GIS Project Issues

Charlene Nielsen, MSc
charlene.nielsen@ualberta.ca
492-9397
B414 BioSci Building

www.biology.ualberta.ca/facilities/gis

Introduction to GIS in Ecology

- What a GIS is not
- What is a GIS?
- GIS is…
- …linking databases to maps
- Types of spatial data
- Data issues in a GIS
- Conducting a GIS analysis
- Common types of spatial analyses
- When (not) to use GIS in ecology
- Advice on using GIS
- Learning options
- Intro to the basics:
  - GIS software
  - File management skills
  - Symbolizing
  - Classifying data

2

What a GIS is not

GPS

Global Positioning System

GPS data can be used in GIS analyses

Static Map

Digital/paper map is an "input" or "product" of GIS

A way to visualize output from GIS analyses

Software

Functions and tools needed to store, analyze, and display geographic information

Requires hardware, data, and personnel in complete system

Database

Set of tables containing data that can be accessed or reassembled in many different ways

Requires link to spatial data

3

What is a GIS?

G

Geographic

Earth Description

3 "W's" of Geography

what is where?

why is it there?

why do you care?

80% of all data has a spatial component

I

Information

Turn data into information

Location data (how many, what kind, where)

Scale of data (local to global)

Data presentation (words, graphs, tables, maps)

S

System

Personnel, Hardware, Software, Data, and Analytical Methods

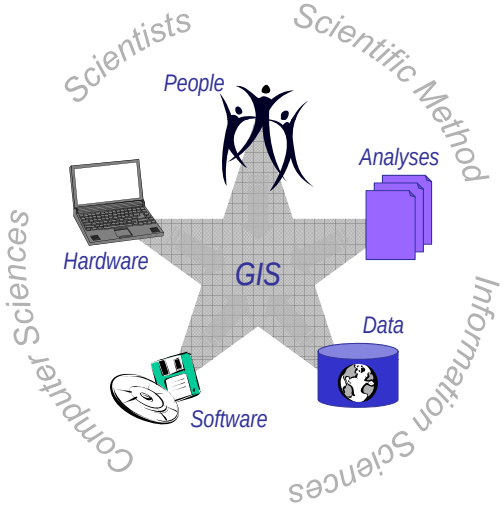
Organized to automate, manage, and deliver information through geographic presentation

4

GIS is…

… a method to visualize, analyze, and display spatial data

… the process of combining layers of information about a place to give you a better understanding of that place

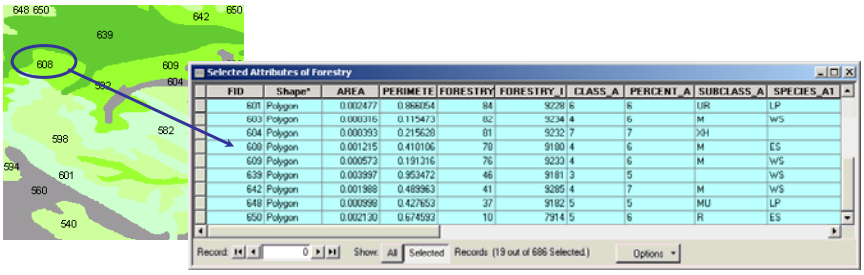


5

…linking databases to maps

A GIS is essentially a database that understands geometry

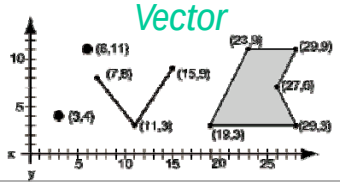
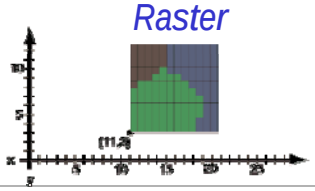
Tables of data can be linked together and joined to vector or raster spatial data if they share a common attribute or ID



