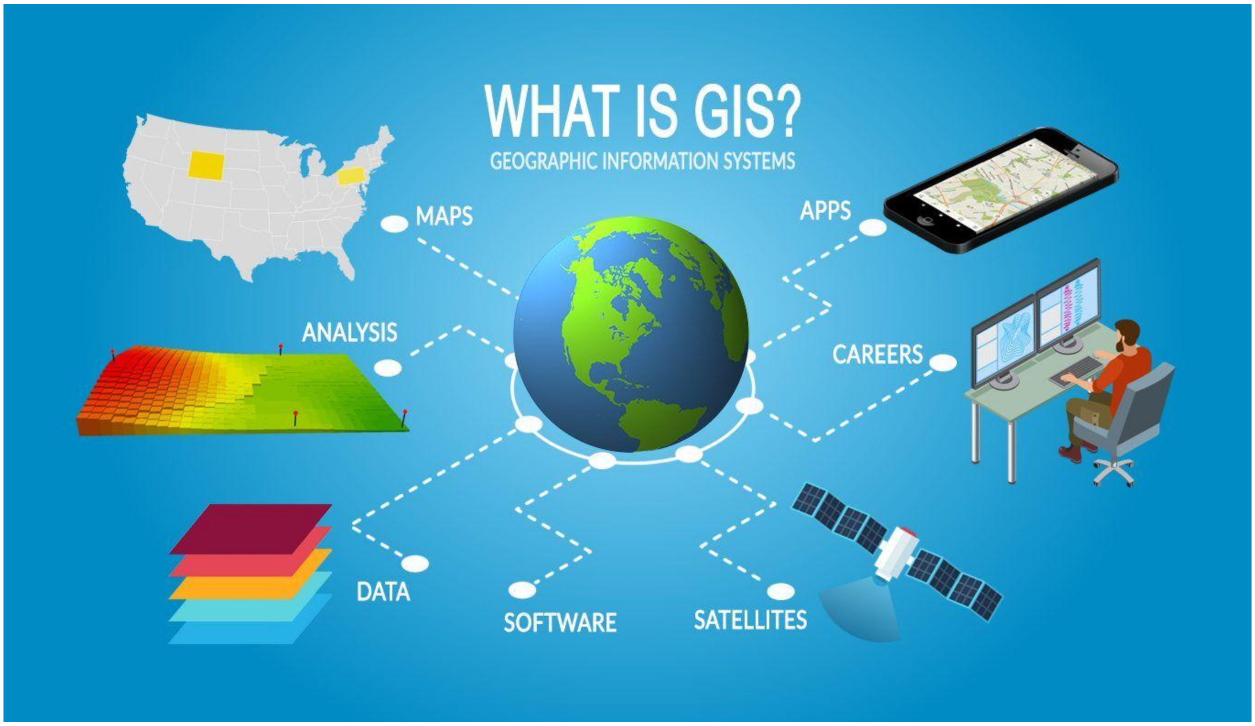
Unit1: Introduction

GIS (Geographic Information System)

GIS stands for Geographic Information System, which is a computer-based system used to capture, store, analyze, manage and present geographical data. It integrates different types of data such as satellite imagery, maps, and other spatial data to provide a comprehensive view of the environment. GIS software enables users to create maps, perform spatial analysis, and identify patterns and relationships in data related to geography. GIS has a wide range of applications in various fields such as urban planning, natural resource management, public health, and emergency response.



History of GIS

GIS, which stands for Geographic Information System, is a technology that allows for the collection, analysis, and visualization of spatial data. The history of GIS can be traced back to the 1960s, when researchers in Canada and the United States began developing computer-based systems to manage and analyze spatial data.

One of the earliest examples of GIS was the Canada Geographic Information System (CGIS), which was developed by the Canadian government in the early

1960s. CGIS was designed to manage land use and resource management data, and it used a mainframe computer to store and analyze the data.

In the 1970s, the development of microcomputers and the growth of the environmental movement led to an increased interest in GIS. During this time, a number of commercial GIS software packages were developed, including ARC/INFO, GRASS, and MapInfo.

The 1980s saw the development of more advanced GIS technologies, including the integration of remote sensing data into GIS systems. In addition, the development of GPS (Global Positioning System) technology made it possible to collect precise location data, which could be integrated into GIS systems.

In the 1990s, the growth of the Internet and the development of web-based GIS technologies made it possible to share and access GIS data from anywhere in the world. This led to the development of online mapping services like Google Maps and Bing Maps, which made GIS technology more accessible to the general public.

Today, GIS technology is used in a wide range of fields, including urban planning, natural resource management, transportation planning, and emergency management. The development of new GIS technologies, including 3D modeling and virtual reality, continues to push the boundaries of what is possible with spatial data.

Functions of GIS

1. **Data Capture**

• Data used in GIS often come from many different sources, are of many types, and are stored in different ways.

• GIS provides tools and a method for the integration of different data into a format to be compared and analysed.

• Data sources are mainly manual digitization/scanning of aerial photographs, paper maps, and existing digital datasets. Remote-sensing satellite imagery and GPS are promising data input sources for GIS.

2. Database Management and Update

• After data are collected and integrated, the GIS must provide facilities which can contain and maintain data.

• Effective data management has many definitions but should include all of the following aspects:

– data security

– data integrity

– data storage

– retrieval, and

– data maintenance abilities.

3. Geographic Analysis

• Data integration and conversion are only a part of the input phase of GIS. What is required next is the ability to interpret and to analyse the collected information quantitatively and qualitatively.

For example, a satellite image can assist an agricultural scientist to project crop yield per hectare for a particular region. For the same region, the scientist also has the rainfall data for the past six months collected through weather station observations. The scientist also has a map of the soils for the region which shows fertility and suitability for agriculture. These point data can be interpolated and what you get is a thematic map showing iso-heights or contour lines of rainfall.

4. Presenting Results

- One of the most exciting aspects of GIS technology is the variety of different ways in which the information can be presented once it has been processed by GIS.
- Traditional methods of tabulating and graphing data can be supplemented by maps and three dimensional images.
- Visual communication is one of the most fascinating aspects of GIS technology and is available in a diverse range of output options.

Applications of GIS

GIS has a wide range of applications in various fields, some of which are:

1. Urban Planning: GIS is used to analyze and manage urban infrastructure, such as transportation networks, housing, and public utilities. It helps in making better decisions for land use planning, zoning, and transportation planning.
2. Natural Resource Management: GIS is used to manage and monitor natural resources such as forests, wildlife, and water resources. It helps in analyzing and mapping ecological systems and identifying areas of conservation and biodiversity.
3. Emergency Management: GIS is used in emergency management for disaster response and recovery. It helps in locating emergency resources, identifying hazards, and developing evacuation plans.
4. Public Health: GIS is used to track and monitor the spread of diseases, analyze health data, and identify high-risk areas. It also helps in identifying access to healthcare facilities and planning for public health interventions.
5. Agriculture: GIS is used in agriculture for precision farming and crop management. It helps in analyzing soil characteristics, identifying suitable areas for crop growth, and monitoring crop growth.
6. Business: GIS is used in business for market analysis, customer segmentation, and site selection. It helps in identifying potential locations for businesses, analyzing demographic data, and understanding consumer behavior.
7. Archaeology: GIS is used in archaeology to map and analyze archaeological sites and landscapes. It helps in identifying areas of historical significance and cultural heritage.

Overall, GIS has a wide range of applications in many fields, and its usefulness continues to grow as more data becomes available and new analytical techniques are developed.

Components of GIS

The major components of GIS include:

1. Hardware: The physical components of a GIS system, including computers, servers, printers, scanners, and other peripherals.
2. Software: The software components of a GIS system, including GIS applications, database management systems, and other software tools.
3. Data: The spatial or geographic data that are collected, stored, and managed in a GIS system. This can include data from maps, aerial photographs, satellite imagery, GPS surveys, and other sources.
4. People: The individuals, who design, build, operate, and use GIS systems. This can include GIS professionals, data analysts, cartographers, and other experts who work with spatial data.
5. Procedures: The documented methods and workflows that are used to collect, process, analyze, and present spatial data in a GIS system. This can include data collection protocols, database design standards, and analytical workflows.
6. Standards: The technical specifications and guidelines that govern the structure, content, and exchange of spatial data in a GIS system. This can include data formatting standards, metadata standards, and data exchange protocols.

Data

Data refers to any collection of information, facts, or statistics that can be analyzed and used to gain insights, make decisions, or solve problems. This information can be in various forms such as text, numbers, images, videos, or sound recordings. Data can be either structured or unstructured, and it can be processed by computers or humans.

In today's digital age, data is everywhere, and it's generated at an unprecedented rate by various sources such as social media, internet search engines, sensors, and other devices. Data is essential in fields such as business, finance, healthcare, education, and many others, as it allows organizations to gain valuable insights and make informed decisions.

However, managing and analyzing large amounts of data can be a complex and challenging task, requiring specialized skills, tools, and techniques. Data scientists, analysts, and engineers use various methods such as data mining, machine learning, and artificial intelligence to extract insights and knowledge from data.

Information

Information refers to the processed or interpreted form of data that is meaningful and useful to the recipient. It is the knowledge or understanding gained from analyzing, organizing, and interpreting data in a way that makes it meaningful and useful.

Information can be presented in various formats, including text, graphics, audio, and video, and it can be communicated through various channels such as books, newspapers, TV, social media, and the internet.

Unlike data, information has a context and a purpose, and it can be used to make decisions, solve problems, or gain insights. For example, a list of customer purchases is data, but when analyzed and presented in a report that shows which products are selling well and which are not, it becomes information that can be used to make business decisions.

Information is essential in many fields, including business, science, education, and healthcare, as it allows us to gain knowledge, solve problems, and make informed decisions.

GIS as Information System

GIS (Geographic Information System) is a type of information system that is designed to capture, store, manipulate, analyze, and display geospatial data. In other words, GIS is a system that allows users to work with spatial data, which is data that has a location or geographic component.

GIS can be thought of as a digital map that contains layers of information. Each layer represents a different type of data, such as roads, buildings, vegetation, or water bodies. By combining and analyzing these layers, GIS can provide valuable insights into spatial relationships and patterns that may not be visible from other types of data.

GIS is designed to handle a variety of data formats, including raster and vector data. It can also work with various types of spatial data, such as GPS data, satellite imagery, and aerial photographs.

As an information system, GIS provides a wide range of tools and capabilities for data management, analysis, and visualization. Some of these tools include spatial query and analysis, geocoding, routing, and spatial modeling.

GIS is used in various fields, such as urban planning, natural resource management, emergency management, and health care, to name a few. With its ability to integrate and analyze large amounts of geospatial data, GIS has become an essential tool for decision-making and planning in many industries

Nature and Sources of GIS Data

GIS (Geographic Information System) data can be classified into two main categories based on their nature and sources:

1. Primary Data: Primary data is the data that is collected specifically for GIS purposes. It is collected using various methods such as GPS (Global Positioning System) survey, remote sensing, and ground surveys. Examples of primary data include land use/cover data, topographic data, soil data, and vegetation data.
2. Secondary Data: Secondary data is the data that is obtained from existing sources. It includes a variety of data types such as census data, road networks, administrative boundaries, and climate data. Secondary data is usually obtained from government agencies, private companies, and research institutions.

The sources of GIS data can be classified into the following categories:

1. Government Agencies: Government agencies such as the United States Geological Survey (USGS), National Aeronautics and Space Administration (NASA), and National Oceanic and Atmospheric Administration (NOAA) are some of the primary sources of GIS data. They provide various types of data such as aerial photographs, satellite images, and topographic maps.
2. Private Companies: Private companies are another source of GIS data. They provide various types of data such as demographic data, traffic data, and real estate data. Examples of private companies that provide GIS data include ESRI, Google, and TomTom.
3. Research Institutions: Research institutions such as universities and non-profit organizations are another source of GIS data. They provide data related to various fields such as environmental data, socio-economic data, and health data.
4. Open Data Sources: There are also several open data sources that provide free GIS data. Examples of open data sources include OpenStreetMap, NASA Earth Observations, and the World Bank.

The nature and sources of GIS data are diverse, and the selection of data sources depends on the application and the scope of the GIS project. It is important to evaluate the quality and accuracy of the data before using it in GIS applications.

Recent trends and future of GIS

Recent trends in GIS (Geographic Information System) technology include:

1. Cloud-based GIS: Cloud-based GIS allows users to access and share data and applications from anywhere, on any device, with reduced hardware and software requirements.
2. Big Data and Machine Learning: The growing volume and diversity of spatial data require new approaches to data analysis and visualization, including the use of machine learning and big data analytics.
3. Mobile GIS: Mobile GIS has become increasingly important, allowing users to collect and analyze data in real-time using mobile devices.
4. Open Data and Open Source GIS: Open data initiatives and open-source GIS software have become popular, allowing for greater collaboration and data sharing across organizations and sectors.
5. 3D GIS and Virtual Reality: 3D GIS and virtual reality technologies are becoming more common, allowing users to create realistic representations of the real world for a wide range of applications.

The future of GIS is expected to continue to grow and evolve. Some possible future developments in GIS technology include:

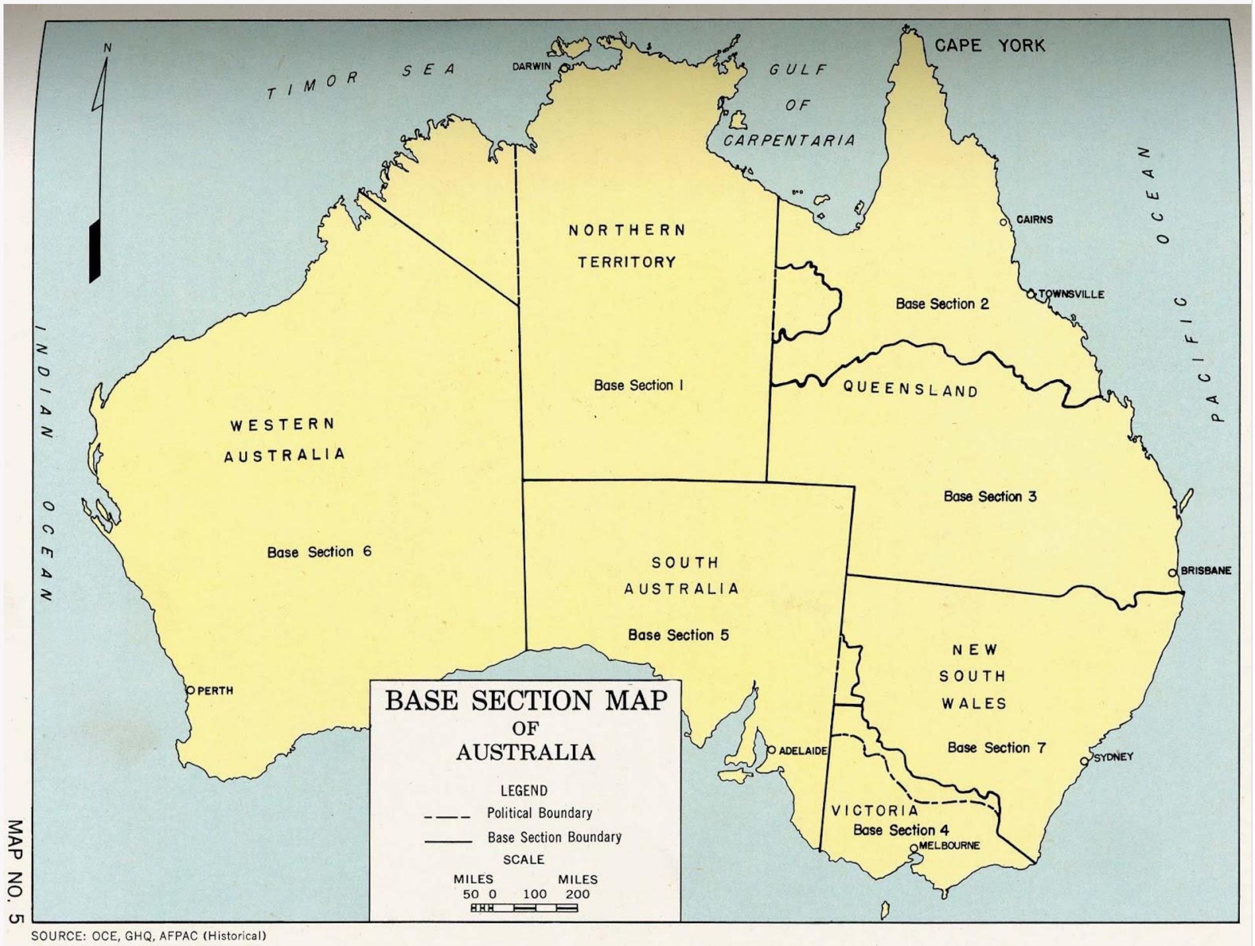
1. Greater use of artificial intelligence and machine learning to analyze and model spatial data.
2. Increased use of real-time data and sensors, such as GPS and IoT devices, to support real-time decision-making.
3. More advanced 3D visualization and virtual reality technologies, allowing users to interact with spatial data in new and innovative ways.
4. Greater integration with other technologies, such as augmented reality, blocks chain and edge computing.
5. Continued growth in cloud-based GIS, allowing for greater collaboration and data sharing across organizations and sectors.

Overall, the future of GIS looks promising, with continued advancements in technology and an ever-growing demand for spatial data and analysis.

Unit 2: GIS and Map

Map and Characteristics(Elements)

A map is a visual representation of the spatial relationships between different features or phenomena. Maps can be classified into different types based on their purpose, scale and content. The following are some of the characteristics of maps:



1. Scale: Maps represent a scaled-down version of the real world. The scale of a map is the ratio between the distance on the map and the corresponding distance on the ground. Maps can be categorized into small-scale maps, medium-scale maps, and large-scale maps, depending on their level of detail.
2. Projection: Maps are created by projecting the three-dimensional surface of the earth onto a two-dimensional surface. Different map projections are used to minimize distortions in size, shape and distance.
3. Symbols: Maps use symbols to represent different features or phenomena. Symbols can include points, lines, and areas, as well as colors, shading, and patterns.

4. Legend: A legend is a key or guide that explains the symbols and colors used on the map.
5. Orientation: Maps have an orientation that shows the direction of north, south, east, and west.
6. Datum: The datum is the reference surface used to measure elevation and coordinate locations on a map.
7. Cartography: The art and science of map-making are called cartography. Cartography involves the design and production of maps, as well as the selection of appropriate scales, projections, and symbols.
8. Purpose: Maps are created for different purposes, such as navigation, land use planning, environmental analysis, and demographic analysis.

Maps are a powerful tool for communicating spatial information, and they play a critical role in decision-making and planning in various fields. Understanding the characteristics of maps is essential for interpreting and using them effectively.

Types of Map

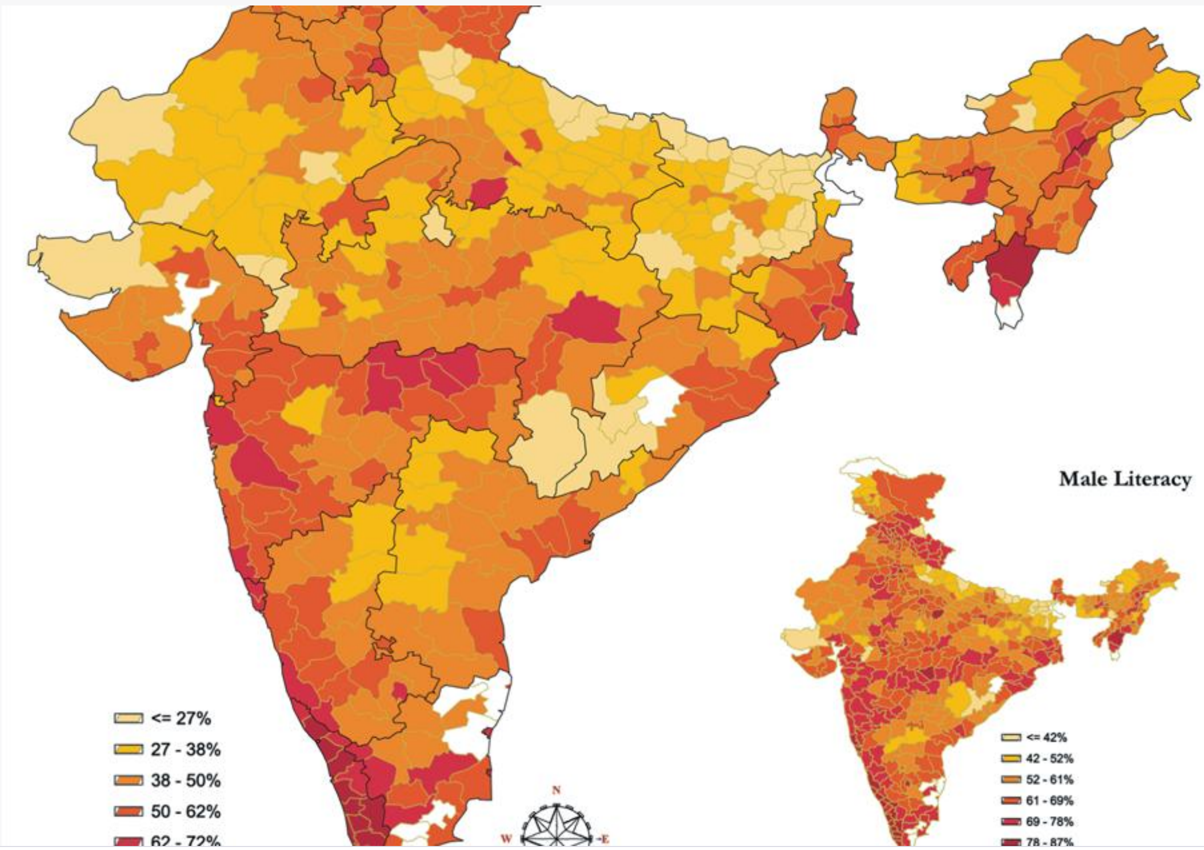
Maps can be classified into several different categories based on their purpose, scale, content, and other characteristics. Here are some general classifications of maps:

1. General reference maps: These maps provide general information about an area, such as political boundaries, roads, rivers, and physical features. They are typically small-scale maps and are used for general reference and navigation.
2. Thematic maps: These maps focus on a specific theme or topic, such as population density, land use, or climate. They are designed to show patterns and relationships in the data.
3. Topographic maps: These maps show the contours and elevations of an area, along with other physical features such as rivers, lakes, and mountains. They are used for outdoor recreation, planning, and resource management.
4. Navigational maps: These maps are used for navigation and include road maps, aviation charts, and nautical charts. They show information such as roads, airports, seaports and other navigational aids.
5. Geologic maps: These maps show the geologic features and formations of an area, including rock types, faults, and other geological features.
6. Land-use maps: These maps show how land is used in an area, such as urban, agricultural, or natural areas. They are used for planning and resource management.

7. Climate maps: These maps show the climate and weather patterns of an area, including temperature, precipitation and wind patterns. They are used for planning and resource management.
8. Historical maps: These maps show the historical changes in an area over time, including changes in borders, land use, and development. They are used for research and historical analysis.

<u>Types of thematic map</u>

Thematic maps are designed to show specific themes or topics, such as population density, land use, or climate. There are several types of thematic maps that are used in different fields. Here are some of the most common types of thematic maps:



1. Choropleth maps: These maps use color shading or patterns to show the distribution or density of a particular theme or topic over an area, such as population density or voting patterns.
2. Dot density maps: These maps use dots to represent the quantity or density of a particular theme or topic over an area, such as the number of people per square mile or the number of accidents in a particular area.

3. Isoline maps: These maps use lines to connect points of equal value, such as elevation or temperature, to show spatial patterns of the theme or topic.
4. Cartograms: These maps use different shapes and sizes to represent the relative values or quantities of a particular theme or topic over an area, such as population or economic activity.
5. Flow maps: These maps show the movement or direction of a particular theme or topic, such as migration or transportation routes, using lines or arrows.
6. 3D maps: These maps use elevation data and other techniques to create a 3D representation of a particular theme or topic, such as terrain or building heights

Maps types on the basis of scale

Maps can be classified into different types based on their scale, which refers to the ratio of the map's size to the size of the area it represents. The three main types of maps based on scale are:

1. Small scale maps: Small scale maps show a large area, such as a country or a continent, but with less detail. These maps are typically used to show relationships between different regions or to provide an overview of a large area. Examples of small scale maps include world maps, country maps, and continent maps.
2. Medium scale maps: Medium scale maps show a smaller area, such as a city or a region, with more detail than small scale maps. These maps are typically used for planning and navigation, and they can show features such as streets, buildings, and topography. Examples of medium scale maps include city maps, state maps, and regional maps.
3. Large scale maps: Large scale maps show a very small area, such as a neighborhood or a building, with a high level of detail. These maps are typically used for detailed planning and analysis, and they can show features such as individual buildings, property lines, and infrastructure. Examples of large scale maps include site plans, floor plans, and zoning maps.

