1

Map Language

1.1 Introduction

The collection of data about the spatial distribution of the significant features of the earth's surface has long been an important part of the activities of organized societies. From the earliest civilizations to modern times, spatial data have been collected by navigators, geographers and surveyors, and rendered into pictorial form by map makers or cartographers. Originally, maps were used to describe far-off places, as an aid for navigation and military strategies (Hodgkiss 1961). During the eighteenth century, many governments realised the advantages of systematic mapping of their lands, and commissioned national government institutions to prepare topographical maps. These institutions are still continuing the mapping work. Many of the developing countries are making all attempts to obtain the status of a developed country. These attempts are based on certain strategies relating to areas like natural resources management and development,information technology, tourism development, infrastructure development,rural development, environmental management, facility management, and e-governance. In order to make an effective study of these thrust and emerging fields, new and innovative technologies have been developed.

In the last two decades innovative technologies have been greatly applied to experimental and operational activities. These technologies have their historical antecedents. For instance, Remote Sensing and GIS have been developed from earlier technologies such as surveying, photogrammetry, cartography, mathematics, and statistics. Laurini and Thompson (1992) adopted the umbrella term "Geomatics" to cover all these disciplines. They stated that the different aspects of each of these areas are necessary for formulating and understanding spatial information systems. The traditional method of storing, analysing and presenting spatial data is the map. The map or spatial language, like any other language, functions as a filter for necessary information to pass through (Witthuhn et al, 1974). It modifies the way we think, observe, and make decisions. Maps are thus the starting point in any analysis and are used in the presentation of results of any operational project. Whether it is remote sensing, photogrammetry, cartography, or GIS, the ultimate output will be the production of high quality, more accurate and clearer map, so that the user finds it easy to make appropriate decisions. Therefore, maps and their production using modern technologies is an essential starting point and they are the necessary tools to explore the characteristics of spatial phenomena. This Chapter is exclusively devoted to providing the fundamental concepts of a map, map scale, various terms used in mapping, review of map projections, and map symbolism.

1.2 Map as Model

The map constitutes the language of simple geography as well as automated geography. As a graphic form of spatial data abstraction it is composed of different grid systems, projections, symbol libraries, methods of simplification and generalisation, and scale.

A map is the representation of the features of the earth drawn to scale. The surface of the map is a reduction of the real scenario. The map is a tool of communication and it has been in use since the days of the primitive man who had to move about constantly in search of food and shelter. A map from any local planning agency provides different kinds of information about the area. This map focuses on the infrastructure and legal descriptions of the property boundaries, existing and planned roadways, the locations of existing and planned utilities such as potable water, electric and gas supplies, and the sanitary sewer system. The planning map may not be of the same scale as the topographic map, the former probably being drawn to a larger scale than the latter. Further, the two may not be necessarily based on the same map projection. For a small area, the approximate scale of the data is probably more important than the details of the map projection. As Robinson et.al (1984) observe, "A map is a very powerful tool and maps are typical reductions which are smaller than the areas they portray". As such, each map must have a defined relationship between what exists in the area and its mapped representation. The scale of a map sets limits on both the type and manner of information that can be portrayed on a map.

1.2.1 Spatial Elements

Spatial objects in the real world can be thought of as occurring in four easily identifiable types namely, points, lines, areas and surfaces (Fig. 1.1). Collectively, these four features, or various permutations and combinations of these four spatial features

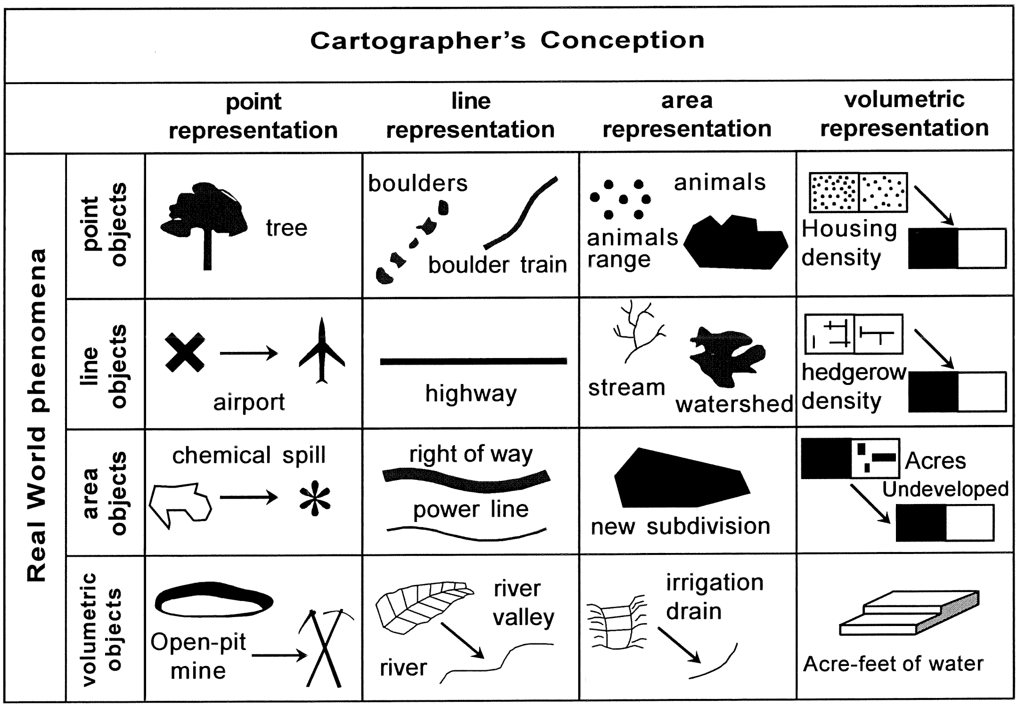


Fig. 1.1 Comparison of real-world phenomena and the cartographer's conception. Point, line, area, and surface features with examples (Source : Demers, 1998).

can form the human phenomena or the spatial real world. Points, lines and areas can be represented by using symbols to depict the real world. Surfaces are represented by any combination of these spatial entities. In general, all the geographic surfaces are in two tangible forms, namely, discrete and continuous. Trees, houses, road intersections and similar items are discrete spatial features. A feature can be termed as discrete, if it occupies a given point in space and time, that is, each feature can be referenced by its locational coordinates. All discrete features are said to have a zero dimensionality but have some spatial dimension. Fig. 1.2 shows the continuous spatial feature. The earth's surface occurs all around as natural features like hills, ridges, cliffs, and trenches, which can be described by citing their locations, the area they occupy, and how they are oriented with the addition of the third dimension. All these are considered continuous surface features. These features are composed of an infinite number of possible height values distributed without interruption across the surface.

