INTRODUCTION AND PRINCIPLES OF GIS

D.S. Rathore
Scientist-F
National Institute of Hydrology, Roorkee–247 667

1.0 INTRODUCTION

GIS is a computer-based system that is used in input, output, storage, manipulation, retrieval and analysis of spatial data. These systems consist of computer hardware and software. GIS are increasingly being used in applications in natural resources, tourism, transportation, trade and commerce etc. GIS is also integrated with modern technology of remote sensing and GPS. GIS are now integral part of hydrological models.

History

GIS has basis in manual overlay operations done as early as in 1912 to 1969. In 1958, computer based cartography initiated in University of Washington, which culminated in development of first general purpose mapping software in 1960s. Canada GIS (CGIS) is also cited as first GIS and was developed. Other GIS developed were SYMAP (1964), SYMVU (1960s), GRID (1960s), ODYSSEY (1970s) etc. In 1960s, US Bureau of the Census created census information of USA on GIS. In 1970s, Environmental System Research Institute (ESRI) was founded. ESRI combined ARC and a relational database management system (RDBMS) INFO to develop ARC INFO GIS package in 1980s. During these decades, software e.g., MAP, Intergraph's CADD and mapping software, MOSS BLM, GRASS (Army Corps of Engineers, National Park Service), SPAN (Tydac Inc.) etc. were developed.

Present day commercial/ open source GIS are Arc GIS, 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, ILWIS and IDRISI have image processing and GIS capabilities.

Advantages and limitations

Using GIS, it is possible to overlay large number of maps. Conventionally, manually overlays are prepared. This process is cumbersome and error prone. In GIS retrieval of information is faster as it is done through computer. Conventionally, maps are browsed to retrieve information. In GIS information retrieval is much easier and is done automatically. In GIS, interactive/ virtual output may be prepared. The virtual output is automatically updated, if the component maps are revised. In convention method, hardcopy output is prepared. Updating of such maps is difficult. The map is required to be redrawn. 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 etc. 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 once digitized and are available to all layers. Once GIS map layers are prepared, any number of maps can be designed. Handling of paper maps is difficult.

Data capture or input is costly in GIS. Commercially available paper maps may be cheaper than GIS layer. This is because of high cost of data capture in GIS. Use of GIS requires investment in computers, software and training. GIS handling requires trained manpower. In GIS data are required to be converted in to native format of GIS software. In suitable import/ export functionality is not available or the format is obscure/ unknown, data may be unusable. GIS software should have proper functionality as desired in an application. For example, in transportation applications, network analysis function should be available. For hydrological modeling, DEM analysis functions should be available.

1.1 SPATIAL DATA

In GIS, maps are called spatial data. Information on paper maps can be input is GIS as spatial data. Example of spatial data are stream network, well locations, villages, cities, topographic contours, spot elevations, roads, land use, soil, geology, hydrological investigation locations, hydrological response units etc. Spatial data are classified in to three types namely area, line and point. Areas are spatial data that are represented as closed figures e.g. forests, lakes, Thiessen polygons etc. Lines are spatial data that are represented as lines e.g. forest boundaries, lake boundaries, contours, stream network, roads etc. Point spatial data are represented as points on maps e.g. well locations, rain gauge stations, spot elevation, villages etc. The data is scale dependent in some instances e.g. on small-scale maps a city will be represented as point data, where as on large-scale map, it will be represented as area data. A closed line data can be converted easily to area data in GIS. For example forest boundary data can be converted to forest land-use map. Point data cannot be converted in to area data. But reverse is true i.e. area data can be converted as point data. For example, if city point data is captured, city polygons cannot be obtained from this point data. City polygon data will be required to be captured separately.

1.1.1 Spatial Data Representations

Spatial data are represented in two mainly ways in GIS namely raster and vector. These data representations can be translated in to one another, albeit with some information loss. In raster, spatial data are structured as grid of cells or pixels. Their row and column numbers addresses the cells. In many distributed hydrological models, processing is spatial data are represented in this form and hydrological computations are also done in this form. This is a native representation for remotely sensed data. Satellite data are captured/ resampled as pixel (picture elements) or grid cell. Thematic maps are prepared from these data through digital image processing. These maps are available in raster representation. In vector model, spatial data are represented as coordinate points. For example, point data is represented as a pair of coordinates. A straight-line data is represented as two pair of coordinates, representing end points of the straight 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.

