8.9 Geographical Information Systems

A Geographic Information System (GIS) is a computer-based system that is used in input, storage, analysis manipulation, retrieval, and output, of spatial data. These systems consist of computer hardware and software. GIS are increasingly being used in applications in natural resources management. These days GIS are also being integrated with remote sensing and GPS.

GIS has origin in manual overlay operations done as early as in 1920s. In 1958, computer based cartography initiated in USA which culminated in development of first general purpose mapping software in 1960s. Canada GIS is also cited as first GIS and was developed around the same time. Present day popular commercial/ open source GIS are: ArcGIS, GRASS, MapInfo, ERDAS, IDRISI, ILWIS etc. GRASS (Geographic Resource Analysis Support System) is high-end open source software. Arc GIS is modular high-end commercial software. ERDAS and ILWIS packages have image processing and GIS capabilities.

8.9.1 Advantages of GIS

- After a map is available in digital form, it is easy to update or modify it.
- Conventionally, overlays are manually prepared. This process is cumbersome and error prone. GIS allows digital overlaying a number of maps.
- Conventionally, maps are browsed to retrieve information. In GIS retrieval of information is very fast, and much easier.
- In conventional method, hardcopy output is prepared. Updating of such maps is difficult and the map is to be redrawn. This process is simple in GIS- the map can be easily edited and printed.
- Annotation is clumsy in hardcopy maps. Thus, while retrieving information ambiguity may arise. Also all features may not be annotated in paper maps. Thus, attribute information, e.g. names for some of the features are lost in paper products. In GIS, information is stored in tables and is linked to geographic features and thus is not limited by availability of annotation space/ color/ symbol, etc.
- If multiple maps are prepared for same area, e.g. watershed, land use, geomorphology, common boundaries are drawn manually and may not match in different maps. In GIS common boundaries are digitized once and are available to all layers. Once GIS map layers are prepared, any number of maps can be designed.
- Storing of a large number of paper maps is difficult; maps deteriorate with time. Digital maps can be stored in a much compact way.

GIS systems are now available for standard computers in practical, low-cost formats. The main cost factor now resides in database compilation, and training and updating of technical staff. Data capture or input in GIS is costly but it is one-time affair. Commercially available paper maps may be cheaper than GIS layer. Use of GIS requires investment in computer hardware, software, and training. GIS handling requires trained manpower. GIS software should
have proper functionality as desired in an application. For example, for hydrological modeling, DEM analysis should be available in the package.

8.9.2 Spatial data representations
Spatial data are represented mainly in two ways in GIS: raster and vector. These data representations can be transformed from one form to the other, albeit with some information loss. In raster, spatial data are structured as grid of cells or pixels. Their row and column numbers address the cells. In many distributed hydrological models, spatial data and hydrological computations are done in this form. This is a native representation for remotely sensed data. Satellite data are captured/resampled as pixels (picture elements) and information is extracted through digital image processing. In vector model, spatial data are represented as coordinate points. For example, point data is represented as a pair of coordinates. A straight-line is represented as two pair of coordinates, representing end points of the line. A curved line is represented as finite line segments. Area data are represented as line data with some additional information e.g. centroid, adjacent areas, etc.

In raster data, points and lines are represented with finite area and finite width and thus is not a natural representation. Lines have stepped or zigzagged appearance. In vector model, points and lines have infinitesimal area and width respectively. Lines are smooth curves. Raster data require large storage space. Vector data require small storage space. Thematic maps prepared from remotely sensed data are available in raster form and are often processed as such. Many hydrological models use both the representations. For example, thematic maps of catchment variables and hydrometeorological measurements are prepared in raster form. Stream network is processed in vector form etc. In raster form, value of many catchment variables is scale dependent. For example, average slope of catchment reduces with increase in raster grid size. In most GIS, the representations coexist. For example, it is better to capture spatial data from conventional thematic maps, through visual interpretation of remotely sensed data in vector form. Thematic maps from digital processing of satellite data may be obtained in raster form.

Topology: Method of representing vector data is called its topology. A line consists of two nodes and one or more vertices. Nodes are end points of the line. Lines also have directions. Thus, nodes are referred as ‘from node’ and ‘to node’ depending on direction of the line. Areas are represented by ‘left area’ and ‘right area’ of each line.

