

Session 4:

Data Formats and

Open-Source GIS

Randy Bucciarelli
randobucci@gmail.com

Outline: Data Formats For GIS

- Introduction
- Tabular formats
- Vector formats
- Raster formats
- Open-Source GIS
- Demonstration
- Project

Class Schedule

Monday	Tuesday	Wednesday	Thursday	Friday
<p>08/05/19</p> <p>Introduction to Geographical Information Systems</p> <p>10:45 am–12:15 am</p>	<p>08/06/19</p> <p>Cartography and Spatial Data Display</p> <p>8:30am – 11:00pm</p>	<p>08/07/19</p> <p>Querying Data for Spatial & Attribute Selections</p> <p>8:30am – 11:00pm</p>	<p>08/08/19</p> <p>Data Formats and Open-Source GIS</p> <p>8:30am – 11:00pm</p>	<p>08/09/19</p> <p>Map Projections and Coordinate Systems</p> <p>8:30am – 11:00pm</p>
<p>08/12/19</p> <p>Spatial Analysis Tool</p> <p>8:30am – 11:00pm</p>	<p>08/13/19</p> <p>Raster and Terrain Analysis</p> <p>8:30 am – 10:00 am</p> <p>Scripps Institution of Oceanography</p> <p>1:00pm – 4:00pm</p>	<p>08/14/19</p> <p>Image Analysis</p> <p>8:30am – 11:00pm</p>	<p>08/15/19</p> <p>Editing Spatial Data and Geocoding</p> <p>8:30am – 11:00pm</p>	<p>08/16/19</p> <p>Web Mapping/ Wrap up</p> <p>8:30am – 11:30am</p>

Introduction

- Data formats for GIS
 - Commercial formats
 - Standardized formats
- You will use many formats:
 - Tabular
 - Vector
 - Raster
 - TIN
 - Network

Available data formats

- Commercial/Proprietary formats
 - ESRI
 - CAD
- Standardized/Open source formats
 - GEOJSON
 - KML
 - TIFF/JPEG (raster)
 - NetCDF

Tabular Data Formats or “Flat Files”

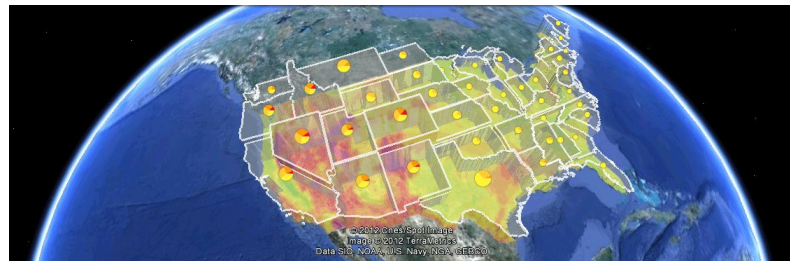
- Excel spreadsheets
- Ascii delimited text:
 - CSV or TSV
- XML Tag based
 - KML
 - GEOJSON
- Spaghetti Data Model

Flat File Example

Name	Group #	Occupation
Watts	1000	Model
Shelton	1000	Chef
Weber	1000	Chef
Tubbs	1001	Musician
Jones	1001	Musician
Carson	1001	Librarian

KML and KMZ Files

- Google Earth Format
- [Keyhole Markup Language](#)
- KML files can display:
 - Points, lines, polygons
 - Images
 - 3-D models
- KMZ is zipped up KML file
 - Contains symbology, legend, etc



```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://earth.google.com/kml/2.1">
  <NetworkLink>
    <name>Real-time Earthquakes</name>
    <open>1</open>
    <LookAt>
      <longitude>-104.1025</longitude>
      <latitude>50.4913</latitude>
      <altitude>0</altitude>
      <range>8384245</range>
      <tilt>0</tilt>
      <heading>-0.2564</heading>
    </LookAt>
    <FlyToView>1</FlyToView>
    <Url>
      <href>http://services.google.com/earth/kmz/r
ealtime_earthquakes.kmz</href>
    </Url>
  </NetworkLink>
</kml>
```