FID	Shape*	AREA	PERIMETE	FORESTRY	FORESTRY_I	CLASS_A	PERCENT_A	SUBCLASS_A	SPECIES_A1
601	Polygon	0.002477	0.866054	84	9228	6	6	UR	LP
603	Polygon	0.000316	0.115473	82	9234	4	6	M	WS
604	Polygon	0.000393	0.215628	81	9232	7	7	XH	
608	Polygon	0.001215	0.410106	78	9180	4	6	M	ES
609	Polygon	0.000573	0.191316	76	9233	4	6	M	WS
639	Polygon	0.003997	0.953472	46	9181	3	5		WS
642	Polygon	0.001988	0.489963	41	9285	4	7	M	WS
648	Polygon	0.000998	0.427653	37	9182	5	5	MU	LP
650	Polygon	0.002130	0.674593	10	7914	5	6	R	ES

6

Types of spatial data

Data Model	Description
Vector 	Models discrete features as **points** (coordinates), **lines** (arcs), and **polygons** (areas) with precise boundaries and shapes with attributes Example file types – vector coverages, shapefiles, CAD drawings
Raster 	Models continuous phenomena in a surface divided into a **regular grid of cells** (pixels) each having an associated attribute value Example file types – grid coverages, satellite imagery, digital photos

7

Data issues in GIS

DATA SOURCES

MAP PROJECTIONS

ATTRIBUTES

METADATA

Investigate and acquire the various sources of data that meet your analysis criteria.

- Digitized and scanned maps
- Databases (tabular data)
- GPS
- Remote sensing/airphotos
- Field sampling of attributes

Consider:

- What you want to do with the data
- Data sharing agreements
- Latest data available
- Geographic extent, scale, etc.

8

Data issues in GIS

DATA SOURCES

MAP PROJECTIONS

ATTRIBUTES

METADATA

For all data layers to match up correctly when drawing them on top of each other or combining them to see relationships, they must be in the same:

- Map Projection – mathematical translation of locations on the globe onto the flat surface of a map
- Coordinate System – specified units used to locate features in two-dimensional space and the origin point of those units

9

Data issues in GIS

DATA SOURCES

MAP PROJECTIONS

ATTRIBUTES

METADATA

Each geographic feature has attributes - particular types or categories of information - that identify, describe, or represent an associated magnitude.

GIS analysis depends partly on the type of attributes:

- Categories = groups of similar things
- Ranks = ordered features
- Counts = actual total number
- Amounts = measurable quantity
- Ratios = proportions and densities

10

Metadata!

… the all-important ancillary information associated with a GIS data layer or coverage that characterizes the data set content, quality, condition and other characteristics:

- date of production and creator's name
- data type: raster or vector
- coordinate system, projection, datum
- subject content and limitations
- attributes (database fields) and data dictionary
- scale: minimum map unit/cell size resolution
- ...and much, much, more as necessary

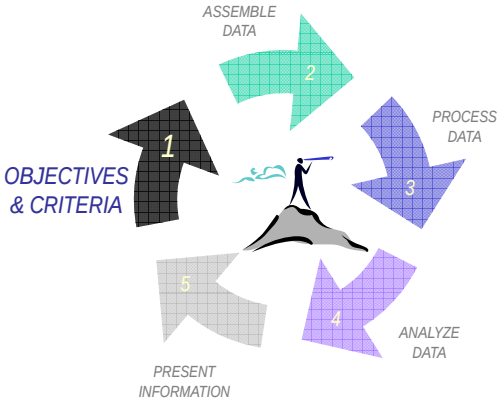
… information that describes, or documents, a geographic dataset to facilitate the intelligent use of it

11

Conducting a GIS analysis - 1

Objectives & Criteria

- Frame your question and define your criteria
- Specify how and who will use your analysis
- Determine the geographic boundaries for the region of interest
- Select the appropriate map scale (resolution, precision, accuracy)

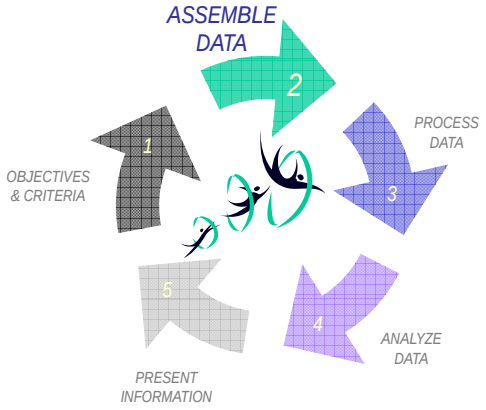


12

Conducting a GIS analysis - 2

Assemble Data

- Acquire existing data
- Create new data (digitize/field work)
- Check for errors and omissions
- Licensing and data sharing agreements
- Understand metadata