Digital Elevation Model (DEM)
Topographic elevation data in GIS are called DEM. These are represented in GIS in various manners namely contours, raster, and TIN (Triangulated Irregular Network). Contours are conventional representations of DEM and are used in topographic maps. Contours are equal elevation lines. Normally, equal interval contours are drawn in topographic maps to represent topography. For example Survey of India (SOI) maps at 1:250,000, 1:50,000 and 1:25,000 scale
have contours at 100, 20 and 10 m elevation interval. DEMs are used to derive topographic information such as slope, aspect and are also used in hydrological calculations, e.g. stream network delineation, topographic index, delineation of catchment area, etc.

**TIN**
In TIN model, elevations at the vertices of triangles are used to compute elevation at interior points of the triangles. Using elevation of the vertices of a triangle, a planar or higher order surface can be fitted. The surface can be used to derive elevation at points inside the triangle. TIN model requires Delaunay triangulation. In this, constituent triangles are as equilateral as possible. Circum-circles of the triangles include no other point of the triangulation. Triangulation is performed first by constructing Voronoi diagram (Thiessen polygons). Points included in adjacent polygons are joined to create Delaunay triangulation. Voronoi diagram is drawn using proximity analysis.

**Interpolation**
Interpolation is a technique of determining unknown value of a variable at location from known values at other locations. Interpolation can be used for any spatial variable, e.g. topographic elevation, pH, SAR, pollutant concentration, groundwater depth and level, population etc. Known values can be at point, line of area locations. Point data can be spot heights, pH, pollutant concentration etc. Line data can be topographic contours, etc. Area data can be population density in regions, etc.

*Thiessen polygons or nearest neighbor*
This is popular method of interpolating rainfall values from point rainfall measurement at rain gauges. Generally, point rainfall stations are limited in number. To determine basin wide average rainfall, the method of interpolation is used. The Thiessen polygon diagram is prepared by proximity analysis. For measuring average rainfall in the catchment, weights for each rain gauge station are area of Thiessen polygon surrounding that station divided by the total catchment area.

Distance weighted averaging
In this averaging, a weight of inverse of distance function is used. Distance function is nth power of distance. Thus, more weight is assigned to stations closer to the interpolation location.

Surface fitting: Here, n-degree polynomial surface is fit between selected known values. The surface can be used, among the other application, to interpolate values.

Krigging: Variation of spatial variables can be partitioned in three components, namely drift or structure, small variations and random noise. First component depicts general trend of the data. Second component represents small variations from the general trend. These variations are random but spatially autocorrelated. Third component depicts random values that are not
spatially autocorrelated. Kriging is best suited for interpolation of pollutant concentration, geological and mining variables, e.g. grade of ores, etc. For these data, single smooth mathematical equations are not suitable. The technique is based on assumption that values in neighbourhood have generally higher correlated. Apart from the estimate of values, error estimates are also provided in kriging technique. In presence of large random noise in data, good semivariogram is not obtained and this results in deterioration in interpolation quality.

Georeferencing: Earth is a three dimensional surface. In maps, this three-dimensional surface is transformed into flat surface. For the transformation, map projections are employed. Locations on the map are drawn using Cartesian coordinates obtained through map projections. Geographic graticules (latitudes and longitudes) are later drawn in maps. Sometimes, Cartesian coordinate grids are also drawn on maps. In georeferencing earth coordinates are assigned to spatial data. Either lat/long or Cartesian coordinates can be used in georeferencing maps in GIS. Cartesian coordinates allow measurements, e.g. area and lengths and are thus frequently used. Geographic coordinates do not allow measurements. A map can also be referenced without using a map projection. In such case, it is difficult to integrate GIS layers obtained from different sources.

A map which is to be geo-referenced is called the source map and the reference map is the map which has known coordinates. Points whose reference coordinates are known and which are clearly identifiable on both the source and reference maps are known as control points. For coordinates of control points in two maps, coefficients of a polynomial transformation equation are estimated.

8.9.3 Map Projection
Map projection is transfer of positions on earth to corresponding points on a flat sheet of paper. Because of the shape of the earth, this transformation involves approximations and is not distortion free. Distortions occur in lengths, angles, shapes and areas. Scale is a ratio of length on map to its counterpart on the earth. Since large size of features on earth surface, a scale is needed to draw these features on a small sheet of paper. Earth shape is assumed spherical or spheroidal. An intermediate plotting surfaces namely cylinder, cone or plane is used in projections. Corresponding projections are called cylindrical, conical and azimuthal respectively.

