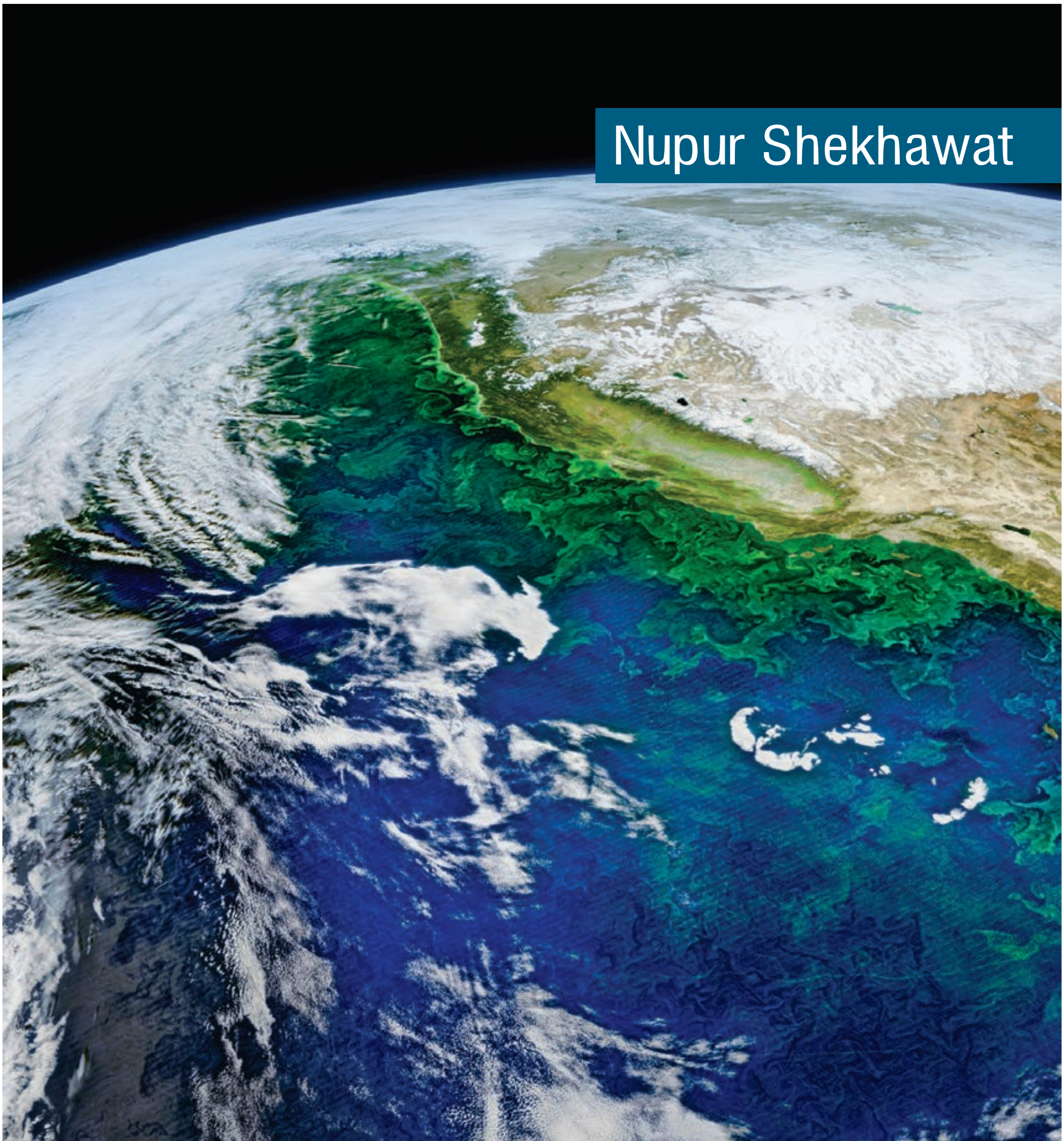
Nupur Shekhawat



Geographic Information Systems and Remote Sensing

Geographic Information Systems and Remote Sensing

Geographic Information Systems and Remote Sensing

Nupur Shekhawat

VIDYA BOOKS

Geographic Information Systems and Remote Sensing
Nupur Shekhawat
ISBN: 978-93-5429-155-5

Published by Vidya Books,
305, Ajit Bhawan,
21 Ansari Road,
Daryaganj, Delhi 110002

Contents

1

Introduction to Geographic Information Systems

The advent of cheap and powerful computers over the last few decades has allowed for the development of innovative software applications for the storage, analysis, and display of geographic data. Many of these applications belong to a group of software known as Geographic Information Systems (GIS). Many definitions have been proposed for what constitutes a GIS. Each of these definitions conforms to the particular task that is being performed. Instead of repeating each of these definitions, Thus, the activities normally carried out on a GIS include:

- The measurement of natural and human made phenomena and processes from a spatial perspective. These measurements emphasize three types of properties commonly associated with these types of systems: elements, attributes, and relationships.
- The storage of measurements in digital form in a computer database. The features can be of three types: points, lines, or areas (polygons).
- The analysis of collected measurements to produce more data and to discover new relationships by numerically manipulating and modelling different pieces of data.
- The depiction of the measured or analysed data in some type of display-maps, graphs, lists, or summary statistics.

The first computerized GIS began its life in 1964 as a project of the Rehabilitation and Development Agency Program within the government of Canada. The Canada Geographic Information System (CGIS) was designed to analyse Canada's national land inventory data to aid in the development of land for agriculture.

The CGIS project was completed in 1971 and the software is still in use today. The CGIS project also involved a number of key innovations that have found their way into the feature set of many subsequent software developments.

From the mid-1960s to 1970s, developments in GIS were mainly occurring at government agencies and at universities. In 1964, Howard Fisher established the Harvard Lab for Computer Graphics where many of the industries early leaders studied. The Harvard Lab produced a number of mainframe GIS applications including: SYMAP (Synagraphic Mapping System), CALFORM, SYMVU, GRID, POLYVRT, and ODYSSEY. ODYSSEY was first modern vector GIS and many of its features would form the basis for future commercial applications.

Automatic Mapping System was developed by the United States Central Intelligence Agency (CIA) in the late 1960s. This project then spawned the CIA's World Data Bank, a collection of coastlines, rivers, and political boundaries, and the CAM software package that created maps at different scales from this data. This development was one of the first systematic map databases.

In 1969, Jack Dangermond, who studied at the Harvard Lab for Computer Graphics, co-founded Environmental Systems Research Institute (ESRI) with his wife Laura. ESRI would become in a few years the dominate force in the GIS marketplace and create ArcInfo and ArcView software. The first conference dealing with GIS took place in 1970 and was organized by Roger Tomlinson (key individual in the development of CGIS) and Duane Marble (professor at Northwestern University and early GIS innovator). Today, numerous conferences dealing with GIS run every year attracting thousands of attendants.

In the 1980s and 1990s, many GIS applications underwent substantial evolution in terms of features and analysis power. Many of these packages were being refined by private companies who could see the future commercial potential of this software. Some of the popular commercial applications launched during this period include: ArcInfo, ArcView, MapInfo, SPANS GIS, PAMAP GIS, INTERGRAPH, and SMALLWORLD. It was also during this period that many GIS applications moved from expensive minicomputer workstations to personal computer hardware.

Components of a GIS

A Geographic Information System combines computer cartography with a database management system. The major components common to a GIS. This diagram suggests that a GIS consists of three subsystems:

(1) an input system that allows for the collection of data to be used and analysed for some purpose;

(2) computer hardware and software systems that store the data, allow for data management and analysis, and can be used to display data manipulations on a computer monitor;

(3) an output system that generates hard copy maps, images, and other types of output.

Two basic types of data are normally entered into a GIS. The first type of data consists of real world phenomena and features that have some kind of spatial dimension. Usually, these data elements are depicted mathematically in the GIS as either points, lines, or polygons that are referenced geographically (or geocoded) to some type of coordinate system. This type data is entered into the GIS by devices like scanners, digitizers, GPS, air photos, and satellite imagery.

The other type of data is sometimes referred to as an attribute. Attributes are pieces of data that are connected or related to the points, lines, or polygons. This attribute data can be analysed to determine patterns of importance. Attribute data is entered directly into a database where it is associated with element data.

Plotted data can be defined as elements because their main purpose is to describe the location of the earthquakes.

For each of the earthquakes, the GIS also has data on their depth. These measurements can be defined as attribute data because they are connected to the plotted earthquake locations shows the attribute earthquake depth organized into three categories: shallow; intermediate; and deep. This analysis indicates a possible relationship between earthquake depth and spatial location-deep earthquakes do not occur at the mid-oceanic ridges.

Within the GIS database a user can enter, analyse, and manipulate data that is associated with some spatial element in the real world. The cartographic software of the GIS enables one to display the geographic information at any scale or projection and as a variety of layers which can be turned on or off. Each layer would show some different aspect of a place on the Earth. These layers could show things like a road network, topography, vegetation cover, streams and water bodies, or the distribution of annual precipitation received.

The difference between element and attribute data can be illustrated. The location of some of the earthquakes that have occurred in the last century. They plotted data points can be defined as elements because their main purpose is to describe the location of the earthquakes.

For each of the earthquakes, the GIS also has data on their depth. These measurements can be defined as attribute data because they are connected to the plotted earthquake locations. The attribute earthquake depth organized into three categories: shallow; intermediate; and deep. This analysis indicates a possible relationship between earthquake depth and spatial location-deep earthquakes do not occur at the mid-oceanic ridges. Within the GIS database a user can enter, analyse, and manipulate data that is associated with some spatial element in the real world. The cartographic software of the GIS enables one to display the geographic information at any scale or projection and as a variety of layers which can be turned on or off. Each layer would show some different aspect of a place on the Earth. These layers could show things like a road network, topography, vegetation cover, streams and water bodies, or the distribution of annual precipitation received. The merges data layers for vegetation community type, glaciers and ice fields, and water bodies (streams, lakes, and ocean).

What is a GIS?

Geographic Information Systems provide a method for integrating and analysing spatial (digital map based) information such as "where is the nearest movie theatre?" alongside related non-spatial information (what movies are playing there?). GIS have three major capabilities (computer mapping, spatial analysis and spatial database) and can operate on a range of platforms (desktop/laptop computer, Internet, PDA, etc). Many people are becoming far more familiar with seeing the results both textually-for example when their phone shows them the nearest pub-and on open map systems such as Google Maps. Where in the past people had to literally use pencils and string on a paper map to find their nearest school, a computer can do this now extremely quickly an accurately, as long as all the information has been entered correctly in the first place.

In a broader context, GIS involves people and often brings a philosophy of change. For example, in 1994, the New York Police Department introduced GIS to locate crime 'hot-spots', analyse underlying problems and devise strategies and solutions to deal with the problems. Since 1993, violent crime has dropped by two-thirds in New York City. This strategy, known as COMPSTAT, has expanded to cities and jurisdictions across the United States and around the world.

GIS Software

One leading GIS software vendor is ESRI, based in Redlands, California, which offers ArcGIS for the desktop, ArcGIS Server for Internet mapping, ArcPad for PDAs and a range of other products and services for developers. Other popular GIS software packages are available from Cadcorp, Intergraph, MapInfo, Manifold and Autodesk. ERDAS Imagine, ENVI, Idrisi, and PCI Geomatica are geared towards remote sensing i.e analysis of satellite/aircraft images.

There are many third-party extensions and utilities for ArcGIS and other GIS and raster software platforms. Currently, open source GIS software options can be chosen from the first OS GIS package GRASS, recent open source options are DIVA

GIS, Quantum GIS, and uDig. There are efforts underway, through the Open GIS Consortium to provide interoperability among spatial data formats and software. The leading contender for spatial data storage is another open source package called PostGIS, which is a spatial extension to the open source database PostgresSQL.

Why is Geospatial Topology Important In A GIS?

Geospatial analysis provides a unique perspective on the world, a tool through which to examine events, patterns, and processes that operate on or near the surface of our planet. To have this tool interact also with the relevant topological information is a must, that is needed to have the complete understanding of the way rivers flow, the weather is affected by the terrain or even how humans and other animals flow or chose to a specific habitat, this kind of information can only be gathered and useful if topological information is present in the GIS.

SPATIAL DATA

Spatial data comes in two major formats, as vector or raster data. The main difference is that a raster is usually a static background picture used to illustrate, whereas a vector is an intelligent ladder of information that can be selected and searched.

Vector (points, lines, polygons)

Vector data often represents anthropogenic (human) features such as roads, buildings, political boundaries (counties, congressional districts, etc), and other features such as lakes and rivers. Vector data is scalable without loss of resolution and is generally represented by XYZ points in a Cartesian frame reference.

Raster (grid/images)

Raster data is pixellated data, and the more pixels that map the data, the better the resolution. However, if raster data is enlarged, it simply enlarges the pixels, which then leads to a loss of resolution. There are many moves being made to make raster data more useable/searchable as it is much faster to collect, unlike vectors where each piece of data usually has a manual input origin.

Raster data is usually derived from satellite imagery or aerial photography (known as remote sensing). Ordinary cameras are only sensitive to visible light. Satellite sensors can capture not only visible light, but also the thermal, microwave, infrared or other types of energy emanating from the Earth's surface. This extra data provides information about sea surface temperatures, vegetation, ozone, etc. Remote sensing is also used to study other planets and extraterrestrial bodies, such as Mars.

AN OVERVIEW OF REMOTE SENSING

Remote sensing is the examination or the gathering of information about a place from a distance. Such examination can occur with devices (e.g. - cameras) based on the ground, and/or sensors or cameras based on ships, aircraft, satellites, or other spacecraft.

Today, the data obtained is usually stored and manipulated using computers. The most common software used in remote sensing is ERDAS Imagine, ESRI, MapInfo, and ERMapper.

A Brief History of Remote Sensing

Modern remote sensing began in 1858 when Gaspard-Felix Tournachon first took aerial photographs of Paris from a hot air balloon. Remote sensing continued to grow from there; one of the first planned uses of remote sensing occurred during the U.S. Civil War when messenger pigeons, kites, and unmanned balloons were flown over enemy territory with cameras attached to them.

The first governmental-organized air photography missions were developed for military surveillance during World Wars I and II but reached a climax during the Cold War.

Today, small remote sensors or cameras are used by law enforcement and the military in both manned and unmanned platforms to gain information about an area. Today's remote sensing imaging also includes infra-red, conventional air photos, and Doppler radar. In addition to these tools, satellites were developed during the late 20th century and are still used today to gain information on a global scale and even information about other planets in the solar system. For example, the Magellan probe is a satellite that has used remote sensing technologies to create topographic maps of Venus.

GIS: DEFINITION

Geographic Information System (GIS) is becoming more & more popular among decision makers as it enables them to quickly refer the GIS outputs which help them in solving problems and making right decisions. Visualization of features, converting data into need-based maps (thematic maps) and capability of providing solutions by taking into account overall scenario of an area are some of the virtues of GIS due to which it is being implemented across a number of sectors and departments (e.g.,transportation, forestry, environment, disaster management, urban planning, health etc). We often feel difficulty to understand and visualize a problem just by seeing data. We are more comfortable with visual representation of a problem that's what GIS is giving to us (that too with the true representation of real-world).

DEFINITION

GIS is made up of three terms-Geographic, Information and System. In literal meaning Geographic Information System is a System containing Information which is geographic in nature. GIS can be defined as-A System which involves collecting/capturing, storing, processing, manipulating, analysing, managing, retrieving and displaying data (information) which is, essentially, referenced to the real-world or the earth (i.e. geographically referenced).

Collection/Capturing

The dataset collected for GIS may be in the form of hard copy maps, satellite images, survey data or other data obtained from other primary and secondary sources. Collection of data depends on the objective of the assignment. Data capturing involves digitization of hard copy maps and satellite images.

Storage

In GIS Storage means not merely storing whatever data we have collected. The collected data is converted in usable GIS format and then finally stored for further use either on computer hard disk or in other storage devices (CD, DVD, magnetic tapes etc).

PROCESSING AND MANIPULATION

The collected and stored dataset is imported and converted into layers. Then required attributes are attached. Then data is processed for refinement, removing errors and preparing it for further GIS-based analysis.

Data manipulation is essential so that it can be represented in proper understandable form.

Analysis

Analysis of GIS data is required to convert it into desired outputs. There are many type of analysis in GIS which is (or are) to be done is objective dependant. The analysis may be statistical, spatial or specialized (like network analysis, utility analysis etc. Need not to say GIS analysis requires skilled professionals.

Management

Data management is essential and very important part of GIS for storing, managing and properly maintaining GIS database.

Retrieval

In GIS, data can be retrieved through SQL or spatial queries. Some software provide tools to retrieve data by simply selecting the features. Retrieval is used for getting information about the features of our interest.

Display

Displaying of final output may in many forms. These may be hard copy printouts, on-screen display of maps, internet-based map display (through Internet Map Servers) or in the form of presentation (like power point).

THE GIS APPROACH TO DECISION MAKING

Data Elements and Models

The interrelations among variables can be incredibly complex. To control the production process, we formulate a model for crop growth that accounts for the dynamic interactions and cause-effect responses of the appropriate factors. A model in this context establishes the functions, sequences, and feedback effects of the determining variables in the vigorous operation of a system.

As evident from the above diagram, we take information from various conventional sources, along with data from Landsat or similar Earth-observing satellites, and often combine it to produce data elements, which we further convolve to produce interpretation maps and other output types. Then these outputs become viable parts of environmental resource planning, site selection, and other outcomes that benefit from modelling.

Interpretative maps are derivative, i.e., they result from decisions to produce new object categories, stemming from combinations of several others.

For example, a map may plot the likelihood of runoff erosion, which we can deduce from maps showing soil properties, topography, rainfall, stream patterns, and other related factors. Variants of this map type are allocative or suitability maps (e.g., transportation routing), evaluative maps (e.g., earthquake damage potential), and predictive maps (e.g., 100-year flood coverage).

Until the 1970s, the traditional way of using geospatial data, making decisions based on combining information sources, and applying models by laying maps on a table was cumbersome. Interpreters inspected these maps individually, and mentally compared data sets-visual and tabular-while seeking to narrow parts of an area to locations that they pinpointed as suitable on several defining maps.

In principle, they could physically overlay two or three maps, treating several relevant themes by placing them on a light table (particularly if they were rephotographed as transparencies) to check on pre-selected, favourable conditions, where patterns on the maps superposed in a positive way.

This is, of course, inherently unwieldy and inefficient, and often the interpreter would simply glance back and forth between maps. One can improve the interrelationship process by laying a grid over each map and extracting information according to some relative merit or weighting, organized by location within the grid. Managing these data, usually in tables, proved labour-intensive, slow, and often ineffective.

Devices were developed to allow more systematic comparisons between maps. These comparisons occurred on a one-to-one basis by superimposing maps and photographs through an optical projector that allowed for differences in scale and even projection. One such device, still in use today, is the Zoom Transfer Scope (ZTS) manufactured by Bausch & Lomb.

The operator mounts a photo or a map at one level of the ZTS and manipulates mirrors and lenses so that its image, projected through prisms, appears to register (align) in the view scope over a second map or photo at the bottom. The operator then draws tracings from the first map/photo on a thin sheet over the second one or takes a photograph using a polaroid camera within the optical train, in which the two views merge.

Obviously, this form of image combination is severely limited to two images. However, one or more additional layers are feasible, if we display map and photo reproductions in other colours, but multiple registration becomes difficult.

GIS Defined

The maturing of the computer age has greatly changed the manner in which we can merge, compare, and manipulate multiple maps and other data sets. Computers and their software have significantly enhanced data handling capacity and flexibility. A powerful new tool, known as the Geographic Information System (GIS) emerged in the 1970s.

Many of those who developed GIS were inspired by the 1969 publication of the classic *Design with Nature* by Ian McHarg (Doubleday/Natural History Press), a leading landscape architect then at the University of Pennsylvania. This document pointed the way to planning and decision-making through comparative, integrated maps and related data types.

Since its inception, GIS has become a major growth industry, now conducted worldwide at the multi-billion-dollar level. It has blossomed into the main way for using maps (novel and practical) in most endeavours that focus and rely on geographically-based data of many kinds. Because remote sensing has routinely provided new images of the Earth's surface, it too has become intertwined with GIS as a means to constantly and inexpensively update some of the GIS data (such as land use and cover). The Association for Geographic Information defines GIS as:

A system for capturing, storing, checking, integrating, manipulating, analysing, and displaying data which are spatially referenced to the Earth.

A simpler working definition is: *A computer-based approach to interpreting maps and images and applying them to problem-solving.* The inclusion of computers to store, process, manipulate, interpret, and display GIS information is the critical ingredient that separate modern GIS from the more conventional (traditional) methods of using maps and correlative data prior to the 1970s. We synopsize the role of GIS in the general planning process for site selection, environmental management, and other geographically-dependent applications.

The driver for this closed-loop operation is the constant need for timely information about human activities and expectations concerning life in the real world.

The specifics underlying those needs define the types and amounts of data/information required. Once GIS users stipulate the specifics, they collect the data from multiformed sources, such as already published maps and tabulations, current field observations, surveys, and aerial/satellite imagery.

In the next step they convert the varied data into computer-compatible formats. The heart of the GIS operation lies within various techniques for analysis that users have devised as GIS evolved. They then present their reports, displays, new maps, statistics, and other kinds of computer-based, information-oriented products to decision-makers. The test of value then happens by applying the results in the same real world that dictated the initial requirements.

DECISION MAKING: SUITABILITY DETERMINATION

We embark now to further explore what GIS does, how it works, what software systems are available, and what typical end products and applications look like. Before moving on, you may wish to gain further insights into GIS by visiting sites on the Internet. One way is to work through some of the entries found through a Web Search Engine. We have picked out a few of the better ones whose links are given here.

First, look at a site developed by an (unidentified; but the author is listed as Dr. Shunji Murai) Japanese source found at this Web mirror site. Geographers at the U.S. Geological Survey prepared the second site, which they call the GIS Tutorial.

Another useful summary is found at the GIS Development site. The leading U.S. software providers are ERDAS with its Imagine and Leica Geosystems programs and ESRI, which also offers training courses. ESRI has produced a glossary appropriate to GIS. A general overview of GIS is found at GisPortal. Lists of GIS organizations and links can be accessed at GIS Resource List and GIS links. Perhaps the best way to appreciate the power of GIS, even before examining the design and function of the data handling system, is to introduce a typical case study and the thinking behind the steps in a site-suitability analysis. The following diagram summarizes the rationale behind such an analysis:

Suppose that three factors or variables, among the attributes that describe a geographic area under consideration, are essential in determining best sites for, say, a land development venture: Vegetation, Topography, and Soils. A data element on a map represents each variable. In this case the map shows the characteristics and distribution of the members or classes within the element theme, e.g., different types of soils and their properties.

We usually designate one map, often a cartographic map, as the base, over which we lay the others, (each then constitutes a data layer) either manually or digitally. Now, some soils, vegetation cover types, and elevations are more favourable than others in specifying their role in the site-selection process. Thus, for a certain intended use, we prefer high areas over low areas.

So, we assign relative heights numerical ratings, say from 1 to 5. We can mark soils with optimum drainage by higher numbers in a scale of 1 to 8. We then subdivide a data element map into cells in a grid. We assign each cell a value based on its thematic rating. We can incorporate other kinds of data, e.g., tables representing some condition in the cells, provided there is some spatial connection. When we combine the maps sharing the same cells, each comprising a data layer, the values sum for each cell (ranging from lowest numbers = worst suited to highest = best suited). In a modern GIS, we do this digitally. The outcome is a map, in which we judge areas with the highest resultant scores the most favourably suited.

CONDUCTING A GIS ANALYSIS

Many organizations have developed software packages for GIS analysis. The most widely used are ARC/INFO (Unix and Windows NT platforms) and ArcView (a smaller PC desktop version) marketed by the Environmental Systems Research Institute (ESRI) of Redlands, California, a company founded in 1969 by Jack and Laura Dangermond. Shane Murnion of Queens University in Belfast, created a tutorial for using ARC/INFO. Behind this table, on which we mount a map, are closely spaced electrical wires arranged in a grid that we can reference as x-y coordinates. The operator places a mobile puck with centered crosshairs over each point on the map and clicks a button to enter its position, along with a numerical code that records its attribute(s) into a computer database. He then moves to the next point, repeats the process, and enter a tie command.

We can capture the information contained in each field by either of two methods of geocoding: *Vector* or *Raster*. In the center panel, the approach creates polygons that approximate the curvature of the field boundary. If that field has irregular boundaries, we may need many lines. Each line consists of two end points (a vector), whose positions we mark by coordinates during digitizing.

Each point has a unique coordinate value. Two ends or node points define a line. We specify a polygon by a series of connecting nodes (any two adjacent lines share a node), which must close (in this writer's experience with digitizing, the main pitfall is that some polygons don't to close, and we must repeat or repair the process). We then identify each polygon by a proper code label (numerical or alphabetical). A look-up table associates the code characters with the attributes they represent. In this way, we can enter all the map fields that are large enough to be conveniently circumscribed, and their category values, into a digital database.

In the raster approach, we manually overlay or scan onto the map a grid array of cells, having some specific size. As shown in the right panel, an irregular polygon then includes a number of cells, completely contained therein.

The system records these cells' locations within the grid and a relevant code number for each data element assigned to them. But some cells straddle field boundaries. A preset rule assigns the cell to one or the other of adjacent fields, usually related to a relevant proportion of either field. The array of cells that comprise a field only approximate the field shape, but for most purposes the inaccuracy is tolerable for making calculations.

Generally, grid cells are larger than the enclosed pixels in pictorial map displays, but the cluster of pixels within a polygon approximates the shape of the field. The relation of cells to pixels makes this raster format well adapted to digital manipulation. The size of a cell depends partly on the internal variability of the represented feature or property. Smaller cells increase accuracy but also require more data storage. Note that multiple data layers referenced to the same grid cell share this spatial dimensionality but have different coded values for the various attributes associated with any given cell.

Data management is sensitive to storage retrieval methods and to file structures. A good management software package should be able to:

I. Scale and rotate coordinate values for "best fit" projection overlays and changes.
II. Convert (interchange) between polygon and grid formats.
III. Permit rapid updating, allowing data changes with relative ease.
IV. Allow for multiple users and multiple interactions between compatible data bases.
V. Retrieve, transform, and combine data elements efficiently.
VI. Search, identify, and route a variety of different data items and score these values with assigned weighted values, to facilitate proximity and routing analysis.
VII. Perform statistical analysis, such as multivariate regression, correlations, etc.
VIII. Overlay one file variable onto another, i.e., map superpositioning.
IX. Measure area, distance, and association between points and fields.
X. Model and simulate, and formulate predictive scenarios, in a fashion that allows for direct interactions between the user group and the computer program.

Developing a GIS can be a costly, complex, and somewhat frustrating experience for the novitiate. We stress that data base design and encoding are major tasks that demand time, skilled personnel, and adequate funds. However, once developed, the information possibilities are exciting, and the intrinsic worth of the output more than compensates for the marginal costs of handling the various kinds of data. In plain language, GIS is a systematic, versatile, and comprehensive way to present, interpret, and recast spatial (geographic) data into intelligible output.

THE PENNSYLVANIA POWER AND LIGHT NEW PLANT SITING PROBLEM

Perhaps GIS analysis will come through more convincingly with a second case study. This case is an applications demonstration, developed at NASA Goddard during the early 1980s, when its Eastern Regional Remote Sensing Applications

Center (ERRSAC) still operated. The “client” was the Pennsylvania Power and Light (PP&L) Company, a public electric utility, whose service areas include parts of central and eastern PA. One service area is the state capital district at Harrisburg on the Susquehanna River that had been receiving much of its power from the infamous Three Mile Island, site of a near-catastrophic nuclear power-plant accident on March 28, 1979. This experience forced PP&L to have special concerns regarding any of its future sitings, for its own power facilities and for its large customers. Specifically, long range planning called for a possible second generating plant to be constructed in the Harrisburg area. Could both GIS and remote sensing inputs help in the decisions involved in site selection?

Although closing remarks are given in the next paragraphs, we urge you to then proceed to the next page, which presents a third, and very detailed and intriguing, case study that was received from an outside contributor at almost the last moment before the present version of the Tutorial was “declared complete”.

Meanwhile, if you are new to GIS, you now should know enough to begin practicing on your own. There is a web site, sponsored by the Research Program for Environmental Planning and Geographic Information Systems (REGIS), a group of geographers at the University of California at Berkeley, that has data sets (maps, aerial photos, and space imagery) of the San Francisco Bay Area, which you can display, combine, and use to output new products. This is the BAGIS program, done with a version of GRASS, called GRASSlinks, developed at UCB by Dr. Susan Huse and her colleagues. It’s fun and a challenge to work with. Access it here (UC Berkeley Regis) site and follow the instructions to build your skills at deriving meaningful maps. Links at this site also guide you to other projects by this group. Still another site built around the San Francisco region is accessed at another Regis Program.

GIS has now been used in tens of thousands of studies. But the Internet has a dearth of good ones that go into enough detail to appraise you of what really can be done.

Still, a search (try entering “GIS + Case Studies”) on the Web will lead you to some examples that at the least hint at the power of this approach. Perhaps the best affordable way to learn to use GIS is to purchase the IDRISI Program, which guides you through most of the essentials and illustrates these with examples.

THE ADVANTAGES/DISADVANTAGES OF GIS

Often when discussing the advantages and disadvantages of a certain commodity, the presentation of the advantages outweigh the disadvantages and a holistic impression is not given. This is not the case however with GIS in archaeology. In an attempt to improve the disciplines use and understanding of GIS, many of the problems or disadvantages associated with the numerous packages available are well represented. This may appear to be a strange approach when considering that the various authors are actually advocates of the technology. In turn this approach tends to promote the current (and future) potential of the use of GIS in archaeology. It is therefore difficult to produce a paper based on the praises of GIS or its uses alone as the problems often outweigh the prospects.

The advantages of using a GIS tend to speak for themselves and have been extensively covered when discussing what a GIS actually is/does. One of the main attractions is the ease at which information can be presented, which is also subject to much criticism. This provides much of the basis for the support of the technology. The extensive data handling capabilities also impress users, providing a depth of knowledge and wide span of information never before experienced by archaeologists using conventional recording methods.

In general the advantages of GIS technology are encountered through the tasks that the various packages can perform. The disadvantages, however tend to receive more publicity amongst the archaeological writings on the subject. There are two main areas identified as presenting the most problems with archaeological adoption of GIS. These relate to the lack of defined data standards within archaeological GIS and the use and abuse of images over data.

The lack of an archaeological data standard is probably the more worrying of the two problems / disadvantages. Many computer illiterate and sceptical archaeologists have reservations about adopting a technology which fails to adhere to any set conventions. The issue of data standards or lack of them (e.g. different regions tend to develop different database fields, formats and data structures which are more often than not highly incompatible), for some presents a major problem. Even at its most basic level, the data standards issue reflects a severe short coming in the potential of widespread GIS use. This problem however, is not restricted to merely GIS packages and often reflects problems with database convention and design rather than fault with the commercial programmes. The database debate is not of much concern to this chapter.

The answer to such a problem appears to be simple - introduce an archaeological data standard which would be cross-compatible. But this would also present specific problems as there are then the difficulties of cultural and linguistic barriers to overcome. This makes the goal of a standard database or GIS 'unity' almost impossible. As far back as 1992, there has been the realisation that a need for a GIS data standard is a pressing issue, "If there is to be in Europe's future an integration of archaeological research it must of course be computer based, and for the computerization to work, it is an intrinsic necessity that we structure and formalize the fundamental way we record and store data."

By allowing a myriad of independently developing databases it becomes impossible for anyone to logically understand all of them. This makes research involving GIS increasingly difficult as the use of database GIS is becoming more popular.

Added to the problem of data standards comes the problem of commercial viability amongst affordable GIS packages. It has been suggested (J. Huggett Pers Comm.) that in the future, commercial GIS software may determine the nature of archaeological research, due to limited design and function. As specific strands of archaeology may require specialist procedures, which are not as commercially viable for inclusion into basic packages, the situation of the 1980's where archaeologists had to 'programme' their own packages / functions may return. Kvamme (1995) similarly addresses this subject but offers a solution by attempting to "maintain a core of high technical expertise within the discipline." It may be that by introducing a specific 'Archaeological Information System', many of these problems would be resolved. But the door would then be open for a number of other 'inter-disciplinary' struggles as to what functions would then become acceptable. This will be discussed further in the section below on the future of GIS.

A further area witnessing much criticism within GIS in archaeology, is that of graphics versus data. Certain archaeologists such as Stephan Shennan warn against the use of "pretty pictures" in place of formal data analysis. This appears to be a problem generated by many 'traditional' archaeologists who believe that the production of well designed computer graphics detracts from the true reason for the use of such a package, statistical spatial analysis. This argument, however, should not have much structure amongst ' morally decent' archaeologists. Nowadays, there should be no need to compromise between effective graphics or convincing statistics, most 'modern' GIS packages should offer the ability to produce both. That is not to say that there are not archaeologists willing to negotiate the statistical results in place of a 'pretty picture'. As Kvamme (1995) comments, "There are already enough examples of GIS studies where one must question whether there is any substance behind the stunning visual effects."

Hopefully, these examples are few and far between and considerably in decline, instead favouring more respectable graphical archaeological results.

It has been implied that most GIS packages do not offer very effective quantitative spatial analysis. This is partly related to the argument of graphics versus statistics. Whereas the majority of current GIS packages are geared towards graphical output and inference from pictures, many archaeologists still want to be able to see the spatial statistics. Some packages attempt to rectify this by providing the ability to produce 'thematic maps'. These thematic maps provide the opportunity to view the statistical data in a simple yet impressive format, often integrating charted data with coordinated mapped outputs.

Much of the criticism directed towards GIS reflects the different approaches to and definitions of GIS as being highly individual perspectives. It is hoped that time and experience will affect the future applications of GIS in archaeology.

THE FUTURE OF GIS

Much of the future of GIS rests with the development of culture, heritage and resource management databases, as this is where the majority of the funding lies. Monument records and heritage management present the largest future potential for GIS within the archaeological community as the analytical aspect is not as pressing here. The much needed improvements with regard to data handling and representation are not as substantial in the management sector. The data standards issue however, is just as important here (if not more) as it is in any other archaeological situation.

The development of a specific Archaeological Information System (AIS) may solve some of the discipline specific problems encountered. There have been attempts to produce such and AIS, such as the ArchaeDATA project, where the emphasis is on “structuring a European archaeological information system”.

Moving away from the purely scientific approach, there has been a call to advance the use of GIS in theoretical models at a more humanistic level. Such an approach has faced the accusation that this is particularly European attempt to incorporate GIS technology and social theory, similar advances, have also emerged in north America recently.

There are numerous more arguments concerning the advantages, disadvantages and future of GIS technology within archaeology, unfortunately this tends to go beyond the scope of this study. What does emerge as a pressing issue that has not been resolved since most of the influential papers were written in the mid 1990’s, is that of data standards and the lack of ability to cope with the mapping of temporal aspects within archaeology. Without the resolutions to such pressing issues in the use of GIS, there will remain to be the standstill that has occurred since the initial development and adoption of such a technology.

2

Introduction to Remote Sensing

Remote sensing is the small-or large-scale acquisition of information of an object or phenomenon, by the use of either recording or real-time sensing device(s) that are wireless, or not in physical or intimate contact with the object (such as by way of aircraft, spacecraft, satellite, buoy, or ship). In practice, remote sensing is the stand-off collection through the use of a variety of devices for gathering information on a given object or area. Thus, Earth observation or weather satellite collection platforms, ocean and atmospheric observing weather buoy platforms, the monitoring of a parolee via an ultrasound identification system, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), X-radiation (X-RAY) and space probes are all examples of remote sensing. In modern usage, the term generally refers to the use of imaging sensor technologies including: instruments found in aircraft and spacecraft as well as those used in electrophysiology, and is distinct from other imaging-related fields such as medical imaging.

OVERVIEW

There are two main types of remote sensing: passive remote sensing and active remote sensing. Passive sensors detect natural radiation that is emitted or reflected by the object or surrounding area being observed. Reflected sunlight is the most common source of radiation measured by passive sensors. Examples of passive remote sensors include film photography, infrared, charge-coupled devices, and radiometers. Active collection, on the other hand, emits energy in order to scan objects and areas whereupon a sensor then detects and measures the radiation that is reflected or backscattered from the target. RADAR is an example of active remote sensing where the time delay between emission and return is measured, establishing the location, height, speed and direction of an object. Remote sensing makes it possible to collect data on dangerous or inaccessible areas. Remote sensing applications include monitoring deforestation in areas such as the Amazon Basin, glacial features in Arctic and Antarctic regions, and depth sounding of coastal and ocean depths. Military collection during the cold war made use of stand-off collection of data about dangerous border areas. Remote sensing also replaces costly and slow data collection on the ground, ensuring in the process that areas or objects are not disturbed.

Orbital platforms collect and transmit data from different parts of the electromagnetic spectrum, which in conjunction with larger scale aerial or ground-based sensing and analysis, provides researchers with enough information to monitor trends such as El Niño and other natural long and short term phenomena. Other uses include different areas of the earth sciences such as natural resource management, agricultural fields such as land usage and conservation, and national security and overhead, ground-based and stand-off collection on border areas.

Data Acquisition Techniques

The basis for multispectral collection and analysis is that of examined areas or objects that reflect or emit radiation that stand out from surrounding areas.

Applications of Remote Sensing Data

- Conventional radar is mostly associated with aerial traffic control, early warning, and certain large scale meteorological data. Doppler radar is used by local law enforcements' monitoring of speed limits and in enhanced meteorological collection such as wind speed and direction within weather systems. Other types of active collection includes plasmas in the ionosphere). Interferometric synthetic aperture radar is used to produce precise digital elevation models of large scale terrain.

- Laser and radar altimeters on satellites have provided a wide range of data. By measuring the bulges of water caused by gravity, they map features on the seafloor to a resolution of a mile or so. By measuring the height and wave-length of ocean waves, the altimeters measure wind speeds and direction, and surface ocean currents and directions.
- Light detection and ranging (LIDAR) is well known in examples of weapon ranging, laser illuminated homing of projectiles. LIDAR is used to detect and measure the concentration of various chemicals in the atmosphere, while airborne LIDAR can be used to measure heights of objects and features on the ground more accurately than with radar technology. Vegetation remote sensing is a principal application of LIDAR.
- Radiometers and photometers are the most common instrument in use, collecting reflected and emitted radiation in a wide range of frequencies. The most common are visible and infrared sensors, followed by microwave, gamma ray and rarely, ultraviolet. They may also be used to detect the emission spectra of various chemicals, providing data on chemical concentrations in the atmosphere.
- Stereographic pairs of aerial photographs have often been used to make topographic maps by imagery and terrain analysts in trafficability and highway departments for potential routes.
- Simultaneous multi-spectral platforms such as Landsat have been in use since the 70's. These thematic mappers take images in multiple wavelengths of electro-magnetic radiation (multi-spectral) and are usually found on Earth observation satellites, including (for example) the Landsat program or the IKONOS satellite. Maps of land cover and land use from thematic mapping can be used to prospect for minerals, detect or monitor land usage, deforestation, and examine the health of indigenous plants and crops, including entire farming regions or forests.

- Within the scope of the combat against desertification, remote sensing allows to follow-up and monitor risk areas in the long term, to determine desertification factors, to support decision-makers in defining relevant measures of environmental management, and to assess their impacts.

Geodetic

- Overhead geodetic collection was first used in aerial submarine detection and gravitational data used in military maps. This data revealed minute perturbations in the Earth's gravitational field (geodesy) that may be used to determine changes in the mass distribution of the Earth, which in turn may be used for geological or hydrological studies.

Acoustic and Near-acoustic

- Sonar: *passive sonar*, listening for the sound made by another object (a vessel, a whale etc); *active sonar*, emitting pulses of sounds and listening for echoes, used for detecting, ranging and measurements of underwater objects and terrain.
- Seismograms taken at different locations can locate and measure earthquakes (after they occur) by comparing the relative intensity and precise timing.

To coordinate a series of large-scale observations, most sensing systems depend on the following: platform location, what time it is, and the rotation and orientation of the sensor. High-end instruments now often use positional information from satellite navigation systems.

The rotation and orientation is often provided within a degree or two with electronic compasses. Compasses can measure not just azimuth (i.e. degrees to magnetic north), but also altitude (degrees above the horizon), since the magnetic field curves into the Earth at different angles at different latitudes.

More exact orientations require gyroscopic-aided orientation, periodically realigned by different methods including navigation from stars or known benchmarks.

Resolution impacts collection and is best explained with the following relationship: less resolution=less detail & larger coverage, More resolution=more detail, less coverage. The skilled management of collection results in cost-effective collection and avoid situations such as the use of multiple high resolution data which tends to clog transmission and storage infrastructure.

DATA PROCESSING

Generally speaking, remote sensing works on the principle of the *inverse problem*. While the object or phenomenon of interest (the state) may not be directly measured, there exists some other variable that can be detected and measured (the observation), which may be related to the object of interest through the use of a data-derived computer model. The common analogy given to describe this is trying to determine the type of animal from its footprints.

For example, while it is impossible to directly measure temperatures in the upper atmosphere, it is possible to measure the spectral emissions from a known chemical species (such as carbon dioxide) in that region. The frequency of the emission may then be related to the temperature in that region via various thermodynamic relations. The quality of remote sensing data consists of its spatial, spectral, radiometric and temporal resolutions.

Spatial Resolution

The size of a pixel that is recorded in a raster image – typically pixels may correspond to square areas ranging in side length from 1 to 1,000 metres (3.3 to 3,300 ft).

Spectral Resolution

The wavelength width of the different frequency bands recorded – usually, this is related to the number of frequency bands recorded by the platform. Current Landsat collection is that of seven bands, including several in the infra-red spectrum, ranging from a spectral resolution of 0.07 to 2.1 µm. The Hyperion sensor on Earth Observing-1 resolves 220 bands from 0.4 to 2.5 ìm, with a spectral resolution of 0.10 to 0.11 ìm per band.

Radiometric Resolution

The number of different intensities of radiation the sensor is able to distinguish. Typically, this ranges from 8 to 14 bits, corresponding to 256 levels of the gray scale and up to 16,384 intensities or “shades” of colour, in each band. It also depends on the instrument noise.

Temporal Resolution

The frequency of flyovers by the satellite or plane, and is only relevant in time-series studies or those requiring an averaged or mosaic image as in deforesting monitoring. This was first used by the intelligence community where repeated coverage revealed changes in infrastructure, the deployment of units or the modification/introduction of equipment. Cloud cover over a given area or object makes it necessary to repeat the collection of said location.

In order to create sensor-based maps, most remote sensing systems expect to extrapolate sensor data in relation to a reference point including distances between known points on the ground.

This depends on the type of sensor used. For example, in conventional photographs, distances are accurate in the center of the image, with the distortion of measurements increasing the farther you get from the center. Another factor is that of the platen against which the film is pressed can cause severe errors when photographs are used to measure ground distances. The step in which this problem is resolved is called georeferencing, and involves computer-aided matching up of points in the image (typically 30 or more points per image) which is extrapolated with the use of an established benchmark, “warping” the image to produce accurate spatial data. As of the early 1990s, most satellite images are sold fully georeferenced.

In addition, images may need to be radiometrically and atmospherically corrected. Radiometric correction gives a scale to the pixel values, e.g. the monochromatic scale of 0 to 255 will be converted to actual radiance values.

Atmospheric correction eliminates atmospheric haze by rescaling each frequency band so that its minimum value (usually realised in water bodies) corresponds to a pixel value of 0. The digitizing of data also make possible to manipulate the data by changing gray-scale values.

Interpretation is the critical process of making sense of the data. The first application was that of aerial photographic collection which used the following process; spatial measurement through the use of a light table in both conventional single or stereographic coverage, added skills such as the use of photogrammetry, the use of photomosaics, repeat coverage, Making use of objects' known dimensions in order to detect modifications.

Image Analysis is the recently developed automated computer-aided application which is in increasing use. Object-Based Image Analysis (OBIA) is a sub-discipline of GIScience devoted to partitioning remote sensing (RS) imagery into meaningful image-objects, and assessing their characteristics through spatial, spectral and temporal scale.

Old data from remote sensing is often valuable because it may provide the only long-term data for a large extent of geography. At the same time, the data is often complex to interpret, and bulky to store.

Modern systems tend to store the data digitally, often with lossless compression. The difficulty with this approach is that the data is fragile, the format may be archaic, and the data may be easy to falsify. One of the best systems for archiving data series is as computer-generated machine-readable ultrafiche, usually in typefonts such as OCR-B, or as digitized half-tone images.

Ultrafiches survive well in standard libraries, with lifetimes of several centuries. They can be created, copied, filed and retrieved by automated systems. They are about as compact as archival magnetic media, and yet can be read by human beings with minimal, standardized equipment.

Data Processing Levels

To facilitate the discussion of data processing in practice, several processing "levels" were first defined in 1986 by NASA as part of its Earth Observing System and steadily adopted since then, both internally at NASA and elsewhere; these definitions are:

Level	Description
0	Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e.g., synchronization frames, communications headers, duplicate data) removed.
1a	Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to the Level 0 data (or if applied, in a manner that level 0 is fully recoverable from level 1a data).
1b	Level 1a data that have been processed to sensor units (e.g., radar backscatter cross section, brightness temperature, etc.); not all instruments have Level 1b data; level 0 data is not recoverable from level 1b data.
2	Derived geophysical variables (e.g., ocean wave height, soil moisture, ice concentration) at the same resolution and location as Level 1 source data.
3	Variables mapped on uniform spacetime grid scales, usually with some completeness and consistency (e.g., missing points interpolated, complete regions mosaicked together from multiple orbits, etc).
4	Model output or results from analyses of lower level data (i.e., variables that were not measured by the instruments but instead are derived from these measurements).

A Level 1 data record is the most fundamental (i.e., highest reversible level) data record that has significant scientific utility, and is the foundation upon which all subsequent data sets are produced. Level 2 is the first level that is directly usable for most scientific applications; its value is much greater than the lower levels. Level 2 data sets tend to be less voluminous than Level 1 data because they have been reduced temporally, spatially, or spectrally. Level 3 data sets are generally smaller than lower level data sets and thus can be dealt with without incurring a great deal of data handling overhead. These data tend to be generally more useful for many applications. The regular spatial and temporal organization of Level 3 datasets makes it feasible to readily combine data from different sources.

HISTORY

Beyond the primitive methods of remote sensing our earliest ancestors used (ex.: standing on a high cliff or tree to view the landscape), the modern discipline arose with the development of flight. The balloonist G. Tournachon (alias Nadar) made photographs of Paris from his balloon in 1858. Messenger pigeons, kites, rockets and unmanned balloons were also used for early images. With the exception of balloons, these first, individual images were not particularly useful for map making or for scientific purposes.

Systematic aerial photography was developed for military surveillance and reconnaissance purposes beginning in World War I and reaching a climax during the Cold War with the use of modified combat aircraft such as the P-51, P-38, RB-66 and the F-4C, or specifically designed collection platforms such as the U2/TR-1, SR-71, A-5 and the OV-1 series both in overhead and stand-off collection. A more recent development is that of increasingly smaller sensor pods such as those used by law enforcement and the military, in both manned and unmanned platforms. The advantage of this approach is that this requires minimal modification to a given airframe. Later imaging technologies would include Infra-red, conventional, doppler and synthetic aperture radar.

The development of artificial satellites in the latter half of the 20th century allowed remote sensing to progress to a global scale as of the end of the Cold War. Instrumentation aboard various Earth observing and weather satellites such as Landsat, the Nimbus and more recent missions such as RADARSAT and UARS provided global measurements of various data for civil, research, and military purposes. Space probes to other planets have also provided the opportunity to conduct remote sensing studies in extraterrestrial environments, synthetic aperture radar aboard the Magellan spacecraft provided detailed topographic maps of Venus, while instruments aboard SOHO allowed studies to be performed on the Sun and the solar wind, just to name a few examples.

Recent developments include, beginning in the 1960s and 1970s with the development of image processing of satellite imagery. Several research groups in Silicon Valley including NASA Ames Research Center, GTE and ESL Inc. developed Fourier transform techniques leading to the first notable enhancement of imagery data. The introduction of online web services for easy access to remote sensing data in the 21st century (mainly low/medium-resolution images), like Google Earth, has made remote sensing more familiar to the big public and has popularized the science.

REMOTE SENSING SOFTWARE

Remote Sensing data is processed and analysed with computer software, known as a remote sensing application. A large number of proprietary and open source applications exist to process remote sensing data. According to an NOAA Sponsored Research by Global Marketing Insights, Inc. the most used applications among Asian academic groups involved in remote sensing are as follows: ERDAS 36% (ERDAS IMAGINE 25% & ERMapper 11%); ESRI 30%; ITT Visual Information Solutions ENVI 17%; MapInfo 17%. Among Western Academic respondents as follows: ESRI 39%, ERDAS IMAGINE 27%, MapInfo 9%, AutoDesk 7%, ITT Visual Information Solutions ENVI 17%. Other important Remote Sensing Software packages include: TNTmips from MicroImages, PCI Geomatica made by PCI Geomatics, the leading remote sensing software package in Canada, IDRISI from Clark Labs, Image Analyst from Intergraph, and the original object based image analysis software eCognition from Definiens. Dragon/ips is one of the oldest remote sensing packages still available, and is in some cases free. Open source remote sensing software includes GRASS GIS, QGIS, OSSIM, Opticks (software) and Orfeo toolbox.

WHAT IS REMOTE SENSING?

We perceive the surrounding world through our five senses. Some senses (touch and taste) require contact of our sensing organs with the objects. However, we acquire much information about our surrounding through the senses of sight and hearing which do not require close contact between the sensing organs and the external objects. In another word, we are performing Remote Sensing all the time.

Generally, Remote sensing refers to the activities of recording/observing/perceiving (sensing) objects or events at far away (remote) places. In remote sensing, the sensors are not in direct contact with the objects or events being observed. The information needs a physical carrier to travel from the objects/events to the sensors through an intervening medium. The electromagnetic radiation is normally used as an information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed. A further step of image analysis and interpretation is required in order to extract useful information from the image. The human visual system is an example of a remote sensing system in this general sense.

In a more restricted sense, remote sensing usually refers to the technology of acquiring information about the earth's surface (land and ocean) and atmosphere using sensors onboard airborne (aircraft, balloons) or spaceborne (satellites, space shuttles) platforms.

SATELLITE REMOTE SENSING

In this CD, you will see many remote sensing images around Asia acquired by earth observation satellites. These remote sensing satellites are equipped with sensors looking down to the earth. They are the "eyes in the sky" constantly observing the earth as they go round in predictable orbits.

EFFECTS OF ATMOSPHERE

In satellite remote sensing of the earth, the sensors are looking through a layer of atmosphere separating the sensors from the Earth's surface being observed. Hence, it is essential to understand the effects of atmosphere on the electromagnetic radiation travelling from the Earth to the sensor through the atmosphere. The atmospheric constituents cause wavelength dependent absorption and scattering of radiation. These effects degrade the quality of images. Some of the atmospheric effects can be corrected before the images are subjected to further analysis and interpretation.

A consequence of atmospheric absorption is that certain wavelength bands in the electromagnetic spectrum are strongly absorbed and effectively blocked by the atmosphere. The wavelength regions in the electromagnetic spectrum usable for remote sensing are determined by their ability to penetrate atmosphere. These regions are known as the atmospheric transmission windows. Remote sensing systems are often designed to operate within one or more of the atmospheric windows.

These windows exist in the microwave region, some wavelength bands in the infrared, the entire visible region and part of the near ultraviolet regions. Although the atmosphere is practically transparent to x-rays and gamma rays, these radiations are not normally used in remote sensing of the earth.

OPTICAL AND INFRARED REMOTE SENSING

In Optical Remote Sensing, optical sensors detect solar radiation reflected or scattered from the earth, forming images resembling photographs taken by a camera high up in space. The wavelength region usually extends from the visible and near infrared (commonly abbreviated as VNIR) to the short-wave infrared (SWIR).

Different materials such as water, soil, vegetation, buildings and roads reflect visible and infrared light in different ways. They have different colours and brightness when seen under the sun. The interpretation of optical images require the knowledge of the spectral reflectance signatures of the various materials (natural or man-made) covering the surface of the earth. There are also infrared sensors measuring the thermal infrared radiation emitted from the earth, from which the land or sea surface temperature can be derived.

