Saleem ya Sir

CE700OE: REMOTE SENSING & GIS (Open Elective - II)

B.Tech. Civil Engg. IV Year I Sem.

L	T/P/D	C
3	0/0/0	3

Course Objectives: The objectives of the course are to

- Know the concepts of Remote Sensing, its interpreting Techniques and concepts of Digital images
- know the concept of Geographical Information System (GIS), coordinate system GIS Data and its types
- Understand the students managing the spatial Data Using GIS.
- Understand Implementation of GIS interface for practical usage.

Course Outcomes: After the completion of the course student should be able to:

- **Describe** different concepts and terms used in Remote Sensing and its data
- Understand the Data conversion and Process in different coordinate systems of GIS interface
- **Evaluate** the accuracy of Data and implementing a GIS
- **Understand the applicability** of RS and GIS for various applications

UNIT – I

Concepts of Remote Sensing Basics of remote sensing- elements involved in remote sensing, electromagnetic spectrum, remote sensing terminology & units, energy resources, energy interactions with earth surface features & atmosphere, atmospheric effects, satellite orbits, Sensor Resolution, types of sensors. Remote Sensing Platforms and Sensors, IRS satellites.

Remote Sensing Data Interpretation Visual interpretation techniques, basic elements, converging evidence, interpretation for terrain evaluation, spectral properties of soil, water and vegetation. Concepts of Digital image processing, image enhancements, qualitative & quantitative analysis and pattern recognition, classification techniques and accuracy estimation.

UNIT- II:

Introduction to GIS: Introduction, History of GIS, GIS Components, GIS Applications in Real life, The Nature of geographic data, Maps, Types of maps, Map scale, Types of scale, Map and Globe, Co-ordinate systems, Map projections, Map transformation, Geo-referencing,

UNIT- III:

Spatial Database Management System: Introduction: Spatial DBMS, Data storage, Database structure models, database management system, entity-relationship model, normalization

Data models and data structures: Introduction, GIS Data model, vector data structure, raster data structure, attribute data, geo-database and metadata,

UNIT- IV:

Spatial Data input and Editing: Data input methods – keyboard entry, digitization, scanning, conversion of existing data, remotely sensed data, errors in data input, Data accuracy, Micro and Macro components of accuracy, sources of error in GIS.

Spatial Analysis: Introduction, topology, spatial analysis, vector data analysis, Network analysis, raster data analysis, Spatial data interpolation techniques

UNIT- V: Implementing a GIS and Applications

Implementing a GIS: Awareness, developing system requirements, evaluation of alternative systems, decision making using GIS

Applications of GIS

GIS based road network planning, Mineral mapping using GIS, Shortest path detection using GIS, Hazard Zonation using remote sensing and GIS, GIS for solving multi criteria problems, GIS for business applications.

TEXT BOOKS

1. Remote Sensing and GIS by Basudeb Bhatta, Oxford University Press, 2nd Edition, 2011.
2. Introduction to Geographic Information systems by Kang-tsung Chang, McGraw Hill Education (Indian Edition), 7th Edition, 2015.
3. Fundamentals of Geographic Information systems by Michael N. Demers, 4th Edition, Wiley Publishers, 2012.

REFERENCE BOOKS

1. Remote Sensing and Image Interpretation by Thomas M. Lillesand and Ralph W. Kiefer, Wiley Publishers, 7th Edition, 2015.\
2. Geographic Information systems – An Introduction by Tor Bernhardsen, Wiley India Publication, 3rd Edition, 2010.
3. Advanced Surveying: Total Station, GIS and Remote Sensing by Satheesh Gopi, R. Sathi Kumar, N. Madhu, Pearson Education, 1st Edition, 2007.
4. Textbook of Remote Sensing and Geographical Information systems by M. Anji Reddy,

Verified By

9/12/2023

Chapter 1: Remote Sensing

Chapter 1:

This chapter briefly discusses the basics of various techniques, viz. the remote sensing, the digital image processing, visual Interpretation and delineation using Geographical Information System (GIS) domain

: Remote Sensing:

"**Remote sensing** is defined as the science and art of obtaining information about an object, area, or phenomenon through the analyses of data acquired by the sensor that is not in direct contact with the target of investigation" (Schultz and England, 2000; Ritchie and Rango, 1996). This can be done by the use of either recording or real-time sensing device(s) mounted on aircraft, spacecraft, satellite, buoy, or ship. During last forty years, space travel has been giving humanity new opportunities, not only to peer into the depths of the cosmos, but also to look at the length and breadth of our own world. Remote sensing enables us to acquire information about a phenomenon, object or surface while at a distance from it. More than just a source of pleasing pictures, these sophisticated techniques now allow scientists to understand the Earth in ways we never before dreamed. From 1960 onwards, since the beginning of spaceflight, an ever-growing body of information is being gathered using space-based remote sensing. Agriculture, meteorology, oceanography, ecology, cartography, botany, geomorphology and geology are just a few of the disciplines, which have been transformed by this technology. Some of the most important environmental issues of our time are only becoming understood because of vast networks of remote sensing devices and data analysis systems. In practice, remote sensing is the standoff collection through the use of a variety of devices for gathering information on a given object or area. The remote sensing is a multi-disciplinary science, which includes a combination of various disciplines such as

Optics, spectroscopy, photography, computers, electronics and telecommunication, etc. All these technologies are integrated to act as one complete system in itself, known as Remote Sensing. In the decades to come, remote sensing will be a key tool for making critical decisions affecting the Earth and its resources.

In a remote sensing system, number of components work as links and each one is important for successful operation. They are listed below (figure 3.1).

- Emission of electromagnetic radiation, or **EMR**
- Transmission of energy from the source to the surface of the earth, interaction with the atmosphere as well as absorption and scattering
- Interaction of **EMR** with the earth's surface: absorption, reflection and emission
- Transmission of energy from the surface to the remote sensor
- Sensor/detector data output
- Data transmission, processing and analysis

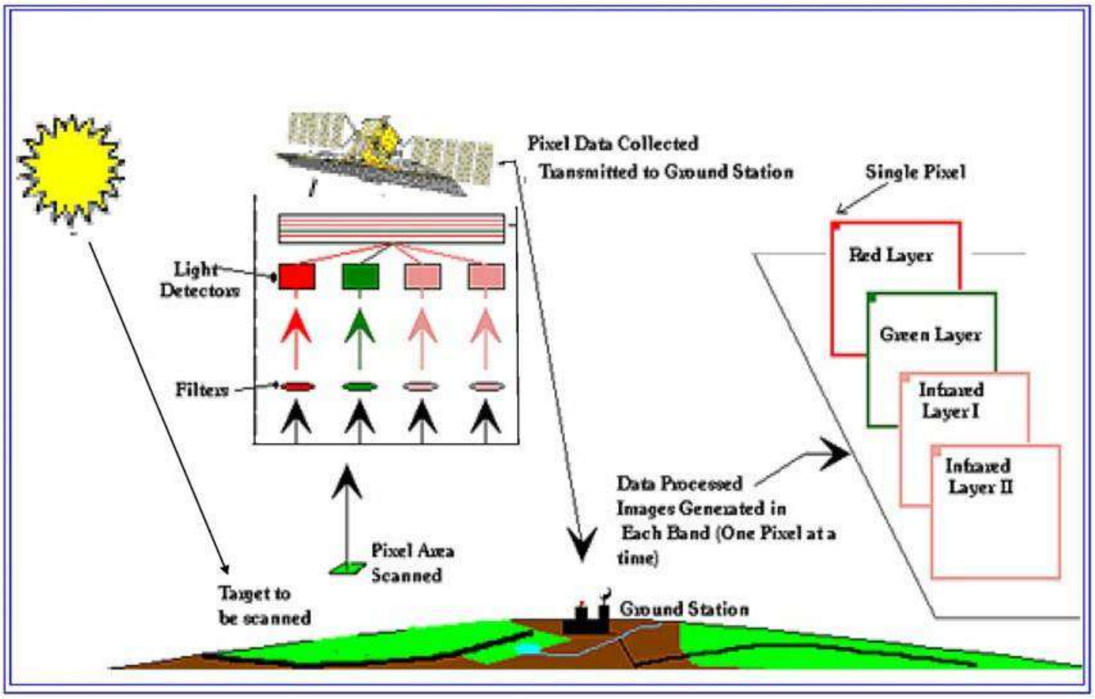


Figure 3.1: Different stages in Remote sensing and data acquisition

(Source: http://www.ansp.org/museum/kye/tech_environment/2001_remote_sensing.php)

: Basic components of Remote Sensing:

The sun is the major source of energy, radiation and illumination. At any given moment, our sun is bombarding the earth with a variety of wavelengths of EMR, including visible light, infrared, radio and microwaves. Detection and discrimination of surface features means detecting and recording of radiant energy reflected or emitted by surface (Joseph, 2004; Lillesand and Kiefer, 1987; 2000). Different features return different amount and kind of energy in different bands of the electromagnetic spectrum, incident upon it. This unique property depends on the property of material (structural, chemical and physical), surface roughness, angle of incidence, intensity and wavelength of radiant energy (Elachi, 1987). Everything in nature has its own unique distribution of reflected, emitted and absorbed radiation. These spectral characteristics, if ingeniously exploited, can be used to distinguish one thing from another or to obtain information about shape, size and other physical and chemical properties. Because the emission and reflection of many different types of EMR can be detected by instruments, they can also be used for remote sensing. Thus, to understand remote sensing, it is important to first understand the basics of EMR.

: Electromagnetic Radiation:

Electromagnetic radiation is one of the fundamental forms of energy in the universe. EMR is a dynamic form of energy that propagates as wave motion at a velocity of light,

I.e. $C=3X10^{10}$ m/s. Electromagnetic energy radiates in accordance with the basic wave theory. This theory describes the EM energy as traveling in a harmonic sinusoidal fashion at the velocity of light.

Here, **$c = v\lambda$ **3.1

Where, v is frequency, λ is wavelength and c is velocity

Although many characteristics of EM energy are easily described by wave theory,

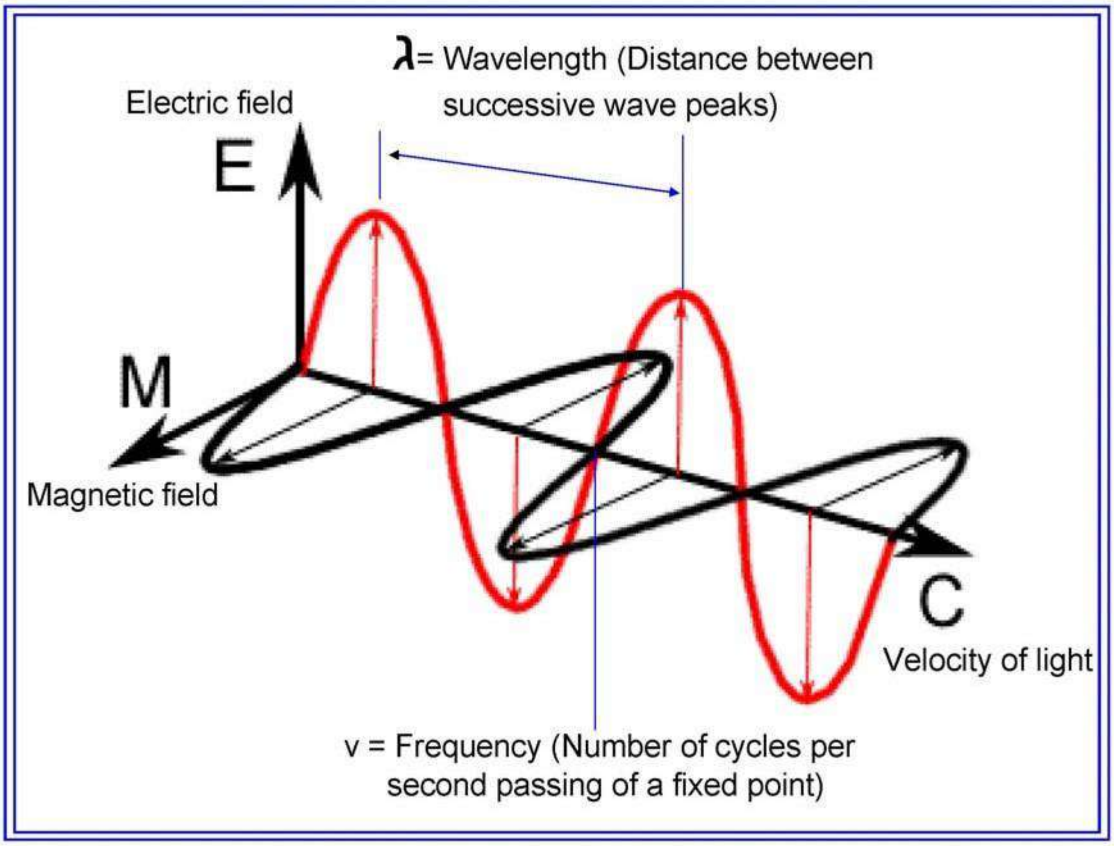


Figure 3.2: An electromagnetic wave. Includes electric wave (E) and magnetic wave (M) at right angles, both perpendiculars to the direction of propogation. (Source: http://www.ccrs.nrcan.gc.ca/resource/tutor/fundam/index_e.php)

another theory known as particle theory offers insight into how electromagnetic energy interacts with matter (figure 3.2). It suggests that EMR is composed of many discrete units called photons/quanta. The energy of quantum is:

$Q = h\, c/\lambda$... 3.2

Where Q is the energy of quantum, h is plank's constant

Hence, from equation 3.1

Q = h cv..3.3

: Electromagnetic Spectrum

Electromagnetic energy is the energy source required to transmit information from the target to the sensor. A crucial medium is described as an electromagnetic spectrum. Many of the basic forms of energy in the universe are related as part of the electromagnetic spectrum. On this spectrum, many forms exist that describe energy in a specific region of the electromagnetic spectrum. These are visible light, radio waves, microwaves, infrared, UV rays, x-rays and gamma rays (figure 3.3).

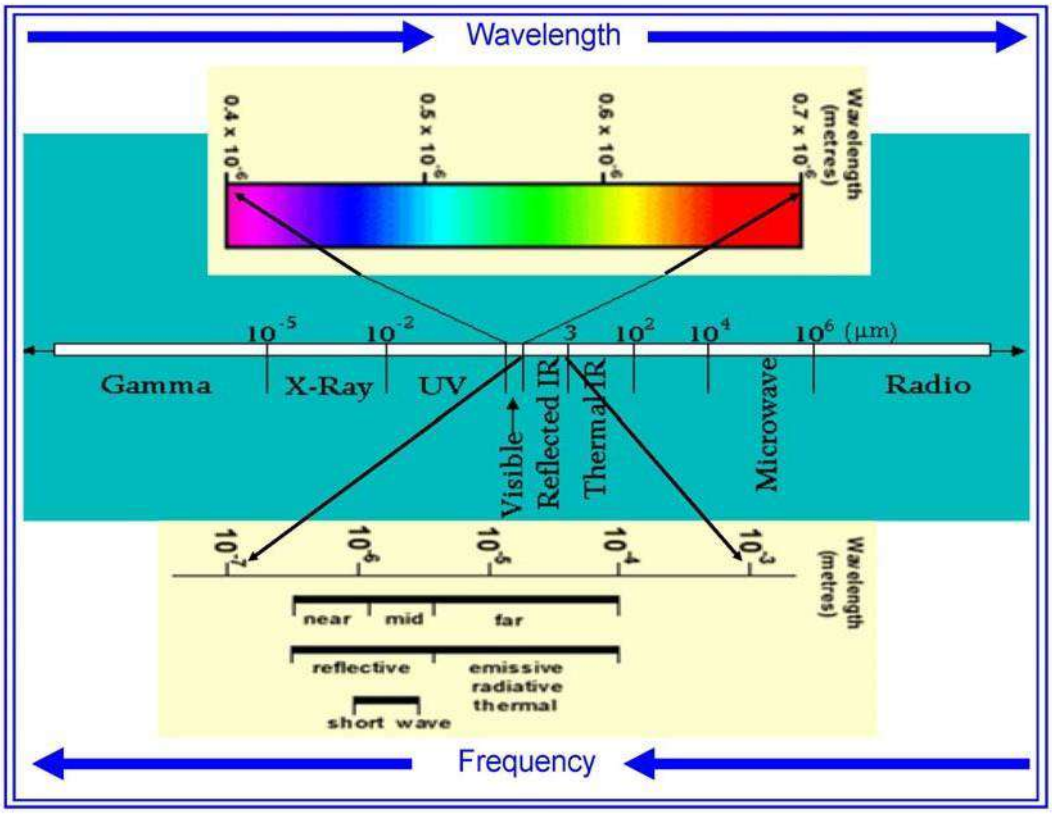


Fig 3.3: Electromagnetic spectrum

(Source: http://hcgl.eng.ohio-state.edu/~ceg603/handouts/IntroRemoteSensing.pdf)

This spectrum is an overview of the continuum of electromagnetic energy from extremely short wavelengths (cosmic gamma rays) to extremely long wavelengths (radio and television waves). Note that as the wavelengths of energy decrease, the frequency increases. These divisions are not absolute and definite; overlapping may occur (table 3.1).

Table 3.1: Major regions of the electromagnetic spectrum

Region Name	**Wavelength**	**Comments**
Gamma Ray	< 0.03 nanometers	Entirely absorbed by the Earth's atmosphere and not available for remote sensing.
X-ray	0.03 to 30 nanometers	Entirely absorbed by the Earth's atmosphere and not available for remote sensing.
Ultraviolet	0.03 to 0.4 micrometers	Wavelengths from 0.03 to 0.3 micrometers absorbed by ozone in the Earth's atmosphere.
Photographic Ultraviolet	0.3 to 0.4 micrometers	Available for remote sensing the Earth. Can be imaged with photographic film.
Visible	0.4 to 0.7 micrometers	Available for remote sensing the Earth. Can be imaged with photographic film.
Reflected Infrared	0.7 to 3.0 micrometers	Available for remote sensing the Earth. Near Infrared 0.7 to 0.9 micrometers. Can be imaged with photographic film.
Thermal Infrared	3.0 to 14 micrometers	Available for remote sensing the Earth. This wavelength cannot be captured with photographic film. Instead, sensors are used to Image this wavelength band.
Microwave or Radar	0.1 to 100 centimeters	Longer wavelengths of this band can pass through clouds, fog, and rain. Images using this band can be made with sensors that actively emit microwaves.
Radio	> 100 centimeters	Not normally used for remote sensing the Earth.

(Source: http://www.physicalgeography.net/fundamentals/2e.html)

: Electromagnetic radiation quantities

- **Radiant energy (Q)** is the energy carried by EMR. Radiant energy causes the detector element of the sensor to respond to EMR in some appropriate manner. Unit of Radiant Energy Q is Joule.

- **Radiant Flux (Φ) (Phi)** is the time rate of the flow of radiant energy. Unit of Radiant flux is Joule/Second or Watt (W).

- **Irradiance (E)** is the Radiant flux intercepted by a plane surface per unit area of the surface. It arrives at the surface from all directions within a hemisphere over the surface. Unit of Irradiance E is W/m^2 or Wm^{-2} (Watt per square meter).

- **Radiance (L)** is defined as the radiant flux per unit solid angle leaving an extended source in a given direction per unit-projected area of the source in that direction. The concept of radiance is intended to correspond to the concept of brightness. The projected area in a direction, which makes an` angle θ (Theta) with the normal to the surface of area A is A cosθ. Unit for Radiance is $Wm^{-2}sr^{-1}$.

- **Spectral Reflectance (ρ(λ))** is the ratio of reflected energy to incident energy as a function of wavelength.

- **Spectral Signature** are the values of the spectral reflectance of objects averaged over different, well defined wavelength intervals comprise the spectral signature of the objects or features by which they can be distinguished.

: Interaction of EMR with atmosphere and earth surface

The sun is the source of radiation and electromagnetic radiation (EMR) from the sun that is reflected by the earth and detected by the satellite or aircraft-borne sensor must pass through the atmosphere twice, once on its journey from the sun to the earth and once after being reflected by the surface of the earth back to the sensor. Interactions of the direct solar radiation and reflected radiation from the target with the atmospheric constituents interfere with the process of remote sensing and are called as "Atmospheric Effects".

The atmospheric constituents scatter and absorb the radiation, modify the radiation reflected from the target by attenuating it. Both scattering and absorption vary in their effect from one part of the spectrum to the other. The solar energy is subjected to modifications by several physical processes as it passes the atmosphere, viz. scattering, absorption and refraction.

- **Atmospheric Scattering:** Scattering is the redirection of EMR by particles suspended in the atmosphere or by larger molecules of atmospheric gases or water vapor. The amount of scattering depends upon the size and abundance of particles, the wavelength of radiation and depth of the atmosphere through which the energy is traveling which varies both in time and over season, thus scattering will be uneven spatially and will vary from time to time. Scattering reduces image contrast and changes spectral signature of ground objects as seen by the sensor.

- **Atmospheric Absorption:** The gas molecules present in the atmosphere strongly absorb the EMR passing through the atmosphere in certain spectral bands. Mainly three gases are responsible for most of the absorption of solar radiation, viz. ozone, carbon dioxide and water vapor. Absorption relatively reduces the amount of light that reaches our eye making the scene look relatively duller.

- **Atmospheric Windows:** The atmosphere selectively transmits energy of certain wavelengths. The spectral bands for which the atmosphere is relatively

Transparent are known as “Atmospheric Windows”. In the visible part transmission is mainly effected by ozone absorption and by molecular scattering (figure 3.4).

- **Refraction:** The phenomenon of refraction that is bending of light at the contact between two media also occurs in the atmosphere as the light passes through the atmospheric layers of varied clarity, humidity and temperature.

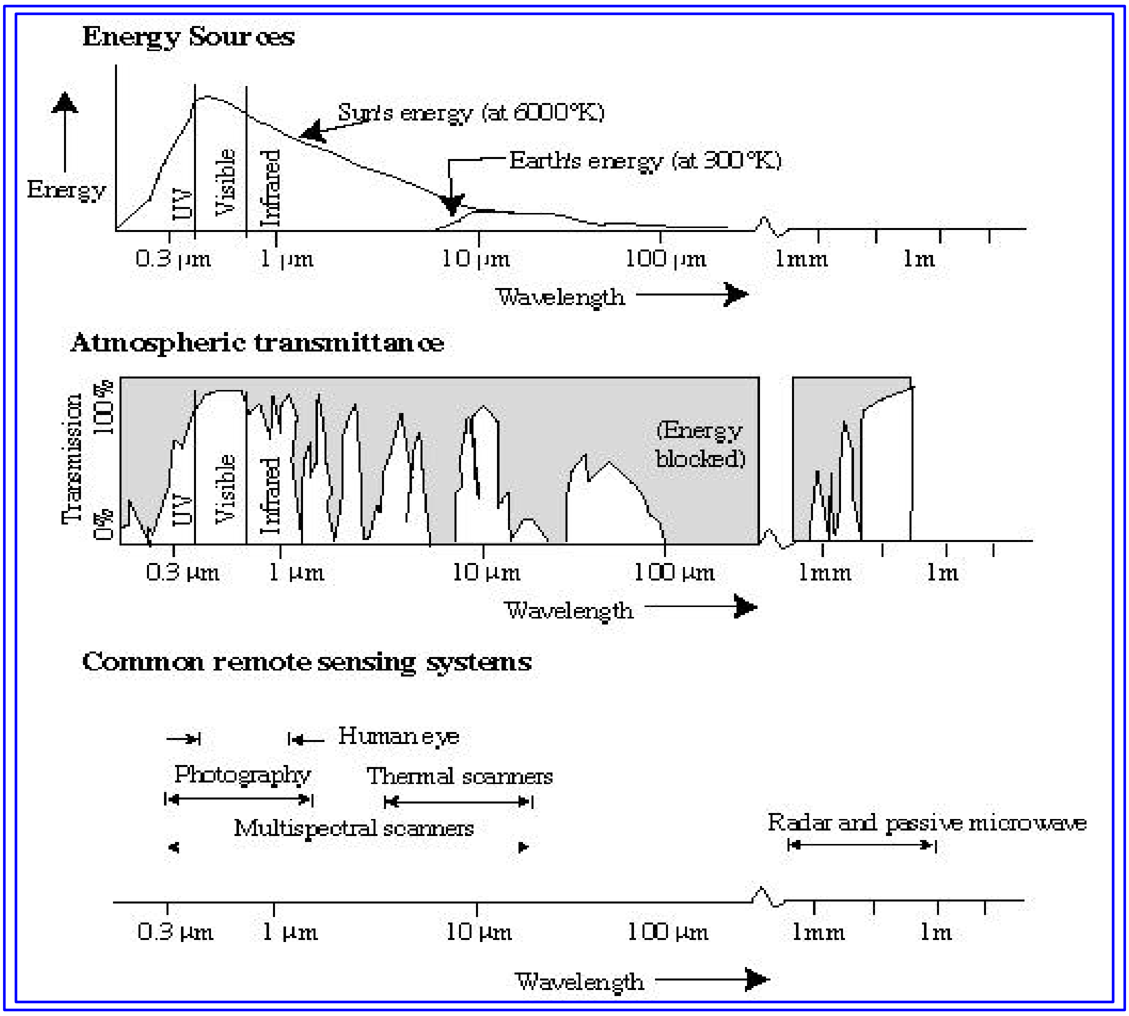


Fig 3.4: Spectral characteristics of energy sources, atmospheric effects and sensing systems wavelength scale are logarithmic.

(Source: www.ucalgary.ca/GEOG/virtual/remoteintro.html)

Radiation from the sun, when incident upon the earth's surface, is reflected by the surface, transmitted into the surface or absorbed and emitted by the surface. The EMR, on the interaction, experiences a number of changes in magnitude, direction, wavelength, polarization and phase. These changes are detected by the remote sensor and enable the interpreter to obtain useful information about the object of interest. The remotely sensed data contain both spatial information (size, shape and orientation) and spectral information (tone, colour and spectral signature). The spectral band from 0.3 µm to 3 µm is known as the reflective region. In this band, the radiation sensed by the sensor is that due to the sun, reflected by the earth's surface.

- **Reflection:** Of all the interactions in the reflective region, surface reflections are the most useful and revealing in remote sensing applications. Reflection occurs when a ray of light is redirected as it strikes a non-transparent surface. The reflection intensity depends on the surface refractive index, absorption coefficient and the angles of incidence and reflection.

- **Transmission:** Transmission of radiation occurs when radiation passes through a substance without significant attenuation. For a given thickness, or depth of a substance, the ability of a medium to transmit energy is measured as transmittance (τ). **τ = transmitted radiation/incident radiation**

: Data detection and output:

The most important component of a remote sensing is the sensor/detector, which, records the variation of radiant energy reflected or emitted by objects or surface material. Different types of sensors are sensitive to different parts of the electromagnetic spectrum. The function of recording system is to convert the energy detected by sensor into a form, which can be perceived. This is done by dividing the incoming energy by beam splitters and filters into different wavelength bands and then converting energy in each

Wavelength band into electrical signal. The electrical signal is processed to give radiometric data for each band, which is recorded in digital format.

: Remote Sensing Systems:

The common remote sensing systems are of two types, Imaging (image forming) and Non-Imaging (nonimage forming). Image-forming systems are again of two types, framing type and scanning type. In scanning type, the information is acquired sequentially from the surface in bits of picture elements or pixels, point-by-point and line-by-line, which may be arranged after acquisition into a frame format.

Remote sensing can be either passive or active. ACTIVE systems have their own source of energy such as RADAR, whereas the PASSIVE systems depend upon external source of illumination such as sun for remote sensing.

: Platforms and Sensors:

The information flows from an object to a sensor in the form of radiation transmitted through the atmosphere. For a sensor to collect and record energy reflected or emitted from a target or surface, it must reside on a stable platform away from the target or surface being observed. Based on its altitude above earth surface, platforms can be classified as ground-borne, air-borne and space-borne.

Space-borne platforms are in space, moving in their orbits around the earth. It is through these space-borne platforms, we get enormous amount of remote sensing data. Depending on their altitude and orbit, these platforms may be divided in two categories:

- **Geostationary satellites:** An equatorial west to east satellite orbiting the earth at an altitude at which it makes one revolution in 24 hours, synchronous with the earth's rotation, hence it gives continuous coverage over the same area day and

Night. These are mainly used for communication and meteorological applications, for e.g. the INSAT satellites.

- **Polar orbiting or Sun-synchronous satellites:** A satellite with inclined north-south orbit track westward at a rate such that it covers each area of the world at a constant time of the day called the local sun time as the satellite moves from north to south. This ensures the similar illumination conditions while acquiring image over a particular area over a series of days (figure 3.5). On the descending pass from north to south, the satellite travels on the sun lit side of the earth, while on ascending pass from south to north it travels on the shadowed side of the earth. Through these satellites, the entire globe is covered on regular basis and gives repetitive coverage on periodic basis. All remote sensing resources satellites may be grouped in this category for e.g. IRS series.

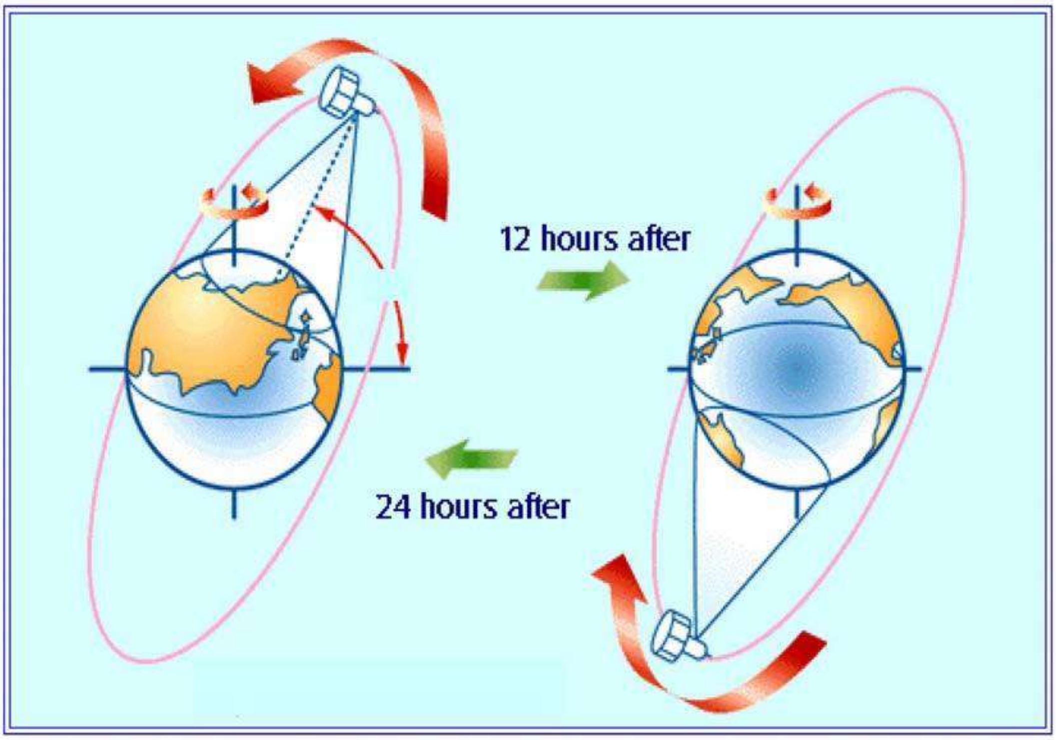


Fig 3.5: Sun synchronous satellites orbiting earth

(Source: http://www.newmediastudio.org/DataDiscovery/Hurr_ED_Center/Satellites_and_Sensors/Polar Orbits/Polar_Orbits.html)

Sensor is the device that gathers energy (EMR or others), converts it into a signal and presents it in a form suitable for obtaining information about the target under investigation. Sensors used for remote sensing can be broadly classified as those operating in Optical Infrared (OIR) region and those operating in microwave region. OIR and microwave sensors can further be subdivided into passive and active depending on the source of energy.

- **Active sensors** use their own source of energy to illuminate earth surface and a part of it is reflected and received to gather information.

- **Passive sensors** do not have their own source of energy but instead receive solar electromagnetic energy reflected from surface or energy emitted from surface itself hence it cannot be used at nighttime, except thermal sensors. Energy that is naturally emitted (e.g. Thermal energy) can be detected day or night, as long as the amount of energy is large enough to be recorded (Sabins, 1997).

: Types of Remote Sensing:

Remote sensing can be broadly classified into three types with respect to the wavelength region and type of sensor involved for data acquisition; viz. Optical (Visible and Reflective Infrared), Thermal Infrared and Microwave. In present study, optical remote sensing technique has been used to achieve the defined objectives.

: Optical Remote Sensing:

Optical remote sensing involves the use of visible part of the EM spectrum. Energy emitted from sun is used for visible and reflective infrared remote sensing. The International Commission on Illumination has defined the visible spectrum to be from
0.38 to 0.79 μm. The human eye has its peak sensitivity at 0.55 microns, which is

Approximately the peak of the emission curve of the sun (Jensen, 1996; Sabins, 1997; Gupta, 1999; 2003). Remote sensing data obtained in the visible and reflective infrared regions mainly depends on the reflectance of objects on the ground surface. Therefore, information about objects can be obtained from the spectral reflectance (Elachi, 1987).

: Thermal Infrared Remote Sensing:

The source of radiant energy used in thermal infrared remote sensing is the object itself, because any object with a normal temperature will emit EM radiation with a peak at about 10 μm (Sabins, 1997; Gupta, 2003).

: Microwave Remote Sensing:

In microwave region, there are two types of microwave remote sensing, passive and active. In passive microwave remote sensing, the microwave radiation emitted from an object is detected, while in active microwave remote sensing, source of EM radiations and detector are placed on the sensor (Elachi, 1988).

Inherent characteristics and spectral signature of objects

In any photographic image forming process, the negative is composed of tiny silver deposits formed by the action light on photosensitive film during exposure. The amount of light received by the various sections of the film depends on the reflection of EMR from various objects. The light, after passing through the optical system, gives rise to different tones and textures.

In visual interpretation, an interpreted is primarily concerned with recognizing changes in tonal values, thereby differentiating an object of a certain reflective characteristic from another. However, he must be aware that the same object under different moisture or illumination conditions and depending on the wavelength of incident energy may reflect a different amount of light. For this reason, a general key, based on tone characteristics of objects, cannot be prepared. In such cases, other characteristics of objects such as their shape, size and pattern, etc., help in their recognition.

Spectral signature is the parameter, which determines the character of the object under observation. This can be defined as a unique pattern of wavelengths radiated/reflected by an object. It can be categorized as:

- **Spectral variation:** Variation in reflectivity and emissivity as a function of wavelength.

- **Spatial variation:** Variation in reflectivity and emissivity with spatial position (i.e. shape, texture and size of the object).

- **Temporal variation:** Variation of emissivity and reflectivity like that in diurnal and seasonal cycle.

- **Polarization variation:** the material in the radiation reflected or emitted by it introduces Variations.

Each of these four features of EMR may be interdependent. A measure of these variations and correlating them with the known features of an object provides signature of the object concerned. The knowledge of the state of polarization of the reflected radiation in addition to spectral signature of various objects in remote sensing adds dimension for

Analysis and interpretation of remote sensing data. These parameters are extremely useful in providing valuable data for discriminating the objects.

: Elements of image interpretation

Recognizing targets is the key to interpretation and information extraction. Observing the differences between targets and their backgrounds involves comparing different targets based on any, or all, of the visual elements of tone, shape, size, pattern, texture, shadow, and association. Identifying targets in remotely sensed images based on these visual elements allows us to further interpret and analyze. The nature of each of these interpretation elements is described below:

: Tone refers to the relative brightness or colour of objects in an image. Generally, tone is the fundamental element for distinguishing between different targets or features. Variations in tone also allow the elements of shape, texture, and pattern of objects to be distinguished.

: Shape refers to the general form, structure, or outline of individual objects. Shape can be a very distinctive clue for interpretation. Straight edge shapes typically represent urban or agricultural targets, while natural features are generally more irregular in shape.

: Size of objects in an image is a function of scale. It is important to assess the size of a target relative to other objects in a scene, as well as the absolute size, to aid in

The interpretation of that target. A quick approximation of target size can direct interpretation to an appropriate result more quickly.

: **Pattern** refers to the spatial arrangement of visibly discernible objects. Typically, an orderly repetition of similar tones and textures will produce a distinctive and ultimately recognizable pattern.

: **Texture** refers to the arrangement and frequency of tonal variation in particular areas of an image. Rough textures would consist of a mottled tone where the grey levels change abruptly in a small area, whereas smooth textures would have very little tonal variation.

: **Shadow** is also helpful in interpretation as it may provide an idea of the profile and relative height of a target or targets, which may make identification easier. However, shadows can also reduce or eliminate interpretation in their area of influence, since targets within shadows are much less (or not at all) discernible from their surroundings.

: **Association** takes into account the relationship between other recognizable objects or features in proximity to the target of interest. The identification of features that one would expect to associate with other features may provide information to facilitate identification.