Comparison

In raster data, points and lines are represented with finite area and finite width and thus is not a natural representation. Lines has jagged or stepped 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 is raster form and are often processed as such. Many hydrological models use both the representation. For example, thematic maps of catchment variables and hydrometeorolgical 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. Thus, results of un calibrated hydrological models will differ at different raster grid size used in parameter derivation. In most of GIS, the representations coexist. For example, it is better to capture spatial data from conventional thematic maps, through visual interpretation of remotely sensed data etc. in vector form. Thematic maps from digital processing of satellite data may be obtained in raster form. Data can be transformed in to one another as and when desired.

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.

1.2 DATA FORMAT

There are many formats prevalent for images, raster and vector GSI data. Image file formats are normally used by non GIS applications. For use in GIS, data in these formats are required to be imported. The image file formats may or may not have projected coordinate system.

Image file formats

GIF (Graphics Interchange Format): This is copyrighted format by ‘CompuServe’ for image data. A data compression scheme similar to lossless LZW compression is used. Georeference information of raster data is not stored.

TIFF (Tagged Image File Format): The format is designed by Adlus and Microsoft. For types of images are supported namely bitlevelm gray scale, color- mapped (256 color) and RGB- based (three layer or band for full color). The compression type can be uncompressed, RLE or modified Huffman. A format for storing georeferenced images is also designed and is called ‘geo- tiff’.

JPEG (Joint Photographic Expert Group): The format uses lossy compression technique. Compression of 20:1 can be achieved. The format is designed for grey- scale and color images. Georeference information of raster data is not stored.

PNG (Portable Network Graphics): This format uses lossless compression and the compression technique used is not a patented technique. This format is suitable for images with line drawings rather than true color images.

GIS file formats

NTF (National Transfer Format): This is developed in United Kingdom 9UK) for transfer of digital cartographic data in 1987 and updated in 1989. It is format for both raster and vector data. Levels are defined from 0 to 5 for raster, simple vector, complex vector, topology and user defined formats.

SDTS (Spatial Data Transfer Specifications): The data exchange format standard is developed by US Digital Cartographic Data Standards Task Force (DCDSTF) in 1988 for both raster and vector data.

e00 (Arc Exchange Format): This is exchange format developed by ESRI for vector data.

Arc coverage: This is a format for vector data developed by ESRI. The format is also used in ERDAS imagine.

Shape: This is a format for vector data developed by ESRI. The format is also used in ERDAS imagine. A shape file is a collection of files with extension shp, shx, dbf, prj etc. File with shp extension stores actural geographc data. File with shx extension stores index. The file with dbf extension is a dbase III file which stores attribute data. The prj file stores projection information and is a ASCII file. Shape files do not store topological information.

Img: This is a format for raster data developed by ERDAS Incorporated (now Lyca).

DXF (Drawing Exchange Format): The format was designed for data exchange between Autocad and other Cad packages. This is a de facto standard for exchange of vector data.

TIGER (Topographically Integrated Geocoding and Referencing) and GBF- DIME: US Bureau of Census developed these formats in respectively 1990 and 1980. Former is a topologically structured format.

DLG (Digital Line Graph): US Geological Survey (USGS) developed the format for transfer of fully structured data from National Cartographic Database. DLG-3 and DLG-E are types of this format.

BSQ/ BIL and BIP: These are generic data formats and are mainly used for remotely sensed data and some times for GIS data e.g. in ILWIS software. In BSQ (Band Sequential) format, each band of remotely sensed data is stored in separate files. In BIL (Band Interleaved by Line) format each band of a scan line are stored in alternate records in single file. In BIP (Band Interleaved by Pixel) group of pixels (e.g. two pixels) are stored for each band, are stored alternately.

1.3 DIGITAL ELEVATION MODEL (DEM)

Topographic elevation data in GIS are called DEM. These are represented in GIS in various manners namely contours, raster, 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 in 1: 250,000, 1:50,000 and 1: 25,000 scale Survey of India (SOI) maps, contours are at 100, 20 and 10 m elevation interval. Ridges, valleys can be interpreted from these maps. DEM in Raster and TIN representations can be used in deriving topographic information such as slope, aspect and can also be used in hydrological calculations e.g. stream network delineation, topographic index, flow routing, up steam contributing area etc. and in turn in hydrological modeling.

1.3.1 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 planner 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. Circumcircles 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.

1.3.2 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.

Voronoi diagram or Thiessen polygons or nearest neighbor

To determine basin wide average rainfall, this method of interpolation is widely used. The diagram is prepared by proximity analysis. The Thiessen polygon map is intersected with the catchment map. Area of a Thiessen polygon corresponding to a raingauge station in this intersected map is used as weight in finding weighted average rainfall for the catchment.