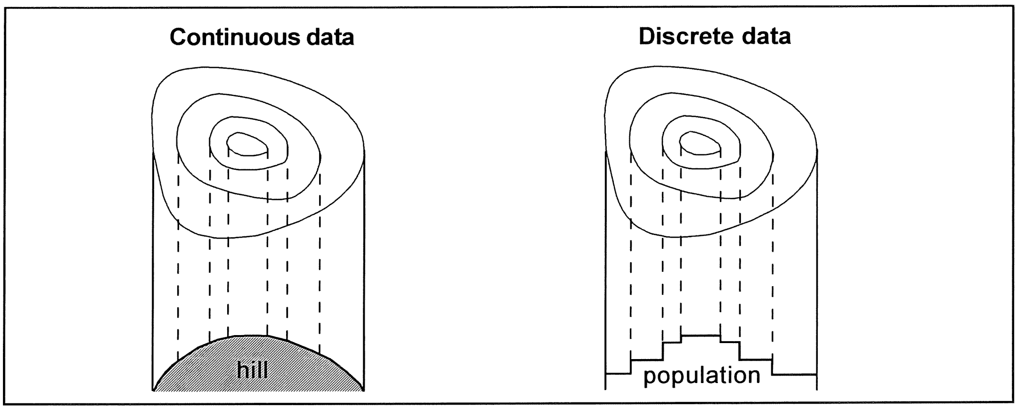


Fig. 1.2 Continuous versus discrete surface.

A topographic map depicts several kinds of information both discrete and continuous. Elevation on the site is portrayed as a series of contour lines. These contour lines provide us with a limited amount of information about the shape of the terrain. Different kinds of man-made features including structures and roadways are typically indicated by lines and shapes. In many cases, the information on this map is five to fifteen years out of date, a common situation resulting from the rate of change of land cover in the area and the cycle of map updates. Each of these different kinds of information, which we may decide to store in various ways, is called a theme.

A topographic map describes the shape, size , position, and relation of the physical features of an area. In addition to mountains, hills, valleys, and rivers, most topographic maps also show the culture of a region, that is, political boundaries, towns, houses, roads, and similar features.

1.2.2 Terminology

The terminology used in describing any kind of spatial/geographic features is discussed in this section. Elevation or altitude is the vertical distance between a given point and the datum plane. Datum plane is the reference surface from which all altitudes on a map are measured. This is usually mean sea level. The height is defined as the vertical difference in elevation between an object and its immediate surroundings. The difference in elevation of an area between tops of hills and bottoms of valleys is known as relief of the terrain. A point of known elevation and position usually indicated on a map by the letters B. M. with the altitude given to the nearest foot is termed as bench mark. On some maps the altitude is given along the contour line (Survey of India Maps).

A map line connecting points representing places on the earth's surface that have the same elevation is called contour line. It thus locates the intersection with the earth's surface of a plane at any arbitrary elevation parallel to the datum plane. Contours represent the vertical or third dimension on a map which otherwise has only two dimensions. They show the shape and size of physical features such as hills and valleys. A depression is indicated by an ordinary contour line except that hachures or short dashes are used on one side and point toward the center of the depression. The difference in elevation represented by adjacent contour lines is termed as contour interval. (Monmonier and Schnell, 1988)

Maps are a very important form of input to a geographical information system, as well as a common means to portray the results of an analysis from a GIS. Like GIS, maps are concerned with two fundamental aspects of reality, locations and attributes. Location represents the position of a point in a two-dimensional space. Attributes at a location are some measure of a qualitative or quantitative characteristic such as land cover, ownership, or precipitation. From these fundamental properties a variety of topology and metric properties of relationship may be identified including distance, direction, connectivity, and proximity.

1.3 Classification of Maps

Maps are thus the cartographer's representation of an area and a graphic representation of selected natural and man-made features of the whole or a part of the earths surface on a flat sheet of paper on a definite scale. Even though there are many different types of maps, all the maps are broadly classified on the basis of two criteria, namely, scale and, contents and purpose. On the basis of the scale, the map may be classified as either a small scale map or a large scale map. Some of large scale maps, are cadastral or revenue maps, utility maps, urban plan maps, transportation or network maps. On the basis of the content, maps are classified either as physical maps considered as small scale maps, or cultural maps. In the process of preparing a map one should remember that inside a GIS, one is likely to encounter a greater variety of maps than one might have expected on the basis of the subject matter. In addition, based on the thematic content of GIS coverages, the maps can be termed as thematic maps like vegetation maps, transportation maps, land use land cover maps, and remotely sensed imagery. These thematic maps will be portrayed as prism maps, choropleth maps (Fig. 1.3), point distribution maps, surface maps, graduated circle maps, and a host of other types. Plate 1 shows land use/land cover hydrogeomorphological map of the Shivanngudem watershed, which is a classical example of a thematic map.

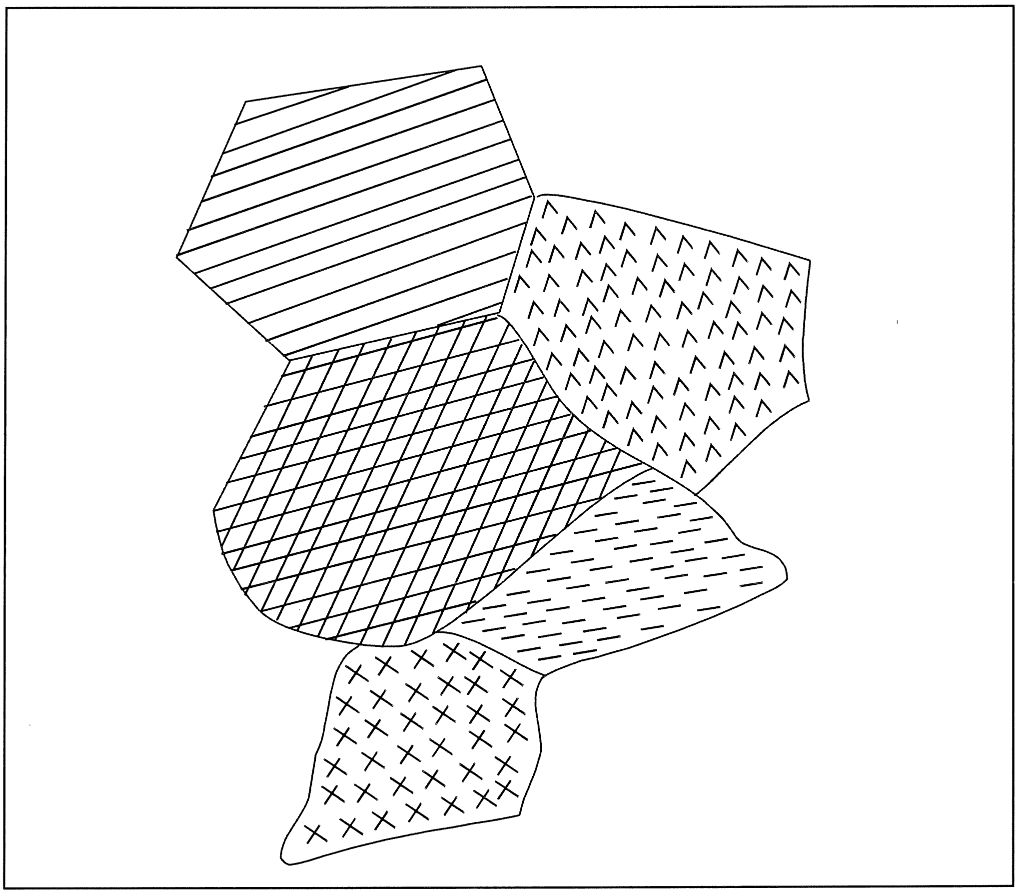


Fig. 1.3 Choropleth map.