: Spectral behavior of surface features

As EMR incidents on earth's surface, behavior of land features is mainly due to the component of the target at that locality. Since each of these components exhibit typical spectral signature influenced by so many other parameters of their own, they are to be considered separately to understand the nature of EMR interaction with each components. Spectral signature or water, vegetation and soil are discussed below (figure 3.6):

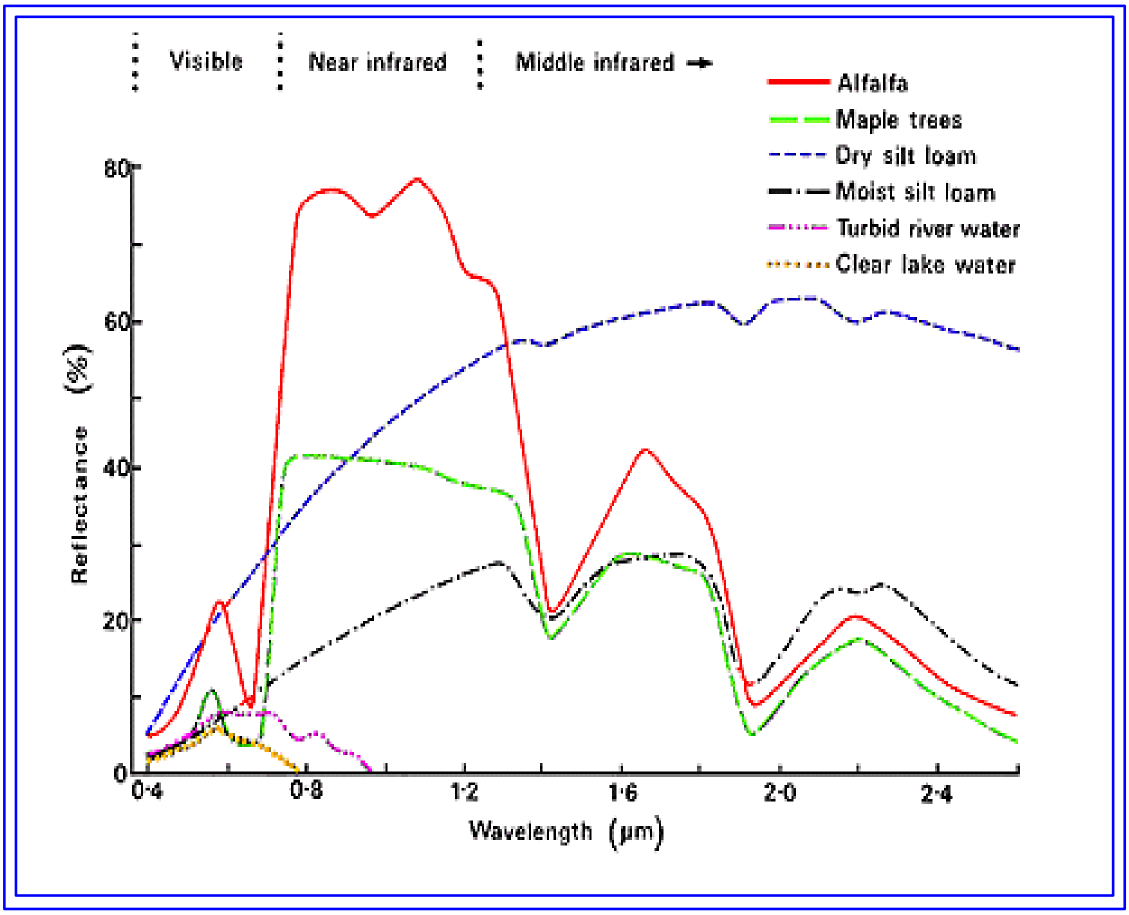


Fig 3.6: Typical spectral reflectance curves for vegetation (two different types), soil (two different types) and water (two different types)

(Source: http://ceos.cnes.fr:8100/cdrom-00/ceosl/irsa/pages/intro2c.htm)

: Spectral reflectance and signature of soil

The majority of the flux incident on a soil surface is reflected or absorbed and little is transmitted. The reflectance properties of the majority of soils are similar, with a positive relationship between reflectance and wavelength, as seen in fig: 3.7. Main factors influence the soil reflectance in remote sensing images are mineral composition, moisture content, organic matter content and soil texture (surface), soil structure and iron oxide content. Size and shape of the soil aggregate also influence the reflectance in the images.

: Effect of mineral composition:

The mineral composition of soils affects the reflectance spectrum. Increasing reflectance of soils occurs from the visible to the shortwave infrared - with absorption bands around 1.4 µm and 1.9 µm related to the amount of moisture in the soil.

: Effect of soil texture, structure and moisture:

The relationship between texture, structure and soil moisture can best be described with reference to two contrasting soil types. A clay soil tends to have strong structure, which leads to rough surface texture, and have high moisture content and as a result have a low diffuse reflectance. In contrast, a sandy soil tends to have a weak structure, which leads to a smooth surface texture, and have low moisture content and as a result have high and often specular reflectance properties. Soil texture (roughness) also affects soil optical properties. In visible wavelengths, the presence of soil moisture considerably reduces the surface reflectance of soil, until the soil is saturated. Reflectance in near and middle infrared wavelengths is negatively related to soil moisture. An increase in soil moisture will result in a rapid decrease in reflectance in water and hydroxyl absorbing wavebands. The effect of water and hydroxyl absorption is more noticeable in clay soil as it has much bound water and very strong hydroxyl absorption properties.

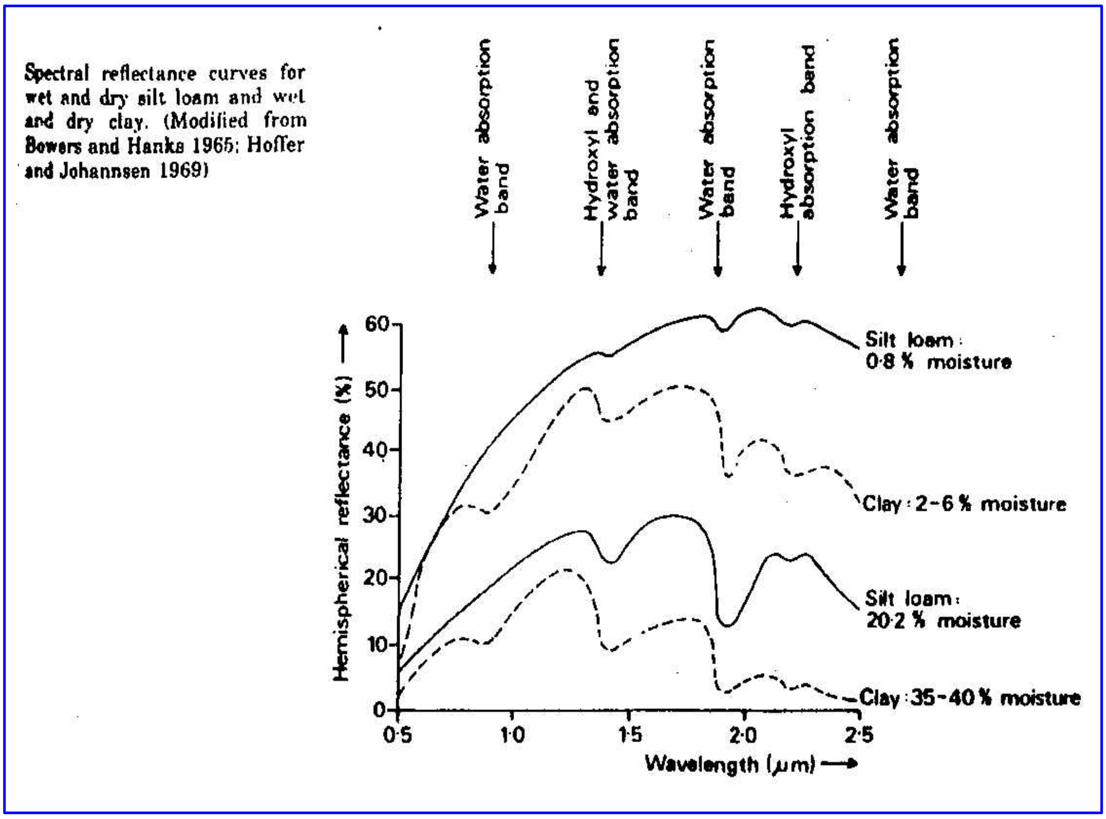


Fig 3.7: effect of soil moisture on soil spectral reflectance

(Source: Tutorials from Indian Institute of Remote Sensing, Dehradun, India)

: Effect of organic matter:

Organic matter may indirectly affect the spectral influence, based on the soil structure and water retention capacity. It is dark and its presence decreases the reflectance from the soil up to an organic matter content of around 4-5% but beyond it hardly effects.

: Effect of iron oxide:

Iron oxide gives many soils their rusty red coloration. Iron oxide selectively reflects red light.

: Effect of size and shape:

Soil aggregate size and shape influence the reflectance properties. If the size of a soil aggregate large in diameter, a decrease in reflection will result.

: Spectral reflectance and signature of vegetation

Components that are involved in classifying vegetation from remote sensing images received from satellites include chemical properties and physical properties recorded for the vegetation (including surface texture, roughness and local slope properties).

Several factors influence the reflectance quality of vegetation on satellite and remote sensing images. These include brightness, greenness and moisture. Brightness is calculated as a weighted sum of all the bands and is defined in the direction of principal variation in soil reflectance. Greenness is orthogonal to brightness and is a contrast between the near infrared and visible bands. It is related to the amount of green vegetation in the scene. Moisture in vegetation will reflect more energy than dry vegetation.

Leaf properties that influence the leaf optical properties are the internal or external structure, age, water status, mineral stresses, and the health of the leaf (figure 3.8). It is important to note that the reflectances of the optical properties of leaves are the same, regardless of the species. What may differ for each leaf, is the typical spectral features recorded for the three main optical spectral domains; leaf pigments, cell structure and water content.

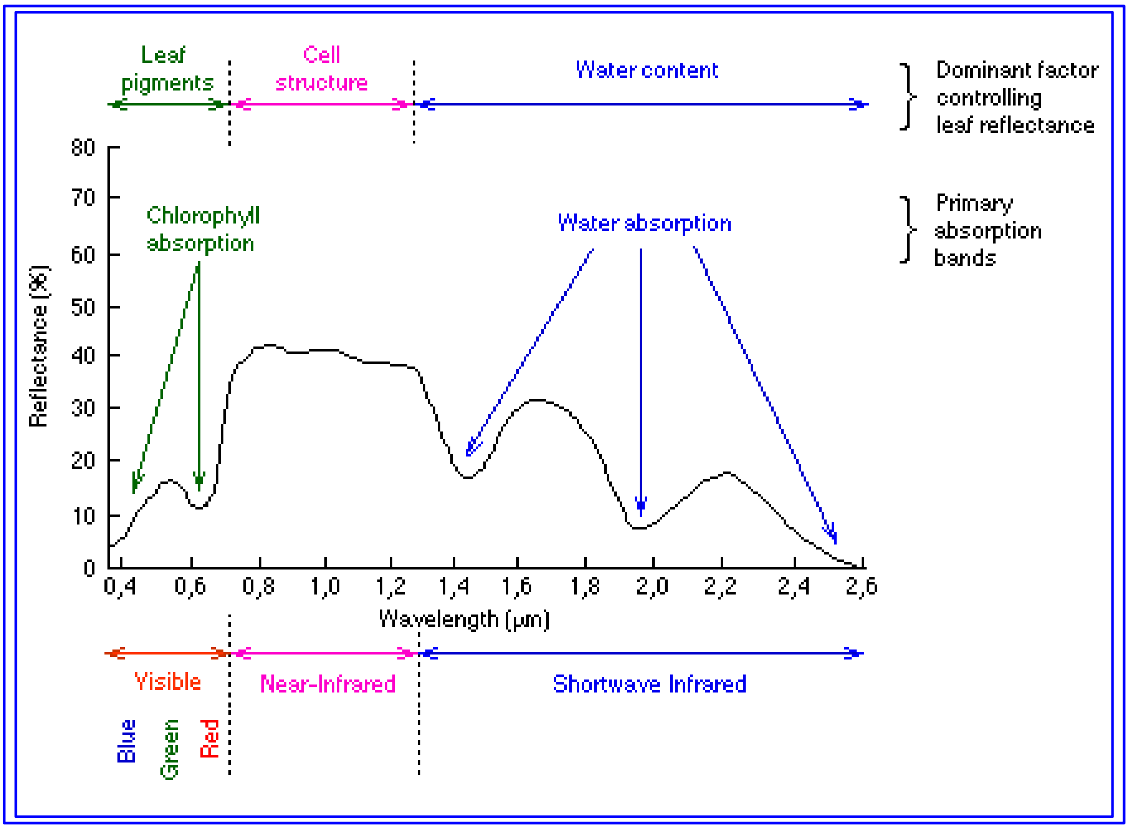


Fig 3.8: Typical spectral response characteristics of green vegetation

(Source: http://www.rsunt.geo.ucsb.edu/rscc/vol2/lec2/2_2.html#21)

Electromagnetic wavelengths affect different parts of plant and trees. These parts include leaves, stems, stalks and limbs of the plants and trees. The length of the wavelengths also plays a role for reflection that occurs. Tree leaves and crop canopies reflect more in the shorter radar wavelengths, while tree trunks and limbs reflect more in the longer wavelengths. The density of the tree or plant canopy will affect the scattering of the wavelengths.

Within the electromagnetic spectrum, bands will produce different levels of reflectance rates. For example, in the visible bands (400 - 700 nm), a lower reflectance will occur as more light will be absorbed by the leaf pigments than reflected. The blue (450 nm) and red (670 nm) wavelengths include two main absorption bands that absorb two main leaf pigments.

The images created by remote sensing will be influenced by these factors: quality, scale and season of photography, film type and background. Other factors that influence vegetation classification are time of day, sun angle, atmospheric haze, clouds, processing errors of transparencies/prints and errors in interpretating the images.

Photographic texture (smoothness and coarseness of images), total contrast or colour, relative sizes of crown images at a given photo scale and topographic location helps to determine the cover types of vegetation.

Different types of images will display diverse characteristics of vegetation. Satellite images can be combined with topographic data (ancillary data) for better interpretation.

: Spectral reflectance and signature of water

The majority of radiation flux incident upon water is not reflected but is either absorbed or transmitted. In visible wavelengths of EMR, little light is absorbed, a small amount, usually below 5% is reflected and the rest is transmitted. Wavelengths in the blue-green portion of the spectrum will have a high transmittance rate. The most distinctive characteristic of water, with reference to spectral reflectance, is the energy absorption at the near-infrared wavelengths (figure 3.9). Water absorbs Infrared strongly, leaving little radiation to be either reflected or transmitted. This results in sharp contrast between any water and land boundaries. The factors, which govern the variability in reflectance of a water body, are the depth of the water, suspended material within the water and surface roughness of the water.

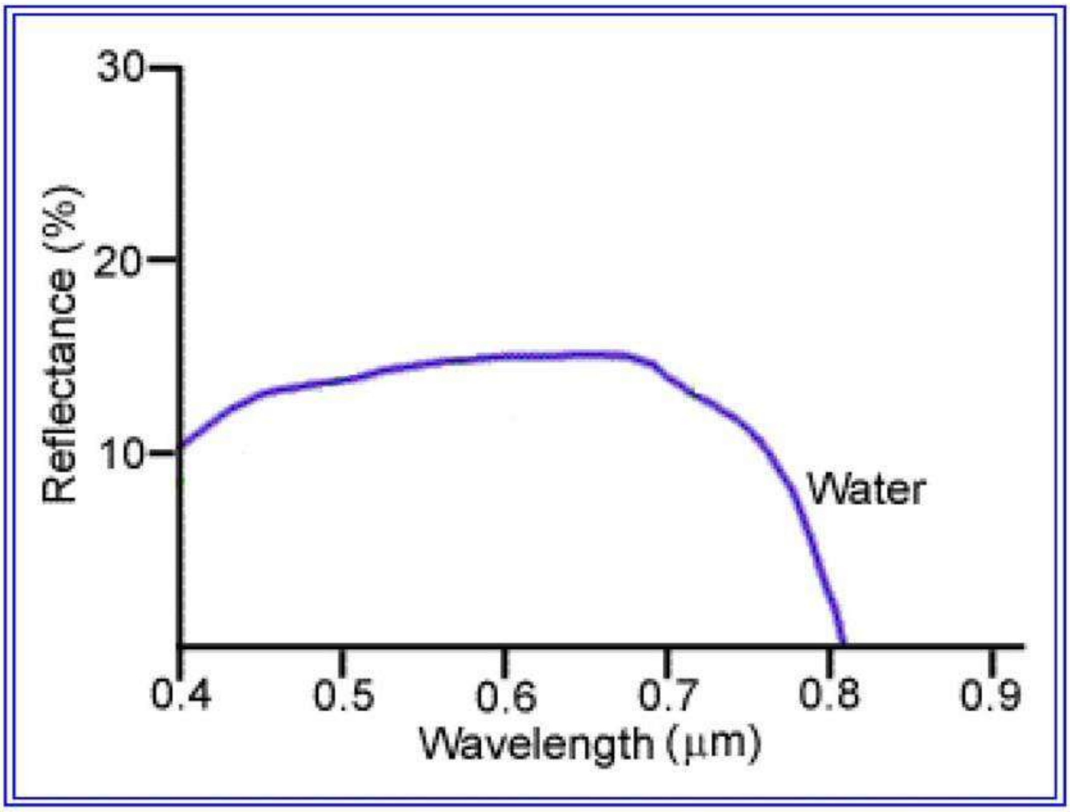


Fig 3.9: Typical spectral reflectance of water

(Source: http://www.ccrs.nrcan.gc.ca/resource/tutor/fundam/index_e.php)

Energy is absorbed by water in wavelengths. The wavelengths will interact in a complex manner with the water and matter (figure 3.10). The reflectance properties of water are a function of the water and the material in the water (organic and/or inorganic material). In shallow water, some of the radiation is reflected not by the water itself but from the bottom of the water body. Therefore, in shallow streams often the underlying material determines the water body's reflectance properties and colour in the FCC. If the water has a large amount of suspended sediment present, than a higher visible reflectance will result compared to clearer waters. Turbulent water will also affect the amount of energy absorbed and transmitted. The amount of chlorophyll will also affect

The amount of water reflectance. An increase in chlorophyll will result in a decrease of blue wavelengths and increase in green wavelengths. Water bodies that contain very high amount of chlorophyll have reflectance properties that resemble, at least in part, those of vegetation with increased green and decreased blue and decreased red reflectance. Temperature variations also affect the reflectance of water throughout the day.

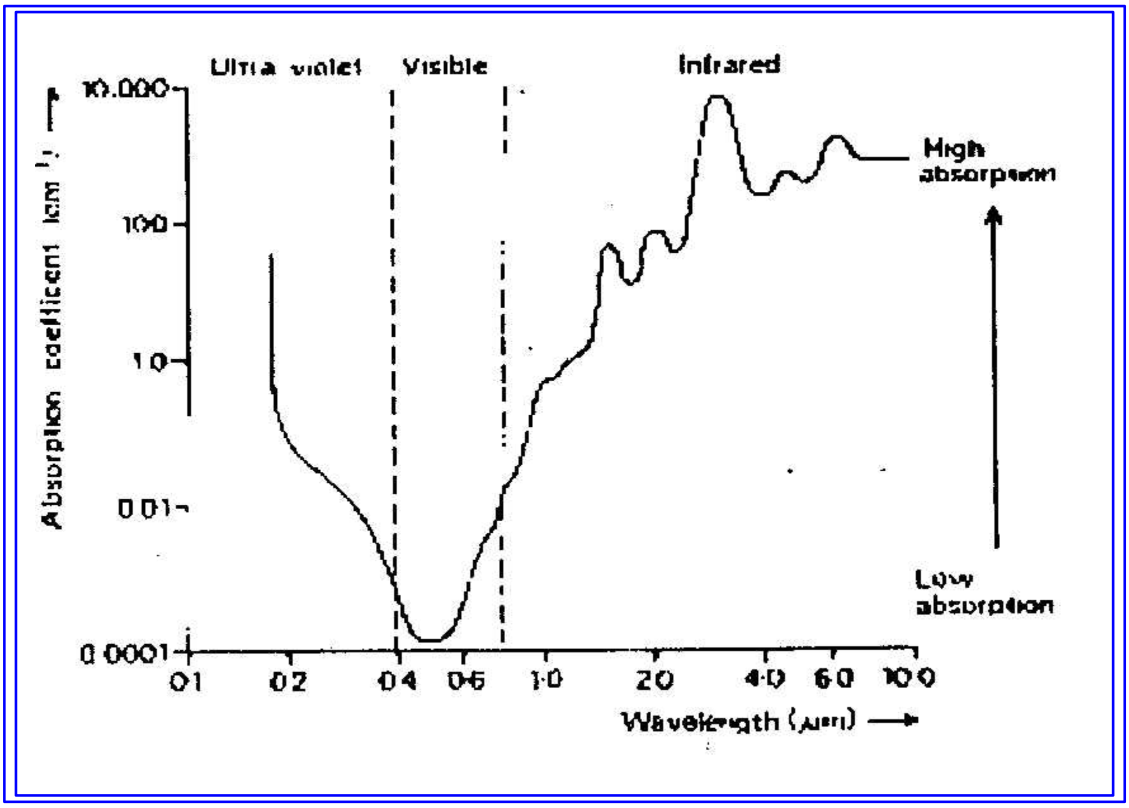


Fig 3.10: Absorption of electromagnetic radiation by sea water

(Source: Tutorials from Indian Institute of Remote Sensing, Dehradun, India)

The roughness of water surface can also affect its reflectance properties. If the surface is smooth then light is reflected, giving very high or very low reflectance, dependent upon the location of the sensor. If the surface is rough then there will be increased scattering at the surface, which in turn will increase the reflectance.

FUNDAMENTALS OF GIS (GEOGRAPHICAL INFORMATION SYSTEM)

UNIT - II

Shaik Saleemmiya,
Assistant Professor,
Dept. of Civil Engineering, Ace Engineering College, Ghatkesar.

IV-I

Topics

1. Geographical Information System
2. Modelling Real World Features Data
3. Data Formats – Spatial and Non-Spatial
4. GIS Components
5. Data Collection and Input
6. Data Conversion
7. Data Management
8. Database Structures, Files; Standard Data Formats
9. Compression Techniques
10. Hardware – Computing, Printing and Scanning Systems
11. Software – Standard Packages like ArcGIS, AutoCAD Map, Map Info etc.

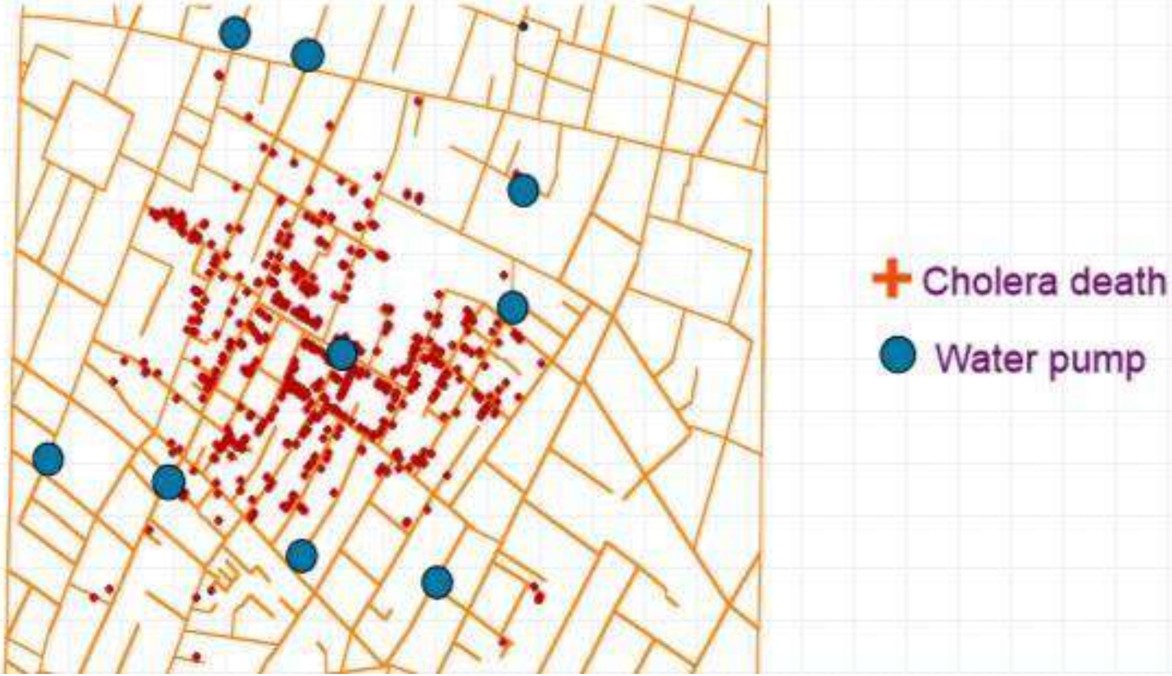
Geographical Information System

What is GIS?

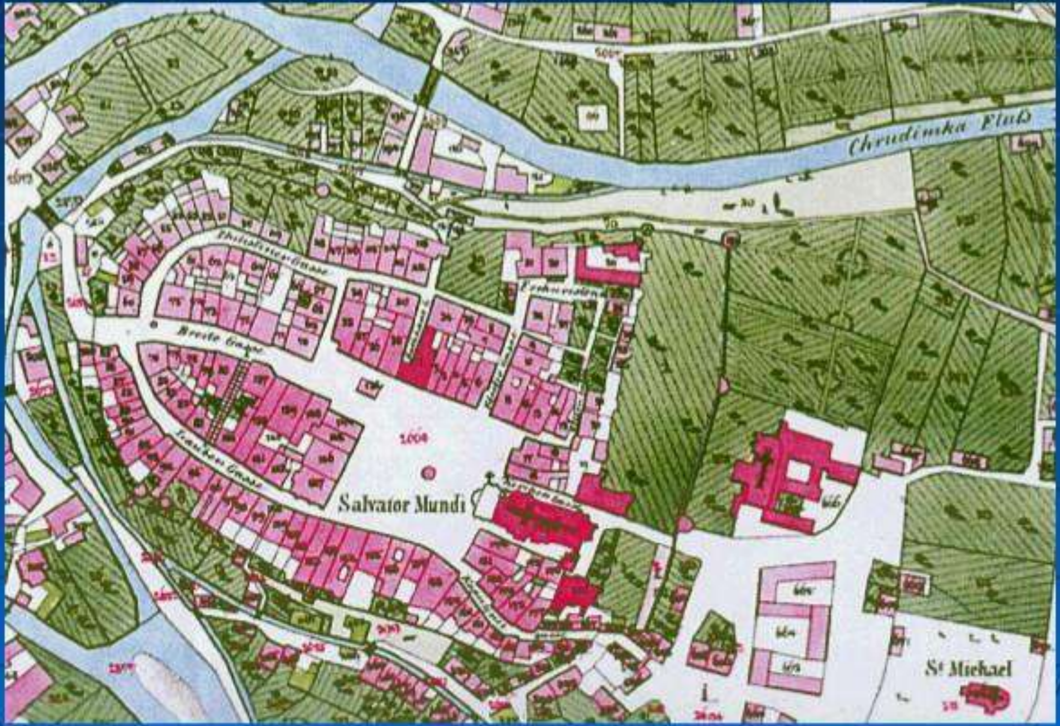
A Geographic Information System (GIS) is a computer system for **capturing, storing, checking, and displaying** data related to positions on Earth's surface.

GIS concept is not new!

- London cholera epidemic 1854

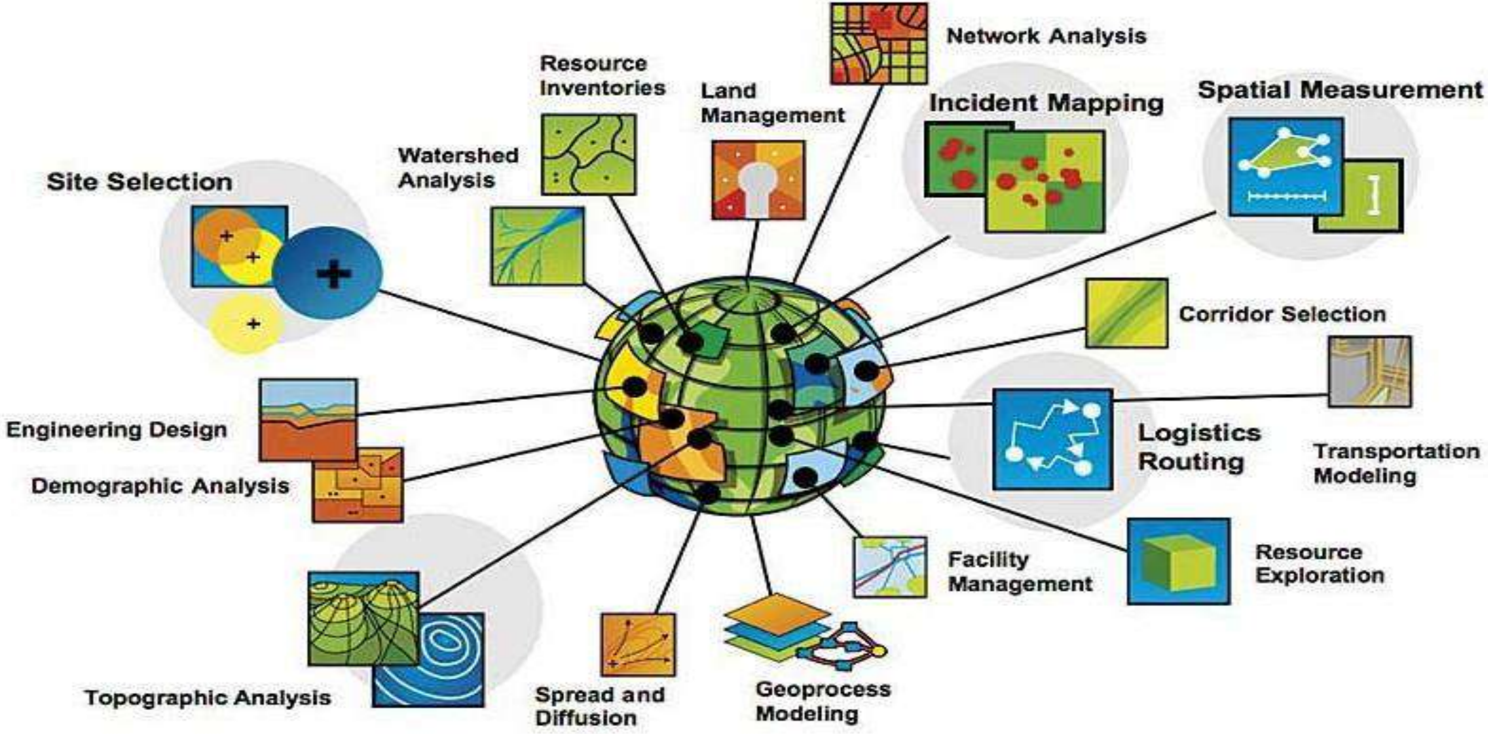


Sample Map
Land Parcel of Village



GIS Is Being Applied Around the World

Across Many Disciplines, Professions, and Organizations



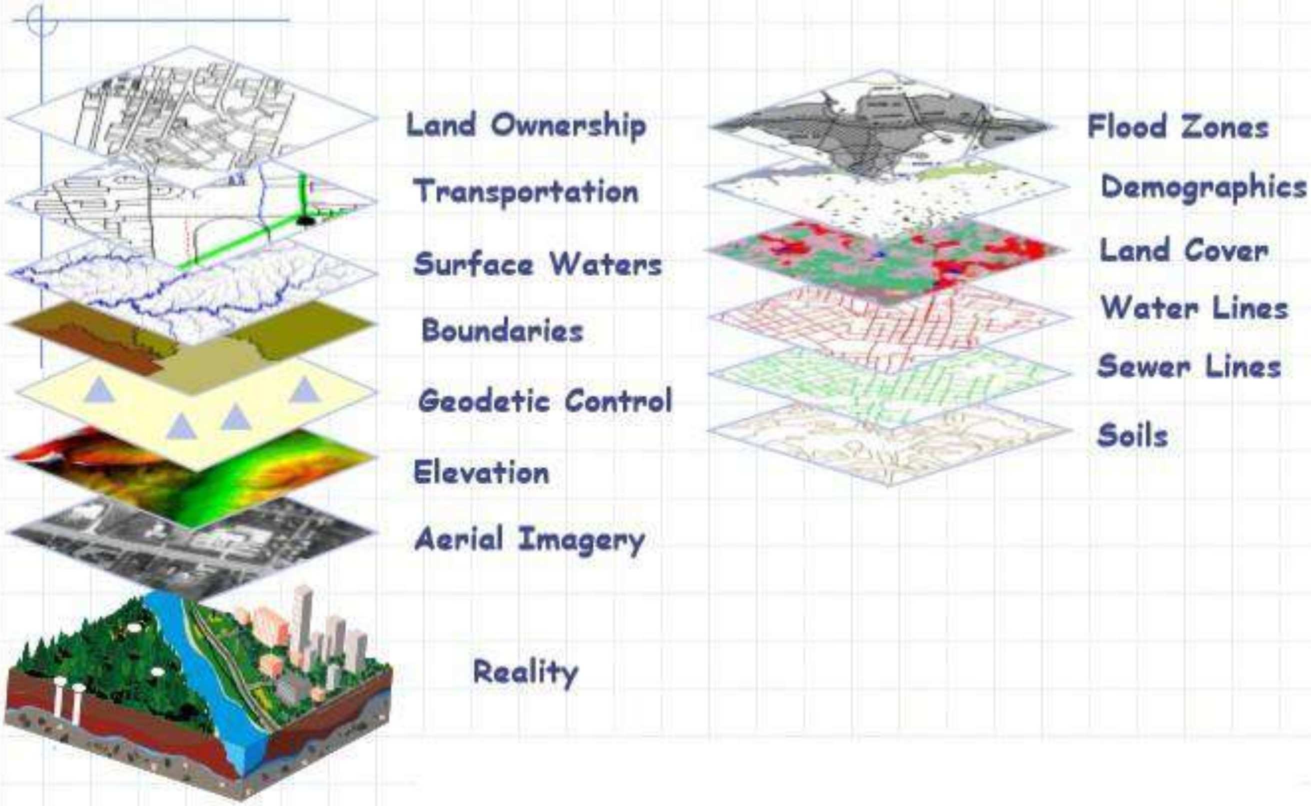
GIS Urban Applications - Examples

- Permit Issuing and Tracking
- Municipal Facilities Management
- Land Records Management
- Optimum Site Selection for
  - schools, hospitals, commercial centers..etc.

Other Applications

- Facilities management
- Marketing and retailing
- Transport/vehicle routing
- Health
- Insurance
- Natural Resources Management
- Water Wells Inventory
- Forestry/Wild Life Management
- Environmental Impact Assessment
- Oil Spill Tracking
- Geology and Oil Exploration
  - and many more....

Layers Concept in GIS



Typical Layers in a GIS Database

- Base map layer includes:
  - Geodetic control points
  - Contour lines
  - Permanent geographic features such as coastlines and rivers
- Land parcels layer

Typical Layers in a GIS Database…

- Aerial Photo or satellite image as background
- Streets network layer
- Infrastructure/Utilities network layers:
  - Electric,
  - Phone,
  - Water, and
  - Sewer

Typical Layers in a GIS Database…

- Planning layers
  - Existing land use
  - Planned land use/zoning
  - Urban growth management lines
  - Vacant land inventory
- Facility location layers - schools, health centers, etc.

Typical Layers in a GIS Database…

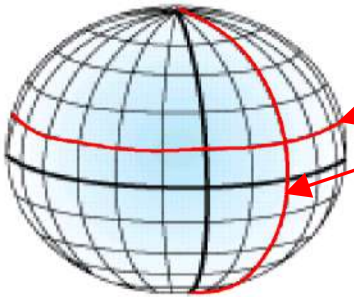
- Boundary layers
  - Municipal boundaries
  - Neighborhood boundaries
- Event location layers:
  - Location of car accidents
  - crime incident maps

Typical Layers in a GIS Database…

- Environmental layers
  - Well sites
  - Contour maps
  - Hazardous waste sites
  - Flood plains/ wetlands

Geographic Information Systems (GIS)

Geography – The study of where features are located on the Earth's surface.



Geographic coordinate systems

latitude

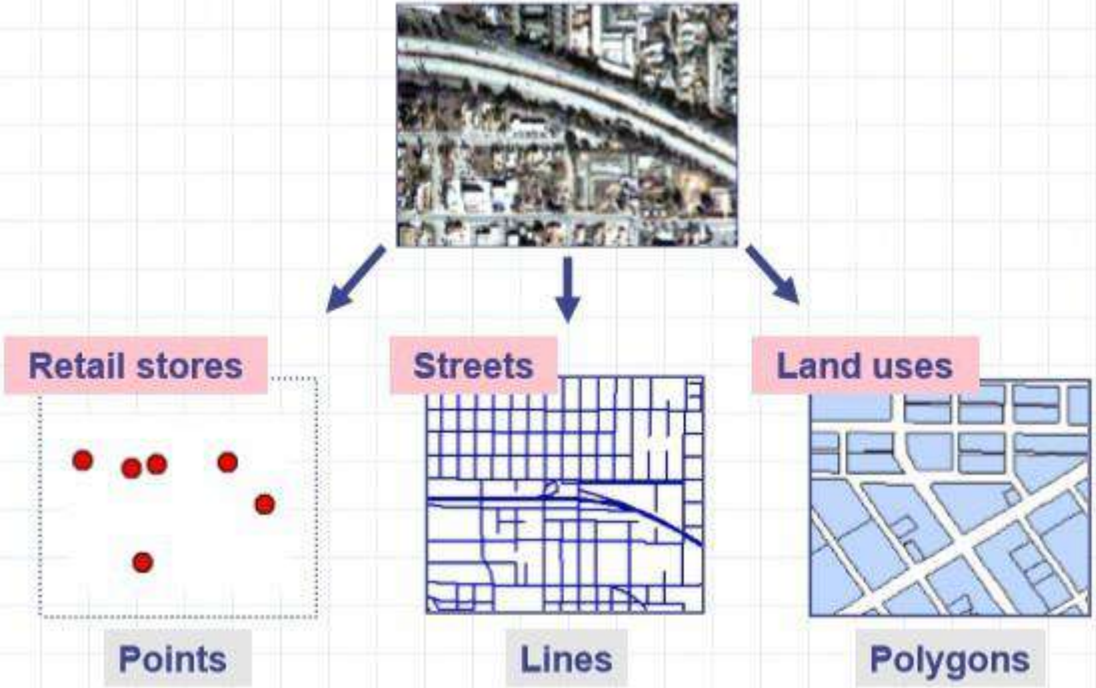
longitude

Represent exact positions on the Earth

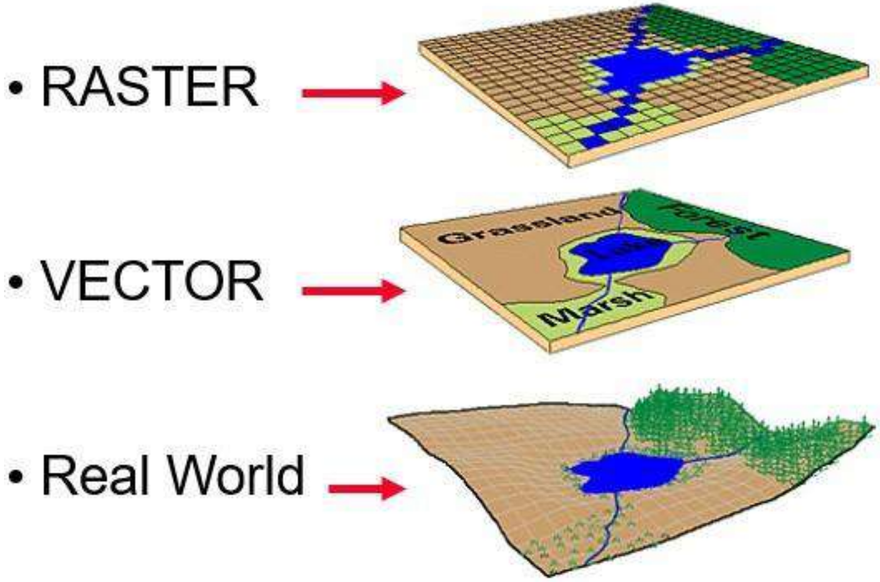
42.3216, -71.089118 (lat. long. for Boston)

Georeferencing / Geocoding – The process of assigning geographiccoordinates to features to represent their location.

