UNIT III: GEOGRAPHICAL INFORMATION SYSTEMS

Session 1

Introduction to Geographical Information Systems

Overview:

1.1. What is a GIS? or Geographical Information System?

- In the past, Geographical Information has been stored and presented in the form of maps.
- A modern GIS is a computer system capable of **assembling, manipulating, and displaying** geographically (spatially) referenced information.
- It is the spatial nature of the information, which makes it different to other information systems.
- The data that goes into the system and the people operating the system are also part of it.
- It is important to note that the geographic data in the system can be analysed (**manipulated).**

The older analogue 'Geographical Information System' is the 'map'.

Spatial scientists and cartographers stored geographical information in the form of maps. This was a means of storing and presenting geographic information in a similar manner to the modern GIS.

The modern GIS is a computerised geographic database system.

The two systems are similar in information content but there is a very big difference with respect to the capabilities of the modern digital GIS. The modern GIS is a geographical database with tools for data analysis and data presentation. A GIS is a tool we use to help us to store and manipulate large datasets and to perform complex operations that would take a human a long time! Algorithms and storage techniques are used for this purpose.

But remember that it is the human being who decides on the structure of the database and the input information for the desired result. Without wise planning on your part regarding the input of data and the storage, you will only end up with large data sets, which will be of no good use!

1.1.1 Who would use a GIS?

Before we start to understand what really is 'geographic data', let us see who needs to use this data in a GIS?

Municipalities use this to maintain large and complex databases that contain street locations, water and sewer lines etc,
Geologists use them to record locations of rock formations and resource information etc.
The military use highly classified databases for many uses.
Emergency services use them for maintaining address and location databases for quick location.
Cartographers use them for map making and storing.
Civil engineers use them for site selection and earthwork manipulations.
Many more!

Operational Applications:

- Utilities management
- Telecommunication
- Transportation
- Emergency management
- Land administration
- Urban planning
- Military applications
- Library management

Social & Environmental applications:

- Local, National & Global applications
- Healthcare applications
- Politics
- Monitoring land-use
- Environmental monitoring and assesment

You will be able to think of many more applications once you understand the function of a GIS!

1.2 What exactly is geographic data?

It is data that is spatially referenced. This means that the data is identified according to the locations. The most obvious type of spatial reference is a map. We can have maps indicating many types of data.

We may need to know details about a piece of land in order to construct some structure. What is the type of soil available? What is the slope of the land? How far is the land from a road? These are questions to ponder on before getting down to decisions.

At national level we need to have data about different areas in order to decide on priority requirements. The different data figures for different items such as the number of doctors serving each district, the number of school going children with the number of schools in a district and so on can be reflected much better using maps. Townships are complicated complexes where a data base is required for many types of information. Location of streets, houses, culverts, electricity, water & sewage line information are just a few.

1.2.1 Spatial data

If we look at a country as a whole there is much to be done in various sectors and regions of the country in order to improve the livelihoods of the people. To make decisions for national priorities and plans, much data needs to be collected, analysed and presented.

Consider information on population statistics for Sri Lanka.

Let us see how we can present this information.

- The data can be in table form.

How can you identify the different areas? In this case you need to go through the data in the tables, laboriously.

- The data can be plotted on a map.

This gives a spatial effect to the presentation. It is a ‘theme’ and we call this type of map a thematic map. This will indicate at a glance, the areas where more work and effort is required to be put in by the government.

See the maps below for some methods of presenting information.

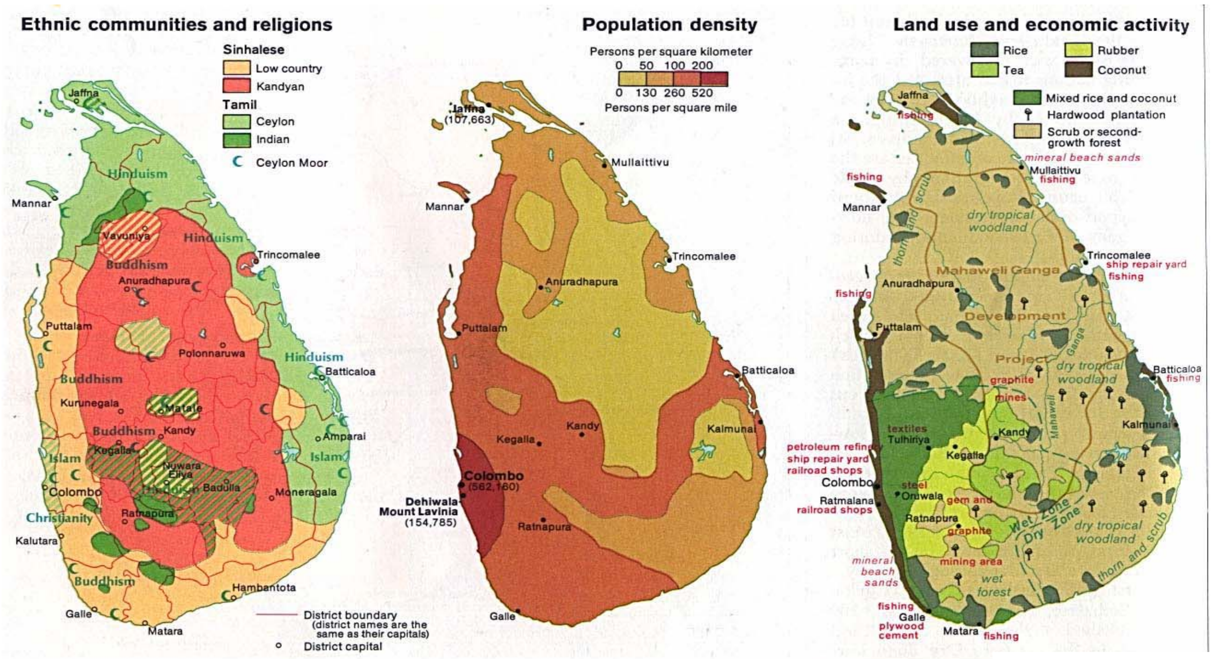


Figure 1.1 Thematic maps of Sri Lanka.

We can give the change in this information with **'time'.** This means that the change in the indices for different areas with time can be observed, recorded and stored.

This is referred to as **'Temporal data'** where the variations with time are indicated for a given data.
In this case the spatial data will also be in temporal form.

A GIS should be able to handle any or any combination of the following types of data.

- Point: Addresses, spot heights etc.
- Line: Contours, streets, rivers etc.
- Polygons/Areas: Municipal boundaries, forests, soil types, land use etc.
- Networks: Streets, water supply etc.
- Tessellations: Electoral boundaries, postal boundaries, (adjacent areas but not overlapping) etc.
- Overlapping regions:

We shall address the above data types in the coming sessions.

1.3 History of Geographical Information as a Science:

1.3.1 Maps & Spatial information:

If you were asked to go to a certain location and you do not have anybody to ask for directions what would you do? Refer a map of course!
A map will give you spatial data and show you where the point of interest is located.

A photograph of an area, taken from a point on the ground, is a map where you see a place as our eyes see it. An aerial photograph will give details of an area as seen from above. A shaded relief map shows how an area looks when sunlight is shining on it from a particular direction. It highlights the hills and

plains of an area. Topographic maps indicate the hills by means of contour lines. Road maps show the routes of highways. 3-D maps show an area in three dimensions.

I am sure you know how a magnetic compass can be used to get the bearings and prepare a map of the route traversed by you.
From very early times up to now, spatial data have been collected by navigators, geographers and surveyors and converted to pictorial form by map-makers or cartographers. With advancing scientific study of the earth, new material needed to be mapped. The spatial data relating to the different areas had to be in a form suitable for study.

Maps give us a better understanding of the Earth. Mapping technologies can be used for the following purposes.

- To know exactly where the place of interest lies on the Earth, by using a coordinate system of latitude and longitude.
- Maps and navigation charts are used to reach the desired destinations.
- A different kind of map for indicating various themes such as maps of rainfall, temperature, population density, etc.

What are the developments in mapping techniques?

Data acquisition techniques have developed to a point where we can now use remote sensing, digital photogrammetry and global positioning systems together with the computer technology and have advanced mapping techniques.

With these developments it has become possible to use mapping technologies in many new applications. Some of these are mapping the structure of the earth's core, mapping the ocean floor and so on.

What are the different types of spatial data and how could they be best laid out?
Let us see what are the data types that need to be mapped. They can be broadly categorised into two types.

- General Purpose Maps, which are topographical maps giving the spatial distribution which is the general layout of the natural form of the earth's surface and man made objects of the area
- Limited Purpose Maps, which indicate natural resources and can be in the form of a number of different maps each one indicating a different resource

These specific purpose maps are also referred to as Thematic Maps because they contain information about a single subject or theme. Examples of themes are, the rainfall pattern, agriculture pattern, population density, industrial estates etc.

Maps indicating the natural resources such as geology, geomorphology, soil science, ecology, land etc are required in the understanding of them. These maps are termed topographic maps. Each of the above resources will be the themes and can be indicated in different thematic maps.

Phenomena vary from place to place, but surveyors were unable to record and process the huge amounts of data that are needed to provide a proper insight of their variation.
To make the data easy to understand, thematic maps are drawn over a simplified topographic base.

Thematic maps can be of many forms. The theme may be qualitative or quantitative.

An example of a qualitative theme is the 'agriculture pattern'.
An example of a quantitative theme is the ' population density'.

Both these types can be represented as a choropleth map. A qualitative choropleth map is one, which shows areas of similar quality parameters separated by boundaries. Typical examples are soil maps showing different soil types and land use maps showing the different land cover.

A quantitative choropleth map will indicate areas of equal values separated by boundaries. Attributes, which can be represented by a continuous quantitative surface are, elevation above mean sea level, distribution of population in numbers etc. Variations between values are shown by isolines or contours.- lines connecting points of equal value.

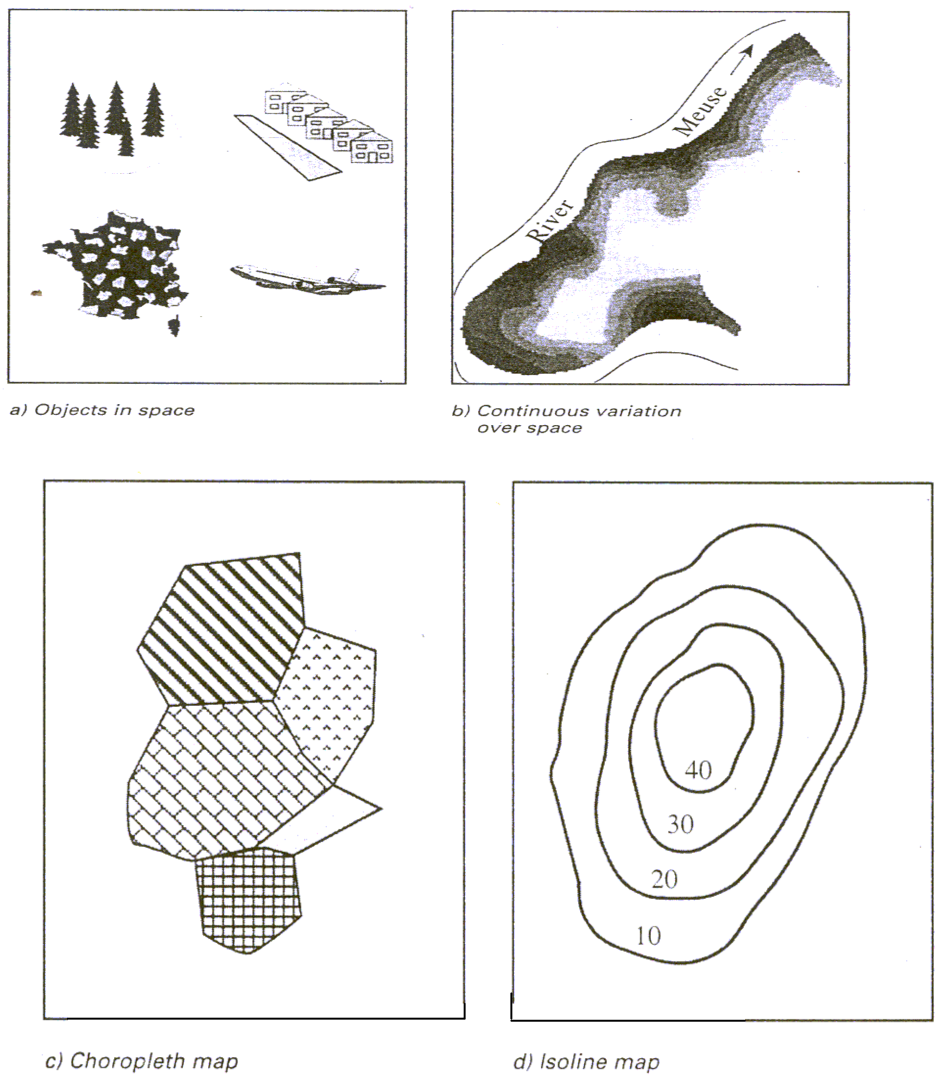


Figure 1.2 Conceptual models and representations of spatial phenomena

1.3.2 The need for Geographical Information

As the population increases what happens to our natural resources? of course they deplete.
So there is a demand on the Earth's resources of land, air, water and raw materials.

With the realisation that better management is required of limited resources, people have become more organised. Complex social and economic patterns have evolved. Human beings have understood the need for better sharing of resources between humans as well as for all forms of life. So we must learn to make the most of our resources. It must be mentioned here that it is not only the natural resources but also the changes brought about through social and technological advances also needs to be monitored constantly.

How do we set about doing all this?
First we need to understand the spatial & temporal patterns of resources and also spatial & temporal processes governing their availability.
Let us see what this means.
We need to know exactly where the resources are located (spatial data).
Data needs to be monitored continuously. For this reason, we need to obtain data of a resource or a phenomenon at different times (temporal data). The temporal process for obtaining rainfall data can be on a daily basis. The temporal process for obtaining population data can be once in ten years.

GIS is all about how we record the information in order to be able to retrieve the data when required.

Let us go back a little in to history now.

Why did the need for data storage arise?

Let us think of land and ownership. People have always had conflict over land & resources. In order to resolve and also prevent such conflicts there was a need to develop laws and codes about land and ownership. So there was a need for records of transactions and agreements over land.

Land ownership, Lot number, extent, land use are some of the information that need to be stored. Storage is in a database.

From ancient times spatial data have been collected by navigators, geographers and surveyors to be recorded in coded, pictorial form by map-makers and cartographers. In Roman times the agrimensores or land surveyors were an important part of the government. With the fall of the Roman Empire the surveying and mapping was swept into the back-ground. It was revived once again in the seventeenth century. By this time it was seen that a mathematical projection system and an accurate set of coordinates improved the measurement and location of land areas. The Cartesian coordinate system with x-y coordinates was used for location.

By this time many governments realised the importance of systematic mapping of their lands.
The Geographical Information society was formed by establishment of national government bodies to produce cadastral and topographical maps of whole countries. The mapping sciences- geodetical surveying, photogrammetry, and cartography – developed to provide very accurately recorded locations.

1.4 Geographical Information

1.4.1 Characteristics of spatial information

For everybody to read and understand records or maps it was necessary to have units of geographic information. Let us see what these are.

1.4.1.1 Symbols and Attributes:

How do cartographers represent real world objects?

They can be represented by means of symbols.

Symbols used are Points, Lines, Polygons.

Symbols used are: Points, Lines, Polygons

A Sri Lanka topographic map contains the following categories of information.

Point:
Schools, hospitals etc.
Lines:
Contours, streams, roads etc.
Polygons:
Land use, Provinces, districts etc.

Attributes:
These any other data relevant to that particular spatial data as indicated by the symbol.

They are the non-geographic characteristics of the geographical entities (spatial information) in a GIS.

> Attributes are the non-geographic data component of geographic data in a GIS.
>
> The usefulness of the GIS is dependent on the content and quantity of attribute data.

Cartographers represent real world objects by accurately drawn point and line symbols illustrating their most important attributes. The spatial data-base is a drawing on a piece of paper. Information is in the form of points, lines or polygons/areas. Geographical entities are given by symbols. Attributes of areas are indicated by uniform colours or shading with varying shades for steepness variations, variation of land cover etc. Mapped size and actual size are related by the scale used for the drawing. eg. 1:10,000 etc. All information is explained in a legend.

Map reading involves interpreting colours, lines and other symbols.
Features are shown as points, lines or areas depending on their size and extent.
Using Map Scales and the coordinates of the symbols, we can obtain the locations of the features.

Point features are represented by a single x,y (2-D) or a x,y,z (3-D) location. Examples are bench marks, electric posts etc.

Line features connect two points. Examples are centre lines of roads, water mains etc.

Area features or polygons have a defined 2 dimensional extent with boundary lines. Examples are land use areas and district divisions.

Three dimensional surfaces are areas with height dimension added to it. Examples are surface terrain represented by contour lines that have an elevation value.

Map scale can be either in the form of 1 cm = 1 Km or this same information given as 1: 100,000

Symbols are explained in the legend.

The Earth is a sphere. How do we represent points on the earth on a 2 dimensional map?
Transformation can be done in many ways.

The projection used for transformation in almost all the maps of Sri Lanka is, the Transverse Mercator Projection.

Figure 1.3 A typical map with a legend

1.4.2 Existing information systems

In order to understand the importance of computer aided information systems let us consider some aspects of existing information systems.

