OSGeo Sample Data Set for research, development and education

Helena Mitasova, North Carolina State University
Markus Neteler, FBK-irst
Motivation

Replacement for 15 year old Spearfish Data

Software testing and development
- compare algorithms accuracy and performance,
- testing new tools, bug fixing

Education and training
- examples for manual pages
- tutorials, courses and workshops
- used in the 3rd edition GRASSBook
Why North Carolina

State has rapid population growth and strong environmental leadership: both stimulate robust mapping programs

- modern, state-of-the-art official data available for free
- FGDC metadata included
- great support from agencies, local - state - federal
- NC One map, WMS compliant web sites
- varied landscape: ocean to mountains, rural and cities
- local presence: better access and knowledge of data
Multiple scales

Hierarchy of scales: State->County->Field/City

State: 800x600km region, 500m resolution raster
County: 20x30 km region, 30-10m resolution raster, most data, scale comparable to Spearfish
Field/city/campus: 800x800m, 1-3m resolution raster
Structure of the data set

Terabytes of data available - creating small and useful data set(s) has been a major challenge

**GRASS LOCATION**: vector and raster maps in GRASS format, State Plane NAD83 (EPSG 3358)

**Data in external formats**: original GDAL/OGR supported formats and coordinate systems (geographic, UTM, State Plane, meters and feet) - teach and test format conversions and coordinate transformations

**Links to web sites, NC One map**: newer and/or higher resolution data, complete spatial extent

*Provide the data integrated in GRASS LOCATION in a different format and coordinate system? If yes, which format(s) suit the best for OSGeo community?*
Geospatial data types I

**Raster:**
- **discrete:** basins, geology
- **continuous:** DEMs, slope

**Vector:**
- polygons, lines, points with attributes
- streets, streams, firestations, lakes

---

GRASS GIS

Open Source Geospatial Foundation
Geospatial data types II

**Imagery:** LANDSAT, MODIS, orthophoto

**Vector:** point clouds

2002 LANDSAT

2001 Orthophoto

3D point cloud
Examples

Hundreds of derived maps can be created from the data set using geospatial analysis and modeling tools

- cost surface /shortest path example
- noise impact
- visibility analysis
- linear reference system
- network analysis
- generalization
- temperature analysis based on MODIS
- lidar terrain modeling and analysis
- watershed analysis and flow pattern
- precipitation modeling
Cost surface: accident

- convert street speed limit data into time (use 5km/hr for off-street areas),
- compute cost surface to reach site of an accident from any point in the area,
- find closest fire station, compute shortest path from it

Can be solved both with raster tools and vector network analysis tools
Noise impact analysis

Noise impact:
overlay highway buffers with developed areas
find schools and number of students impacted by noise
Point data with attributes

Schools: capacity (blue symbols), number of mobile units (yellow-red balls)
Viewshed analysis

- computed from bare ground lidar based elevation model, 30m resolution
- can be compared with the results from SRTM, higher resolution DEM and DSM
- can be overlayed with Points of Interest map
LRS: Linear Referencing System

Build LRS: linear referencing system for bus lines

Tasks:
- bus stop relocation
- distance to stop along the route

See real-time location of buses at: http://ncsu.transloc-inc.com/
Network analysis

Find shortest path between two hospitals based on cost (1/speed) and distance
Planimetry combines many layers: buildings, parking lots, fences, roads, ... Features can be extracted using SQL and represented using 3D vectors.
Google SOC project testing

Generalization:
- contours (thin the points to improve interpolation),
- streams, polygons

Daniel Bundala, student
Wolf Bergenheim, mentor

Reduce points from 40,000 to 3,000
interpolate 8-times faster, without artifacts
MODIS analysis: temperature

MODIS time series data analysis, 400 MB of data - one year, twice a day

GRASS GIS
Open Source Geospatial Foundation
Lidar: point clouds, DEMs, DSMs

- Multiple return point cloud
- 10m DEM derived from lidar
- Bare ground DEM and first return DSM

GRASS GIS
Open Source Geospatial Foundation
Point elevation data for interpolation:

- random elevations
- lidar-based profiles
- contour points

a: Voronoi polygons
b: IDW
c: linear contour-based
d: spline (RST)
Topographic analysis

- 1m res. lidar DSM
- 10m res. lidar DEM
- 30m SRTM
- 500m DEM

Topographic analysis from different DEMs and DSMs:
- slope
- aspect
- profile curvature
- tangential curvature

GRASS GIS
Open Source Geospatial Foundation
Flow analysis

- Single flow direction D8
- Dinf
- Multiple flow direction D8
- Bivariate continuity equation

SFD D8: extracted streams, watersheds

GRASS GIS
Open Source Geospatial Foundation
Sediment transport modeling

High resolution DEM, soils and land cover data:

process-based simulations

soil detachment sediment flow net erosion and deposition

GRASS GIS

Open Source Geospatial Foundation
Precipitation interpolation

Annual and monthly precipitation normals for 137 stations

Statewide 500m DEM

Bivariate RST interpolation
Trivariate RST interpolation adds impact of elevation

Point data with annual precipitation shown using size and color

Volume model used in trivariate interpolation

GRASS GIS

Open Source Geospatial Foundation
Future plans

Data set will evolve both in terms of its organization and data

- thematic mapsets will be added:
  3D urban (NCSU campus, Raleigh, Charlotte)
  3D multitemporal coast data,
  imagery,
  real time data: point and webcam (meteorology, bus locations, traffic, construction, monitoring)
- online browsing and analysis through pyWPS
- examples in the book, tutorials, manual pages, courses

Thanks to Julia Harrell, Doug Newcomb, and Silvia Terziotti for their continuing help with the OSGeo sample dataset
Download

Download web site
http://www.grassbook.org/data_menu3rd.php

Data set license:
Creative Commons Attribution-ShareAlike (CC-BY-SA)

OSGeo wiki web site - needs update

Summary list:

Vector data: administrative boundaries, census data, zip codes, fire stations, hospitals, roads and railroads, public schools and colleges, bus routes, points of interest, precipitation, hydrography maps, geodetic points, soils and geological maps, bare earth and multiple-return lidar.

Raster data: elevation (NED 3arc-sec, SRTM-V1 30m, lidar derived DEMs at 1m and 6m), slope, aspect, watershed basins, geology, and land use. The resolution of raster maps is 500m, 30m, 10m, and 1m.

Imagery data: 1m resolution orthophoto, several LANDSAT-TM5/7 scenes and a MODIS daily Land Surface Temperature (LST) time series.