GeoJSON

- Extension of JSON (Javascript Object Notation)
- Used for sending, receiving, and storing data
- Wraps data in objects:
 - Name/Value pairs
 - Points, polygons, lines
- Example: [Plenario](#)



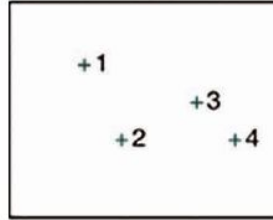
GeoJSON is a format for encoding a variety of geographic data structures.

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [125.6, 10.1]
  },
  "properties": {
    "name": "Dinagat Islands"
  }
}
```

GeoJSON supports the following geometry types: Point, LineString, Polygon, MultiPoint, MultiLineString, and MultiPolygon. Geometric objects with additional properties are Feature objects. Sets of features are contained by FeatureCollection objects.

Spaghetti Data Model

Points

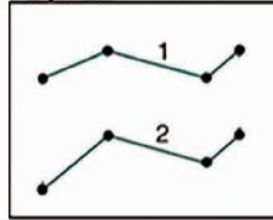


Point number

(x,y) coordinates

1	(2,4)
2	(3,2)
3	(5,3)
4	(6,2)

Polylines

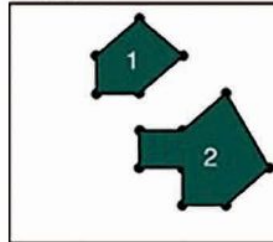


Polyline number

(x,y) coordinates

1	(1,5) (3,6) (6,5) (7,6)
2	(1,1) (3,3) (6,2) (7,3)

Polygons

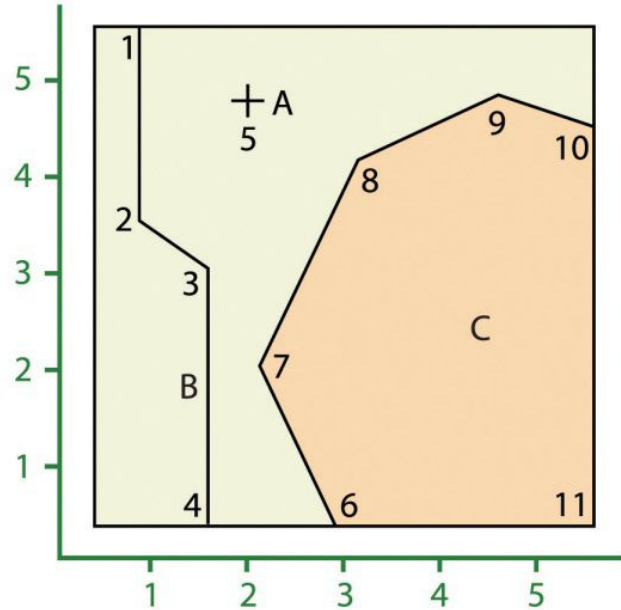


Polygon number