We have seen that cartographers have been using paper as information systems over the years.
What are the problems associated with the storage of spatial information in the form of paper?
Some of them are;

1. The data layout on paper, has to be understandable and representable. In order to do this original data obtained from the field has to be reduced in volume or be classified. Hence much detail has to be left out.
2. Map has to be extremely accurate.
3. There is a large volume of information to be inserted on any map. This means that for more detailed information on certain areas, each of these areas has to be re-represented in several separate map sheets per area.

(Think of the following problem in this case. Your area of interest lies exactly between two map sheets!)

4. Once the data is put into a map on paper, a part of it cannot be easily retrieved for use on another map. You need to redraw the area in order to combine with other spatial data.
5. A paper map is a static, qualitative document. Quantitative spatial analysis with respect to the area often cannot be performed without collecting additional information about the area.

Now let us see how much of geographical information needs to be recorded and stored in the present times and how we set about doing it.

Can we continue with the paper representations or should we have better recording and manipulating systems?

Yes we do need better recording systems.
We also need to do some analysis of the geographical data.

Today's geographical information is basically, location of and interactions between, well defined objects in space like trees on a forest, houses on a street, aeroplanes en route to destinations or administrative units. Other reasons for the need of maps are the understanding of the natural resources- geology, geophysical geodesy, geomorphology, soil science, ecology, land etc.

The 20^{th} century saw a great demand for maps of topography and natural resources. Stereo aerial photography and remotely sensed imagery have led to large area mapping. Also these technologies have given earth resource scientists the data for detailed mapping of geology, soil, ecology and land use.
Maps are essential for resource exploitation and management.

But why is the study of land evaluation so important?

- It is because we need to match the land requirements for producing food and to sustain the population with respect to environment, soil, water, etc. Earth scientists are constantly engaged in studying the Earth's resources.

This need for spatial data & spatial analysis is important not only for earth scientists. Urban planners and cadastral agencies need detailed information about the distribution of land and resources in towns and cities. Civil engineers need to plan the routes of roads and canals, evaluate volumes of cut & fill and to estimate construction costs. Police departments need to know the spatial distribution of crime types. Medical personnel need to know the distribution of sickness & disease. Business community will need to know the distribution of market potential. And also very importantly, the basic utilities such as water, gas, electricity, telephone lines, sewerage systems, all need to be recorded and manipulated in map form.

Compilation of data and the publication of a printed map is costly and time consuming. To extract one theme from a general purpose topographic map will be expensive if drawn by hand. The expense may not justify the use. Especially not if the map is not going to be valid for some years.
How often do maps need to be redone or updated?
Depending on the attribute represented on it, it may be in a number of years, months, days or even a few hours!
Think of weather charts. Imagine drawing them by hand every few hours!

A surveyor will record the information at the time of the survey. If needed periodically he has to redo the survey!
We have seen how the aerial photograph gives us information over time. We also know how the satellite image gives us much more information of landscape change over time.
How important are these photographs and images of the landscape to us!
They give an indication of the areas of desertification, erosion etc. They can also show us the weather systems.
We know that satellites provide us with digital images, which are streams of data on magnetic tapes.
But remember when using images, that image data has to be ground verified and a certain amount of field survey is essential for proper interpretation.
And also remember the images have to be located properly with respect to a proper geodetic grid.
Otherwise you cannot relate the area on the image to a definite place!

So having seen how difficult it is to redo work by hand at short time intervals, we come to the obvious solution.
Computer assisted mapping.
How did it come about? Let us see!
As we were saying before, the need for monitoring the resources meant the monitoring of the environmental characteristics of landform, geology, soil, vegetation and water.

It has been found that several maps of the same area drawn on transparent sheets can be laid one on top of each other and combined and integrated. They are viewed by laying the copies on a light table.

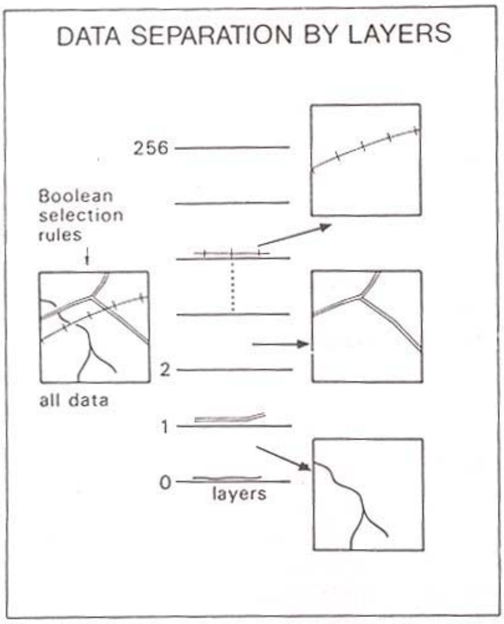


Figure 1.4 Overlay

- **It was this overlay concept which really led to computer aided mapping.**

Once the importance of the overlay concept of maps was realised, this gradually led to using the computer to make simple maps by printing statistical values on a grid of plain paper. The data was analysed and manipulated to produce choropleth or isoline interpolations. The results could be displayed in many ways using the overprinting of lineprinter characters to produce suitable gray scales.
What is important is that the grid-cell (raster) mapping program allowed the user to do on the computer, what was done with transparent overlays.

Note: You will see how the computer can be used for GIS functions by using Arc Explorer and ArcView software when you come for the practical sessions.

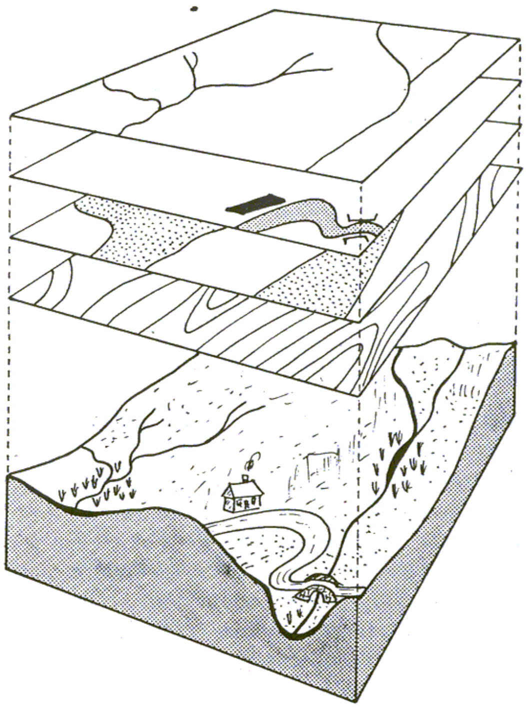


Figure 1.5: The ‘overlay’ concept-
The real world is portrayed by a series of overlays.
Each overlay is recorded with one aspect of reality.
The overlays can consist of soil types, topography, roads, rivers etc.

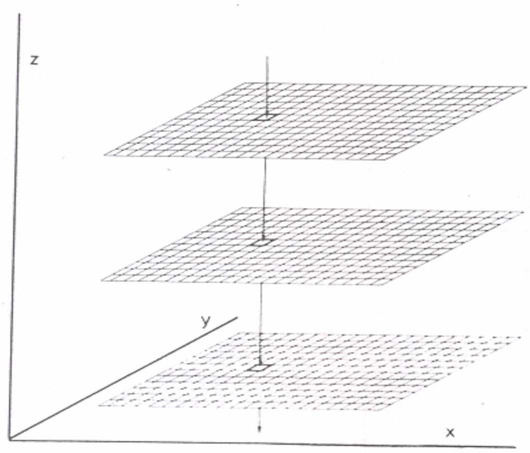


Figure 1.6: Three dimensional arrays used for coding map overlays in raster database structures

- At first the computer method of studying maps was only found to be quicker than the manual operation.

- But later it was realised that planning studies such as urban planning, ecological studies etc. proved to be possible with computer models, which was not really possible by manual methods.

How was this possible?

- GIS overlaying of different maps made it also possible to <u>analyse geographical data using the computer.</u>

- Spatial and logical analysis on mapped data was possible using the computer.

The GIS is not only a computer assisted mapping or database management system.
It is possible to computationally analyse the data as well.

Many computer programs were developed at this stage for quick and cheap analysis of grid data.
By 1977 Rhind was able to present the following reasons for using computers in cartography.

1. To make existing maps more quickly
2. To make existing maps more cheaply
3. To make maps for specific user needs
4. To make map production possible where skilled staff are not available
5. To allow experimentation with different graphical representations of the same data
6. To facilitate map making and updating when the data are already in digital form
7. To facilitate analyses of data that demand interaction between statistical analyses and mapping
8. To minimize the use of the printed map as a data store and thereby to minimize the effects of classification and generalisation on the quality of the data.
9. To create maps that are difficult to make by hand For example 3-D maps or stereoscopic maps
10. To create maps in which selection and generalisation procedures are explicitly defined and consistently executed.
11. Introduction of automation can lead to a review of the whole map-making process, which can also lead to savings and improvements.

What can we say about the further development of the computer-based mapping?

The capital cost involved was immense! The computer hardware and the software required, proved to be very expensive.
Computerised mapping did not cut down the costs of mapping! Even reproducing maps manually for transparent overlays was much cheaper!
Then what did this system of mapping really achieve? Was it only to produce maps more quickly?

So people did not even like to try the system!

Technology developed so fast and the computer hardware changed with it and managers could not cope with the changes. It was really difficult to see how this new technology was going to address the fundamental problems of mapping. Many used it only for accurate drafting, scanning contours etc.

Many cartographers did not realise that having mapped data in digital form meant that you have a powerful data-base that could be used for the analysis of many important spatial problems!

So although a technological breakthrough had occurred in geographical mapping, with the advent of digital mapping using the computer, people were not geared to make use of it!

As stated by computer scientist Joseph Weizenbaum, “this is really a matter of automation of existing manual techniques without a parallel conceptual development of the subject matter itself!”

It is true even today.

- Most development in the use of GIS has been directed to automating existing manual methods rather than exploring new ways of handling spatial data!

So there is still much room for research in this area of study.

There were two main trends in the application of computer methods to mapping in the 1960's and the 1970's.

- One was the automation of existing tasks with accent on cartographic accuracy and visual quality
- The other was with the accent on spatial analysis but at the expense of good graphical results.

What about the quantitative description of data? It was not possible until recently due to the following reasons.

(i) the large amounts of data involved and
(ii) the lack of appropriate mathematical tools for describing spatial variation quantitatively

It was begun in the 1930 's and 1940's with development in statistical methods and time series analysis. Later in the 1960's with the digital computer. This led to the real development of conceptual methods of spatial analysis, quantitative thematic mapping and spatial analysis.

Today Geographical Information Systems (GIS) leads to logical and numerical modelling. The application of statistical methods to spatial data, has become routine and is of great relevance.

A GIS is designed to accept large volumes of spatial data derived from a variety of sources including remote sensors.
Remote sensing inputs do not represent the primary source of data. They do represent in many instances, one of the following.

(i) either a potential source of new data elements for the GIS, which must be made to mesh with the existing GIS data definitions,
(ii) or an alternative form of data capture for one or more well defined data elements currently incorporated within the GIS.

The integration of remote sensing data inputs is judged by GIS operators by their compatibility with the existing data sources which are normally digitised data based on existing maps. It is important to understand this if remote sensing data is to be used in operational geographical information systems.

So what has GIS done?

Let us summarise what GIS has done?

- GIS has created a means of interconnection between Remote Sensing, Earthbound Survey and Cartography.
- It has also provided a means of manipulating vast amounts of data very useful for planning purposes.

1.5 Geographical Information Systems (GIS):

Spatial information is very important in modern society for planning, marketing, and the development of the 'Information Society'.
By the 1990's GIS became appreciated for the following reasons;

- Awareness that you need much spatial information to be manipulated efficiently.
- Computer technology provided huge amounts of processing power and data storage capacity on modestly priced personal computers. This made it possible for more people to use it.
- Many computers were made connected by electronic net works allowing expensive data and software to be shared.
- Standardisation and interfaces between database programs and other computer programs made it much easier to handle large amounts of data
- The automation of the existing methods became somewhat uniform resulting in a limited number of commercial systems dominating the market place.

1.5.1 Definitions of GIS

- **Toolbox-based definitions.**

A powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world. (Burrough 1986)

A system for capturing, storing, checking, manipulating, analysing and displaying data which are spatially referenced to the Earth. (Department of Environment 1987)
An information technology which stores, analyses, and displays both spatial and non-spatial data. (Parker 1988)

- **Database definitions.**

A database system in which most of the data are spatially indexed, and upon which a set of procedures is operated in order to answer queries about spatial entities in the database. (Smith 1987)

Any manual or computer based set of procedures used to store and manipulate geographically referenced data. (Aronoff 1989)

- **Organisation-based definitions.**

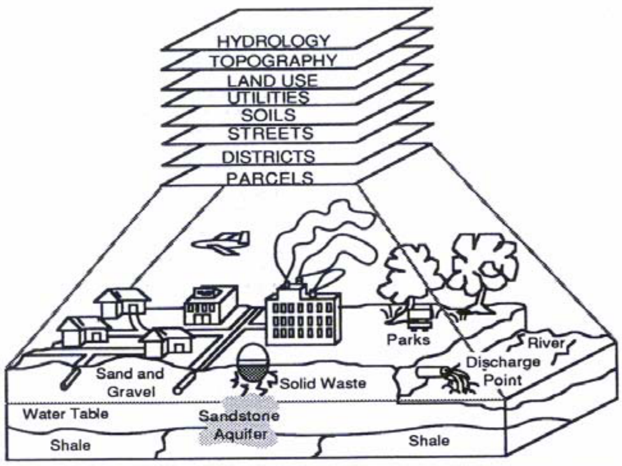
An automated set of functions that provides professionals with advanced capabilities for the storage, retrieval, manipulation and display of geographically located data. (Ozemoy, Smith, and Sicherman 1981)

An institutional entity, reflecting an organisational structure that integrates technology with a database, expertise and continuing financial support over time. (Crater 1989)

A decision support system involving the integration of spatially referenced data in a problem solving environment. (Cowen 1988)

- **A simple definition**

A computer system capable of holding and using data describing places on the Earth's surface.



The real world consists of many geographies which can be represented as a number of related data layers.

Figure 1.7 Real world data layers

1.5.2 Spatial Operations & Data Linkage

Spatial operations
By now it should be clear to you that a GIS is not only a database but also permits spatial operations on the data.

For example say we have a table of data giving the towns in Sri Lanka with their locational coordinates and the population of each town.

- aspatial query - What is the population of each city? Answering this question is simple and does not require location data or any other computation.

- spatial query - What are the towns situated within 50 Km of each other? In order to answer this question you need to manipulate the data and demarcate boundaries that lie within 50 km of each town.

Data linkage

A GIS will link different data sets.
Why is data linkage so important?

You may have several data sets for a given area. Each data set can be analysed/mapped separately. OR the data sets can be combined. By bringing the data sets together, you add value to the database.

To do this combining operation you need a GIS.
(explained in session 3)

1.5.3 What are the questions that can be answered using a GIS?

There are five generic questions that a GIS can answer.

- **Location** - What is existing at a particular location?

- **Condition** - Where can we find a location satisfying certain conditions?

- **Trends** - What has changed at a particular location over time? Where can we find locations with given changes in conditions? Seeks to find the differences within an area over time.

- **Patterns** - What are the spatial patterns that exist? How many schools are there within certain radius of a town?

- **Modelling** - What happens if some change takes place? Eg. A new road is added to a highway or a dam is located across a catchment area.

1.5.4 Active Domain for GIS: Taken from Geographical Information Systems, Burrough

Producers, Types and Applications of Geographical Information
The main producers and sources Topographical mapping: National Mapping Agencies, private Mapping Companies, Land Registration and Cadastre Hydrographic Mapping Military Organisations Remote Sensing companies and satellite agencies Natural resource surveys: Geologists, Hydrologists, Physical Geographers & Soil Scientists, Land Evaluators, Ecologists & Biogeographers, Meteorologists & Climatologists, Oceanographers
The main types of geographical data available Topographic maps at a wide range of scales Satellite and airborne scanner images and photographs Administrative boundaries: Census tracts and census data; Postcode areas, Statistical data on people, land cover, land use at a wide range of levels Data from marketing surveys Data on utilities (gas, water, electricity lines, cables) and their locations Data on rocks, water, soil, atmosphere, biological activity, natural hazards, and disasters collected for a wide range of spatial and temporal levels of resolution

Some current applications are given in the table below.

Agriculture	Monitoring and management from farm to National levels
Archaeology	Site description and scenario evaluation
Environment	Monitoring, modelling and management for land degradation; land evaluation and rural planning; landslides; desertification; water quality and quantity; plagues; air quality; weather and climate modelling and prediction
Epidemiology and Health	Location of disease in relation to environmental factors
Forestry	Management, planning and optimising extraction and replanting
Emergency services	Optimising fire, police and ambulance routing; improved understanding of crime and its location
Navigation	Air, sea and land
Marketing	Site location and target groups; optimising goods delivery

Real Estate	Legal aspects of the cadastre, property values in relation to location, insurance
Regional/local planning	Development of plans, costing, maintenance, management
Road and rail	Planning & management
Site evaluation and costing	Cut & fill, computing volumes of materials
Social studies	Analysis of demographic movements and developments
Tourism	Location and management of facilities and attractions
Utilities	Location, management and planning of water, drains, gas. Electricity, telephone, cable services

Session 2

Basic functions of Geographical Information Systems

Aims
To give an understanding of the components of a GIS and the four basic functions of a GIS.

Objectives:
At the end of the session you will be able to explain components of a GIS and the following processes.

- data capture
- data management
- spatial analysis
- presenting the results

2.1 Components of a Geographical Information System

There are three important components in a GIS.

