



PDHonline Course L154 (5 PDH)

Data in GIS

Instructor: Steve Ramroop, Ph.D.

2020

PDH Online | PDH Center

5272 Meadow Estates Drive
Fairfax, VA 22030-6658
Phone: 703-988-0088
www.PDHonline.com

An Approved Continuing Education Provider

Slide 1

Lecture 2 Content

■ Geographic Information Systems (GIS)

Data in GIS – Acquisition and input

L154 - Data in GIS Lecture 2
Dr. Steve Ramroop

1

This lecture is a continuation of the data in GIS topics identified in the course description, that is data in GIS – Acquisition and input. In this lecture we discuss some data source issues.

Slide 2

■ Data source issues

- Georeference**
- Accuracy**
- Use**
- Sharing**

L154 - Data in GIS Lecture 2
Dr. Steve Ramroop

2

This slide shows the content of lecture. Each one of the data source issues will be discussed in this lecture in further detail.

Slide 3

- **GIS database creation is a complex operation**
- **Raw data is available in many different analogue and digital forms**
- **Need to also decide appropriate data sets that is suitable for the given GIS application without having to be expensive**
- **Need to consider the projection systems used to capture and store data sets to be used in GIS**

L154 - Data in GIS Lecture 2
Dr. Steve Ramroop

3

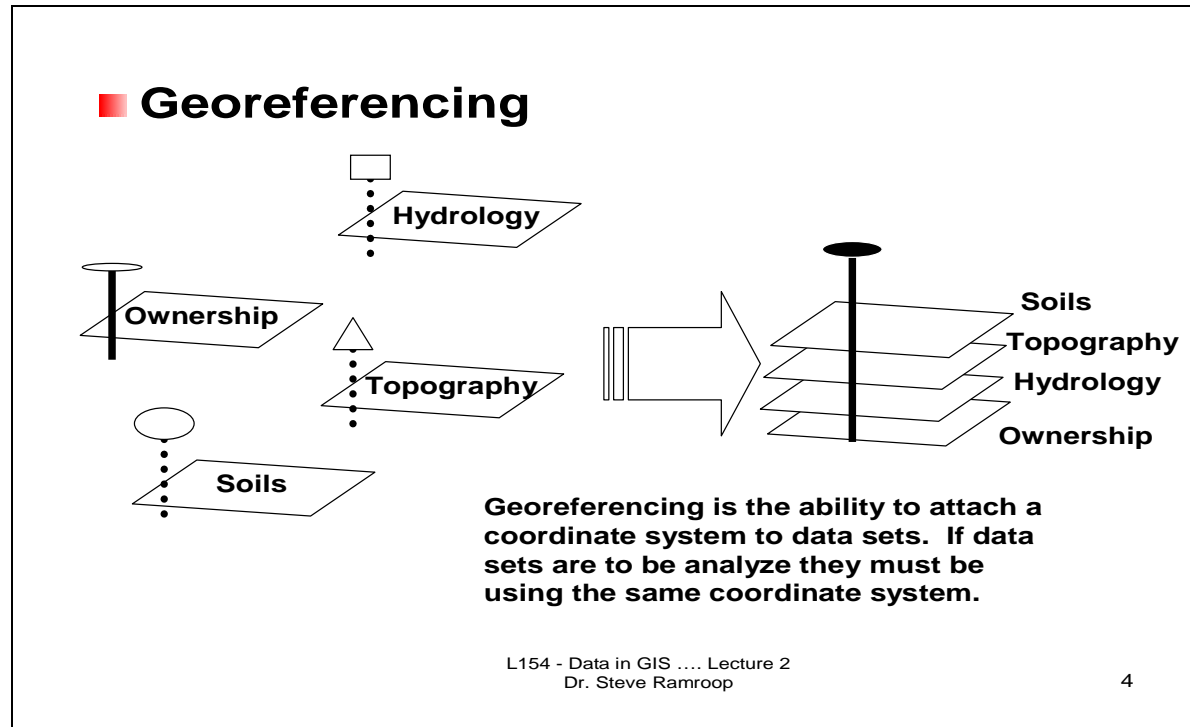
This slide gives some general remarks regarding the collection of GIS data to create GIS databases. These remarks are related to the issues regarding data sources.