MICROWAVE REMOTE SENSING

There are some remote sensing satellites which carry passive or active microwave sensors. The active sensors emit pulses of microwave radiation to illuminate the areas to be imaged. Images of the earth surface are formed by measuring the microwave energy scattered by the ground or sea back to the sensors.

These satellites carry their own "flashlight" emitting microwaves to illuminate their targets. The images can thus be acquired day and night. Microwaves have an additional advantage as they can penetrate clouds. Images can be acquired even when there are clouds covering the earth surface.

A microwave imaging system which can produce high resolution image of the Earth is the synthetic aperture radar (SAR). The intensity in a SAR image depends on the amount of microwave backscattered by the target and received by the SAR antenna. Since the physical mechanisms responsible for this backscatter is different for microwave, compared to visible/ infrared radiation, the interpretation of SAR images requires the knowledge of how microwaves interact with the targets.

REMOTE SENSING IMAGES

Remote sensing images are normally in the form of digital images. In order to extract useful information from the images, image processing techniques may be employed to enhance the image to help visual interpretation, and to correct or restore the image if the image has been subjected to geometric distortion, blurring or degradation by other factors.

There are many image analysis techniques available and the methods used depend on the requirements of the specific problem concerned. In many cases, image segmentation and classification algorithms are used to delineate different areas in an image into thematic classes. The resulting product is a thematic map of the study area. This thematic map can be combined with other databases of the test area for further analysis and utilization.

VISUAL SYSTEM

The human visual system is an example of a remote sensing system in the general sense. The sensors in this example are the two types of photosensitive cells, known as the cones and the rods, at the retina of the eyes. The cones are responsible for colour vision. There are three types of cones, each being sensitive to one of the red, green, and blue regions of the visible spectrum. Thus, it is not coincidental that the modern computer display monitors make use of the same three primary colours to generate a multitude of colours for displaying colour images. The cones are insensitive under low light illumination condition, when their jobs are taken over by the rods. The rods are sensitive only to the total light intensity. Hence, everything appears in shades of grey when there is insufficient light.

As the objects/events being observed are located far away from the eyes, the information needs a carrier to travel from the object to the eyes. In this case, the information carrier is the visible light, a part of the electromagnetic spectrum. The objects reflect/scatter the ambient light falling onto them. Part of the scattered light is intercepted by the eyes, forming an image on the retina after passing through the optical system of the eyes. The signals generated at the retina are carried via the nerve fibres to the brain, the central processing unit (CPU) of the visual system. These signals are processed and interpreted at the brain, with the aid of previous experiences.

When operating in this mode, the visual system is an example of a “Passive Remote Sensing” system which depends on an external source of energy to operate. We all know that this system won’t work in darkness. However, we can still see at night if we provide our own source of illumination by carrying a flashlight and shining the beam towards the object we want to observe. In this case, we are performing “Active Remote Sensing”, by supplying our own source of energy for illuminating the objects.

THE PLANET EARTH

The planet Earth is the third planet in the solar system located at a mean distance of about 1.50×10^8 km from the sun, with a mass of 5.97×10^{24} kg. Descriptions of the shape of the earth have evolved from the flat-earth model, spherical model to the currently accepted ellipsoidal model derived from accurate ground surveying and satellite measurements. A number of reference ellipsoids have been defined for use in identifying the three dimensional coordinates (i.e. position in space) of a point on or above the earth surface for the purpose of surveying, mapping and navigation. The reference ellipsoid in the World Geodetic System 1984 (WGS-84) commonly used in satellite Global Positioning System (GPS) has the following parameters:

- Equatorial Radius = 6378.1370 km
- Polar Radius = 6356.7523 km

The earth's crust is the outermost layer of the earth's land surface. About 29.1% of the earth's crust area is above sea level. The rest is covered by water. A layer of gaseous atmosphere envelopes the earth's surface.

THE EARTH'S ATMOSPHERE

The earth's surface is covered by a layer of atmosphere consisting of a mixture of gases and other solid and liquid particles. The gaseous materials extend to several hundred kilometres in altitude, though there is no well defined boundary for the upper limit of the atmosphere. The first 80 km of the atmosphere contains more than 99% of the total mass of the earth's atmosphere.

Vertical Structure of the Atmosphere

The vertical profile of the atmosphere is divided into four layers: troposphere, stratosphere, mesosphere and thermosphere. The tops of these layers are known as the tropopause, stratopause, mesopause and thermopause, respectively.

Troposphere: This layer is characterized by a decrease in temperature with respect to height, at a rate of about 6.5^{o}C per kilometer, up to a height of about 10 km. All the weather activities (water vapour, clouds, precipitation) are confined to this layer. A layer of aerosol particles normally exists near to the earth surface. The aerosol concentration decreases nearly exponentially with height, with a characteristic height of about 2 km.

Stratosphere: The temperature at the lower 20 km of the stratosphere is approximately constant, after which the temperature increases with height, up to an altitude of about 50 km. Ozone exists mainly at the stratopause. The troposphere and the stratosphere together account for more than 99% of the total mass of the atmosphere.

Mesosphere: The temperature decreases in this layer from an altitude of about 50 km to 85 km.

Thermosphere: This layer extends from about 85 km upward to several hundred kilometres. The temperature may range from 500 K to 2000 K. The gases exist mainly in the form of thin plasma, i.e. they are ionized due to bombardment by solar ultraviolet radiation and energetic cosmic rays.

The term upper atmosphere usually refers to the region of the atmosphere above the troposphere. Many remote sensing satellites follow the near polar sun-synchronous orbits at a height around 800 km, which is well above the thermopause.

Atmospheric Constituents

The atmosphere consists of the following components:

Permanent Gases: They are gases present in nearly constant concentration, with little spatial variation. About 78% by volume of the atmosphere is nitrogen while the life-sustaining oxygen occupies 21%. The remaining one percent consists of the inert gases, carbon dioxide and other gases.

Gases with Variable Concentration: The concentration of these gases may vary greatly over space and time. They consist of water vapour, ozone, nitrogeneous and sulphurous compounds.

Solid and liquid particulates: Other than the gases, the atmosphere also contains solid and liquid particles such as aerosols, water droplets and ice crystals. These particles may congregate to form clouds and haze.

ELECTROMAGNETIC RADIATION

Electromagnetic waves are energy transported through space in the form of periodic disturbances of electric and magnetic fields. All electromagnetic waves travel through space at the same speed, $c = 2.99792458 \times 10^8$ m/s, commonly known as the speed of light. An electromagnetic wave is characterized by a frequency and a wavelength. These two quantities are related to the speed of light by the equation,

speed of light = frequency × wavelength

The frequency (and hence, the wavelength) of an electromagnetic wave depends on its source. There is a wide range of frequency encountered in our physical world, ranging from the low frequency of the electric waves generated by the power transmission lines to the very high frequency of the gamma rays originating from the atomic nuclei. This wide frequency range of electromagnetic waves constitute the Electromagnetic Spectrum.

THE ELECTROMAGNETIC SPECTRUM

The electromagnetic spectrum can be divided into several wavelength (frequency) regions, among which only a narrow band from about 400 to 700 nm is visible to the human eyes. Note that there is no sharp boundary between these regions.

Wavelength units: 1 mm = 1000 µm; 1 µm = 1000 nm.

- Radio Waves: 10 cm to 10 km wavelength
- Microwaves: 1 mm to 1 m wavelength. The microwaves are further divided into different frequency (wavelength) bands: (1 GHz = 10^9 Hz)
  - P band: 0.3-1 GHz (30-100 cm)
  - L band: 1-2 GHz (15-30 cm)
  - S band: 2-4 GHz (7.5-15 cm)
  - C band: 4-8 GHz (3.8-7.5 cm)
  - X band: 8-12.5 GHz (2.4-3.8 cm)
  - Ku band: 12.5-18 GHz (1.7-2.4 cm)
  - K band: 18-26.5 GHz (1.1-1.7 cm)
  - Ka band: 26.5-40 GHz (0.75-1.1 cm)
- Infrared: 0.7 to 300 µm wavelength. This region is further divided into the following bands:
  - Near Infrared (NIR): 0.7 to 1.5 µm.
  - Short Wavelength Infrared (SWIR): 1.5 to 3 µm
  - Mid Wavelength Infrared (MWIR): 3 to 8 µm
  - Long Wanelength Infrared (LWIR): 8 to 15 µm

  - Far Infrared (FIR): longer than 15 µm

The NIR and SWIR are also known as the Reflected Infrared, referring to the main infrared component of the solar radiation reflected from the earth's surface. The MWIR and LWIR are the Thermal Infrared.

- Visible Light: This narrow band of electromagnetic radiation extends from about 400 nm (violet) to about 700 nm (red). The various colour components of the visible spectrum fall roughly within the following wavelength regions:
  - Red: 610-700 nm
  - Orange: 590-610 nm
  - Yellow: 570-590 nm
  - Green: 500-570 nm
  - Blue: 450-500 nm
  - Indigo: 430-450 nm
  - Violet: 400-430 nm
- Ultraviolet: 3 to 400 nm
- X-Rays and Gamma Rays

Photons

According to quantum physics, the energy of an electromagnetic wave is quantized, i.e. it can only exist in discrete amount. The basic unit of energy for an electromagnetic wave is called a photon. The energy *E* of a photon is proportional to the wave frequency *f*,

$E = h f,$

where the constant of proportionality h is the Planck's Constant,

$h = 6.626 \times 10^{-34}$ J s.

EFFECTS OF ATMOSPHERE

When electromagnetic radiation travels through the atmosphere, it may be absorbed or scattered by the constituent particles of the atmosphere. Molecular absorption converts the radiation energy into excitation energy of the molecules. Scattering redistributes the energy of the incident beam to all directions. The overall effect is the removal of energy from the incident radiation. The various effects of absorption and scattering are outlined in the following sections.

Atmospheric Transmission Windows

Each type of molecule has its own set of absorption bands in various parts of the electromagnetic spectrum. As a result, only the wavelength regions outside the main absorption bands of the atmospheric gases can be used for remote sensing. These regions are known as the Atmospheric Transmission Windows.

The wavelength bands used in remote sensing systems are usually designed to fall within these windows to minimize the atmospheric absorption effects. These windows are found in the visible, near-infrared, certain bands in thermal infrared and the microwave regions.

Effects of Atmospheric Absorption

Atmospheric absorption affects mainly the visible and infrared bands. Optical remote sensing depends on solar radiation as the source of illumination. Absorption reduces the solar radiance within the absorption bands of the atmospheric gases. The reflected radiance is also attenuated after passing through the atmosphere. This attenuation is wavelength dependent. Hence, atmospheric absorption will alter the apparent spectral signature of the target being observed.

Effects of Atmospheric Scattering

Atmospheric scatterring is important only in the visible and near infrared regions. Scattering of radiation by the constituent gases and aerosols in the atmosphere causes degradation of the remotely sensed images. Most noticeably, the solar radiation scattered by the atmosphere towards the sensor without first reaching the ground produces a hazy appearance of the image. This effect is particularly severe in the blue end of the visible spectrum due to the stronger Rayleigh Scattering for shorter wavelength radiation.

Furthermore, the light from a target outside the field of view of the sensor may be scattered into the field of view of the sensor. This effect is known as the adjacency effect. Near to the boundary between two regions of different brightness, the adjacency effect results in an increase in the apparent brightness of the darker region while the apparent brightness of the brighter region is reduced. Scattering also produces blurring of the targets in remotely sensed images due to spreading of the reflected radiation by scattering, resulting in a reduced resolution image.

Airborne Remote Sensing

In airborne remote sensing, downward or sideward looking sensors are mounted on an aircraft to obtain images of the earth's surface. An advantage of airborne remote sensing, compared to satellite remote sensing, is the capability of offering very high spatial resolution images (20 cm or less). The disadvantages are low coverage area and high cost per unit area of ground coverage. It is not cost-effective to map a large area using an airborne remote sensing system. Airborne remote sensing missions are often carried out as one-time operations, whereas earth observation satellites offer the possibility of continuous monitoring of the earth.

Analog aerial photography, videography, and digital photography are commonly used in airborne remote sensing. Synthetic Aperture Radar imaging is also carried out on airborne platforms.

Analog photography is capable of providing high spatial resolution. The interpretation of analog aerial photographs is usually done visually by experienced analysts. The photographs may be digitized using a scanning device for computer-assisted analysis.

Digital photography permits real-time transmission of the remotely sensed data to a ground station for immediate analysis. The digital images can be analysed and interpreted with the aid of a computer.

In spaceborne remote sensing, sensors are mounted on-board a spacecraft (space shuttle or satellite) orbiting the earth. At present, there are several remote sensing satellites providing imagery for research and operational applications. Spaceborne remote sensing provides the following advantages:

- Large area coverage;
- Frequent and repetitive coverage of an area of interest;
- Quantitative measurement of ground features using radiometrically calibrated sensors;
- Semiautomated computerised processing and analysis;
- Relatively lower cost per unit area of coverage.

Satellite imagery has a generally lower resolution compared to aerial photography. However, very high resolution imagery (up to 1-m resolution) is now commercially available to civilian users with the successful launch of the IKONOS-2 satellite in September 24, 1999.

Satellite Orbit

A satellite follows a generally elliptical orbit around the earth. The time taken to complete one revolution of the orbit is called the orbital period. The satellite traces out a path on the earth surface, called its ground track, as it moves across the sky. As the earth below is rotating, the satellite traces out a different path on the ground in each subsequent cycle. Remote sensing satellites are often launched into special orbits such that the satellite repeats its path after a fixed time interval. This time interval is called the repeat cycle of the satellite.

GEOSTATIONARY ORBIT

If a satellite follows an orbit parallel to the equator in the same direction as the earth's rotation and with the same period of 24 hours, the satellite will appear stationary with respect to the earth surface. This orbit is a geostationary orbit. Satellites in the geostationary orbits are located at a high altitude of 36,000 km. These orbits enable a satellite to always view the same area on the earth. A large area of the earth can also be covered by the satellite. The geostationary orbits are commonly used by meteorological satellites.

Near Polar Orbit

A near polar orbit is one with the orbital plane inclined at a small angle with respect to the earth's rotation axis. A satellite following a properly designed near polar orbit passes close to the poles and is able to cover nearly the whole earth surface in a repeat cycle.

SUN SYNCHRONOUS ORBIT

Earth observation satellites usually follow the sun synchronousorbits. A sun synchronous orbit is a nearpolar orbit whose altitude is such that the satellite will alwayspass over a location at a given latitude at the same local solartime. In this way, the same solarillumination condition (except for seasonal variation) can be achieved for the images of a given location taken by the satellite.

REMOTE SENSING SATELLITES

Several remote sensing satellites are currently available, providing imagery suitable for various types of applications.

Each of these satellite-sensor platform is characterised by the wavelength bands employed in image acquisition, spatial resolution of the sensor, the coverage area and the temporal coverge, i.e. how frequent a given location on the earth surface can be imaged by the imaging system.

In terms of the spatial resolution, the satellite imaging systems can be classified into:

- Low resolution systems (approx. 1 km or more)
- Medium resolution systems (approx. 100 m to 1 km)
- High resolution systems (approx. 5 m to 100 m)
- Very high resolution systems (approx. 5 m or less)

In terms of the spectral regions used in data acquisition, the satellite imaging systems can be classified into:

- Optical imaging systems (include visible, near infrared, and shortwave infrared systems)
- Thermal imaging systems
- Synthetic aperture radar (SAR) imaging systems

Optical/thermal imaging systems can be classified according to the number of spectral bands used:

- Monospectral or panchromatic (single wavelength band, “black-and-white”, grey-scale image) systems

- Multispectral (several spectral bands) systems
- Superspectral (tens of spectral bands) systems
- Hyperspectral (hundreds of spectral bands) systems

Synthetic aperture radar imaging systems can be classified according to the combination of frequency bands and polarization modes used in data acquisition, e.g.:

- Single frequency (L-band, or C-band, or X-band)
- Multiple frequency (Combination of two or more frequency bands)
- Single polarization (VV, or HH, or HV)
- Multiple polarization (Combination of two or more polarization modes)

ANALOG AND DIGITAL IMAGES

An image is a two-dimensional representation of objects in a real scene. Remote sensing images are representations of parts of the earth surface as seen from space. The images may be analog or digital. Aerial photographs are examples of analog images while satellite images acquired using electronic sensors are examples of digital images.

Pixels

A digital image comprises of a two dimensional array of individual picture elements called pixels arranged in columns and rows. Each pixel represents an area on the Earth's surface. A pixel has an intensity value and a location address in the two dimensional image.

The intensity value represents the measured physical quantity such as the solar radiance in a given wavelength band reflected from the ground, emitted infrared radiation or backscattered radar intensity.

This value is normally the average value for the whole ground area covered by the pixel. The intensity of a pixel is digitised and recorded as a digital number. Due to the finite storage capacity, a digital number is stored with a finite number of bits (binary digits). The number of bits determine the radiometric resolution of the image. For example, an 8-bit digital number ranges from 0 to 255 (i.e. 2^8-1), while a 11-bit digital number ranges from 0 to 2047. The detected intensity value needs to be scaled and quantized to fit within this range of value.

In a Radiometrically Calibrated image, the actual intensity value can be derived from the pixel digital number. The address of a pixel is denoted by its row and column coordinates in the two-dimensional image. There is a one-to-one correspondence between the column-row address of a pixel and the geographical coordinates (e.g. Longitude, latitude) of the imaged location. In order to be useful, the exact geographical location of each pixel on the ground must be derivable from its row and column indices, given the imaging geometry and the satellite orbit parameters.

Multilayer Image

Several types of measurement may be made from the ground area covered by a single pixel. Each type of measurement forms an image which carry some specific information about the area. By “stacking” these images from the same area together, a multilayer image is formed. Each component image is a layer in the multilayer image.

Multilayer images can also be formed by combining images obtained from different sensors, and other subsidiary data. For example, a multilayer image may consist of three layers from a SPOT multispectral image, a layer of ERS synthetic aperture radar image, and perhaps a layer consisting of the digital elevation map of the area being studied.

MULTISPECTRAL IMAGE

A multispectral image consists of a few image layers, each layer represents an image acquired at a particular wavelength band. For example, the SPOT HRV sensor operating in the multispectral mode detects radiations in three wavelength bands: the green (500-590 nm), red (610-680 nm) and near infrared (790-890 nm) bands.

A single SPOT multispectral scene consists of three intensity images in the three wavelength bands. In this case, each pixel of the scene has three intensity values corresponding to the three bands. A multispectral IKONOS image consists of four bands: Blue, Green, Red and Near Infrared, while a landsat TM multispectral image consists of seven bands: blue, green, red, near-IR bands, two SWIR bands, and a thermal IR band.

Superspectral Image

The more recent satellite sensors are capable of acquiring images at many more wavelength bands. For example, the MODIS sensor on-board the NASA's TERRA satellite consists of 36 spectral bands, covering the wavelength regions ranging from the visible, near infrared, short-wave infrared to the thermal infrared. The bands have narrower bandwidths, enabling the finer spectral characteristics of the targets to be captured by the sensor. The term "superspectral" has been coined to describe such sensors.

Hyperspectral Image

A hyperspectral image consists of about a hundred or more contiguous spectral bands. The characteristic spectrum of the target pixel is acquired in a hyperspectral image.

The precise spectral information contained in a hyperspectral image enables better characterisation and identification of targets. Hyperspectral images have potential applications in such fields as precision agriculture (e.g. monitoring the types, health, moisture status and maturity of crops), coastal management (e.g. monitoring of phytoplanktons, pollution, bathymetry changes). Currently, hyperspectral imagery is not commercially available from satellites. There are experimental satellite-sensors that acquire hyperspectral imagery for scientific investigation (e.g. NASA's Hyperion sensor on-board the EO1 satellite, CHRIS sensor onboard ESA's PRABO satellite).

Spatial Resolution

Spatial resolution refers to the size of the smallest object that can be resolved on the ground. In a digital image, the resolution is limited by the pixel size, i.e. the smallest resolvable object cannot be smaller than the pixel size. The intrinsic resolution of an imaging system is determined primarily by the instantaneous field of view (IFOV) of the sensor, which is a measure of the ground area viewed by a single detector element in a given instant in time. However this intrinsic resolution can often be degraded by other factors which introduce blurring of the image, such as improper focusing, atmospheric scattering and target motion. The pixel size is determined by the sampling distance.

Spatial Resolution and Pixel Size

The image resolution and pixel size are often used interchangeably. In realiaty, they are not equivalent. An image sampled at a small pixel size does not necessarily has a high resolution. The following three images illustrate this point. The first image is a SPOT image of 10 m pixel size. It was derived by merging a SPOT panchromatic image of 10 m resolution with a SPOT multispectral image of 20 m resolution. The merging procedure "colours" the panchromtic image using the colours derived from the multispectral image. The effective resolution is thus determined by the resolution of the panchromatic image, which is 10 m. This image is further processed to degrade the resolution while maintaining the same pixel size. The next two images are the blurred versions of the image with larger resolution size, but still digitized at the same pixel size of 10 m. Even though they have the same pixel size as the first image, they do not have the same resolution.

Radiometric Resolution

Radiometric Resolution refers to the smallest change in intensity level that can be detected by the sensing system. The intrinsic radiometric resolution of a sensing system depends on the signal to noise ratio of the detector. In a digital image, the radiometric resolution is limited by the number of discrete quantization levels used to digitize the continuous intensity value.

Data Volume

The volume of the digital data can potentially be large for multispectral data, as a given area is covered in many different wavelength bands. For example, a 3-band multispectral SPOT image covers an area of about 60 x 60 km^2 on the ground with a pixel separation of 20 m. So there are about 3000 x 3000 pixels per image.

Each pixel intensity in each band is coded using an 8-bit (i.e. 1 byte) digital number, giving a total of about 27 million bytes per image. In comparison, the panchromatic data has only one band. Thus, panchromatic systems are normally designed to give a higher spatial resolution than the multispectral system. For example, a SPOT panchromatic scene has the same coverage of about 60 x 60 km^2 but the pixel size is 10 m, giving about 6000 x 6000 pixels and a total of about 36 million bytes per image. If a multispectral SPOT scene is digitized also at 10 m pixel size, the data volume will be 108 million bytes.

For very high spatial resolution imagery, such as the one acquired by the IKONOS satellite, the data volume is even more significant. For example, an IKONOS 4-band multispectral image at 4-m pixel size covering an area of 10 km by 10 km, digitized at 11 bits (stored at 16 bits), has a data volume of 4 x 2500 x 2500 x 2 bytes, or 50 million bytes per image. A 1-m resolution panchromatic image covering the same area would have a data volume of 200 million bytes per image.

The images taken by a remote sensing satellite is transmitted to Earth through telecommunication. The bandwidth of the telecommunication channel sets a limit to the data volume for a scene taken by the imaging system. Ideally, it is desirable to have a high spatial resolution image with many spectral bands covering a wide area. In reality, depending on the intended application, spatial resolution may have to be compromised to accommodate a larger number of spectral bands, or a wide area coverage. A small number of spectral bands or a smaller area of coverage may be accepted to allow high spatial resolution imaging.

Optical remote sensing makes use of visible, near infrared and short-waveinfrared sensors to form images of the earth’s surface by detecting thesolar radiation reflected from targets on the ground. Different materials reflect and absorb differently at different wavelengths. Thus, the targets can be differentiated by their spectral reflectance signatures in the remotely sensed images. Optical remote sensing systems are classified into the following types, depending on the number of spectral bands used in the imaging process.

- Panchromatic imaging system: The sensor is a single channel detector sensitive to radiation within a broad wavelength range. If the wavelength range coincide with the visible range, then the resulting image resembles a “black-and-white” photograph taken from space. The physical quantity being measured is the apparent brightness of the targets. The spectral information or “colour” of the targets is lost. Examples of panchromatic imaging systems are:
  - IKONOS PAN
  - SPOT HRV-PAN
- Multispectral imaging system: The sensor is a multichannel detector with a few spectral bands. Each channel is sensitive to radiation within a narrow wavelength band. The resulting image is a multilayer image which contains both the brightness and spectral (colour) information of the targets being observed. Examples of multispectral systems are:
  - LANDSAT MSS
  - LANDSAT TM
  - SPOT HRV-XS
  - IKONOS MS
- Superspectral Imaging Systems: A superspectral imaging sensor has many more spectral channels (typically >10) than a multispectral sensor. The bands have narrower bandwidths, enabling the finer spectral characteristics of the targets to be captured by the sensor. Examples of superspectral systems are:
  - MODIS
  - MERIS
- Hyperspectral Imaging Systems: A hyperspectral imaging system is also known as an “imaging

spectrometer". it acquires images in about a hundred or more contiguous spectral bands. The precise spectral information contained in a hyperspectral image enables better characterisation and identification of targets. Hyperspectral images have potential applications in such fields as precision agriculture (e.g. monitoring the types, health, moisture status and maturity of crops), coastal management (e.g. monitoring of phytoplanktons, pollution, bathymetry changes). An example of a hyperspectral system is:

- o Hyperion on EO1 satellite.

Solar Irradiation

Optical remote sensing depends on the sun as the sole source of illumination.

The solar irradiation spectrum above the atmosphere can be modeled by a black body radiation spectrum having a source temperature of 5900 K, with a peak irradiation located at about 500 nm wavelength. Physical measurement of the solar irradiance has also been performed using ground based and spaceborne sensors.

After passing through the atmosphere, the solar irradiation spectrum at the ground is modulated by the atmospheric transmission windows. Significant energy remains only within the wavelength range from about 0.25 to 3 µm.

SOLAR IRRADIATION SPECTRA

When solar radiation hits a target surface, it may be transmitted, absorbed or reflected. Different materials reflect and absorb differently at different wavelengths. The reflectance spectrum of a material is a plot of the fraction of radiation reflected as a function of the incident wavelength and serves as a unique signature for the material.

In principle, a material can be identified from its spectral reflectance signature if the sensing system has sufficient spectral resolution to distinguish its spectrum from those of other materials. This premise provides the basis for multispectral remote sensing.

The reflectance of clear water is generally low. However, the reflectance is maximum at the blue end of the spectrum and decreases as wavelength increases. Hence, clear water appears dark-bluish. Turbid water has some sediment suspension which increases the reflectance in the red end of the spectrum, accounting for its brownish appearance. The reflectance of bare soil generally depends on its composition. In the example shown, the reflectance increases monotonically with increasing wavelength. Hence, it should appear yellowish-red to the eye.

Vegetation has a unique spectral signature which enables it to be distinguished readily from other types of land cover in an optical/near-infrared image.

The reflectance is low in both the blue and red regions of the spectrum, due to absorption by chlorophyll for photosynthesis.

It has a peak at the green region which gives rise to the green colour of vegetation. In the near infrared (NIR) region, the reflectance is much higher than that in the visible band due to the cellular structure in the leaves. Hence, vegetation can be identified by the high NIR but generally low visible reflectances.

This property has been used in early reconnaisance missions during war times for “camouflage detection”. The shape of the reflectance spectrum can be used for identification of vegetation type. They exhibit the generally characteristics of high NIR but low visible reflectances. Vegetation 1 has higher reflectance in the visible region but lower reflectance in the NIR region. For the same vegetation type, the reflectance spectrum also depends on other factors such as the leaf moisture content and health of the plants.

The reflectance of vegetation in the SWIR region (e.g. band 5 of Landsat TM and band 4 of SPOT 4 sensors) is more varied, depending on the types of plants and the plant’s water content. Water has strong absorption bands around 1.45, 1.95 and 2.50 µm. Outside these absorption bands in the SWIR region, reflectance of leaves generally increases when leaf liquid water content decreases.

This property can be used for identifying tree types and plant conditions from remote sensing images. The SWIR band can be used in detecting plant drought stress and delineating burnt areas and fire-affected vegetation.

The SWIR band is also sensitive to the thermal radiation emitted by intense fires, and hence can be used to detect active fires, especially during night-time when the background interference from SWIR in reflected sunlight is absent.

REMOTE SENSING

Remote Sensing is the science and art of acquiring information (spectral, spatial, temporal) about material objects, area, or phenomenon, without coming into physical contact with the objects, or area, or phenomenon under investigation. Without direct contact, some means of transferring information through space must be utilised.

In remote sensing, information transfer is accomplished by use of electromagnetic radiation (EMR). EMR is a form of energy that reveals its presence by the observable effects it produces when it strikes the matter. EMR is considered to span the spectrum of wavelengths from 10-10 mm to cosmic rays up to 1010 mm, the broadcast wavelengths, which extend from 0.30-15 mm.

Types

1. In respect to the type of Energy Resources: Passive Remote Sensing: Makes use of sensors that detect the reflected or emitted electro-magnetic radiation from natural sources.

   Active remote Sensing: Makes use of sensors that detect reflected responses from objects that are irradiated from artificially-generated energy sources, such as radar.

2. In respect to Wavelength Regions:

   Remote Sensing is classified into three types in respect to the wavelength regions

   - Visible and Reflective Infrared Remote Sensing.
   - Thermal Infrared Remote Sensing.
   - Microwave Remote Sensing.

BANDS USED IN REMOTE SENSING

Emission of EMR (Electo-Magnetic Radiation) from gases is due to atoms and molecules in the gas. Atoms consist of a positively charged nucleus surrounded by orbiting electrons, which have discrete energy states. Transition of electrons from one energy state to the other leads to emission of radiation at discrete wavelengths.

The resulting spectrum is called line spectrum. Molecules possess rotational and vibrational energy states. Transition between which leads to emission of radiation in a band spectrum. The wavelengths, which are emitted by atoms/molecules, are also the ones, which are absorbed by them. Emission from solids and liquids occurs when they are heated and results in a continuous spectrum. This is called thermal emission and it is an important source of EMR from the viewpoint of remote sensing.

The Electro-Magnetic Radiation (EMR), which is reflected or emitted from an object, is the usual source of Remote Sensing data. However, any medium, such as gravity or magnetic fields, can be used in remote sensing. Remote Sensing Technology makes use of the wide range Electro-Magnetic Spectrum (EMS) from a very short wave "Gamma Ray" to a very long 'Radio Wave'.

Wavelength regions of electro-magnetic radiation have different names ranging from Gamma ray, X-ray, Ultraviolet (UV), Visible light, Infrared (IR) to Radio Wave, in order from the shorter wavelengths. The optical wavelength region, an important region for remote sensing applications, is further subdivided as follows:

Name	***Wavelength (mm)***
Optical wavelength	0.30-15.0
Reflective	0.38-3.00
1. Portion Visible	0.38-0.72
2. Near IR	0.72-1.30
3. Middle IR	1.30-3.00
Far IR (Thermal, Emissive)	7.00-15.0

Microwave region (1mm to 1m) is another portion of EM spectrum that is frequently used to gather valuable remote sensing information.

Spectral Characteristics vis-à-vis different systems. The sunlight transmission through the atmosphere is effected by absorption and scattering of atmospheric molecules and aerosols. This reduction of the sunlight's intensity s called extinction.

The interaction and the interdependence between the primary sources of Electro-magnetic energy, the atmospheric windows through which source energy may be transmitted to and from the earth's surface features, and the spectral sensitivity of the sensors available to detect and record the energy. One cannot select the sensors to be used in any given remote-sensing task arbitrarily; one must instead consider.

1. the available spectral sensitivity of the sensors,
2. the presence or absence of atmospheric windows in the spectral range(s) in which one wishes to sense,
3. the source, magnitude, and spectral composition of the energy availabe in these ranges.

Ultimately, however, the choice of spectral range of the sensor must be based on the manner in which the energy interacts with the features under investigation.

ENERGY INTERACTIONS, SPECTRAL REFLECTANCE AND COLOUR READABILITY

All matter is composed of atoms and molecules with particular compositions. Therefore, matter will emit or absorb electro-magnetic radiation on a particular wavelength with respect to the inner state.

All matter reflects, absorbs, penetrates and emits Electro-magnetic radiation in a unique way. Electro-magnetic radiation through the atmosphere to and from matters on the earth's surface are reflected, scattered, diffracted, refracted, absorbed, transmitted and dispersed. For example, the reason why a leaf looks green is that the chlorophyll absorbs blue and red spectra and reflects the green. The unique characteristics of matter are called spectral characteristics.

Spectral Reflectance & Colour Readability

Two points about the above given relationship (expressed in the form of equation) should be noted.

1. The proportions of energy reflected, absorbed, and transmitted will vary for different earth features, depending upon their material type and conditions. These differences permit us to distinguish different features on an image.
2. The wavelength dependency means that, even within a given feature type, the proportion of reflected, absorbed, and transmitted energy will vary at different wavelengths.

Thus, two features may be distinguishable in one spectral range and be very different on another wavelength brand. Within the visible portion of the spectrum, these spectral variations result in the visual effect called COLOUR. For example we call blue objects 'blue' when they reflect highly in the 'green' spectral region, and so on. Thus the eye uses spectral variations in the magnitude of reflected energy to discriminate between various objects. A graph of the spectral reflectance of an object as a function of wavelength is called a spectral reflectance curve.

The configuration of spectral reflectance curves provides insight characteristics of an object and has a strong influence on the choice of wavelength region(s) in which remote sensing data are acquired for a particular application. The highly generalized spectral reflectance curves of deciduous and coniferous trees.

It should be noted that the curve for each of these object types is plotted as a 'ribbon' (or 'envelope') of values, not as a single line. This is because spectral reflectances vary somewhat within a given material class. That is, the spectral reflectance of one deciduous tree species and another will never be identical. Nor will the spectral reflectance of trees of the same species ever be exactly equal.

Types of earth feature:

- Green vegetation

- Soil
- Water

The average reflectance curves compiled by measuring large sample features. It should be noted how distinctive the curves are for each feature. In general, the configuration of these curves is an indicator of the type and condition of the features to which they apply. Although the reflectance of individual features will vary considerably above and below the average, these curves demonstrate some fundamental points concerning spectral reflectance.

MAJOR COMPONENTS OF REMOTE SENSING TECHNOLOGY

The following are major components of Remote sensing System:

1. Energy Source.
2. Passive System: sun, irradiance from earth's materials.
3. Active System: irradiance from artificially generated energy sources such as radar.
4. Platforms: (Vehicle to carry the sensor) (truck, aircraft, space shuttle, satellite, etc).
5. Sensors: (Device to detect electro-magnetic radiation) (camera, scanner, etc).
6. Detectors: (Handling signal data) (photographic, digital, etc).
7. Processing: (Handling Signal data) (photographic, digital etc).
8. Institutionalisation: (Organisation for execution at all stages of remote-sensing technology: international and national orrganisations, centres, universities, etc).

Platforms

The vehicles or carriers for remote sensors are called the platforms. Typical platforms are satellites and aircraft, but they can also include radio-controlled aeroplanes, balloons kits for low altitude remote sensing, as well as ladder trucks or 'cherry pickers' for ground investigations. The key factor for the selection of a platform is the altitude that determines the ground resolution and which is also dependent on the instantaneous field of view (IFOV) of the sensor on board the platform.

Resolution

In general resolution is defined as the ability of an entire remote-sensing system, including lens antennae, display, exposure, processing, and other factors, to render a sharply defined image. Resolution of a remote-sensing is of different types.

1. Spectral Resolution: of a remote sensing instrument (sensor) is determined by the band-widths of the Electro-magnetic radiation of the channels used. High spectral resolution, thus, is achieved by narrow bandwidths width, collectively, are likely to provide a more accurate spectral signature for discrete objects than broad bandwidth.
2. Radiometric Resolution: is determined by the number of discrete levels into which signals may be divided.
3. Spatial Resolution: in terms of the geometric properties of the imaging system, is usually described as the instantaneous field of view (IFOV). The IFOV is defined as the maximum angle of view in which a sensor can effectively detect electro-magnetic energy.
4. Temporal Resolution: is related to the repetitive coverage of the ground by the remote-sensing system. The temporal resolution of Landsat 4/5 is sixteen days.

AN IDEAL REMOTE SENSING SYSTEM

Having introduced some basic concepts, we now have the necessary elements to conceptualize an ideal remote sensing system. In doing so, we can then appreciate some of the problems encountered in the design and application of the various real remote-sensing systems examined in subsequent chapters.

The basic components of an ideal remote-sensing system. These include the following components.

- A uniform energy source. This source will provide energy over all wavelengths, at a constant, known, high level of output, irrespective of time and place.
- A non-interfering atmosphere. This will be an atmosphere that will not modify the energy from the source in any manner, whether that energy is on its way to earth's surface or coming from it. Again, ideally this will hold irrespective of wavelength, time, place, and sensing altitude involved.
- A series of unique energy/matter interaction at the earth's surface. These interactions will generate reflected and/or emitted signals that are not only selective in respect to wavelengths, but also are known, invariant, and unique to each and every earth surface feature type and subtype of interest.
- A super sensor. This will be a sensor, highly sensitive to all wavelengths, yielding spatially detailed data on the absolute brightness (or radiance) from a scene (a function of wavelength), throughout the spectrum. This super sensor will be simple and reliable, require, virtually no power or space, and be accurate and economical to operate.
- A real-time data handling system. In this system, the instant the radiance versus wavelength response over a terrain element is generated, it will be processed into an interpretable format and recognized as being unique to the particular terrain element from which it comes. This processing will be performed nearly instantaneously (real time), providing timely information. Because of the consistent nature of the energy/matter interactions, there will be no need for reference data in the analytical procedure. The derived data will provide insight into the physical-chemical-biological state of each feature of interest.
- Multiple data users. These people will have comprehensive knowledge of both their respective disciplines and of remote-sensing data acquisition and analysis techniques. The same set of data will become

various forms of information for different users, because of their vast knowledge about the particular earth resources being used.

REMOTE SENSING SATELLITES

A satellite with remote sensors to observe the earth is called a remote-sensing satellite, or earth observation satellite. Remote-Sensing Satellites are characterised by their altitude, orbit and sensor.

TRIOS Series (1960-1965)

The Television and Infrared Observation Satelites. NOAA It is the first generation of National Oceanic and Atmospheric Administration satellites and was as the first operation operational remote sensing satellite system.

The third generation NOAA satellites are also successfully used for vegetation monitoring, apart from meteorological monitoring.

It is equipped with Advanced Very High Resolution Radiometer (AVHRR) sensors, and is established at an altitude of 850 km. In polar orbit. GMS Geo-synchronous meteorological satellite. It is established at an altitude of 36,000 km, and its main purpose is meteorological observations.

Landsat is established at an altitude of 700 Kms is a polar orbit and is used mainly for land area observation. Other remote sensing satellite series in operations are: SPOT, MOS, JERS, ESR, RADARSAT, IRS etc.

TYPES OF REMOTE SENSING DATA

The types of remote sensing data vary but each plays a significant role in the ability to analyze an area from some distance away. The first way to gather remote sensing data is through radar.

Its most important uses are for air traffic control and the detection of storms or other potential disasters. In addition, Doppler radar is a common type of radar used in detecting meteorological data but is also used by law enforcement to monitor traffic and driving speeds. Other types of radar are also used to create digital models of elevation.

Another type of remote sensing data comes from lasers. These are often used in conjunction with radar altimeters on satellites to measure things like wind speeds and their direction and the direction of ocean currents.

These altimeters are also useful in seafloor mapping in that they are capable of measuring bulges of water caused by gravity and the varied seafloor topography. These varied ocean heights can then be measured and analysed to create seafloor maps.

Also common in remote sensing is LIDAR - Light Detection and Ranging. This is most famously used for weapons ranging but can also be used to measure chemicals in the atmosphere and heights of objects on the ground.

Other types of remote sensing data include stereographic pairs created from multiple air photos (often used to view features in 3-D and/or make topographic maps), radiometers and photometers which collect emitted radiation common in infra-red photos, and air photo data obtained by earth-viewing satellites such as those found in the Landsat program.

Applications of Remote Sensing

As with its varied types of data, the specific applications of remote sensing are diverse as well. However, remote sensing is mainly conducted for image processing and interpretation. Image processing allows things like air photos and satellite images to be manipulated so they fit various project uses and/or to create maps. By using image interpretation in remote sensing an area can be studied without being physically present there.

The processing and interpretation of remote sensing images also has specific uses within various fields of study. In geology, for instance, remote sensing can be applied to analyze and map large, remote areas. Remote sensing interpretation also makes it easy for geologists in this case to identify an area's rock types, geomorphology, and changes from natural events such as a flood or landslide.

Remote sensing is also helpful in studying vegetation types. Interpretation of remote sensing images allows physical and biogeographers, ecologists, those studying agriculture, and foresters to easily detect what vegetation is present in certain areas, its growth potential, and sometimes what conditions are conducive to its being there.

Additionally, those studying urban and other land use applications are also concerned with remote sensing because it allows them to easily pick out which land uses are present in an area. This can then be used as data in city planning applications and the study of species habitat, for example.

Finally, remote sensing plays a significant role in GIS. Its images are used as the input data for the raster-based digital elevation models (abbreviated as DEMs) - a common type of data used in GIS. The air photos taken during remote sensing applications are also used during GIS digitizing to create polygons, which are later put into shapefiles to create maps.

Because of its varied applications and ability to allow users to collect, interpret, and manipulate data over large often not easily accessible and sometimes dangerous areas, remote sensing has become a useful tool for all geographers, regardless of their concentration.

THE REGIONAL CENTRE FOR TOPOGRAPHIC SCIENCE, CARTOGRAPHY, REMOTE SENSING AND GIS

Resources. A laboratory with state of the art equipment and software + 1 mobile lab.

Hardware is made of topographic instruments (1 non mobile GPS, 1 mobile station and 1 total station both having a precision of 5 mm; the mobile lab. has also 4 medium precision GPSs and one laptop for data gathering and procesing on the field), a LAN of 6 pentium III machines, 1 Internet map-server, 6 printers, 1 A0 plotter, 1 A0 scanner, 1 digital camera with accessories, 1 digitizing tablet (size A1).

Software is made of: Arc/Info 8, 2 licenses, ArcView LabKit 8.1 (including Spatial Analyst, Network Analyst, 3D Analyst) 6 licenses, Image Analyst 1 license, ER Mapper 2 licenses, OrthoImage 1 license, ERDAS 1 license, OrthoBase 1 license, Idrisi 1 licence including Carta Linx.

Research activity. The members of the MURC are involved mainly in GIS and Remote Sensing projects. MSc and PhD students also participate in many of them.

Services. We perform work for many of industries, including government agencies, oil and gas pipeline companies and utility providers.

Each and every job we do receives our utmost attention to assure accurate and quality work performed on time and at resonable costs.

GIS and Digital Cartography – Data Capture, Conversion & Mapping Services

- Data Conversion based on assessment process
- Conversion of existing hardcopy maps to digital form
- Client data delivered in desired coordinate system, formatted to be compatible with most of software programs
- Comprehensive GPS field mapping and facility inventory
- Geo-referencing of existing digital map data for compatibility with scanned aerialphotography
- Geo-coding digital maps for address matching
- DEM data creation

GIS and Digital Cartography – Data Maintenance

- Digital maps and data, custom hardcopy maps
- Work with client to design on-going plan to keep map data and database current and up-to-date
- Serve as client's personal mapping department, providing map and data maintenance services
- Modify databases to meet the client's changing needs

GIS and Digital Cartography – Consulting

- Database model design and development
- Data conversion plan design
- GIS implementation plan design
- Training plan for core users
- Enterprise – wide GIS implementation plan

GIS and Digital Cartography – Training

Classes offered include : Introduction to GIS, Introduction to ArcView®, Advanced ArcView®, Introduction to Arc/Info®. Classes may be tailored on request. Classes offer comprehensive, hands-on agenda teaching functional use of the software, as well as an overview of GIS technology.

Remote Sensing – Training

Classes offered include: Introduction to Remote Sensing, Introduction to ER Mapper®, Advanced ER Mapper®.

Environmental Modeling and Territorial Planning with GIS and Remote Sensing:

- Hydrological Modeling
- Environmental Hazard Assessment
- Environmental Impact Modeling
- Environmental Monitoring
- Regional Planning

3

GIS Data Types and Data Models

Although the two terms, data and information, are often used indiscriminately, they both have a specific meaning. Data can be described as different observations, which are collected and stored. Information is that data, which is useful in answering queries or solving a problem. Digitizing a large number of maps provides a large amount of data after hours of painstaking works, but the data can only render useful information if it is used in analysis.

GIS DATA TYPES

Attribute Data

The attributes refer to the properties of spatial entities. They are often referred to as non-spatial data since they do not in themselves represent location information. This type of data describes characteristics of the spatial features. These characteristics can be quantitative and/or qualitative in nature. Attribute data is often referred to as tabular data.

Spatial Data

Geographic position refers to the fact that each feature has a location that must be specified in a unique way. To specify the position in an absolute way a coordinate system is used. For small areas, the simplest coordinate system is the regular square grid. For larger areas, certain approved cartographic projections are commonly used. Internationally there are many different coordinate systems in use. This Locational information is provided in maps by using Points, Lines and Polygons. These geometric descriptions are the basic data elements of a map. Thus spatial data describes the absolute and relative location of geographic features.

The coordinate location of a forest would be spatial data, while the characteristics of that forest, e.g. cover group, dominant species, crown closure, height, etc., would be attribute data. Other data types, in particular image and multimedia data, have become more prevalent with changing technology. Depending on the specific content of the data, image data may be considered either spatial, e.g. photographs, animation, movies, etc., or attribute, e.g. sound, descriptions, narration's, etc.

GIS DATA MODELS

A GIS is based on data, hence there must be a data model that has to be followed to standardize procedures.

They are :

1. Spatial Data Models
2. Attribute Data Models

Spatial Data Models

Traditionally spatial data has been stored and presented in the form of a map. Three basic types of spatial data models have evolved for storing geographic data digitally. These are referred to as:

Raster Vector Image

The selection of a particular data model, vector or raster, is dependent on the source and type of data, as well as the intended use of the data. Certain analytical procedures require raster data while others are better suited to vector data.

Raster Data Formats

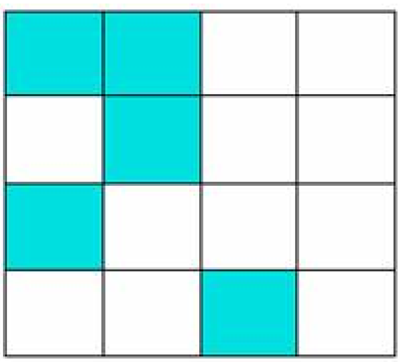
A simple raster data set is a regular grid of cells divided into rows and columns. In a raster data set, data values for a given parameter are stored in each cell – these values may represent an elevation in meters above sea level, a land use class, a plant biomass in grams per square meter, and so forth. The spatial resolution of the raster data set is determined by the size of the cell. For example, Landsat TM satellite imagery data are raster data that are corrected to have a cell size of approximately 30 meters on a side. However, spatial resolution can be much finer, or much coarser than 30 meters. In general, spatial resolution is a function of the data collection techniques used, and the desired outcomes.

The size of cells in a tessellated data structure is selected on the basis of the data accuracy and the resolution needed by the user. There is no explicit coding of geographic coordinates required since that is implicit in the layout of the cell A raster data structure is in fact a matrix where any coordinate can be quickly calculated if the origin point is known, and the size of the grid cells is known. Since grid-cells can be handled as two-dimensional arrays in computer encoding many analytical operations are easy to program. This makes tessellated data structures a popular choice for many GIS software. Topology is not a relevant concept with tessellated structures since adjacency and connectivity are implicit in the location of a particular cell in the data matrix.

Since geographic data is rarely distinguished by regularly spaced shapes, cells must be classified as to the most common attribute for the cell. The problem of determining the proper resolution for a particular data layer can be a concern. If one selects too coarse a cell size then data may be overly generalized. If one selects too fine a cell size then too many cells may be created resulting in a large data volume, slower processing times, and a more cumbersome data set. As well, one can imply accuracy greater than that of the original data capture process and this may result in some erroneous results during analysis. As well, since most data is captured in a vector format, e.g. digitizing, data must be converted to the raster data structure. This is called vector-raster conversion. Most GIS software allows the user to define the raster grid (cell) size for vector-raster conversion. It is imperative that the original scale, e.g. accuracy, of the data be known prior to conversion. The accuracy of the data, often referred to as the resolution, should determine the cell size of the output raster map during conversion. Most raster based GIS software requires that the raster cell contain only a single discrete value. Accordingly, a data layer, e.g. forest inventory stands, may be broken down into a series of raster maps, each representing an attribute type, e.g. a species map, a height map, a density map, etc. These are often referred to as one attribute maps. This is in contrast to most conventional vector data models that maintain data as multiple attribute maps.

0	0	1	1
1	0	1	1
0	1	1	1
1	1	0	1

A Simple Raster Data Set: Each cell in the raster is assigned a single data value. In the above example simple binary data values have been used meaning that the possibilities are limited to two digit numbers – either 0 or 1. This is an example of a 1-bit raster data file. Mathematically, there are only two possibilities for each pixel, 0 or 1. By contrast in an 8-bit data file, there are 256 possibilities of data values for each pixel. In the above example, the computer "sees" the cells that contain 0 as "turned off", while the cells that contain 1 as "turned on".



The horizontal dimension of raster data is often oriented parallel to the east-west direction. Following image processing convention, raster cells are numbered beginning on the left margin of the raster.

Further, the positions of cells in the vertical dimension are numbered starting from the top – or northern boundary. Thus, the origin of the raster is in the upper left corner.

This location is most often referenced (1,1). It is important to note that this referencing system is different from more traditional geo-referencing systems that are based on Cartesian geometry where the origin is in the lower left corner, and the origin is typically referenced as (0,0).

1,1	1,2	1,3	1,4
2,1	2,2	2,3	2,4
3,1	3,2	3,3	3,4
4,1	4,2	4,3	4,4

0,3	1,3	2,3	3,3
0,2	1,2	2,2	3,2
0,1	1,1	2,1	3,1
0,0	1,0	2,0	3,0

The selection of a particular data structure can provide advantages during the analysis stage.

For example, the vector data model does not handle continuous data, e.g. elevation, very well while the raster data model is more ideally suited for this type of analysis.

Accordingly, the raster structure does not handle linear data analysis, e.g. shortest path, very well while vector systems do. It is important for the user to understand that there are certain advantages and disadvantages to each data model.

Advantages of Raster Data

1. The geographic location of each cell is implied by its position in the cell matrix. Accordingly, other than an origin point, e.g. bottom left corner, no geographic coordinates are stored.
2. Due to the nature of the data storage technique data analysis is usually easy to program and quick to perform.

3. The inherent nature of raster maps, e.g. one attribute maps, is ideally suited for mathematical modelling and quantitative analysis.
4. Discrete data, e.g. forestry stands, is accommodated equally well as continuous data, e.g. elevation data, and facilitates the integrating of the two data types.

Disadvantages of Raster Data

1. The cell size determines the resolution at which the data is represented.
2. It is especially difficult to adequately represent linear features depending on the cell resolution. Accordingly, network linkages are difficult to establish.
3. Processing of associated attribute data may be cumbersome if large amounts of data exists. Raster maps inherently reflect only one attribute or characteristic for an area.
4. Since most input data is in vector form, data must undergo vector-to-raster conversion. Besides increased processing requirements this may introduce data integrity concerns due to generalization and choice of inappropriate cell size.

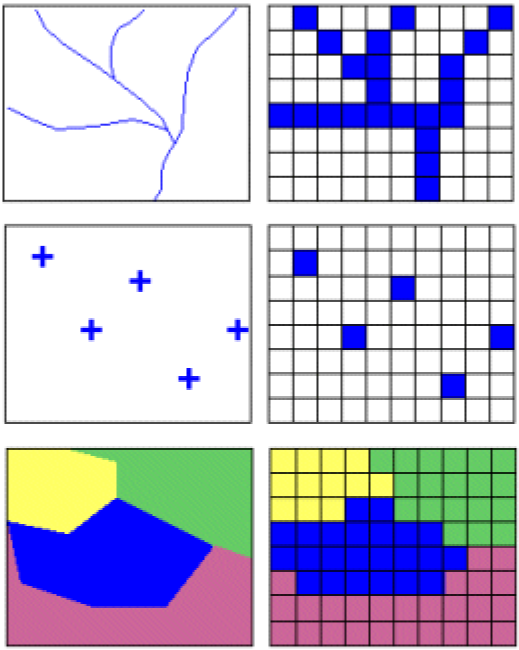
Vector Data Models

The vector data model is based upon vectors as opposed to space occupancy of raster data structures. The fundamental primitiveof the vector model is a point.