- Computer hardware
- Sets of application software modules
- Skilled people to manage it

2.1.1 Computer hardware:

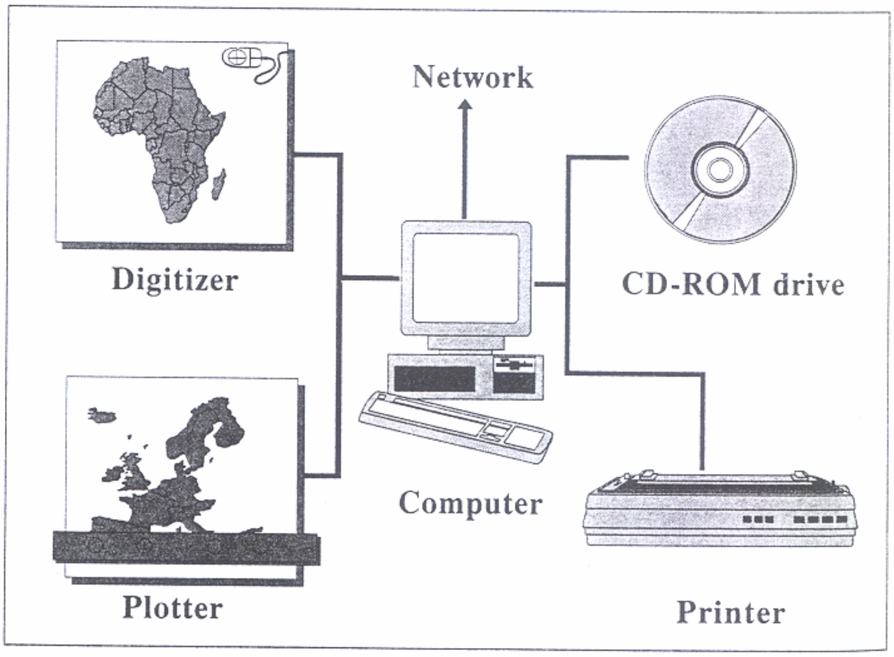


Figure 2.1: Major hardware components of a geographical information system

The basic Computer Hardware components are given in the table below.

Table 2.1

Item	**Usage**
Hard disk drive	Storing data and programs
Digital tape cassettes, Optical CD-ROM's etc.	Storage of data
Digitizer or Scanner	Converts maps and documents into digital form
Plotter, Printer or any other display device	Gives the output of data processing
Local & Global electronic network with either of the following; 1. Optical fibre data lines 2. Telephone lines with 'modem'	Provides Inter-Computer communication
Computer screen, Keyboard & mouse or other pointing device	To control the computer and the peripherals such as the digitizer, plotter, printer etc. which are linked to the computer.

2.1.2 GIS Software

The software modules can be grouped as follows;

- Data input and verification
- Data storage and database management
- Data output and presentation
- Data transformation
- Interaction with the user

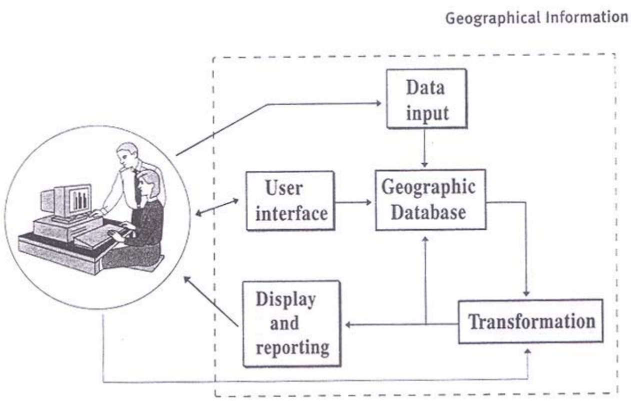


Figure 2.2: The main software components of a geographical information system

The software can be grouped as shown in the diagram above.
The following GIS softwares will be introduced to you during the practical sessions.
Arc Explorer, ArcView

2.2 Basic functions of a GIS

2.2.1 Data capture, data input & verification:

Geographic data is the information about the earth's surface and features on it. Feeding the data or the information into the system is referred to as 'data capture'. Before we see how we can capture data, let us first understand what geographic data really is and the different forms in which this data can be represented.

2.2.1.1 What is geographic data?

There are two important components of geographic data. The geographic position and attributes or properties related to that position.

Geographic data = Geographic position + attributes/properties

Geographic position gives the location by using a coordinate system.

Attributes refer to properties of spatial entities. They are as follows,

(i) identity (town, road, residence etc.),
(ii) ordinal (ranking such as class 1,2 etc.),
(iii) scale (elevation, length, width etc.).

Attributes are non-spatial data.

Raster and vector data
A GIS database will store spatial feature data in a raster or vector format.

Vector data
In the vector format, positions are stored in the form of coordinates. (x,y and sometimes z). A point is described by a single x,y coordinate pair and by its name or label.
Although a line is actually an infinite set of points, in practice a line is described by straight-line segments, each segment described by a set of coordinate pairs and of course the name or label.
An area also called a polygon, is described by a set of coordinate pairs and by its name or label with the difference that the coordinate pairs at the beginning and the end are the same.

A vector format represents the location and shape of features and boundaries precisely provided they are accurate at the point of input.

Raster data
The grid-based format generalises map features as cells or pixels in a grid matrix.
The space is defined by a matrix of points or cells organised into rows and columns. If the rows and columns are numbered, the position of each element can be specified by the column and row numbers. These can be linked to coordinate positions through the introduction of a coordinate system.

Each cell has an attribute value (a number) that represents a geographical phenomenon or nominal data such as elevation, land use class etc. The fineness of the grid or in other words the size of the cells in the grid matrix, will determine the level of detail in which map features can be represented.

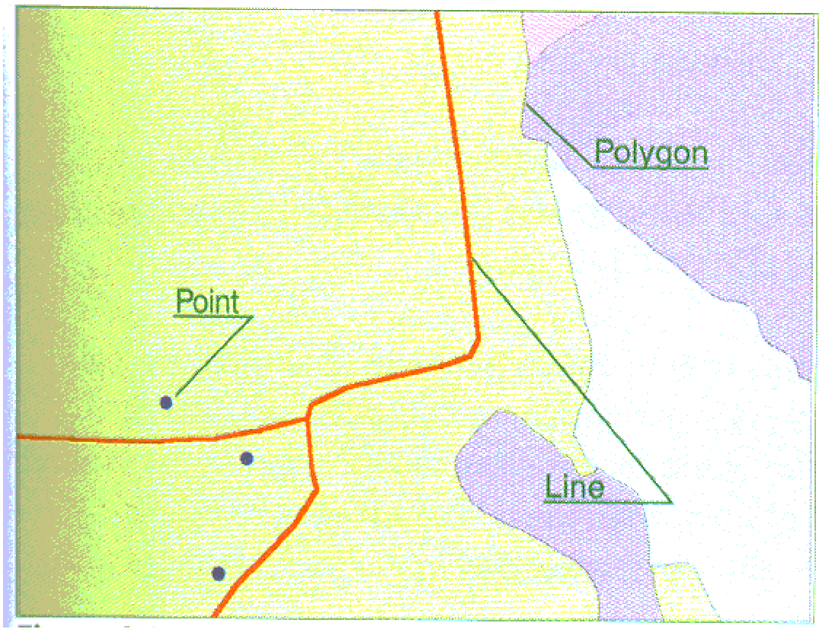


Figure 12.3: Vector format

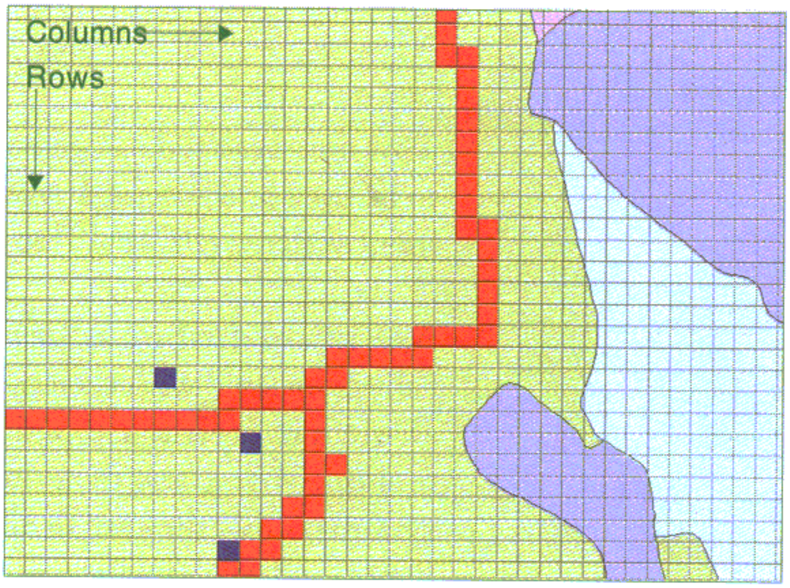


Figure 2.4: Raster format

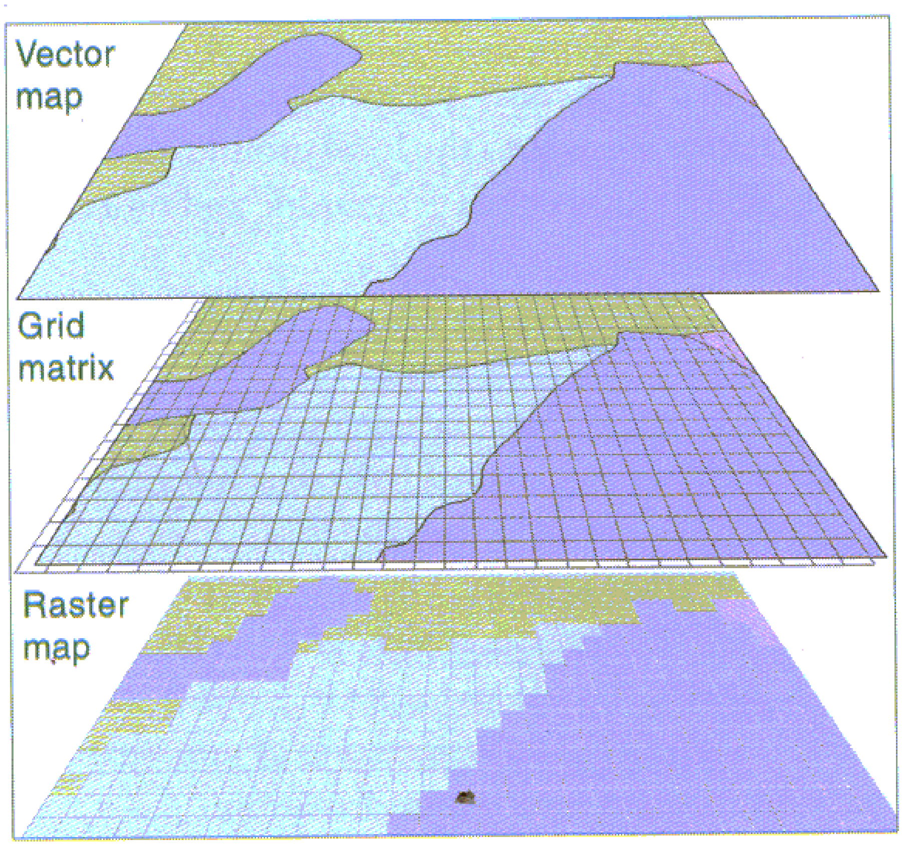


Figure 2.5: Vector-Raster relationship

Analogue and digital data
Data is available in the analogue or digital form.

What are the analogue data?
Analogue data are a physical product displaying information visually on paper.
Maps on paper, aerial photos and reports are some forms of analogue data.

What are digital data?
Digital data are information in computer readable form.
Digital photographs, satellite image data, GPS and existing computer data-bases are some forms of digital data. Digital data can be obtained from digital cameras, satellite sensors, using the GPS etc. You can also get digital data by digitising existing maps using digitizers.

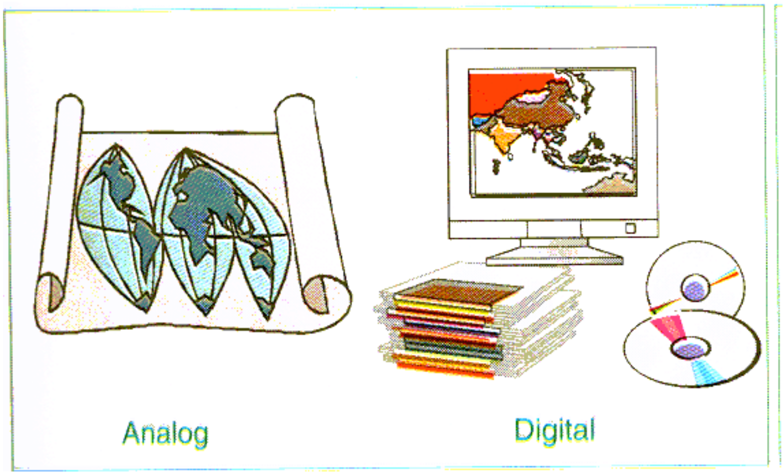


Figure 2.6: Analogue and digital data

2.2.1.2 Data input

Data input covers all aspects of capturing spatial data and converting the data into a form that can be processed using the computer.

We need to encode the data and write them to the database.

There are two aspects of the data that need to be considered separately in geographical information systems.

1. Positional or geographical data necessary to define where the graphic or cartographic features occur- which means that we need to know the coordinates
2. Associated attributes that record what the cartographic features represent – these are any other relevant data related to a feature

What is the main difference between Geographical Information Systems and conventional automated cartography?

- Geographical Information Systems - It is possible to process spatial data and non-spatial data in relation to cartographic features

- Automated cartography – In this case non-spatial data only relate to colour, line type, symbolism etc.

Methods of Capturing Spatial Data;

Capturing Analogue data:

- From existing maps
- Aerial photos
- Reports

Capturing Digital data:

- Sensors on aircraft & satellites - Scanners will record the data to be processed into photographs or images

- Field observations - GPS

- Digitising existing maps
- Existing digital database

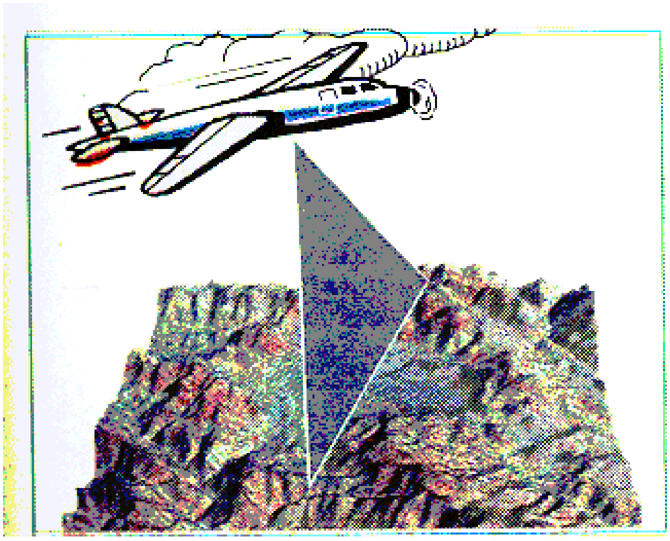


Figure 2.6: Capturing data via aircraft

Table 2.2

Data capturing methods	Process
Photogrammetric compilation	Source is aerial photography. The process involves using specialised equipment (a stereoplotter) to project overlapping aerial photos so that a viewer can see a 3-D picture of the terrain.
Digitising	Converts analogue data into digital format. Sources are images and maps. Digitising involves tracing with a precise cross hair in the digitising cursor features on a source map that is taped to the digitising tablet and instructing the computer to accept the location and type of the feature.
Map scanning	Optical scanning systems automatically capture map features, text and symbols as individual cells or pixels and produce an automated product in rastor format. Creating an intelligent GIS database from a scanned map will require vectorising the rastor data and manual entry of attribute data.
Satellite data	Earth resource satellites are a main source of data for GIS systems. High resolution data will provide panchromatic and multispectral data in 1m to 3m range.

Field data collection	Capture of field data is used to compile utility inventories, land-use inventories etc. Electronic survey systems and global positioning systems (GPS) have greatly improved field data collection. Electronic distance measurement data can be directly entered into a GIS. A GPS will provide coordinates in the field that can be loaded to a GIS.
Tabular data entry	Attributes in table form can be converted to digital form
Document scanning	Rastor files of documents such as site photographs etc can be created using smaller format scanners.
Translation of existing digital data	Many GIS packages have programmes that translate mapping data between Raster/Vector platforms.

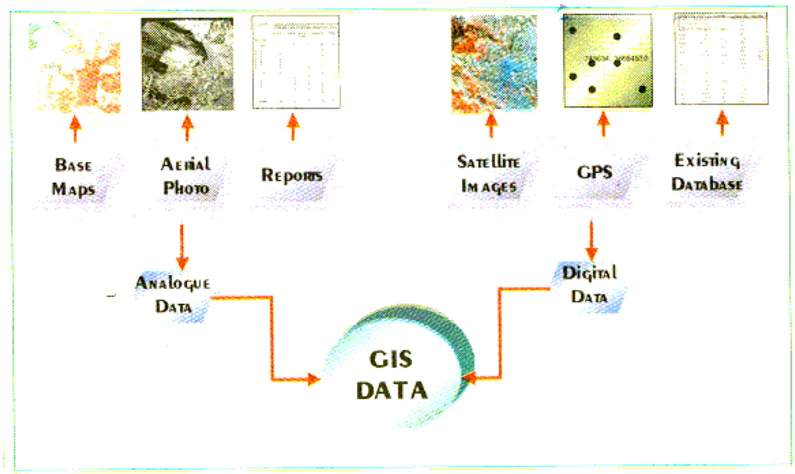


Figure 2.7: Methods of capturing analogue and digital data

Methods of converting this data to a standard digital form;

- Interactive computer screen & mouse
- Digitizer
- Word processors & spreadsheet programs
- Stereo plotters
- Scanners for converting maps & photographic images
- Using devices to read data from magnetic media such as tapes or CD-ROMs

Imagine doing a conversion manually!
There are many ways in which remote sensing data can be integrated into a GIS for analytical purposes. Some examples are as follows;

- Land cover maps or vegetation maps classified from remote-sensing data can be overlaid on to other geographic data to enable analysis for environmental monitoring and its change.

