Foss 4G Workshop on Open source Geo spatial tools
Coveners

TV Ramachandra	**T G Sitharam**	**M S Mohan Kumar**	**S Narendra Prasad**
Chairman, OSGeo India, Software Development committee	Chairman CiSTUP, IISc	Secretary KSCST	Chairman, OSGeo India, Education and Training Committee

Geoinformatics lab

Centre for *Infrastructure*, Sustainable Transportation and Urban Planning (C*i*STUP)

Indian Institute of Science

Bangalore 560 012, India

http://cistup.iisc.ernet.in

http://wgbis.ces.iisc.ernet.in/foss

CiSTUP

OSGeo - India
Your Open Source Compasss

Foss 4G: Workshop on Open source Geo spatial tools

Workshop Overview

The growth of GIS is hindered by highly priced proprietary softwares and unavailability of public geodata. Many of such licensed software remain unutilized in various government offices because they require either extensive training or easy availability on multiple locations. All this can be changed by the use of Free/Open source software tools. This is significant since all government data and most educational institutions work in local language media. FOSS GIS or Free/open source Geospatial softwares are more powerful and usable for developing systems where more and more public participation is sought. In this regard the proposed workshop during 1 - 5th May 2011 helps to understand the potential and limitations of Open source geospatial tools.

Venue: C*i*STUP Lecture Hall, C*i*STUP, IISc, Bangalore-560012

1st - 5th May 2011 (5 days)

(C*i*STUP jointly with KSCST, OSGEO, IIT (Hyderabad), IIRS and SACON)

Background

CiSTUP, Centre for infrastructure, Sustainable Transportation and Urban Planning was established at IISc (http://cistup.iisc.ernet.in) with the financial aid from the Government of Karnataka to develop a unique expertise having an interface in the areas of infrastructure, sustainable transportation and urban planning. CiSTUP is actively engaged in several activities such as basic and applied research and development, academic activities, training programmes, workshops and consultancy projects in the areas of infrastructure, sustainable transportation and urban planning. Geoinformatics lab has been set up at CiSTUP with the state of the art gadgets in spatial and temporal analysis.

KSCST, the Karnataka State Council for Science and Technology (KSCST) is actively involved in Natural Resources Data Management System (NRDMS) program (http://kscst.org.in/nrdms.html) - a multi-disciplinary and multi-institutional program aimed at developing methodologies for building and promoting the use of spatial data management and analysis technologies in local area planning.

OSGEO, the Open Source Geospatial Foundation (http://osgeo.org) created to support and builds the highest-quality open source geospatial software. The goal is to encourage the use and collaborative development of community-led projects. The OSGeo-India is actively involved in the capacity building in Open Source Geospatial software encouraging student and researchers by conducting lectures and workshops in collaboration with the academic institutions (http://wiki.osgeo.org/wiki/India_Chapter_Report_2007). The India OSGeo Chapter is aiding individuals and institutions interested in Open Source Geospatial Solutions and related issues, covering all activities associated with the application, development and promotion of Open Source Geospatial solutions in India.

IIT (Hyderabad), the International Institute of Information Technology, Hyderabad (IIIT-H) is an autonomous university founded in 1998. It was set up as a not-for-profit public private partnership (NPPP) and is the first IIIT to be set up (under this model) in India. The Government of Andhra Pradesh lent support to the institute by grant of land and buildings. IIIT-H was set up as a research university focused on the core areas of Information Technology, such as *Computer Science, Electronics and Communications,* and their applications in other domains.

IIRS, **Indian Institute of Remote Sensing (IIRS)** under **National Remote Sensing Centre**, **Department of Space**, Govt. of India is a premier training and educational institute set up for developing trained professional in the field of Remote Sensing, Geoinformatics and GPS Technology for Natural Resources, Environmental and Disaster Management. The main area of the function of the Institute is capacity building through technology transfer among user community, education at post-graduate level in the application of Remote Sensing and Geoinformatics for Natural Resources Management and promote research in Remote Sensing and Geoinformatics.

SACON, the Sálim Ali Centre for Ornithology and Natural History was formally inaugurated on 5th June 1990 and registered as a society under the Society Registration Act 1860. SACON, an autonomous organization is a national center for studies in Ornithology and Natural History. The center was named befittingly after Dr. Sálim Ali in appreciation of his lifelong services to India's bird life and conservation of natural resources.

RESOURCE PERSONS

1st - 5th May 2011 (5 days)
(C*i*STUP jointly with KSCST, OSGEO, IIIT (Hyderabad), IIRS and SACON)

Sl.No	Date	Resource Person	Address
1	1/ 05/2011	T V Ramachandra	Geoinformatics Lab, C*i*STUP, IISc, Bangalore
2	1/ 05/2011	Bharath H Aithal	Energy & Wetlands Research Group, CES, IISc, Bangalore
3	1/ 05/2011	Uttam Kumar	Energy & Wetlands Research Group, CES, IISc, Bangalore
4	2/ 05/2011	Bharath Setturu	Energy & Wetlands Research Group, CES, IISc, Bangalore
5	2/ 05/2011	Anindita Dasgupta	Energy & Wetlands Research Group, CES, IISc, Bangalore
6	2/ 05/2011	S. Narendra Prasad	Senior Principal Scientist, Salim Ali Centre for Ornithology & Natural History, Hyderabad
7	2/ 05/2011	Sameer Saran	Geoinformatics Division, Indian Institute of Remote Sensing, Department of Space, Govt of India, 4 Kalidas Road, Dehradun
8	3/ 05/2011	Harish Karnatak	Scientist, Geo Informatics Division, National Remote Sensing Centre, ISRO, Balanagar, Hyderabad
9	3/ 05/2011	Santosh	Salim Ali Centre for Ornithology & Natural History, Hyderabad
10	3/ 05/2011	Achary	Salim Ali Centre for Ornithology & Natural History, Hyderabad
11	4/ 05/2011	Uma Shama	GeoGraphics Laboratory, Bridgewater State College, USA
12	4/ 05/2011	P M. Balamanikavelu	Head, Bhuvan Cell, National Remote Sensing Centre, Hyderabad
13	4/ 05/2011	Arulraj M	Lead Developer, Bhuvan, National Remote Sensing Centre, Hyderabad
14	4/ 05/2011	Chiranjit Mukhopadhyay	Management Studies, IISc, Bangalore

Introduction to GIS & Remote sensing

Many of our decisions depend on the details of our immediate surroundings, and require information about specific places on the Earth's surface. In this regard, recent developments in information technologies have opened a vast potential in communication, analysis of spatial and temporal data. Data representing the real world can be stored and processed so that they can be presented later in a simplified form to suite specific needs. Such information is called geographical because it helps us to distinguish one place from another and to make decisions for one place that are appropriate for that location. Geographical information allows us to apply general principles to the specific conditions of each location, allows us to track what is happening at any place, and helps us to understand how one place differs from another. Spatial information is essential for effective planning and decision-making at regional, national and global levels.

The geographical information in the form of maps (based on field surveys), photos taken from aircraft (aerial photography), and images collected from the space borne platforms (satellite) can be represented in digital form, this opens an enormous range of possibilities for communication, analysis, modeling, and accurate decision making, but a degree of approximation.

GIS can be defined as computerized information storage processing and retrieval system that has hardware, software specially designed to cope with geographically referenced spatial data. Collective name for such system is geographical information systems, (GISs). Processing geographical information includes:

- Techniques to input geographical information, converting the information to digital form

- Technique for sorting such information in a compact format on computer disks, and other digital storage media

- Methods for automated analysis for geographical data, to search for the patterns, combine different kinds of data, make measurements find optimum sites or routes, and a host of other tasks

- Methods to predict the outcome of various scenarios, such as the effects of climate change on vegetation

- Techniques for display of data in the form of maps, images and other kinds of display

- Capabilities for output of results in the form of numbers and tables.

Elements of GIS:

Components of geographical data are *Spatial and Attribute Database, Cartographic Display System, Map Digitizing System, Database Management System, Geographic Analysis System, Statistical analysis system and Decision support system.* The linkages among these components are illustrated in Figure 1.1.
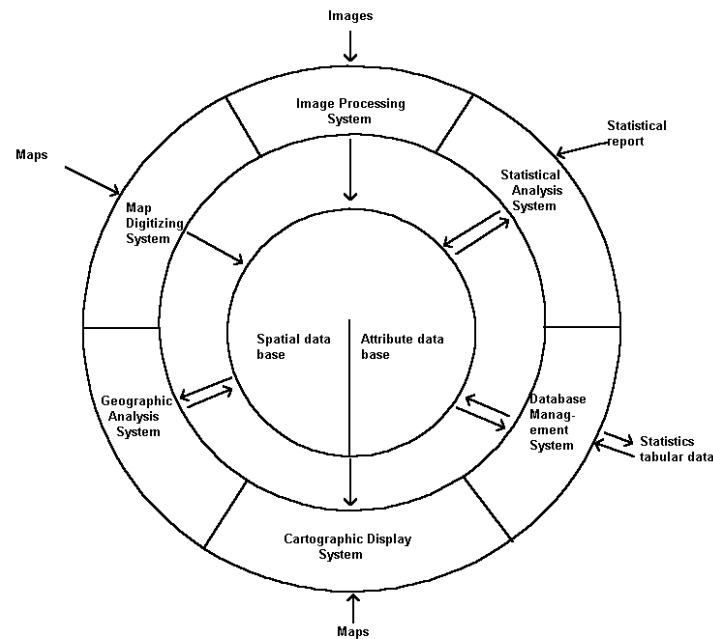
Figure 1.1: Components of GIS.

i)   *Spatial and Attribute Database:* Central to the system is the database – a collection of maps and associated information in digital form. Since the database is concerned with earth surface features, it is seen to comprise of two elements – a spatial database describing the geology (shape and position) of the earth surface features, and an attribute database describing the characteristics or quantities of these features. Thus, for example, we might have a property parcel defined in the spatial database and qualities such as its land use, owner, property valuation, etc. in the attribute database.

ii)   *Cartographic Display System:* Surrounding the central database, we have a series of software components. The most basic of these is the cartographic display system. The cartographic display system allows one to take selected elements of the database and produce map output on the screen or some hardcopy device such as printer or plotter.

iii) ***Map Digitizing System:*** After cartographic display, the next most essential element is a Map Digitization System. With a map digitizing system, one can take existing paper maps and convert them into digital form, thus further developing the database. In the most common method of digitizing, one attaches the paper map to a digitizing tablet or board and then traces the features of interest with a stylus according to the procedures required for digitizing. Many maps digitizing system also allows for some editing of the digitized data. Scanners can also be used to digitize data such as aerial photographs. The results is a graphic image, rather than the outlines of features that are created with a variety of standard graphics file formats for export. These files are then imported into the GIS. Computer assisted design (CAD) and Coordinate Geometry (COGO) are two examples of software systems that provide the ability to add digitized map information to the database, in addition to providing cartographic display capabilities.

iv) ***Database Management System***: The next logical component in a GIS is Database Management System (DBMS), which is used to input, manage and analyze attribute information along with then spatial data. GIS thus typically incorporates a variety of utilities to manage the spatial and attribute components of the geographic data. DBMS aids to enter attribute data, such as tabular information and statistics, and subsequently extract specialized tabulations and statistical summaries to provide new tabular reports. The DBMS provides the ability to analyze attribute data. Many map analysis have no true spatial component, and for these a DBMS will often function quite well. For example, we might inquire of the system to find all property parcels where the head of the household is single but with one or more child dependents, and to produce a spatial map. Software that provides cartographic display, map digitizing, and database query capabilities are often referred to as Automated Mapping and Facilities Management (AM/FM) system.

v) ***Geographic Analysis System:*** Up to this point, we have described a very powerful set of capabilities that the GIS offer, the ability to digitize spatial data and attach attribute to the features stored; to analyze these data based on those attribute; and to map to the result. But on inclusion geographic analysis system, we extend the capabilities of the traditional database query to include the ability to analyze data based on their location. Perhaps the simplest example of this is to consider what happens when we are concerned with the joint occurrence of features with different geographies. For example, suppose we want to find all areas of residential land on bedrock types associated with high levels of radon gas. A traditional DBMS cannot solve this problem because bedrock types and landuse divisions simply do not share the same geography. Traditional database query is fine as long as we are taking about attributes belonging to the same features. But when the features are different, it cannot cope. For this we need a GIS. In fact, it is this ability to compare different feature based on their common geographic occurrence that is the hallmark of GIS. This analysis is accomplished by the process of overlay, thus named because it is identical in character to overlaying transparent maps of the two entity groups on top of one another. Like the DBMS, the Geographic Analysis System as highlighted in Figure 1.1 has a two-way interaction with the database; the process is distinctly analytical in character. Thus while it may access data from the database, it may equally contribute the results of that analysis as a new addition to the database. For example we might look

for joint occurrence of lands on steep slopes with erodable soil under agriculture and call the results based on existing data and set of specific relations. Thus the analytic capabilities of the Geographic Analysis System and the DBMS play a vital role in extending the database through the addition of knowledge of relationships between features.

vi) **Image Processing System:** In addition to these essential GIS elements, remotely sensed image and specialized statistical analysis are also important. This we will discuss in the subsequent sections.

vii) **Statistical analysis system:** GIS incorporates a series of specialized routines for analyzing the statistical description of spatial data and for inferences drawn from statistical procedures.

viii) **Decision support system (DSS):** Decision support constitutes a vital function of a GIS. It helps in the construction of multi-criteria suitability maps, and address allocation decisions when there is multiple objectives involved while accounting for error in the process. Used in conjunction with the other components of the system, DSS provides a powerful tool in decision-making for resource allocation.

Map Data Representation

A Geographic Information System stores two types of data that are found on a map—the geographic definitions of earth surface features and the attributes or qualities that those features possess. Most systems use nearly one or a combination of both the fundamental map representation techniques: vector and raster.

Vector:

This refers to the spatial data represented in the form of point, line or polygon depending on the feature of interest (and scale). With vector representation, the boundaries or the course of the features are defined by a series of points that, when joined with straight lines, form the graphic representation of that feature. The points themselves are encoded with a pair of numbers giving the X and Y coordinates in systems such as latitude/ longitude, etc. The attributes of features are then stored in the database management system (DBMS). For example, a vector map of property parcels might be tied to an attribute database of information containing the address, owner's name, property valuation and land use. The link between these two data files can be a simple identifier number that is given to each feature in the map (Figure: 1.2).

Raster: In this case, the graphic representation of features and the attributes they possess are merged into unified data files. In fact, we typically do not define features at all. Rather, the study area is subdivided into a fine mesh of grid cells in which we record the condition or attribute of the earth's surface at that point (Figure 1.2). Each cell has a numeric value (often

referred as digital number or spectral signature), representing a feature identifier, a qualitative attribute code or a quantitative attribute value. For example, a cell could have the value "6" to indicate that it belongs to District 6 (a feature identifier), or that it is covered by soil type 6 (a qualitative attribute), or that it is 6 meters above sea level (a quantitative value). Although the data we store in these grid cells do not necessarily refer to phenomena that can be seen in the environment, the data grids themselves can be thought of as images or layers, each depicting one type of information over the mapped region. This information can be made visible through the use of a raster display. In a raster display, such as the screen on your computer, there is also a grid of small cells called pixels (or picture elements). The word pixel is a contraction of the term picture element. Pixels vary in their color, shape or gray tone depending on features in the object. To make an image, the cell values in the data grid are used to regulate directly the graphic appearance of their corresponding pixels. Thus in a raster system, the data directly controls the visible form we see. *Raster versus Vector:* Raster systems are typically data intensive since they must record data at every cell location regardless of whether that cell holds information that is of interest or not. However, the advantage is that geographical space is uniformly defined in a simple and predictable fashion. As a result, raster systems have substantially more analytical power than their vector

Vector   Raster
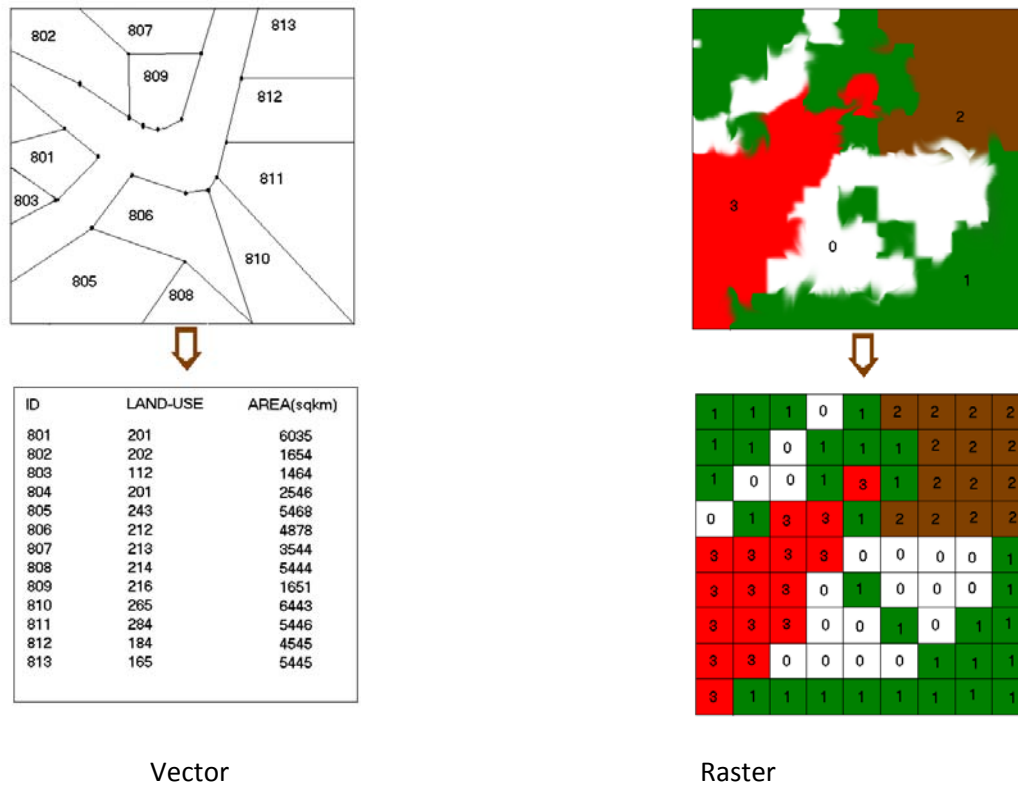
Figure1.2

counterparts in the analysis of continuous space and are thus ideally suited to the study of data that are continuously changing over space such as terrain, vegetation biomass, rainfall and the like. The second advantage of raster is that its structure closely matches the architecture of digital computers.

As a result, raster systems tend to be very rapid in the evaluation of problems that involve various mathematical combinations of the data in multiple layers. Hence they are excellent for evaluating environmental models such as soil erosion potential and forest management suitability. In addition, since satellite imagery employs a raster structure, most raster systems can easily incorporate these data, and some provide full image processing capabilities.

While raster systems are predominantly analysis oriented, vector systems tend to be more database management oriented. Vector systems are quite efficient in their storage of map data because they only store the boundaries of features and not that which is inside those boundaries. Because the graphic representation of features is directly linked to the attribute database, vector systems usually allow one to roam around the graphic display with a mouse and query the attributes associated with a displayed feature, such as the distance between points or along lines, the areas of regions defined on the screen, and so on. In addition, they can produce simple thematic maps of database queries.

Compared to their raster counterparts, vector systems do not have as extensive a range of capabilities for analyses over continuous space. They do, however, excel at problems concerning movements over a network and can undertake the most fundamental of GIS operations that will be sketched out below. For many, it is simple database management functions and excellent mapping capabilities that make vector systems attractive. Because of the close affinity between the logic of vector representation and traditional map production, a pen plotter can be driven to produce a map that is in distinguishable from that produce by traditional means. As a result, vector systems are very popular in municipal applications where issues of engineering map production and database management predominate.

Geographic database concepts:

Regardless of the logic used for spatial representation, raster and vector, we begin to see that a geographic database as a Complete database for a given region and is organized in a fashion similar to a collection of maps. Vector systems come closest to this logic with what are known as coverages. Map like collection that contain the geographic definition of a set of features and their associated attributes tables. However, they differ from maps in two ways. First, each will typically contain information on only a single feature types, such property parcels, soil polygons, and the like. Second, they may contain a whole series of attributes that pertain to those features, such as a set of census information for city blocks.

Raster system also uses this map like logic, but usually divides data sets into unitary layers. A layer contains all the data for a single attribute. Thus one might have a soil layer, a road layer and a land-use layer.

There are subtle differences, for all intents and purposes, raster layer and vector coverage can be thought of as simply different manifestations of the same concepts as the organization of the database into elementary map-like themes. Layers and coverage differ from traditional paper maps, however, in an important way. When a map is digitized, scale differences are

removed. The digital data may be displayed or printed at any scale. More importantly, digital data layers that were derived from maps of different scale, but covering the same geographic area, may be combined.

GIS provide utilities for changing the projection and reference system of digital layers. This allows multiple layers, digitized from maps having various projections and reference system, to be converted to a common system.

With the ability to manage differences of scale, projection and reference system, layers can be merged with ease, elimination a problem that has traditionally hampered planning activities with maps. It is important to note, however, that the issue of resolution of the information in the data layers remains. Although features digitized from a poster sized world map could be combined in a GIS with features digitized from very large-scale local map, such as a city street map, this would normally not be done. The level of accuracy and detail of the digital data can be as good as that of the original maps.

Georeferencing:

 All spatial data files in GIS are georeferenced. Georeferencing refers to the location of a layer or coverage in the space as a definition by a known coordinate referring system. With raster images, a common form of georeferencing is to indicate the reference system, the reference units and the coordinate positions of the left, right, top, and bottom edges of the image. The same is true of the vector data files, although the left, right, top and bottom edges now refer to what is commonly called the bounding rectangle of the coverage; rectangle which defines the limit of the mapped area (corners of a feature). This information is particularly important in an integrated GIS since it allows raster and vector files to be related to one another in a reliable and meaningful way. It is also vital for the referencing of the data values to actual positions on the ground.

GIS Applicability:

The society is so complex, and their activities so interwoven, that no problem can be considered in isolation or without regard for the full range of its interconnections. For example, a new housing development will affect the local school system. The volume of city traffic put constraints on the maintenance of buried pipe networks, affecting health. The action needed to solve such a problems are best taken on the basis of standardized information that can be combined in many ways to serve many users. GISs have this capability.

Environmental and resource management: Decision making is becoming increasing complex as dwindling natural resources and more demanding economic priorities diminish the chances of today's decision being right tomorrow. Furthermore, environmental awareness is constantly increasing among the general public, particularly among the younger generation. To help us map and monitor changes, and plan appropriate responses that take account of the

complex interactions of the Earth system, many countries now have comprehensive programs to capture and archive information on the existing natural resources and known sources of pollution, using technologies such as satellite remote sensing and GIS. The data may be used both to expose conflicts and to examine environmental impacts and even simulate the causes and the alternative will become possible.

Planning and development

The planning and development of new housing, roads, and industrial facilities require data on the terrain and other geographical information. Development often involves building on marginal terrain, increasing the density of the building in the areas already built up, or both. Yet the new structures must fit within the existing technical infrastructure; here computerization is a great aid. One of the benefits GIS holds for such projects is a mineralization of disruption to the existing infrastructure.

Escalating construction costs have made the optimizing of building and road location extremely important. Minimizing blasting and earthmoving are significant aspect of minimizing costs. Flexibility is vital: plans should be amenable to rapid changes as decisions are made. The influence of special interest groups and individual citizens require that initial plans be presented effectively and in a manner that is easily understood. Simplified, visualized plans are instrumental in conveying both the content of the scheme and the nature of any likely impact on those concerned.

Management and public services:

In modern societies, decisions should be made quickly, using reliable data, even though there may be many differing viewpoints to consider and large amount of information to process. Today, the impact of development decisions is ever greater, involving conflicts between society and individuals, or between development and preservation. Information must therefore be readily available to decision makers; the majority of such information is likely to be geographical in nature, and best handled using GIS.

Overviews of administrative units and properties are crucial in the development of both virgin terrain and built-up area, in both developed and developing nations. In many countries, property registration is extensive: even in smaller states, 2 to 3 million properties maybe involved. Moreover, property is also an economic factor in taxation and security for loans; so comprehensive overviews are essential to a well-ordered society. Computerized registers based on GIS technology are now well established in many countries.

Land transportation:

In many countries, the greater part of transportation has shifted from rail to road, at the same time, the use of private vehicle has greatly increased. These developments have created traffic problems, which cause loss of time and money. Large goods are now transported by road. In most countries the annual costs of traffic accidents have become extremely high.

The automobile industry is now investing heavily in the development of driver information system, and several systems are now in the market. In principle, all of them involve simple GIS function with digital maps and supplementary information.

Remote sensing

Remote sensing brings up the events of recording, observing, and sensing objects or events in faraway (remote) places. In remote sensing, the sensors are not in direct contact with the objects or events being observed. Electromagnetic radiation normally is used as the information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed.

Remote sensing can define in two forms-

Active

Passive

Passive remote sensing systems record the reflected energy of electromagnetic radiation or the emitted energy from the earth, such as cameras and thermal infrared detectors.

Active remote sensing systems send out their own energy and record the reflected portion of that energy from the earth's surface, such as radar imaging systems.

Remote Sensing involves four basic inputs: 1. The Target; 2. The Platform; 3. The Sensor(s); and 4. The Signal (usually electromagnetic radiation or acoustical waves)
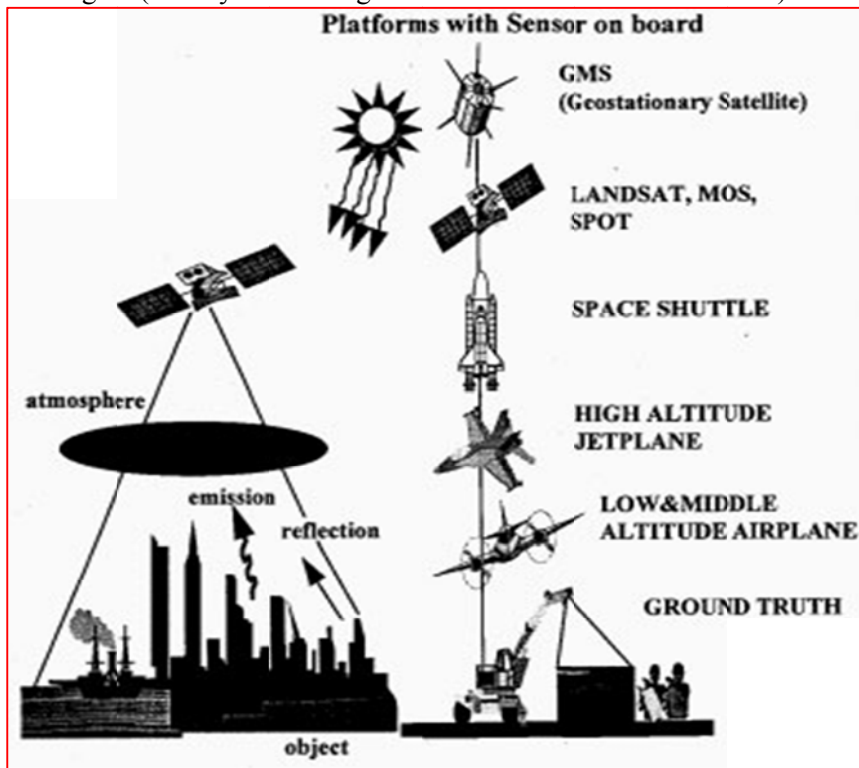
Principles of Electromagnetic Radiation

Electromagnetic radiation is a form of energy with the properties of a wave, and its major source is the sun. Solar energy traveling in the form of waves at the speed of light (denoted as *c* and equals to $3*10^8 ms{-1}$) is known as the *electromagnetic spectrum*. The wavelength is λ the distance between successive crests of the waves. The frequency μ is the

number of oscillations completed per second. Wavelength and frequency are related by the following equation:

$C = \lambda * \mu$
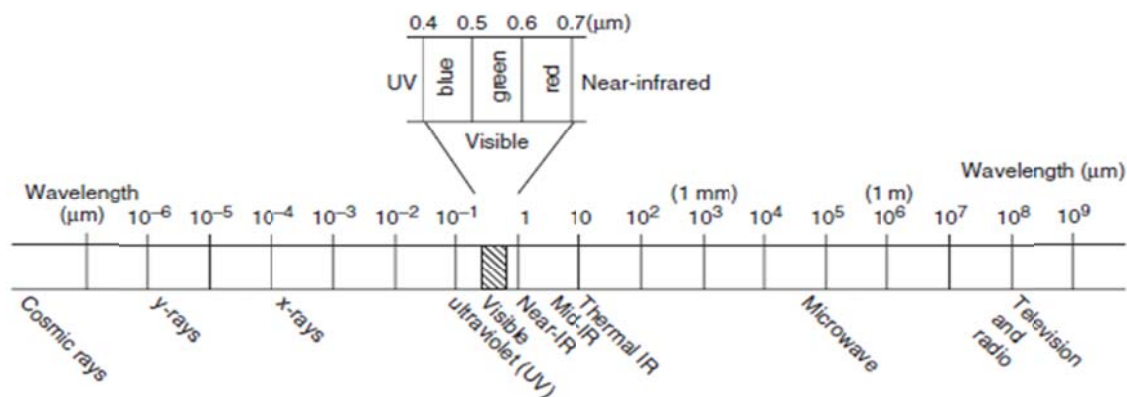
The electromagnetic spectrum, despite being seen as a continuum of wavelengths and frequencies, is divided into different portions by scientific convention (Fig. 1.1). Major divisions of the electromagnetic spectrum, ranging from short-wavelength, high-frequency waves to long-wavelength, low-frequency waves, include gamma rays, x-rays, ultraviolet (UV) radiation, visible light, infrared (IR) radiation, microwave radiation, and radio waves. The visible spectrum, commonly known as the *rainbow of colors* we see as visible light (sunlight), is the portion of the electromagnetic spectrum with wavelengths between 400 and 700 billionths of a meter (0.4–0.7 m). Although it is a narrow spectrum, the visible spectrum has a great utility in satellite remote sensing and for the identification of different objects by their visible colors in photography. The IR spectrum is the region of electromagnetic radiation that extends from the visible region to about 1 mm (in wavelength). Infrared waves can be further partitioned into the near-IR, mid-IR, and far-IR spectrum, which include thermal radiation. IR radiation can be measured by using electronic detectors. IR images obtained by sensors can yield important information on the health of crops and can help in visualizing forest fires even when they are enveloped in an opaque curtain of smoke. Microwave radiation has a wavelength ranging from approximately 1 mm to 30 cm. Microwaves are emitted from the earth, from objects such as cars and planes, and from the atmosphere. These microwaves can be detected to provide information, such as the temperature of the object that emitted the microwave. Because their wavelengths are so long, the energy available is quite small compared with visible and IR wavelengths. Therefore, the fields of view must be large enough to detect sufficient energy to record a signal. Most passive microwave sensors thus are characterized by low spatial resolution. Active microwave sensing systems (e.g., radar) provide their own source of microwave radiation to illuminate the targets on the ground.

Remote sensing, geographic information system (GIS) and global positioning system (GPS) provides extremely useful tools for environmental and natural resources management. They are widely recognized as supporting tools for the planning, monitoring, and management of the appropriate utilization of resources at the country, regional and global levels.

FOSS for Geoinformatics (FOSS4G)

Ramachandra T V and Uttam Kumar

Geoinformatics Laboratory
Centre for Infrastructure, Sustainable Transport and Urban Planning
Indian Institute of Science, Bangalore 560 012
E Mail: cestvr@cistup.iisc.ernet.in

Geoinformatics constitute a vital component of information science for addressing the problems of geography, geosciences and related branches of engineering. This domain combines geospatial analysis and modeling, development of geospatial databases, information systems design, human-computer interaction and both wired and wireless networking technologies. Geoinformatics uses geocomputation for analysing geoinformation. One of the major applications of geoinformatics in recent times is the study of variation in landscapes over multiple spatial and temporal scales encompassing a variety of domains – land use and land cover change, climate change, water resources, urban development, natural disaster mitigation, etc. Geoinformatics include geographic information systems, spatial decision support systems, global positioning systems (GPS), and remote sensing.

Geographic Information Systems (GIS) are increasingly being used as the principal tool for digital exploration of variation in landscapes, as they provide the necessary functions for spatial data collection, management, analysis and representation (Turner et al., 2001; Longley et al., 2005; Steiniger and Weibel, 2009; Ramachandra et al., 2004).These tools

provide new and critical ways of understanding our earth and its biogeochemical cycle. GIS software used for this kind of studies fulfill several GIS functionality including:

i.) Ensure world wide development, advancement and application of solutions.
ii.) Allow studying of data, methods and algorithm implementation.
iii.) Furthermore, developed models and algorithms need not be reimplemented by others in order to continue research or validate previous results.

Apart from these, researchers should have access to libraries of the original models for analysis, validation, development and implementation (Steiniger and Hay, 2009; Jolma et al., 2008b) for further improvement and customisation depending on the local requirement.

Over the last years the paradigm of Free and Open Source Software (FOSS) development has taken root in the GIS community, resulting in the creation of several sophisticated GIS software projects whose aim is to develop free software for numerous purposes. GIS software fulfilling the specific requirements have been distributed with licenses that grant more freedoms of use and that support openness, such as licenses used by FOSS GIS projects (for example: http://grass.itc.it/; http://wgbis.ces.iisc.ernet.in/grass).

FOSS have proved to be promising tools that allow us to see and change the software codeswritten in any programming language. FOSS is generally synonymous with free softwareand open source software, and describes similar development models, but with differing cultures and philosophies. Because of the way it is licensed, it has the potential to be legally given away for free or for very little cost and copied and shared with others.