1.4 Map Scale

By necessity, the process of representing geographic features on a sheet of paper involves the reduction of these features. The ratio between the reduced depiction on the map and the geographical features in the real world is known as the map scale, that is the ratio of the distance between two points on the map and the corresponding distance on the ground. The scale may be expressed in three ways and the pictorial representation of these three types is shown in Fig. 1.4.

Fractional scale : If two points are 1 km apart in the field, they may be represented on the map as separated by some fraction of that distance, say 1 cm. In this instance, the scale is 1 cm to a kilometer. There are 100,000 cm in 1 km; so this scale can be expressed as the fraction or ratio of 1:100,000. Many topographic maps of the United States Geological Survey have a scale of 1:62,500; and many recent maps have a scale of 1:31,250, and others of 1:24,000. In India, commonly used fractional map scales are 1:1,00,000,00; 1:250,000, 1:50,000; 1:25,000 and 1:10,000. The method of representing this type of scale is called Representation Fraction (RF) method.

Graphic scale : This scale is a line printed on the map and divided into units that are equivalent to some distance such as 1 km or 1 mile. The measured ground distance appears directly on the map in graphical representation.

Verbal scale : This is an expression in common speech, such as, "four centimeters to the kilometer", "an inch to a mile". This common method of expressing a scale has the advantage of being easily understood by most map users.

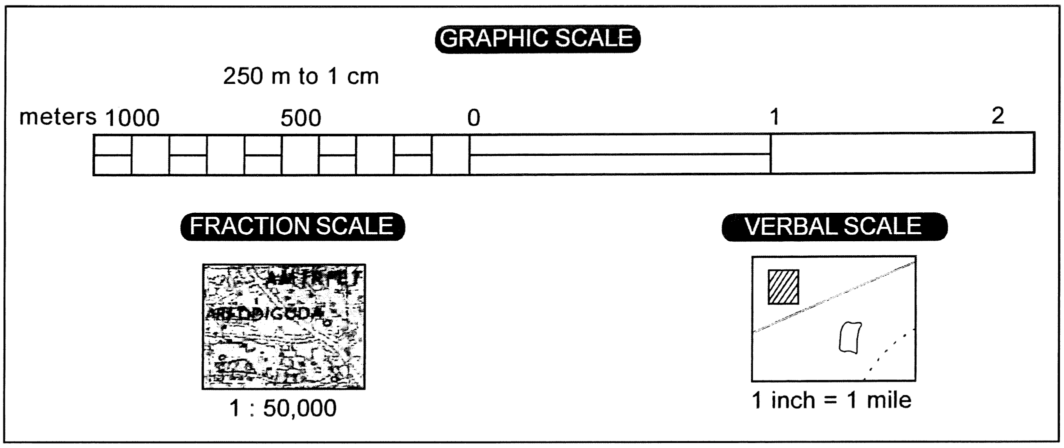


Fig. 1.4 Three different types of scale representation.

The ratio and map scale are inversely proportional. Therefore, 1:1,00,000 (large ratio) map is considered a small scale. In this instance, the scale is 1 cm to a kilometer. As there are 100,000 cm in 1 km, this scale can be expressed as the fraction or ratio of 1:10,000 (small ratio) map is considered a large scale. The small scale maps depict large tracts of lands (such as continents or countries) usually with a limited level of detail and a simple symbology. Large scale maps can depict small areas (such as cities) with a richness of detail and a complex symbology.

The terms 'small scale' and 'large scale' are in common use. A simple example helps illustrate the difference. Consider a field of 100 meters on a side. On a map of 1:10000 scale, the field is drawn 1 centimeter on a side. On a map of 1:1,000,000 scale, the field is drawn 0.1 millimeter on a side. The field appears larger on the 1:10000 scale map; we call this a large-scale map. Conversely, the field appears smaller on the 1:1,000,000 scale map, and we call this a small-scale map. Alternately, if we have a small area of the earth's surface on a page, we have a large-scale map; if we have a large area of the earth's surface on a page we have a small-scale map. Fig. 1.5 shows the effect of scale. On 1:50,000 scale map (large), even small knicks can be mapped where as on 1:250,000 scale the mapping of knicks is practically impossible.

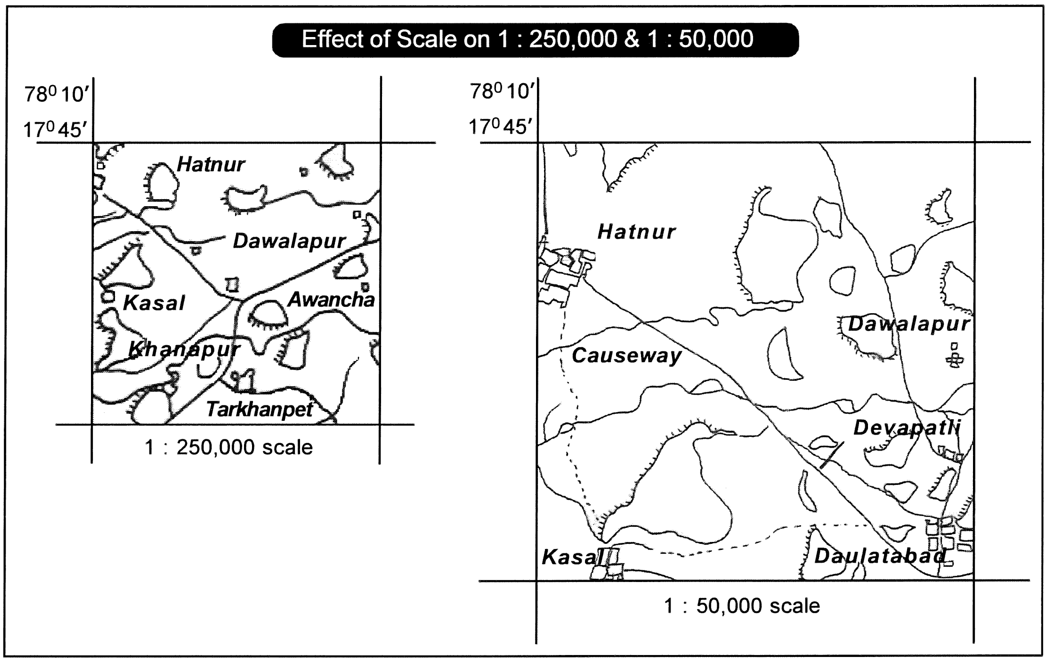


Fig. 1.5 Effect of Scale-1:250,000 vs 1:50,000

Scale values are normally written as dimensionless numbers, indicating that the measurements on the map and the earth are in the same units. A scale of 1:25,000, pronounced one to twenty five thousand, indicates that one unit of distance on a map corresponds to 25,000 of the same units on the ground. Thus, one centimeter on the map refers to 25,000 centimeters (or 250 meters) on the earth. This is exactly the same as one inch on the map corresponding to 25,000 inches (or approximately 2,080 feet) on the earth. Scale always refers to linear horizontal distances, and not measurements of area or elevation,

An explanation of the symbols used on topographic maps is printed on the bottom of each topographic sheet along the margin, and for other maps on a separate legend sheet. In general, culture (artificial works) is shown in black. All water features, such as streams, swamps and glaciers are shown in blue. Relief is shown by contours in brown. Red may be used to indicate main highways, and green overprints may be used to designate areas of woods, orchards, vineyards, and scrub.