The various objects are created by connecting the points with straight lines, but some systems allow the points to be connected using arcs of circles.

The areas are defined in this model by sets of lines. The term polygon is synonymous with area in vector databases because of the se of straight-line connections between points.

Very large vector databases have been built for facilitating different purposes as vectors dominate in various different fields such as transportation, utility and marketing applications.



Vector Raster

Several different vector data models exist, however only two are commonly used in GIS data storage. The topologic data structure is often referred to as an *intelligent data structure* because spatial relationships between geographic features are easily derived when using them. Primarily for this reason the topologic model is the dominant vector data structure currently used in GIS technology. Many of the complex data analysis functions cannot effectively be undertaken without a topologic vector data structure.

The secondary vector data structure that is common among GIS software is the computer-aided drafting (CAD) data structure. This structure consists of listing elements, not features, defined by strings of vertices, to define geographic features, e.g. points, lines, or areas. There is considerable redundancy with this data model since the boundary segment between two polygons can be stored twice, once for each feature. The CAD structure emerged from the development of computer graphics systems without specific considerations of processing geographic features. Accordingly, since features, e.g. polygons, are self-contained and independent, questions about the adjacency of features can be difficult to answer. The CAD vector model lacks the definition of spatial relationships between features that is defined by the topologic data model.

Advantages of Vector Data

1. Data can be represented at its original resolution without generalization.
2. Graphic output is usually more aesthetically pleasing.
3. Since most data, e.g. hard copy maps are in vector form, no conversion is required.
4. Accurate geographic location of data is maintained.
5. Allows for efficient encoding of topology, and as a result more efficient operations that require topological information, e.g. proximity, network analysis.

Disadvantages of Vector Data

1. The location of each vertex needs to be stored explicitly.
2. Algorithms for manipulative and analysis functions are complex and may be processing intensive.
3. Continuous data, such as elevation data, is not effectively represented in vector form.
4. Spatial analysis and filtering within polygons is impossible.

Image Data Format

Image data is most often used to represent graphic or pictorial data. The term *image* inherently reflects a graphic representation, and in the *GIS world*, differs significantly from raster data. Most often, image data is used to store remotely sensed imagery, e.g. satellite scenes or orthophotos, or ancillary graphics such as photographs, scanned plan documents, etc. Image data is typically used in GIS systems as background display data (if the image has been rectified and georeferenced); or as a graphic attribute. Remote sensing software makes use of image data for image classification and processing. Typically, this data must be converted into a raster format (and perhaps vector) to be used analytically with the GIS.

ATTRIBUTE DATA MODELS (DBMS MODELS USED IN GIS)

A separate data model is used to store and maintain attribute data for GIS software. These data models may exist internally within the GIS software, or may be reflected in external commercial Database Management Software (DBMS). A variety of different data models exist for the storage and management of attribute data. The most common are:

Tabular Model

The simple tabular model stores attribute data as sequential data files with fixed formats (or comma delimited for ASCII data), for the location of attribute values in a predefined record structure. This type of data model is outdated in the GIS arena. It lacks any method of checking data integrity, as well as being inefficient with respect to data storage, e.g. limited indexing capabilityfor attributes or records, etc.

Hierarchical Model

The hierarchical database organizes data in a *tree* structure. Data is structured downward in a *hierarchy* of tables. Any level in the hierarchy can have unlimited *children*, but any *child* can have only one *parent*. Hierarchical DBMS have not gained any noticeable acceptance for use within GIS. They are oriented for data sets that are very stable, where primary relationships among the data change infrequently or never at all. Also, the limitation on the number of parents that an element may have is not always conducive to actual geographic phenomenon.

Network Model

The network database organizes data in a network or *plex* structure. Any column in a plex structure can be linked to any other. Like a tree structure, a plex structure can be described in terms of *parents* and *children*. This model allows for children to have more than one parent.

Relational Model

The relational database organizes data in *tables*. Each table, is identified by a unique table name, and is organized by *rows* and *columns*. Each column within a table also has a unique name. Columns store the values for a specific attribute, e.g.

cover group, tree height. Rows represent one record in the table. In a GIS each row is usually linked to a separate spatial feature, e.g. a forestry stand. Accordingly, each row would be comprised of several columns, each column containing a specific value for that geographic feature.

Data is often stored in several tables. Tables can be joined or referenced to each other by common columns (relational fields). Usually the common column is an identification number for a selected geographic feature, e.g. a forestry stand polygon number. This identification number acts as the *primary key* for the table. The ability to join tables through use of a common column is the essence of the relational model. Such relational joins are usually ad hoc in nature and form the basis of for querying in a relational GIS product. Unlike the other previously discussed database types, relationships are implicit in the character of the data as opposed to explicit characteristics of the database set up. The relational database model is the most widely accepted for managing the attributes of geographic data.

The relational DBMS is attractive because of it's:

- Simplicity in organization and data modelling.
- Flexibility-data can be manipulated in an ad hoc manner by joining tables.
- Efficiency of storage-proper design of data tables can reduce redundancy.
- Queries do not need to take into account the internal organization of data.

The relational DBMS has emerged as the dominant commercial data management tool in GIS implementation and application.

Object Oriented Model

The object-oriented database model manages data through *objects*. An object is a collection of data elements and operations that together are considered a single entity.

The object-oriented database is a relatively new model. This approach has the attraction that querying is very natural, as features can be bundled together with attributes at the database administrator's discretion. To date, only a few GIS packages are promoting the use of this attribute data model.

However, initial impressions indicate that this approach may hold many operational benefits with respect to geographic data processing. Fulfilment of this promise with a commercial GIS product remains to be seen.

FUNDAMENTALS OF GIS

Mapping Concepts, Features & Properties

A map represents geographic features or other spatial phenomena by graphically conveying information about locations and attributes.

Locational information describes the position of particular geographic features on the Earth's surface, as well as the spatial relationship between features, such as the shortest path from a fire station to a library, the proximity of competing businesses, and so on. Attribute information describes characteristics of the geographic features represented, such as the feature type, its name or number and quantitative information such as its area or length.

Thus the basic objective of mapping is to provide:

- descriptions of geographic phenomenon
- spatial and non spatial information
- map features like Point, Line, & Polygon

Map Features

Locational information is usually represented by points for features such as wells and telephone pole locations, lines for features such as streams, pipelines and contour lines and areas for features such as lakes, counties and census tracts.

Point Feature

A point feature represents as single location. It defines a map object too small to show as a line or area feature. A special symbol of label usually depicts a point location.

Line Feature

A line feature is a set of connected, ordered coordinates representing the linear shape of a map object that may be too narrow to display as an area such as a road or feature with no width such as a contour line.

AREA FEATURE

Map Characteristics

In addition to feature locations and their attributes, the other technical characteristics that define maps and their use includes:

- Map Scale
- Map Accuracy
- Map Extent and
- Data Base Extent

Scale

To show a portion of the Earth's surface on a map, the scale must be sufficiently adjusted to cover the objective. Map scale or the extent of reduction is expressed as a ratio. The unit on the left indicates distance on the map and the number on the right indicates distance on the ground. The following three statements show the same scale.

1 inch = 2.000 feet => 1 inch = 24.000 inches => 1:24.000

The latter is known as a representative fraction (RF) because the amounts on either side of the colon are equivalent: that is 1:24.000 means 1 inch equals 24.000 inches or 1 foot equals 24.000 feet or 1 meter equals 24.000 meters and so on.

Map scale indicates how much the given area has been reduced. For the same size map, features on a small-scale map (1:1,000,0000) will be smaller than those on a large-scale map (1:1,200).

A map with less detail is said to be of a smaller scale than one with more detail. Cartographers often divide scales into three different categories.

Small-scale maps have scales smaller than 1 : 1,000,000 and are used for maps of wide areas where not much detail is required.

Medium-scale maps have scales between 1 : 75,000 and 1 : 1,000,000.

Large-scale maps have scales larger than 1 : 75,000. They are used in applications where detailed map features are required.

So each scale represents a different tradeoff. With a small-scale map, you'll be able to show a large area without much detail. On a large-scale map, you'll be able to show a lot of detail but not for a large area. The small-scale map can show a large area because it reduces the area so much that the large-scale map can only show a portion of one street, but in such detail that you can see shapes of the houses.

To convert this statement to a representative fraction, the units of measure on both the sides being compared must be the same. For this example, both measurements will be in meters.

To do this:

1. Convert 1.6 inches into meters.

   1.6 inches x 0.0254 meters/inch = 0.04 meters.

2. Let us suppose that.

   0.04 units on the map = 10,000 units on the ground.

Then, you can now state the scale as a representative fraction (RF): 0.04:10,000.

Though it is a valid statement of scale, most cartographers may find it clumsy. Traditionally, the first number in the representative fraction is made equal to 1:

0.04/0.04 = 1 units on the map = 10,000/0.04 units on the ground.

1 unit on the map = 250,000 units on the ground.

SCALE IN DIGITAL MAPS

With digital maps, the traditional concept of scale in terms of distance does not apply because digital maps do not remain fixed in size. They can be displayed or plotted at any possible magnification. Yet we still speak of the scale of a digital map. In digital mapping, the term scale is used to indicate the scale of the materials from which the map was made. For example, if a digital map is said to have a scale of 1:100,000, it was made from a 1:100,000-scale paper map.

However, a digital map's scale still allows you to make some educated guesses about its contents because, generally, digital maps retain the same accuracy and characteristics as their source maps. So it is still true that a large-scale digital map will usually be more accurate and less general than a small-scale digital map. Because the display size of a computer-based map is not fixed, users are often tempted to blow up maps to very large sizes. For example, a 1:100,000-scale map can easily be plotted at a size of 1:24,000 or even 1:2,000-but it usually is not a good idea to do so. It encourages the user to make measurements that the underlying data does not support. You cannot measure positions to the nearest foot if your map is only accurate to the nearest mile. You will end up looking for information that does not exist.

MAP RESOLUTION

Map resolution refers to how accurately the location and shape of map features can be depicted for a given map scale. Scale affects resolution. In a larger-scale map, the resolution of features more closely matches real-world features because the extent of reduction from ground to map is less. As map scale decrease, the map resolution diminishes because features must be smoothed and simplified, or not shown at all.

MAP ACCURACY

Many factors besides resolution, influence how accurately features can be depicted, including the quality of source data, the map scale, your drafting skill and the width of lines drawn on the ground. A fine drafting pen will draw line’s 1/100 of an inch wide. Such a line represents a corridor on the ground, which is almost 53 feet wide. In addition to this, human drafting errors will occur and can be compounded by the quality of your source maps and materials. A map accurate for one purpose is often inaccurate for others since accuracy is determined by the needs of the project as much as it is by the map itself.

Some measurements of a map’s accuracy are discussed below.

- Absolute accuracy of a map refers to the relationship between a geographic position on a map (a street corner, for instance) and its real-world position measured on the surface of the earth. Absolute accuracy is primarily important for complex data requirements such as those for surveying and engineering-based applications.
- Relative accuracy refers to the displacement between two points on a map (both distance and angle), compared

to the displacement of those same points in the real world. Relative accuracy is often more important and easier to obtain than absolute accuracy because users rarely need to know absolute positions. More often, they need to find a position relative to some known landmark, which is what relative accuracy provides. Users with simple data requirements generally need only relative accuracy.

- Attribute accuracy refers to the precision of the attribute database linked to the map's features. For example, if the map shows road classifications, are they correct? If it shows street addresses, how accurate are they? Attribute accuracy is most important to users with complex data requirements.
- A map's Currency refers to how up-to-date it is. Currency is usually expressed in terms of a revision date, but this information is not always easy to find.
- A map is Complete if it includes all the features a user would expect it to contain. For example, does a street map contain all the streets? Completeness and currency usually are related because a map becomes less complete as it gets older.

The most important issue to remember about map accuracy is that the more accurate the map, the more it costs in time and money to develop. For example, digital maps with coordinate accuracy of about 100 feet can be purchased inexpensively. If 1-foot accuracy is required, a custom survey is often the only way to get it, which drives up data-acquisition costs by many orders of magnitude and can significantly delay project implementation-by months or even years. Therefore, too much accuracy can be as detrimental to the success of a GIS project as too little. Rather than focusing on the project's benefits, a sponsoring organization may focus on the costs that result from a level of accuracy not justified for the project. Project support inevitably erodes when its original objectives are forgotten in a flurry of cost analyses.

A far better strategy is to start the project with whatever data is readily available and sufficient to support initial objectives. Once the GIS is up and running, producing useful results, project scope can be expanded. The quality of its data can be improved as required. Even though no maps are entirely accurate, they are still useful for decision-making and analysis. How ever, it is important to consider map accuracy to ensure that your data is not used inappropriately.

Any number of factors can cause error. Note these sources can have at cumulative effect.

E = f(f) + f(1) + f(e) + f(d) + f(a) + f(m) + f(rms) + f(mp) + u

Where,

f = flattening the round Earth onto a two-dimensional surface (transformation from spherical to planar geometry)

I = accurately measuring location on Earth (correct project and datum information).

c = cartographic interpretation (correct interpretation of features).

d = drafting error (accuracy in tracing of features and width of drafting pen).

a = analog to digital conversion (digitizing board calibration).

m = media stability (warping and stretching, folding. Wrinkling of map).

p = digitizing processor error (accuracy of cursor placement)

rms = Root Mean Square (registration accuracy of ties).

mp = machine precision (coordinate rounding by computer in storing and transforming).

u = additional unexplained source error.

MAP EXTENT

The aerial extent of map is the area on the Earth's surface represented on the map. It is the limit of the area covered, usually defined by rectangle just large enough to include all mapped features. The size of the study area depends on the map scale. The smaller the scale the larger the area covered.

Database Extent

A critical first step in building a geographic database is defining its extent. The aerial extent of a database is the limit of the area of interest for your GIS project. This usually includes the areas directly affected by your organization's responsibility (such as assigned administrative units) as well as surrounding areas that either influence or are influenced by relevant activities in the administrative area.

Data Automation

Map features are logically organized into a set of layers or themes of information. A base map can be organized into layers such as streams, soils, wells or boundaries. Map data, regardless of how a spatial database will be applied, is collected, automated and updated as series of adjacent map sheets or aerial photograph. Here each sheet is mounted on the digitizer and digitized, one sheet at a time.

In order to be able to combine these smaller sheets into larger units or study areas, the co-ordinates of coverage must be transformed into a single common co-ordinate system. Once in a common co-ordinate system, attributes are associated with features. Then as needed map sheets for layer are edge matched and joined into a single coverage for your study area.

TYPES OF INFORMATION IN A DIGITAL MAP

Any digital map is capable of storing much more information than a paper map of the same area, but it's generally not clear at first glance just what sort of information the map includes. For example, more information is usually available in a digital map than what you see on-screen. And evaluating a given data set simply by looking at the screen can be difficult: What part of the image is contained in the data and what part is created by the GIS program's interpretation of the data? You must understand the types of data in your map so you can use it appropriately.

Three general types of information can be included in digital maps:

- Geographic information, which provides the position and shapes of specific geographic features.
- Attribute information, which provides additional non-graphic information about each feature.
- Display information, which describes how the features will appear on the screen.

Some digital maps do not contain all three types of information. For example, raster maps usually do not include attribute information, and many vector data sources do not include display information.

Geographic Information

The geographic information in a digital map provides the position and shape of each map feature. For example, a road map's geographic information is the location of each road on the map. In a vector map, a feature's position is normally expressed as sets of X, Y pairs or X, Y, Z triples, using the coordinate system defined for the map. Most vector geographic information systems support three fundamental geometric objects:

- *Point:* A single pair of coordinates.
- *Line:* Two or more points in a specific sequence.
- *Polygon:* An area enclosed by a line.

Some systems also support more complex entities, such as regions, circles, ellipses, arcs, and curves.

Attribute Information

Attribute data describes specific map features but is not inherently graphic. For example, an attribute associated with a road might be its name or the date it was last paved. Attributes are often stored in database files kept separately from the graphic portion of the map. Attributes pertain only to vector maps; they are seldom associated with raster images. GIS software packages maintain internal links tying each graphical map entity to its attribute information. The nature of these links varies widely across systems. In some, the link is implicit, and the user has no control over it. Other systems have explicit links that the user can modify. Links in these systems take the form of database keys. Each map feature has a key value stored with it; the key identifies the specific database record that contains the feature's attribute information.

Display Information

The display information in a digital-map data set describes how the map is to be displayed or plotted. Common display information includes feature colours, line widths and line types (solid, dashed, dotted, single, or double); how the names of roads and other features are shown on the map; and whether or not lakes, parks, or other area features are colour coded. However, many users do not consider the quality of display information when they evaluate a data set. Yet map display strongly affects the information you and your audience can obtain from the map-no matter how simple or complex the project. A technically flawless, but unattractive or hard-to-read map will not achieve the goal of conveying information easily to the user.

Cartographic Appeal

Clearly, how a map looks-especially if it is being used in a presentation-determines its effectiveness. Appropriate colour choices, linetypes, and so on add the professional look you want and make the map easier to interpret. Since display information often is not included in the source data set or is filtered out by conversion software, you may need to add it yourself or purchase the map from a vendor who does it for you. Map display information should convey the meaning of its underlying attribute data. Various enhancements will increase a map's usefulness and cartographic appeal.

- Feature Colours and Linetypes. Colours and line representations should be chosen to make the map's meaning clear. For example, using double-line roads can be quite helpful. Many GIS data sets only include road centerline information. Actual road width is not given. So maps with centerlines only can look like spider webs, which is visually unappealing. Some software and conversion systems can draw roads as double lines, with distance between lines varying according to road type. Centerlines can be included, if necessary. Double-line maps are appropriate for detailed studies of small areas, such as subdivisions, or maps where right-of-way information is important.
- Naming Roads. Naming, or labeling, roads are important for proper map interpretation. This information should be legible, positioned in the center of the road or offset from the center, and drawn at intervals suited to the scale of the final map or its purpose.

- Landmark Symbols. A good set of symbols should be used to indicate landmarks, such as hospitals, schools, churches, and cemeteries. The symbols should be sized appropriately in relation to map scale.
- Polygon Fills. Polygon features, such as lakes or parks, should be filled with an appropriate colour or hatch pattern.
- Zoom Layer Control. If the GIS software platform permits, map layers should be set up so that detailed, high-density information only appears when the user zooms in for a close-up of part of the map. For example, when a large area is displayed, only the major roads should appear; for a smaller area, both major and minor roads should appear.

Layering

Most GIS software has a system of layers, which can be used to divide a large map into manageable pieces. For example, all roads could be on one layer and all hydrographic features on another. Major layers can be further classified into sub-layers, such as different types of roads-highways, city streets, and so on. Layer names are particularly important in CAD-based mapping and GIS programs, which have excellent tools for handling them. Some digital maps are layered according to the numeric feature-classification codes found in their source data sets. For example, a major road might be on the 170-201 layer. However, this type of system is not very useful. A well-thought-out layering scheme can make any data set much easier to use because it allows the user to control the features with which you want to work. A good layering standard has layer names that are mnemonic (suggest their meanings) and hierarchical (have a structured classification scheme that makes it easy to choose general or specific classes).

For example, a map could have its roads on a layer called RD, its railroads on a layer called RR, its road bridges on a layer called RD-BRIDGE, and its railroad bridges on a layer called RR-BRIDGE. This scheme is mnemonic because it is easy to tell a layer's contents from its name, and it's hierarchical because the user can easily select all the roads, railroads, bridges, road bridges, or railroad bridges.

MAPS AND MAP ANALYSIS

Automated Mapping

Computer Aided Mapping has its limitations. Goal of GIS is not only to prepare a good map but also perform map analysis. Maps are the main source of data for GIS. GIS, though an accurate mapping tool, requires error management. MAP is a representation on a medium of a selected material or abstract material in relation to the surface of the earth (defined by Cartographic association). Maps originated from mathematics. The term Map is often used in mathematics to convey the motion of transferring the information from one form to another just as Cartographers transfer information from the surface of the earth to a sheet of paper. Map is used in a loose fashion to refer to any manual display of information particularly if it is abstract, generalised or schematic.

Process involved in the production of Maps:

- Selection of few features of the real world.
- Classification of selected features in to groups eg. Railway in to different lines. Classification depends upon the purpose.
- Simplification of jaggered lines like the coast lines.
- Exaggeration of features.
- Symbolisation to represent different classes of features.

Drawing Digitization of Maps

Maps can be broadly classified in to two groups:

1. Topographical maps.
2. Thematic maps.

Topographical Maps

It is a reference map showing the outline of selected man-made and natural features of the earth. It often acts as a frame for other features Topography refers to the shape of surface represented by contours or shading. It also shows lands, railway and other prominent features.

Thematic Maps

Thematic maps are an important source of GIS information. These are tools to communicate geographical concepts such as

Density of population, Climate, movement of goods and people, land use etc. It has many classifications.

DATA MODEL

A data model is used to organize data. A data model captures the cardinality and referential integrity rules needed to ensure that the data is of good quality for the users. A data model has 3 uses in an application which are getting data in, integrating data and getting data out. A data model is also used as a communication tool for teams to communicate within the team on how data is organized and between teams.

Overview

Managing large quantities of structured and unstructured data is a primary function of information systems. Data models describe structured data for storage in data management systems such as relational databases. They typically do not describe unstructured data, such as word processing documents, email messages, pictures, digital audio, and video.

The Role of Data Models

The main aim of data models is to support the development of information systems by providing the definition and format of data. According to West and Fowler (1999) "if this is done consistently across systems then compatibility of data can be achieved. If the same data structures are used to store and access data then different applications can share data. However, systems and interfaces often cost more than they should, to build, operate, and maintain. They may also constrain the business rather than support it. A major cause is that the quality of the data models implemented in systems and interfaces is poor".

- "Business rules, specific to how things are done in a particular place, are often fixed in the structure of a data model. This means that small changes in the way business is conducted lead to large changes in computer systems and interfaces".
- "Entity types are often not identified, or incorrectly identified. This can lead to replication of data, data structure, and functionality, together with the attendant

costs of that duplication in development and maintenance".

- "Data models for different systems are arbitrarily different. The result of this is that complex interfaces are required between systems that share data. These interfaces can account for between 25-70% of the cost of current systems".
- "Data cannot be shared electronically with customers and suppliers, because the structure and meaning of data has not been standardised. For example, engineering design data and drawings for process plant are still sometimes exchanged on paper".

The reason for these problems is a lack of standards that will ensure that data models will both meet business needs and be consistent.

Three Perspectives

A data model *instance* may be one of three kinds according to ANSI in 1975:

- Conceptual schema□: describes the semantics of a domain, being the scope of the model. For example, it may be a model of the interest area of an organization or industry. This consists of entity classes, representing kinds of things of significance in the domain, and relationships assertions about associations between pairs of entity classes. A conceptual schema specifies the kinds of facts or propositions that can be expressed using the model. In that sense, it defines the allowed expressions in an artificial 'language' with a scope that is limited by the scope of the model. The use of conceptual schema has evolved to become a powerful communication tool with business users. Often called a subject area model (SAM) or high-level data model (HDM), this model is used to communicate core data concepts, rules, and definitions to a business user as part of an overall application development or enterprise initiative. The number of objects should be very small and focused on key concepts. Try to limit this model to one page, although for extremely large organizations or complex projects, the model might span two or more pages.

- Logical schema□: describes the semantics, as represented by a particular data manipulation technology. This consists of descriptions of tables and columns, object oriented classes, and XML tags, among other things.
- Physical schema□: describes the physical means by which data are stored. This is concerned with partitions, CPUs, tablespaces, and the like.

The significance of this approach, according to ANSI, is that it allows the three perspectives to be relatively independent of each other. Storage technology can change without affecting either the logical or the conceptual model. The table/column structure can change without (necessarily) affecting the conceptual model. In each case, of course, the structures must remain consistent with the other model. The table/column structure may be different from a direct translation of the entity classes and attributes, but it must ultimately carry out the objectives of the conceptual entity class structure. Early phases of many software development projects emphasize the design of a conceptual data model. Such a design can be detailed into a logical data model. In later stages, this model may be translated into physical data model. However, it is also possible to implement a conceptual model directly.

History

One of the earliest pioneering works in modelling information systems has been done by Young and Kent (1958), who argued for “a precise and abstract way of specifying the informational and time characteristics of a data processing problem”. They wanted to create “a notation that should enable the analyst to organize the problem around any piece of hardware”.

Their work was a first effort to create an abstract specification and invariant basis for designing different alternative implementations using different hardware components. A next step in IS modelling was taken by CODASYL, an IT industry consortium formed in 1959, who essentially aimed at the same thing as Young and Kent: the development of “a proper structure for machine independent problem definition language, at the system level of data processing”.

This led to the development of a specific IS information algebra. In the 1960s data modelling gained more significance with the initiation of the management information system (MIS) concept.

According to Leondes (2002), “during that time, the information system provided the data and information for management purposes. The first generation database system, called Integrated Data Store (IDS), was designed by Charles Bachman at General Electric. Two famous database models, the network data model and the hierarchical data model, were proposed during this period of time”. Towards the end of the 1960s Edgar F. Codd worked out his theories of data arrangement, and proposed the relational model for database management based on first-order predicate logic.

In the 1970s entity relationship modelling emerged as a new type of conceptual data modelling, originally proposed in 1976 by Peter Chen.

Entity relationship models were being used in the first stage of information system design during the requirements analysis to describe information needs or the type of information that is to be stored in a database. This technique can describe any ontology, i.e., an overview and classification of concepts and their relationships, for a certain area of interest. In the 1970s G.M. Nijssen developed “Natural Language Information Analysis Method” (NIAM) method, and developed this in the 1980s in cooperation with Terry Halpin into Object-Role Modelling (ORM).

Further in the 1980s according to Jan L. Harrington (2000) “the development of the object-oriented paradigm brought about a fundamental change in the way we look at data and the procedures that operate on data.

Traditionally, data and procedures have been stored separately: the data and their relationship in a database, the procedures in an application program. Object orientation, however, combined an entity’s procedure with its data.”

TYPES OF DATA MODELS

Database Model

A database model is a theory or specification describing how a database is structured and used. Several such models have been suggested. Common models include:

- Flat model: This may not strictly qualify as a data model. The flat (or table) model consists of a single, two-dimensional array of data elements, where all members of a given column are assumed to be similar values, and all members of a row are assumed to be related to one another.
- Hierarchical model: In this model data is organized into a tree-like structure, implying a single upward link in each record to describe the nesting, and a sort field to keep the records in a particular order in each same-level list.
- Network model: This model organizes data using two fundamental constructs, called records and sets. Records contain fields, and sets define one-to-many relationships between records: one owner, many members.
- Relational model: is a database model based on first-order predicate logic. Its core idea is to describe a database as a collection of predicates over a finite set of predicate variables, describing constraints on the possible values and combinations of values.
- Object-relational model: Similar to a relational database model, but objects, classes and inheritance are directly supported in database schemas and in the query language.
- Star schema is the simplest style of data warehouse schema. The star schema consists of a few “fact tables” (possibly only one, justifying the name) referencing any number of “dimension tables”. The star schema is considered an important special case of the snowflake schema.

Data Structure Diagram

A data structure diagram (DSD) is a diagram and data model used to describe conceptual data models by providing graphical notations which document entities and their relationships, and the constraints that binds them. The basic graphic elements of DSDs are boxes, representing entities, and arrows, representing relationships. Data structure diagrams are most useful for documenting complex data entities.

Data structure diagrams are an extension of the entity-relationship model (ER model). In DSDs, attributes are specified inside the entity boxes rather than outside of them, while relationships are drawn as boxes composed of attributes which specify the constraints that bind entities together. The E-R model, while robust, doesn't provide a way to specify the constraints between relationships, and becomes visually cumbersome when representing entities with several attributes. DSDs differ from the ER model in that the ER model focuses on the relationships between different entities, whereas DSDs focus on the relationships of the elements within an entity and enable users to fully see the links and relationships between each entity. There are several styles for representing data structure diagrams, with the notable difference in the manner of defining cardinality. The choices are between arrow heads, inverted arrow heads (crow's feet), or numerical representation of the cardinality.

Entity-relationship Model

An entity-relationship model (ERM) is an abstract conceptual data model (or semantic data model) used in software engineering to represent structured data. There are several notations used for ERMs.

Geographic Data Model

A data model in Geographic information systems is a mathematical construct for representing geographic objects or surfaces as data. For example,

- the vector data model represents geography as collections of points, lines, and polygons;
- the raster data model represent geography as cell matrixes that store numeric values;

- and the Triangulated irregular network (TIN) data model represents geography as sets of contiguous, nonoverlapping triangles.

Generic Data Model

Generic data models are generalizations of conventional data models. They define standardised general relation types, together with the kinds of things that may be related by such a relation type. Generic data models are developed as an approach to solve some shortcomings of conventional data models.

For example, different modelers usually produce different conventional data models of the same domain. This can lead to difficulty in bringing the models of different people together and is an obstacle for data exchange and data integration. Invariably, however, this difference is attributable to different levels of abstraction in the models and differences in the kinds of facts that can be instantiated (the semantic expression capabilities of the models).

The modelers need to communicate and agree on certain elements which are to be rendered more concretely, in order to make the differences less significant.

Semantic Data Model

A semantic data model in software engineering is a technique to define the meaning of data within the context of its interrelationships with other data. A semantic data model is an abstraction which defines how the stored symbols relate to the real world. A semantic data model is sometimes called a conceptual data model.

The logical data structure of a database management system (DBMS), whether hierarchical, network, or relational, cannot totally satisfy the requirements for a conceptual definition of data because it is limited in scope and biased toward the implementation strategy employed by the DBMS. Therefore, the need to define data from a conceptual view has led to the development of semantic data modelling techniques. That is, techniques to define the meaning of data within the context of its interrelationships with other data.The real world, in terms of resources, ideas, events, etc., are symbolically defined within physical data stores. A semantic data model is an abstraction which defines how the stored symbols relate to the real world. Thus, the model must be a true representation of the real world.

DATA MODEL TOPICS

Data Architecture

Data architecture is the design of data for use in defining the target state and the subsequent planning needed to hit the target state. It is usually one of several architecture domains that form the pillars of an enterprise architecture or solution architecture. A data architecture describes the data structures used by a business and/or its applications. There are descriptions of data in storage and data in motion; descriptions of data stores, data groups and data items; and mappings of those data artifacts to data qualities, applications, locations etc. Essential to realizing the target state, Data architecture describes how data is processed, stored, and utilized in a given system. It provides criteria for data processing operations that make it possible to design data flows and also control the flow of data in the system.

Data Modelling

Data modelling in software engineering is the process of creating a data model by applying formal data model descriptions using data modelling techniques. Data modelling is a technique for defining business requirements for a database. It is sometimes called *database modelling* because a data model is eventually implemented in a database. The way data models are developed and used today. A conceptual data model is developed based on the data requirements for the application that is being developed, perhaps in the context of an activity model. The data model will normally consist of entity types, attributes, relationships, integrity rules, and the definitions of those objects. This is then used as the start point for interface or database design.

Data Properties

Some important properties of data for which requirements need to be met are:

- definition-related properties;
  - *relevance*: the usefulness of the data in the context of your business.
  - *clarity*: the availability of a clear and shared definition for the data.
  - *consistency*: the compatibility of the same type of data from different sources.
- content-related properties;
  - *timeliness*: the availability of data at the time required and how up to date that data is.
  - *accuracy*: how close to the truth the data is.
- properties related to both definition and content;
  - *completeness*: how much of the required data is available.
  - *accessibility*: where, how, and to whom the data is available or not available (e.g. security).
  - *cost*: the cost incurred in obtaining the data, and making it available for use.

Data Organization

Another kind of data model describes how to organize data using a database management system or other data management technology. It describes, for example, relational tables and columns or object-oriented classes and attributes. Such a data model is sometimes referred to as the *physical data model*, but in the original ANSI three schema architecture, it is called “logical”. In that architecture, the physical model describes the storage media (cylinders, tracks, and tablespaces). Ideally, this model is derived from the more conceptual data model described. It may differ, however, to account for constraints like processing capacity and usage patterns.

While *data analysis* is a common term for data modelling, the activity actually has more in common with the ideas and methods of *synthesis* (inferring general concepts from particular instances) than it does with *analysis* (identifying component concepts from more general ones). {*Presumably we call ourselves systems analysts because no one can say systems synthesists.*} Data modelling strives to bring the data structures of interest together into a cohesive, inseparable, whole by eliminating unnecessary data redundancies and by relating data structures with relationships.

A different approach is through the use of adaptive systems such as artificial neural networks that can autonomously create implicit models of data.

Data Structure

A data structure is a way of storing data in a computer so that it can be used efficiently. It is an organization of mathematical and logical concepts of data. Often a carefully chosen data structure will allow the most efficient algorithm to be used.

The choice of the data structure often begins from the choice of an abstract data type. A data model describes the structure of the data within a given domain and, by implication, the underlying structure of that domain itself. This means that a data model in fact specifies a dedicated *grammar* for a dedicated artificial language for that domain. A data model represents classes of entities (kinds of things) about which a company wishes to hold information, the attributes of that information, and relationships among those entities and (often implicit) relationships among those attributes. The model describes the organization of the data to some extent irrespective of how data might be represented in a computer system.

The entities represented by a data model can be the tangible entities, but models that include such concrete entity classes tend to change over time. Robust data models often identify abstractions of such entities. For example, a data model might include an entity class called “Person”, representing all the people who interact with an organization. Such an abstract entity class is typically more appropriate than ones called “Vendor” or “Employee”, which identify specific roles played by those people.

Data Model Theory

The term data model can have two meanings:

1. A data model *theory*, i.e. a formal description of how data may be structured and accessed.
2. A data model *instance*, i.e. applying a data model *theory* to create a practical data model *instance* for some particular application.

A data model theory has three main components:

- The structural part: a collection of data structures which are used to create databases representing the entities or objects modeled by the database.
- The integrity part: a collection of rules governing the constraints placed on these data structures to ensure structural integrity.
- The manipulation part: a collection of operators which can be applied to the data structures, to update and query the data contained in the database.

For example, in the relational model, the structural part is based on a modified concept of the mathematical relation; the integrity part is expressed in first-order logic and the manipulation part is expressed using the relational algebra, tuple calculus and domain calculus.

A data model instance is created by applying a data model theory. This is typically done to solve some business enterprise requirement. Business requirements are normally captured by a semantic logical data model. This is transformed into a physical data model instance from which is generated a physical database.

For example, a data modeler may use a data modelling tool to create an entity-relationship model of the corporate data repository of some business enterprise. This model is transformed into a relational model, which in turn generates a relational database.

Patterns

Patterns are common data modelling structures that occur in many data models.

RELATED MODELS

Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system. It differs from the flowchart as it shows the *data* flow instead of the *control* flow of the program. A data flow diagram can also be used for the visualization of data processing. Data flow diagrams were invented by Larry Constantine, the original developer of structured design, based on Martin and Estrin's "data flow graph" model of computation. It is common practice to draw a context-level Data flow diagram first which shows the interaction between the system and outside entities. The DFD is designed to show how a system is divided into smaller portions and to highlight the flow of data between those parts. This context-level Data flow diagram is then "exploded" to show more detail of the system being modeled.

Information Model

An Information model is not a type of data model, but more or less an alternative model. Within the field of software engineering both a data model and an information model can be abstract, formal representations of entity types that includes their properties, relationships and the operations that can be performed on them. The entity types in the model may be kinds of real-world objects, such as devices in a network, or they may themselves be abstract, such as for the entities used in a billing system. Typically, they are used to model a constrained domain that can be described by a closed set of entity types, properties, relationships and operations.

According to Lee (1999) an information model is a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. It can provide sharable, stable, and organized structure of information requirements for the domain context. More in general the term *information model* is used for models of individual things, such as facilities, buildings, process plants, etc. In those cases the concept is specialised to Facility Information Model, Building Information Model, Plant

Information Model, etc. Such an information model is an integration of a model of the facility with the data and documents about the facility.

An information model provides formalism to the description of a problem domain without constraining how that description is mapped to an actual implementation in software. There may be many mappings of the information model. Such mappings are called data models, irrespective of whether they are object models (e.g. using UML), entity relationship models or XML schemas.

Object Model

An object model in computer science is a collection of objects or classes through which a program can examine and manipulate some specific parts of its world. In other words, the object-oriented interface to some service or system. Such an interface is said to be the *object model of* the represented service or system.

For example, the Document Object Model (DOM) is a collection of objects that represent a page in a web browser, used by script programs to examine and dynamically change the page. There is a Microsoft Excel object model for controlling Microsoft Excel from another program, and the ASCOM Telescope Driver is an object model for controlling an astronomical telescope.

In computing the term *object model* has a distinct second meaning of the general properties of objects in a specific computer programming language, technology, notation or methodology that uses them. For example, the *Java object model*, the *COM object model*, or *the object model of OMT*. Such object models are usually defined using concepts such as class, message, inheritance, polymorphism, and encapsulation. There is an extensive literature on formalized object models as a subset of the formal semantics of programming languages.

Object-Role Model

Object-Role Modelling (ORM) is a method for conceptual modelling, and can be used as a tool for information and rules analysis.

Object-Role Modelling is a fact-oriented method for performing systems analysis at the conceptual level. The quality of a database application depends critically on its design. To help ensure correctness, clarity, adaptability and productivity, information systems are best specified first at the conceptual level, using concepts and language that people can readily understand.

The conceptual design may include data, process and behavioral perspectives, and the actual DBMS used to implement the design might be based on one of many logical data models (relational, hierarchic, network, object-oriented etc).

Unified Modelling Language Models

The Unified Modelling Language (UML) is a standardized general-purpose modelling language in the field of software engineering. It is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system. The Unified Modelling Language offers a standard way to write a system's blueprints, including:

- Conceptual things such as business processes and system functions
- Concrete things such as programming language statements, database schemas
- Reusable software components

UML offers a mix of functional models, data models, and database models.

4

GIS, Remote Sensing and Image Processing in Urban Systems

The Geographic Information System (GIS) is a computer-assisted system for acquisition, storage, analysis and display of geographic data. GIS allows for creating, maintaining and querying electronic databases of information normally displayed on maps. These databases are spatially oriented, the fundamental integrating element being their position on the earth's surface. This system consists of a set of computerised tools and procedures that can be used to effectively encode, store, retrieve, overlay, correlate, manipulate, analyse, query, and display land-related information. They also facilitate the selection and transfer of data to application specific analytical models capable of assessing the impact of alternatives on the chosen environment. The underlying foundation of sound GIS is an effective digital map database, tied to an accurate horizontal control survey framework.

The spatial data generally is in the form of maps, which could be showing topography, geology, soil types, forest and vegetation, land use, water resource availability etc., stored as layers in a digital form. Integrating many layers of data in a computer can easily generate new thematic maps. Thus, a GIS has a database of multiple information layers that can be manipulated to evaluate relationships among the selected elements. GIS can create maps, integrate information, visualise scenarios, solve complicated problems, present powerful ideas, and develop effective solutions. GIS works with two fundamentally different types of geographic models, the “vector” model and the “raster” model. Raster organises spatial features in regular spaced grid of pixels, while the vector data structure organises spatial feature by the set of vectors, which are specified by starting point co-ordinates. A single x, y co-ordinates, can describe the location of a point feature, such as a location of boreholes. Point features are represented as vectors without length and direction. Linear features such as roads and rivers can be stored as point co-ordinates. Polygon features, such as land parcels and river catchments, can be stored as a closed loop of co-ordinates. Compared to a line designated in a raster format, a vector line is 1-d and has no width associated with it. The vector model is extremely useful for describing discrete features, but less useful for describing continuously varying features such as soil top.

ADVANTAGES OF VECTOR TYPE DATA

The Vector Storage Type uses Less Storage Space

I*t supports greater precision in the computation and processing of spatial features.*

The smallest feature in a raster data structure is represented by a single pixel.

The raster model has evolved to model continuous features. A raster image comprises a collection of grid cells rather like a scanned map or a picture. Both the vector and raster models for storing geographic data have unique advantage and disadvantages.

Advantages of Raster Data Type

Provides better representation of continuous surfaces.

Map overlays are efficiently processed if thematic layers are coded in a simple raster structure.

Because the raster grid defines pixels that are constant in shape, spatial relationships among pixels are constant and easily traceable.

GIS has been touted as a great boon to engineering, science, planning, and decision-making in every field. Some of the noteworthy applications of GIS are:

Map generation;

Calculation of land use;

Analysis of optimal land use allocations;

Determining changes over time – Temporal Analyses;

Route guidance and planning;

Targeted marketing;

Habitat prediction;

Ecosystem simulation/Environmental modelling.

REMOTE SENSING

Remote sensing refers to obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation. The eyes are an excellent example of a remote sensing device. With this, it is possible to gather information about the surroundings or even reading the text as in this case. However, this simple definition of remote sensing is more commonly associated with the gauging of interactions between the earth surface and the electromagnetic energy. These days the data gathering of the earth surface is enabled with various sensors that are able to efficiently absorb reflected energy from the earth's surface. The satellites with onboard sensors, play an important role in data capture.

Remote sensing systems are a very important source of information for GIS, as they provide access to spatio-temporal information on surface processes on scales ranging from regional to global. A wide range of environmental parameters can be measured including land use, vegetation types, surface temperatures, soil types, precipitation, phytoplankton, turbidity, surface elevation and geology. Remote sensing and GIS aid immensely in urban sprawl studies.

In the case of a combined application, an efficient, even though more complex, approach is the integration of remote sensing data processing, GIS analyses, database manipulation and models into a single analysis system. Such an integrated analysis, monitoring and forecasting system based on GIS and database management system technologies requires the analyst to understand not only the problem but also the available technologies yet without being a computer expert.

The integration of GIS and remote sensing with the aid of models and additional database management systems (DBMS) is the technically most advanced and applicable approach today.

The remote sensing applications are growing very rapidly with the availability of high-resolution data from the state of the art satellites like IRS-1C/1D/P4 and Landsat. The advancement in computer hardware and software in the area of remote sensing also enhances remote sensing applications. IRS-1C/1D/P4 provides data with 5.8 m resolution in panchromatic mode giving more information of the ground area covered. The remote sensing satellites with high-resolution sensors and wide coverage capabilities will provide the data with better resolution, coverage and revisit (once in 24 days for IRS 1C) to meet the growing applications needs. Many applications like crop acreage and yield estimation, drought monitoring and assessment, flood mapping, wasteland mapping, mineral prospects, forest resource survey etc., have become an integral part of the resources management system. These resource management systems need the data to be transferred in real time or near real time for processing.

The land use classification is primarily to understand the spatial distribution of various land features and plan accordingly for optimum utilisation of the land with least effects on the associated systems. The pattern and extent of land use is influenced mainly by two factors – physical and anthropogenic. Physical factors include topography, climate and soils, which set the broad limits upon the capabilities of the land, and the anthropogenic factors are, density, occupation of the people, socio-economic institutions, the technological level, and infrastructure facilities. GIS and remote sensing collectively help in understanding and undertaking these applications effectively.

IMAGE PROCESSING–RESTORATION, ENHANCEMENT, CLASSIFICATION, TRANSFORMATION

The digital *image processing* is largely concerned with four basic operations: image restoration, image enhancement, image classification, and image transformation. The image restoration is concerned with the correction and calibration of images in order to achieve as faithful representation of the earth surface as possible.

Image enhancement is predominantly concerned with the modification of images to optimise their appearance to the visual system.

Image classification refers to the computer-assisted interpretation of images that is vital to GIS. The *image transformation* refers to the derivation of new imagery as a result of some mathematical treatment of the raw image bands.

The operation of *image restoration* is to correct the distorted image data to create a more faithful representation of the original scene. This normally involves the initial processing of raw image data to correct for geometric distortions, to calibrate the data radiometrically, and to eliminate the noise present in the data. Image rectification and restoration are also termed as pre-processing operations.

Enhancement is concerned with the modification of images to make them more suited to the capabilities of human vision. Regardless of the extent of digital intervention, visual analysis invariably plays a very strong role in all aspects of remote sensing. Enhancement of the imagery can be done by the histogram equalisation method or linear saturation method before analysis.

Digital image classification is the process of assigning a pixel (or groups of pixels) of remote sensing image to a land cover or land use class. The objective of image classification is to classify each pixel into one class (crisp or hard classification) or to associate the pixel with many classes (fuzzy or soft classification). The classification techniques are categorised based on the training process – supervised and unsupervised classification.

Supervised classification has three distinct stages namely training, allocation and testing. Training is the identification of a sample of pixels of known class membership gathered from reference data such as ground truth, existing maps and aerial photographs. In the second stage, the training pixels are used to derive various statistics for each land cover class and so are correspondingly assigned as signature. In the third stage, the pixels are allocated to the same class with which they show greatest similarity based on the signature files.

Unsupervised classification techniques share a common intent to uncover the major land cover classes that exist in the image, without prior knowledge of what they might be. Such procedures often come under cluster analysis, since they search for clusters of pixels with similar reflectance values. Unlike the supervised classification, only major land classes are separated as clusters, while smaller classes may be ignored. The decision for the number of clusters can be based on the histogram analysis of the reflectance values. The most prominent number of classes as seen in the histogram can be considered as the number of clusters.

The Indian Remote Sensing (IRS) satellite's Linear Imaging Self Scanning Sensor (LISS) imagery contains four bands. National Remote Sensing Agency (NRSA) distributes the satellite data for India. This will have image in Band Interleaved by Line (BIL) format i.e., this file contains first line from first band, first line from second band, first line from third band and first line from fourth band in one interleaved line and in second interleaved line it contains second line from first band, second line from second band, second line from third band, second line from fourth band and so on. Band extraction is implemented to separate these bands. Imagery obtained from the satellites will have geometric errors due to the nature of motion of satellite and high altitude of sensing platform. Prominent Ground Control Points (GCPs) from toposheets (which is always correct) are taken to rectify geometric errors. This procedure is also called as geo-correction/geo-rectification.

Image processing, neural network and other techniques are used to analyse the satellite imagery. The decision rule based on geometric shapes, sizes, and patterns present in the data is termed as Spatial Pattern Recognition. Similarly, visual interpretation is done on satellite imagery by considering the elements of image interpretation such as, shape, size, tone, texture, pattern and size for pattern recognition. The pattern recognition of urban sprawl is identified after classification of the remote sensed images for the built-up area and is then further analysed.

DECISION SUPPORT SYSTEM

In recent years, considerable interest has been focused on the use of GIS as a decision support system. For some, this role consists of simply informing the decision making process. However, it is more likely in the realm of resource allocation that the greatest contribution can be made with the aid of GIS and remote sensing.

The use of GIS as a direct extension of the human decision-making process, most particularly in the context of resource allocation decisions, is indeed a great challenge and an important milestone. With the incorporation of many software tools to GIS for multi-criteria and multi-objective decision-making, an area that can broadly be termed decision strategy analysis there seems to be no bounds for the application of GIS.

Closely associated with the decision strategy analysis is the uncertainty management. Uncertainty is not considered as a problem with data, but it is an inherent characteristic of the decision making process. With the increasing pressures on the resource allocation process, the need to recognise uncertainty as a fact of the decision making process has to be understood and carefully assessed. Uncertainty management thus lies at the very heart of effective decision-making and constitutes a very special role for the software systems that support GIS.

The decision support is based on a choice between alternatives arising under a given set of criterion for a given objective. A criterion is some basis for a decision that can be measured and evaluated. Criterion can be of two kinds: factors and constraints, and this can pertain either to attributes of the individual or to an entire decision set.

In this case the objective being to urbanise; constraints include the already existing built-up area, road-rail network, water bodies, etc., where there is no scope for further sprawl; and factors include the components of population growth rate, population density and proximity to the highway and cities.

The decision support system evaluates these sets of data using multi-criteria evaluation. This predicts the possibilities of sprawl in the subsequent years using the current and historical data giving the output images for the objective mentioned.

GEOGRAPHIC INFORMATION SYSTEM: GIS

Geographic information system (GIS) technology can be used for scientific investigations, resource management, and development planning. For example, a GIS might allow emergency planners to easily calculate emergency response times in the event of a natural disaster, or a GIS might be used to find wetlands that need protection from pollution.

What is a GIS?

A GIS is a computer system capable of capturing, storing, analysing, and displaying geographically referenced information; that is, data identified according to location. Practitioners also define a GIS as including the procedures, operating personnel, and spatial data that go into the system.

How does a GIS Work?

The power of a GIS comes from the ability to relate different information in a spatial context and to reach a conclusion about this relationship. Most of the information we have about our world contains a location reference, placing that information at some point on the globe. When rainfall information is collected, it is important to know where the rainfall is located. This is done by using a location reference system, such as longitude and latitude, and perhaps elevation. Comparing the rainfall information with other information, such as the location of marshes across the landscape, may show that certain marshes receive little rainfall. This fact may indicate that these marshes are likely to dry up, and this inference can help us make the most appropriate decisions about how humans should interact with the marsh. A GIS, therefore, can reveal important new information that leads to better decisionmaking.

Many computer databases that can be directly entered into a GIS are being produced by Federal, State, tribal, and local governments, private companies, academia, and nonprofit organizations. Different kinds of data in map form can be entered into a GIS. A GIS can also convert existing digital information, which may not yet be in map form, into forms it can recognize and use. For example, digital satellite images can be analysed to produce a map of digital information about land use and land cover. Likewise, census or hydrologic tabular data can be converted to a maplike form and serve as layers of thematic information in a GIS.

Data Capture

How can a GIS use the information in a map? If the data to be used are not already in digital form, that is, in a form the computer can recognize, various techniques can capture the information. Maps can be digitized by hand-tracing with a computer mouse on the screen or on a digitizing tablet to collect the coordinates of features. Electronic scanners can also convert maps to digits. Cordinates from Global Positioning System (GPS) receivers can also be uploaded into a GIS. A GIS can be used to emphasize the spatial relationships among the objects being mapped. While a computer-aided mapping system may represent a road simply as a line, a GIS may also recognize that road as the boundary between wetland and urban development between two census statistical areas.

Data capture—putting the information into the system—involves identifying the objects on the map, their absolute location on the Earth's surface, and their spatial relationships. Software tools that automatically extract features from satellite images or aerial photographs are gradually replacing what has traditionally been a time-consuming capture process. Objects are identified in a series of attribute tables—the "information" part of a GIS. Spatial relationships, such as whether features intersect or whether they are adjacent, are the key to all GIS-based analysis.

Data Integration

A GIS makes it possible to link, or integrate, information that is difficult to associate through any other means. Thus, a GIS can use combinations of mapped variables to build and analyse new variables. Data integration is the linking of information in different forms through a GIS.

For example, using GIS technology, it is possible to combine agricultural records with hydrography data to determine which streams will carry certain levels of fertilizer runoff. Agricultural records can indicate how much pesticide has been applied to a parcel of land. By locating these parcels and intersecting them with streams, the GIS can be used to predict the amount of nutrient runoff in each stream. Then as streams converge, the total loads can be calculated downstream where the stream enters a lake.

Projection and Registration

A property ownership map might be at a different scale than a soils map. Map information in a GIS must be manipulated so that it registers, or fits, with information gathered from other maps. Before the digital data can be analysed, they may have to undergo other manipulations—*projection conversions,* for example—that integrate them into a GIS. Projection is a fundamental component of mapmaking. A projection is a mathematical means of transferring information from the Earth's three-dimensional, curved surface to a two-dimensional medium—paper or a computer screen. Different projections are used for different types of maps because each projection is particularly appropriate for certain uses. For example, a projection that accurately represents the shapes of the continents will distort their relative sizes.

Since much of the information in a GIS comes from existing maps, a GIS uses the processing power of the computer to transform digital information, gathered from sources with different projections, to a common projection. An elevation image classified from a satellite image of Minnesota exists in a different scale and projection than the lines on the digital file of the State and province boundaries. The elevation image has been reprojected to match the projection and scale of the State and province boundaries.

Data Structures

Can a land use map be related to a satellite image, a timely indicator of land use? Yes, but because digital data are collected and stored in different ways, the two data sources may not be entirely compatible. Therefore, a GIS must be able to convert data from one structure to another. Satellite image data that have been interpreted by a computer to produce a land use map can be "read into" the GIS in raster format. Raster data files consist of rows of uniform cells coded according to data values. An example is land cover classification. Raster files can be manipulated quickly by the computer, but they are often less detailed and may be less visually appealing than vector data files, which can approximate the appearance of more traditional hand-drafted maps. Vector digital data have been captured as points, lines (a series of point coordinates), or areas (shapes bounded by lines). An example of data typically held in a vector file would be the property boundaries for a particular housing subdivision.