- Image data can be also used as image maps, with an overlay of political boundaries, roads, rivers, etc. Such an image map can be successfully used for visual interpretation.

Data Sources:

The acquisition and input of geospatial data is about 80% of the total GIS project cost. Hence it is important to understand that the data sources for data acquisition have to be very carefully selected. Selection will depend on the specific purpose. We have already listed the methods of capturing spatial analog and digital data. Let us review them.

1. Analog maps

Types of maps used are,

- topographic maps with contours and other terrain features and
- thematic maps with defined object classes

Analog maps are digitised by digitizers manually or by scanners semi-autimatically.

2. Aerial photographs

There are two types of aerial photogrammetry.
a. Analytical photogrammetry - A stereo pair of analog photographs is set up in a stereo plotter and the operator will manually read terrain features through the stereo photogrammetric plotter also called analytical plotter.
b. Digital photogrammetry - A digital photogrammetric work station is used to convert aerial films into high resolution digital image data.

3. Global Positioning System

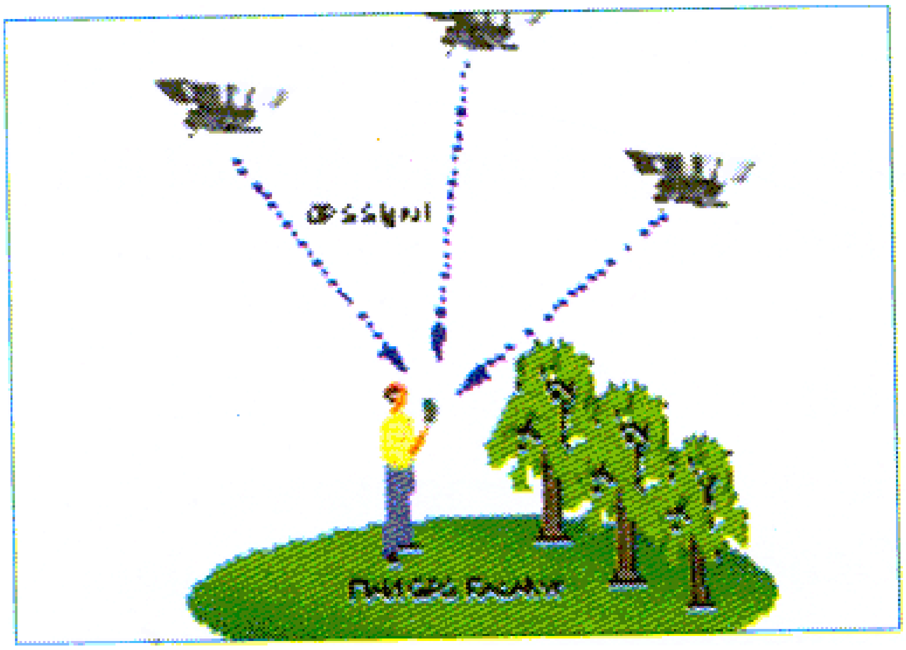


Figure 2.8: GPS

The Global Positioning System is extremely useful in precise positioning of geospatial data and collection of data in the field. Hence it is a valuable tool in GIS data collection, surveying and mapping.

We may have read about the GPS before (Block 1 Sessions) but let us very briefly review the operation of a GPS for recollection.

The GPS is divided into three major components. They are the control segment, the space segment and the user segment. All three segments are required to perform positional determination.

- The control segment consists of five monitoring stations spread over the globe.
- The space segment consists of the constellation of Earth-orbiting satellites.
- The user segment consists of all Earth based GPS receivers. (See Unit I session 6 of RS and GIS course for more details)

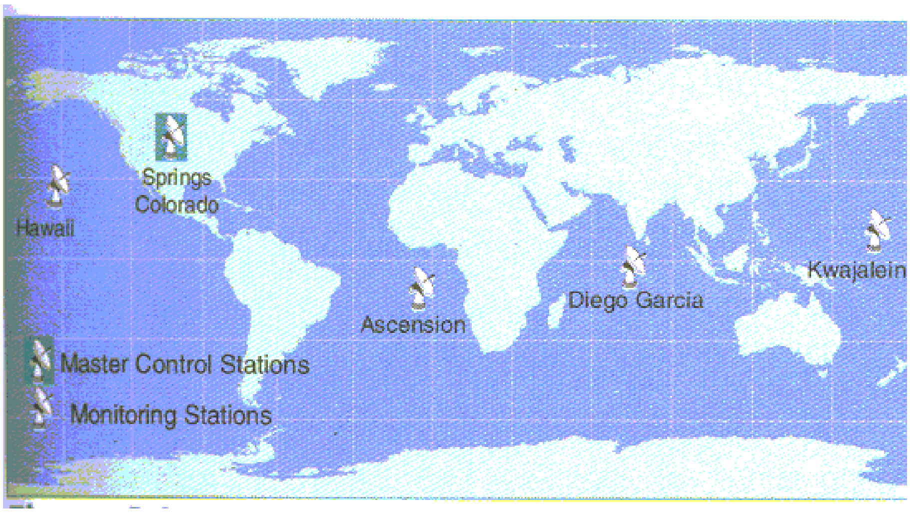


Figure 2.9: Control segment

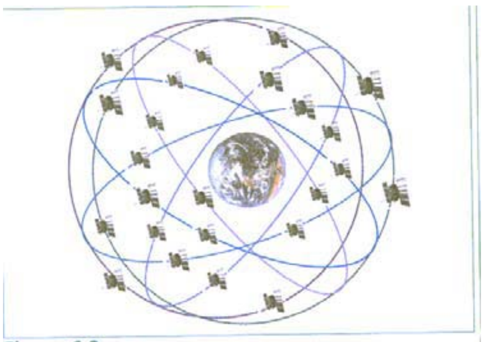


Figure 2.10 Space segment

Figure 2.11: User segment – GPS receiver

The GPS uses satellites and computers to calculate positions anywhere on earth based on satellite ranging.

Using the GPS for field data collection, data such as longitude, latitude height (and sometimes time), can be recorded instantly.

We can integrate GPS positioning in GIS.

GPS is an effective tool for GIS locational data capture. The GPS can be easily linked to a laptop computer in the field, or the data fed into a PC, and with appropriate software, users can place all their data on a common base with little distortion. GPS is also used in remote sensing methods such as photogrammetry and aerial scanning.

GPS is very advantageous in the construction of accurate and timely GIS databases.

4. Ground survey with GPS

This modern method of ground survey is a very accurate method of collecting spatial data. Total Station is used together with GPS. But this can be too expensive for covering large areas.

5. Reports & publications

Social & economic data can be obtained from reports.

6. Satellite Remote Sensing

Digital image data can be obtained of the terrain surface using satellite imagery. Images are taken in the visible, infrared and microwave regions of the electromagnetic spectrum.

Satellite image data are available for land use classification, digital elevation (DEM) model generation, updating spatial networks such as highway, irrigation etc. Image map scale will be around 1:50,000 to 1:100,000. High resolution satellite imagery with ground resolution of 1-3 meters will produce 1:25,000 topographic maps. As the spatial resolution of modern systems increase even further, it will become possible to produce more accurate maps.

Multi spectral bands including visible, near-infrared (NIR) and /or thermal infrared (TIR) are most commonly used for production of land use map, soil map, geological map, agricultural map, forest map etc at scales of 1:50,000 to 1:250,000.
Synthetic Aperture Radar (SAR) can be used in all, weather conditions since it can penetrate clouds.

Table below indicates an outline of some types of satellite data with their possible use.

Name of satellite (country)	Spectral Bands	Main purpose of satellite data
Landsat TM (USA) SPOT (France) ERS (European Community)	Visual, near infrared & thermal infrared Synthetic Aperture Radar	Land cover Topography Land cover Wind Wave

		Ice
JERS (Japan)	Synthetic Aperture Radar	Land cover Geology Terrain features
IRS (India)		Land cover Topography
Radarsat (Canada)	Synthetic Aperture Radar	Ice Snow

In addition to the above, High Resolution satellites images such as taken from IKONOS and Quickbird, have a great potential for urban GIS applications.

Data input to a geographical information system can be best described under three headings

1. entering the spatial data (digitising)
2. entering the non-spatial, associated attributes
3. linking the spatial to the non-spatial data

But it is important to remember that any errors in the database have to be avoided!
In order to avoid errors the data needs proper verification.

2.2.1.3 Data errors

Errors can arise during the encoding and input of spatial and non-spatial data. These errors can be grouped as follows:

- incomplete spatial data,
- spatial data are in the wrong place or at the wrong scale,
- spatial data are distorted,
- spatial data are linked to the wrong non-spatial data,
- incomplete non-spatial data.

The data may also need to be reduced in volume.

The best way to check the correctness of the input spatial data is to get the computer to draw them out again. The two maps can be placed one over the other and viewed on a light table.

Non-spatial data can be checked by printing out the files of non-spatial data and going through them visually.

2.2.2 Methods of data storage & database management:

What does the geographic data consist of?

Complex entities representing objects on the Earth's surface will consist of elements such as points, lines, polygons/areas in a Geographic Database.

Remember that building a digital geographical database is very time consuming and also very expensive!

So it is important to have a proper means of storage in a permanent storage medium.
Digital databases for topographical, cadastral and environmental mapping should last about 25 years.

The geographic components are structured and organised in such a way so as to fulfil the user requirements and to enable easy handling by the computer.

The computer program used to organise the database is known as a Database Management System. The following aspects of the geographical data are important in the functioning of the GIS.

- The Location of the geographical data
- Topology or the linkages and the
- Attributes or the non-spatial data,

We shall read more about database management in the sessions on data structures.
How can the data be organised to handle the input information in an efficient manner?

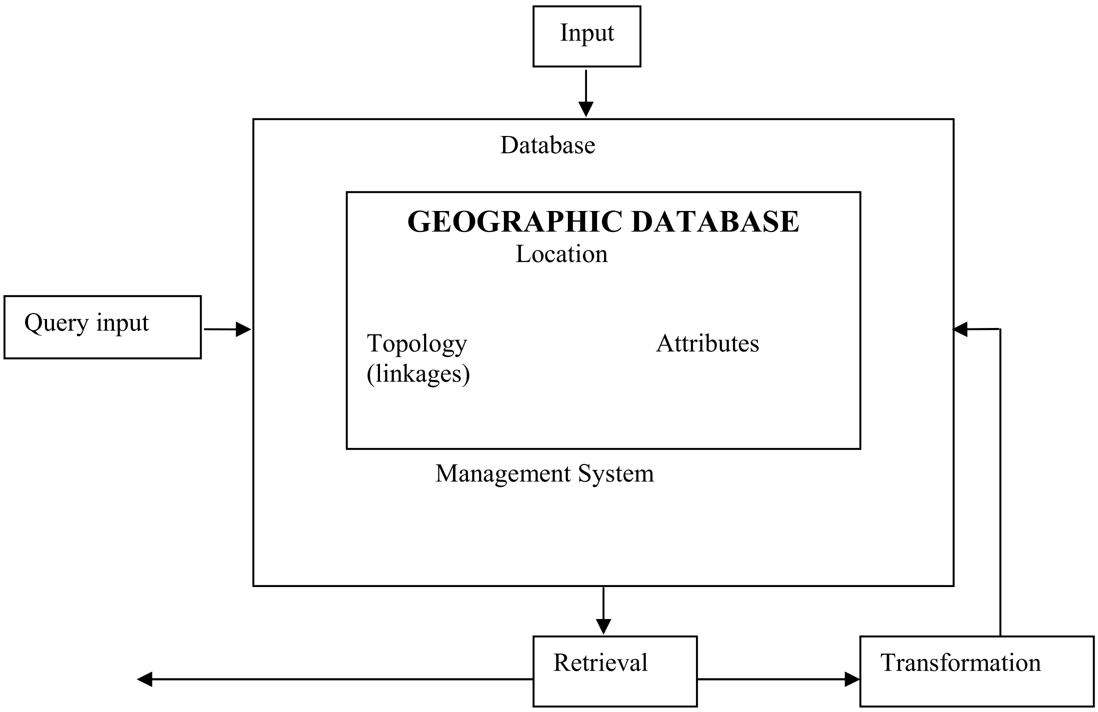


Figure 12.12: The components of the geographic database.

2.2.3 Spatial analysis

Data Transformation

A GIS can perform spatial analysis. Spatial analysis deals with geographic patterns in data and relationships between features. For example it can give an answer to questions such as, ‘how much of urbanisation has taken place over the years?’ or ‘how mush of forest is being destroyed every year?’ and so on.

Spatial analysis functions

Functions are as follows;

- Simple database query
- Arithmetic and logical operations
- Complicated model analysis

Database query is used to retrieve attribute data. You can either click on the feature or use conditional statements for complex queries. The conditional statement can involve the following operators.
(i) Boolean (logical) operators – and, or, not. xor (exclusive of or)

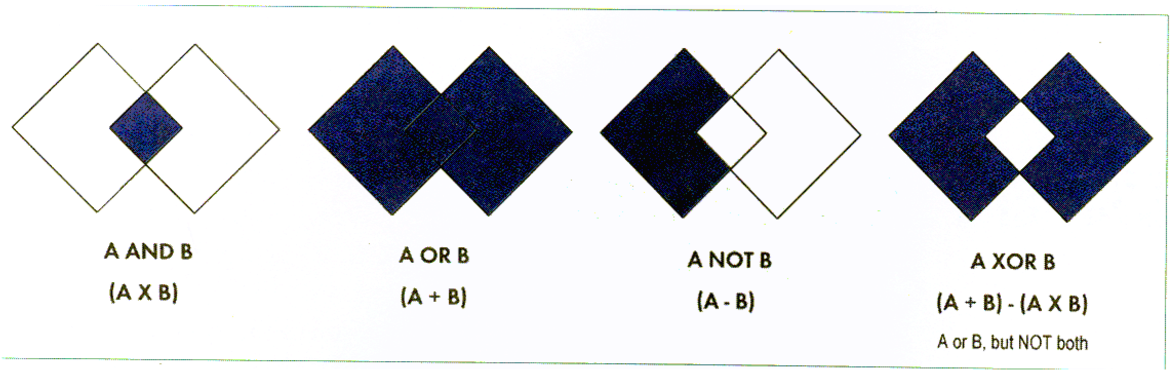


Figure 2.13: Boolean operations

(ii) Relational (conditional) operators - =, >, <, >< (not equal to)

Two types of operations are involved in data transformation.

1. Transformations required for the following:
   a. to remove errors from the data
   b. to bring the data up-to-date
   c. to match the data to other data sets
2. The methods of data analysis required in order to achieve the results required from the GIS.

Can you remember the different aspects of the geographical data?
Yes, they are;

- The Spatial details
- The Topology or the linkages
- The Non-spatial related data

Transformations can operate on the spatial, topological and non-spatial aspects of the data. They can operate separately on these aspects or in combination.

What are these transformations?

Scale changing
Fitting data to new projections
Logical retrieval of data
Calculation of areas & perimeters
Others

The transformations listed in the box are generally found in many kinds of GISs. But some of the 'other' transformations can be very application specific.

It is not intended to deal with transformation functions in detail in the sessions in this course. If you are interested in knowing more about transformation functions required for geographical or spatial modelling, you need to refer additional literature on GIS.

Reclassification
Reclassification means to reassign thematic values. An example of reclassification is, to classify an elevation map into classes with intervals of 100m.

Overlay
A GIS operates on the overlay concept. Overlay is a spatial operation that combines various geographic layers to generate new information.
Overlay is performed using arithmetic, Boolean, and relational operators. It is performed in both vector and rastor domains.

Vector overlay
Features represented by vectors, and which are in many different files, are overlain and combined. Logical rules can be applied to determine how the maps are combined. Vector overlay can be performed on various types of map features. Polygon-on-polygon, line-on-polygon, point-on-polygon are just a few types of overlay. In the overlay process the attribute data associated with each type of feature are also merged. Hence the resulting combined map will contain all the attribute data.

Rastor overlay
Rastor overlay is performed using arithmetic or Boolean operators. Map layers are combined mathematically. The map algebraic functions use mathematical expressions to create new combined rastor layers. The three types of operators are used; Arithmetic, Boolean and Relational.