Distance weighted averaging

In distance weighted averaging, a weight of inverse of distance function is used. Distance function is nth power of distance. Thus, a higher weight is assigned to values closer to the interpolation location. At any point values are estimated as weighted sum of known values at selected locations. The selected point can be as follows:

- All points within a given range
- Specified number of closest points
- Specified number of closest points within quadrants/ Octants etc.

Surface fitting

A n- degree polynomial surface can be fit between selected known values. The points on this surface represent interpolate values. The points to be used for interpolation can be selected in similar way as that in distance weighted averaging method.

Krigging

Kriging is a statistical technique called best linear unbiased estimator (BLUE). Spatial variables have 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 depict random values that are not spatially autocorrelated. Kriging technique is best suited for interpolation of pollutant concentration, geological and mining variables e.g. grade of ores etc. In these data, single smooth mathematical equations are not suitable. The technique is based on assumption that values in neighbourhood have generally higher correlated. For example elevation in plain area is generally lower than that for hills and varies less abruptly. 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.

Semivariogram

Semivariogram is a plot of semivariances and distances of the samples for which the semivariance corresponds. Semivariogram is also referred as variogram. When variogram for all separation distances are plotted, the resulted variogram plot is called raw variogram. For ‘n’ data points, the resulted points in the raw variogram will be ‘n*(n-1)/2’. Raw variogram show cloud of points. A representative variogram or experimental variogram is more useful for fitting theoretical models and doing kriging. In computation of this, the separation distances are grouped. Plotting positions for these groups are mean, median, or middle of the class intervals. Distance intervals for these groups at smaller distances can be smaller.

For mathematical formulation of kriging interpolation, a theoretical models is fit for experimental variogram. Models e.g. Gaussian, exponential, power etc. are used. Gaussian model has parabolic nature at origin. Exponential and power models are linear at the origin. Parameters of the model are sill, range and nugget. The sill is equal to the variance of the data. Due to experimental error and micro level variations, experimental variogram may not pass through origin. This property is modeled by Nugget. It is modeled through a discontinuity at the origin.

Validation

‘Jack- knifing’ can validate the model. A subset of data is used in kriging. The complementary datasets of these subsets are used in validation of selected model. Residuals and standard errors are estimated at all discrete locations. The mean and variance of these residuals should be theoretically 0 and 1 respectively. Normally acceptable deviations of these measures from their true values may be utmost 0.15 to 0.20.

Linear contour interpolation

In linear interpolation from contours, distance map is estimated from contours. Based on distances at an interpolation point towards two nearest points, value is interpolated linearly. The value at the contours is retained in the final map. Distance function can be used in estimating distances.

Ray method

In this method rays are drawn in four, eight or sixteen directions from interpolation points. A value of the contours at points where the rays intersects the contours is determined. Using two nearest values, interpolation is done. Average of all interpolated values provide required interpolation.

1.4 GIS OPERATIONS

1.4.1 Input

Digitization

Digitization is done on- screen to create/ edit GIS objects in vector format.

Digitization errors for area objects

Following errors occur while digitization of area objects:

Dangle: Additional lines are some times digitized that are not areas. These lines appear as dangles.

Overshoot: Areas are closed figures. Sometimes, line is extended beyond closed figure. These are called overshoots.

Undershoot: Sometimes, line stop before closing of the area. These are called undershoots.

Self-overlap: A line crossing itself is called self-overlap.

Intersection: Lines cross each other without a node at the intersection.

Other errors are missing lines, incorrect labels etc.

Data import

Input data are, some times, available in GIS, image formats. These data are converted to native format of GIS through import utility.

1.4.2 Storage

Geographic data are stored in GIS is native format of GIS. For one data, many computer files are created. These files are copied, renamed, deleted wihin GIS. These operations can also be done outside GIS. Attribute data are stored in DBMS. Attribute data are managed within GIS or through DBMS software. Attribute data are linked to geographic objects. External databases can also be connected to geographic data. Data types of geographic objects and their attributes are bit, byte, integer, real, double, text etc.

1.4.3 Analysis

Data analysis involves operations with geographic data and their attributes to obtain derived information, generate query, statistics etc.

Statistics: Statistics e.g. count, length, area, perimeter, shape, centroid, rose diagram etc. of 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 watersheds 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 USLE namely R, K, L, S, C and P can be prepared and multiplied using multiplication operation. This operation will multiply these factors for all cells and provide long-term average annual soil erosion for each cell.

Logical operations: Logical operations namely or, and, not, xor can be performed on maps. 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.

Overlay: 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.

