Statistical perspectives on spatial social science

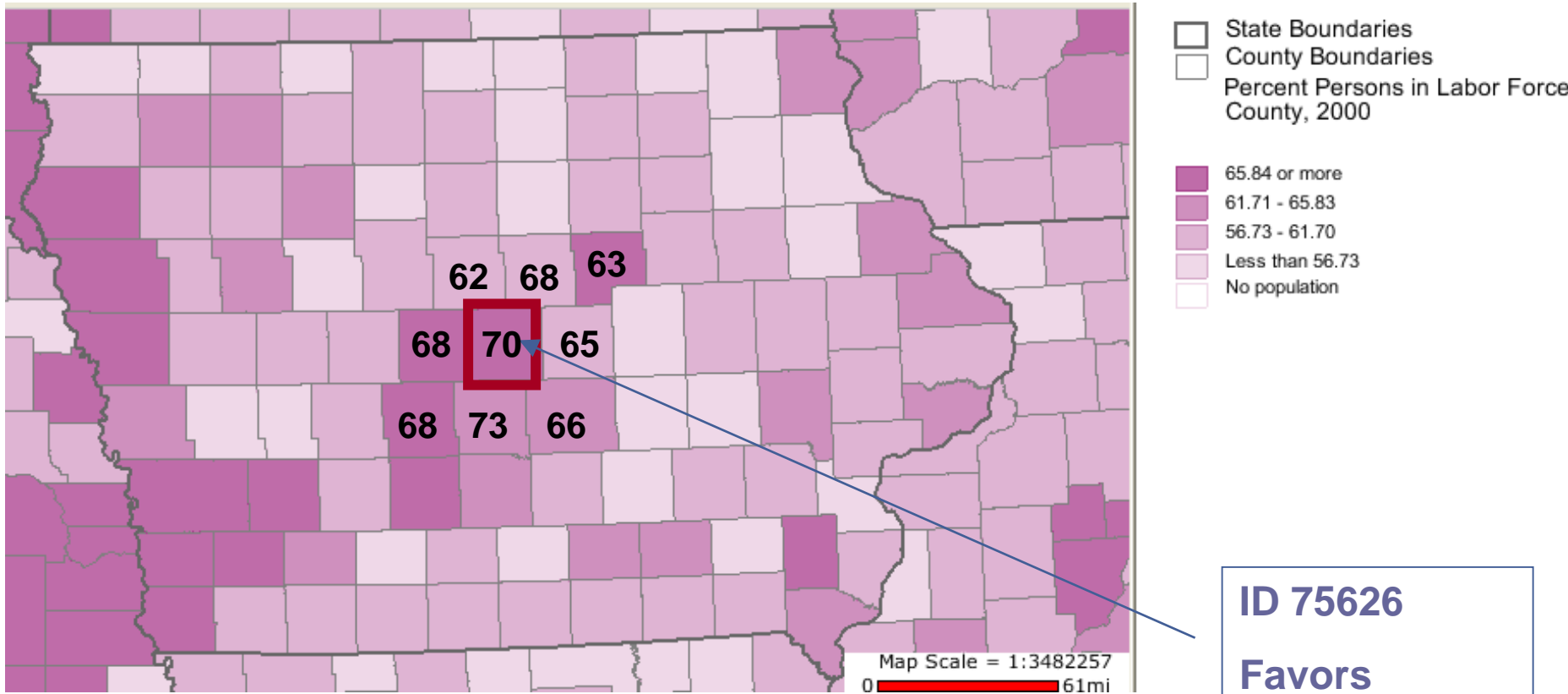# Discussion

Sarah Nusser   (nusser@iastate.edu)