FOSS, F/OSS or FLOSS (for Free/Libre/Open Source Software) is liberally licensed to grant the right of users to study, change, and improve its design through the availability of its source code. It has more scope for being available in multiple languages and for being adapted or tweaked to particular needs. This can be very useful for students, researchers, teachers, scientists wanting to use legal software that is appropriate to their needs and fits within their modest budgets. This approach has gained both momentum and acceptance as the potential benefits have been increasingly recognised by many (Steiniger and Hay, 2009).This is proving to be the boon to researchers from economically disadvantaged countries.

The open source definition is used by the Open Source Initiative to determine whether or not a software license can be considered open source. Under the open source definition, licenses must meet the following ten conditions in order to be considered open source licenses:

Free redistribution: the software can be freely given away or sold. (This was intended to expand sharing and use of the software on a legal basis.)
Source code: the source code must either be included or freely obtainable. (Without source code, making changes or modifications can be impossible.)

- Derived works: redistribution of modifications must be allowed. (To allow legal sharing and to permit new features or repairs.)
- Integrity of the author's source code: licenses may require that modifications are redistributed only as patches.
- No discrimination against persons or groups: no one can be locked out.
- No discrimination against fields of endeavor: commercial users cannot be excluded.
- Distribution of license: the rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.
  License must not be specific to a product: the program cannot be licensed only as
- part of a larger distribution.
  License must not restrict other software: the license cannot insist that any other
- software it is distributed with must also be open source.
  License must be technology neutral: no click-wrap licenses or other medium-
- specific ways of accepting the license must be required.

There is a distinction between open source software and free software. Open source software are those, for which the human-readable source code is made available under a copyright license (or arrangement such as the public domain) that meets the Open Source Definition. This permits users to use, change and improve the software, and to redistribute it in modified or unmodified form. It is often developed in a collaborative manner in a public domain. Public domain comprises the body of knowledge and innovation (especially creative works such as writing and inventions) in relation to which no person or other legal entity can establish or maintain proprietary interests within a particular legal jurisdiction. This body of information and creativity is considered to be part of a common cultural and intellectual heritage, which, in general, anyone may use or exploit, whether for commercial or non-commercial purposes. Public domain software is not protected by copyright and may be copied and used without payment (http://www.fsf.org/; wikipedia).

On the other hand, free software is software that can be used, studied, and modified without restriction, and which can be copied and redistributed in modified or unmodified form either without restriction, or with restrictions only to ensure that further recipients can also do these things. To make these acts possible, the human readable form of the program (called the source code) must be made available. The source code can be placed in the public domain, accompanied by a software license saying that the copyright holder permits these acts (a free software licence), etc. (http://www.fsf.org/).

The first formal definition of free software states that software is free software if people who receive a copy of the software have the following four freedoms:

1.) Freedom 0: The freedom to run the program for any purpose.
2.) Freedom 1: The freedom to study and modify the program.
3.) Freedom 2: The freedom to copy the program so you can help your neighbor.

4.) Freedom 3: The freedom to improve the program, and release your improvements to the public, so that the whole community benefits.

Freedoms 1 and 3 require source code to be available because studying and modifying software without its source code is highly impractical (http://www.fsf.org/).

Proprietary software has restrictions on copying and modifying as enforced by the proprietor. Restrictions on modification and copying are sought by either legal or technical means or sometimes both. Technical means include releasing machine-readable binaries to users and withholding the human-readable source code. Legal means can involve software licensing, copyright, and patent law.

Copyleft is a form of licensing and may be used to modify copyrights for works such as computer software, documents, etc. In general, copyright law allows an author to prohibit others from reproducing, adapting, or distributing copies of the author's work. In contrast, an author may, through a copyleft licensing scheme, give every person who receives a copy of a work permission to reproduce, adapt or distribute the work as long as any resulting copies or adaptations are also bound by the same copyleft licensing scheme. A widely used and originating copyleft license is the GNU General Public License.

The GNU General Public License (GNU GPL or simply GPL) is a widely used free software license, originally written by Richard Stallman for the GNU project. It is the license used by the Linux kernel. The GPL is the most popular and well-known example of the type of strong copyleft license that requires derived works to be available under the samecopyleft. Under this philosophy, the GPL is said to grant the recipients of a computer program the rights of the free software definition and uses copyleft to ensure the freedoms are preserved, even when the work is changed or added to. This is in distinction to permissive free software licences, of which the BSD (Berkeley Software Distribution) licenses are the standard examples. The GNU Lesser General Public License (LGPL) is a modified, more permissive, version of the GPL, intended for some software libraries.

FOSS has become an essential component in geoinformatics research. Many free and open source software are available that facilitate customisation, provide good support via forums and email lists and have up-to-date documentation. Among the many GIS tools that are frequently used are Desktop GIS, Mobile GIS, Remote Sensing and Image Processing software, GIS extensions and libraries, Spatial Database Management Systems, Map Server and Geostatistical tools (Steiniger and Hay, 2009).

Next, we present a non-comprehensive list of the FOSS commonly used in GIS applications along with their web address for further references.

	Application	Software	References
1	GIS	Sav GIS	http://www.savgis.org
2		Forestry GIS	http://www.forestpal.com/fgis.html
3		GRASS	http://wgbis.ces..isc.ernet.in/grass
4		QGIS	http://qgis.org
5		ImageILWIS	http://ilwis.org
6		Processing /uDIG	http://udig.refractions.net
7		Raster, Vector SAGA	http://saga-gis.org
8		Analysis	http://openjump.org
9		OpenJUMP	http://mapwindow.org
10		MapWindow	http://gvsig.gva.es
11		gvGIS	http://www.lvc.ele.puc-rio.br/projects/
12		InterImage	http://www.landserf.org
13		Landserf	http://www.ossim.org
14		OSSIM	http://grass.itc.it/grass70/manuals/html70_user/r.li.html
		Landscaper.li	
		(GRASS)	
		Analysis	http://www.umass.edu/landeco/research/fragstats/
15		Fragstats	fragstats.html
16	Spatial DBMS	MySQL	http://www.mysql.org
		PostGIS	for http://postgis.refractions.net
17		Postgre SQL	
	Web MapServer	GeoServer	http:// www.geoserver.org
		MapServer	http:// www.mapserver.org
18		R	http://www.r-project.org/
19	Statistical software	Gstat	http://www.gstat.org
20		Past	http://folk.uio.no/ohammer/past/
21		GeoDa	http://geodacenter.asu.edu/software
22		GeoVista	http://www.geovistastudio.psu.edu
23	Exploratory Data Analysis	STARS	http://regionalanalysislab.org/index.php/Main/STARS
24		GDAL/OGR	http://gdal.osgeo.org
25		Generic	
		Mapping Tool	http://gmt.soest.hawaii.edu/
26		JAMA/GNU	
27		JTS Topology	
		Suite	http://math.nist.gov/javanumerics/jama/
28		LUPOLib	http://tsusiatsoftware.net/jts/main.html
29		OpenBugs	http://www.ufz.de/index.php?en=4302
30		Sextane	http://mathstat.helsinki.fi/openbugs/
		TerraLib	http://forge.osor.eu/projects/sextante/
31	Libraries	MASON	http://www.terralib.org/
32		Rapast	http://cs.gmu.edu/~eclab/projects/mason/
33		Simphony	http://repast.sourceforge.net/
34		SWARM	http:// www.swarm.org
35		Netlogo	http://ccl.northwestern.edu/netlogo/
36			
37			

38	Multi Agent	(Open-)Start Logo	http://education.mit.edu/openstarlogo/
39		OBEUS	http://www.tau.ac.il/~bennya/research1.html
40		Drupal	http://drupal.org/
41		Joomla	http://www.joomla.org/
42		Atutor	http://www.atutor.ca/
43		Open Modeller	
44		Desktop GARP	http://openmodeller.sourceforge.net/
45	CMS	Maxent	http://www.cs.princeton.edu/~schapire/maxent/
46		Landisview	http://www.nhm.ku.edu/desktopgarp/index.html
47	Modelling	PCRaster	
48		SANET	http://sourceforge.net/projects/landisview
49		S-Distance	http://pcraster.geo.uu.nl
			http://ua.t.u-tokyo.ac.jp/okabelab/atsu/ sanet/ http://www.prd.uth.gr/res_labs/spatial_analysis/software/SdHome_en.asp
50	Miscellaneous	TAS	http://www.uoguelph.ca/~hydrogeo/TAS/
51		Harvest	http://www.nrs.fs.fed.us/tools/harvest/
52		Qrule	

Proprietary software licenses impose several restrictions on the use of software such as not allowing users to distribute the software or to install it on a second computer or to give it to others. The licenses also prohibit a reverse engineering of the software and modifying the software. If the source code is not available, then it is not possible to study how algorithms are implemented and it is not possible to improve the software. FOSS licenses, such as the GPL and the LGPL, explicitly allow users to study, modify and re-distribute software. Consequently the following three benefits of FOSS have been identified(Steiniger and Bocher, in press):

1. FOSS avoids 'reinventing the wheel',
2. in terms of the source code, FOSS provides the best 'documentation' available, and
3. users can adapt the software to their own needs without restrictions.

All three points are essential for research, when considering that research should not be limited by the functionality that is provided by the software, research experiments need to be repeatable and reproducible, and research can progress faster when models can be analysed, validated, and improved directly, i.e. based on source code, without the problem of misinterpretation, as may be the case when knowledge is obtained/interpreted from articles (Steiniger and Hay, 2009). In addition to these general research advantages, the use of FOSS licenses have enhanced education and knowledge transfer, particularly in developing countries that don't have the (financial) resources. Students and researchers can freely and legally download the software and study the algorithms. Finally it benefits society in general, as the use of free software licenses can facilitate the application of new technologies and knowledge thatenables a sustainable use of resources (Jolma et al., 2008b). If such unified software development and research efforts could be initiated then we see great potential to accelerategeoinformatics research world wide.

References:
• Jolma, A., Ames, D.P., Horning, N., et, al., 2008a. Free and open source geospatial tools for environmental modelling and management. In: Jakeman, A.J., Voinov, A.A., Rizzoli, A.E., Serena, H.C. (Eds.), Environmental Modelling, Software and Decision Support. Elsevier, Amsterdam, pp. 163–180.
• Jolma, A., Ames, D.P., Horning, N., et, al., 2008b. Environmental modeling using open source tools. In: Shekhar, S., Xiong, H. (Eds.), Encyclopedia of GIS. Springer, New York, pp. 275–279.
• Longley, P.A., Goodchild, M.F., Maguire, D.J., and Rhind, D.W., 2005. Geographic information systems and science, 2nd ed. Wiley, Chichester.
• Steiniger, S., and Bocher, E., in press. An overview on current free and open source desktop GIS developments. International Journal of Geographical Information Science.doi:10.1080/13658810802634956.
• Steiniger, S., and Hay, G. J., 2009. Free and open source geographic information tools for landscape ecology. Ecological Informatics, 4, 183-195.
• Steiniger, S., and Weibel, R., 2009. GIS software — a description in 1000 words. Available from: http://www.geo.unizh.ch/publications/sstein/gissoftware_steiniger2008.pdf.
• Turner, M.G., Gardner, R.H., and O'Neill, R.V., 2001. Landscape ecology in theory and practice pattern and processes. Springer, New-York.

Useful links:

http://wgbis.ces.iisc.ernet.in/foss
http://www.opensourcegis.org
http://www.spatialserver.net/osgis
http://www.spatialanalysisonline.com
http://www.ai-geostats.org
http://ces.iisc.ernet.in/grass Mirror site for GRASS in India
http://ces.iisc.ernet.in/biodiversity Geoinformatics Applications
http://ces.iisc.ernet.in/energy

Introduction to Linux and Installing Ubuntu

Linux is an operating system: a series of programs that let you interact with your computer and run other programs.Linux is modeled on the UNIX operating system. From the start, Linux was designed to be a multi-tasking, multi-user system. These facts are enough to make Linux different from other well-known operating systems. However, Linux is even more different than you might imagine. Linux has a capability to be installed on any type computer hardware. Linux is a leading server operating system, and runs the 10 fastest supercomputers in the world (21st fastest can be found @SERC, IISc).

The development of Linux is one of the most prominent examples of free and open source software collaboration; typically all the underlying source code can be used, freely modified, and redistributed, both commercially and non-commercially, by anyone under licenses such as the GNU General Public License. Typically Linux is packaged in a format known as a Linux distribution for desktop and server use. Some popular mainstream Linux distributions include Debian (and its derivatives such as Ubuntu), Red hat, Fedora and openSUSE. Linux distributions include the Linux kernel and supporting utilities and libraries to fulfill the distribution's intended use.

A distribution oriented toward desktop use may include the X Window System, the GNOME and KDE desktop environments. A distribution intended to run as a server may omit any graphical environment from the standard install and instead include other software such as the Apache HTTP Server and a SSH server like OpenSSH. Because Linux is freely redistributable, it is possible for anyone to create a distribution for any intended use. Commonly used applications with desktop Linux systems include the Mozilla Firefox web browser, the OpenOffice.org office application suite.

The main supporting user space system tools and libraries from the GNU Project (announced in 1983 by Richard Stallman) are the basis for the Free Software Foundation's preferred name GNU/Linux.

- Ubuntu is a complete Open source - desktop Linux operating system, freely available with both community and professional support. And people should have the freedom to customize and alter their software in whatever way they see fit for their Use.
- Ubuntu will always be free of charge, and there is no extra fee for the "enterprise edition", we make our very best work available to everyone on the same Free terms.

- Ubuntu includes the very best in translations and accessibility infrastructure that the Free Software community has to offer, to make Ubuntu usable by as many people as possible.
- Ubuntu is shipped in stable and regular release cycles; a new release will be shipped every six months. You can use the current stable release or the current development release. A release will be supported for 18 months.
- Ubuntu is entirely committed to the principles of open source software development; we encourage people to use open source software, improve it and pass it on.
- Ubuntu is suitable for both desktop and server use. The current Ubuntu release supports Intel x86 (IBM-compatible PC), AMD64 (Hammer) and PowerPC (Apple iBook and Powerbook, G4 and G5) architectures.

Sponsorship by Canonical

The Ubuntu Project is sponsored by Canonical Ltd. Canonical will not charge license fees for Ubuntu, now or at any stage in the future. Canonical's business model is to provide technical support and professional services related to Ubuntu. We encourage more companies also to offer support for Ubuntu, and will list those that do on the Support pages of this web site.

Debian:

Debian is an all-volunteer organization dedicated to developing free software and promoting the ideals of the Free Software community.

Ubuntu and Debian

Ubuntu and Debian are distinct but parallel and closely linked systems. The Ubuntu project seeks to complement the Debian project in the following areas:

Package selection
Ubuntu does not provide security updates and professional support for every package available in the open source world, but selects a complete set of packages making up a solid and comprehensive desktop system and provides support for that set of packages. For users that want access to every known package, Ubuntu provides a "universe" component (set of packages) where users of Ubuntu systems install the latest version of any package that is not in the supported set. Most of the packages in Ubuntu universe are also in Debian, although there are other sources for universe too. See the Ubuntu Components page for more detail on the structure of the Ubuntu web distribution.

Development community
Many Ubuntu developers are also recognized members of the Debian community. They continue to stay active in contributing to Debian both in the course of their work on Ubuntu and directly in Debian.

Getting Ubuntu

Ubuntu can be easily downloaded based on the architecture of the system in use. Ubuntu parent website http://www.ubuntu.com/ support various mirrors located in various countries to download various version of Ubuntu CD/DVD. Burn the ISO image into the media.

If your machine doesn't support CD booting, but you do have a CD, you can use an alternative strategy such as hard disk,usb stick, net boot, or manually loading the kernel from the CD to initially boot the system installer. The files you need for booting by another means are also on the CD; the Ubuntu network archive and CD folder organization are identical. So when archive file paths are given below for particular files you need for booting, look for those files in the same directories and subdirectories on your CD.

Once the installer is booted, it will be able to obtain all the other files it needs from the CD.If you don't have a CD, then you will need to download the installer system files and place them on the hard disk or usb stick or a connected computer so they can be used to boot the installer.

System Requirements

Ubuntu does not impose hardware requirements beyond the requirements of the Linux kernel and the GNU tool-sets. Therefore, any architecture or platform to which the Linux kernel, libc, **gcc**, etc. have been ported, and for which an Ubuntu port exists, can run Ubuntu.

Supported Architectures

Ubuntu 10.04 onwards supports three major architectures and several variations of each architecture known as "flavors". Three other architectures (HP PA-RISC, Intel ia64, and IBM/Motorola PowerPC) have unofficial ports.

Architecture	Ubuntu Designation	Sub-architecture	Flavor		
Intel x86-based	i386				
AMD64 & Intel EM64T	amd64				
HP PA-RISC	hppa	PA-RISC 1.1	32		
		PA-RISC 2.0	64		
Intel IA-64	ia64				
IBM/Motorola PowerPC sun4v	powerpc	PowerMac	pmac		
		Sun SPARC	sparc	sun4u	sparc64

INSTALLING UBUNTU

1. System must need a minimum of 2.5gb Hard disk space for installing Linux, 256MB ram.
2. This can be checked by Pressing F2 at startup of the system and checking the BIOS
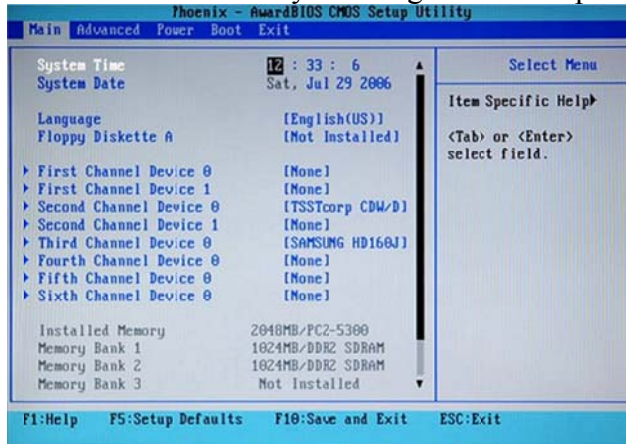
3. If u want have a dual boot(ex: Windows and Linux, Install windows first and then linux preferably in last hard drive/partition)
4. Set the boot parameters in the bios menu while starting your computer.(normally F2)
5. Go to boot devices priority and change priority as First priority as CD/DVD drive and later 2nd priority as Hard disk 3rd as Usb and so on.
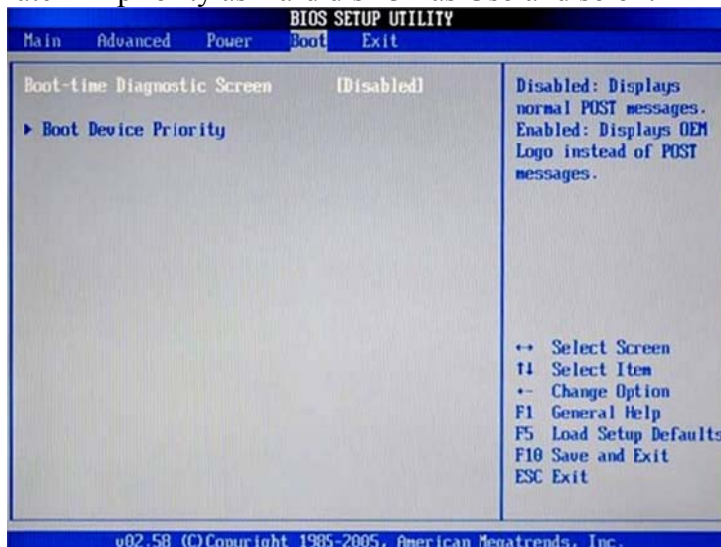
6. Exit saving Changes
7. Insert UBUNTU cd in your CDROM and wait until it boots and loads the necessary boot files.
8. At first UBUNTU will ask for certain information about Language, Place and date.
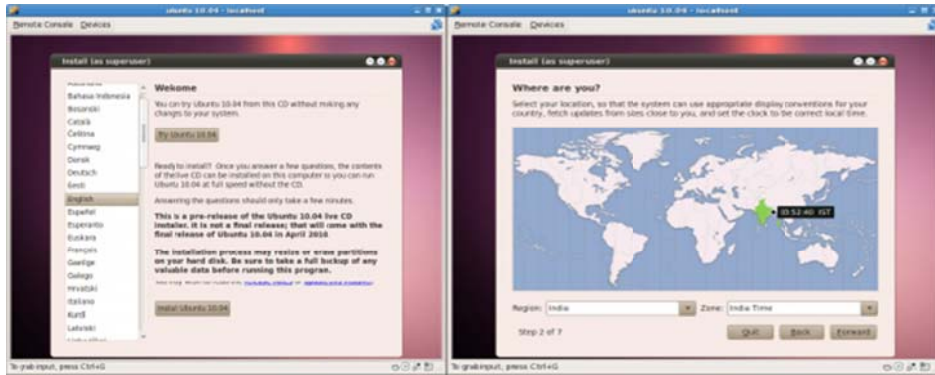
9. Enter the Keyboard layout in next step(default : US English)

10. Enter little personal information such as your name, User id and password in this step.

11. Now Ubuntu will return to the partitioning screen.
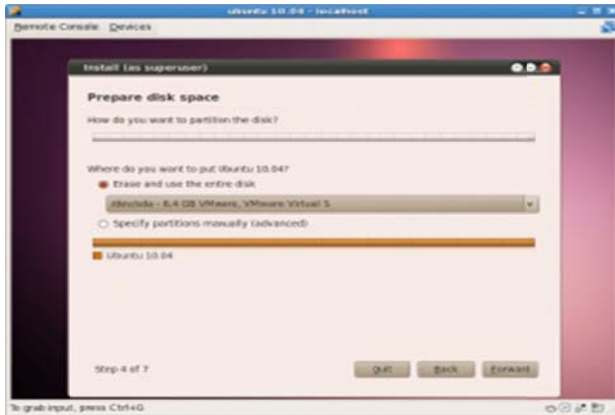
12. There are two options AUTO partition and Manual(Advanced) partition(Recommended).

13. Auto partition will take partition by itself with unfilled space. Manual Partition is for the user to partition based on his usage and requirement.

14. Let us consider installing in system having 100gb of space with 2gb Ram

15. In manual partition Root ("/") will be given some space(about 20 gb) Swap space will be allotted based on ram(in this example its 2gb ram, Hence allot about 6gb space)

16. Allot some space to Home folder "/home" as 50gb so that the files can be easily accessed (acts as sec. drive)

17. Allot rest to /usr/local so that any libraries can installed with ease(installing libraries in /usr/local will help in access of libraries by all the programs easy and access to other users easily)

18. Press next and installation starts with formatting the space that has been specified.

19. With this, you have successfully installed Linux and is a time for new world of Ubuntuing! Happy Time with Ubuntu Linux.
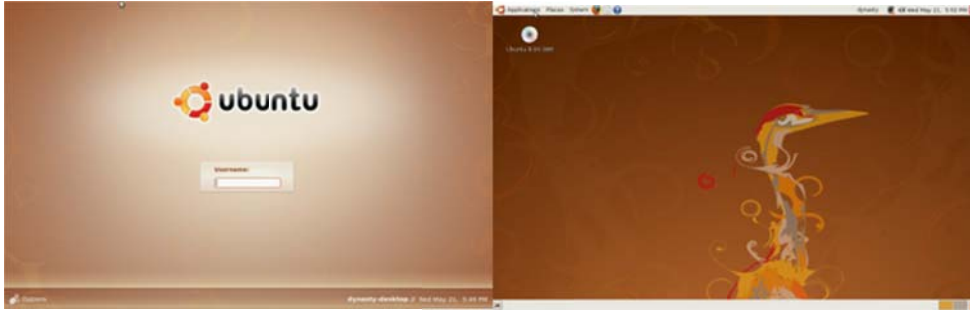
Change the root password at first use and maintain the password

This can be done by going to terminal and typing SU PASSWD and enter the new password twice

SUPPORT OPEN SOURCE SOFTWARES!

References

http://www.ubuntu.com

Hands-on Tutorial for GRASS GIS Beginners

GRASS GIS

Energy and Wetlands Research Group,
Centre for Ecological Sciences,
Indian Institute of Science
Bangalore – 560012

http://wgbis.ces.iisc.ernet.in/grass
http://wgbis.ces.iisc.ernet.in/foss

email: grass@ces.iisc.ernet.in, energy@ces.iisc.ernet.in
_____

GRASS Team at IISc

Dr. T. V. Ramachandra	cestvr@ces.iisc.ernet.in
Uttam Kumar	uttam@ces.iisc.ernet.in
S. N. Prasad	snarendra.prasad@gmail.com
N. V. Joshi	nvjoshi@ces.iisc.ernet.in
Bharath H. Aithal	bharath@ces.iisc.ernet.in
Anindita Dasgupta	anindita_dasgupta@ces.iisc.ernet.in
Setturu Bharath	settur@ces.iisc.ernet.in
Vishnu Bajpei	vishnu@ces.iisc.ernet.in

Useful websites for GRASS GIS references

1) http://grass.osgeo.org/wiki/Installation_Guide
2) http://grass.osgeo.org/grass64/manuals/html64_user/helptext.html
3) http://grass.osgeo.org/wiki/Importing_data
4) http://grass.osgeo.org/gdp/tutorials.php
5) http://ces.iisc.ernet.in/energy
6) http://wgbis.ces.iisc.ernet.in/energy/paper/researchpaper.html#rg

Log on to http://grass.osgeo.org/
Click on the Download Tab

Click on GNU/Linux for free download of GRASS GIS

Click on to Download older GRASS versions

This is what you get.

Scroll down to the bottom of the page. And click on Ubuntu.

This Screen appears after clicking the button.

Click on karmic option.

This is the window what comes next.

Scroll down and click on the i386 shown below.

Then click to the link mirror.pn.gov/ubuntu/

Then save the file.

This appears on the screen when the download is complete and file is saved.

Then double click on grass 6.4.0-re5-2i386.deb. This will be downloaded in the /home/user/Downloads directory.
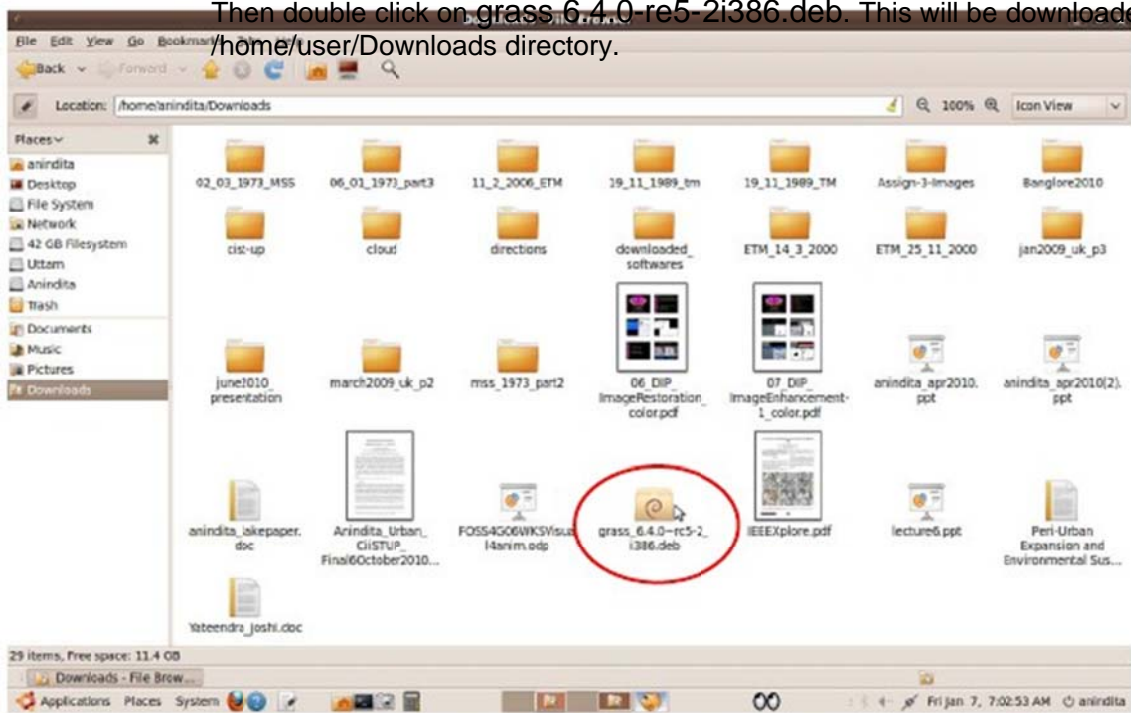
Click on Install Package.

Then installation starts.

After the installation is finished click on Close button.
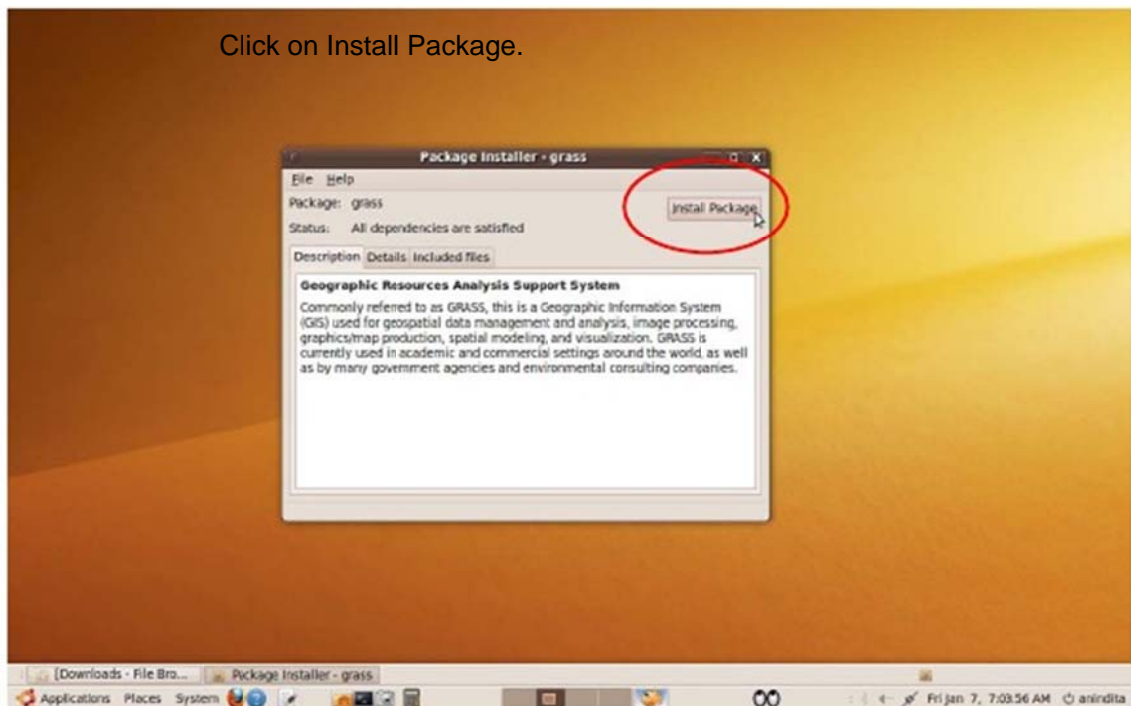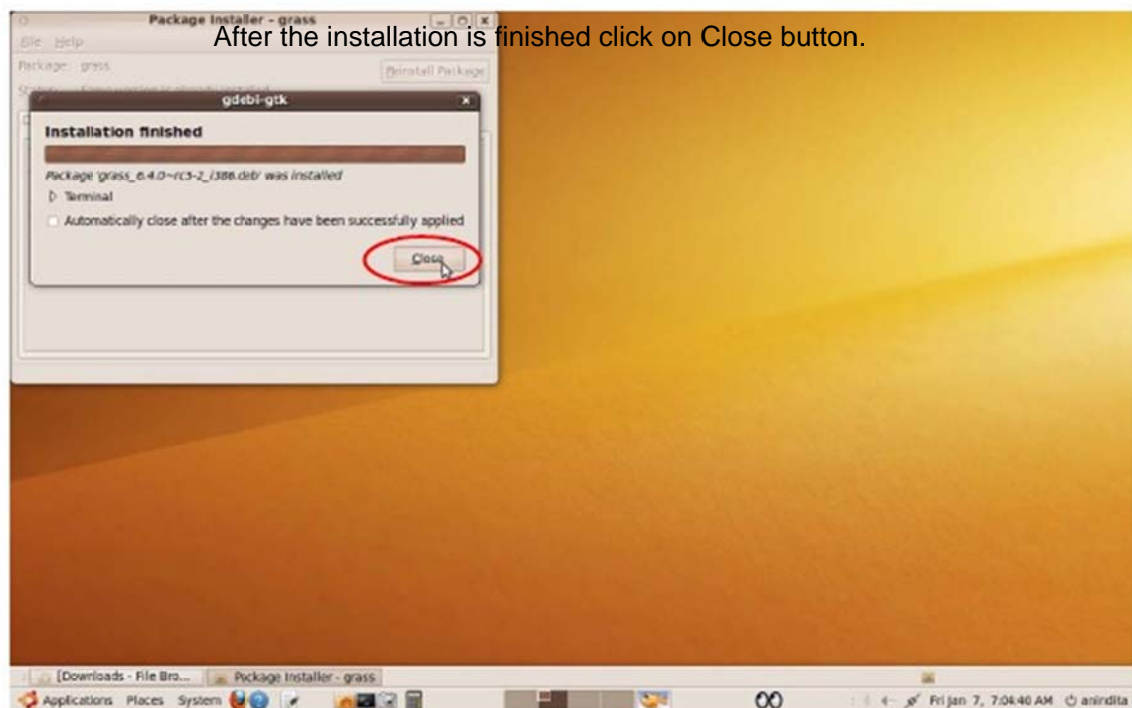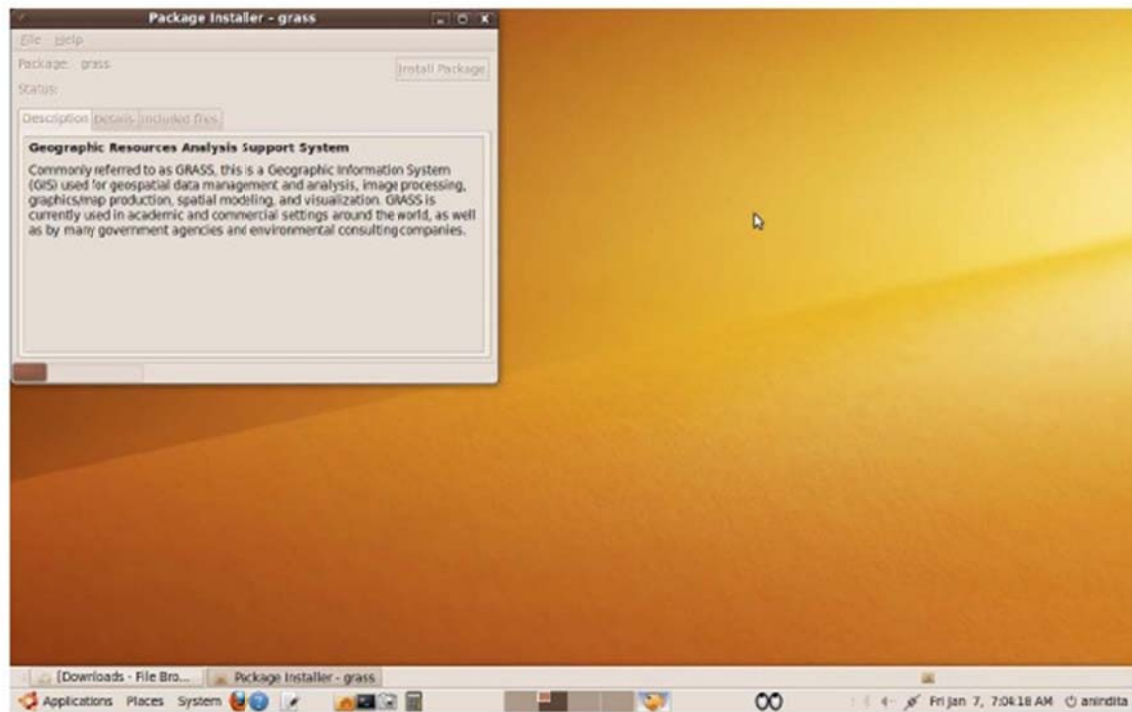
Congratulations. Installation is complete.