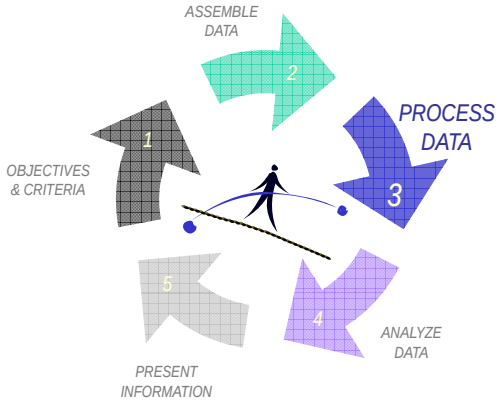


13

Conducting a GIS analysis - 3

Process Data

- Import/convert files
- Define coordinate systems and match projections
- Extract/modify existing features
- Update attribute information
- Generate symbology and explore data
- Document metadata

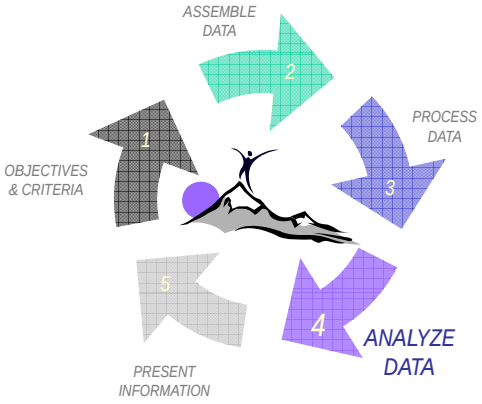


14

Conducting a GIS analysis - 4

Analyze Data

- Often, more than one method to get the information you need
- Select by location or by attribute, create buffers, overlay layers
- Do calculations, apply complex modelling
- See "Common types of spatial analyses"

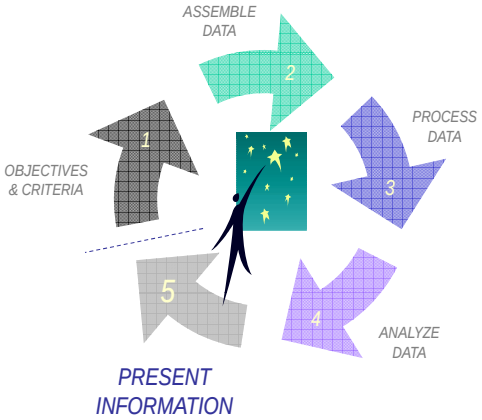


15

Conducting a GIS analysis - 5

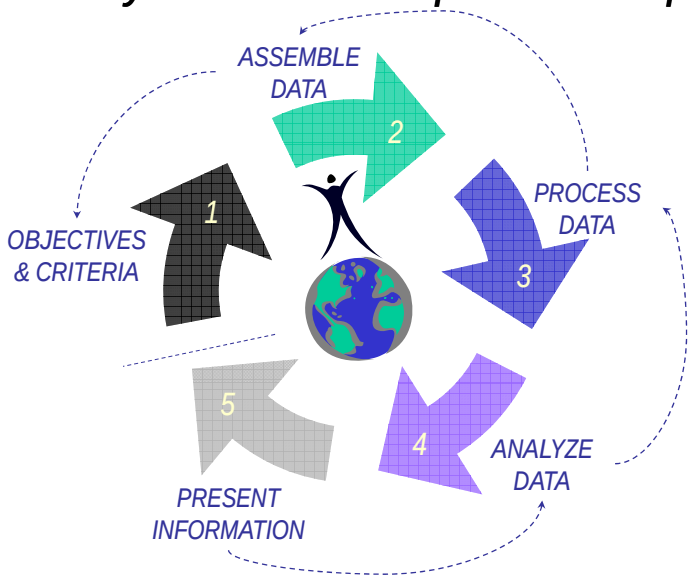
Present Information

- Interpret the output
- Decide whether the information is valid or useful, or if you should revisit previous steps in your analysis
- Design map layout
- Create graphs
- Export tables
- Write reports



16

You may need to revisit previous steps!