The scale of a map is important because it affects the level of detail that can be shown and the accuracy of the information presented. For example, a small scale map may not show individual streets or buildings, but it can provide a broad overview of a large area. A large scale map, on the other hand, can show detailed features but may not be suitable for showing relationships between different regions or areas.

Advantages of Map

Maps have a variety of advantages, some of which include:

1. Visual Representation: Maps provide a visual representation of data that is easy to interpret and understand. By using colors, symbols, and labels, maps can convey complex information in a simple, intuitive way.
2. Spatial Analysis: Maps allow for spatial analysis, which means that data can be analyzed based on its location. This is particularly useful for understanding patterns and trends in data that are related to geography, such as population density, land use, or natural resource distribution.
3. Navigation: Maps are essential for navigation, whether you are driving, hiking, or traveling by any other means. Maps can help you to plan a route, find your way, and avoid getting lost.
4. Planning: Maps are also useful for planning, whether it is for urban planning, resource management, or disaster response. By mapping out the landscape and understanding how different factors interact, planners can make informed decisions that benefit communities and the environment.
5. Education: Maps can be used as educational tools to teach geography, history, and other subjects. By visualizing data and making it more accessible, maps can help to engage students and make learning more enjoyable.
6. Communication: Maps are a universal language that can be used to communicate with people from all over the world. By using maps to convey information, people can communicate complex ideas without needing to know each other's language.

Disadvantages of Map

While maps have many advantages, there are also some disadvantages to consider. Some of these include:

1. Simplification: Maps often simplify complex data to make it easier to understand. While this can be useful, it can also lead to oversimplification and distortion of the data. For example, a map that shows only the average temperature of an area may not accurately represent the temperature variations that occur throughout the day.
2. Misinterpretation: Maps can be misinterpreted or misunderstood if they are not labeled or presented clearly. Users may misinterpret symbols or labels, or they may not understand the scale of the map or the projection used.
3. Bias: Maps can be biased or influenced by the creator's perspective. For example, a map that emphasizes one political or cultural group may ignore or downplay the presence of other groups in the same area.
4. Outdated Information: Maps may become outdated if they are not updated regularly. Changes in the landscape, such as the construction of new buildings or the destruction of natural resources, may not be reflected on the map.
5. Cost: Creating and maintaining accurate maps can be expensive, especially for large or complex areas. Access to some types of mapping data, such as satellite imagery, can also be expensive and limited.

6. Privacy Concerns: Maps that include personal data or sensitive information, such as the location of military bases or the home addresses of individuals, may raise privacy concerns.

Mapping Concept and Techniques

Mapping is the process of creating a representation of a physical or abstract space. This representation can take many forms, including charts, diagrams, graphs, or digital maps. Mapping is used in a wide variety of fields, such as geography, cartography, data science, and computer science. It involves the collection of data, its analysis, and the creation of a visual representation of that data.

In geography and cartography, mapping refers to the creation of maps that show the physical features of the earth's surface, such as mountains, rivers, and oceans, as well as human-made features like roads, buildings, and cities. These maps can be used to help people navigate, understand the landscape, and plan for future development.

In data science and computer science, mapping involves the process of converting data from one format to another. This can involve transforming data into a different format for visualization, analysis, or use in a particular software or application.

Mapping is also used in fields such as biology, where it is used to create diagrams of gene sequences, and in social science, where it can be used to map social networks and relationships.

Why Mapping Concept?

Mapping concept is an important process for a number of reasons:

1. Understanding the big picture: By mapping out the important concepts within a field, you can gain a better understanding of the big picture and how different pieces of information fit together. This can help you develop a more holistic understanding of a topic, which can be especially important when working on complex projects or solving difficult problems.
2. Identifying knowledge gaps: Mapping concept can help you identify areas where you may be lacking knowledge or where further research is needed. By identifying these gaps, you can develop a plan to fill them and improve your understanding of the subject.

3. Effective communication: Mapping concept can help you communicate complex ideas more effectively. By identifying the most important concepts and their relationships, you can communicate information in a way that is easy to understand and digest.
4. Improved decision-making: By mapping out important concepts, you can make better-informed decisions. This is because you have a more complete understanding of the subject matter and can see how different pieces of information fit together.

A. Digital Techniques

Digital mapping techniques involve the use of digital tools and technologies to create, analyze, and visualize maps and geographic data. These techniques have revolutionized the field of mapping, making it possible to create detailed, dynamic maps with rich layers of information.

Some of the key digital mapping techniques include:

1. Geographic Information Systems (GIS): GIS is a software tool used to capture, analyze, and visualize geographic data. It allows users to create digital maps with multiple layers of information, such as population data, land use, and transportation networks.
2. Remote Sensing: Remote sensing involves the use of satellite or airborne sensors to collect data about the earth's surface. This data can be used to create digital maps that show features such as vegetation cover, land use, and water resources.
3. Global Positioning System (GPS): GPS is a system of satellites and receivers used to determine the precise location of objects on the earth's surface. GPS technology can be used to create accurate digital maps and to track the movement of objects or people.
4. Web Mapping: Web mapping refers to the use of the internet to access and share digital maps. Web mapping applications, such as Google Maps and OpenStreetMap, allow users to view and interact with digital maps from anywhere in the world.
5. Data Visualization: Data visualization techniques can be used to create dynamic, interactive maps that allow users to explore and analyze complex data sets. These maps can be used to identify patterns and trends in data that might be difficult to see in other formats.

Digital mapping techniques are used in a wide range of fields, from urban planning and environmental management to marketing and business intelligence. They offer a powerful set of tools for analyzing and understanding geographic data and for communicating that information to others.

B. Analog Mapping Techniques

Analog mapping techniques refer to the traditional methods of creating maps before the advent of digital technology. These techniques involve the use of manual tools and instruments to draw and represent geographic features. Some of the key analog mapping techniques include:

1. Cartography: Cartography is the art and science of creating maps. Traditional cartographers used manual tools such as pens, pencils, and brushes to draw maps by hand. They also used specialized instruments such as compasses, protractors, and scales to ensure accuracy in their drawings.
2. Aerial Photography: Aerial photography involves taking photographs of the earth's surface from an airplane or helicopter. These photographs were used to create maps and to identify features such as land use, vegetation cover, and topography.
3. Topographic Surveying: Topographic surveying involves the use of surveying equipment to measure and record the elevations and contours of the earth's surface. This information was used to create maps that showed the shape and features of the land.
4. Field Mapping: Field mapping involves physically visiting a site and making observations and measurements of the landscape. This information was used to create detailed maps that showed features such as vegetation cover, soil types, and water resources.
5. Printing: Once a map was created, it was typically reproduced using printing techniques such as lithography or engraving. This allowed the map to be distributed and shared with others.

Analog mapping techniques are still used today, particularly in fields such as art and illustration. They offer a more tactile and hands-on approach to map-making and can be useful in creating unique and visually appealing representations of geographic features. However, digital mapping techniques have largely replaced analog methods in most applications due to their greater efficiency and accuracy.

Differences between Digital and Analog Mapping Techniques

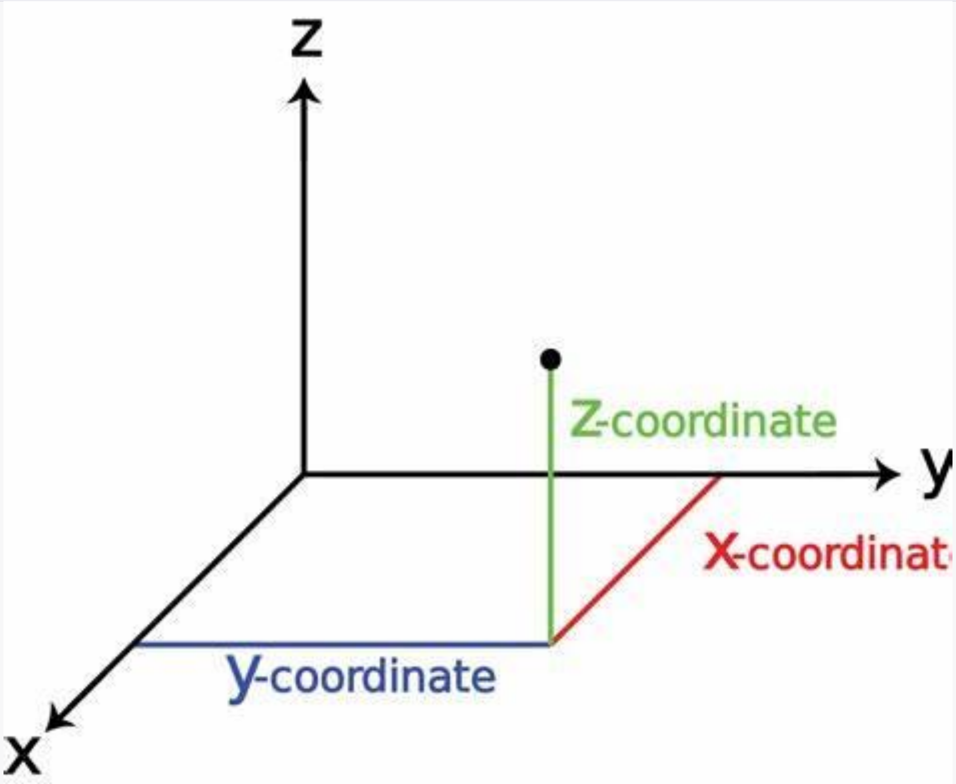
Digital and analog mapping techniques are two different methods of creating and representing maps. The main differences between the two are:

1. Technology Used: Digital mapping uses computer technology and software to create and manipulate maps, while analog mapping uses physical tools such as pens, pencils, and drafting tools to create maps.
2. Precision and Accuracy: Digital mapping is generally more precise and accurate than analog mapping. Digital maps can be created with exact measurements and can be zoomed in to show details, while analog maps may have limitations in terms of precision and accuracy.
3. Time and Effort: Digital mapping can be created quickly and easily using software, while analog mapping can be time-consuming and labor-intensive. Analog maps require more manual labor to create and may need to be redrawn if changes are made, while digital maps can be easily edited and updated.
4. Data Integration: Digital maps can integrate data from various sources, such as satellite imagery, GPS data, and sensor data, to create detailed and accurate maps. Analog maps are limited to the data that can be manually collected and incorporated.
5. Reproduction: Digital maps can be easily reproduced and shared electronically, while analog maps need to be physically reproduced, which can be time-consuming and costly.

Digital mapping is more precise, efficient, and flexible in terms of data integration and reproduction, while analog mapping may have more of an artistic feel and require more manual labor but may have limitations in precision and accuracy.

Coordinate

A coordinate is a numerical value that represents the position of a point in a space, usually in reference to a specific coordinate system. Coordinates are commonly used in mathematics, physics, engineering, and many other fields to describe the location of objects in two or three-dimensional space. The most commonly used coordinate systems are the Cartesian coordinate system (also known as rectangular coordinates) and the polar coordinate system. In the Cartesian coordinate system, coordinates are given as ordered pairs (x, y) where x represents the horizontal position and y represents the vertical position of the point. In the polar coordinate system, coordinates are given as (r, θ) where r represents the distance of the point from the origin and θ represents the angle between the positive x-axis and the line connecting the origin to the point.



Coordinate System

A coordinate system is a mathematical framework that allows us to define and describe the position of objects in space. It is a set of rules or conventions used to represent spatial positions of objects. The most commonly used coordinate systems are the Cartesian coordinate system, the polar coordinate system, and the spherical coordinate system.

In the Cartesian coordinate system, a point is represented by an ordered pair (x, y) of real numbers, where x represents the horizontal position and y represents the vertical position of the point. The origin of the Cartesian coordinate system is the point (0, 0).

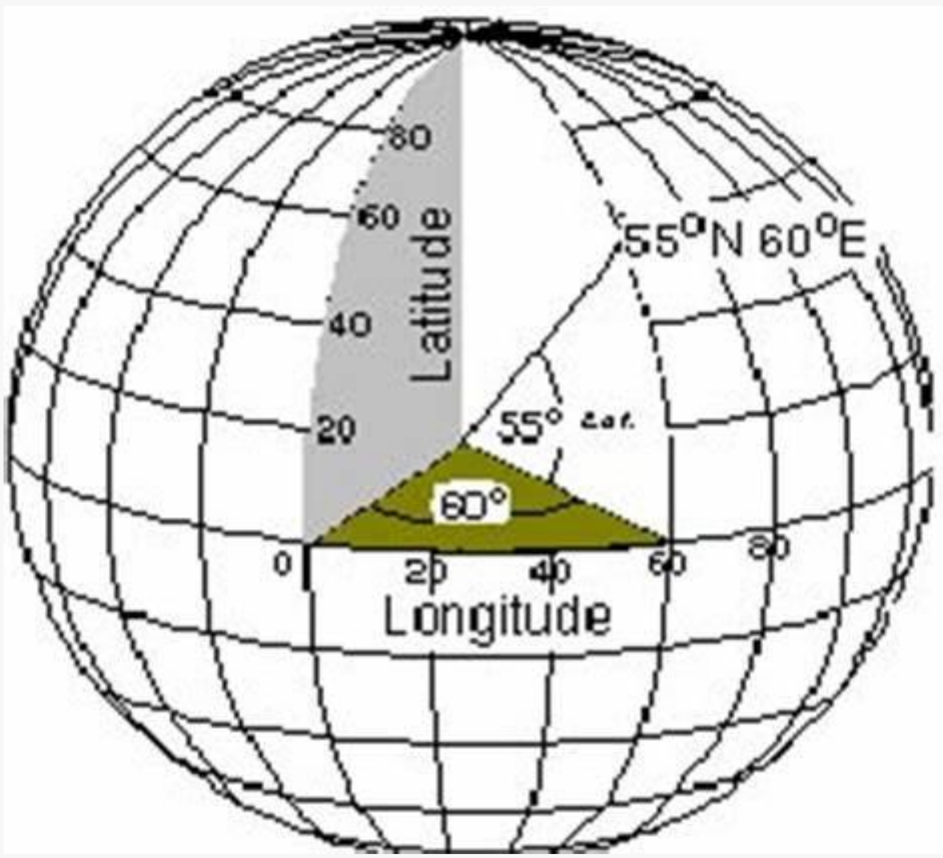
In the polar coordinate system, a point is represented by an ordered pair (r, θ), where r represents the distance of the point from the origin and θ represents the angle between the positive x-axis and the line connecting the origin to the point.

In the spherical coordinate system, a point is represented by an ordered triple (r, θ, φ), where r represents the distance of the point from the origin, θ represents the azimuth angle (i.e., the angle between the positive x-axis and the projection of the line connecting the origin to the point onto the xy-plane), and φ represents the polar angle (i.e., the angle between the positive z-axis and the line connecting the origin to the point).

Coordinate systems are used in many areas of mathematics, physics, engineering, and other fields to describe the positions of objects and to perform calculations involving spatial positions.

Geographic Coordinate system

A geographic coordinate system (GCS) is a coordinate system that uses latitude and longitude to represent positions on the earth's surface. It is a system of reference lines and grids used to locate any point on the surface of the earth. The GCS is used by cartographers, geographers, surveyors, and other professionals who work with spatial data related to the earth's surface.



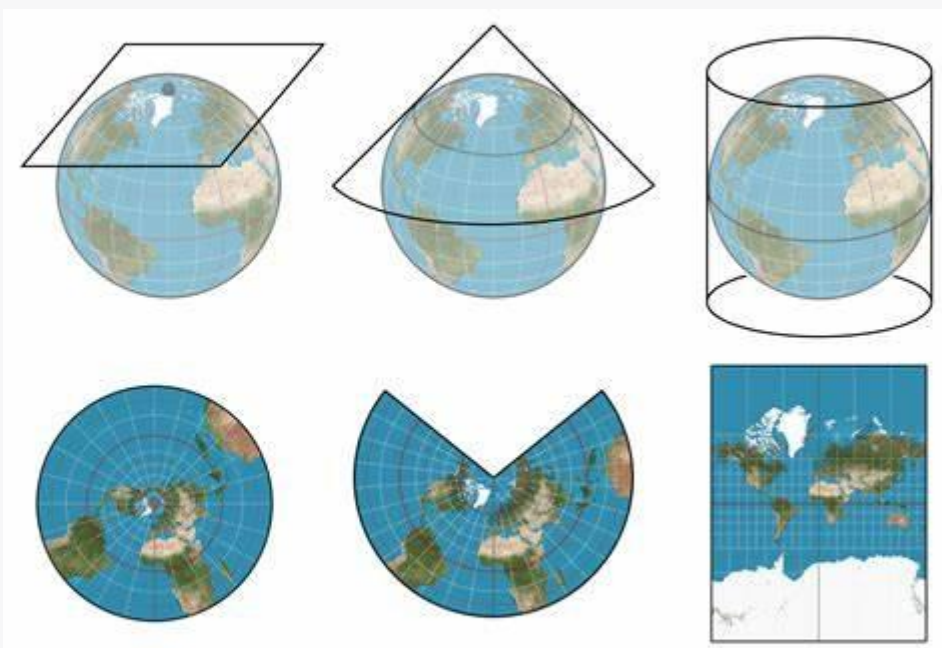
In the GCS, latitude measures the north-south position of a point on the earth's surface relative to the equator, while longitude measures the east-west position relative to the Prime Meridian, which passes through Greenwich, England. Latitude is measured in degrees north or south of the equator, with the equator itself being at 0 degrees latitude. Longitude is measured in degrees east or west of the Prime Meridian, with the Prime Meridian itself being at 0 degrees longitude.

The GCS is essential for mapping, navigation, and other applications that require precise location information. It allows us to identify any point on the earth's surface using a unique set of coordinates. Different countries and organizations may use different GCS, but the most commonly used one is the World Geodetic System 1984 (WGS84), which is used as the standard for global positioning system (GPS) technology.

Map Projection

Map projection is the process of representing the three-dimensional surface of the earth on a two-dimensional map. Since the earth's surface is curved, a flat map inevitably involves some degree of distortion. Map projections are designed to minimize or eliminate this distortion, depending on the purpose of the map.

There are many different types of map projections, but they can be broadly classified into three categories:



1. Cylindrical projections: These projections wrap a cylinder around the earth's surface, with the cylinder tangent to the equator. Examples include the Mercator

projection, which is commonly used for navigation, and the Miller projection, which distorts the poles less than the Mercator projection.

2. Conic projections: These projections wrap a cone around the earth's surface, with the cone tangent to a single latitude line. Examples include the Albers projection, which is often used for choropleth maps of the United States, and the Lambert conformal conic projection, which is used for large-scale maps of mid-latitudes.
3. Azimuthal projections: These projections project the earth's surface onto a plane that is tangent to a single point on the earth's surface. Examples include the stereographic projection, which is often used for maps of polar regions, and the orthographic projection, which shows one hemisphere in true size and shape.

Each type of projection has its own advantages and limitations, depending on the purpose of the map and the area being depicted. Mapmakers must carefully select an appropriate projection that balances accuracy and distortion for their specific needs.

WGS 84 Coordinate system

WGS 84 (World Geodetic System 1984) is a global coordinate system used for geospatial data. It was developed by the United States Department of Defense and the National Geospatial-Intelligence Agency, and it is used by many organizations around the world, including GPS satellite navigation systems.

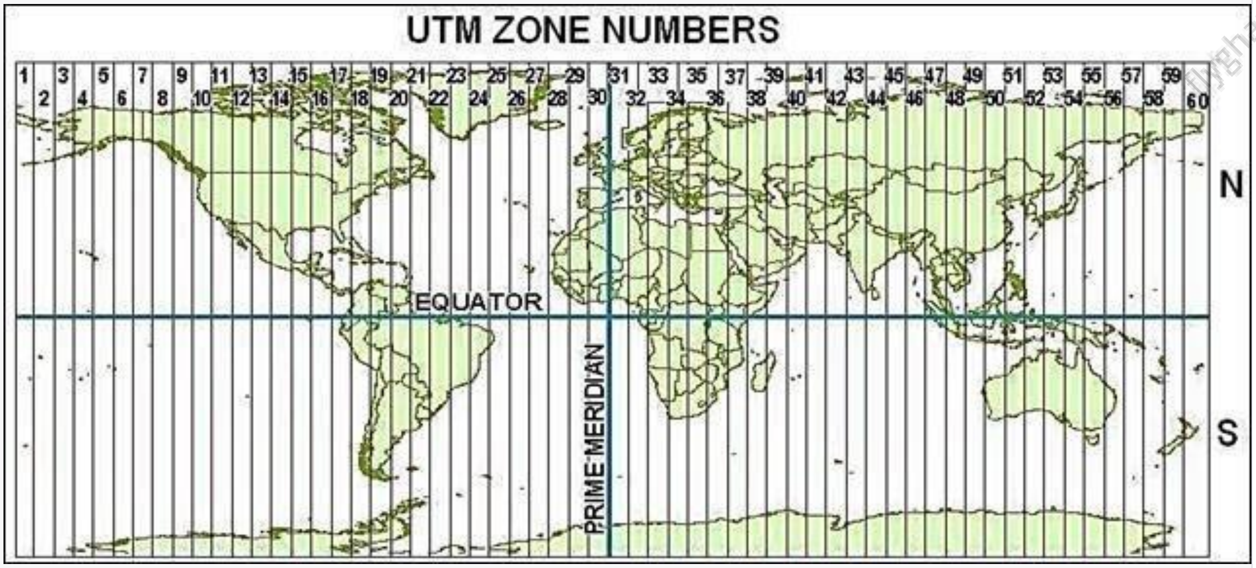
The WGS 84 coordinate system is a geographic coordinate system that uses latitude and longitude to define locations on the Earth's surface. The origin of the coordinate system is the Earth's center of mass, and the equator and the prime meridian are used as the reference lines. The equator is defined as 0 degrees latitude, while the prime meridian is defined as 0 degrees longitude.

The WGS 84 coordinate system is based on a spheroid that approximates the shape of the Earth. The spheroid has a semi-major axis of 6,378,137 meters and a semi-minor axis of 6,356,752.3 meters, and it is slightly flattened at the poles and bulging at the equator. This shape is known as an oblate spheroid.

The WGS 84 coordinate system is widely used in GIS applications, particularly for global positioning and navigation systems, because it provides a standard reference for locations on the Earth's surface. It is also used as the standard coordinate system for many online maps and mapping software, including Google Maps and OpenStreetMap.

UTM Projection System

The Universal Transverse Mercator (UTM) projection system is a cylindrical projection that divides the earth's surface into 60 zones, each 6 degrees of longitude wide. The UTM system is commonly used for large-scale mapping and navigation in many countries around the world.



In the UTM system, each zone is projected onto a flat surface, resulting in a series of rectangular grids that can be easily measured and analyzed. The UTM projection minimizes distortion within each zone, resulting in a more accurate representation of the earth's surface than some other types of projections.

The UTM system uses a Cartesian coordinate system, with each point on the earth's surface represented by a pair of coordinates (eastings and northings) in meters. This allows for precise measurements of distance, direction, and area.

One of the advantages of the UTM system is its ease of use with modern navigation and mapping technologies, such as GPS and GIS. These technologies can accurately locate and map features on the earth's surface within the UTM coordinate system, allowing for more accurate and efficient analysis and decision-making.

Overall, the UTM projection system is a widely used and effective tool for large-scale mapping and navigation, particularly in areas with well-defined longitudinal boundaries such as many parts of Europe and North America.

MUTM Projection System

Map projection is a vital component of modern cartography and plays an essential role in accurately representing the geography of a country or region on a two-dimensional plane. In Nepal, the Universal Transverse Mercator (UTM) and Modified UTM (MUTM) projection systems are commonly used for mapping purposes. In this article, we will explore these two systems and discuss their significance in Nepali cartography.