1.5 Spatial Referencing System

The first important spatial concept in mapping technology is to locate objects with respect to some reference system. The system must have a structured mechanism to communicate the location of each object under study. The characteristics that a referencing system should possess include stability, the ability to show points, lines and areas and the ability to measure length, size and shape (Dale and McLaughlin, 1988). There are several methods of spatial referencing systems and they can be grouped into three categories, namely, geographic coordinate systems, rectangular coordinate systems and non-coordinate systems. In geographic coordinate systems, the coordinates of any location on the earth surface can be defined by latitude and longitude. Lines of longitude, called meridians are, drawn from pole to pole. The starting point for these lines called the prime meridian runs through Greenwich (Fig. 1.6).

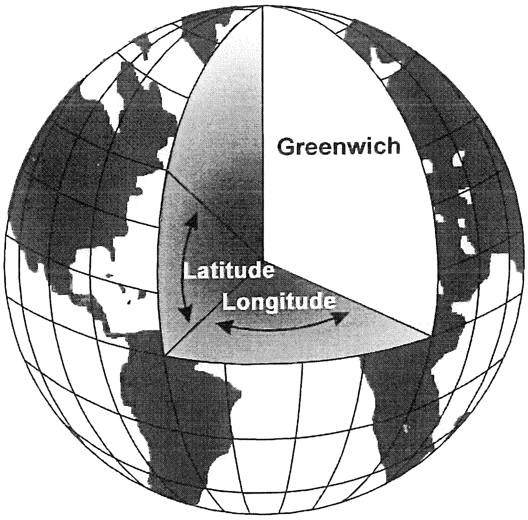


Fig. 1.6 The geographic grid.

The prime meridian is the starting or zero point for angular measurements, east and west. The lines of latitude lie at right angles to lines of longitude, and run parallel to one another. That is, each line of latitude represents the circle rounding the globe. Each circle will have a definite circumference and area depending on where it lies relative to the two poles. The circle of greatest circumference is called the equator and will be at equidistant from the poles. This type of location is called absolute location and it gives a definitive, measurable, and fixed point in space. This system is also called spherical grid system. This spherical grid system is produced by slicing the entire globe and placing two sets of imaginary lines around the earth. The first set of lines starts at the middle of the earth or equator called parallels, and circle the globe from east to west. These parallels are called latitudes. The second set of lines, called meridians, are drawn from pole to pole. These are called longitudes. Simply, it can be stated that the system of angular measurements allows us to state the absolute position of any location on the surface of the earth while calculating the degrees of latitude north or south of the equator, and the degrees of longitude east and west of the prime meridian.

Most of the spatial data available by means of remote sensing systems or any other sources of data for use in GIS are in two-dimensional form. This coordinate referencing system to locate any object point is called rectangular coordinate system. The location of any point on the earth's surface with reference to rectangular coordinate system is generally termed as relative position. The third coordinate system, namely, non-coordinate system, provides spatial references using a descriptive code rather than a coordinate, such as, postal codes, which are numeric in nature. Some countries use the non-coordinate system which is alpha numeric as in the case of UK and Canadian postal codes. This type of reference system is of particular importance for GIS users. Public land survey systems of Western United states is a classical example of this non coordinate-referencing system (Heywood et. al 2000).

For instance, in relation to prominent features a point may be referred to as being so many kilometers in a given direction from a city, mountain, lake, river mouth, or any other easily located feature on the map.

1.6 Map Projections

Map projection is a basic principle of map making in that when projected on to a flat map, objects on the earth's surface are distorted in some way, either in size, shape or in relative location (Maling, 1980). When the information is digitised from a map, the recorded locations will be often based on a rectangular coordinate system determined by the position of the map on the digitising table (star and Eastes, 1999). In order to determine the true earth locations of these digitised entities, it is necessary to devise the mathematical transformation required to convert these rectangular coordinates into the positions on the curved surface of the earth as represented on the map. Mathematical formulae to convert map units into latitude and longitude are available for most common projections (Snyder, 1987). Such mathematical transformation functions are normally built into projection as it is mathematically produced and is a two-fold process. First by, an obvious scale change converts the actual globe to a reference globe based on the desired scale. Secondly, the reference globe is mathematically projected on to the flat surface (Robinson et. al,1995). In this process of projection there is a change in scale. The representative fraction for the reference globe called the principle scale, can be calculated by dividing the earth's radius by the radius of the globe. The scale divided by the principle scale, is by definition 1.0 at every location on the reference globe. The process of transformation of three-dimensional space into a two dimensional map inevitably distorts at least one of the properties, namely, shape, area, distance or direction, and often more than one. Therefore, the scale factor will differ in different places on the map (Robinson et. al, 1995).

Map projection properties can be evaluated by means of applying three principal cartographic criteria, namely, conformity orthomorphic projections, equivalent projections, and equidistant projections (peuckar and chrisman, 1975). The projection that retains the property of maintaining correct angular correspondence can be preserved, and this is called angular conformity, conformal, or orthomorphic projection. The conformal type of projection results in distortions of areas leading to incorrect measurements. The projections by which areas can be preserved are called equal area or equivalent projections, the scale factor being equal to 1.0 mm. The projections by which the distances are preserved are known as equidistant projections. These three criteria are basic and mutually exclusive and other properties have only a peripheral importance. In fact, there is no ideal map projection, but only the best representation for a given purpose can be achieved.

A special emphasis is laid on transforming the satellite data on to a map. One of the requirements of the remotely sensed data is its ability to process an image from a generic coordinate system on to a projected coordinate system. Projecting imagery from line and pixel coordinates to the Universal Transverse Mercator (UTM) is an example in this regard. This is particularly critical when different trends of information from a Geographical Information Systems (GIS) are to be combined. The imagery must be accurately projected and rectified. By applying relevant map projections, a few basic concepts essential to the understanding of a map projections (Fig. 1.7) are reviewed in this section.

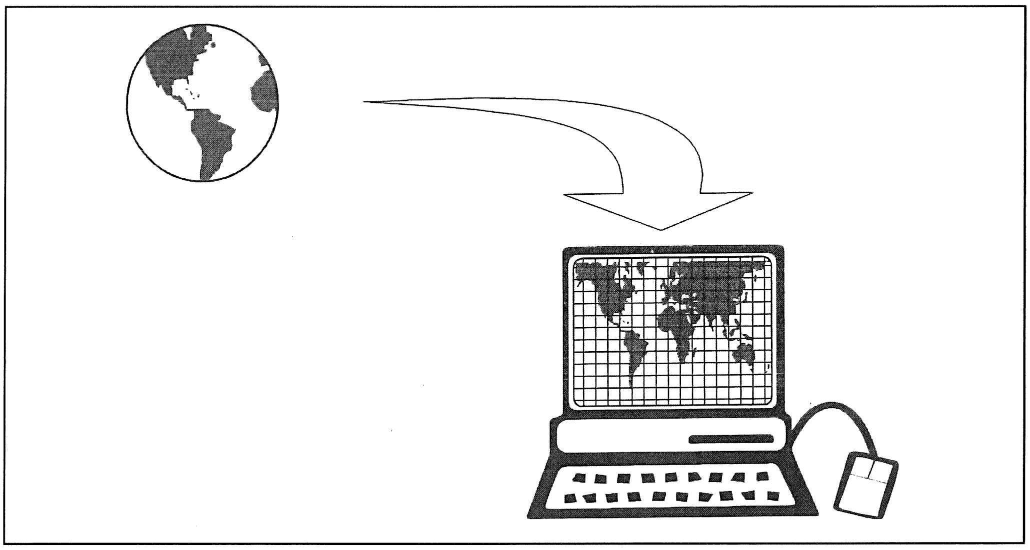


Fig. 1.7 Concept of Map Projections.