Features Location / Spatial Representation: Points, Lines, and Polygons



Modelling Real World Features Data



Types of Data

1. Attribute data:

Says what a feature is

Ex.: statistics, text, images, sound, etc.

2. Spatial data:

Says where the feature is

- Vector data: Discrete data
  - Points
  - Lines
  - Polygons (zones or areas)
- Raster data: Continuous data

Two fundamental types of data

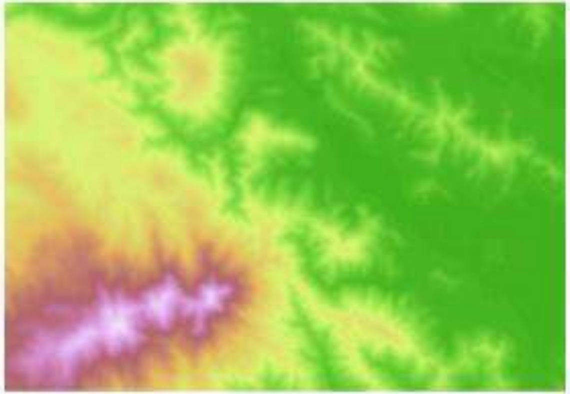
- **Vector**
  - A series of x,y coordinates
  - For discrete data represented as points, lines, polygons
- **Raster**
  - Grid and cells
  - For continuous data such as elevation, slope, surfaces

- A Desktop GIS should be able to handle both types of data effectively!

Data Representation (Raster)

A spatial data model that defines space as an **array of equally sized cells** arranged in rows and columns. Each cell contains an **attribute value** and **location coordinates**.

Raster as Satellite Imagery



Raster as Elevation Surface

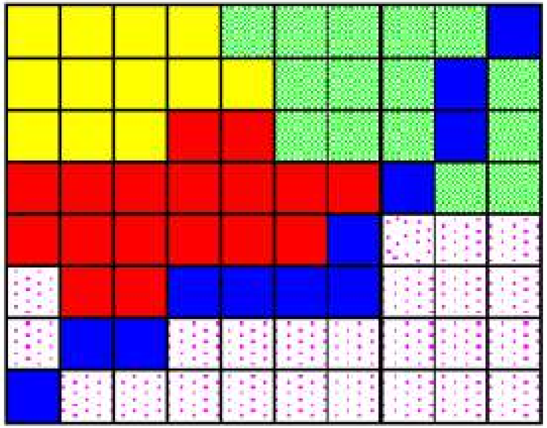
Raster Data

Stores images as rows and columns of numbers with a Digital Value/Number (DN) for each cell.

Units are usually represented as square grid cells that are uniform in size.

Data is classified as "*continuous*" (such as in an image), or "*thematic*" (where each cell denotes a feature type.

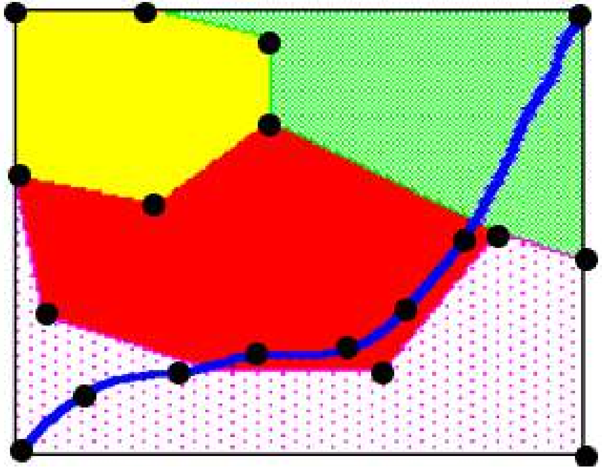
Numerous data formats (TIFF, GIF, ERDAS.img etc)



Vector Data

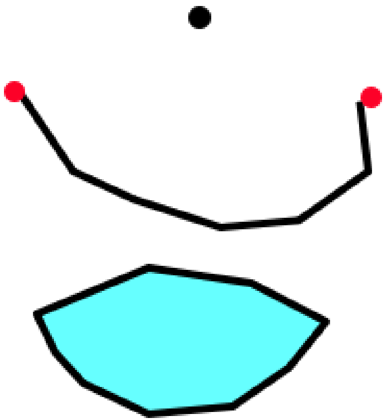
Allows user to specify specific spatial locations and assumes that geographic space is continuous, not broken up into discrete grid squares

We store features as sets of X,Y coordinate pairs.



Entity Representations

We typically represent objects in space as three distinct spatial elements:



Points - simplest element

Lines (arcs) - set of connected points

Polygons - set of connected lines

We use these three spatial elements to represent real world features and attach locational information to them.

Raster vs Vector

Raster Advantages

The most common data format

Easy to perform mathematical and overlay operations

Satellite information is easily incorporated

Better represents “continuous”- type data

Vector Advantages

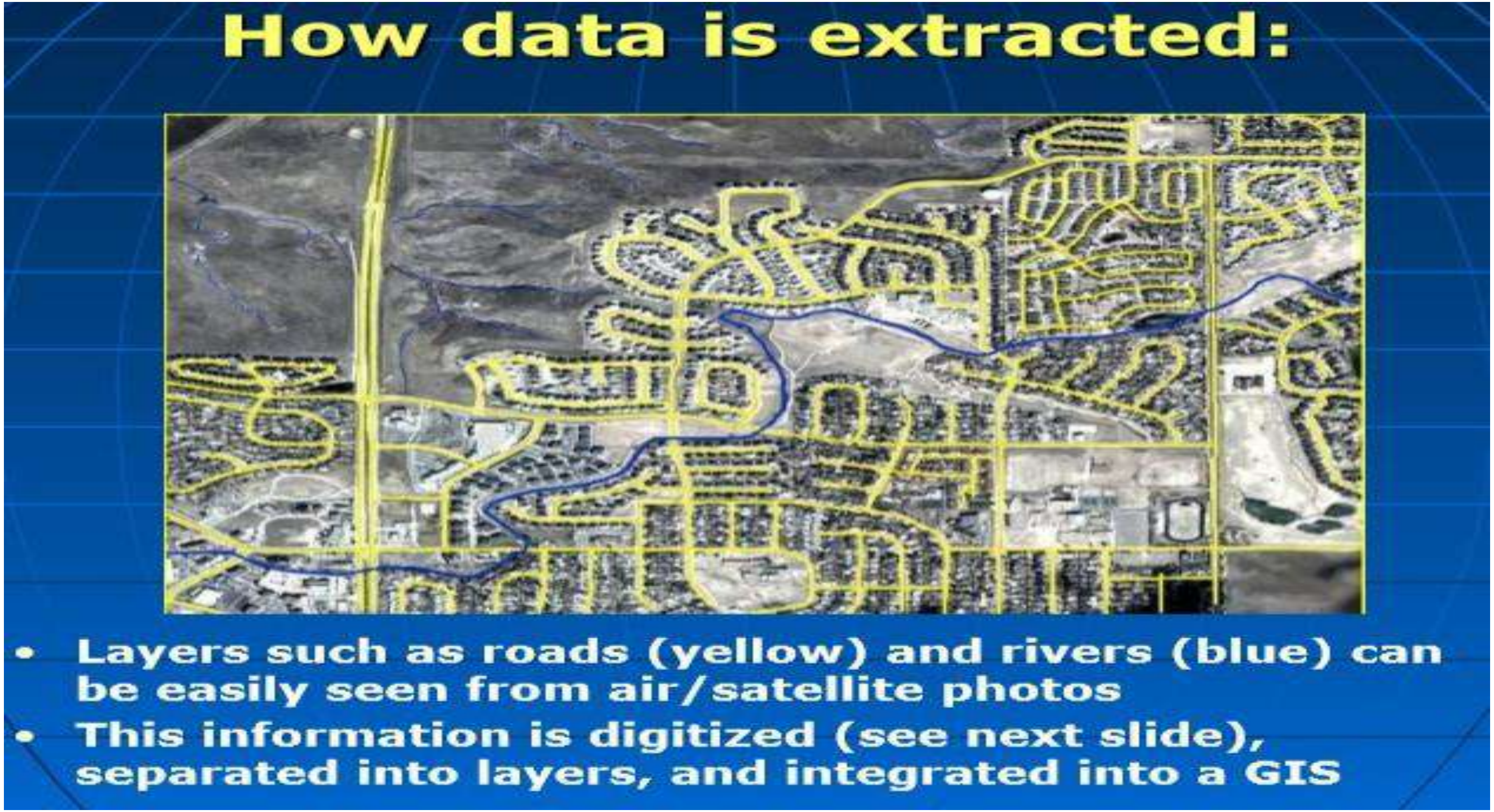
Accurate positional information that is best for storing discrete thematic features (e.g., roads, shorelines, sea-bed features.

Compact data storage requirements

Can associate unlimited numbers of attributes with specific features

Components of GIS

Data Collection



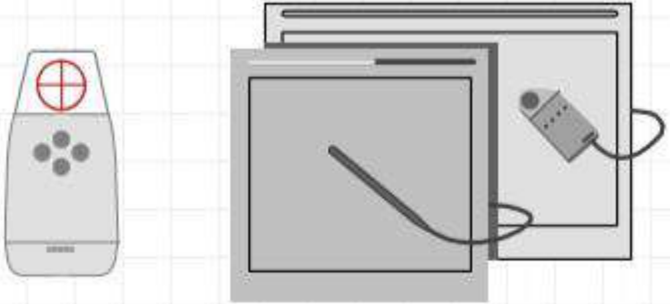
Data Collection and Input

Data Input Systems

- Manual Digitizing
- Automated Scanning

Manual Digitizing

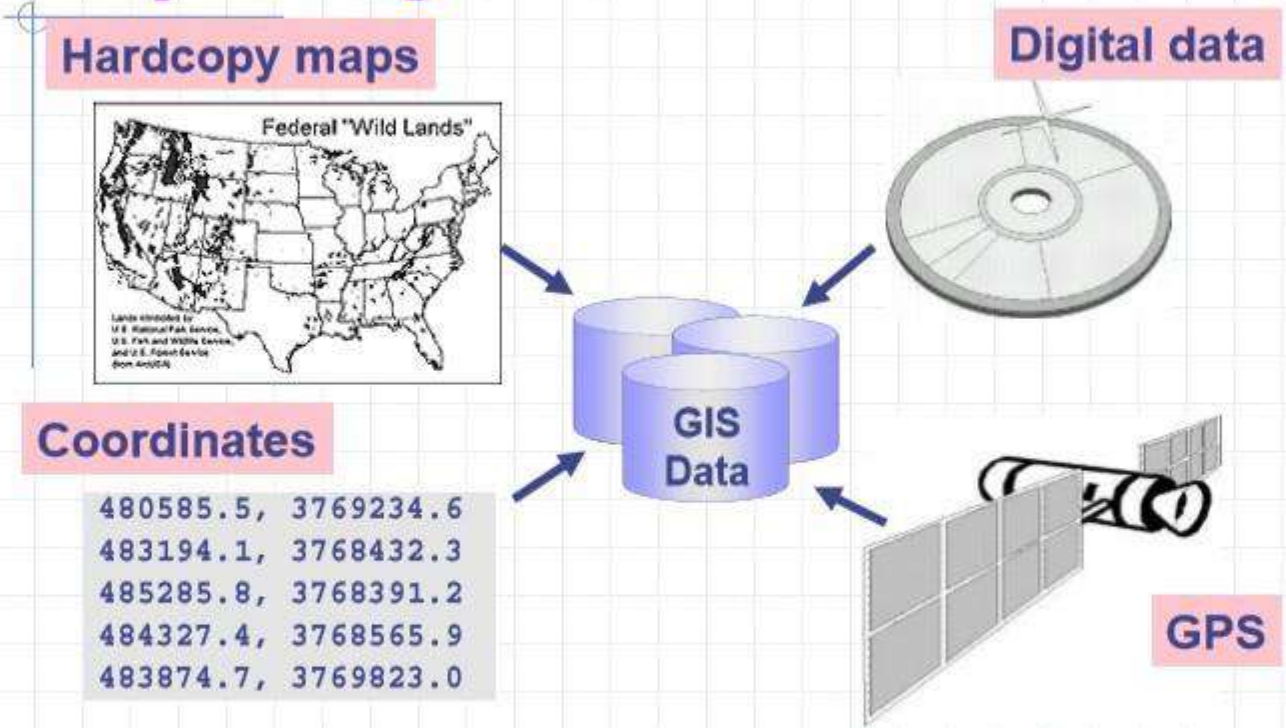
- Done through a digitizer
- Coordinates entered through a cursor
- Slow and costly
- Accurate results



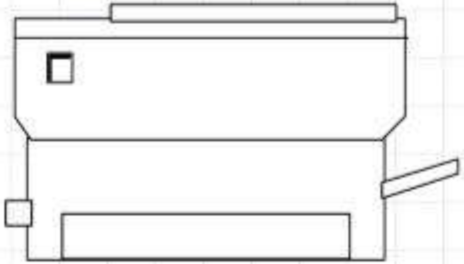
Automated Scanning

- Maps scanned through a scanner
- Faster and lower cost
- Requires substantial amount of editing
- Lower accuracy

Capturing data

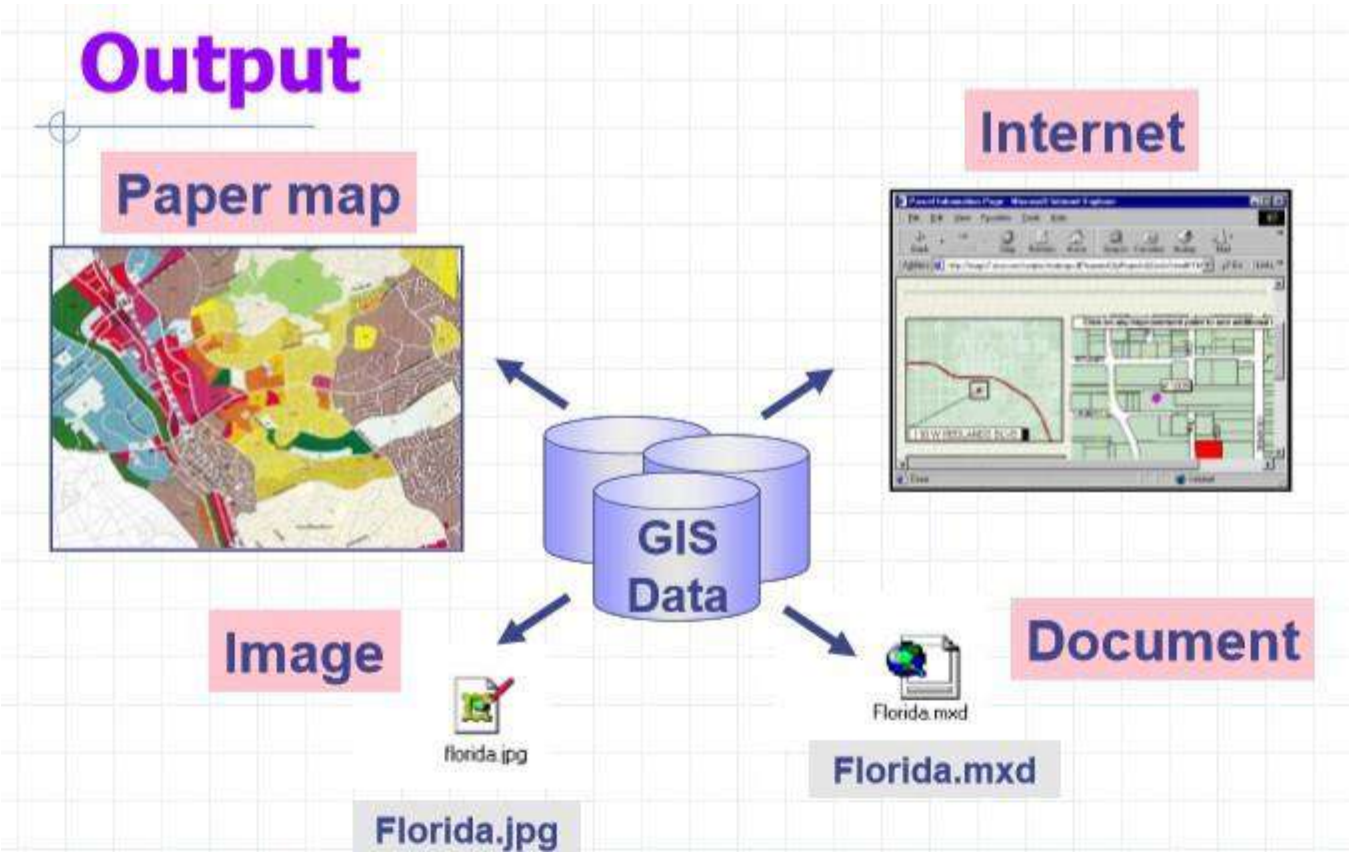


Data Output Peripherals



Printers

Plotters


Output
Paper map
Internet
GIS
Data
Image
Document
florida.jpg
Florida.mxd
Florida.mxd
Florida.jpg

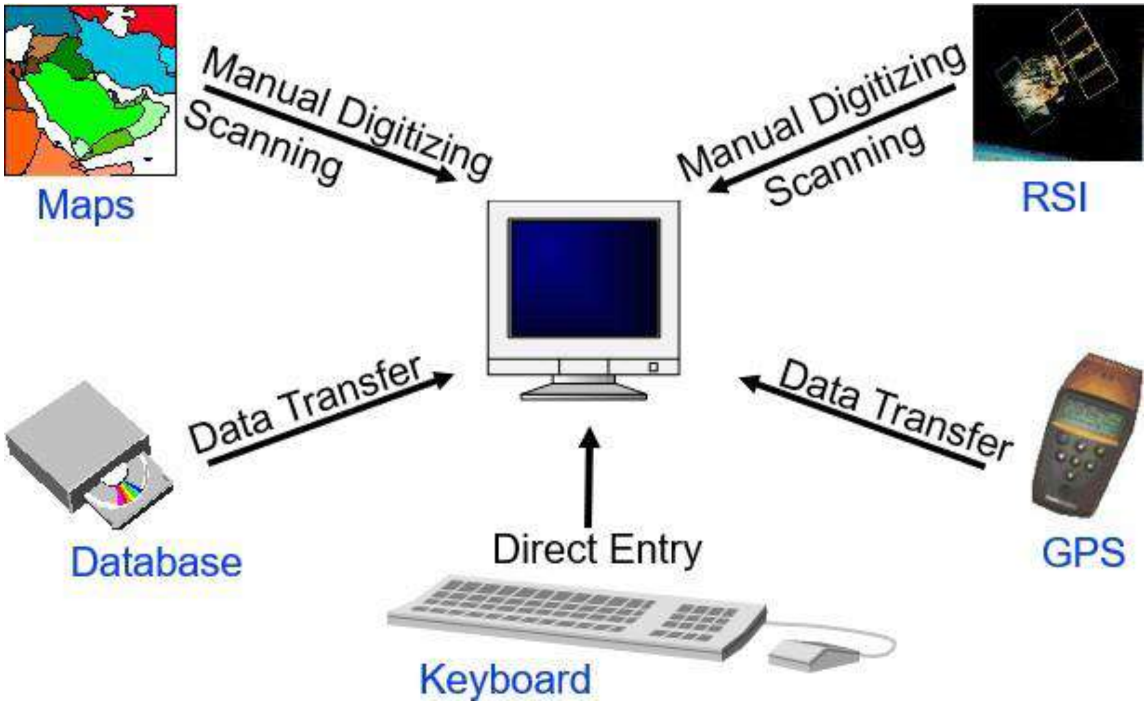
Commercial GIS Packages

- ArcGIS (ESRI, USA)
- GeoMedia (InterGraph, USA)
- GeoGraphics (Bentley Corp., USA)
- MapInfo (MapInfo Corp., USA)
- STAR (STAR Company, Belgium)
- Many Others

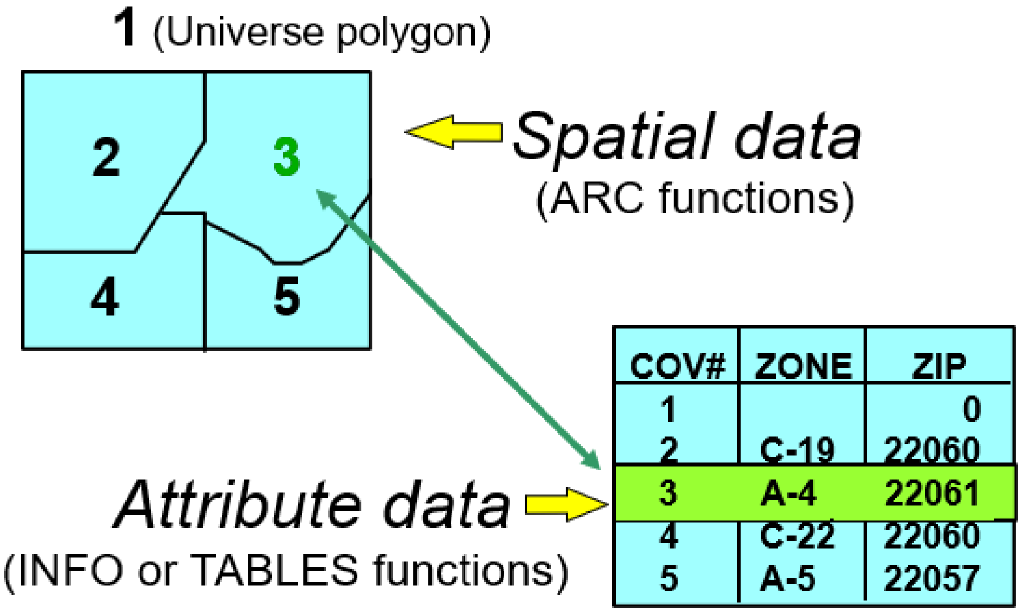
GIS Industry & Market

- Average growth around 15% per year over the last two decades
- U.S. vendors dominate world market
- Business customers rapidly increasing

Data Collection



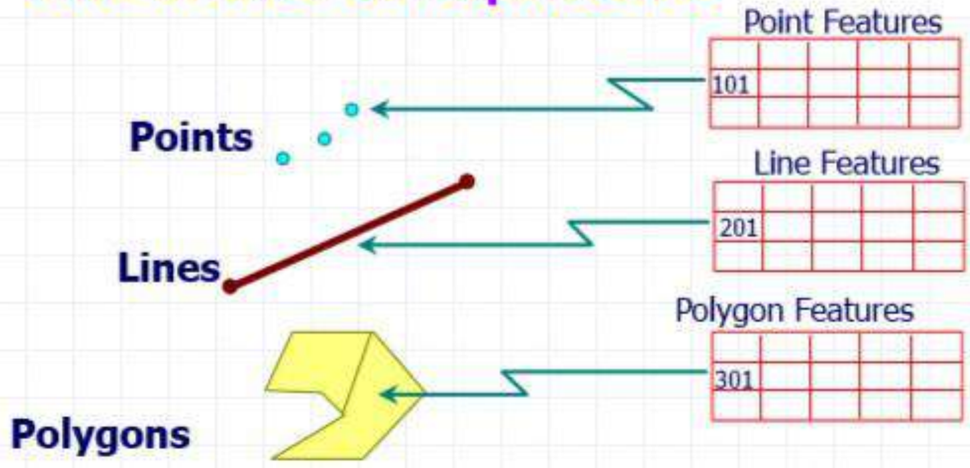
Database Structures, Files; Standard Data Formats



COV#	ZONE	ZIP
1		0
2	C-19	22060
3	A-4	22061
4	C-22	22060
5	A-5	22057

Features Location & Descriptive Attributes

One-to-One Correspondence

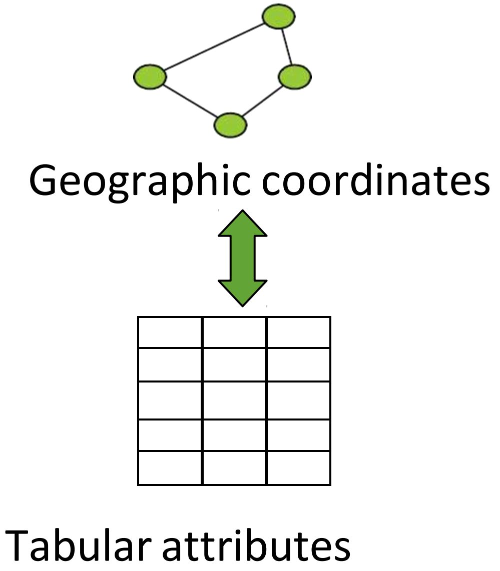


Features Location/Spatial Representation

Descriptive Attributes

Data Model

A **geographic data model** is a structure for organizing geospatial data so that it can be easily stored and retrieved.

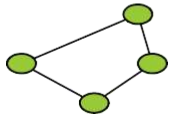


Geodatabase model

Stores geographic coordinates as one attribute(shape) in a relational database table

Uses MS Access for “Personal Geodatabase” (single user)

Uses Oracle, SQL Server, MySQL or other commercial relational databases for “Enterpris Geodatabases” (many simultaneous users)



Shape		
e		

Common GIS Data Format

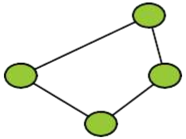
Vector Data Model

- **Shapefile** (.shp, .shx & .dbf)
- Geography Markup Language **GML** (.gml)
- Personal Geodatabase (.mdb)
- **File Geodatabase** (.gdb)

Raster Data Model

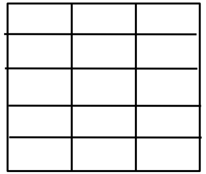
- Enhanced Compressed Wavelet **ECW** (.ecw)
- **GeoTIFF** = TIFF file with GeoReference (.tif)
- MultiResolution Seamless Image Database **MrSID** (.sid)
- **JPEG2000** = JPEG file with GeoReference (.jpg)

File-based Data Models



Arc

Geographic coordinates and attributes are stored in separate but linked files



Info

Coverages

- Developed for workstation Arc/Info ~ 1980
- Complex structure, proprietary format
- Attributes in Info tables

Shapefiles

- Developed for ArcView ~ 1993
- Simpler structure in public domain
- Attributes in dBase (.dbf) tables

Introduction to Data Compression

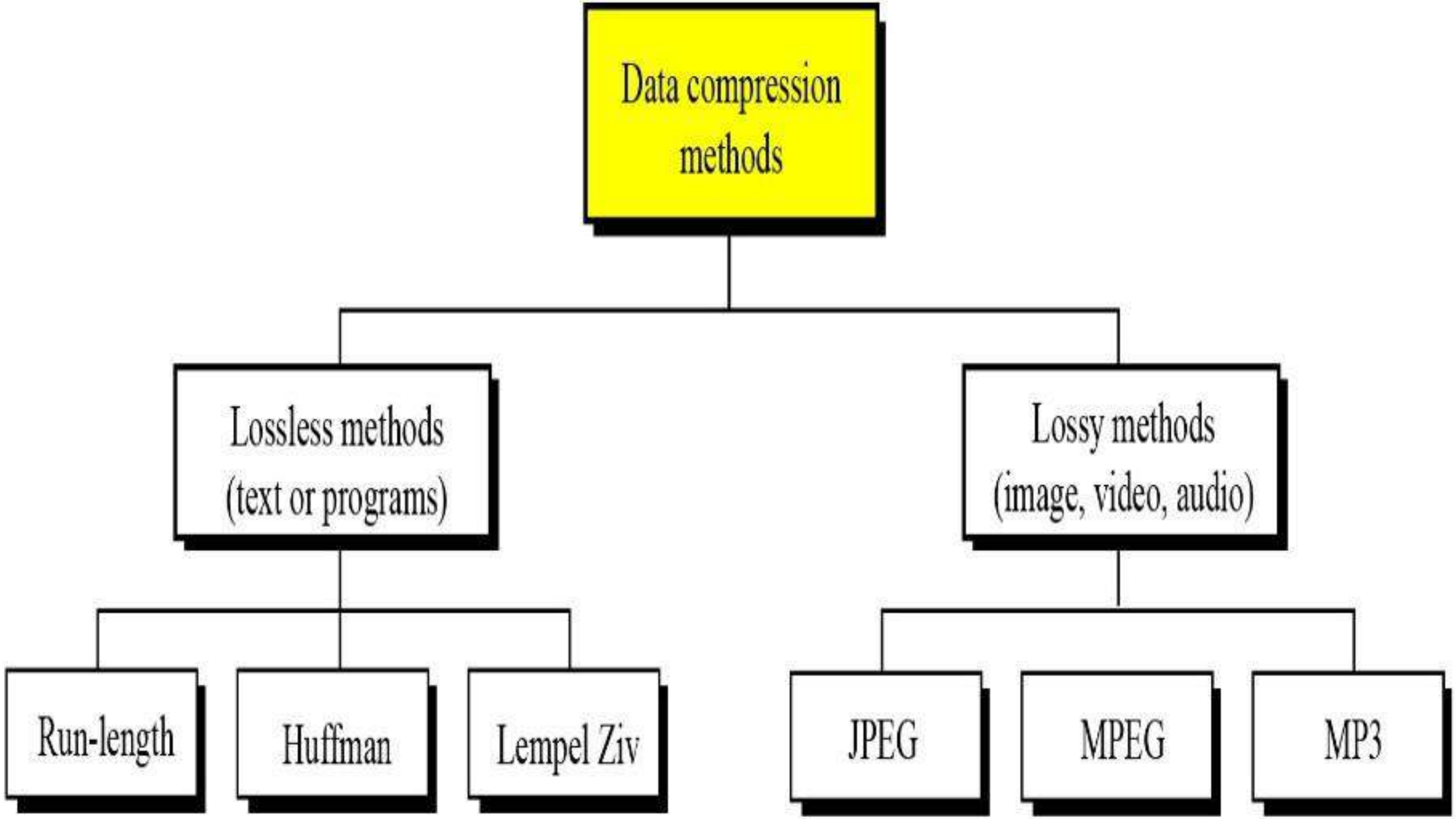
Data compression has important application in the areas of data transmission and data storage.

Many data processing applications are required for storage of large volumes of data, and the number of such applications is constantly increasing as the use of computers extends to new disciplines.

Compressed data reduces storage and/or communication costs.

Why data compression?

- Decrease storage requirements
- Strict bandwidth bound of communication channels
- Effective use of communication bandwidth; reduce amount of data without loosing information
- Improves program execution speed by reducing disk access times
- Increase data security
- Multimedia data on information superhighway
- Growing number of 'killer' applications (digital library, websites like Google, medical imaging, satellite images, journals, newspaper, online radio/movie/video, etc.)



Data compression methods

GIS Functions

- **Capture:** Paper maps, digital data, coordinates, GPS, all together = gis data
- **Store:** Vector (database consisting of points, lines and polygons) or raster formats (gridcells or pixels)
- **Query:** Identifying specific features, identify features based on conditions (counties with population greater than 500,000)
- **Analyze:** Proximity (parcels 100 feet from the road), overlay, network analysis (howlinear features are connected
- **Display:** Maps, graphs, Reports
- **Output:** Paper maps, image

Spatial Data Manipulation and Analysis

- Common Manipulation
  - Query
  - Reclassification
  - Map Projection changes
- Spatial Analysis
  - Buffering
  - Overlay
  - Network

INTRODUCTION TO MAPS (GEOGRAPHICAL INFORMATION SYSTEM)

UNIT – II

Shaik Saleemmiya,

Assistant Professor,

Dept. of Civil Engineering, Ace Engineering College, Ghatkesar

Syllabus

1. Mapping Concepts
2. Analysis with paper based Maps
3. Limitations of Paper based Maps
4. Computer Aided Cartography History and Development
5. GIS Definition
6. Advantage of Digital Maps

Topics

1. Mapping Concepts
2. Analysis with paper based Maps
3. Limitations of Paper based Maps
4. Computer Aided Cartography
5. Introduction to GIS
6. Advantage of Digital Maps

Cartography

ART / SCIENCE OF CREATING MAPS

Cartographic design uses symbology and typography to convey geographical information in a manner which is both appealing to the eye and easily understood.

A map symbol is a graphic design which represents a map feature and its characteristics.

Mapping Concepts

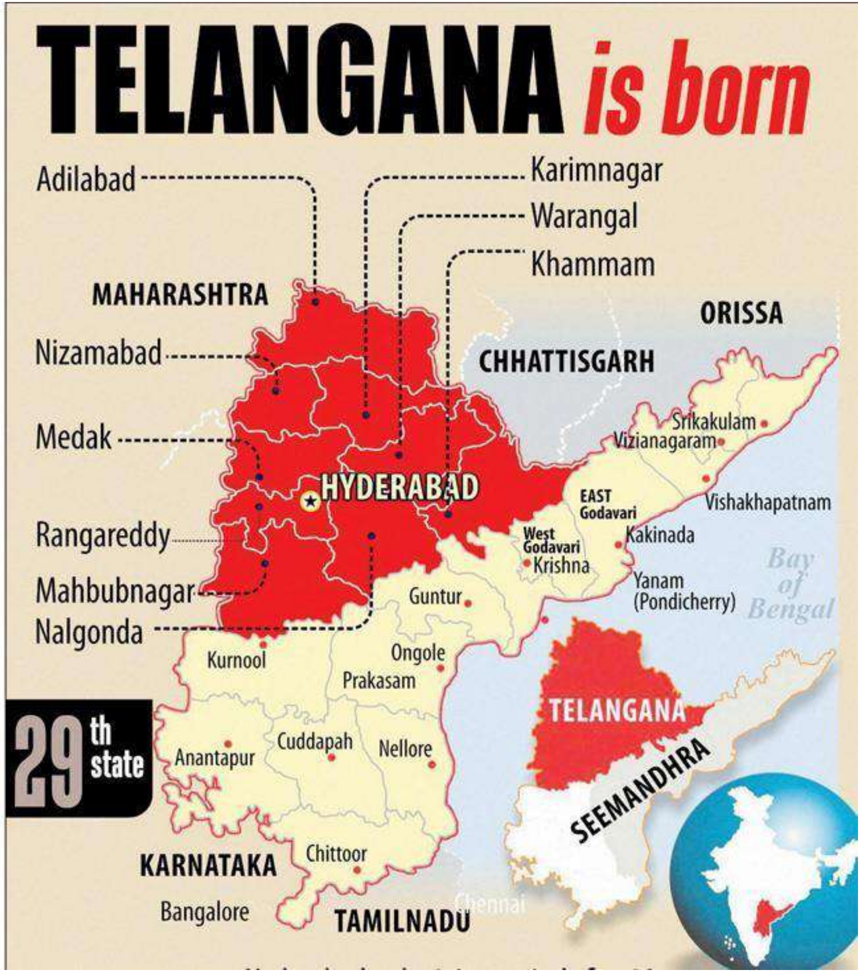
A map is a two dimensional representation of earth surface which uses graphics to convey geographical information. It describes the geographical location of features and the relationship between them.