PART – II

STARTING GRASS 64

Click on the terminal

Terminal opens.
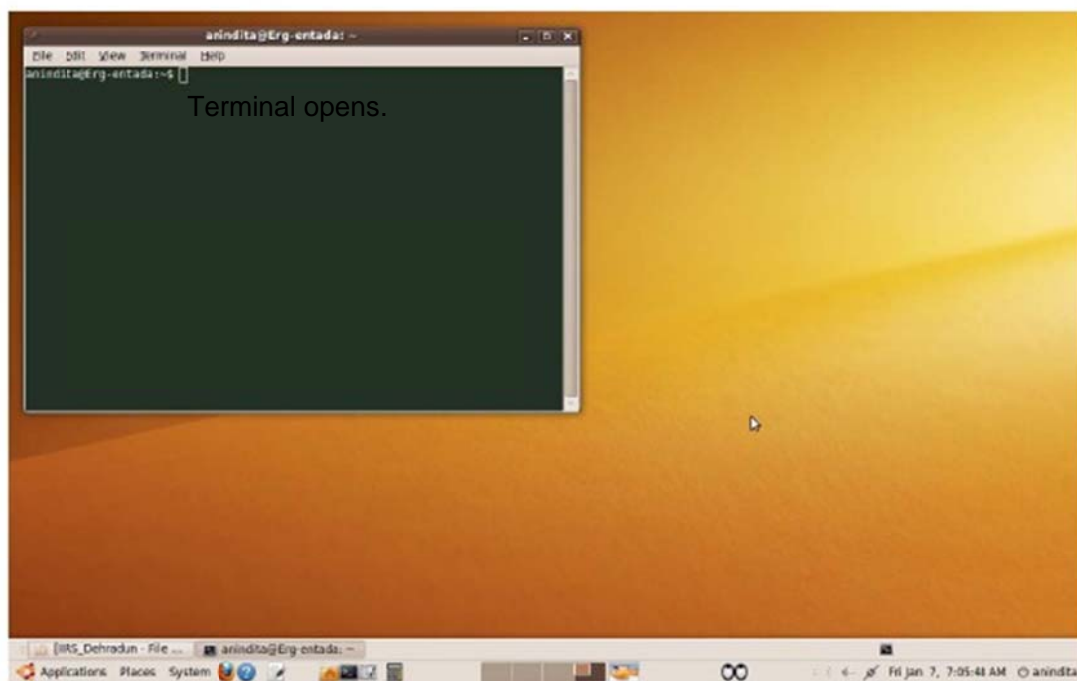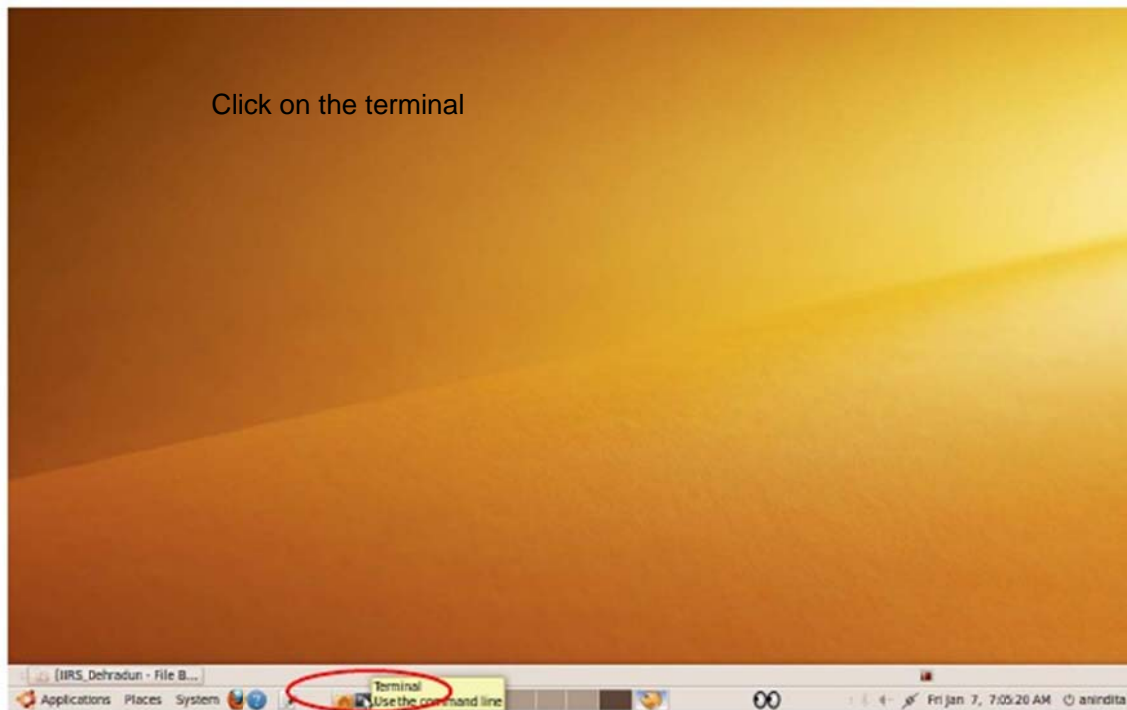
Type grass64 or type grass and press Tab button on you key
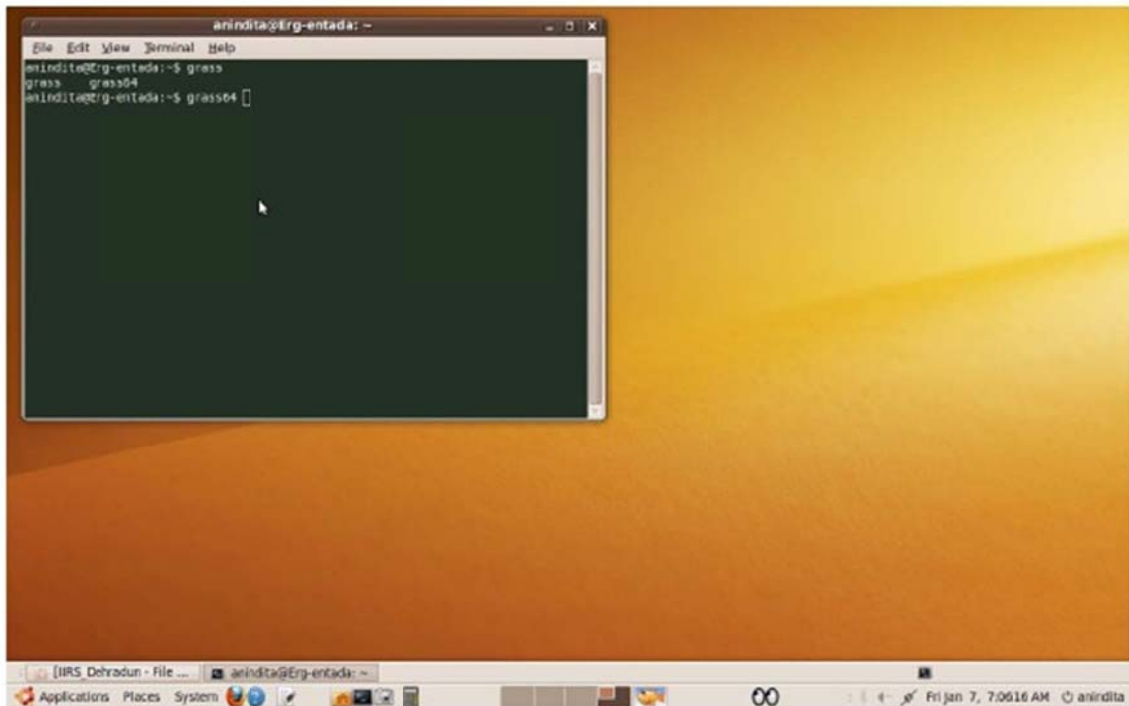
Grass GIS version 64 is ready for operation.
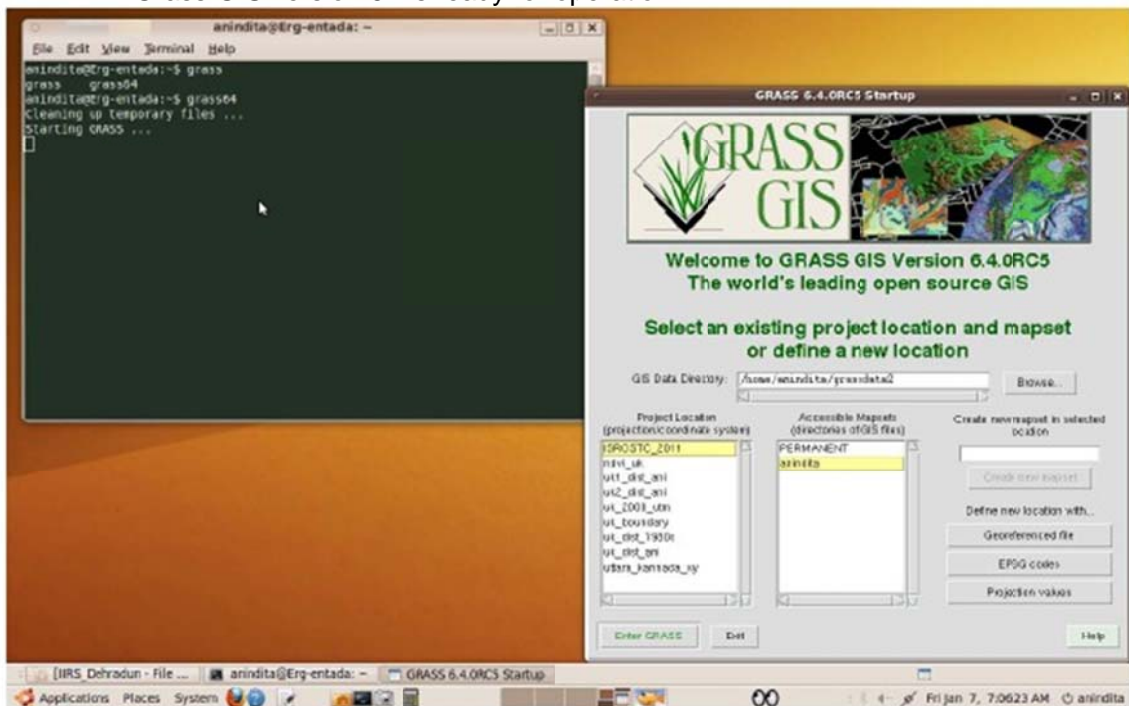
PART – III

DOWNLOADING GRASS DATASETS

To download default data from Grass GIS Website: goto http://grass.osgeo.org/
Click on Download Tab

Click on the Sample datasets. This is Spearfish, South Dakota, USA data set.

Click on Spearfish

Click on Full tar gz (135MB). Additional information is available at Data set description and Quick usage tutorial.

Save file.

PART – IV
START GRASS 64

These are the downloaded GRASS data sets.

Extract here.

Start grass64 by typing grass 64 on terminal
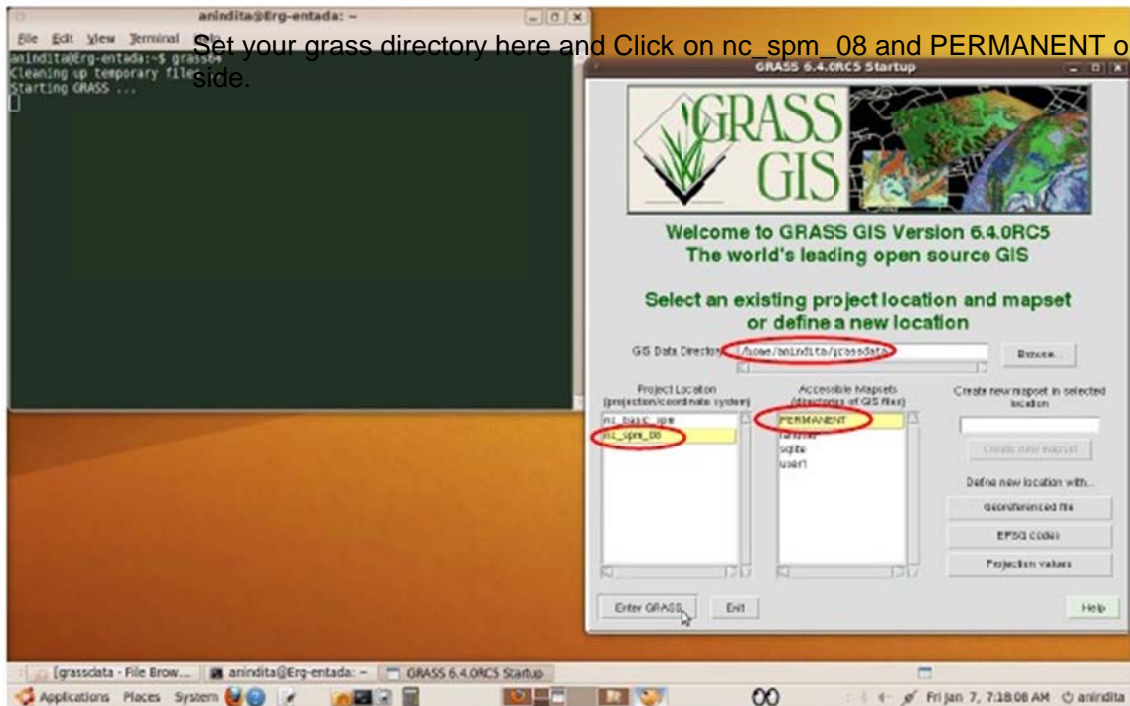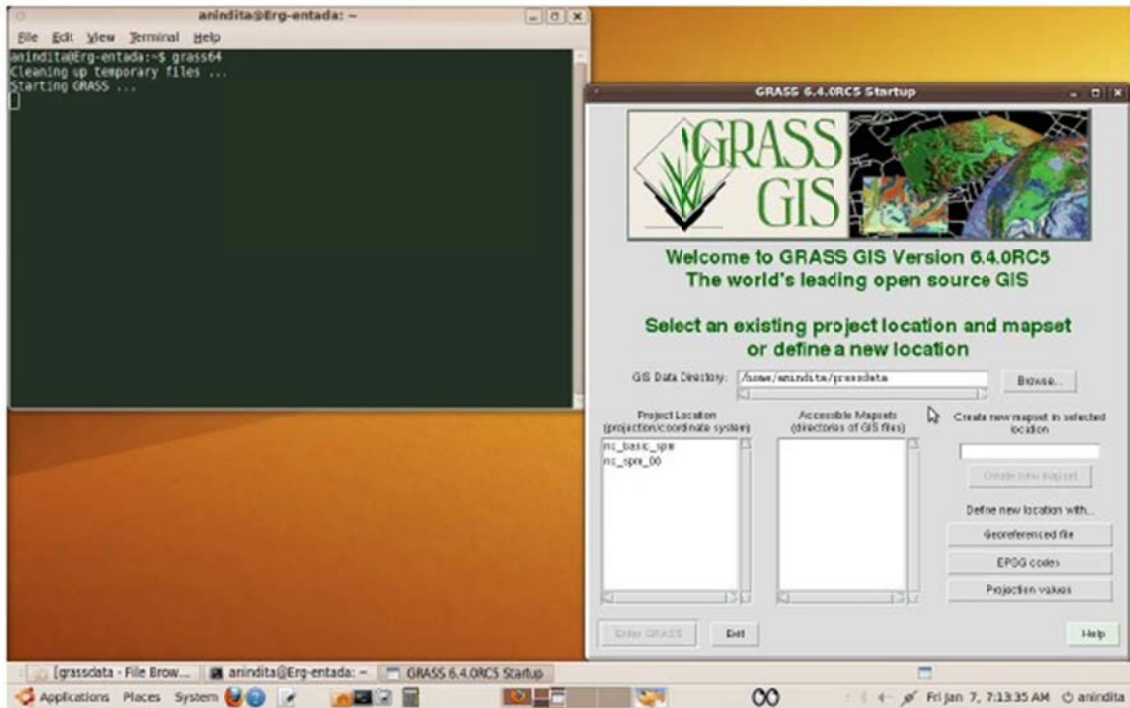
Set your grass directory here and Click on nc_spm_08 and PERMANENT on right side.

This is what appears on screen as Grass GIS display is opened.
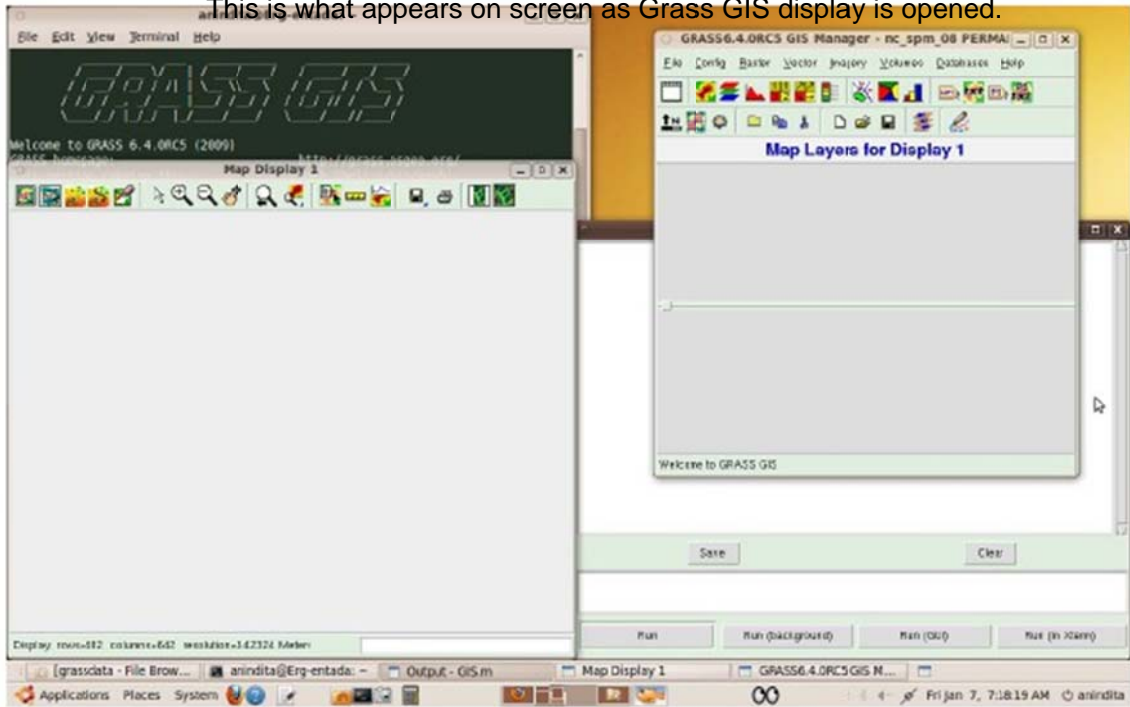
PART – V
WORKING WITH RASTER DATA

Click on add Raster layer.
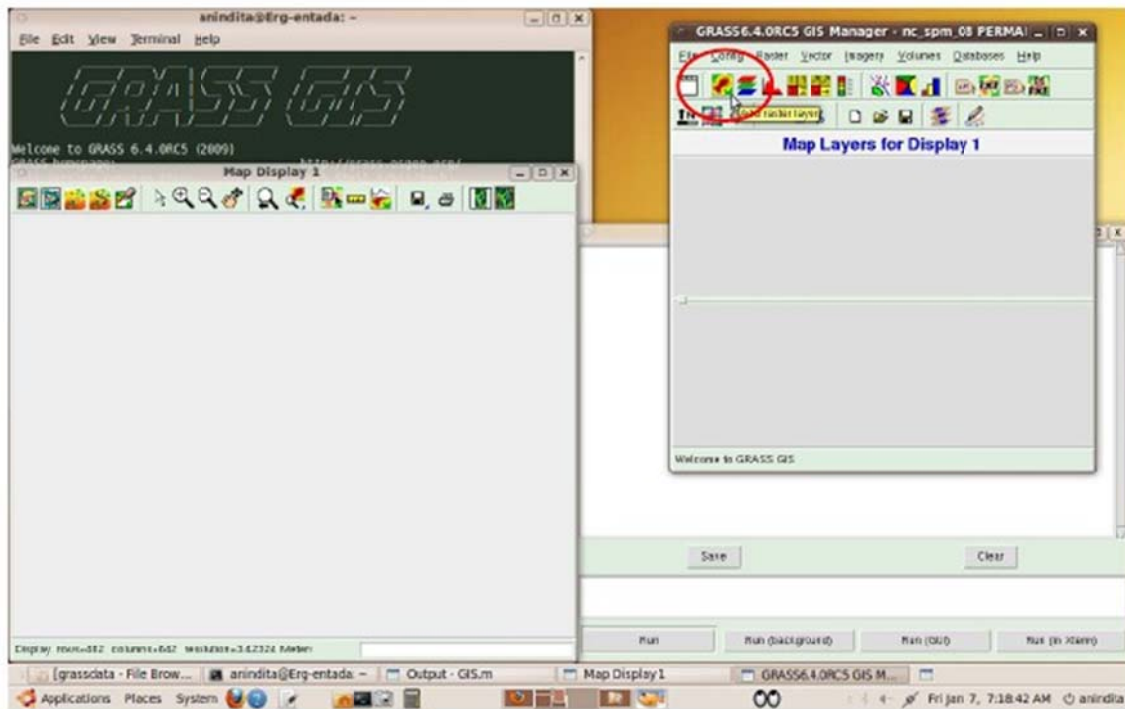
Click on raster map to display.

This is what comes up next.
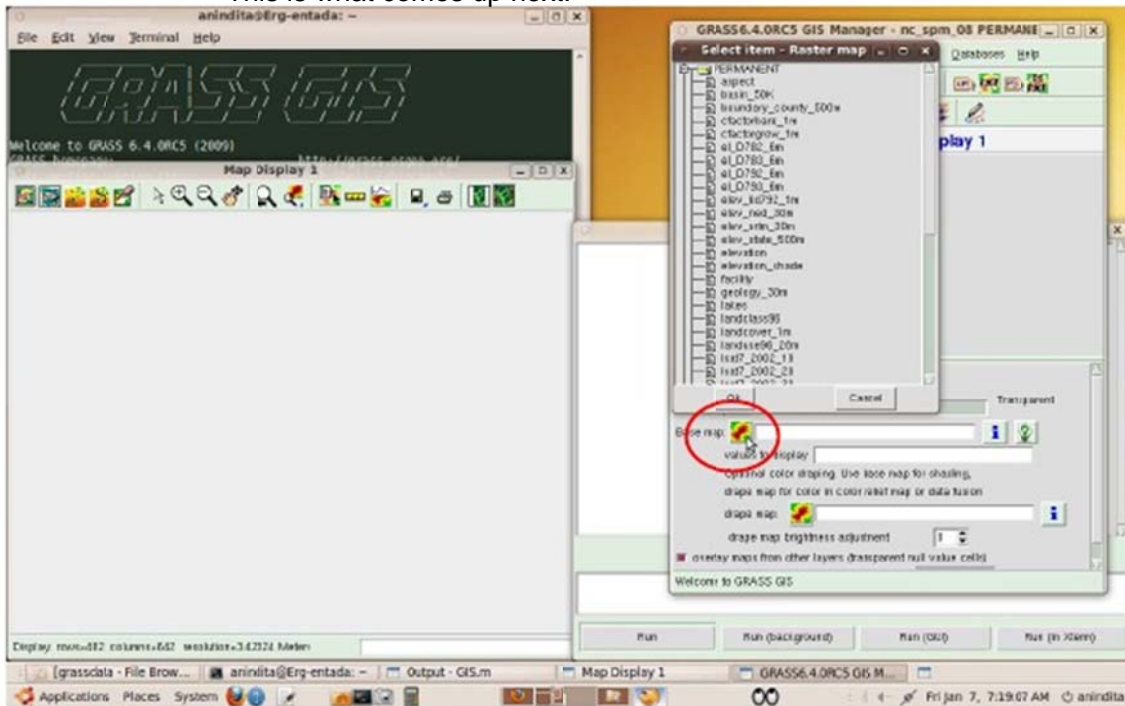
Go to aspect of Select item-Raster map and click Ok.
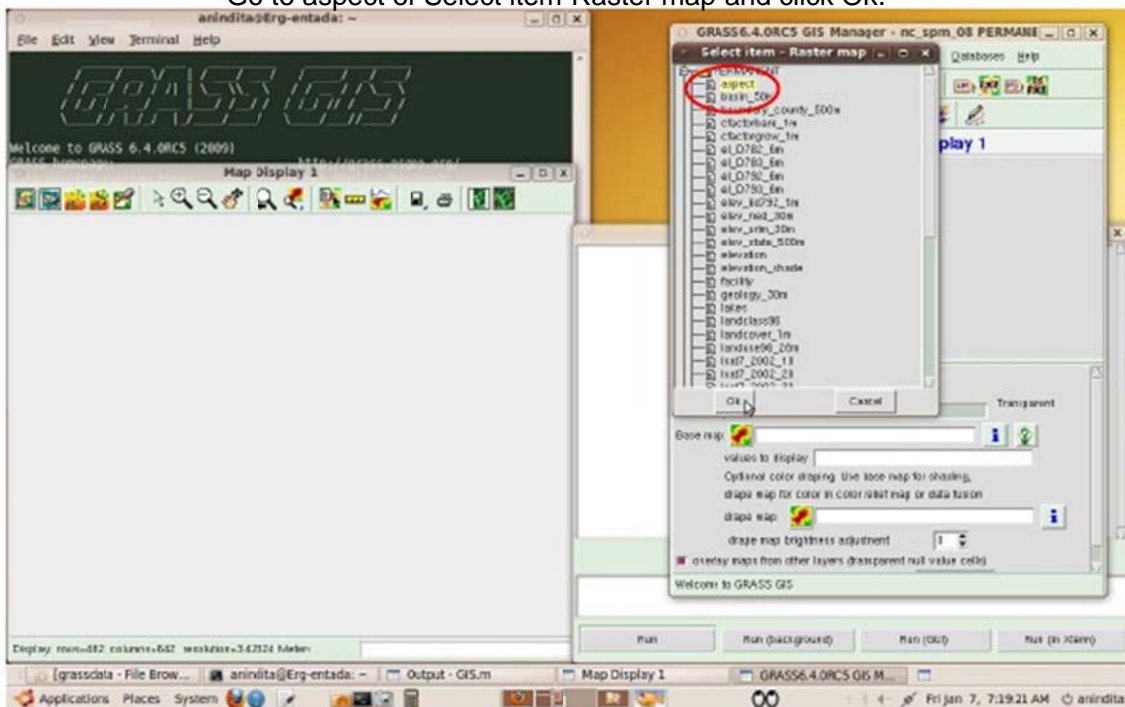
Click on the map display.
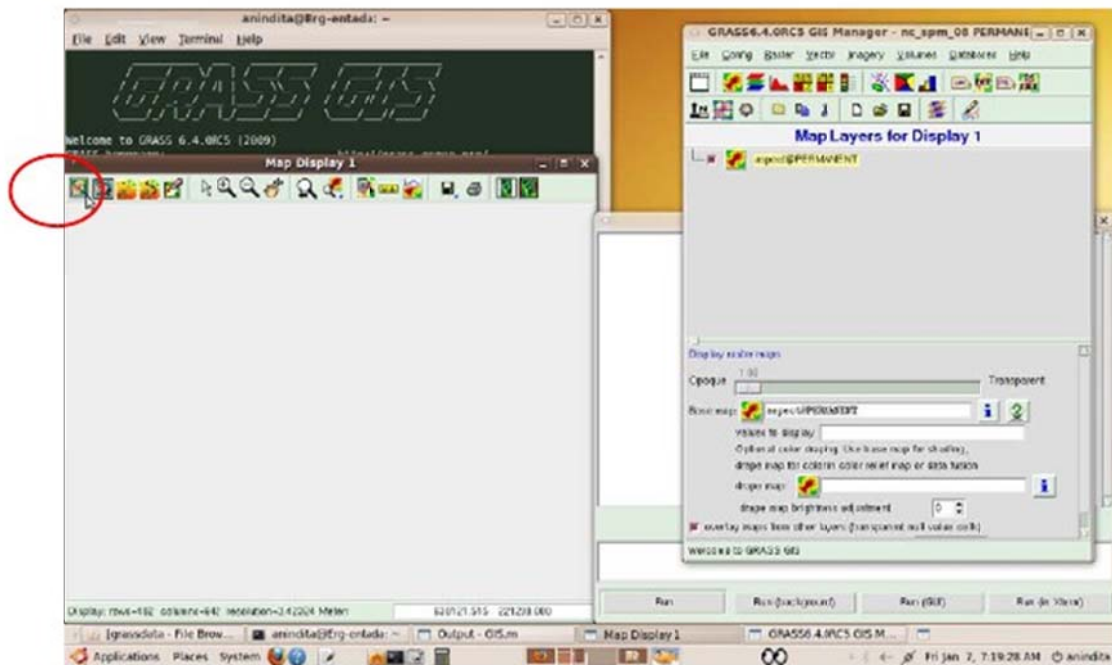
The image will be shown in the map display. The other utilities – zoom in, zoom out, save display (in jpeg, tiff), pan, measurement of length segment, printing are in the top menu of the displayed image.
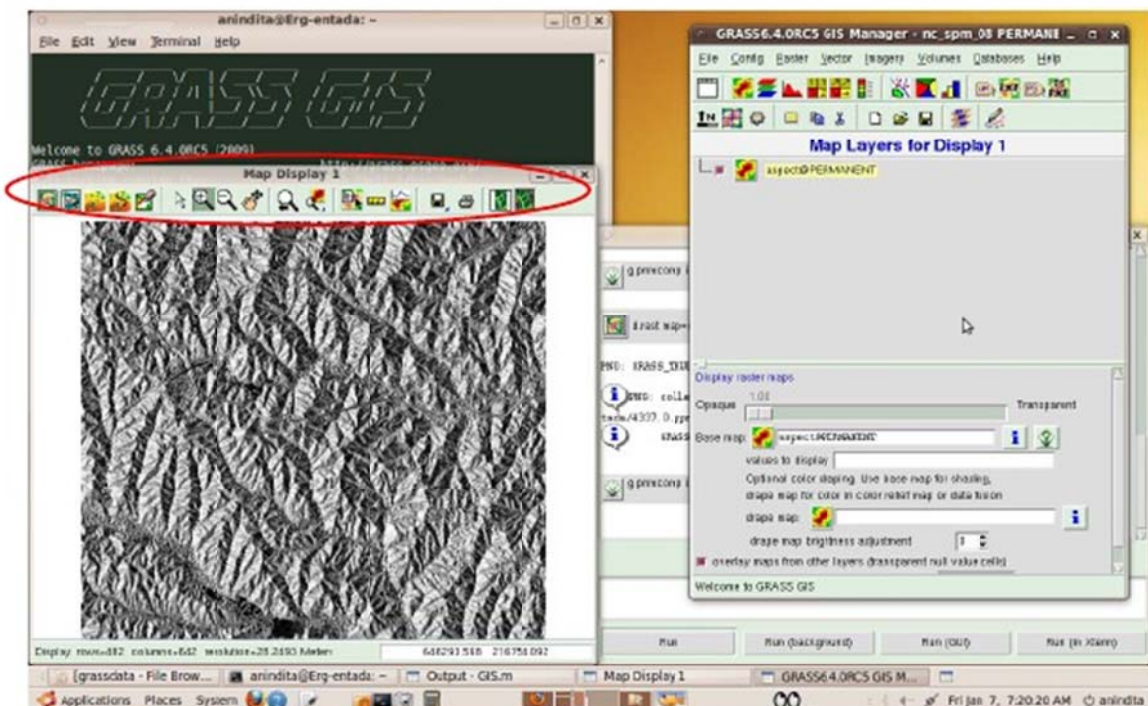
Now we will display a classified land use map. Display Landclass96 rastermap. Then to add legend, add raster legend layer to get the legend as shown below.
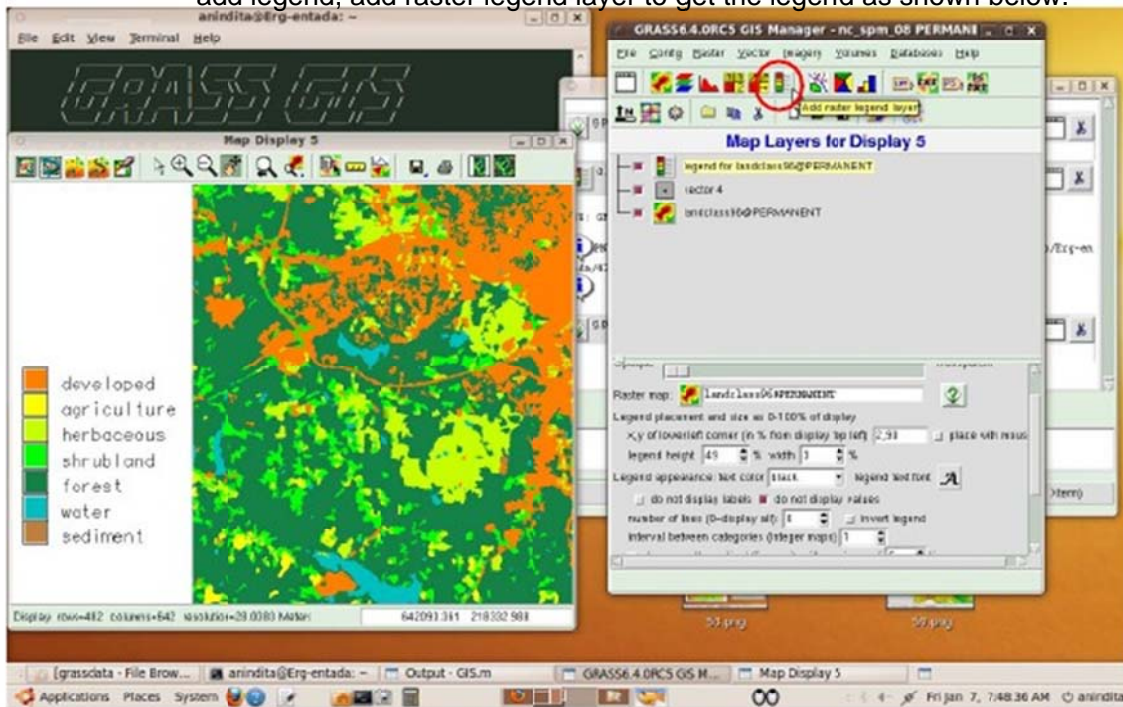
PART – VI

WORKING WITH 3D DATA

Click on the elevation map to see it on the display screen.

