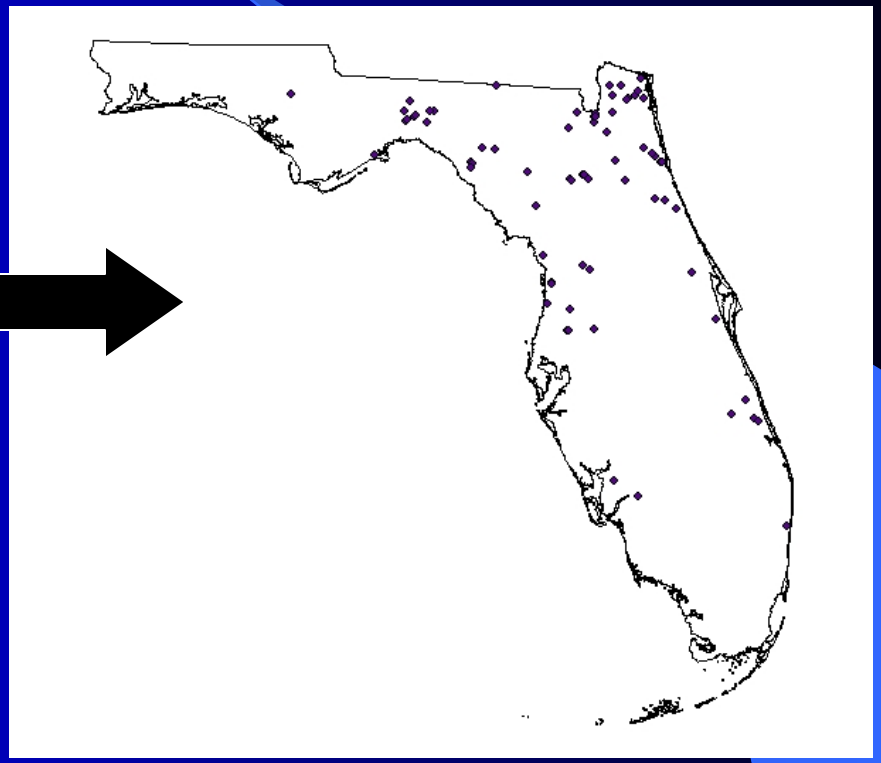
Exactly ‘x’ cases from the total ‘y’ cases