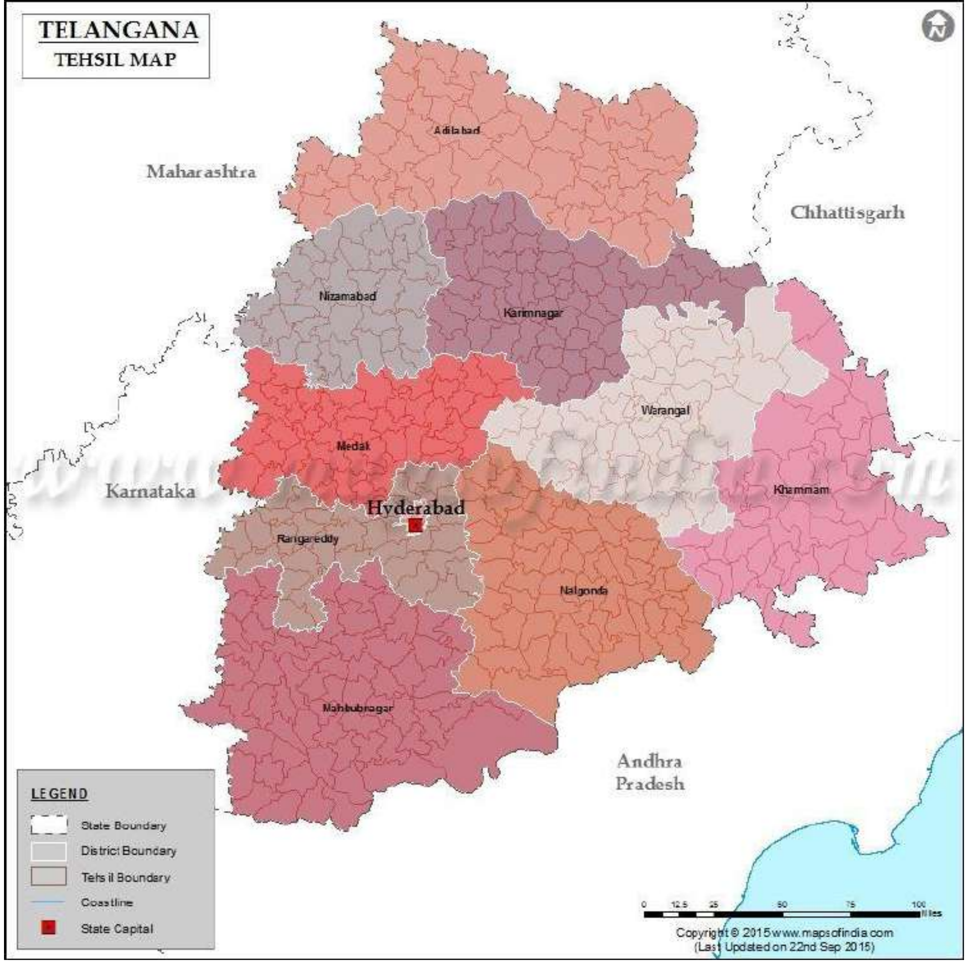
Most maps have keys and scales that help the reader understand the map. The key explains symbols and colors used on the map, so you know what they represent. Scales tell you the relationship between the size of the map and the size of the placein real life.

REPRESENTATION OF EARTH

Globes

Maps


TELANGANA is born
Adilabad
Karimnagar
Warangal
Khammam
MAHARASHTRA
ORISSA
CHHATTISGARH
Nizamabad
Medak
HYDERABAD
Rangareddy
Mahbubnagar
Nalgonda
Srikakulam
Vizianagaram
Vishakhapatnam
EAST Godavari
West Godavari
Kakinada
Krishna
Guntur
Yanam (Pondicherry)
Bay of Bengal
Kurnool
Ongole
Prakasam
29 th state
Anantapur
Cuddapah
Nellore
TELANGANA
SEEMANDHRA
KARNATAKA
Chittoor
Bangalore
TAMILNADU
Chennai


TELANGANA
TEHSIL MAP
Adilabad
Maharashtra
Chhattisgarh
Nizamabad
Karimnagar
Warangal
Medak
Karnataka
Khammam
Hyderabad
Rangareddy
Nalgonda
Mahbubnagar
Andhra Pradesh
LEGEND
State Boundary
District Boundary
Tehsil Boundary
Coastline
State Capital
Copyright © 2015 www.mapsofindia.com
(Last Updated on 22nd Sep 2015)

Map Scale

Map Scale

The size of earth is too big to be represented as it is on a map. To represent the whole earth or part thereof on a small map, the concept of scale is used.

Representation of Scale:

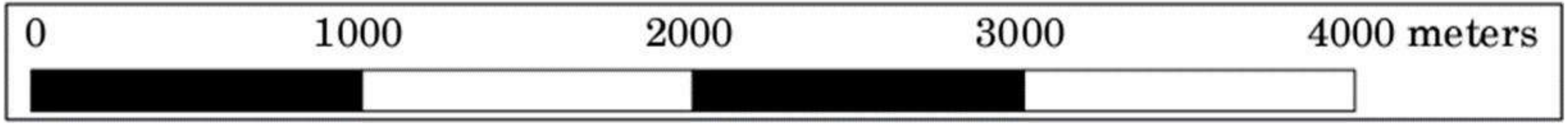
1. Statement: Expressed in words such as '1 centimeter to the kilometer' which means one centimeter on the map represents one kilometer on the ground.

2. Representative fraction: It is expressed in fraction. If the scale is 1:50000, it means that 1 unit of measurement on map represents 50,000 units on ground. It is also known as numerical scale.

3. Graphic: The scale is shown in form of a strip, where the strip is divided into a number of equal parts andis marked to show what these divisions represent on actual ground. It is also known as plain scale or linear scale.

a) (1 centimeter represents 250 meters)

b) 1: 25 000

c) 

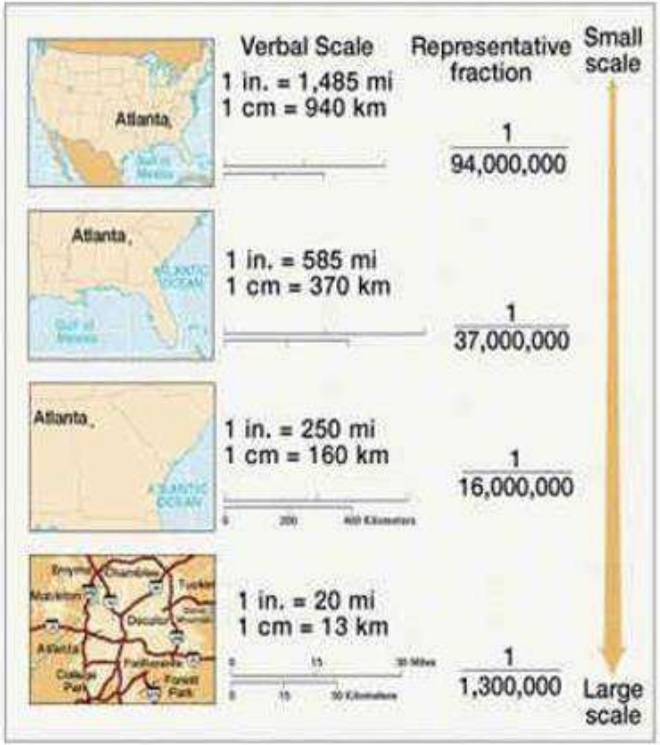
A qualitative nomenclature used to express the relative scale of a map consists of small or large scale. The qualitative differences of these two scales are:

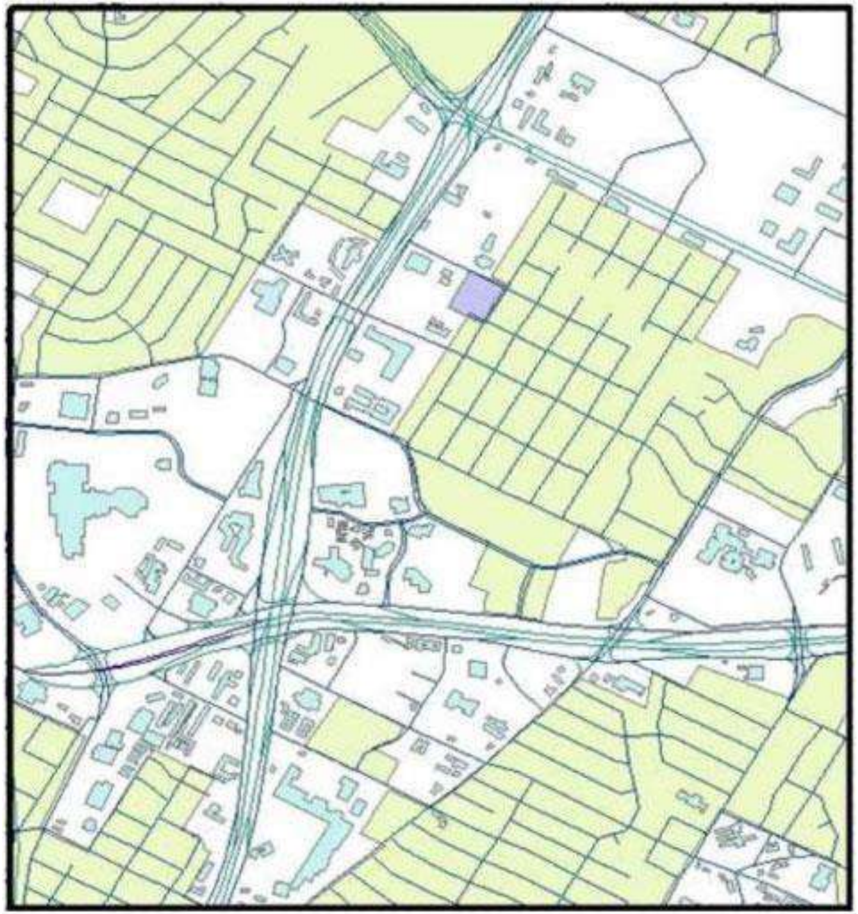
Large scale map	Small scale map
Area covered is small	Area covered is large
Details are more, objects are seen as large	Less details, objects are seen as small

To understand the difference we can take help of representative fraction method.

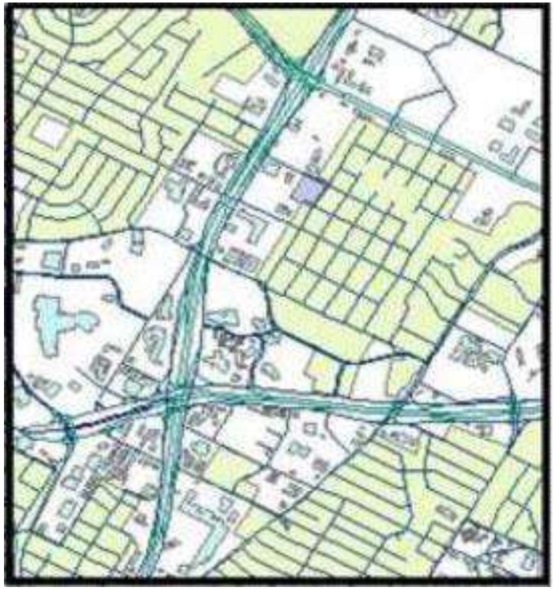
A map on a scale of 1:5,000 means that the size of the objects on map is 1/5000 of their size on the ground. Similarly, a map on a scale of 1:50,000 mean that the size of the objects on map is 1/50000 of their size on the ground.
As the map on the scale of 1:5000 shall show the objects on map bigger (as well as smaller objects not visible earlier shall become visible) than those shown by 1:50,000 scale map therefore it is a large scale map.

Large Scale Vs. Small Scale





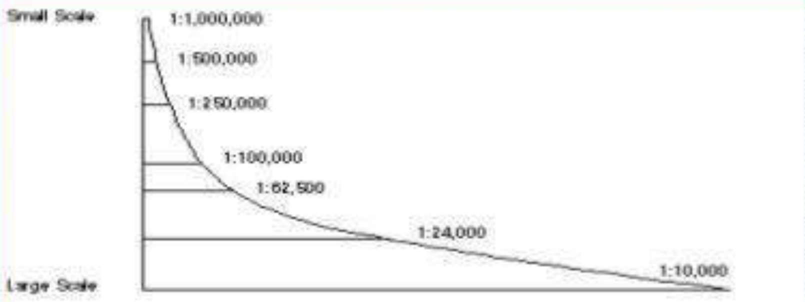
(a) 1:25,000 [Large scale]



(b) 1:50,000 [Small scale]

Map Resolution

- It refers how accurately location and shape of features can be depicted for a given map scale.
- Scale affects resolution



Each line represents one kilometer at its respective scale.

Types of Maps

The maps can be classified on the following criteria:

1. Scale: Scale is important for correct representation of geographical features and phenomenon. Different features require different scales for their display. For example preparation of a cadastral map of a village and the soil map of a state would use different scale for representing the information. According to scale, maps can be classified as follows:
   a) Cadastral Maps - These maps register the ownership of land property
   b) Topographical Maps - These maps are prepared on fairly large scale and are based on precise survey
   c) Chorographical/Atlas - Drawn on a very small scale, atlas maps give a generalized view of physical, climatic and economic conditions of different regions of the world. The scale of atlas map is generally greater than 1:1000,000.
2. Purpose
   a) Natural Maps
   b) Cultural Maps

Types of Maps

Maps based on Purpose

1. Natural Maps - These maps represent natural features and the processes associated with them
    a) Astronomical map
    b) Geological map
    c) Relief map
    d) Climate map
    e) Vegetation map
    f) Soil map
2. Cultural Maps
    a) Political map
    b) Military map
    c) Historical map
    d) Social map
    e) Land-utilization map
    f) Communication map
    g) Population map

Natural Maps

These maps represent natural features and the processes associated with them. Given below is the list of some such maps:

1. Astronomical map: It refers to the cartographic representation of the heavenly bodies suchas galaxies, stars, planets, moon etc.
2. Geological map: A map that represents the distribution of different type of rocks andsurficial deposits on the Earth.
3. Relief map: A map that depicts the terrain and indicates the bulges and the depressions present on the surface.
4. Climate map: A climate map is a depiction of prevailing weather patterns in a given area. These maps can show daily weather conditions, average monthly or seasonal weather conditions of an area.
5. Vegetation map: It shows the natural flora of an area.
6. Soil map: A soil map describes the soil cover present in an area.

Cultural Maps

These maps tell about the cultural patterns designed over the surface of the earth. They describe the activities of man and related processes. Given below is the list of such maps:

1. Political map: A map that shows the boundaries of states, boundaries between different political units of the world or of a particular country which mark the areas of respective political jurisdiction
2. Military map: A military map contains information about routes, points, security and battle plans.
3. Historical map: A map having historical events symbolized on it.
4. Social map: A map giving information about the tribes, languages and religions of an area.
5. Land-utilization map: A map describing the land and the ongoing activities on it.
6. Communication map: A map showing means of communication such as railways, road, airways etc.
7. Population map: A map showing distribution of human beings over an area.

The map and the globe are similar in a manner that both of them represent earth (on particular scales) but there also exist a few differences between them, which are enumerated below:

Globe	Map
Three dimensional representation of earth in the form of a sphere	Two dimensional representation of earth in the form of a flat surface
Impossible to see all the countries of the world at a glance as only half of the globe can be seen at a time	All countries of the world can be seen on a world map at a glance
The shape and size of geographical features is correctly represented.	Due to projection there are distortions in shape and size of geographical features.
Accurate tracing of the maps is not possible due to the curvature of the globe	Maps can be accurately traced
A part of earth can't be separately represented on the globe	A part of earth can be separately represented on the map
Inconvenient to carry	Easy to carry

Cartographic Principles & Design

Cartography is the art and science of map making.
Cartographers make a huge contribution in making the maps more meaningful and understandable.

Major principles of cartography:
Data Classification - It is important to have a good understanding of the data which needs to be represented on a map. The scales of measurement used are described below:

a) Nominal Scale - It only satisfies the identity property
b) Ordinal Scale - This scale has the properties of both identity and magnitude but the intervalbetween any two values is indeterminate.
c) Interval Scale - This scale has the properties of identity, magnitude and equal intervals
d) Ratio Scale - The ratio scale has the properties of identity, magnitude, equal interval andabsolute zero

Map Elements

Title (Subject or theme of the map)

Mapping Feature (The area to be represented in a map)

Scale

North Arrow

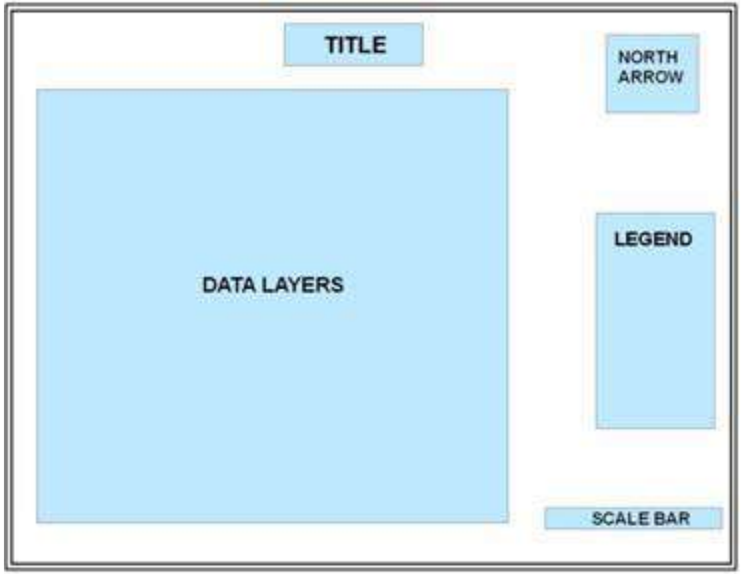
Legend (The colours and symbols used in the map to specify the objects)

Labels (Name, count, etc.)

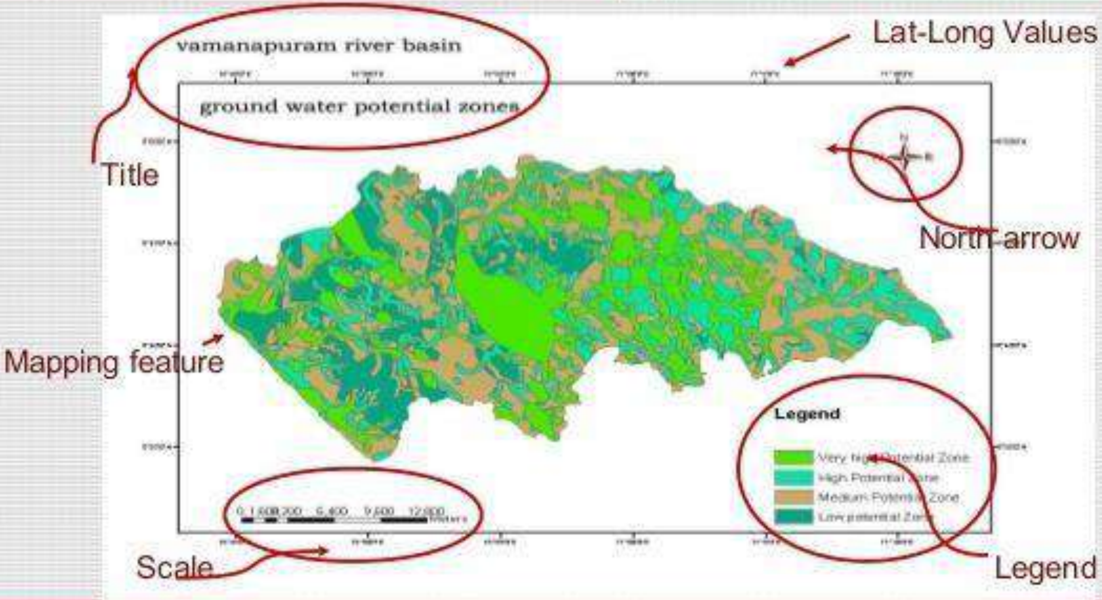
Map Layout

The process of map composition starts with preparing a layout for the map. Apart from the data, a map has certain other things that make map a package of effective and clear communication. These provide critical information to users and are known as map elements.

A layout specifies the space and positions for different map elements such as neat lines, title, North arrow, scale bar, legend etc.



Map Elements

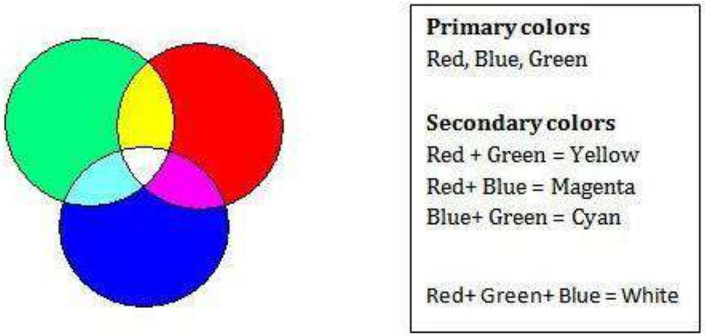


Map Characteristics

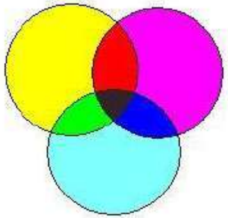
- Map Scale
- Map Accuracy
- Map Resolution
- Map Extent
- Database Extent

Map Elements

1. Color - The aim of filling colors in a map is to make visual distinction among various features thus making map more decipherable. Generally, dark colors mean more and represent high values of the attribute under study
    a) Additive Color Model
    b) Subtractive Color Model - A subtractive color model explains the mixing of dyes, inks etc. to create a range of colors.
2. Text
3. Placement of Labels
4. Symbols
    a) Shape
    b) Size
    c) Combination of Symbol and Color

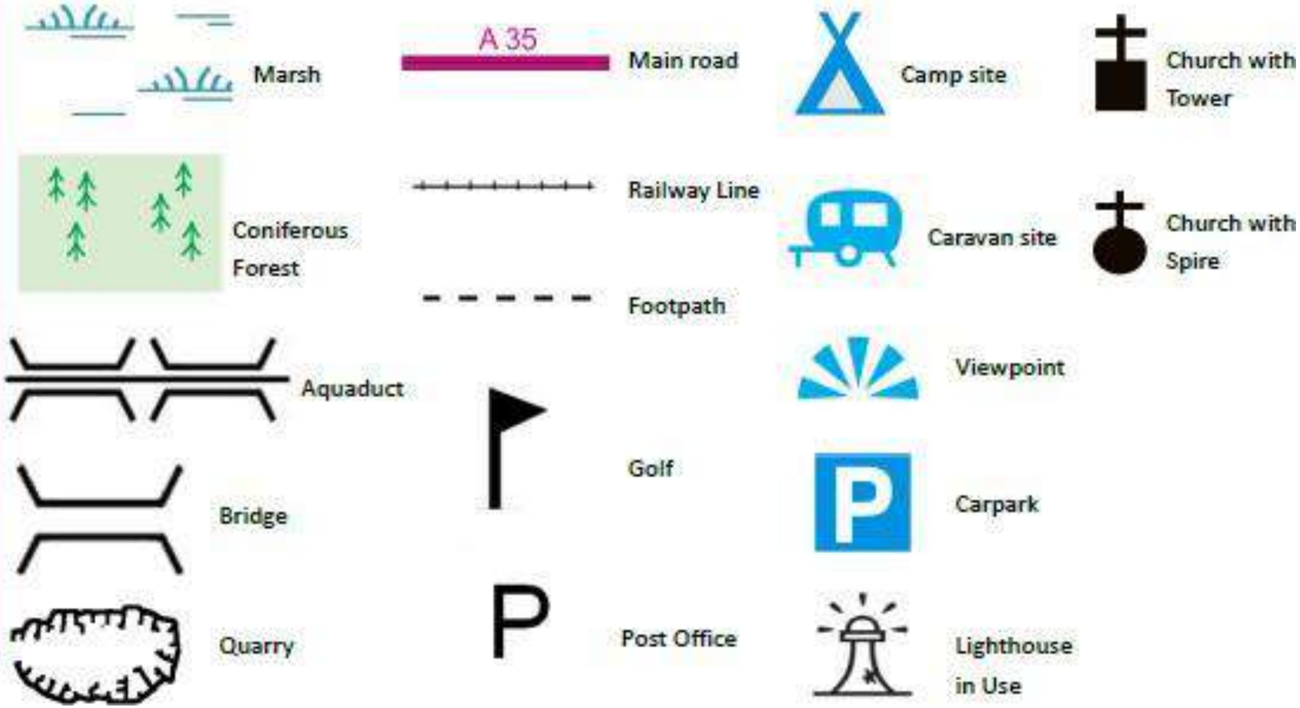


Additive color mixing

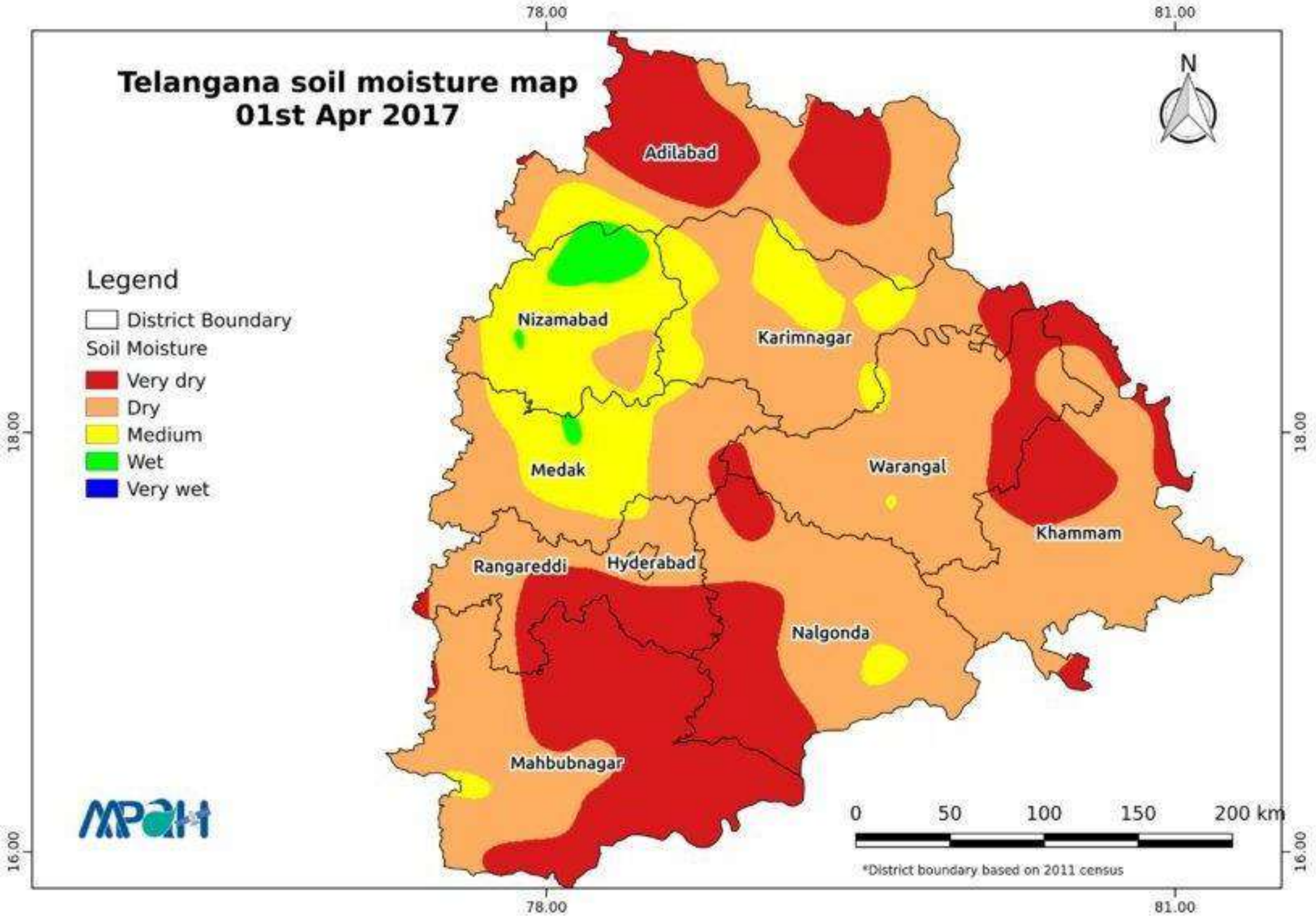


Subtractive color mixing

LEGEND



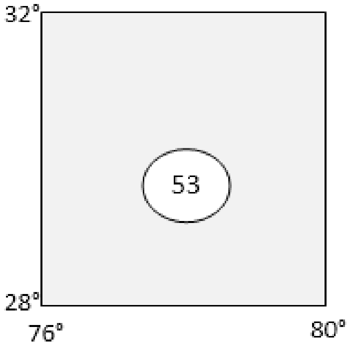
Marsh
Coniferous Forest
Aquaduct
Bridge
Quarry
A 35
Main road
Railway Line
Footpath
Golf
Post Office
Camp site
Caravan site
Viewpoint
Carpark
Lighthouse in Use
Church with Tower
Church with Spire



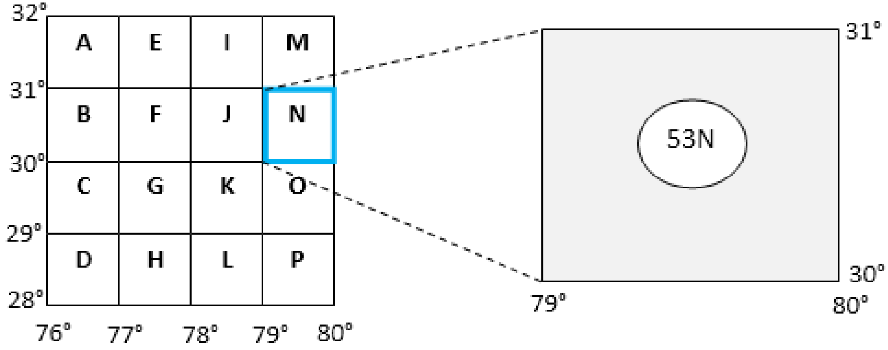
Telangana soil moisture map
01st Apr 2017
Legend
District Boundary
Soil Moisture
Very dry
Dry
Medium
Wet
Very wet
Adilabad
Nizamabad
Karimnagar
Medak
Warangal
Khammam
Rangareddi
Hyderabad
Nalgonda
Mahbubnagar
N
0
50
100
150
200 km
*District boundary based on 2011 census
78.00
81.00
18.00
16.00

Toposheet Indexing

Survey of India produces the topographic maps of India. These maps are produced at different scales. In order to identify a map of a particular area, a numbering system has been adopted by the Survey of India.

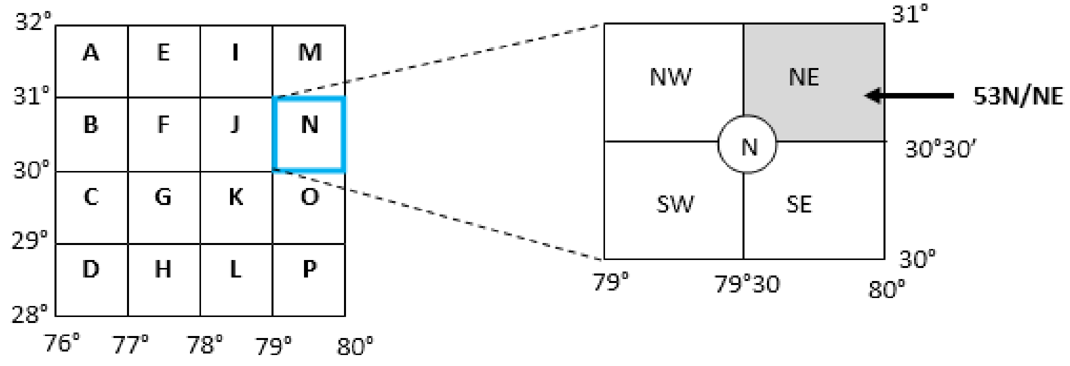


A sheet of 4° × 4° (scale: 1: 1,000,000)



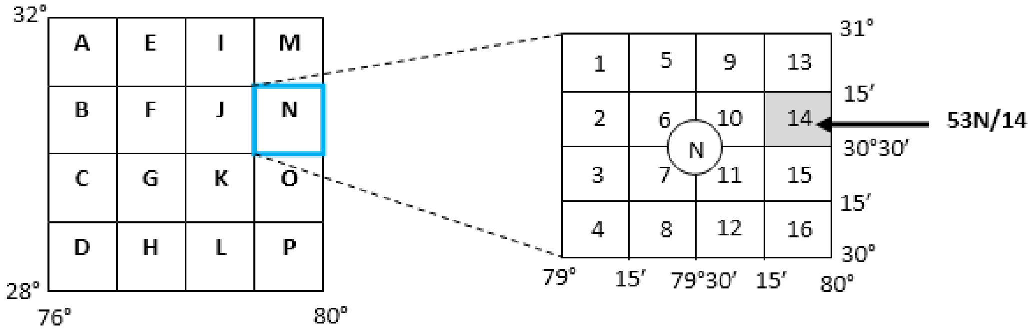
A sheet of 4° × 4° (scale: 1: 1,000,000)

A sheet of 1° × 1° (scale: 1: 250, 000)



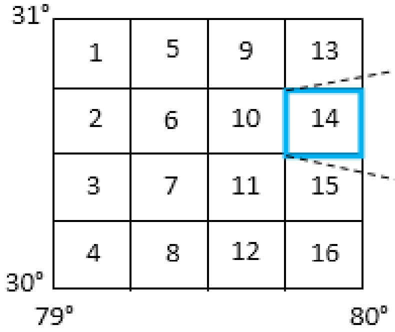
A sheet of 4° × 4° (scale: 1: 1,000,000)

A sheet of 30′ × 30′ (scale: 1:100,000)

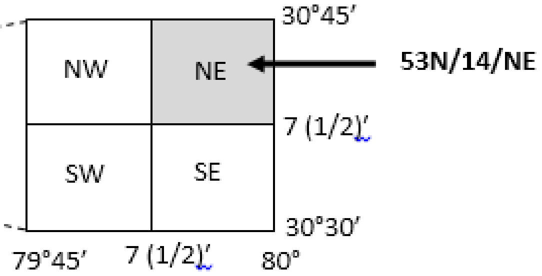


A sheet of 4° × 4° (scale: 1: 1,000,000)

A sheet of 15′ × 15′ (scale: 1: 50,000)



A sheet of 1° × 1° (scale: 1: 250, 000)



A sheet of 7(1/2)′ × 7(1/2)′ (Scale: 1: 25,000)

Toposheets

- It depicts the elevation of an area as contours.
- These are large scale maps.
- A region is divided into number of smaller sections for detailed study.
- May form a part of a district or more than two.
- The official agency preparing maps in India is The Survey Of India.
- Because of its national importance the topographic maps of strategic areas are restricted for sale.
- Some of the scales used in the topographic sheets are 1:250,000, 1:50,000, 1:25,000

Limitations of Paper based Maps

1. Time consuming: They require complex interpretations when reading them.
2. Good quality paper maps may be hard to find.
3. Map printing errors
4. Paper maps can easily be damaged
5. Paper maps are limited: Paper maps only show limited areas on a single map and you willneed several maps if you are visiting several destinations.
6. It is archaic: Paper maps are an old way of representing an area on a map and may be difficult to understand
7. Paper maps are not complete: Paper maps are never complete and one may never get acomplete paper map of an area. This is because features and landscapes keep on changing.
8. Paper maps are biased: Paper maps are designed and printed by people whose perceptionof an area may be subject to personal bias.

Limitations of Paper based Maps

9. Paper maps show limited features:

Paper maps show limited features of a given area on the paper because they only concentrate on one aspect of the landscape.

10. Paper maps are complicated to understand:

Paper maps are difficult to understand because most of the features are represented in symbols.

11. Paper maps only use symbols:

Paper maps rely solely on symbols to represent important features on the map andthis may be a problem during interpretation.

12. Difficult to show elevation:

It is difficult to show elevation on a paper map because of the limited representation features.

Digital Maps

Digital mapping (also called digital cartography) is the process by which a collection of data is compiled and formatted into a virtual image.

The primary function of this technology is to produce maps that give accurate representations of a particular area, detailing major road arteries and other points ofinterest.

Digital Cartography involves the creation and analyses of maps through the use of computers. As such, Geographical Information Systems (GIS) are an integral tool for Cartography.

A geographic information system (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data.

Difference between Digital Map and a Traditional Map

- Both GIS and traditional maps represent the real world. One way of thinking about it is that the GIS is the data while the map is the picture.

- The end result of GIS analysis is often a map, or cartographic output, but itis only one of many output types.

- Once created, a traditional map cannot be changed. However, as an integrated system, a GIS allows users to ask new questions of a databaseand visualize answers.

Advantages of Digital Maps

1. They can be downloaded for free: Digital maps are not sold in physical shops but instead they are available online and can be downloaded for free.
2. Storage requires digital space: Digital maps do not require physical space for storage.They are stored in digital format and therefore require digital space.
3. Shows all features including time and the actual building: Digital maps do not depend on symbols to represent features. Instead, they show the actual features and the time.
4. Can be easily updated: Digital maps are mostly real time representations of an areaand can therefore be updated easily because the changes will be updated automatically.
5. It is dynamic: Digital maps are dynamic which means one can choose to view the previous versions of the represented area unlike paper maps.

Advantages of Digital Maps

6. Digital maps can represent all features at the same time: Digital maps can be used to represent all features of a given area at the same time. They allow filteringfor specific features but can also show the entire area.
7. Digital maps are not limited to any area: Digital maps are not limited to show onlya specific area based on scale. They can be widened to show the entire area.
8. May not require special skills since it shows the real life object: Interpreting digitalmaps is easier than paper maps since the features are real representations of the real world objects.
9. Good at showing area overlays from various angles: Digital maps do not only show 2D representations of an area. They can be used to show the 3D angle of thearea and also show the area overlays.
10. Digital maps are always up to date

GIS Projections

❶ Maps are flat but they represent curved surfaces. Transforming 3-D space onto a 2-D map is called a projection.

❷ Projections are mathematical expressions which convert data from a geographic location (lat, long) on a sphere or spheroid to a representative location on a flat surface.