To transfer the image of the earth and its irregularities on to the plane surface of a map, three factors are involved, namely, a geoid, an ellipsoid or a datum with ellipsoid, and a projection. The geoid is a rendition of an irregular spheroidal shape. The variations in gravity are accounted for at this level. The observations made on the geoid are then transfered to a regular geometric reference surface, the ellipsoid. Many countries and organisations have calculated a variety of ellipsoids over the years. Variation in ellipsoid calculations are in part due to different observations on the geoid from different points upon the earth. The geographical relationships of the ellipsoid, still in a three-dimensional form, are transformed into two-dimensional plane of a map by a process called 'map projection' or simply projection. As illustrated in Fig. 1.8, the vast majority of projections are based upon cones, cylinders and planes. Each of these formats has advantages and disadvantages in terms of distortions and accuracy. Every flat map misrepresents the surface of the earth in some way. No map can rival a globe in truly representing the surface of the entire earth. However , a map or parts of a map can show one or more, but never all-of the following: True shapes, true directions, true distances, true areas.

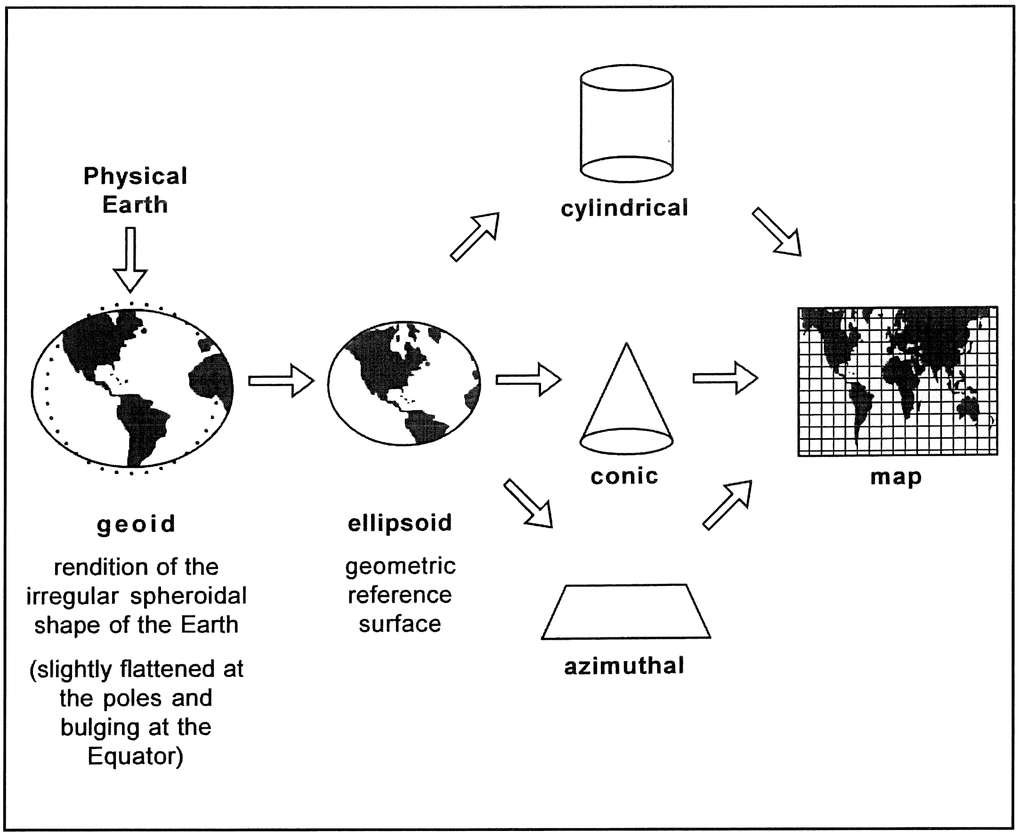


Fig 1.8 Geoid Ellipsoid - Projection : Relationship.

1.6.1 Grouping of Map Projections

All the map projections are grouped into four main families. They are,

(i) the family of planar projections

(ii) the family of cylindrical projections

(iii) the family of conical projections, and

(iv) the family of azimuthal projections.

If we wrap a sheet of paper round the globe in the form of a cylinder, transfer the geographical features of the globe on to it, and then unroll the sheet and lay of on a flat surface, we would achieve a cylindrical projection. The resulting graticule would be rectangular. In conic projection, if we repeat the above process, by wrapping the sheet of paper round the globe in the form of a cone, the resulting graticule would be fan shaped. The cone can be either tangent to a chosen parallel or it may intersect the plane along two parallels. Conic projections are specially suited to mapping areas having east-west extents, such as, Canada, USA and China. If a sheet of paper is laid tangent to a point on the globe and transfer the geographical features of the globe on to it, we would achieve azimuthal projections that appear as straight lines intersecting at the designated centre point, and parallels that appear as concentric circles round the centre point. A combination of any two of the above projections forms an hybrid projection.

The classification of map projections should follow a standard pattern so that any regular projection can be described by a set of criteria, and, conversely, a set of criteria will define a regular projection. Thus a classification scheme may follow a number of criteria subdivided into classes. Adler (1968) have named five basic criteria, as follows :

(i) Nature of the projection surface as defined by geometry,

(ii) Coincidence or contact of the projection surface with the datum surface,

(iii) Position or alignment of the projection surface with relation to the datum surface,

(iv) Properties of cartographic requirements, and

(v) Mode of generation of datum surface and coordinate systems.

The datum surface of the earth is usually an ellipsoid of revolution, but sometimes it is also approximated by a sphere. Although this assumption that the earth is a sphere can be used for small-scale maps to maintain the accuracy, for large scale maps the earth must be treated as a spheroid. Coordinate systems are necessary for the expression of position of points upon the surface, be it on an ellipsoid or a sphere or a plane. For the ellipsoid or the sphere the system of longitude and latitude is expressed in degrees, minutes and seconds of arc. For the plane a system of rectangular X and Y coordinates, sometimes referred to as Northings and Eastings, is usually applicable. Some of the commonly used map projections are described below.

1.7 Commonly used Map Projections and their Comparison

Though several conventional and non conventional map projections exist and are used, only a few projections with specific advantages are considered here.

1.7.1 Mercator

This is used for navigation for maps of equatorial regions. Any straight line on map is a thumb line (line of constant direction). Directions along a thumb line are true between any two points on a map, but a thumb line usually is not the shortest distance between points. Distances are true only along equator and are reasonably correct. Special scales can be used to measure distances along other parallels. Two particular parallels can be made correct in scale instead of the equator. Areas and shapes of large areas are distorted. Distortion increases as distance increases from the equator and is extreme in polar regions. The map, however, is conformal in that angles and shapes within any small area (such as that shown by a USGS topographic map) are essentially true.