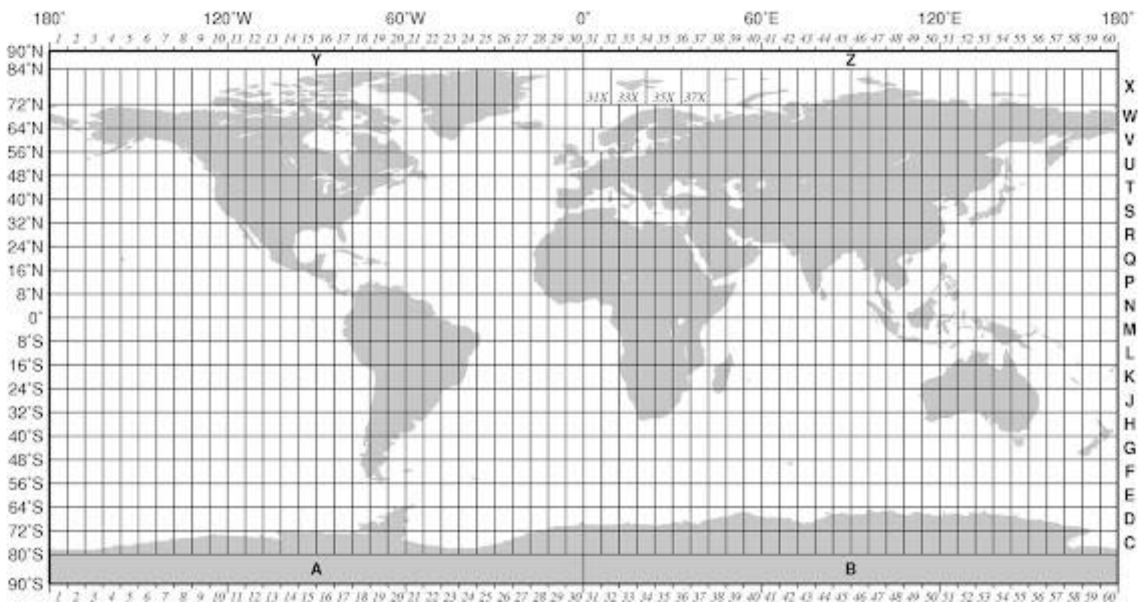


Fig:- UTM projection

The UTM system is a conformal cylindrical projection that preserves shape and angle but can distort length and breadth. This system divides the earth into 60 zones of 6 degrees each, with a central meridian for each zone. Nepal falls under zones 44 degrees N and 45 degrees N, and the central meridian for Nepal has a scaling factor of 0.9996. The UTM projection system is widely used for surveying and mapping in Nepal.

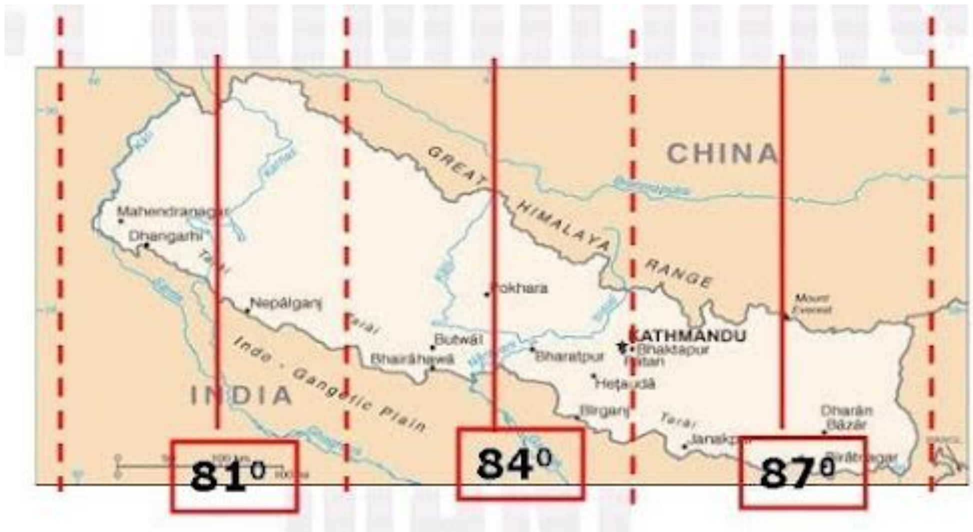


Fig:-MUTM (Modified UTM)

To improve the accuracy of cadastral surveys and other purposes, Nepal uses the Modified UTM (MUTM) projection system. The MUTM system divides the earth into 120 zones of 3 degrees each, with Nepal having a central meridian of 81 degrees, 84 degrees, and 87 degrees. The central meridian for Nepal typically has a scale factor of 0.9999, with a false easting of 500,000 meters to keep all coordinates within the country positive. The false northing at the Equator is 0 meters.

MUTM is a modified version of the UTM projection system that provides increased accuracy and precision for mapping and surveying. It is especially useful for Nepali cartography as it allows for more precise representation of the country's topography, which is characterized by diverse terrain, including high-altitude mountain ranges, hills, and plains.

In conclusion, map projection systems are crucial for accurate cartography, and the UTM and MUTM systems are essential for mapping and surveying in Nepal. The UTM system is widely used, but the MUTM system offers increased accuracy and precision for Nepali cartography. Understanding these projection systems is critical for those involved in cartography and surveying in Nepal

Unit 3: GIS data sources and Structure

Capturing GIS data

Capturing GIS data involves the process of collecting and inputting spatial and attribute data into a GIS software. Here are some of the methods used for capturing GIS data:

1. Global Positioning System (GPS) survey: GPS survey is a method of collecting spatial data using a GPS receiver that can capture latitude, longitude, and elevation of a point on the earth's surface. GPS survey can be done using handheld GPS devices, drones, or other surveying equipment.
2. Aerial photography and remote sensing: Aerial photography and remote sensing are methods of collecting spatial data using aircraft or satellite imagery. These methods can capture high-resolution images and data on land cover, terrain, and other features that can be used to create GIS maps.
3. Field survey and data collection: Field survey involves the collection of spatial and attributes data through direct observation and measurement of features on the ground. This method can be used to capture data on features such as roads, buildings, and land cover.
4. Digitization: Digitization is a process of converting data from a paper map or other hardcopy source into digital format using a scanner or digitizing tablet. This method can be used to capture data on features such as property boundaries, road networks, and other infrastructure.
5. Data conversion: Data conversion involves converting data from different formats into a GIS-compatible format. This method can be used to import data from sources such as spreadsheets, databases, and other software into a GIS.

Overall, capturing GIS data involves the process of collecting and inputting spatial and attributes data using a variety of methods, which can be tailored to the specific needs of a GIS project.

GPS Concepts

GPS stands for Global Positioning System. It is a network of satellites orbiting the earth that provides precise location and time information to users around the world. GPS was developed by the United States Department of Defense and became fully operational in 1995.

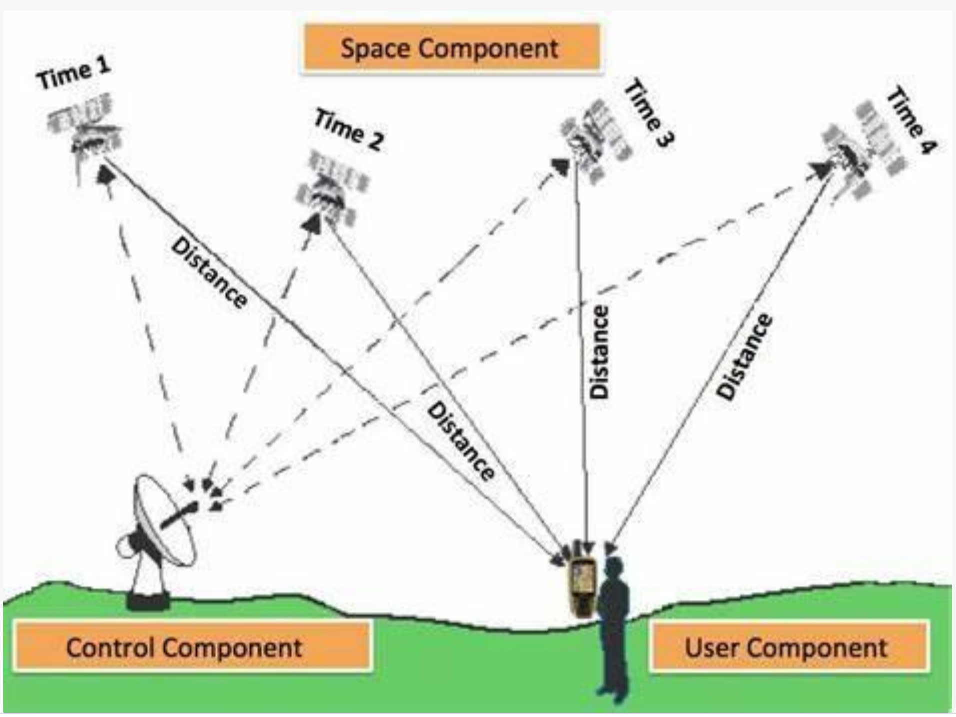
GPS technology works by using a network of at least 24 satellites that constantly transmit signals to Earth. GPS receivers on the ground pick up these signals and use them to calculate the receiver's location, velocity, and time. By receiving signals from multiple satellites, the GPS receiver can determine its precise location with a high degree of accuracy.

GPS technology has a wide range of applications in fields such as navigation, surveying, geology, and meteorology. It is used in cars and airplanes for navigation, in agriculture for precision farming, and in search and rescue operations to locate missing persons or aircraft.

GPS technology has also led to the development of location-based services such as geotagging, which allows users to attach location data to photos and social media posts. GPS is now also commonly integrated into mobile devices such as smartphones and tablets, making it a ubiquitous technology for location-based services and applications.

GPS Componenets

GPS (Global Positioning System) is a satellite-based navigation system that uses a network of satellites to determine the location, speed, and direction of a receiver on or near the Earth's surface. The GPS system is composed of three main components:



1. Satellites(Space Segment): There are currently over 30 active GPS satellites orbiting the Earth at an altitude of approximately 20,000 kilometers. The satellites

are placed into six different orbital planes and are evenly spaced around each plane. The satellites broadcast signals containing time and location data that are received by GPS receivers on the ground.

2. Control stations: There are several control stations on the ground that monitor and maintain the GPS satellites. The control stations send signals to the satellites to update their orbital and time data, and to correct any errors in the satellite's signal.
3. GPS receivers (User Segment): GPS receivers are the devices that receive the signals from the GPS satellites and calculate the user's location, speed, and direction. The receiver uses the time data in the satellite signals to determine the distance to each satellite, and then uses that information to calculate the user's location using a process called trilateration. GPS receivers can be built into devices such as smartphones, cars, and aircraft, or can be standalone devices.

Remote Sensing

Remote sensing is the process of acquiring information about the Earth's surface without physically being in direct contact with it. It involves the use of sensors that are mounted on aircraft or satellites to detect and measure the physical properties of the Earth's surface and atmosphere. Remote sensing technology can be used to collect information about a variety of features and phenomena, including land cover and land use, vegetation, water resources, climate, and geology.

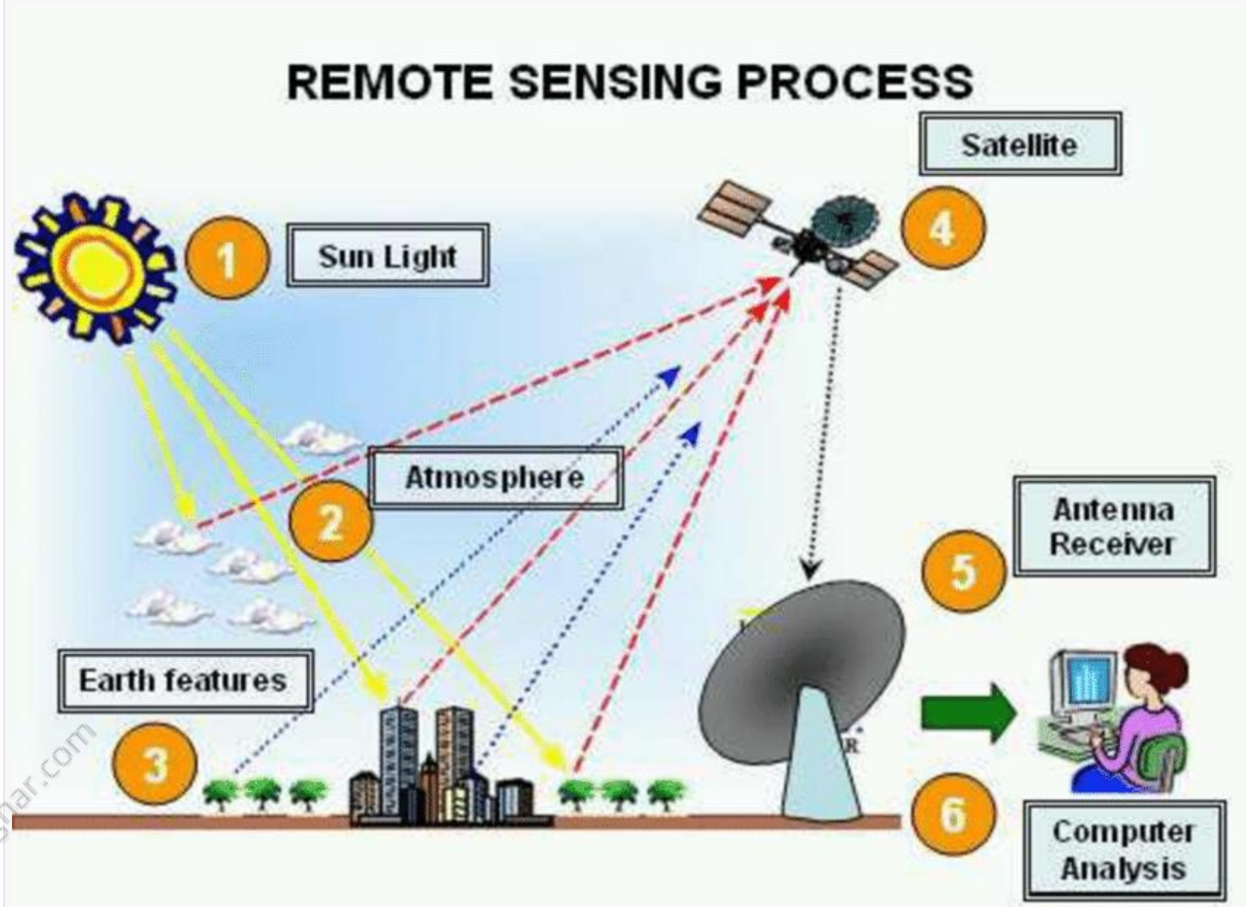
Remote sensing data can be collected in a variety of forms, including images, digital elevation models, and other types of data. These data can be processed and analyzed to extract information about the Earth's surface, and can be used for a wide range of applications, such as:

1. Land use and land cover mapping: Remote sensing data can be used to map and monitor land use and land cover changes over time. This information is useful for urban planning, agriculture, forestry, and natural resource management.
2. Vegetation monitoring: Remote sensing data can be used to monitor vegetation health, productivity, and distribution. This information is useful for crop management, forest management, and conservation.
3. Water resource management: Remote sensing data can be used to monitor water resources, including the location and extent of water bodies, water quality, and changes in water levels. This information is useful for water resource management, flood forecasting, and drought monitoring.
4. Disaster management: Remote sensing data can be used to assess the impact of natural disasters such as floods, fires, and earthquakes, and to aid in disaster response and recovery efforts.

5. Climate studies: Remote sensing data can be used to monitor and model climate patterns, such as temperature, precipitation, and wind. This information is useful for climate research, weather forecasting, and predicting the impact of climate change.

Principle of Remote Sensing

The principle of remote sensing involves the detection and interpretation of information about an object or area without physically contacting it. Remote sensing relies on the interaction of electromagnetic radiation with matter. The basic principle of remote sensing is that different materials reflect, absorb or transmit electromagnetic radiation differently. By analyzing these differences, remote sensing can provide information about the physical and chemical properties of the object or area being studied. The following are the key principles of remote sensing:



1. Energy source: Remote sensing relies on the use of an external energy source, such as the sun or a man-made source, to illuminate the object or area being studied.

2. Interaction with matter: The energy from the external source interacts with the matter in the object or area, which causes the matter to reflect, absorb, or transmit the energy.
3. Transmission of energy: The energy that is not absorbed or reflected by the object or area is transmitted through the atmosphere and captured by sensors on the ground or on a satellite.
4. Detection and recording: The sensors detect and record the energy that has been transmitted and reflected, producing digital or analog images or other data products.
5. Analysis and interpretation: The digital or analog images and data products are analyzed and interpreted to extract information about the object or area being studied.
6. Application: The information extracted from remote sensing can be used for a variety of applications, such as mapping land use and land cover, monitoring vegetation health, assessing water resources, and detecting changes in the environment.

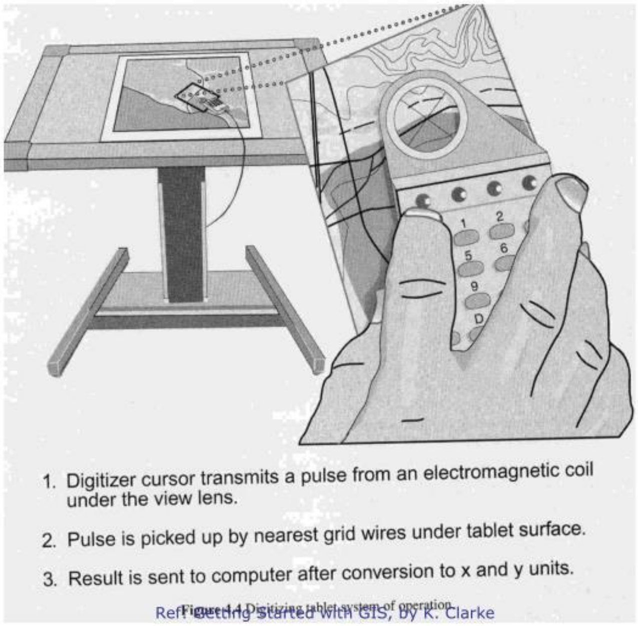
Digitization

Digitization in GIS (Geographic Information Systems) refers to the process of converting analog or physical geographic data into digital form that can be processed and analyzed using computer software. Digitization is a fundamental step in GIS data creation, as it involves converting geographic features such as points, lines, and polygons into digital representations that can be stored, manipulated, and analyzed in a GIS.

There are two main methods of digitization in GIS:

1. Manual Digitization: This method involves using a digitizing tablet or a mouse to trace the geographic features from a physical map, aerial photograph, or other analog sources, and then creating corresponding digital representations of those features using GIS software. Manual digitization requires careful attention to detail to ensure accuracy and can be time-consuming.

Manual Digitizing



2. Automated Digitization: This method involves using advanced software algorithms and techniques to automatically extract geographic features from digital imagery such as satellite images, LiDAR data, or other digital sources. Automated digitization can be more efficient and faster compared to manual digitization, but it may require additional processing and verification steps to ensure accuracy.

Digitization in GIS has numerous benefits, including improved data accuracy, ease of data storage and retrieval, ability to perform spatial analysis and modeling, and efficient data sharing and dissemination. It has revolutionized the way geographic data is collected, stored, and analyzed, enabling better decision-making in fields such as urban planning, environmental management, emergency response, transportation planning, and many others. However, it also comes with challenges such as data quality assurance, potential errors in digitization, and the need for ongoing data maintenance and updates. Overall, digitization has greatly advanced the field of GIS and has become an essential process in modern geographic data management. So, digitization plays a crucial role in enhancing the capabilities of GIS and enabling more effective use of spatial data in various applications.

Data Conversion

Data conversion in Geographic Information Systems (GIS) refers to the process of changing data from one format, coordinate system, or projection to another in order to facilitate its use in a different context or application. GIS data conversion is a crucial step in many GIS workflows, as different GIS software and systems may require data to be in specific formats or projections for proper analysis, visualization, or integration with other data sources.

There are several common types of data conversion in GIS, including:

1. Format conversion: This involves converting data from one file format to another. For example, converting a shapefile to a geoJSON file or converting a raster image to a grid format.
2. Coordinate system conversion: This involves converting data from one coordinate system to another. Different geographic coordinate systems use different reference datums, ellipsoids, and units of measurement, so converting data from one coordinate system to another ensures that it is accurately positioned on the Earth's surface.
3. Projection conversion: This involves converting data from one map projection to another. Map projections are mathematical methods used to represent the three-dimensional Earth on a two-dimensional map, and different map projections have different properties and distortions. Projection conversion ensures that data is displayed accurately on a map and preserves its spatial integrity.
4. Data attribute conversion: This involves converting data attributes or properties from one data type or schema to another. For example, converting a text field to a numeric field, or merging data from different sources with different attribute schemas.

Data conversion in GIS can be performed using GIS software tools, such as ArcGIS, QGIS, or specialized data conversion tools. It requires careful consideration of the accuracy, precision, and integrity of the data being converted, as well as an understanding of the target format or system requirements. Properly executed data conversion is essential to ensure that GIS data is compatible, accurate, and reliable for analysis, visualization, and decision-making in GIS applications. So, it is important to understand the characteristics and requirements of the data being converted, as well as the capabilities and limitations of the tools and software used for data conversion in GIS. Consulting with experienced GIS professionals and following best practices for data conversion can help ensure accurate and reliable results. In some cases, manual data transformation may be required to address unique data characteristics or complex conversion requirements. Careful validation and quality assurance processes should also be applied to ensure the integrity of the converted data. Overall, data conversion is a critical step in the GIS data management process, enabling interoperability and integration of diverse data sources in GIS applications. It is important to carefully plan, execute, and validate data conversion processes to ensure accurate, reliable, and meaningful results in GIS analysis and decision-making. Properly executed data conversion is essential to ensure that GIS data is compatible, accurate, and reliable for analysis, visualization, and decision-making in GIS applications. So, it is important to understand the characteristics and requirements of the data being converted, as well as the capabilities and limitations of the tools and software used for data conversion in GIS. Consulting with experienced GIS professionals and following best practices for data conversion can help ensure accurate and reliable results. In some cases, manual data transformation may be required to address unique data characteristics or complex conversion requirements. Careful validation and quality assurance processes should also be applied to ensure the integrity of the converted data.

Rasterization and Vectorization

Rasterization and vectorization are two processes used in GIS to convert data between raster and vector formats. Raster data represents geographic features as a grid of cells, where each cell contains a value that represents a specific attribute or characteristic of the feature. Vector data, on the other hand, represents geographic features as points, lines, or polygons with associated attributes.

1. Rasterization: Rasterization is the process of converting vector data into raster data. In this process, vector features such as points, lines, or polygons are converted into a raster grid where each cell represents a specific geographic location and contains a value that represents a specific attribute or characteristic of

the feature. Rasterization is commonly used when working with satellite imagery, aerial photographs, or other continuous data sources that are inherently raster-based.

For example, if you have a vector dataset representing land use boundaries (e.g., polygons) and you want to create a raster dataset that represents land use classes with each cell in the raster representing a specific land use class (e.g., forest, urban, agriculture), you can use rasterization to convert the vector polygons into a raster grid with attribute values assigned to each cell.

2. Vectorization: Vectorization is the process of converting raster data into vector data. In this process, raster data such as satellite images or scanned maps are converted into vector features such as points, lines, or polygons. Vectorization is commonly used when you need to extract features such as roads, buildings, or other discrete objects from raster data and represent them as vector features.

For example, if you have a raster image of a map that shows roads as pixels and you want to create a vector dataset that represents the roads as lines with associated attributes (e.g., road type, speed limit), you can use vectorization to convert the raster roads into vector lines with attributes assigned to each feature.

Rasterization and vectorization are important processes in GIS as they enable the conversion of data between raster and vector formats, allowing for different types of data analysis, visualization, and processing. However, it's important to note that these processes can introduce some loss of information or accuracy, depending on the resolution and complexity of the data being converted, and the methods used. Therefore, it's important to carefully consider the characteristics of the data and the specific requirements of the GIS analysis or application when performing rasterization or vectorization. Additionally, using appropriate tools and techniques, and validating the results are important steps to ensure accurate and reliable data conversion in GIS. Overall, rasterization and vectorization are valuable tools that allow GIS practitioners to convert data between different formats to meet the specific needs of their analysis or application.

Spatial and Non-Spatial data

In GIS, data can be broadly classified into two categories: spatial and non-spatial data.

Spatial data refers to information that can be tied to a specific geographic location or spatial extent. This includes information such as geographic coordinates, elevation, land use, and topography. Spatial data is usually represented as points, lines, polygons, or raster data, which can be displayed on a map or used for spatial analysis. Spatial data can be collected using methods such as GPS survey, aerial photography, remote sensing, and field survey.

Non-spatial data, on the other hand, refers to information that is not tied to a specific geographic location. This includes data such as demographic information, socio-economic data, and land ownership information. Non-spatial data can be in the form of tables, spreadsheets, text files, or other formats. Non-spatial data can also be used in GIS by linking it to spatial data using a unique identifier, such as a parcel or address.

In summary, spatial data is information that has a specific location or geographic extent, while non-spatial data is information that is not tied to a specific location. Both types of data can be used in GIS to perform spatial analysis, mapping, and other applications.

Images

Images are visual representations of objects, scenes, or ideas, typically captured through photography, digital rendering, or other forms of visual art. They can be two-dimensional or three-dimensional, and can come in various formats such as JPEG, PNG, BMP, GIF, or TIFF. Images are used for a wide variety of purposes, such as advertising, storytelling, education, and entertainment. They can also be used in scientific research, medical imaging, and other technical fields. Images are an important means of communication and can convey information, evoke emotions, and inspire creativity.