❸ Projection **always causes distortion** in one OR more ways: shape, area, distance, direction. Therefore, one must choose which characteristic to be accurate at the expense of the others

Properties of Map Projection

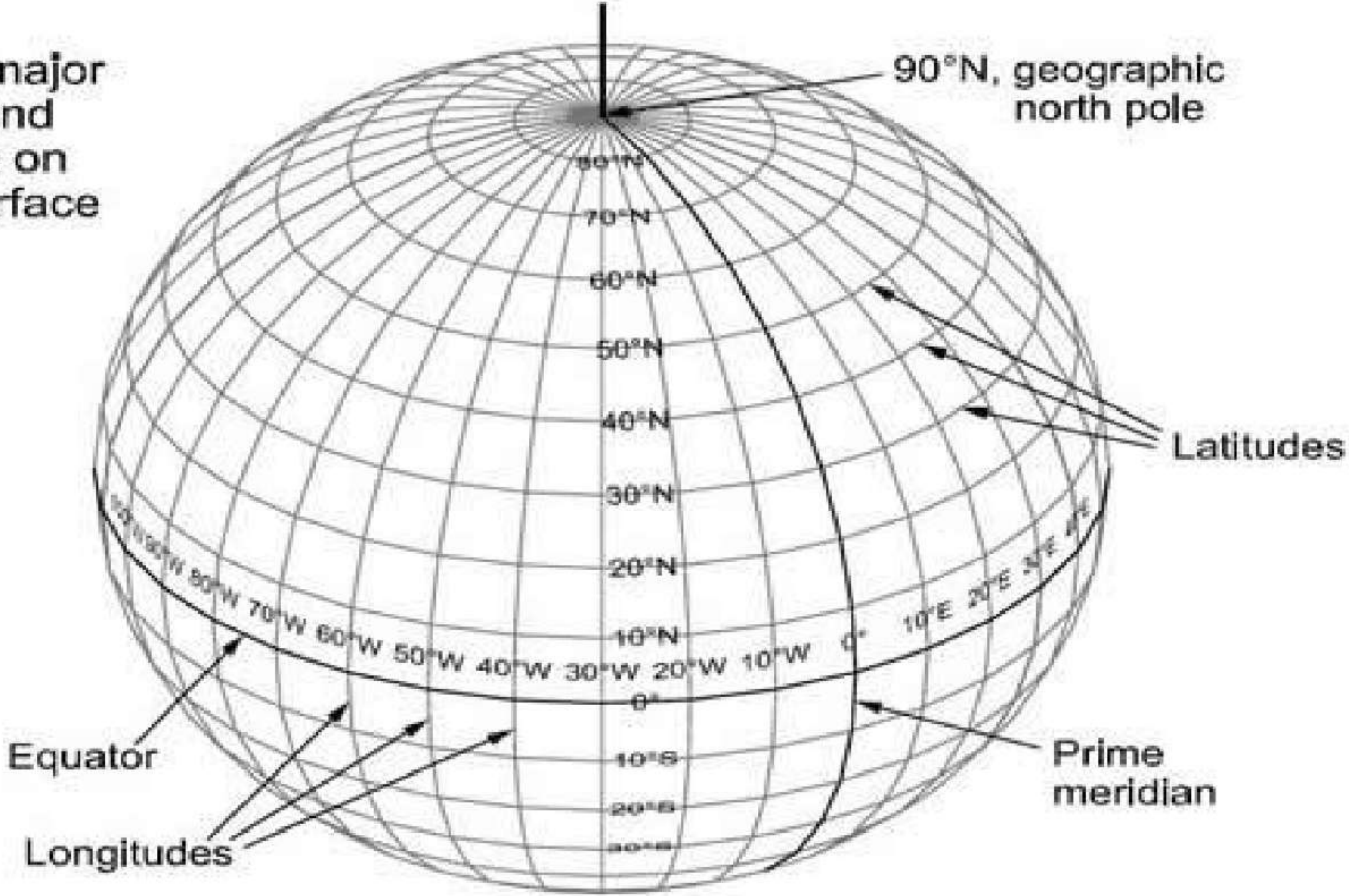
- **Conformal projections** - **Shape**
  (Preserve local shape)

- **Equal-area projections** - **Area**
  (Preserve the area of displayed features)

- **Equidistant projections** - **Distance**
- (Preserve the distance between certain points)

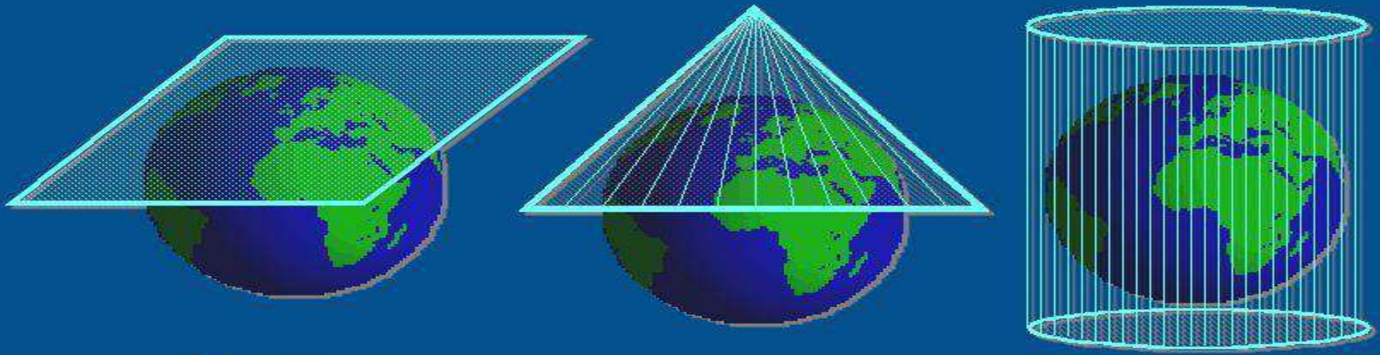
- **True-direction projections** - **Direction**
- (Maintain the directions)

Selected major latitudes and longitudes on Earth's surface



Projections

- The earth is "projected" from an imaginary light source in its center onto a surface, typically a plate, cone, or cylinder.



Projected coordinate system (PCS)

Geographic coordinate system (GCS)

Unit III. Spatial Database Management System

Introduction

Data in a GIS represent a simplified view of physical *entities*, the roads, mountains, accident locations, or other features we wish to identify. Data include information on the spatial location and extent of the entities, and information on their nonspatial properties.

Each entity is represented by a *spatial feature* or *cartographic object* in the GIS, and so there is an entity–object correspondence. Because every computer system has limits, only a subset of essential characteristics is recorded. As illustrated in Figure 2-1, we may represent landcovers in a region by a set of polygons. These polygons are associated with a set of essential characteristics that define each landcover, perhaps vegetation type, ownership, or landuse.

Essential characteristics are subjectively chosen by the spatial data developer. The essential characteristics of a forest would be different in the eyes of a logger than those of a conservation officer, a hunter, or a hiker. Objects are abstract representations of reality that we store in a spatial database, and they are imperfect representations because we can only record a subset of the characteristics of any entity. No one abstraction is universally better than any other, and the goal of the GIS developer is to define objects that support the

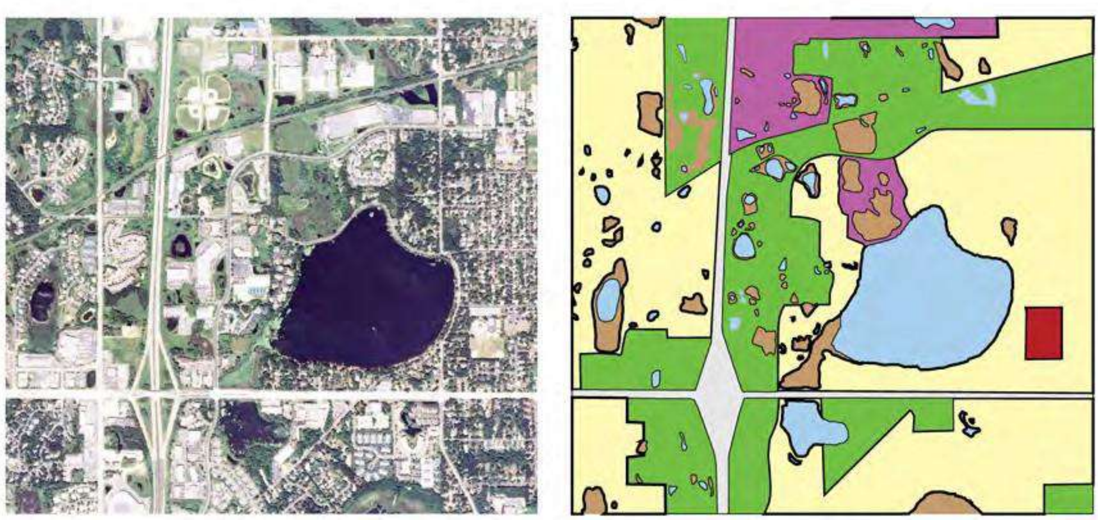


Figure 2-1: A physical entity is represented by a spatial object in a GIS. Here, lakes (dark areas) and other land cover types are represented by polygons.

intended use at the desired level of detail and accuracy.

A *spatial data model* (Figure 2-2) may be defined as the objects in a spatial database plus the relationships among them. The term "model" is fraught with ambiguity because it is used in many disciplines to describe many things. Here the purpose of a spatial data model is to provide a formal means of representing and manipulating spatially referenced information. In Figure 2-1, our data model consists of two parts. The first part is a set of polygons (closed areas) recording the edges of distinct land uses, and the second part is a set of numbers or letters associated with each polygon. The data model may be considered the most recognizable level in our computer abstraction of the real world. Data structures and binary machine code are successively less recognizable, but more computer-compatible, forms of the spatial data.

Coordinates are used to define the spatial location and extent of geographic objects. A coordinate most often consists of a pair or triplet of numbers that specify location in relation to a point of origin. The coordinates quantify the distance from the origin when measured along standard directions. Single or groups of coordinates are organized to represent the shapes and boundaries that define the objects. Coordinate information is an important part of the data model, and models differ in how they represent these coordinates. Coordinates are usually expressed in one of many standard coordinate systems. The coordinate systems are usually based upon standardized map projections (discussed in Chapter 3) that unambiguously define the coordinate values for every point in an area.

Typically, attribute data complement coordinate data to define cartographic objects (Figure 2-3). Attribute data are collected and referenced to each object. These

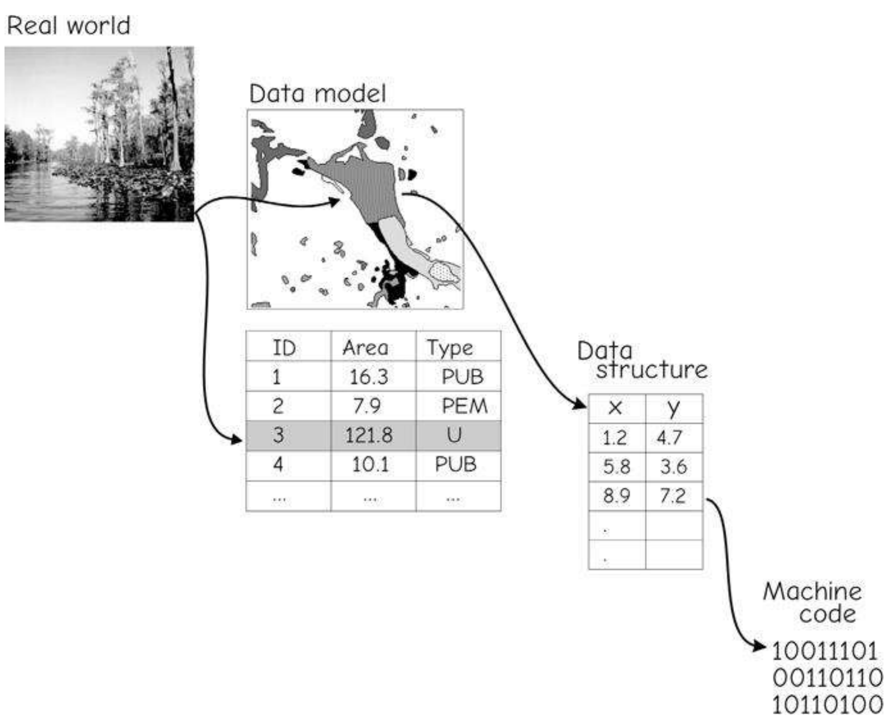


Figure 2-2: Levels of abstraction in the representation of spatial entities. The real world is represented in successively more machine-compatible but humanly obscure forms.

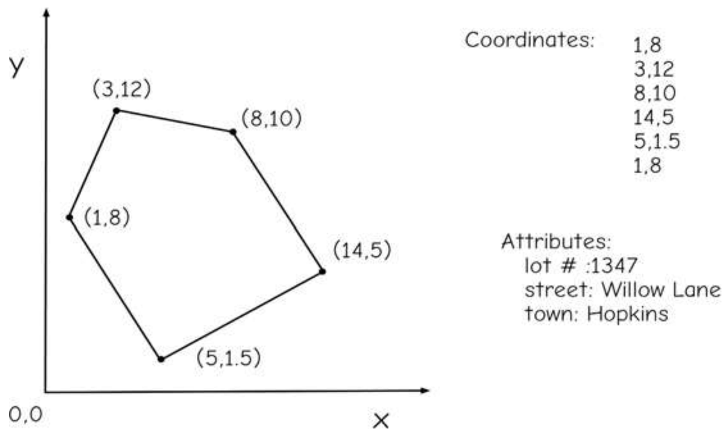


Figure 2-3: Coordinate and attribute data are used to represent entities.

attribute data record the non-spatial components of an object, such as a name, color, pH, or cash value. Keys, labels, or other indices are used so that the coordinate and attribute data may be viewed, related, and manipulated together.

Most conceptualizations view the world as a set of layers (Figure 2-4). Each layer organizes the spatial and attribute data for a given set of cartographic objects in the region of interest. These are often referred to as *thematic layers*. As an example, consider a GIS database that includes a soils data layer, a population data layer, an elevation data layer, and a roads data layer. The roads layer contains only roads data, including the location and properties of roads in the analysis area (Figure 2-4). There are no data regarding the location and properties of any other geographic entities in the roads layer. Information on soils, population, and elevation are contained in their respective data layers. Through analyses we may combine data to create a new data layer; for example, we may identify areas that have high elevation and join this information with the soils data. This combination may create a new data layer with a new, composite soils/elevation variable.

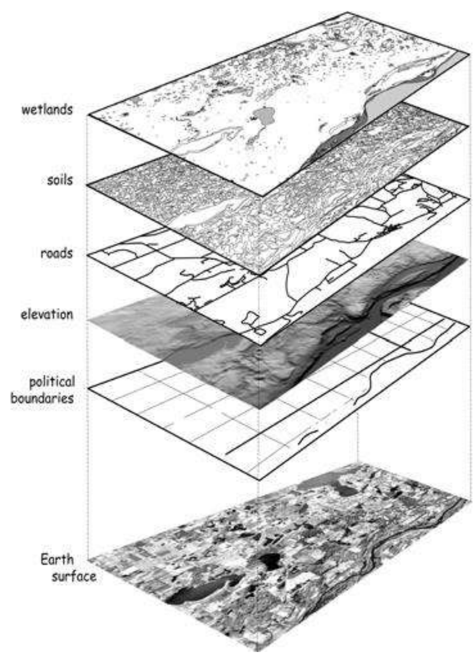


Figure 2-4: Spatial data are often stored as separate thematic layers, with objects grouped based on a set of properties, e.g., water, roads, or land cover, or some other agreed-upon set.

Attribute Data and Types

Attribute data are used to record the nonspatial characteristics of an entity. Attributes, also called *items* or *variables,* may be envisioned as a list of characteristics that describe the features we represent in a GIS. Color, depth, weight, owner, vegetation type, or land use are examples of variables that may appear as attributes. Attributes have values; for example, a fire hydrant may be colored red, yellow, or orange, have 1 to 4 flanges, and a rating of high, medium, or low.

any data layer prior to development of a layer or its use in analysis.

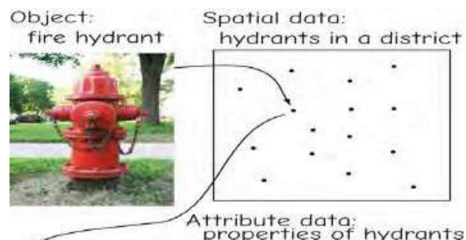
Attributes are often presented in tables and arranged in rows and columns

As noted in Chapter 1, a finite number of characteristics is selected to represent a set of features. Sometimes only one attribute is stored for each location, for example, a simple layer that contains smog concentration in the atmosphere. Attributes in a city layer may include the population, total municipal area, the name of the current mayor, and the county and state containing the city. These attributes may be sufficient for some of the intended uses of the layer, but there are many other attributes that may be used to characterized a city. The number and type of attributes stored for a feature are by definition a subset of those possible, and there is no set that is universally superior to other potential sets. The set of attributes reflects the needs of the data developer and users, and should be carefully assessed for

(Figure 2-17). Each row corresponds to a spatial object, and each column corresponds to an attribute. Tables are often organized and managed using a specialized computer pro- gram called a database management system (DBMS, described more fully in Chapter 8).

Attributes may be in many forms, but all attributes can be categorized as nominal, ordinal, or interval/ratio attributes.

Nominal attributes are variables that provide descriptive information about an object. The color is recorded for each hydrant in Figure 2-17. Other examples of nominal data are vegetation type, a city name, the owner of a parcel, or soil series. There is no implied order, size, or quantitative information contained in nominal attributes.



ID	Flanges	Color	Rating
1	3	red	high
2	3	red	high
3	2	green	low
4	3	red	mid
.	.	.	.

Figure 2-17: Attributes are typically envisioned as arranged by columns and rows, with objects arranged in rows, and attributes aligned in columns.

Nominal attributes may also be images, film clips, audio recordings, or other descriptive information, for example, GIS for real estate often have images of the buildings as part of the database. Image, video, or sound recordings stored as attributes are sometimes referred to as “BLOBs” for *binary large objects*.

Ordinal attributes imply a ranking or order by their values. An ordinal attribute may be descriptive, such as high, mid, or low, or it may be numeric; for example, an erosion class may be given a value from 1 to 10. The order reflects only rank, and not the scale. An ordinal value of four has a higher rank than a value of two, but we can’t infer that the attribute value is twice as large, because we can’t assume the scale is linear.

Interval/ratio attributes are used for numeric items where both rank order and absolute difference in magnitudes are represented, for example, the number of flanges in the second column of Figure 2-17. These data are often recorded as real numbers on a linear scale. Area, length, weight, height, or depth are a few examples of attributes that are represented by interval/ratio variables.

Items have a *domain*, a range of values they may take. Colors might be restricted to red, yellow, and green; cardinal direction to north, south, east, or west; and size to all positive real numbers.

Common Spatial Data Models

Spatial data models begin with a conceptualization. Consider a road map suitable for use at a statewide or provincial level. This map is based on a conceptualization that defines roads as lines. These lines connect cities and towns that are shown as discrete points or polygons on the map. Road properties may include only the road type, for example, an interstate highway, county road, or some other type of road. The roads have a width represented by the drawing symbol on the map, however, this width, when scaled, may not represent the true road width. This conceptualization identifies each road as a linear feature that fits into a small number of categories. All state highways are represented equally although they may vary. Some may have wide shoulders, others not, or dividing barriers of concrete, versus a broad vegetated median. We realize that differences exist within our conceptualization.

There are two main conceptualizations used for digital spatial data. The first conceptualization defines discrete objects using a *vector data model.* Vector data models use discrete elements such as points, lines, and polygons to represent the geometry of real-world entities (Figure 2-18).

Farm fields, roads, wetlands, cities, and census tracts are examples of entities that are often represented by discrete objects. Points are often used to define the locations of “small” objects such as wells, buildings, or ponds. Lines may be used to represent linear objects, for example, rivers or roads, or to identify the boundary between what is a part of the object and what is not a part of the object. We may map landcover for a region of interest, and we categorize discrete areas as a uniform landcover type. A forest may share an edge with a pasture, and this boundary is represented by lines. The boundaries between two polygons may not be discrete on the ground. For example, a forest edge may grade into a mix of trees and grass, then to pasture; however in the vector conceptualization, a line between two landcover types will be drawn to indicate a discrete, abrupt transition. Lines and points have coordinate locations, but points have no dimension, and lines have no dimension perpendicular to their direction. Area features may be defined by a closed, connected set of lines.

The second common conceptualization identifies and represents grid cells for a given region of interest. This conceptualization employs a *raster* data model (Figure 2-18). Raster cells are arrayed in a row and column pattern to provide “wall-to-wall” coverage of a study region. Cell values are used to represent the type or quality of mapped variables. The raster model is used most commonly with variables that may change continuously across a region. Elevation, mean temperature, slope, average rainfall, cumulative ozone exposure, or soil moisture are examples of phenomena that are often represented as continuous fields. Raster representations are also sometimes used to represent discrete features, for example, class maps of vegetation or political units.

Data models are at times interchangeable in that many phenomena may be represented with either the vector or raster approach. For example, elevation may be represented as a surface (continuous field) or as a series of lines representing contours of equal elevation (discrete objects). Data may be converted from one conceptual view to another; for example, the location of contour lines (lines of equal elevation) may be determined by evaluating the raster surface, or a raster data layer may be derived from a set of contour lines. These conversions entail some costs both computationally and perhaps in data accuracy.

The decision to use either a raster or vector conceptualization often depends on the most frequent operations performed. Slope is more easily determined when elevation is represented in a raster data set. However, discrete contours are often the preferred format for printed maps, so the discrete conceptualization of a vector data model may be preferred for this application. The best data model for a given application depends on the most common operations, the experiences and views of the GIS users, the form of available data, and the influence of the data model on data quality.

Other, less common data models are sometimes used. A triangulated irregular network (TIN) is an example of such a data model, employed to represent surfaces, such as elevations, through a combination of point, line, and area features. Variants or other representations related to raster data models also exist. We will introduce and discuss less common data models later in this and other chapters.

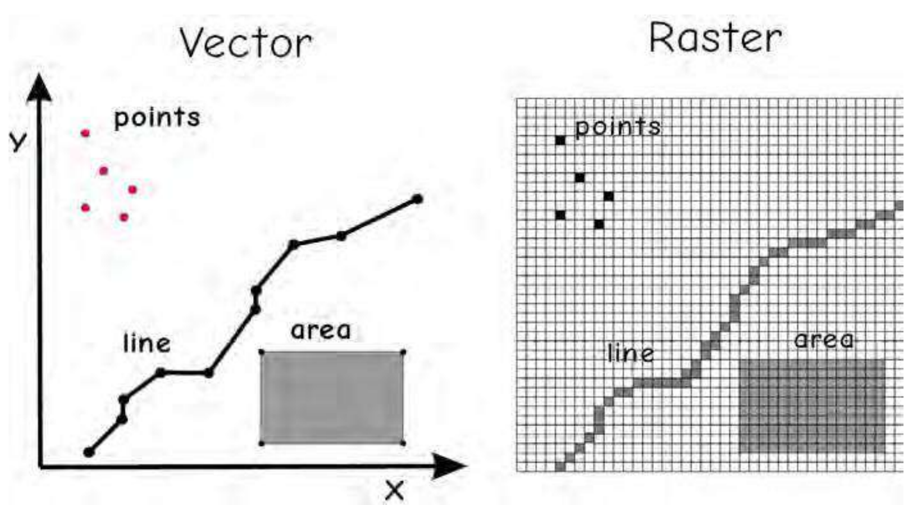


Figure 2-18: Vector and raster data models.

Vector Data Models

A vector data model uses sets of coordinates and associated attribute data to define discrete objects. Groups of coordinates define the location and boundaries of discrete objects, and these coordinate data plus their associated attributes are used to create vector objects representing the real-world entities (Figure 2-19).

There are three basic types of vector objects: points, lines, and polygons (Figure 2-20). A point uses a single coordinate pair to represent the location of an entity that is considered to have no dimension. Gas wells, light poles, accident location, and survey points are examples of entities often represented as point objects. Some of these have real physical dimension, but for the purposes of the GIS users they may be represented as points. In effect, this means the size or dimension of the entity is not important, only its location.

Attribute data are attached to each point, and these attribute data record the important nonspatial characteristics of the point entities. When using a point to represent a light pole, important attribute information might be the height of the pole, the type of light and power source, and the last date the pole was serviced.

Linear features, often referred to as lines or *arcs*, are represented as lines when using vector data models. Lines are most often represented as an ordered set of coordinate pairs. Each line is made up of line segments that run between adjacent coordinates in the ordered set (Figure 2-20). A long, straight line may be represented by two coordinate pairs, one at the start and one at the end of the line. Curved linear entities are most often represented as a collection of short, straight, line segments, although curved lines are at times represented by a mathematical equation describing a geometric shape. Lines typically have a starting point, an ending point, and intermediate points to represent the shape of the linear entity. Starting points and ending points for a line are sometimes referred to as *nodes*, while intermediate points in a line are referred to as *vertices* (Figure 2-20). Attributes may be attached to the whole line, line segments, or to nodes and vertices along the lines.

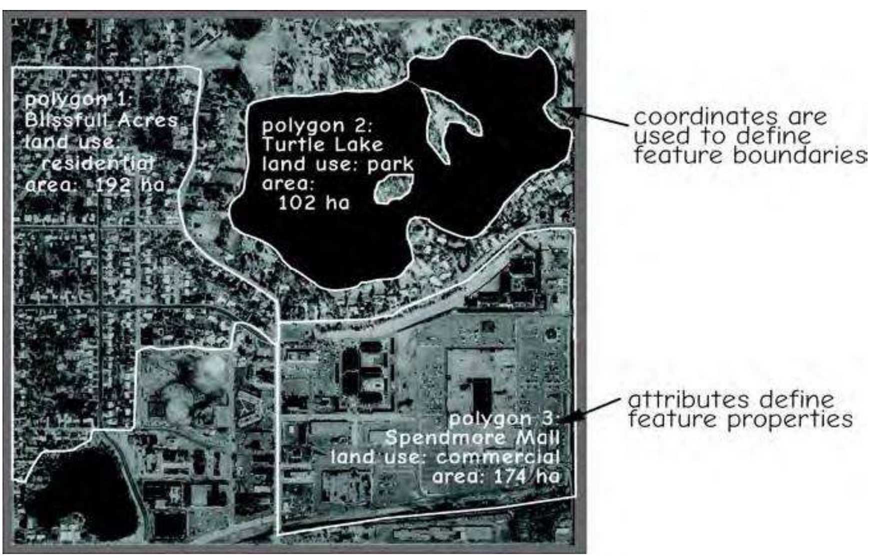


Figure 2-19: Coordinates define spatial location and shape. Attributes record the important non-spatial characteristics of features in a vector data model.

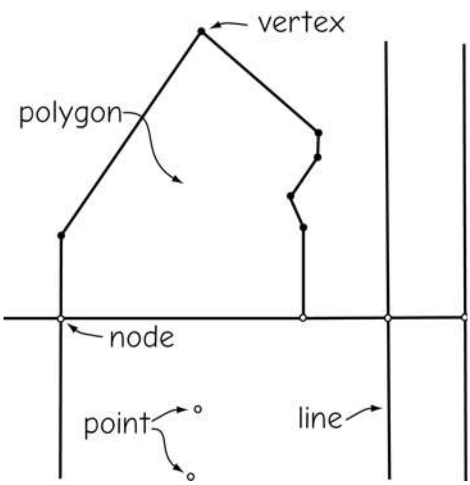


Figure 2-20: Points, nodes, and vertices define points, line, and polygon features in a vector data model.

Area entities are most often represented by closed polygons. These polygons are formed by a set of connected lines, either one line with an ending point that connects back to the starting point, or as a set of lines connected start-to-end. Polygons have an interior region and may entirely enclose other polygons in this region. Polygons may be adjacent to other polygons and thus share "bordering" or "edge" lines with other polygons. Attribute data such as area, perimeter, landcover type, or county name may be linked to each polygon.

Note that there is no uniformly superior way to represent features. Some feature types may appear to be more "naturally" represented one way: manhole covers as points, roads as lines, and parks as polygons. However, in a very detailed data set, the manhole covers may be represented as circles, and both edges of the roads may be drawn and the roads represented as polygons. The representation depends as much on the detail, accuracy, and intended use of the data set as our common conception or general shape of the objects.

A single set of features may be represented differently, depending on the interests and purposes of the GIS users (Figure 2-21). A point layer may be chosen when general feature position is needed, for example, general building location (Figure 2-21, top). Other users may be interested in the outline of the feature and so require representation by lines, while polygon representations may

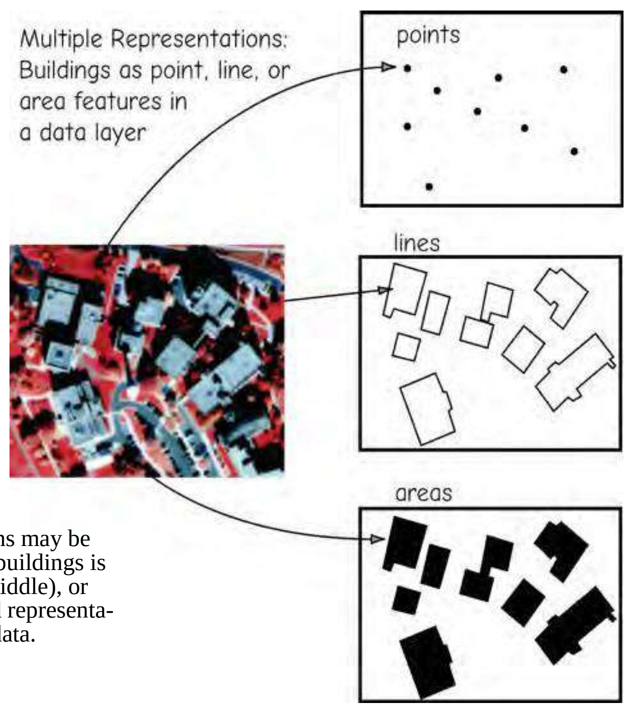


Figure 2-21: Alternate conceptualizations may be used to represent features. Here a set of buildings is represented by either point (top), line (middle), or polygon features (bottom). The preferred representation depends on the intended use of the data.

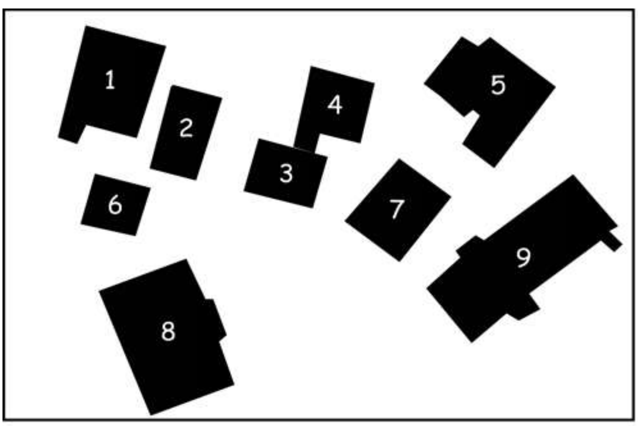
be preferred for other applications (Figure 2-21 mid and lower, respectively). Our intended use often determines our conceptual model and hence vector type used to represent a feature.

In the most common vector models, there is an attribute table associated with each vector layer, and a single row in the table corresponding to each feature in the data layer (Figure 2-22).This table holds the attributes of interest for each feature.

Each column holds an attribute value for a given feature. All values in a column have the same type, so for any given column, all entries might be ordinal, or interval ratio, or a BLOB, or some other defined type. An identifier value, or ID, is typically included, and this value is often unique within the table, with an unrepeated value assigned for each row and corresponding feature (Figure 2-22).

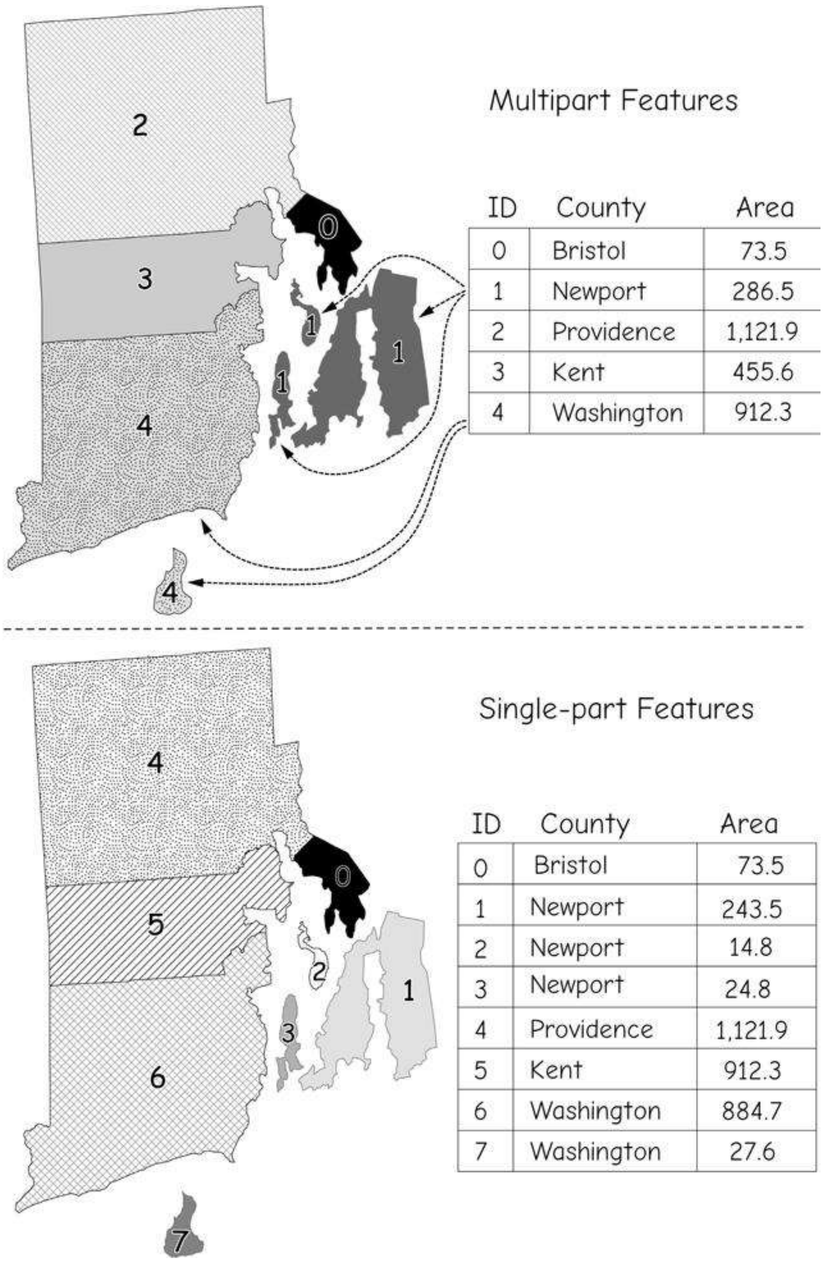
Vector layers sometimes have a "many to one" relationship between geographic features and table rows (Figure 2-23). In these instances, many spatially distinct features are matched with a row, and the row attributes apply to all the distinct features. This is common when representing islands, groups of buildings, or other clusters of features that make up a perceived "whole thing." These are sometimes referred to as *multipart features*, because multiple geographic objects may correspond to one row. .

Multipart features may also be used when there are extremely large volumes of data, for example, when millions of point



ID	Building Name	Floors	Roof Type
1	Hodson Hall	6.0	flat, sealed tar
2	Borlaug Hall	5.5	pitched 9/12, tile
3	Guilford Technology Bldg.	4.0	flat, gasket
4	Shop Annex	2.5	flat, sealed tar
5	Animal Sciences Bldg.	1.0	pitched 12/12, tile
6	Administration Bldg.	14.0	pitched 6/12, metal
7	Climate Sciences Center	6.0	flat, sealed tar
8	Grantham Tower	1.0	pitched, 9/12, tile
9	Biological Sciences Bldg.	9.0	pitched 12/12, tile

Figure 2-22: An example of the most common vector data model structure. Geographic features correspond to rows in a table with an identifier (ID) and a set of attributes arrayed in columns.



Multipart Features

ID	County	Area
0	Bristol	73.5
1	Newport	286.5
2	Providence	1,121.9
3	Kent	455.6
4	Washington	912.3

Single-part Features

ID	County	Area
0	Bristol	73.5
1	Newport	243.5
2	Newport	14.8
3	Newport	24.8
4	Providence	1,121.9
5	Kent	912.3
6	Washington	884.7
7	Washington	27.6

Figure 2-23: Example of multipart and single-part features. Here, counties for Rhode Island, a state in the Eastern U.S.A., are shown with one table entry for each county (top), with multipart features in a layer, and with one table entry for each distinct polygon (bottom), with only single-part features in the layer. Note that calculations, analysis, and interpretation may differ for multipart v.s. single-part features.

observations are collected automatically with laser scanners or similar devices. Tables are often slower to process than point features, and so reducing the table size by grouping into multi-part features may shorten many operations.

Care is warranted when converting multipart features to single-part features. The most common problems arise for aggregate variables in polygon layers, such as total counts. For example, population data are often delivered by census areas such as states. Many states, such as Hawaii, have several parts and are represented by a multipart shape. The population is associated with the aggregated set of polygons corresponding to a state (Figure 2-24). When converted to single-part shapes, the attributes are often copied for each component polygon. In our example all single-part polygons will be assigned the attribute values for the multipart feature, in effect repeating counts for each part. Each polygon will have the population count associated with it, so any aggregation or calculation based on population will likely be in error, sometimes substantially so.

Attributes for converted shapes may be corrected. If component data are available, they can be assigned to each of the single-part features. If not, then some weighting scheme may be available, for example, if there is a correlation between area and count. Until they are reviewed and appropriately adjusted, single-part attributes derived from multipart features should be used with caution.

Polygon Inclusions and Boundary Generalization

Vector data frequently exhibit two characteristics, polygon inclusions and boundary generalization. These characteristics are oft-ignored, but may affect the use of vector data, occasionally with dire consequences. These concepts must be understood, their presence evaluated, and effects weighed in the use of vector data sets.