Center for Survey Statistics and Methodology

Department of Statistics

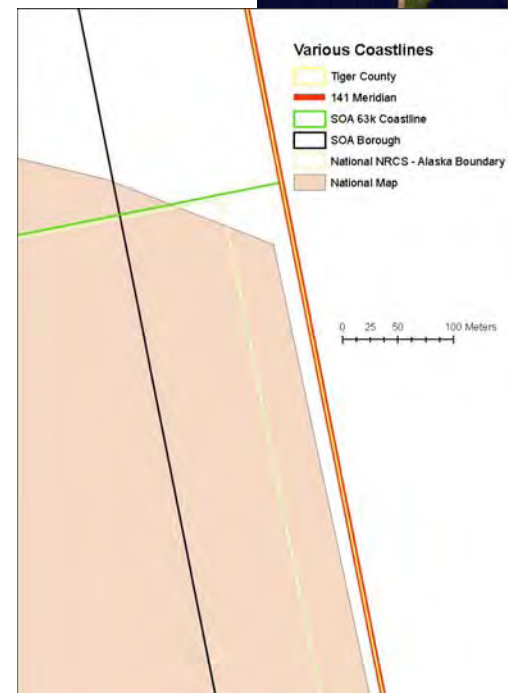Iowa State University

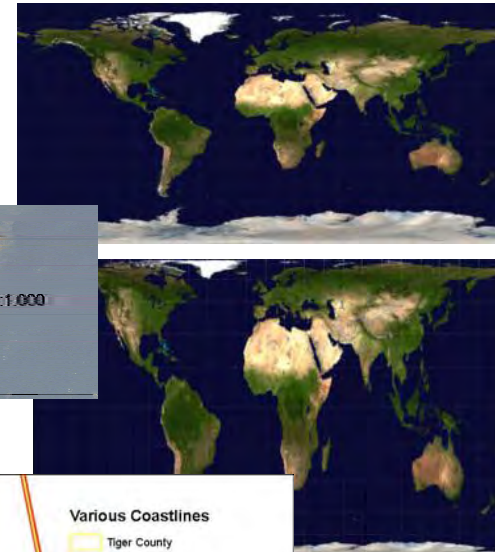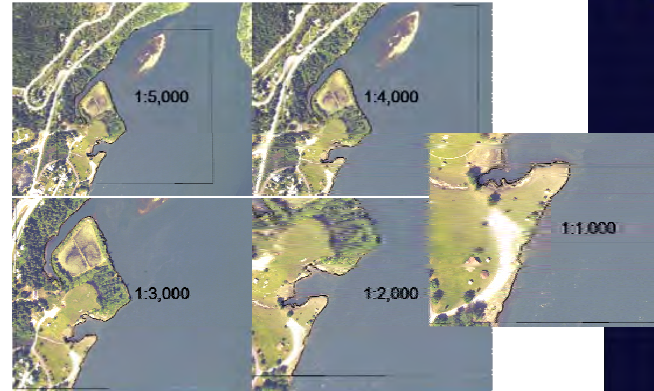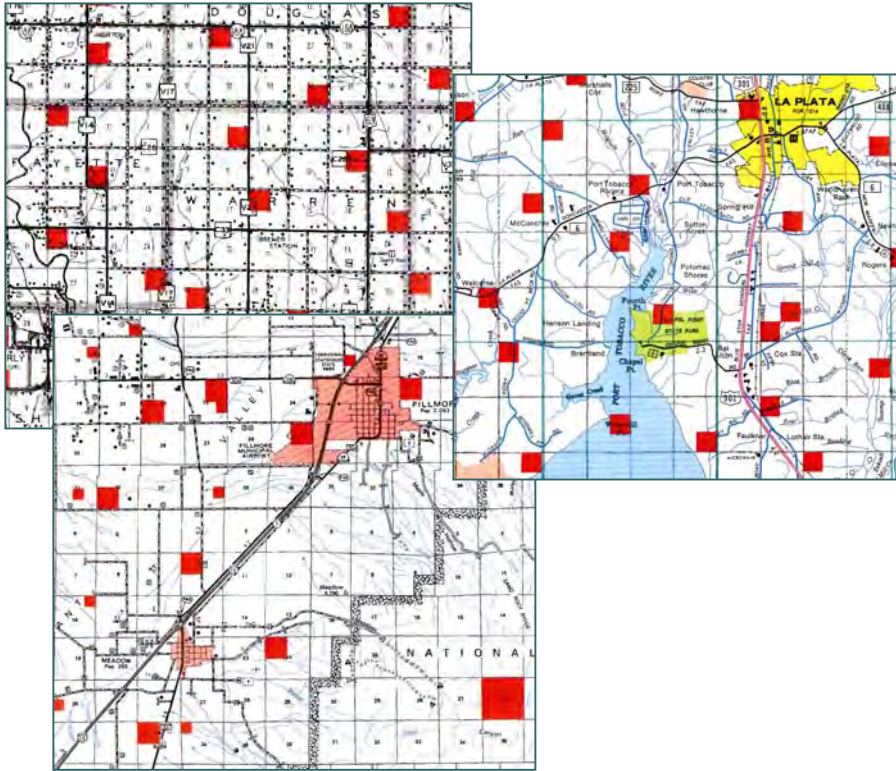Morris Hansen Lecture, WSS Nov 6, 2006

Linking geographic information

Black numbers from Census 2000 labor force data

Map of labor force intervals from second party web-based map utility with Census 2000 data