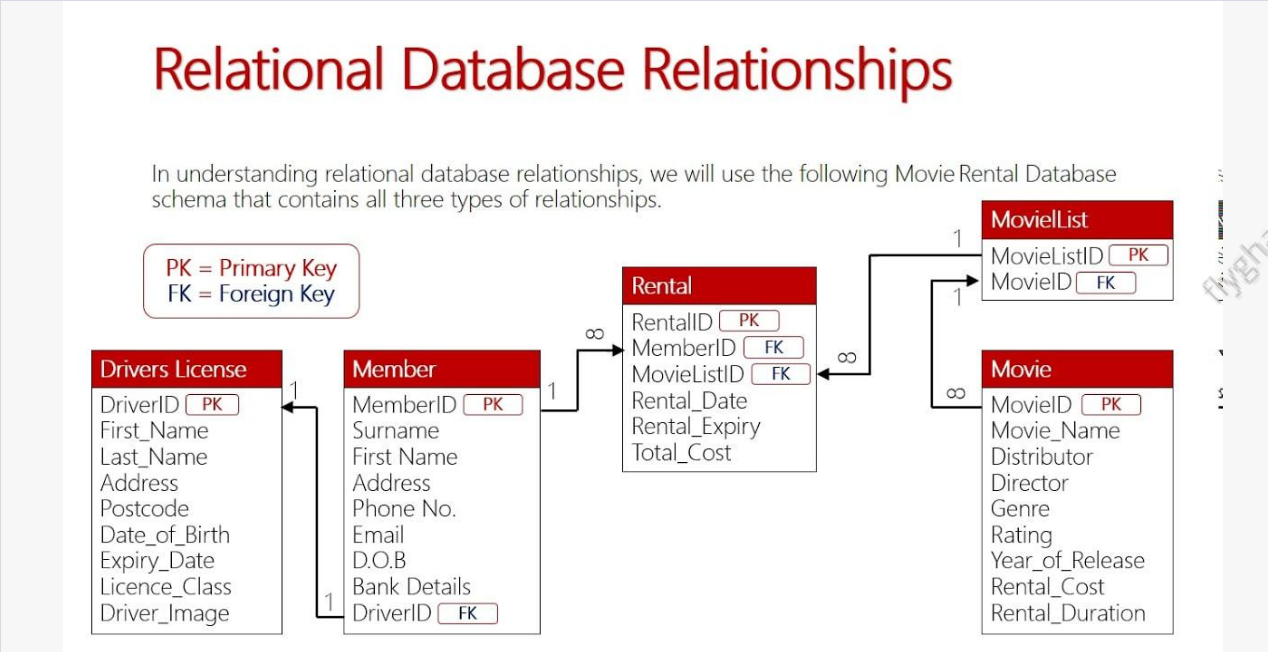
Image: NASA Landsat 8

Databases

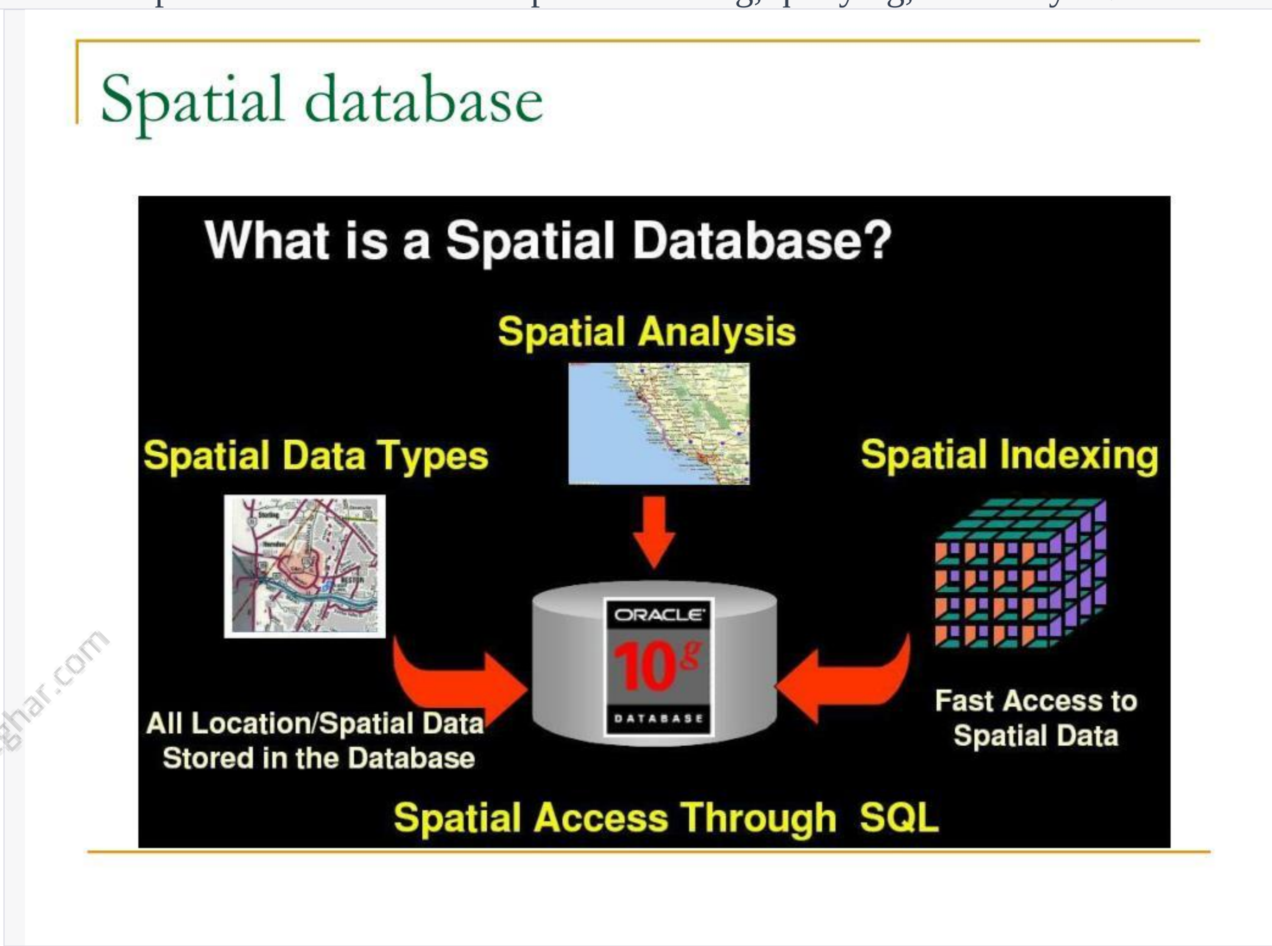
A database is a structured collection of data that is organized and stored in a way that allows for efficient storage, retrieval, and manipulation of data. In GIS, databases are used to store and manage spatial and attribute data, and they can be accessed and analyzed using GIS software.

There are several types of databases commonly used in GIS:

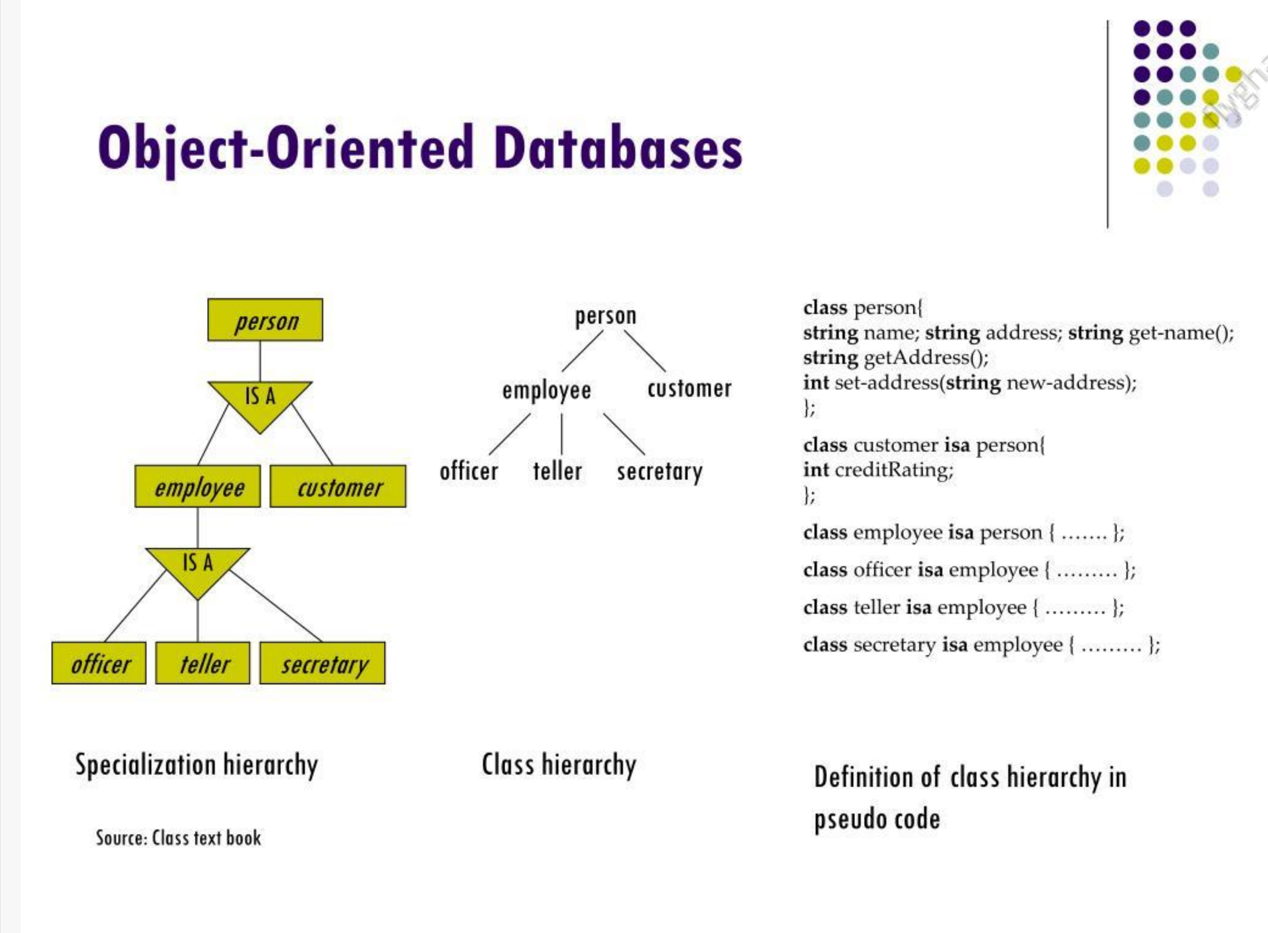
1. Relational database: This is a database that organizes data into tables with rows and columns. Each row represents a record, and each column represents an attribute or field. Relational databases are widely used in GIS because they can store both spatial and non-spatial data, and they allow for powerful queries and analysis.



2. Spatial database: This is a type of database that is specifically designed to store and manage spatial data. Spatial databases can store vector and raster data, and they include specialized functions for spatial indexing, querying, and analysis.



3. Object-oriented database: This is a type of database that stores data as objects, which can include spatial and non-spatial data. Object-oriented databases are well-suited for complex data structures, and they can be used in GIS for storing and managing data such as 3D models, terrain data, and other spatially-enabled objects.



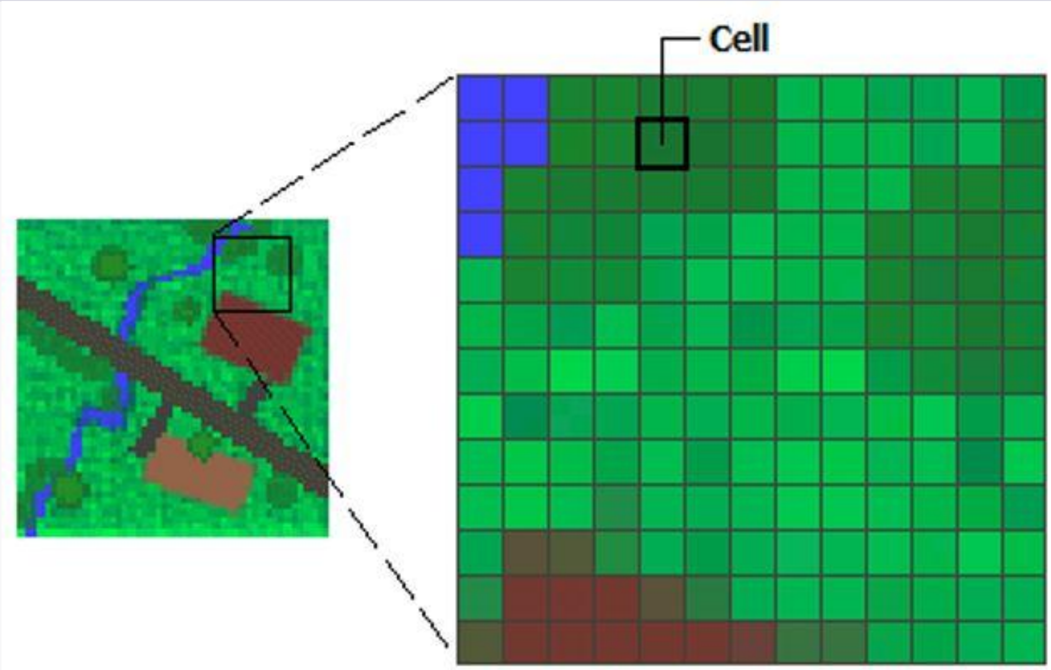
4. SQL database: This is a type of database that is designed for large-scale data storage and processing. SQL databases are often used for big data and real-time data processing applications, and they can be used in GIS for storing and managing large volumes of spatial data.

Data Structure in GIS

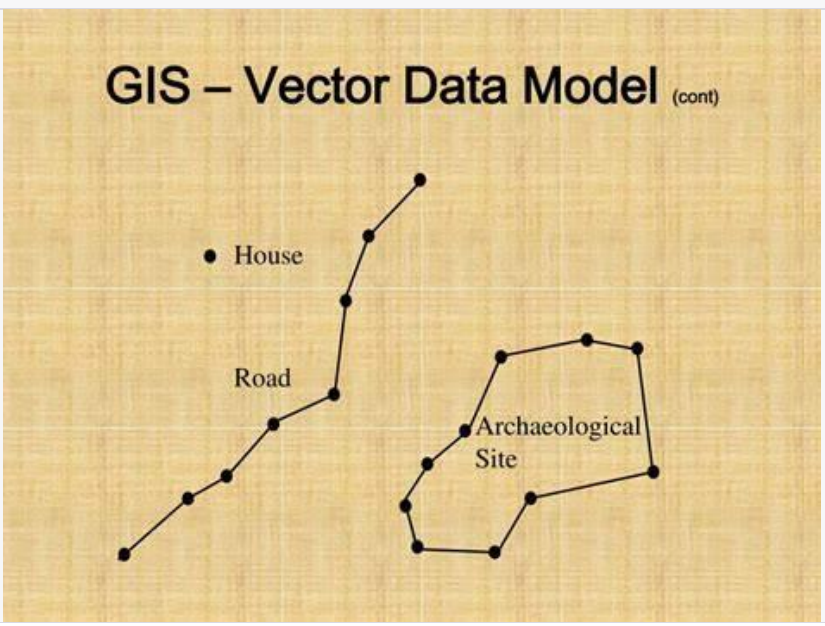
Data structure in GIS refers to the way spatial and non-spatial data are organized and stored within a GIS. There are two main types of data structures used in GIS: raster and vector.

1. Raster data structure: Raster data consists of a grid of cells or pixels, each with a unique value that represents a specific attribute of the data. Raster data is commonly used to represent continuous data such as elevation, temperature, and

satellite imagery. The size and resolution of the grid determine the level of detail that can be represented, and the type of data analysis that can be performed.



2. Vector data structure: Vector data represents spatial features using points, lines, and polygons. Each feature is represented by a set of coordinates that define its location, and attributes that describe its properties. Vector data is used to represent discrete features such as roads, buildings, and land parcels. Vector data can be stored as points, lines, or polygons, and can be easily edited, updated, and analyzed.



In addition to these two main data structures, GIS databases also include tables and metadata that describe the properties of the spatial data. These tables can be used to join and relate different data sets, and to perform advanced queries and analysis.

Advantages and dis-advantages of raster and vector

Raster and vector data are two main types of digital data used in GIS (Geographic Information System). Both have their advantages and disadvantages.

Advantages of raster data:

1. Easy to understand: Raster data are easy to understand and use because they resemble a photograph or image.
2. Good for continuous data: Raster data can represent continuous data such as elevation or temperature.
3. Easy to manipulate: Raster data are easy to manipulate with image processing software.
4. Efficient storage: Raster data are compact and efficient to store.

Disadvantages of raster data:

1. Large file size: Raster data can have a large file size, making it difficult to store and transfer.
2. Limited resolution: Raster data have a limited resolution and can become pixelated when zoomed in too far.
3. Not suitable for precise measurements: Raster data are not suitable for precise measurements and can introduce errors in calculations.

Advantages of vector data:

1. High accuracy: Vector data are accurate and precise for measurements and analysis.
2. Efficient storage: Vector data are compact and efficient to store, especially for large datasets.
3. Easy to edit: Vector data are easy to edit, such as changing attribute data.
4. Small file size: Vector data have small file sizes compared to raster data.

Disadvantages of vector data:

1. Limited to discrete data: Vector data are limited to discrete data such as points, lines, and polygons.
2. Difficult to understand: Vector data can be difficult to understand for those not familiar with the GIS software.
3. Complex to manipulate: Vector data are complex to manipulate and can require a lot of work for large datasets.

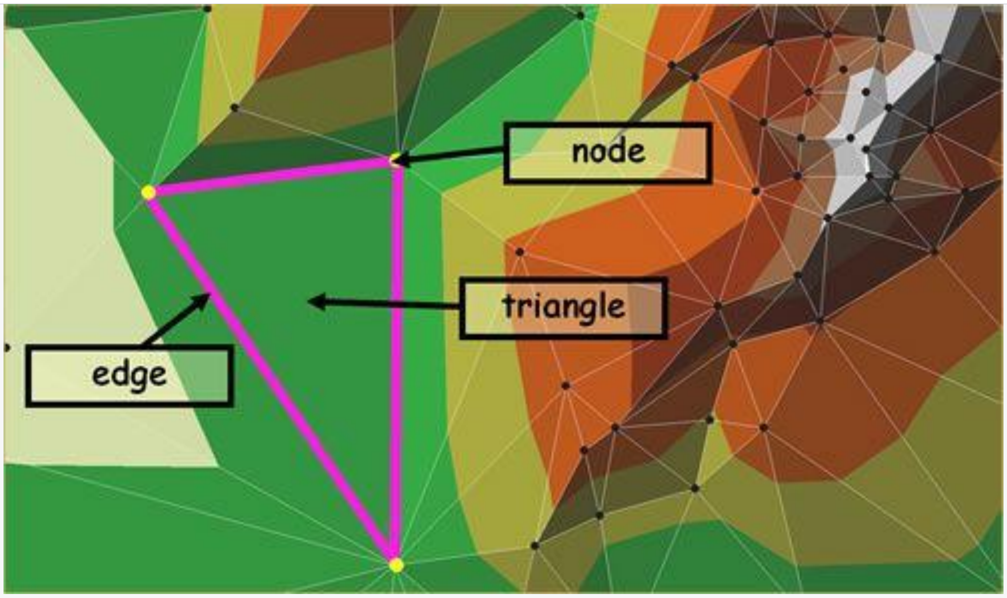
TIN

TIN (Triangular Irregular Network) is a data structure used in GIS to represent continuous surfaces such as terrain models, elevation data, and 3D models. TIN is a vector-based data structure that divides a surface into a set of non-overlapping triangles, each defined by three vertices. The vertices of the triangles are represented as 3D points, and the triangles are connected to form a continuous surface.

TIN data structure is commonly used for terrain analysis, hydrology, and geology applications. TINs can be created from contour lines, spot elevation points, and other types of elevation data. The TIN algorithm creates a set of non-overlapping triangles that best fit the elevation data, with the smallest number of triangles possible.

One of the main advantages of the TIN data structure is that it can represent complex surfaces and irregular terrain features. TINs can also be used to interpolate elevation values for points that fall within a triangle, allowing for detailed elevation analysis and modeling.

TIN data structure can be used in conjunction with other data structures, such as raster and vector data, to create comprehensive GIS models. For example, a TIN may be combined with a vector layer of building footprints to create a 3D model of a city, or with a raster layer of soil types to analyze the slope and aspect of different areas of a landscape.



GIS Data Modelling

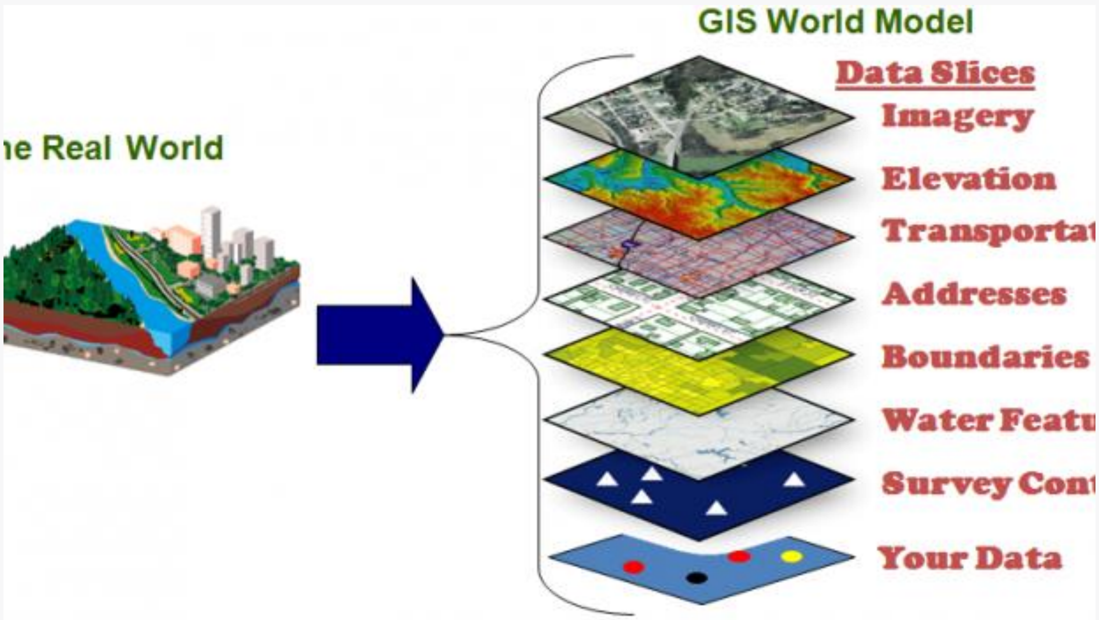
GIS data modeling is the process of creating a conceptual representation of real-world objects, phenomena, and events in a GIS. This involves identifying the relevant data elements and relationships between them, and organizing them in a way that supports spatial analysis and decision-making.

GIS data modeling can be divided into two main types: spatial data modeling and attribute data modeling.

The process of designing geospatial data models is similar to data modeling in general, at least in its overall pattern. For example, it can be segmented into three distinct levels of model abstraction:[10]

- Conceptual data model, a high-level specification of how information is organized in the mind and in enterprise processes, without regard to the restrictions of GIS and other computer systems. It is common to develop and represent a conceptual model visually using tools such as an entity-relationship model.
- Logical data model, a broad strategy for how to represent the conceptual model in the computer, sometimes novel but often within the framework of existing software, hardware, and standards. The unified modeling language (UML), specifically the class diagram, is commonly used for visually developing logical and physical models.

- Physical data model, the detailed specification of how data will be structured in memory or in files.



1. Spatial Data Modeling: Spatial data modeling is the process of defining the spatial relationships between objects in the real world. This involves creating spatial data layers, such as point, line, and polygon layers, and defining the spatial relationships between them. For example, a point layer of fire hydrants may be related to a polygon layer of fire response zones, with each zone containing multiple hydrants.
2. Attribute Data Modeling: Attribute data modeling is the process of defining the non-spatial characteristics of objects in the real world. This involves identifying the attributes of the objects, such as name, address, type, and owner, and defining the relationships between them. For example, a building may have attributes such as its address, the year it was built, the number of floors, and its owner.

Once the spatial and attribute data models are defined, they can be combined to create a comprehensive GIS data model. This involves creating a database schema that defines the structure of the data, including tables, fields, and relationships between them. The data model can then be populated with data, which can be analyzed and visualized using GIS software.

GIS data modeling is important because it allows for the creation of accurate and comprehensive GIS databases that can support spatial analysis and decision-making. A well-designed GIS data model can also help to ensure data quality, consistency, and interoperability, which are critical for effective GIS applications.