Data restructuring can be performed by a GIS to convert data between different formats. For example, a GIS can be used to convert a satellite image map to a vector structure by generating lines around all cells with the same classification, while determining the spatial relationships of the cell, such as adjacency or inclusion.

Data Modelling

It is impossible to collect data over every square meter of the Earth's surface. Therefore, samples must be taken at discrete locations. A GIS can be used to depict two-and three-dimensional characteristics of the Earth's surface, subsurface, and atmosphere from points where samples have been collected. For example, a GIS can quickly generate a map with isolines that indicate the pH of soil from test points. Such a map can be thought of as a soil pH contour map. Many sophisticated methods can estimate the characteristics of surfaces from a limited number of point measurements. Two-and three-dimensional contour maps created from the surface modelling of sample points from pH measurements can be analysed together with any other map in a GIS covering the area.

The way maps and other data have been stored or filed as layers of information in a GIS makes it possible to perform complex analyses. A crosshair pointer (top) can be used to point at a location stored in a GIS. The bottom illustration depicts a computer screen containing the kind of information stored about the location—for example, the latitude, longitude, projection, coordinates, closeness to wells, sources of production, roads, and slopes of land.

Information Retrieval

What do you know about the swampy area at the end of your street? With a GIS you can "point" at a location, object, or area on the screen and retrieve recorded information about it from offscreen files. Using scanned aerial photographs as a visual guide, you can ask a GIS about the geology or hydrology of the area or even about how close a swamp is to the end of a street. This type of analysis allows you to draw conclusions about the swamp's environmental sensitivity.

Topological Modelling

Have there ever been gas stations or factories that operated next to the swamp? Were any of these uphill from and within 2 miles of the swamp? A GIS can recognize and analyse the spatial relationships among mapped phenomena. Conditions of adjacency (what is next to what), containment (what is enclosed by what), and proximity (how close something is to something else) can be determined with a GIS.

Networks

When nutrients from farmland are running off into streams, it is important to know in which direction the streams flow and which streams empty into other streams. This is done by using a linear network. It allows the computer to determine how the nutrients are transported downstream. Additional information on water volume and speed throughout the spatial network can help the GIS determine how long it will take the nutrients to travel downstream. A GIS can simulate the movement of materials along a network of lines. These illustrations show the route of pollutants through a stream system. Flow directions are indicated by arrows.

Overlay

Using maps of wetlands, slopes, streams, land use, and soils, the GIS might produce a new map layer or overlay that ranks the wetlands according to their relative sensitivity to damage from nutrient runoff.

Data Output

A critical component of a GIS is its ability to produce graphics on the screen or on paper to convey the results of analyses to the people who make decisions about resources. Wall maps, Internet-ready maps, interactive maps, and other graphics can be generated, allowing the decisionmakers to visualize and thereby understand the results of analyses or simulations of potential events.

Framework for Cooperation

The use of a GIS can encourage cooperation and communication among the organizations involved in environmental protection, planning, and resource management. The collection of data for a GIS is costly. Data collection can require very specialized computer equipment and technical expertise. Standard data formats ease the exchange of digital information among users of different systems. Standardization helps to stretch data collection funds further by allowing data sharing, and, in many cases, gives users access to data that they could not otherwise collect for economic or technical reasons. Organizations such as the University Consortium for Geographic Information Science and the Federal Geographic Data Committee seek to encourage standardization efforts.

GIS THROUGH HISTORY

Some 35,000 years ago, Cro-Magnon hunters drew pictures of the animals they hunted on the walls of caves near Lascaux, France. Associated with the animal drawings are track lines and tallies thought to depict migration routes. These early records followed the two-element structure of modern geographic information systems (GIS): a graphic file linked to an attribute database.

Today, biologists use collar transmitters and satellite receivers to track the migration routes of caribou and polar bears to help design programs to protect the animals. In a GIS, the migration routes were indicated by different colours for each month for 21 months. Researchers then used the GIS to superimpose the migration routes on maps of oil development plans to determine the potential for interference with the animals.

Mapmaking

Researchers are working to incorporate the mapmaking processes of traditional cartographers into GIS technology for the automated production of maps. One of the most common products of a GIS is a map. Maps are generally easy to make using a GIS and they are often the most effective means of communicating the results of the GIS process. Therefore, the GIS is usually a prolific producer of maps. The users of a GIS must be concerned with the quality of the maps produced because the GIS normally does not regulate common cartographic principles. One of these principles is the concept of generalization, which deals with the content and detail of information at various scales. The GIS user can change scale at the push of a button, but controlling content and detail is often not so easy. Mapmakers have long recognized that content and detail need to change as the scale of the map changes. For example, the State of New Jersey can be mapped at various scales, from the small scale of 1:500,000 to the larger scale of 1:250,000 and the yet larger scale of 1:100,000,but each scale requires an appropriate level of generalization.

Site Selection

The U.S. Geological Survey (USGS), in a cooperative project with the Connecticut Department of Natural Resources, digitized more than 40 map layers for the areas covered by the USGS Broad Brook and Ellington 7.5-minute topographic quadrangle maps. This information can be combined and manipulated in a GIS to address planning and natural resource issues. GIS information was used to locate a potential site for a new water well within half a mile of the Somers Water Company service area. To prepare the analysis, cartographers stored digital maps of the water service areas in the GIS. They used the proximity function in the GIS to draw a half-mile buffer zone around the water company service area. This buffer zone was the “window” used to view and combine the various map coverages relevant to the well site selection.

The land use and land cover map for the two areas shows that the area is partly developed. A GIS was used to select undeveloped areas from the land use and land cover map as the first step in finding well sites. The developed areas were eliminated from further consideration. The quality of water in Connecticut streams is closely monitored. Some of the streams in the study area were known to be unusable as drinking water sources. To avoid pulling water from these streams into the wells, 100-meter buffer zones were created around the unsuitable streams using the GIS, and the zones were plotted on the map. The areas in blue have the characteristics desired for a water well site.

Point sources of pollution are recorded by the Connecticut Department of Natural Resources. These records consist of a location and a text description of the pollutant. To avoid these toxic areas, a buffer zone of 500 meters was established around each point. This information was combined with the previous two map layers to produce a new map of areas suitable for well sites. Points sources of pollution in the water service area are identified and entered into a GIS.

The map of surficial geology shows the earth materials that lie above bedrock. Since the area under consideration in Connecticut is covered by glacial deposits, the surface consists largely of sand and gravel, with some glacial till and fine-grained sediments. Of these materials, sand and gravel are the most likely to store water that could be tapped with wells. Areas underlain by sand and gravel were selected from the surficial geology map. They were combined with the results of the previous selections to produce a map consisting of: (1) sites in underdeveloped areas underlain by sand and gravel, (2) more than 500 meters from point sources of pollution, and (3) more than 100 meters from unsuitable streams. A map that shows the thickness of saturated sediments was created by using the GIS to subtract the bedrock elevation from the surface elevation. For this analysis, areas having more than 40 feet of saturated sediments were selected and combined with the previous overlays.

The resulting site selection map shows areas that are undeveloped, are situated outside the buffered pollution areas, and are underlain by 40 feet or more of water-saturated sand and gravel. Because of map resolution and the limits of precision in digitizing, the very small polygons (areas) may not have all of the characteristics analysed, so another GIS function was used to screen out areas smaller than 10 acres. The final six sites are displayed with the road and stream network and selected place names for use in the field.

Potential water well sites, roads, streams and place names. The process illustrated by this site selection analysis has been used for many common applications, including transportation planning and waste disposal site location. The technique is particularly useful when several physical factors must be considered and integrated over a large area.

Emergency Response Planning

A GIS was used to combine road network and earth science information to analyse the effect of an earthquake on the response time of fire and rescue squads. The area covered by the USGS Sugar House 7.5-minute topographic quadrangle map was selected for the study because it includes both undeveloped areas in the mountains and a part of Salt Lake City. Detailed earth science information was available for the entire region. The road network from a USGS digital line graph includes information on the types of roads, which range from rough trails to divided highways. The locations of fire stations were plotted on the road network. A GIS function called network analysis was used to calculate the time necessary for emergency vehicles to travel from the fire stations to different areas of the city. The network analysis function considers two elements: (1) distance from the fire station, and (2) speed of travel based on the type of road. The analysis shows that under normal conditions, most of the area within the city will be served in less than 7 minutes and 30 seconds because of the distribution and density of fire stations and the continuous network of roads.

The accompanying illustration depicts the blockage of the road network that would result from an earthquake, assuming that any road crossing the fault trace would become impassable. The primary effect on emergency response time would occur in neighbourhoods west of the fault trace, where travel times from the fire stations would be noticeably lengthened. Road network of area covered by the Sugar House quadrangle plotted from USGS digital line graph data, indicating the locations of fire stations and travel times of emergency vehicles. Areas in blue can receive service within 2½minutes, area in green within 5 minutes, areas in yellow within 7½ minutes, and areas in magenta within 10 minutes. Areas in white cannot receive service within 10 minutes.

After faulting, initial model. Network analysis in a GIS produces a map of travel times from the stations after faulting.

The fault is in red. Emergency response times have increased for areas west of the fault. The Salt Lake City area lies on lake sediments of varying thicknesses.

These sediments range from clay to sand and gravel, and most are water-saturated. In an earthquake, these materials may momentarily lose their ability to support surface structures, including roads.

The potential for this phenomenon, known as liquefaction, is shown in a composite map portraying the inferred relative stability of the land surface during an earthquake. Areas near the fault and underlain by thick, loosely consolidated, water-saturated sediments will suffer the most intense surface motion during an earthquake. Areas on the mountain front with thin surface sediments will experience less additional ground acceleration. The map of liquefaction potential was combined with the road network analysis to show the additional effect of liquefaction on response times.

The final map shows that areas near the fault, as well as those underlain by thick, water-saturated sediments, are subject to more road disruptions and slower emergency response than are other areas of the city. Map of potential ground l liquefaction during an earthquake. The least stable areas are shown by yellows and oranges, the most stable by grays and browns. After faulting, final model. A map showing the effect of an earthquake on emergency travel times is reduced by combining the liquefaction potential information.

THREE-DIMENSIONAL GIS

To more realistically analyse the effect of the Earth's terrain, we use three-dimensional models within a GIS. A GIS can display the Earth in realistic, three-dimensional perspective views and animations that convey information more effectively and to wider audiences than traditional, two-dimensional, static maps.

The U.S. Forest Service was offered a land swap by a mining company seeking development rights to a mineral deposit in Arizona's Prescott National Forest. Using a GIS, the USGS and the U.S. Forest Service created perspective views of the area to depict the terrain as it would appear after mining. To assess the potential hazard of landslides both on land and underwater, the USGS generated a three-dimensional image of the San Francisco Bay area. It created the image by mosaicking eight scenes of natural colour composite Landsat 7 enhanced thematic mapper imagery on California fault data using approximately 700 digital elevation models at 1:24,000 scale.

Graphic Display Techniques

Traditional maps are abstractions of the real world; each map is a sampling of important elements portrayed on a sheet of paper with symbols to represent physical objects. People who use maps must interpret these symbols. Topographic maps show the shape of the land surface with contour lines. Graphic display techniques in GISs make relationships among map elements more visible, heightening one's ability to extract and analyse information.

Two types of data were combined in a GIS to produce a perspective view of a part of San Mateo County, Calif. The digital elevation model, consisting of surface elevations recorded on a 30-meter horizontal grid, shows high elevations as white and low elevations as black. The accompanying Landsat thematic mapper image shows a false-colour infrared image of the same area in 30-meter pixels, or picture elements and combine the two images to produce the three-dimentional image.

Visualization

Maps have traditionally been used to explore the Earth. GIS technology has enhanced the efficiency and analytical power of traditional cartography. As the scientific community recognizes the environmental consequences of human activity, GIS technology is becoming an essential tool in the effort to understand the process of global change. Map and satellite information sources can be combined in models that simulate the interactions of complex natural systems. Through a process known as visualization, a GIS can be used to produce images—not just maps, but drawings, animations, and other cartographic products. These images allow researchers to view their subjects in ways that they never could before. The images often are helpful in conveying the technical concepts of a GIS to nonscientists.

Adding the Element of Time

The condition of the Earth's surface, atmosphere, and subsurface can be examined by feeding satellite data into a GIS. GIS technology gives researchers the ability to examine the variations in Earth processes over days, months, and years. As an example, the changes in vegetation vigor through a growing season can be animated to determine when drought was most extensive in a particular region. The resulting normalized vegetation index represents a rough measure of plant health. Working with two variables over time will allow researchers to detect regional differences in the lag between a decline in rainfall and its effect on vegetation. The satellite sensor used in this analysis is the advanced very high resolution radiometer (AVHRR), which detects the amounts of energy reflected from the Earth's surface at a 1-kilometer resolution twice a day. Other sensors provide spatial resolutions of less than 1 meter.

Serving GIS over the Internet

Through Internet map server technology, spatial data can be accessed and analysed over the Internet. For example, current wildfire perimeters are displayed with a standard web browser, allowing fire managers to better respond to fires while in the field and helping homeowners to take precautionary measures.

The Future of GIS

Environmental studies, geography, geology, planning, business marketing, and other disciplines have benefitted from GIS tools and methods. Together with cartography, remote sensing, global positioning systems, photogrammetry, and geography, the GIS has evolved into a discipline with its own research base known as geographic information sciences. An active GIS market has resulted in lower costs and continual improvements in GIS hardware, software, and data. These developments will lead to a much wider application of the technology throughout government, business, and industry. GIS and related technology will help analyse large datasets, allowing a better understanding of terrestrial processes and human activities to improve economic vitality and environmental quality.

GIS AND REMOTE SENSING IN URBAN PLANNING, IMPLEMENTATION AND MONITORING

At the threshold of 21st century it is an interesting question to ask ourselves, if 20th century belonged to technological advancement, then what will the major event of 21st century? Last century witnessed a trickle of urbanisation and emergence of metropolises. Never in the history of cities have been stakes related to the controlling the world ' s urban development been as crucial for the future of humanity. This is true not only in terms of quantity, as city dwellers are expected to account for roughly half the worlds population by the year 2000, but also in terms of quality, since the bulk of economic, social, scientific, technological and cultural development will take place in cities and in particular in largest conurbation known as Metropolises. The increasing urbanisation of the world's population is inescapable and irreversible. This goes to show that it is necessary and fundamental for policy makers to make technologies like GIS and Remote Sensing imperative for the urban planning.

Biggest challenge of 21st century will be to understand phenomenon of urban agglomeration. Most of the 20th centuries concept and visions about cities are being questioned. Traditional approaches and techniques designed for town and cities may turn out to be inadequate tools to deal with metropolises. New approaches needs invented, other actions/methods needs to be incorporated in the existing work methods. From the vast experience of the past, a solution for the future planning methods has to be derived.

The objective of this paper is to confine its search for new techniques to deal with large urban agglomeration and application of GIS and Remote sensing technique at the various stages of planning, implementation and monitoring of the urban projects. One of the established practice to search for new techniques is to identify the problems faced by the planners in the past. From the experience of last two decades in planning and development of large urban projects like Rohini, Dwarka and Narela Sub-City, it is learnt by the planners that transformation of paper plan on the ground takes much more than simply preparation of the plan. For example, the area of the each planned project is ranging between 2000 hact to 5000 hact. Identification of land for land acquisition, actual physical possession of land and development of roads etc. takes considerable time. During this incubation time ground realities have changed and planning proposals do suffer. Therefore, rate of implementation of the plans is below the desired level. One of the major lacuna identified by the planners is non-availability of base-maps, absence of any method for updating the base maps and the absence of any monitoring mechanism. Broadly the problems can be described in the following manner:-

a. In most of the metropolitan areas Survey of India base maps are available. However, these maps are 5 to 10 years old which create problem for making accurate planning proposals and predictions.

b. Even after the base maps are prepared and proposals are placed on it, the difficulty arises due to absence of any reliable method for updating available base maps.

c. The fringe area where most of urban development is taking place i.e. Urban Fringe area is very dynamic is terms of land development, so a device/method is required to detect changes on continuous basis, so that the planning proposals can be re-adjusted and re-shaped as per the ground realities.

d. Physical planning largely depends upon inputs from the physical survey, various social and physical requirements and site conditions. The planners based on his training and intuition, works out the planning proposals. However it has been observed that GIS and Remote Sensing can provide intelligence for formulation of the proposals.

e. The most important aspect which is generally overlooked is to co-relate or transfer planning proposals on revenue maps. The planners face difficulty in understanding the plans prepared, by revenue staff as these plans are on different scale and projection. Similarly revenue staff is unable to superimpose planning proposals on revenue maps manually. However it has been observed that superimposition of any two ‘digital maps’ is easy as digital maps are scale free. Thus an interface between planning basemap and revenue maps can be established.

Once problems faced by the planners while preparing basemaps etc. are identified, then search for solutions leads to mapping related technologies which are at our disposal. GIS and Remote Sensing are mapping related technologies and it is capable to overcome problems discussed in the foregoing paragraphs.

Application of these technologies has given way to innovative methods in planning process-

a. Digitisation of planning basemaps and various layoutplan has facilitated updating of basemaps wherever changes have taken place in terms of land development etc. Digital maps provides flexibility as digital maps are scale free. Superimposition of any two digital maps which are on two different scales is feasible. This capability of digital maps facilitates insertion of fresh survey or modified maps into existing basemaps. Similarly superimposition of revenue maps on basemaps with reasonable accuracy is great advantage compared to manually done jobs.

b. Computerization within the planning department helps to disseminate information from one section to another section also. This may result in better horizontal co ordination between various sections and also remove redundant data from the database.

c. Experience shows that Remote Sensing techniques are extremely useful for change detection analysis and selection of sites/areas for specific purposes like land acquisition etc. The use of satellite imageries in Rohini Project has demonstrated systematic analysis of land available for acquisition and monitoring of areas which are prone to encroachment etc.

d. Since all this information/maps is available in digital format, co-relating various layers of information i.e. satellite imagery, planning maps and revenue maps on common feature is feasible, with the help of GIS & Image processing software available in the market. Such superimposed maps provides valuable information/ inputs for planning, implementation and land management.

Applications of GIS and Remote Sensing technology have been demonstrated in large urban projects like Rohini and Dwarka sub city of Delhi Development Authority.

The Rohini Project in the north and Dwarka Project in the south west Delhi is having about one million population each Planning, implementation and land management of such large projects demands continuous monitoring of plan proposals and vast tracts of acquired land.

Experience shows that these techniques are useful to detect-

a. Large scale encroachments.
b. Large scale conversion of agriculture land into non agriculture land.
c. Vacant pockets of agriculture land for acquisition.
d. Identification of revenue boundaries of villages on satellite imageries.
e. Large scale change detection in status of terrain is possible with help of satellite imageries.

THE IMPORTANCE OF PROJECT PLANNING

GIS projects are expensive in terms of both time and money. Municipal GIS and facilities management projects developed by utilities may take a decade or more to bring on-line at a cost of tens or hundreds of millions of dollars.

Careful planning at the outset, as well as during the project, can help to avoid costly mistakes. It also provides assurance that a GIS will accomplish it goals on schedule and within budget.

There is a temptation, when a new technology like GIS becomes available, to improvise a solution to its use, that is to get started without considering where the project will lead. The greatest danger is that decisions made in haste or on the spur of the moment will have to be reversed later or will prove too costly to implement, meaning a GIS project may have to be abandoned. To avoid disappointing experiences like these, GIS professionals have developed a well-defined planning methodology often referred to as *project lifecycle*. Lifecycle planning involves setting goals, defining targets, establishing schedules, and estimating budgets for an entire GIS project.

The original impetus for developing effective lifecycle planning was cost containment. For many decades, the rationale for implementing new information technologies was that, in the long run, such projects would reduce the cost of business operations.

This optimistic appraisal of the benefits of information technologies has not borne out in the American economy during the past two decades. In almost all cases, adopting new information technologies has added to the cost of business operations without producing a corresponding increase in traditional measures of labour productivity.

This does not mean that information technologies have been a failure. Rather, these systems allow users to accomplish a greater range of varied and complex tasks, but at a higher cost. Users are not so much doing their previous work at faster speeds, but assuming new tasks offered by the new technologies. Support staff once satisfied with producing in-house documents may now be tempted to issue them using desktop publishing software or on-line in the Worldwide Web.

Cartographers once satisfied with producing discrete utility maps for individual construction projects may be tempted to create an encompassing map and GIS database containing maintenance records for an entire city. It is generally recognized that, for the foreseeable future, most information technologies projects will have to be justified on the basis of a “do more, pay more” philosophy.

This means that effective lifecycle planning is all the more important. In the past, projected existing costs could be used as a baseline against which improvements could be measured. If the cost curve for new information technologies is always above the baseline, then greater care must be exerted in setting goals, establishing targets, and estimating budgets. There is far too great a danger that, in the absence of such checks and balances, a project may grow out of control.

THE VALUE OF A PROBLEM-SOLVING APPROACH

Lifecycle planning is really a process of practical problem solving applied to all aspects of a GIS development project.

Particular care is exerted in defining the nature of a problem or new requirement, estimating the costs and feasibility of proceeding, and developing a solution. This process should not be abridged; each step is important to the overall process.

OTHER PLANNING AND DATABASE ISSUES

The project planning cycle outlines a process, but the issues that must be addessed at each stage of this process will vary considerably from organization to organization. Some topics are of critical importance to large municipal, state, and private AM/FM applications, but less so for research applications of limited scope. Among the issues that must be addressed in large GIS projects are:

Security

The security of data is always a concern in large GIS projects. But there is more to security than protecting data from malicious tampering or theft. Security also means that data is protected from system crashes, major catastrophes, and inappropriate uses.

As a result, security must be considered at many levels and must anticipate many potential problems. GIS data maintained by government agencies often presents difficult challenges for security. While some sorts of data must be made publicly accessible under open records laws, other types are protected from scrutiny. If both types are maintained within a single system, managing appropriate access can be difficult. Distribution of data across open networks is always a matter of concern.

Documentation

Most major GIS datasets will outlive the people who create them. Unless all the steps involved in coding and creating a dataset are documented, this information will be lost as staff retire or move to new positions. Documentation must begin at the very start of GIS project and continue through its life. It is best, perhaps, to actually assign a permanent staff to documentation to make sure that the necessary information is saved and revised in a timely fashion.

Data Integrity and Accuracy

When mistakes are discovered in a GIS database, there must be a well-defined procedure for their correction (and for documenting these corrections).

Furthermore, although many users may have to use the information stored in a GIS database, not all of these users should be permitted to make changes. Maintaining the integrity of the different layers of data in a comprehensive GIS database can be a challenging task.

A city's water utility may need to look at GIS data about right-of-ways for power and cable utilities, but it should not be allowed to change this data. Responsibility for changing and correcting data in the different layers must be clearly demarcated among different agencies and offices.

Synchronization of Usage

GIS datasets employed in government or by utilities will have many users. One portion of the dataset may be in demand simultaneously by several users as well as by staff charged with updating and adding new information. Making sure that all users have access to current data whenever they need it can be a difficult challenge for GIS design. Uncontrolled usage may be confusing to all users, but the greatest danger is that users may actually find themselves interfering with the project workflow or even undoing one another's work.

Update Responsibility

Some GIS datasets will never be "complete." Cities and utility territories keep growing and changing and the database must be constantly updated to reflect these changes. But these changes occur on varying schedules and at varying speeds. Procedures must be developed to record, check, and enter these changes in the GIS database. Furthermore, it may be important to maintain a record of the original data. In large GIS projects, updating the database may be the responsibility of a full-time staff.

Minimization of Redundancy

In large GIS projects, every byte counts. If a database is maintained for 30-50 years, every blank field and every duplicated byte of information will incur storage costs for the full length of the project. Not only will wasted storage space waste money, it will also slow performance. This is why in large, long-term GIS projects, great attention to devoted to packing data as economically as possible and reducing duplication of information.

Data Independence and Upgrade Paths

A GIS database will almost always outlive the hardware and software that is used to create it. Computer hardware has a useable life of 2-5 years, software is sometimes upgraded several times a year. If a GIS database is totally dependent on a single hardware platform or a single software system, it too will have to be upgraded just as often. Therefore, it is best to create a database that is as independent as possible of hardware and software.

Through careful planning and design, data can be transferred as ASCII files or in some metadata or exchange format from system to system. There is nothing worse than having data held in a proprietary vendor-supported format and then finding that the vendor has changed or abandoned that format. In this way, GIS designers should think ahead to possible upgrade paths for their database. It is notoriously difficult to predict what will happen next in the world of computers and information technology. To minimize possible problems, thought should be given to making the GIS database as independent as possible of the underlying software and hardware.

Privacy

Safeguards on personal privacy have become a great concern over the past decade, particularly with the rise of the internet and web. These concerns arise to two principal situations. The first, is the hacking into, accidentally release, or inappropriate disclosure of privileged information which can compromise an individual's privacy with respect to medical conditions, financial situation, sexual, political, religious beliefs & values and other privileged personal information. The second is the ease with which information and computer technologies permit the creation of information "mosaics" or personal profiles from small pieces of seemingly innocuous, non-confidential data.

5

GIS, GPS and Remote Sensing Technologies in Extension Services

INTRODUCTION

The rapid development of spatial technologies in recent years has made available new tools and capabilities to Extension services and clientele for management of spatial data. In particular, the evolution of geographic information systems (GIS), the global positioning system (GPS), and remote sensing (RS) technologies has enabled the collection and analysis of field data in ways that were not possible before the advent of the computer. How can potential users with little or no experience with GIS-GPS-RS technologies determine if they would be useful for their applications? How do potential users learn about these technologies? Once a need is established, what potential pitfalls or problems should the user know to avoid? This chapter describes some uses of GIS-GPS-RS in agricultural and resource management applications, provides a roadmap for becoming familiar with the technologies, and makes recommendations for implementation.

SPATIAL TECHNOLOGIES

Geographic Information Systems

GIS applications enable the storage, management, and analysis of large quantities of spatially distributed data. These data are associated with their respective geographic features.

For example, water quality data would be associated with a sampling site, represented by a point. Data on crop yields might be associated with fields or experimental plots, represented on a map by polygons. A GIS can manage different data types occupying the same geographic space. For example, a biological control agent and its prey may be distributed in different abundances across a variety of plant types in an experimental plot. Although predator, prey, and plants occupy the same geographic region, they can be mapped as distinct and separate features.

The ability to depict different, spatially coincident features is not unique to a GIS, as various computer aided drafting (CAD) applications can achieve the same result. The power of a GIS lies in its ability to analyse relationships between features and their associated data. This analytical ability results in the generation of new information, as patterns and spatial relationships are revealed.

The Global Positioning System

GPS technology has provided an indispensable tool for management of agricultural and natural resources. GPS is a satellite-and ground-based radio navigation and locational system that enables the user to determine very accurate locations on the surface of the Earth. Although GPS is a complex and sophisticated technology, user interfaces have evolved to become very accessible to the non-technical user. Simple and inexpensive GPS units are available with accuracies of 10 to 20 meters, and more sophisticated precision agriculture systems can obtain centimeter level accuracies.

Remote Sensing

Remote sensing technologies are used to gather information about the surface of the earth from a distant platform, usually a satellite or airborne sensor. Most remotely sensed data used for mapping and spatial analysis is collected as reflected electromagnetic radiation, which is processed into a digital image that can be overlaid with other spatial data. Reflected radiation in the infrared part of the electromagnetic spectrum, which is invisible to the human eye, is of particular importance for vegetation studies. For example, chlorophyll strongly absorbs blue (0.48 mm) and red (0.68 mm) wavelength radiation and reflects near-infrared radiation (0.75-1.35 mm). Leaf vacuole water absorbs radiation in the infrared region from 1.35-2.5 mm. The spectral properties of vegetation in different parts of the spectrum can be interpreted to reveal information about the health and status of crops, rangelands, forests and other types of vegetation.

Applications

The uses of GIS, GPS, and RS technologies, either individually or in combination, span a broad range of applications and degrees of complexity. Simple applications might involve determining the location of sampling sites, plotting maps for use in the field, or examining the distribution of soil types in relation to yields and productivity. More complex applications take advantage of the analytical capabilities of GIS and RS software. These might include vegetation classification for predicting crop yield or environmental impacts, modelling of surface water drainage patterns, or tracking animal migration patterns.

Precision Agriculture

GIS-GPS-RS technologies are used in combination for precision farming and site-specific crop management. Precision farming techniques are employed to increase yield, reduce production costs, and minimize negative impacts to the environment. Using GIS analytical capabilities, variable parameters that can affect agricultural production can be evaluated. These parameters include yield variability, physical parameters of the field, soil chemical and physical properties, crop variability (e.g., density, height, nutrient stress, water stress, chlorophyll content), anomalous factors (e.g., weed, insect, and disease infestation, wind damage), and variations in management practices (e.g., tillage practices, crop seeding rate, fertilizer and pesticide application, irrigation patterns and frequency). Site-specific data, such as soil characteristics, fertility and nutrient data, topographic and drainage characteristics, yield data, harvester-mounted yield sensor data, and remotely-sensed vegetation indices, are collected from different sources and stored and managed in a spatial database, either contained within the GIS or connected to the GIS from an external source. The analytical power of a GIS is applied to the data to identify patterns in the field (e.g., areas of greater or lesser yield; correlations between yield and topography or characteristics such as nutrient concentrations or drainage).

Once patterns and correlations are elucidated, management practices can be modified to optimize yield and production costs, and minimize environmental impacts caused by excessive applications of fertilizers and pesticides. Site-specific applications of fertilizers, pesticides and other applications can be implemented by dividing a field into smaller management zones that are more homogeneous in properties of interest than the field as a whole.

Forest Management

Spatial technologies are well suited for applications to resource management issues. The ability to interface GIS with relational databases enables integration of large data sets and many variables to support management decisions (e.g., Arvanitis, Ramachandran, Brackett, Rasoul, & Du, 2000). One example is the Florida Agroforestry Decision Support System (FADSS). FADSS is a GIS application that integrates geographically linked data on climate and soil characteristics in the state of Florida with a database of over 500 trees and 50 tree attributes. FADSS enables landowners, farmers and extension agents to make management decisions based on site-specific and tree-specific information.

Habitat Analysis

The modelling capabilities of GIS can be combined with remotely sensed landscape imagery to evaluate the effects of management practices and to assist resource managers and public decision makers in making informed decisions. For example, a GIS-enabled program, VVF, was developed to assess the suitability of a landscape as a species habitat. VVF integrates user-selected environmental variables to produce habitat suitability maps, and enables the user to create habitat suitability models for a specified area. Another model, LEEMATH (Landscape Evaluation of Effects of Management Activities on Timber and Habitat), evaluates both economic and ecological effects of alternative management strategies on timber production and habitat quality.

Data Analysis and Display

The spatial visualization capabilities of GIS technology interfaced with a relational database provide an effective method for analysing and displaying the impacts of Extension education and outreach projects. This application was demonstrated in the Florida Yards & Neighborhood (FY&N) program developed by the University of Florida Extension to teach homeowners and landowners how to reduce non-point source pollution and storm water runoff and protect the environment through landscape practices they exercise in their own yards.

Homeowners filled out surveys both before and after receiving training in landscaping methods. Responses to questions concerning landscape practices were rated as good, fair, or poor, and statistical analysis was conducted on before and after scores for each landscape practice using a relational database interfaced with GIS software. Geospatial analysis of the extent of homeowner/landowner adoption of these best management practices taught by the program enabled assessment of impact by acreage and location, identification of areas needing greater emphasis, tracking of change, and the ability for policymakers to see impacts in map format.

Where to Start

To the uninitiated Extension specialist, the complexity and vast array of potential applications can be confusing and intimidating. Because the applications of GIS-GPS-RS cut across a great many disciplines, chances are good that these technologies can be beneficial in your own area of expertise. How do potential users with little prior knowledge identify specific ways in which they can be useful in their own work?

The decision must begin with a process of self-education. This includes gaining an understanding of the basic concepts of the technologies and doing a careful evaluation of your own needs and the needs of your clientele. A mapping gateway for resource managers and provides information on the fundamentals of GIS-GPS-RS technologies and data and provides numerous links to other sources of information, tools, utilities, data, and software applications. Another useful resource is ESRI's Virtual Campus online education and training Web site. This site offers course both free and inexpensive modules on the use of ESRI's GIS software, as well as a courses on how to go about planning and implementing a GIS.

Once you have researched the potential for GIS-GPS-RS technologies in your field, it is important to become familiar with the workings and capabilities of different software applications and equipment technologies before a decision is made about implementation. A number of software companies have free data readers and browsers that provide an opportunity to examine and use some of the functionality of their software packages. Some examples include ESRI's Arcexplorer and Leica Geosystems' ViewFinder. Many software companies will also provide time-limited trial copies of software packages to allow the user to evaluate the applications before purchasing. If your needs are limited to obtaining and viewing images and GIS data layers and performing simple analysis functions, then one or several of these free data viewer programs may be sufficient for your purposes. Using ESRI's ArcExplorer, for example, you can view, identify, locate and query geographic and attribute data; create thematic maps; and perform basic statistical analysis.

If you need to create or edit new data layers, or perform some basic analysis and data conversion functions, several freeware applications are available that may fulfill your requirements. One of these applications, Forestry GIS (fGIS™), was developed for operational field managers in the natural resources by the University of Wisconsin. fGIS is freely downloadable and can be used to edit existing GIS data, digitize new data layers, query and search spatial data, build customized data views and create maps. The application also includes utilities for working with database tables, transforming spatial data to different coordinate systems, and designing diagrams. Another freeware GIS application, DIVA-GIS was designed for mapping and analysing the distributions of species. DIVA-GIS can create, edit, and transform GIS data files, and has capabilities for various statistical and biological modelling functions.

If you need to perform more sophisticated spatial analytical functions, then you will most likely need to purchase commercial software with more functionality. A number of different GIS and RS software packages are commercially available, each with different features and functionality. Some of these applications can be expensive for organizations with a limited budget. If your organization is affiliated with an educational institution, then a special educational price may be available. Although ESRI GIS software products and file formats are probably the most common for manipulating and distributing GIS data, most other GIS software applications can translate data from and to these file types.

GPS EQUIPMENT SELECTION

Recent advances, refinements, and expansion of GPS technology have provided a broad array of choices to users. The Global Positioning System was developed by the U.S. Department of Defense for military applications and consists of a number of continuously orbiting satellites that transmit low power radio signals.

Ground-based receivers can use these signals to calculate a location on the surface of the Earth with a high degree of accuracy and precision. In general, obtaining higher degrees of accuracy requires the use of more complex, and therefore more expensive, equipment. The type of equipment selected depends on a number of considerations, including the degree of accuracy required by the user, budget considerations, ease of use, and working conditions (e.g., is waterproof equipment required?). The issue of instrument accuracy is one that appears to cause some concern among new users. Many users seem to assume that an inexpensive ($100-$200) handheld unit is not able to deliver the necessary accuracy and precision. The fact is that even inexpensive units are capable of attaining good accuracies.

This has not always been the case. Prior to May, 2000, the horizontal accuracy of locations from non-corrected data obtained using the Global Positioning System was limited to at least 100 m. This limited accuracy was due to the effects of selective availability (i.e., artificial signal degradation) applied by the US Government. On May 1, 2000, selective availability was turned off by Presidential order, and greatly improved accuracies are now possible from even inexpensive handheld GPS units. Studies have demonstrated that accuracies on the order of 10-20 m can be obtained from typical stand-alone GPS units.

If accuracies less than 10 meters are needed, then differential correction of data (DGPS) is required. There are several ways of obtaining differential corrections. One method requires a base station receiver or beacon placed at a known location, which then transmits corrections in real time to a roving receiver via a ground-or satellite-based radio signal. Another method is to obtain pre-recorded correction files for post processing. Files can be obtained from commercial and governmental agencies. The increased accuracy of DGPS data comes at an increased cost, and the user can expect to pay significantly more for equipment. A good overview of GPS can be found on The Geographer's Craft Web site, developed by the University of Colorado at Boulder.

Things to Keep in Mind

From our collective experience in using and teaching GIS and remote sensing for agricultural sciences, landscape design, urban forestry, geology, and Extension services, we have prepared a list of guidelines that may be helpful when considering or implementing spatial technologies for your program.

- Educate Yourself: GIS GPS RS technologies have rapidly become more accessible, less expensive, and more sophisticated. As a result of the relatively fast evolution of geospatial technologies, many professionals may either be unaware of their capabilities or may have an obsolete understanding of their potential and current implementation. It is important for potential users to educate themselves before investing in equipment and software.
- Clarify Your Needs: Make sure there is a clear need for GIS-GPS-RS technologies. Lack of understanding can lead users to overestimate the usefulness of geospatial technologies. Using these technologies requires a broad understanding of many different concepts, including map projections and coordinate systems, data types and formats, computer literacy, and proper documentation of data. If all you require is the ability to make maps or locate features, and don't need sophisticated spatial analysis capabilities, then you may not need a full-

featured GIS package. In many instances, conventional methods of data collection, analysis, and presentation are more appropriate and efficient

- Know Your Users: Carefully consider the needs of the intended users. Do you need technical support for your own staff, or do you want to create deliverables for your clientele? Do the intended users have a high or low degree of technical savvy? Are they teachable or not? Applications should be kept as simple as possible for your needs.
- Be Realistic in Your Expectations: It has been our own experience that it is impractical to expect all members of your staff or faculty to learn to use GIS GPS RS technologies. Workshops we have held for this purpose have been poorly attended, despite enthusiasm expressed by the would be attendees. Before investing in infrastructure, it may be wise to consider if your work could be farmed out to a consulting agency. There are now many commercial and independent contractors doing geospatial consulting work. In the long run, hiring a consultant may be more cost efficient.
- Maintain Spatial Integrity of Your Data: One of the most frustrating aspects of working with geospatial data is dealing with different geographic coordinate systems and map projections. Because the Earth is not a perfect spheroid, numerous different projection systems have been devised to transfer points from an irregular curved surface to a plane surface. Different projections and coordinate systems are used for different purposes. For example, the State Plane coordinate system is used for many surveying applications, whereas a Transverse Mercator projection is useful for showing equatorial and mid-latitiutde continental regions. When data are stored and distributed in different projections, they must be reprojected so that all layers will plot in the same coordinate space. It is extremely important to carefully keep track of both the original and reprojected systems.
- Document Your Data: Developing metadata documentation of your spatial data cannot be emphasized

enough. Without proper documentation of coordinate and projection systems your data may be useless to both you and others. A commonly used method for preparing metadata documentation has been developed by the Federal Geographic Data Committee (FGDC) and is described in the Content Standard for Digital Geospatial Metadata. It is very important that users of GIS data understand the documentation procedure *before* using and creating data.

- Organization Is Key: Another important point to keep in mind before establishing and working with a GIS database is that your data can quickly become very disorganized. You will find that you will accumulate a large number of files as well as different versions of the same data (for example, in different projections systems). It is vitally important to establish a system for organizing data from the beginning.

6

Technology of Remote Sensing and Agriculture Resources

Agriculture resources are among the most important renewable, dynamic natural resources. Comprehensive, reliable and timely information on agricultural resources is very much necessary for a country like India whose mainstay of the economy is agriculture. Agriculture survey are presently conducted throughtout the nation in order to gather information and associated statistics on crops, rangeland, livestock and other related agricultural resources. These information of data are most importance for the implementation of effective management decisions at local, panchayat and district levels. In fact, agricultural survey is a backbone of planning and allocation of the limited resources to different sectors of the economy. With increasing population pressure throughout the nation and the concomitant need for increased agricultural production (food and fiber crops as well as livestock) there is a definite need for improved management of the nation agricultural resources. In order to accomplish this, it is first necessary to obtain reliable data on not only the types, but also the quality, quantity and location of these resources. The remote sensing techniques has been and it will continue to, a very important factor in the improvement of the present systems of acquiring and generating agricultural data.

REMOTE SENSING AND ITS IMPORTANCE IN AGRICULTURAL SURVEY

Remote sensing is nothing but a means to get the reliable information about an object without being in physical contact with the object. It is on the observation of an object by a device separated from it by some distance utilizing the characteristics response of different objects to emissions in the electromagnetic energy is measured in a number of spectral bands for the purpose of identification of the object.

In such study single tabular form of data or map data is not sufficient enough which can provide can be, combined with information's obtained from existing maps and tabular data.

- Remote Sensing techniques using various plate form has provide its utility in agricultural survey.
- Satellite data provides the actual synoptic view of large are at a time, which is not possible from conventional survey methods.
- The process of data acquisition and analysis is very fast through Geographic Information System (GIS) as compared to conventional methods.

Remote Sensing techniques have a unique capability of recording data in visible as well as invisible (i.e. ultraviolet, reflected infrared, thermal infrared and microwave etc.) part of electromagnetic spectrum. Therefore certain phenomenon, which cannot be seen by human eye, can be observed through remote sensing techniques i.e. the trees, which are affected by disease, or insect attack can be detected by remote sensing techniques much before human eyes see them.

Present system of Generating Agricultural Data and its Problems

The present system of agricultural data is collected throughout the nation. The main responsibility of collection agricultural survey lies on the Director of Land Records, Director of agriculture and District Statistical Office under the Ministry of Agriculture. These data are collected not only on a local but also some extent of district and state level. The associate of agricultural survey on crops (crop production, type of crop and crop yield), range land (condition of range, forest type, water quality, types of irrigation system and soil characteristics) and livestock (livestock population, sex of animal, types of farm and distribution of animals).

The basic problems in this survey are;

- Reliability of data
- Cost and benefits
- Timeless
- Incomplete sample frame and sample size
- Methods of selection
- Measurement of area
- Non sampling errors
- Gap in geographical coverage
- Non availability of statistics at disaggregated level

Remote Sensing techniques make it use before the remote sensing data may provide solution to these particular problems of agricultural survey.

Remote Sensing Techniques for Agricultural Survey

The given factors influenced the use of remote sensing in agricultural surveys; via 1. Characteristics of the agricultural landscape 2. Characteristic of EMR on Agricultural survey. Detection, identification, measurement and monitoring of agricultural phenomena are predicated on the assumption that agricultural landscape features (e.g. crops, livestock, crop infestations and soil anomalies) have consistently identifiable signatures on the type of remote sensing data.

Some of the parameters which may cause these identifiable signatures include crop type, state of maturity, crop density, crop geometry, crop vigor, crop moisture, crop temperature, soil moisture, soil temperature. An image analysis can correlate a certain signature with one of these many characteristics. Remote Sensing techniques in agriculture survey which affect the signature on remote sensing imagery. It is important to consider briefly the significance of choosing the appropriate sensor system, as well as the scale and resolution requirements that will yield optimum benefits for objectives of agricultural survey.

SIGNATURE IN REMOTE SENSING

The knowledge of spectral signatures is essential for exploiting the potential of remote sensing techniques. This knowledge enables one to identify and classify the objects of agricultural resources. It is also required for interpretation of all remotely sensed data, especially in agricultural resource data whether the interpretation is carried out visually or using digital techniques. It also helps us in specifying requirements for any remote sensing mission e.g. which optimal wave length bands to be used or which type of sensor will be best suited for a particular task (agricultural survey). All objects of agricultural resource on the surface of the earth have characteristic spectral signatures. The average spectral reflectance curves (or) spectral signatures for three typical earth's features; vegetation, soil and water.

The spectral reflectance curves for vigorous vegetation manifests the "Peak-and valley" configuration. The valleys in the visible portion of the spectrum are indicative of pigments in plant leaves. Dips in reflectance that can be sent at wavelengths of 0.65 mm, 1.4 mm and 1.9 mm are attributable to absorption of water by leaves. The soil curves show a more regular variation of reflectance. Factors that evidently affect soil reflectance are moisture content, soil texture, surface roughness and presence of organic matter. The water curves shows that from about 0.5 mm, reduction in reflectance with increasing wavelength, so that in the near infrared range, the reflectance of deep clear water is virtually zero (Mather, 1987) However, the spectral reflectance of water is significantly affected by the presence of dissolved and suspended organic and inorganic material and by the depth of the water body. Determinations of spectral signatures implies basic under standing of interaction of electromagnetic radiation with agricultural resources objects. This is also necessary for analysing and designing sensor systems for agricultural survey.

Sensor systems in Remote Sensing

In remote sensing the acquisition of data is depending upon the sensor system used. Various remote sensing platforms (Aircraft, Satellite) are equipped with different sensor systems. Sensor is a device that receives electromagnetic radiation, converts it into a signal and presents it in a form suitable of obtaining information about the land or earth resource as used by an information gathering system. Sensor can be grouped, either on the basic of energy source. They are as classified.

Active sensor: An active sensor operates by emitting its own energy, which is needed to detect the various phenomena (e.g. RADAR, camera with a flash gun).

Passive sensor: The operation of passive sensor is dependent of the existing sources of energy, like sun (e.g. photographic systems, multispectral scanners).

The given sensor system of camera are in agricultural survey.

Photographic Cameras

The photographic system, having conventional camera with black and white photography, is the oldest and probably, so far, the most widely used sensor for recording information about ground object. Photographic cameras have been successfully used in aircraft platform remote sensing. In this system, the information is limited to size and shape, as the films used are sensitive only to visible region of spectrum. The response of black & white films is about 0.4-0.7 mm for infrared imagery, films with response extending up to 0.9 mm are available.

Return Beam Vidicon (RBV)

This is very similar to a television camera. In such a system, a fixed camera lens on a photosensitive semi-transparent sheet forms the ground image. This image is created on the surface as electrical change or potential. The TV cameras are the best example of high resolution, operated in space for resource survey was the RBV used in LAND SAT series. On LAND SAT I, II and III RBV cameras were used, each corresponding to a different wavelength band 0.475-0.585 mm (green), 0.580-0.690 mm (red) and 0.690-0.830 mm (near infrared). The Indian experimental remote sensing satellite, Bhaskara-I and II carried a two-band TV camera system, Multispectral imagery was produced in LAND SAT and Bhaskara by using separate camera tubes of each band and selecting the spectral band with appropriate filters.

Optical-mechanical Scanners

This imaging system has the advantage that any set of desired spectral bands can be selected with appropriate filter and detector combinations. The mostly widely used sensor in this category is the MSS on LAND SAT series. MSS has four spectral bands, covering form 0.5-to 1.1 mm region. MSS operates on the principle of scanning successive lines at right angles to the flight path by means of a rotation or oscillating optical system. The radiation levels along the lines are recorded by appropriate sensor elements. When used in the visible band, the collected light can be split by the optics and separately filtered and recorded, giving simultaneous multispectral recording from the one instrument. MSS can record in any part of ultraviolet to near IR window. They are use also in the thermal IR windows.

Radar and Microwave Sensors

The acquisition of data in microwave region has been possible since 1950s but its application to natural resources is considerably less developed, as compared to the visible and IR image interpretations. Microwave sensors have distinct advantages because they are unaffected by atmospheric conditions and are thus able to penetrated smoke, clouds, haze and snow. Under this system, Plan Position Indicator (PPI), Side Looking Air borne Radar (SLAR) and Synthetic Aperture Radar (SAR) can be grouped. These systems offer day and night as well as all weather capability and ability to penetrate a cover of vegetation.

Advance Remote Sensors

Linear Imaging and Self Scanning Sensors (LISS) are the advanced imaging systems. This type of scanning sensor are used an array of solid-state devices. The array may be made of photo-diodes, phototransistors or Charge-Coupled Devices (CCDs). In the LISS, the optics focuses a strip of terrain in the cross-track into the sensor array. The image from each detector is stored and shifted out sequentially to receive a video signal. The SPOT (Satellite Probatorie d' Observation de la Terra) and IRS (Indian Remote Sensing Satellite) series carry such solid-state sensor systems, which are also known as push-broom scanners. The IRS IC most advanced satellite, carries an improved sensor system. Besides carrying a sophisticated LISS-III camera, it has a Panchromatic camera (PAN) and a Wide Field Sensor (WiFS). The PAN has been designed to provide data with a spatial resolution of 5.8m in stereo mode, with a ground swath of 70km, whereas WiFS provides data in two spectral bands, with a spectral resolution of 188m and a ground swath of 180km.

ELECTROMAGNETIC REMOTE SENSING PROCESS

Agricultural resources data are collected y aircraft and satellite-mounted instruments, which receive reflected energy from target in some frequency of the electromagnetic spectrum. The process involved in electromagnetic remote sensing system namely, data acquisition and data analysis are outlined below and a schematic diagram of electromagnetic remote sensing process.

Data Acquisition

The data acquisition process comprises of the following distinct elements, which are necessary in agricultural survey

- Energy sources
- Propagation of energy through the atmosphere
- Energy interactions with earth's surface features
- Air borne/Space borne sensors to record the reflected energy
- Generation of sensor data in the form of pictures or digital information

Data Analysis

The data analysis process involves examining the data using various viewing instrument to analyse pictorial data, which is called the visual interpretation technique and computer to analyse digital data, a process known as digital analysis.

Reference Data

Reference data also called ground truth area an essential part of remote sensing data processing. It is used to analyse and interpret remotely sensed data, to calibrate a sensor, and to verify information extracted from remote sensing data. The spectrum relevant to remote sensing. The most common form of remote sensing was Aerial photography, in which used the visible light section of the electromagnetic spectrum. Newer sensors can acquire data in this and other sections of the electromagnetic spectrum, such as the non-visible infrared and near infrared wavelength, as will as microwaves used for radar. Many of these sensors can acquire several section of the spectrum concurrently and may be termed multispectral scanners. The electromagnetic wavelength bands with their bands with their utility in remote sensing.

ADVANTAGES OF REMOTE SENSING TECHNIQUES IN AGRICULTURAL SURVEY

With the primary aim of improving the present means of generating agricultural data, a number of specific advantages may result form the use of remote sensing techniques. Vantage point. Because the agricultural landscape depends upon the sun as a source of energy, it is exposed to the aerial view and, consequently, is ideally suited or remote sensing techniques.

Coverage. With the use of high-altitude sensor platforms, it is now possible to record extensive areas on a single image. The advent of high-flying aircraft and satellites, single high quality images covering thousand of square miles.

Permanent record. After an image is obtained, it serves as a permanent record of a landscape at a point in time which agriculture changes can be monitored and evaluated.

Mapping Base. Certain types of remote sensing imagery are, in essence, pictorial maps of the landscape and after rectification (if needed), allow for precise measurement (such as field acreages) to be made on the imagery, obviating time-consuming on the ground surveys.

These images may also aid ground data sampling by serving as a base map for location agriculture features while in the field, and also as a base for the selection of ground sampling point or areas.

Cost savings. The costs are relatively small when compared with the benefits, which can be obtained form interpretation of satellite imagery.

Real-time capability. The rapidly with which imagery can be obtained and interpreted may help to eliminate the lock of timeliness which plagues, so many agricultural survey.

Other advantages of Remote Sensing:

- Easy data acquisition over inaccessible area

- Data acquisition at different scales and resolutions
- The images are analysed in the laboratory, thus reducing the amount of fieldwork
- Colour composites can be produced from three individual band images, which provide better details of the area then a single band image or aerial photograph
- Stereo-satellite data may be used for three-dimensional studies. At present, all advantages listed above have been demonstrated either operationally or experimentally

APPLICATION OF REMOTE SENSING TECHNIQUES FOR AGRICULTURAL SURVEY

The specific application of remote sensing techniques can be used for i) detection ii) identification iii) measurement iv) monitoring of agricultural phenomena.

Area of Specific Applications

a) Applicable to crop survey

1. Crop identification
2. Crop acreage
3. Crop vigor
4. Crop density
5. Crop maturity
6. Growth rates
7. Yield forecasting
8. Actual yield
9. Soil fertility
10. Effects of fertilizes
11. Soil toxicity
12. Soil moisture
13. Water quality
14. Irrigation requirement
15. Insect infestations
16. Disease infestations
17. Water availability
18. Location of canals

b) Applicable to range survey:

1. Delineation of forest types
2. Condition of range
3. Carrying capacity
4. Forage
5. Time of seasonal change
6. Location of water
7. Water quality
8. Soil fertility
9. Soil moisture
10. Insect infestations
11. Wildlife inventory

c) Applicable to livestock survey:

1. Cattle population
2. Sheep population
3. Pig population
4. Poultry Population
5. Age sex distribution
6. Distribution of animals
7. Animal behavior
8. Disease identification
9. Types of farm buildings

The use of remote sensing technology has been rapidly expanded for the development of key sectors. This paper highlights the fact that the remote sensing techniques will continue to be very important factor in the improvement of present system of acquiring agricultural data. The remote sensing provides various platforms for agricultural survey. Satellite imagery has unique ability to provide the actual synoptic vies of large area at a time, which is not possible for conventional survey methods and also the process of data acquisition and analysis are very fast through GIS (Geographic Information System) as compared to the conventional methods. The different features of agriculture are acquired by characteristic, spectral reflectance, spectral signature of agriculture and associated phenomena through EMR. In General the research paper emphasizes the utmost need of timeliness and accuracy of the output generated by remote sensing techniques and its calibration with ground-truth and other information systems like aerial photography and satellite imagery etc. Further, the importance of remote sensing with special reference to agricultural sector involving crop acreage, crop production, rangeland and livestock.

