GEOG 176A: Introduction to Geographic Information Systems

Lecture 08: Getting the Map into the Computer
(Obtaining Geospatial Data for a GIS)
(chapter 4)

Rui Zhu
Overview

● How to obtain geospatial data for a GIS?
  ○ Consume existing data
  ○ Create new data

● How to evaluate the accuracy of geospatial data?
Overview

- How to obtain geospatial data for a GIS?
  - Consume existing data
  - Create new data

- How to evaluate the accuracy of geospatial data?
Existing spatial data collections

- Traditionally, geospatial data were collected for each project, which **cost** most in a GIS project
- Cost: money, time and labor → Can we **reuse** existing data?
- **Data sharing**: one-time cost but reused many times in different projects
- Initial data sharing: **individual websites**
  - Hard to find
  - **Maintenance** cost
- The emergence of data portals (**Geoportals**)
  - One-stop style data access
  - Centralized management
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Geoportals - Data.gov

https://www.data.gov
Geoportals - INSPIRE

http://inspire-geoportal.ec.europa.eu/
Geoportals - The National Map (USGS)

https://nationalmap.gov/
National Map Viewer - LULC
National Map Viewer - DEM
Seamless download

- Identify a point or a polygon
- Server “cuts out” and mosaics data
- Direct download / Sends an email with limited time access to an ftp site
- Data are compressed and structured by layers
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Most geoportals are federal funded

In addition to Data.gov, USGS the National Map and INSPIRE geoportals, there are several other federal funded geospatial data sources:

- U.S. Bureau of the Census
- National Oceanic and Atmospheric Administration (NOAA)
- National Aeronautics and Space Administration (NASA)
U.S. Bureau of the Census - TIGER

https://www.census.gov/geo/maps-data/data/tiger.html

TIGER Products

- Topologically Integrated Geographic Encoding and Referencing
- TIGER products are spatial extracts from the Census Bureau’s MAF/TIGER database, containing features such as roads, railroads, rivers, as well as legal and statistical geographic areas. The Census Bureau offers several file types and an online mapping application. Our products are:
  - TIGERLine Shapefiles - New 2017 Shapefiles
  - TIGERLine Geodatabases
  - TIGERLine with Selected Demographic and Economic Data
  - Cartographic Boundary Shapefiles
  - KML - Cartographic Boundary Files
  - TIGERWeb

25 Years and Counting
- TIGER Story Map (Part 1)
- Happy 25th Anniversary TIGER

TIGER Data and Product FAQs

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
<th>File Format</th>
<th>Type of Data</th>
<th>Level of Detail</th>
<th>Descriptive Attributes</th>
<th>Vintages Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIGERLine Shapefiles</td>
<td>Most mapping projects this is our most comprehensive dataset.</td>
<td>Shapefiles (.shp) and database files (.dbf)</td>
<td>Boundaries, roads, address information, water features, and more</td>
<td>Full detail (not generalized)</td>
<td>Extensive</td>
<td>2000-2017, CD 113</td>
</tr>
<tr>
<td>TIGERLine Geodatabases</td>
<td>Useful for users needing national datasets or all major boundaries by state.</td>
<td>Geodatabase (.gdb)</td>
<td>Boundaries, roads, address information, water features, and more</td>
<td>Full detail (not generalized)</td>
<td>Limited</td>
<td>2000-2017</td>
</tr>
<tr>
<td>KML - Cartographic Boundary Files</td>
<td>Viewing data or creating maps using Google Earth, Google Maps, or other platforms that use KML</td>
<td>KML (.kmz)</td>
<td>Selected boundaries</td>
<td>Less detail (generalized)</td>
<td>Limited</td>
<td>2013-2017</td>
</tr>
<tr>
<td>TIGERWeb</td>
<td>Viewing spatial data online or streaming to your mapping application.</td>
<td>Interactive viewer, HTML data files, plus REST and WMS map services</td>
<td>Boundaries, roads, address information, water features, and more</td>
<td>Detailed</td>
<td>Extensive</td>
<td>Current, 2015 ACS, 2014 ACS, 2010</td>
</tr>
</tbody>
</table>
U.S. Bureau of the Census - Thematic Maps

Thematic Maps

Thematic maps are data maps of a specific subject or for a specific purpose. Statistical thematic maps include a variety of different map types such as choropleth or shaded maps, dot maps, proportional symbol maps, and isarithmic maps. Special purpose maps generally concentrate on a single theme, such as the Centers of the Population Centers of the U.S. between 1790 and 2000.

Business

Centers of Population

Maps of the mean and median centers of population from the 2010 Census are available on our Centers of Population page. The centers of population from Census 2000 can be found by following the link under Centers of Population from Previous Censuses.