Run g.region for region setting.
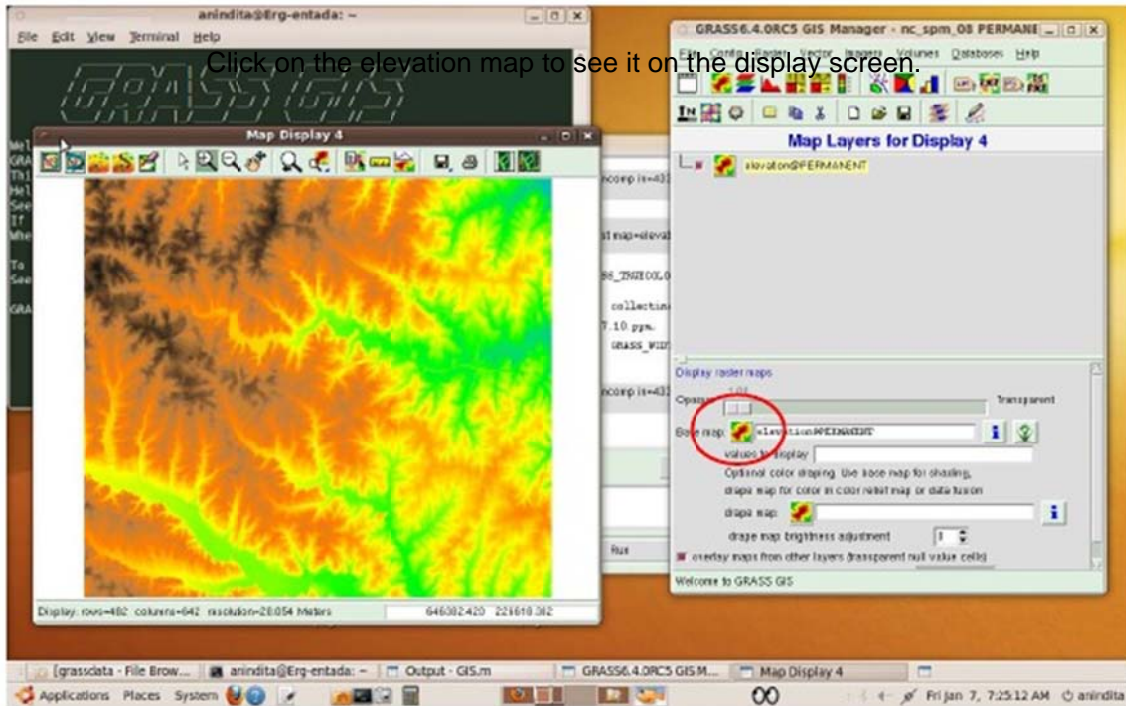
The region has been set.
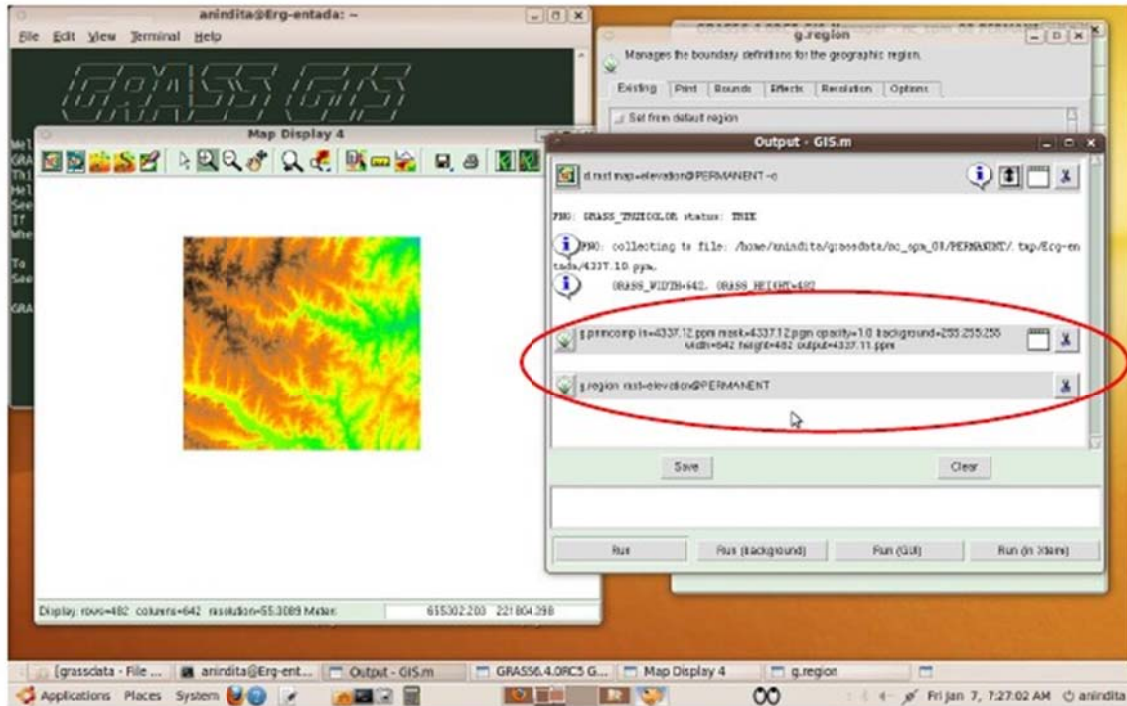
Click here on NVIZ.
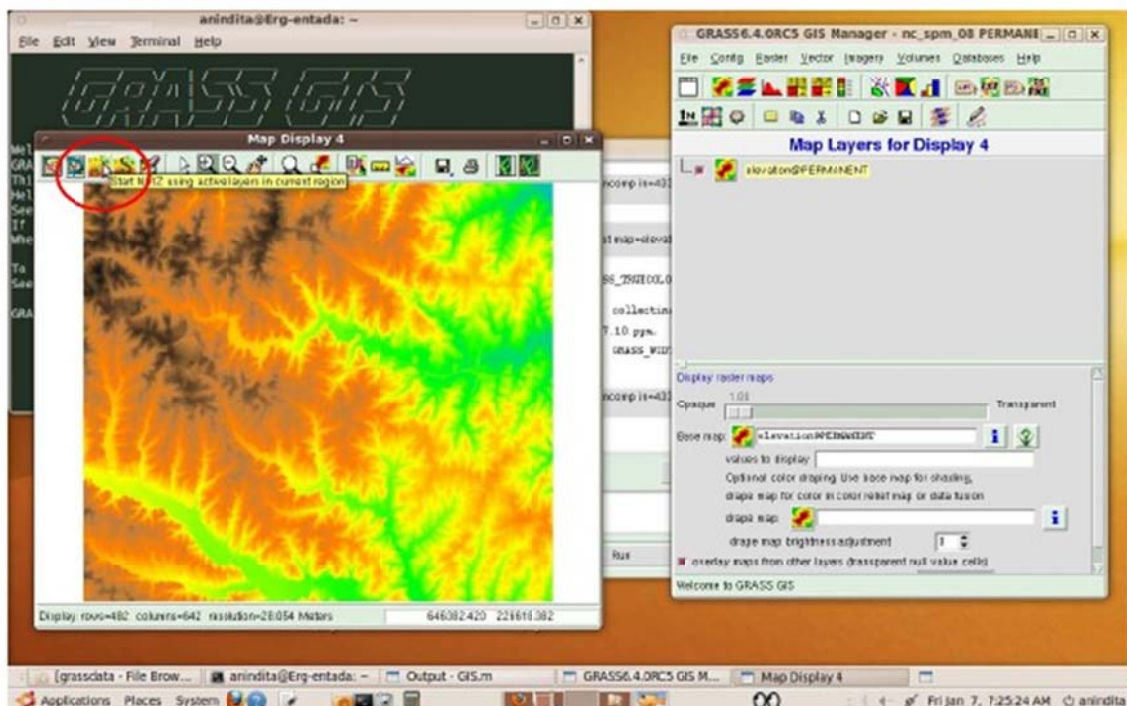
After clicking on NVIZ this is what appears.
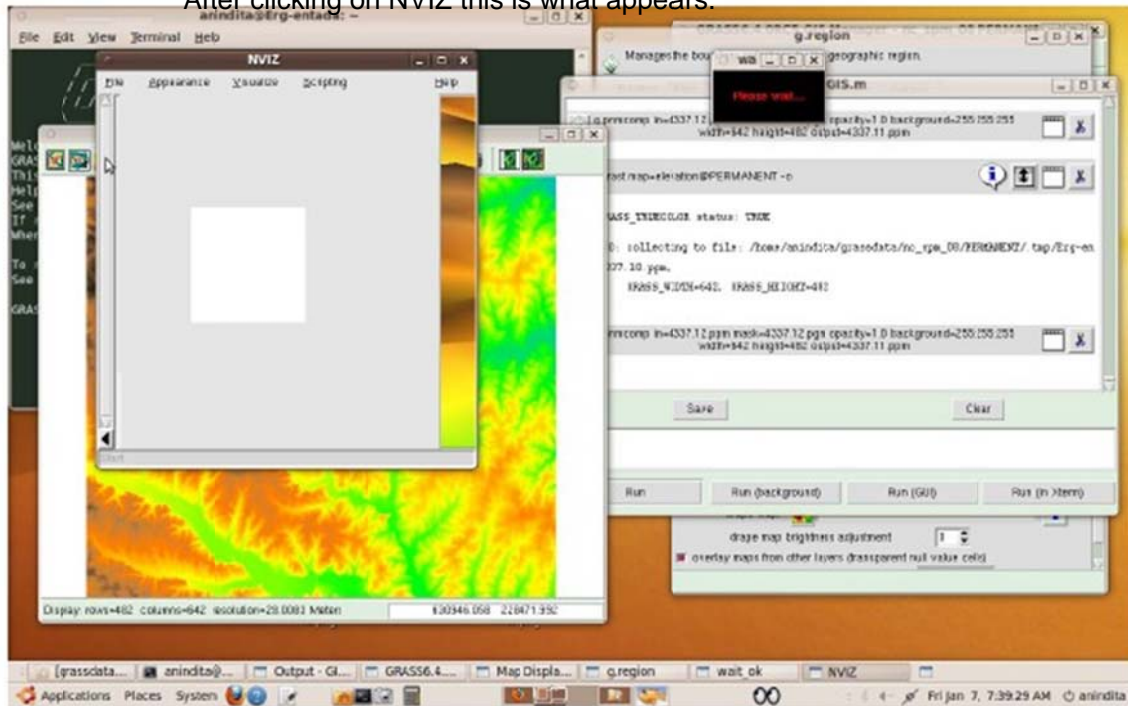
This is what we get as output.
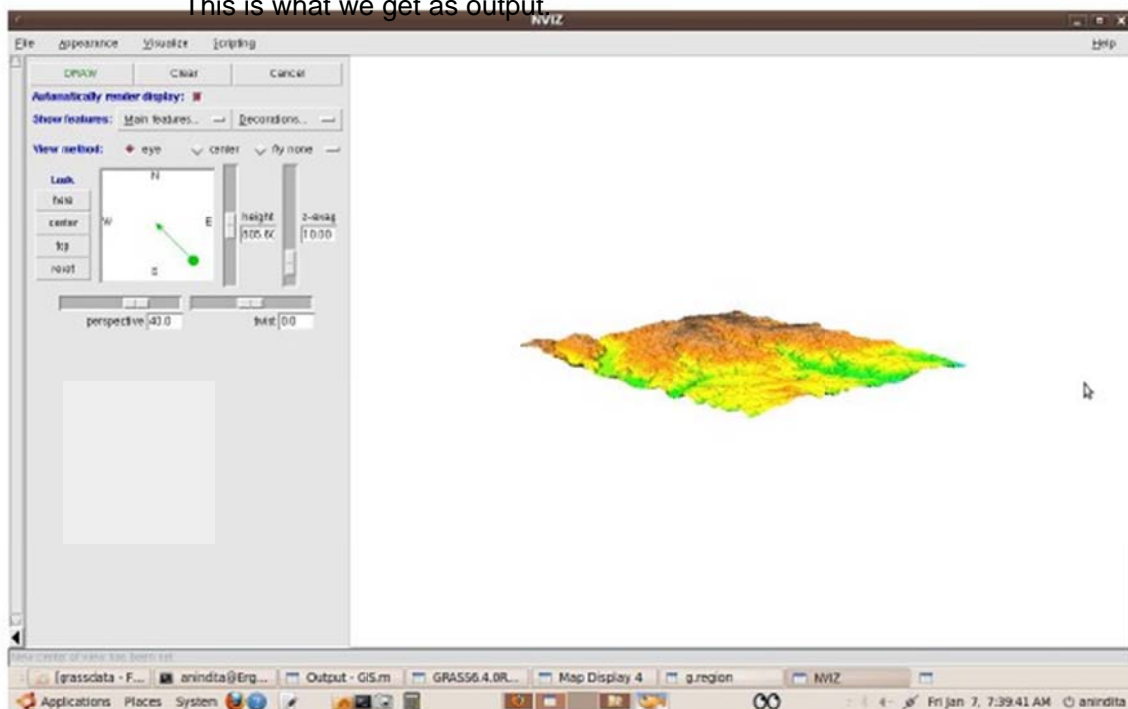
Click on draw and change the directions, exaggeration as you wish.
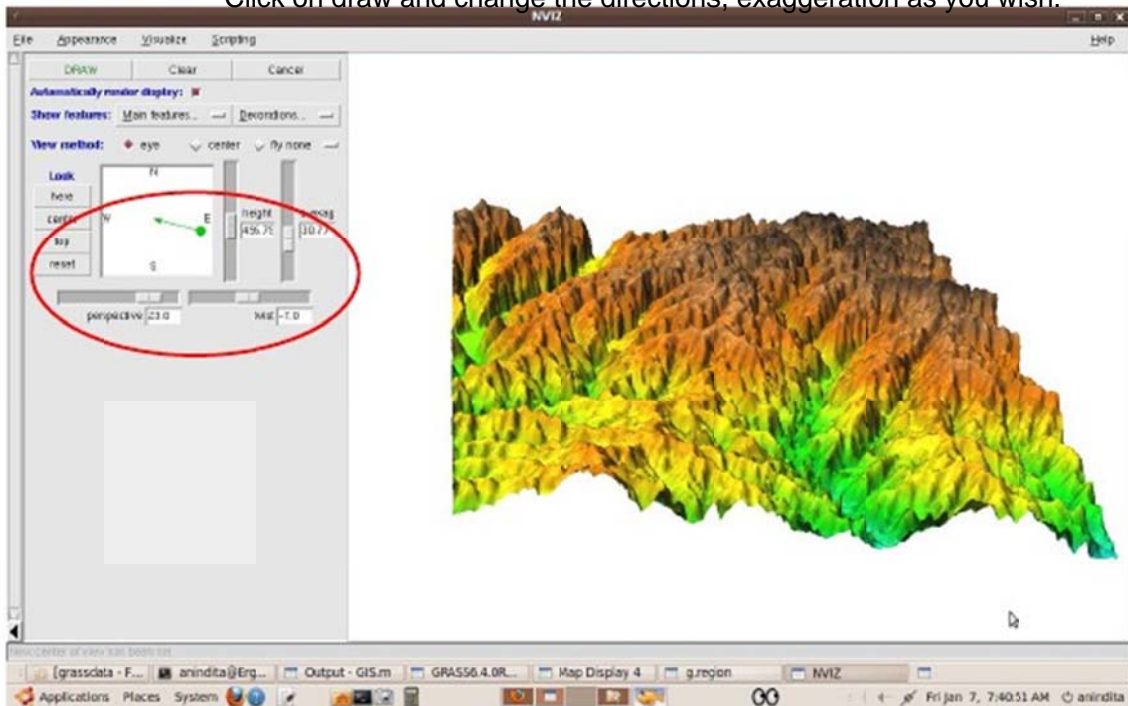
There is option to decrease the light intensity. Here you can see the effect of lighting from different angles.
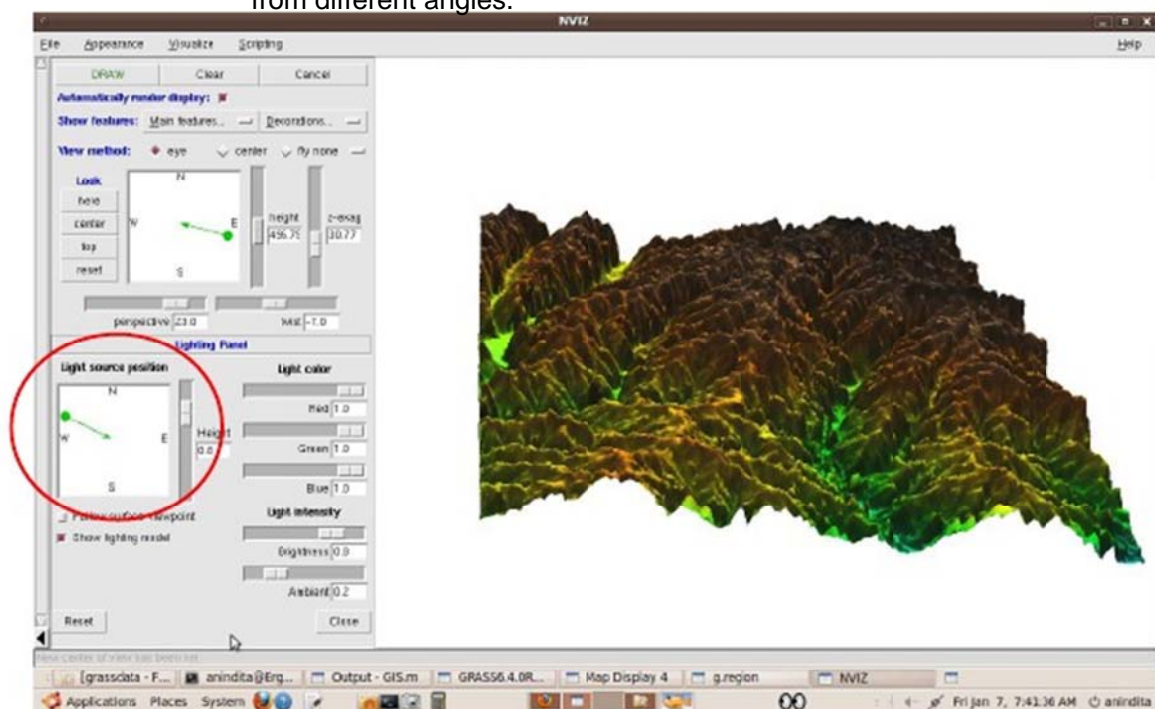
Increasing light intensity from front.
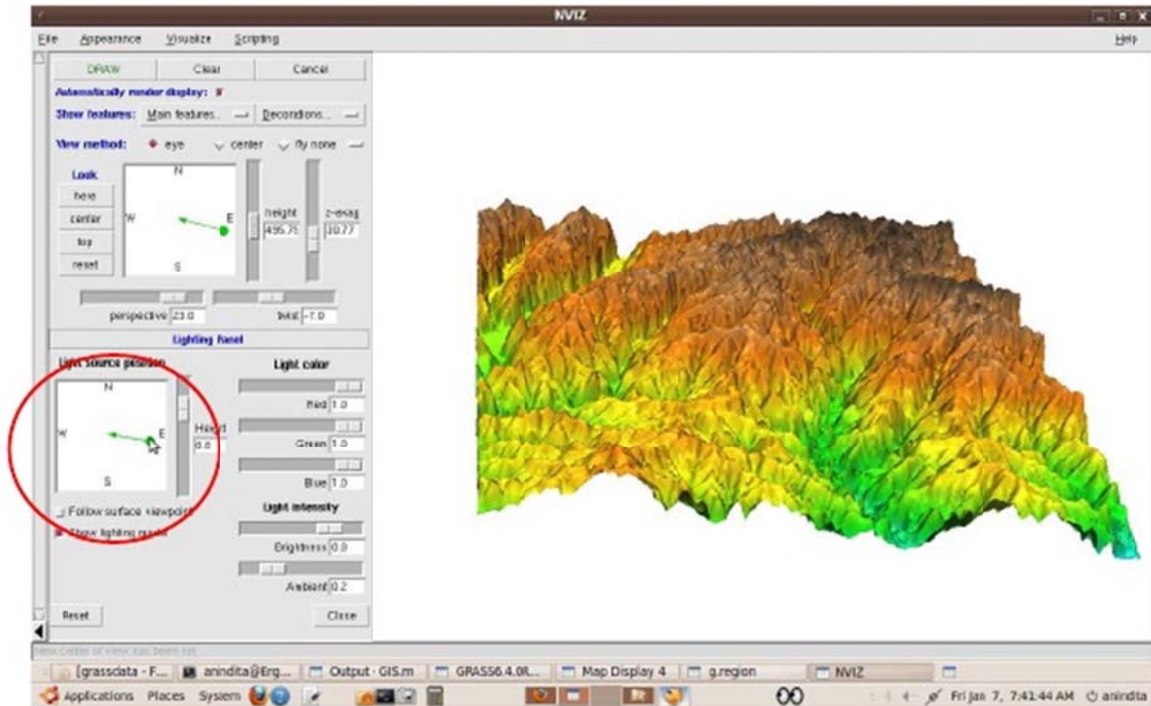
We can add vector layer from the map browser. Here water and road layer is added.
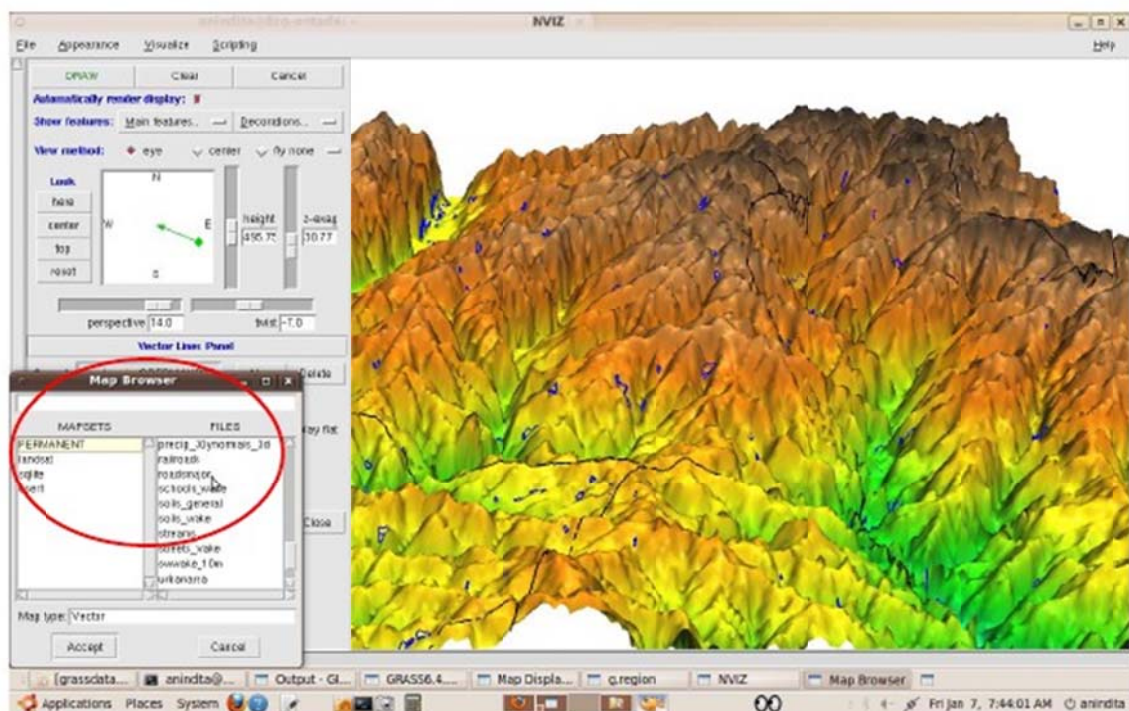
And change the direction to see a better view.

WORKING WITH VECTOR DATA

Open a window.

Click on add vector button.

Click on the vector map to display.