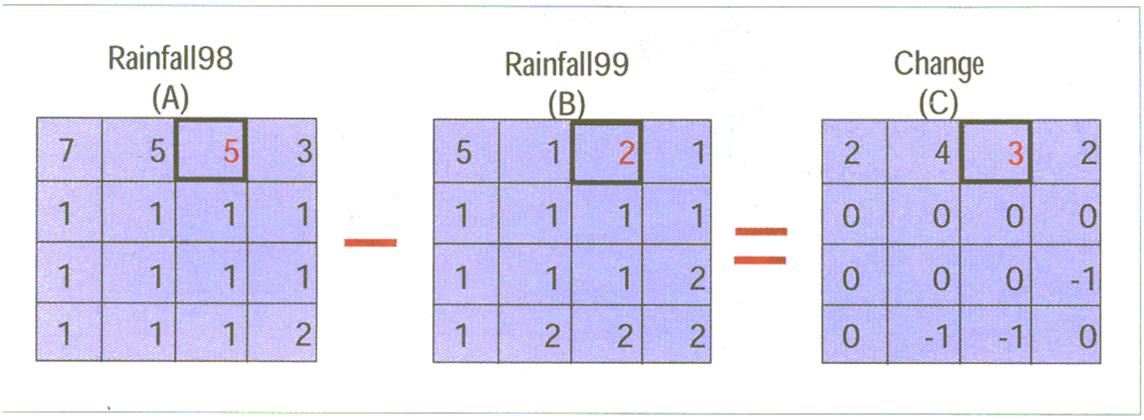


Figure2.14: Raster overlay using arithmetic operators

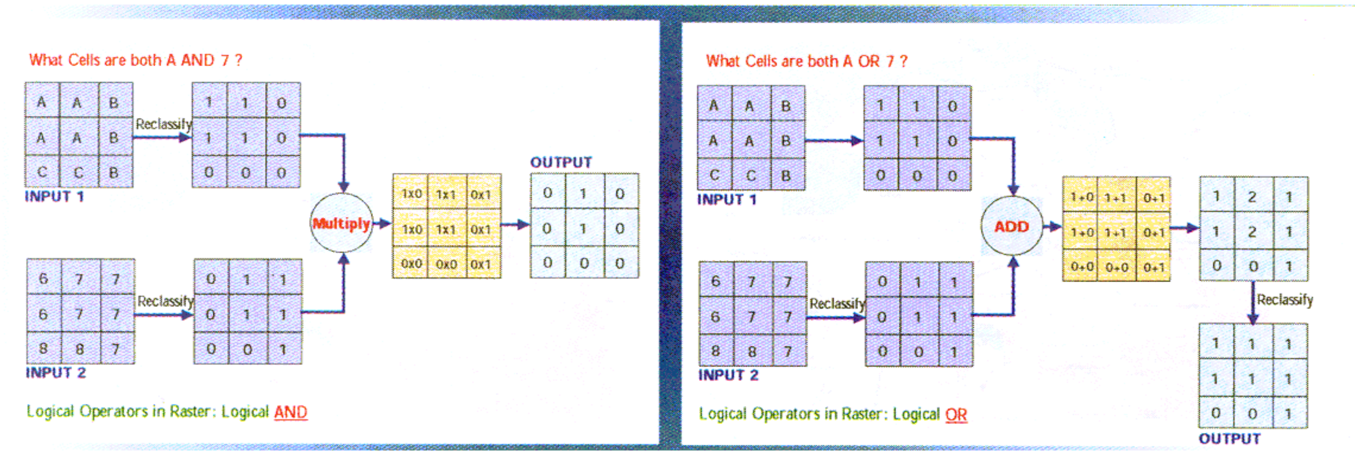


Figure 2.15: Raster overlay using 'and' and 'or' logical operators

Proximity analysis
Proximity analysis is the measurement of distances from points, lines and polygons. An example is to create a buffer zone giving an area with equal distance from a point, line or polygon.

Network analysis
Network analysis is used for the study of movement of resources or utilities from one location to another through a network. A network analysis can be used to identify the critical path or the shortest route for efficient functioning.

A GIS is always required to retrieve data and then to transform the data. So all geographical systems are made up of many complex programs. These can perform all kinds of operations. Hence most geographical information systems provide a range of interfaces by which the user can interact with the system. The simplest are the menu-driven commands. There are also systems where you could type simple commands by means of a command language interpreter, which is really a 'search' command.

But sometimes it may become necessary to write your own computer programs in order to perform certain operations. So some GIS systems provide Macro Languages which are simplified formal programming languages that may be used to link many basic applications together.

It is only now that Geographical Information Systems are becoming more and more important for data & query input and the writing of models for data analysis.

So this science of information has an expandable future to be explored by you!!

2.2.4 Methods of data output & Presentation:

Presenting the information output from a geographical database is a most important function. Proper visualisation of the results helps in taking decisions for planning of resources.

In order to make decisions on locational problems we need to present the information suitably. This will make it easier to take correct decisions regarding important geographical factors. Examples are location of schools, hospitals etc on district population results, location of unstable slopes for stabilising based on soil and topography overlays and many more.

Analysis of spatial data leaves you with resulting combined data in the database. How do we convert this data into graphics?

Graphics will be in the form of maps.
The type of map will depend on the following;
Who is going to use it?
What is its purpose?
What is its content?
What is the scale of the map?
What is the map projection used?
What is the accuracy of the data?

Having done so many manipulations on the digital data related to Earth features, don't you think it is important to display the result as best as possible?

Output data should be understandable. You should also be able to transfer the data to another computer screen.

We shall now see the methods of data output and presentation.

Results can be presented in any of the following forms;

Maps Tables Figures (graphs & charts)

People-compatible outputs of the above presentations are in the form of display on the computer screen, hard copy outputs through printers & plotters.

Computer-compatible outputs are as information recorded on magnetic media in digital form and can be read into another system or be transmitted via telephone lines by means of a modem or optical cables.

The visual display units (VDUs) and plotters can display data in the rastor scanning techniques or vector line-drawing techniques. (see next session for details)

Vector display devices:
Pen plotters – Pen plotters are essentially pen holders and all information is drawn on paper by a series of line-drawing commands.
Vector screens or 'storage displays' – This is similar to above but the paper is replaced by a screen and the pen by a beam of electrons.

Rastor display devices:
Rastor plotters – The simplest rastor plotter is the lineprinter or printing terminal.
Rastor display screens – Here the images are built up from lines scanned across the screen

The interactive graphics workstation:
This is a set of computer hardware for controlling a computer mapping or design system. It is used for interactive design work, for detailed editing and for checking the map before it is printed.

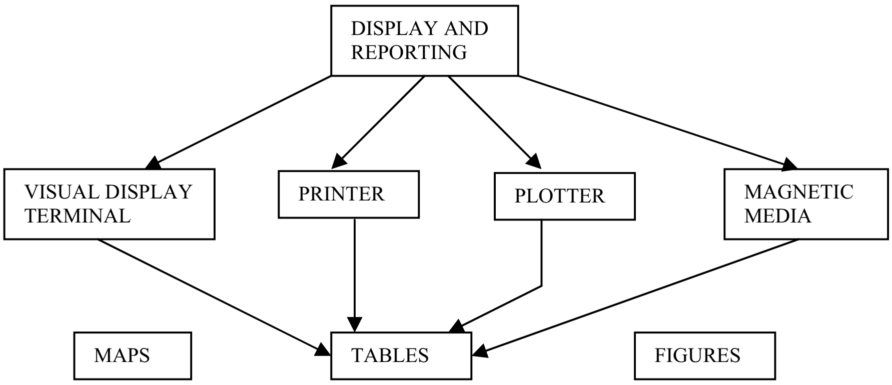


Figure 2.16: Data output

Mapping methods
Various map types can be used for presenting information.

Table 2.3

Mapping method	**Description**
Chorochromatic maps	Nominal values are used for areas with different colours. No hierarchy or order is conveyed.
Choropleth maps	Values are rendered for areas. A hierarchy or order can be perceived between the classes.
Isoline maps	Isolines connect the points with an equal value. They show in which way the value is increasing or decreasing.
Nominal point data maps	Nominal data for point locations can be represented by symbols of different shapes.
Absolute proportional maps	Discrete absolute values for points or areas are represented by proportional symbols. Different values are represented by symbols differing in size.
Diagram maps	Line diagrams, bar graphs, histograms or pie graphs are used on maps to allow comparison between figures or to visualise temporal trends.
Dot maps	Proportional symbol maps which represent point data through symbols. Each denote the same quantity.
Flowline maps	Flowline maps simulate movement using arrow symbols.
Statistical surfaces	These are the three dimensional representation of qualitative data such as used in choropleth and isoline maps.

Computers are being used more and more for visualisation of spatial information.
The results are electronic atlases, cartographic animations, and multimedia systems.

Multimedia allows for interactive integration of sound, animation, text and video. This new technology makes it possible to link geographical information such as text documents, photographs in a GIS database, or a video clip of the landscape of the study area.

Maps are also found online on the internet. Maps can represent spatial data, they can function as an index of spatial data, a preview of data and as a search engine to locate spatial data.

2.2.4 Implementing a GIS:

To properly run a GIS you need hardware, software, data, people and policies and procedures.

We have seen what the hardware and software components are. GIS software runs on a wide range of hardware types, from centralised computer servers to desktop computers. GIS software provides the functions and tools needed to store, analyse and display geographic information. Key software components are as follows.

- Tools for the input and manipulation of geographic information
- A database management system (DBMS)
- Tools that support geographic query, analysis and visualisation
- A graphical user interface (GUI) for easy access to tools

Data for the GIS is the most costly component.

Building the database takes the most time, costs and database management. The data collected by different organisations should be shared in a manner so that there is no duplication of data collection. This is essential for efficiency. There should be willingness to share data. GIS as a technology will be viable and cost effective only of data are available at an affordable cost.

A GIS needs many types of specialists to run it. Computer and information technology experts are required for system support. Domain specialists are required for application specific models and analyses.

Geographical information systems designed for different organisations will have their different functional requirements.
How do we choose the right GIS?
The right GIS for a particular implementation will involve matching the GIS needs to the functionality demanded by the type of application required by the organisation.

Session 3

Geographical Information

3.1 Describing geographic phenomena

We have understood by now that, geographically (spatially) referenced information is described as geographic information.

Geographical information is made up of two components.

- Location information of geographic entities
- Their characteristics or attributes

3.1.1 Geographic data

Geographic location is expressed in a two-dimensional frame, by two coordinate axes on a horizontal plane. They can be X and Y. The horizontal coordinate axes are defined along the east-west and north-south directions.

One vertical axis (Z), together with the X and Y coordinates will define the location in a three-dimensional frame. The vertical axis will represent the elevation of the point. The vertical dimension may be defined by any other attribute as well. For example temperature and rainfall can replace elevation as the vertical dimension. The third dimension can be used to generate a three-dimensional surface, which shows the continuous variation of a value attached to the third dimension.

Most geographic phenomena can be described under the three major feature (entity) types.
They are as follows:

- points
- lines
- polygons (areas) and

Point, line or polygon data can be used to generate a three-dimensional surface. When elevation is used as the third dimension, the surface generated is called the digital elevation model (DEM) or digital terrain model (DEM).

Major feature (entity) types: **Lines, Polygons, Points**

These are the three data primitives

Any feature (entity) can be reduced to these three data primitives.

In addition to the geographic information, a GIS would also include non-geographic or **attribute data.**

3.1.2 Non-Geographic or Attribute data

A GIS also includes non-spatial or attribute data, which represents some property of the spatial entity.

In other words attributes are the characteristics of the geographical entities. They form the non-geographic data components of a GIS.

Attribute Data Types in a Spatial Information System fall into one of the following categories.

Attribute Data Type	Description	Examples
1. Nominal	Qualitative with discrete states but no specific order. No numerical value.	Rock type (granite, marble, gneiss) Soil type (alluvial,
2. Ordinal	Variables are listed as discrete classes and are ordered. Numerical value can be assigned to each state to show the magnitude. The value assigned will give the relative magnitude of the attribute but the distance between states is not defined.	Soil drainage (poorly drained, moderately drained, well drained) Numerical codes can be assigned as 1,2,3
3. Interval	Has a natural sequence and the distance between values has a meaning. Data zero is arbitrary	Temperature in C or F scales. The Zero in these scales is arbitrary.
4. Ratio	Same as interval but has natural zero.	Temperature in Kelvin. Here 0°K means absolute zero.
5. Narrative descriptions	Identification by given names	Place names on maps such as Colombo, Nawala etc.

Given below are geographic entities as indicated by point, line and polygon, with some of their possible attributes.

Feature (entity)	Attributes
Mountain (Point) Stream (Line) District (Polygon)	Name, height above MSL etc. Name, length, discharge etc. Name, area, population etc.

3.2 Geographical Space

Data models are used for describing geographical phenomena. Real world phenomena need to be well represented by these conceptual models.

We must remember that GIS is only a tool, which helps you to manipulate the data related to the real world and to people. So how do we obtain this real world data?

The first step is to observe the world and to be able to understand the phenomena that are either fixed or change with time.

We can observe the real world in two ways.

- either directly with what we see with our eyes, by manually analysing statistically located samples or
- with the use of aerial photography, and remotely sensed images.

- ***The most important thing is for you to be able to observe the world and decide what needs to be done. It is your decision that is important.***
- ***The use that you make of the computer systems will only help you with manipulations of data, which will be very difficult manually.***
- ***The digital systems will also provide the results in presentable forms, which will help you to make appropriate decisions easily.***

So be wise when using a GIS!

We spoke of using conceptual models for describing real world geographical phenomena.
Let us see what these are.

There can be two descriptors to represent the real world.

What is present and **where** it is present.

Phenomena are recognised and described in terms of either 'objects' or 'entities'

How do you reference the phenomena in space?
It can be a

- geometrically exact location or a
- relative location.

The geometrically exact location is given by projecting the Earths surface onto a flat surface using a standard system of spheroids. The coordinate system used may be based on an internationally accepted projection using latitudes and longitudes or it can be a national grid based on a local origin.

(For the maps prepared by the Survey Department of Sri-Lanka, the projection is the Transverse Mercator Projection and the origin of the Metric Grid System is 200,000metres South and 200,000 metres West of Pidurutalagala. We shall not study projections in this course. You may refer a text book on Surveying for information.)

The relative location can be with relation to any point of reference.

What is Geographical Phenomena?

They are 'What is present in space and where it is present'.

How do you describe data models?

Data models can be **objects** or **entities**

How do you reference the phenomena in space?

It can be a geometrically exact location or a relative location.

3.2.1 Conceptual models used to perceive geographical phenomena:

How can we formalise the representation of space and spatial properties?

When we consider any space such as a room or a landscape, how can we describe all what is in that space?

Let us think of conceptual models to represent the data.
We can think of the following two extreme possibilities. The data can be identified,

1. As separate entities, which are described by their attributes or properties.
2. As a continuous mathematical function or field.

Entities:
Space is occupied by 'objects' or **'entities'** as we shall call them from now on.
(note that 'object-orientation' will refer to a way of structuring data in the computer or in a computer program)
First you must recognise the entity. Is it a house or a tree or a river?
The next step is to list the attributes, define the boundaries and the location of the entity.

Continuous fields:
Space is occupied in terms of Continuous Cartesian Coordinates in two or three dimensions (four if 'time' is included). The attribute is assumed to vary smoothly and continuously over that space. When there are clusters of like attribute values in geographical space or time, these zones will be recognised as 'things'. For example if we consider the elevations above mean sea level, clusters of like elevation values will become 'mountain peaks'.

Is it an entity or a continuous field?
How do we decide on an entity model or continuous field approach when the entities are seen as sets of extreme attribute values clustered in geographical space?

Think of a set of mountain peaks. If we decide on the entity model we can record who climbed each mountain-top but we cannot record the information for computing the slopes of the hills. But if we decide on the continuous fields approach, calculation of slopes can be performed as the first derivative of the surface.

The choice of the approach, whether 'entity' or 'continuous field' will depend on the application.

3.2.2 Geographical data models: Entities and Continuous Fields

We have seen the conceptual models with which we explained the objects in space.

Now let us see how they can be formalised as geographical data models.

Geographical data models can describe phenomena in space as follows;

- Phenomena can be uniquely identified
- It is possible to identify parts for analysis and communication
- Attributes attached to the phenomena can be measured or specified
- Geographical coordinates can be registered
- The level of resolution, which will be an indication of the accuracy of the observation or measurement, will be given by the way the data was collected.

There are basically three types of abstract models in geographic space.

- **Entity or object model-a discrete model**
- **Continuous field model**
- **Network based model- a discrete model**

We have already spoken of the 'entity' or 'object' model, which is a discrete model and secondly, the 'continuous field' model. The third type of model in geographic space, the 'network' based models are similar to the object based model.

The geographic phenomena in the network model are also discrete. The interaction between the different objects is given by the network model. Phenomena such as traffic movement, telecommunication networks, water supply lines etc. are examples of models where the flow paths are of interest.

- **Geographic space can be represented by three types of abstract models.**
- **These models represent spatial phenomena as objects, fields and networks.**

- **A continuous field model is a view where space is everywhere defined. A three dimensional terrain is an example of a continuous geographic space. (continuous field)**

- **A entity or object oriented (discrete) model is a view where space is not everywhere defined. A road network is an example of a discrete (entity-oriented) space.**

Most phenomena in space such as buildings, roads, agricultural fields, etc. are described as entities and further specified by attributes and geographical location. This can be further sub divided according to one of the three basic geographical data primitives, namely a point, a line or an area.

These are the fundamental units of the vector data model.

A method of representing continuous fields is by means of tessellations of regular-shaped polygons, by using sets of pixels.
The tessellations are reasonable approximations of reality and it is assumed that operations such as differentiability which can be applied to continuous mathematical functions can also be applied to these discretised approximations.