SPACE INPUTS FOR PRECISION AGRICULTURE

Precision agriculture, newly emerging agricultural management concept, embodies the convergence of biotechnologies & other agricultural technologies with space and informatics. With a goal to achieve the quantum jump in agricultural productivity, reduced cost of cultivation, diversified and resilient agricultural systems, the precision agriculture plays catalytic role in order to achieve a common ground between environmental and economic goals. It is basically designed to optimize agricultural inputs viz., fertilizers, pesticides, water etc, in tune with micro-level/field requirements. Optimization is focused on increased yields, reduced cost of cultivation and to minimized environmental impacts through location-specific management.

The success stories pertaining to Precision Agriculture have mainly drawn from the developed countries; wherein Agriculture is characterized by highly mechanized and automated systems, and is driven by market forces and has been professionally managed enterprise. Taking into account the predominance of fragmented land holdings, heterogeneity of crops and livestock and concept of farm families in the rural conditions, the model of Precision Agriculture representing the typical Indian Agricultural scenario is yet to evolve. While the ecological integrity of farming systems is an imperative need, it is equally important to extend the access of information and market to the small and marginalized farmers. The Precision Agriculture model for India while addressing these issues should provides an innovative route for sustainable agriculture in globalize and liberalized economy.

TECHNOLOGY ELEMENTS

Recent advances in technology for variable rate technology (VRT), with concurrent advances in remote sensing, GIS &

GPS, and the developments taken place in crop simulation modelling, have provided enough opportunities and scope to take-up proto-type Precision Agriculture experiments. The VRT applies production inputs at rates appropriate to soil and crop conditions within micro-level field conditions. The VRT systems have been demonstrated for several materials, including herbicides, fertilizers, insecticides and seeds.

Role of space technologies becomes more crucial in order to address the spatial variability of soils and crops across the various scales of mapping. The space technology inputs also capture the vulnerability and dynamism of agricultural systems. The developments in space-borne imaging sensors, particularly their spatial, spectral and temporal resolutions are well characterized to capture these features. While high spatial resolution images enable mapping and monitoring the structural attributes of agro-ecosystems, high spectral resolution or hyper-spectral imaging addresses their functionalities. The high temporal resolution captures the dynamisms of agro-ecosystems.

The use of remote sensing, GIS and GPS for crop monitoring, condition assessment and yield modelling has already been well established. Crop simulation models (CSMs) provide potential production under the different scenario of constraints, including weather, soils, crops, cultural practices etc. The conjunctive use of VTR, remote sensing, GIS, GPS and CSMs provides technological framework for Precision Agriculture.

SPACE APPLICATIONS TO AGRICULTURE

Agriculture has been at the top of our priorities for space applications. The road map of agricultural applications started with the first remote sensing experiment on identification of coconut root-wilt diseases in Kerala using Infrared aerial photography-way back in 1969-70. Since then, space applications to agriculture sector have touched almost all the segments of agricultural ecosystems. These include the mapping and monitoring of major crops, soils/degraded lands, command areas, wastelands, surface & ground water, floods & drought, and watersheds.

Agricultural statistics that provides the vital informatics base to agriculture sector originates from age-old village level Patwari system. It often moves upward with its inherent subjectivity and bias. There are hundreds of crucial decisions at different levels, which are taken purely based on this. In order to strengthen the foundations of agricultural statistics in the country, a remote sensing based the nationwide mission called pre-harvest Crop Acreage and Production Estimation (CAPE) was launched in late 80s.

Covering all the major cereals, pulses and oilseeds, CAPE provides in-season crop statistics with 90/90 accuracy at state level. The CAPE could thus provide the scientific basis of agricultural statistics and is transitioning further to yet another institutional destination entitled Forecasting Agricultural Output using Space Agro-meteorology and Land-based Observations (FASAL) within Ministry of Agriculture itself. Synthesizing the state-of-the-art in econometrics, agricultural meteorology and remote sensing based modelling, FASAL envisages the multiple productions forecasting of the major crops with improved accuracy to the extent of 95/95 criteria, timeliness and scope in terms of covering the whole country. The FASAL could be used to provide scientific solutions facilitating crop insurance, bridging the gaps between crop production and post-harvest technology, pricing and policy decisions.

Extensions of irrigation, genetically improved crops and use of inorganic fertilizers have been the harbingers of Green Revolution in the country. Those multi-purpose irrigation commands, which accelerated the irrigation networks in 70s & 80s, are now characterized by depleted irrigation efficiency, water logging and salinity.

At the behest of Ministry of Water Resources, remote sensing based command areas inventory has been taken up in the priority irrigation commands to enhance the water use efficiency. In the similar line, a coarse scale land degradation mapping was carried out at behest of Ministry of Agriculture. It is important to mention here that in some of critically affected areas of Gangetic plains of Uttar Pradesh, land degradation to the extent of village level has been mapped out. Based on this, several land reclamation efforts-including the World Bank supported programmes, have been taken up to restore the health of soils. In coming years, with the support from Ministry of Agriculture, ISRO/DOS is planning to launch Nationwide

Land Degradation Mapping Mission at the scale 1: 50,000, so as to delineate land degradation at Village level.

By virtue of the unique combinations of IRS satellites with varied resolutions and capabilities, the space segment in the country is in a position to monitor agricultural drought and recurrent floods more efficiently. Capturing the flood events with the damages to the Villages and crops has formed the scientific basis for damage assessment and thereby relief operations at various levels. This year, for example, flood inundation maps in Orissa were generated depicting the marooned areas during the different flood waves and disseminated to the user community within hours through the Internet. Space inputs to agricultural drought monitoring is yet another aspect to target the districts and taluks on the basis of severity of drought.

Department of Space at the behest of Ministry of Rural Development has carried out nation-wide wasteland mapping on 1: 50,000 scale using the IRS data. The Wasteland Atlas of India covering the wasteland statistics of entire country has been brought out. These maps would help in retrieving the information at village/watershed (500 ha) levels, for implementation of wastelands/watershed development programmes. The uniqueness of wasteland mapping has also been the creation of digital database with village as well as watershed boundaries.

With the hydro-geomorphological maps prepared under the National Drinking Water Mission, and subsequently to the Rajiv Gandhi National Drinking Water Mission, satellite remote sensing has really made a dent in reaching its potentials down the line to grassroots. Search for groundwater, particularly in areas with consolidated and semi-consolidated rock formations, considered more difficult from the point of view of exploration as well as recharging of groundwater, is considerably aided by the use of the hydro-geomorphological maps. These maps are extensively used for locating prospective groundwater sites around problem habitations in the country, as a part of the 'scientific-source finding'.

At the behest of the Planning Commission, scientific inputs were generated from remote sensing on land use/land cover of entire country to evolve agro-climatic zonation based planning. These applications have moved further to large scale mapping across the problem regions of the country (especially the rain-fed regions), towards micro-level decision making for sustainable development. Under remote sensing based Integrated Mission for Sustainable Development Mission (IMSD), land and water resources development plans were carried out for about 80 mha. in 175 districts of the country. Implementation of these plans in certain areas has enriched the ground water potential, increased the cropping intensity along with the net returns from the fields. It is important to note that these experiments were conducted in the terrain of low productivity, mainly dry lands/rain-fed areas, predominantly wastelands, where the scope of doubling food production is the main issue. Emphasis is being laid in operational utilization of this vast database for ground level implementation by concerned user agencies.

The success of these experiments lies not only due to the applications of space inputs alone, but largely due to the synergistic interfaces among grass root level farmers, administrators/policy makers and multi-disciplinary scientific communities. With operationalisation of the National Natural Resources Management System (NNRMS) in the country, with DOS as the nodal agency, the space, ground and user segments have been properly tuned to respond to the challenges of sustainable crop production in the country.

The INSAT-VHRR observations on clouds, cyclone depressions, surface radiance and monsoon parameters in the Indian Ocean have really strengthened the agro-meteorological services in the country. To disseminate the appropriate locale-specific agricultural packages, there are several programmes such as 'Krishi Darshan' being beamed through all the regional channels. An operational system of public instruction has been established since 1995, using the INSAT based Training and Development Communication Channel (TDCC) for disseminating the improved agricultural practices, training primary school teachers, Panchayat Raj, elected representatives, Anganwadi workers, wasteland development functionaries etc. The Jhabua Development Communication Project (JDCP), undertaken by ISRO in Jhabua district of MP, has been providing communication support to the developmental activities and also interactive training to the development functionaries, besides empowering the poor and marginalised tribal.

A pilot project sponsored by the Planning Commission, a single window information service provider concept-called Agro-climatic Planning and Information Bank (APIB) has been developed for the farmers of drought prone areas of Karnataka. The APIB is an effort to empower the farmers and reduce their vulnerability to crop failures. Moving to the farmer's cooperatives, ISRO has taken up jointly with IFFCO-a major initiative on developing remote sensing and GIS based decision support system for fertilizer movements in the priority areas. Finally, space applications over last three decades in agriculture have made this sector information rich and have addressed the various issues encompassing the different sectors of agriculture.

Indian Precision Agriculture

Agriculture in India, as we see today, is at the crossroads.

1. On the one hand there are depletions of ecological foundations of the agro-ecosystems, as reflected in terms of increasing land degradation, depletion of water resources and rising trends of floods, drought and crop pests and diseases. On the other hand, there is imperative socio-economic need to have enhanced productivity per units of land, water and time.
2. At present, 3 ha of rain fed areas produce cereal grain equivalent to that produced in 1 ha. of irrigated. Out of 142 million ha. Net sown areas, 92 million ha. are under rain-fed agriculture in the county.
3. From equity point of view, even the record agricultural production of more than 200 Mt is unable to address food security issue. A close to 60 Mt food grains in the storehouses of Food Corporation of India (FCI) is beyond the affordability and access to the poor and marginalized in many pockets of the country.
4. Globally, there are challenges arising from the Globalization especially the impact of WTO regime on small and marginalized farmers.
5. Some other unforeseen challenges could be anticipated global warming scenario and its possible impact on

diverse agro-ecosystems in terms of alterations in traditional crop belts, micro-level perturbations in hydrologic cycle and more uncertain crop-weather interactions etc.

At this stage, agriculture needs new paradigms to deal with the present situation. The strategy lies in integration of the dynamic information and scientific knowledge into the management of agro-ecosystems, and thereby optimizing the radiation, water and nutrient usages. Agriculture has to transition from high inputs material inputs to the optimum level, through the appropriate use of information, knowledge and strategies for efficient resource usages. In such case, productivity of agricultural may not be the function of the quantum of agricultural input use alone, but will include information, knowledge and efficiency while managing the agricultural practices.

With the fragmentations of land holdings and predominance of small and marginal farmers, our agricultural systems are basically dis-aggregated farm families. Fundamentally, Precision Agriculture aims at a dis-aggregated micro-level farm management strategy with intense information inputs addressing the variations of soils, crops, water, chemicals, market access etc.

Taking into the present state of Agriculture in the country, Precision Agriculture is absolutely essential in order to address poverty alleviation and food security to a very large cross-section of the population. For example, at present, the average Orissa farmer produces slightly more than one ton of rice per hectare, and keeps little below one tone for his family. But if he can produce four tons, then he has three tons to sell, and more cash in hand. The smaller the farm, the greater the need for a marketable surplus.

Defining Precision Agriculture in Indian Context Taking into account the typical characteristics of Indian Agriculture, the definition of Precision Agriculture, therefore, must encompass the strategy and framework to achieve higher productivity, reduced cost of cultivation by optimization of inputs, and diversified & resilient agricultural systems. All these goals are to be achieved within the typical constraints of India's agro-ecosystems. While the depletion of ecological foundations of farming systems needs to be arrested, the access of information, credit, agricultural inputs and market to the small and marginal farmers are equally crucial. The Precision Agriculture model for must be derive encompassing all these issues, within a broad framework of addressing the negatives of globalization and achieving sustainable agriculture in the long run.

The definition of Precision Agriculture has to be micro-level contextual and should capture the local variabilities, vulnerabilities and dynamisms of agro-ecosystems. For example, Agriculture in Punjab, Haryana & Western UP is characterized by higher productivity (about 4 t/ha), higher use of inputs (irrigation ~ 96%; fertilizer consumption 0.158 t/ha), higher cropping intensities and predominance of medium and large farmers.

Sustainability of such agro-ecosystems has been a cause of concern in the recent times. The model for Precision Agriculture for such ecosystems may focus on sustainability through the optimization of agricultural inputs and thereby reduction of cost of cultivation, and cropping system analysis.

Agriculture in Southern part of the country-especially Andhra and Tamilnadu is moderate yield (around 2 t/ha), inputs (irrigation ~ 55%; fertilizer consumption 0.12 t/ha) and cropping system based, with predominance of small and marginal farmers. Precision Agriculture for such system should aim at enhancing the productivity based on in situ soil & water conservation.

The Eastern India especially Orissa, though high potential agricultural system is unfortunately characterized by low inputs (irrigation ~ 25%, fertilizer use 0.025 t/ha), low yield (~1.2t/ha) and low cropping intensity, with predominance of marginal and small farmers. A model for Precision Agriculture for Orissa could focus more on poverty alleviation and food security through enhanced productivity and low cost of cultivation. The models for Precision Agriculture in Indian context could therefore be the micro-level and contextual-addressing the empirical issues in the diverse agro-ecosystems of the country. The access and outreach of best practices in Precision Agriculture-as demonstrated by contextual models, to the large cross-sections of the farming community would then help in cultivating the new paradigms in age-old agriculture enterprise in the country.

Precision Agriculture models are not complete, unless the parameters related to empowerment of the farmers; especially small and marginal farmers are integrated. In this context, ISRO has also initiated Gramsat project in Orissa. In the line of JDCP, the Gramsat project aims at empowering the people especially the poor and marginalized, by awareness building and access to information and services. Towards this, a network of one-way video and two-way audio Village Information Kiosks is being developed in the selected blocks of Orissa. The same networks are also planned to facilitate e-governance in the region. Precision Agriculture model should present a synergy that could lead to a holistic mission, focused on agricultural development with the backdrop of present issues and challenges.

7

Precision Farming: Dreams and Realities for Indian Agriculture

Ever since the man appeared on the earth, he has been harnessing the natural resources to meet his basic requirements. Reference to soil, water and air as basic resources, their management and means to keep them pure are mentioned in the Vedas, Upanishads and in ancient Hindu literature. The phenomenal increase in population of both man and animal in the last century and fast growing industrialization and urbanization in last few decades have overstrained the natural resource base, which are getting degraded much faster than ever before. Thus, the attention of whole world is focused on how to increase production to feed the burgeoning population and the question uppermost in every ones mind is “Can we produce enough food in a sustainable manner without damage to the natural resource base?”

Belying all predictions made to the contrary, India could achieve unprecedented increase in the food grain production as a result of expansion of irrigation and technological advancement in agriculture. While it has been a satisfying experience, Indian agriculture would need a new vision to make rapid progress in the ensuing millennium. To achieve the required growth will not be easy as some of the existing production systems are based on unsustainable use of the resources. The signs of fatigue in the natural resources have already appeared which is a cause for serious concern to the planners, decision-makers and researchers alike.

Precision Agriculture is the application of technologies and principles to manage spatial and temporal variability associated with all aspects of agricultural production for improving production and environmental quality. The success in precision agriculture depends on the accurate assessment of the variability, its management and evaluation in space-time continuum in crop production. The agronomic feasibility of precision agriculture has been intuitive, depending largely on the application of traditional arrangement recommendations at finer scales. The agronomic success of precision agriculture has been quite convincing in crops like sugar beet, sugarcane, tea and coffee. The potential for economic, environmental and social benefits of precision agriculture is largely unrealized because the space-time continuum of crop production has not been adequately addressed.

Successful implementation of precision agriculture depends on numerous factors, including the extent to which conditions within a field are known and manage, the adequacy of input recommendation and the degree of application control. The enabling technologies of precision agriculture can be grouped in to fine major categories: Computers, Global Positioning System (GPS), Geographic Information System (GIS), Remote Sensing (RS) and Application control. Aspects of precision agriculture encompass a broad array of topics including variability of the soil resource base, weather, plant genetics, crop diversity, machinery performance and most physical, chemical and biological inputs used in crop production. Precision agriculture must fit the needs and capabilities of the farmer and must be profitable.

CONTENT ISSUES AND THE GATEWAY DEBATE

Despite all the natural advantages, India's productivity of food grains per hectare is no more than three-fourths of the world average and less than half of that in agriculturally advanced countries, per capita food grain availability even after the Green Revolution, has been less than two thirds of the world average. Only five states in India, namely Himachal Pradesh, Punjab, Haryana, Uttar Pradesh and Madhya Pradesh – produce more grain than their populations can consume. The combined population of the five states is less than one-third of the total of the country. More than two thirds of the population lives in states that are still food-deficit. This requires transport of lakhs of tonnes of food grain, involving high costs and pilferage. Our effort should have been to make all the states self-sufficient with respect to food grains and if some disturbances occurred due to unnoticed natural calamities the nation must be in an ever ready position to mitigate such challenging tasks.

The Indian green revolution is also associated with negative ecological/environmental consequences. The status of Indian environment shows that, in India, about 182 million ha of the country's total geographical area of 328.7 million ha is affected by land degradation of this 141.33 million ha are due to water erosion, 11.50 million ha due to wind erosion and 12.63 and 13.24 million ha due to water logging and chemical deterioration (salinization and loss of nutrients) respectively. On the other end, India shares 17 per cent of world's population with 2.5 per cent of geographical area, 1 per cent of gross world product, 4 per cent of world carbon emission and hardly 2 per cent of world forest area. The Indian status on environment is though not alarming when compared to developed countries, it gives an early warning to take appropriate precautionary measures. The growth rate of grain production during the Ninth plan has been less than the population growth rate. The poor agricultural performance has not been because of the vagaries of monsoon. From 1997-98 to 2002-03, rainfall was between 92 and 106 per cent of the normal. Per capita availability of grain and per capita calorie intake, which were less than the minimum required for adequate nutrition, have further declined. According to Human Development Report 2003, the percentage of the undernourished in India, which was 21 a few years ago, has now reached 24.

The Government claims that India has emerged as the seventh largest exporter of food grains in the world. This is nothing to be proud of, if we take into account that the total Indian grain export in 2002-03 did not add up to even 4 per cent of the total world exports; and the value of our grain exports did not add up to even the value of our imports of vegetable oils and pulses. The more crucial question, however, is whether it is morally justifiable to export grain when 24 per cent of the population remains under-nourished. The decline in agricultural growth and increase in rural poverty have been due to the long persisting government indifference towards the farm sector, which is evident from plan outlays on Agricultural and Allied activities, Rural Development and Irrigation, which added up to 37.1 per cent of the total during the first plan were brought down to only 19.4 per cent during the Ninth plan. However, the main reason, for poor performance of the farm sector has been the long persisting adverse terms of trade policies for agriculturists in addition to the mismanagement of natural resources leads to ever ending crisis.

THE NEED FOR PRECISION FARMING

The 'Green revolution' of 1960's has made our country self sufficient in food production. In 1947, the country produced a little over six million tonnes of wheat, in 1999; our farmers harvested over 72 million tonnes, taking the country to the second position in wheat production in the world. The production of food grains in five decades, has increased more than three fold, the yield during this period has increased more than two folds. All this has been possible due to high input application, like increase in fertilization, irrigation, pesticides, higher use of HYV's, increase in cropping intensity and increase in mechanization of agriculture.

Fatigue of Green Revolution

Green revolution of course contributed a lot. However, even with the spectacular growth in the agriculture, the productivity levels of many major crops are for below than expectation. We have not achieved even the lowest level of potential productivity of Indian high yielding varieties, whereas the worlds highest productive country have crop yield levels significantly higher than the upper limit of the potential of Indian HYV's. Even the crop yields of India's agriculturally rich state like Punjab is far below than the average yield of many high productive countries.

Natural Resource Degradation

The green revolution is also associated with negative ecological/environmental consequences. The status of Indian environment shows that, in India, about 182 million ha of the country's total geographical area of 328.7 million ha is affected by land degradation of this 141.33 million ha are due to water erosion, 11.50 million ha due to wind erosion and 12.63 and 13.24 million ha are due to water logging and chemical deterioration (salinisation and loss of nutrients) respectively. On the other end, India shares 17 per cent of world's population, 1 per cent of gross world product, 4 per cent of world carbon emission, 3.6 per cent of CO2 emission intensity and 2 per cent of world forest area. One of the major reasons for this status of environment is the population growth of 2.2 per cent in 1970 – 2000. The Indian status on environment is, though not alarming when compared to developed countries, gives an early warning.

In this context, there is a need to convert this green revolution into an evergreen revolution, which will be triggered by farming systems approach that can help to produce more from the available land, water and labour resources, without either ecological or social harm. Since precision farming, proposes to prescribe tailor made management practices, it can help to serve this purpose.

THE BASIC COMPONENTS OF PRECISION FARMING

Precision farming basically depends on measurement and understanding of variability, the main components of precision farming system must address the variability. Precision farming technology enabled, information based and decision focused, the components include, (the enabling technologies) Remote Sensing (RS), Geographical Information System (GIS), Global Positioning System (GPS), Soil Testing, Yield Monitors and Variable Rate Technology. Precision farming requires the requisition, management, analysis and output of large amount of spatial and temporal data. Mobile computing systems were needed to function on the go in farming operations because desktop systems in the farm office were not sufficient. Because precision farming is concerned with spatial and temporal variability and it is information based and decision focused. It is the spatial analysis capabilities of GIS that enable precision agriculture. GPS, DGPS has greatly enabled precision farming and of great importance to precision farming, particularly for guidance and digital evaluation modelling position accuracies at the centimeter level are possible in DGPS receivers. Accurate guidance and navigation systems will allow for farming operations at height and under unfavorable weather conditions even.

In India, we have all these technologies available and they can be implemented through agricultural training centers by giving training to agriculture officers in these technologies.

BASIC STEPS IN PRECISION FARMING

The basic steps in precision farming are,

i. Assessing variation
ii. Managing variation
iii. Evaluation

The available technologies enable us in understanding the variability and by giving site specific agronomic recommendations we can manage the variability that make precision agriculture viable. And finally evaluation must be an integral part of any precision farming system. The detailed steps involved in each process are clearly depicted in a diagramme.

Assessing Variability

Assessing variability is the critical first step in precision farming. Since it is clear that one cannot manage what one does not know. Factors and the processes that regulate or control the crop performance in terms of yield vary in space and time. Quantifying the variability of these factors and processes and determining when and where different combinations are responsible for the spatial and temporal variation in crop yield is the challenge for precision agriculture.

Techniques for assessing spatial variability are readily available and have been applied extensively in precision agriculture. The major part of precision agriculture lies in assessing to spatial variability. Techniques for assessing temporal variability also exist but the simultaneous reporting a spatial and temporal variation is rare. We need both the spatial and temporal statistics. We can observe the variability in yield of a crop in space but we cannot predict the reasons for the variability. It needs the observations at crop growth and development over the growing season, which is nothing but the temporal variation. Hence, we need both the space and time statistics to apply the precision farming techniques. But this is not common to all the variability/factor that dictate crop yield. Some variables are more produced in space rather with time, making them more conducive to current forms of precision management.

Managing Variability

Once variation is adequately assessed, farmers must match agronomic inputs to known conditions employing management recommendations. Those are site specific and use accurate applications control equipment. We can use the technology most effectively.

In site-specific variability management. We can use GPS instrument, so that the site specificity is pronounced and management will be easy and economical. While taking the soil/plant samples, we have to note the sample site coordinates and further we can use the same for management. This results in effective use of inputs and avoids any wastage and this is what we are looking for.

The potential for improved precision in soil fertility management combined with increased precision in application control make precise soil fertility management as attractive, but largely unproven alternative to uniform field management. For successful implementation, the concept of precision soil fertility management requires that within-field variability exists and is accurately identified and reliably interpreted, that variability influences crop yield, crop quality and for the environment. There fore inputs can be applied accurately.

The higher the spatial dependence of a manageable soil property, the higher the potential for precision management and the greater its potential value.

The degree of difficulty, however, increases as the temporal component of spatial variability increases. Applying this hypothesis to soil fertility would support that Phosphorus and Potassium fertility are very conducive to precision management because temporal variability is low. For N, the temporal component of variability can be larger than its spatial component, making precision N management much more difficult in some cases.

Evaluation

There are three important issues regarding precision agriculture evaluation.

- Economics
- Environment
- Technology transfer

The most important fact regarding the analysis of profitability of precision agriculture is that the value comes from the application of the data and not from the use of the technology. Potential improvements in environmental quality are often cited as a reason for using precision agriculture. Reduced agrochemical use, higher nutrient use efficiencies, increased efficiency of managed inputs and increased production of soils from degradation are frequently cited as potential benefits to the environment. Enabling technologies can make precision agriculture feasible, agronomic principles and decision rules can make it applicable and enhanced production efficiency or other forms of value can make it profitable.

The term technology transfer could imply that precision agriculture occurs when individuals or firms simply acquire and use the enabling technologies. While precision agriculture does involve the application of enabling technologies and agronomic principles to manage spatial and temporal variability, the key term is manage. Much of the attention in what is called technology transfer has focused on how to communicate with the farmer. These issues associated with the managerial capability of the operator, the spatial distribution of infrastructure and the compatibility of technology to individual farms will change radically as precision agriculture continues to develop.

Technology Transition

Precision agriculture is dependent on the existence of variability in either or both product quantity and quality. If this variability does not exist then a uniform management system is both the cheapest and most effective management strategy and precision farming is redundant. Thus, in precision farming, “Variability of production and quality equals opportunity”. Having said this, the nature of the variation is also important in determining the potential for PA in a system. For example the magnitude of the variability may be too small to be economically feasible to manage. Alternatively the variability may be highly randomized across the production system making it impossible to manage with current technology. Finally the variability may due to a constraint that is not manageable. Thus the implementation of precision farming is limited by the ability of current variable rate technology (VRT-machinery/technology that allows for differential management of a production system) to cope with the highly variable sites and the economic inability to produce returns from sites with low variability using precision farming (VRT).

PRESENT SCENARIO

Though precision farming is very much talked about in developed countries, it is still at a very nascent stage in developing countries, including India. Space Application center, ISRO, in collaboration with Central Potato Research Institute, Shimla, has initiated a study on exploring the role remote sensing for precision farming.

The study on precision agriculture has already been initiated in India, in many research institutes. Space Application Center (ISRO), Ahmedabad has started experiment in the Central Potato Research Station farm at Jalahandhar, Punjab to study the role of remote sensing in mapping the variability with respect to space and time. M S Swaminathan Research Foundation, Chennai, in collaboration with NABARD, has adopted a village in Dindigul district of Tamil Nadu for variable rate input application. Indian Agricultural Research Institute has drawn up a plan to do precision farming experiments in the institutes' farm. Project Directorate for Cropping Systems Research (PDCSR), Modipuram and Meerut (UP) in collaboration with Central Institute of Agricultural Engineering (CIAE), Bhopal also initiated variable rate input application in different cropping systems. In coming few years precision farming may help the Indian farmers to harvest the fruits of frontier technologies without compromising the quality of land.

Prospects

Precision farming, though in many cases a proven technology is still mostly restricted to developed (American and European) countries. Except for a few (Wang, 2001), there is not much literature to show the scope of its implementation in India.

We feel that, one of the major problems is the small field size. In India more than 57.8 per cent of operational holdings have size less than 1 ha. However, in the major agricultural states of Punjab, Rajastan, Haryana and Gujarat there are more than 20 per cent of agricultural lands have operational holding size of more than 4 ha. These are individual field sizes. However, when we consider contiguous field with same crop (mostly under similar management practices) the field (rather simulated field) sizes are large. Using aerial data, has found that in Patiala district of Punjab, more than 50 per cent of contiguous field sizes are larger than 15 ha. These contiguous fields can be considered a single field for the purpose of implementation of precision farming.

There is a scope of implementing precision farming for major food-grain crops such as rice, wheat, especially in the states of Punjab and Haryana. However many horticultural crops in India, which are high profit making crops, offer wide scope for precision farming.

MISCONCEPTIONS ABOUT PRECISION AGRICULTURE

There are several mistaken preconceptions about precision agriculture.

a). Precision agriculture is a cropping rather than an agricultural concept. This is due to cropping systems, in particular broad-acre cropping, being the face and driving force of PA technology. However precision farming concepts are applicable to all agricultural sectors from animals to fisheries to forestry. In fact it might be argued that precision farming concepts are more advanced in the dairy industry where the “site” becomes an individual animal, which is recorded, traced and fed individually to optimize production. These industries are just as concerned with improved productivity and quality decreased environmental impact and better risk management as the cropping industry however precision farming concepts have yet to be applied on the same scale in these areas. For example a grazer’s use of advance warning meteorological data and market predictions to estimate fodder reserves and plan livestock numbers is a form of precision farming.

b). Precision agriculture in cropping equals yield mapping. Yield mapping is a crucial step and the wealth of information farmers are able to obtain from a yield map makes them very valuable. However they are only a stepping-stone in a precision farming management system. The bigger agronomic hurdle lies in retrieving the information in the yield map and using it to improve the production system. The advance of PA adoption (usefulness) in this country is may soon be bottlenecked at this point due to the lack of decision support systems (DSS) to help agronomists and farmers understand their yield maps. Yield maps may not tell the whole story either with other data sources, e.g. crop quality and soil maps, economic indicators or weather predictions, proving further information necessary for correct agronomic interpretations.

c). Precision agriculture equals sustainable agriculture. Precision agriculture is a tool to make agriculture more sustainable however it is not the total answer. Precision farming aims at maximum production efficiency with minimum environmental impact. Currently it is the potential for improved productivity (and profitability) that is driving precision farming rather than the more serious issue of long-term sustainability. Precision farming will not fix problems such as erosion and salinity by itself although it will help to reduce the risk of these problems occurring. Sensible sustainable practices still need to be used in conjunction with precision farming.

Obstacles

There are many obstacles to adoption of precision farming in developing countries in general and India in particular. Some are common to those in other regions but the others are specific to Indian conditions are as follows.

- Culture and perceptions of the users
- Small farm size
- Lack of success stories

- Heterogeneity of cropping systems and market imperfections
- Land ownership, infrastructure and institutional constraints
- Lack of local technical expertise
- Knowledge and technical gaps
- Data availability, quality and costs

Opportunities

Despite the many obstacles listed earlier, business opportunities for precision farming technologies including GIS, GPS, RS and yield monitor systems are immense in many developing countries. The scope for funding new hardware, software and consulting industries related to precision agriculture is gradually widening. In Japan, the market in the next 5 years is estimated at about US $ 100 billion for GIS and about US $ 50 billion for GPS and RS (Srinivasan, 2001). Punjab and Haryana states in India, where farm mechanization is more common than in others, may be the first to adopt precision farming on a large scale.

Recently, the governments of certain Asian countries initiated special efforts to promote precision farming. In Japan, the Ministry of Agriculture has allocated special funds for research on remote sensing applications of precision farming. A quasi-governmental institute “Bio-oriented Technology Research Advancement Institute (BRAIN)” is also funding research on precision farming. In Malaysia, the Malaysian Agricultural Research and Development Institute (MARDI) is promoting research on precision farming of upland rice. In other countries, the private sector, which holds or leases a large acreage, is likely to adopt precision farming sooner than the small holders.

Precision farming is useful in many situations in developing countries. Rice, wheat, sugar beet, onion, potato and cotton among the field crops and apple, grape, tea, coffee and oil palm among horticultural crops are perhaps the most relevant. Some have a very high value per acre, making excellent cases for site-specific management. For all these crops, yield mapping is the first step to determine the precise locations of the highest and lowest yield areas of the field. Researchers at Kyoto University recently developed a two-row rice harvester for determining yields on a micro plot basis.

Precision farming can bring several benefits to the sugar beet industry in Hokkaido, where the marketing system was changed in 1986 from a quantity (fresh weight) to quality (sugar yield) basis. Heavy N fertilization, a common practice here, results in excessive N levels and a decreased sugar concentration. It is now possible to estimate sugar and amide N concentration in leaves using reflectance in visible bands and root yield using reflectance in visible and infrared bands. Incorporation of such data into a GIS along with precise positioning of non-uniform areas using GPS can be used to vary fertilizer dose within a field, thereby improving productivity. Although weighing conveyor technology has been known for some time, effective yield measurement still remains the main problem in crops such as sugar beet, onion and potato. Further, the preparation of product quality maps for these crops is as important as yield maps. In India, a few researchers in the private sector initiated studies on precision agriculture in high value crops like cotton, coffee and tea. In cotton, remote sensing coupled with GIS can assist in improved precision of insect pest management and harvesting. In Sri Lanka, researchers at the Tea Research Institute are examining precision management of soil organic carbon.

In so far as dairy farming in Asia is concerned, precision farming techniques can help in improving efficiency of methods, timing and rate of application of animal wastes leading to high application efficiency and low environmental pollution. While considering soil and climatic conditions. For instance, factors determining the risk of NO3 leaching, release of N2O through denitrification and contamination of surface and ground water by runoff can be mapped and analysed. Likewise, poorly managed areas in grass lands can be identified and the optimum period for cutting on a plot basis determined.

Nutrient stress management is another area where precision farming can help Indian farmers. Most cultivated soils in India are acidic and spatial variation in pH is high. Detecting nutrient stresses using remote sensing and combining data in a GIS can help in site-specific applications of fertilizers and soil amendments such as lime, manure, compost, gypsum and sulphur. This in turn would increase fertilizer use efficiency and reduce nutrient losses. In semi arid and arid tropics, precision technologies can help growers in scheduling irrigation more profitably by varying the timing, amounts and placement of water. For example, drip irrigation, coupled with information from remotely sensed stress conditions (e.g., canopy temperature) can increase the effective use of applied water from 60 to 95 per cent there by, reducing runoff from 23 to 1 per cent and deep percolation from 18 to 4 per cent.

Pests and diseases cause huge losses to Indian crops. If remote sensing can help in detecting small problem areas caused by pathogens, timing of applications of fungicides can be optimized. Recent studies in Japan show that pre-visual crop stress or incipient crop damage can be detected using radio-controlled aircraft and near-infrared narrow-band sensors. Like wise, airborne video data and GIS have been shown to effectively detect and map black fly infestations in citrus orchards, making it possible to achieve precision in pest control. Perennial weeds, which are usually position-specific and grow in concentrated areas, are also a major problem in developing countries. Remote sensing combined with GIS and GPS can help in site-specific weed management. Although through cost-benefit analysis has not been done yet, the possible use of precision technologies in managing the environmental side effects of farming and reducing pollution is appealing.

Strategies

Precision farming is still only a concept in many developing countries and strategic support from the public and private sectors is essential to promote its rapid adoption. Successful adoption, however, comprises at least three phases including exploration, analysis and execution. Data on crop yield, soil variables, weather and other characteristics are collected and mapped in the exploratory stage, which is important for increasing the awareness among farmers of long term benefits. The approaches to data collection and mapping must, therefore, reflect local needs and resources.

In the analysis stage, factors limiting the potential yield in various areas within a field and their interrelationships are examined using GIS-based statistical modelling. Sadler et al. (1998) showed that quantitatively important yield variation may occur over distances as short as 10m, however, only some factors such as soil structure, water status, pH, nutrient levels, weeds, pests and diseases can be controlled but not the others (soil texture, weather, topography). After determining the significance of each source of variability to profitability of a particular crop and relative importance of each controllable factor, management actions can be prioritized. It must be remembered that in some low yielding areas, the reason for poor yields may be the lack of sufficient soil nutrients in the first place. In such cases, application beyond just replenishment is necessary.

Lastly, execution phase includes variable application of inputs or cultural operations. In most developing countries of Asia. However, it is not always necessary and/or possible to use variable rate applicators. Efforts must, therefore, initially focus on limiting indiscriminate use of inputs in conventional methods. Once the economic and environmental benefits are known widely, variable rate technology would be rapidly implemented at least in high value crops. To spur adoption of precision farming methods in developing countries, pilot demonstration projects must be conducted at various growers locations by involving farmers in all stages of the project. The pilot projects must attempt to answer the grower’s needs and emphasize the operational implementation of technology and complete analysis of the costs and savings involved. Documentation of pilot projects would help in examining the operational weaknesses and identification of remedial measures. The projects can be used to train innovative farmers and early adopters, expose the neighboring nonparticipating farmers to the new technologies, and show the usefulness of the technology for short and long-term management.

The role of agricultural input suppliers, extension advisors and consultants in the spread of these technologies is vital. For instance, public agencies should consider supplying free data such as remotely sensed imagery to the universities and research institutes involved in precision farming research. Also, professional societies of agronomy, agricultural informatics, and engineering must provide training guidance in the use of technologies. The involvement of inter/disciplinary teams is essential in this. Small farm size will not be a major constraint, if the technologies are available through consulting, custom and rental services. The role of agricultural cooperatives is important in dissemination of precision farming technologies to small farmers. If precision farming is considered a series of discrete services: map generation, targeted scouting, it is possible to fit these services within the structure of a progressive agricultural cooperative in each developing country.

Changes in agricultural policies are also necessary to promote the adoption of precision farming. There are basically two policy approaches: regulatory policies and market based policies. The former refer to environmental regulations on the use of farm inputs and later refer to taxes and financial incentives aimed at encouraging growers to efficiently use farm inputs. In most developing countries the lack of penalties for pollutant generation has partly contributed to an excessive use of inputs.

Subsidies on inputs and outputs and mechanisms that prevent the price system from rationing limited resources are also common. The latter include state-guaranteed crop prices, tariffs, import quotas, export subsidies. Inputs such as water and fossil fuels are usually sold at prices that are well below the real resource cost of their use, which consists not only production costs but also includes scarcity value and costs of pollution. In such cases, the formulation of policies that reflect the real scarcity value of natural resources and penalize pollution and policies such as green payments for farmers adopting techniques that would lower environmental costs can promote the adoption of precision farming technologies.

In most developing countries, the pollution effects of agriculture have been largely ignored so far because of inability to effectively monitor such effects. The advent of precision farming, and the computerization of input and output flows, will now enable such monitoring. Higher taxes on pollution farms are often recommended, but there is strong opposition to the implementation of the polluter-pays-principle concept in most countries including India. At the same time, some consumers in India would like to see a drastic reduction in the use of pesticides and fertilizers, and are willing to pay as much as 4 to 6 times the normal price for produce such as organic vegetables, soybean and wheat. When the price elasticity of input use is low and the input costs are only a small part of the total production expenditure, as in the case of fertilizers and pesticides. Very high taxes are required to reduce their use adequately. Given the unfeasibility of such high taxes, a hybrid policy may be implemented for controlling pollution. A tax-free quota of N can be combined with taxes on additional N use.

At the research level, many issues remain to be resolved. Although some progress has been made at Space Application Center, Ahmedabad, yield monitors for small farm conditions are yet to be developed. The development of standards for the hardware and software (image transfer formats and GPS transfer formats, map projection formats) is another issue. Crop models and decision support systems must be improved by considering local resources. Data for calibration of models must be made available to increase their accuracy and/or predictability. The ability to finance a creative information venture in agriculture will affect the speed of diffusion of precision farming technologies. Commercial banks, as well as other sources of funding, have to be educated regarding the potential of precision farming. In many developing countries, it may be worthwhile to develop programmes of subsidized credit to enable R&D activities on precision farming.

OPERATIONALIZATION OF PRECISION FARMING

"Agriculture is the backbone of the Indian Economy"-said Mahatma Gandhi five decades ago. Even today, as we enter the new millennium, the situation is still the same, with almost the entire economy being sustained by agriculture, which is the mainstay of the villages. Not only the economy, but also every one of us looks up to agriculture for our sustenance too. Therefore, it is no surprise if agriculture gets the celebrity status in the name of Precision farming (PF). In recent times, the researchers in the field have been busy formulating methodologies and fabricating new implements for precision farming. It is here the challenge arises considering the implementation of the technology at various levels in the Global community. The need of the hour is not application of the technology but the adoption of appropriate technology, which would suit the particular level of the global community. In India, the farming practices are too haphazard and non-scientific and hence need some forethought before implementing any new technology. It is here that this paper comes out with a cost-benefit analysis to prove that precision farming is possible in India.

Applications of agricultural inputs at uniform rates across the field without due regard to in-field variations in soil fertility and crop conditions does not yield desirable results in terms of crop yield. The management of in-field variability in soil fertility and crop conditions for improving the crop production and minimizing the environmental impact is the crux of precision farming. Thus, the information on spatial variability in soil fertility status and crop conditions is a pre-requisite for adoption of precision farming.

Space technology including global positioning system (GPS) and GIS holds good promise in deriving information on soil attributes and crop yield, and allows monitoring seasonally-variable soil and crop characteristics, namely soil moisture, crop phenology, growth, evapotranspiration, nutrient deficiency, crop disease, and weed and insect infestation, which, in turn, help in optimizing inputs and maximizing crop yield and income. Though widely adopted in developed countries, the adoption of precision farming in India is yet to take a firm ground primarily due to its unique pattern of land holdings, poor infrastructure, lack of farmers' inclination to take risk, socio-economic and demographic conditions. The aim of this paper is to suggest measures for the implementation of this novel technique in the country with a greater emphasis on the systematic approach towards its operationalisation.

Objectives

1. To explain the feasibility of precision farming technology with emphasis on seed spacing, tillage, etc.
2. To set up a DGPS network all around the country and achieving few centimeters accuracy for the purpose of Site-Specific Management (SSM) in Precision Farming.
3. To analyse the cost and benefit in terms of Indian farmer's income-expenditure.

Precision Farming

The conventional agronomic practices follow a standard management option for a large area irrespective of the variability occurring within and among the field. For decades now, the farmers have been applying fertilizers based on recommendations emanating from research and field trials under specific agro-climatic conditions. Since soil-nutrient, characteristics vary not only from one region to another, but also from field to field. Even within a field, there is a need to take into account such variability while applying fertilizers to a particular crop. Consideration of in-field variations in soil fertility and crop conditions and matching the agricultural inputs like seed, fertilizer, irrigation, insecticide, pesticide, etc. in order to optimize the input or maximizing the crop yield from a given quantum of input, is referred to as precision farming or precision agriculture or precision crop management. The term “precision farming” means carefully tailoring the soil and crop management to fit the different conditions found in each field. It is defined as the application of technologies and principles to manage spatial and temporal variability associated with all aspects of agricultural production.. It is also referred to as “prescription farming”, “site specific farming” or “variable rate technology.”

By catering to this variability, called precision farming, one can improve the productivity or reduce the cost of production and diminish the chance of environmental degradation caused by excess use of inputs. Thus, mapping and analysis of within field variability is an essential input for precision crop management. Thus, PF involves acquiring the variations in crop or soil properties, mapping, and analysing the variations, adopting suitable management techniques to maximize the yield. Farmers have been applying fertilizers based on recommendations emanating from research and field trials under specific agro-climatic conditions, which have been extrapolated to a regional level. Since soil nutrient characteristics vary not only between regions and between farms but also from plot to plot, and within a field or plot, there is a need to take into account such variability while applying fertilizers to a particular crop. Consideration of in-field/plot variations in soil fertility and crop conditions and matching the agricultural inputs like seed, fertilizer, irrigation, insecticide, pesticide, etc. in order to optimize the input or maximizing the crop yield from a given quantum of input, is referred to as precision farming or precision agriculture or precision crop management.

The information for variability map can be obtained from soil tests for nutrient availability, yield monitors for crop yield, soil samples for organic matter content, information in soil maps, or ground conductivity meters for soil moisture. Generally, the fields are manually sampled along a regular grid and the analysed results of the samples are interpolated using geostatistical techniques.

These techniques are time consuming, labour intensive and in many cases destructive especially, for agricultural situation in India. With small size of landholdings and low income of farmers, the adoption of this methodology in its present form is not feasible. Various workers have shown the advantages of using remote sensing technology to obtain spatially and temporally variable information for precision farming. In an earlier work, Ray et al. (2001) have shown the usefulness of IRS merged data in mapping the variability.

Components of precision farming:

1. Remote Sensing
2. GIS
3. DGPS
4. Variable rate Applicator

Remote Sensing

This is for Data acquisition of the farms to find the soil, vegetation and other parameters that are amenable for remote sensing. Remote sensing techniques play an important role in precision farming by providing continuous acquired data of agricultural crops. Remote sensors image vegetation, which is growing on different soil types with different water availability, substrate, impact of cultivation, and relief. These differences influence the state of the plants and cause heterogeneous regions within single fields. Hence, the heterogeneous vegetation acts as an interface between soil and remote sensing information, because vegetation parameters describing the state of the plants can be deduced from remote sensing imagery. The analysis of the variability occurring within the field was carried out by measuring soil and plant parameters through conventional methods as well as through spectral techniques using ground truth spectroradiometer (350-1800 nm) and satellite data. Ray et al. (2001), in general, have reviewed by Moran et al. (1997) and for Indian conditions the potentials of remote sensing in providing information required for precision farming.

GEOGRAPHICAL INFORMATION SYSTEM

The Geographic Information System (GIS) contributes significantly to precision farming by allowing presentation of spatial data in the form of a map. In addition, GIS forms an ideal platform for the storage and management of model input data and the presentation of model results, which the process model provides.

DGPS

'Do the right thing, in the right place and in right time'

This is where GPS comes into picture. In addition, the accuracy, which is the important factor in PF, demands for DGPS (Differential Global Positioning System). GPS makes use of a series of military satellites that identify the location of farm equipment within a meter of an actual site in the field. The value of knowing a precise location within inches is that:

1. Locations of soil samples and the laboratory results can be compared to a soil map,
2. Fertilizer and pesticides can be prescribed to fit soil properties (clay and organic matter content) and soil conditions (relief and drainage),
3. Tillage adjustments can be made as one finds various conditions across the field,
4. One can monitor and record yield data as one goes across the field.

The Global Positioning System (GPS) technology provides accurate positioning system necessary for field implementation of variable rate technology (VRT). The Internet makes possible the development of a mechanism for effective farm management using remote sensing.

Variable Rate Applicator

The variable rate applicator has three components:

1. Control computer
2. Locator
3. Actuator

The control computer coordinates the field operation. It has a map of desired activity as a function of geographic location. It receives the equipment's current location from the locator, which has a GPS in it, and decides what to do based upon the map in its memory or data storage.

It then issues the command to the actuator, which does the input application.

PRACTICAL PROBLEMS IN INDIAN AGRICULTURE

Precision Agriculture has been mostly confined to developed countries. Reasons of limitations of its implementation in developing countries like India are:

a. Small land holdings,
b. Heterogeneity of cropping systems and market imperfections,
c. Lack of technical expertise knowledge and technology (India spends only 0.3% of its agricultural Gross Domestic Product in Research and Development),
d. High cost.

In India, major problem is the small field size. More than 58 percent of operational holdings in the country have size less than 1ha. Only in the states of Punjab, Rajasthan, Haryana and Gujarat more than 20 per cent of agricultural lands have operational holding size of more than four hectare.

There is a scope of implementing precision Agriculture for crops like, rice and wheat especially in the states of Punjab and Haryana. Commercial as well as Horticultural crops shows a wider scope for precision Agriculture.

In India, broadly two types of agriculture viz., high input agriculture characterized by the provision of assured irrigation and other agricultural inputs, and subsistence farming, which is confined mostly to rain-fed, or dry land regions, are prevalent. Nearly two-third arable land in India is rain-fed. The crop yields are very low (˜ 1 t ha-1) and very good potential exists for increasing productivity of rain-fed Cropping systems.

Methodology

Using a GPS along with a digital drainage map, the farmer is able to apply these pesticides in a safer manner. The spraying equipment can be preprogrammed automatically to turn off when it reaches the distance limitation or zone of the drainage feature. Additionally, farmers can preprogram the rate of pesticide or fertilizer to be applied so that only the amount needed determined by the soil condition is applied varying this rate from one area of the field to another. This saves money and allows for safer use of these materials.

Data collected by the GPS operations can be automatically recorded with the GIS program. Remotely sensed data can be analysed and added to the GIS using soil maps, digital terrain and field operations information as ground truth. Additionally the farmers are able to record observations through the growing season such as weed growth, unusual plant stress or coloring and growth conditions. This can be used to guide further field operations like spraying, fertilizing and irrigating and it is part of the permanent record. Precision farming will make a strong impact on the way farmers manage their farm operations in the future.

Seed Spacing

The yield from the seeds is maximum when adequate care is taken to adequately space them to enable each plant the optimum requisites for its growth, namely the soil nutrients, water, sunlight and protection against pest infestation. This is best accomplished by varying the seeding rate even within a single piece of land according to the soil conditions such as texture, organic matter and available soil moisture. One would plant fewer seeds in sandy soil as compared to silt loam soils because of less available moisture. The lower seed population usually has larger heads (ears) of harvested seeds providing for a maximum yield. Since soils vary even across an individual farm field, the ability to change seeding rates as one goes across the field allows the farmer to maximize this seeding rate according to the soil conditions. A computerized soil map of a specific field on a computer fitted on the tractor along with a

GPS can tell farmers where they are in the field allowing the opportunity to adjust this seeding rate as they go across their fields.

STEPS TO BE TAKEN FOR IMPLEMENTING PF

In the present existent situation, the potential of precision agriculture in India is limited by the lack of appropriate measurement and analysis techniques for agronomically important factors (National Research Council, 1997). High accuracy sensing and data management tools must be developed and validated to support both research and production. The limitation in data quality/availability has become a major obstacle in the demonstration and adoption of the precision technologies. The adoption of precision agriculture needs combined efforts on behalf of scientists, farmers and the government. The following methodology could be adopted in order to operationalise precision farming in the country.

1. Creation of multidisciplinary teams involving agricultural scientists in various fields, engineers, manufacturers and economists to study the overall scope of precision agriculture.
2. Formation of farmer's co-operatives since many of the precision agriculture tools are costly (GIS, GPS, RS, etc).
3. Government legislation restraining farmers using indiscriminate farm inputs and thereby causing ecological/environmental imbalance would induce the farmer to go for alternative approach.
4. Pilot study should be conducted on farmer's field to show the results of precision agriculture implementation.
5. Creating awareness amongst farmers about consequences of applying imbalanced doses of farm inputs like irrigation, fertilizers, insecticides and pesticides.

Realizing the potential of space technology in precision farming, the Department of Space, Government of India has initiated eight pilot studies in well-managed agricultural farms of the ICRISAT, the Indian Council of Agricultural Research and the Agricultural Universities, as well as in farmers' fields.

The pilot studies aim at delineating homogeneous zones with respect to soil fertility and crop yield, estimation of potential yield, yield gap analysis, monitoring seasonally-variable soil and crop conditions using optical and microwave sensor data, and matching the farm inputs to bridge the gap between potential and actual yield through Spatial Decision Support Systems (SDSS). The test sites are spread over a fairly large area across a cross section of agro-climatic zones of the Indian sub-continent, and cover some of the important crops like wheat, rice, sorghum, pigeon pea, chickpea, soybean and groundnut.

The next step would be to generate detailed-level information on soil resources addressing potentials and limitations of individual fields since except for states like Punjab, Haryana, Madhya Pradesh and Maharashtra where fields size is quite large, practically individual field could be treated as a homogenous management unit for the purpose of precision farming.