Stratified Random Sampling for the Historic Data

Selecting Sites from Historic Data





Random Selection Using SPSS

Importing data for each code

↓

Data

↓

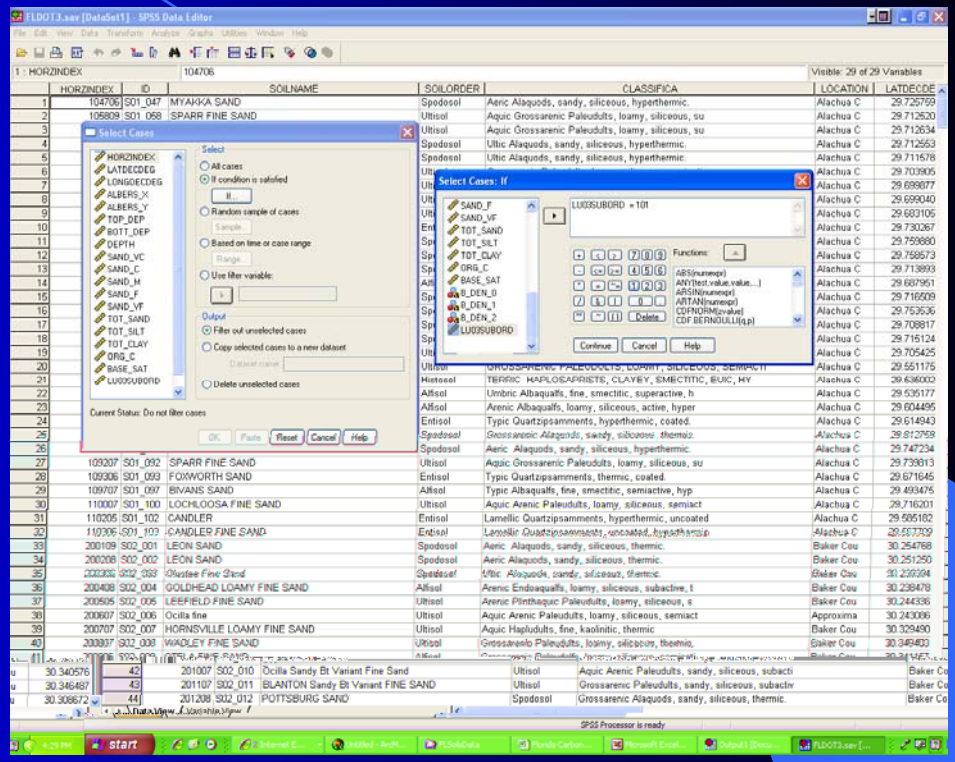
Select Cases

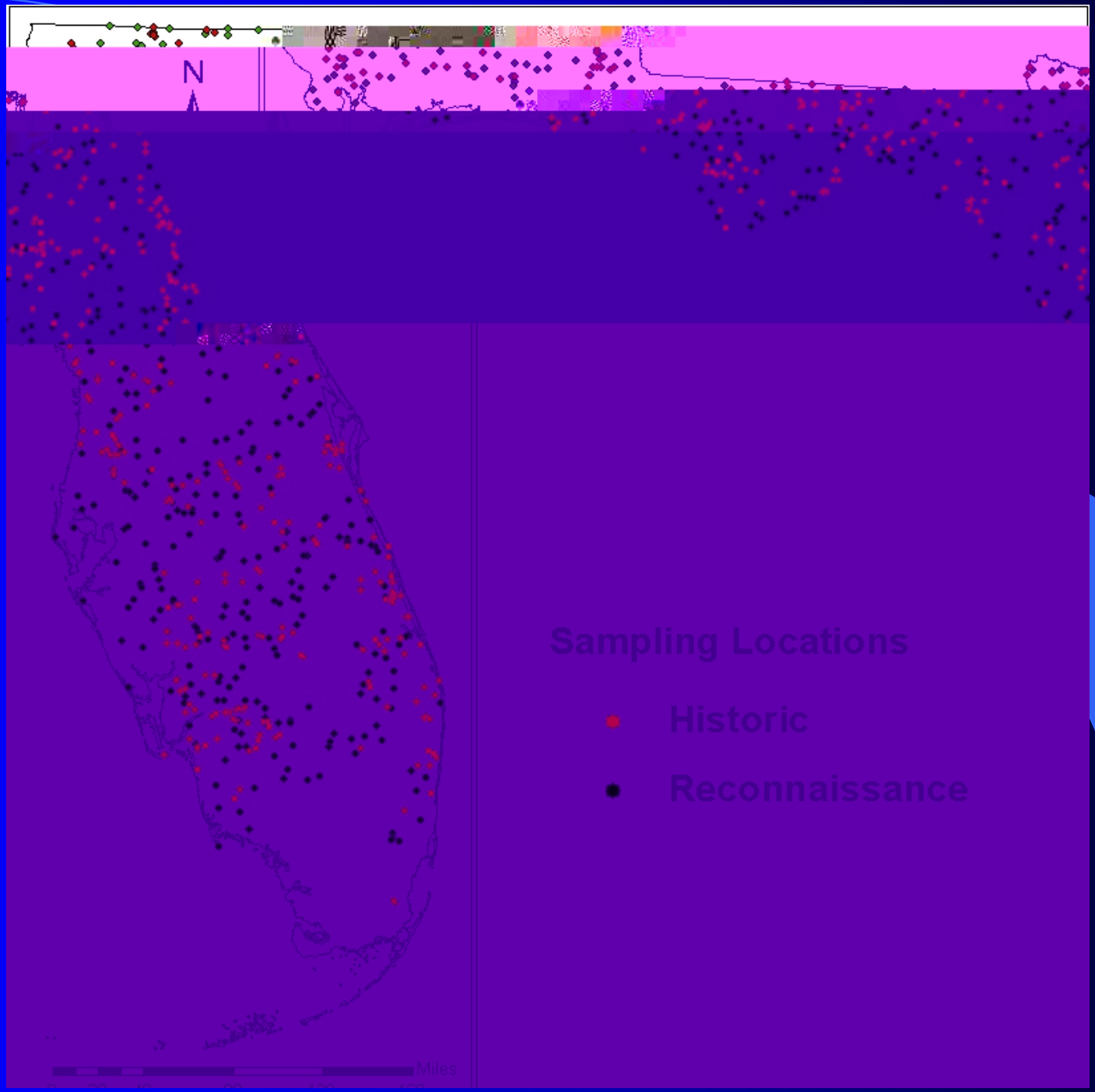
↓

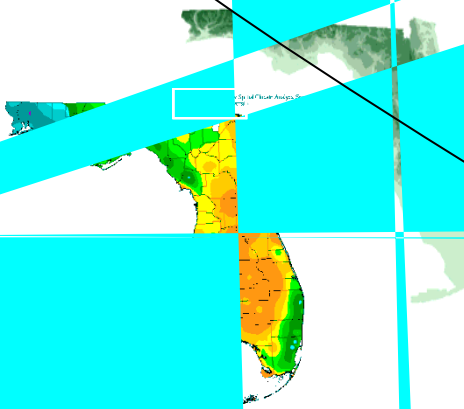
Random sample of cases

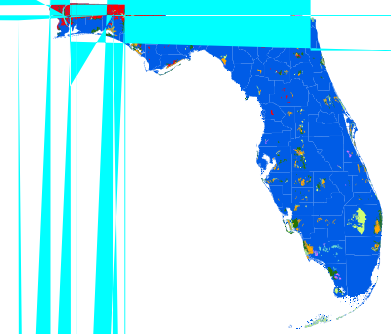
↓

Exactly ‘x’ cases from the total ‘y’ cases




N
Sampling Locations
Historic
Reconnaissance
Miles

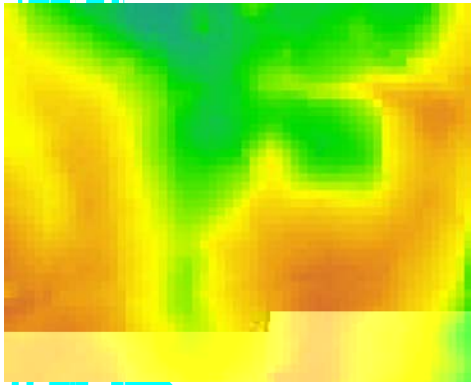


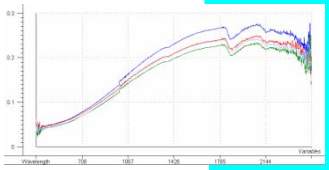