2

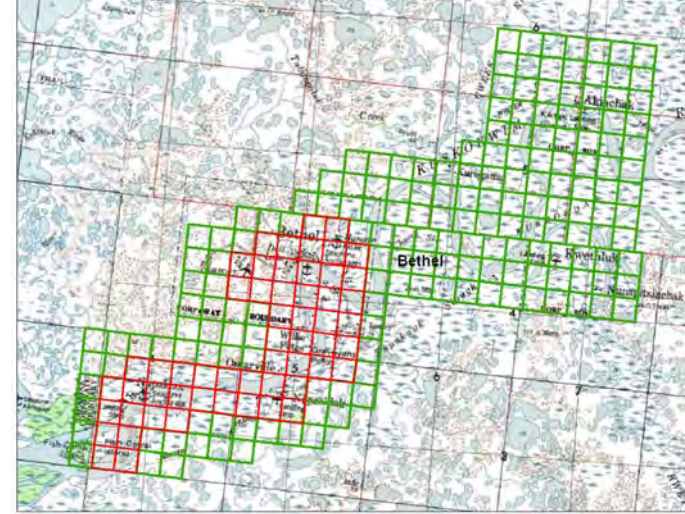# Understanding geospatial data and their errors

- Attributes (analysis variables)
- Location (link)
  - Coordinate systems, projections
  - Scale, resolution
  - Errors in location
- Conflicting data, classifications
  - Metadata?

Various Coastlines
- Tiger County
- 141 Meridian
- SOA 63k Coastline
- SOA Borough
- National NRCS - Alaska Boundary
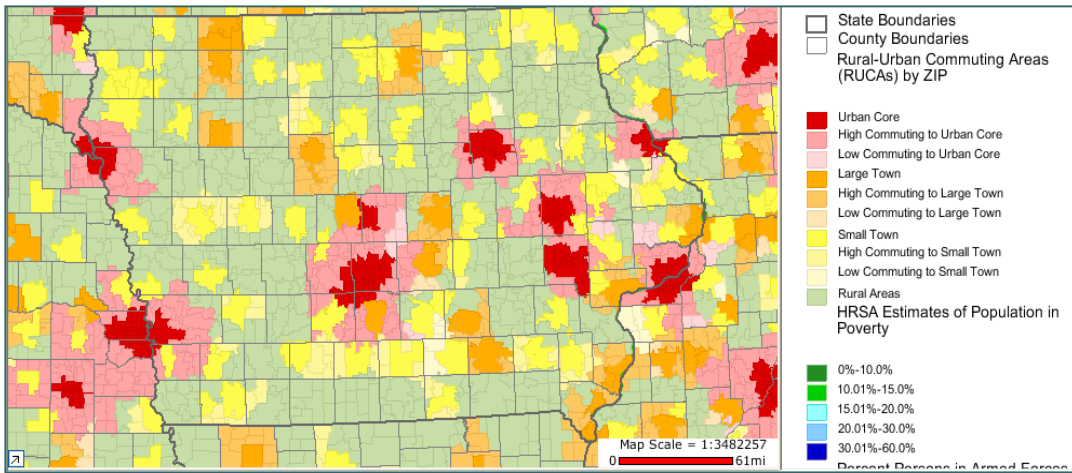- National Map