A GIS database is very complex and depends upon the database model and the data model used to represent graphic features (the drawing) and its attributes (the drawing descriptive information). Raw data is the original observed data collected out in the field.

The quantity of data sets needs consideration in light of the intended GIS application. A balance is sought between the amount of data sets needed and the amount of money available to satisfy the intended GIS application.

Also, each data set has a coordinate system that is defined by projection system. Typically, to work with data sets of the same area of interest would require all of the data sets to be defined or transformed into the same projection system.

Slide 4



This slide shows the reality of acquiring data sets from different sources and usually would be defined using different projection systems. All of the data sets would be georeferenced but to work with them collectively all of the data sets need to be transformed into the same coordinate system. In the above slide we have four data sets that have different coordinate systems, and in order for us to collectively analyze the data for a given application, we will need to transform the different coordinate systems into one common coordinate system.

Slide 5

- **Most important in a GIS is that spatial data are positioned on a common reference**
- **Various coordinate systems are available and is particularly important for small scale mapping where the earth's curvature introduces distortions**
- **Coordinate systems are based upon the following surfaces (discussed later in lecture 5 of this course – L154):**
 - **Plane**
 - **Cylinder**
 - **Cone**

This slide gives more general remarks about what are the essential requirements of data sets before they can be used in a GIS application.

The coordinate systems discussed later in lecture 5 are an important consideration. Two approaches can be taken. Either collect all data sets using the same projection system or perform transformations between a projection system of a source data set to a destination data set projection system. The destination data set projection system will be the common projection system which will be used by all of the data sets of the intended GIS application.

Slide 6

– Large scale mapping, a convenient coordinate system is that of a plane

Cylindrical projections → Tropical regions

Conical projections → Temperate regions

Azimuthal projections → Polar regions

Depending upon the scale of the map different projection surfaces can be adopted.

- Cylindrical projections are typically used to map Tropical Regions.
- Conical projections are typically used to map Temperate Regions.
- Azimuthal (or plane) projections are typically used to map Polar Regions.

Slide 7

- **If there is no georeferencing, then GIS analyses are halted**
- **If there is little similarity between data layers, errors will be introduced such as:**
 - **Slivers, new polygons, incorrect distances and directions, and so on**
- **Methods used to georeference:**
 - **Capture and store data sets using a common projection system**
 - **Transformation of different projection systems to a common projection system**

L154 - Data in GIS Lecture 2
Dr. Steve Ramroop

7

If there is no projection system then the analysis done will not have realistic coordinates and the use of a GIS will be insignificant. When overlaying similar areas which have different projection systems the output will introduce polygons that are caused by the mis-alignment of same lines. When these lines do not align, then small polygons will be created due to the mis-alignment. These polygons are called slivers.

This slide also identifies two methods of how data is georeferenced. Either define the common coordinate system at the beginning of the GIS application development or later transform from one known coordinate system to another.

Slide 8

- **Aerial surveys are georeferenced by using ground control points**
- **GPS data sets are georeferenced by:**
 - **Distance between satellites (at least 3)**
 - **Atomic clock correction**
 - **Positional information of control points**

All three are required to compute latitude/longitude, grid reference, and altitude of GPS receiver

- **Remote sensing data are georeferenced using ground control and projection system parameters such as the spheroid**

This slide shows the how aerial surveys, GPS, and Remote Sensing data are georeferenced. The role of ground control, and the mathematical formulae representation are major considerations.

For remote sensing done by satellites, the spherical nature of the earth must be taken into consideration using the mathematics that defines the shape of the earth. The earth can be defines as a spheroid, ellipsoid, or a geoid. There are various mathematical approaches in defining the earth and each has their accuracy limitations.