Income and Poverty

Miscellaneous

Population Change

Population Density

Population Distribution

Population Estimates

Population Profiles

Total Population

Urban Areas

Veterans of the United States

[NOTE: This remote file contains labels. The remote file is available from Adobe. This symbol [+] indicates a link to a non-government web site. Our linking to these sites does not constitute an endorsement of any products, services or the information found on them. Once you link to another site you are subject to the policies of the new site.]
NOAA

https://www.ncdc.noaa.gov/data-access
NOAA - Satellite Data Access

https://www.ncdc.noaa.gov/data-access/satellite-data/satellite-data-access-datasets

Satellite Data Access by Datasets

Satellite data listed by parameter being observed. Access to long-term, high-quality products is supported by NOAA’s Climate Data Record (CDR) Program.

Atmospheric Data

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Spatial Information</th>
<th>Temporal Information</th>
<th>Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISCCP Cloud Data</td>
<td>The International Satellite Cloud Climatology Project (ISCCP) provides global cloud information at many resolutions (10 to 110 km) and time scales (3 hourly to monthly) derived from geostationary and polar orbiting satellite instruments.</td>
<td>Global; varying resolution</td>
<td>Global Product (GPP)</td>
<td>Access Information Email</td>
</tr>
<tr>
<td>SSMI-SSMIS Hydrological Products</td>
<td>Monthly and pentad averaged SSMI and SSMIS products include precipitation, cloud liquid water, total precipitable water, snow cover, and sea ice extent. These products are useful for evaluating the mean climate state, its interannual and seasonal variations, and the detection of anomalies associated with large-scale (e.g., ENSO, Arctic Oscillation) and regional climatic variations. A time series of the entire SSMI and SSMIS archive includes data from July 1987 to the present.</td>
<td>Global; 1 km and 2.5 km</td>
<td>Monthly and pentad; 1987-present</td>
<td>FTP</td>
</tr>
<tr>
<td>GOES Aerosol</td>
<td>Aerosol Optical Depth (AOD) data using GOES visible channel imagery. Available data include 4 km, CONUS and half-hourly, 06/2006-2018.</td>
<td>4 km, CONUS</td>
<td>half-hourly, 06/2006-2018</td>
<td>HAS</td>
</tr>
</tbody>
</table>

GEOG176A - Summer 2018
NASA - GDEM

https://asterweb.jpl.nasa.gov/gdem.asp

The Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA) jointly announced the release of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 2 (GDEM V2) on October 17, 2011.

The first version of the ASTER GDEM, released in June 2009, was generated using stereo-pair images collected by the ASTER instrument onboard Terra. ASTER GDEM coverage spans from 63 degrees north latitude to 63 degrees south, encompassing 96 percent of Earth's landmass.

The improved GDEM V2 adds 260,000 additional stereo-pairs, improving coverage and reducing the occurrence of artifacts. The refined production algorithm provides improved spatial resolution, increased horizontal and vertical accuracy, and superior water body coverage and detection. The ASTER GDEM V2 maintains the GeoTIFF format and the same gridding and tile structure as V1, with 30-meter postings and 1 x 1 degree tiles.

Version 2 shows significant improvements over the previous release. However, users are advised that the data contains anomalies and artifacts that will impede effectiveness for use in certain applications. The data are provided "as is," and neither NASA nor METI/Japan Space Systems (J-SpaceSystems) will be responsible for any damages resulting from use of the data.

As a contribution from METI and NASA to the Global Earth Observation System of Systems (GEOS). ASTER GDEM V2 data are available free of charge to users worldwide from the Land Processes Distributed Active Archive Center (LP DAAC).

The GDEM is available for download from NASA Earthdata and LP DAAC Global Data Explorer.

This ASTER product is available at no charge for any user pursuant to an agreement between METI and NASA. For more information about the GDEM, see the Validation Report: ASTER GDEM V2 Validation Summary Report.

Below are a couple of browse images derived from ASTER GDEM V1 globe map data sets. The full data set can be downloaded from the links above.
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
State and local agencies

Santa Barbara county:

https://www.countyofsb.org/pwd/surveyordownloads.sbc
State and local agencies

Santa Barbara County Surveyor Information System

https://secure.simplelayers.com/?do=start&project=939&application=CG3Viewer&embedded=1&query_url=0&context=2
State and local agencies

Santa Barbara Map Analysis and Printing System

https://www.santabarbaracaca.gov/services/maps/default.asp
State and local agencies

Los Angeles GeoHub

http://geohub.lacity.org/
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Map libraries

UCSB Map & Imagery Laboratory Collections

https://www.library.ucsb.edu/mil/collections
Map libraries

Aerial Photography Information

https://www.library.ucsb.edu/src/airphotos/aerial-photography-information
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g. Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Commercial data suppliers - Maps

Applied Geographic Solution

Maps.com

tomtom
**Commercial data suppliers - Remote sensing images**

**QuickBird Satellite Sensor (0.65m) - Decommissioned**

DigitalGlobe's QuickBird satellite sensor was successfully launched October 18, 2001 at Vandenberg Air Force Base, California, USA.