Area sampling

Area sample of quarter sections
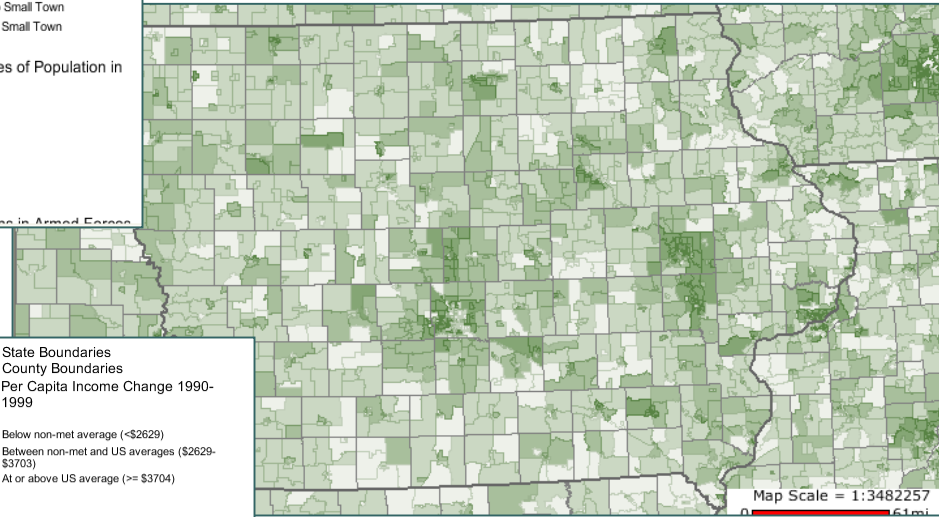on county transportation map
(ISU 1970s)
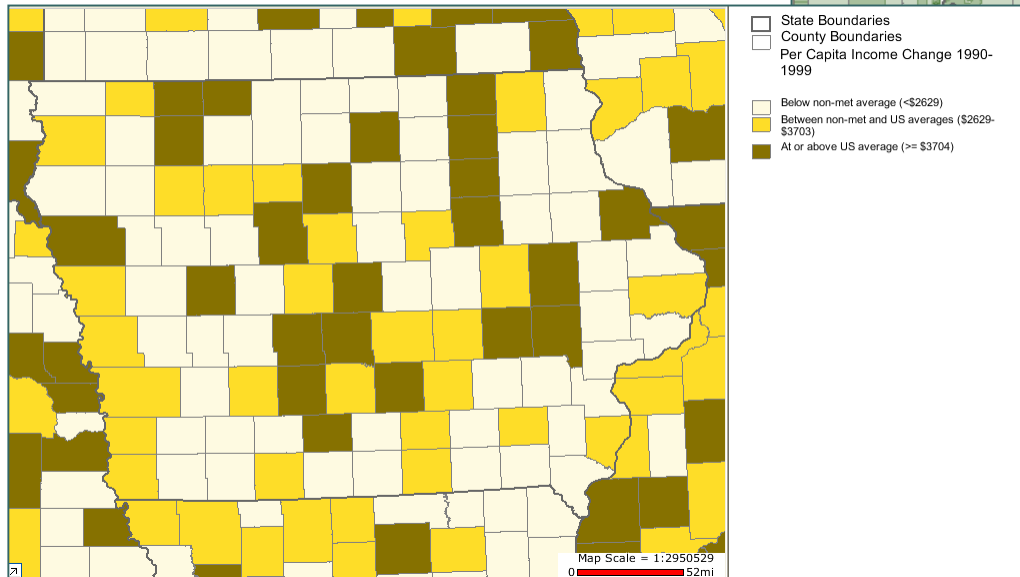

Digital GIS area sample frame
(ISU 2005)

Aerial photo – using visual features
for sampling structures (NASS)

4

Commuting areas
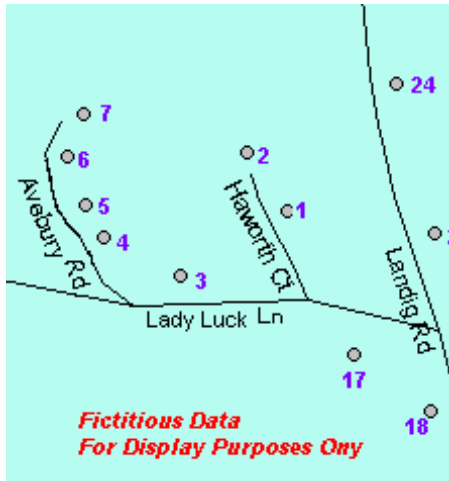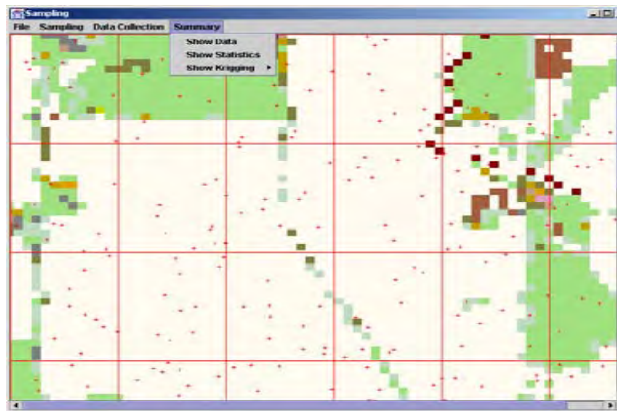

Education level

Change in income in the 1990s

5

Coordinates & sample units

Geo-coded addresses?

Real-time in-field sampling
(Mukhopadhyay 2003)

NRI: true location is
not GPS coordinate,
but on base image
for data collection

Planning & navigation

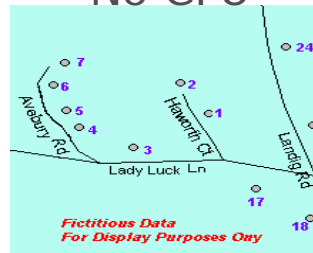Pilot studies of GPS, maps, navigation aids (BLS, Census Bureau 2001-2003)