Table: Discrete data models for spatial data (you will read more about these in the session on Data Structures)

Vector representations of exact entities	Tessellations of continuous fields
Non-topological structures (loose points and lines 'spaghetti')	Regular triangular, square or hexagonal grid (square pixels = rastor)
Simple topology with linked lines A drainage net or utility infrastructure is an example	Irregular tessellation: Thiessen polygons
Complex topology with linked lines and nested structures Examples are linked polygons	Triangular Irregular Nets (TIN)
Complex topology of object orientation with internal structures and relations	Finite Elements
--------	Nested regular cells/ quadtrees irregular nesting

Both the ‘entity’ and ‘tessellation’ models assume that the phenomena can be specified exactly in terms of both their attributes and spatial position.

Geographic phenomena are set in a specific time-frame. Several layers of data may be used to represent the time dimension. One example is, the land use pattern of a given area. The land use pattern can be found through analysis of remotely sensed data.   By analyzing image data of images taken at specific times over the years, we can prepare the land use plans for the area. The variation of land use with time will be represented in several layers, each layer representing a situation at a particular time.

We can see that a point in space can have a time dimension.

> Continuous space model: (example is a three dimensional terrain)
>
> Think of a point in three-dimensional space. This point is represented by (x,y,z). It can also have a time dimension (t) . It can have other attributes as well (a).
>
> How can we represent this point?   This point can be given by {a,x,y,z,t}

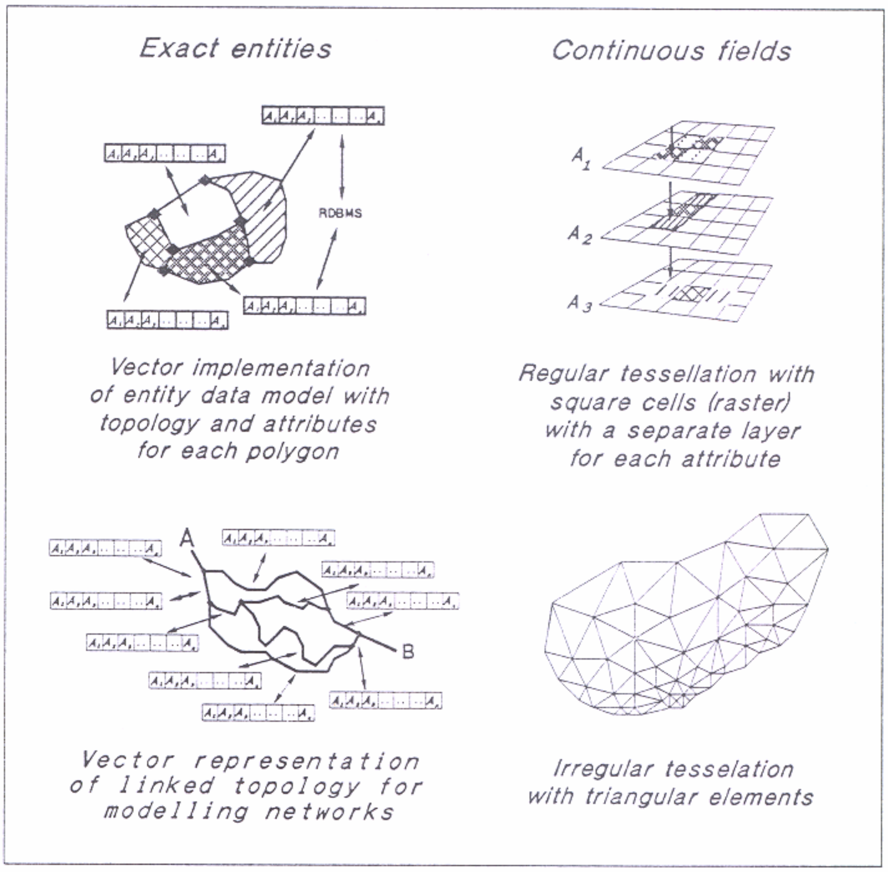


Figure 2.4. The encoding of exact objects (entities) and continuous fields in different data models. (*a*) top left: vector representation of crisp polygons; (*b*) top right—raster model of continuous fields; (*c*) bottom left—vector representation of linked lines; (*d*) bottom right—Delaunay triangulation of a continuous field

Figure 3.1: The encoding of 'entities' in vector data models and 'continuous fields' in tessellation data models.

> Entity-oriented (discrete) model: (example is a road network)
>
> In this model, space is not everywhere defined. A point in space can have a time dimension. (t). It can have one or many attributes (a). It will have entities (e). It will have positional information.
>
> How can we represent this point? This point can be given by {e,a,s,t}
>
> Polygon layers are also discrete models since their boundaries are abrupt.

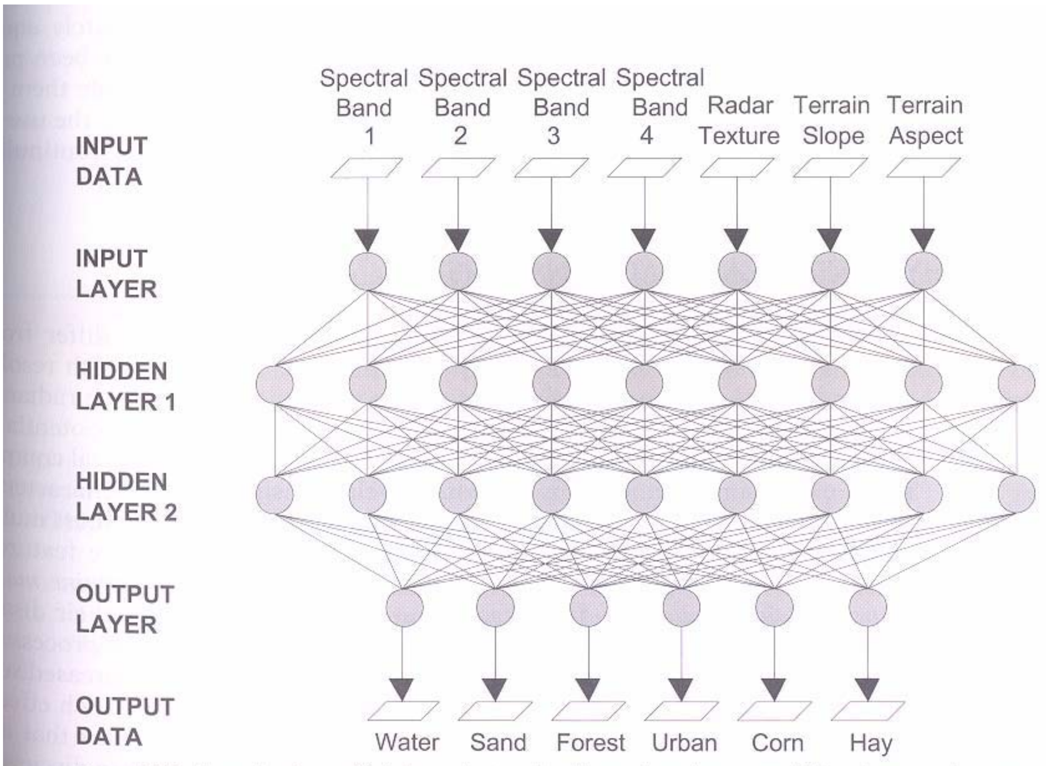


Figure 7.60 Example of an artificial neural network with one input layer, two hidden layers, and one output layer.

Figure 3.2: Example of an artificial neutral network with one input layer, two hidden layers, and one output layer.

3.2.3 Definition of a map

What is a map?
How are data represented in maps on paper?
How are data represented in geographical information systems?

A map is a set of points, lines, and areas that are defined both by their location in space with reference to a coordinate system and by their non-spatial attributes. A map is usually represented in two dimensions. The map legend links the non-spatial attributes to the spatial entities. Non-spatial attributes are indicated visually by colours, symbols or shading. Regions by colour, objects by symbols and so on.
In geographical information systems, the non-spatial attributes are coded on a form that can be used in data analysis. In this case a region will be a set of pixels, which form areas or polygons and described by a single legend unit. A region may be made up of several discrete occurrences, which may be uniform or which contain polygons belonging to regions of another kind. The eye can distinguish topological relationships shown in fig. d-f, but when using the computer for analysis, the relationships have to be indicated in digital form.

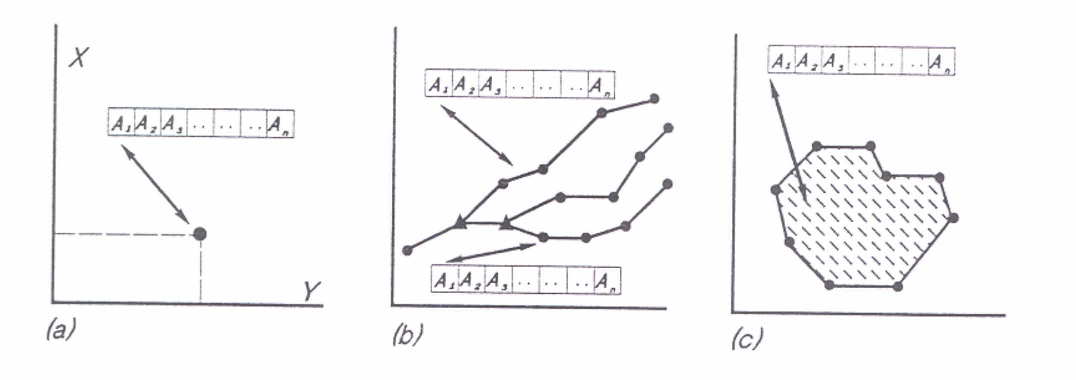


Figure 3.3: A map indicating points (a) lines (b) and polygons/areas(c).

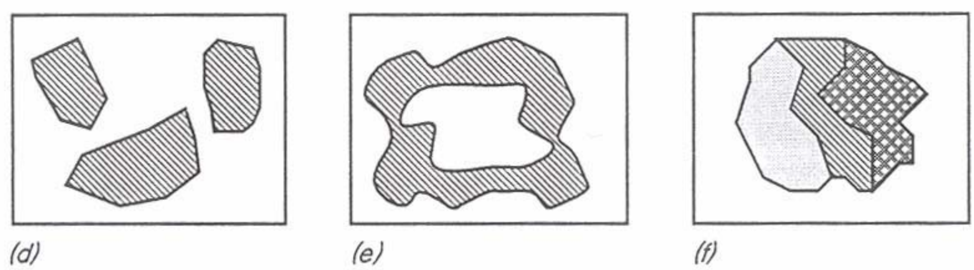


Figure 2.3. The fundamental geographical primitives of points, lines, and polygons

Figure 3.4: A map indicating regions of different types.
Disjoint region (d)
perforated region (e)
three different regions (f)

3.2.3.1 Scale and resolution of a map or image

We have seen how a view of the real world is represented by a map drawn to scale.

How do we determine the scale to be used for a particular map?

We need to see what information is available and consider the intended use of the map.

So we can now see that the information content of a map depends on its scale. Large scale maps contain more information about a unit area than a small-scale map.

Scale of a map or image is the ratio between map distance and the ground distance.

Resolution of a map is another concept related to the scale. Resolution will define the smallest size of entities that we can distinguish on a map or image. Usually larger the scale of the map, higher the resolution. As the number of observations per given area increases, the resolution also increases. The required resolution depends on the size of the object to be studied.

The Resolution of a map/image will define the smallest size of an entity that we can distinguish on it.

To compare and analyse information on maps the maps under study should be of the same scale. But generally the information is given on map sheets or images of varying scale. How can we get all the maps/images to one scale?
Reduction and enlargement of maps can be done using a GIS and this is called map generalization. But since the maps which are scale changed in this manner do not change in the information content to suit the scale change, we have to be cautious in this exercise. It is always better to use maps of nearly similar scale to use for analysis in a GIS.

- Be careful in generalizing maps of different scales before analysis in a GIS.
- Always try to use maps of nearly similar scale for computations in a GIS.

3.3 Geographical data types

Some important properties of geographical/spatial entities:

- Spatial entities are phenomena that cannot be sub-divided into like units. A district can be divided into villages, but the new units will not be districts. This means that a sub division will result in the loss of the entity's identity.
- Spatial entities can be mapped. These are rivers, roads etc. Spatial entities can also be statistically defined. These are population zones, soil types etc.
- Spatial units are homogeneous only in terms of the properties that are used to define the unit. These units may not be internally homogeneous entities in all their properties. A unit mapped as paddy land may not be homogeneous with respect to the soil type within the unit. Hence it will be heterogeneous with respect to soil.
- A spatial unit may or may not be variable with time.

Information about geographical/spatial entities can be categorised into five classes (Laurini & Thompson)

(1) An identifier (name, code number or other)
(2) Location coordinates
(3) Character of the entity/attributes
(4) The function/role, behaviour/change over time
(5) Spatial properties of the entity (shape, size, volume, topological characteristics such as relative position & adjacency)

Location of geographic entities is represented in the database by one or many pairs of coordinates depending on the feature type.

Feature	Coordinate pairs
Point	One
Line	Two or more
Polygon	Three or more

A point is represented as a pair of X-Y or X-Y-Z coordinates.

A line is a string of X-Y coordinate pairs.

A polygon is a string of X-Y coordinate pairs where the start and end nodes are identical.

How do we decide on the entity type (point, line or polygon), for a particular geographic phenomena?

Let us take a 'town' as an example. The way a town is represented on a map will depend on the scale of the map. A town may be represented as a point on a world map and as a polygon on a one-inch topographic map.
Similarly if we consider a 'road' it will be represented by a line on a world map and as a polygon giving the width of the road in another map.

A point entity is dimensionless. The line and polygon features described above are two dimensional. There are also three dimensional surfaces which can be considered as another type of feature (entity).

Example of a Point Entity:

Information about Geographical/spatial entity	Data
Identifier	Open University, Nawala Campus
Location	X-Y coordinates
Character of entity: Number of Faculties Number of students	 4 20,000
Role/function of entity	Education (in this case it will not change with time)
Spatial properties: Nearest town Land area	 Colombo

3.3.1 Types of Spatial Entities

3.3.1.1 Spatial entities can be classified on the basis of their dimensionality.

Dimension	Entity Type	Example
Zero-dimension	Point entity(no area)	Buildings, Towns
One-dimension	Line entity(no width)	Streams, Roads
Two-dimension	Polygon entity(closed area)	Land use, Administrative divisions
Three-dimension	Surface entity ,Volume	Terrain
Four-dimension	Space-Time	Landscape evolution,

Most GIS software can handle three-dimensional data.

Three dimensional nature of the terrain is represented on paper by contours or shaded relief. GIS software used for 3-D modeling can convert 2-D surface into a 3-D terrain model. (Digital Elevation Model or DEM).

Applications of DEMs

- What can these models be used for?

They can be used for calculating terrain parameters such as slope, aspect, inter-visibility, volume of material etc.

- Where can these be applied?

They can be very valuable for terrain studies, such as geological/geomorphological mapping, military planning, development activities and modeling landscape processes.

The fourth dimension, Time, cannot be handled by the most of the present software.
The change in geographical phenomena over time is an important factor to be taken into consideration in planning of future activities.

3.3.1.2 Spatial entities can be classified based on <u>topology</u>

- What is topology?

Topology is the linkage or adjacency of the geographical information with respect to the surrounding.

Topologically spatial entities can fall into one of the following three categories: (see figure below)

1. Simple polygons - a single unit of space bounded by three or more lines
2. Complex Polygon - a polygon containing one or more holes
3. Regions - consisting of two or more (simple or complex) separate polygons

Figure 3.5 : Topological Spatial units(workshop manual)

3.3.1.3 Five types of geographical data as identified by Laurini and Thompson

1. Real data: Entities that are directly observable terrain conditions such as rivers, drainage basins, cliffs, buildings, roads, etc.
2. Captured data: Data captured using physical devices such as cameras, satellite instruments, weather instruments etc.
3. Interpreted data: Data obtained by interpreting the observations from instruments. An example is seismic signals.
4. Encoded data: Data from maps. Most of the data used at present in GIS come from digitized maps.
5. Structured data: This is data that come in tables. Attribute data of most types fall into this category.

3.4 Data storage and operations of a GIS

What are the methods of storage of data in a GIS?

Layers: Various themes of information are stored separately in what are called 'layers'. It is similar to drawing on a transparency and overlaying them.

? Figure 3.5 A GIS analysis procedure (This one is for studying potential soil erosion)

- Rastor data model: The area is divided into small regular blocks, usually squares which are called pixels. Each pixel will have a value attached to it. Each variable in the data set will be indicated in a different layer.

- Vector data model: Geographic objects are described in terms of geometric elements such as points, lines, polygons, areas and volumes. Different data sets will be stored in different layers.

The different layers in a GIS are stored in the form of what is termed as **'shape files'.**

3.4.1 Operations of a GIS

You can observe the following operations using a GIS software, when you come for the practical session.

- Point-in polygon queries: When you click the 'mouse' while inside the region of a computerised map, the program must perform a point-in-polygon test to find out what you are interested in.
- Proximity-based queries: For example if you need to know the number of schools lying within a certain distance from a point, you can create a buffer zone to obtain the result.
- Network queries: To get at the optimal routes for travel, water pipeline or any other. It can also be network load queries where you try to find out the response of a drainage system to heavy rainfall etc.
- Thematic maps: These are methods of presenting information for quick and easy reference. Just having a table of numbers may very often not indicate some important point.

Spatial Operations & Data Linkage

Spatial Operations

By now it should be clear to you that a GIS is not only a database but also permits spatial operations on the data.

For example say we have a table of data giving the towns in Sri Lanka with their locational coordinates and the population of each town.

- aspatial query - What is the population of each city? Answering this question is simple and does not require location data or any other computation.

- spatial query - What are the towns situated within 50 Km of each other? In order to answer this question you need to manipulate the data and demarcate boundaries that lie within 50 km of each town.