1.7.2 Transverse Mercator

This is used by USGS for many quadrangle maps at scales from 1:24,000 to 1:250,000. Such maps can be joined at their edges only if they are in the same zone with one central meridian. Transverse Mercator is also used for mapping large areas that are mainly north-south in extent. Distances are true only along the central meridian selected by the map maker, or else along two lines parallel to it, and all distances, directions, shapes, and areas are reasonably accurate. Distortion of distances, directions, and size or area increases rapidly outside the defined distance. Since the map is conformal, shapes and angles within any small area are essentially true. Graticule spacing increases away from central meridian. Equator is straight. Other parallels are complex curves concave towards the nearest pole. The Central meridian and each meridian 90^0 from it are straight. Other meridians are complex curves concave toward central meridian.

This projection is a transverse cylindrical case, in which the scale will be kept exact along the central meridian and along the equator. The mapped area may be extended without limit in the north-south direction. This projection is also an orthomorphic projection with small shapes and angles maintained accurately. The scale distortions are systematic and can be predetermined.

1.7.3 Oblique Mercator

This is used to show regions along a great circle other than the equator or a meridian. These regions have their general extent oblique to the equator. This kind of map can be made to show as a straight line, the shortest distance between any two pre-selected points along the selected great circle.

Distances are true only along the great circle (the line of tangency for this projection), or along two lines parallel to it. Distances, directions, areas, and shapes are fairly accurate within 15^0 of the great circle. Distortion of areas, distances, and shapes increases away from the great circle. It is excessive toward the edges of a world map except near the path of the great circle. The map is conformal, but not perspective, of equal area, or equidistant. Graticule spacing increases away from the great circle, but conformity is retained. Both poles can be shown. Equator and other parallels are complex curves concave towards the nearest pole. Two meridians 180 apart are straight lines, all others being complex curves concave towards the great circle.

1.7.4 Polyconic Projection

This projection has generally been accepted as the best for a small, regular shaped area, such as, the standard quadrangle. Survey of India uses this projection for making topographical maps of Scale 1:250,000 and more. Although this projection is not conformal, the scale is not uniform, shapes and areas not being retained exactly. It comes closer to compliance with most of these projections. It cannot be used on a large area without noticeable distortion, and although two or three adjacent sheets can be matched continuously in one direction, two or more strips cannot be matched for any great distance without developing gaps along the abuting edges.

1.7.5 Lambert Conical Orthomorphic Projection

This projection portrays a portion of the earth's surface on the developed surface of a secant cone. It is used along the parallel of latitude at orthomorphic projection with two standard parallels by countries having predominant east-west directions for topographical mapping. In India we use this projection for geographical maps. When using two standard parallels, in the area between these parallels, the map scale will be too small and in the area outside the parallels it will be too great. The mapped area may be extended without limit in the east-west direction, but restricted within narrow limits in the north-south direction. The principal advantages of this projection is that the scale distortions are systematic and can be predetermined, and that a map sheet of any portion of a zone can be matched perfectly with an adjoining sheet in the same zone. However, a map of one zone cannot be matched with a map of an adjoining zone.

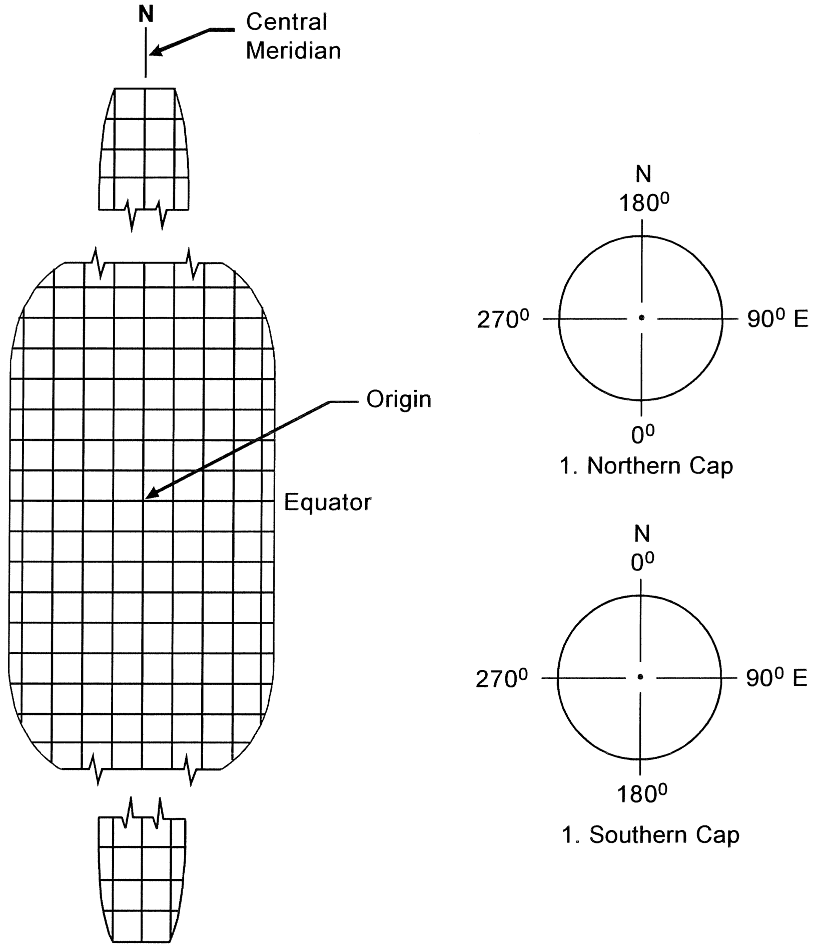
1.8 Grid Systems

Though the map projections discussed in the foregoing paragraphs are generally used by the National Mapping Agencies for preparation of various kinds of maps and charts, grid maps are also used mainly in large scale maps. A Grid system has certain advantages over the use of geographical coordinates. Firstly, every grid square is of the same size and shape. Secondly, linear values can be used rather than angular values. A rectangular grid system is required for the army, and for engineering and cadastral maps, as the computations of reference azimuth and survey coordinates distance can be easily done in this system, and a brief, precise, unique and convenient map references can be given. Two of the important grid systems that are essential to our studies are Lambert Grid system used in India and Universal Transverse Mercator (UTM) coordinate system used in more than a hundred countries the world over.

1.8.1 Lambert Grid System for India

The grid system adopted for India is the Lambert Grid system. The grid squares in this system can be superimposed on any conformal projection. However, it is only on the Lambert conformal conical projection that they represent equal and perfect squares. On the polyconic maps of Survey of India, the grid squares appear as squares, but they may differ by as much as 1 in 824.