GIS Database Design

GIS database design involves creating a structured framework for organizing and managing spatial and attributes data in a GIS. The design of a GIS database should support efficient data storage, retrieval, and analysis, and should be flexible enough to accommodate changing data needs over time.

Here is an example of GIS database design for a city's transportation network:

1. Identify the Data Requirements: The first step in GIS database design is to identify the data requirements for the transportation network. This may include information about roads, highways, bridges, tunnels, and other transportation infrastructure, as well as data about traffic volumes, accidents, and other factors that affect the transportation network.
2. Define the Data Model: Based on the data requirements, a conceptual data model can be created to define the relationships between different types of data. For example, roads may be represented as a network of interconnected lines, with attributes such as road type, speed limit, and traffic volume.
3. Design the Database Schema: Using the data model, a database schema can be designed to specify the structure of the database, including tables, fields, and relationships between them. For example, a database schema may include tables for roads, intersections, traffic volumes, and accidents, with relationships between the tables to allow for spatial and attribute queries.
4. Implement the Database: Once the database schema is designed, it can be implemented using a database management system (DBMS) such as PostgreSQL or Oracle. This involves creating tables, defining fields, and setting up relationships between tables.
5. Populate the Database: With the database schema in place, data can be added to the database through a variety of methods, including manual data entry, digitization, or import from other data sources.
6. Query and Analyze the Data: Once the database is populated with data, it can be queried and analyzed using GIS software. For example, the database can be used to generate maps of the transportation network, perform spatial and attribute queries to identify areas of congestion or high accident rates, and visualize the results using charts and graphs.

In conclusion, GIS database design is a critical component of any GIS project, and involves identifying data requirements, defining data models, designing database schemas, implementing the database, populating the database, and querying and analyzing the data. By following these steps, a well-designed GIS database can provide a powerful tool for managing and analyzing spatial and attribute data.

Query

A query is a request or question submitted to a system or database in order to retrieve specific information. In the context of computing and information technology, a query is typically a command or statement written in a specific programming language or query language that is used to retrieve, manipulate, or modify data stored in a database. Queries can take many forms and can be used for a wide range of purposes, from simple searches for specific pieces of information to complex data analysis and processing tasks. Some common types of queries include SELECT queries, which retrieve data from a database, and UPDATE queries, which modify existing data in a database.

Unit 4: Spatial Data Modelling and Analysis

Spatial data modeling

Spatial data modeling refers to the process of creating a mathematical representation of spatial phenomena in the real world. This representation can then be used to analyze and visualize spatial patterns and relationships.

Spatial data are what drive a GIS. Every functionality that makes a GIS separate from another analytical environment is rooted in the spatially explicit nature of the data.

Spatial data are often referred to as **layers**, **coverages**, or **layers**. We will use the term **layers** from this point on, since this is the recognized term used in ArcGIS. Layers represent, in a special digital storage format, features on, above, or below the surface of the earth. Depending on the type of features they represent, and the purpose to which the data will be applied, layers will be one of 2 major types.

Vector data represent features as discrete points, lines, and polygons.

Raster data represent the landscape as a rectangular matrix of square cells.

Depending on the type of problem that needs to be solved, the type of maps that need to be made, and the data source, either raster or vector, or a combination of the two can be used. Each data model has strengths and weaknesses in terms of functionality and representation. As you get more experience with GIS, you will be able to determine which data type to use for a particular application.

Data Structures

The 2 basic data structures in any fully-functional GIS are:

- Vector, e.g.,
  - ArcInfo Coverages
  - ArcGIS Shape Files
  - CAD (AutoCAD DXF & DWG, or MicroStation DGN files)
  - ASCII coordinate data

- Raster, e.g.,
  - ArcInfo Grids
  - Images
  - Digital Elevation Models (DEMs)

- generic raster datasets

Vector data are composed of:

- points,
- lines, and
- polygons

These features are the basic features in a vector-based GIS, such as ArcGIS 9. The basic spatial data model is known as "arc-node topology." One of the strengths of the vector data model is that it can be used to render geographic features with great precision.. However, this comes at the cost of greater complexity in data structures, which sometimes translates to slow processing speed. Most of the features you see on printed maps are represented with vector data.

Points represent discrete locations on the ground. Either these are true points, such as the point marked by a USGS brass cap, such as a section corner, or they may be virtual points, based on the scale of representation. For example, a city's location on a driving map of the United States is represented by a point, even though in reality a city has area. As the map's scale increases, the city will soon appear as a polygon. Beyond a certain scale of zoom (i.e., when the map's extent is completely within the city), there will be no representation of the city at all; it will simply be the background of the map.

Here is a view of the Puget Sound area with airports. The airports are stored as points within the GIS.

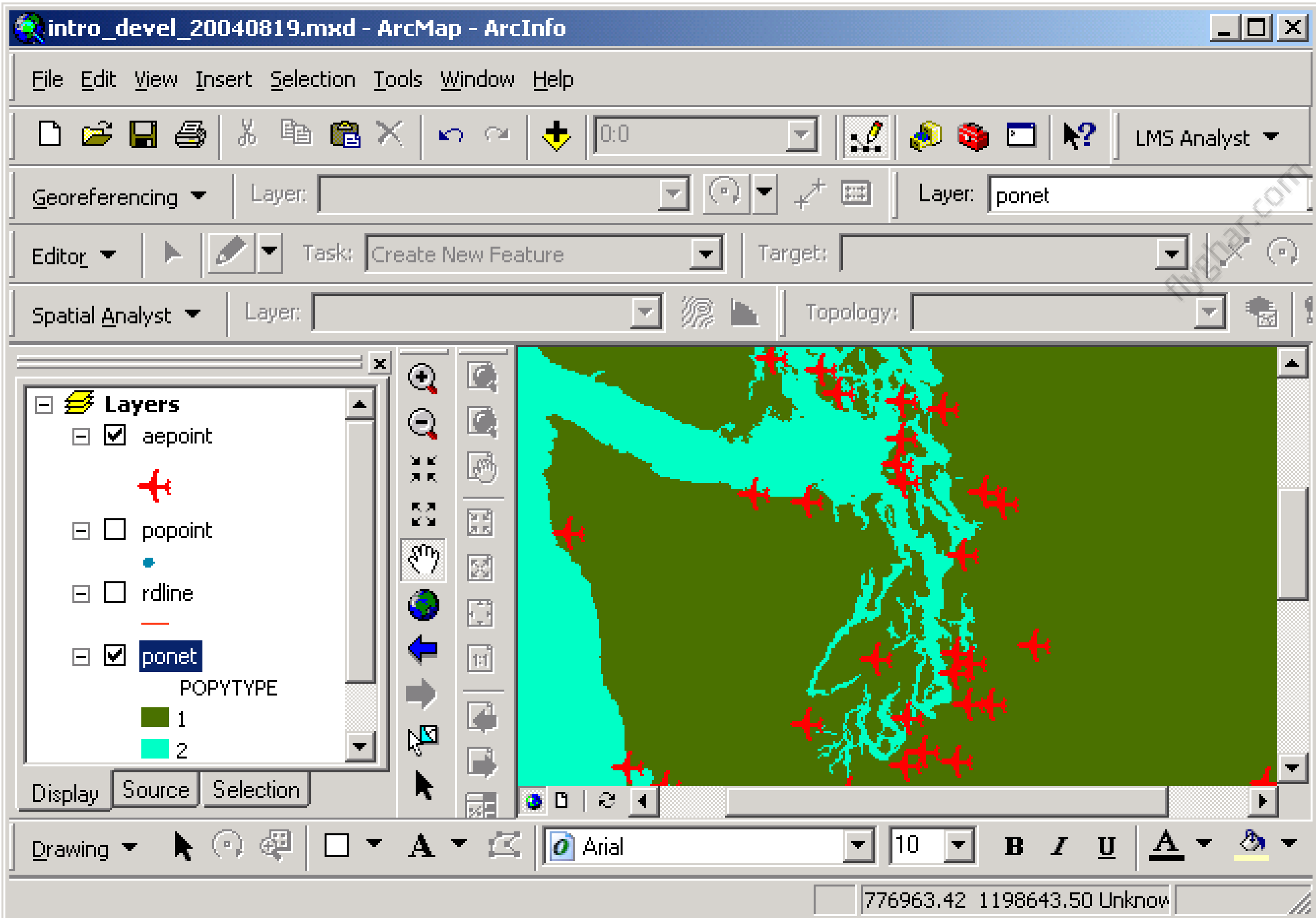
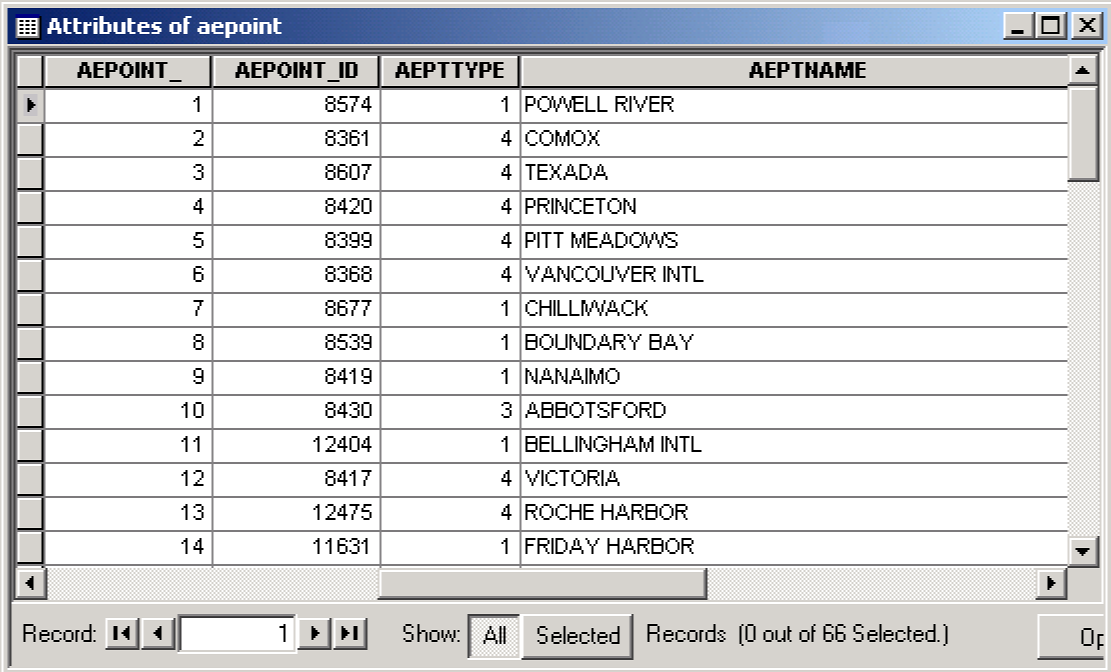


Figure 1

There will be more discussion on attribute tables in the next section, the Relational Database Model, but the tabular data model is worth mentioning here. Each layer is a combination of the coordinate (vector) data, and an attribute table containing a record for each vector feature. The records hold attributes for the feature, such as city name, sampling point number, or radio tower frequency. In this example, airports represented as points, and are associated with their name as well as with other codes in the point layer attribute table.



AEPOINT_	AEPOINT_ID	AEPTTYPE	AEPTNAME
1	8574	1	POWELL RIVER
2	8361	4	COMOX
3	8607	4	TEXADA
4	8420	4	PRINCETON
5	8399	4	PITT MEADOWS
6	8368	4	VANCOUVER INTL
7	8677	1	CHILLIWACK
8	8539	1	BOUNDARY BAY
9	8419	1	NANAIMO
10	8430	3	ABBOTSFORD
11	12404	1	BELLINGHAM INTL
12	8417	4	VICTORIA
13	12475	4	ROCHE HARBOR
14	11631	1	FRIDAY HARBOR

Figure 2

Lines represent linear features, such as rivers, roads and transmission cables. Here are major roads in the Puget Sound region, along with line attributes. In ArcGIS, lines are also known as "arcs," hence the name "ArcGIS." Each line is composed of a number of different coordinates, which make up the shape of the line, as well as the tabular record for the line vector feature.

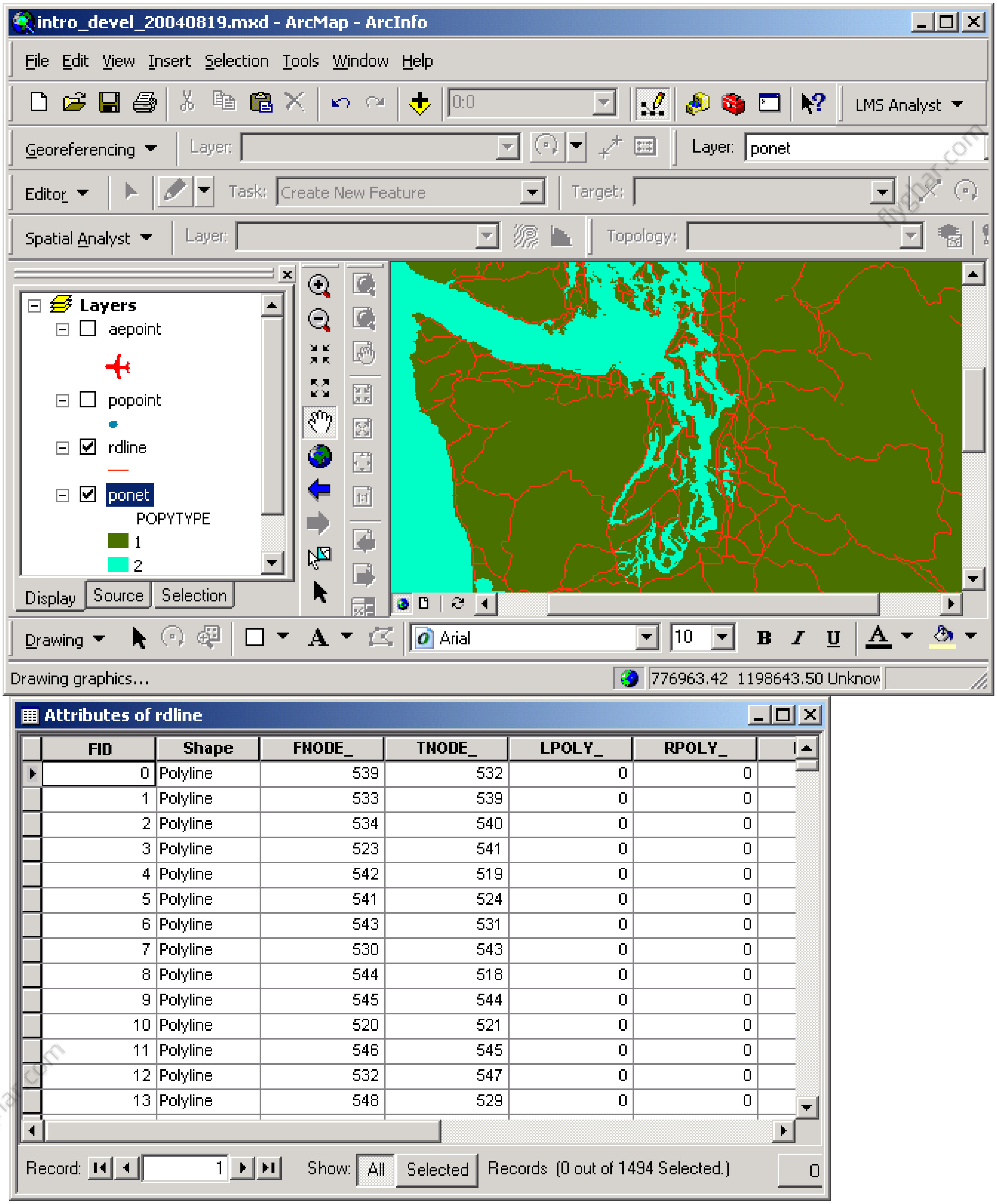

intro_devel_20040819.mxd - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help
LMS Analyst
Georeferencing
Layer:
Layer: ponet
Editor
Task: Create New Feature
Target:
Spatial Analyst
Topology:
Layers
aepoint
popoint
rdline
ponet
POPYTYPE
1
2
Display
Source
Selection
Drawing
Arial
10
Drawing graphics...
776963.42 1198643.50 Unknow
Attributes of rdline
FID Shape FNODE_ TNODE_ LPOLY_ RPOLY_
0 Polyline 539 532 0 0
1 Polyline 533 539 0 0
2 Polyline 534 540 0 0
3 Polyline 523 541 0 0
4 Polyline 542 519 0 0
5 Polyline 541 524 0 0
6 Polyline 543 531 0 0
7 Polyline 530 543 0 0
8 Polyline 544 518 0 0
9 Polyline 545 544 0 0
10 Polyline 520 521 0 0
11 Polyline 546 545 0 0
12 Polyline 532 547 0 0
13 Polyline 548 529 0 0
Record: 1
Show: All Selected
Records (0 out of 1494 Selected.)
0

Figure 3

Arcs are composed of nodes and vertices. Arcs begin and end at nodes, and may have 0 or more vertices between the nodes. The vertices define the shape of the arc along its length. Arcs which connect to each other will share a common node.

Sometimes several arcs are connected and share common attributes. These arcs may appear to be a continuous line, but in reality, the "line" is composed of multiple features, each with its own record.

The arc-node topology data model is central to many ArcGIS vector operations. Arcs are represented with starting and ending nodes, which imparts directionality to the arcs. In the image below, arc #1 starts at **node** #2 and ends at **node** #1, passing through several vertices along its way. Each of the nodes and vertices is stored with coordinate values representing real-world locations in a real-world coordinate system (e.g., longitude/latitude angles, State Plane feet, or UTM meters). These coordinate values represent locations you could locate using the numeric values printed on the graticule on the edges of a paper map. The storage of real-world coordinate values for features stored in the GIS is known as **georeferencing**. If features stored in the GIS are referenced to real world locations, the features are said to be **geoferenced**.

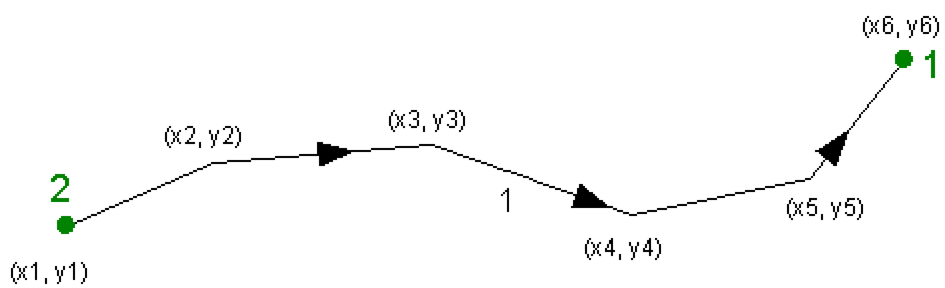


Figure 4

Note that each node and vertex has particular X and Y coordinates values. These X and Y coordinates form the basis of the location of features in coordinate (map) space. All vector features (points, lines, and polygons) are composed of locations defined by particular coordinate values. In the ArcGIS shapefile data model, the coordinate values on points, nodes, and vertices are stored within the dataset as "hidden" values on a feature-by-feature basis.

Polygons form bounded areas. In the point and line datasets shown above, the land masses, islands, and water features are represented as polygons. Polygons are formed by bounding arcs, which keep track of the location of each polygon, as shown in this image:

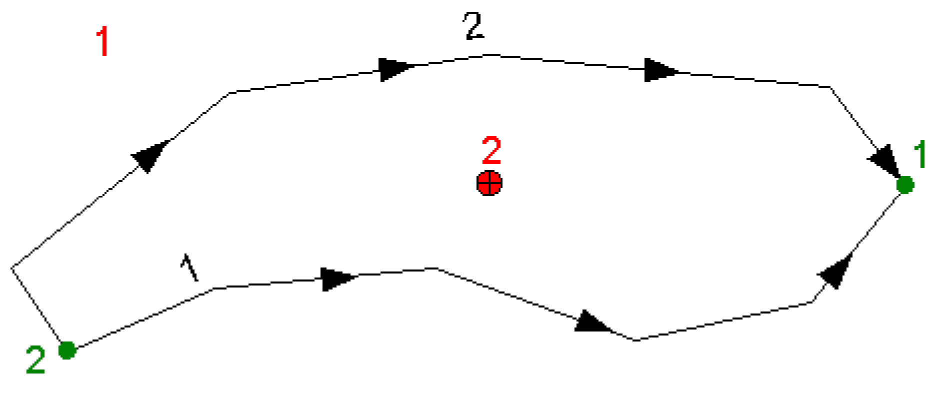


Figure 5

Polygon #1 is the "universe polygon," representing the outside of all other polygons. Polygon #2 (which is also denoted by **labelpoint** 2) is bounded by **arcs** 1 and 2. (Although all of these features have coordinates, they are not displayed here as they were in the previous image.)

Now consider arc #2. It starts at node #2, and ends at node #1. Following its direction, on its left-hand side is polygon #1 (which needs no label as the "universe" polygon), and on its right side is polygon #2 (denoted by **labelpoint** 2).

Continuing with this pattern, look at the same dataset with one added polygon:

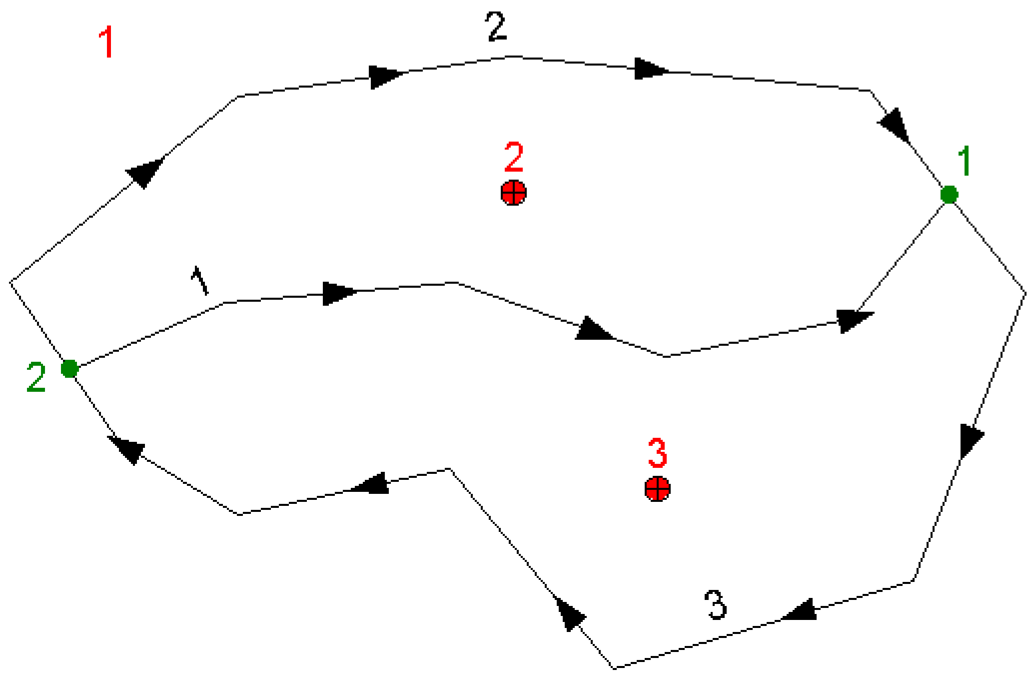


Figure 6

You can see now that managing vector datasets is complex. Each **node**, **label point**, **vertex**, **line**, must be stored with explicit coordinates, and the software also needs to keep track of the spatial relationships of each feature as well as the relational link with the tabular data. Imagine a dataset composed of several hundred thousand polygons; this dataset will take a lot of computing resources to manage.

Tics manage the georeferencing of almost all ArcInfo vector datasets. While tics are ***not*** a part of the native ArcGIS shapefile data specification, they are central to the ArcInfo vector data model, and because many of our data sources are ArcInfo coverage, this feature deserves to be mentioned. Tics are registered to ground coordinates, and all other data in the coverage are in turn referenced to the tics (which in turn makes the features referenced to ground coordinates).

Below is an image that shows all of the major features of an ArcInfo polygon dataset (technically known as an "ArcInfo coverage"). Tics are shown in cyan; arcs in black; nodes in green, and polygon label points in red.