Reclassification: Information of geographic object is changed in reclassification. For example soil series map may be changed to soil pH map.

Classification: Classification converts values in to 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, isobath, isohyete, isotherm, isobar etc., which represent topographic elevation, groundwater table, rainfall, temperature, pressure etc.

Distance: Distance from a geographic object is estimated. Diagonal distances are nearly 1.4 times that of horizontal and vertical distances.
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, predominant, 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.

1.4.4 Output

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 only 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 to the output.

1.5 GIS SOFTWARE

There are many GIS software available both in public domain (free of cost) and as commercial software. Examples of public domain/ open source software are GRASS, OSSIM, QGIS, ILWIS (since July 1, 2007) etc. Commercial software are Arc Info, Map Info, ERDAS Imagine, Geomatica etc. Among these, Arc Info is high end software. Several software e.g. R2V, Surfer etc. have specific GIS functionalities. R2V is useful software for data input from scanned images. Scanned images are converted to vector format using the software. Surfer has many GIS function e.g. interpolation, contouring, output, three-dimensional visualization etc. It is widely used as a spatial data interface for Modflow groundwater modeling software. Some of the main stream GIS software are described below:

ILWIS (Integrated Land and Water Information System)

The software has both image processing and GIS capabilities. It has scripting language. Application/ interface development environment is not available. It supports raster format and limited support for vector format is available. Some of the functions available are topological digitization, polygonizsation, raster to vector conversion (polygon), projection transformation (both water and vector), hydrological processing of DEM etc.

The software will be available as free software since July 1, 2007 under GPL. It is no longer supported by ITC, The Netherlands from this year.

GRASS (Geographic Resources Analysis Support System)
It is a modular system and was developed by USA- CERL (US Army Construction Engineering Research Laboratory). Since 1999, it is available under GNU GPL as pen source software. It has many of the capabilities as described above for ILWIS. Interfaces for hydrological models have been developed for the software. It has interface with QGIS software for data input. It is developed under cygwin/ X- window environment under MS Windows.

OSSIM (Open Source Software Image Map)
It has image processing, GIS, photogrammetric capabilities. It was funded by several US government agencies. The program was developed in C++ with GUI and command line utilities.

QGIS (Quantum GIS)
GIS is an open source GIS and support several vector, raster and database formats e.g. shape, geotiff etc. The GIS has capabilities to visualize, manage, edit, analyse data, and compose printable maps. Current version is Enceladus 1.4.0. For editing topological editing for line and polygon data is available.

Arc GIS
Arc GIS software is developed by ESRI and is modular, high end commercial software. The software has several versions namely Arc Reader (free), Arc View, Arc Editor and Arc Info. For each version several extensions are available seperately. Applications/ interfaces can be developed using the software objects provided. Various programming languages e.g. Visual Basin, .Net platform etc. may be used in developing these applications. For hydrological software, several extensions are developed by the third parties and are available from them. It simplifies creation of data layers for hydrological software.

Map Info
Map Info is a vector based GIS and is suitable for projects where attributes data are required to be manipulated.

1.6 DATA BASE MANAGEMENT SYSTEM (DBMS)

Data bases are collection of data put in an organised manner, specifically put in a tabular form. An example of database is watershed database, which contains watershed name, code, area, perimeter, areas under different land use, average rainfall, runoff yield at different dependability, groundwater potential, population, water demand for domestic, cattle, industrial purposes, average slope, stream length, stream slope etc. The information may be collected or derived from various sources e.g. topographic maps, census, remotely sensed, hydrometeorological data etc. Information may be primary or secondary.

For storing, editing, manipulation, retrieval and output of data bases special software are available. These softwares are called ‘data base management system’ or DBMS. Examples of DBMS software are MS Access, MS SQL, MySQL, Oracle etc.

In DBMS, information for any object, e.g. a watershed, is stored in a single row or record. Each column contains single information or attribute. For example a column named ‘watershed_area’ will contain watershed area information only. It is not possible to store along with area, specific notes e.g. source from which area was obtained. For this purpose, another column may be created in the data base where notes may be stored. Also, it is not possible to record multiple information e.g. watershed area from different sources, in a single column. This requires separate columns, e.g. ‘W_area_topo’, W_area_report” etc.

GIS requires storing of attributes of spatial objects. DBMS are required to handle these data. DBMS functionality can be coded within GIS software or external commercial DBMS may be utilized by GIS. For example, ILWIS software itself handles its attribute data. Arc Info originally utilized ‘Info’ commercial DBMS. Now, a commercial DBMS of user’s choice may be utilized in Arc Info.