Distortions occur in projecting earth surface on to intermediate plotting surface. Ideally, areas distances, directions, angles and shapes should be preserved. In reality, few of these properties are preserved. Based on application, choice is made as to which properties are to be preserved and appropriate map projection is selected. In areal distortions area of a figure may increase or reduce. In linear distortion length and its curvature may change. In angular distortion an angle may increase or decrease. In shape distortion, a square may become parallelogram, rectangle or may have curved boundaries or both. A point may be distorted in to a line. In equal area projections, area of a figure is preserved. In the process distortions are introduced in
distances and angles or shape of figures. In conformal projection, shape is preserved. In this process, areas figures are distorted. Projections with these contrasting properties are called equal area and conformal projections respectively. With different orientation of intermediate plotting surface, it is possible to obtain different projections. Azimuthal projections are called polar, equatorial or oblique depending on point of contact of plane falling on poles, equator or at intermediate latitude. For conic and cylindrical projection based on orientation of axis of the plotting surfaces, the projections are classified. When the axis coincides with earth’s polar axis, perpendicular to it and lying in equatorial plane and is oblique to it, the projections are called regular or equatorial, transverse and oblique respectively. The plotting surfaces can also be tangential or secant to the earth surface. Normally for projection, mathematical approach is used.

8.9.4 GIS Operations
Input of data in a GIS database is either by digitization or import. Digitization can be done on-screen or by a digitizer to create/ edit GIS objects in vector format. Input data are, sometimes, available in GIS image formats. These data are converted to native format of GIS.

Storage: Geographic data are stored in GIS the native format of GIS. For one spatial data, many computer files are created which contain different information. Attribute data are stored in Data Base Management System and are linked to geographic objects. Storage of data in the form of layers looks very attractive from water resources data. Different types of data, such as soil and land cover, are stored in different map layers. GIS permits analysis of single or multiple layers and various layers can be overlaid, one on top of another. From a water resources point of view, spatial variation of data is important, e.g., the variation of soil hydraulic properties.

8.9.5 Spatial Data Analysis
Data analysis involves operations with geographic data and their attributes to obtain derived information, generate query, statistics etc. broad categories and operations therein are as follows.

Statistics: for example, count, length, area, perimeter, shape, centroid, etc. for geographic objects can be derived in GIS. For continuous surfaces, average, standard deviation, maximum, minimum, etc. are derived. Summary operation produces zonal statistics for a map. For example, land use statistics for watershed in a basin can be generated.

Mathematical operations: Mathematical operations, e.g., addition, subtraction, multiplication, division, exponential, logarithm, absolute, truncation, round off, negative, trigonometric operations can be performed in GIS. For example various component maps in the universal soil loss equation, namely R, K, L, S, C and P can be prepared as different largess and multiplied using multiplication operation.

Logical operations: Logical operations, namely or, and, not, xor can be performed on maps.
Figure 8.8 shows logical operations. For example, landuse= agriculture and pH >= 8 will result in salt affected agriculture area.

Conditional: If-then-else conditional operational can be performed on maps. For example, ‘if 50 < return period <=100 and land use= residential, then vulnerability= high else vulnerability=low’ condition gives flood vulnerability map.

![Figure 8.8 Logical operations commonly used in GIS.](image)

Conditional: In this operation, all combinations of classes in two maps are obtained in the resulted map. For example, overlay of soil hydrological soil group and land use/ cover map will provide soil-cover complex map.

Classification: It converts values into interval. A continuous surface is input and area map is output for the operation. In the output area map, isolines, i.e., line of equal values, enclose the area. Examples of various isolines are contours, isohyete, isotherm, isobar etc., which represent topographic elevation, rainfall, temperature, pressure, respectively. In reclassification; information of geographic object is changed. For example soil series map may be changed to soil pH map.

Search/ buffer: The operation is similar to distance, except that at a specified distance an area geographic object is created.

Neighborhood: Information in eight neighbor, their locations and statistics, e.g. mean, mode, median, minimum, maximum, standard deviation, coefficient of variation, etc. are extracted.
Aggregate: Cell size of raster maps can be changed in fractions of half, one fourth etc. using functions e.g. mean, predominant, minimum and maximum.