Navigation aids appear to improve performance, but performance also related to ability to mentally manipulate / visualize spatial information
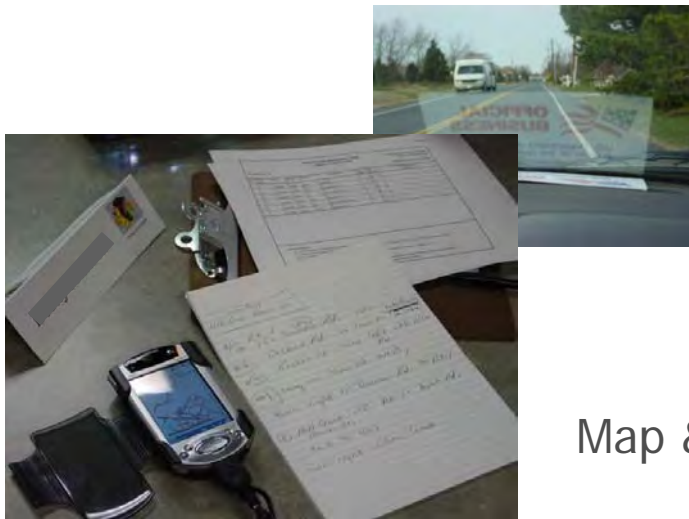
No GPS   GPS

No route info

Map & text route info
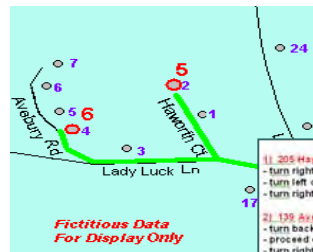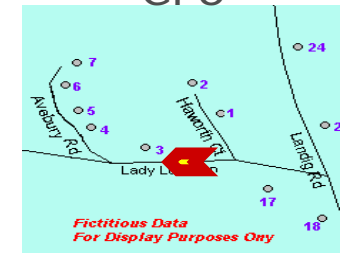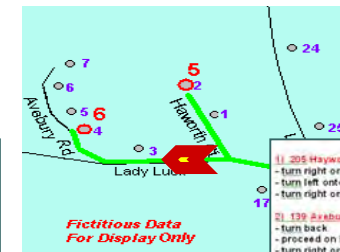
Collecting geospatial data objects
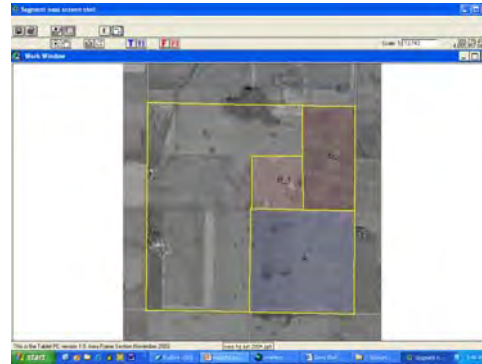
Using digital images & GIS to collect field boundaries from farm operators (NASS 2003)

Digitized fields from **interview** and **office** digitizing were consistent with one another

Which area measure?

2005 Adams County, Iowa

Categories
- Corn
- Soybeans
- Sunflowers
- All Small Grains
- Other crops
- Fallow/Idle
- Non-Ag/Pasture
- Woods
- Clouds
- Urban
- Water
- Wetlands