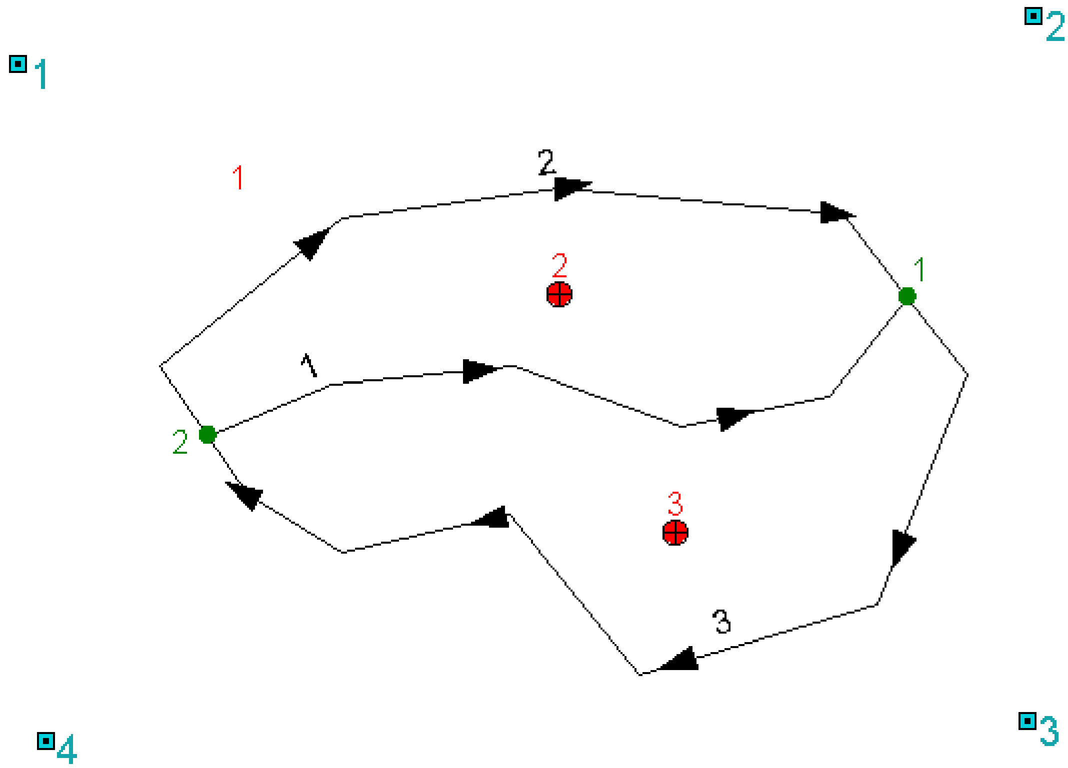


Figure 7

Although these spatial relationships are complex, you, as the user, do not need to keep track of these relationships; the software keeps track of all this for you. While a human can easily see the relationships among all of these features, the software lacks intelligence, and must therefore explicitly store all of the locations and relationships of features in digital files.

ArcGIS can automatically store line length and polygon area and perimeter in the units of measure defined by the user. These units can be feet, meters, or whatever units the user needs. The topic of units will be covered in more detail in the section on map projection.

In ArcGIS there are 3 major types of vector data source files:

- ESRI geodatabases
- ESRI shape files
- ArcInfo coverages

ESRI **geodatabases** are a relatively new format. Geodatabases are databases stored in Microsoft Access (for the "personal" geodatabase), as a special collection of files (for the "file-based" geodatabase), or higher-end applications (e.g. SQL Server, Oracle, Informix). A geodatabase stores all features and related tables, as well as other files, within a single or distributed database format. There are several advantages to using geodatabases rather than other storage formats: portability, integrity, validation of allowed data values, storage of data relationships as part of the data structure, topological rules, etc. At ArcGIS, it is possible to create geoprocessing models for complex analyses, as well as toolboxes containing custom tools, and have these stored in a geodatabase. In terms of the basic representation of spatial features (points, lines, and polygons), and their spatial referencing, the geodatabase functions the same as the other formats. To find out more about geodatabases, read the ArcGIS help topics. ESRI also publishes a textbook on data modeling that is focused around the geodatabase.

ESRI **Shape Files** are used mainly in ArcView 3.x and ArcGIS, although supported in other software as well. Because of the simple data and file structure, shapefiles draw very quickly in ArcGIS 10. Shapefiles can be fully managed (created, edited, and deleted) within ArcGIS 10's environment. Shapefile data files can also be managed using operating system tools, such as the Windows Explorer. The shapefile standard is public, so any software can be made to read or write shapefiles.

A single shapefile represents features that are either point, line, or polygon in spatial data type. If you create a shapefile, you need to choose what feature type you want at the time of creation.

Spatial referencing of shapefiles is enforced by maintaining explicit X and Y coordinates for each point or vertex in the layer. Typically, this is done at the time of data creation, where a new dataset is drawn in reference to existing datasets that are already georeferenced.

For each shapefile there exist at least 3 files, the shape data (stored in the **.shp** file), an associated dBASE (relational database) table (stored in the **.dbf** file), and a spatial index (stored in the **.shx** file). For each feature within the shapefile, there is an associated record within the attribute table. We will cover the operating system's file structure for shapefiles later in the quarter during the module on Project and Data Management.

ArcInfo **Coverages** are a basic implementation of the vector arc-node topological model as shown in the cartoon-like images above. Like shapefiles, coverages also have associated database tables, with a one-to-one feature-to-record relationship.

As you can see in the schematic for the ArcInfo polygon coverage, the coverage can be a multi-feature, or "polymorphic" dataset, composed of polygons, arcs, nodes, label points, and tics. This is due to the original way features were modeled in some of the first GIS software applications. It is a complex and sometimes unwieldy way to store data, but because many systems still use ArcInfo as the main GIS software, the coverage will be around for quite some time.

Because of the complex structure of some coverage data, coverages can take a long time to draw. Also, due to intricacies of the file system storage model, coverages cannot be managed fully (i.e., created, edited, deleted) without the use of ArcInfo software. Operating system tools for file management (e.g., the Windows Explorer) cannot manage ArcInfo datasets without corrupting them. For these reasons, ArcInfo coverages are often used as sources in ArcGIS, but they are frequently converted to the shapefile format for other uses.

Spatial referencing of ArcInfo coverages is enforced by tics, and all other features within the coverage are spatially referenced relative to the tics.

Most of the datasets we will be using throughout the term are ArcInfo coverages. Within ArcGIS, there are only minor differences between the functionality of the shapefile and the coverage.

The ArcInfo dataset file structure causes all sorts of problems for ArcGIS 10 users. Some of these problems are avoidable, while some are not. "Dealing" with ArcInfo datasets in ArcGIS is covered in Project and Data Management.

For polygon features, shapefiles and coverages have very different spatial data structures. The greatest differences can be shown comparing how polygons are stored. While coverages use the standard arc-node topology data structure, in which adjacent polygons share common bounding arcs, shapefiles store each feature as a separate object.

Here is a close up view of a few adjacent coverage polygons. When one bounding arc is moved, both polygons are affected.

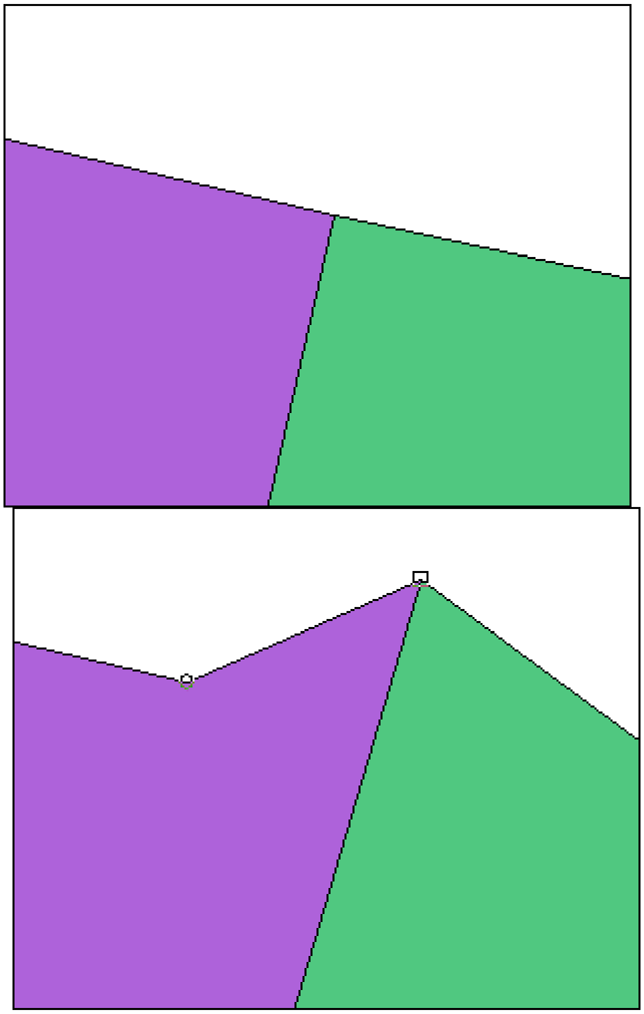


Figure 8

Compare this to a shapefile, in which different polygons can be moved without affecting adjacent polygons:

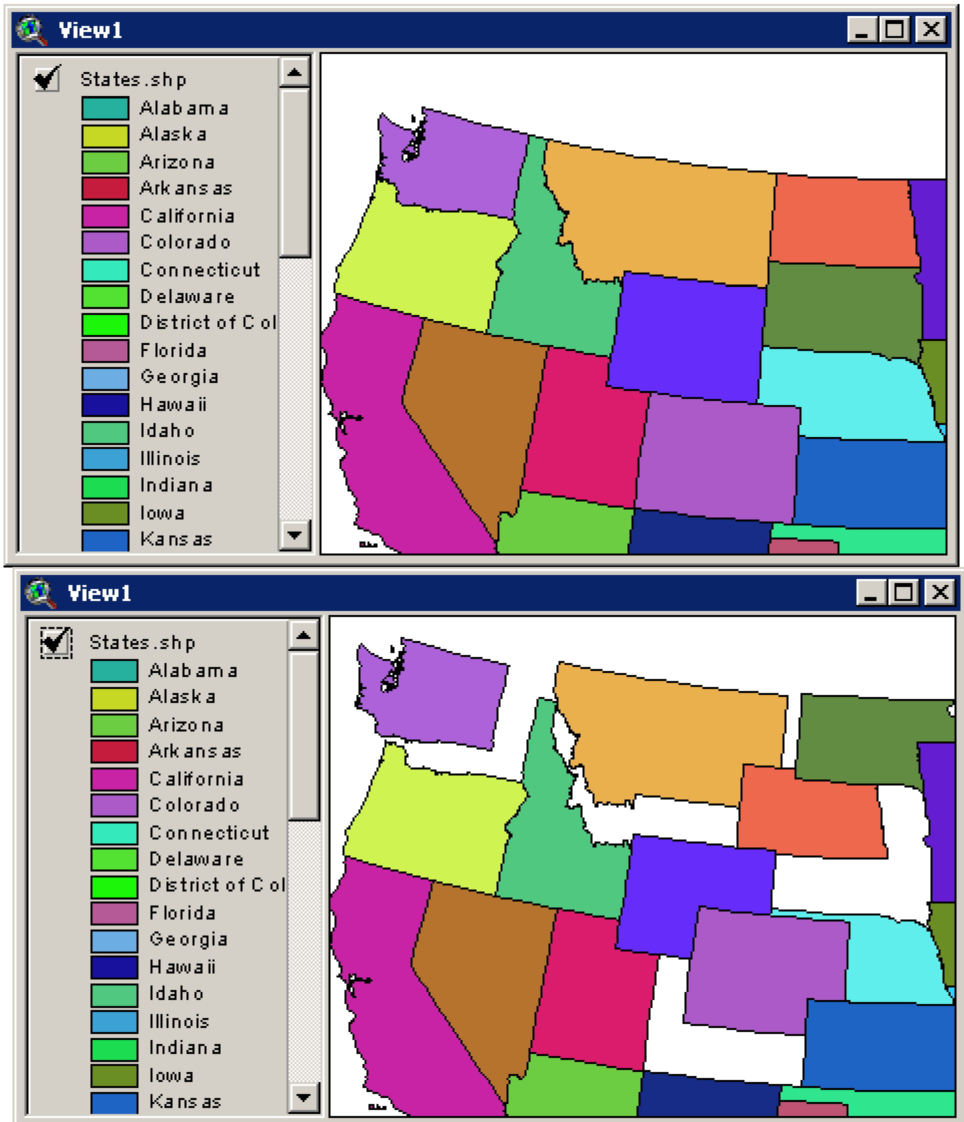


Figure 9

The geodatabase can function in either of the ways shown above. While polygonal features in a geodatabase are stored as individual "rings" (as in Figure 8), it is possible to define topological relationships so that adjacent features are not able to be edited independently (as in Figure 9).

ASCII Coordinate Data files may also be used in ArcGIS. Point layers can be created from files containing single records for individual points. The source files which single records for individual points, where each record contains X and Y coordinates, as well as any other optional attribute data. In this example, there is a different record for each point representing a populated place in the dataset.

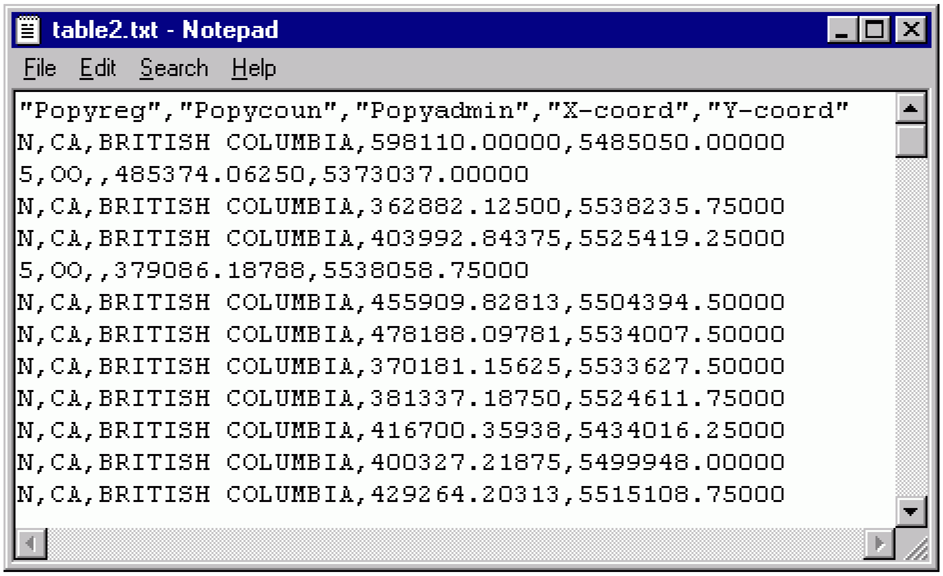


table2.txt - Notepad

File Edit Search Help

```
"Popyreg","Popycoun","Popyadmin","X-coord","Y-coord"
N,CA,BRITISH COLUMBIA,598110.00000,5485050.00000
5,OO,,485374.06250,5373037.00000
N,CA,BRITISH COLUMBIA,362882.12500,5538235.75000
N,CA,BRITISH COLUMBIA,403992.84375,5525419.25000
5,OO,,379086.18788,5538058.75000
N,CA,BRITISH COLUMBIA,455909.82813,5504394.50000
N,CA,BRITISH COLUMBIA,478188.09781,5534007.50000
N,CA,BRITISH COLUMBIA,370181.15625,5533627.50000
N,CA,BRITISH COLUMBIA,381337.18750,5524611.75000
N,CA,BRITISH COLUMBIA,416700.35938,5434016.25000
N,CA,BRITISH COLUMBIA,400327.21875,5499948.00000
N,CA,BRITISH COLUMBIA,429264.20313,5515108.75000
```


The coordinates can be represented as points on a map.

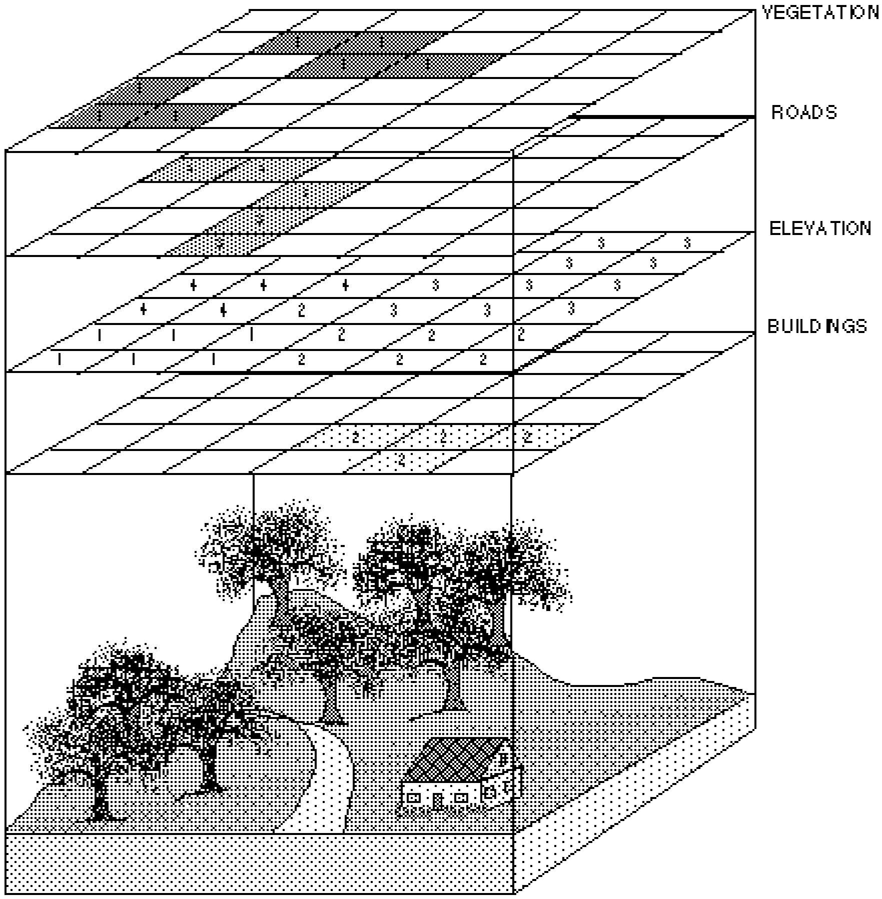
Vector data scale dependency

For all vector datasets, you should always consider the scale dependency of spatial data. When should an airport be represented as a point, and when should it be a polygon? If you are measuring the distance from major cities to their airports, then the cities and airports would be best represented as points. However, if you are planning wetland mitigation for an addition to an airport, then the airport boundary would be better represented as a polygon.

Raster datasets are composed of rectangular arrays of regularly spaced square grid cells. Each cell has a value, representing a property or attribute of interest. While any type of geographic data can be stored in raster format, raster datasets are especially suited to the representation of continuous, rather than discrete, data. Some examples of continuous data are:

- oil depth across an open-water oil spill
- soil pH
- reflectance in a certain band in the electromagnetic spectrum
- elevation
- landform aspect (compass bearing of steepest downward descent)
- salinity of a water body

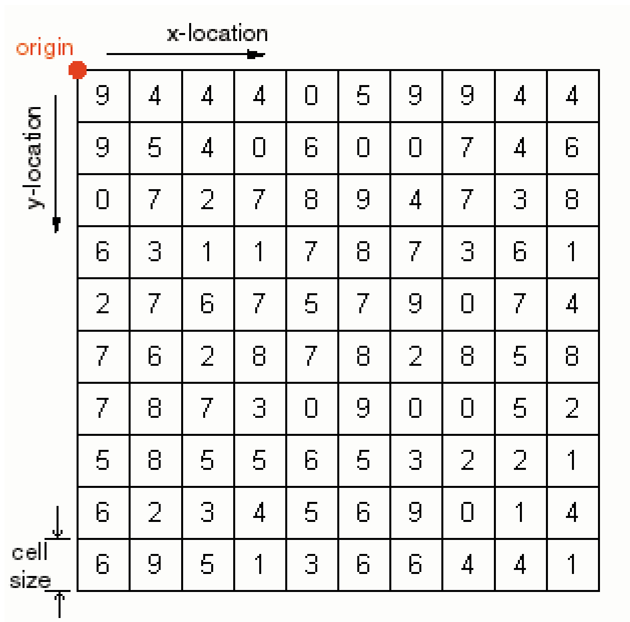
Here is a diagrammatic model of how raster datasets represent real-world features:



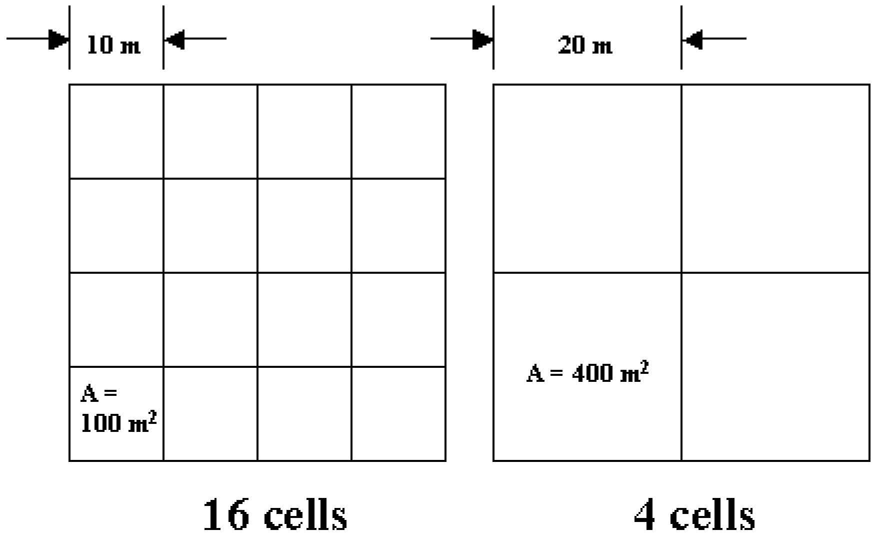
Generally, cells are assigned a single numeric value, but with GRID (a proprietary ArcInfo data format) layers, cell values can also contain additional text and numeric attributes. ArcInfo format grids are the native raster dataset for ArcGIS as well as ArcInfo. In the above diagram, each feature type on the landscape (buildings, elevation, roads, vegetation) is represented in its own raster layer. Note that each raster layer has cells with numbers.

- For the buildings layer, all cell values are 2 (in this case, 2 is a code for houses; other buildings would be encoded with a different value).
- For the elevation layer, the cell value is the elevation at the center of the cell.
- For roads, a value of 3 indicates a road (other road features, e.g., highways, would have a different code).
- For vegetation, trees have a value of 1. In this example, grass is treated as a background value and has no data value (although it could have been given a different numeric value).

All raster datasets are spatially referenced by a very simple method: only one corner of the raster layer is georeferenced. Because cell size is constant in both X and Y directions, cell locations are referenced by row/column designations, rather than with explicit coordinates for the location of each cell's center. This image shows the upper-left corner as the grid origin, with arrows representing the X and Y location of the cells. Different raster file formats may have an origin located at the lower left rather than at the upper left. Each cell or pixel contains a value representing some numerical phenomenon, or a code use for referencing to a non-numerical value.



Whereas with vector data, each point, node, and vertex has an explicit and absolute coordinate location, raster cells are georeferenced relative to the layer's coordinate origin. This speeds up processing time immensely in comparison to certain types of vector data processing. However, the file sizes of raster datasets can be very large in comparison to vector datasets representing the same phenomenon for the same spatial area. Also, there is a geometric relationship between raster resolution and file size. A raster dataset with cells half as large (e.g., 10 m on a side instead of 20 m on a side) may take up 4 times as much storage space, because it takes four 10 m cells to fit in the space of a single 20 m cell. The following image shows the difference in cell sizes, area, and number of cells for two configurations of the same total area:



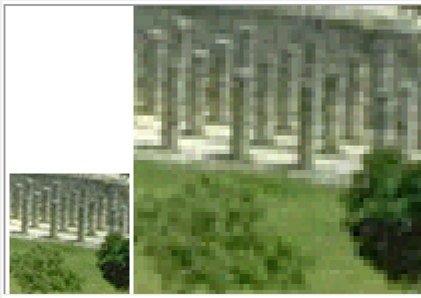
Cells may either have a value (0-infinity) or no value (null, or no data). The difference between these is important. Null values mean that data either fall outside the study area boundary, or that data were either not collected or not available for those cells. In general, when null cells are used in analysis, the output value at the same cell location is also a null value. Grid datasets can store either integer or floating-point (decimal) data values; though some other data formats can only store integer values. Typical simple image data will have strict limits on the number of unique cell values (typically 0-255).

Pixel or **cell**? All raster datasets are stored in similar formats. You will want to know the difference between a pixel and a cell, even though they are functionally equivalent. A pixel (short for PICture ELement) represents the smallest resolvable "piece" of a scanned image, whereas a cell represents a user-defined area representing a phenomenon. A pixel is always a cell, but a cell is not always a pixel.