Select the vector file.
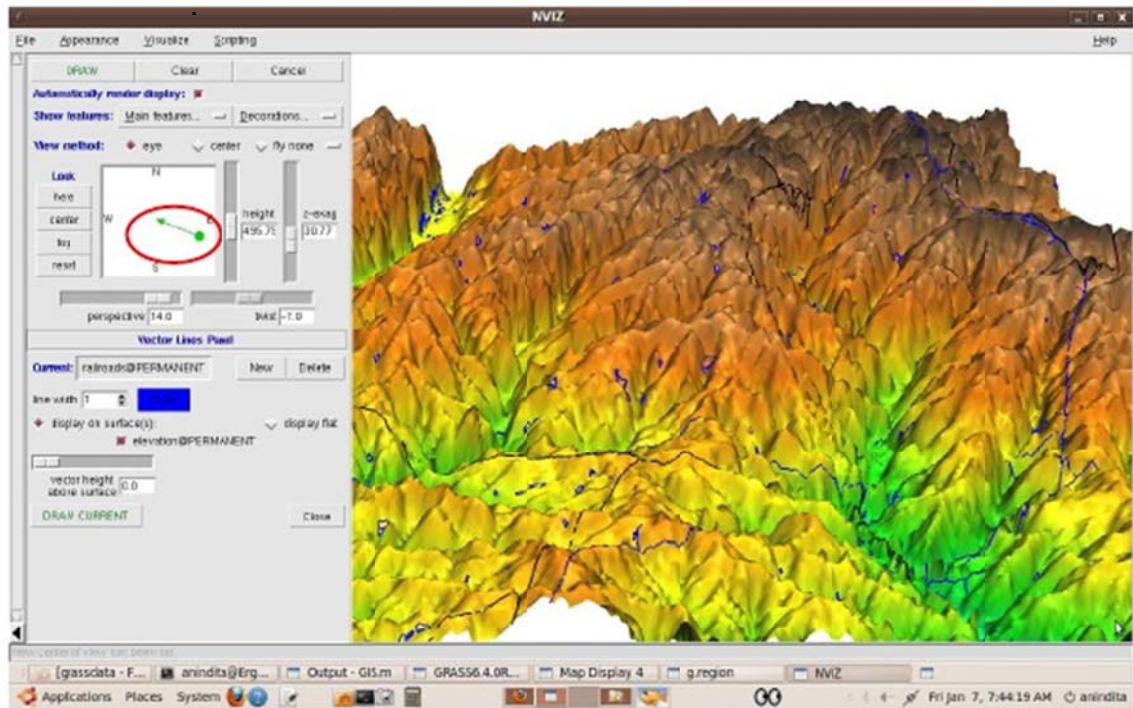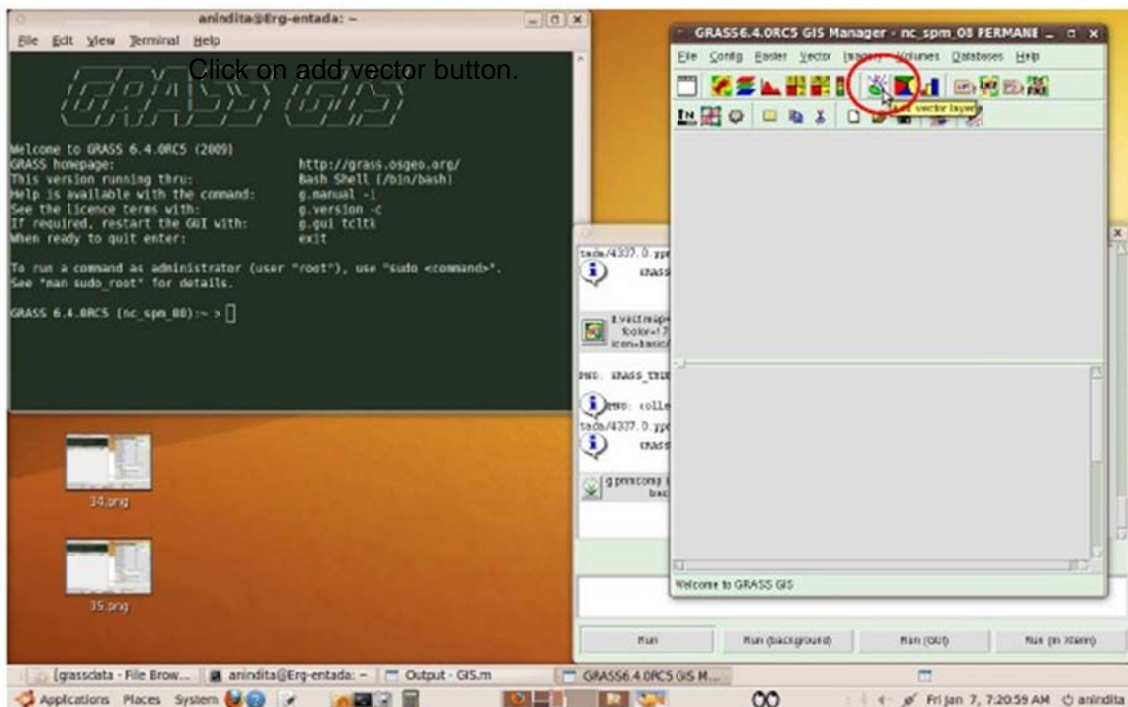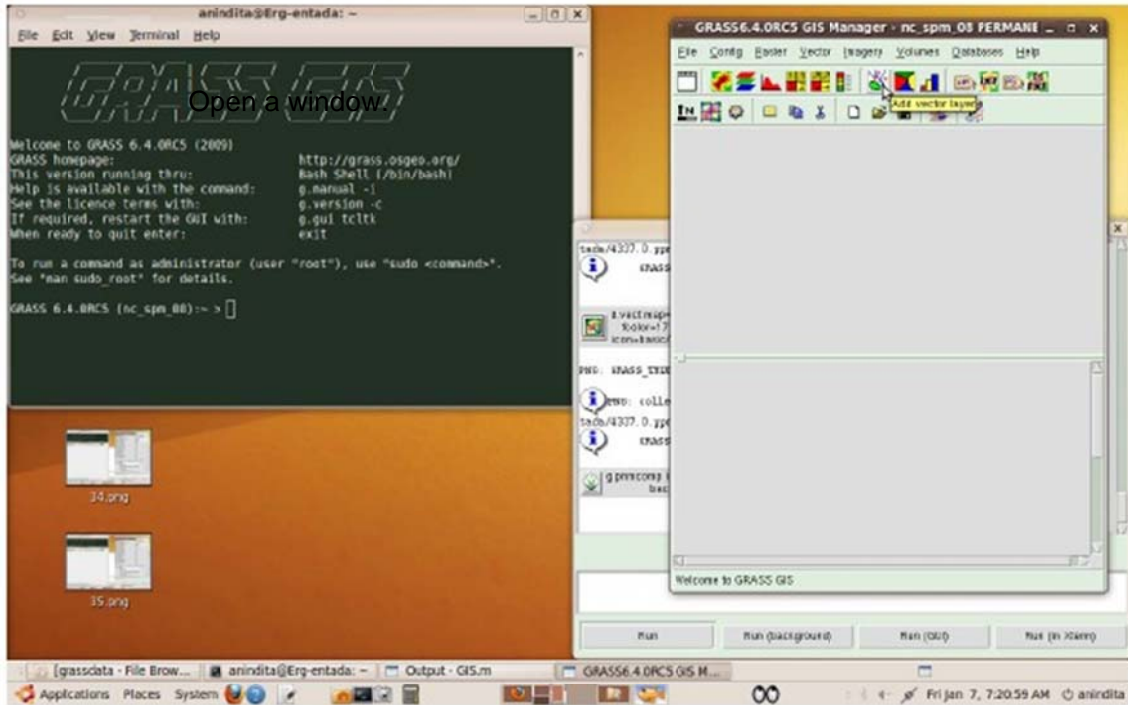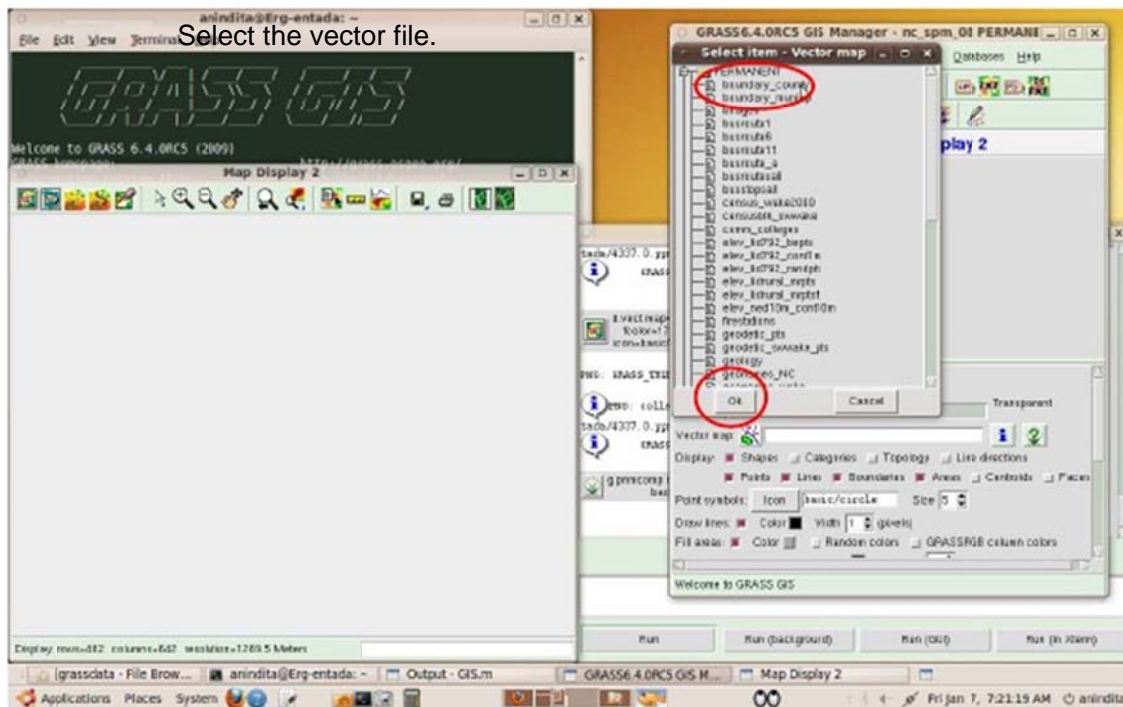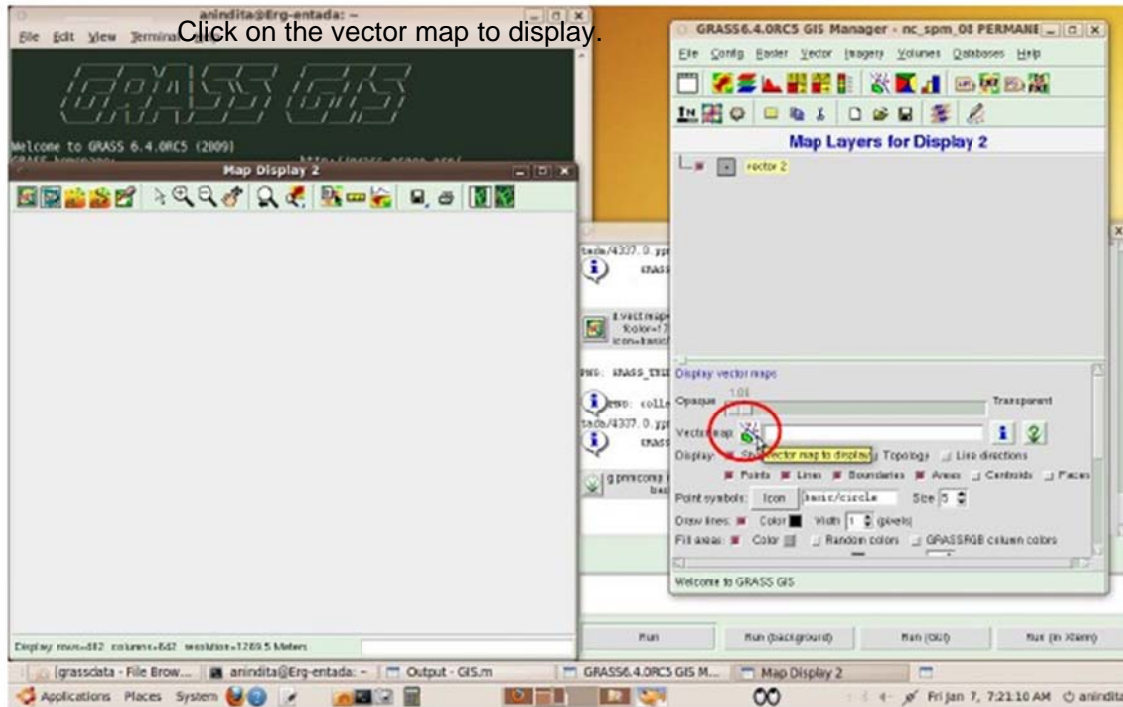
Click on display active layers.

We can see the vector map on the screen.
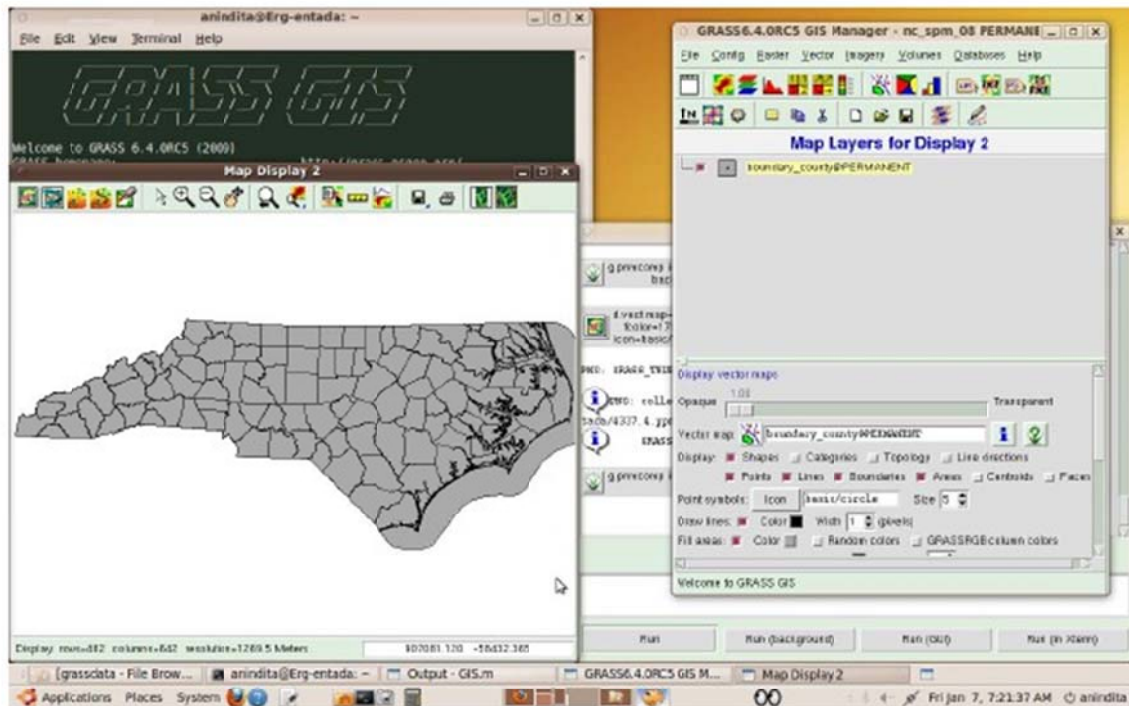
To query: Click on the Query button to obtain additional information from database.

To measure line segment length. Click on the Scale bar.

Draw line to measure the distance between one point to the other and see the length on the output file.
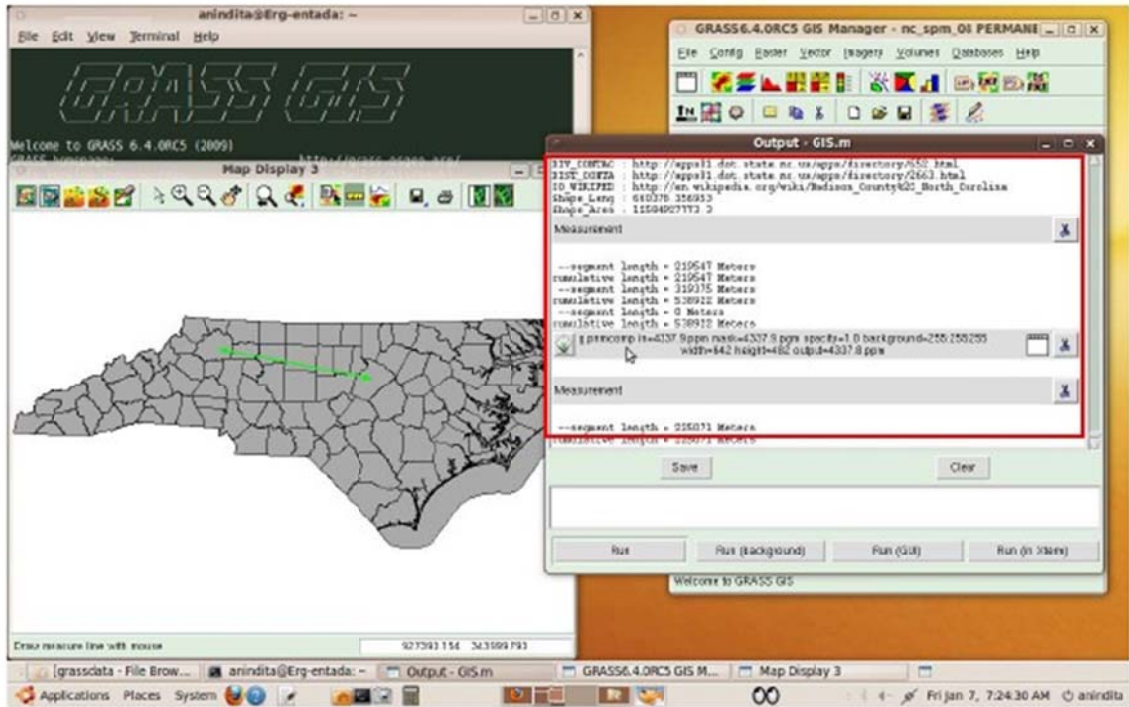
WORKING WITH OWN DATA: CREATING A NEW LOCATION

After having a practice on default data, now, we can start working with our own data for our study area. Before that, we need to create a Location and Mapset. Location can be created in

1) X, Y coordinate system.
2) Latitude, Longitude coordinate system.
3) UTM system.

We start with creating X, Y location.
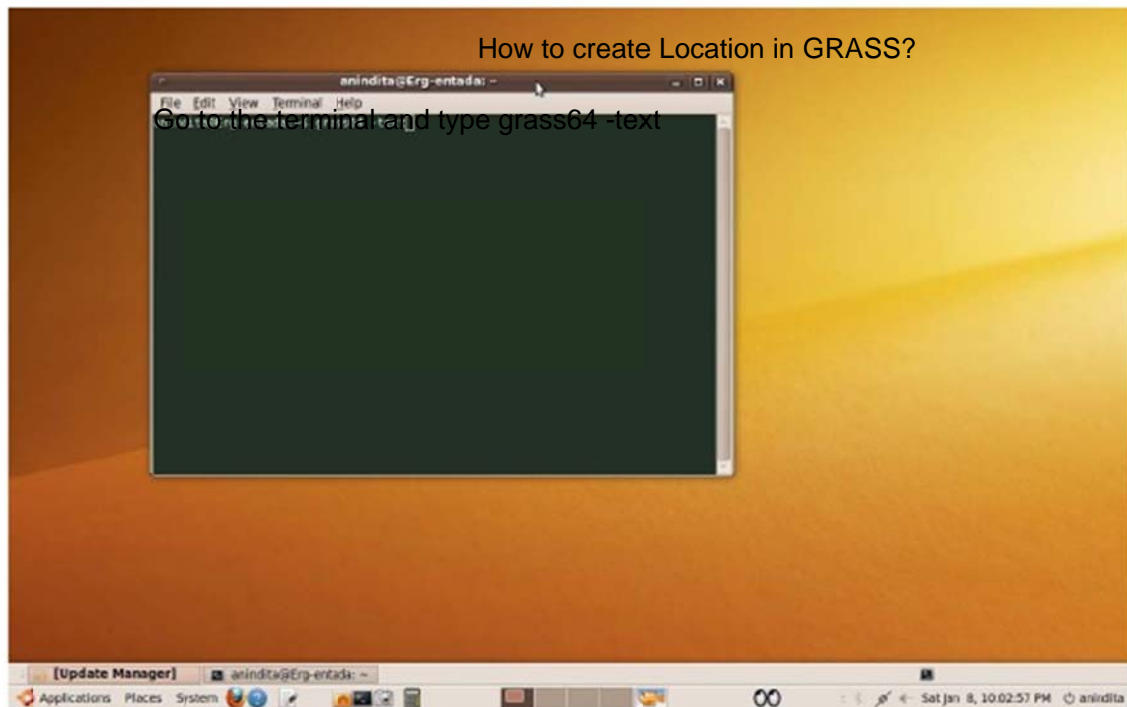
Here we have to enter the Location and Mapset name. We shall set LOCATION: Bangalore and MAPSET: username. We can create a folder (directory) to keep all the data. Let us name that folder as "grassdata". Then set the default database directory (DATABASE: /home/user/grassdata).

After pressing <ESC> and <ENTER> together we get this on screen.
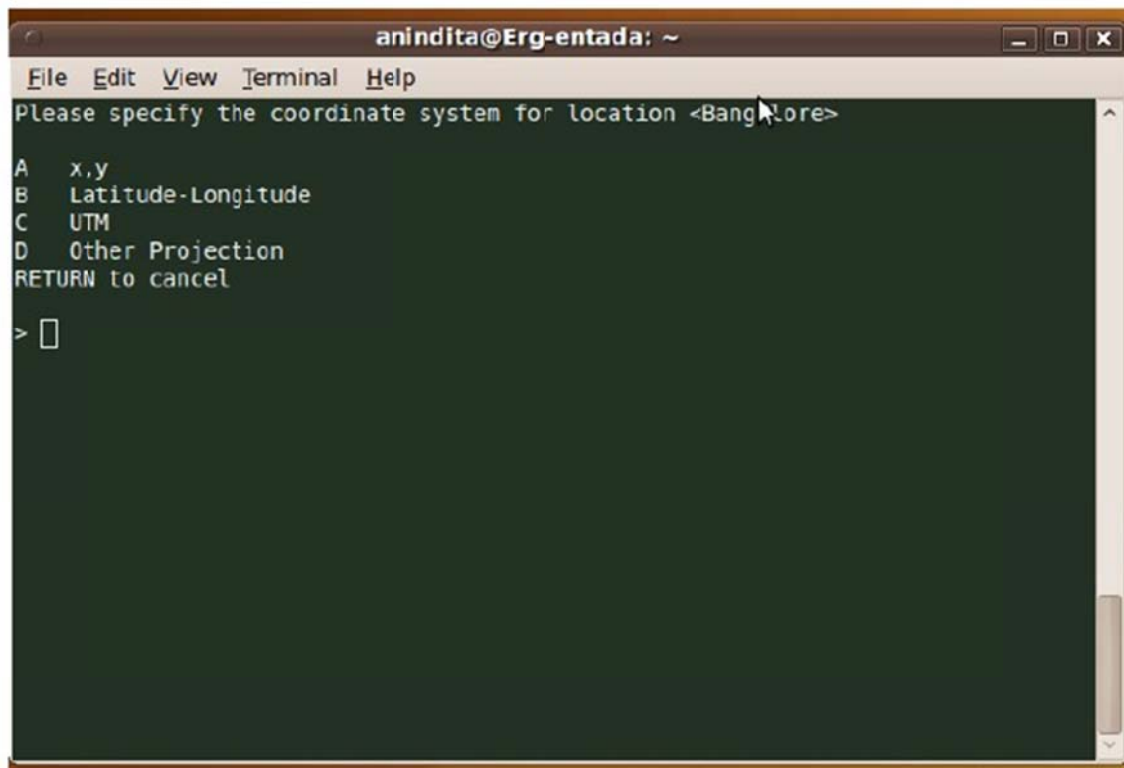
After typing y when we press enter, we get this.
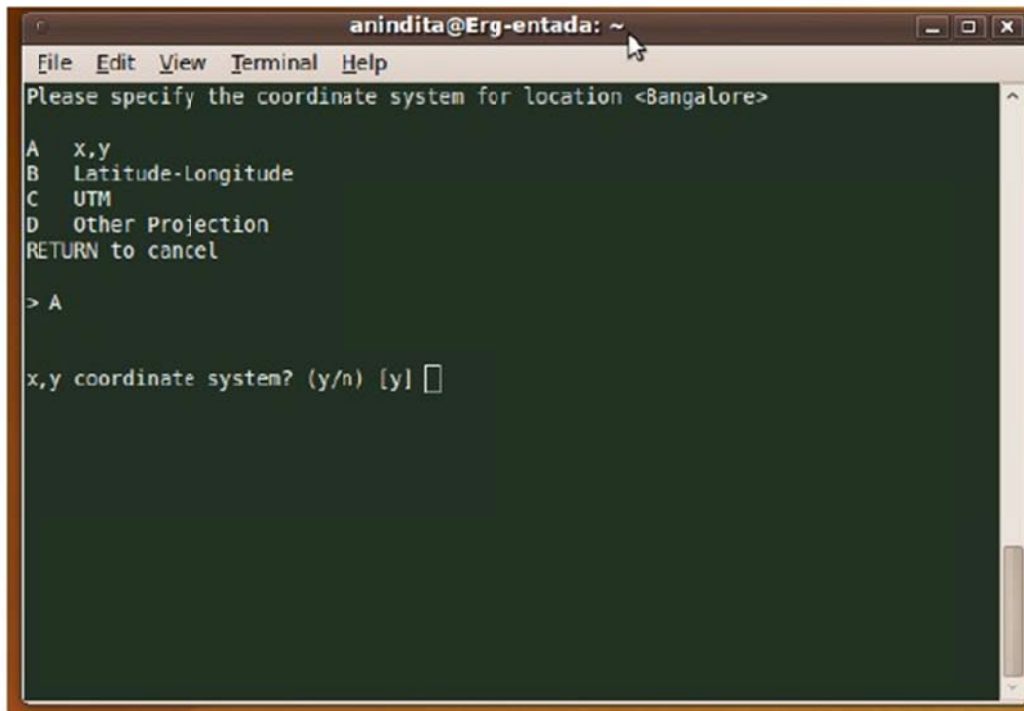
Again after typing y we get this screen.

To create x, y location, type A on the text box and press y.
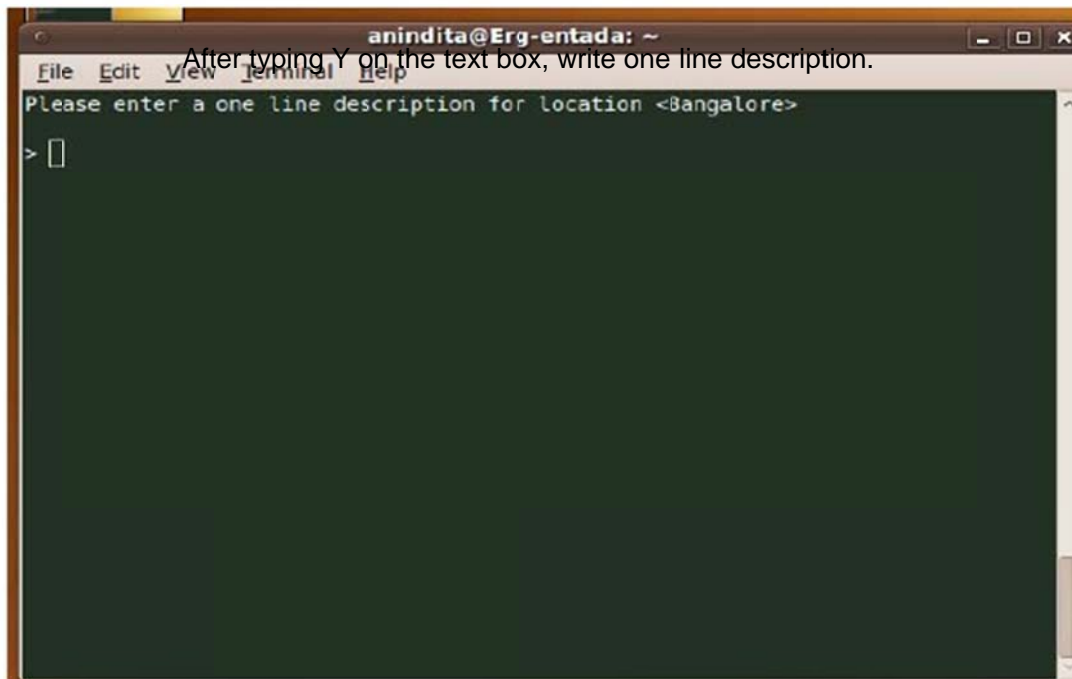

```
anindita@Erg-entada: ~
File  Edit  View  Terminal  Help
Please specify the coordinate system for location <Bangalore>

A    x,y
B    Latitude-Longitude
C    UTM
D    Other Projection
RETURN to cancel

> A


x,y coordinate system? (y/n) [y] []
```

After typing Y on the text box, write one line description.


```
anindita@Erg-entada: ~
File  Edit  View  Terminal  Help
Please enter a one line description for location <Bangalore>

> []
```

Type y after completing the description.

Press <ESC><ENTER> together to continue.

Type y to accept the region.

```
projection: 0 (x,y)
zone: 0
  north:      1
  south:      0
  east:       1
  west:       0

  e-w res:    1
  n-s res:    1

total rows: 1
total cols: 1
total cells: 1


Do you accept this region? (y/n) [y] > □
```


x, y location is created!!!! Press Enter and Press Ctrl + C to exit GRASS.

```
projection: 0 (x,y)
zone: 0
  north:      1
  south:      0
  east:       1
  west:       0

  e-w res:    1
  n-s res:    1

total rows: 1
total cols: 1
total cells: 1


Do you accept this region? (y/n) [y] > y
LOCATION <Bangalore> created!

Hit RETURN -->□
```

Now we will import Landsat data (ETM+ 8 bands) to GRASS. These data are downloaded from GLCF (Global Land Cover Facility). To import a file run r.in.gdal from terminal or goto File -> Import-> Raster-> Multiple formats using gdal.

Browse the downloaded file from Download directory or wherever you have kept the files.

Give the output file name and go to the menu options.

Click on override projection and Run to import the image.
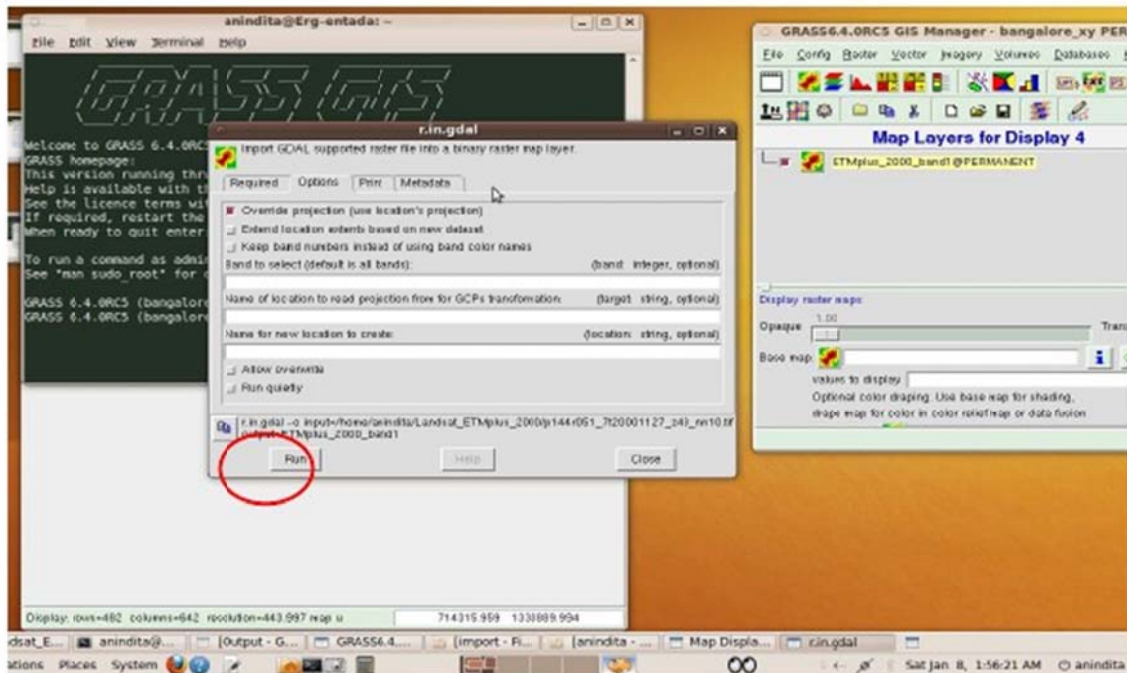
File is imported and we can see it on map display by clicking on it.
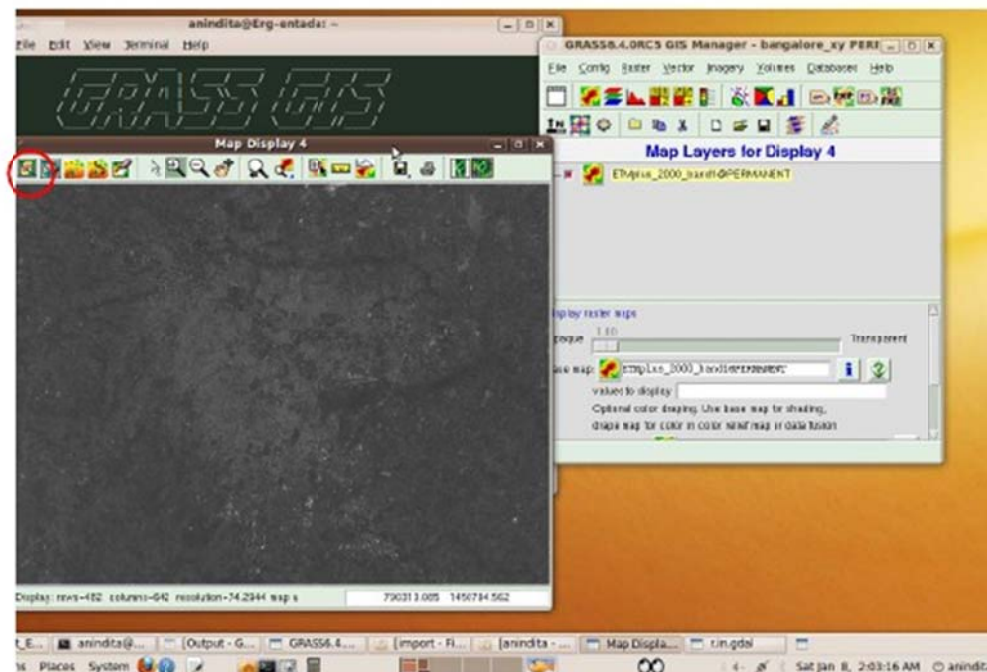
IMPORTING IRS 1C/1D/P6 LISS-III/IV Data INTO GRASS

Use the command r.in.gdal to import IRS data or go to File, Import raster map, multiple format using gdal from the selection.
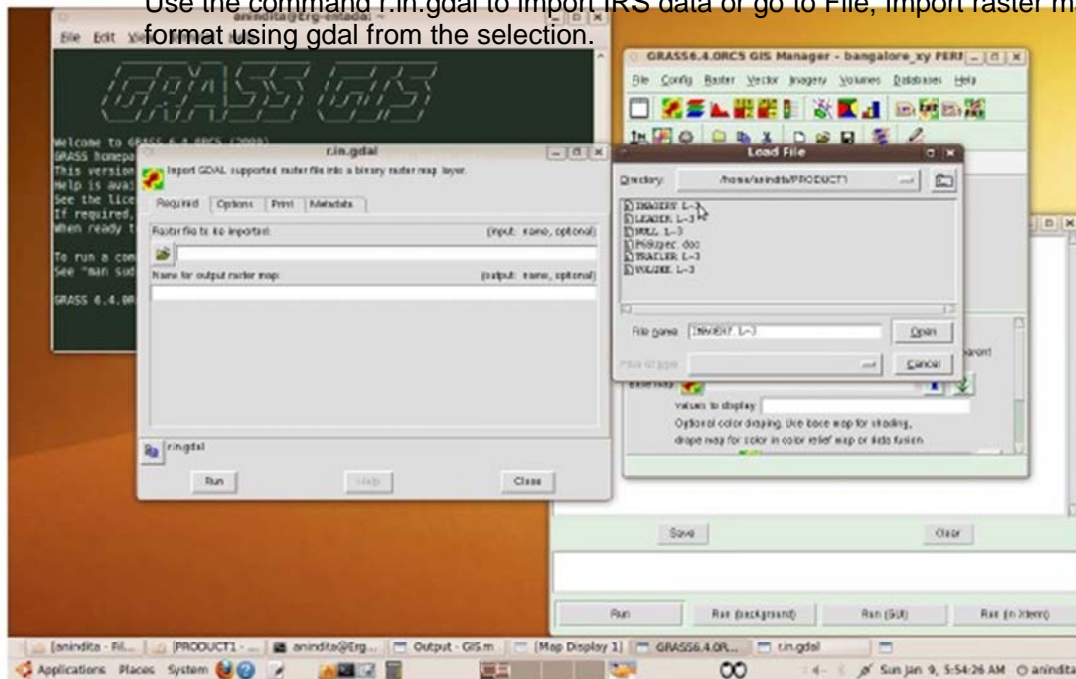
Give the input and output file name and click on Run.

This appears on screen when the program has ran completely.