Table or relation

A data base may contain several tables. For a watershed, there may be several tables, e.g. for watershed characteristics, dams salient features, conservation agencies etc.

Primary key

A primary key is one or more columns in a table used to identify a row. Primary key has unique values. For example, watershed code attribute may be used as primary key in watershed characteristics table. Since this attribute has unique value, any row can be identified using this attribute. Watershed names can also be used as ‘primary key’ provided they are unique. Many times, primary key may be generated automatically by the DBMS software.

Relations

A data base is organised in to several tables. These tables have relation between them. The relation may be of several types, namely one- to- one, one- to- many or many- to- many.

One- to- one: Watersheds are delineated at the dam location. In such case, watershed characteristics table and dam information table will have one- to- one relation.

One- to- many: In a basin, there may be several dams. In such case, basin characteristics and dam salient feature tables will have one- to- many relationship i.e. for one basin there can be several dams.

Many- to- many: In a command area and watershed data base, there can be many- to- many relationship i.e. a command may spread across several watershed or a watershed may be under several commands.

Join

Join is an operation used to extract information across several tables. For examples, to get information of dams located within several basins, tables of basin and dam will be joined. Join operation depends on type of relationship.

One- to- one: In this type of relation, identification column of any one table can be added in to another table. This column is used in the join operation.

One- to- many: In this type of relation, identification column of the table of ‘one’ side of the relation is added in to the table on ‘many’ side of relation. For example to join table of basin (one) to dam (many), the identification column of basin is placed in the dam table.

Many- to- many: In this type of relationship, an intermediate table is prepared. This table contains, identification column of both the tables, which have this relationship type. There after all three tables (to tables to be joined and the intermediate table) are joined using these identification numbers.

Standard Query language (SQL)

SQL is a scripting language used for defining and processing data bases. These may be executed as commands within DBMS. In DBMS software, visual interfaces are also provided, which are equivalent to SQL commands. The SQL can also be used in scripting languages, e.g. VBScript etc., or languages, e.g. java, C# etc. SQL was developed in 1970. ANSI adopted the SQL as standard in 1992. This standard is called SQL- 92.

1.7 APPLICATIONS

Groundwater studies

Groundwater depth and quality is studied in GIS through kriging interpolation technique. Water quality variables e.g. EC, RSC, HCO3 etc. can be interpolated using the technique. In these applications, sample locations can be important. For example, samples are taken from working tube wells, which may be in general of good quality and poor water-quality is under represented. Range can be of order of 10 to 100 km. Models can be exponential, spherical etc.

Groundwater potential and quality can be mapped in GIS environment. Various layers namely slope, geology, hydromorphogeology, 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 namely magnesium, incrustation problem, TH, TDS can be integrated to obtain quality map in incrustation and corrosion problem areas.

Land degradation

Irrigated agriculture areas often face problems of water logging and salinity. The problems are refereed as twin problem as waterlogging leads to soil salinization in long run. Identification of extent of the affected area is pre requisite for reclamation. Criteria adopted for water logging are given in Table 1.1. Criteria for salinity and sodicity are listed in Table 1.2.

Area with surface pondage and moist soil can be delineated easily using remote sensing data. Water has black tone in standard FCC in visible and near IR bands. Most soil has dark signature in these imageries. Shallow water table conditions often are not detected using optical remotely sensed data unless its expression is visible on the surface of the earth. Areas where yield is affected can be monitored.

Table 1.1 Criteria adopted for waterlogging

Waterlogging	National Commission on Agriculture (1976)	Ministry of Water Resources, Go1 (1991)
Waterlogged/ Critical	Water table < 1.5 m	Water table < 2 m
Potentially waterlogged		Water table 2-3 m

Safe area		Water table > 3 m

Table 1.2 Criteria for soil salinity/ sodicity

Degree of salinity/ sodicity	Salinity EC (dSm^{-1})	Sodicity	
		pH	ESP
Slight	4-8	8.2-9.0	< 15
Moderate	8-25	9.0-9.8	15-40
Strong	>25	>9.8	>40

Saline areas possess salt efflorescence on the surface. Due to this, saline areas have bright appearance in optical remote sensing. Sodic area have different signature than saline areas. In sodic areas, the infiltration is very less and thus water gets stagnant in the areas and thus the area can be identified through surface pondage and moist soil. Both, saline and sodic areas have poor growth of vegetation. Waterlogged areas can also be delineated using GIS technique using water depth map. The map is processed to correct any discrepancies in depths (e.g. negative values). The maps can be utilized to classify areas as waterlogged/ critical, potential waterlogged and safe.