CSSM
Iowa State University

USDA
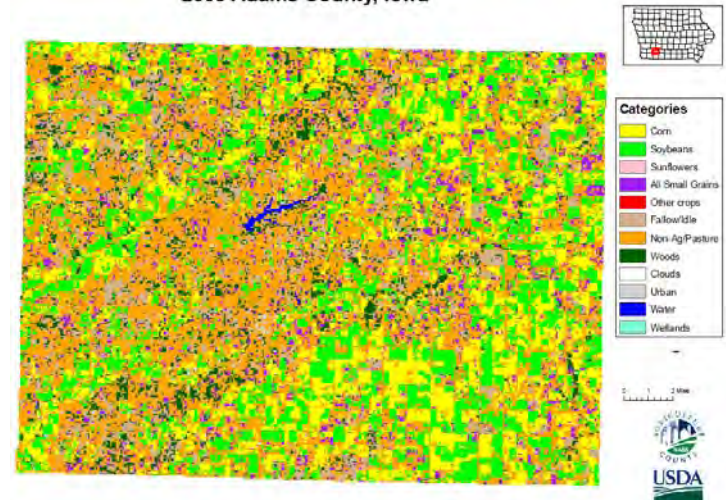
GIS survey instruments (and systems)
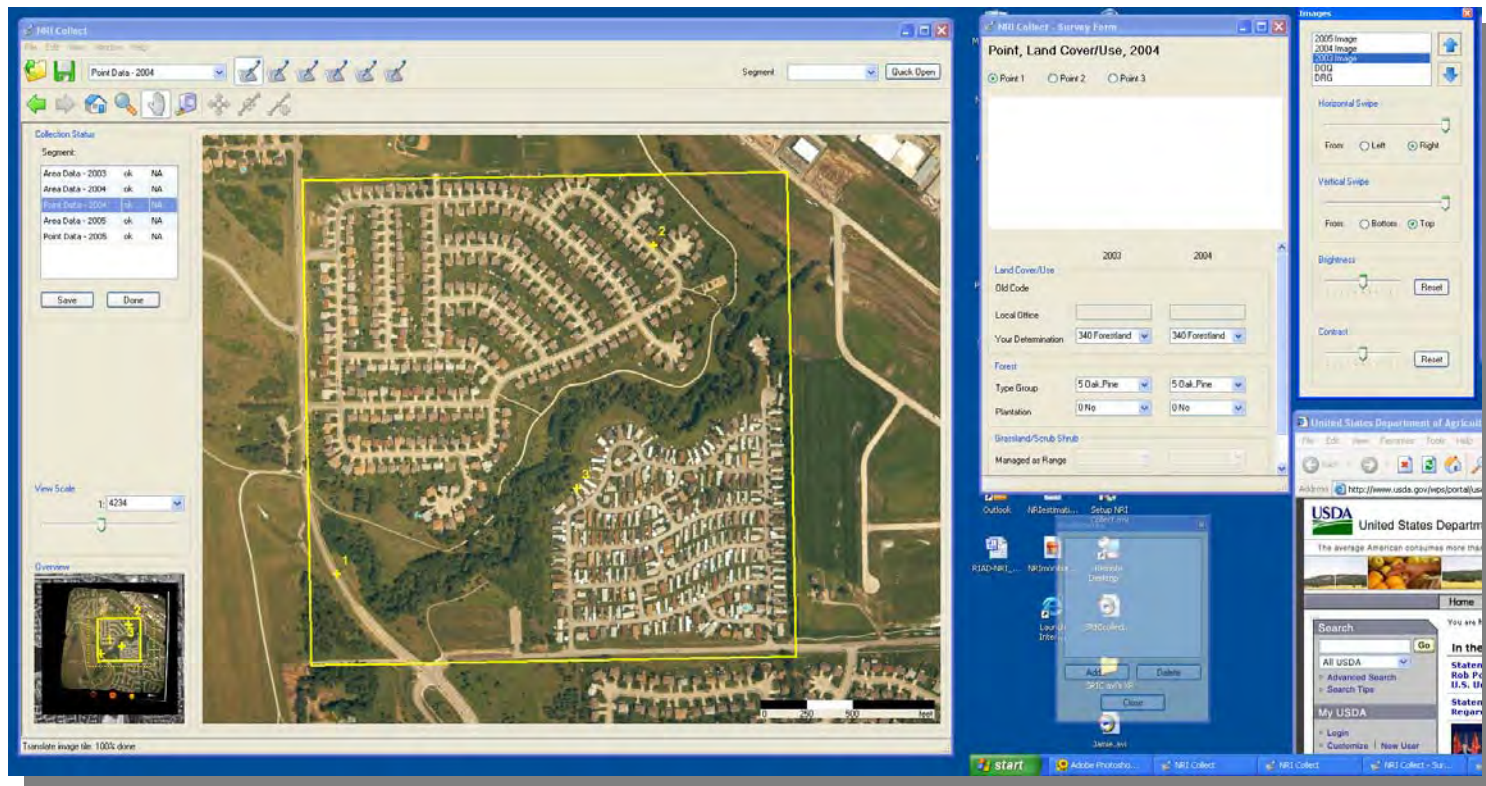
Custom software with graphics acceleration, simplified tool set, survey process control, access to multiple resources   (ISU / NRCS 2006)

CSSM
Iowa State University

Collecting data over time

- Measuring change from one year to the next requires a stable coordinate base for polygons
- Start with orthoimage
- Register next orthoimage to stable features on the ground to maintain base for change

Registered (2D)   Orthorectified (3D)