A series of eight Lambert Grids cover India and Burma which are named as I, IIA, IIB, IIIA, IIIB, IVA, IVB and O. The N -S extent of the grid is limited to 8 degrees in order to limit the maximum scale error to 1 in 824 and the E - W extent limited to 16 degrees owing to limitations on the grid convergence -- the deviation of the grid north from true (geographical) north. The grid origins are spaced 7 degrees apart in latitude with 1 degree overlap at borders, and are assigned arbitrary grid coordinates of 1 million grid meters north and 3 million grid meters East. The scale factor is 823/824 along the central parallel and 824/823 along the edges. (Fig. 1.9)



(a) A UTM Grid Zone (b) UPS Grid Zones

Fig 1.9 Coordinate system for UTM grid.

1.8.2 Universal Transverse Mercator (UTM) Grid

For the UTM grid the world is divided into 60 north to south zones. Each zone covers a 6^0 - wide strip of Longitude. The maximum extent of the zone was chosen to minimise distortion. The zones are numbered consecutively, beginning at zone 1 between 180^0 and 174^0 west to zone 60 between 174^0 and 180^0 east longitude. Each zone is then divided into 19 segments with an 8^0 difference in latitude plus an additional segment at the extreme north with a 12^0 difference in latitude. The rows of these segments are lettered from south to north by the letters C through X (with omission of the letters I and O to avoid confusion). By specifying a letter and a number, each element in the UTM system is uniquely identified.

Within each zone , the location of a point is made by specifying its X coordinate, the easting and its Y coordinate, the northing. The unit of the coordinates is meters. The central meridian of each zone is assigned the false easting value of 500,000 m. This is done to avoid negative numbers at the western edge of the zone. The easting numbers increase from west to east. For north-south values in the north hemisphere, the equator has a false northing value of 0 m, and the northing decrease towards the south pole. The coordinate system for the UTM grid is illustrated in (Fig.1.9). Of course, the scale of the grid must be consistent with the scale of the base transverse mercator projection.

For the UTM grid the mapping cylinder is secant to the earth. The local scale factor for the central meridian, halfway between zone boundary meridians is 0.9996. This limits maximum variations to 1 in 1000. No less than 100 countries in the world have accepted this projection, as they must be finding it very satisfactory for both mapping and rectangular referencing. A few of the advantage are :

(i) It is a conformal projection and hence the direction errors are minimal within a selected UTM zone.

(ii) It has continuity over sizable areas coupled with a minimal number of zones. In case of India it involves only 6 UTM zones.

(iii) Scale errors caused by the projections do not exceed the specified tolerance, that is, 1/2,500.

(iv) Unique referencing is possible in a plane coordinate system for all zones.

(v) Transformation formulae from one zone to another are uniform throughout the system (assuming one reference ellipsoid). This helps in easy programming for the computer eliminating the need for auxillary tables based on manual computations.

(vi) Meridionial convergence does not exceed five degrees.

In addition to the above quoted advantages, modern satellite based-geodetic measurement techniques like Global Positioning System provide directly the coordinate in UTM projection avoiding the hustles of conversion from one coordinate system to another. In the area of digital mapping also, most of the GIS software today handles the cartographic data in a UTM projection with the possibility of converting the coordinates from one projection system to another.

The basic disadvantage of the UTM lies in the maximum projections scale distortion which reaches 4 parts in 10,000. The main reason for this is the choice of zone width, 6 degrees in the case of UTM.

1.9 Computer in Map Production

By 1977, however, the experience of using computers in map making considerably advanced. Rhind (1977) was able to present a list of reasons for using computers in the process of making maps. These reasons are listed below.

(i) To make existing maps more at a quicker pace.

(ii) To make existing maps at a cheaper rate.

(iii) To make maps for specific user needs.

(iv) To make map production possible in situations where skilled staff are unavailable.

(v) To allow experimentation with different graphical representations of the same data.

(vi) To facilitate map making and updating when the data are already in digital form.

(vii) To facilitate analysis of data that demand interaction between statistical analysis and mapping .

(viii) To minimize the use of the printed map as a data store and thereby to minimize the effects of classification and generalization on the quality of the data.

(ix) To create maps like the 3-D type.

(x) To create maps in which selection and generalization procedures are explicitly defined and consistently executed.

(xi) Introduction of automation can lead to a review of the whole map-making process, which can also lead to savings and improvements.

During the 1960's and 1970's there were two main trends in the application of computer methods for mapping. One was the automation of existing tasks, with an accent on cartographic accuracy and visual quality, and the other with the accent on spatial analysis but at the expense of good graphical results.

1.10 Digital Database in a GIS

The essential component of a GIS is its information database. This database is developed by capturing information from different sources like topographical maps, thematic maps, and cadastral maps, in analog form, remotely sensed images in digital form and so on. Whatever may be the source of information, they need to be integrated on to a common projection and registration.

The general technique of converting the analog map to digital form is either by manual digitisation or by scanning. Manual digitisation is the commonly used technique for including analog data in a digital database. The use of remotely sensed data which is available in digital form for maintaining the dynamicity of the database for GIS, has become very common. Map projection aspects involved in these two techniques are discussed in the following paragraphs.

1.10.1 Digitiser Units Vs Real-world Coordinates

After determining the projection of the data source and deciding upon a projection in which to store the database, we need to project data into a real-world coordinate system. Registering all layers to common coordinate system ensures data integrity during spatial joins.

Coverages may be digitised in digitiser units or real-world units. The digitiser is based on a type of rectangular coordinate system with its origin in the lower-left corner. On the digitiser surface, moving one inch up or down covers the same distance as moving one inch left or right. Anywhere on the table's surface, an inch is an inch. Whether the unit of measurement be inches or centimeters, when associated with digitising, it is called a digitiser unit.

Location and distance are the key features in mapping geographic information. The real world has a curved surface and is often measured in feet or meters. Both feet and meters are standard, and the system of spherical coordinates (Global Reference System) is used to reference system location on the surface of the globe with points on a map. The units of this reference system are degrees of latitude and longitude. But the distance represented by a degree depends upon its location on the globe which is the global reference system and not a rectangular coordinate system. Thus, a coverage can be digitised in meters, but not in degrees. The advantages and disadvantages of digitising a coverage in either digitiser units or in real-world units are outlined in Table 1.1.

Table 1.1 Advantages and disadvantages of digitiser units and Real-world units

S.No.	Digitiser units	Real-world units
(i)	Easy to create edit plots at scale of source map	Maps need to be plotted at a precise scale to overlay edit plots
(ii)	Digitising staff has less to learn and understand	Digitising staff should understand transformation and projection concepts
(iii)	Coverages are not spatially referenced and cannot be displayed simultaneously	Allow multiple coverages to be shown, such as background or adjacent coverages.
(iv)	Inconvenient for update	Usually used for update
(v)	Less concern over bad information	Must have correct projection parameters
(vi)	Don't know whether initial digitising of tics was accurate	RMS error indicates actual tic accuracy in real-world units.

A common step in quality assurance when initially developing a database is to compare a digitised file with its source map. This is most commonly done by creating an edit plot and overlaying it on a light table. If a map is digitised in real-world units, it may have been stretched and scaled so that it will no longer register accurately with the source map even if the file was digitised accurately. If coverages are to be digitised in real-world units, the digitising staff should understand how to project and transform coverages. This naturally requires some knowledge of projection concepts.