Slide 9

■ Accuracy

- **This is the measure of the exactness of data sets to the true value**
- **Variable accuracy of data sets are due to:**
 - **Method of capture**
 - **Resolution and scale**
 - **Instruments used to capture the data sets**
 - **Intended use of the data sets**
 - **Validation and checking procedures**

The second issue discussed in this lecture is accuracy. It is a measure to its true value. There are various reasons which influence the variable accuracy of data sets. A few of these reasons are listed on this slide. They are self explanatory and you are required to expand on each of the reasons.

Slide 10

- **All collected data sets that has been checked and validated, can be used in a GIS**
- **However, in a collection of data sets, not all can be used in a GIS for a given application because of the accuracy of the data**

Continuing the issue with accuracy, this slide indicates that data sets need to be checked and validated out in the field which is a measure of accuracy.

Note that one GIS application may not be suitable for another GIS application. This typically implies that, all data sets for one application may not be suitable for another. The reason for this is the accuracy of the data sets will influence the quality. For example the data used in by a tourist will not be suitable for a LIS.

Slide 11

■ Use

- **Data sets used in one GIS application may not be useable in another GIS application because of:**
 - **Content : refers to the area and the theme presented in the data set**
 - **Accuracy : discussed earlier**
 - **Scale : defines the level of detail presented in the data set**
 - **Quality : gives an indication of the limitations of the data set which are due to content, accuracy, and scale**

L154 - Data in GIS Lecture 2
Dr. Steve Ramroop

11

The third data source issue is “Use”. This slide explores and extends the use of data from the last slide which explored the issue of ‘Accuracy’. The use of data sets can be variable due to content, accuracy, scale, and quality. Each of these is presented on this slide.

Slide 12

■ Data collectors**Two types of data collectors:****– Systematic data collectors**

- Responsible for maintaining national or regional data interest for the public. For example: Human settlements, natural resources, etc.

– Ad hoc data collectors

- Responsible for data sets which are project specific. For example: private surveyors, civil engineers, market researchers, etc.

Another data issue is Data Collectors. **NOTE:** The data collectors referred to here are NOT the people who collect the data but more at a higher national level.

Two types are identified. The ideal is to have systematic data collectors but in reality we typically have to deal with Ad hoc data collectors because organizations tend to be autonomous. This of course introduces many problems when data is shared between organizations. Systematic data collectors supports national and international data sharing while ad hoc data collectors work in a vacuum with no insight in support of further data sharing strategies.

Slide 13

■ Metadata

- Defined as *“Data about data”*
- Assists by informing data users of the type of data by describing its content, quality, use, lineage (history), and ownership
- Much research is in this area

This is an area of continuous research interest. Metadata provides a method for describing geographic data sets so that interested users can locate suitable data sets. The metadata describes each data set. Many metadata standards exist and only recently (2003) there is an international standard which is used by all GIS users.

Slide 14

– **Metadata provides the user with the ability to search for data sets based upon:**

- **The currency of data**
- **Length of record**
- **Scale of data**
- **Georeference system**
- **Data collection technique**
- **Data quality**
- **Classification and interpolation methods**
- **And so on....**

The items listed on this slide are examples of metadata attributes in which interested users can search for their suitable data sets. Each metadata standard describes GIS data as shown in the slide. The international metadata standard has over 200 metadata attributes which can be used to describe GIS data.

Slide 15

■ Sharing

– Issues considered are shown in the table below:

Geodata	Data formats
Geoservice	Royalties
Standards	Ownership
Metadata	Accuracy
Data structures	Quality

Students: You are required to write short notes on each of the issues listed in the table.

This slide identifies some of the issues regarding the sharing of data among organizations. Students need to find out online more about the issues listed within the context of data set sharing in GIS.

Slide 16

... The End ...