Time 1

Time 2

CSSM
Iowa State University
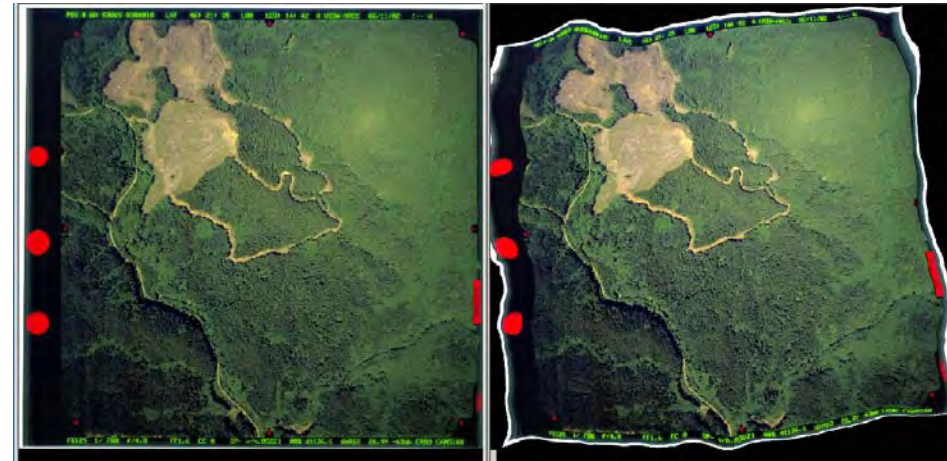
GPS-derived objects

- Coordinates of sample units
  - Link sample units to other geocoded features, e.g., point source hazards
- GPS paths (space and time)
  - Quality monitoring for interviewers
  - Data for travel surveys
  - Scope of travel and activity for elderly (Hicks 2004)
  - Obtaining speed in a seatbelt survey as input to weights (Westat 2006)

CSSM
Iowa State University
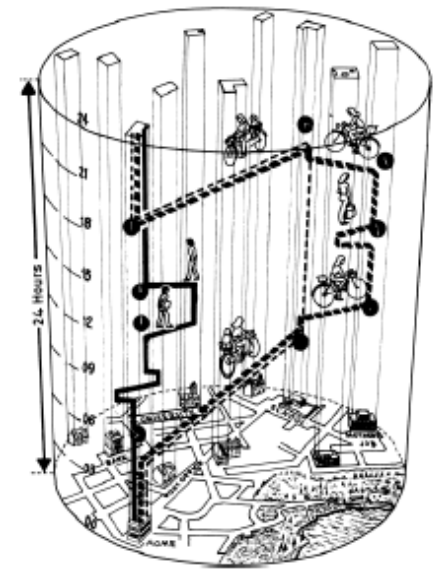
Imputation / weighting to reflect local conditions

GIS coverage of
federal land and large water

- Control totals for land cover/use categories from wall-to-wall coverage

- Imputation of records ★ to represent known patterns within small domains

Watershed inside county
shows federal land,
but no segments / points
intersect with the federal land