Click on the map display to see the imported bands.
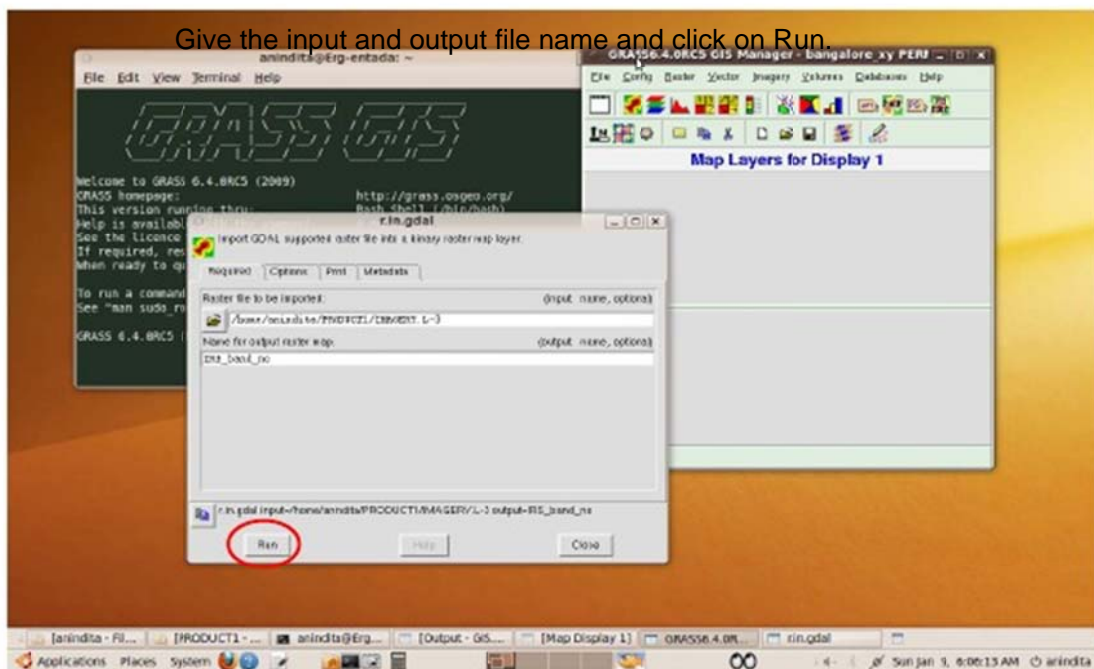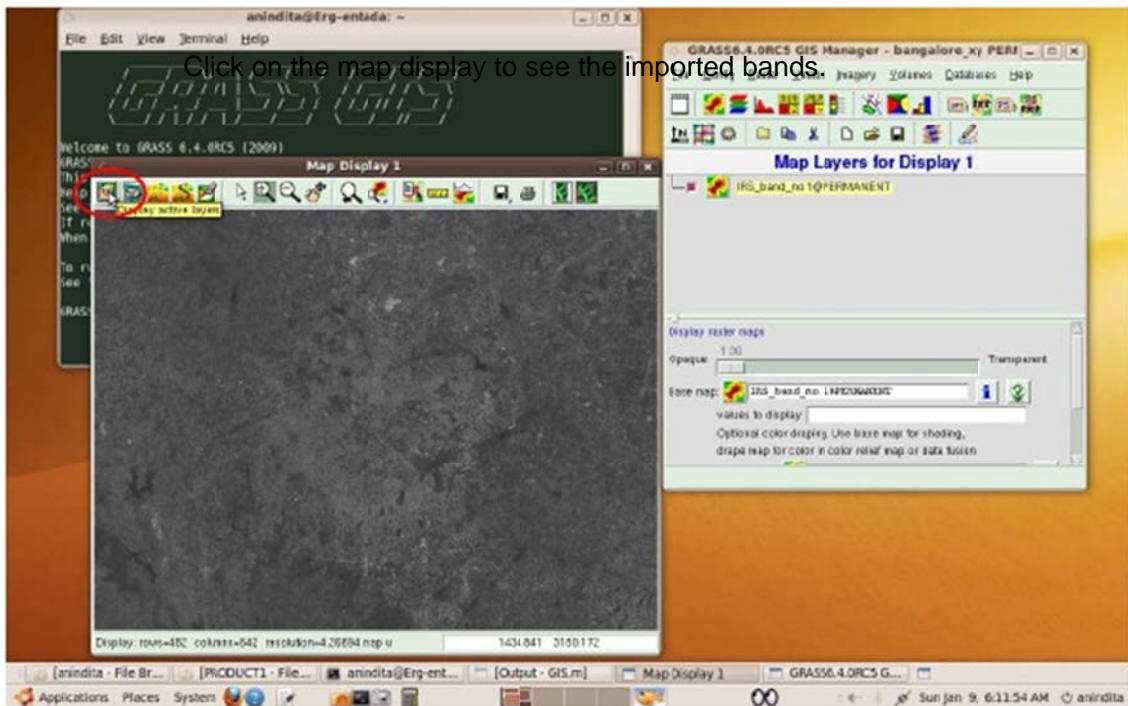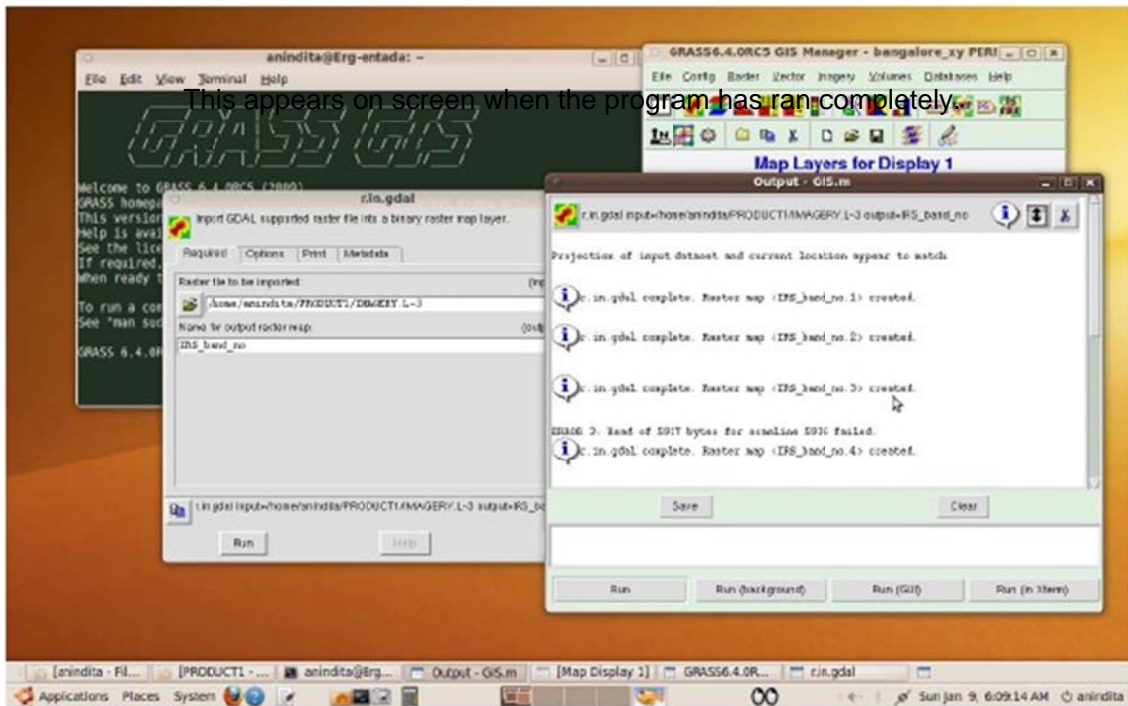
You can import vector data from other softwares in many formats (such as dxf, mif, shape file, kml, etc.) See the Import options in File menu. We use the command v.in.dxf to import dxf vector file.

See the output dxf vector file by clicking vector map display.
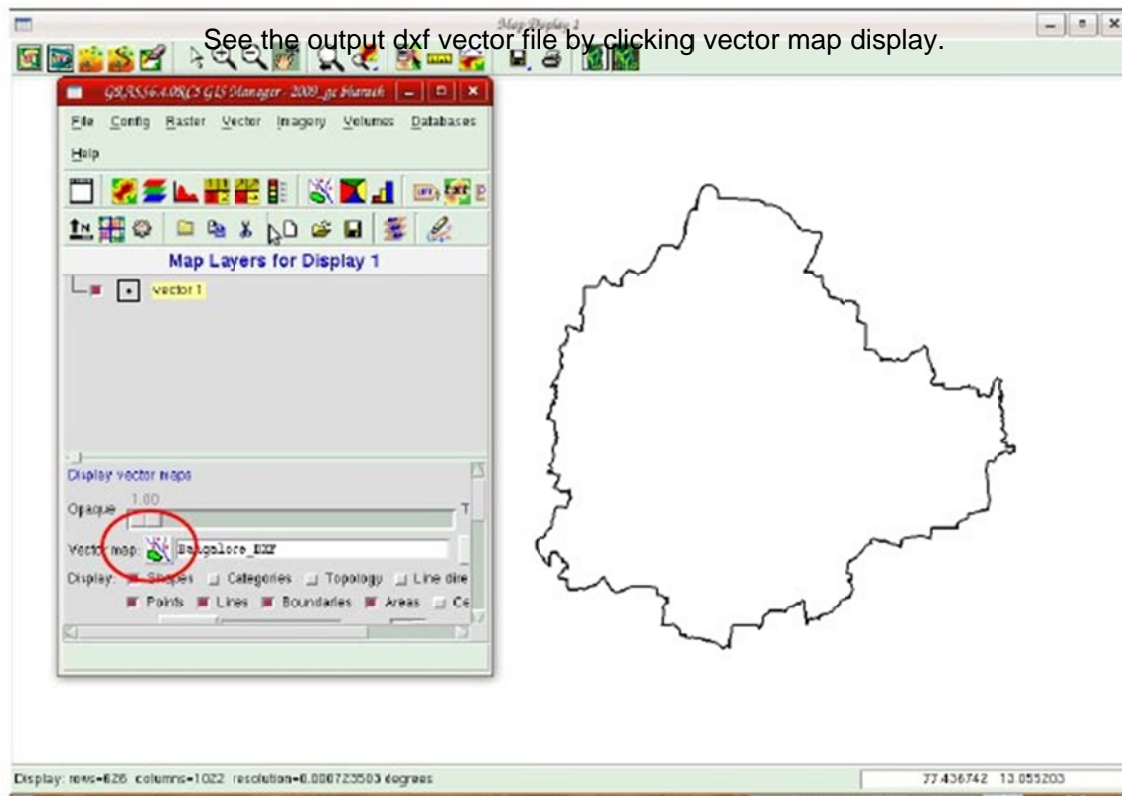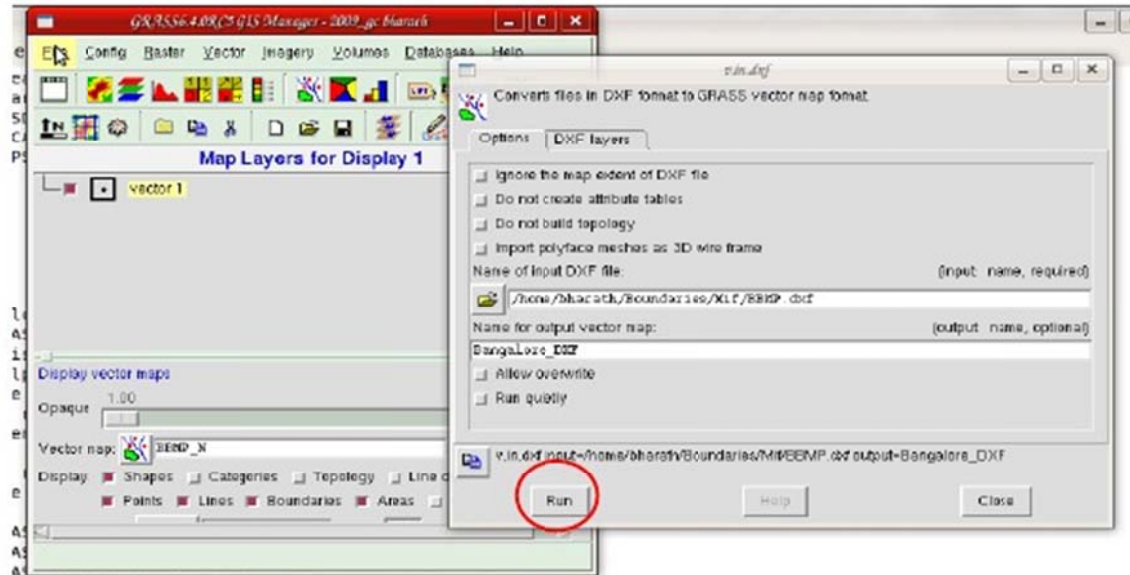
We use the command v.in.ogr to import kml (Obtained from Googleearth digitization) vector file and then click on Run.

Display the kml output vector file by clicking vector map display.
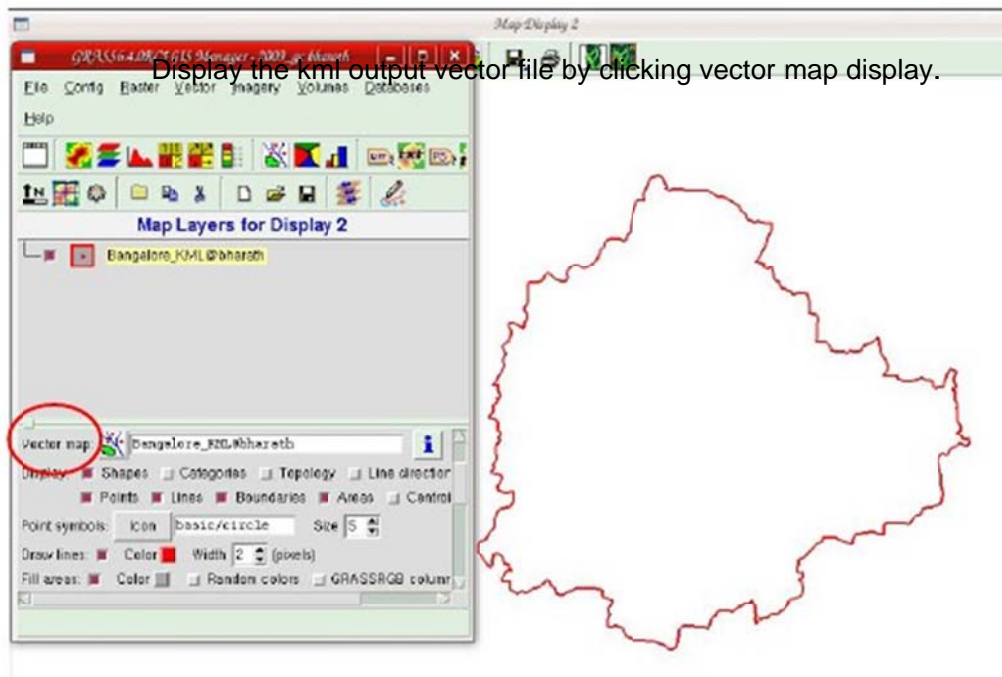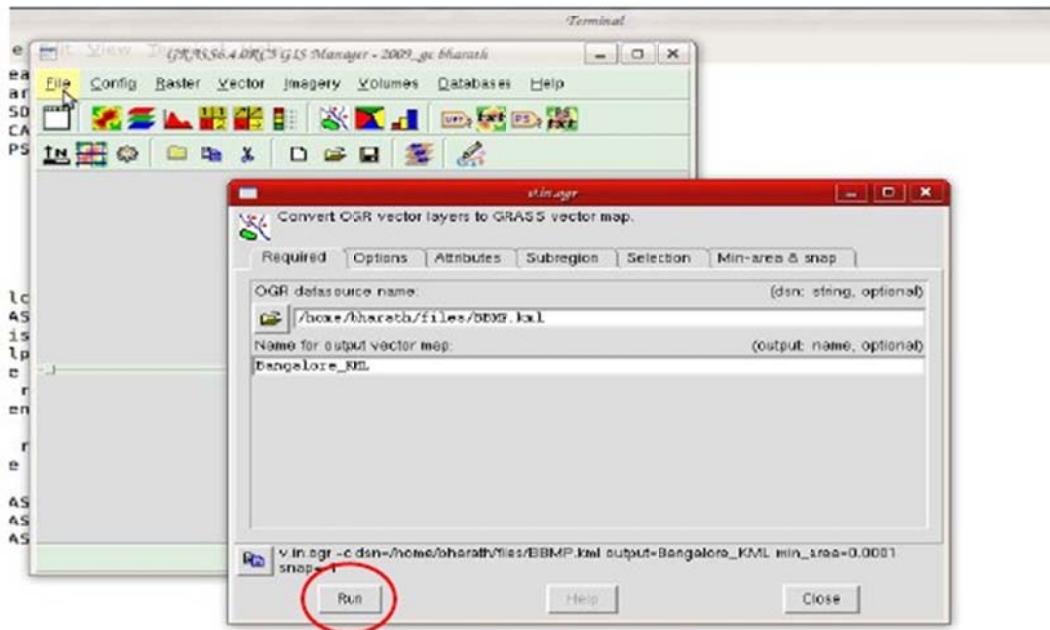
We use the command v.in.ogr to import shape vector file and then click on Run.
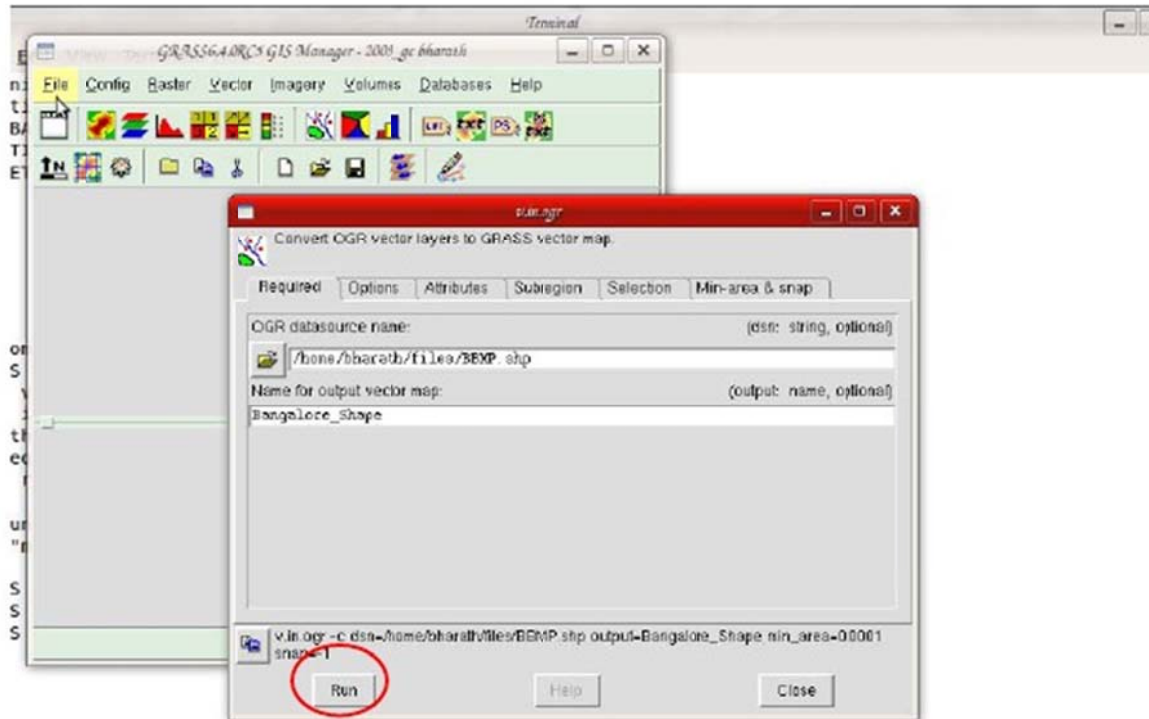
Clicking to vector map display to see the shape output vector file.
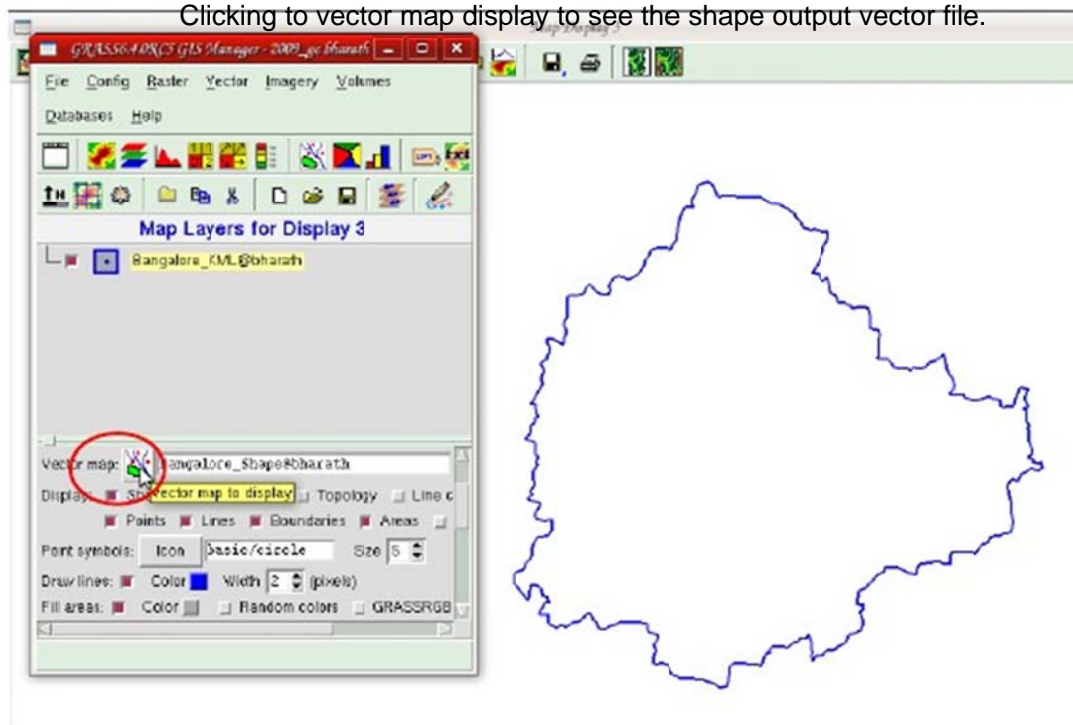
PART – X

_____

IMAGE PROCESSING – GENERATING FALSE COLOUR
COMPOSITE FROM LANDSAT and IRS LISS-III MULTI-
SPECTRAL MULTISPECTRAL IMAGERY

Type the command i.landsat.rgb on textbox or goto Imagery Tab on the main menu of GRASS to select the same option from GUI for creating False colour composite.
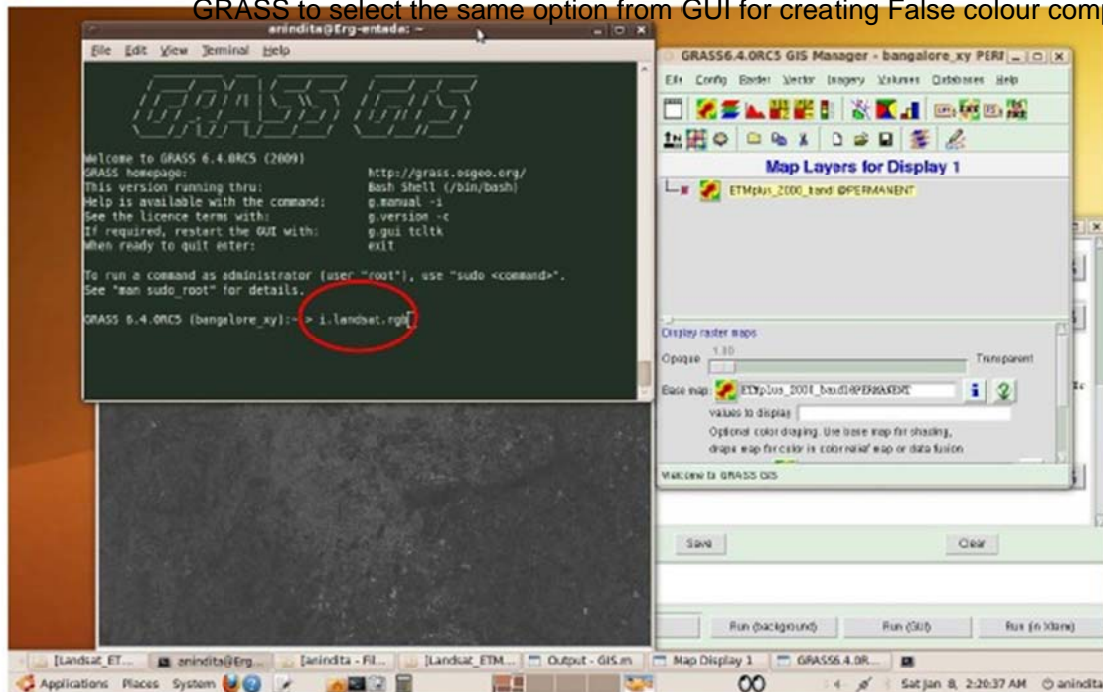
This is what appears.
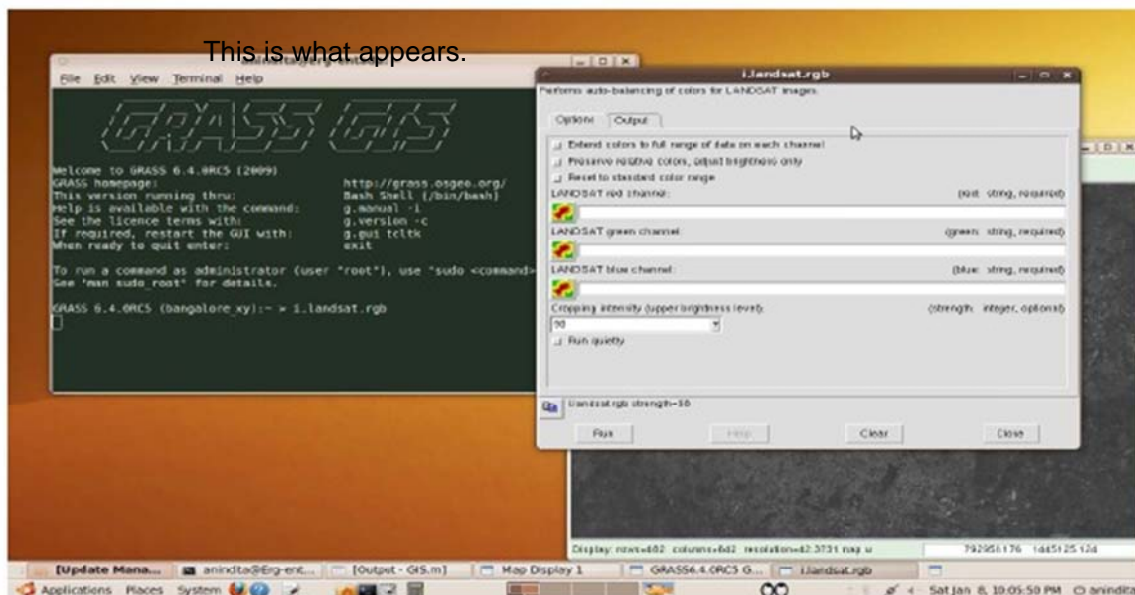
In LANDSAT red channel put ETMplus_2000_band4, LANDSAT green channel put ETMplus_2000_band3, LANDSAT blue channel put ETMplus_2000_band3 and click on Run.
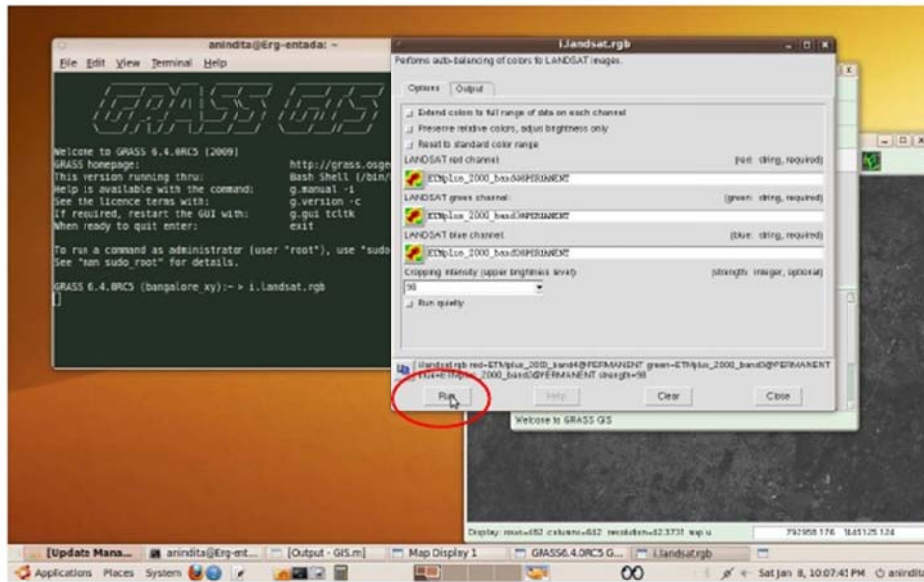
The program has ran.
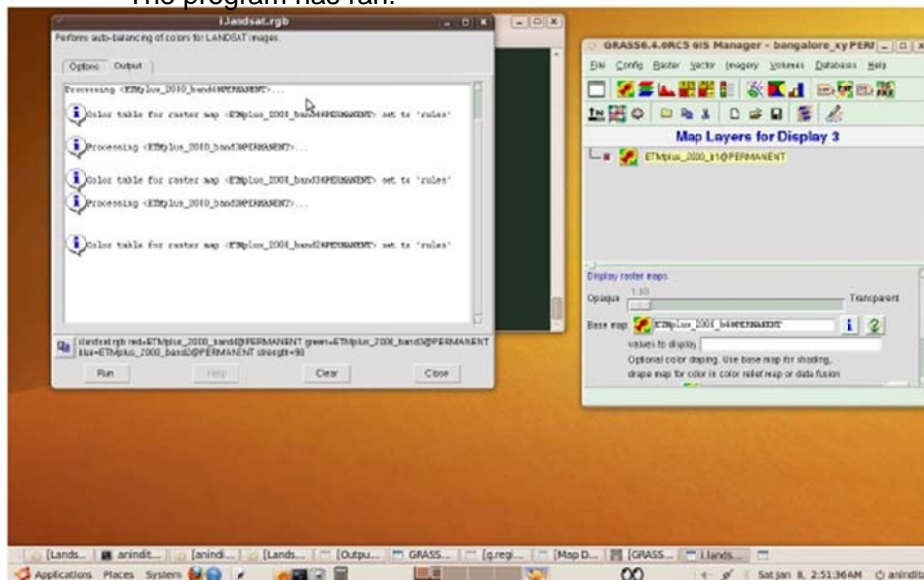
Type the command r.composite in the text box or browse from GRASS main Menu.
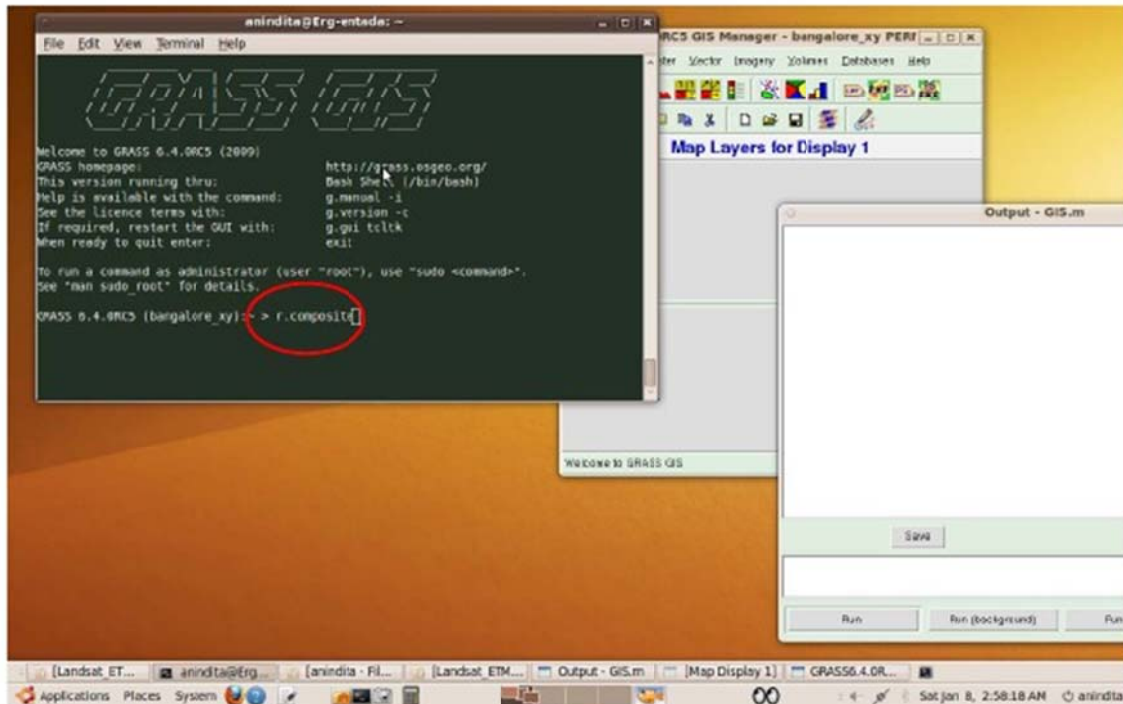
Name of raster map to be used for <red> put ETMplus_2000_band4.
Name of raster map to be used for <green> put ETMplus_2000_band3.
Name of raster map to be used for <blue> put ETMplus_2000_band2.
Give name to the output file and click on Run.
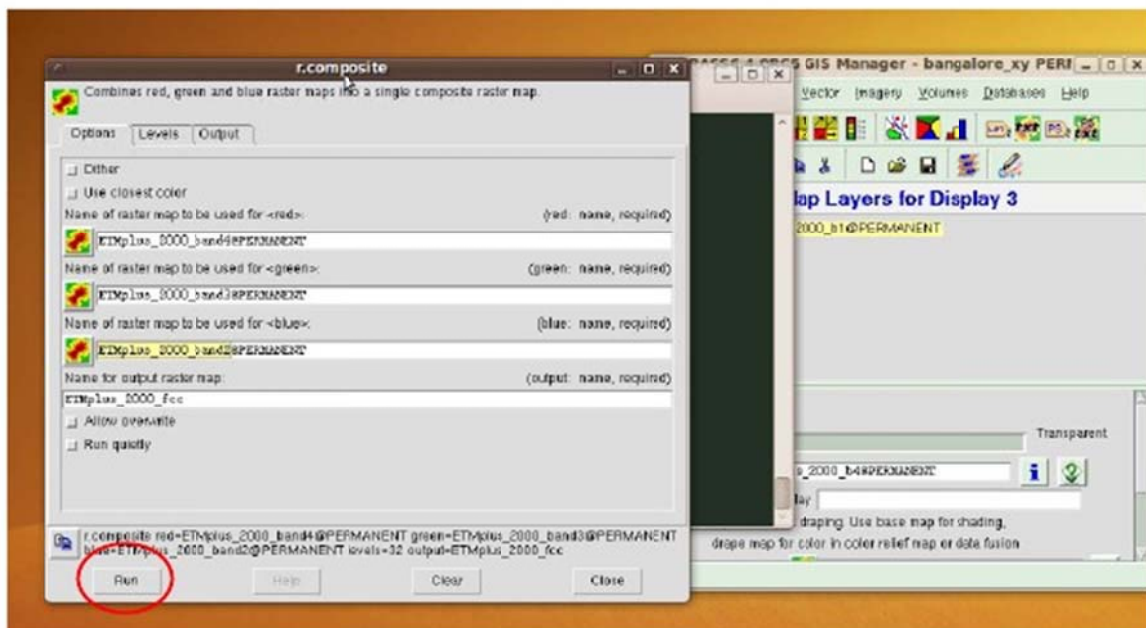
The program has ran.
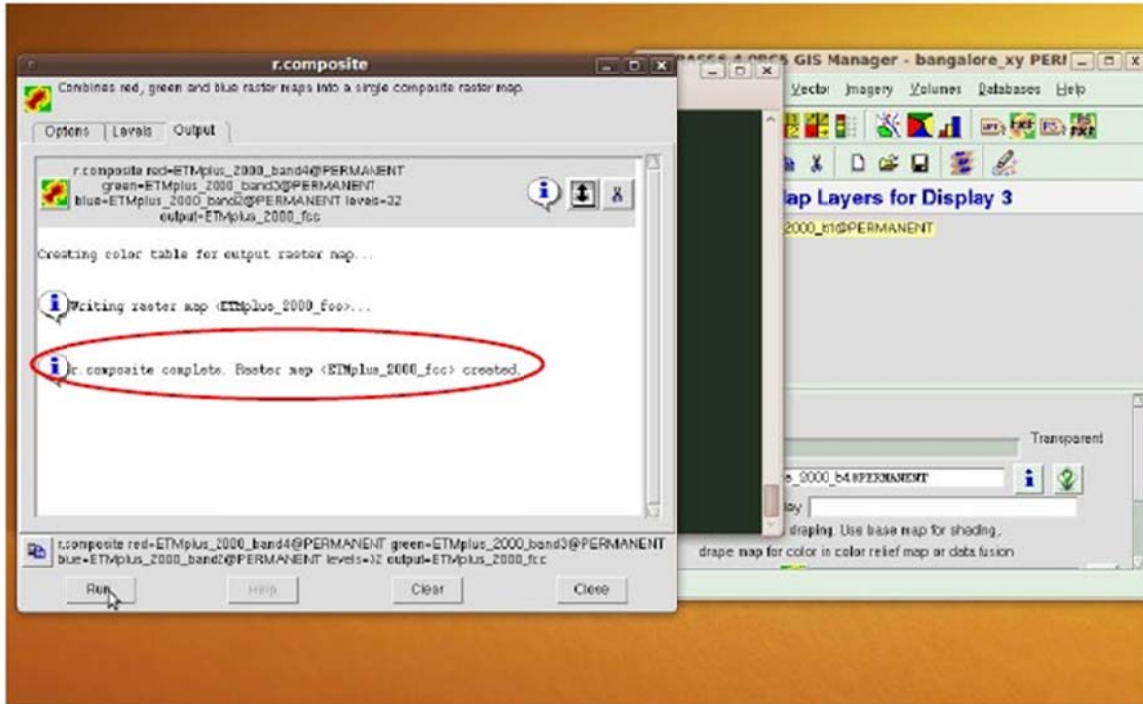
Display Landsat ETM False colour composite on the map display.
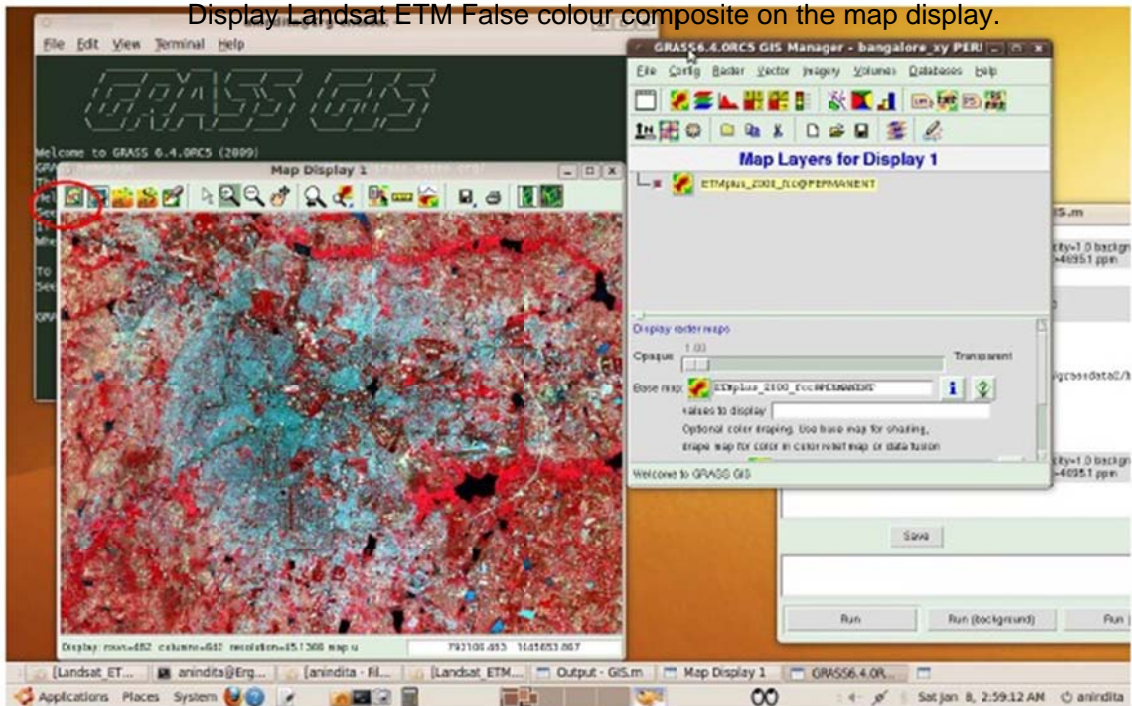
Display IRS False colour composite on the map display.
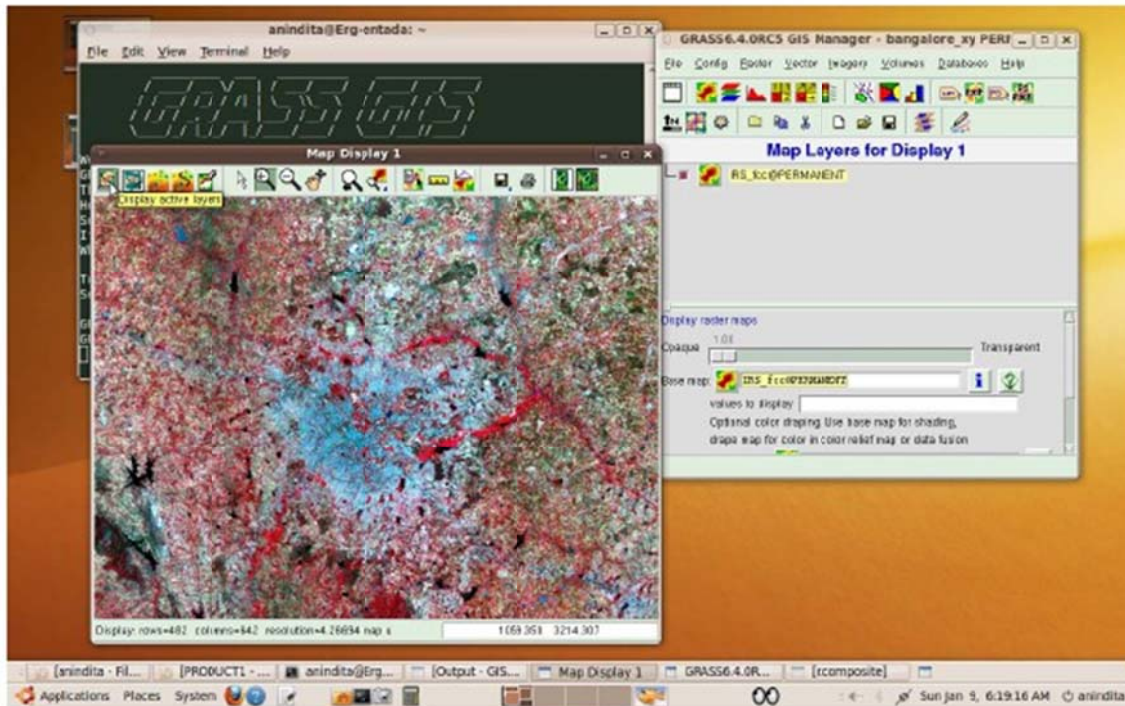
Tutorial for r.li program in GRASS