17

Common types of spatial analyses

Where things are

- Find places that have the features you are looking for
- Find patterns by looking at the distribution of features

The most and the least

- Find places that meet certain criteria and take action
- To see the relationships between places

Density

- Measure the number of features using a uniform areal unit, such as hectares or square meters, to clearly see the distribution

What's inside

- Monitor what is happening within a specific area

What's nearby

- Find out what is occurring within a set distance of a feature

Change

- To see where and how things move over a period of time
- Anticipate future need
- To see the impact before and after an action or event

18

When (not) to use GIS in ecology

PROs:

- *Combine, measure, and analyze data reliably and consistently*
- *Alternative sources of spatial data*
- *Facilitates broad-scale study of relatively inaccessible areas*

Alternatives to GIS include:

- *Databases*
- *Digital cartography*
- *Spreadsheets*
- *Image analysis*
- *Statistical packages, etc.*

CONs:

- *Lack of digital data*
- *Lack of time for data collection and entry*
- *Lack of experience and familiarity with the software*
- *False precision (obscuring sources of error)*
- *A technology-led approach*
- *Over-investment in data irrelevant to decision making*

19

Advice on using GIS

- *Keep it **simple***
- *Use **existing data** where possible*
- *Plan ahead and conceptualize*
- *Document the methods and sources of data used*
- *Check results to determine if output is valid/useful*
- *Consult with experienced GIS users on database management, data needs, GIS procedures*
- *Gain experience and familiarity with the software AND the concepts*

20

Learning options

UofA credit courses

- *BIOL 471/571 Landscape Ecology*
- *Other Departments have various levels of excellent GIS instruction*
  - *http://www.biology.ualberta.ca/facilities/gis/?Page=769*

Informal education

- *ESRI Virtual Campus: **http://training.esri.com***
  - *www.biology.ualberta.ca/facilities/gis/index.php?Page=484#virtualcampus*
- *ESRI ArcGIS user manuals have tutorial chapters*
- *GIS Short Courses in Ecology*
  - *Free, non-credit, hands-on demonstrations of GIS tools and applications*
  - *Currently 9 courses, with more being developed*

21

Intro to the basics…

- *GIS software*
- *File management skills*
- *Symbolizing*
- *Classifying*

*Always **keep a GIS notebook** beside your computer to jot down:*

notes on sources (e.g. website) of your data ~ informal metadata ~ details on any processing you perform ~ information on useful help files ~ step-by-step instructions on the analyses you execute to answer your research questions ~ etc.

This will fast become an invaluable and personal "GIS bible" for you!

22

GIS software

The software used in the short courses (B118) and installed in the GIS Research Lab (B418) is ArcGIS by Environmental Systems Research Institute (ESRI)

~$185 (prairiessales@esricanada.com) ESRI Canada

~$195 (gordana.brouilette@ualberta.ca) UofA AICT

One-year student license

ArcView and all extensions for ArcGIS

www.esri.com

23

File management skills

GIS file formats

www.grime.net/gistools/a-z.htm

Bio_print and Shared_GIS servers

Managing GIS data files (storing, transferring, naming)

- ~~*Windows Explorer*~~
- *ArcCatalog*

No spaces!