2000

2000

2000,
2005

2000,
2005

2000,
2005
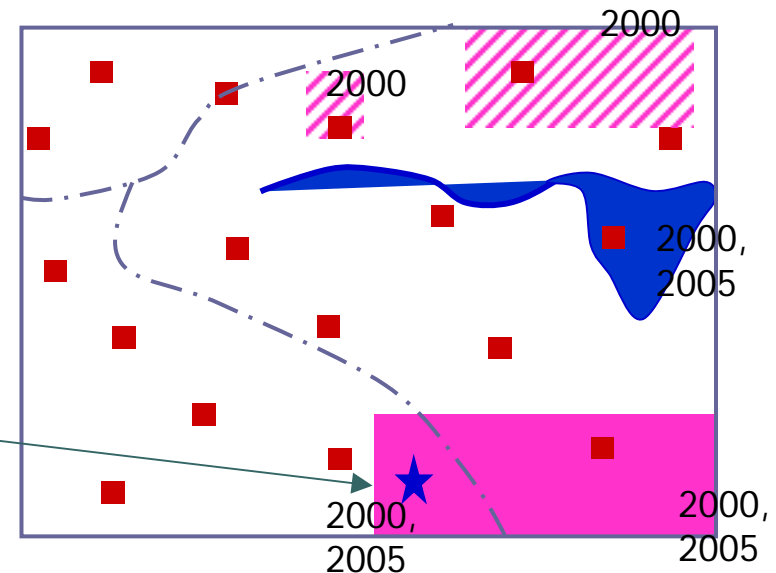
Disclosure limitation

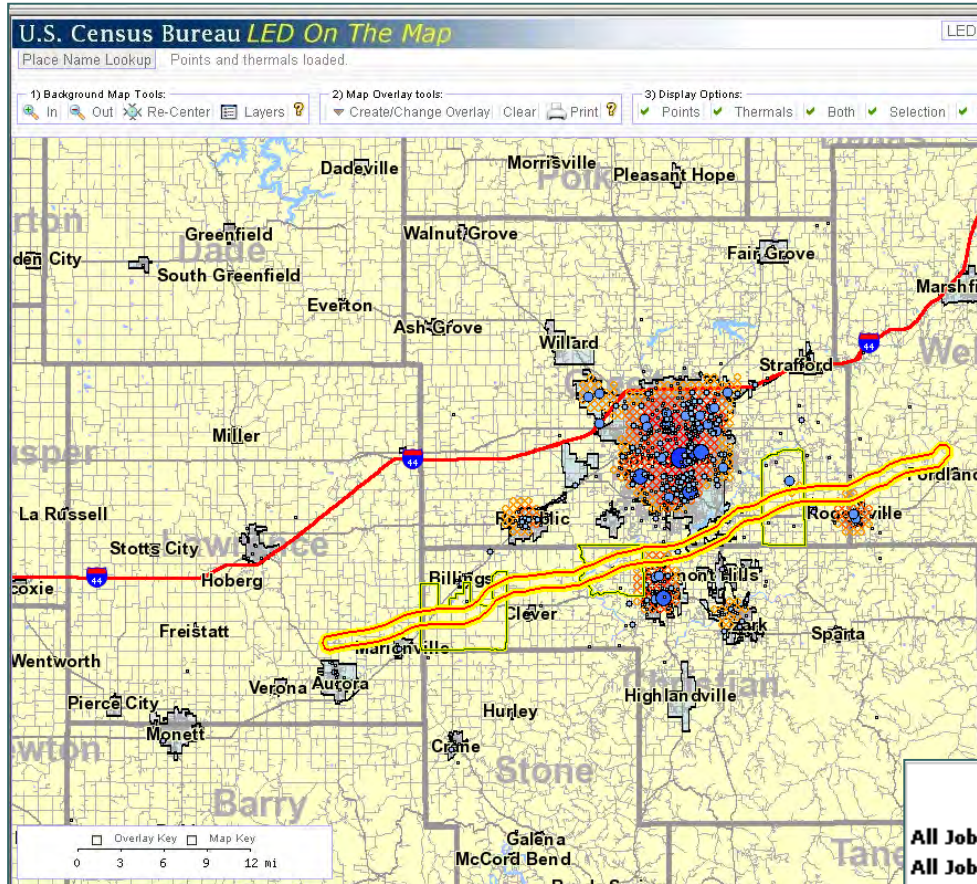| Workers by Industry Type (2-digit NAICS)	2003	
	Count	Share
* Agriculture, Forestry, Fishing and Hunting	5	0.3%
* Mining	1	0.1%
* Utilities	3	0.2%
* Construction	134	8.6%
* Manufacturing	195	12.5%
* Wholesale Trade	132	8.5%
* Retail Trade	230	14.8%

QWI Indicators - Private Sector Jobs	2003:Q2
* Employment (Beginning-of-2nd quarter)	248
* Employment, Stable Jobs	210
* Separations, Stable Jobs	24
* New Hires, Stable Jobs	21
* Firm Job Gain	38
* Firm Job Loss	21
* Employment (reference quarter)	326
* Average Monthly Earnings, Stable Jobs	$ 2,283
* Average Monthly Earnings Separations from Stable Jobs	$ 975
* Average Monthly Earnings, New Hires, Stable Jobs	$ 2,128

* Other Services (excluding Public Administration)	74	4.8%
* Public Administration	0	0%

Resident Held Jobs by Category	2003	
	Count	Share
All Jobs	1,943	100.0%
All Jobs (Private Sector Only)	1,644	84.6%
All Primary Jobs (Worker's highest paying job)	1,843	94.9%
All Primary Jobs (Private Sector Only) and Baseline Count of Workers	1,557	80.1%

Workers by Earnings Paid	2003	
	Count	Share
$1,200 per month or less	452	29%
$1,201 to $3,400 per month	674	43.3%
More then $3,400 per month	431	27.7%

Where do people who live in the path of the tornado *work*?
(Census Bureau's On the Map application, 2006)

Concluding remarks

- Vast opportunities to redesign survey processes and create new and richer quantitative measures and data products
- Need to build a stronger bridge between statisticians and geographic information scientists
- Methodological research
    - Develop methods for incorporating geospatial data into sampling and estimation
    - Develop methods for deriving meaningful quantitative measures from geospatial data objects, particularly over time
    - Integrate cognitive theories on variation in how humans work with spatial information in data collection
    - Quantify errors with geospatial data sources and geographic features recorded during survey data collection
    - Develop systems components: metadata models, geospatial data models, GIS software for integrating and manipulating quantitative research data

CSSM

Iowa State University

Statistical perspectives on spatial social science

Discussion

Sarah Nusser   (nusser@iastate.edu)

Center for Survey Statistics and Methodology

Department of Statistics

Iowa State University

Morris Hansen Lecture, WSS Nov 6, 2006