Polygon inclusions are areas in a polygon that are different from the rest of the polygon, but still part of the polygon. Inclusions occur because we typically assume an area represented by a polygon is homogeneous, but this assumption may be wrong, as illustrated in Figure 2-25. The figure shows a vector polygon layer representing raised landscaping beds (a). The general attributes for the polygon may be coded; for example, the surface type may be recorded as cedar mulch. The area noted in Figure 2-25b shows a walkway that is an inclusion in a raised bed. This walkway has a concrete sur-

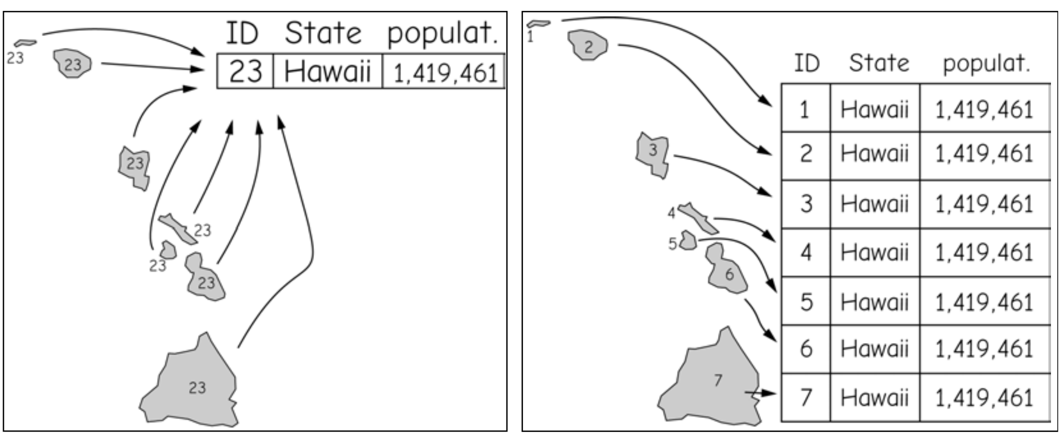


Figure 2-24: Multipart to single-part conversion may lead to errors in subsequent analysis because attributes may be copied from the original, multi-part cluster (left, above), to each single-part component (above right). Density, sums, or other derived variables often should be re-calculated for single-part features, but often are not, resulting in errors.

face. Hence, this walkway is an unresolved inclusion within the polygon.

One solution creates a polygon for each inclusion. This often is not done because it may take too much effort to identify and collect the boundary location of each inclusion, and there typically is some lower limit, or *minimum mapping unit*, on the size of objects we care to record in our data. Inclusions are present in some form in many polygon data layers.

Boundary generalization is the incomplete representation of boundary locations. This problem stems from the typical way we represent linear and area features in vector data sets. As shown in Figure 2-25c, polygon boundaries are represented as a set of connected, short, straight-line segments. The segments are a means to trace the position of line features, or the boundaries separating area features. For curved lines, these straight line segments may be viewed as a sampling of the true curve, and there is typically some deviation of the line segment from the "true" curved boundary. The amount of generalization depends on many factors, and should be so small as to be unimportant for any intended use of the spatial data. However, since many data sets may have unforeseen uses, or may be obtained from a third party, the boundary generalization should be recognized and evaluated relative to the specific requirements of any given spatial analysis. There are additional forms of generalization in spatial data, and these are described more thoroughly in Chapter 4.

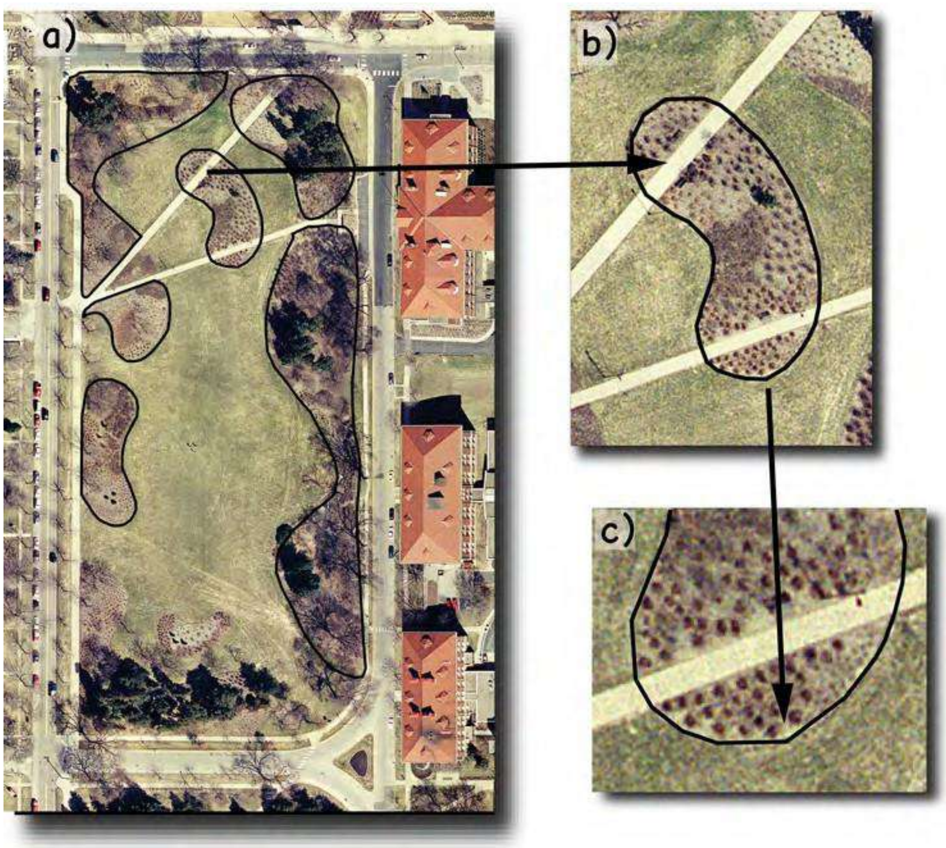


Figure 2-25: Examples of polygon inclusions (sidewalk inclusion in flower bed shown in a and b), and boundary generalization (c) in a vector data model. These approximations typically occur as a consequence of adopting a vector representation, and their impacts must be considered when using vector data.

Vector Topology

Vector data often contain *vector topology*, enforcing strict connectivity and recording adjacency and planarity. Early systems employed a spaghetti data model (Figure 2-26a), in which lines may not intersect when they should, and may overlap without connecting. The spaghetti model severely limits spatial data analysis and is little used except for very basic data entry or translation. Topological models create an intersection and place a node at each line crossing, record connectivity and adjacency, and maintain information on the relationships between and among points, lines, and polygons in spatial data. This greatly improves the speed, accuracy, and utility of many spatial data operations.

Topological properties are conserved when converting vector data among common coordinate systems, a common practice in GIS analysis (described in Chapter 3). Polygon adjacency is an example of a topologically invariant property, because the list of neighbors for any given polygon does not change during geometric stretching or bending (Figure 2-26, b and c). These relationships may be recorded separately from the coordinate data.

Topological vector models may vary, and enforce particular types of topological relationships. *Planar topology* requires that all features occur on a two-dimensional surface. There can be no overlaps among lines or polygons in the same layer (Figure 2-27). When planar topology is enforced, lines may not cross over or under other lines. At each line crossing there must be an intersection.

The left side of Figure 2-27 shows nonplanar graphs. In the top left figure, four line segments coincide. At some locations the lines intersect at a node, shown as white-filled circles, but at some locations a line passes over or under another line segment. These lines are nonplanar. The top right of Figure 2-27 shows planar topology enforced for these same four line segments. Nodes are found at each line crossing.

Polygons can also be nonplanar, as shown at the bottom left of Figure 2-27. Two polygons overlap slightly at an edge. This may be due to an error; for example, the two polygons share a boundary but have been recorded with an overlap, or there may be two areas that overlap in some way. If topological planarity is enforced, these two polygons must be resolved into three separate, nonoverlapping polygons. Nodes are placed

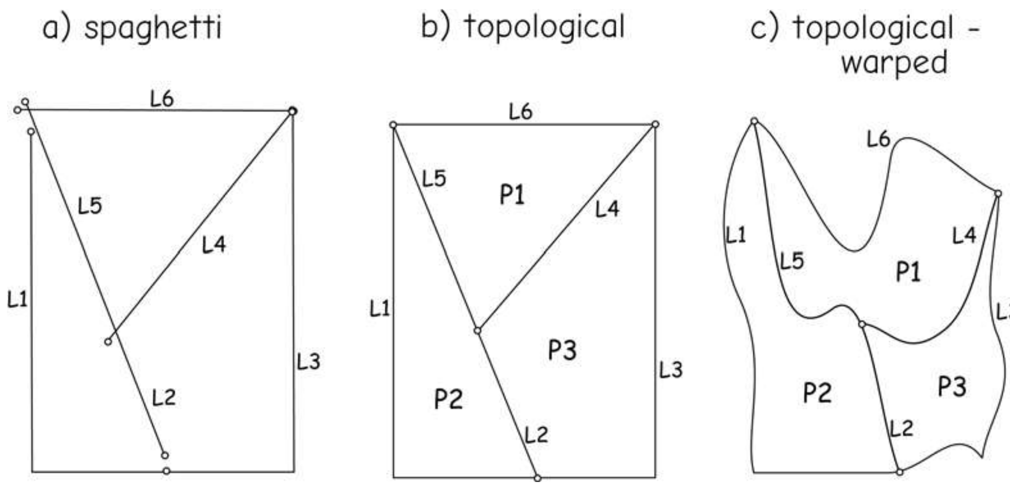


Figure 2-26: Spaghetti (a), topological (b), and topological warped (c) vector data. Figures b and c are topologically identical because they have the same connectivity and adjacency.

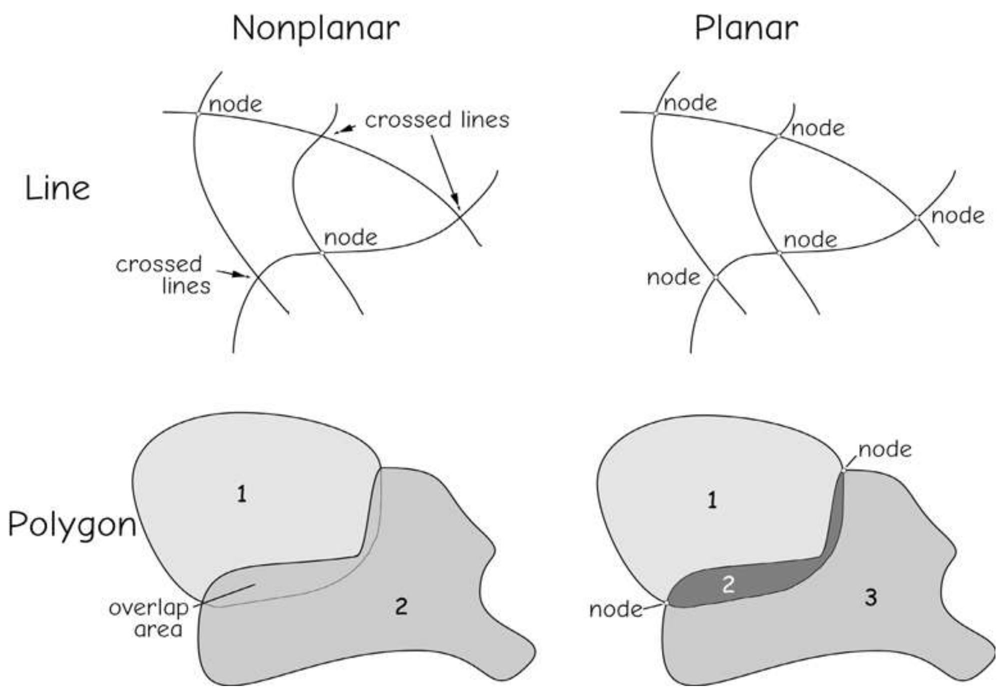


Figure 2-27: Nonplanar and planar topology in lines and polygons.

at the intersections of the polygon boundaries (lower right, Figure 2-27).

There are additional topological constructs besides planarity that may be specified. For example, polygons may be exhaustive, in that there are no gaps, holes, or "islands" allowed. Line direction may be recorded, so that a "from" and "to" node are identified in each line. Directionality aids the representation of river or street networks, where there may be a natural flow direction.

There is no single, uniform set of topological relationships that are included in all topological data models. Different vendors have incorporated different topological information in their data structures. Planar topology is often included, as are representations of *adjacency* (which polygons are next to which) and *connectivity* (which lines connect to which).

Some GIS software create and maintain detailed topological relationships in their data. This results in more complex and perhaps larger data structures, but access is often faster, and topology provides more consistent, "cleaner" data. Other systems maintain little topological information in the data structures, but compute and act upon topology as needed during specific processing.

Topology may also be specified between layers, because we may wish to enforce spatial relationships between entities that are stored separately. As an example, consider a data layer that stores property lines (cadastral data), and a housing data layer that stores building footprints (Figure 2-28). Rules may be specified that prevent polygons in the housing data layer from crossing property lines in the cadastral data layer. This would indicate a building that crosses a property line. Most such instances occur as a result of small errors in data entry or misalignment among data layers. Topological restrictions between two data layers avoid these inconsistencies. Exceptions may be

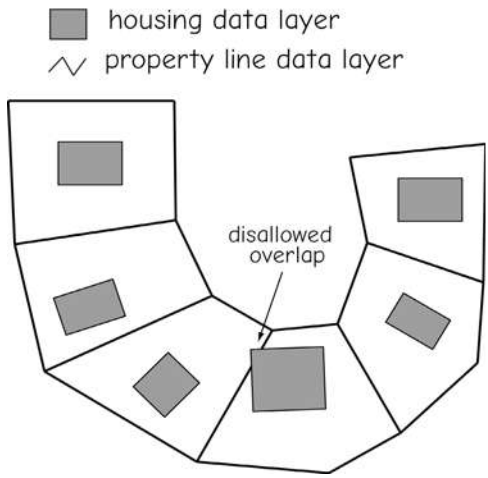


Figure 2-28: Topological rules may be enforced across data layers. Here, rules may be specified to avoid overlap between objects in different layers.

granted in those few cases when a building truly does cross property lines.

There are many other types of topological constraints that may be enforced, both within and between layers. *Dangles*, lines that do not connect to other lines, may be proscribed, or limited to be greater or less than some threshold length. Lines and points may be required to coincide, for example, water pumps as points in one data layer and water pipes as lines in another, or lines in separate layers may be required to intersect or be coincident. While these topological rules add complexity to vector data sets, they may also improve the logical consistency and value of these data.

Topological vector models often use codes and tables to record topology. As described above, nodes are the starting and ending points of lines. Each node and line is given a unique identifier. Sequences of nodes and lines are recorded as a list of identifiers, and point, line, and polygon topology recorded in a set of tables. The vector features and tables in Figure 2-29 illustrate one form of this topological coding.

Many GIS software systems are written such that the topological coding is not visible to users, nor directly accessible by them. Tools are provided to ensure the topology is created and maintained, that is, there may be directives that require that polygons in two layers do not overlap, or to ensure planarity for all line crossings. However, the topological tables these commands build are often quite large, complex, and linked in an obscure way, and therefore hidden from users.

Point topology is often quite simple. Points are typically independent of each other, so they may be recorded as individual identifiers, perhaps with coordinates included, and in no particular order (Figure 2-29, top).

Line topology typically includes substantial structure and identifies at a minimum the beginning and ending points of each line (Figure 2-29, middle). Variables record the topology and may be organized in a table. These variables may include a line identifier, the starting node, and the ending node for each line. In addition, lines may be assigned a direction, and the polygons to the left and right of the lines recorded. In most cases left and right are defined in relation to the direction of travel from the starting node to the ending node.

Polygon topology may also be defined by tables (Figure 2-29, bottom). The tables may record the polygon identifiers and the list of connected lines that define the polygon. Edge lines are often recorded in sequential order. The lines for a polygon form a closed loop and thus, the starting node of the first line in the list also serves as the ending node for the last line in the list. Note that there may be a “background” polygon defined by the outside area. This background polygon is not a closed polygon like all the rest; however it may be defined for consistency and to provide entries in the topology tables.

Topological vector models greatly enhance many vector data operations. Adjacency analyses are reduced to a “table lookup”, a quick and easy operation in most software systems. For example, an analyst may want to identify all polygons adjacent to a city. Assume the city is represented as a single polygon. Adjacency analysis reduces to 1) scanning the polygon topology table to find the polygon labeled "city" and reading the list of lines that bound the polygon, and 2) scanning this list of lines for the city polygon, accumulating a list of all left and right polygons. Polygons adjacent to the city may be identified from this list. List searches on topological tables are typically much faster than searches involving coordinate data.

Topological vector models also enhance many other spatial data operations. Network and other connectivity analyses are concerned with the flow of resources through defined pathways. Topological vector models explicitly record the connections of a set of pathways and so facilitate network analyses. Overlay operations are also enhanced when using topological vector models. The mechanics of overlay operations are discussed in greater detail in Chapter 9; however, we will state here that they involve identifying line adjacency, intersection, and resultant polygon formation. The interior and exterior regions of existing and new polygons must be determined, and these regions depend on polygon topology. Hence, topological data are useful in many spatial analyses.

Topological data models often have an advantage of smaller file sizes, largely because coordinate data are recorded once. For example, a nontopological approach often stores polygon boundaries twice. Lines 52 and 53 at the bottom of Figure 2-29 will be recorded for both polygon A and polygon B. Long, complex boundaries in polygon data sets may double their size. This increases both storage requirements and processing.

There are limitations and disadvantages to topological vector models. First, there are

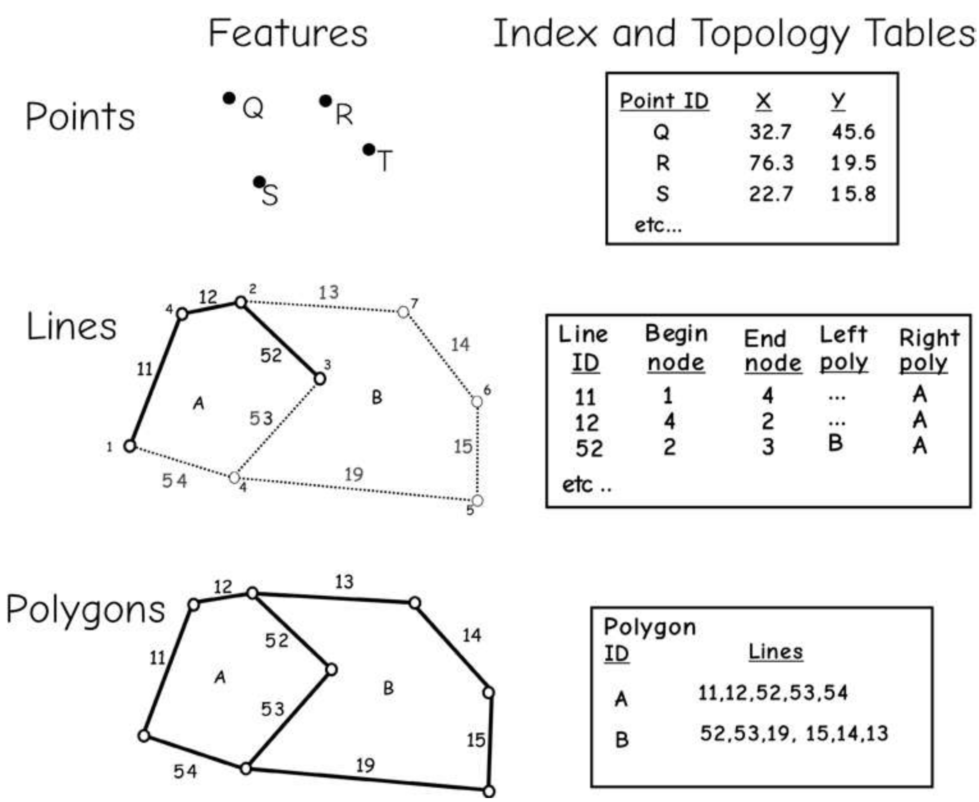


Figure 2-29: An example of vector features and corresponding topology tables. Information on the adjacency, connectivity, and other spatial relationships may be stored in topology tables, and joined to features by indices, here represented by values in the ID columns.

computational costs in defining the topological structure of a vector data layer. Software must determine the connectivity and adjacency information, assign codes, and build the topological tables. Computational costs are typically quite modest with current computer technologies.

Second, the data must be very "clean", in that all lines must begin and end with a node, all lines must connect correctly, and all polygons must be closed. Unconnected lines or unclosed polygons will cause errors during analyses. Significant human effort may be required to ensure clean vector data because each line and polygon must be checked. Software may help by flagging or fixing "dangling" nodes that do not connect to other nodes, and by automatically identifying all polygons. Each dangling node and polygon may then be checked, and edited as needed to correct errors.

Limitations and the extra editing are far outweighed by the gains in efficiency and analytical capabilities provided by topological vector models. Many current vector GIS packages use topological vector models in some form.

Vector Features, Tables, and Structures

Topological vector models are commonly used to define spatial features in a data layer. As we described earlier in this chapter, geographic features are associated with nonspatial attributes. Typically a table is used to organize the attributes, and there is a linkage between rows in the attribute table and the spatial components of a data layer (Figure 2-30, top). In most GIS software, we can most easily view the tables and a graphic representation of the spatial data as a linked table and digital map (Figure 2-30, top).

There is commonly a *one-to-one linkage* between each entry in the attribute table and each feature in the data layer. This means for each feature in the data layer there is one and only one entry in the attribute table. Occasionally there may be layers with a many-to-one relationship between attribute table entries and multiple features in a data layer.

Most GIS employ underlying file structures to organize components of the spatial data. An example organization is shown in the bottom half of Figure 2-30, where the topological elements are recorded in a linked set of tables, here, one each for the polygons, lines, and nodes and vertices. Most GIS maintain the spatial and topological data as a single or cluster of linked files. This internal file structure is often insulated from direct manipulation by the GIS user, but underlies nearly all spatial data manipulations. A user may directly edit or otherwise manipulate table values, usually with the exception of the ID, and the underlying topology and coordinate data are accessed via requests to display, change, or analyze the spatial data components. Data layers may also include additional information (not shown) on the origin, region covered, date of creation, edit history, coordinate system, or other characteristics of a data set.

Note that not all GIS store coordinate and topological data in non-tabular file structures. Coordinates, points, lines, polygons, and other composite features may be stored in tables similar to attribute tables. It is premature to discuss the details of these *spatially enabled* databases, because they are based on something called a *relational data model*, described in detail in Chapter 8. Faster computers support this generally more flexible approach, allowing simpler and more transparent access across different types of GIS software.

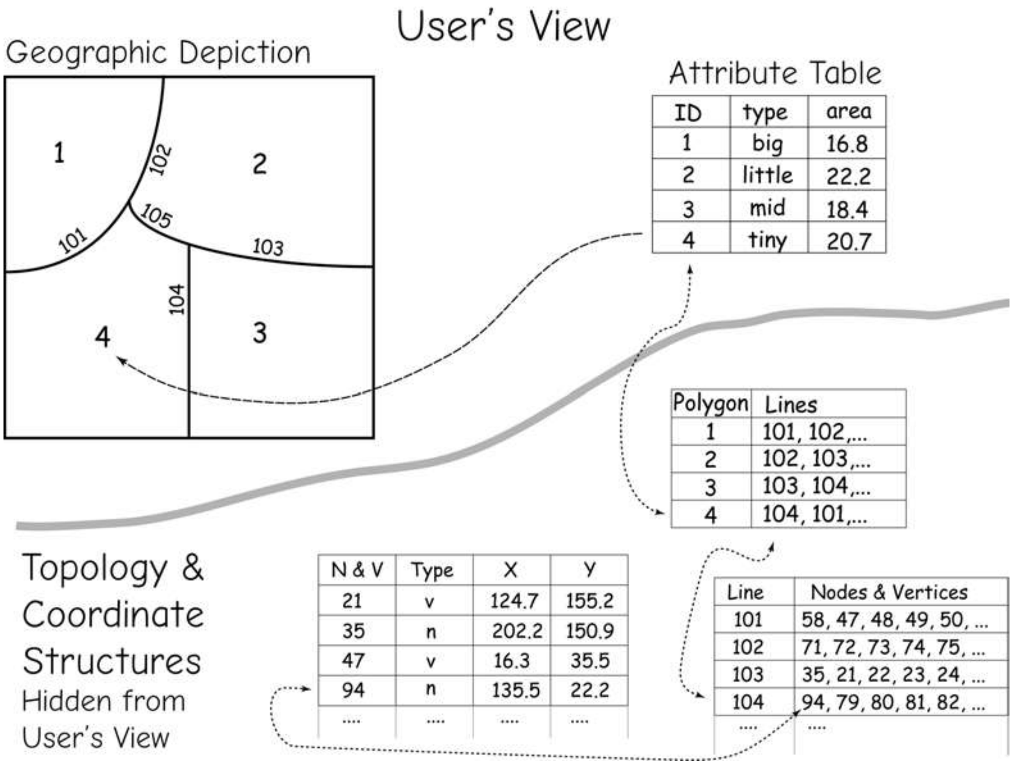


Figure 2-30: Features in a topological data layer typically have a one-to-one relationship with entries in an associated attribute table. The attribute table typically contains a column with a unique identifier, or ID, for each feature. Topology and coordinate data are often hidden from the user, but linked to the attribute and geographic features through pointers and index variables, described in the Data and File Structures section, later in this chapter.

Raster Data Models

Models and Cells

Raster data models define the world as a regular set of cells in a grid pattern (Figure 2-31). Typically these cells are square and evenly spaced in the x and y directions. The phenomena or entities of interest are represented by attribute values associated with each cell location.

Raster data models are the natural means to represent "continuous" spatial features or phenomena. Elevation, precipitation, slope, and pollutant concentration are examples of continuous spatial variables. These variables characteristically show significant changes in value over broad areas. The gradients can be quite steep (e.g., at cliffs), gentle (long, sloping ridges), or quite variable (rolling hills). Raster data models depict these gradients by changes in the values associated with each cell.

Raster data sets have a *cell dimension*, defining the edge length for each square cell (Figure 2-31). For example, the cell dimension may be specified as a square 30 meters on each side. The cells are usually oriented parallel to the x and y directions, and the coordinates of a corner location are specified.

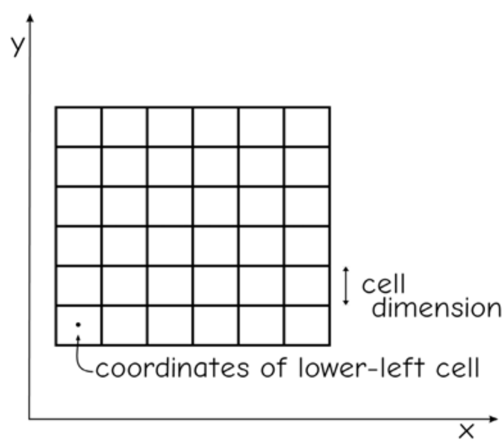


Figure 2-31: Important defining characteristics of a raster data model.

When the cells are square and aligned with the coordinate axes, the calculation of a cell location is a simple process of counting and multiplication. A cell location may be calculated from the cell size, known corner coordinates, and cell row and column number. For example, if we know the lower-left cell coordinate, all other cell coordinates may be determined by the formulas:

$$N_{cell} = N_{lower\text{-}left} + row * cell\ size \qquad (2.3)$$

$$E_{cell} = E_{lower\text{-}left} + column * cell\ size \qquad (2.4)$$

where N is the coordinate in the north direction (y), E is the coordinate in the east direction (x), and the row and column are counted starting with zero from the lower left cell.

There is often a trade-off between spatial detail and data volume in raster data sets. The number of cells needed to cover a given area increases four times when the cell size is cut in half (Figure 2-32). Smaller cells provide greater spatial detail, but at the cost of larger data sets.

The cell dimension also affects the spatial precision of the data set, and hence positional accuracy. The cell coordinate is usually defined at a point in the center of the cell. The coordinate applies to the entire area covered by the cell. Positional accuracy is typically expected to be no better than approximately one-half the cell size. No matter the true location of a feature, coordinates are truncated or rounded up to the nearest cell center coordinate. Thus, the cell size should be no more than twice the desired accuracy and precision for the data layer represented in the raster, and often it is specified to be smaller.

Each raster cell represents a given area on the ground and is assigned a value that may be considered to apply to the entire cell.

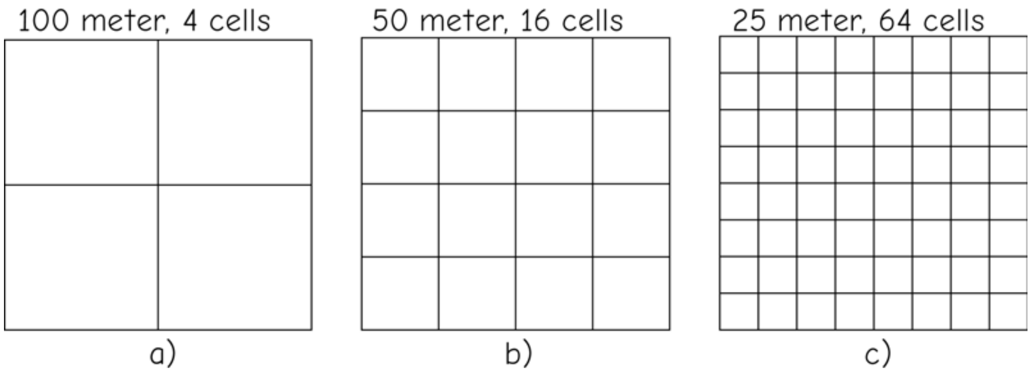


Figure 2-32: The number of cells in a raster data set depends on the cell size. For a given area, a linear decrease in cell size causes an exponential increase in cell number, e.g., halving the cell size causes a four fold increase in cell number.

If the variable is uniform across the raster cell, the value will be correct over the cell. However, under most conditions there is within-cell variation, and the raster cell value represents the average, central, or most common value found in the cell. Consider a raster data set representing annual weekly income with a cell dimension that is 300 meters (980 feet) on a side. Further suppose that there is a raster cell with a value of 710. The entire 300 by 300 meters area is considered to have this value of 710 pesos per week. There may be many households within the raster cell that do not earn exactly 710 pesos per week. However, the 710 pesos may be the average, the highest point, or some other representative value for the area covered by the cell. While raster cells often represent the average or the value measured at the center of the cell, they may also represent the median, maximum, or another statistic for the cell area.

An alternative interpretation of the raster cell applies the value to the central point of the cell. Consider a raster grid containing elevation values. Cells may be specified as 200 meters square, and an elevation value assigned to each square. A cell with a value of 8000 meters (26,200 feet) may be assumed to have that value at the center of the cell, but this value will not be assumed to apply to the entire cell.

A raster data model may also be used to represent discrete data (Figure 2-33), for example, to represent landcover in an area. Raster cells typically hold numeric or single-letter alphabetic characters. A coding scheme defines what land cover type the discrete values signify. Each code may be found at many raster cells.

Raster cell values may be assigned and interpreted in at least seven different ways (Table 2-1). We have described three: a raster cell as a point physical value (elevation), as a statistical value (average income), and as discrete data (landcover). Raster values may also be used to represent points and

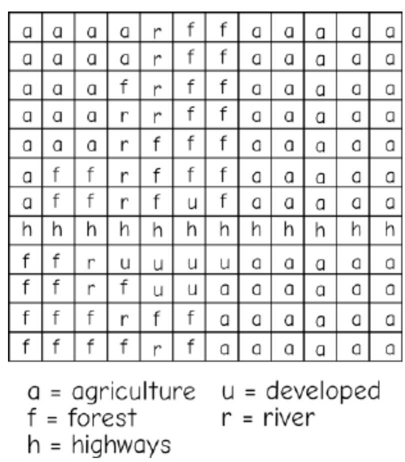


Figure 2-33: Discrete or categorical data may be represented by codes in a raster data layer.

Table 2-1: Types of data represented by raster cell values (from L. Usery, pers. comm.).

Data Type	Description	Example
point ID	alpha-numeric ID of closest point	hospital
line ID	alpha-numeric ID of closest line	nearest road
contiguous region ID	alpha-numeric ID for dominant region	state
class code	alpha-numeric code for general class	vegetation type
table ID	numeric position in a table	row
physical analog	numeric value representing surface value	elevation
statistical value	numeric value from a statistical function	population density

lines, as the IDs of lines or points that occur closest to the cell center.

Point and line assignment to raster cells may be complicated when there are multiple features within a single cell. For example, when light poles are represented in a raster data layer, cell value assignment is straightforward when there is only one light in a cell (Figure 2-34, near A). When there are multiple poles in a single cell there is some ambiguity, or generalization in the assignment (Figure 2-34, near B). One common solution represents one feature from the group, and retains information on the attributes and characteristics of that feature. This entails some data loss. Another solution is to reduce the raster cell size so that there are no multiple features in a cell. This may result in impractically large data sets. More complex schemes may record multiple instances of features in a cell, but these then may slow access or otherwise decrease the utility that comes from the simple raster structure.

Similar problems may occur when there are multiple line segments within a raster cell, for example, when linear features such as roads are represented in a raster data set. When two or more roads meet, they will do so within a raster cell, and some set of attributes must be assigned (Figure 2-34C). Since attributes are assigned by cells, some precedence must be established, with one line given priority over others.

Raster cell assignment also may be complicated when representing what we typically think of as discrete, uniform areas. Consider the area in Figure 2-35. We wish to represent this area with a raster data layer,

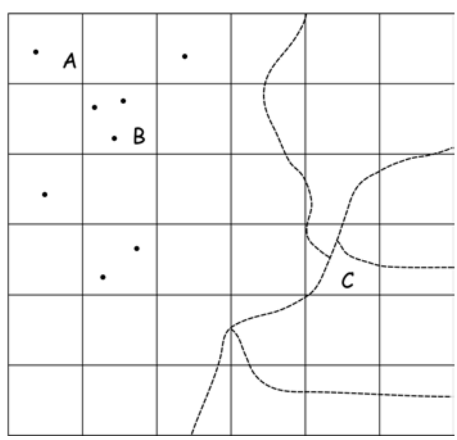


Figure 2-34: Raster cell assignment requires decisions when multiple objects occur in the same cell.

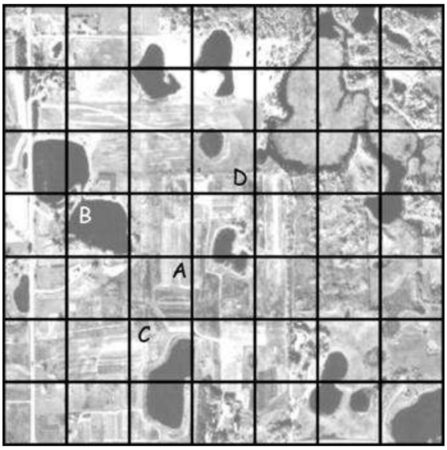


Figure 2-35: Raster cell assignment with mixed landscapes. Upland areas are lighter greys, water the darkest greys.

with cells assigned to one of two class codes, one each for land or water. Water bodies appear as darker areas in the image, and the raster grid is shown overlain. Cells may contain substantial areas of both land and water, and the proportion of each class may span from zero to 100 percent. Some cells are purely one class and the assignment is unambiguous; for example, the cell labelled A in the Figure 2-35 contains only land. Others are unambiguous, such as cell B (water) or D (land). Some are nearly equal in their proportion of land and water, as in cell C.

One common method to assign classes for mixed cells is called "winner-take-all". The cell is assigned the class of the largest area type. Cells A, C, and D would be assigned the land class, cell B the water class. Another option applies preference in cell assignment. If any of an "important" type is found, then the cell is assigned that value, regardless of the proportion. If we specify a preference for water, then cells B, C, and D in Figure 2-35 would be assigned the water type, and cell A the land type.

Regardless of the assignment method used, Figure 2-35 illustrates two phenomena when discrete objects are represented using a raster data model. First, some areas that are not the assigned class are included in some rasters cells. These "inclusions" are inevitable because cells must be assigned to a discrete class, the cell boundaries are rigidly assigned, and the class boundaries on the ground rarely line up with the cell boundaries. Some mixed cells occur in nearly all raster layers. The GIS user must acknowledge these inclusions, and consider their impact on the intended spatial analyses.

Second, differences in class assignment rules may substantially alter the data layer, as shown in our simple example. In more complex landscapes, there will be more potential cell types, which may increase the assignment sensitivity. Decreasing the raster cell size reduces the significance of classes in the assignment rule, but at the cost of increased data volumes.

The occurrence of more than one line or point within a raster cell can result in similar assignment problems. If two points occur, then which point ID is assigned? If two lines occur, then which line ID should be assigned? Some rule must be developed; for example, the point that falls nearest the center may be assigned, or the line with the longest segment within the raster cell. Similar to when area features are assigned to rasters, "inclusions" and dependence on the class assignment rules affect the output.

Raster Features and Attribute Tables

Raster layers may also have associated attribute tables. This is most common when nominal data are represented, but may also be used with ordinal or interval/ratio data. Just as with topological vector data, features in the raster layer may be linked to rows in an attribute table, and these rows may describe the essential nonspatial characteristics of the features.

Figure 2-36a and b show data represented in a raster model. Figure 2-36a shows a raster data set that maintains a one-to-one relationship between raster cells and in the data table. An additional column, cell-ID, must be added to uniquely identify each raster location. The corresponding attributes IDorg, class, and area are repeated for each cell. Note that the area values are the same for all cells and thus all rows in the table.

A one-to-one correspondence is rarely used with raster data sets because it often would require an unmanageably large size of attribute table. This small example results in 100 rows for the attribute table, but we often use raster data sets with billions of cells. If we insist on a one-to-one cell/attribute relationship, the table may become too large. Even simple processes such as sorting, searching, or subsetting records become prohibitively time consuming. Display and redraw rates become low, reducing the utility of these data, and decreasing the likelihood that GIS will be effectively applied.