Using a state-of-the-art IKONOS sensor (PDF), QuickBird collected image data to 0.65m pixel resolution degree of detail. This satellite was an excellent source of environmental data useful for analyses of changes in land usage, agricultural and forest climates. QuickBird's imaging capabilities had applied to a host of industries, including oil and gas exploration and production (E&P), engineering and construction and environmental studies.

Sample Images

https://lta.cr.usgs.gov/UCDP

---

**WorldView-1 Satellite Sensor (0.46m)**

WorldView-1 satellite sensor was successfully launched from Vandenberg Air Force Base, California, U.S.A., at 11:33 Hrs Pacific Daylight Time (PDT) on September 18th, 2007. Watch video of WorldView-1 Satellite Launch.

The WorldView-1 offers a high-capacity, panchromatic imaging system which features 0.46m resolution imagery. Operating at an altitude of 464 kilometers, WorldView-1 satellite has an average revisit time of 1.7 days and is capable of collecting up to 750,000 square kilometers (290,000 square miles) per day of high-resolution imagery. The satellite is also equipped with state-of-the-art geo-location capabilities and exhibits stunning agility with rapid targeting and efficient in-track stereo collection.

Sample Images
Existing geospatial data sources

- Geoportals (e.g., data.gov)
- Federal agencies (e.g., USGS, NOAA, NASA)
- State and local agencies (e.g., Santa Barbara County)
- Map libraries (e.g., UCSB map library)
- Commercial data suppliers (e.g., Rand McNally, Thompson, TomTom, map.com, Navteq)
- Volunteered geographic information (e.g., OpenStreetMap)
Volunteered Geographic Information

- Data are provided by not only authorities but also the general public (i.e., citizens)
- Volunteered Geographic Information - VGI (Goodchild, 2007)
VGI - OpenStreetMap

https://www.openstreetmap.org
VGI - Location-based social networks (LBSN)

Create Magical Real-World Moments for Your Users

Create Magical Real-World Moments for Your Users

Filtering Tweets by location

https://developer.foursquare.com/
https://developer.yahoo.com/developers

GEOG176A - Summer 2018
GeoPrivacy

Although there are more and more data becoming available to public, the privacy of these geographic information is a huge issue!
Overview

- How to obtain geospatial data for a GIS?
  - Consume existing data
  - Create new data

- How to evaluate the accuracy of geospatial data?
How to create new geospatial data?

- To be captured from paper maps (geocoding)
  - Digitizing
  - Scanning
- To be collected from field work
How to create new geospatial data?

- To be captured from paper maps (geocoding)
  - Digitizing
  - Scanning
- To be collected from field work
Digitizing

- Captures spatial data by \textit{tracking} lines from a paper map using hand
- Uses an \textit{electronically-sensitive} tablet
- Result is \textit{a string of points with (x, y) values} \rightarrow vector data
Digitizing tablet

1. Digitizer cursor transmits a pulse from an electromagnetic coil under the view lens.
2. Pulse is picked up by nearest grid wires under tablet surface.
3. Result is sent to computer after conversion to x and y units.
Some common digitizing errors

- Slivers
- Duplicate lines
- Duplicate nodes
- Unended lines
- Gap
- ...

GAP AND SLIVER ERRORS IN DIGITIZED POLYGONS
How to create new geospatial data?

● To be captured from paper maps (geocoding)
  ○ Digitizing
  ○ Scanning

● To be collected from field work
Scanning

- Scan a paper map into an image
- Sanning result is a grid of pixels --> raster data
Scanning examples

DPI: dots per inch

--> note the impact of changing dpi
On-screen (heads-up) digitizing

- Sometimes it is also called "digitizing" (should be differentiated from the previous "digitizing")
  - Already have data in digital format, but they are raster
  - Converts raster images into vector data
  - Can be based on scanned maps or remote sensing images

Use the raster data as background, and then draw points, lines and polygons

Automatic Raster-to Vector
How to create new geospatial data?

- To be captured from paper maps
  - Digitizing
  - Scanning
- To be collected from field work
Field data collection
GPS data collection using hand-held GPS/smartphones
Geocoding

- Geocoding is the conversion of geospatial information into digital form
  - E.g., digitizing and scanning
- Another example of geocoding --> address matching
Geocoding - Address matching

The process:
Flat file of addresses → TIGER line file → address locator → points with geographic coordinates and attributes
Geocoding - Address matching

https://google-developers.appspot.com/maps/documentation/utils/geocoder/
Data errors / uncertainties

● Geospatial data can contain errors

● Errors can be:
  ○ in the geometries, locations, attributes values, and semantics
  ○ random or systematic

● Sources of errors
  ○ Malfunction of measurement devices
  ○ Data transformation errors
  ○ ...

● It is the responsibility of data producers to reduce data errors, but the GIS users must understand errors as well.

● Data errors can affect all other spatial operations in GIS
To reduce data errors

- Use **topology** for vector data
- Generate relevant **metadata**
  - record the accuracy
  - record provenance (i.e., source)
Example of data errors

Occasionally TIGER Data are Inaccurate or Imprecise