Also see http://grass.osgeo.org/gdp/landscape/r_le_manual5.pdf

Creation of configuration file

Type the command on the terminal - **r.li.setup.**
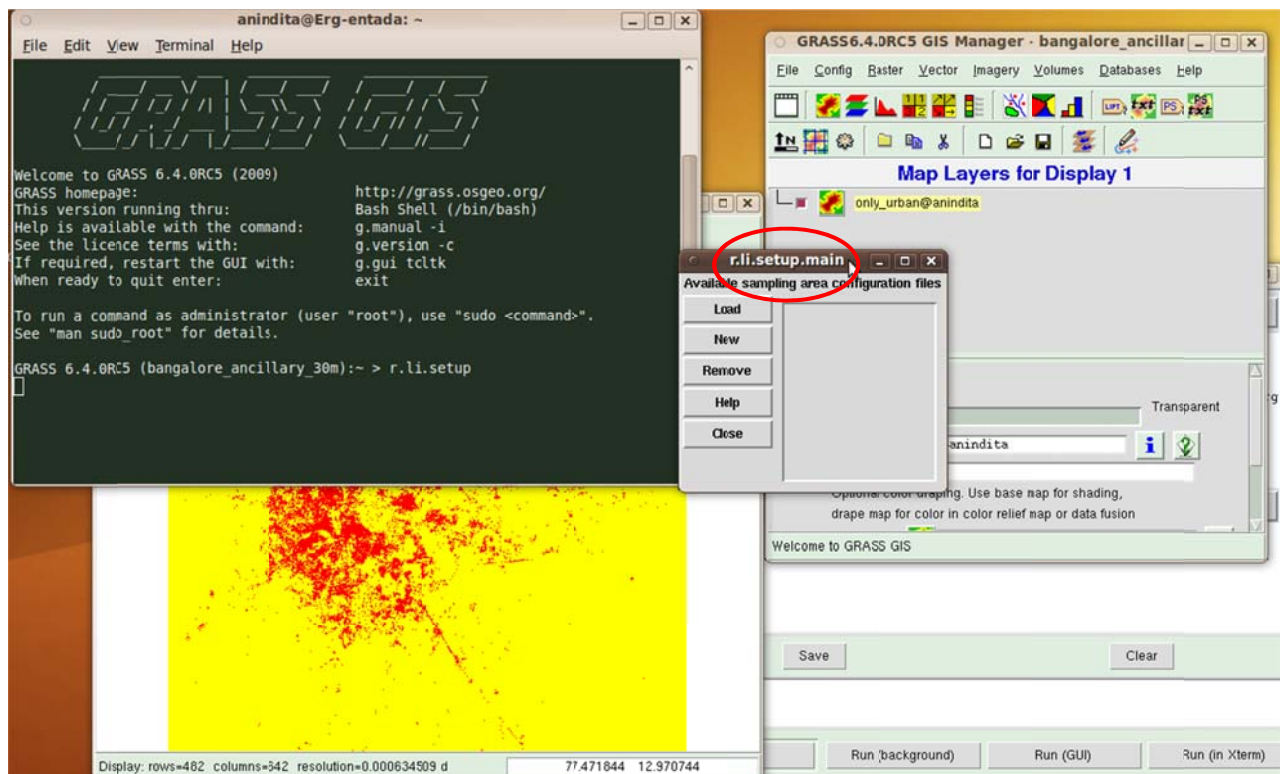
When r.li.setup window is displayed then click on **NEW.**
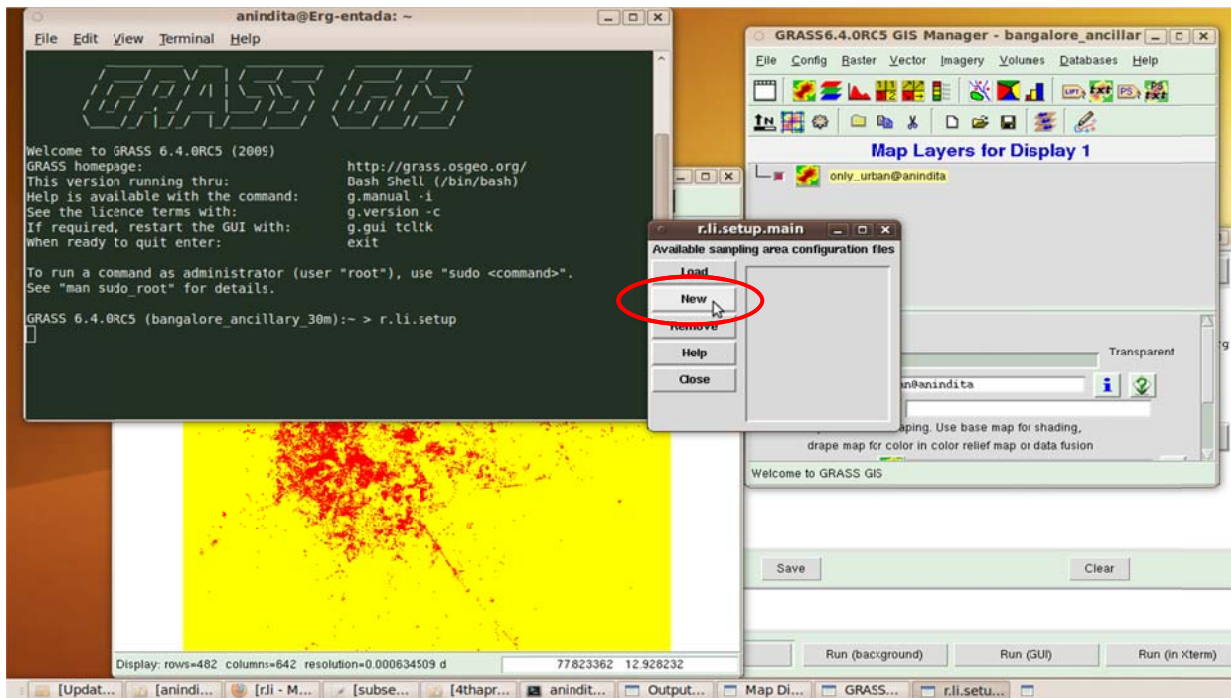
When the new configuration window is displayed then enter the **Configuration file name** and **name of the raster map**.
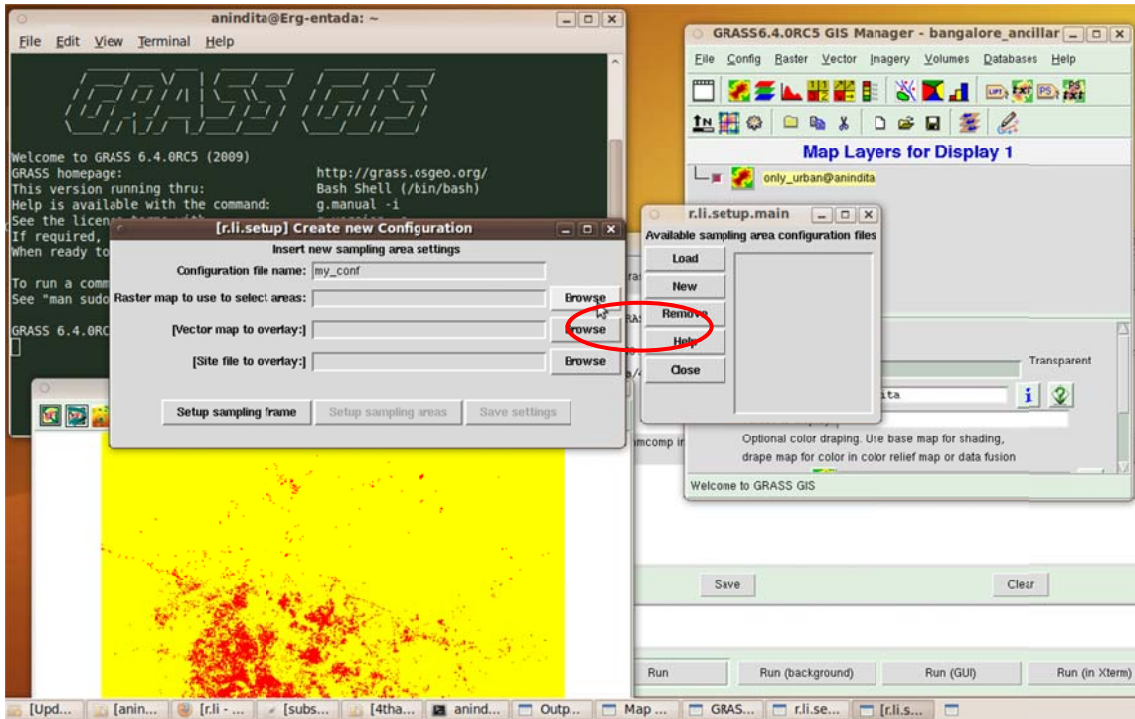
Click on **Setup sampling frame.**
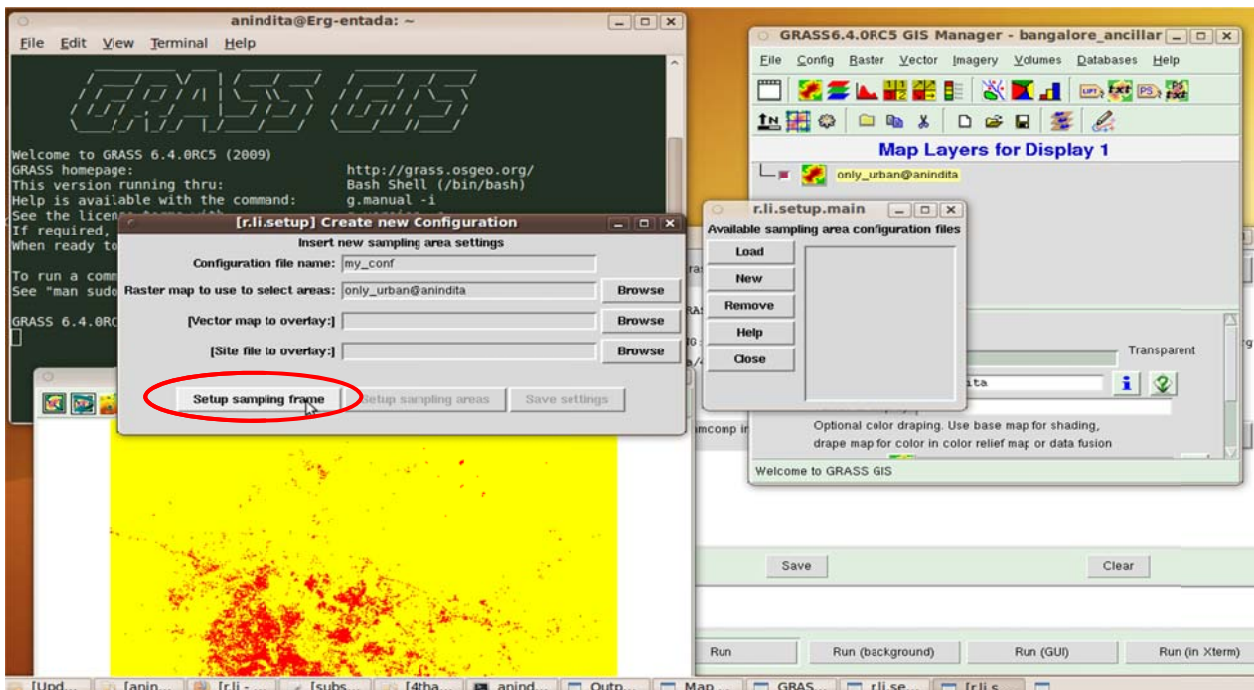
Then select **Whole maplayer** and click **OK.**
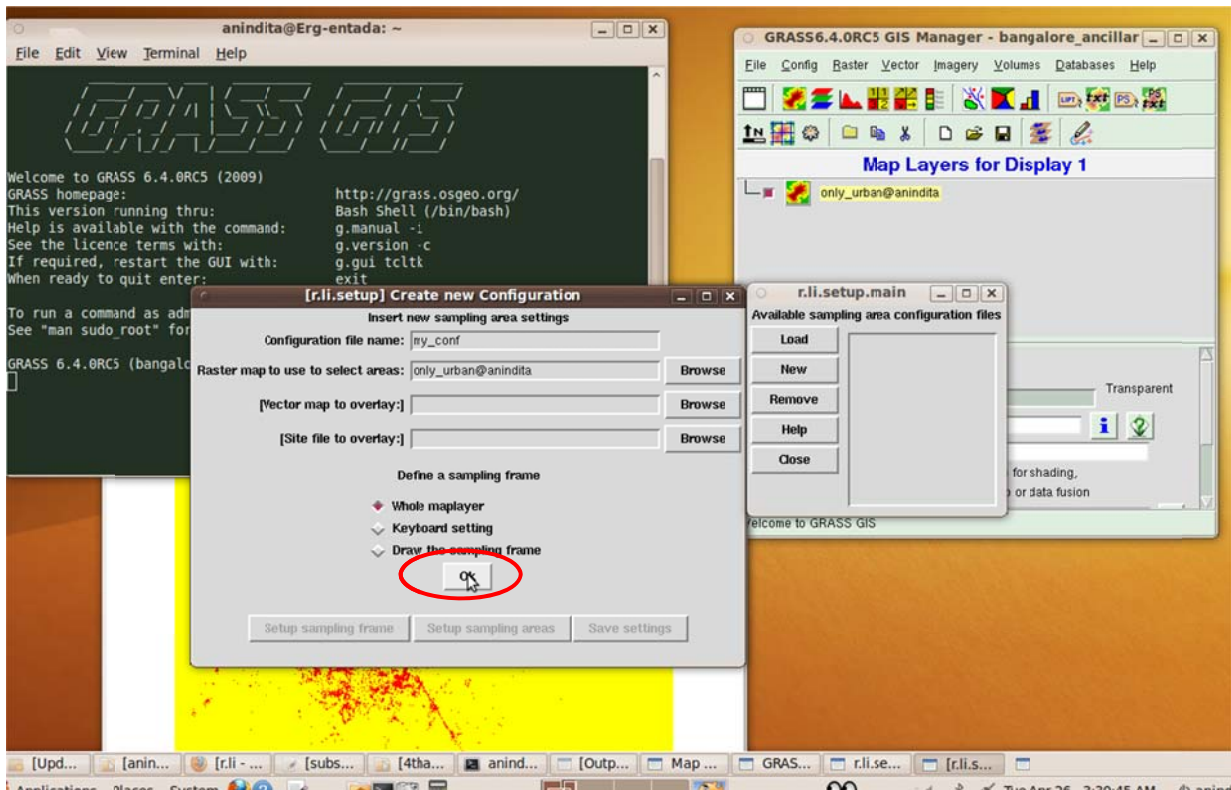
The whole maplayer is set as sampling frame.
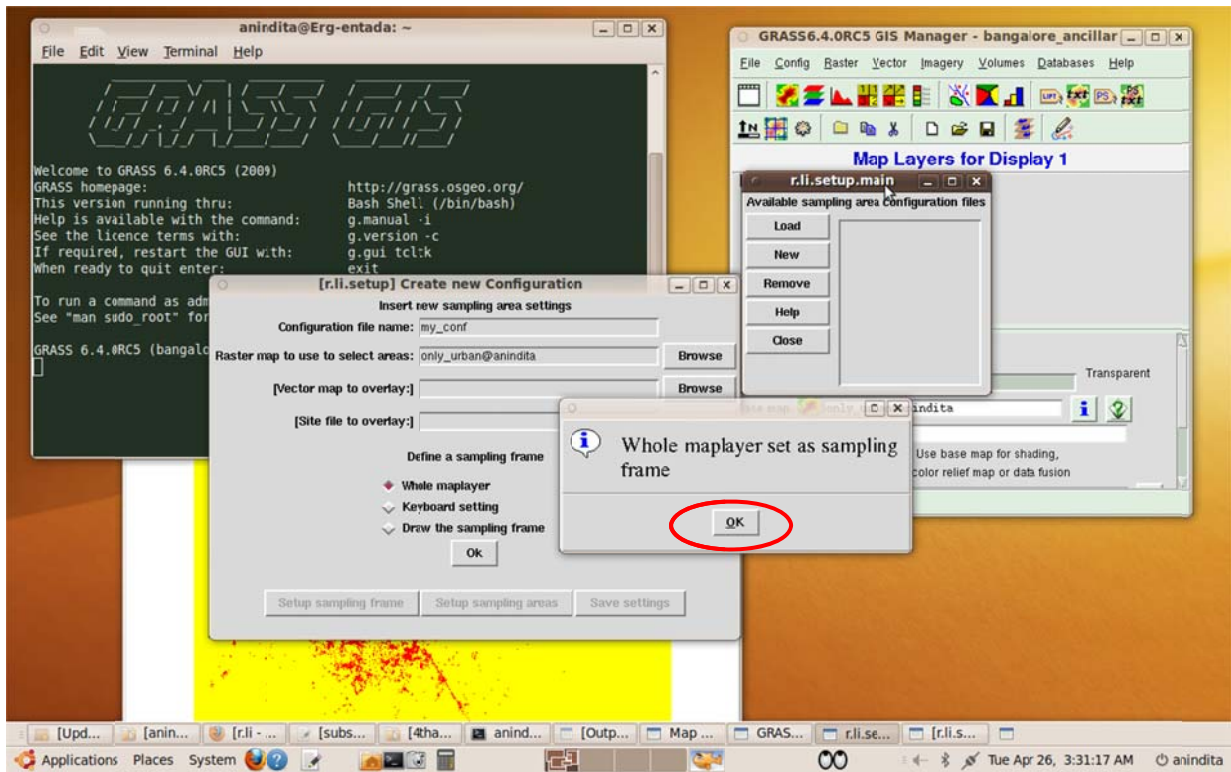
Click on **Setup sampling areas**.
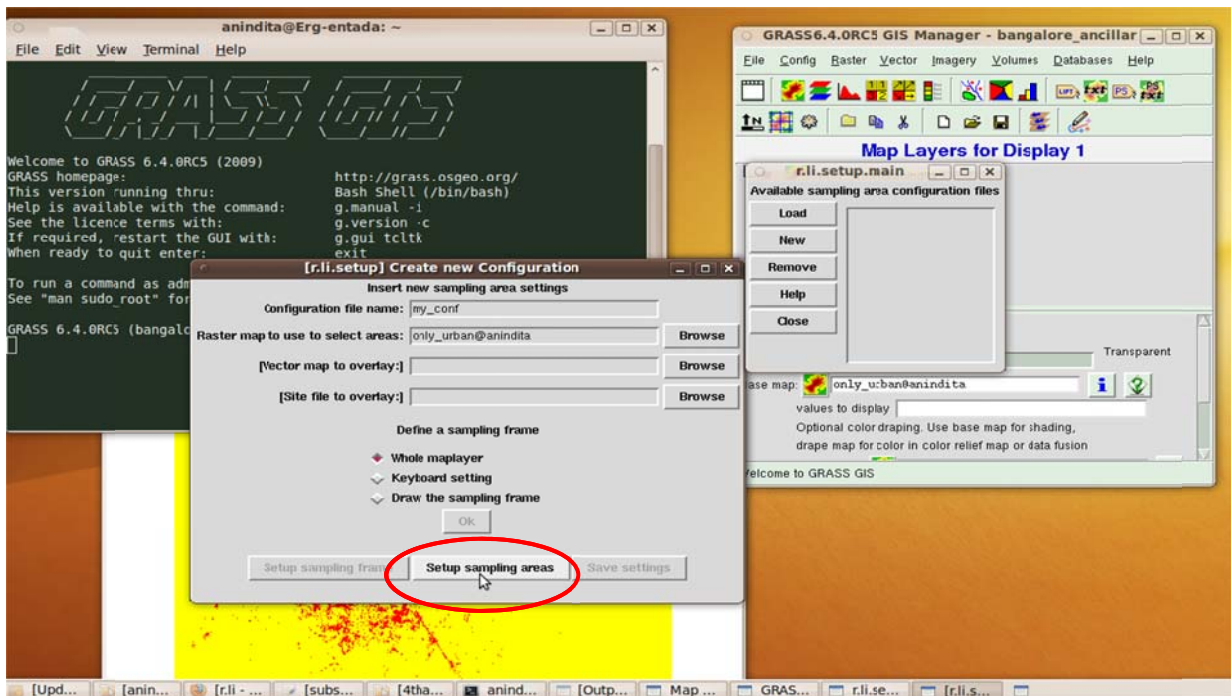
Select **Moving window** and click **OK.**
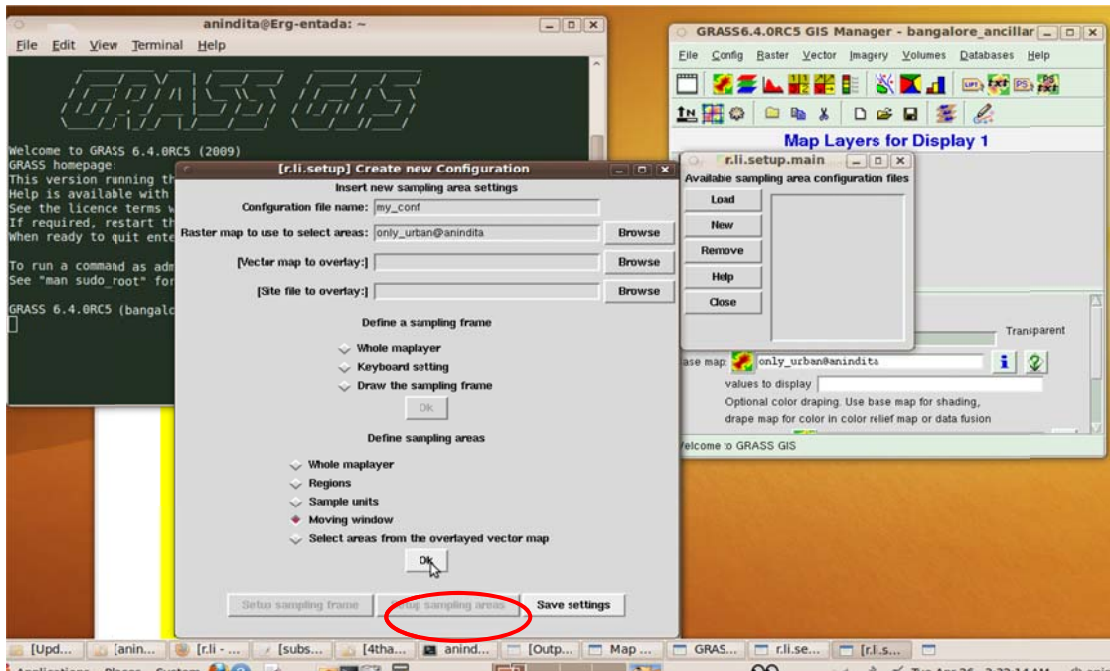
Click on **Use keyboard to enter moving window dimension**.
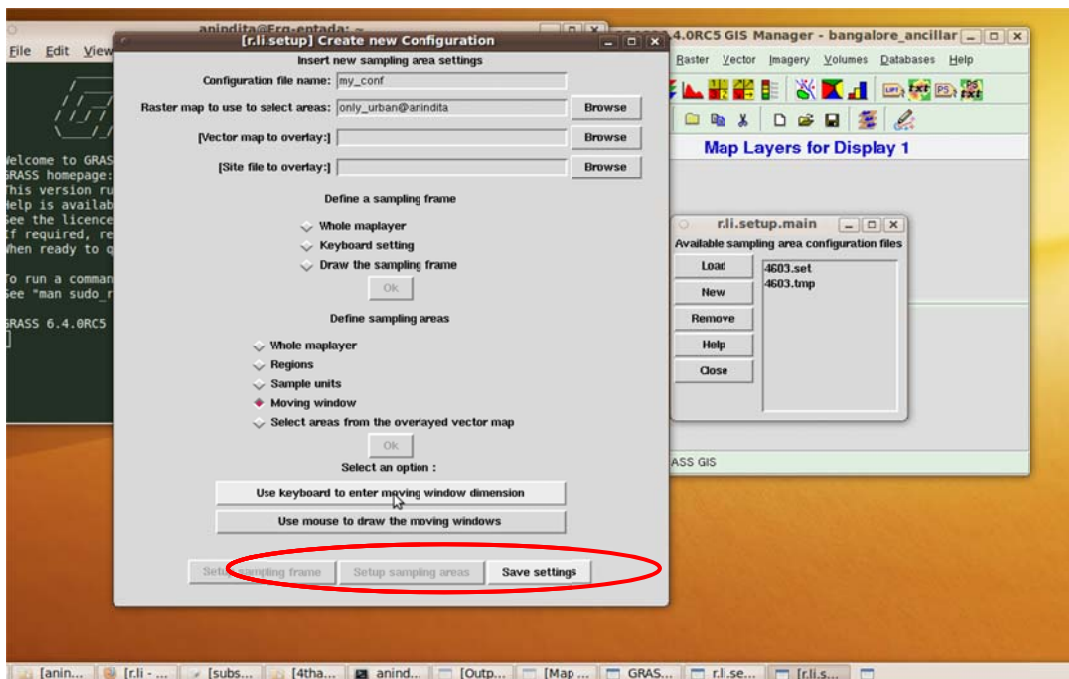
Select **Rectangle** and enter 5 on **height** and **width** sizes.
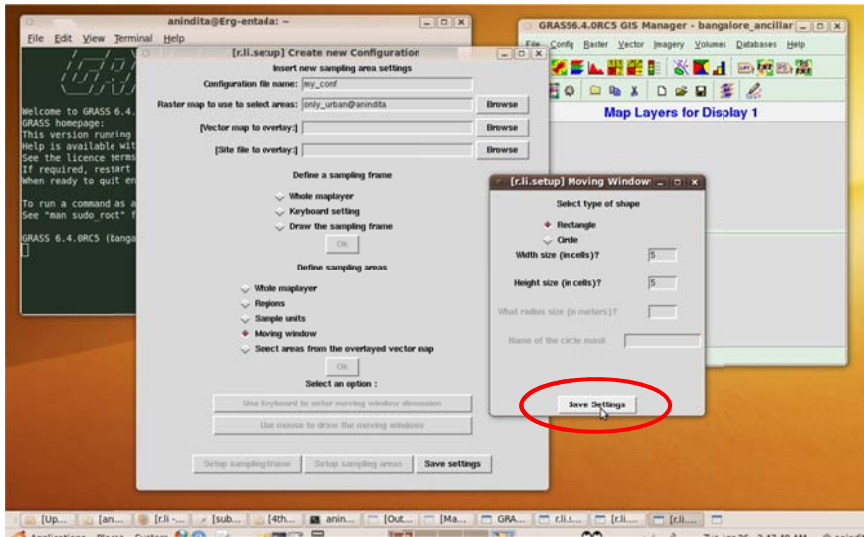
Click on **Save settings.**
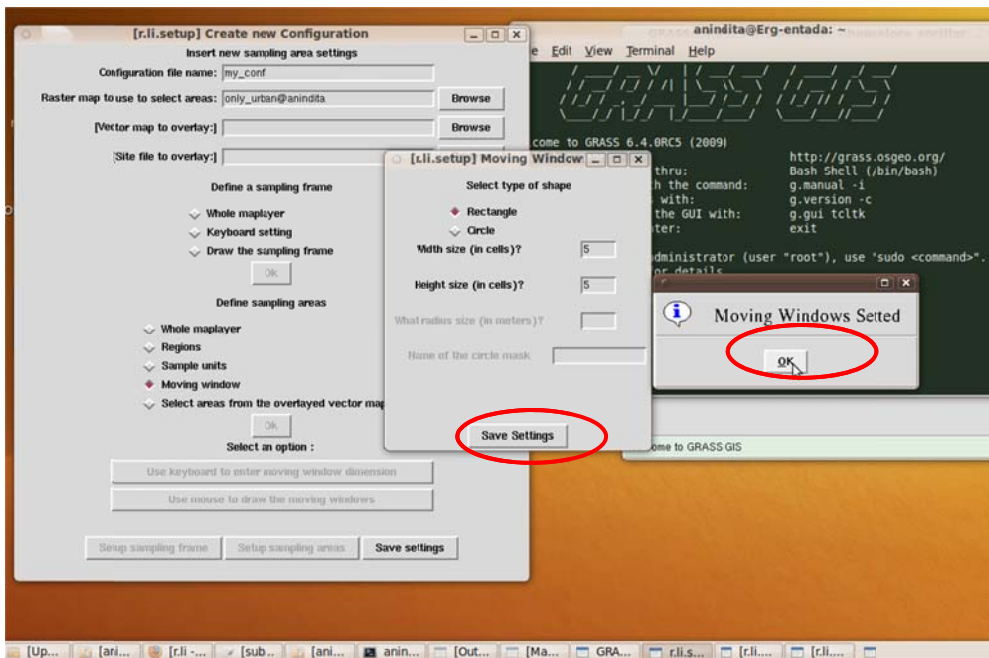
New configuration file **my_conf** is created!!!