To avoid these problems, a many-to-one relationship is usually allowed between the raster cells and the attribute table (Figure 2-36b). Many raster cells may refer to a single row in the attribute column. This substantially reduces the size of the attribute table for most data sets although it does so at the cost of some spatial ambiguity. There may be multiple, noncontiguous patches for a specific type. For example, the upper left and lower right portion of the raster data set in Figure 2-36b are both of class 10. Both are recognized as distinct features in the vec-

a) Raster, one-to-one

A	A	A	A	B	B	B	B	B	B
A	A	A	A	B	B	B	B	B	B
A	A	A	A	B	B	B	B	B	B
A	A	A	B	B	B	B	B	B	B
A	A	A	C	C	B	B	B	B	B
C	C	C	C	C	D	D	D	D	D
C	C	C	C	C	D	D	D	D	D
C	C	C	C	C	D	D	D	D	D
C	C	C	C	C	D	D	D	E	E
C	C	C	C	C	D	D	E	E	E

attribute table
(cell 1 is upper-left corner)

cell-ID	IDorg	class	area
1	A	10	0.8
2	A	10	0.8
3	A	10	0.8
4	A	10	0.8
5	B	11	0.8
6	B	11	0.8
7	B	11	0.8
⋮	⋮	⋮	⋮
100	E	10	0.8

b) Raster, many-to-one

10	10	10	10	11	11	11	11	11	11
10	10	10	10	11	11	11	11	11	11
10	10	10	10	11	11	11	11	11	11
10	10	10	11	11	11	11	11	11	11
10	10	10	15	15	11	11	11	11	11
15	15	15	15	15	21	21	21	21	21
15	15	15	15	15	21	21	21	21	21
15	15	15	15	15	21	21	21	21	21
15	15	15	15	15	21	21	21	10	10
15	15	15	15	15	21	21	10	10	10

attribute table

class	area
10	18.4
11	24.0
15	21.6
21	13.6

Figure 2-36: Raster data models rarely maintain this one-to-one relationship between cells and attributes (a), because table access and performance usually suffer. A many-to-one relationship between cells and table rows is adopted more often (b).

tor and one-to-one raster representation, but are represented by the same attribute entry in the many-to-one raster representation. This reduces the size of the attribute table, but at the cost of reducing the flexibility of the attribute table. Many-to-one relationships effectively create multipart areas. The data for the represented variable may be summarized by class; however, these classes may or may not be spatially contiguous.

An alternative is to maintain the one-to-one relationship, but to index all the raster cells in a contiguous group, thereby reducing the number of rows in the attribute table. This requires software to develop and maintain the indices, and to create them and reconstitute the indexing after spatial operations. These indexing schemes add overhead and increase data model complexity, thereby removing one of the advantages of raster data sets over vector data sets.

A Comparison of Raster and Vector Data Models

The question often arises, "Which are better, raster or vector data models?" The answer is neither and both. Neither of the two classes of data models is better in all conditions or for all data. Both have advantages and disadvantages relative to each other and to additional, more complex data models. In some instances it is preferable to maintain data in a raster model, and in others in a vector model. Most data may be represented in both, and may be converted among data models. As an example, land cover may be represented as a set of polygons in a vector data model or as a set of identifiers in each cell in a raster grid. The choice often depends on a number of factors, including the predominant type of data (discrete or continuous), the expected types of analyses, available storage, the main sources of input data, and the expertise of the human operators.

Raster data models exhibit several advantages relative to vector data models. First, raster data models are particularly suit-

Table 2-2: A comparison of raster and vector data models**.**

Characteristic	Raster	Vector
data structure	usually simple	usually complex
storage requirements	larger for most data sets with-out compression	smaller for most data sets
coordinate conversion	may be slow due to data volumes, and require resampling	simple
analysis	easy for continuous data, sim- ple for many layer combinations	preferred for network analyses, many other spatial operations more complex
spatial precision	floor set by cell size	limited only by positional measurements
accessibility	easy to modify or program, dueto simple data structure	often complex
display and output	good for images, but discrete features may show "stairstep" edges	maplike, with continuous curves, poor for images

able for representing themes or phenomena that change frequently in space. Each raster cell may contain a value different than its neighbors. Thus trends as well as more rapid variability may be represented.

Raster data structures are generally simpler than vector data models, particularly when a fixed cell size is used. Most raster models store cells as sets of rows, with cells organized from left to right, and rows stored from top to bottom. This organization is quite easy to code in an array structure in most computer languages.

Raster data models also facilitate easy overlays, at least relative to vector models. Each raster cell in a layer occupies a given position corresponding to a given location on the Earth's surface. Data in different layers align cell-to-cell over this position. Thus, overlay involves locating the desired grid cell in each data layer and comparing the values found for the given cell location. This cell look-up is quite rapid in most raster data structures, and thus layer overlay is quite simple and rapid when using a raster data model.

Finally, raster data structures are the most practical method for storing, displaying, and manipulating digital image data, such as aerial photographs and satellite imagery. Digital image data are an important source of information when building, viewing, and analyzing spatial databases. Image display and analysis are based on raster operations to sharpen details on the image, specify the brightness, contrast, and colors for display, and to aid in the extraction of information.

Vector data models provide some advantages relative to raster data models. First, vector models often lead to more compact data storage, particularly for discrete objects. Large homogenous regions are recorded by the coordinate boundaries in a vector data model. These regions are recorded as a set of cells in a raster data model. The perimeter grows more slowly than the area for most feature shapes, so the amount of data required to represent an area increases much more rapidly with a raster data model. Vector data are much more compact than raster data for most themes and levels of spatial detail.

Vector data are a more natural means for representing networks and other connected linear features. Vector data by their nature store information on intersections (nodes) and the linkages between them (lines). Traffic volume, speed, timing, and other factors may be associated with lines and intersections to model many kinds of networks.

Vector data models are easily presented in a preferred map format. Humans are familiar with continuous line and rounded curve representations in hand- or machine-drawn maps, and vector-based maps show these curves. Raster data often show a “stair-step” edge for curved boundaries, particularly when the cell resolution is large relative to the resolution at which the raster is displayed. Vector data may be plotted with more visually appealing continuous lines and rounded edges.

Vector data models facilitate the calculation and storage of topological information. Topological information aids in performing adjacency, connectivity, and other analyses in an efficient manner. Topological information also allows some forms of automated error and ambiguity detection, leading to improved data quality.

Conversion Between Raster and Vector Models

Spatial data may be converted between raster and vector data models. Vector-to-raster conversion involves assigning a cell value for each position occupied by vector features. Vector point features are typically assumed to have no dimension. Points in a raster data set must be represented by a value in a raster cell, so points have at least the dimension of the raster cell after conversion from vector-to-raster models. Points are usually assigned to the cell containing the point coordinate. The cell in which the point resides is given a number or other code identifying the point feature occurring at the cell location. If the cell size is too large, two or more vector points may fall in the same cell, and either an ambiguous cell identifier assigned, or a more complex numbering and assignment scheme implemented. Typically a cell size is chosen such that the diagonal cell dimension is smaller than the distance between the two closest point features.

Vector line features in a data layer may also be converted to a raster data model. Raster cells may be coded using different criteria. One simple method assigns a value to a cell if a vector line intersects with any part of the cell (Figure 2-37a, left). This ensures the maintenance of connected lines in the raster form of the data. This assignment rule often leads to wider than appropriate lines because several adjacent cells may be assigned as part of the line, particularly when the line meanders near cell edges. Other assignment rules may be applied, for example, assigning a cell as occupied by a line only when the cell center is near a vector line segment (Figure 2-37a, right). "Near" may be defined as some sub-cell distance, for instance, 1/3 the cell width. Lines passing through the corner of a cell will not be recorded as in the cell. This may lead to thinner linear features in the raster data set, but often at the cost of line discontinuities.

The output from vector-to-raster conversion depends on the algorithm used, even though you use the same input. This brings up an important point to remember when applying any spatial operation. The output often depends in subtle ways on the spatial operation. What appear to be quite small dif-

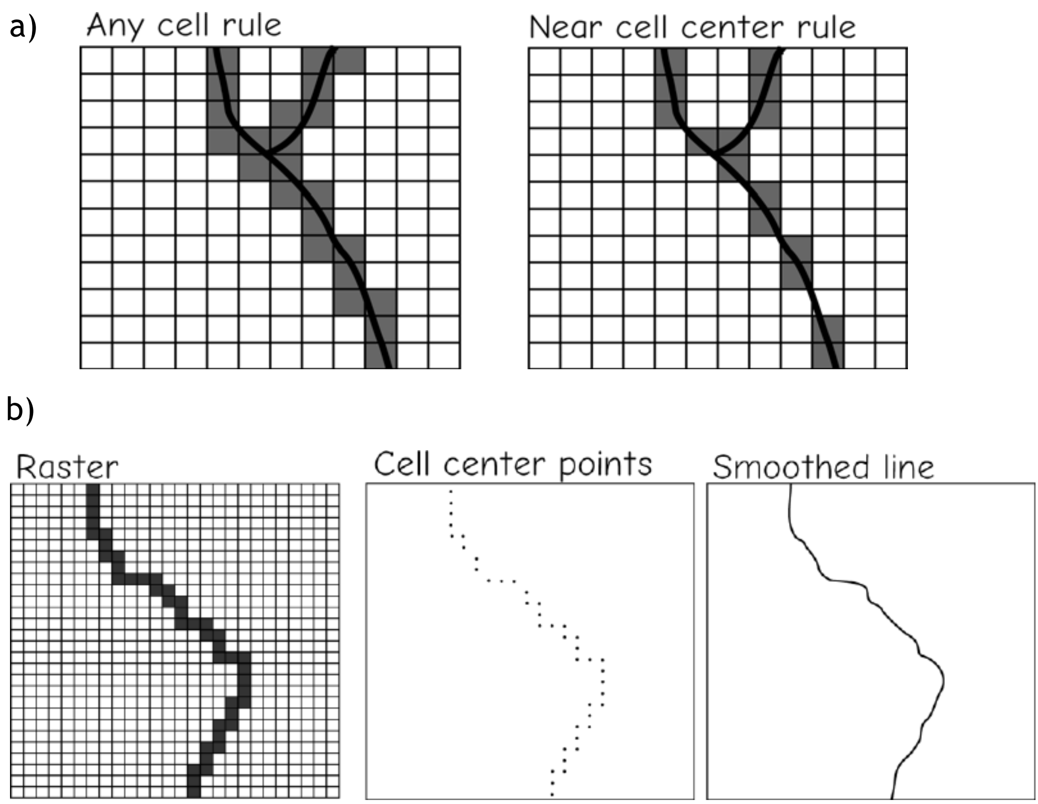


Figure 2-37: Vector-to-raster conversion (a) and raster-to-vector conversion (b). In a, cells are assigned in a raster if they intersect with a converted vector. The left and right panels show how two assignment rules result in different raster coding near lines. Panels in b show how raster data may be converted to vector formats, and may involve line smoothing or other operations to remove the "stair-step" effect.

ferences in the algorithm or key defining parameters may lead to quite different results. The ease of spatial manipulation in a GIS provides a powerful and often easy-to-use set of tools. The GIS user should bear in mind that these tools can be more efficient at producing errors as well as more efficient at providing correct results. Until sufficient experience is obtained with a suite of algorithms, in this case vector-to-raster conversion, small, controlled tests should be performed to verify the accuracy of a given method or set of constraining parameters.

Up to this point we have covered vector-to-raster data conversion. Data may also be converted in the opposite direction, from raster to vector data. Point, line, or area features represented by raster cells are converted to corresponding vector data coordinates and structures. Point features are represented as single raster cells. Each vector point feature is usually assigned the coordinate of the corresponding cell center.

Linear features represented in a raster environment may be converted to vector lines. Conversion to vector lines typically involves identifying the continuous connected set of grid cells that form the line. Cell centers are typically taken as the locations of vertices along the line (Figure 2-37b). Lines may then be "smoothed" using a mathematical algorithm to remove the "stair-step" effect.

Data Compression

We often compress spatial data files because they are large. Data compression reduces file size while maintaining the information contained in the file. Compression algorithms may be "lossless," in that all information is maintained during compression, or "lossy," in that some information is lost. A lossless compression algorithm will produce an exact copy of the original when it is applied and then the appropriate decompression algorithm applied. A lossy algorithm will alter the data when it is applied and the appropriate decompression algorithm applied. Lossy algorithms are most often used with image data, where substantial degradation still leaves a useful image, and are uncommonly applied to thematic spatial data, where any data degradation is typically not tolerated.

Data compression is most often applied to discrete raster data, for example, when representing polygon or area information in a raster GIS. There are redundant data elements in raster representations of large homogenous areas. Each raster cell within a homogenous area will have the same code as most or all of the adjacent cells. Data compression algorithms remove much of this redundancy.

Run length coding is a common data compression method. This compression technique is based on recording sequential runs of raster cell values. Each run is recorded as the value found in the set of adjacent cells and the run length, or number of cells with the same value. Seven sequential cells of type A might be listed as A7 instead of AAAAAAA. Thus, seven cells would be represented by two characters. Consider the data recorded in Figure 2-45, where each line of raster cells is represented by a set of run-length codes. In general, run length coding reduces data volume, as shown for the top three rows in Figure 2-45.

Raster

9	9	6	6	6	6	6	7
6	6	6	6	6	6	6	6
9	9	6	6	6	6	7	7
9	8	9	6	6	7	7	5

Run length codes

2:9, 5:6, 1:7

8:6

2:9, 4:6, 2:7

1:9, 1:8, 1:9, 2:6, 2:7, 1:5

Figure 2-45: Run-length coding is a common and relatively simple method for compressing raster data. The left number in the run-length pair is the number of cells in the run, and the right is the cell value. Thus, the 2:9 listed at the start of the first line indicates a run of length two for the cell value 9.

Note that in some instances run-length coding increases the data volume, most often when there are no long runs. This occurs in the last line of Figure 2-45, where frequent changes in adjacent cell values result in many short runs. However, for most thematic data sets containing area information, run-length coding substantially reduces the size of raster data sets.

There is also some data access cost in run-length coding. Standard raster data access involves simply counting the number of cells across a row to locate a given cell. To locate a cell in run-length coding we must sum along the run-length codes to identify a cell position. This is typically a minor additional cost, but in some applications the trade-off between speed and data volume may be objectionable.

Quad tree representations are another raster compression method. Quad trees are similar to run-length codings in that they are most often used to compress raster data sets when representing area features. Quad trees may be thought of as a raster data structure with a variable spatial resolution. Raster cell sizes are combined and adjusted within the data layer to fit into each specific area feature (Figure 2-46). Large raster cells that fit entirely into one uniform area are assigned the value corresponding to that area; for example, the three largest cells in Figure 2-46 are all assigned the value a. Successively smaller cells are then fit, halving the cell dimension at each iteration, again fitting the largest cell that will fit in each uniform area. This is illustrated in the top-left corner of Figure 2-46. Successively smaller cells are defined by splitting "mixed cells" into four quadrants, and assigning the values a or b to uniform areas. This is repeated down to the smallest cell size that is needed to represent uniform areas at the required detail.

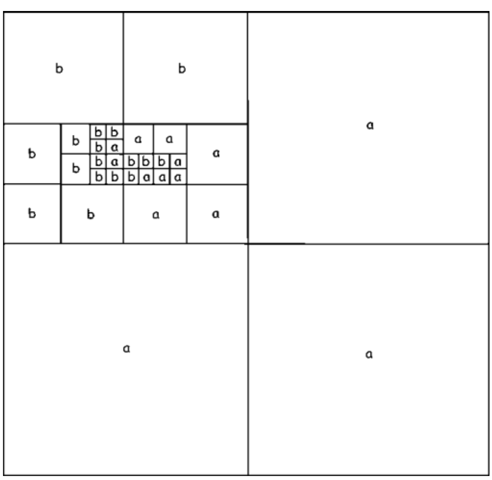


Figure 2-46: Quad tree compression.

The varying cell size in a quad tree representation requires more sophisticated indexing than simple raster data sets. Pointers are used to link data elements in a tree-like structure, hence the name quad trees. There are many ways to structure the data pointers, from large to small, or by dividing quadrants, and these methods are beyond the scope of an introductory text. Further information on the structure of quad trees may be found in the references at the end of this chapter.

There are many other data compression methods that are commonly applied. LZW is a lossless compression method commonly applied to image and raster data sets, particularly GIF images and TIFF formats. JPEG and wavelet compression algorithms are often applied to reduce the size of spatial data, particularly image or other data, although as implemented these are lossy algorithms. Generic bit- and byte-level compression methods may be applied to any files for compression or communications. There is usually some cost in time to the compression and decompression.

Raster Pyramids

We sometimes intentionally increase the size of our raster data sets without increasing the resolution in a process known as *pyramiding*. We create pyramids to increase display speeds when viewed at small scales ("zoomed out"). Long redraw times often hinder use of large data sets, particularly when panning frequently. When displayed at very small scales, the cell size of a data set may be smaller than the resolution of the computer screen. A raster data set 1,000,000 pixels across has 1000 times the data that can be displayed on a monitor with 1000 pixel horizontal resolution. However, display software must wade through all 1,000,000 data elements in a row to pick the 1 cell in 1000 to display. While clever software can help, there are limits to how much we can speed up the redraws.

Pyramiding in effect saves subsampled copies of the cells at various resolutions. In our example above, pyramids may do the equivalent of saving every two, every four, every 10, every 30, and every 100 cells, all within the same raster data set. The software then compares the display scale to the dimensions of the data set, and chooses the most appropriate cell resolution to display. Redraws are much faster, and transparent to the user.

Note that we say pyramids "in effect" save copies of cells at various resolutions. This is the simplest method, but often not the most efficient for space or speed of access. Sophisticated indexing may be used to point to the cells at the appropriate resolutions.

Note that pyramiding comes at a cost, both in the size and complexity of the raster data set. Indexing schemes complicate the simple raster data structure, and the software must be able to navigate the indexing scheme. Already large raster data sets may be inflated from a few percent to several times, although in practice it is typically less than a doubling of size.

Common File Formats

A small number of file formats are commonly used to store and transfer spatial data. Some of these file structures arose from distribution formats adopted by governmental departments or agencies. Others are based on formats specified by software vendors, and some have been devised by standards-making bodies. Some knowledge of the types, properties, and common naming conventions of these file formats is helpful to the GIS practitioner.

Common geographic data formats may be placed into three large classes: raster, vector, and attribute. Raster formats may be further split into single-band and multi-band file types. Multiband raster data sets are most often used to store and distribute image data, while single-band raster data sets are used to store both single-band images and nonimage spatial data. Table 2-3 summarizes some of the most common spatial data formats.

Most GIS softwares provide some utility for data import and export from standard formats, as well as formats specific to the data types in the purview of the software; for example, a terrain analysis package would likely support a broad range of raster data formats.

There is an open source initiative and tool by the name of GDAL, for Geospatial Data Abstraction Library, with provides a cross-platform utility to translate among many common vector and raster file formats. Command-line tools are described and downloadable through links at www.gdal.org. This free utility is flexible, and often can be used to extend the reach of commercial packages, by first using GDAL to convert files from unsupported to supported types, and then importing these into the target software. GDAL source code is available under an Open Geospatial Foundation license, and so have been incorporated into many other open source tools.

Table 2-3: Common formats for spatial data.

Type and source	Extension	Characteristics (R=Raster, V=Vector, A=Attribute, I=Image)
Comma Separated Value	.csv	Common ASCII text format used to distribute attributeand often vector information (A, V).
DXF, AutoDesk	.dxf	Drawing exchange file, an ASCII or binary file for exchanging spatial data (V)
DWG, Autodesk	.dwg	Native binary file used by AutoDesk to store geo-graphic data and drawings in AutoCAD (V)
Geodatabase, ESRI	.gdb, .mdb	ESRI container for many data types (R, V, A, I)
GeoJSON, open standard	.json, .geojson	Open standard for representing and displaying simple geographic features (V, A).
GeoPackage, open standard	.gpkg	Open standard for representing vector and raster data, compatible with SQlite
GeoTIFF, open stan-dard	.TIF, .TFF	An extension for georeferencing Aldus-Adobe public domain TIFF format (R)
GPX, open standard	.gpx	A specification based on XML for basic GNSS data
Imagine, ERDAS	.img	Multiband capable image format (R)
Interchange, ESRI	.e00	ASCII text file for vector and identifying attribute data (V).
Keyhole Markup Lan-guage, Google	.KML	XML extension for displaying and annotating featuresand images (V, I, A)
LAS, ASPRS	.LAS	Laser point cloud data storage (V)
shapefile, ESRI	.shp, .shx, .dbf, .prj, and others	Three or more binary files that include the vector coor-dinate, attribute, and other information (V)
TIGER, U.S. Census	tgrxxyyy, stfzz	Set of files by U.S. census areas, xx is a state code,yyy an area code, zz numbers for various file types
MIF/MID, MapInfo	.mif, .mid	Map Interchange File, vector and raster data transportfrom MapInfo (V,R)
NetCDF, OGC	.cdf	Machine-independent data formats for scientific dataarrays (R, A, I)
NLAPS, NASA	various in a directory	Image data from various Landsat satellites, in a speci-fied directory structure (I, R)
SDTS, U.S. Govern-ment	none	Spatial Data Transfer Standard, specifies the spatial objects, attributes, reference system (R,V, A)

Summary

In this chapter we have described the main ways of conceptualizing spatial entities, and of representing these entities as spatial features in a computer. We commonly employ two conceptualizations, also called spatial data models: a raster data model and a vector data model. Both models use a combination of coordinates, defined in a Cartesian or spherical system, and attributes, to represent our spatial features. Features are usually segregated by thematic type in layers.

Vector data models describe the world as a set of point, line, and area features. Attributes may be associated with each feature. A vector data model splits that world into discrete features, and often supports topological relationships. Vector models are most often used to represent features that are considered discrete, and are compatible with vector maps, a common output form.

Raster data models are based on grid cells and represent the world as a "checkerboard," with uniform values within each cell. A raster data model is a natural choice for representing features that vary continuously across space, such as temperature or precipitation. Data may be converted between raster and vector data models.

We use data structures and computer codes to represent our conceptualizations in more abstract, but computer-compatible forms. These structures may be optimized to reduce storage space and increase access speed, or to enhance processing based on the nature of our spatial data.

Study Questions

- How is an entity different from a cartographic object?

- Describe the successive levels of abstraction when representing real-world spa- tial phenomena on a computer. Why are there multiple levels, instead of just one levelin a spatial data representation?

- Define a data model and describe three primary differences between the two most commonly used data models.

- Characterize the following lists as nominal, ordinal, or interval/ratio:a) 1.1, 5.7, -23.2, 0.4, 6.67

b) green, red, blue, yellow, sepia

c) white, light grey, dark grey, black

d) extra small, small, medium, large, extra large

e) forest, woodland, grassland, bare soil

f) 1, 2, 3, 4, 5, 6, 7.

- Characterize the following lists as nominal, ordinal, or interval/ratio:

a) Spurs, Citizens, Reds, Hornets, Baggies, Toffees, Potters

b) pinch, handful, bucket, bushel, truckload

c) 6.2, 7.8, 1.1, 0.5, 19.3

d) gram, kilogram, metric ton

e) Mexico, Canada, Argentina, Guyana, Martinique.

f) small, smaller, smallest

- Complete the following coordinate conversion table, converting the listed points from degrees-minutes-seconds (DMS) to decimal degrees (DD), or from DD to DMS. See Figure 2-11 for the conversion formula.

Point	DMS	Decimal Degrees
1	36°45'12"	36.75333
2	114°58'2"	
3	85°19'7"	
4		14.00917
5		275.00001
6		0.99528
7	183°19'22"	

- Complete the following coordinate conversion table, converting the listed points from degrees-minutes-seconds (DMS) to decimal degrees (DD), or from DD to DMS. See Figure 2-11 for the conversion formula.

Point	DMS	Decimal Degrees
1	97°45'10"	97.75278
2	122°10'2"	
3	15°0'12"	
4		322.19861
5		152.65583
6		5.75
7	23°12'50"	

- What is topology, and why is it important? What is planar topology, and when might nonplanar be more useful than planar topology?

- What are the respective advantages and disadvantages of vector data models vs. raster data models?

- Under what conditions are mixed cells a problem in raster data models? Inwhat ways may the problem of mixed cells be addressed?

- Indicate which of the following are allowable geographic coordinates:

a) N45 45’ 45”   b) longitude -127.34795   c) S96 12’ 33”

d) E 66 15’ 60”   e) W -12 23’ 55”   f) N 56.9999

- Indicate which of the following are allowable geographic coordinates:

a) N145 45’12”   b) latitude -62.34795   c) S110 52’ 43”

d) S 49 15’ 60”   e) N 89 59’ 59”   f) S -46.6000

- Express the following base 10 numbers in binary notation:

a) 2   b) 8   c) 9   d) 17

e) 0   f) 128   g) 22   h) 19

- Express the following base 10 numbers in binary notation:

a) 1   b) 23   c) 256   d) 4

e) 11   f) 10   g) 3   h) 20

- Express the following binary numbers in base 10 notation:

a) 0101   b) 0001   c) 1111   d) 00101101

e) 1101   f) 1011   g) 10000001   h) 11111111

- Express the following binary numbers in base 10 notation:

a) 1110   b) 1001   c) 0011   d) 10000101

e) 1000   f) 1010   g) 10010001   h) 11110000

- The following figure shows change in raster resolution, combining four smallcells on the left to create and output for each corresponding larger cell on the right. Fill in the two rasters on the right, for the interval/ratio data (top), and the nominal data (bottom). Assume null values are not ignored, and a majority rule for nominal data.

2	2	1	1
2	2	2	2
1	2	1	2
4	5	3	null

→

a	a	a	a
a	a	a	b
a	b	a	b
c	c	c	d

→

- The following figure shows change in raster resolution, combining four smallcells on the left to create and output for each corresponding larger cell on the right. Fill in the two rasters on the right, for the interval/ratio data (top), and the nominal data (bottom). Assume null values are not ignored, and a majority rule for nominal data.

2	2	1	2
null	2	2	2
1	null	4	3
3	null	3	4

→

a	a	b	a
a	b	c	b
null	c	c	c
c	c	c	c

→

- What is a triangulated irregular network?

- Why do we use binary numbers in computers?

- Why do we need to compress data? Which are most commonly compressed, raster data or vector data? Why?

- What are pointers when used in the context of spatial data, and how are they helpful in organizing spatial data?

- What are the main concepts behind object data models, and how do they differfrom other data models?

- Write the run length coding for each of the rows in this raster:

b	b	a	a	a	c	a	a	a
c	c	b	b	d	d	d	a	a
b	b	b	b	b	b	b	b	b
e	c	f	b	a	d	f	b	a
a	s	a	f	f	f	b	b	a

- Write the run length coding for each of the rows in this raster:

c	c	c	c	a	a	a	a	a
a	a	b	b	d	d	d	a	a
e	e	e	f	f	f	f	f	e
a	a	a	a	a	a	a	a	a
c	c	a	a	a	b	f	d	e

SPATIAL ANALYSIS AND MODELING (GEOGRAPHICAL INFORMATION SYSTEM)

UNIT – IV

Dec 2023

Shaik Saleemmiya, Civil Dept, Ace Engg College

Spatial Analysis

Most GIS systems are for the purpose of representing and describing features of the real world. Spatial databases perform this function

- Points, lines, polygons concepts for representation
- Coordinate systems as fundamental properties of spatial data
- geographic file formats for storage

Spatial Analysis involves gaining an understanding of the patterns, and associated cause and effect processes, underlying the features which have been described in order to

- Make better decisions
- Understand the phenomena as a goal in itself

Spatial Analysis

Types of Spatial Analysis:

1. Spatial data manipulation:

Classic GIS capabilities - Spatial queries & measurement, buffering, map layer overlay

2. Spatial data analysis:

Descriptive and exploratory, Visualization through data manipulation and mapping

3. Spatial statistical analysis:

Hypothesis testing; Are data “to be expected” or are they “unexpected” relative to some statistical model, usually of a random process

4. Spatial modeling:

Prediction, Constructing models (of processes) to predict spatial outcomes (patterns) What if analyses

Spatial Analysis

Types of Spatial Analysis:

1. Vector Spatial Analysis - Extracting portions of data is an effective means of isolating specific areas for further processing or data analysis
   a) Extraction
      i. Clip
      ii. Split
      iii. Select
      iv. Erase
   b) Overlay
      i. Point-in-Polygon
      ii. Line-in-Polygon
      iii. Polygon-in-Polygon
   c) Proximity
      i. Buffer
      ii. Near
      iii. Point Distance

Spatial Analysis

2. Raster Spatial Analysis

a) Surface Analysis
   a) Slope
   b) Aspect
   c) Contour
   d) Hill Shade
   e) Viewshed
b) Local Functions and Statistics
c) Neighborhood Functions and Statistics
d) Zonal Functions
   a) Distance
   b) Straight-line
   c) Shortest path

Spatial Analysis

Spatial Data Criteria:

- X-Y Coordinate System
- Shape
- Area/Size
- Perimeter
- Distance
- Neighborhood

Spatial Analysis Functions

- *Measurements*
- *Query*
- *Extraction*
- *Proximity*
- *Classification*
- *Topology*
- *Network analysis …*

MEASUREMENT

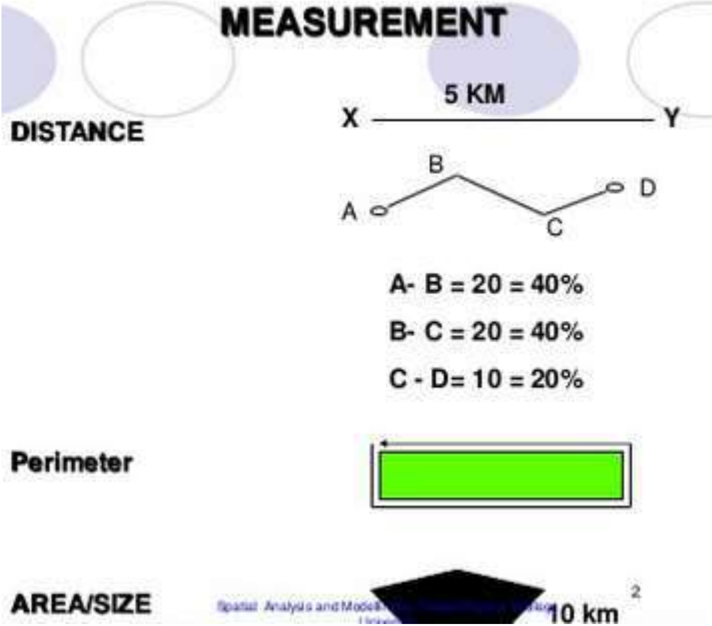
- ***Measurements*** are simple numerical values that describe aspects of geographic data.
- Measurement functions in GIS includes
  - Distance,
  - Perimeter and
  - Area

- Many types of interrogation ask for measurements
  - ✓ we might want to know the total area of a parcel of land, or the
  - ✓ distance between two points, or the length of a stretch of road and in principle all of these measurements are obtainable by simple calculations inside GIS.

Comparable measurements by hand from maps can be very tedious and error-prone.

Measuring Area/Perimeter

- Many vector systems automatically generate area and perimeter measurements for polygon features and store these values in prescribed attribute fields.
- The systems that do not have this automatic function do provide a way for you to generate area and perimeter and store the results in user-defined fields.
- Once calculated and stored, you can select multiple polygon features and sum their area and perimeter



QUERY

- *Queries* are the most basic of analysis operations, in which the GIS is used to answer simple questions posed by the user.
- No changes occur in the database, and no new data are produced with these type of selection

- The operations vary from simple and well-defined queries like 'how many houses are found within 1 km of this point', to vaguer questions like 'which is the closest city to Los Angeles going north', where the response may depend on the system's ability to understand what the user means by 'going north'

Attribute Query (Boolean Selection)

- It involves picking features based on query expressions, which use
  - Boolean algebra (and, or, not),
  - set algebra (>, <, =, >=, <=),
  - arithmetic operators (=, -, *, /), and
  - user-defined values.

Simply put, the GIS compares the values in an attribute field with a query expression that you define

- For example, if you want to select every restaurant whose price is considered inexpensive, you would use a query expression like "PRICE = $"
- where "**PRICE**" is the attribute field under investigation,
- "**=**" is the set algebra operator, and "**$**" is the value

Spatial Selection (Spatial Searches/query)

- While attribute queries select features by sorting through records in a data file, spatial selection chooses features from the map interface.
- In most cases, it selects features from one layer that fall within or touch an edge of polygon features in a second layer (or an interactively drawn graphic polygon).

- There are many types of spatial selection like point in polygon, it is a spatial operation in which points from one feature dataset are overlaid on the polygons of another to determine which points are contained within the polygons.

Extraction

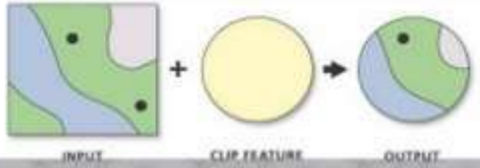
- Extracting portions of data helps to isolate specific areas for further processing or data analysis.
- Similar to queries and selection sets, extraction functions can reduce the size of datasets and/or facilitate more complex interpretation.
- Queries and selection sets also allow to isolate portions of a dataset
- Extraction techniques differ in that these portions of data are isolated in a permanent way - through the creation of new data layers

- GIS software packages provide a suite of tools to extract data, the most useful being, clip,select, and split

Clip

- Extracts input features that overlay the clip features.
- Working much like a cookie-cutter.
- This is particularly useful for creating a new feature class—also referred to as study area or area of interest (AOI)—that contains a geographic subset of the features in another, larger feature class

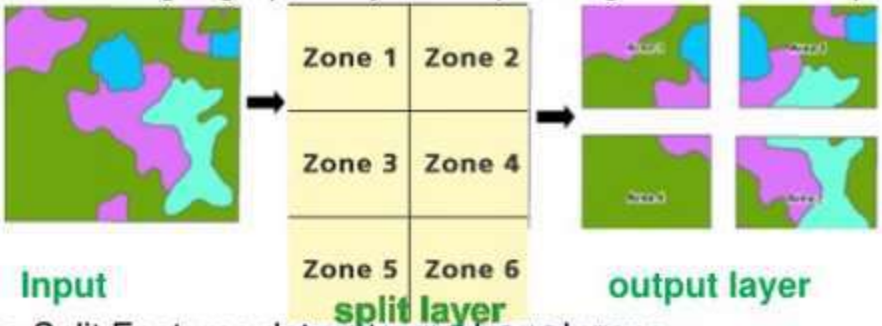
- Use this tool to cut out a piece of one feature class using one or more of the features in another feature class as a cookie cutter



Clip is useful for developing a subset of features from a series of existing data layers to match a common boundary.

Split

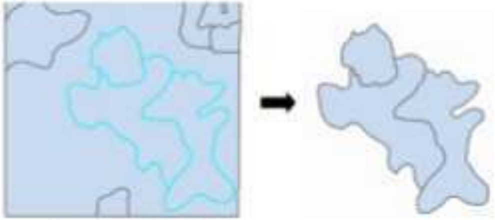
- Split is used to divide an input layer into two or more independent layers:
  - based on geographically corresponding features in a split layer



The Split Features dataset must be polygons.

Select

- Extracts features from an input feature class or input feature layer, typically using a select or Structured Query Language (SQL) expression and stores them in an output feature class.



Selected features in the Input layer

Output

Eg. Urban planner might wish to look at only double-line streets in the particular municipality of interest.
In this case, he or she would execute a selection query to extract only those desired features to a new layer

CLASSIFICATION

- Classification is the procedure of identifying a set of features as belonging to a group and defining patterns.
- Some form of classification function is provided in every GIS.
- Classification is important because it defines patterns.
- One of the important functions of a GIS is to assist in recognizing new patterns.
- Classification is done using single data layers, as well as with multiple data layers as part of an overlay operation.

- When you perform a classification, you group similar features into classes by assigning the same symbol to each member of the class.
- Aggregating features into classes allows you to spot patterns in the data more easily.
- The definition of a class range determines which features fall into that class and affect the appearance of the map.