(x,y) coordinates

1	(2,4) (2,5) (3,6) (4,5) (3,4) (2,4)
2	(3,2) (3,3) (4,3) (5,4) (6,2) (5,1) (4,1) (4,2) (3,2)

Spaghetti Data Model



A,1 (identifier of the point and the number of vertex)

2, 4.7

B, 4 (identifier of the line and number of vertex)

1, 5.6

1, 3.5

1.6, 3

1.6, 0.4

C, 6 (identifier of polygon and number of vertex)

3, 0.4 (coordinates of the first vertex)

2.2, 2

3, 4.3

4.8, 4.8

5.6, 4.4

5.6, 0.4

3, 0.4 (coordinates of the first vertex again)

Vector Formats

- Esri Shapefiles
 - Most common and popular format
- Esri Geodatabases
 - Does all file management work for you
- US Census Bureau TIGER/Line
 - Topologically Integrated Geographic Encoding and Referencing system
- AutoCAD DXF
 - Drawing Exchange Format

ESRI Shapefile

- Stores geometry and attributes of vector layers
- Made up of several files
- Field names limited to 10 characters
- White paper released to public

Shapefile File Types

File Extension	Purpose
SHP*	Feature geometry
SHX*	Index format for the feature geometry
DBF*	Feature attribute information in dBASE IV format