Zip GIS data files prior to e-mailing

www.zipgenius.it

Database practices (MS Excel vs. MS Access):

http://www.microsoft.com/education/howto.mspx

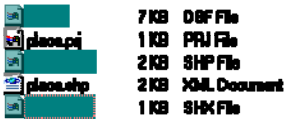
24

GIS file formats

Shapefile

A **vector** data storage format that stores the shape and location (*.shp), attributes (*.dbf), and lookup index (*.shx) of geographic features in a set of related files having the same prefix that must be kept together in the same directory

WINDOWS EXPLORER



Coverage

A **vector OR raster** "folder-based" data storage format that stores geographic features (as arcs, nodes, polygons, and label points OR grid cells) in a self-named folder and the associated attributes and spatial relationships in an info folder

WINDOWS EXPLORER

25

GIS file formats

E00

An interchange file format that is used to transfer **vector OR raster** coverages between computers (*. E00)

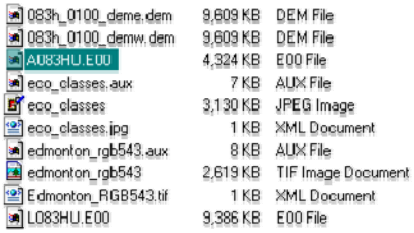
Requires importing using a specific application or utility with your GIS software

Image

A **raster** representation or description of a scene usually produced by an optical or electronic device; e.g. satellite data, digital aerial photographs, and scanned maps

Many file formats – TIFF, JPEG, BIL, DEM, IMG, SID, etc. – some require importing using a special extension

WINDOWS EXPLORER



26

GIS file formats

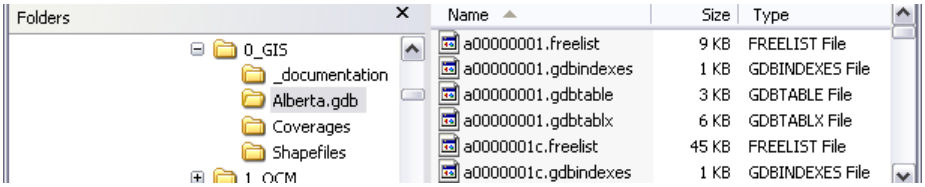
Geodatabase

A "container" that stores a collection of datasets as a folder with a name ending in .gdb (file geodatabase) or .mdb (personal geodatabase).

Vector datasets are stored as feature classes. Other objects, such as **raster** datasets and **tables**, are also stored in the geodatabase.

It is the recommended native data format for ArcGIS stored and managed in a file system folder .

WINDOWS EXPLORER



27

References

Crosier, Scott, Bob Booth, Kathy Dalton, Andy Mitchell, and Kristin Clark. 2004. Getting Started with ArcGIS. Environmental Systems Research Institute, Inc.

ESRI. 2004. GIS.com - Your Internet Guide to GIS (Geographic Information Systems). ONLINE: www.gis.com

Longley, Paul A., Michael F. Goodchild, David J. Maguire,and David W. Rhind. 2001. Geographic Information Systems and Science. John Wiley & Sons, Ltd. Chichester UK.

Mitchell, Andy. 1999. The ESRI Guide to GIS Analysis. Volume 1: Geographic Patterns and Relationships. Environmental Systems Research Institute, Inc.

Wadsworth, Richard and Jo Treweek. 1999. Geographical Information Systems for Ecology: An Introduction. Addison Wesley Longman Ltd.

28

Guidelines for symbolizing your data

Choice and arrangement of symbols are key to making your map easy to understand and depends on type of map:

- **Reference map** = show location of features; e.g. atlas, topographic map, study area map in a publication
- **Thematic map** = show structure and distribution of phenomena; e.g. population, climate, densities of organisms

Effective symbols take advantage of common associations that people make:

- **Color** = blue for water and green for vegetation
- **Symbol size** = thick line for a busier or more important road than one drawn with a thinner line

29

Guidelines for symbolizing your data

You may symbolize all of a layer's features in two ways:

- With the same symbol
  - Distinguishes features in different layers; e.g. GPS vs telemetry locations
- Based on an attribute
  - Indicate different types of features in the same layer; e.g. vegetation areas such as forest, grassland, and water

The type of symbology you use to create a thematic map depends on whether an attribute's values are:

- Categories
  - e.g. ecoregion, wetland type, habitat rating
- Quantities
  - e.g. population, number of observations, length

