Changes in cultivated cropland soil in Iowa

Andreea L. Erciulescu

Iowa State University
Statistics Department

April, 2014

Outline

- Motivation and problem description
- Available data
- Data processing steps
- Results
- Conclusions and future work

National Resources Inventory (NRI)

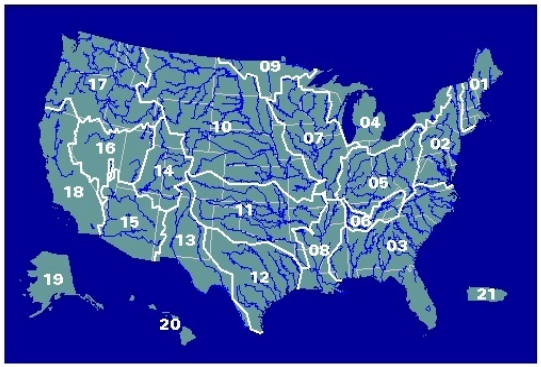
- annual survey conducted collaboratively by USDA NRCS (Natural Resources Conservation Services) and ISU Center for Survey Statistics and Methodology (CSSM)
- to provide status and trend estimates for natural resources on nonfederal lands in US
    - Example of such estimates are soil erosion estimates in relation to land characteristics and programs.

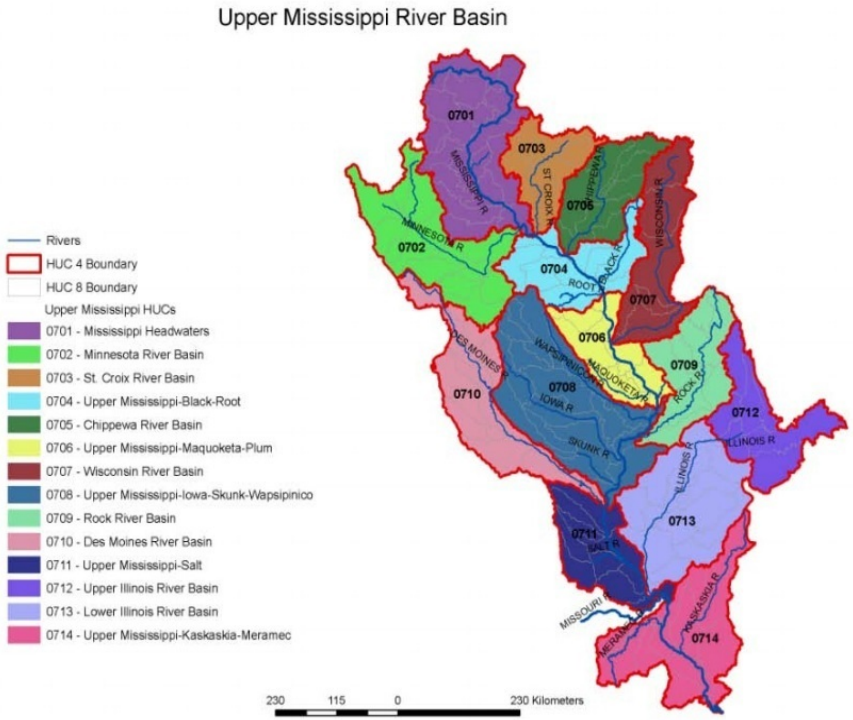
Conservation Effects Assessment Project (CEAP)

- series of surveys intended to quantify environmental effects of conservation prctices and programs by hydrologic unit codes (HUCs)
- CEAP sample is a subset of the NRI points classified as cultivated cropland
  - Farmer interviews about on-farm practices
  - Hydrologic, climate and soil databases
  - APEX model

United States territory divion into HUCs

Local concerns regarding the existance of nitrates in drinking water, particularly in Des Moines





Data - available

Crop Data Layer (CDL)
The CDL data is available at http://nassgeodata.gmu.edu/CropScape/ in the form of Tagged Image File (.tif) Format. We are interested in the state of Iowa data, available for the years of 2003-2007. The information consists of pixel counts and acreage values for different category of cropland data. Each of the category has an associated value (code), for example 1 stands for Corn and 5 stands for Soybean. A complete list of cat-
egory codes, class names and colors for the USDA NASS CDL is available at http://www.nass.usda.gov/research/Cropland/docs/CDL$_2$013$_c$*rosswalk*.*htm*.

Data - available

- Public Land Survey System (PLSS)
  The PLSS data can be found on
  http://http://www.geocommunicator.gov/GeoComm/lsis$_h$*ome*/*home*/*ir*
  in the form of shapefiles. Information is available at both state and county levels.
- GIS data on hydrologic basins
  The GIS data can be found at
  ftp://ftp.igsb.uiowa.edu/gis$_l$*ibrary*/*basins*/ in the form of shapefiles. Information is available for the entire Des Moines River basin.
- Atlas of historical countyu boundaries (AtlasHCB)
  The AtlasHCB data is available at
  http://publications.newberry.org/ahcbp/pages/Iowa.html in the form of shapefiles. Information os available for the entire state of Iowa.
- Census Topologically Integrated Geographic Encoding and Referencing database (Tigerweb)
  The Census Tigerweb data is available at
  http://tigerweb.geo.census.gov/tigerwebmain for both national and

Data processing - CDL

- download from http://nassgeodata.gmu.edu/CropScape/, years 2003-2007
- raster graphics images, spatial data structures that divide the US teritory into pixels that store crop information. This type of data are referred to as a grid, contrasted with vector data is used to represent points, lines, polygons. The dimensions of files are very large
- Raster package in R
  - uses sp package
  - S4 method
  - - the raster values from the files and to convert the cell numbers to coordinates and back

2003 CDL data - Raster object attributes

- read the values for the region of interest using cell numbers and coordinates (xy) in the extraction method from the cellFromXY

Data processing - GIS, PLSS, AtlasHCB data

- download and read in using maptool library in R
- extract polygon information from the shapefiles
- universal transverse mercador (UTM) and we need to convert it to the longitude-latitude, then to the CRS with the appropriate raster characteristics

Data processing - Tigerweb

- pull Iowa data and Hydrologic data from web using the XML library in R
- select the data on Des Moines River and 'create' the watershed region to get a significant number of points

Results

Table: Proportions of crop by category, by year

year	corn	soybean	...	...	...	....	...	...	...
2003	0.0111	0.0609	0.0335	0.0220	0.0171	0.0254	0.0566	0.0840	0.0964
2004	0.0292	0.0534	0.0442	0.0425	0.0568	0.0374	0.0442	0.0367	0.0448
2005	0.0313	0.0659	0.0512	0.0494	0.0729	0.0411	0.0521	0.0449	0.0627
2006	0.0292	0.0534	0.0442	0.0425	0.0568	0.0374	0.0442	0.0367	0.0448
2007	0.0313	0.0659	0.0512	0.0494	0.0729	0.0411	0.0521	0.0449	0.0627

Conclusions and future work

Challenges

- big data
- coordinates in different measurement system
- raster package uses sp package
- different sources of data, need to choose useful, realible and (hopefully) not big data

Future work

- use the CDL data and the sampled CEAP points (data not publicly available) to compute crop estimates
- shiny app ?

End

Thank you for your attention!