Query: Query is done by attributes or geometry. In query by attribute, a logical expression is written in attributes and result is obtained. For example land use=agriculture will select/display agriculture areas. In query by geometry, objects are selected on screen to view their attributes. After the requisite data are stored in a GIS database, it is easy to answer complex queries like what areas in the catchment have forest on shallow soil with 3% slope?

Output: GIS output may be obtained as paper maps or as a picture which may be shared or printed or inserted in a report or presentation. The output maps may contain various cartographic elements namely title, legend, graticules or grids, north arrow, scale, annotations, notes, etc. In one output more than one GIS layer may be included apart from cartographic elements. When design is saved, it only contains reference to the layers. Thus, if a layer is modified and designed output map is opened at a later time, the changes are reflected in the output.

The topography of a river basin may be represented in two different ways: as a digital elevation model or as a triangulated irregular network (TIN). The digital elevation model is a grid of elevation values that has regular spacing while TIN is a series of points linked into triangular surfaces that approximate the surface. The spacing of points in TIN are non-uniform, which allows points to be located on critical terrain features, roads or river banks. The accuracy of such digital terrain models depends on the source of the data, the point density and distribution, and other related data used in their development. Conventional contour maps may be prepared from a digital elevation model or TIN. Orthophotos are images of the landscape from which features can be referenced to one another. They are digital images produced by processing aerial photography to geodetic control elevation data to remove all sources of distortion. The image has the properties of scale and accuracy associated with a map. Such images can be derived using airborne or satellite sensors.

8.9.6 Applications
Several hydrological software now have a GIS interface. GIS interfaces have been developed for hydrological models such as HEC models, SWAT, etc. ArcInfo extensions such as Spatial Analyst, 3D analyst, are useful for hydrological tasks. Script languages are also available to write interfaces in GIS, for example, Arc Macro Language is the script language in ArcInfo.

GIS techniques facilitate input of spatial data to hydrological model. GIS are being incorporated in hydrological models to extract and format distributed watershed data. Use of DEM permits complete physiographic and hydrological depiction of basins. The efficiency of handling large volumes of data means that more comprehensive and detailed maps, isolines and
themes can be prepared. This is a significant improvement in water resources management as map preparation is often time-consuming and expensive.

GIS are becoming very common in the field of water resources assessment and management. Many tasks of data collection, compilation, and interpretation can be facilitated by means of GIS. In network planning and design, the ability to map quickly and display surface water and related stations enables a more effective integration to take place. Network maps, showing basins or stations selected according to record quality, watershed, or operational characteristics, can be used for both short-term and long-term planning.

Groundwater potential and quality can be studied in GIS environment. Various layers namely slope, geology, distances to drainage channel, tanks and lineaments, depth to water table, depth of weathered zone can be overlaid and integrated on GIS environment to obtain groundwater potential map. Similarly, layers of water quality variables may be created to obtain quality map.

8.10 Closure
RS and GIS are powerful tools for spatial data collection and management. Spatial information and its attributes can be stored, analysed and output efficiently through GIS. It has many applications in natural resources and infrastructure management. GIS also has application in varied industries namely power, transportation etc. Whereas in many applications, it can be independently used, it also helpful in hydrological modeling in terms of data preparation. For such applications, varieties of interfaces, stand-alone programs, embedded programs exist. These increase productivity and reproducibility in modeling.

<table>
<thead>
<tr>
<th>Table 8.1 Map projections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projection</td>
</tr>
<tr>
<td>Equal area</td>
</tr>
<tr>
<td>Albers, one standard parallel</td>
</tr>
<tr>
<td>Albers, two standard parallels</td>
</tr>
<tr>
<td>Lambert zenithal</td>
</tr>
<tr>
<td>Cylindrical equal area</td>
</tr>
<tr>
<td>Conformal</td>
</tr>
<tr>
<td>Lambert conformal, one standard parallel</td>
</tr>
<tr>
<td>Lambert conformal, two standard parallel</td>
</tr>
<tr>
<td>Stereographic</td>
</tr>
<tr>
<td>Mercator (e.g. Universal Transverse Mercator or UTM)</td>
</tr>
</tbody>
</table>

References