PRECISION FARMING IN INDIAN AGRICULTURAL SCENARIO

Agriculture is a very important sector for the sustained growth of the Indian economy. About 70 per cent of the rural households and 8 per cent of urban households are still principally dependent on agriculture for employment. Since some three-quarters of the population live in rural areas, a majority of households thus depend principally on this sector. Though, industrialization of the Indian economy has adversely affected the share of agriculture in the GDP, the fact cannot be ignored that India has undergone a series of successful agricultural revolutions-starting with the ‘green’ revolution in wheat and rice in the 1960’s and 1970’s, the ‘white’ revolution in milk to the ‘yellow’ revolution in oilseeds in 1980’s. As a result, India has achieved self-sufficiency in agriculture. Applications of agricultural inputs at uniform rates across the field without due regard to in-field variations in soil fertility and crop conditions does not yield desirable results in terms of crop yield. The management of in-field variability in soil fertility and crop conditions for improving the crop production and minimizing the environmental impact is the crux of precision farming. Geographically, India is widely distributed into several agro-climatic zones, and the information need for the farming systems in these areas is entirely different. Integrating the application of available technologies to realize farmers' goals requires a systems approach to farming. The concept of fully automated villages was a dream till few months back. But the reality has come to true with the rapid growth of information and communication technology in the world scenario. The wired villages and info villages have shown that Information can be disseminated in more useful manner and as farmers need.

PRECISION AGRICULTURE

Precision agriculture is an agricultural system that has the potential of dramatically changing agriculture in this 21st century. Precision agriculture lends it self to most agricultural applications and can be implemented at whatever levels are required. Precision agriculture is based on information technology, which enables the producer to collect information and data for better decision making. Precision agriculture is a pro-active approach that reduces some of the risk and variables common to agriculture. Precision agriculture is more environmentally sound and is and integral part in sustaining natural resources. To better understand the need for an accurate definition of precision agriculture lets look at how precision agriculture is being considered. Precision agriculture is considered a concept, management strategy, and even a philosophy. It is said, "Precision agriculture is a phrase that captures the imagination of many concerned with the production of food, feed, and fiber." The concept of precision agriculture offers the promise of increasing productivity while decreasing production cost and minimizing environmental impacts. Precision agriculture conjures up images of farmers overcoming the elements with computerized machinery that is precisely controlled via satellites and local sensors and using planning software that accurately predicts crop development. This image has been called the future of agriculture.

Precision farming is characterised by a number of sophisticated tools that assist in monitoring variation and managing inputs. These include:

- Global Positioning System (GPS) – a referencing device capable of identifying sites within a field;

- Sensors and dataloggers – crop, soil and climate information can be monitored at a high frequency using these technologies;
- Geographic Information Systems (GIS) – maps of these attributes can be generated and analysed using simple browsers or complex models.

GPS ROLE IN PRECISION AGRICULTURE

Precision Agriculture is doing the right thing, at the right place, at the right time. Knowing the right thing to do may involve all kinds of high tech equipments and fancy statistics or other analysis. Doing the right thing however starts with good managers and good operators doing a good job of using common tools such as planters, fertilizer applicators, harvesters and whatever else might be needed.

The use of GPS in Agriculture is limited but it is fair to expect wide spread use of GPS in future. Recently a GPS-based crop duster (precision GPS Helicopter), which can spray an area as small as 4 X 4 mtr. is attracting great attention. Some progressive farmers are now beginning to use GPS for recording observations. Such as weed growth, unusual plant stress, colouring and growth conditions, which can then be mapped with a GIS programmes. In the years to come, GPS system role in precision agriculture may help the Indian farmers to harvest the fruits of frontier technologies without compromising the quality of land and produce.

STATUS OF PRECISION FARMING IN INDIA

Hence, the first thing that comes to mind is that, this system is not for developing countries, especially India, where the farmers are poor, farming is mostly subsistent and the land holding size is small. But, this is far from the truth as this approach has a large potential for improving the agricultural production in developing world. Imagine this situation where a farmer goes to his field with a GPS (Global Positioning System)-guided tractor. The GPS senses the exact location of tractor within the field. It sends signals to the computer fixed on to the tractor, which has a Geographical Information System (GIS), storing the soil nutrient requirement map in it. The GIS, in consultation with a Decision Support System would decide what is the exact requirement of fertilisers for that location. It then commands a variable rate fertiliser applicator, which is again attached with the tractor, to apply the exact dosage at the precise location of farm.

But, this is what precision farming means to large growers in the highly developed parts of the globe. To make it clearer, Precision Farming is the system of matching of resource application and agronomic practices with soil attributes and crop requirements since they vary across a field.

Tata Kisan Kendra : The concept of precision farming being implemented by the TKKs has the potential to catapult rural India from the bullock-cart age into the new era of satellites and IT. TCL's extension services, brought to farmers through the TKKs, use remote-sensing technology to analyse soil, inform about crop health, pest attacks and coverage of various crops predicting the final output. This helps farmers adapt quickly to changing conditions. The result: healthier crops, higher yields and enhanced incomes for farmers.

Government organization: Precision Agriculture models are not complete, unless the parameters related to empowerment of the farmers; especially small and marginal farmers are integrated. Now it is the turn of good news to the Indian farming community.

Some of the research institutes such as Space Applications Centre (ISRO), M.S. Swaminathan Research Foundation, Chennai, Indian Agricultural Research Institute, New Delhi, and Project Directorate of Cropping Systems Research, Modipuram, had started working in this direction and in soon it will help the Indian farmers harvest the fruits of frontier technologies without compromising on the quality of land. According to the Exim Bank officials, though the research and development on PF is currently at a nascent stage in the country, the efforts being put on by the four research institutes were expected to turn the green revolution into an evergreen revolution.

In this context, ISRO has also initiated Gramsat project in Orissa. In the line of JDCP, the Gramsat project aims at empowering the people especially the poor and marginalized, by awareness building and access to information and services.

Towards this, a network of one-way video and two-way audio Forecasting the yield of mono and multiple crops is being done at NRSA.

Acreage estimates and crop inventory is being done during Kharif and Rabi seasons for Rice, which is the major crop grown in our India. Other crops like Banana, Chillies, Cotton, Maize, Sugarcane and Tobacco are also being inventoried. Satellite data can also delineate different crops that are grownin the same area, and an inventory of each of the crops can be done.

Remote Sensing and Sensors for PF

Precision farming needs information about mean characteristics of small, relatively homogeneous management zones. This is for Data acquisition of the farms to find the soil, vegetation and other parameters that are amenable for remote sensing. Remote sensing techniques play an important role in precision farming by providing continuous acquired data of agricultural crops. Remote sensors image vegetation, which is growing on different soil types with different water availability, substrate, impact of cultivation, and relief Sensors use for the following applications:

- Soil Properties Sensing: Soil Texture, Structure, and Physical Condition Soil Moisture; Soil Nutrients.
- Crop Sensing: Plant Population; Crop Stress and Nutrient Status.
- Yield Monitoring Systems: Crop Yield; Harvest Swath Width; Crop Moisture.
- Variable Rate Technology Systems: Fertilizer flow; Weed detection, pressure sensors.

SOFTWARE SOLUTION FOR PRECISION FARMING

Linux is a good platform for doing this research-oriented work. Much of the analysis can be translated into such mainstream topics as signal processing or multi-dimensional statistics. Some of the best software for exploring software in these topics is the product of government and university research and is “free”—an important quality in tight budgets. GRASS, xldlas and Santis are three packages which are very helpful in precision farming.

ECONOMICS FEASIBILITY OF PRECISION FARMING

Unlike some new technologies, there is no clear answer as to whether or not PA is economical beneficial. Indian agriculture condition :

1. On the one hand there are depletions of ecological foundations of the agro-ecosystems, as reflected in terms of increasing land degradation, depletion of water resources and rising trends of floods, drought and crop pests and diseases. On the other hand, there is imperative socio-economic need to have enhanced productivity per units of land, water and time.
2. At present, 3 ha of rain fed areas produce cereal grain equivalent to that produced in 1 ha. of irrigated. Out of 142 million ha. Net sown areas, 92 million ha. are under rain-fed agriculture in the county.
3. From equity point of view, even the record agricultural production of more than 200 Mt is unable to address food security issue. A close to 60 Mt food grains in the storehouses of Food Corporation of India (FCI) is beyond the affordability and access to the poor and marginalized in many pockets of the country.
4. Globally, there are challenges arising from the Globalization especially the impact of WTO regime on small and marginalized farmers.
5. Some other unforeseen challenges could be anticipated global warming scenario and its possible impact on diverse agro-ecosystems in terms of alterations in traditional crop belts, micro-level perturbations in hydrologic cycle and more uncertain crop-weather interactions etc.

While some studies have reported positive returns to variable-rate technology (VRT), others have reported costs higher than returns or no significant difference in returns.

- Precision Agriculture is a system, not a single piece of equipment or technology. A GPS by it self has little value to farmer. However, when combined with a yield monitor or a VRT, it may have value.
- Returns may be positive if costs can be spread over many applications. Specialized equipment, which has

limited uses, has greater risks associated with it than equipment that has many uses. A multi-use tractor will likely pay for itself sooner than a new, single-use machine.

- Precision agriculture may not return on low-valued commodities as it does on high-valued specialty crops i.e. high revenue for grape than the wheat and paddy.

GPS controlled tractor guidance systems may affect when and how tractors are operated. Precision farming is useful in many situations in developing countries. Rice, wheat, sugar beet, onion, potato and cotton among the field crops and apple, grape, tea, coffee and oil palm among horticultural crops are perhaps the most relevant. Some have a very high value per acre, making excellent cases for site-specific management. For all these crops, yield mapping is the first step to determine the precise locations of the highest and lowest yield areas of the field. Researchers at Kyoto University recently developed a two-row rice harvester for determining yields on a micro plot basis.

Subsidies on inputs and outputs and mechanisms that prevent the price system from rationing limited resources are also common. The latter include state-guaranteed crop prices, tariffs, import quotas, export subsidies. Inputs such as water and fossil fuels are usually sold at prices that are well below the real resource cost of their use, which consists not only production costs but also includes scarcity value and costs of pollution. In such cases, the formulation of policies that reflect the real scarcity value of natural resources and penalize pollution and policies such as green payments for farmers adopting techniques that would lower environmental costs can promote the adoption of precision farming technologies.

8

Training and Programmes in Remote Sensing and GIS

INTRODUCTION

One of the important spin offs of the space research is satellite based remote sensing which facilitates synoptic view repetitive coverage and provides multispectral and multiresolution data. It enables identification, delineation and mapping of natural resources on regional basis. Integration of remote sensing with Geographic Information system Technology has enhanced its capabilities in the areas of resources management.

The application areas are increasing dramatically. It is massive task to transfer these new technologies to end users for operational purpose. GIS are often seen as a multipurpose technology which can satisfy the needs of a variety of users. The process of transferring this technology from laboratory to real world organizations is complex and problematic. It includes users, senior management, technical personnel and computer specialists.

The implementation of GIS tends to be a highly resources intensive activity. The experiences of the case studies indicate that regardless of the balance between in house and external development, the financial cost, staff time and variety of skills required are considerable if utilization is to be achieved. The training must provide users with sufficient knowledge to be able to handle geographic information as well as the computer technology.

The last decade has witnessed a steady increase in the number of training courses and post graduate programs being offered in educational institutions and R&D organizations. The target audience vary from high level administrators and professional scientists to application oriented personnel and technicians. If senior decision makers receive a formal training in these emerging technologies, they will be prepared and motivated to accelerate the use of these areas in their organizations. They will use their positions of responsibility to initiate or promote programmes incorporating these technologies and will welcome innovation from young officers. Another interesting aspect on transfer of these technologies to user departments is the availability of low cost computer systems and a software to process the spatial data. In the initial stages, many user departments establishing the infrastructure. However with the advent of microcomputers and the availability of low cost software packages, it is becoming possible for the user departments to take advantage to take advantage of these technologies.

LEVELS OF TRAINING

The past experience in training and education of remote sensing and GIS technologies has indicated that three levels of training programs are required to satisfy the needs of three distinct groups.

i. *A short course of about 1 or 3 weeks duration:* This course is meant for the decision makers, planners, manager and administrators. It introduces the principles of remote sensing and geographic information systems to the professionals who do not have any prior exposure to these fields. Both the capabilities and limitations of these technologies are brought out during this course.

ii. *A middle level course of 1 or 2 months duration.* This course is suitable or middle level officers or scientists/ manager, who supervise the activities of remote sensing and GIS in their organization. The course generally covers principles of remote sensing, visual and digital interpretation of remote sensing data, concept of GIS, use of spatial analysis, integration of remote sensing and GIs and practical demonstrations.

iii. *An in depth course of 1 or 2 semesters duration.* This course is tailored for working level scientists and technicians. They must get the capabilities to establish the infrastructure and carry out operational programmes relating to remote sensing and GIS analysis. The course content includes principles of remote sensing and GIS, in depth detail of visual and digital interpretation of remote sensing data, spatial data structures and analysis, integration of remote sensing and GIS for resources applications along with hands on training on GIS packages and digital image processing. Each candidate carries out a case study and submits a dissertation/project report at the end of the programme.

Majority of the potential users want to understand the general principles, techniques and the advantages of these technologies before undertaking operational programmes. Also the participants may not be able to be away from their place of employment for long durations.

The center pf studies in Resources Engineerings, Indian Institute of Technology, Bombay is conducting short courses of one week duration in Geographic Information system and Remote Sensing. General framework of the course comprises of 15 lectures and 3 practical sessions. Each lecture is of one and half hours duration and the practical session is of 3 hours duration. In the lecture, the expert speaks for about an hours leaving 30 minutes for closer participants. In the practical sessions, the participants are provided hands-on training of GIS packages including remote sensing data analysis for resources applications. This gives an opportunity for closer interaction between the users and technical specialists. If implementation and utilization have to be achieved, the interests of the users must be paramount.

The number of participants is normally limited to 15 n an attempt to maximise the interaction between the instructors and participants. Also it helps to provide better hands on training to the participants in the practical session. Moreover when the number of participants is less, the instructors are able to tailor the courses as per clients requirements. The participants are provided the lecture notes and the handouts for each practical session. A PC based GIS package GRAM

(Geo-Referenced Area Management) developed by the centre is used for providing hands on training in map digitalization, geo-registration, digital image processing, terrain modelling and spatial analysis. Case studies are illustrated to demonstrate the capabilities of remote sensing and GIS technology for resources management.

The 4 week training programme on GIs has been designed in a modular approach. The first week covers introduction to geographic information systems, remote sensing and data base management systems. As the participants of these short courses come with varied backgrounds, the first week lectures become a prerequisite to bring uniformity among the participants. The second week specializes on spatial data structures, spatial modelling and spatial decision support systems. The third week covers integration of remote sensing growth GUS and application of GIS to water resources, agricultures, forestry, terrain evaluation, mineral resources, urban/regional planning environmental impact assessment etc. During the last week, transfer of GIS technology to the use requirements and the organizational issues for effective transfer. In addition for the benefit of R and D organizations, open areas of research in GIS are covered. The core curriculum prepared by NCGIA, USA on Introduction, Technical Issues and Application Issues in GIS can be a base material for designing any GIS training any GIS training course.

Theory lectures are given in the morning sessions while hands on training is provided in the afternoon sessions. The participants select the case studies depending on their specializations and carry out the complete exercised using GRAM GIS Package. In addition, exposure to other GIS packages namely Arc/Info, Intergraph, SPANS, PAMAP, ILWIS etc are provided. Often users face unforeseen difficulties with the software they purchased. The problems include lack of compatibility between software packages, in liability to undertake simple tasks and extremely cumbersome commands. Hence exposure to various packages will give the users an idea to select the package suitable for their application studies.

An evaluation questionnaire is circulated to the participants on the completion of the course to get feed back from them. The questionnaire asks the participants to evaluate course content, lecturing, practical exercises, course materials and other logistic facilities. The comments and suggestions help to improve the course module and content. Generally, participants are more interested impractical session and case studies compared to theory. They like to spend longer hours in hands on training and less time in listening to lectures.

MAPS, REMOTE SENSING AND GIS

A map can be simply defined as a graphic representation of the real world. This representation is always an abstraction of reality. Because of the infinite nature of our Universe it is impossible to capture all of the complexity found in the real world. For example, topographic maps abstract the three-dimensional real world at a reduced scale on a two-dimensional plane of paper.

Maps are used to display both cultural and physical features of the environment. Standard topographic maps show a variety of information including roads, land-use classification, elevation, rivers and other water bodies, political boundaries, and the identification of houses and other types of buildings. Some maps are created with very specific goals in mind. A weather map showing the location of low and high pressure centers and fronts over most of North America. The intended purpose of this map is considerably more specialized than a topographic map.

The art of map construction is called cartography. People who work in this field of knowledge are called cartographers. The construction and use of maps has a long history. Some academics believe that the earliest maps date back to the fifth or sixth century BC. Even in these early maps, the main goal of this tool was to communicate information. Early maps were quite subjective in their presentation of spatial information. Maps became more objective with the dawn of Western science. The application of scientific method into cartography made maps more ordered and accurate. Today, the art of map making is quite a sophisticated science employing methods from cartography, engineering, computer science, mathematics, and psychology.

Cartographers classify maps into two broad categories: reference maps and thematic maps. Reference maps normally show natural and human-made objects from the geographical environment with an emphasis on location. Examples of general reference maps include maps found in atlases and topographic maps. Thematic maps are used to display the geographical distribution of one phenomenon or the spatial associations that occur between a number of phenomena.

MAP PROJECTION

The shape of the Earth's surface can be described as an ellipsoid. An ellipsoid is a three-dimensional shape that departs slightly from a purely spherical form. The Earth takes this form because rotation causes the region near the equator to bulge outward to space. The angular motion caused by the Earth spinning on its axis also forces the polar regions on the globe to be somewhat flattened.

Representing the true shape of the Earth's surface on a map creates some problems, especially when this depiction is illustrated on a two-dimensional surface. To overcome these problems, cartographers have developed a number of standardized transformation processes for the creation of two-dimensional maps. All of these transformation processes create some type of distortion artifact. The nature of this distortion is related to how the transformation process modifies specific geographic properties of the map. Some of the geographic properties affected by projection distortion include: distance; area; straight line direction between points on the Earth; and the bearing of cardinal points from locations on our planet.

The first two-dimensional projection shows the Earth's surface as viewed from space. This orthographic projection distorts distance, shape, and the size of areas. Another serious limitation of this projection is that only a portion of the Earth's surface can be viewed at any one time.

The second illustration displays a Mercator projection of the Earth. On a Mercator projection, the north-south scale increases from the equator at the same rate as the corresponding east-west scale. As a result of this feature, angles drawn on this type of map are correct. Distortion on a Mercator map increases at an increasing rate as one moves toward higher latitudes. Mercator maps are used in navigation because a line drawn between two points of the Earth has true direction.

However, this line may not represent the shortest distance between these points.

The Gall-Peters projection was developed to correct some of the distortion found in the Mercator system. The Mercator projection causes area to be gradually distorted from the equator to the poles. This distortion makes middle and high latitude countries to be bigger than they are in reality. The Gall-Peters projection corrects this distortion making the area occupied by the world's nations more comparable.

The Miller Cylindrical projection is another common two-dimensional map used to represent the entire Earth in a rectangular area. In this project, the Earth is mathematically projected onto a cylinder tangent at the equator. This projection in then unrolled to produce a flat two-dimensional representation of the Earth's surface. This projection reduces some of the scale exaggeration present in the Mercator map. However, the Miller Cylindrical projection describes shapes and areas with considerable distortion and directions are true only along the equator.

The Mollweide projection improves on the Robinson projection and has less area distortion. The final projection presented presents areas on a map that are proportional to the same areas on the actual surface of the Earth. However, this Sinusoidal Equal-Area projection suffers from distance, shape, and direction distortions.

MAP SCALE

Maps are rarely drawn at the same scale as the real world. Most maps are made at a scale that is much smaller than the area of the actual surface being depicted. The amount of reduction that has taken place is normally identified somewhere on the map. This measurement is commonly referred to as the map scale. Conceptually, we can think of map scale as the ratio between the distance between any two points on the map compared to the actual ground distance represented. This concept can also be expressed mathematically as:

$$\text{Map Scale} = \frac{\text{Map Distance}}{\text{Earth Distance}}$$

On most maps, the map scale is represented by a simple fraction or ratio. This type of description of a map's scale is called a representative fraction. For example, a map where one unit (centimeter, meter, inch, kilometer, etc.) on the illustration represents 1,000,000 of these same units on the actual surface of the Earth would have a representative fraction of 1/1,000,000 (fraction) or 1:1,000,000 (ratio). Of these mathematical representations of scale, the ratio form is most commonly found on maps.

Scale can also be described on a map by a verbal statement. For example, 1:1,000,000 could be verbally described as "1 centimeter on the map equals 10 kilometres on the Earth's surface" or "1 inch represents approximately 16 miles".

Most maps also use graphic scale to describe the distance relationships between the map and the real world. In a graphic scale, an illustration is used to depict distances on the map in common units of measurement. Graphic scales are quite useful because they can be used to measure distances on a map quickly.

Maps are often described, in a relative sense, as being either small scale or large scale. We have maps representing an area of the world at scales of 1:100,000, 1:50,000, and 1:25,000. Of this group, the map drawn at 1:100,000 has the smallest scale relative to the other two maps. The map with the largest scale is map C which is drawn at a scale of 1:25,000.

INTRODUCTION TO REMOTE SENSING

Remote sensing can be defined as the collection of data about an object from a distance. Humans and many other types of animals accomplish this task with aid of eyes or by the sense of smell or hearing.

Geographers use the technique of remote sensing to monitor or measure phenomena found in the Earth's lithosphere, biosphere, hydrosphere, and atmosphere. Remote sensing of the environment by geographers is usually done with the help of mechanical devices known as remote sensors. These gadgets have a greatly improved ability to receive and record information about an object without any physical contact. Often, these sensors are positioned away from the object of interest by using helicopters, planes, and satellites. Most sensing devices record information about an object by measuring an object's transmission of electromagnetic energy from reflecting and radiating surfaces.

Remote sensing imagery has many applications in mapping land-use and cover, agriculture, soils mapping, forestry, city planning, archaeological investigations, military observation, and geomorphological surveying, among other uses. For example, foresters use aerial photographs for preparing forest cover maps, locating possible access roads, and measuring quantities of trees harvested. Specialized photography using colour infrared film has also been used to detect disease and insect damage in forest trees.

The simplest form of remote sensing uses photographic cameras to record information from *visible* or *near* infrared wavelengths. In the late 1800s, cameras were positioned above the Earth's surface in balloons or kites to take oblique aerial photographs of the landscape.

During World War I, aerial photography played an important role in gathering information about the position and movements of enemy troops. These photographs were often taken from airplanes. After the war, civilian use of aerial photography from airplanes began with the systematic vertical imaging of large areas of Canada, the United States, and Europe. Many of these images were used to construct topographic and other types of reference maps of the natural and human-made features found on the Earth's surface.

SATELLITE REMOTE SENSING

In the 1960s, a revolution in remote sensing technology began with the deployment of space satellites.

From their high vantage-point, satellites have a greatly extended view of the Earth's surface. The first meteorological satellite, TIROS-1, was launched by the United States using an Atlas rocket on April 1, 1960. This early weather satellite used vidicon cameras to scan wide areas of the Earth's surface. Early satellite remote sensors did not use conventional film to produce their images. Instead, the sensors digitally capture the images using a device similar to a television camera. Once captured, this data is then transmitted electronically to receiving stations found on the Earth's surface.

Today, the GOES (Geostationary Operational Environmental Satellite) system of satellites provides most of the remotely sensed weather information for North America. To cover the complete continent and adjacent oceans two satellites are employed in a geostationary orbit.

The western half of North America and the eastern Pacific Ocean is monitored by GOES-10, which is directly above the equator and 135° West longitude. The eastern half of North America and the western Atlantic are cover by GOES-8. The GOES-8 satellite is located overhead of the equator and 75° West longitude.

Advanced sensors aboard the GOES satellite produce a continuous data stream so images can be viewed at any instance. The imaging sensor produces visible and infrared images of the Earth's terrestrial surface and oceans. Infrared images can depict weather conditions even during the night. Another sensor aboard the satellite can determine vertical temperature profiles, vertical moisture profiles, total precipitable water, and atmospheric stability.

In the 1970s, the second revolution in remote sensing technology began with the deployment of the Landsat satellites. Since this 1972, several generations of Landsat satellites with their Multispectral Scanners (MSS) have been providing continuous coverage of the Earth for almost 30 years.

Current, Landsat satellites orbit the Earth's surface at an altitude of approximately 700 kilometres. Spatial resolution of objects on the ground surface is 79 x 56 meters. Complete coverage of the globe requires 233 orbits and occurs every 16 days.

The Multispectral Scanner records a zone of the Earth's surface that is 185 kilometres wide in four wavelength bands: band 4 at 0.5 to 0.6 micrometers, band 5 at 0.6 to 0.7 micrometers, band 6 at 0.7 to 0.8 micrometers, and band 7 at 0.8 to 1.1 micrometers. Bands 4 and 5 receive the green and red wavelengths in the visible light range of the electromagnetic spectrum.

The last two bands image near-infrared wavelengths. A second sensing system was added to Landsat satellites launched after 1982. This imaging system, known as the Thematic Mapper, records seven wavelength bands from the visible to far-infrared portions of the electromagnetic spectrum. In addition, the ground resolution of this sensor was enhanced to 30 x 20 meters. This modification allows for greatly improved clarity of imaged objects.

The usefulness of satellites for remote sensing has resulted in several other organizations launching their own devices. In France, the SPOT (*Satellite Pour l'Observation de la Terre*) satellite program has launched five satellites since 1986. Since 1986, SPOT satellites have produced more than 10 million images.

SPOT satellites use two different sensing systems to image the planet. One sensing system produces black and white panchromatic images from the visible band (0.51 to 0.73 micrometers) with a ground resolution of 10 x 10 meters. The other sensing device is multispectral capturing green, red, and reflected infrared bands at 20 x 20 meters. SPOT-5, which was launched in 2002, is much improved from the first four versions of SPOT satellites. SPOT-5 has a maximum ground resolution of 2.5 x 2.5 meters in both panchromatic mode and multispectral operation.

Radarsat-1 was launched by the Canadian Space Agency in November, 1995. As a remote sensing device, Radarsat is quite different from the Landsat and SPOT satellites. Radarsat is an active remote sensing system that transmits and receives microwave radiation. Landsat and SPOT sensors passively measure reflected radiation at wavelengths roughly equivalent to those detected by our eyes.

Radarsat's microwave energy penetrates clouds, rain, dust, or haze and produces images regardless of the Sun's illumination allowing it to image in darkness. Radarsat images have a resolution between 8 to 100 meters. This sensor has found important applications in crop monitoring, defence surveillance, disaster monitoring, geologic resource mapping, sea-ice mapping and monitoring, oil slick detection, and digital elevation modelling.

PRINCIPLES OF OBJECT IDENTIFICATION

Most people have no problem identifying objects from photographs taken from an oblique angle. Such views are natural to the human eye and are part of our everyday experience. However, most remotely sensed images are taken from an overhead or vertical perspective and from distances quite removed from ground level. Both of these circumstances make the interpretation of natural and human-made objects somewhat difficult. In addition, images obtained from devices that receive and capture electromagnetic wavelengths outside human vision can present views that are quite unfamiliar.

To overcome the potential difficulties involved in image recognition, professional image interpreters use a number of characteristics to help them identify remotely sensed objects. Some of these characteristics include:

Shape: this characteristic alone may serve to identify many objects. Examples include the long linear lines of highways, the intersecting runways of an airfield, the perfectly rectangular shape of buildings, or the recognizable shape of an outdoor baseball diamond.

Size: noting the relative and absolute sizes of objects is important in their identification. The scale of the image determines the absolute size of an object. As a result, it is very important to recognize the scale of the image to be analysed.

Image Tone or *Colour*: all objects reflect or emit specific signatures of electromagnetic radiation. In most cases, related types of objects emit or reflect similar wavelengths of radiation. Also, the types of recording device and recording media produce images that are reflective of their sensitivity to particular range of radiation. As a result, the interpreter must be aware of how the object being viewed will appear on the image examined. For example, on colour infrared images vegetation has a colour that ranges from pink to red rather than the usual tones of green.

Pattern: many objects arrange themselves in typical patterns. This is especially true of human-made phenomena. For example, orchards have a systematic arrangement imposed by a farmer, while natural vegetation usually has a random or chaotic pattern.

Shadow: shadows can sometimes be used to get a different view of an object. For example, an overhead photograph of a towering smokestack or a radio transmission tower normally presents an identification problem.

This difficulty can be over come by photographing these objects at Sun angles that cast shadows. These shadows then display the shape of the object on the ground. Shadows can also be a problem to interpreters because they often conceal things found on the Earth's surface.

Texture: imaged objects display some degree of coarseness or smoothness. This characteristic can sometimes be useful in object interpretation. For example, we would normally expect to see textural differences when comparing an area of grass with a field corn. Texture, just like object size, is directly related to the scale of the image.

DATA FORMATS FOR DIGITAL SATELLITE IMAGERY

Digital data from the various satellite systems supplied to the user in the form of computer readable tapes or CD-ROM. As no worldwide standard for the storage and transfer of remotely sensed data has been agreed upon, though the CEOS (Committee on Earth Observation Satellites) format is becoming accepted as the standard. Digital remote sensing data are often organised using one of the three common formats used to organise image data.

For an instance an image consisting of four spectral channels, which can be visualised as four superimposed images, with corresponding pixels in one band registering exactly to those in the other bands. These common formats are:

- Band Interleaved by Pixel (BIP)
- Band Interleaved by Line (BIL)
- Band Sequential (BQ)

Digital image analysis is usually conducted using Raster data structures - each image is treated as an array of values. It offers advantages for manipulation of pixel values by image processing system, as it is easy to find and locate pixels and their values.

Disadvantages becomes apparent when one needs to represent the array of pixels as discrete patches or regions, where as Vector data structures uses polygonal patches and their boundaries as fundamental units for analysis and manipulation. Though vector format is not appropriate to for digital analysis of remotely sensed data.

9

Global Positioning System

The Global Positioning System (GPS) is a space-based global navigation satellite system (GNSS) that provides reliable location and time information in all weather and at all times and anywhere on or near the Earth when and where there is an unobstructed line of sight to four or more GPS satellites.

It is maintained by the United States government and is freely accessible by anyone with a GPS receiver. In addition to GPS other systems are in use or under development. The Russian GLObal NAvigation Satellite System (GLONASS) was for use by the Russian military only until 2007. There are also the planned Chinese Compass navigation system and Galileo positioning system of the European Union (EU). GPS was created and realized by the U.S. Department of Defense (DOD) and was originally run with 24 satellites. It was established in 1973 to overcome the limitations of previous navigation systems.

History

The first satellite navigation system, Transit, used by the United States Navy, was first successfully tested in 1960. It used a constellation of five satellites and could provide a navigational fix approximately once per hour. In 1967, the U.S. Navy developed the Timation satellite that proved the ability to place accurate clocks in space, a technology that GPS relies upon. In the 1970s, the ground-based Omega Navigation System, based on phase comparison of signal transmission from pairs of stations, became the first worldwide radio navigation system. However, limitations of these systems drove the need for a more universal navigation solution with greater accuracy.

While there were wide needs for accurate navigation in military and civilian sectors, almost none of those were seen as justification for the billions of dollars it would cost in research, development, deployment, and operation for a complex constellation of navigation satellites. However during the Cold War arms race, the nuclear threat to the very existence of the United States was the one need that did justify this cost in the view of the US Congress. This deterrent effect is why GPS was funded. The nuclear triad consisted of the US Navy's submarine-launched ballistic missiles (SLBMs) along with the US Air Force's strategic bombers and intercontinental ballistic missiles (ICBMs). Considered vital to the nuclear deterrence posture, accurate determination of the SLBM launch position was a force multiplier.

Precise navigation would enable US submarines to get an accurate fix of their positions prior to launching their SLBMs. The US Air Force with two-thirds of the nuclear triad also had requirements for a more accurate and reliable navigation system. The Navy and Air Force were developing their own technologies in parallel to solve what was essentially the same problem. To increase the survivability of ICBMs, there was a proposal to use mobile launch platforms so the need to fix the launch position had similarity to the SLBM situation.

In 1960, the Air Force proposed a radio-navigation system called MOSAIC (Mobile System for Accurate ICBM Control) that was essentially a 3-D LORAN. A follow-on study called Project 57 was worked in 1963 and it was "in this study that the GPS concept was born." That same year the concept was pursued as Project 621B, which had "many of the attributes that you now see in GPS" and promised increased accuracy for Air Force bombers as well as ICBMs. Updates from the Navy Transit system were too slow for the high speeds that the Air Force operated at. The Navy Research Laboratory continued advancements with their Timation (Time Navigation) satellites, first launched in 1967, and with the third one in 1974 carrying the first atomic clock put into orbit.

With these parallel developments out of the 1960s, it was realized that a superior system could be developed by synthesizing the best technologies from 621B, Transit, Timation, and SECOR in a multi-service program.

Over the Labour Day weekend in 1973, a meeting of about 12 military officers at the Pentagon discussed the creation of a *Defense Navigation Satellite System (DNSS)*. It was at this meeting that "the real synthesis that became GPS was created." Later that year, the DNSS program was named *Navstar*. With the individual satellites being associated with the name Navstar (as with the predecessors Transit and Timation), a more fully encompassing name was used to identify the constellation of Navstar satellites, *Navstar-GPS*, which was later shortened simply to GPS.

After Korean Air Lines Flight 007, carrying 269 people, was shot down in 1983 after straying into the USSR's prohibited airspace, in the vicinity of Sakhalin and Moneron Islands, President Ronald Reagan issued a directive making GPS freely available for civilian use, once it was sufficiently developed, as a common good. The first satellite was launched in 1989, and the 24th satellite was launched in 1994.

Initially, the highest quality signal was reserved for military use, and the signal available for civilian use was intentionally degraded ("Selective Availability", SA). This changed with U.S. President Bill Clinton ordering Selective Availability turned off at midnight May 1, 2000, improving the precision of civilian GPS from 100 meters (about 300 feet) to 20 meters (about 65 feet). The U.S. military by then had the ability to deny GPS service to potential adversaries on a regional basis. GPS is owned and operated by the U.S. Government as a national resource. Department of Defense (DOD) is the steward of GPS. *Interagency GPS Executive Board (IGEB)* oversaw GPS policy matters from 1996 to 2004. After that the *National Space-Based Positioning, Navigation and Timing Executive Committee* was established by presidential directive in 2004 to advise and coordinate federal departments and agencies on matters concerning the GPS and related systems. The executive committee is chaired jointly by the deputy secretaries of defense and transportation. Its membership includes equivalent-level officials from the departments of state, commerce, and homeland security, the joint chiefs of staff, and NASA. Components of the executive office of the president participate as observers to the executive committee, and the FCC chairman participates as a liaison.

DOD is required by law to "maintain a Standard Positioning Service (as defined in the federal radio navigation plan and the standard positioning service signal specification) that will be available on a continuous, worldwide basis," and "develop measures to prevent hostile use of GPS and its augmentations without unduly disrupting or degrading civilian uses."

Summary of satellites

Block	***Launch Period***	***Satellite launches***				***Currently in orbit and healthy***
		Suc-cess	***Fail-ure***	***In prep-aration***	***Plan-ned***	
I	1978–1985	10	1	0	0	0
II	1989–1990	9	0	0	0	0
IIA	1990–1997	19	0	0	0	10
IIR	1997–2004	12	1	0	0	12
IIR-M	2005–2009	8	0	0	0	7
IIF	2010–2011	1	0	11	0	1
IIIA	2014–?	0	0	0	12	0
IIIB		0	0	0	8	0
IIIC		0	0	0	16	0
Total	59	2	11	36	30	

Timeline and Modernization

- In 1972, the U.S. Air Force Central Inertial Guidance Test Facility (Holloman AFB), conducted developmental flight tests of two prototype GPS receivers over White Sands Missile Range, using ground-based pseudo-satellites.
- In 1978, the first experimental Block-I GPS satellite was launched.
- In 1983, after Soviet interceptor aircraft shot down the civilian airliner KAL 007 that strayed into prohibited airspace due to navigational errors, killing all 269 people on board, U.S. President Ronald Reagan announced that GPS would be made available for civilian uses once it was completed.
- By 1985, ten more experimental Block-I satellites had been launched to validate the concept.

- On February 14, 1989, the first modern Block-II satellite was launched.
- The Gulf War from 1990 to 1992, was the first conflict where GPS was widely used.
- In 1992, the 2nd Space Wing, which originally managed the system, was de-activated and replaced by the 50th Space Wing.
- By December 1993, GPS achieved initial operational capability.
- By January 17, 1994 a complete constellation of 24 satellites was in orbit.
- Full Operational Capability was declared by NAVSTAR in April 1995.
- In 1996, recognizing the importance of GPS to civilian users as well as military users, U.S. President Bill Clinton issued a policy directive declaring GPS to be a dual-use system and establishing an Interagency GPS Executive Board to manage it as a national asset.
- In 1998, U.S. Vice President Al Gore announced plans to upgrade GPS with two new civilian signals for enhanced user accuracy and reliability, particularly with respect to aviation safety and in 2000 the U.S. Congress authorized the effort, referring to it as *GPS III*.
- In 1998, GPS technology was inducted into the Space Foundation Space Technology Hall of Fame.
- On May 2, 2000 “Selective Availability” was discontinued as a result of the 1996 executive order, allowing users to receive a non-degraded signal globally.
- In 2004, the United States Government signed an agreement with the European Community establishing cooperation related to GPS and Europe’s planned Galileo system.
- In 2004, U.S. President George W. Bush updated the national policy and replaced the executive board with the National Executive Committee for Space-Based Positioning, Navigation, and Timing.
- November 2004, QUALCOMM announced successful tests of assisted GPS for mobile phones.

- In 2005, the first modernized GPS satellite was launched and began transmitting a second civilian signal (L2C) for enhanced user performance.
- On September 14, 2007, the aging mainframe-based Ground Segment Control System was transferred to the new Architecture Evolution Plan.
- On May 19, 2009, the U. S. Government Accountability Office issued a report warning that some GPS satellites could fail as soon as 2010.
- On May 21, 2009, the Air Force Space Command allayed fears of GPS failure saying "There's only a small risk we will not continue to exceed our performance standard."
- On January 11, 2010, an update of ground control systems caused a software incompatibility with 8000 to 10000 military receivers manufactured by a division of Trimble Navigation Limited of Sunnyvale, Calif.
- The most recent launch was on May 28, 2010. The oldest GPS satellite still in operation was launched on November 26, 1990, and became operational on December 10, 1990.

Structure

GPS consists of three parts: the space segment, the control segment, and the user segment. The U.S. Air Force develops, maintains, and operates the space and control segments. GPS satellites broadcast signals from space, which each GPS receiver uses to calculate its three-dimensional location (latitude, longitude, and altitude) plus the current time.

The space segment is composed of 24 to 32 satellites in medium Earth orbit and also includes the payload adapters to the boosters required to launch them into orbit. The control segment is composed of a master control station, an alternate master control station, and a host of dedicated and shared ground antennas and monitor stations. The user segment is composed of hundreds of thousands of U.S. and allied military users of the secure GPS Precise Positioning Service, and tens of millions of civil, commercial, and scientific users of the Standard Positioning Service.

Applications

While originally a military project, GPS is considered a *dual-use* technology, meaning it has significant military and civilian applications.

GPS has become a widely used and useful tool for commerce, scientific uses, tracking and surveillance. GPS's accurate timing facilitates everyday activities such as banking, mobile phone operations, and even the control of power grids. Farmers, surveyors, geologists and countless others perform their work more efficiently, safely, economically, and accurately.

Civilian

Many civilian applications use one or more of GPS's three basic components: absolute location, relative movement, and time transfer.

- Cellular telephony: Clock synchronization enables time transfer, which is critical for synchronizing its spreading codes with other base stations to facilitate inter-cell handoff and support hybrid GPS/cellular position detection for mobile emergency calls and other applications. The first handsets with integrated GPS launched in the late 1990s. The U.S. Federal Communications Commission (FCC) mandated the feature in 2002 so emergency services could locate 911 callers. Third-party software developers later gained access to GPS APIs from Nextel upon launch, followed by Sprint in 2006, and Verizon soon thereafter.
- Disaster relief/emergency services: Depend upon GPS for location and timing capabilities.
- Geofencing: Vehicle tracking systems, person tracking systems, and pet tracking systems use GPS to locate a vehicle, person, or pet. These devices attach to the vehicle, person, or the pet collar. The application provides 24/7 tracking and mobile or Internet updates should the trackee leave a designated area.
- Geotagging: Applying location coordinates to digital objects such as photographs and other documents for purposes such as creating map overlays.
- GPS Aircraft Tracking.

- GPS tours: Location determines which content to display; for instance, information about an approaching point of interest is displayed.
- Map-making: Both civilian and military cartographers use GPS extensively.
- Navigation: Navigators value digitally precise velocity and orientation measurements.
- Phasor measurement units: GPS enables highly accurate timestamping of power system measurements, making it possible to compute phasors.
- Recreation: For example, geocaching, geodashing, GPS drawing and waymarking.
- Surveying: Surveyors use absolute locations to make maps and determine property boundaries.
- Tectonics: GPS enables direct fault motion measurement in earthquakes.

RESTRICTIONS ON CIVILIAN USE

The U.S. Government controls the export of some civilian receivers. All GPS receivers capable of functioning above 18 kilometres (11 mi) altitude and 515 metres per second (1,001 kn) are classified as munitions (weapons) for which U.S. State Department export licenses are required. These limits attempt to prevent use of a receiver in a ballistic missile. They would not prevent use in a cruise missile since their altitudes and speeds are similar to those of ordinary aircraft. This rule applies even to otherwise purely civilian units that only receive the L1 frequency and the C/A (Clear/Acquisition) code and cannot correct for Selective Availability (SA), etc. Disabling operation above these limits exempts the receiver from classification as a munition. Vendor interpretations differ. The rule targets operation given the combination of altitude and speed, while some receivers stop operating even when stationary. This has caused problems with some amateur radio balloon launches, which regularly reach 30 kilometres (19 mi).

Military

As of 2009, military applications of GPS include:

- Navigation: GPS allows soldiers to find objectives, even in the dark or in unfamiliar territory, and to coordinate troop and supply movement. In the US armed forces, commanders use the *Commanders Digital Assistant* and lower ranks use the *Soldier Digital Assistant*.

- Target tracking: Various military weapons systems use GPS to track potential ground and air targets before flagging them as hostile. These weapon systems pass target coordinates to precision-guided munitions to allow them to engage targets accurately. Military aircraft, particularly in air-to-ground roles, use GPS to find targets (for example, gun camera video from AH-1 Cobras in Iraq show GPS co-ordinates that can be viewed with special software.
- Missile and projectile guidance: GPS allows accurate targeting of various military weapons including ICBMs, cruise missiles and precision-guided munitions. Artillery projectiles. Embedded GPS receivers able to withstand accelerations of 12,000 *g* or about 118 km/s^2 have been developed for use in 155 millimetres (6.1 in) howitzers.
- Search and Rescue: Downed pilots can be located faster if their position is known.
- Reconnaissance: Patrol movement can be managed more closely.
- GPS satellites carry a set of nuclear detonation detectors consisting of an optical sensor (Y-sensor), an X-ray sensor, a dosimeter, and an electromagnetic pulse (EMP) sensor (W-sensor), which form a major portion of the United States Nuclear Detonation Detection System.

Awards

Two GPS developers received the National Academy of Engineering Charles Stark Draper Prize for 2003:

- Ivan Getting, emeritus president of The Aerospace Corporation and engineer at the Massachusetts Institute of Technology, established the basis for GPS, improving on the World War II land-based radio system called LORAN (*Lo*ng-range *R*adio *A*id to *N*avigation).
- Bradford Parkinson, professor of aeronautics and astronautics at Stanford University, conceived the present satellite-based system in the early 1960s and

developed it in conjunction with the U.S. Air Force. Parkinson served twenty-one years in the Air Force, from 1957 to 1978, and retired with the rank of colonel.

GPS developer Roger L. Easton received the National Medal of Technology on February 13, 2006. On February 10, 1993, the National Aeronautic Association selected the GPS Team as winners of the 1992 Robert J. Collier Trophy, the nation's most prestigious aviation award. This team combines researchers from the Naval Research Laboratory, the U.S. Air Force, the Aerospace Corporation, Rockwell International Corporation, and IBM Federal Systems Company. The citation honors them "for the most significant development for safe and efficient navigation and surveillance of air and spacecraft since the introduction of radio navigation 50 years ago."

BASIC CONCEPT OF GPS

A GPS receiver calculates its position by precisely timing the signals sent by GPS satellites high above the Earth. Each satellite continually transmits messages that include:

- the time the message was transmitted
- precise orbital information (the ephemeris)
- the general system health and rough orbits of all GPS satellites (the almanac)

The receiver uses the messages it receives to determine the transit time of each message and computes the distance to each satellite. These distances along with the satellites' locations are used with the possible aid of trilateration, depending on which algorithm is used, to compute the position of the receiver. This position is then displayed, perhaps with a moving map display or latitude and longitude; elevation information may be included. Many GPS units show derived information such as direction and speed, calculated from position changes.

Three satellites might seem enough to solve for position, since space has three dimensions and a position near the Earth's surface can be assumed. However, even a very small clock error multiplied by the very large speed of light — the speed at which satellite signals propagate — results in a large positional error. Therefore receivers use four or more satellites to solve for the receiver's location and time. The very accurately computed time is effectively hidden by most GPS applications, which use only the location. A few specialized GPS applications do however use the time; these include time transfer, traffic signal timing, and synchronization of cell phone base stations.

Although four satellites are required for normal operation, fewer apply in special cases. If one variable is already known, a receiver can determine its position using only three satellites. For example, a ship or aircraft may have known elevation. Some GPS receivers may use additional clues or assumptions (such as reusing the last known altitude, dead reckoning, inertial navigation, or including information from the vehicle computer) to give a less accurate (degraded) position when fewer than four satellites are visible.

Position Calculation Introduction

To provide an introductory description of how a GPS receiver works, error effects are deferred to a later section. Using messages received from a minimum of four visible satellites, a GPS receiver is able to determine the times sent and then the satellite positions corresponding to these times sent. The x, y, and z components of position, and the time sent, are designated as $[x_i, y_i, z_i, t_i]$ where the subscript *i* is the satellite number and has the value 1, 2, 3, or 4. Knowing the indicated time the message was received t_r, the GPS receiver can compute the transit time of the message as $(t_r - t_i)$. Assuming the message traveled at the speed of light, *c*, the distance traveled or pseudorange, p_i can be computed as $(t_r - t_i)c$.

A satellite's position and pseudorange define a sphere, centered on the satellite with radius equal to the pseudorange. The position of the receiver is somewhere on the surface of this sphere. Thus with four satellites, the indicated position of the GPS receiver is at or near the intersection of the surfaces of four spheres. In the ideal case of no errors, the GPS receiver would be at a precise intersection of the four surfaces. The intersection of a third spherical surface with the first two will be its intersection with that circle; in most cases of practical interest, this means they intersect at two points. For automobiles and other near-earth vehicles, the correct position of the GPS receiver is the intersection closest to the Earth's surface. For space vehicles, the intersection farthest from Earth may be the correct one. The correct position for the GPS receiver is also the intersection closest to the surface of the sphere corresponding to the fourth satellite.

Correcting a GPS Receiver Clock

One of the most significant error sources is the GPS receiver's clock.

Because of the very large value of the speed of light, *c*, the estimated distances from the GPS receiver to the satellites, the pseudoranges, are very sensitive to errors in the GPS receiver clock; for example an error of one microsecond (0.000 001 second) corresponds to an error of 300 metres (980 ft). This suggests that an extremely accurate and expensive clock is required for the GPS receiver to work. Since manufacturers prefer to build inexpensive GPS receivers for mass markets, the solution for this dilemma is based on the way sphere surfaces intersect in the GPS problem.

It is likely that the surfaces of the three spheres intersect, since the circle of intersection of the first two spheres is normally quite large, and thus the third sphere surface is likely to intersect this large circle. It is very unlikely that the surface of the sphere corresponding to the fourth satellite will intersect either of the two points of intersection of the first three, since any clock error could cause it to miss intersecting a point. However, the distance from the valid estimate of GPS receiver position to the surface of the sphere corresponding to the fourth satellite can be used to compute a clock correction. Let r_4 denote the distance from the valid estimate of GPS receiver position to the fourth satellite and let p_4 denote the pseudorange of the fourth satellite. Let da = $r_4 - p_4$. da is the distance from the computed GPS receiver position to the surface of the sphere corresponding to the fourth satellite. Thus the quotient, b = da/c, provides an estimate of:

(correct time) – (time indicated by the receiver's on-board clock),

and the GPS receiver clock can be advanced if b is positive or delayed if is negative. However, it should be kept in mind that a less simple function of may be needed to estimate the time error in an iterative algorithm as discussed in the Navigation section.

SYSTEM SEGMENTATION

The current GPS consists of three major segments. These are the space segment (SS), a control segment (CS), and a user segment (US).

Space Segment

The space segment (SS) is composed of the orbiting GPS satellites, or Space Vehicles (SV) in GPS parlance. The GPS design originally called for 24 SVs, eight each in three circular orbital planes, but this was modified to six planes with four satellites each. The orbital planes are centered on the Earth, not rotating with respect to the distant stars. The six planes have approximately 55° inclination (tilt relative to Earth's equator) and are separated by 60° right ascension of the ascending node (angle along the equator from a reference point to the orbit's intersection). The orbits are arranged so that at least six satellites are always within line of sight from almost everywhere on Earth's surface. The result of this objective is that the four satellites are not evenly spaced (90 degrees) apart within each orbit. In general terms, the angular difference between satellites in each orbit is 30, 105, 120, and 105 degrees apart which, of course, sum to 360 degrees.

Orbiting at an altitude of approximately 20,200 kilometres (about 12,550 miles or 10,900 nautical miles; orbital radius of approximately 26,600 km (about 16,500 mi or 14,400 NM)), each SV makes two complete orbits each sidereal day, repeating the same ground track each day. This was very helpful during development, since even with just four satellites, correct alignment means all four are visible from one spot for a few hours each day. For military operations, the ground track repeat can be used to ensure good coverage in combat zones.

As of March 2008, there are 31 actively broadcasting satellites in the GPS constellation, and two older, retired from active service satellites kept in the constellation as orbital spares. The additional satellites improve the precision of GPS receiver calculations by providing redundant measurements. With the increased number of satellites, the constellation was changed to a nonuniform arrangement. Such an arrangement was shown to improve reliability and availability of the system, relative to a uniform system, when multiple satellites fail.

About eight satellites are visible from any point on the ground at any one time.

Control Segment

The control segment is composed of:

1. a master control station (MCS),
2. an alternate master control station,
3. four dedicated ground antennas,
4. six dedicated monitor stations.

The MCS can also access U.S. Air Force Satellite Control Network (AFSCN) ground antennas (for additional command and control capability) and NGA (National Geospatial-Intelligence Agency) monitor stations. The flight paths of the satellites are tracked by dedicated U.S. Air Force monitoring stations in Hawaii, Kwajalein, Ascension Island, Diego Garcia, Colorado Springs, Colorado and Cape Canaveral, along with shared NGA monitor stations operated in England, Argentina, Ecuador, Bahrain, Australia and Washington DC.

The tracking information is sent to the Air Force Space Command's MCS at Schriever Air Force Base 25 km (16 miles) ESE of Colorado Springs, which is operated by the 2nd Space Operations Squadron (2 SOPS) of the United States Air Force (USAF).

Then 2 SOPS contacts each GPS satellite regularly with a navigational update using dedicated or shared (AFSCN) ground antennas (GPS dedicated ground antennas are located at Kwajalein, Ascension Island, Diego Garcia, and Cape Canaveral). These updates synchronize the atomic clocks on board the satellites to within a few nanoseconds of each other, and adjust the ephemeris of each satellite's internal orbital model. The updates are created by a Kalman filter, which uses inputs from the ground monitoring stations, space weather information, and various other inputs.

Satellite maneuvers are not precise by GPS standards. So to change the orbit of a satellite, the satellite must be marked *unhealthy*, so receivers will not use it in their calculation. Then the maneuver can be carried out, and the resulting orbit tracked from the ground. Then the new ephemeris is uploaded and the satellite marked healthy again.

User Segment

The user segment is composed of hundreds of thousands of U.S. and allied military users of the secure GPS Precise Positioning Service, and tens of millions of civil, commercial and scientific users of the Standard Positioning Service. In general, GPS receivers are composed of an antenna, tuned to the frequencies transmitted by the satellites, receiver-processors, and a highly stable clock (often a crystal oscillator). They may also include a display for providing location and speed information to the user. A receiver is often described by its number of channels: this signifies how many satellites it can monitor simultaneously. Originally limited to four or five, this has progressively increased over the years so that, as of 2007, receivers typically have between 12 and 20 channels.