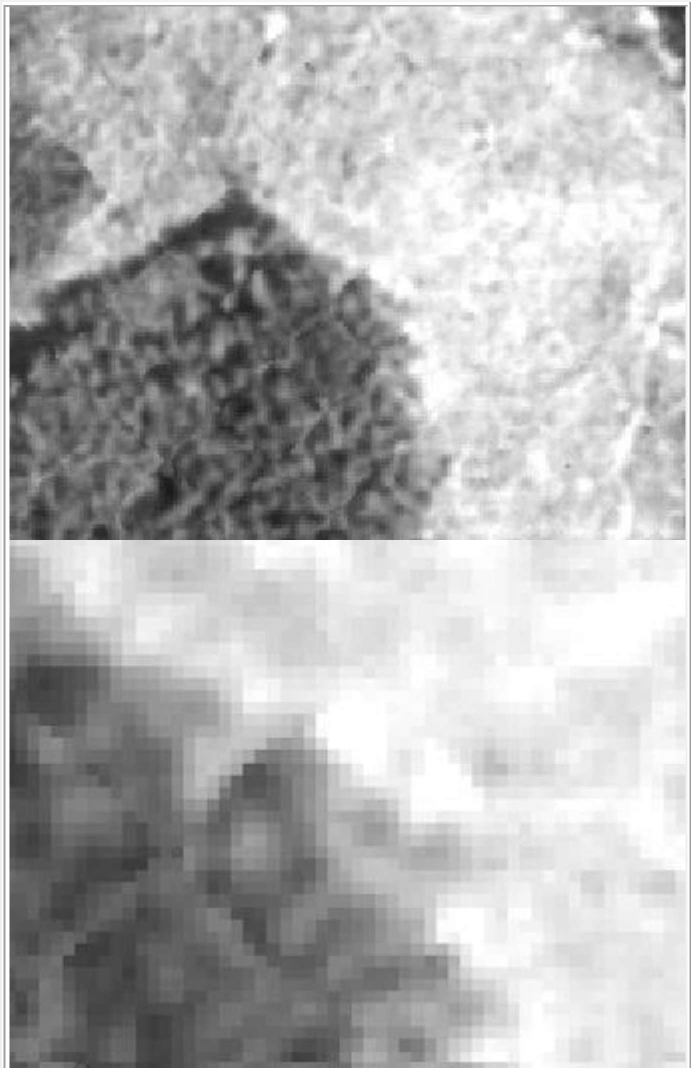
There are many types of raster data you may be familiar with:

- grids (ArcGIS & ArcInfo specific)
- graphical images (TIFF, JPEG, BMP, GIF, etc.)
- USGS DEM (Digital Elevation Model)
- remotely-sensed images (Landsat, SPOT, AVIRIS, AVHRR, Imagine IMG, digital orthophotos)

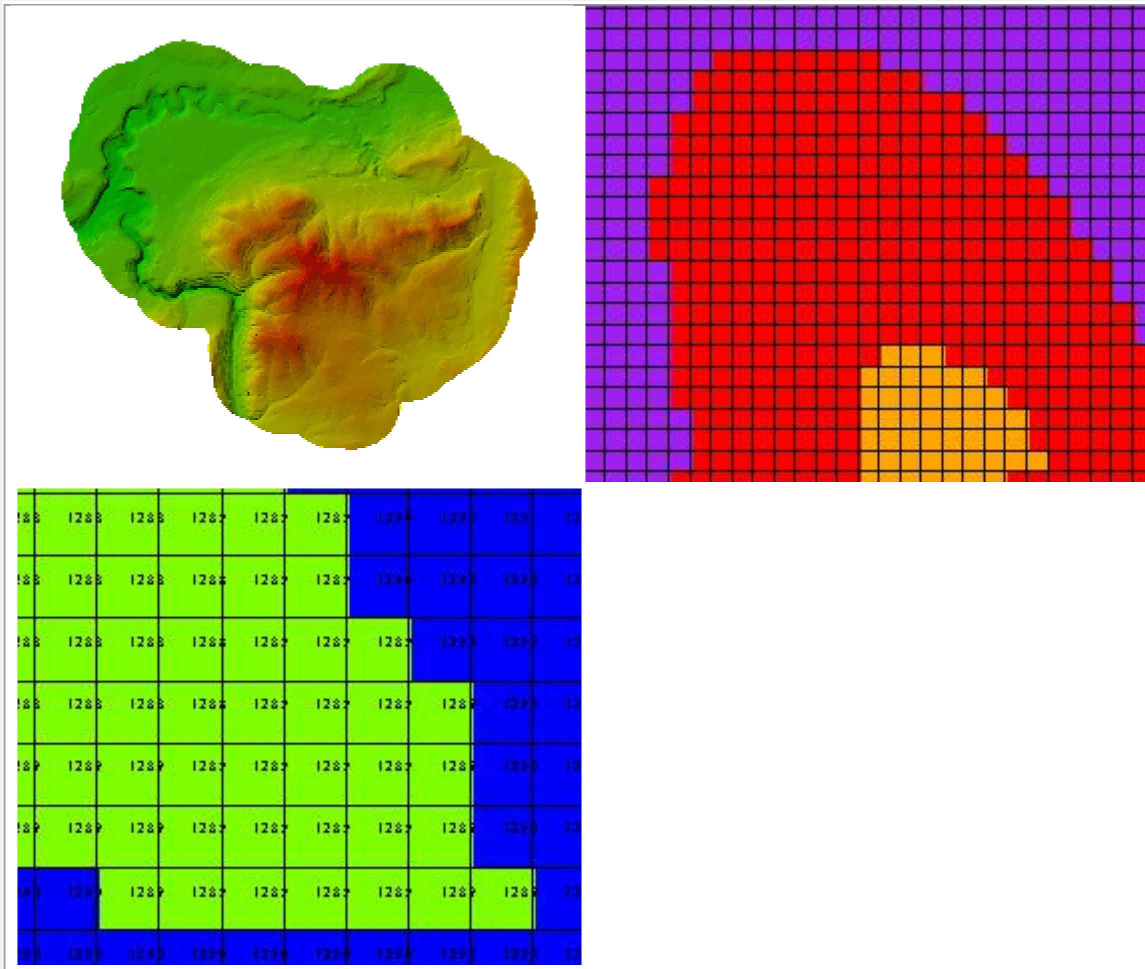
All raster datasets have essentially the same tessellated structure. Here are few graphical examples of raster data. Note that each image, when zoomed in, shows the same pixilated structure.



A simple scanned color photographic image (GIF format).



Some 8-bit digital orthophotos (in BIL, "band interleave by line" format).



Three differently scaled views of an ArcInfo format elevation grid, showing cell outlines and elevation values.

Most of the raster datasets we will use are single-band, which means that they contain a single "layer" of data. The data can represent elevation, slope, or reflectance (in the case of the black-and-white digital orthophotos).

These single-band images are viewed with a color mapping, so that the cell value is associated with a particular color. For the orthophotos, the color map is a 256-value grayscale ramp. Other raster data, such as elevation models, can be mapped to color ramps that display elevation ranges, as shown in the image directly above. Most GIF files have a limit of 256 unique values (this is known as 8-bit data, because 2^8 = 256).

Multi-band raster data (such as RGB images or satellite images) are generally displayed with a mixture of red, green, and blue values for each different band in the image.

1. Identify the Spatial Phenomenon: In this example, the spatial phenomenon is a river network, consisting of a series of interconnected streams, creeks, and rivers.
2. Define the Spatial Model: Based on the spatial phenomenon, a spatial model can be created to represent the river network. One way to do this is to use a network model, where the river network is represented as a set of nodes (where streams and rivers intersect) and edges (where streams and rivers flow between nodes).
3. Define the Attributes: In addition to the spatial model, attributes can be defined to describe the river network, such as stream order (the hierarchy of streams and rivers in the network), stream length, stream slope, and stream flow rate.
4. Create the Spatial Data Model: Using the network model and attribute information, a spatial data model can be created to represent the river network. This might involve creating a set of tables or feature classes in a GIS software package, such as ArcGIS or QGIS.
5. Populate the Model with Data: With the spatial data model in place, data can be added to the model to represent the river network. This could include digitizing streams and rivers from aerial photos or satellite imagery, or importing data from other sources such as government agencies or non-profit organizations.
6. Analyze and Visualize the Data: Once the data is added to the spatial data model, it can be analyzed and visualized using GIS software. For example, the model could be used to generate maps of the river network, calculate stream order, estimate water flow rates, or identify areas of the network that are at risk of flooding.

In conclusion, spatial data modelling involves creating a mathematical representation of spatial phenomena, such as a river network, using a combination of spatial models, attributes, and GIS software tools. By following these steps, analysts and researchers can gain insight into complex spatial relationships and patterns that are difficult to observe directly in the real world.

Vector based analysis

Vector-based analysis in GIS involves using vector data, such as points, lines, and polygons, to perform spatial analysis. Here are some common vector-based analysis techniques in GIS:

1. Overlay Analysis: Overlay analysis involves overlaying two or more vector data layers to create a new layer that combines the attributes and spatial relationships of the original layers. This can be used to identify areas where multiple features overlap or where different features meet, such as where roads intersect with rivers.

2. Buffer Analysis: Buffer analysis involves creating a buffer around a vector feature, such as a point or polygon. The buffer is typically a defined distance from the feature and can be used to identify areas within a certain proximity to the feature. For example, buffer analysis can be used to identify areas within a certain distance of a school or hospital.
3. Network Analysis: Network analysis involves using vector data to model transportation networks, such as roads or railroads, and analyzing the network to determine the shortest or fastest routes between two points. This can be used for applications such as route planning or logistics optimization.
4. Density Analysis: Density analysis involves using vector data to create a density surface that shows the density of features in a given area. This can be used to identify areas of high or low population density, for example, or to identify areas where certain types of features, such as trees or buildings, are concentrated.
5. Spatial Query: Spatial query involves querying a vector data layer to select features based on their spatial relationship to other features. For example, spatial query can be used to select all the points within a certain distance of a road, or all the polygons that intersect with a river.

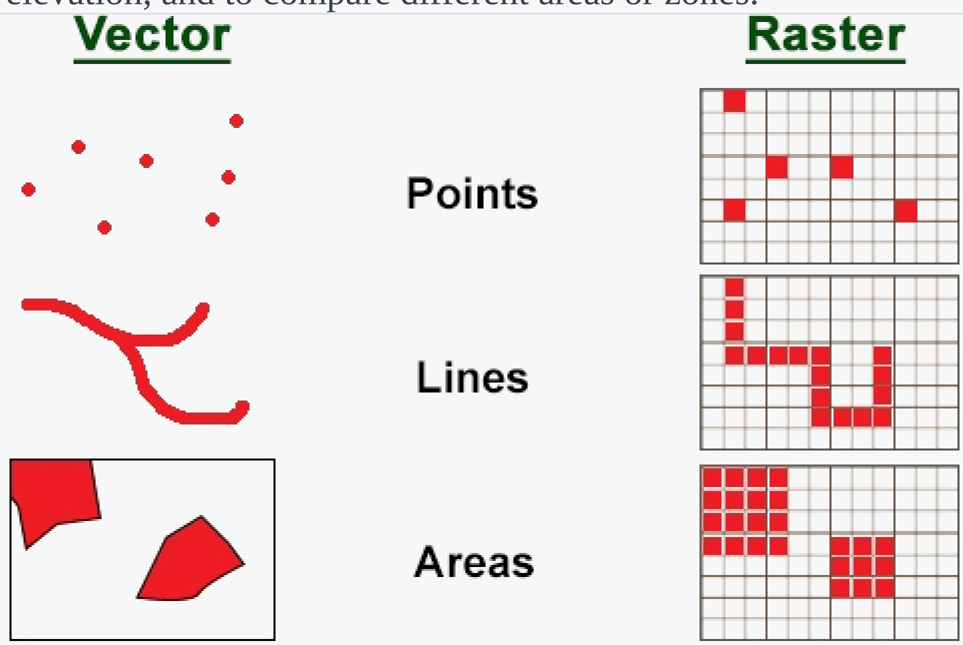
Raster based analysis

Raster-based analysis in GIS involves using raster data, such as satellite imagery or digital elevation models, to perform spatial analysis. Here are some common raster-based analysis techniques in GIS:

1. Classification: Classification involves grouping pixels in a raster image into different classes based on their spectral or radiometric properties. This can be used to identify land cover types, such as forest, agriculture, or urban areas, or to identify other features, such as water bodies or roads.
2. Reclassification: Reclassification involves changing the values of the cells in a raster dataset to create a new dataset with different classes or values. This can be used to simplify a raster dataset or to create a new dataset with different classifications, such as changing all cells with a value of 1 to a value of 0.
3. Distance Analysis: Distance analysis involves calculating the distance between cells in a raster dataset, either in a straight line or along a path that takes into account other factors, such as slope or land cover. This can be used to identify areas within a certain distance of a feature, such as a road or river, or to identify areas that are most accessible or least accessible.
4. Surface Analysis: Surface analysis involves calculating properties of a raster surface, such as slope, aspect, or curvature. This can be used to identify areas that

are most susceptible to erosion, for example, or to identify areas that are suitable for different types of development.

5. Zonal Statistics: Zonal statistics involves calculating statistics, such as mean or maximum, for groups of cells within a larger raster dataset. This can be used to identify areas that have similar characteristics, such as similar land cover or elevation, and to compare different areas or zones.



Vector and Raster based operation tools

Vector and raster data are the two primary types of digital data used in GIS (Geographic Information System), and each has its own set of operation tools.

Vector-based operation tools:

1. Buffering: Creates a buffer zone around a set of features based on a specified distance.
2. Clipping: Removes portions of one layer that overlap with another layer.
3. Union: Combines two or more layers into a single layer by merging their attributes and geometries.
4. Intersect: Creates a new layer from the overlapping portions of two or more layers.
5. Dissolve: Combines features that share the same attribute values into a single feature.
6. Spatial Join: Joins attributes from one layer to another based on their spatial relationship.

Raster-based operation tools:

1. Re-sampling: Changes the resolution of a raster dataset by interpolating new pixel values based on existing values.
2. Mosaic: Combines two or more raster datasets into a single raster dataset.
3. Clip: Removes portions of a raster dataset that fall outside a specified boundary.
4. Mask: Hides portions of a raster dataset based on a mask layer.
5. Reclassify: Changes the values of a raster dataset based on a new classification scheme.
6. Slope Analysis: Calculates the slope of a terrain based on a digital elevation model (DEM).

Overall, both vector and raster-based operation tools are essential for analyzing geographic

Unit 5: GIS data modeling and Creating Maps

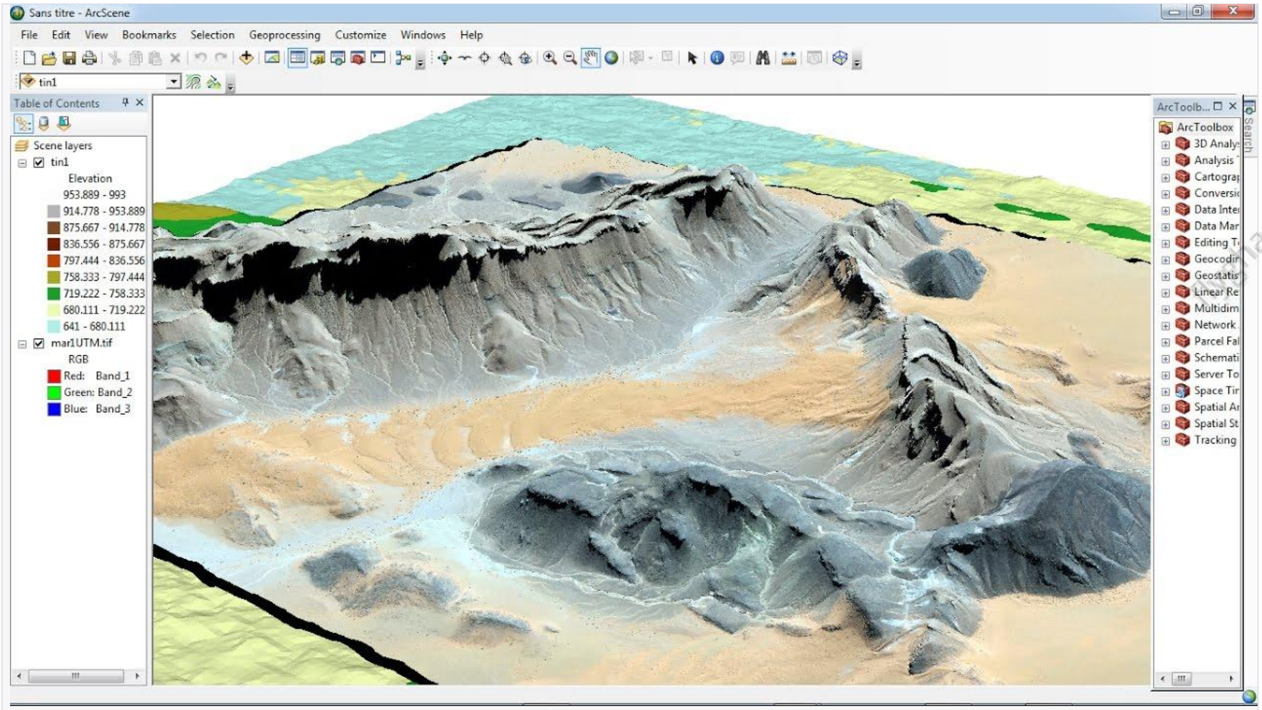
Surface Modeling

Surface modeling is the process of creating a digital representation of the surface of the earth or other objects in three-dimensional space. In GIS, surface modeling is typically done using raster or TIN data structures.

Raster-based surface modeling involves using a grid of cells to represent the surface, with each cell having a value that corresponds to the elevation or other attribute of the surface at that location. Common raster-based surface modeling techniques include interpolation, which involves estimating values for cells based on surrounding values, and hill shading, which involves using light and shadow to create a three-dimensional appearance of the surface.

TIN-based surface modeling involves creating a triangulated irregular network (TIN) that connects a set of discrete points, such as elevation measurements. The TIN is then used to interpolate values for the surface at locations between the points. TIN-based surface modeling can provide more accurate representations of surfaces with irregular shapes or features than raster-based approaches.

Surface modeling is used in a wide range of applications in GIS, including terrain analysis, hydrological modeling, and urban planning. By creating accurate representations of surfaces, GIS users can better understand the spatial relationships and patterns of the data they are analyzing, and make more informed decisions based on that understanding.



Surface Modeling Steps

Surface modeling in GIS (Geographic Information System) involves creating a 3D representation of the terrain surface using elevation data. Here are the general steps involved in surface modeling:

1. Data collection: The first step in surface modeling is collecting elevation data. This can be done using various techniques, such as airborne LiDAR, satellite imagery, or ground surveys.
2. Data preparation: The elevation data needs to be processed and prepared for surface modeling. This can involve cleaning the data, removing outliers, and filling in missing values.
3. Surface creation: Once the elevation data is prepared, the surface model can be created using surface modeling tools such as interpolation, triangulation, or contouring. This process involves creating a continuous surface representation of the terrain based on the elevation data.
4. Analysis: Once the surface model is created, it can be analyzed using various techniques such as slope analysis, aspect analysis, and visibility analysis.
5. Visualization: The final step in surface modeling is visualizing the results. This can be done using 3D visualization tools, such as fly-through, or by creating 2D maps that display the terrain surface and analysis results.

Surface Modeling Types

In Geographic Information Systems (GIS), surface modeling is used to represent the shape of the Earth's surface and to model changes to the surface over time. Here are some of the most common types of surface modeling used in GIS:

1. Digital Elevation Models (DEM): DEM is a digital representation of the elevation of the Earth's surface. It is created by collecting elevation data from various sources and then interpolating the data to create a continuous surface.
2. TIN (Triangular Irregular Network): TIN is a surface modeling technique that uses a network of triangles to represent the surface of the Earth. It is often used to create 3D terrain models.
3. Isolines: Isolines are lines that connect points of equal value, such as points of equal elevation. Isolines are often used in contour maps to represent changes in elevation over a large area.
4. Kriging: Kriging is a statistical technique used to interpolate data values from a set of points to create a continuous surface. It is often used to create maps of surface characteristics such as soil pH or water quality.
5. Interpolation: Interpolation is a general term used to describe a range of techniques used to estimate values for locations that are not directly measured. In GIS, interpolation techniques are often used to create continuous surfaces from point data.

These are some of the most common surface modeling techniques used in GIS. Each technique has its own strengths and weaknesses, and the choice of which one to use depends on the specific requirements of the project.

Advanteges of Modeling

There are several advantages to using modeling in various fields, including:

1. Improved understanding: Modeling allows for a better understanding of complex systems or phenomena. By creating a simplified version of a system or process, researchers can more easily identify the key variables and interactions that contribute to the outcome.
2. Prediction: Models can be used to make predictions about future events or outcomes. By inputting data and variables into a model, researchers can project what might happen under different scenarios or conditions.
3. Optimization: Models can be used to optimize processes or systems. By identifying the key variables and interactions, researchers can identify ways to improve efficiency, reduce costs, or enhance performance.

4. Cost-effective: Modeling can be a cost-effective way to test hypotheses or explore new ideas. Rather than conducting expensive experiments or surveys, researchers can create a model to test different scenarios and identify the most promising avenues for further investigation.
5. Risk reduction: Modeling can be used to assess risks and potential consequences. By simulating different scenarios, researchers can identify potential risks and develop strategies to mitigate or avoid them.
6. Communication: Models can be used to communicate complex ideas or data to a wider audience. By creating visual representations of data or processes, researchers can help non-experts understand complex concepts more easily.

Disadvantages of Modeling

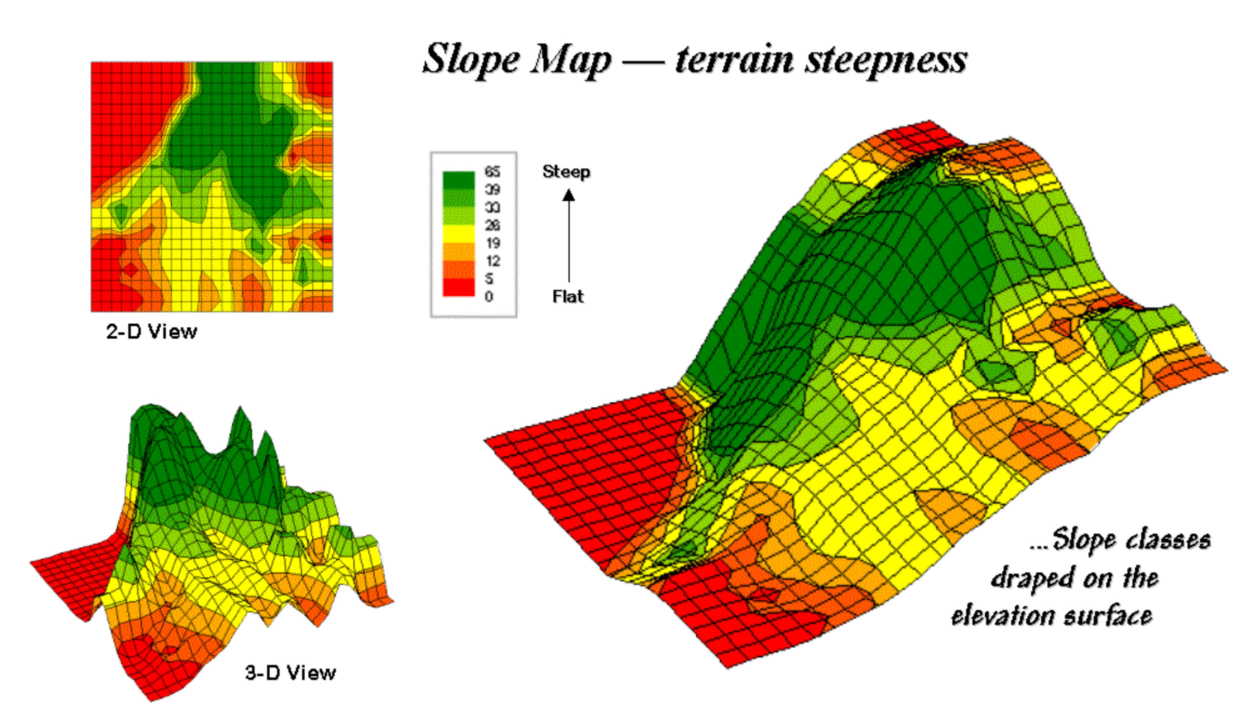
While spatial modeling has many advantages, there are also some disadvantages to consider, including:

1. Data limitations: Spatial modeling requires large amounts of data, which can be difficult to obtain or may not be available for certain regions or time periods. This can limit the accuracy and scope of the model.
2. Model complexity: Spatial modeling can be complex, particularly for those without advanced statistical or computational skills. Developing and running a spatial model requires a significant amount of technical expertise and resources.
3. Uncertainty: Spatial models are subject to uncertainty, particularly when dealing with complex systems or processes. Uncertainty can arise from incomplete or inaccurate data, as well as from the assumptions and simplifications made in the model.
4. Overreliance on technology: Spatial modeling relies heavily on technology, which can be a disadvantage if the technology fails or becomes obsolete. This can limit the usefulness and longevity of the model.
5. Lack of transparency: Spatial models can be difficult for non-experts to understand, which can limit their usefulness in decision-making processes. The assumptions and simplifications made in the model may not be transparent, making it difficult to assess the model's accuracy and reliability.
6. Interpretation: Spatial models are only as good as the data and assumptions used to create them. Interpretation of results can be difficult, and it may be unclear how to apply the results to real-world situations.