PRJ	Projection information
SBN and SBX	Spatial index of the features
FBN and FBX	Read-only spatial index of the features
AIN and AIH	Attribute information for active fields in the table
IXS	Geocoding index for read-write shapefiles
MXS	Geocoding index for read-write shapefiles with ODB format
ATX	Attribute index used in ArcGIS 8 and later
SHP.XML	Metadata in XML format
CPG	Code page specifications for identifying character encoding
* Indicates mandatory files	

ESRI Geodatabase

- Hybrid file format:
 - Vector and Raster
- Collection of datasets within single database
- Native data structure for ArcGIS
- Seamlessly Manages different file formats

Raster File Formats

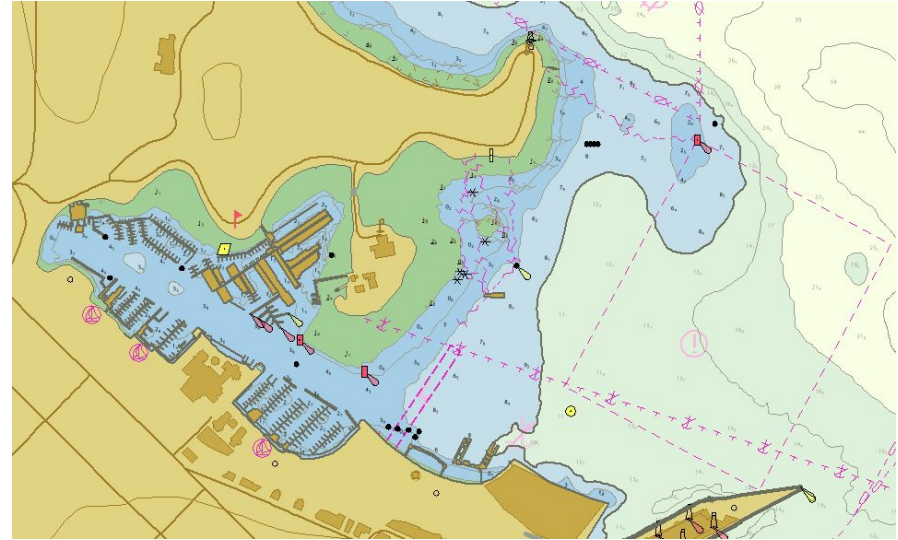
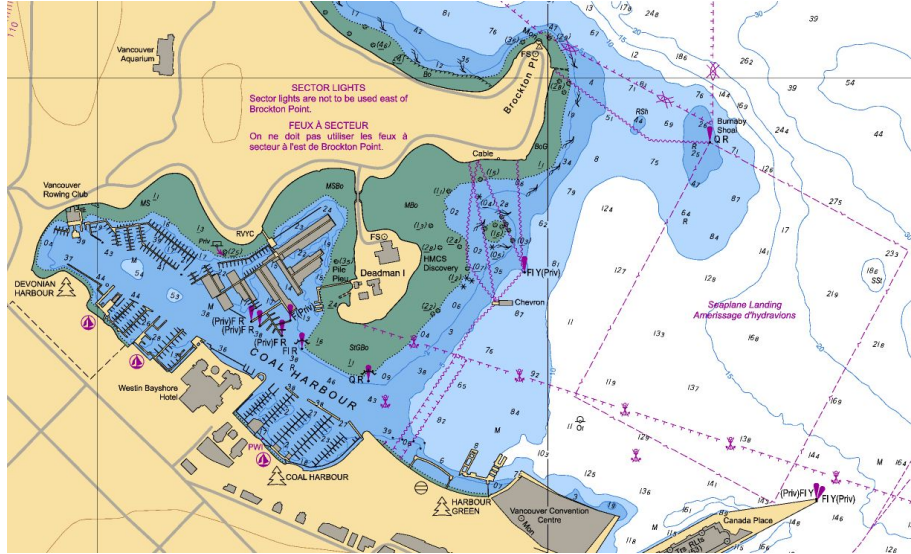
- Esri Grid
- Georeferenced Images:
 - JPEG, TIFF, and PNG
 - MrSID (Multiresolution Seamless Image Database)
 - DRG (Digital Raster Graphic)
- USGS DEM (Digital Elevation Model)
- GRIB
- NetCDF

Raster to Vector Conversion

Digital Raster Graphic

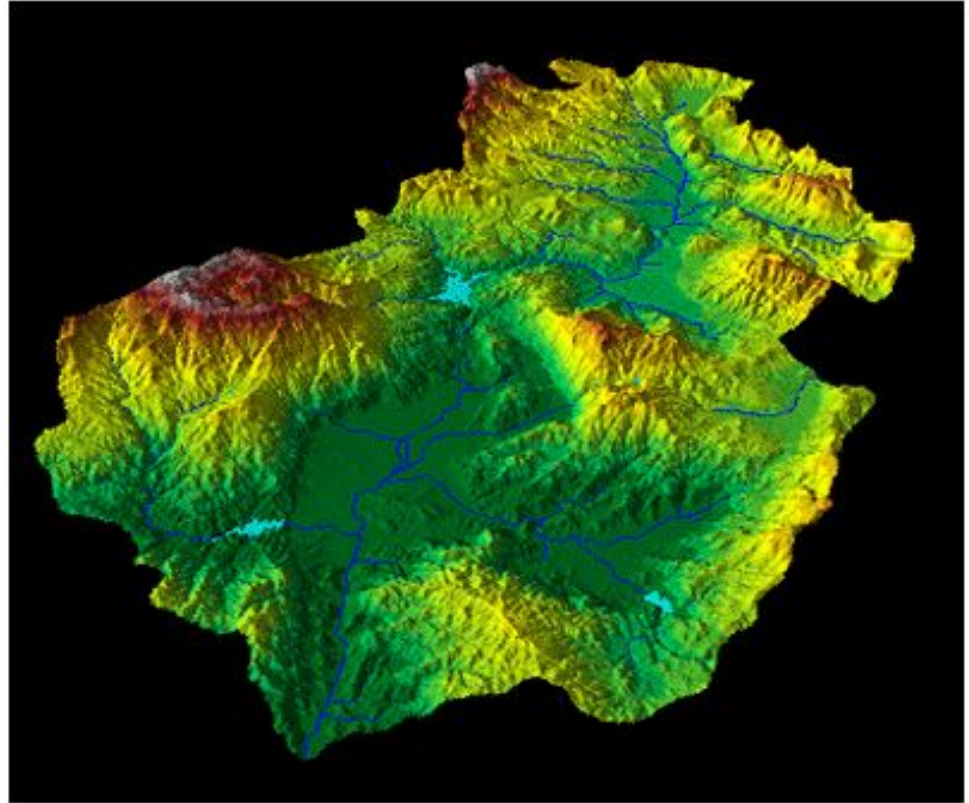