GPS receivers may include an input for differential corrections, using the RTCM SC-104 format. This is typically in the form of an RS-232 port at 4,800 bit/s speed. Data is actually sent at a much lower rate, which limits the accuracy of the signal sent using RTCM. Receivers with internal DGPS receivers can outperform those using external RTCM data. As of 2006, even low-cost units commonly include Wide Area Augmentation System (WAAS) receivers.

Many GPS receivers can relay position data to a PC or other device using the NMEA 0183 protocol. Although this protocol is officially defined by the National Marine Electronics Association (NMEA), references to this protocol have been compiled from public records, allowing open source tools like gpsd to read the protocol without violating intellectual property laws. Other proprietary protocols exist as well, such as the SiRF and MTK protocols. Receivers can interface with other devices using methods including a serial connection, USB, or Bluetooth.

COMMUNICATION

The navigational signals transmitted by GPS satellites encode a variety of information including satellite positions, the state of the internal clocks, and the health of the network. These signals are transmitted on two separate carrier frequencies that are common to all satellites in the network. Two different encodings are used, a public encoding that enables lower resolution navigation, and an encrypted encoding used by the U.S. military.

Message Format

GPS message format	
Subframes	**Description**
1	Satellite clock, GPS time relationship
2–3	Ephemeris (precise satellite orbit)
4–5	Almanac component (satellite network synopsys, error correction)

Each GPS satellite continuously broadcasts a *navigation message* at a rate of 50 bits per second. Each complete message is composed of 30-second frames, distinct groupings of 1,500 bits of information. Each frame is further subdivided into 5 subframes of length 6 seconds and with 300 bits each. Each subframe contains 10 words of 30 bits with length 0.6 seconds each. Each 30 second frame begins precisely on the minute or half minute as indicated by the atomic clock on each satellite. The first part of the message encodes the week number and the time within the week, as well as the data about the health of the satellite.

The second part of the message, the *ephemeris*, provides the precise orbit for the satellite. The last part of the message, the *almanac*, contains coarse orbit and status information for all satellites in the network as well as data related to error correction.

All satellites broadcast at the same frequencies. Signals are encoded using code division multiple access (CDMA) allowing messages from individual satellites to be distinguished from each other based on unique encodings for each satellite (which the receiver must be aware of). Two distinct types of CDMA encodings are used: the coarse/acquisition (C/A) code, which is accessible by the general public, and the precise (P) code, that is encrypted so that only the U.S. military can access it.

The ephemeris is updated every 2 hours and is generally valid for 4 hours, with provisions for updates every 6 hours or longer in non-nominal conditions. The almanac is updated typically every 24 hours. Additionally data for a few weeks following is uploaded in case of transmission updates that delay data upload.

Satellite Frequencies

All satellites broadcast at the same two frequencies, 1.57542 GHz (L1 signal) and 1.2276 GHz (L2 signal). The satellite network uses a CDMA spread-spectrum technique where the low-bitrate message data is encoded with a high-rate pseudo-random (PRN) sequence that is different for each satellite. The receiver must be aware of the PRN codes for each satellite to reconstruct the actual message data. The C/A code, for civilian use, transmits data at 1.023 million chips per second, whereas the P code, for U.S. military use, transmits at 10.23 million chips per second. The L1 carrier is modulated by both the C/A and P codes, while the L2 carrier is only modulated by the P code. The P code can be encrypted as a so-called P(Y) code which is only available to military equipment with a proper decryption key. Both the C/A and P(Y) codes impart the precise time-of-day to the user. GPS Modernization added a third frequency, 1.17645 GHZ (L5 signal). The L5 consists of two carrier components that are in phase quadrature with each other. Each carrier component is bi-phase shift key (BPSK) modulated by a separate bit train.

Demodulation and Decoding

Since all of the satellite signals are modulated onto the same L1 carrier frequency, there is a need to separate the signals after demodulation. This is done by assigning each satellite a unique binary sequence known as a Gold code. The signals are decoded, after demodulation, using addition of the Gold codes corresponding to the satellites monitored by the receiver.

If the almanac information has previously been acquired, the receiver picks which satellites to listen for by their PRNs, unique numbers in the range 1 through 32. If the almanac information is not in memory, the receiver enters a search mode until a lock is obtained on one of the satellites. To obtain a lock, it is necessary that there be an unobstructed line of sight from the receiver to the satellite. The receiver can then acquire the almanac and determine the satellites it should listen for. As it detects each satellite's signal, it identifies it by its distinct

C/A code pattern. There can be a delay of up to 30 seconds before the first estimate of position because of the need to read the ephemeris data. Processing of the navigation message enables the determination of the time of transmission and the satellite position at this time.

NAVIGATION EQUATIONS

The receiver uses messages received from four satellites to determine the satellite positions and time sent. The *x, y,* and *z* components of position and the time sent are designated as where the subscript *i* denotes the satellite and has the value 1, 2, 3, or 4. Knowing when the message was received, the receiver computes the message's transit time as. Assuming the message traveled at the speed of light (c) the distance traveled, is. Knowing the distance from receiver to satellite and the satellite's position implies that the receiver is on the surface of a sphere centered at the satellite's position. Thus the receiver is at or near the intersection of the surfaces of four spheres. In the ideal case of no errors, the receiver is at the intersection of the surfaces of four spheres. Excluding the unrealistic case (for GPS purposes) of two coincident spheres, the surfaces of two intersecting spheres is either a point (if they merely touch) or a circle as depicted in the illustration below. Two of the points at which the surfaces of the spheres intersect are clearly marked. The distance between these two points is the diameter of the circle of intersection.

This can be seen more clearly by considering a side view of the intersecting spheres. The symmetry of the spheres. A view from any horizontal direction would look exactly the same. Therefore the diameter as seen from all directions is the same and thus the surfaces actually do intersect in a circle. Having found that two sphere surfaces intersect in a circle, we now consider how the intersection of the first two sphere surfaces, the circle, intersect with the third sphere. A circle and sphere surface intersect at zero, one or two points. For the GPS problem we are concerned with the case of two points of intersection. Trilateration algebraically confirms this geometric observation. The ambiguity of two points of intersection of three sphere surfaces can be resolved by noting which point is closest to the fourth sphere surface.

Let denote the clock error or bias, the amount by which the receiver's clock is off. The receiver has four unknowns, the three components of GPS receiver position and the clock bias. The equation of the sphere surfaces are given by:

Another useful form of these equations is in terms of *pseudoranges*, which are the approximate ranges based on the receiver clock's uncorrected time so that. Then the equations becomes:

Methods of Solution of Navigation Equations

- Bancroft's method is perhaps the most important method of solving the navigation equations since it involves an algebraic as opposed to numerical method. The method requires at least four satellites but more can be used. Two numerical methods of computing GPS receiver position and clock bias are (1) by using trilateration and one dimensional numerical root finding and (2) multidimensional Newton-Raphson calculations.
- The receiver can solve by trilateration and one dimensional numerical root finding. Trilateration determines the intersection of the surfaces of three spheres. In the usual case of two intersections, the point nearest the surface of the sphere corresponding to the fourth satellite is chosen. The Earth's surface can also sometimes be used instead, especially by civilian GPS receivers, since it is illegal in the United States to track vehicles more than 60,000 feet (18,000 m) in altitude. Let da denote the signed magnitude of the vector from the receiver position to the fourth satellite (i.e. da = r4-p4) as defined in the section "Clock correction". da is a function of the correction since the correction changes the satellite transmission times and thus the pseudoranges. The notation, da(correction) denotes this function. The problem is to determine the correction such that.

  This is the familiar problem of finding the zeroes of a one dimensional non-linear function of a scalar variable. Iterative numerical methods, such as those found in the chapter on root finding in *Numerical Recipes* can solve this type of problem. One advantage of this method is that it involves one dimensional as opposed to multidimensional numerical root finding.

- Alternatively, multidimensional root finding method such as Newton-Raphson method can be used. The approach is to linearize around an approximate solution, say from iteration *k,* then solve four linear equations derived from the quadratic equations above to obtain. The Newton-Raphson method is more rapidly convergent than other methods of numerical root finding. A disadvantage of this multidimensional root finding method as compared to single dimensional root findiing is that, "There are no good general methods for solving systems of more than one nonlinear equations."
- When more than four satellites are available, the calculation can use the four best or more than four, considering number of channels, processing capability, and geometric dilution of precision (GDOP). Using more than four is an over-determined system of equations with no unique solution, which must be solved by least-squares or a similar technique. If all visible satellites are used, the results are as good as or better than using the four best. Errors can be estimated through the residuals. With each combination of four or more satellites, a GDOP factor can be calculated, based on the relative sky directions of the satellites used. As more satellites are picked up, pseudoranges from various 4-way combinations can be processed to add more estimates to the location and clock offset. The receiver then takes the weighted average of these positions and clock offsets. After the final location and time are calculated, the location is expressed in a specific coordinate system such as latitude and longitude, using the WGS 84 geodetic datum or a country-specific system.
- Finally, results from other positioning systems such as GLONASS or the upcoming Galileo can be incorporated or used to check the result. (By design, these systems use the same frequency bands, so much of the receiver circuitry can be shared, though the decoding is different).

ERROR SOURCES AND ANALYSIS

The positioning data provided directly by the satellites is extremely precise but there are many factors that can make the errors in the data non-trivial. In situations where high accuracy is necessary, understanding and compensating for these sources of error is important. Sources of error include atmospheric distortion (predominantly in the ionosphere), satellite clock inaccuracies, and the travel delays of the satellite signals. The analysis of errors in the information reported by the Global Positioning System, a space-based satellite system for navigation, is important to estimating the accuracy of position estimates and correcting for the errors. The Global Positioning System (GPS) was created by the United States Department of Defense (DOD) in the 1970s. It has come to be widely used for navigation both by the U.S. military and the general public.

ACCURACY ENHANCEMENT AND SURVEYING

Augmentation

Integrating external information into the calculation process can materially improve accuracy. Such augmentation systems are generally named or described based on how the information arrives. Some systems transmit additional error information (such as clock drift, ephemera, or ionospheric delay), others characterize prior errors, while a third group provides additional navigational or vehicle information. Examples of augmentation systems include the Wide Area Augmentation System (WAAS), European Geostationary Navigation Overlay Service (EGNOS), Differential GPS, Inertial Navigation Systems (INS) and Assisted GPS.

Precise Monitoring

Accuracy can be improved through precise monitoring and measurement of existing GPS signals in additional or alternate ways. The largest remaining error is usually the unpredictable delay through the ionosphere. The spacecraft broadcast ionospheric model parameters, but errors remain. This is one reason GPS spacecraft transmit on at least two frequencies, L1 and L2. Ionospheric delay is a well-defined function of frequency and the total electron content (TEC) along the path, so measuring the arrival time difference between the frequencies determines TEC and thus the precise ionospheric delay at each frequency.

Military receivers can decode the P(Y)-code transmitted on both L1 and L2. Without decryption keys, it is still possible to use a *codeless* technique to compare the P(Y) codes on L1 and L2 to gain much of the same error information. However, this technique is slow, so it is currently available only on specialized surveying equipment. In the future, additional civilian codes are expected to be transmitted on the L2 and L5 frequencies. Then all users will be able to perform dual-frequency measurements and directly compute ionospheric delay errors. A second form of precise monitoring is called *Carrier-Phase Enhancement* (CPGPS). This corrects the error that arises because the pulse transition of the PRN is not instantaneous, and thus the correlation (satellite-receiver sequence matching) operation is imperfect. CPGPS uses the L1 carrier wave, which has a period of which is about one-thousandth of the C/A Gold code bit period of, to act as an additional clock signal and resolve the uncertainty. The phase difference error in the normal GPS amounts to 2–3 metres (6.6–9.8 ft) of ambiguity. CPGPS working to within 1% of perfect transition reduces this error to 3 centimeters (1.2 in) of ambiguity. By eliminating this error source, CPGPS coupled with DGPS normally realizes between 20–30 centimetres (7.9–12 in) of absolute accuracy.

Relative Kinematic Positioning (RKP) is a third alternative for a precise GPS-based positioning system. In this approach, determination of range signal can be resolved to a precision of less than 10 centimeters (3.9 in). This is done by resolving the number of cycles in which the signal is transmitted and received by the receiver. This can be accomplished by using a combination of differential GPS (DGPS) correction data, transmitting GPS signal phase information and ambiguity resolution techniques via statistical tests—possibly with processing in real-time (real-time kinematic positioning, RTK).

Timekeeping

While most clocks are synchronized to Coordinated Universal Time (UTC), the atomic clocks on the satellites are set to *GPS time*. The difference is that GPS time is not corrected to match the rotation of the Earth, so it does not contain leap seconds or other corrections that are periodically added to UTC. GPS time was set to match Coordinated Universal Time (UTC) in 1980, but has since diverged. The lack of corrections means that GPS time remains at a constant offset with International Atomic Time (TAI) (TAI-GPS = 19 seconds on November 21, 2010). Periodic corrections are performed on the on-board clocks to correct relativistic effects and keep them synchronized with ground clocks. The GPS navigation message includes the difference between GPS time and UTC, which as of 2009 is 15 seconds due to the leap second added to UTC December 31, 2008. Receivers subtract this offset from GPS time to calculate UTC and specific timezone values. New GPS units may not show the correct UTC time until after receiving the UTC offset message. The GPS-UTC offset field can accommodate 255 leap seconds (eight bits) which, given the current rate of change of the Earth's rotation (with one leap second introduced approximately every 18 months), should be sufficient to last until approximately the year 2300.

As opposed to the year, month, and day format of the Gregorian calendar, the GPS date is expressed as a week number and a seconds-into-week number. The week number is transmitted as a ten-bit field in the C/A and P(Y) navigation messages, and so it becomes zero again every 1,024 weeks (19.6 years). GPS week zero started at 00:00:00 UTC (00:00:19 TAI) on January 6, 1980, and the week number became zero again for the first time at 23:59:47 UTC on August 21, 1999 (00:00:19 TAI on August 22, 1999). To determine the current Gregorian date, a GPS receiver must be provided with the approximate date (to within 3,584 days) to correctly translate the GPS date signal. To address this concern the modernized GPS navigation message uses a 13-bit field, which only repeats every 8,192 weeks (157 years), thus lasting until the year 2137 (157 years after GPS week zero).

10

Geographic and Demographic Setting

India is a Country of great diversity with a wide range of landform types, including major mountain ranges, deserts, rich agricultural plains, and hilly jungle regions. Indeed, the term *Indian subcontinent* aptly describes the enormous extent of the earth's surface that India occupies, and any attempt to generalize about its physiography is inaccurate. Diversity is also evident in the geographical distribution of India's ethnic and linguistic groups. In ancient times, the major river valleys of the Indo-Gangetic Plain of South Asia were among the great cradles of civilization in Asia, as were the valleys of the Tigris and Euphrates rivers in West Asia and the Huang He (Yellow River) in East Asia. As a result of thousands of years of cultural and political expansion and amalgamation, contemporary India has come to include many different natural and cultural regions.

The Himalayas (and the nations of Nepal and Bhutan) form India's northern frontier with China. Pakistan borders India to the west and Bangladesh (formerly East Pakistan) to the east. Although both were formerly part of the British Indian Empire, India and Pakistan became separate countries in 1947 and East Pakistan became independent Bangladesh in 1971. The boundaries of the Indian polity are not fully demarcated because of regional ethnic and political disputes and are the source of occasional tensions.

When the 1991 national census was taken, India's population was approximately 846.3 million. The annual population growth rate from 1981 to 1991 was 2 percent. Accounting for only 2.4

percent of the world's landmass, India is home to 16 percent of the world's population. Every sixth person in the world in the early 1990s was an Indian. It is generally assumed that India's population will surpass the 1 billion mark some time before the next census in 2001. In July 1995, the population was estimated at 936.5 million.

Some 38 percent of all Indians were officially listed as living below the poverty line in fiscal year 1991. This number represented an increase from the low mark of 26 percent in FY 1989, but the rise was believed to be only temporary by some observers. Although government-sponsored health clinics are widely available in the mid-1990s, their emphasis is on curative techniques rather than preventive medicine.

However, the lack of such basic amenities as safe, potable water for much of the population is indicative of the severity of health problems. This situation has traditionally led most Indians to have large families as their only form of insurance against sickness and for their care in old age. Although family planning programs are becoming integrated with the programs of urban and rural health clinics, no official birth control programs have widespread support. The severity of the acquired immune deficiency syndrome (AIDS) epidemic in India has become increasingly apparent to health specialists, but local awareness of the causes of and ways to prevent the spread of AIDS is growing slowly.

Although many public schools are inadequate, improvements to the education system overall have been substantial since 1947. In the mid-1990s, however, only about 50 percent of children between the ages of six and fourteen are enrolled in schools. The goal of compulsory and free primary and middle school education is embodied in the Indian constitution but has been elusive. The National Policy on Education of 1986 sought to institutionalize universal primary education by setting 1990 as a target date for the education of all children up to eleven years of age. The ability of India's education system to meet this goal has been constrained by lack of adequate financial resources. Important achievements have been made, however, with implementation of the nonformal education system and adult education programs. Whereas public education is generally below standard, education standards in private schools are very high. There also are high standards among the elite institutions in the higher education system.

Principal Regions

India's total land mass is 2,973,190 square kilometres and is divided into three main geological regions: the Indo-Gangetic Plain, the Himalayas, and the Peninsula region. The Indo-Gangetic Plain and those portions of the Himalayas within India are collectively known as North India. South India consists of the peninsular region, often termed simply the Peninsula. On the basis of its physiography, India is divided into ten regions: the Indo-Gangetic Plain, the northern mountains of the Himalayas, the Central Highlands, the Deccan or Peninsular Plateau, the East Coast (Coromandel Coast in the south), the West Coast (Konkan, Kankara, and Malabar coasts), the Great Indian Desert (a geographic feature known as the Thar Desert in Pakistan) and the Rann of Kutch, the valley of the Brahmaputra in Assam, the northeastern hill ranges surrounding the Assam Valley, and the islands of the Arabian Sea and the Bay of Bengal.

Indo-Gangetic Plain

In social and economic terms, the Indo-Gangetic Plain is the most important region of India. The plain is a great alluvial crescent stretching from the Indus River system in Pakistan to the Punjab Plain (in both Pakistan and India) and the Haryana Plain to the delta of the Ganga (or Ganges) in Bangladesh (where it is called the Padma). Topographically the plain is homogeneous, with only floodplain bluffs and other related features of river erosion and changes in river channels forming important natural features.

Two narrow terrain belts, collectively known as the Terai, constitute the northern boundary of the Indo-Gangetic Plain. Where the foothills of the Himalayas encounter the plain, small hills known locally as *ghar* (meaning house in Hindi) have been formed by coarse sands and pebbles deposited by mountain streams. Groundwater from these areas flows on the surface where the plains begin and converts large areas along the rivers into swamps. The southern boundary of the plain begins along the edge of the Great Indian Desert in the state of

Rajasthan and continues east along the base of the hills of the Central Highlands to the Bay of Bengal. The hills, varying in elevation from 300 to 1,200 meters, lie on a general east-west axis. The Central Highlands are divided into northern and southern parts. The northern part is centered on the Aravalli Range of eastern Rajasthan. In the northern part of the state of Madhya Pradesh, the Malwa Plateau comprises the southern part of the Central Highlands and merges with the Vindhya Range to the south. The main rivers that flow through the southern part of the plain—the Narmada, the Tapti, and the Mahanadi—delineate North India from South India.

Some geographers subdivide the Indo-Gangetic Plain into three parts: the Indus Valley (mostly in Pakistan), the Punjab (divided between India and Pakistan) and Haryana plains, and the middle and lower Ganga. These regional distinctions are based primarily on the availability of water. By another definition, the Indo-Gangetic Plain is divided into two drainage basins by the Delhi Ridge; the western part consists of the Punjab Plain and the Haryana Plain, and the eastern part consists of the Ganga-Brahmaputra drainage systems. This divide is only 300 meters above sea level, contributing to the perception that the Indo-Gangetic Plain appears to be continuous between the two drainage basins. The Punjab Plain is centered in the land between five rivers: the Jhelum, the Chenab, the Ravi, the Beas, and the Sutlej.

Both the Punjab and Haryana plains are irrigated with water from the Ravi, Beas, and Sutlej rivers. The irrigation projects emanating from these rivers have led to a decrease in the flow of water reaching the lower drainage areas in the state of Punjab in India and the Indus Valley in Pakistan. The benefits that increased irrigation has brought to farmers in the state of Haryana are controversial in light of the effects that irrigation has had on agricultural life in the Punjab areas of both India and Pakistan.

The middle Ganga extends from the Yamuna River in the west to the state of West Bengal in the east. The lower Ganga and the Assam Valley are more lush and verdant than the middle Ganga. The lower Ganga is centered in West Bengal from which it flows into Bangladesh and, after joining the Jamuna (as the lower reaches of the Brahmaputra are known in Bangladesh), forms the delta of the Ganga. The Brahmaputra (meaning son of Brahma) rises in Tibet (China's Xizang Autonomous Region) as the Yarlung Zangbo River, flows through Arunachal Pradesh and Assam, and then crosses into Bangladesh. Average annual rainfall increases moving west to east from approximately 600 millimetres in the Punjab Plain to 1,500 millimetres around the lower Ganga and Brahmaputra.

The Himalayas

The Himalayas, the highest mountain range in the world, extend along the northern frontiers of Pakistan, India, Nepal, Bhutan, and Burma. They were formed geologically as a result of the collision of the Indian subcontinent with Asia. This process of plate tectonics is ongoing, and the gradual northward drift of the Indian subcontinent still causes earthquakes. Lesser ranges jut southward from the main body of the Himalayas at both the eastern and western ends.

The Himalayan system, about 2,400 kilometres in length and varying in width from 240 to 330 kilometres, is made up of three parallel ranges—the Greater Himalayas, the Lesser Himalayas, and the Outer Himalayas—sometimes collectively called the Great Himalayan Range. The Greater Himalayas, or northern range, average approximately 6,000 meters in height and contain the three highest mountains on earth: Mount Everest (8,796 meters) on the China-Nepal border; K2 (8,611 meters, also known as Mount Godwin-Austen, and in China as Qogir Feng) in an area claimed by India, Pakistan, and China; and Kanchenjunga (8,598 meters) on the India-Nepal border.

Many major mountains are located entirely within India, such as Nanda Devi (7,817 meters) in the state of Uttar Pradesh. The snow line averages 4,500 to 6,000 meters on the southern side of the Greater Himalayas and 5,500 to 6,000 on the northern side. Because of climatic conditions, the snow line in the eastern Himalayas averages 4,300 meters, while in the western Himalayas it averages 5,800 meters.

The Lesser Himalayas, located in northwestern India in the states of Himachal Pradesh and Uttar Pradesh, in north-central India in the state of Sikkim, and in northeastern India in the state of Arunachal Pradesh, range from 1,500 to 5,000 meters in height. Located in the Lesser Himalayas are the hill stations of Shimla (Simla) and Darjiling (Darjeeling). During the colonial period, these and other hill stations were used by the British as summer retreats to escape the intense heat of the plains. It is in this transitional vegetation zone that the contrasts between the bare southern slopes and the forested northern slopes become most noticeable.

The Outer or Southern Himalayas, averaging 900 to 1,200 meters in elevation, lie between the Lesser Himalayas and the Indo-Gangetic Plain. In Himachal Pradesh and Uttar Pradesh, this southernmost range is often referred to as the Siwalik Hills. It is possible to identify a fourth, and northernmost range, known as the Trans-Himalaya.

This range is located entirely on the Qinghai-Xizang Plateau, north of the great west-to-east trending valley of the Yarlung Zangbo River. Although the Trans-Himalaya Range is divided from the Great Himalayan Range for most of its length, it merges with the Great Himalayan Range in the western section—the Karakoram Range—where India, Pakistan, and China meet.

The southern slopes of each of the Himalayan ranges are too steep to accumulate snow or support much tree life; the northern slopes generally are forested below the snow line. Between the ranges are extensive high plateaus, deep gorges, and fertile valleys, such as the vales of Kashmir and Kulu. The Himalayas serve a very important purpose. They provide a physical screen within which the monsoon system operates and are the source of the great river systems that water the alluvial plains below. As a result of erosion, the rivers coming from the mountains carry vast quantities of silt that enrich the plains.

The area of northeastern India adjacent to Burma and Bangladesh consists of numerous hill tracts, averaging between 1,000 and 2,000 meters in elevation, that are not associated with the eastern part of the Himalayas in Arunachal Pradesh. The Naga Hills, rising to heights of more than 3,000 meters, form the watershed between India and Burma. The Mizo Hills are the southern part of the northeastern ranges in India. The Garo, Khasi, and Jaintia hills are centered in the state of Meghalaya and, isolated from the northeastern ranges, divide the Assam Valley from Bangladesh to the south and west.

The Peninsula

The Peninsula proper is an old, geologically stable region with an average elevation between 300 and 1,800 meters. The Vindhya Range constitutes the main dividing line between the geological regions of the Indo-Gangetic Plain and the Peninsula. This range lies north of the Narmada River, and when viewed from there, it is possible to discern the prominent escarpments that rise between 800 and 1,400 meters. The Vindhya Range defines the north-central and northwestern boundary of the Peninsula, and the Chota Nagpur Plateau of southern Bihar forms the northeastern boundary. The uplifting of the plateau of the central Peninsula and its eastward tilt formed the Western Ghats, a line of hills running from the Tapti River south to the tip of the Peninsula. The Eastern Ghats mark the eastern end of the plateau; they begin in the hills of the Mahanadi River basin and converge with the Western Ghats at the Peninsula's southern tip.

The interior of the Peninsula, south of the Narmada River, often termed the Deccan Plateau or simply the Deccan (from the Sanskrit *daksina*, meaning south), is a series of plateaus topped by rolling hills and intersected by many rivers. The plateau averages roughly 300 to 750 meters in elevation. Its major rivers—the Godavari, the Krishna, and the Kaveri—rise in the Western Ghats and flow eastward into the Bay of Bengal.

The coastal plain borders the plateau. On the northwestern side, it is characterized by tidal marshes, drowned valleys, and estuaries; and in the south by lagoons, marshes, and beach ridges. Coastal plains on the eastern side are wider than those in the west; they are focused on large river deltas that serve as the centers of human settlement.

Offshore Islands

India's offshore islands, constituting roughly one-quarter of 1 percent of the nation's territory, lie in two groups located off the east and west coasts. The northernmost point of the union territory of the Andaman and Nicobar Islands lies 1,100 kilometres southeast of Calcutta. Situated in the Bay of Bengal in a chain stretching some 800 kilometres, the Andaman Islands comprise 204 islands and islets, and their topography is characterized by hills and narrow valleys. Although their location is tropical, the climate of the islands is tempered by sea breezes; rainfall is irregular. The Nicobar Islands, which are south of the Andaman Islands, comprise nineteen islands, some with flat, coral-covered surfaces and others with hills. The islands have a nearly equatorial climate, heavy rainfall, and high temperatures. The union territory of Lakshadweep (the name means 100,000 islands) in the Arabian Sea, comprises—from north to south—the Amindivi, Laccadive, Cannanore, and Minicoy islands. The islands, only ten of which are inhabited, are spread throughout an area of approximately 77,000 square kilometres. The islands are low-lying coral-based formations capable of limited cultivation.

Coasts and Borders

India has 7,000 kilometres of seacoast and shares 14,000 kilometres of land frontier with six nations: Pakistan, China, Nepal, Bhutan, Bangladesh, and Burma. India claims a twelve-nautical-mile territorial sea and an exclusive economic zone of 200 nautical miles. The territorial seas total 314,400 square kilometres. In the mid-1990s, India had boundary disagreements with Pakistan, China, and Bangladesh; border distances are therefore approximations. The partition of India in 1947 established two India-Pakistan frontiers: one on the west and one on the east.

Disputes over the state of Jammu and Kashmir led to hostilities between India and Pakistan in 1947. The January 1, 1949, cease-fire arranged by the United Nations (UN) divided control of Kashmir. India controls Jammu, the Vale of Kashmir, and the capital, Srinagar, while Pakistan controls the mountainous area to the northwest. Neither side accepts a divided Kashmir as a permanent solution. India regards as illegal the 1963 China-Pakistan border agreement, which ceded to China a portion of Pakistani-controlled Kashmir. The two sides also dispute the Siachen Glacier near the Karakoram Pass. Further India-Pakistan hostilities in the 1965 war were settled through the Soviet-brokered Tashkent Declaration.

In 1968 an international tribunal settled the dispute over the Rann of Kutch, a region of salt flats that is submerged for six months of the year in the state of Gujarat. The following year, a new border was demarcated that recognized Pakistan’s claim to about 10 percent of the area. In 1992 India completed fencing most of the 547-kilometer-long section of the boundary between the Indian state of Punjab and the Pakistani province of Punjab. This measure was undertaken because of the continuing unrest in the region caused by both ethnic and religious disputes among the local Indian population and infiltrators from both sides of the frontier. The more rugged terrain north of Punjab along the entire cease-fire line between India and Pakistan in Jammu and Kashmir continues to be subject to infiltration and local strife.

The 2,000-kilometer-long border with China has eastern, central, and western sections. In the western section, the border regions of Jammu and Kashmir have been the scene of conflicting claims since the nineteenth century. China has not accepted India's definitions of the boundary and has carried out defense and economic activities in parts of eastern Kashmir since the 1950s. In the 1960s, China finished construction of a motor road across Aksai Chin (a region under dispute between India and China), the main transportation route linking China's Xinjiang-Uygur Autonomous Region and Tibet. In the eastern section, the China-India boundary follows the McMahon Line laid down in 1914 by Sir Arthur Henry McMahon, the British plenipotentiary to a conference of Indian, British, and Chinese representatives at Simla (now known as Shimla, Himachal Pradesh). The Simla Convention, as the agreement is known, set the boundary between India and Tibet. Although the British and Tibetan representatives signed the agreement on July 3, 1914, the Chinese delegate declined to sign. The line agreed to by Britain and Tibet generally follows the crest of the eastern Himalayas from Bhutan to Burma. It serves as a legal boundary, although the Chinese have never formally accepted it. China continued to claim roughly the entire area of Arunachal Pradesh south of the McMahon Line in the early 1990s. In 1962 China and India fought a brief border war in this region, and China occupied certain areas south of the line for several months. India and China took a major step toward resolving their border disputes in 1981 by opening negotiations on the issue. Agreements and talks held in 1993 and 1995 eased tensions along the India-China border. Sikkim, which became an Indian state in 1975, forms the small central section of India's northern border and lies between Nepal and Bhutan. India's border with Bangladesh is essentially the same as it was before East Pakistan became Bangladesh in 1971. Some minor disputes continued to occur over the size and number of the numerous enclaves each country had on either side of the border. These enclaves were established during the period from 1661 to 1712 during fighting between the Mughal Empire and the principality of Cooch Behar. This complex pattern of enclaves was preserved by the British administration and passed on intact to India and Pakistan.

The 1,300-kilometer frontier with Burma has been delimited but not completely demarcated. On March 10, 1967, the Indian and Burmese governments signed a bilateral treaty delimiting the boundary in detail. India also has a maritime boundary with Burma in the area of the northern Andaman Islands and Burma's Coco Islands in the Bay of Bengal. India's borders with Nepal and Bhutan have remained unchanged since the days of British rule. In 1977 India signed an accord with Indonesia demarcating the entire maritime boundary between the two countries. One year earlier, a similar accord was signed with the Maldives.

Rivers

The country's rivers are classified as Himalayan, peninsular, coastal, and inland-drainage basin rivers. Himalayan rivers are snow fed and maintain a high to medium rate of flow throughout the year. The heavy annual average rainfall levels in the Himalayan catchment areas further add to their rates of flow. During the monsoon months of June to September, the catchment areas are prone to flooding. The volume of the rain-fed peninsular rivers also increases. Coastal streams, especially in the west, are short and episodic. Rivers of the inland system, centered in western Rajasthan state, are few and frequently disappear in years of scant rainfall. The majority of the South Asia's major rivers flow through broad, shallow valleys and drain into the Bay of Bengal.

The Ganga River basin, India's largest, includes approximately 25 percent of the nation's area; it is bounded by the Himalayas in the north and the Vindhya Range to the south. The Ganga has its source in the glaciers of the Greater

Himalayas, which form the frontier between India and Tibet in northwestern Uttar Pradesh. Many Indians believe that the legendary source of the Ganga, and several other important Asian rivers, lies in the sacred Mapam Yumco Lake (known to the Indians as Manasarowar Lake) of western Tibet located approximately 75 kilometres northeast of the India-China-Nepal tripoint. In the northern part of the Ganga River basin, practically all of the tributaries of the Ganga are perennial streams. However, in the southern part, located in the states of Rajasthan and Madhya Pradesh, many of the tributaries are not perennial.

The Brahmaputra has the greatest volume of water of all the rivers in India because of heavy annual rainfall levels in its catchment basin. At Dibrugarh the annual rainfall averages 2,800 millimetres, and at Shillong it averages 2,430 millimetres. Rising in Tibet, the Brahmaputra flows south into Arunachal Pradesh after breaking through the Great Himalayan Range and dropping rapidly in elevation. It continues to fall through gorges impassable by man in Arunachal Pradesh until finally entering the Assam Valley where it meanders westward on its way to joining the Ganga in Bangladesh.

The Mahanadi, rising in the state of Madhya Pradesh, is an important river in the state of Orissa. In the upper drainage basin of the Mahanadi, which is centered on the Chhattisgarh Plain, periodic droughts contrast with the situation in the delta region where floods may damage the crops in what is known as the rice bowl of Orissa. Hirakud Dam, constructed in the middle reaches of the Mahanadi, has helped in alleviating these adverse effects by creating a reservoir. The source of the Godavari is northeast of Bombay (Mumbai in the local Marathi language) in the state of Maharashtra, and the river follows a southeasterly course for 1,400 kilometres to its mouth on the Andhra Pradesh coast. The Godavari River basin area is second in size only to the Ganga; its delta on the east coast is also one of the country’s main rice-growing areas. It is known as the “Ganga of the South,” but its discharge, despite the large catchment area, is moderate because of the medium levels of annual rainfall, for example, about 700 millimetres at Nasik and 1,000 millimetres at Nizamabad. The Krishna rises in the Western Ghats and flows east into the Bay of Bengal. It has a poor flow because of low levels of rainfall in its catchment area—660 millimetres annually at Pune. Despite its low discharge, the Krishna is the third longest river in India.

The source of the Kaveri is in the state of Karnataka, and the river flows southeastward. The waters of the river have been a source of irrigation since antiquity; in the early 1990s, an estimated 95 percent of the Kaveri was diverted for agricultural use before emptying into the Bay of Bengal. The delta of the Kaveri is so mature that the main river has almost lost its link with the sea, as the Kollidam, the distributary of the Kaveri, bears most of the flow.

The Narmada and the Tapti are the only major rivers that flow into the Arabian Sea. The Narmada rises in Madhya Pradesh and crosses the state, passing swiftly through a narrow valley between the Vindhya Range and spurs of the Satpura Range. It flows into the Gulf of Khambhat (or Cambay). The shorter Tapti follows a generally parallel course, between eighty kilometres and 160 kilometres to the south of the Narmada, flowing through the states of Maharashtra and Gujarat on its way into the Gulf of Khambhat.

Harnessing the waters of the major rivers that flow from the Himalayas is an issue of great concern in Nepal, India, and Bangladesh. Issues of flood control, drought prevention, hydroelectric power generation, job creation, and environmental quality—but also traditional lifestyles and cultural continuities—are at stake as these countries grapple with the political realities, both domestic and international, of altering the flow of the Ganga and Brahmaputra. Although India, Nepal, and Bangladesh seek to alleviate problems through cooperation over Himalayan rivers, irrigation projects altering the flow of Punjab-area rivers are likely to continue to be an irritant between India and Pakistan—countries between which cooperation is less likely to occur—in the second half of the 1990s. Internally, large dam projects, such as one on the Narmada River, are also controversial.

Climate

The Himalayas isolate South Asia from the rest of Asia. South of these mountains, the climate, like the terrain, is highly diverse, but some geographers give it an overall, one-word characterization—violent. What geographers have in mind is the abruptness of change and the intensity of effect when change occurs—the onset of the monsoon rains, sudden flooding, rapid erosion, extremes of temperature, tropical storms, and unpredictable fluctuations in rainfall. Broadly speaking, agriculture in India is constantly challenged by weather uncertainty.

It is possible to identify seasons, although these do not occur uniformly throughout South Asia. The Indian Meteorological Service divides the year into four seasons: the relatively dry, cool winter from December through February; the dry, hot summer from March through May; the southwest monsoon from June through September when the predominating southwest maritime winds bring rains to most of the country; and the northeast, or retreating, monsoon of October and November.

The southwest monsoon blows in from sea to land. The southwest monsoon usually breaks on the west coast early in June and reaches most of South Asia by the first week in July. Because of the critical importance of monsoon rainfall to agricultural production, predictions of the monsoon's arrival date are eagerly watched by government planners and agronomists who need to determine the optimal dates for plantings.

Theories about why monsoons occur vary. Conventionally, scientists have attributed monsoons to thermal changes in the Asian landmass. Contemporary theory cites other factors—the barrier of the Himalayas and the sun's northward tilt (which shifts the jet stream north). The hot air that rises over South Asia during April and May creates low-pressure areas into which the cooler, moisture-bearing winds from the Indian Ocean flow. These circumstances set off a rush of moisture-rich air from the southern seas over South Asia.

The southwest monsoon occurs in two branches. After breaking on the southern part of the Peninsula in early June, the branch known as the Arabian Sea monsoon reaches Bombay around June 10, and it has settled over most of South Asia by late June, bringing cooler but more humid weather. The other branch, known as the Bay of Bengal monsoon, moves northward in the Bay of Bengal and spreads over most of Assam by the first week of June. On encountering the barrier of the Great Himalayan Range, it is deflected westward along the Indo-Gangetic Plain toward New Delhi. Thereafter the two branches merge as a single current bringing rains to the remaining parts of North India in July.

The withdrawal of the monsoon is a far more gradual process than its onset. It usually withdraws from northwest India by the beginning of October and from the remaining parts of the country by the end of November.

During this period, the northeast winds contribute to the formation of the northeast monsoon over the southern half of the Peninsula in October. It is also known as the retreating monsoon because it follows in the wake of the southwest monsoon. The states of Tamil Nadu, Karnataka, and Kerala receive most of their rainfall from the northeast monsoon during November and December. However, 80 percent of the country receives most of its rainfall from the southwest monsoon from June to September. South Asia is subject to a wide range of climates—from the subfreezing Himalayan winters to the tropical climate of the Coromandel Coast and from the damp, rainy climate in the states of Assam and West Bengal to the arid Great Indian Desert. Based on precipitation and temperature, experts define seven climatic regions: the Himalayas, Assam and West Bengal, the Indo-Gangetic Plain, the Western Ghats and coast, the Deccan (the interior of the Peninsula south of the Narmada River), and the Eastern Ghats and coast.

In the Himalayan region, climate varies with altitude. At about 2,000 meters, the average summer temperature is near 18°C; at 4,500 meters, it is rarely above 0°C. In the valleys, summer temperatures reach between 32°C and 38°C. The eastern Himalayas receive as much as 1,000 to 2,000 millimetres more precipitation than do the Western Himalayas, and floods are common.

Assam and West Bengal are extremely wet and humid. The southeastern part of the state of Meghalaya has the world's highest average annual rainfall, some 10,900 millimetres. The Indo-Gangetic Plain has a varied climatic pattern. Rainfall and temperature ranges vary significantly between the eastern and western extremes. In the Peninsula region, the Western Ghats and the adjoining coast receive heavy rains during the southwest monsoon. Rainfall in the peninsular interior averages about 650 millimetres a year, although there is considerable variation in different localities and from year to year. The Eastern Ghats receive less rainfall than the western coast. Rainfall there ranges between 900 and 1,300 millimetres annually.

The northern Deccan region, bounded by the Western Ghats, the Vindhya Range and the Narmada River to the north, and the Eastern Ghats, receives most of its annual rainfall during the summer monsoon season. The southern Deccan area is in a "rain shadow" and receives only fifty to 1,000 millimetres of rainfall a year. Temperature ranges are wide—from some 15°C to 38°C—making this one of India's most comfortable climatic areas. Throughout most of non-Himalayan India, the heat can be oppressive and sometimes, such as was experienced in 1994 and 1995, literally can be a killer. Hot, relatively dry weather is the norm before the southwest monsoons, which, along with heavy rains and high humidity, bring cloud cover that lowers temperatures slightly. Temperatures reach the upper 30s°C and can reach as high as 48°C during the day in the premonsoon months.

Earthquakes

India has experienced some of the world's most devastating earthquakes. Some 19,000 people died in Kangra District, northeastern Himachal Pradesh, in April 1905, and more than 30,000 died in Maharashtra and Andhra Pradesh in September 1993. Although resulting in less extensive loss of life, major earthquakes occurred in Assam in 1950 (more than 1,500 killed) and in Uttarkashi District, Uttar Pradesh, in 1991 (1,600 killed).

STRUCTURE AND DYNAMICS

The 1991 final census count gave India a total population of 846,302,688. However, estimates of India's population vary widely. According to the Population Division of the United Nations Department of International Economic and Social Affairs, the population had already reached 866 million in 1991. The Population Division of the United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) projected 896.5 million by mid-1993 with a 1.9 percent annual growth rate. The United States Bureau of the Census, assuming an annual population growth rate of 1.8 percent, put India's population in July 1995 at 936,545,814. These higher projections merit attention in light of the fact that the Planning Commission had estimated a figure of 844 million for 1991 while preparing the Eighth Five-Year Plan. India accounts for some 2.4 percent of the world's landmass but is home to about 16 percent of the global population. The magnitude of the annual increase in population can be seen in the fact that India adds almost the total population of Australia or Sri Lanka every year. A 1992 study of India's population notes that India has more people than all of Africa and also more than North America and South America together. Between 1947 and 1991, India's population more than doubled.

Throughout the twentieth century, India has been in the midst of a demographic transition. At the beginning of the century, endemic disease, periodic epidemics, and famines kept the death rate high enough to balance out the high birth rate. Between 1911 and 1920, the birth and death rates were virtually equal—about forty-eight births and forty-eight deaths per 1,000 population. The increasing impact of curative and preventive medicine (especially mass inoculations) brought a steady decline in the death rate. By the mid-1990s, the estimated birth rate had fallen to twenty-eight per 1,000, and the estimated death rate had fallen to ten per 1,000. Clearly, the future configuration of India's population (indeed the future of India itself) depends on what happens to the birth rate. Even the most optimistic projections do not suggest that the birth rate could drop below twenty per 1,000 before the year 2000. India's population is likely to exceed the 1 billion mark before the 2001 census.

The upward population spiral began in the 1920s and is reflected in intercensal growth increments. South Asia's population increased roughly 5 percent between 1901 and 1911 and actually declined slightly in the next decade. Population increased some 10 percent in the period from 1921 to 1931 and 13 to 14 percent in the 1930s and 1940s. Between 1951 and 1961, the population rose 21.5 percent. Between 1961 and 1971, the country's population increased by 24.8 percent.

Thereafter a slight slowing of the increase was experienced: from 1971 to 1981, the population increased by 24.7 percent, and from 1981 to 1991, by 23.9 percent.

Population density has risen concomitantly with the massive increases in population. In 1901 India counted some seventy-seven persons per square kilometer; in 1981 there were 216 persons per square kilometer; by 1991 there were 267 persons per square kilometer—up almost 25 percent from the 1981 population density. India's average population density is higher than that of any other nation of comparable size. The highest densities are not only in heavily urbanized regions but also in areas that are mostly agricultural.

Population growth in the years between 1950 and 1970 centered on areas of new irrigation projects, areas subject to refugee resettlement, and regions of urban expansion. Areas where population did not increase at a rate approaching the national average were those facing the most severe economic hardships, overpopulated rural areas, and regions with low levels of urbanization.

The 1991 census, which was carried out under the direction of the Registrar General and Census Commissioner of India (part of the Ministry of Home Affairs), in keeping with the previous two censuses, used the term *urban agglomerations*. An urban agglomeration forms a continuous urban spread and consists of a city or town and its urban outgrowth outside the statutory limits. Or, an urban agglomerate may be two or more adjoining cities or towns and their outgrowths. A university campus or military base located on the outskirts of a city or town, which often increases the actual urban area of that city or town, is an example of an urban agglomeration. In India urban agglomerations with a population of 1 million or more—there were twenty-four in 1991—are referred to as metropolitan areas. Places with a population of 100,000 or more are termed "cities" as compared with "towns," which have a population of less than 100,000. Including the metropolitan areas, there were 299 urban agglomerations with more than 100,000 population in 1991. These large urban agglomerations are designated as Class I urban units. There were five other classes of urban agglomerations, towns, and villages based on the size of their populations: Class II (50,000 to 99,999), Class III (20,000 to

49,999), Class IV (10,000 to 19,999), Class V (5,000 to 9,999), and Class VI.

The results of the 1991 census revealed that around 221 million, or 26.1 percent, of Indian's population lived in urban areas. Of this total, about 138 million people, or 16 percent, lived in the 299 urban agglomerations. In 1991 the twenty-four metropolitan cities accounted for 51 percent of India's total population living in Class I urban centers, with Bombay and Calcutta the largest at 12.6 million and 10.9 million, respectively.

In the early 1990s, growth was the most dramatic in the cities of central and southern India. About twenty cities in those two regions experienced a growth rate of more than 100 percent between 1981 and 1991. Areas subject to an influx of refugees also experienced noticeable demographic changes. Refugees from Bangladesh, Burma, and Sri Lanka contributed substantially to population growth in the regions in which they settled. Less dramatic population increases occurred in areas where Tibetan refugee settlements were founded after the Chinese annexation of Tibet in the 1950s.

The majority of districts had urban populations ranging on average from 15 to 40 percent in 1991. According to the 1991 census, urban clusters predominated in the upper part of the Indo-Gangetic Plain; in the Punjab and Haryana plains, and in part of western Uttar Pradesh. The lower part of the Indo-Gangetic Plain in southeastern Bihar, southern West Bengal, and northern Orissa also experienced increased urbanization. Similar increases occurred in the western coastal state of Gujarat and the union territory of Daman and Diu. In the Central Highlands in Madhya Pradesh and Maharashtra, urbanization was most noticeable in the river basins and adjacent plateau regions of the Mahanadi, Narmada, and Tapti rivers. The coastal plains and river deltas of the east and west coasts also showed increased levels of urbanization.

The hilly, inaccessible regions of the Peninsular Plateau, the northeast, and the Himalayas remain sparsely settled. As a general rule, the lower the population density and the more remote the region, the more likely it is to count a substantial portion of tribal people among its population. Urbanization in some sparsely settled regions is more developed than would seem warranted at first glance at their limited natural resources. Areas of western India that were formerly princely states (in Gujarat and the desert regions of Rajasthan) have substantial urban centers that originated as political-administrative centers and since independence have continued to exercise hegemony over their hinterlands.

The vast majority of Indians, nearly 625 million, or 73.9 percent, in 1991 lived in what are called villages of less than 5,000 people or in scattered hamlets and other rural settlements. The states with proportionately the greatest rural populations in 1991 were the states of Assam (88.9 percent), Sikkim (90.9 percent) and Himachal Pradesh (91.3 percent), and the tiny union territory of Dadra and Nagar Haveli (91.5 percent). Those with the smallest rural populations proportionately were the states of Gujarat (65.5 percent), Maharashtra (61.3 percent), Goa (58.9 percent), and Mizoram (53.9 percent). Most of the other states and the union territory of the Andaman and Nicobar Islands were near the national average.

Two other categories of population that are closely scrutinized by the national census are the Scheduled Castes and Scheduled Tribes. The greatest concentrations of Scheduled Caste members in 1991 lived in the states of Andhra Pradesh (10.5 million, or nearly 16 percent of the state's population), Tamil Nadu (10.7 million, or 19 percent), Bihar (12.5 million, or 14 percent), West Bengal (16 million, or 24 percent), and Uttar Pradesh (29.3 million, or 21 percent). Together, these and other Scheduled Caste members comprised about 139 million people, or more than 16 percent of the total population of India. Scheduled Tribe members represented only 8 percent of the total population (about 68 million). They were found in 1991 in the greatest numbers in Orissa (7 million, or 23 percent of the state's population), Maharashtra (7.3 million, or 9 percent), and Madhya Pradesh (15.3 million, or 23 percent). In proportion, however, the populations of states in the northeast had the greatest concentrations of Scheduled Tribe members. For example, 31 percent of the population of Tripura, 34 percent of Manipur, 64 percent of Arunachal Pradesh, 86 percent of Meghalaya, 88 percent of Nagaland, and 95 percent of Mizoram were Scheduled Tribe members. Other heavy concentrations were found in Dadra and Nagar Haveli, 79 percent of which was composed of Scheduled Tribe members, and Lakshadweep, with 94 percent of its population being Scheduled Tribe members.

POPULATION PROJECTIONS

The Registrar General and Census Commissioner of India (both positions are held by the same person) oversees an ongoing intercensal effort to help maintain accurate annual estimates of population. The projection method used in the mid-1980s to predict the 1991 population, which was accurate enough to come within 3 million (843 million) of the official, final census count in 1991 (846 million), was based on the Sample Registration System. The system employed birth and death rates from each of the twenty-five states, six union territories, and one national capital territory plus statistical data on effective contraceptive use. Assuming a 1.7 percent error rate, India's projection for 1991 was close to those made by the World Bank and the UN.

Projections of future population growth prepared by the Registrar General, assuming the highest level of fertility, show decreasing growth rates: 1.8 percent by 2001, 1.3 percent by 2011, and 0.9 percent by 2021. These rates of growth, however, will put India's population above 1.0 billion in 2001, at 1.2 billion in 2011, and at 1.3 billion in 2021. ESCAP projections published in 1993 were close to those made by India: nearly 1.2 billion by 2010, still considerably less than the 2010 population projection for China of 1.4 billion. In 1992 the Washington-based Population Reference Bureau had a similar projection to ESCAP's for India's population in 2010 and projected nearly 1.4 billion by 2025 (nearly the same as projected for 2025 by the United Nations Department of International Economic and Social Affairs). According to other UN projections, India's population may stabilize at around 1.7 billion by 2060. Such projections also show an increasingly aging population, with 76 million (8 percent of the population) age sixty and above in 2001, 102 million (9 percent) in 2011, and 137 million (11 percent) in 2021. These figures coincide closely with those estimated by the United States Bureau of the Census, which also projected that whereas the median age was twenty-two in 1992, it was expected to increase to twenty-nine by 2020, placing the median age in India well above all of its South Asian neighbors except Sri Lanka.

THE LOCAL MAPPING CONTEXT

Traditionally maps have acted as a form of literal and abstract representation. The standard map is a precise top-down cartographic representation of a geographic terrain, a visualization of place. Fundamentally, maps are used to provide a view of data that is manageable for particular groups and uses. They are designed to be easily understood and represent selected information that is scaled down for ease of use. Similarly, a looser definition of cartography may allow us to consider that any form of data abstraction or representation based on a location is a map. This framing has seen maps being used as a medium to communicate ideas beyond the scope of physical geography, such as the amount of pollution in a given neighbourhood. More recently, the environmental movement has adopted mapping as a form of communication. Green Map System (1999), an organization based in America, in particular, formed around the notion of almost exclusively using maps to further the cause of environmentalism.

Green Map System encourages communities to gather data about their green facilities and spaces. Communities are encouraged to map toxic hotspots, good places to view stars, or green businesses for instance. The end result of this community process-orientated mapping is that Green Maps typically present a variety of ecology-related points of interest on a map.

Similarly, Parish Maps is led by a British environmental arts organization called Common Ground. The project calls for a communal mapping of villages, towns and cities. Here people participate in mapping what concerns them about their place in order to protect their local distinctiveness (Common Ground 1983). The boundaries of the map are determined from the outset by the Parish boundaries limiting the project to producing static maps of UK villages (Crouch and Matless 1996: 237–9). Whereas Parish Maps set up boundaries by use of their terminology and by often veering towards a rosy-type Ordnance Survey representation of place (Crouch and Matless 1996), Green Maps set boundaries in terms of what type of data constitutes 'green' data. Both favour quantitative data to the exclusion of qualitative data and have a tendency to freeze information by restricting people from adding to the picture.