30

Classifying data

- Feature quantities are typically represented on a map by creating **classes** (groups of features with similar numeric values) and assigning a different symbol to each class
  - Commonly symbolized as **graduated size** and **graduated color** to aid in visualizing geographic distribution patterns
- When symbolizing features based on quantities **you decide**:
  - How many classes to have
    - generally fewer is better ~ 3-7 classes but this depends on colour scheme
  - What method to use
  - What kind of symbology to use
    - i.e. Graduated colors or graduated symbols

31

Guidelines for classifying your data

- The best classification scheme for a given layer depends on:
  - Purpose of the map
  - Characteristics of the data
  - Cartographic considerations
    - How easily the resulting map can be understood
- Two approaches to choosing a classification method:
  - Let the data inform your decision
    - Examine the histogram and resulting maps to help guide your choice
  - Choose a classification scheme first
    - Let the attribute values fall where they may when you have a scientific or statistical reason for using a particular classification method with particular data, or you have predetermined standards or criteria that dictate the method or number of classes

32

Classification methods

METHOD:

Natural breaks

WHEN TO USE:

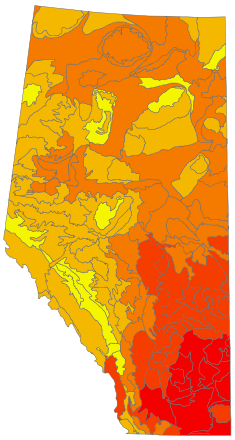
Attributes are distributed unevenly across the overall range of values

NUMBER OF CLASSES

Choose a number that best reflects the natural groups of attributes you want to show

5 Natural Breaks

- 0.0 - 273.5
- 273.6 - 434.6
- 434.7 - 557.5
- 557.6 - 716.6
- 716.7 - 943.3



33

Classification methods

METHOD:

Equal Interval

WHEN TO USE:

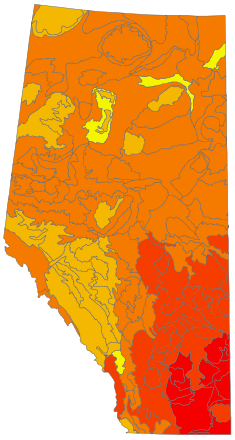
All classes should have the same range

NUMBER OF CLASSES

Choose a number that produces an easily understood interval (2, 50, 1000, etc.) or, the number of classes that produces a map with your intended message

5 Equal Interval

- 0.0 - 188.7
- 188.8 - 377.3
- 377.4 - 566.0
- 566.1 - 754.6
- 754.7 - 943.3



34

Classification methods

METHOD:

Quantile

WHEN TO USE:

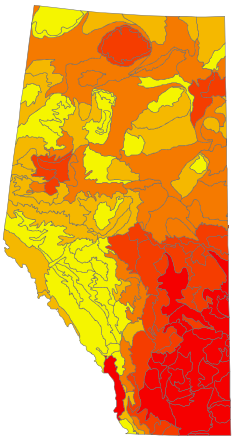
Attributes are distributed evenly distribution across the range of values

NUMBER OF CLASSES

Choose a number that makes sense for the purpose of your map

5 Quantile

- 0.0 - 385.7
- 385.8 - 449.9
- 450.0 - 499.5
- 499.6 - 630.6
- 630.7 - 943.3



35

Classification methods

METHOD:

Standard Deviation

WHEN TO USE:

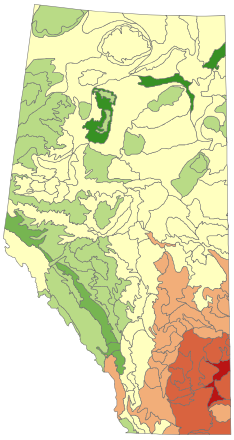
Show how much a feature's attribute value varies from the mean

NUMBER OF CLASSES

Choose the interval size and ArcMap calculates the mean value and the standard deviations from the mean to use in creating class breaks

1 Standard Deviation

- < -2.5 SD
- -2.5 - -1.5 SD
- -1.5 - -0.5 SD
- -0.5 - 0.5 SD
- 0.5 - 1.5 SD
- 1.5 - 2.5 SD
- > 2.5 SD



36