Esri Shapefile



Digital Elevation Models

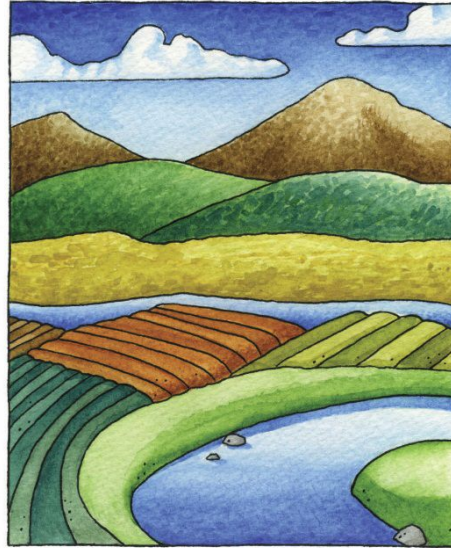
- Raster representation of the earth surface
- Cells contain continuous elevation values
- Accuracy determined by raster resolution



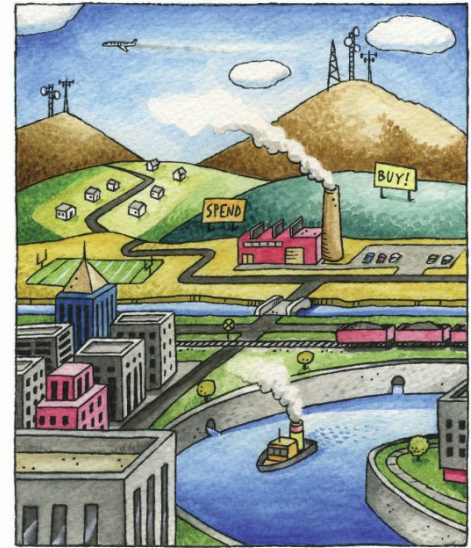
Digital Elevation Models

- Digital Terrain Models (DTM)
 - Bare earth model
- Digital Surface Model (DSM)
 - Landscape features (e.g. buildings, trees)

DSM



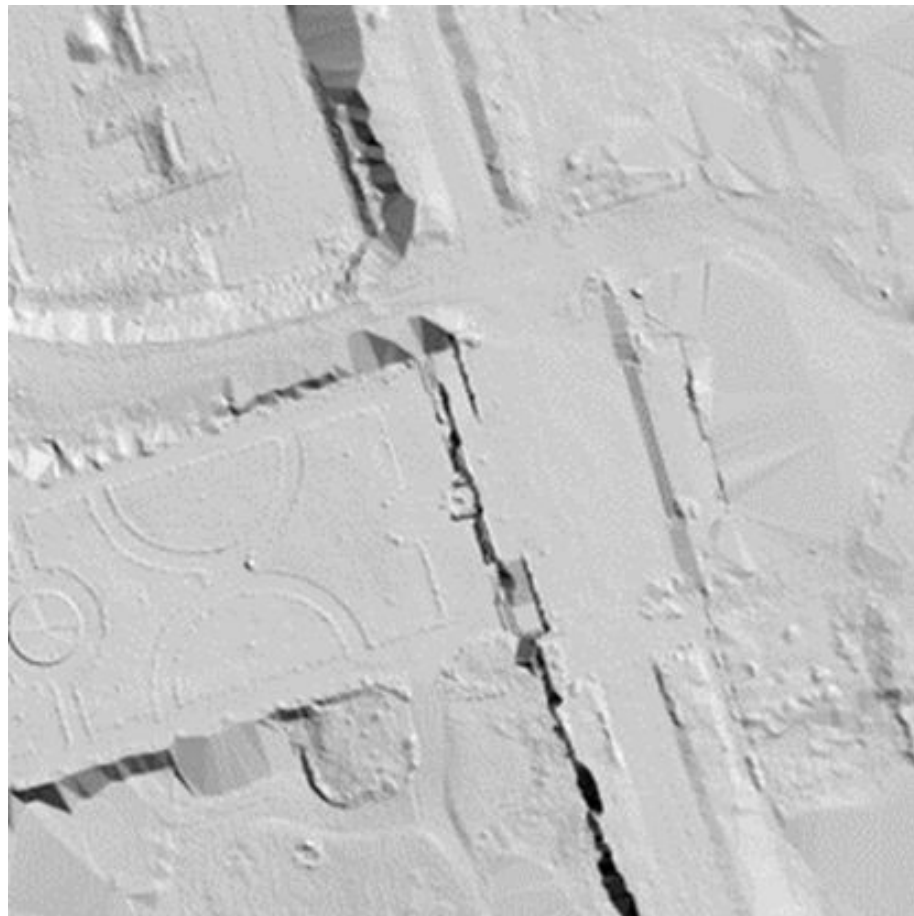
DTM



Digital Surface Model



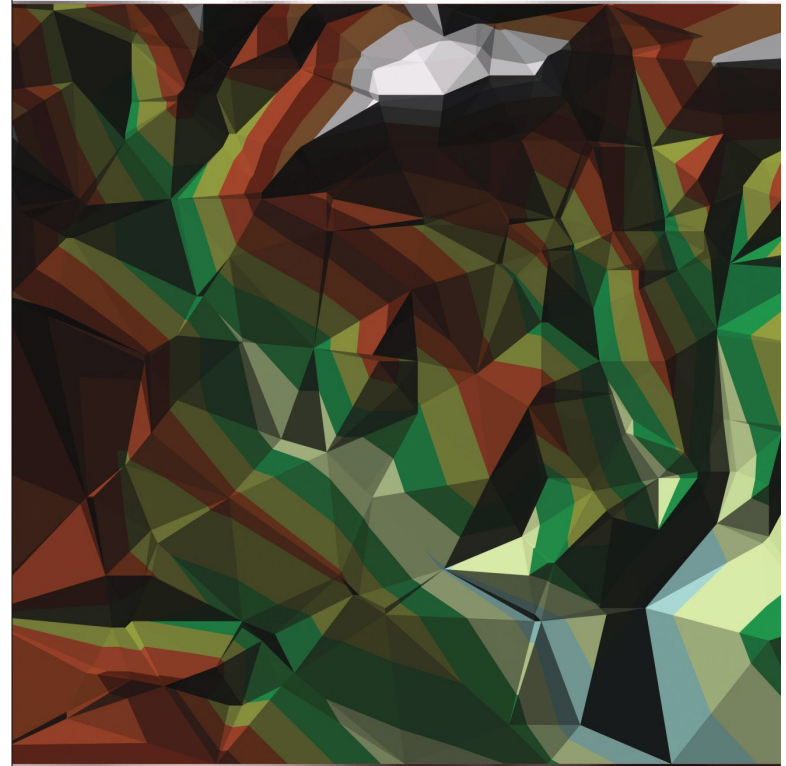
Digital Terrain Model



TIN - Triangulated Irregular Network

TIN Example

- Structure to represent surface elevations
- Non-overlapping triangles
- Derive elevation, slope, aspect, and surface area
- Advantages over rasters
 - Smaller file sizes
 - Delineate flood plains



Source: Essentials of GIS (Campbell, 2011)

Outline: Data Formats For GIS

- Introduction
- Tabular formats
- Vector formats
- Raster formats
- Open-Source GIS
- Demonstration
- Project

GDAL - Geospatial Data Abstraction Library

- Open source GIS tool
 - Library for reading and writing geospatial formats
 - “Swiss Army Knife” does it all: raster and vector
 - Check out [website](#) or [wiki](#)
 - Called from programming languages
 - Part of OSGeo



OSGeo - [Open Source Geospatial Foundation](#)

- Open Data
 - Free and open to geospatial data
- Open Education
 - Creation of teaching materials
- Open Science
 - Sharing data used for research
- Open Standards
 - Promote interoperability between applications, organizations, and research fields



OSGeo - Desktop Mapping Solutions



QGIS: The leading Free and Open Source Desktop GIS.



Marble: Desktop globe; pan around



gvSIG Desktop: Powerful and user-friendly



GRASSGIS

GRASS: Geographic Resource Analysis Support Software

OSGeo - Mobile Mapping Solutions



OpenLayers: Makes it easy to put a dynamic map on the web



GC2/Vidi: Platform for building spatial data and web GIS



Leaflet: Javascript library for mobile-friendly interactive maps



Geopaparazzi: Tool for very fast qualitative engineering GIS