The ability to display adjacent and background coverage can be quite helpful, particularly for updating maps. This is not possible when maps are digitised in digitiser units. Coverages should not be edited, cleaned, built, or buffered, nor should any spatial analysis be performed when they are stored in reference units (latitude-longitude). The algorithms that perform snapping functions using a measurement of length or area are based upon Cartesian coordinates. The length of a line of latitude between two meridians varies with latitude , and the area is confusing when measured in square degrees.

1.11 Linkage of GIS to Remote Sensing

This book aims to provide an introduction to the theoretical and technical principles that need to be understood to work effectively and critically with GIS. Today maps are not just made using GIS, but the infrastructure of utilities in the streets of our towns will be held in a GIS. Your taxis and emergency services may be guided to their destination using satellite-linked spatial systems; the foresters and farmers will be monitoring their standing crops with spatial information systems; natural resources management developments, and environmental management strategies may be compared with the integration of satellite data while processing results. To use GIS technology, for the above allied application areas, a huge GIS database on geographical features is to be created. Creating such a database is a complex operation which may involve data capture, verification and structuring process. Because raw geographical data are available in many different analogue or digital forms like maps, aerial photographs, and satellite images, a spatial database can be built in several, not mutually exclusive ways, such as acquiring data in digital form from a data supplier, digitising existing analogue data, carrying out field survey of geographic entities, and interpolating from point observations to continuous surfaces.

Images derived from optical and digital remote sensing systems mounted in aircraft and satellites provide much spatial information and major data as an input to GIS. Remote sensing data are a major source of data for the mapping of resources like geology, forestry, water resources, land use and land cover. Integration of the two technologies, remote sensing and GIS, can be used to develop decision support systems for a planner or decision maker. Remotely sensed images can be used for two purposes, as a source of spatial data within GIS and using the functionality of GIS in processing remotely sensed data in both pictorial and digital modes.

Since digital remote sensing images are collected in a raster format, digital images are inherently compatible spatially with other sources of information in a raster domain. Because of this, "raw" images can be directly and easily included as layers in a raster-based GIS. Similarly, such image processing procedures as automated land cover classification result in the creation of interpreted or derived data files in a raster format. These derived data are again inherently compatible with the other sources of data represented in a raster format.

Remote sensing images need not be digital in format to be of any value in a GIS environment. Visual interpretation of hardcopy images is used extensively to locate specific features and conditions, which are then subsequently geocoded for inclusion in a GIS. At the same time, the information resident in a GIS can also be used to aid in a visual or digital image interpretation process. For example, GIS information on elevation, slope, and aspect might be used to aid in the classification of forest types appearing in images acquired over areas of high relief. Thus, the interaction between remote sensing and GIS techniques is two-way in nature.

Remote sensing images including the information extracted from such images, along with GPS data, have become primary data sources for modern GIS. Indeed, the boundaries between remote sensing, GIS, and GPS technology have become blurred, and these combined fields will continue to revolutionise the inventory, monitoring, and managing natural resources on a day-to-day basis. Similarly, these technologies are assisting us in modeling and understanding biophysical process at all scales. They are also permitting us to develop and communicate cause-and-effect "what-if" scenarios in a spatial context in ways never before possible.

The importance of remote sensing, GIS, GPS, and related information technologies in the professional careers of today's students involved in measuring, studying, and managing earth resources cannot be over-stated.

Hence, in recent years, remote sensing has become a powerful source of spatial data as an input for GIS through which a detailed map can be generated with the help of other collateral data derived from several other sources. There are, two methods of extracting data for GIS from the remote sensing data. They are, visual interpretation of satellite imageries in pictorial format, and computer processing of remotely sensed digital data. The output of either of these analysis methods can be considered an input for GIS for any kind of application. Fig. 1.10 shows an overview of the linkage of remote sensing and GIS.

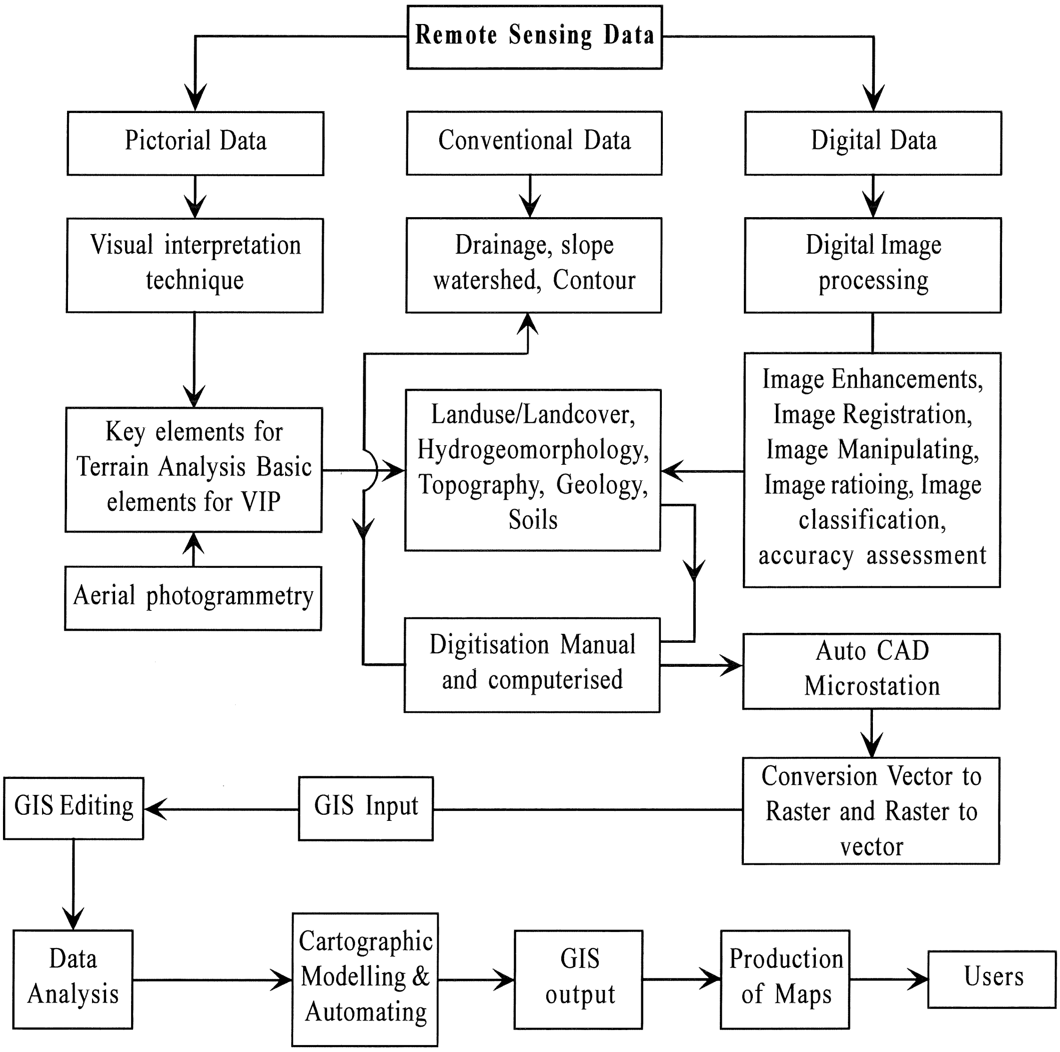


Fig 1.10 Overview of the linkage of remote sensing and GIS.