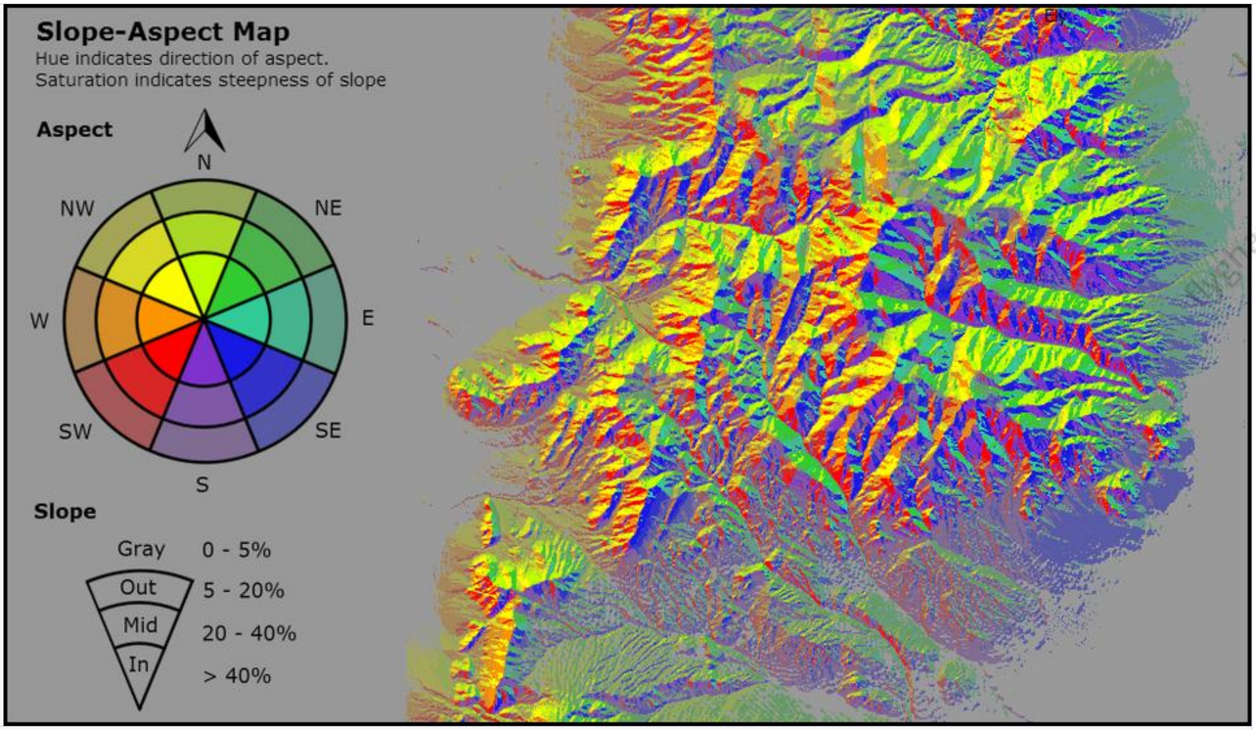
Surface Analysis

Surface analysis is a technique used in GIS (Geographic Information System) to study and understand the characteristics of the Earth's surface. It involves using surface modeling techniques to create a 3D representation of the terrain surface, and then performing various analyses on the surface to gain insights into its properties and characteristics. Some common surface analysis techniques include:

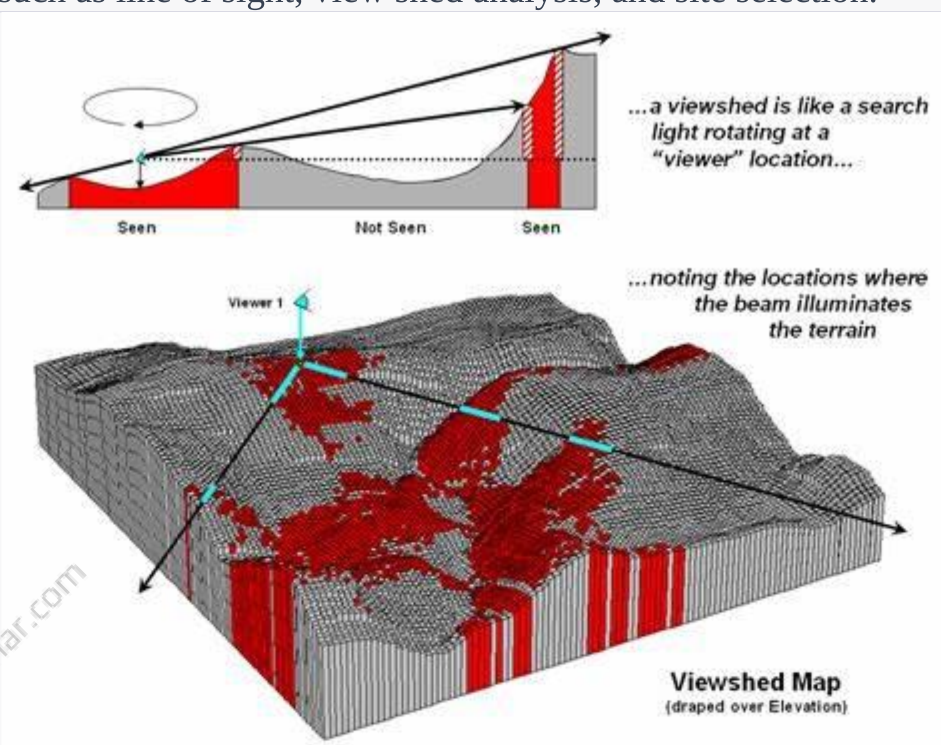
1. Slope analysis: This involves calculating the slope of the terrain surface at each point in the study area. Slope analysis is useful for understanding the steepness of the land, which can impact factors such as erosion, vegetation growth, and water runoff.



2. Aspect analysis: This involves calculating the aspect of the terrain surface at each point in the study area. Aspect refers to the direction that a slope faces, and is important for understanding factors such as solar radiation and temperature variation.



3. Visibility analysis: This involves calculating what areas of the terrain surface are visible from a given point. Visibility analysis is useful for understanding factors such as line of sight, view shed analysis, and site selection.



4. Terrain ruggedness index analysis: This involves calculating the terrain ruggedness index (TRI) of the terrain surface at each point in the study area. The TRI is a

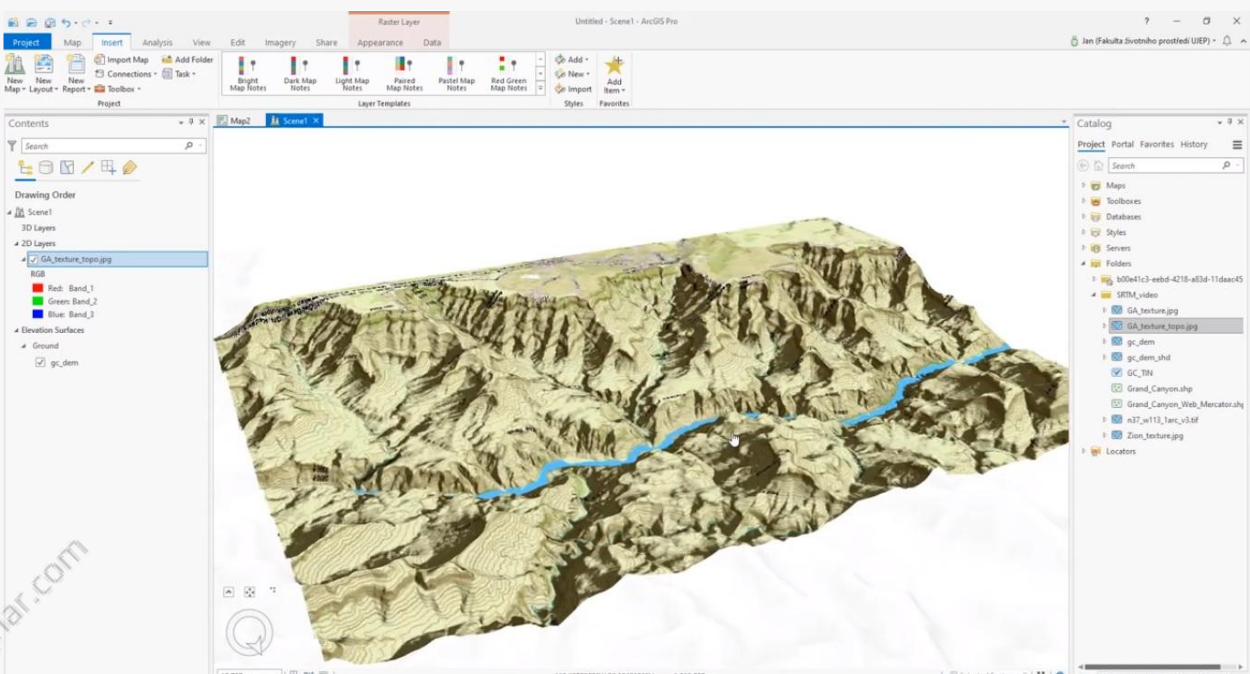
measure of how rough or uneven the terrain surface is, and is useful for understanding factors such as accessibility and wildlife habitat.

5. Landform classification: This involves classifying the terrain surface into different landforms, such as hills, valleys, and ridges. Landform classification is useful for understanding the physical characteristics of the land, which can impact factors such as vegetation growth and erosion.

<u>Hillshade analysis</u>

Hillshade analysis is a GIS technique used to create a shaded relief image of a terrain model, based on the angle and direction of the light source. The result is a three-dimensional representation of the terrain, with shadow and highlight areas that simulate the effect of sunlight on the landscape.

Hillshade analysis can be used for a variety of applications, including cartography, landscape visualization, and natural resource management. It is often used in conjunction with other GIS analysis techniques, such as slope analysis and viewshed analysis, to create more detailed and informative maps.



The process of hillshade analysis involves the following steps:

1. Select a terrain model: The first step in hillshade analysis is to select a terrain model, such as a digital elevation model (DEM), that represents the topography of the study area.
2. Define light source: The next step is to define the angle and direction of the light source, which will determine the shadows and highlights in the final hillshade image. This can be done manually or using a GIS software tool.
3. Calculate hillshade: Once the light source is defined, the GIS software can calculate the hillshade for each pixel in the terrain model. This involves calculating the slope and aspect of each pixel, and simulating the effect of the light source on the terrain.
4. Adjust hillshade parameters: The final step in hillshade analysis is to adjust the hillshade parameters, such as brightness and contrast, to optimize the visual appearance of the hillshade image. This can be done using a GIS software tool or image processing software.

Hillshade analysis can provide valuable information about the terrain, such as the orientation and shape of landforms, the presence of slopes and ridges, and the distribution of light and shade. However, it is important to keep in mind that hillshade analysis is an artificial representation of the terrain, and should be used in conjunction with other GIS analysis techniques and field data to ensure accurate interpretation and analysis.

Viewshade analysis

Viewshed analysis is a GIS technique used to determine the visibility of a point or area from a specified vantage point or set of vantage points. It can be used to identify areas that are visible or hidden from a particular location, and to assess the potential visual impact of development projects, land use changes, or other environmental factors.

Viewshed analysis involves the following steps:

1. Choose the observer point: The first step in viewshed analysis is to choose the observer point or set of observer points. These points represent the location from which the visibility will be analyzed.
2. Define the visibility range: The next step is to define the visibility range, which represents the maximum distance that an object can be seen from the observer point. This can be based on factors such as terrain elevation, vegetation cover, and atmospheric conditions.
3. Generate a digital elevation model: A digital elevation model (DEM) of the study area is needed to perform viewshed analysis. This model represents the topography

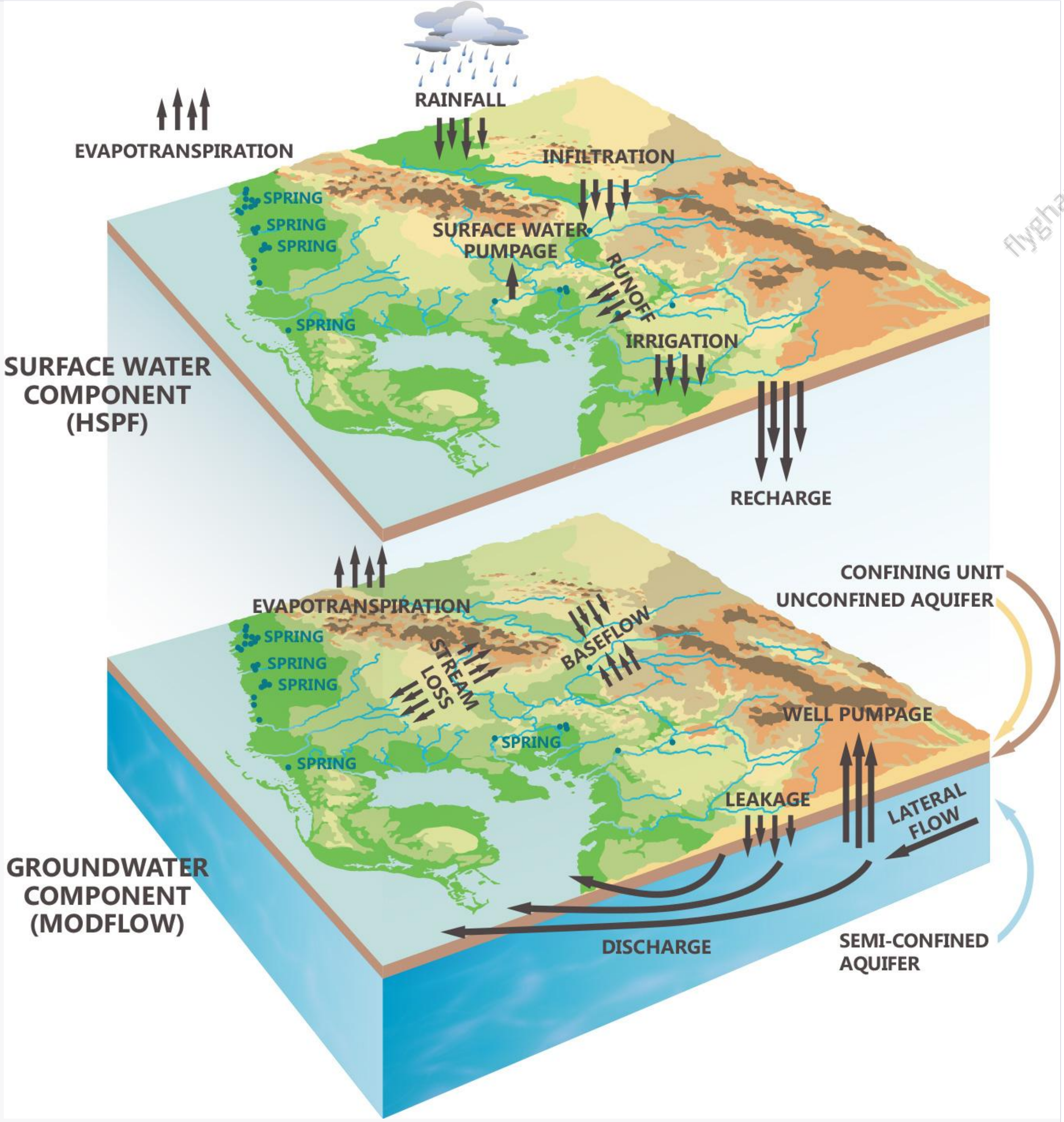
of the study area, and can be obtained from various sources such as aerial photography, LiDAR, or satellite imagery.

4. Perform viewshed analysis: Using the observer point and visibility range, the GIS software can perform viewshed analysis to identify areas that are visible or hidden from the observer point. The resulting viewshed map shows areas that are visible in green and areas that are hidden in red.
5. Interpret the results: The final step is to interpret the results of the viewshed analysis in the context of the study objectives. This may involve assessing the potential visual impact of proposed developments or land use changes, identifying potential safety hazards, or assessing the suitability of a location for a particular activity.

Viewshed analysis is a powerful tool for assessing the visibility of a location, and can be used in a wide range of applications such as urban planning, transportation planning, and natural resource management. However, it is important to keep in mind that viewshed analysis is based on a model of the terrain, and may not accurately represent the actual visibility conditions in the field.

Hydrological Modeling

Hydrological modeling in GIS involves using digital elevation models, surface water networks, rainfall data, and other geospatial data to simulate and predict the movement of water across a landscape. This type of modeling is used to understand and predict the behavior of rivers, streams, and other water bodies, as well as the impact of human activities and climate change on the water cycle.



There are several steps involved in hydrological modeling in GIS:

1. Data collection: The first step in hydrological modeling is to gather data on the landscape and water systems being modeled. This includes elevation data, stream networks, rainfall data, land cover data, and other relevant information.
2. Pre-processing: Once the data is collected, it must be pre-processed to prepare it for modeling. This involves tasks such as removing errors and outliers, filling data gaps, and converting data into a format suitable for modeling.

3. Model development: The next step is to develop a hydrological model that simulates the movement of water across the landscape. There are several types of hydrological models, including rainfall-runoff models, river routing models, and groundwater models.
4. Calibration: Once the model is developed, it must be calibrated to ensure that it accurately simulates the behavior of the water systems being modeled. This involves adjusting model parameters and comparing model output to observed data.
5. Analysis: Finally, the calibrated model can be used to analyze different scenarios and predict the impact of changes in the landscape or water systems. For example, hydrological modeling can be used to predict the impact of land use changes on water quality or the impact of climate change on the water cycle.

Hydrological modeling in GIS is a powerful tool for understanding and managing water resources. It can be used for a wide range of applications, including flood forecasting, water supply planning, and environmental management.

Hydrological Modeling Types

Hydrological modeling is the process of creating a mathematical representation of the hydrological cycle, which involves the movement of water through the atmosphere, across the land surface, and underground. The aim of hydrological modeling is to simulate the flow of water in a river basin or watershed, predict how it will respond to changes in land use, climate, and water management practices, and provide decision support for water resources management.

There are several types of hydrological models that can be used for this purpose, ranging from simple empirical models to complex physically-based models. Here are some of the most common types of hydrological models:

1. Empirical models: These are based on statistical relationships between input and output variables, such as rainfall and stream flow. Empirical models are relatively simple to use and require less data than other models, but they may be less accurate.
2. Conceptual models: These are based on simplified representations of the hydrological cycle, such as the Soil Conservation Service (SCS) curve number method. Conceptual models are useful for estimating runoff from ungauged catchments and for scenario analysis, but they may be less accurate than physically-based models.
3. Distributed hydrological models: These are based on a detailed representation of the hydrological cycle, including the spatial variation of precipitation, soil

properties, and land use. Distributed models can simulate the hydrological response of a catchment to changes in climate and land use, but they require a large amount of data and processing power.

4. Physically-based models: These are based on the fundamental principles of hydrology, such as the mass and energy balance equations. Physically-based models are the most accurate type of hydrological models, but they require a high level of expertise and a large amount of data.

Designing and printing the map

The steps involved in designing a map can vary depending on the purpose and scope of the map, but generally include the following:

1. Define the purpose and audience: Determine the purpose of the map, and identify the audience that the map will be designed for. This will help determine the content, style, and level of detail that the map should include.
2. Gather data: Collect and compile the data that will be used to create the map. This can include spatial data, such as GIS layers or satellite imagery, as well as non-spatial data such as demographic or economic statistics.
3. Choose the appropriate projection and coordinate system: Select the appropriate projection and coordinate system that will best represent the spatial data in the map.
4. Select the map scale: Choose an appropriate map scale that will allow the data to be displayed at a level of detail that is appropriate for the purpose and audience of the map.
5. Choose the map elements: Determine what elements will be included on the map, such as title, legend, scale bar, and north arrow.
6. Layout the map: Arrange the map elements in a visually appealing and effective way, considering the hierarchy of information and the visual balance of the map.
7. Design the map: Select an appropriate color scheme, fonts, symbols, and other design elements that will enhance the clarity and visual appeal of the map.
8. Label the map: Add labels to the map to provide context and information about the features being mapped.
9. Check for accuracy: Ensure that the map is accurate and free of errors, and that all data is properly labeled and represented.
10. Print or share the map: Print the map or share it electronically, depending on the intended use and audience.

Map printing process

The map printing process typically involves the following steps:

1. Prepare the map: The first step in the map printing process is to prepare the map for printing. This may involve selecting the appropriate scale, layout, and colors, as well as adding labels, legends, and other annotations.
2. Select the printing method: There are several printing methods available for map printing, including lithography, digital printing, offset printing, screen printing, and engraving. The choice of printing method will depend on factors such as the quality and quantity of maps required, the type of material to be printed on, and the budget available for printing.
3. Print the map: Once the map is prepared and the printing method selected, the map can be printed. This may involve sending the digital file to a printer or producing plates for lithography, offset printing, or engraving.
4. Finishing: After the map is printed, it may require finishing touches such as trimming, binding, or mounting. This will depend on the intended use of the map and the preferences of the end-user.
5. Quality control: Before the final product is released, it is important to perform quality control checks to ensure that the map is accurate and free of errors. This may involve visual inspection, comparison with source data, and testing for color accuracy.

<u>Map printing techniques</u>

Map printing techniques refer to the methods used to produce a physical copy of a map for distribution or display. Here are some of the most common map printing techniques:

1. Lithography: This is a printing technique that uses a flat stone or metal plate to transfer an image onto paper or another material. Lithography is often used for high-quality map printing because it produces sharp, detailed images with accurate colors.
2. Digital printing: This is a printing technique that uses digital files to produce printed materials. Digital printing is a flexible and efficient method for map printing, allowing for on-demand printing and customization of maps.
3. Offset printing: This is a printing technique that uses plates to transfer an image onto paper or another material. Offset printing is commonly used for large print runs of maps, as it is a cost-effective method for producing high-quality maps.
4. Screen printing: This is a printing technique that uses a stencil to transfer ink onto a surface. Screen printing is often used for printing maps on non-traditional materials, such as cloth or plastic.

5. Engraving: This is a printing technique that involves carving an image onto a metal plate, which is then used to transfer the image onto paper or another material. Engraving is a high-quality, precise method for map printing, but it is expensive and time-consuming.

Advantages and dis-advantages of digital and traditional map(analog) preparation techniques

Digital and analog map preparation techniques each have their own advantages and disadvantages.

Advantages of digital map preparation techniques:

1. Faster and more efficient: Digital map preparation techniques can save a significant amount of time and effort compared to analog techniques, as they allow for faster and more efficient data entry, analysis, and mapping.
2. Easy to edit and update: Digital maps can be easily edited and updated using software tools, making it easier to correct errors or incorporate new information.
3. Greater accuracy: Digital map preparation techniques can provide greater accuracy and precision in data analysis and mapping, as they allow for more detailed and complex calculations and measurements.
4. Easier to store and share: Digital maps can be stored and shared easily using electronic files, which can be accessed and used from anywhere with an internet connection.

Disadvantages of digital map preparation techniques:

1. Requires specialized skills: Digital map preparation techniques require specialized skills and knowledge of software tools, which may not be accessible or affordable for all users.
2. May be expensive: The cost of digital map preparation tools and software can be high, especially for advanced or specialized applications.
3. Vulnerable to technological issues: Digital maps are vulnerable to technological issues, such as software crashes or hardware failures, which can result in data loss or corruption.

Advantages of analog map preparation techniques:

1. Low cost: Analog map preparation techniques can be low cost, as they require only basic materials such as paper, pencils, and rulers.

2. Easy to understand: Analog maps can be easier to understand for non-experts, as they use familiar visual cues such as lines and symbols to represent data.
3. More tactile: Analog maps can be more tactile and engaging than digital maps, as they allow for a more hands-on approach to data analysis and mapping.

Disadvantages of analog map preparation techniques:

1. Time-consuming: Analog map preparation techniques can be time-consuming, as they require manual data entry, analysis, and mapping.
2. Difficult to edit and update: Analog maps can be difficult to edit and update, as changes may require manual adjustments or the creation of a new map.
3. Limited accuracy: Analog map preparation techniques may be less accurate or precise than digital techniques, as they are limited by the precision of manual measurements and calculations.