Data linkage

A GIS will link different data sets.

Why is data linkage so important?

Link between Entities and Attributes: (figures from workshop manual)

Figure 3.6: Point Entity and link with attributes

Figure 3.7: Line Entity and link with attributes

Figure 3.8: Polygon Entity and link with attributes

You may have several data sets for a given area. Each data set can be analysed/mapped separately. OR the data sets can be combined. By bringing the data sets together, you add value to the database.

To do this combining operation you need a GIS.

Exact matching:
Say you have information in one computer file about many geographic features eg. districts of Sri Lanka. It can give the population in each district.
You may also have additional information in another file about the same set of features. It can be the number of schools in the district.
The operation to bring them together can be achieved by using a key common to both files. The record in each file with the same district is extracted and the two are joined and stored in another file.

District	Population
Colombo	3,000,000
Kandy	800,000
Gampaha	1,000.500
Kurunegala	400,000
Kegalle	600,000

District	Number of Schools
Colombo	500
Kandy	400
Gampaha	400
Kurunegala	300
Kegalle	350

These two tables are joined as follows.

District	Population	Number of Schools
Colombo	3,000,000	500
Kandy	800,000	400
Gampaha	1,000.500	400
Kurunegala	400,000	300
Kegalle	600,000	350

Hierarchical matching:

For certain types of information, that for the whole district may have to be obtained by adding the data obtained for different grama-sevaka divisions within the district. So then we have to first add the data of subsets of data to arrive at the data for the district. *Here the smaller grama sevaka areas will fit exactly within the larger district boundaries. This is called hierarchical matching.*

Once the district data is obtained through this matching, an exact match can then be performed.

Fuzzy matching:
On many occasions the boundaries of the smaller areas do not match those of the larger ones. This is especially true when dealing with environmental data. For example crop boundaries, usually defined by field edges, rarely match the boundary between types of soil. If your objective is to determine the most productive soil for a particular crop, you need to overlay the data sets of crop productivity and soil types, and compute the crop productivity for each type of soil.
Fig of overlay

A GIS can do these overlay/computing operations because it uses geography or space as the common key between the data sets. Information is linked only if it relates to the same geographical area.

GIS Session 4

Geographical Data Structures

Overview:

We have said that GIS deals with geographically referenced or in other words spatially referenced data sets.
GIS also deals with attribute data related to the above data.

The two data types are linked by means of key variables.

Now let us see how the data can be arranged and manipulated in the computer.

4.1 Geographical data in the computer
Database structures

What are the forms of capturing digital spatial data, which can be used in the computer?

There are many methods of capturing digital spatial data.
The methods are as follows:

- Digitizing of existing analog maps
- Aerial photographs (analytical photogrammetry, digital photogrammetry)
- Global Positioning System (GPS)
- Ground survey using Total Station and GPS
- Reports and publications
- Satellite Remote Sensing

The GIS has to perform the efficient storage, retrieval and manipulation of this data. It is here that the difficulty lies!

There are certain requirements, which any type of database system must satisfy. They are;

(i) Provide a standardised format and a set of standards and procedures for recording data
(ii) Allow efficient storage and retrieval for all the intended data
(iii) Provide independence of the storage format of the data from all programs that use the storage format of the data
(iv) Ensure non-redundancy of the data
(v) Protect the data from accidental loss through hardware failure or human error
(vi) Provide ease of system use
(vii) Provide economy of system use

Now we shall try to understand the concept of data structures.
Phenomena in space can be perceived by us as geographical data, and stored in a manner we are used to, such as the maps we have been using over the years.
But this representation may not be the best way to structure the data. Data manipulation and retrieval becomes easier with well-structured data sets.

Whatever way we decide to represent the data we have to take into account the fact that the computer will address the data only in a certain format. So we have to convert the data to that format.

Let us first see how data sets can be organised.

What do you think should be the essential features of any data storage system?

- **It should be easy to access the data and to make cross-references.**

This will involve a database management system. Computer programs are used to control data input, output, storage and retrieval from a digital database. We must see how data can be organised in files and how these files can be accessed easily.

4.1.1 File structures:

If we need to organise our data in such a way so as to be able to retrieve it quickly. How do we set about it? Let us examine some types of data files.

(i) Simple lists of data – Here you just add the data to the bottom of the list. If you have n items on the list, it will take (n+1)/2 search operations to find the item you want.

(ii) Ordered sequential files – Same as having in alphabetical order. You have to create room to input a new item, but retrieving is much easier than in the earlier case. Files can be accessed by binary search procedure, which takes $\log_2 (n+1)$ steps for a list of n items.

(iii) Indexed files – Here we have direct files and inverted files. In a direct file, the data items in the files provide the main order of the file. The location of items in the main file can also be specified according to topic, which is given in a second file known as the inverted file. If we compare with a standard text book, this will be like looking up the subject index to get the required pages. See table for Indexed files.

Table 2.1 Indexed files

Direct files

Index		File item
Item key	Record No.	
A	1	A_1
B	n_a+1	A_2
C	n_a+n_b+1	⋮
.	.	B_1
.	.	⋮
.	.	C_1
		⋮

Inverted files

Soil profile number	Attributes					
	S	pH	De	Dr	T	E
1	A	4	deep	good	sandy	—
2	B	5	shallow	good	clay	yes
3	C	6	shallow	poor	sandy	no
4	D	7	deep	good	clay	yes
5	E	4	deep	poor	clay	no
6	F	5	shallow	poor	clay	no

S = Series, De = Depth, Dr = Drainage, T = Texture, E = Erosion.

Index (inverted file)

Topic	Profiles (sequential numbers in original file)					
Deep	1			4	5	
Shallow		2	3			6
Good drainage	1	2		4		
Poor drainage			3		5	6
Sandy	1		3			
Clay		2		4	5	6
Eroded		2		4		

Table 4.1: Indexed files

Indexed files provide rapid access to databases. But addition and deletion of files is not easy. In the case of direct files the file and index has to be updated and in the case of inverted files the index has to be updated each time new data is added or data is deleted. File modification becomes an expensive process.

Now let us look at some database structures, which will help access data from one or more files easily.

4.2 Database structures:

There are basically three types of database structure. They are as follows;

- Hierarchical
- Network
- Relational

In all these types of database structures the data are written in the form of records.

4.2.1 Hierarchical Data Structure:

A hierarchical system is used when the data have a one-to-many relation. Data can be in a ‘tree’ system. Travel within the database is restricted to the paths up and down the ‘branches’.

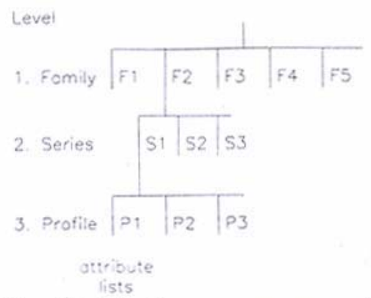


Fig. 2.2 Hierarchical database structure commonly used in soil science.

Figure 4.1: Hierarchical database structure commonly used in soil science

Hierarchical systems assume that each part of the hierarchy can be reached using a set of discriminating criteria (key) that fully describes the data structure. It is also assumed that there is good correlation between the key attributes, which are the discriminating criteria, and the associated attributes.

Data access is easy for key attributes but very difficult for associated attributes. This type of database is good for banks, airlines etc. but is not suitable for systems where the retrieval cannot be based on a rigid hierarchy such as say for environmental data.

Another disadvantage is that large index files have to be maintained and certain attribute values may have to be repeated many times, leading to data redundancy. This leads to unnecessary storage and access costs.

4.2.2 Network systems:

A rapid linkage between data is needed very often which cannot be provided by the hierarchical structure. Examples are when adjacent items in a map or figure need to be linked together. The actual data about their coordinates may be written in a different part of the database. In such a situation a network system is useful.

Consider the figure below. Figure (a) gives a map of two adjacent polygons defined by a set of lines, one of which is common to both. The ends of the lines are defined by coordinates.
Think of how you would draw a hierarchical data structure for this case. You can see how much of data becomes redundant in the process.

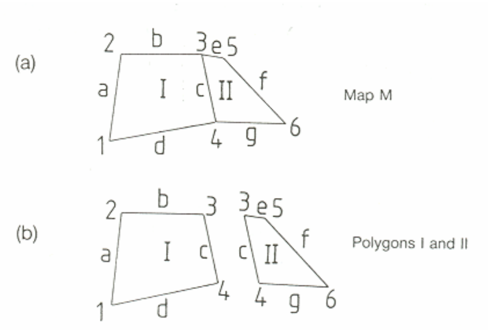


Figure 4.2: A map M consisting of two polygons

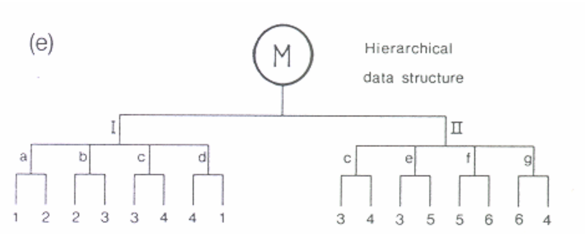


Figure 4.3 : A hierarchical data structure for the adjacent polygons

These problems are avoided by the compact network structure shown in figure below. Each line and each coordinate need appear only once.

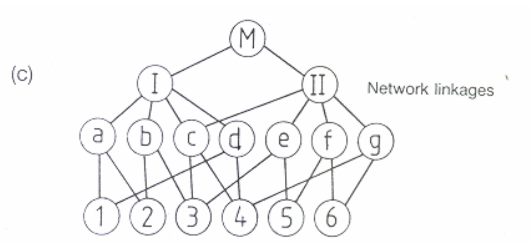


Figure 4.4 : Network linkages for the adjacent polygons.

Networks having a ring pointer structure (arrow networks) are often used when navigating around complex topological structures. Here the relationships are specified and there is no data redundancy. But the database can get overloaded with the pointer system and also each time a change is made the pointer system has to be altered.

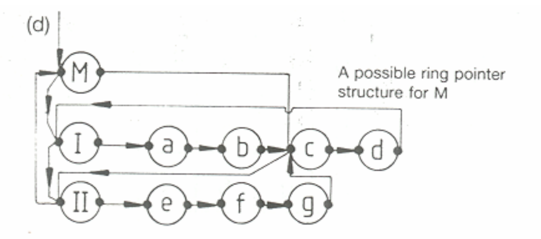


Figure 4.5: A possible arrow network for adjacent polygons.

4.2.3 Relational Database Structures:

In this type of database structure the data are stored in simple records, known as *tuples,* containing an ordered set of attribute values that are grouped together in two-dimensional tables, known as *relations.*

Map

M	I	II

Polygon

I	a	b	c	d
II	c	e	f	g

Lines

I	a	1	2
I	b	2	3
I	c	3	4
I	d	4	1
II	e	3	5
II	f	5	6
II	g	6	4
II	c	4	3

Fig. 2.4 A relational data structure for the map M.

Figure 4.6: A relational data structure for the adjacent polygons.

Each table or relation is usually a separate file. When extracting the data from the files, the user defines the relationship of the data that is appropriate for the query. The required data form may not be already present in the database. Boolean logic and mathematical operations are used to formulate the queries. The relational data structure is very flexible and can meet the demands of all queries. But there has to be a sequential search through the data and this can take time.

Adding and removing data is done by adding and removing a tuple.

- Record structures:

Note that in all the database structures described so far, the data are written in the form of records.

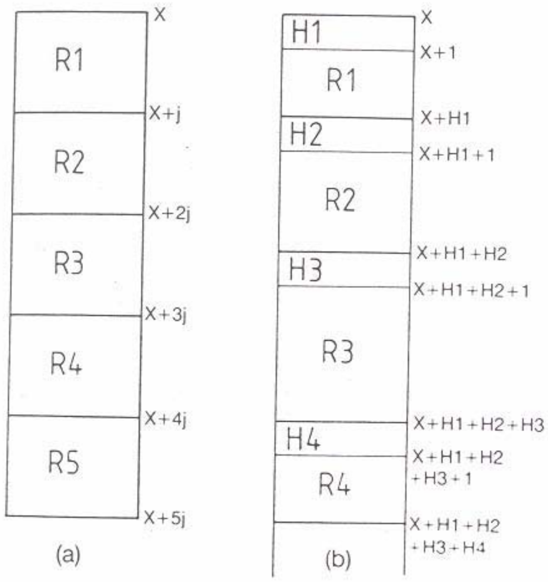


Fig. 2.5 Two kinds of sequential list: (a) fixed length records; (b) variable length records including a 'header' H to record data about the record itself.

Figure 4.7: Two kinds of sequential records. (a) fixed length records (b) variable length records including a 'header' H to record data about the record itself.

4.3 Computer representations of geographical data: Raster & Vector format

Two data structures are used to store geographic data. They are,

- Raster and
- Vector

Characteristics of Raster data structures:

- Treats space as a continuous surface and hence it is coverage oriented
- Divides the map into grid cells and only the grid cell value is stored
- Requires more memory space
- Data is stored as a matrix with the origin of the data layer at the top left hand corner.

Characteristics of Vector data structures:

- Identifies the geographical entities and hence if feature oriented
- The x-y coordinates of the lines and points in the coverage are stored
- Requires only a fraction of memory for storage
- The x-y coordinates define the geographical entities explicitly and the origin is at the bottom left hand corner.

4.3.1 Raster Data Structure

Raster data structures can be used for all types of geographic entities.

In a raster data structure the data is stored in a two-dimensional array (data matrix) with rows and columns. Data values are assigned to every cell. The accuracy of raster maps depend on the cell size. The smaller the cell size the greater is the accuracy. But since this will result in a greater number of cells to cover a given area, the amount of data to be stored will be increased. The storage requirement is directly proportional to the number of cells in the image.

The cell value can take the following forms:

- 1 or 0
- A discrete class value (For example, a different value for each soil class)
- A value representing the magnitude of the cell (Example, elevations)

The size of the raster cell determines the size of the smallest feature that can be represented. When deciding the minimum mapping unit, a rule of thumb is to use a raster cell half the length of the smallest feature you wish to record.

The Raster data structure is a quadtree structure.

4.3.2 Vector Data Structure

Vector format gives more precise locational information about geographical entities and requires less storage space.

There are several vector data structures. Whole polygon structure, Dime structure, Arc-Node structure, Relational structure and Digital line graph (DLG) are some of the structures.

The table below gives a comparison of Raster and Vector models.
Table 4.

	Raster model	Vector model
Advantages	• It is a simple data structure • Overlay operations are easily & efficiently implemented • High spatial variability is efficiently represented in a raster format • Manipulation & enhancement of digital images can be performed effectively in the raster domain	• Provides a more compact data structure than the raster model • Provides efficient encoding of topology & hence suitable for operations that require topological information, such as network analysis • The vector model is more suited to supporting graphics that closely approximate hand drawn maps
Disadvantages	• Data structure is less compact & data compression techniques are required • Topological relationships are more difficult to represent • The output will have a blocky appearance. To reduce this effect the cell size has to decrease. But the increased cells will then result in large files of storage	• A more complex data structure than a simple raster • Overlay operations are more difficult to implement • The representation of high spatial variability is inefficient • Manipulation & enhancement of digital images cannot be effectively done in the vector domain

The figure below shows how the real world can be represented as Raster and Vector data models.

Figure 4.8: from Leka's sheet

The figures below show two ways in which a chair can be explicitly or implicitly represented in a computer.

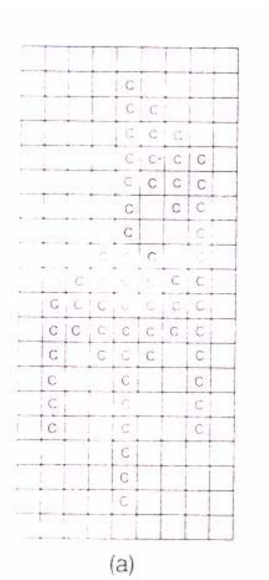


Figure 4.9: An image of a chair explicitly represented in a rastor or grid-cell format.

Each cell will be represented by a numerical value or a colour or gray scale.

The Rastor data structure for the chair would be;

Chair attribute → symbol/colour → grid cell

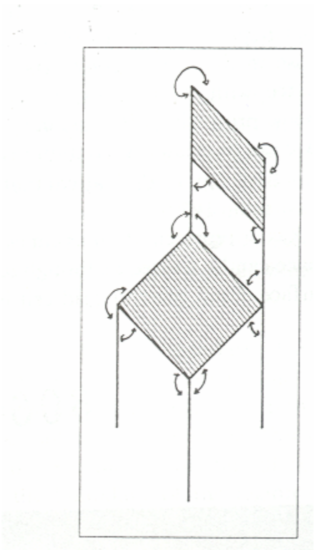


Figure 4.10: An image of a chair implicitly represented in a vector format.

In this format a set of lines are used. The lines will have a starting point and end point and some form of connectivity. The starting and end points of the lines define vectors that represent the form of the chair and pointers between the lines indicate to the computer how the lines link together to form the chair. The Vector data structure for the chair would be;

Chair attribute → set of vectors → connectivity

Let us see what the differences are between these two representations by identifying some important aspects in the implicit, vector representation.

(i) Fewer numbers are needed. Less storage required. (Compare the 11 pairs of coordinates used with the 60 grid cells in the rastor system)
(ii) Aesthetically more pleasing
(iii) The connectivity information allows directed spatial searches to be made over the chair.
(iv) One disadvantage is that when data is updated the coordinates and the connectivity has to be rebuilt.

What are the points to be taken note of in the rastor representation?