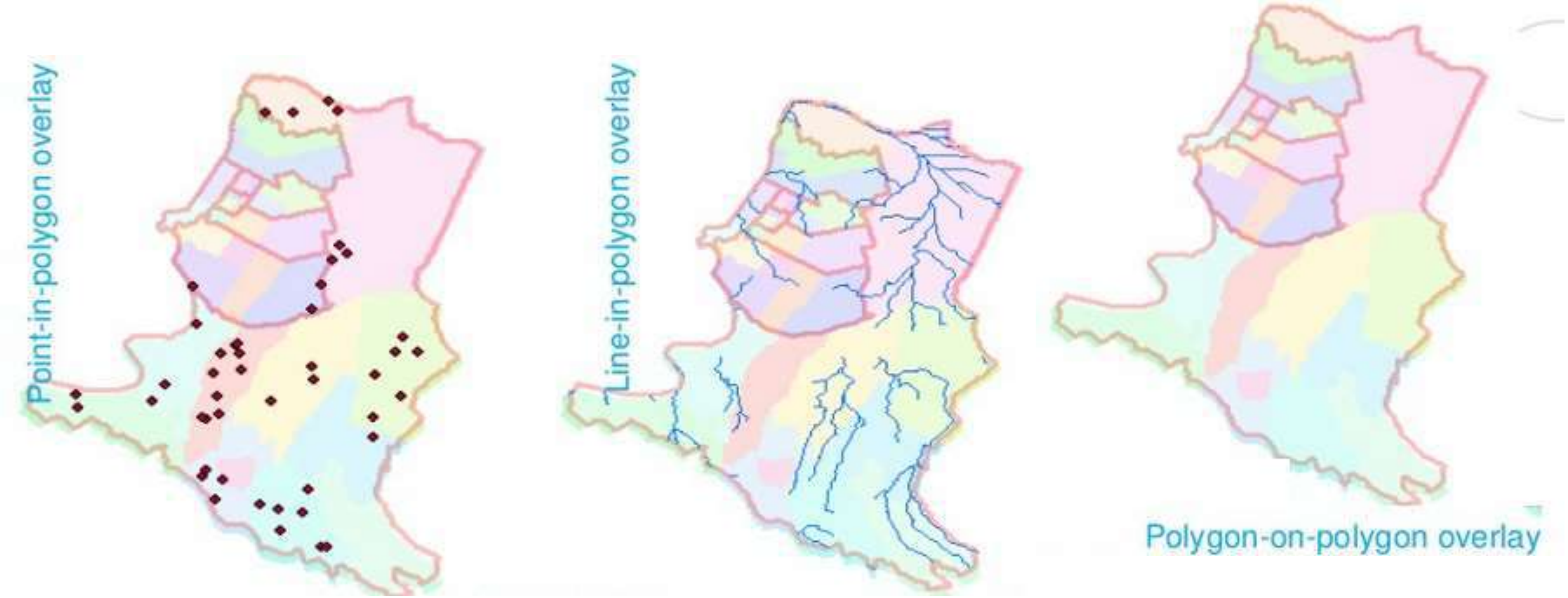
OVERLAY ANALYSIS

- Overlay is one of the most common and powerful GIS functions.
- It investigates the spatial association of features by "vertically stacking" feature layers to investigate geographic patterns and determine locations that meet specific criteria.
- An overlay operation combines the geometries and attributes of two feature layers to create the output
- Feature layers to be overlaid must be spatially registered and based on the same coordinate systems

Feature type and overlay

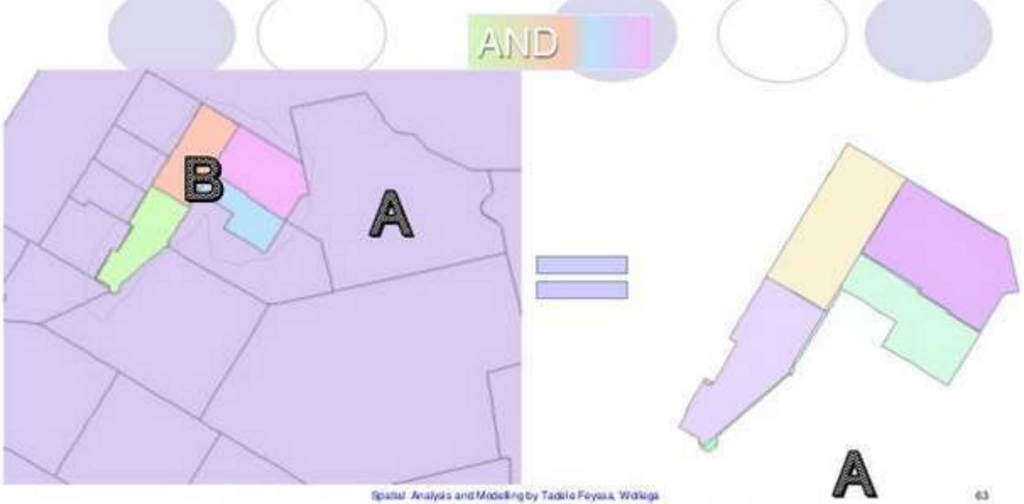
- There are two group of overlay operations
- The first group uses two polygon layers as input
- The second group uses one polygon layer and other layer which may contain points or lines

- Overlay operation can be classified as:
  - Point-in-polygon overlay
  - Line-in-polygon overlay
  - Polygon-on-polygon overlay



Overlay Methods

- Overlay methods are based on the Boolean connectors AND, OR and XOR
- Intersect uses the AND connector
- Union uses the OR connector.
- Differences uses XOR connector
- Union preserves all features from the inputs
- The area extent of the output combines the area extents of both input layers



- Union requires that both input layers be polygon layers
- Intersect preserves only those features that fall within the area extent common to the inputs
- The input layers may contain different feature types although in most cases one of them is a point, line or polygon

Application of Overlay

- An overlay operation combines features and attributes from the input layers
- The overlay output is useful for query and modelling purposes.
- For example a company who is looking a parcel that is zoned a commercial area, not subject for flooding and not more than a mile from heavy duty road may use overlay method to identify the area

NEIGHBORHOOD FUNCTIONS

- Neighborhood operations consider the characteristics of neighboring areas around a specific location.
- The principle here is to find out the characteristics of the vicinity, here called neighborhood, of a location.
- After all, many suitability questions, for instance, depend not only on what is at the location, but also on what is near the location.
- Thus, the GIS must allow us 'to look around locally'.

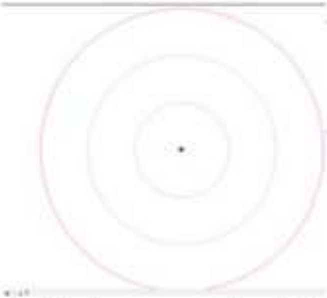
- For example, our target is Nekemte Hospital. Its neighborhood can be defined as
  - an area within 2 km distance, as the crow flies, or
  - an area within 2 km travel distance, or
  - all roads within 500 m travel distance, or
  - all other clinics within 15 minutes travel time, or
  - all residential areas, for which the clinic is the closest clinic

Proximity computation

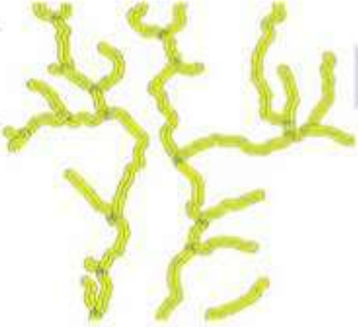
- In proximity computations, we use geometric distance to define the neighborhood of one or more target locations.
- All GIS programs provide some neighborhood analyses, which include buffering, interpolation, Theissen polygons, and various topographic functions.
- The most common and useful technique is buffer zone generation.
- Another technique based on geometric distance that is Thiessen polygon generation.

Buffering

- Buffering works based on proximity concept
- Feature for buffering may be points, lines or polygons
- Buffering around point create a circle
- Around lines a series of elongated buffer zones around each line segment
- A buffer around a polygon creates an extended area from the polygon boundaries



Buffering around **point**



Buffering Around **lines**

Buffering around a **polygon**

- Buffering uses distance measurements from selected features
- We must know the measurement unit of features we are dealing with

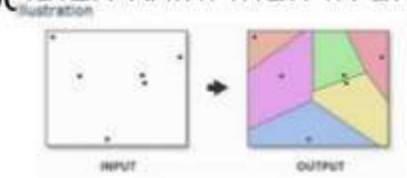
Application of Buffering

- Buffering creates a buffer zone data set
- A buffer zone often treated as a protection zone and is used for planning and regulatory purposes
- A city may require a buffer zone of 500m for alcohol trading from school
- A 30m buffer zone along river bank may needed to protect a river
- A buffer zone may be treated as a neutral zone and as a tool for conflict resolution
- Buffering zone also used for identifying suitable sites for different purposes
- Buffering also can be applied for sampling methods.

EG A stream network can be buffered at regular distances to analyse vegetation variations as one moves away from the stream

Thiessen Polygons (Analysis)

- It Creates Thiessen polygons from point input features.
- Each Thiessen polygon contains only a single point input feature. Any location within a Thiessen polygon is closer to its associated point than to any other point input feature.



This tool is used to divide the area covered by the point input features into Thiessen or proximal zones. These zones represent full areas where any location within the zone is closer to its associated input point than to any other

NETWORK ANALYSIS

Network is *any system of interconnected linear features*
A network is a system of interconnected elements, such as edges (lines) and connecting junctions (points), that represent possible routes from one location to another

What is network analysis?

Solving problems involving networks

Its goal is efficiency – Saving time and money.

Network Analysis

- Network analyses involve analyzing the flow of networks—a connected set of lines and point nodes.
- These linear networks most often represent features such as rivers, transportation corridors (roads, railroads, and even flight paths), and utilities (electric, telephone, television, sewer, water, gas).
- Point nodes usually represent pickup or destination sites, clients, transformers, valves, and intersections. People, water, consumer packages, kilowatts, and many other resources flow to and from nodes along linear features.
  - For example, a street segment might only provide flow in one direction (a one-way street) and at a certain speed.
  - Network analysis tools help you analyze the “cost” of moving through the network.
  - “cost” can be money, time, distance, or effort.
  - The three major types of network analyses include route selection (optimal path or shortest path), resource allocation, and network modeling.

Spatial Modeling

Introduction

- One of many uses of GIS analysis tools is to build models.
- What is a model?
- A **model** is a simplified representation of a phenomena or a system.
- A map is a model.
- So are the vector and raster data models for representing spatial features.

Modeling

- Models are used in many different ways, from simulations of how the world works, to evaluations of planning scenarios, to the creation of indicators of suitability or vulnerability
- Model is a simplification of reality in be viewed as a model.

- In the field of GIS, modelling provide understanding of the way the world works with sufficient precision and accuracy to allow prediction and confident decision-making.
- Modeling concern the way in which analyses are carried out using standard functionality

Application

- Road construction
- Construction in mountainous areas is complex engineering task, Cost factors, such as the number of tunnels & bridges to be constructed,
- Volume of rock & soil to be removed.
- GIS can help to compute such costs on the basis of an up-to-date DEM and soil map.

Basic Elements of GIS Modeling

Classification of GIS Models

- Descriptive or Prescriptive
- Deterministic or Stochastic
- Static or Dynamic
- Deductive or Inductive

A model may be descriptive or prescriptive.

- Descriptive model describes the existing conditions of spatial data
- Prescriptive model offers a prediction of what the conditions could be or should be.
- Eg If we use maps as analogies, a vegetation map would represent a descriptive model and a potential natural vegetation map, a prescriptive model.

- Spatial analysis could be either for prescriptive or predictive applications.

Prescriptive model:

- Used for planning & site selection.
- This involve the use of criteria & parameters to quantify environmental, economic & social factors.
- The model enumerates a number of conditions to be met.

Predictive model:

- A forecast is made of the likelihood of future events.
- Various spatial data layers used (raster or vector).
- Analytical questions, such as why or what if.
- It is intended to construct models and perform predictions.

  E.g. Pollution, erosion, landslides.

descriptive or prescriptive model...

The vegetation map shows existing vegetation, whereas the potential natural vegetation map predicts the vegetation that could occupy a site without disturbance or climate change

A model may be deterministic or stochastic...

- Both deterministic and stochastic models are mathematical models represented by equations with parameters and variables.
- A stochastic model considers the presence of some randomness in one or more of its parameters or variables, but a deterministic model does not

A model may be static or dynamic…

- A dynamic model emphasizes:
  - the changes of spatial data and the interactions between variables,
- Whereas a static model deals with the state of spatial data
  - at a given time
- Many environmental models such as groundwater pollution and soil water distribution are best studied as dynamic models

A model may be deductive or inductive...

- A deductive model represents the conclusion derived from a set of premises.
- These premises are often based on scientific theories or physical laws
- An inductive model represents the conclusion derived from empirical data and observations.
- Eg To assess the potential for a landslide one can use a deductive model based on laws in physics or use an inductive model based on recorded data from past landslides

The Modeling Process

- The development of a model follows a series of steps.
- 1st step is to define the goals of the model
- This is similar to defining a research problem
  - What is the phenomenon to be modeled?
  - Why is the model necessary?
  - What spatial and time scales are appropriate for the model?
- One can use a conceptual or schematic model to show the essential structure of the model

2nd step is to break down the model into elements and to define the properties of each element and the interactions between the elements

- A flowchart may used as a useful tool for linking the elements
- Also at this step, one will gather mathemat**ical equations of the model and use tools in a** GIS to carry out the **computation**

3rd step is the implementation and calibration of the model

- Data are needed for running and calibrating the model

4th Validate the model

Digital Elevation Model (DEM)

- The land surface has been the object of mapping, modelling and analysis for several years
- Map makers devised many techniques for terrain mapping; such as contouring, hill shading, 3D-view perspectives, etc.
- Geomorphologists have developed measures of the land surface which include slope, aspect and surface curvature

- Terrain mapping is no more a subject of specialists.
- GIS has made it relatively easy to incorporate them in many application areas
- Slope and aspect play a regular role in hydrological modelling, snow cover evaluation, soil mapping land slide area delineation and soil erosion modelling

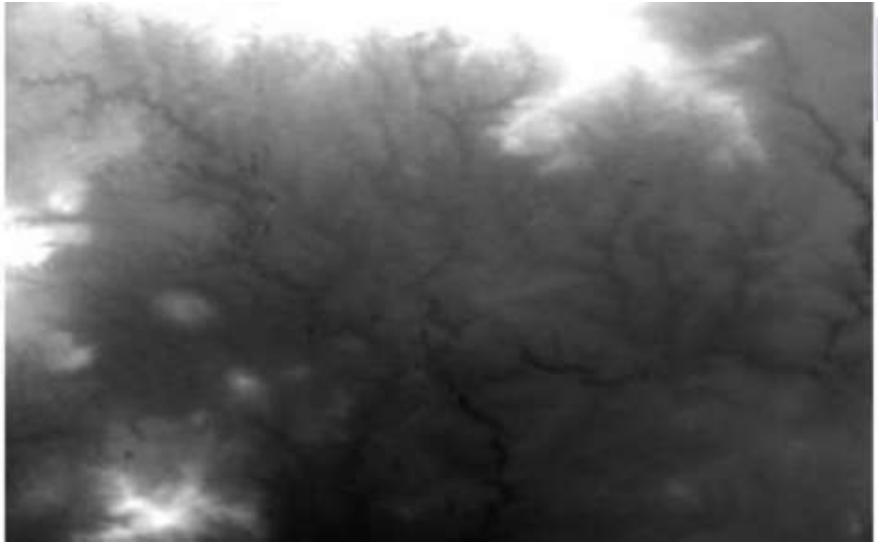
- Most GIS packages treat elevation data (Z values) as attribute data at point or cell location
- However, in 3-D model an additional coordinate to x and y
- In raster data modelling the z value corresponds to the cell value
- In vector data modelling the z value will be stored as attribute data

Digital elevation model and TIN

- Digital elevation model is a data input for terrain mapping
- The two most common data input for terrain mapping and analysis are DEM and TIN
- Digital elevation model (DEM) is based on raster data analysis
- Triangulated irregular network (TIN) is based on vector based data analysis
- We can't use TIN and DEM together but we can change TIN to DEM or DEM to TIN

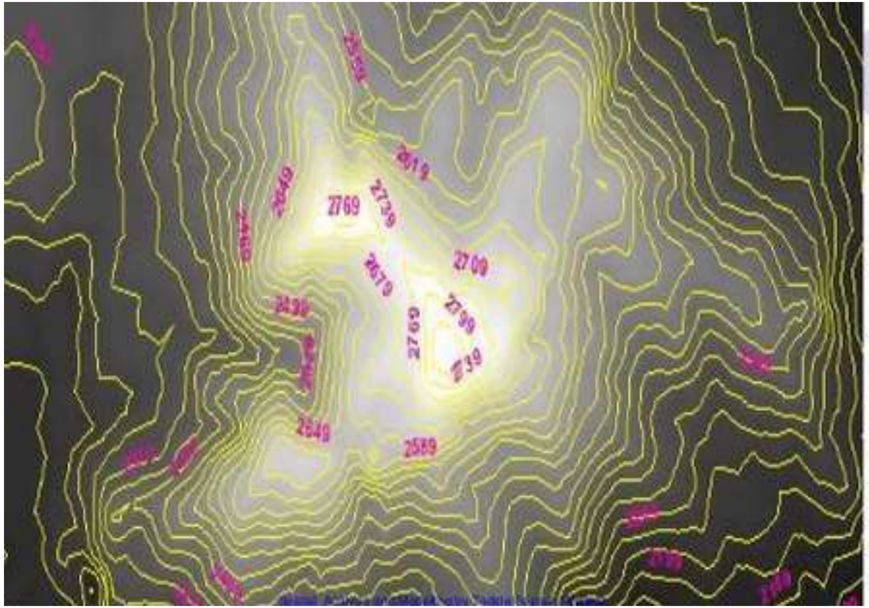
DEM

- A DEM represents a regular array of elevation points
- A point based DEM must be converted before using this data for terrain mapping
- This conversion simply places each point in the DEM at the centre of a cell in elevation raster
- Therefore DEM and elevation raster can be used interchangeably



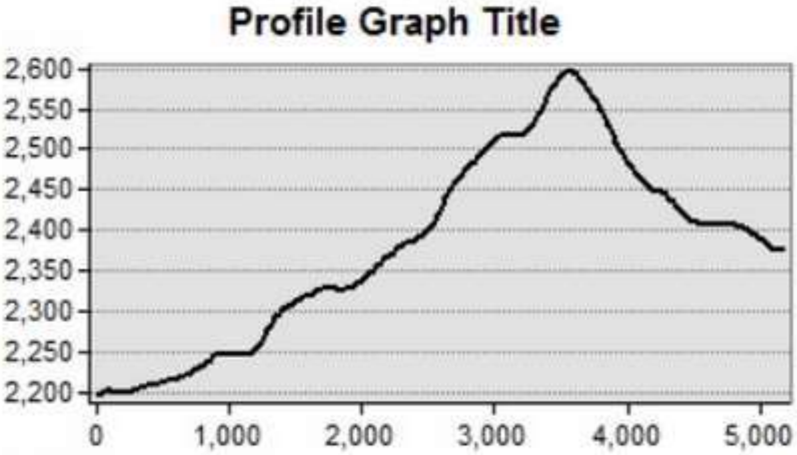
Terrain Mapping

- There are different kinds of terrain mapping
- Contouring is the most common for the terrain mapping it has contour interval and base contour
- The arrangement and pattern of contour lines reflect the topography
- Contour lines don't intersect one an other and will not stop in the middle of the map



Vertical Profiling

- A vertical profile shows changes in elevation along a line such as a road or a stream



Hill Shading

- Hill shading simulates how the terrain looks with interaction between sunlight and surface features
- A mountain slope directly facing incoming light will be very bright
- A slope opposite to it is will be dark.
- Four factors control the visual effects of hill shading

- The sun's azimuth is the direction of the incoming light ranging from 0 (due north to 360 in a clockwise direction
- Typically the default for the sun's azimuth is 315 with
- The sun's altitude is the angle of the incoming light measured above the horizon between 0 and 90 degree

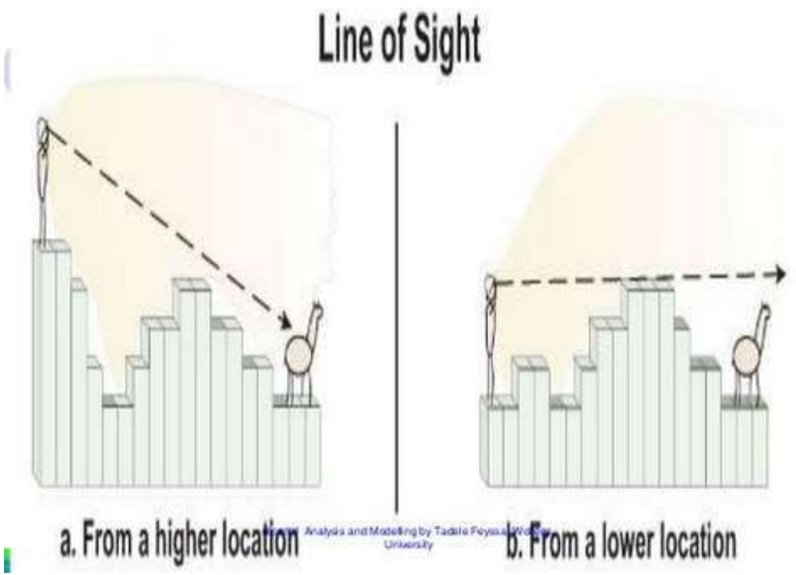
- The other two factors are slope and aspect
- Slope ranges between 0 and 90 degrees
- Aspect ranges between 0 and 360 degrees
- Slope measures the rate of change of elevation at a surface location
- It can be expressed in percent or degree
- Aspect is the directional measures slope

- Aspect starts with 0 degree at North pole and measures clockwise and ends with 360 degree

Viewshed Analysis

- Viewshed refers to the portion of the land surface that is visible from one or more viewpoints
- The process for deriving viesheds is called viewshed or visibility analysis
- A veiwshed analysis requires two input datasets
- The first is usually point layer containing one or more view points such as a layer containing communication tower
- The second input is DEM (an elevation raster or a TIN which represents the land surface

- The line-of-sight operation is the basis for veiwshed analysis
- The line of sight also called sight-line, connects the viewpoint and the target
- GIS can display a sightline with symbols for the visible and invisible portions a long the sightline



Application areas of viewshed

- Urban Planning
- Cell phone tower placement
- Location for wind turbines
- Conservation projects
- Military purpose
- Many more!

Watershed Analysis

- A watershed refers to an area defined by topographic divides that drains surface water to a common outlet
- A watershed is often used as a unit area for the management and planning of water and other resources
- Watershed analysis refers to the process of using DEMs and following water flow to delineate stream networks and watersheds

- Traditionally watershed boundary can be drawn manually onto a topographic map
- The person who draws the boundaries uses topographic features on the map to determine where a divide is located
- Today computer based watershed analysis can do this job in a fraction of time
- Delineation of watershed can tale place at different spatial scales

Application of water shed analysis

- Water shed analysis is mainly used for natural resource management
- Floodplain Management
- Land Use Planning
- Invasive Species
- River Function
- Water Supply

Unit V

Implementing a GIS:

Certainly, here are some key points to consider regarding awareness on GIS (Geographic Information System):

1. **Definition:**
   - GIS stands for Geographic Information System, a technology that integrates spatial data (maps) with attribute data to analyze, interpret, and visualize geographical patterns and relationships.
2. **Components of GIS:**
   - **Spatial Data:** Information related to the location and shape of features on the Earth's surface.
   - **Attribute Data:** Non-spatial information associated with the spatial features.
   - **Software:** Tools and applications used for creating, managing, analyzing, and visualizing spatial data.
   - **Hardware:** Equipment and devices required to run GIS software effectively.
3. **Applications of GIS:**
   - GIS is utilized in various fields, including urban planning, environmental management, natural resource exploration, agriculture, disaster management, epidemiology, and more.
   - It helps in decision-making processes by providing spatial insights and aiding in data-driven analysis.
4. **Data Layers:**
   - GIS organizes information into layers, allowing users to overlay and analyze different datasets simultaneously. This layering system enhances the understanding of spatial relationships.
5. **Mapping and Visualization:**
   - GIS enables the creation of maps and visual representations of data. This helps in conveying complex spatial information in a comprehensible manner.
6. **Spatial Analysis:**
   - GIS allows for sophisticated spatial analysis, such as proximity analysis, spatial statistics, and network analysis. This helps in deriving meaningful patterns and trends from spatial data.
7. **Data Collection Methods:**
   - GIS data can be collected through various methods, including satellite imagery, GPS surveys, remote sensing, and digitization of existing maps.
8. **Role in Decision Making:**
   - GIS facilitates better decision-making by providing a spatial context to information. It allows users to explore relationships, identify patterns, and make informed choices based on geographical data.
9. **Open Source GIS:**
   - Apart from commercial GIS software, there are open-source alternatives like QGIS that provide powerful GIS capabilities. These tools contribute to making GIS accessible to a broader audience.

10. **Challenges and Considerations:**
    - Challenges in GIS include data accuracy, interoperability, and the need for skilled professionals. Addressing these challenges is crucial for successful GIS implementation.
11. **Career Opportunities:**
    - With the growing importance of spatial data, there is an increasing demand for GIS professionals. Careers in GIS include GIS analysts, cartographers, remote sensing specialists, and GIS developers.
12. **Environmental and Social Impact:**
    - GIS plays a significant role in addressing environmental issues, managing natural resources, and understanding social patterns. It contributes to sustainable development and informed decision-making.

Creating awareness about GIS involves highlighting its versatility, impact, and potential benefits across various sectors. As GIS continues to evolve, staying informed about its capabilities and applications becomes essential for individuals and organizations.

GIS based Road network Planning:

Geographic Information System (GIS) applications play a crucial role in road network planning by providing tools to analyze, manage, and visualize spatial data. Here are several key ways GIS is applied in road network planning:

1. **Site Selection and Feasibility Studies:**
    - GIS aids in identifying suitable locations for new roads or expansions by considering factors such as topography, land use, environmental constraints, and existing infrastructure. Feasibility studies benefit from spatial analysis to assess the impact and cost-effectiveness of potential road projects.
2. **Traffic Analysis and Management:**
    - GIS is used to model and analyze traffic patterns, predict congestion, and optimize road capacity. By integrating real-time traffic data, GIS helps in developing efficient traffic management strategies, including signal timings and route planning.
3. **Terrain and Geotechnical Analysis:**
    - GIS tools assist in evaluating the terrain and geotechnical characteristics of the area where roads are planned. This includes analyzing slope stability, soil types, and geological features to ensure safe and cost-effective road construction.
4. **Environmental Impact Assessment:**
    - GIS is employed to assess the environmental impact of road projects. It helps in identifying sensitive ecological areas, water bodies, and habitats that need to be preserved or mitigated. This aids in sustainable road planning and design.
5. **Land Use and Zoning:**
    - GIS integrates land use and zoning information to understand how road networks interact with different types of developments. This knowledge is crucial for optimizing road layouts, minimizing disruptions, and supporting urban planning initiatives.
6. **Infrastructure Inventory and Asset Management:**

- GIS is utilized to create comprehensive inventories of existing road infrastructure. This includes mapping road segments, bridges, intersections, and other elements. Asset management through GIS ensures efficient maintenance and upgrades.

7. **Emergency Response Planning:**
   - GIS supports emergency response planning by providing spatial information on road networks. This helps authorities identify evacuation routes, emergency service access points, and areas prone to natural disasters.
8. **Public Engagement and Communication:**
   - GIS enables effective communication with the public and stakeholders by creating interactive maps and visualizations. This helps in conveying complex information about road projects, fostering transparency, and gathering feedback.
9. **Cost-Benefit Analysis:**
   - GIS aids in performing cost-benefit analysis for road projects by integrating various spatial factors. This includes evaluating the economic impact, travel time savings, and environmental benefits of proposed road improvements.
10. **Maintenance Planning:**
    - GIS assists in planning and scheduling routine maintenance activities by providing insights into the condition of roads and infrastructure. This proactive approach ensures the longevity and safety of the road network.
11. **Integration with Other Planning Systems:**
    - GIS integrates with other planning systems, such as urban planning and transportation planning, to create a holistic view of the built environment. This collaborative approach enhances the effectiveness of road network planning.

In summary, GIS applications in road network planning contribute to informed decision-making, efficient resource allocation, and the development of sustainable and resilient transportation infrastructure.

Mineral Mapping using GIS:

Geographic Information System (GIS) applications play a crucial role in mineral mapping, providing a powerful tool for the exploration, management, and analysis of mineral resources. Here are some key ways GIS is applied in mineral mapping:

1. **Satellite Remote Sensing:**
   - GIS is used to process and analyze satellite imagery for mineral exploration. Remote sensing data can identify spectral signatures associated with different minerals, helping in the detection and mapping of mineral deposits.
2. **Geological Mapping:**
   - GIS enables the creation of detailed geological maps by integrating geological data such as rock types, structures, and stratigraphy. This aids geologists in understanding the spatial distribution of minerals and identifying potential mineral-rich areas.
3. **Topographic Analysis:**
   - GIS tools are utilized for topographic analysis to identify terrain features and landforms associated with mineralization. Understanding the landscape helps in narrowing down potential areas for exploration.

4. **Mineral Prospectivity Mapping:**
   - GIS is employed to generate mineral prospectivity maps by integrating various geological, geochemical, and geophysical datasets. These maps help in prioritizing areas with higher potential for mineral deposits.
5. **Geochemical Analysis:**
   - GIS is used to manage and analyze geochemical data, including soil samples, rock samples, and water samples. By mapping and spatially analyzing geochemical anomalies, geologists can identify areas with elevated mineral concentrations.
6. **Geophysical Data Integration:**
   - GIS integrates geophysical data, such as magnetic and gravity surveys, to identify subsurface structures and anomalies associated with mineral deposits. This aids in refining exploration targets.
7. **3D Modeling:**
   - GIS facilitates the creation of three-dimensional geological models that represent the subsurface distribution of minerals. This aids in visualizing the spatial relationships between different geological features.
8. **Land Use Planning:**
   - GIS is used to assess the compatibility of mineral exploration or mining activities with existing land use plans. This involves analyzing environmental and social factors to minimize the impact on surrounding areas.
9. **Environmental Impact Assessment (EIA):**
   - GIS supports the EIA process by providing spatial information on potential environmental impacts of mineral extraction. This includes mapping sensitive ecosystems, water bodies, and areas of cultural significance.
10. **Infrastructure Planning:**
    - GIS helps in planning the infrastructure required for mineral extraction, such as roads, power lines, and water supply. This ensures efficient development and logistics for mining operations.
11. **Land Management and Permitting:**
    - GIS is utilized in land management by mapping land ownership, lease agreements, and permitting requirements. This aids in complying with regulations and managing the legal aspects of mineral exploration and mining.
12. **Monitoring and Reclamation:**
    - GIS is applied in monitoring the progress of mining activities and assessing the success of reclamation efforts. This helps in maintaining environmental sustainability and meeting regulatory requirements.

In summary, GIS applications in mineral mapping provide a comprehensive and spatially informed approach to mineral exploration, helping in the discovery, assessment, and sustainable management of mineral resources.

Shortest path detection using GIS

Detecting the shortest path using Geographic Information System (GIS) involves leveraging spatial data and network analysis tools. Here is a general outline of the steps involved in shortest path detection using GIS:

1. **Data Preparation:**
    - Acquire or create spatial data representing the road network. This may include information on roads, intersections, and their attributes (such as speed limits).
    - Ensure the data is in a suitable GIS format (e.g., shapefiles or geodatabases) and is properly georeferenced.
2. **Network Dataset Creation:**
    - In GIS software (e.g., ArcGIS, QGIS), create a network dataset from the road network data. This involves defining connectivity rules, one-way streets, turn restrictions, and assigning travel costs based on attributes like road length or speed limits.
3. **Define Source and Destination Points:**
    - Specify the starting point (source) and the destination point for which you want to find the shortest path. These points should be located on the road network.
4. **Network Analysis Setup:**
    - Use the network analysis tools in your GIS software to set up a network analysis layer. Define parameters such as impedance (travel cost), analysis type (shortest path), and any additional constraints (e.g., avoiding certain roads or areas).
5. **Perform Network Analysis:**
    - Run the network analysis tool to find the shortest path between the specified source and destination points. The GIS software will use algorithms such as Dijkstra's or A* to calculate the optimal route based on the defined parameters.
6. **Visualization:**
    - Visualize the results on a map to see the shortest path. The output may include the route geometry, directions, and any relevant information about the path, such as travel time or distance.
7. **Analysis Refinement:**
    - Depending on the GIS software and tools used, you may have options to refine the analysis further. This could include considering traffic conditions, time of day, or optimizing for multiple objectives (e.g., shortest time vs. shortest distance).
8. **Output and Reporting:**
    - Generate reports or export the results for further analysis or presentation. This could include creating printable maps, sharing digital files, or integrating the results into other applications.
9. **Iterative Analysis:**
    - Shortest path detection can be an iterative process, especially in dynamic environments. Periodically update the network data and re-run analyses to account for changes in the road network or other factors.

Hazard Zonation using Remote Sensing and GIS

Hazard zonation using remote sensing and Geographic Information System (GIS) involves the application of satellite imagery and spatial analysis tools to identify and map areas prone to natural hazards. Here's a concise explanation of the process:

1. **Objective Definition:**
    - Clearly define the natural hazard of interest, whether it's earthquakes, landslides, floods, or another type.
2. **Data Collection:**
    - Gather relevant geospatial data, including satellite imagery, digital elevation models (DEMs), land cover data, geological maps, and other pertinent information.
3. **Preprocessing:**
    - Preprocess satellite imagery to enhance its quality, correct for atmospheric conditions, and ensure accurate spatial representation.
4. **Topographic Analysis:**
    - Use DEM data to analyze topographic features such as slope, aspect, and elevation. Steep slopes, low-lying areas, or specific slope orientations may be more susceptible to certain hazards.
5. **Land Cover Classification:**
    - Employ satellite imagery and machine learning techniques to classify land cover types. Different land cover characteristics influence the susceptibility to various hazards.
6. **Geological and Seismic Data Integration:**
    - Integrate geological maps and seismic data to identify fault lines, seismic activity zones, or geological features contributing to hazards.
7. **Vulnerability Assessment:**
    - Assess the vulnerability of structures and infrastructure in the study area, considering factors such as building types and land use.
8. **GIS Analysis:**
    - Use GIS tools to overlay and analyze data layers, combining information on topography, land cover, geology, and vulnerability to identify areas at varying levels of risk.
9. **Zonation Map Creation:**
    - Generate hazard zonation maps that categorize the study area into zones based on the level of risk or susceptibility to the identified hazard.
10. **Validation and Calibration:**
    - Validate the hazard zonation map against historical hazard events. If needed, calibrate the model based on observed data to improve accuracy.
11. **Documentation and Reporting:**
    - Document the methodology, data sources, and results. Create reports and visualizations to effectively communicate hazard zonation to stakeholders, emergency responders, and the public.
12. **Periodic Updating:**
    - Regularly update the hazard zonation analysis with new data to account for changes in the landscape and maintain the accuracy of the risk assessment.

Hazard zonation using remote sensing and GIS provides a systematic and spatially informed approach to understanding and mitigating the impact of natural hazards in a given area.

GIS for solving multi criteria problems

Geographic Information System (GIS) is a powerful tool for solving multi-criteria problems, where decision-making involves considering multiple factors or criteria simultaneously. Here's an explanation of how GIS can be used for addressing multi-criteria problems:

1. **Data Integration:**
    - GIS allows the integration of diverse spatial datasets into a unified platform. These datasets can represent various criteria such as land use, elevation, population density, infrastructure, and environmental variables.
2. **Layer Overlay Analysis:**
    - One of the fundamental capabilities of GIS is layer overlay analysis. This involves overlaying different map layers to identify areas where multiple criteria intersect. For example, determining suitable locations for a new facility by considering factors like proximity to roads, avoiding environmentally sensitive areas, and being close to population centers.
3. **Weighted Overlay Analysis:**
    - GIS enables the assignment of weights to different criteria based on their relative importance. Weighted overlay analysis combines these weighted layers to produce a composite map that reflects the overall suitability or desirability of locations.
4. **Spatial Analysis Tools:**
    - GIS provides a variety of spatial analysis tools for assessing relationships, proximity, connectivity, and other spatial patterns. These tools help in understanding how different criteria interact and influence each other.
5. **Decision Support Systems:**
    - GIS-based decision support systems (DSS) are designed to assist decision-makers in evaluating alternative solutions considering multiple criteria. These systems often include optimization algorithms and modeling tools to find the best compromise among conflicting criteria.
6. **Scenario Analysis:**
    - GIS allows the creation and analysis of different scenarios by manipulating criteria layers. Decision-makers can explore the implications of changes in criteria weights or introduce new criteria to understand their impact on the overall decision.
7. **Spatial Statistics:**
    - GIS incorporates spatial statistical techniques to analyze the spatial distribution and patterns of criteria. This is particularly useful for identifying clusters, hotspots, or trends that may influence decision outcomes.
8. **Multi-Criteria Decision Analysis (MCDA):**
    - MCDA is a systematic approach facilitated by GIS to evaluate and compare alternative solutions based on multiple criteria. GIS tools help in aggregating, normalizing, and synthesizing diverse data to support decision-making.
9. **Risk Assessment:**
    - GIS is employed for risk assessment by considering various criteria contributing to potential risks. This can include analyzing factors such as proximity to natural hazards, vulnerability, and exposure to create risk maps.

10. **Environmental Impact Assessment (EIA):**
    - In scenarios where decisions have environmental implications, GIS is widely used for conducting EIAs. It helps in evaluating and mitigating the impact of projects by considering a range of environmental criteria.
11. **Visualization:**
    - GIS enables the creation of visualizations and maps that effectively communicate the spatial relationships among different criteria, making it easier for stakeholders to comprehend complex decision-making scenarios.

GIS for business applications

Geographic Information System (GIS) finds diverse applications in the business domain, offering spatial insights to enhance decision-making and operational efficiency. Here's a brief explanation of GIS for business applications:

1. **Location Intelligence:**
    - GIS enables businesses to analyze and visualize data based on geographic location. This location intelligence helps in understanding spatial patterns, market trends, and customer behavior.
2. **Site Selection:**
    - Businesses use GIS for optimal site selection by considering factors such as proximity to customers, competitors, transportation networks, and demographic characteristics.
3. **Supply Chain Management:**
    - GIS aids in optimizing supply chain routes, warehouse locations, and distribution networks. Businesses can analyze transportation costs, delivery times, and inventory management based on spatial data.
4. **Market Analysis:**
    - GIS helps businesses conduct market analysis by mapping customer locations, identifying target demographics, and assessing market potential. This supports strategic marketing and sales efforts.
5. **Customer Relationship Management (CRM):**
    - Integrating GIS with CRM systems allows businesses to visualize customer data on maps. This spatial perspective aids in understanding customer locations, preferences, and tailoring marketing strategies accordingly.
6. **Competitor Analysis:**
    - GIS facilitates competitor analysis by mapping competitors' locations and assessing their impact on the market. This information assists businesses in strategic planning and differentiation.
7. **Risk Management:**
    - Businesses use GIS for risk management by mapping and analyzing potential risks such as natural disasters, supply chain disruptions, or geopolitical factors. This allows for proactive risk mitigation strategies.

8. **Facility Management:**
    - GIS assists in managing facilities by mapping infrastructure, tracking maintenance needs, and optimizing space utilization. This is valuable for industries like real estate, utilities, and retail.
9. **Field Service Optimization:**
    - For businesses with field service operations, GIS helps optimize routes, assign tasks based on location, and improve overall efficiency in managing on-site activities.
10. **Environmental Impact Assessment:**
    - Industries can use GIS for assessing and mitigating environmental impacts. It aids in compliance with regulations and demonstrates a commitment to sustainable business practices.
11. **Network Planning:**
    - GIS supports network planning for telecommunications, utilities, and transportation industries. It helps optimize the layout of networks, plan for expansions, and enhance infrastructure resilience.
12. **Data Visualization and Reporting:**
    - GIS provides businesses with powerful tools for visualizing data on maps, making it easier to communicate complex information. Dashboards and reports with spatial insights aid in decision-making.

In summary, GIS plays a crucial role in business applications by providing a spatial perspective that enhances decision-making processes, improves operational efficiency, and contributes to strategic planning across various industries.