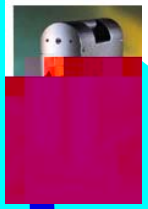
+

F
s

+





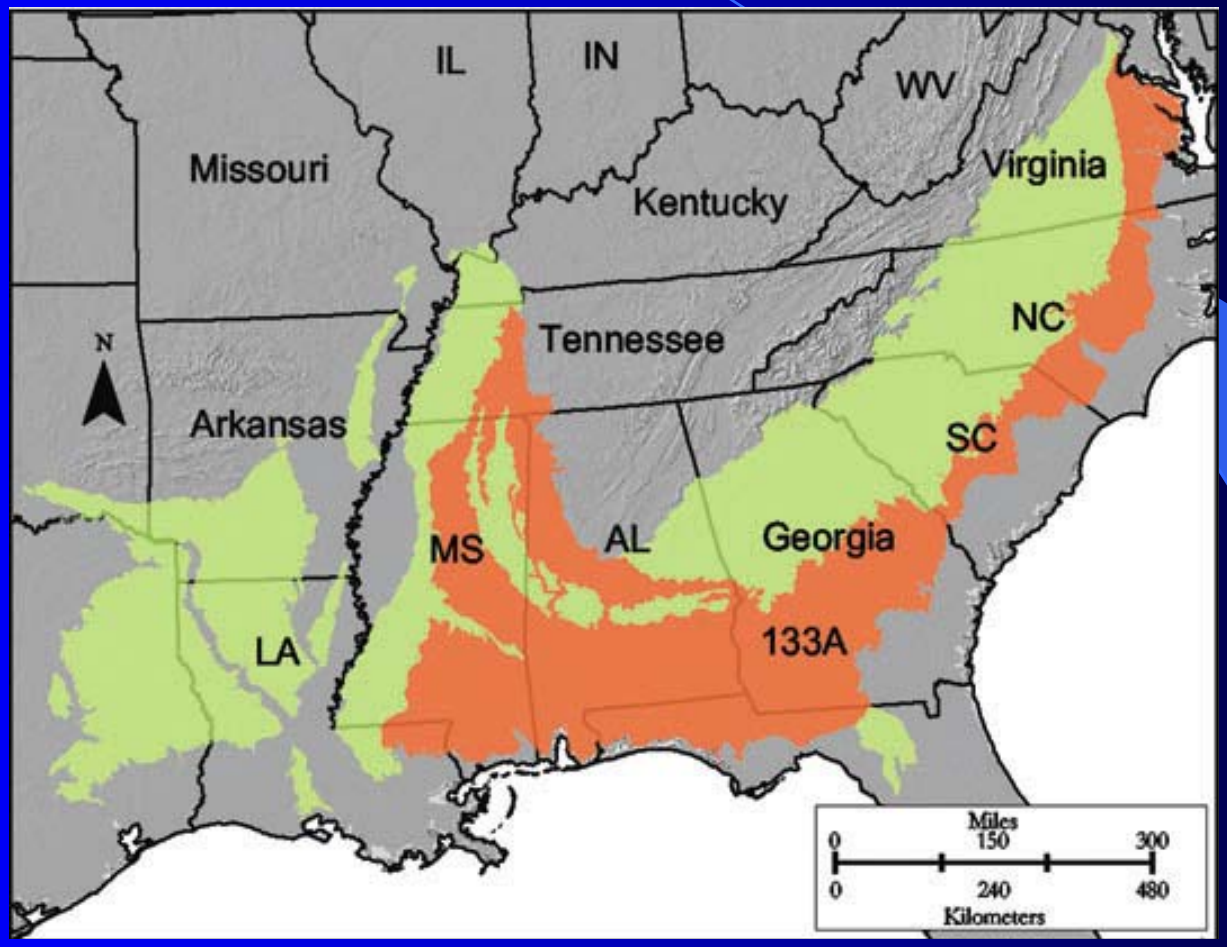


Applications

Other large scale projects

- **Strategic soil sampling across Major Land Resource Areas**
- **The USDA NRI topic areas Global Climate Change, Soil Processes etc.**
- **Tool in landscape SOC assessments, supporting the mission of USDA to improve soil quality and reduce the loss of SOC as a result of changing land uses.**

Use of other input layers



e.g. DEM for MLRA 133A - Southern Coastal Plain

Other Soil Properties of Significance

- **Land use and geology layers for soil phosphorus**
- **Rainfall and land use for risk of nitrate contamination**

Other Advantages of the Sampling Technique

- Applicable for large areas by overcoming the 2GB processing limit of ArcGIS
- Any spatial data layer can be used for developing the sampling design
- The method offers precision and flexibility for developing even more complex sampling design
- The layers developed in the process offer valuable information about the population