(i) Changing the shape or size of the chair is much easier.
(ii) Updating data is easier where you have to only fill or delete grid cells.

From the example of the chair we have just read about, we can see that there are at least two fundamental ways of representing topological data.

Rastor representation-
- Set of cells located by coordinates.
- Each cell is independently addressed with the value of an attribute.

Vector representation-
- Three main geographical entities, Points, Lines and Polygons (Areas).
- Points are similar to cells but they do not cover areas.
- Lines and areas are sets of interconnected coordinates that can be linked to their given attributes.

Most modern interactive computer-aided design and mapping systems work with vector-structured data bases but use colour rastor displays and vector plotters.

Vector data structures represent a more or less direct translation of the map. But the development and utilisation of the rastor data structure has been largely driven by graphic input/output hardware technology. Remote sensing devices of many kinds, as well as scanners for high volume map-digitisation, record spatial data in narrow strips across the data surface. The rastor structure is more compatible with modern input/output hardware. The rastor format preserves many intrinsic spatial interrelationships among the data elements (as contrasted to vector formats) and makes data retrieval on the basis of these spatial relationships a more straightforward task.

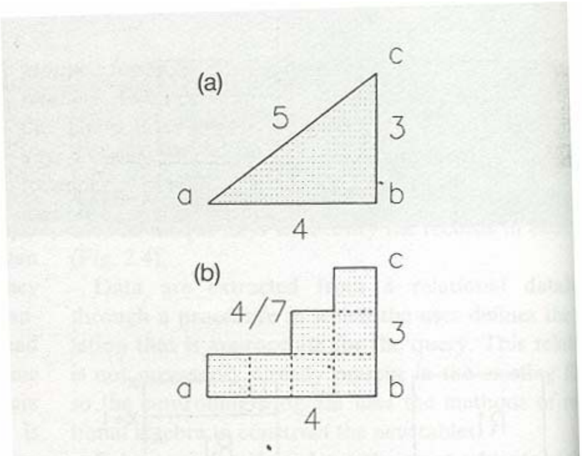


Fig. 2.7 Raster coding can affect estimates of distance and area because of the quantizing effect of the cells.

Fig 4.11:Raster coding can affect estimates of distances and areas because of the quantising effect of the cells

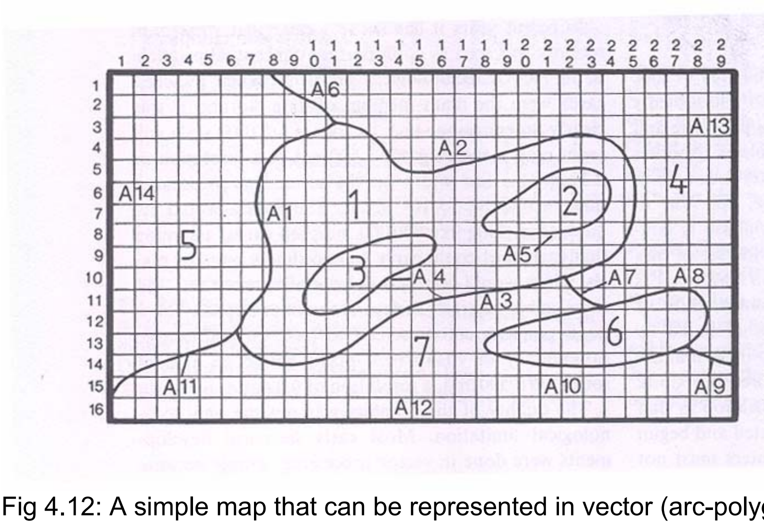


Fig. 2.28 A simple map that can be represented in vector (arc-polygon) or raster form.

Fig 4.12: A simple map that can be represented in vector (arc-polygon) or raster form

4.4 Data types

The attributes of entities can be expressed using different 'data types'.
The different data types are;

- Boolean
- Nominal
- Ordinal
- Integer
- Real

Integer & Real data are known as scalar data.
Real data includes decimals.

Logical operations can be carried out with all data types.
The kind of data analysis is governed by the data types used in the data model.
The limits of accuracy of arithmetical operations is limited by the length of the computer word used to record the numbers.

The following table indicates the data types with some uses.

Data type	Allowed values	Allowed operations	Used as
Boolean	0 or 1	Logical & indicator operations True vs false	
Nominal	Any names	Logical operations, classification & identification	
Ordinal	Numbers from 0 to ∞	Logical & ranking operations, comparisons of magnitude	
Integer	Whole numbers from -∞ to +∞	Logical operations, integer arithmetic	Topological linkages Geographical co-ordinates Arithmetical operations
Real	Real numbers(with decimals)from -∞ to +∞	All logical & numerical operations	Geographical co-ordinates Differentiable continuous surfaces Arithmetical operations
Topological	Whole numbers	Indicate links between entities	

4.5 Data Manipulation and Display in the GIS:

Any given GIS must be capable of performing a series of manipulations upon the spatial data held in its files. A typical set of manipulation capabilities is listed below.

Typical map data handling capabilities. (After Tomlinson & Boyle)

Data Manipulation:

Reclassify – attributes
Generalisation

- Dissolving & merging
- Line smoothening
- Complex generalisation

Interpolation

- Centroid allocation
- Contouring
- Scale change
- Distortion elimination-linear(rubber sheeting)
- Projection change

Generation

- Points
- Lines
- Polygons
  - Simple five-sided polygons
  - Irregular polygons with islands
  - Circles
  - Grid cell nets
  - Latitude & longitude lattices
  - Other

Data extraction

- Search & identification
  - Attributes
  - Shapes
- Measurement
  - Number of items
  - Distances(straight line between points, along convoluted lines)
  - Size of areas
  - Angle direction
  - Volume(cubic measure)

Comparison

- Intersection-overlay
  - Point-in-polygon
  - Polygon-on-polygon(grid cell on polygon, circle on polygon)
  - Other
- Juxtaposition(proximity)
  - Shortest route
  - Nearest neighbor
  - Line of sight
- Contiguity
- Connectivity
- Complex space-attribute-time correlation

Interpretation

- Determination of optimum location
- Determination of suitability
- Determination of desirability

A GIS basically consists of a data structure and a selection of algorithms that are matched to the data structures. Most operational GIS use vector data structures. So vector based algorithms have to be used.

But lack of knowledge on vector based algorithms becomes a major stumbling block in the development of GIS.

Rastor algorithms are less used in a GIS. But you will remember that rastor algorithms are used in image-processing.

It is not clear at this stage whether it is better to base a GIS design on vector based GIS or rastor based GIS. Research is needed in this area of application.

4.6 Data Modeling and Spatial analysis:

There are direct links between the data model, the data type used to represent a geographical phenomenon and the kinds of analysis that can be carried out.

The following situations can arise between these fundamental axioms.

Situation under consideration	Model to be used
Location & form of the entity is unchanging. Attributes can change. Location and form needs to be known accurately	Vector model
Attributes fixed Entity changes form or shape Position fixed (This can be compared to the drying up of a lake)	Rastor model of a continuous field (vector model will require redefinition of the boundary when area of lake changes)
Attributes vary Entity can change position but not form Or separate parts of entity are linked together	Object oriented model
No clear entities	Discretised continuous field

4.6.1 Examples of the use of data models:

Data Models and Axioms

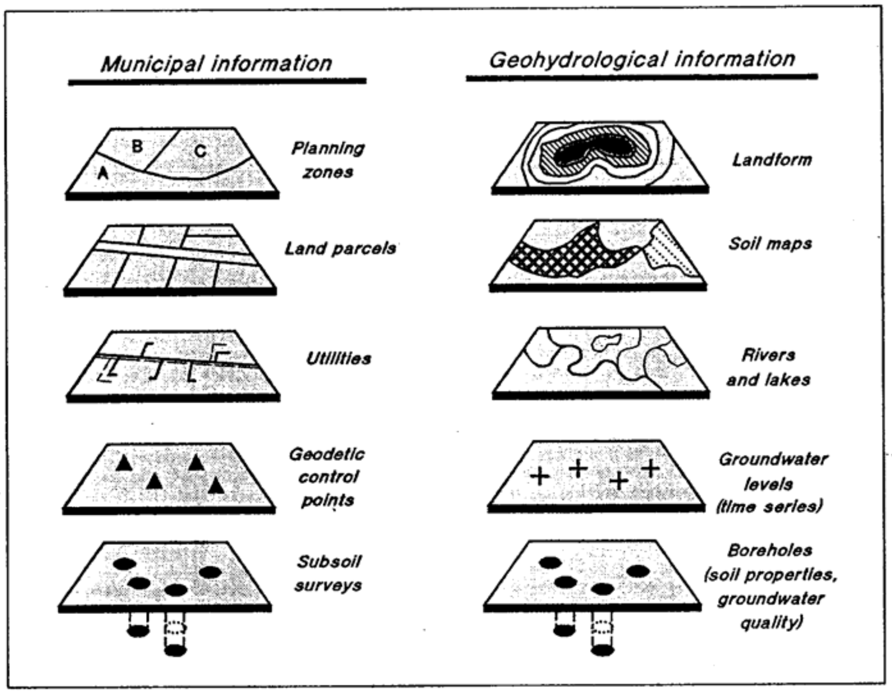


Figure 2.2. Examples of the different kinds of geographical data collected for different purposes by persons from different disciplines

Figure 4.13: Examples of different kinds of geographic data collected for various purposes by persons from different disciplines.

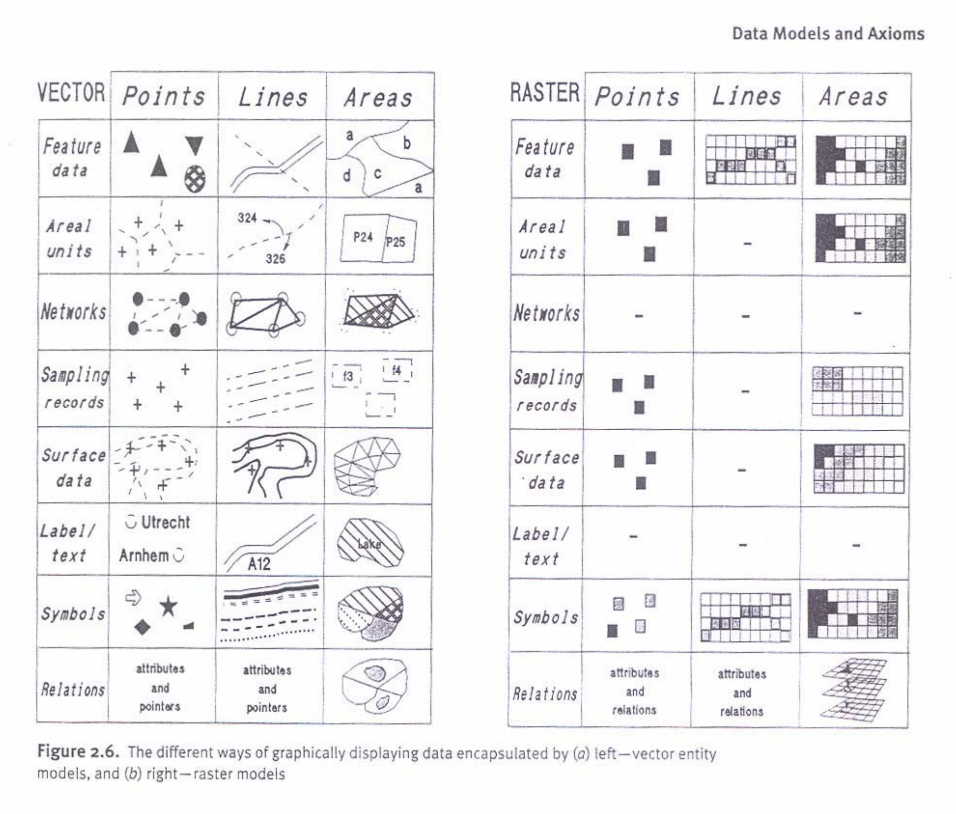


Figure 4.14: The different ways of graphically displaying data as (a) vector entity models (b) rastor models.

How do we decide on the choice of the data model we should use for a specific application?

Let us consider some examples.
Refer to figure on encoding of exact objects and continuous fields in different data models, for deciding on the model.

(i) Cadastre (land registry)
The important issue is to maintain records of division and ownership of land.

What are the facts to be on record?
The location, area & extent of the land and its attributes, address, and information about transactions and legal matters.

From the figure we can see that the vector implementation of entity data model will work well for this case. Topology and attributes will be allocated for each polygon. Due to the polygon representation the boundaries may be shared by adjacent parcels of land. As shown earlier this saves duplication in the database and links the boundaries into a topologically sound polygon net which can handle both adjacency and inclusions.

(ii) Utility networks
Utility networks comprise of the following types supplies and disposal using pipe and wiring systems; Water, gas, electricity, telephone, cable television, waste water & solid waste disposal

What has to be recorded in these systems?

(a)The attributes of a given network such as what it carries, what kind of wire or pipe, age of the system, name of contractor who installed it etc.
(b) The location of the network so that it can be located easily and persons digging in the street will not damage it
(c) Information on how different parts of the net are connected together

The ideal data model for this case will be a model of topologically connected lines (entities) that are described by attributes.

(iii) Land cover databases (land use models)

Land cover determination and creating a data model for same requires several steps;
First, you have to define exactly what is meant by the class of cover.
Secondly, you need to decide how you are going to recognise them.
Thirdly, you must choose a survey method such as point sample survey or remote sensing scanner in a satellite, to collect the data to be interpreted.

We can consider two cases of data model based on two assumptions.

1. We can assume that the classes of land cover are distinctly separable on the ground. If so, the simple polygon representation can be used with adjacent boundaries. This will give a choropleth map. So what you have to do is, define the classes, identify and map the boundaries and attach attributes to each class similar to the polygon net model for cadastre mapping.
2. Consider the case where the lines of demarcation on the ground are not clear and you have to interpolate to unobserved areas based on 'ground truth' in order to arrive at the classification. We can either decide to interpolate to crisply delineated areas of land using the choropleth model based on the polygon. Or we can interpolate to a discretised surface such as a regular grid. In this case the land cover map consists of sets of pixels with attributes indicating the land cover class to which they belong. If we use remotely sensed data to identify land cover we automatically work with a gridded discretisation of continuous space because that is how the satellite scanner works.

This example shows how two different data models can be used for land cover mapping. Depending on the how the data has been collected and how the mapped data will be used, we can decide on a rastor or vector format.

(iv) Soil maps

A soil map is required in order to identify different boundaries of soil classes in geographical space. Important soil differences may be indicated by abrupt, clearly observable physiographic features such as changes in lithology, drainage or breaks of slope which we call 'primary boundaries'. We can also interpret differences in soil classification in the data space, and these are termed 'secondary boundaries'.
In reality soil boundaries can be sharp, gradual or diffuse.

So how can we create a model for soil profiles?

Most published soil maps use the entity model based on the vector polygon as the geographical primitive. Polygons are defined in terms of their soil class, which is assumed to be homogeneous over the unit. Boundaries are thin lines implying abrupt changes of soil cover over very short distances. Polygons can be included inside polygons. Although this model is practical and is somewhat similar to the cadastre model, it ignores the fact that there can be spatial variation due to soil forming processes.
An alternative to the discrete polygon data model for soil is to assume that soil properties vary gradually over the landscape. The soil is sampled at a series of locations and attributes are determined for those samples. The simplest data model then becomes the geographical point with

the values of the associated attributes. From this simple data model new data models of continuous spatial variation may be created by interpolation.
(compare this with the 'indirect' method of contouring, when locating contours)
The data models for describing soil as a continuous variable are, in principle, very similar to those used for the land elevation surfaces. Sets of discrete contour lines can be used to link zones of equal attribute values. Alternatively attribute values can be interpolated to locations on a regular grid. This is the rastor model of space.

(v) Hydrology

What are the different types of hydrological applications? What will be required of a model?

It is the modelling of the transport of water and materials over space and time.
Attributes will change and also the form of critical patterns such as water bodies.
Changes can occur in the following;
Water levels in rivers, reservoirs and lakes, which also leads to change in the location of the boundary between water and land.
Geometry and location of water bodies
A flood can open up new channels changing both topology and location.

A vector data model is not suitable in the case of hydrology. Can you think why?
In a vector model, changes in geometry will mean changing the coordinate and topological data in polygon networks. This involves considerable computations.

What is more suitable is a data model based on 'object orientation' in which primitive entities are linked together in functional groups. The internal structure of the data model permits action on one component of the group to be passed automatically to other parts. So actually such a data model will not only contain geographical location, geometry, topology and attributes but also information on how all these react to change.
Hydrological data models use the concept of continuous variation for transport of material but these have not been included in GISs.

Another way of dealing with transport of material over a surface, is by using the rastor data model of continuous variation to which the surface topology has been added or derived by computation.

4.6.2 Nine factors to be considered in choosing data models for spatial analysis using a GIS:

1. Is the real world situation/phenomena under study simple or complex?
2. Are the kinds of entities used to describe the situation/phenomena detailed or generalised?
3. Is the data type used to record attributes Boolean, nominal, ordinal, integer, real or topological?
4. Do the entities in the database represent objects that can be described exactly, or are these objects complex and possibly somewhat vague? Are their properties exact, deterministic or stochastic?
5. Do the database entities represent discrete physical things or continuous fields?
6. Are the attributes of database entities obtained by complete enumeration or by sampling?
7. Will the database be used for descriptive, administrative, or analytical purposes?
8. Will the users require logical, empirical, or process-based models to derive new information from the database and hence make inferences about the real world?
9. Is the process under consideration static or dynamic?