A plethora of other local mapping initiatives have come about, such as indigenous communities fusing their traditional mapmaking techniques with other mapping processes to fight for their rights – and with success. These call for democratic mapping processes and have attempted to reframe who and what a map is for. Indigenous maps are often made, used, re-made and used again in a communal setting. Both indigenous and standard maps can become powerful dynamic educational and decision-making tools.

Participatory geographical information system projects have burgeoned recently and focus on ensuring the voice, and so the map, of 'the other' is heard/seen. Such projects gather and input data with the community using GIS software. Data sets can then be modelled geographically to raise community issues to influence policy-making. Qualitative data is also being gathered in this way to ensure that issues are not excluded from the map and so that sophisticated data sets can represent realistic views of how people are operating in a place. Modern technology is set to revolutionize the production and distribution of maps further than this. The global Web has made the means of production available to almost everyone with access to it. Wikipedia is a prime example of an evolving knowledge resource based on online community data editing, while Google Maps allow the creation of custom-made maps, substantially lowering technical difficulties of map creation.

The participatory mapping methodologies but some reject proprietary mapping software. For instance, Google Maps is founded on the open source software approach, allowing free access, adaptation and re-distribution of software without any or few copyright restrictions. This free and open access to all must be seen as a much more sustainable, holistic approach to mapmaking. Within the field of Google Maps this open, integrated, community-led approach to developing a project has much to offer, in particular to groups who wish to mobilize a community wanting to feel invested in a movement, as is the case with Green Map System.

The Project

Theirwork is an online open map. Open source software that drives the map is available for anyone to use or re-appropriate, rejecting a proprietary approach. Loe Pool in Cornwall, Britain (the county's largest natural lake) is the first area to be mapped by the software. While mapmaking is at the centre of the project and is used to ground the collected data, it is also used to root the project in real-time space. *theirwork* works closely with end-users, who are treated as co-developers by walking, talking and recording in its landscape. The mapmaking it seeks to produce is grounded in multiple perspectives; therefore multiple voices and autonomous experiences are documented via first person sensory experience and through a community's felt experience of landscape. The project is open, inclusive and non-hierarchical in both form and content. The software (form) and data collection (content) are symbiotic and mutually supportive in terms of 'openness'. *theirwork* software rejects a top-down system of classification or taxonomy and adopts instead a system of crowd-sourced labelling, or what has been dubbed folksonomy. Regarding authoritative and hierarchical taxonomic systems as disempowering, the folksonomic approach enables the *theirwork* participant, who works online, to collaboratively generate open-ended labels for mapped data.

Using open methodological frameworks, *theirwork* ensures that the development, production and dissemination of local definitions of place are gathered and visualized through soft (qualitative) and hard (quantitative) data collection, without any restriction on re-use. Importantly, such innovations guarantee that local definitions of a place are presented using sustaining, rejuvenating software. Foregoing other top-down systems that often produce hegemonic systems and organizations (such as copyrighted Ordnance Survey maps and copyrighted Geographical Information Systems (GIS) data), *theirwork* innovates and builds upon the movement called Green Mapmaking.

CRITIQUING GREEN MAPMAKING

Their work adopts three cognate disciplines: psychophysical geography, phenomenology and ethnography. These complementary approaches have created a methodological framework, through which open data are sourced and collected. Ethnographic methodology ensures multiple voices construct the map: the phenomenological approach ensures autonomous experiences are documented via first person sensory experience, and through a community's felt experience of landscape. Last, a psychophysical geographic approach ensures the map is emotive and deeply personal. All three approaches ensure the map is grounded in locality, subjectivity and a lived experience of place. A common discourse exists among cognate genres committed to plurality, locality and subjective interrelations of body-landscape. Immediately, the terms Green Mapmaking, indigenous mapmaking and bioregional mapping come to mind, each advancing and augmenting current mapmaking praxis. *Their work* is situated within all the above, but seeks to advance the area of soft data collection and challenges existing software that is used by many mapmakers – as although existing maps often work on the principle of open content and sharing, many use closed systems of software and licensing production to make their maps.

In line with bioregional mapmaking and the writings of Ben Whelan, *theirwork* calls 'the community into the process of mapmaking' where 'the charted landscape is filled with the stories of its dwellers and an intimate knowledge of their ecosystem'. Whelan's call, radical and compassionate, seeks to deepen 'the communion between human and nature' and create maps 'that can accommodate multiple levels of reality'.

Bioregional mapmaking's allegiance to the non-human world grew directly out of various genres of indigenous mapmaking– all wayfinders deeply connected to the landscape. Mapmaking that calls for a human appreciation and protection of the landscape and its indigenous species, has in turn created a genre of urban bioregional mapmaking called Green Mapmaking. *theirwork* is situated within Whelan's inclusive discourse of Green Mapmaking praxis; for example, *theirwork* 'seeks to energize local knowledge and mobilize citizens into action' in order to address 'greenness'.

Green Mapmaking at first glance appears an inclusive term, because it allows communities to shape their own picture of the present and future, by supplying toolkits that encourage them to chart their natural and cultural environment. These toolkits centre on a set of global Green Map Icons that the community must use in order to label their project a Green Map. However, this model still operates through a structure of exclusivity. Although these toolkits are a marked improvement upon Ordnance Survey maps and other traditional mapmaking systems, they are still partly exclusive in terms of creation, access and usage – in short, they restrict innovation. First, the structural and visual boundaries of Green Maps are often defined by criteria, which in turn are usually defined by a steering committee. Second, software is difficult to use when a community want to reproduce icons digitally. In an online environment, due to copyright restrictions, icons are difficult to use. Third, data is hard and lacks qualitative insight. Definitions of a locality tend to be shaped by hard data collection only, because data is often fitted into this icon set. Last, icons, although a powerful visualization tool, are aligned in our cultural memory with traditional topographical maps and their boundaries.

Green Mapmaking is situated within the wider problematic discourse of sustainability. For example, *theirwork* was born out of a concern for the environment, fuelled and shaped by an escalating political rhetoric that centres on the concept of sustainability. The European government is attempting to translate and implement sustainability through a practice-based legislative process, whereby industry is forced to comply with greening initiatives and general 'lay audiences' are targeted by local government bodies to construct social well-being and encourage new communities via new initiatives.

The concept of a Green Map was created in New York over a decade ago. At first, Green Mapmakers did not directly use sustainability and its associated terminology. However, bioregional mapmakers and Green Map System started to use the word to situate their work within a wider socio-context. 'The impetus for creating and teaching these new skills of sustainability [and mapping] are coming from residents in scores of places who refuse to see their social and ecological capital either under-utilized or squandered'.

Here in the UK, different fields of knowledge work to gain funds that will help them address the social, environmental or economic aspects of sustainability, but few agree about what the concept means in its entirety and even fewer are able to implement it in practice:

> Problems arise in part because the sustainability of the human enterprise in the broadest sense depends on technological, economic, political, and cultural factors as well as on environmental ones and in part because practitioners in the different relevant fields see different parts of the picture, typically think in terms of different time scales, and often use the same words to mean different things. (Daily *et al*. 1995: no pagination).

Theirwork recognizes the confusion and disparity that surrounds the word sustainability. Most importantly, *theirwork* believes the term sustainability exists and operates within a number of governmental hegemonic discourses, i.e. the term itself is continually produced within legislative power structures. For example, Agenda21 officers were situated in each UK district council by the late 1990s. Their job was to help find sustainable solutions to problems within their local community. In contrast, *theirwork* does not centre mapmaking praxis on generic or legislative definitions of sustainability, but rather encourages dialogue that supports the re-formation of self, community and place. *theirwork* does not seek to overturn generic understandings of sustainability, but rather seeks a more complex understanding and proliferation of the term via local 'grounded' definitions. *theirwork* therefore builds on Green Mapmaking and sustainable discourses, but has created innovative strategies within the genre of bioregional mapping, particularly in the following areas: mapping software, online access and the gathering of soft data.

POSSIBLE SOLUTIONS, COMING FROM OPEN SOURCE

Having identified fundamental problems and restrictions inherent in existing models of Green Mapmaking, the question of how to define an alternative framework presents itself. The flexibility provided by Internet mapping has already been explained, and in the case of *theirwork* was seen as the most likely medium to allow the type of open-ended activities that traditional Green Maps cannot.

There is obviously an established body of work in the field of Web-based mapmaking that requires critical appraisal. First, however, it is necessary to survey the wider terrain of computing and take stock of what influences can be garnered from its politics and philosophies. An immediate parallel can be drawn between the wider green movement, from which Green Mapmaking emerged, and the open source software movement. In an effort to establish a more holistic and sustainable approach to mapmaking in general, it was deemed necessary to focus on each of the constituent parts that the framework takes, and ensure that the approach is consistent and self-propagating. Hence, a focus on the ideologies of software development was central to the maturation of *theirwork* as a coherent movement.

The open source movement at its core stands for the development of source code (the algorithms and computer logic written by computer programmers to create software) in a completely open and free way. Pragmatically, this manifests itself as a methodology of making code freely available to anyone who may wish to access it for any purpose, unconditionally. Concurrently, open source is for many a philosophical approach to software development, and is seen as the only truly sustainable approach to software develop□ment. Open source code may be shared, studied, copied, reused, modified, built upon and redistributed in any way. As such this model has made possible innumerable software projects that would otherwise have been almost impossible to realize (the most popular examples include the Linux operating system and the Firefox Web browser, both used by millions of computer users).

The possibilities of the model are highlighted by open source evangelist Eric S. Raymond in his seminal 1997 essay *The Cathedral and the Bazaar*, in which he compares the development approach to 'a great babbling bazaar of differing agendas and approaches', all of which create a finished product that could never have been designed or executed by a single architect.

In today's world of corporate global software giants, whose billions are based upon the materiality and inaccessibility of code, this can seem to be a revolutionary set of ideas. Yet in both its execution as a model for making possible new forms of collaborative work, and its philosophical underpinnings of sustainability and openness, it is an essential component in and influence upon a computer-based mapping solution.

In the earliest planning stages of the project, it was resolved that in order to improve upon the existing framework of Green Mapmaking, the entire back-to-front process of it should be executed in line with the ideals of the project as a whole. The ongoing development of the tool, the process of creation, was taken to be as important as the final artefact, and consistent with the ideology that drove its inception. Thus, although the use of open source software was in line with the spirit of *theirwork*, this alone was not sufficient to constitute a holistic approach. At every stage of development, decisions were consciously influenced by the desire to create a project that would at every turn reinforce itself. As this concept developed, the approach came to be labelled 'sustainable software', and drew together influencing characteristics from a number of disparate fields, combining select strands from each into what was hoped to be a coherent whole.

The architectural method of adaptive design, that of designing and building to ensure that a system retains enough inherent flexibility to be modified (or even encourage modification) that had not originally been considered, was an influence on the planning of the project outcome. The Slow Food move□ment, which encourages a change of pace and even lifestyle in order to reassess priorities and values, was another. Apart from open source licensing, the object-oriented approach to writing computer code, which ensures that each part of the code is modular and easily replaceable, was an influence from the arena of technology. The copyleft and Creative Commons (2008) movements that eschew the traditional concepts of information ownership in favour of a more liberal approach to content sharing (of which more later) were studied carefully. Some of these ideas were adapted quite literally, but were also taken as philosophical or political approaches, helping to shape the concept of sustainable software.

With a driving ideology defined (or as clearly defined as any set of ideas which have at their core the intention to be as flexible as possible), the question of how to actually implement the project naturally arose. It was decided at an early stage to make the software Web-based to allow for a process of rapid development and iteration and allow a maximum number of potential participants. Another, more pragmatic, reason was to facilitate the fact that the two main contributors to the project live in different countries; almost all communication was carried out via a combination of email, phone and instant messaging.

For reasons that should be obvious from the influence of open source, it was decided not to pay for the right to use commercial mapping software. One of the next obvious approaches when creating Web-based maps is to use an already-available service, such as Google Maps. It is relatively simple to create what is known as a Google Maps mashup; that is, taking an existing map, and overlaying one's own data on to that map. As an immediate technical solution, a Google Maps mashup would appear to be the easiest option. However, close inspection of Google's terms and conditions revealed that the licensing it bore did not meet the strict guidelines that had already been established in relation to software licensing for the project. Nor did any existing open source mapping toolkit meet the needs of the project. It was eventually decided to build a custom software solution, and make it available to the public as open source software. It should be noted, however, that a number of existing open source toolkits were used to create smaller parts of the tool, combining to create a new whole. Without the ability to reuse and adapt the code that already existed within the ecosystem of the open source community, it would have been practically unfeasible to develop such a complex system.

Creating a base image for the map (i.e. a top-down view of the lake where the project was piloted, on which to plot the data) was not an easy process. Again, licensing restrictions proved a point of contention; the now-controversial laws surrounding Ordnance Survey data meant that purchasing the map data for the lake was ideologically and financially out of the question. Although there are nascent communities such as OpenStreetMap (2004) currently endeavouring to make geodata freely available in the UK, no efforts existed in the geographic location that *theirwork* focused on. There was no pre-existing, freely available data on which to build. In order to obtain the data, GPS units were used to record the track points of a walk around the perimeter of the lake, and specific points of interest were marked along the way. The result was a matrix of latitude and longitude GPS coordinates, which were then loaded onto a computer, where pre-existing open source software was used to generate a simple line drawing of the lake's outline. This outline was then annotated by hand to create a defined background map on which data points could be plotted. This was a laborious and technical process, and represents one of the major remaining obstacles to the breakthrough and popularization of people-powered mapping; it will inevitably be overcome by the proliferation of user-friendly convergent hardware that integrates GPS with popular consumer recording devices, such as cameras.

The fully developed beta version of the software consists of a Web-based Google Maps-like interface, by which the user can interact with a map of the lake. A number of data points that have already been added by other users are overlaid. What makes *theirwork* slightly different from other mapping software is the ability for users to immediately add their own points of interest to the map directly at any time.

They may also edit existing points to improve them as they see fit. This open model of community data editing is taken directly from the wiki model (the best known example of which is the online encyclopaedia Wikipedia), in which participants may add or edit any page on the website. This distributed model of content creation can work remarkably well in some cases, and is surprisingly capable of 'self-healing' in cases of vandalism, whereby a subsequent user notices and immediately rectifies an existing error.

At the same time, a completely open data system such as this could make for a chaotic set of data, if not presented in a logical manner. The question arises: what is a sustainable model of group data classification? Green Maps have encountered the problem that their maps can be too narrow in subject if a strong editorial control is exerted, and too chaotic and unstructured if free rein is permitted. How can people be empowered to add whatever type of data they wish to the map, but also have a coherent picture emerging from the map as a whole?

Fortunately, computers are adept at taking a lot of information and shuffling it, or slicing and dicing it, in any way. Many websites with user-generated content have experienced a similar problem recently, attempting to classify an open data set without imposing structure. The aim is to somehow capture (to paraphrase a book title on this topic) the Wisdom of Crowds (Surowiecki 2004), and allow an emergent picture to develop from the teeming mass of individual actions happening within a system.

The solution here is to reject a top-down system of classification, or taxonomy, and adopt instead a system of labelling, or what has been dubbed folksonomy. This involves rejecting any notions of hierarchical classification, and allowing users to tag their data with keywords that describe it instead. A data point has many keywords pinned on to it, instead of being placed into a single category.

This actually opens up the process considerably, and leads to a much more creative way of adding data. Users now have the freedom to use the map in ways that the map designers may never have even conceived. The map becomes an adaptive, open-ended, and sustainable ecosystem of data.

At the data output stage, when trying to discover or extract all of the data that a user is interested in, they do not dig down into a category to find the relevant items, but rather filter out all items by keyword.

This may be thought of as viewing a cross-section or slice of all data, except that even within this single slice, there exists a lot more information still to be mined; many more strata of keywords that may line up, or move off in a different direction. The whole experience makes for a much richer data process. This approach works well for open data in mapping, as it means that we can dismiss concerns about misclassified information, or editorial control, and concentrate on extracting a meaningful signal from the rich information set.

This opens up a route for an entirely new type of emergent, community-developed map creation that coherently represents the combined impressions of an unrelated group of self-interested actors, and conveys a truly distributed simulation of a geographic space.

DEVELOPING OPEN DATA, OUT OF PLACE-BASED MAPPING

The voice of the participant, rather than the voice of the researcher, will be heard best when participants not only provide the data to be analysed, but when they also contribute to the questions that frame the research and contribute to the way data are analysed.

In terms of mapmaking praxis, construction of the map has been an entirely de-centred process and authoritative models of data collection and transcription have been overturned. The application of ethnographic methodology ensures multiple voices construct the map. Within this work the relationship of emotion, memory, and sensory engagement with the landscape was mapped. First, data was sourced while walking, talking and recording with participants on the landscape. After an initial recruitment period and focus session, each co-developer chose a location for a 'one-to-one' walk that in some way was connected to the lake. Co-developers chose the date and time – some brought their binoculars or dog along, others even brought 'somebody else along'. The co-developers were helped in tracking the walk; sites of interest, objects, plants and animals, favourite places, memory spots and stories connected to the place. Places were noted using cameras, notebooks, a GPS unit and a dictaphone. A framework of open and closed questions was asked. Answers to open questions, such as 'What do you feel about the lake?' were geo-tagged. The 'type of walk' (their special walk) became an integral part of data collection and data analysis. These processes helped capture the walk and created a supplementary resource to each recorded conversation that took place, which was then transcribed.

In the spirit of ethnographic methodology, transcription and coding of data was a mutually inclusive activity (all information was verified with co-developers). Some of the codes that developed from the walks were words such as rocks, water, agriculture, birds, meditation, trees, fields, memories, fish and events. In a paper-based workshop, co-developers then jointly discussed the codes and each shared their record of the walk from memory. Memories were added to the discussed codes. Importantly here, qualitative data became coded by the co-developers and not by some distant and 'removed' ethnographer.

To this end, paper-based tags were ready-made for the map interface. A sort of starter kit had been created, effectively introducing co-developers to tagging or folksonomy. Qualitative coding methodologies in turn introduced the community to the art of good folksonomy. This is an important issue, because it deepened the practice of folksonomy and helped to reflect on it in a practical and academic manner.

A computer workshop then tested the beta version of the sustainable software. Each co-developer put marks on the map, using latitude and longitude figures supplied from the archive of walks and paper-based workshop.

They tagged their marks efficiently and with ease, having been introduced to the concept of folksonomy in the paper workshop. When things started taking shape onscreen the mood in the workshop room was electrifying. Everyone watched their places appear on the map – and all the efforts and concepts that must at times have seemed utterly puzzling started to make sense and finally paid off.

Technical problems were fixed as and when they arose. Co-developers' views, feelings and ideas for the future of the software, as well as ideas for new data, were taken into account. Outdoor events have since become interchangeable with ongoing paper and computer workshops. All types of place-based mapping happenings have been called for by the co-developers, and are enabling the gathering of data that was not pre-determined. For instance, moth migration nights, stargazing gatherings, butterfly balls, drawing picnics and plastic bag counts have taken place. The geo-coded data is challenging how the base map could look and function, and is drawing in experts in the field of flora and fauna and qualitative research.

At present, when data is added, co-developers either leave trails of red dots where they have been recording a walk or they add to pools of information where groups have gathered. For instance, a moth and bat night focused on three spots at the mouth of the lake, and became like three micro-maps of fascinating creatures and facts. These red marks could become a sea of pictures, telling a tale of moths in this area. At the moment pictures are uploaded to a separate space, to a group Flickr (2004) account. As funding is applied to develop the project further, the co-developers will become involved in the application, asserting what they think should be developed next. Their priorities so far are: to make the base map more sophisticated; to make their map entries about each place more editable; to style the forum and to be able to guest blog (at the moment the project initiators and a few others are blogging).

DIMENSIONS ACROSS WHICH MAP THEORY IS CONSTITUTED

A useful way of starting to understand how and why map theory varies is to explore some of the dimensions across which philosophical debate is made. Some important binary distinctions that strongly influence views on the epistemological and ontological status of mapping: judging a philosophy against these distinctions provides an often unspoken set of rules for knowing the world, or in our case, for arguing about the status of mapping.

These distinctions are clearly related to each other. An emphasis upon the map as representation, for example, is also often strongly associated with the quest for general explanation, with a progressive search for order, with Cartesian distinctions between the map and the territory it claims to represent, with rationality, and indeed with the very act of setting up dualistic categories.

By exploring how these dimensions work we can begin to rethink mapping and explain the complex variety of approaches described later in this chapter.

The mind–body distinction is often a fundamental influence on how people think about the world. If the mind is conceptualized as separate from the body then instrumental reason becomes possible: the map can be separated from the messy and subjective contingencies that flow from an embodied view of mapping.

As such, science and reason become possible and a god□like view from nowhere can represent the world in an objective fashion, like a uniform topographic survey. On the other hand assuming a unity of mind and body and emphasizing the idea of embodied knowing focuses attention on different, more hybrid and subjective qualities of mapping, rendering problematic distinctions between the observer and observed.

Table : Rules for knowing the world: binary opposites around which ideas coalesce.

Mind	***Body***	***Structure***	***Agency***
Empirical	Theoretical	Process	Form
Absolute	Relative	Production	Consumption
Nomothetic	Ideographic	Representation	Practice
Ideological	Material	Functional	Symbolic
Subjective	Objective	Immutable	Fluid
Essence	Immanence	Text	Context
Static	Becoming	Map	Territory

The question of whether geographic knowledge is unique or whether the world might be subject to more general theorizing also has fundamental implications for mapping. An ideographic emphasis on uniqueness has frequently pervaded theorizing about mapping in the history of cartography: if each map were different, and described a unique place, searching for general principles that might govern design, or explain use would be doomed to fail. Instead, mapping becomes the ultimate expression of descriptive endeavour, an empirical technique for documenting difference.

Artistic approaches to mapping that privilege the subjective may be strongly compatible with this kind of interpretation. On the other hand a more nomothetic approach, which emphasizes laws and denies idiosyncratic difference risks reifying artificially theorized models or generalizations while at the same time offering the possibility of scientific universalization. Many of the approaches described in the chapters by Goodchild and Gartner in this volume subscribe to this quest for order. Debate continues around the nature of map generalization and whether mapping is holistic or fragmentary, stochastic or regular, invariant or contingent, natural or cultural, objective or subjective, functional or symbolic, and so on. It is clear, however, that since World War II a number of different scientific orthodoxies have pervaded the world of Western academic cartographic research which almost all trade on the notion of searching for a common, universal approach. Yet, paradoxically, everyday ideas of geography and mapping as ideographic and empirical survive.

The idea of viewing maps as texts, discourses or practices emerged in the late 1980s, in stark opposition to the more practical and technologically driven search for generalization. These new theoretical ways of understanding mapping often emphasized the discursive power of the medium, stressing deconstruction, and the social and cultural work that cartography achieves. Here, the power of mapping becomes a more important consideration than the empirical search for verifiable generalization, and the chapters in this volume by Crampton, Harris and Hazen, and Propen consider some of these alternative approaches.

Structural explanations of the significance of mapping have also strongly influenced understandings of maps. Insights drawn might stem from class relations, from cultural practice, from psychoanalysis, or linguistics: for example, semiotic approaches to mapping have been a powerful and influential way of approaching the medium and its messages for academic researchers. There is an ongoing debate in relation to mapping over how the agency of an individual might be reconciled with this kind of approach, given that structural approaches often posit fundamental and inevitable forces underpinning all maps. There is also a continuing debate over the philosophical basis of the structural critique. For example, is it grounded in a materialist view of the world, or in a more ideological reading of the human condition. The distinction between forces producing the world and the forces consuming it also has a strong resonance in philosophical debates around mapping. The cultural turn in academic geography encouraged a growing emphasis on the contexts in which maps operate, encouraging a shift away from theorizing about production and towards philosophies of mapping grounded in consumption. Here, the map reader becomes as important as the mapmaker. Technological change that reduced the significance of barriers to accessing data, and the democratization of cartographic practice have also encouraged this changed emphasis. Associated with this shift has been the increasingly nuanced drift towards poststructuralist ways of knowing the world, which distrust all-encompassing knowledge claims. Instead of a belief in absolute space, or a socially constructed world, an alternative way of understanding mapping has emphasized relativity and contingency in a universe where notions of reality come to be replaced by simulation and in which the play of images replaces visual work, or in which speed of change itself gains agency.

REPRESENTATIONAL CARTOGRAPHY

Maps as Truth

It is usually accepted that cartography as a scientific endeavour and industry seeks to represent as faithfully as possible the spatial arrangements of phenomena on the surface of the earth. The science of cartography aims to accurately capture relevant features and their spatial relations and to re-present a scaled abstraction of that through the medium of a map. Maps seek to be truth documents; they represent the world as it really is with a known degree of precision. Cartography as an academic and scientific pursuit then largely consists of theorizing how best to represent and communicate that truth (through new devices, e.g. choropleth maps, contour lines; through the use of colour; through ways that match how people may think, e.g. drawing on cognitive science).

This quest for producing truth documents has been the preoccupation for Western cartographers since the late Middle Ages, and especially with the need for accurate maps with respect to navigation, fighting wars and regulating property ownership. It was only in the 1950s, however, that the first sustained attempts began to emerge in the US to reposition and remould academic cartography as an entirely scientific pursuit. Up until then the history of cartography was a story of progress.

Over time maps had become more and more precise, cartographic knowledge improved, and implicitly it was assumed that everything could be known and mapped within a Cartesian framework. The artefact and individual innovation were what mattered. Space, following Kant, became conceived as a container with an explicit geometry that was filled with people and things, and cartography sought to represent that geometry. Scientific principles of collecting and mapping data emerged, but cartography was often seen as much of an art as a science, the product of the individual skill and eye of the cartographer. Mapping science was practical and applied and numerous small advances built a discipline.

In the latter part of the twentieth century, US scholar Arthur Robinson and his collaborators sought to re-cast cartography, focusing in particular on systematically detailing map design principles with the map user in mind. His aim was to create a science of cartography that would produce what he termed 'map effectiveness' – that is, maps that capture and portray relevant information in a way that the map reader can analyse and interpret (cf. Robinson and Petchenik 1976). Robinson suggested that an instrumental approach to mapping grounded in experimental psychology might be the best way for cartography to gain intellectual respectability and develop a rigorously derived and empirically tested body of generalizations appropriate for growing the new subject scientifically. Robinson adopted a view of the mind as an information-processing device. Drawing upon Claude Shannon's work in information theory, complexity of meaning was simplified into an approach focusing on input, transfer and output of information about the world. Social context was deemed to be irrelevant; the world existed independent of the observer and maps sought only to map the world. The cartographer was separate from the user and optimal maps could be produced to meet different needs.

The aims of the cartographer were normative – to reduce error in the representation and to increase map effectiveness through good design. Research thus sought to improve map designs by carefully controlled scientific experimentation that focused on issues such as how to represent location, direction and distance; how to select information; how best to symbolize these data; how to combine these symbols together; and what kind of map to publish. Framed by an empiricist ideology, the research agenda of cartography then was to reduce signal distortion in the communication of data to users. Art and beauty had no place in this functional cartographic universe.

Out of this context in the late 1960s and 1970s emerged an increasingly sophisticated series of attempts to develop and position cartographic communication models as the dominant theoretical framework to direct academic research. Communication models encouraged researchers to look beyond a functional analysis of map design, exploring filters that might hinder the encoding and decoding of spatial information. For researchers such as Grant Head (1984) or Hansgeorg Schlichtmann (1979) the map artefact became the focus of study, with an emphasis on the semiotic power of the map as opposed to its functional capacity, while Christopher Board (1981) showed how the map could be conceived as a conceptual, as well as a functional, model of the world.

As models of cartographic communication multiplied so attention also increasingly focused on the map reader, with cognitive research seeking to understand how maps worked, in the sense of how readers interpreted and employed the knowledge maps sought to convey. Drawing on behavioural geography, it was assumed that map reading depended in large part upon cognitive structures and processes and research sought to understand how people came to know the world around them and how they made choices and decisions based on that knowledge. This approach is exemplified in the work of Reginald Golledge (1999), Robert Lloyd (2005) and Cynthia Brewer. Here the map user is conceived as an apolitical recipient of knowledge and the cartographer as a technician striving to deliver spatially precise, value-free representations that were the product of carefully controlled laboratory-based experiments that gradually and incrementally improved cartographic knowledge and praxis.

Most research investigated the filters in the centre of this system concerned with the cartographers' design practice, and the initial stages of readers extracting information from the map. Little work addressed either what should be mapped or how mapping was employed socially because this was beyond the philosophical remit for valid research.

Other strands of scientific research into mapping emphasized the technologies that might be employed. Waldo Tobler's (1976) analytical cartography emerged in the early 1970s, offering a purely mathematical way of knowing the world, and laying the foundations for the emergence of geographic information science. This analytical approach sought progress through the application of mathematical models and the subsequent application of technology so as to create new conceptual bases for mapping the world. Over time, conceptual and technically driven developments in computer graphics, computation and user interfaces have begun to fundamentally transmute the role of the map from a finished product to a situation where the map is displayed within a visual toolbox to be used interactively for exploratory data analysis (typically with the interlinking of multiple representations such as statistical charts, three-dimensional plots, tables and so on). This changing conceptualization of the map is at the heart of the emerging field of geovisualization, which in the last decade or so has been one of the leading areas of applied cartographic research. Although distinctly positivist epistemologies underlie most of the geovisualization research, some have tried to open up the scope of visualization in more politically progressive directions, for example, Craine and Aitken's chapter in this volume, which considers the emotional energy latent in cinematic qualities of maps, and Kwan's (2007) work in fusing geospatial technologies with feminist theory to map affect and emotional geographies.

In other contexts different theoretical positions were adopted. For example, the French disciplinary tradition was much less influenced by Robinsonian functionalism and empirical research. Semiotic approaches were much more influential in this context, and may be traced back to the influential theories of Jacques Bertin. In 1967 Bertin derived from first principles a set of visual variables that might be manipulated by designers concerned with the effective design of mapping and other visualizations.

By the mid 1980s the cartographic communication model as an organizing framework for academic research was beginning to wane. Technological changes rendered problematic a single authoritative view of the world at a time when data were becoming much more readily available, and when technologies for the manipulation and dissemination of mapping were also being significantly changed.

Users could become mappers and many possible mappings could be made. Digital mapping technologies separated display from printing and removed the constraint of fixed specifications. GIS increasingly supplanted many technical aspects of cartographic compilation and production.

Digital position, elevation and attribute data could be captured from remotely sensed sources, and easily stored and manipulated in a digital form. Imagery could be generated to provide frequent updates of changing contexts. Maps could become animated. From the late 1990s the Internet has allowed maps to be evermore widely shared and disseminated at low cost. Mapping needed to be understood as much more of a process than was possible in communication models.

In the face of these profound challenges a second dominant approach to mapping research had replaced cartographic communication by the mid 1990s as the scientific orthodoxy. The linear inevitability of communication was supplanted by a multifaceted and multilayered merging of cognitive and semiotic approaches, centred on representational theory, and strongly influenced by the work of Alan MacEachren (1995).

Articulating ideas grounded in Peircean semiotics, this approach recognized the need for a much less literal and functional positioning of maps. The iconic diagrammatic description of this approach is the notion of 'cartography cubed'. The dimensions of interactivity, the kind of knowledge, and the social nature of the process show the three key ways in which scientific understanding has been repositioned.

Mapping can now be investigated as collaborative, the social context beyond map reading per se can be charted, and the process of knowing explored. And mapping is one of many kinds of visualization. However, mapping is still about revealing truth through a scientific approach reliant upon Western ways of seeing and upon technologies of vision; it still depends upon scientific experimentation and a representational view of the world.

Maps as Social Constructions

The view that cartography produces maps of truth in an objective, neutral, scientific fashion has been challenged by a number of scholars. In the late 1980s, the work of Brian Harley began to question how mapping operated as a powerful discourse, challenging the scientific orthodoxy of cartographic research. He proposed a new research agenda concerned with the roles maps play in different societies, arguing that maps often reinforce the status quo or the interests of the powerful, and that we should investigate the historical and social context in which mapping has been employed. In this view cartography was not necessarily what cartographers said it was. Instead, Harley argued that we could only understand the history of cartography if we interrogate the forces at play around mapping.

Harley (1989) drew on the ideas of Michel Foucault among others to argue that the process of mapping was not a neutral, objective pursuit but rather was one laden with power. He contended that the process of mapping consists of creating, rather than simply revealing, knowledge. In the process of creation many subjective decisions are made about what to include, how the map will look and what the map is seeking to communicate. As such, Harley noted, maps are imbued with the values and judgements of the individuals who construct them and they are undeniably a reflection of the culture in which those individuals live. Maps are typically the products of privileged and formalized knowledges and they also tend to produce certain kinds of knowledge about the world. And in this sense, maps are the products of power and they produce power. In contrast to the scientific view that positions maps in essentialist terms, Harley cast maps as social constructions; as expressions of power/knowledge. Others, such as Denis Wood (1992) and John Pickles (2004), have extensively demonstrated this power/knowledge revealing the ideology inherent in maps (or their 'second text') and how maps 'lie' (or at least provide selective stories while denying their selectivity) due to the choices and decisions that have to be made during their creation, and through how they are read by users.

This social constructivist critique sometimes also articulated structural explanations for mapping, which sought understanding beneath the apparent surface of observable evidence. For example, David Harvey's (1989) Marxist analysis of the role of mapping in time–space compression examined the role of global images in the expansion of European colonial powers, and situated these as reflections of a changing mode of production. Drawing on linguistic structural thought Denis Wood (1992) employed Barthean semiotics to persuasively argue that the power of maps lay in the interests they represented. Mapping in this view always has a political purpose, and this 'interest' often leads to people being pushed 'off the map'. Wood argued that mapping works through a shared cultural reading of a number of different codes in every map, which may be analysed in a semiotic process to reveal the power behind the map. These interests all too often led to subjugation, oppression, control and inequality. Through economic relations, legal evidence, governance or social practice the power of maps continues to be used to control.

It has been argued that many of the social roles played by cartographic knowledge stem from the modernist project, and that a mapping mentality is integral to the modernist enterprise itself. By examining different categories across which power might be articulated contextual studies can reveal how maps reflect but also constitute different kinds of political relation. Colonialism, property ownership, national identity, race, military power, bureaucracy and gender have all been theorized as playing key roles in mapping relations.

For example, local knowledge has been translated into tools to serve the needs of the colonizer, with new territories scripted as blank spaces, empty and available for the civilizing Western explorer to claim, name, subjugate and colonize (Edney 1997). Projection and design have been used to naturalize the political process of imperial control and sell imperial values to citizens at home. The continuing progress of colonial adventures is mapped out nowadays in our news broadcasts and on the Internet, but the imperial rhetoric of control, governance, management of territory and creation of new imperial landscapes remains the same. The colonial project relies on the map, and in turn the map relies on colonial aspirations.

The work by Harley, Wood, Harvey and others set the groundwork for work since the 1990s that has been labelled critical cartography and with respect to wider geospatial technologies, critical GIS. Critical cartography is avowedly political in its analysis of mapping praxis seeking to deconstruct the work of spatial representations in the world and the science that produces them. It is, however, decidedly not against maps, but rather seeks to appreciate the diverse ways in which maps are produced and used by different individuals and groups. From such a perspective there is no one ‘right way’ to produce maps, but their makers need to be sensitive to politics and context of their making and use. For some theorists this means moving beyond thinking of maps as representations to try to conceive of a post-representational cartography.

POST-REPRESENTATIONAL CARTOGRAPHY

From Ontic Knowledge to Ontology

Despite the obvious advances of the various social constructivist approaches in rethinking maps, more recent work has sought to further refine cartographic thought and to construct post-representational theories of mapping. Here, scholars are concerned that the critique developed by Harley and others did not go far enough in rethinking the ontological bases for cartography, which for them has too long been straitjacketed by representational thinking.

Denis Wood (1993) and Jeremy Crampton (2003) outline, Harley's application of Foucault to cartography is limited. Harley's observations, although opening a new view onto cartography, stopped short of following Foucault's line of inquiry to its logical conclusion. Instead, Crampton (2003: 7) argues that Harley's writings 'remained mired in the modernist conception of maps as documents charged with "confessing" the truth of the landscape'. In other words, Harley believed that the truth of the landscape could still be revealed if one took account of the ideology inherent in the representation. The problem was not the map per se, but 'the bad things people *did* with maps'; the map conveys an inherent truth as the map remains ideologically neutral, with ideology bound to the subject of the map and not the map itself. Harley's strategy was then to identify the politics of representation in order to circumnavigate them (to reveal the truth lurking underneath), not fully appreciating, as with Foucault's observations, that there is no escaping the entangling of power/knowledge.

Crampton's solution to the limitations of Harley's social constructivist thinking is to extend the use of Foucault and to draw on the ideas of Heidegger and other critical cartographers such as Edney (1993). In short, Crampton (2003: 7) outlines a 'non-confessional understanding of spatial representation' wherein maps instead of 'being interpreted as objects at a distance from the world, regarding that world from nowhere, that they be understood as being in the world, as open to the disclosure of things'. Such a shift, Crampton argues, necessitates a move from understanding cartography as a set of ontic knowledges to examining its ontological terms. Ontic knowledge consists of the examination of how a topic should proceed from within its own framework where the ontological assumptions about how the world can be known and measured are implicitly secure and beyond doubt (Crampton 2003). In other words, there is a core foundational knowledge – a taken for granted ontology – that unquestioningly underpins ontic knowledge.

With respect to cartography this foundational ontology is that the world can be objectively and truthfully mapped using scientific techniques that capture and display spatial information. Cartography in these terms is purely technical and develops by asking self-referential, procedural questions of itself that aim to refine and improve how maps are designed and communicate (Crampton gives the examples of what colour scheme to use, the effects of scale, how maps are used historically and politically). In these terms a book like Robinson *et al*. (1995) is a technical manual that does not question the ontological assumptions of the form of mapping advocated, rather it is a 'how to do "proper" cartography' book that in itself perpetuates the security of cartography's ontic knowledge. In this sense, Harley's questioning of maps is also ontical, as his project sought to highlight the ideology inherent in maps (and thus expose the truth hidden underneath) rather than to question the project of mapping per se; 'it provided an epistemological avenue into the map, but still left open the question of the ontology of the map'. In contrast, Crampton details that examining cartography ontologically consists of questioning the project of cartography itself.

Such a view leads to Crampton, following Edney (1993), to argue for the development of a non-progressivist history of cartography; the development of a historical ontology that rather than being teleological (wherein a monolithic view of the history of cartographic practices is adopted that sees cartography on a single path leading to more and more complete, accurate and truthful maps) is contingent and relational (wherein mapping – and truth – is seen as contingent on the social, cultural and technical relations at particular times and places). Maps from this perspective are historical products operating within 'a certain horizon of possibilities'. It thus follows that maps created in the present are products of the here-and-now, no better than maps of previous generations, but rather different to them.

Defining a map is dependent on when and where the map was created, as what constitutes a map has changed over time. For Crampton (2003: 51) this means that a politics of mapping should move beyond a ‘critique of existing maps’ to consist of ‘a more sweeping project of examining and breaking through the boundaries on how maps are, and our projects and practices with them’; it is about exploring the ‘being of maps’; how maps are conceptually framed in order to make sense of the world. Several other cartographic theorists have been following similar lines of enquiry to Crampton in seeking to transfer map theory from ontic knowledge to ontology and it is to them that we now turn.

Maps as Inscriptions

John Pickles (2004) has sought to extend cartographic theory beyond ontic status by conceiving of maps as inscriptions as opposed to representations or constructions. His work focuses on ‘the work that maps do, how they act to shape our understanding of the world, and how they code that world’. As such his aim is to chart the ‘practices, institutions and discourses’ of maps and their social roles within historical, social and political contexts using a poststructural framework that understands maps as complex, multivocal and contested, and which rejects the notion of some ‘truth’ that can be uncovered by exposing ideological intent. Pickles’ detailed argument unpicks the science of representation, calling for a post-representational cartography that understands maps not as mirrors of nature, but as producers of nature. To paraphrase Heisenberg (1959, cited in Pickles 2004), Pickles argues that cartography does not simply describe and explain the world; it is part of the interplay between the world and ourselves; it describes the world as exposed to our method of questioning.

For Pickles, maps work neither denotatively (shaped by the cartographic representation, labelling, embedded with other material such as explanatory text, etc.) or connotatively (what the mapper brings to the representation in terms of skills, knowledges, etc.) but as a fusion of the two. Pickles thus proposes a hermeneutic approach that interprets maps as unstable and complex texts, texts that are not authored or read in simple ways. Rather than a determinate reading of the power of maps that seeks to uncover in a literal sense the authorial and ideological intent of a map (who made the map and for what purpose), Pickles expresses caution in fixing responsibility in such a manner, recognizing the multiple, institutional and contextual nature of mapping. Similarly, the power of maps is diffuse, reliant on actors embedded in contexts to mobilize their *potential* effects: 'All texts are... embedded within chains of signification: meaning is dialogic, polyphonic and multivocal – open to, and demanding of us, a process of ceaseless contextualization and recontextualization'.

Alongside a hermeneutic analysis of maps, Pickles proposes that a post-representational cartography consists of the writing of denaturalized histories of cartography and the production of de-ontologized cartography. Denatural□ized histories reveal the historicizing and contextualizing conditions that have shaped cartographic practices to 'explore the ways in which particular machines, disciplines, styles of reasoning, conceptual systems, bodies of knowledge, social actors of different scales... and so forth, have been aligned at particular times and particular places'. In other words, they consist of genealogies of how cartography has been naturalized and institutionalized across space and time as particular forms of scientific practices and knowledge. A de-ontologized cartography is on the one hand about accepting counter-mappings as having equal ontological status as scientific cartographic (that there are many valid, cartographic ontologies), and on the other, deconstructing, reading differently, and reconfiguring scientific cartography.

Maps as Propositions

Like Pickles, Crampton and others, Wood and Fels (2008) extend the notion of a map as social construction to argue that the map itself, its very make□up and construction – its self-presentation and design, its symbol set and categorisation, its attendant text and supporting discourse – is ideologically loaded to convey particular messages. A map does not simply represent the world; it produces the world. They argue that maps produce the world by making propositions that are placed in the space of the map. Maps achieve their work by exclaiming such propositions and Wood and Fels define this process as one of 'posting' information on map. Posting is the means by which an attribute is recognized as valid (e.g. some class of the natural world) and is spatialized. It is the means by which the *nature* of maps (is – category) and the nature of *maps* (there – sign) conjoin to create a unified spatial ontology (this is there).

Wood and Fels argue that the power of this spatial propositional framework is affirmed through its call to authority – by being an objective reference object that is prescriptive not descriptive. So the map produces and reaffirms territory rather than just describing it.

Authority is conveyed through what they term the paramap. A paramap is the combination of perimap and epimap. The perimap consists of the production surrounding a map: the quality of the paper, the professionalism of the design, the title, legend, scale, cartouches, its presentation and so on. The epimap consists of the discourse circulating a map designed to shape its reception: advertisements, letters to reviewers, endorsements, lectures, articles, etc. Together, the perimap and epimap work to position the map in a certain way and to lend it the authority to do work in the world.

Because maps are prescriptive systems of propositions, Wood and Fels contend that map creation should not solely be about presenting information through attractive spatial representations as advocated by the majority of cartographic textbooks (which borrow heavily from graphic design traditions). Instead they suggest map design should be about the 'construction of meaning as a basis for action'.

They propose turning to cognitive linguistics to rethink map design as a form of 'cognitive cartographics'. Cognitive linguistics examines the ways in which words activate neural assemblages and open up 'thinking spaces' in the mind within which meaning is constructed by linking present information with past knowledge. They contend that maps perform like words, by firing up thinking spaces. Employing cognitive cartographics, they suggest, will create a non-representational approach to map design focused on the construction of meaning rather than graphic design and the nature of signs. It will also enable cartographic theory to move beyond the compartmentalized thinking that has divided mapmaking from map use by providing a more holistic framework. In other words, both map design and map reading can be understood through a cognitive cartographics framework. These ideas are developed in Krygier and Wood's chapter in this volume.

Maps as Immutable Mobiles and Actants

In his book *Science in Action* Bruno Latour (1987) used the example of cartography to explore how the cultures and mechanisms involved in production of Western scientific knowledge gained their power and authority to make truth claims about the world that in turn are employed to do work in the world. He cogently argued that the assemblage of cartographic theory, mapping technologies (e.g. quadrants, sextants, log books, marine clocks, rulers, etc.), and disciplinary regimes of trade and service (e.g. sea captains all taught the same principles and practices of surveying, recording and bringing back spatial data) worked together to enable information from distant places to be accumulated in a cyclical and systematic fashion and for maps to enable appropriate action at a distance.

As the scientific basis of mapmaking and map use became conventionalized, Latour argues that maps increasingly took on the status of immutable mobiles. That is, the mechanisms used to generate cartographic information and the form maps took (in terms of scale, legend, symbols, projection, etc.) became familiar and standardized through protocols so that the map became a stable, combinable and transferable form of knowledge that is portable across space and time. As such, a map produced in South America by Argentinian cartographers is decipherable to someone from another country because it shares common principles that render it legible. Moreover, spatial data transported from South America in the form of latitude and longitude can be used to update charts of the area or be combined with other information, despite the fact that the cartographer is unlikely to have ever visited the area they are mapping.

Mapping then is seemingly transformed into a 'universal' scientific practice and maps become mobile and immutable artefacts through which the world can be known and a vehicle through which spatial knowledge can be transported into new contexts. What is mapped, how it is mapped, and the power of maps is the result of Western science's ability to set the parameters and to dominate the debate about legitimate forms of knowledge. As Latour notes, however, cartographic theory and praxis is seemingly immutable in nature because it disciplines its practitioners and silences other local mapping knowledges. And yet, immutable Western cartographic practice is itself similarly the product of localized practices that are deemed appropriate within a limited circle of practitioners and mapping agencies, who exercise powerful claims to scientific objectivity and truth. The immutability of maps is then at one level a powerful illusion, but one that readily does work in the world.

Latour contends that the immutability, combinability and mobility of maps allowed exploration, trade and ultimately colonialism to develop by allowing control to be exerted from afar and knowledges about new territories to be effectively transported globally. Maps became a vital part in the cycle of knowledge accumulation that allowed explorers to '*bring* the lands *back with them*' and to successfully send others in their footsteps. Latour thus argues that the European cartographers of the Renaissance produced centres of calculation (key sites of cartographic practice) that came to dominate the world. In so doing, maps he suggests do not simply represent space at a particular time, but produce new spaces – times. Maps open up new possibilities – such as international trade and territorial conquest – and thus create new geographies and histories.

To understand maps then, Latour suggests that it is necessary to unpick the cultures, technologies and mechanics of how a particular form of mapping came to gain immutability and mobility to reveal its contingencies and relationalities. Following on from his work, the development of Actor-Network Theory (ANT) in science studies has provided a framework for considering how maps work in concert with other actants and actors to transform the world. ANT involves the tracing out of the context and instruments of mapping – its assemblage – not just cartographic praxis. For example, understanding the road system, Latour argues, cannot be fully realized by looking at infrastructure and vehicles alone, it also needs to consider civil engineering, plans of roads, standards for signage, garages, mechanics, drivers, political lobbying, funding, spare parts and so on.

Maps do not have meaning or action on their own; they are part of an assemblage of people, discursive processes and material things. They are deployed in an actor-network of practices rather than existing as de-corporalized, a priori, non-ideological knowledge objects.

ANT then seeks to provide a broader and richer understanding of the creation of maps through particular actor-networks (e.g. a national mapping agency) and the use of maps as actants within various actor-networks (e.g. land conservation) by considering the diverse, day-to-day practices of, and the interactions and the circulation of ideas and power between, various actors (people, texts, objects and money). In so doing, ANT identifies the nature of 'boundary objects' (objects such as technical standards that enable the sharing of information across networks), 'centres of calculation' (locations such as mapping agencies where observations are accumulated, synthesized and analysed), 'inscription devices' (technical artefacts that record and translate information such as tables of coordinates or satellite imagery), 'obligatory points of passage' (a site in a network that exerts control and influence such as government department), 'programs of action' (the resources required for an actor to perform certain roles) and 'trials by strength' (how competing visions and processes within the network compete for superiority).

From this perspective, the stories of mapping always need to be considered as historically contingent actor-networks; as timed, placed, cultured and negotiated; a Web of interacting possibilities in which the world is complex and nothing is inevitable. The focus shifts from what the map represents to how it is produced and how it produces work in the world.

From Ontology to Ontogenesis: Maps as Practices

In recent years, there has been a move towards considering cartography from a relational perspective, treating maps not as unified representations but as constellations of ongoing processes. Here it is recognized that maps are produced and used through multiple sets of practices. Spatial data are surveyed, processed and cleaned; geometric shapes are drafted, revised, updated, copied, digitized and scanned; information is selected for inclusion, generalized and symbolized. A map is then worked upon by the world and does work in the world. It might be folded or rolled, converted to another file format, embedded in other media; it might be packaged, marketed, sold, bought, used, stored, collected, re-used, thrown away or recycled; it might be read in different ways in different contexts; it might be employed to plan a journey, make money, play a game or teach moral values. Mapmaking and map use are understood as processual in nature, being both embodied and dynamic.

Mapping can then be conceptualized as a suite of cultural practices involving action and affects. This kind of approach reflects a philosophical shift towards performance and mobility and away from essence and material stability. This rethinking of cartography is supported by historical and contemporary work. Researchers concerned with historical contexts increasingly stress the interplay between place, times, actions and ideas.

Mapping in different cultures reflects multiple traditions including an internal or cognitive set of behaviours involving thinking about space; a material culture in which mapping is recorded as an artefact or object; and a performance tradition where space may be enacted through gesture, ritual, song, speech dance or poetry. In any cultural context there will be a different blend of these elements. Interpreting mapping then means considering the context in which mapping takes place; the way it is invoked as part of diverse practices to do work in the world. Instead of focusing on artefacts, aesthetics, human agency, or the politics of mapping, research focuses on how maps are constituted in and through diverse, discursive and material processes.

Arguments presently emerging in the literature extend both the notion of maps as processes and the ontological thought underpinning cartography by problematizing the ontological security enjoyed by maps. The idea that a map represents spatial truth might have been challenged and rethought in a number of different ways, but a map is nonetheless understood as a coherent, stable product – *a* map; a map has an undeniable essence that can be interrogated and from which one can derive understanding. Moreover, the maps and mapping practices maintain and reinforce dualities with respect to their conceptualization – production–consumption, author–reader, design– use, representation–practice, map–space. This position has been rejected by those adopting performative and ontogenetic understandings of mapping. Maps rather are understood as always in a state of becoming; as always mapping; as simultaneously being produced *and* consumed, authored *and* read, designed *and* used, serving as a representation *and* practice; as mutually constituting map/space in a dyadic relationship.

TYPES OF MAPS

Reference Maps: Topographic

Reference or navigational maps are created to help you navigate over the earth surface. These kinds of maps show you where particular places are located and can be used to navigate you way to them. A street map or the common highway road map falls into this category. Physical geographers use topographic maps to show the locations of landscape features on the earth.

Topographic maps illustrate the horizontal and vertical positions (relief) of land surface features. Topographic maps use contour lines to show elevation (height above sea level). Contour lines connect points of equal elevation above a specified reference, usually as sea level. The heavy brown contour lines with the elevation printed on them are called index contours. Intermediate contours are the lighter brown lines between index contours. Sometimes dashed lines called supplemental contours are used in areas of very low relief. Benchmarks are locations where the elevation has been surveyed. Benchmarks are denoted on a map with the letters "BM", "X" or a triangle with the elevation printed beside.

Not only are natural features like mountains, valleys, streams and glaciers portrayed, but cultural features as well, e.g., houses, schools, streets, urbanized area. Take a look at the topographic symbol sheets provided by the United States Geological Survey to get an idea of the information provided on them. Colours and shading can also be used to illustrate relief. Shaded relief maps are great at giving us the over all shape of the surface, but can't help us much in determining the elevation of a particular place. The National Geographic

Society is noted for its excellent map products (and magazine too). Satellite maps are particularly useful to illustrate relief.

Thematic Maps

Thematic maps are used to communicate geographic concepts like the distribution of densities, spatial relationships, magnitudes, movements etc. World climate or soils maps are notable examples of thematic maps. Graduated circles indicate the area over which the earthquakes were felt. This map was created using a geographic information system which has the capability of overlying different kinds of spatial data to show the relationships between them.