The details of the metrics are given in

http://wgbis.ces.iisc.ernet.in/grass/grass64/manuals/html64_user/r.li.html

Type these commands on the **GRASS** terminal in order to get various spatial metrics such as MPS (Mean Patch Size), Patch density, Edge density, etc.

1. d.mon start=x0 (to start a monitor)

2. r.le.trace: Computes the number of patches along with their area and perimeter.

3. r.le.trace -pt map=urban_only out=urban_only_patch

4. r.li.mps: Calculates mean patch size index on a raster map, using a 4 neighbour algorithm
    r.li.mps map=urban_only conf=my_conf output=urban_only_mps

    A map of NULL values is considered to have zero patches. If you want to have null values instead run

    r.null setnull=0 map=urban_only_mps

5. r.li.patchdensity: Calculates patch density index on a raster map, using a 4 neighbour algorithm
    r.li.patchdensity map=urban_only conf=my_conf output=urban_only_patchden

    A map of NULL values is considered to have zero patches.
    If you want to have null values instead run

    r.null setnull=0 map=urban_only_patchden

    r.null setnull=-1 map=urban_only_patchden (To remove negative values from the image).

6. r.li.padcv: Calculates coefficient of variation of patch area on a raster map r.li.padcv map=urban_only conf=Ani_conf output=urban_only_padcv   r.null setnull=-1 map=urban_only_padcv

7. r.li.edgedensity: Calculates edge density index on a raster map, using a 4 neighbour algorithm   r.li.edgedensity map=urban_only conf=my_conf output=urban_only_edgedens r.null setnull=-1 map=urban_only_edgedens

QGIS OVERVIEW

Quantum gis is a GIS tool for manipulating geographical data, 3-D analysis, statistical analysis. QGIS is Open Source software and it is free of cost.It is an official project of OSGEO (Open Source Geospatial Foundation).Available under GNU General Public License.QGIS is multiplatformgis that runs on Linux, Unix, Mac OSX, and Windows having supportfor vector, raster, and database formats. It supports many common spatial data formats (e.g. ESRIShape File, geotiff). QGIS supports Plug-ins to do things like display tracks fromthe GPS units. Qgis is a *modular architecture,* which can be extended by modules for new data sources and plugins.

1. History:

Gary Sherman in February 2002 started to develop a giswith a wide range of data stores.The first release came on July 19, 2002. The first release supported only Post GIS layers.The current version is Qgis 1.6.0.
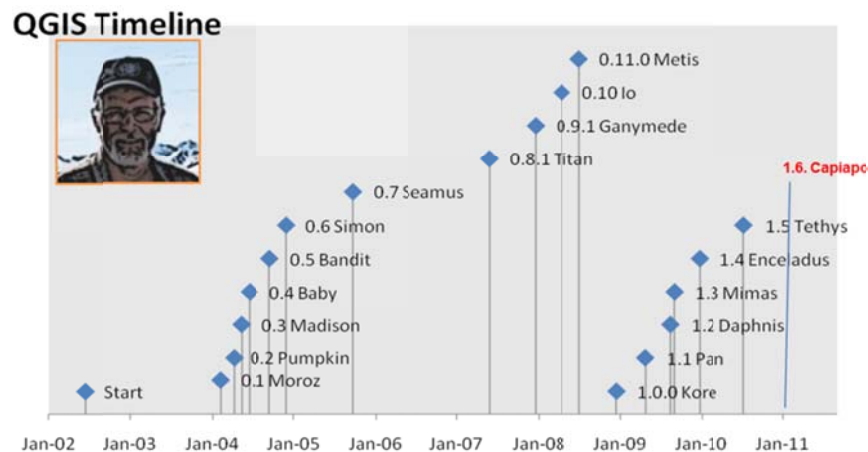
2. Getting Started

2.1. Installation

Building QGIS can be possible in two ways one is from source and other is binaries. The Installation instructions are distributed with the QGIS source code and also available at http://qgis.org.Standard installer packages are available for Windows and Mac OS X. For GNU/Linux binary packages are also available. Get the latest information on binary packages

at the QGIS website.

For Ubuntu follow the steps shown below

Click on System→ Administration→ Softwaresources. Enter pass word; a dialogue box will be displayed click on Other Software; this will show the package information which is included in the system.
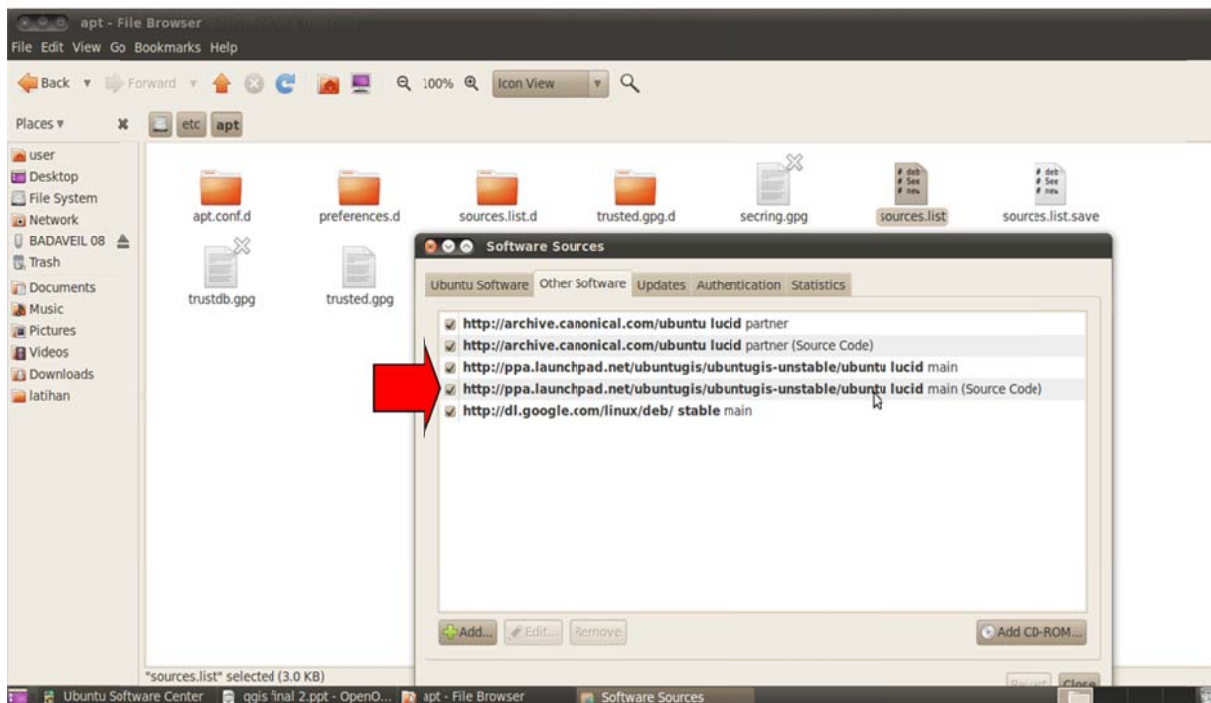
Click on Synaptic package Manager to install Qgis and other sources.
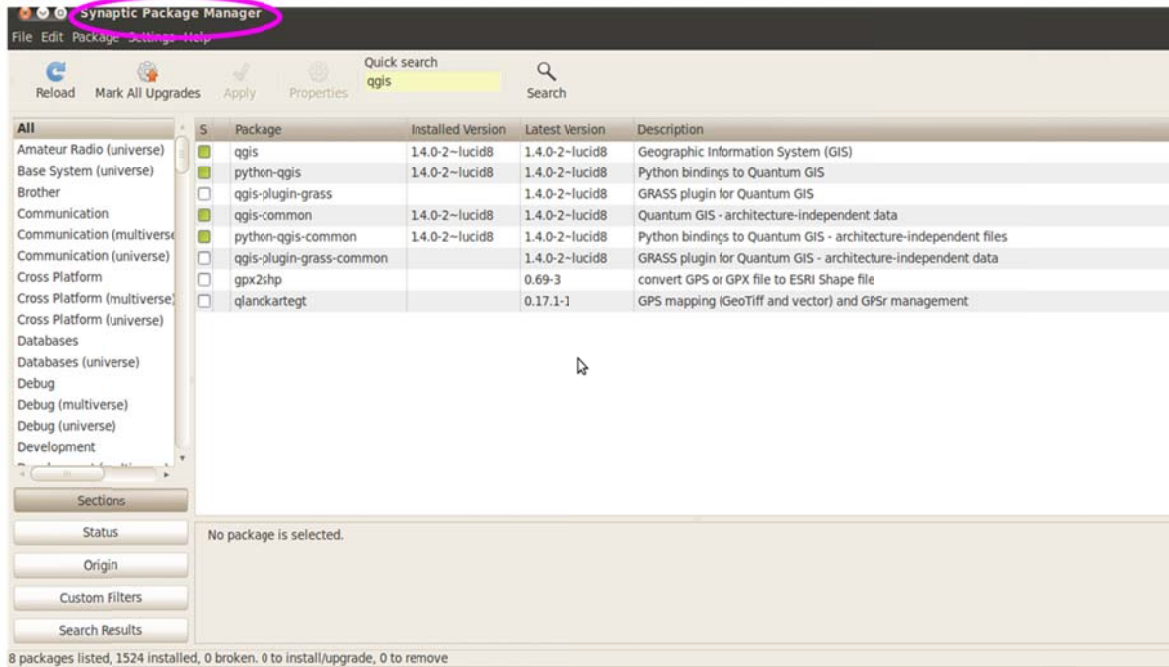
Or

- First Go to **System -> Administration -> Software Sources**
  (enter your password)
  "**Third Party Software**" Tab
  Click "**+Add**" When it prompts you for a APT Line, enter:
- deb http://ppa.launchpad.net/qgis/ubuntu/ jaunty main

  click on "**Add Source**"
  Then click "**Reload**" to update
- Then type into the terminal:

Code:

```
~$ sudo apt-get install qgis
```

(It should install)
if it doesn't then try this:
Code:

```
~$ sudo apt-get update
```

Code:

```
~$ sudo apt-get install qgis
```

Code:

~$ qgis

It opens Qgis session. Or you can find it automatically added to programs:

Applications -> Education -> Quantum GIS

2.2 Starting QGIS

To start QGIS, choose Qgis from the menu, Applications→ Education→Qgis. Now the QGIS map canvas and an empty legend will be appeared.

Exploring the GUI interface:

GUI shows 6 different menu bars they are

1. Menu bar   2.Tool bar   3.Legend   4.Overview map   5.Map canvas
6. Status bar
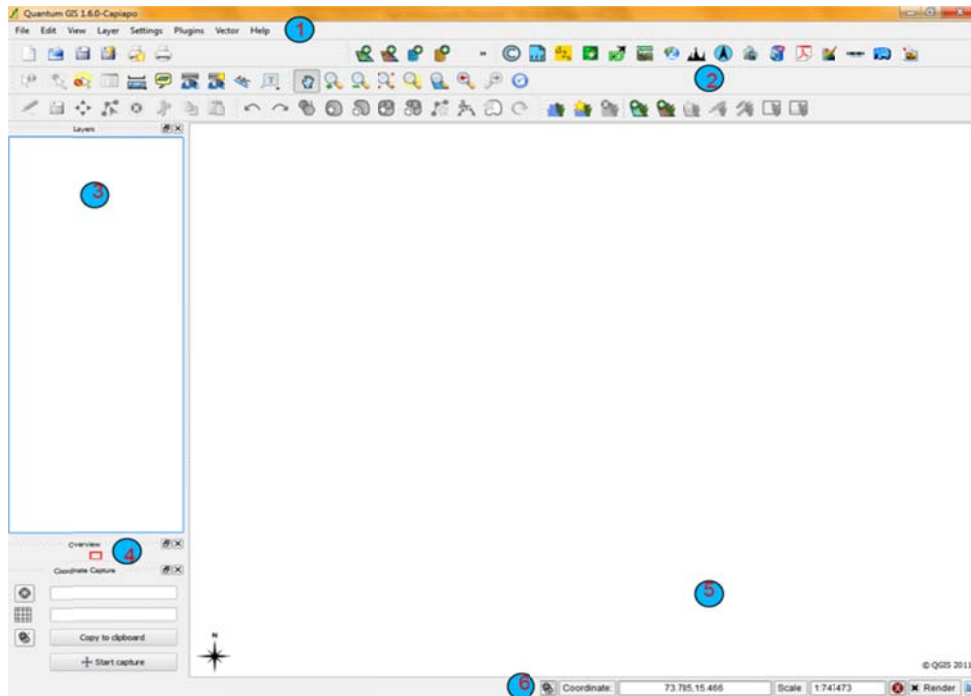
→The menu bar provides access to numerous QGIS features using a standard hierarchical menu. The most menu options have a corresponding tool and vice-versa; the menus are not organized quite like the toolbars. The toolbar containing the tool is listed after each menu option as a checkbox entry.

→The toolbars offers access to most of the similar functions as the menus, plus additional tools for interacting with the map. Each toolbar item has popup help available. Hold themouse over the particular icon, a short description of the tool's purpose will be

displayed. Every menu bar can be moved around and can be switched off using right mouse button context.

→The map legend area sets the visibility (used to show or hide the layer by check box)and z-ordering of layers. Z-ordering means that layers listed nearer the top of the legend are drawn over layers listed lower down in the legend.

→ QGIS - maps are displayed in map canvas area.The loaded vector and raster layers will be displayed in this window. The map view can be panned, zoomed in and out. The map view and the legend are strongly bound to each other - the maps in view reflect changes you make in the legend area.

→The map overview panel provides a full extent view of layers added to it, within the view is a red rectangle showing the current map extent. This allows to quickly determining which area of the map is in current view.

→The status bar shows the current position in map coordinates (e.g. meters or decimal degrees) as the mouse pointer is moved across the map view. A progress bar in the status bar shows progress of rendering as each layer is drawn to the map view. If a new plugin or a plugin update is available, a message will be shown in the status bar. At the far right of the status bar is a projector icon; by clicking it opens the projection properties.

2.3 Working with Vector Data:

QGIS uses the GDAL/OGR library to read and write vector data formats. QGIS supports numerous vector data formats, including those supported by the OGR library data provider plugin, such as ESRI shapefiles,kml etc.QGIS also supports PostGIS layers in a PostgreSQL database using the Postgre SQL data provider plugin. Support for additional data types (ex. delimited text) is provided by additional data provider plugins. To load a vector file click on ![Add Vector Layer] available in tool bar or ctrl+shift+v, a dialogue box will be displayed allows to traverse through the file system and load a shape file or other formats.
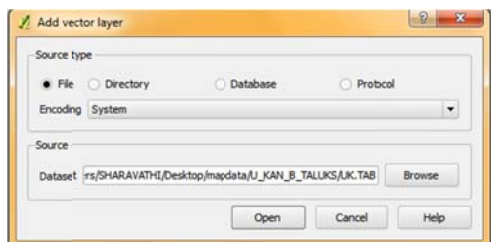
The vector file will be loaded and displayed in map canvas.
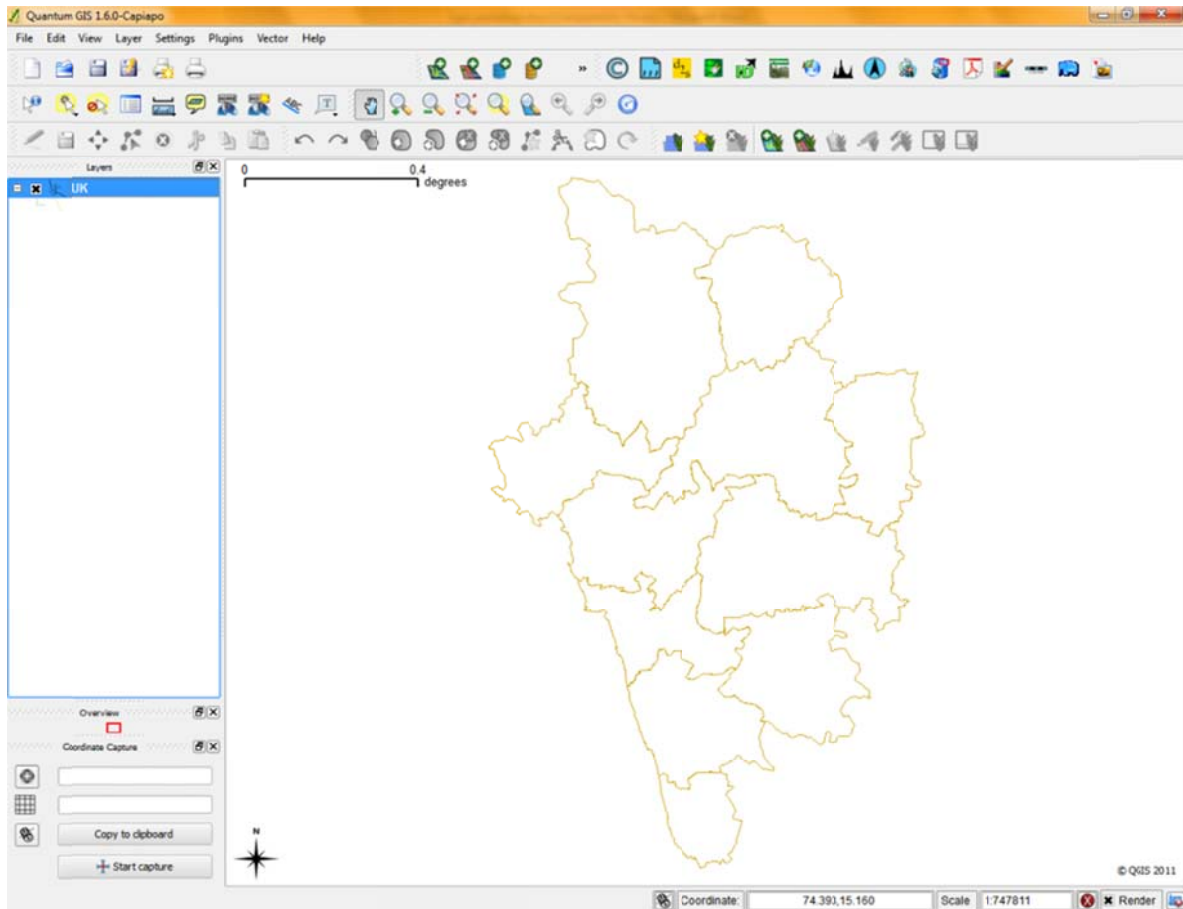
To improve the performance of vector file, righ click on the layer name select properties. The properties dialogue box will be opened with multiple options i.e Symbology tab, Lables, Attributes, General, Metadata, Actions, Diagram overlay.
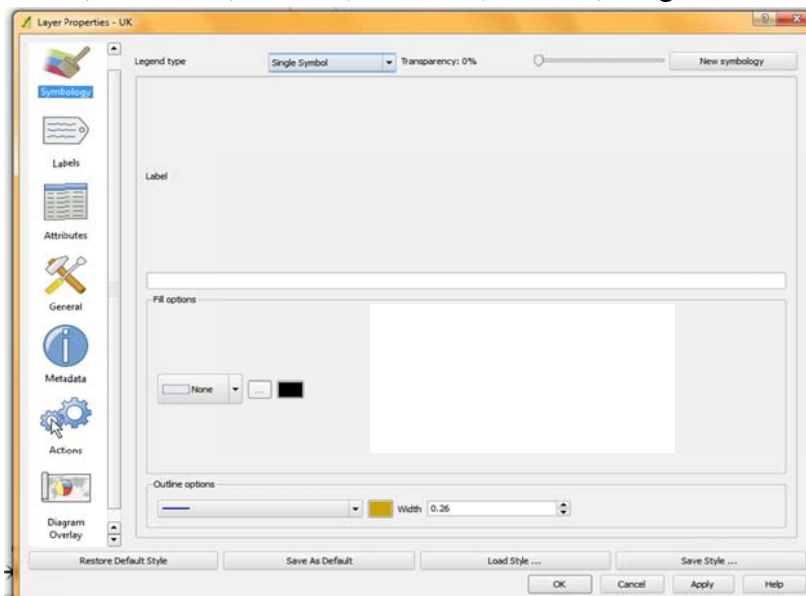
➢ QGIS supports a number of symbology renderers to control how vector features are displayed. Single symbol - a single style is applied to every object in the layer. Graduatedsymbol - objects within the layer are displayed with different symbols classified by the values of a particular field.Continuous color - objects within the layer are displayed with a spread of colours classified by the numerical values within a specified field. Unique value - objects are classified by the unique values within a specified field with each valuehaving a different symbol.

➢ The Labels tab allows to enable labeling features and control a number of options related to fonts,placement, style, alignment and buffering.

➢ The Attributes tab allows to manipulate the attributes of the selected dataset.The buttonsNew Column and Delete Column can be used, when the dataset is Editing mode. The OGR library supports toadd new columns, but not to remove them, if GDAL version >= 1.6 installed then the deletion is allowed.

➢ The General tab permits to change the display name, set scale dependent rendering options, view or change the projection of the specific vector layer.

➢ The Metadata tab contains general information about the layer whichis not yet editable. It includes the type andlocation, number of features, feature type, and the editing capabilities. The Extents section, providing layerextent information, and the Layer Spatial Reference System section, providing information about the projection ofthe layer.

➢ QGIS provides the facility to implement an action based on the attributes of a feature. Actions are useful to perform a search based on an attribute valueor view a web page based on oneor more values in vector layer.

➢ The Diagram tab used to overlaya graphicsto a vector layer.It allows overlyingpie charts, bar charts.

2.3Working with Raster Data

The raster data in GIS is usually remote sensing data such as aerial photography or satellite imagery and modelled data such as an elevation matrix. In QGISthe GDAL libraryis used for raster implementation.QGIS supports the following raster formats,
– Arc/Info ASCII Grid   – GRASS Raster   – Geotiff   – JPEG   –   Erdas Imagine
– USGS ASCII DEM
Raster layers are loaded either by clicking on the Load Raster icon or by selecting the View Add Raster Layer menu option. More than one layer can be loaded at the same time by holding down the Control or Shift key and clicking on multiple items in the dialog. Once a

raster layer is loaded in the map legend by clicking the right mouse buttonopens a dialog to set raster properties for the layer.

Improvethe performance by right click on layer and change properties by adding the different options. The properties include – Symbology, Transparency, Colormap, General, Metadata, Pyramids, and Histogram.
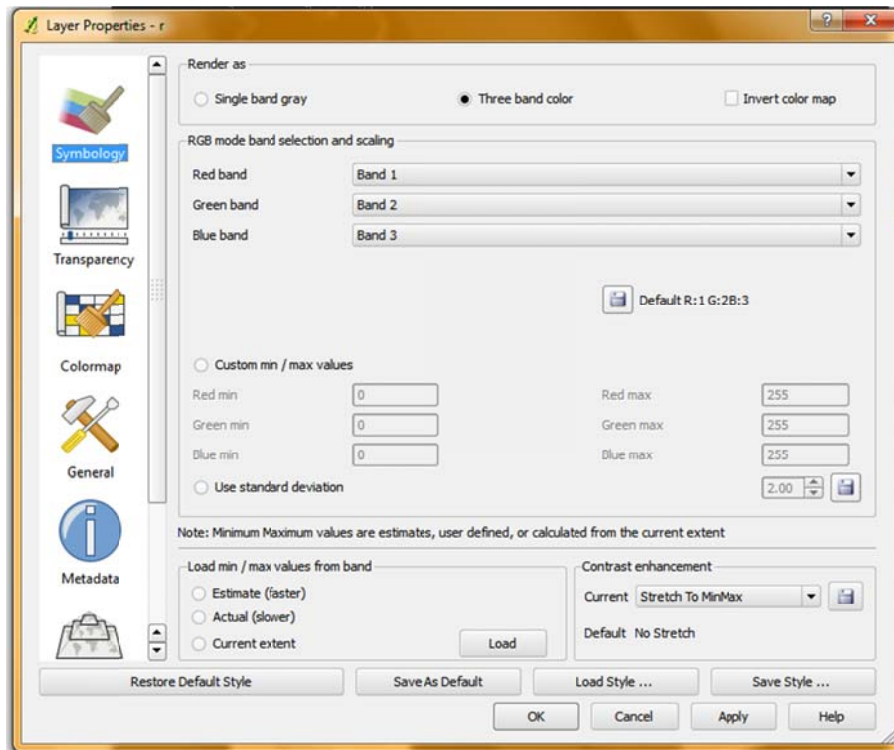
- ❖ Qgis renders the raster layers in two different forms.
- ➢ Gray scale (one band of the image will be rendered as gray or in pseudocolours)
- ➢ Three band colour (three bands from the image will be rendered, each band representing the red, green or blue component that will be used to create a colour image)

Within both render types you can invert the colour output using the Invert colour map checkbox.
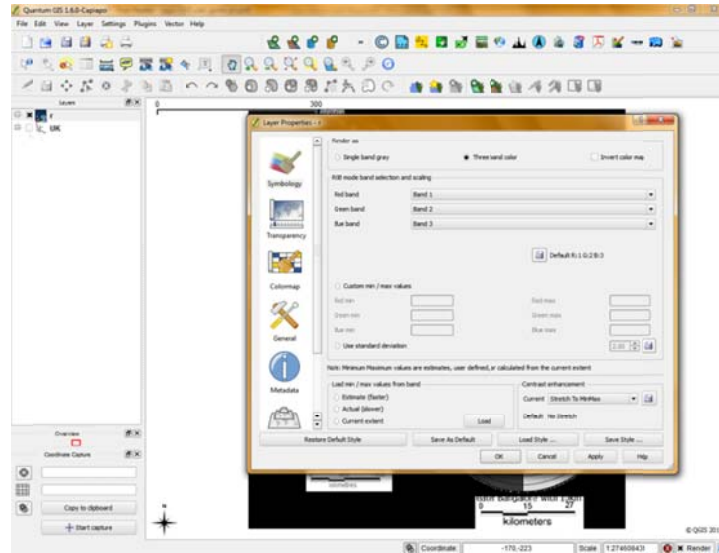
❖ QGIS has the ability to display each raster layer at varying transparency levels. Use the transparency slider to indicate to what extent the underlying layers (if any) should be visible though the current raster layer. This is very useful, to overlay more than one raster layer.Ex: elevation map overlaid by a classified raster map. This will make the look of the map more three dimensional.

❖ The Colormap tab is only available, when a single-band-rendering is selected within the tabSymbology

❖ Displays basic information about the selected raster, including the layer source and display name in the legend (which can be modified). The spatial reference system is printed here as a PROJ.4-string. This can be modified by hitting the change button.
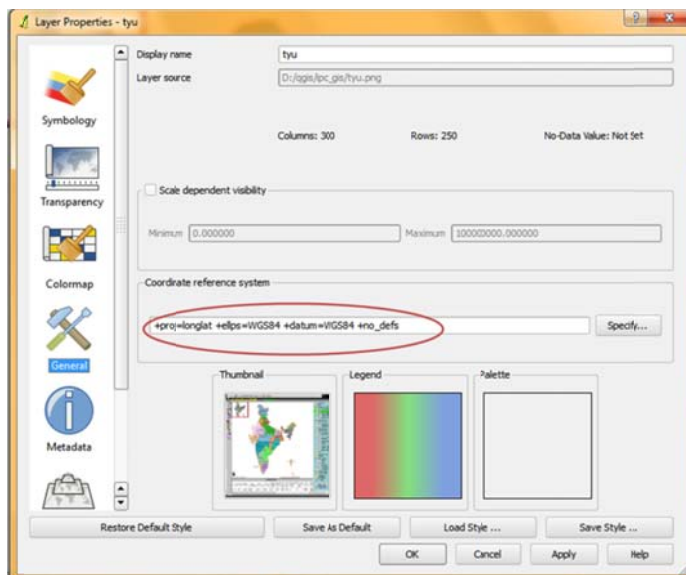
❖ Displays information about the raster layer, including statistics about each band in the current raster layer.

❖ Large resolution raster layers can slow navigation in QGIS. By creating lower resolution copies of the data (pyramids), performance can be considerably improved as QGIS selects the most suitable resolution to use depending on the level of zoom.Disadvantage with this is it alters the original data.

❖ The Histogram tab allows you to view the distribution of the bands or colours in the raster. The distribution of the histogram image can be saved for further references.
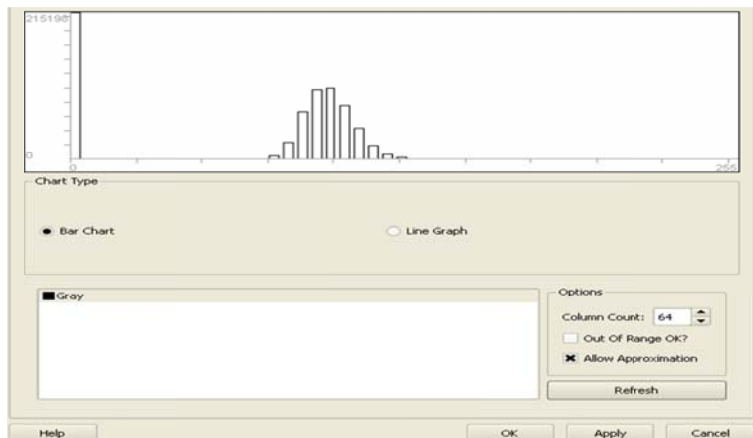
Raster Calculator

The Raster Calculator in the Layer menu allows performing calculations on basis of existing raster pixel values. The results are written to a new raster layer with a GDAL supported format. Ex from the raster layer of India classified image, to separate the Bangalore layer multiply the Bangalore boundary raster layer can results the classified output file from the whole.

Geo referencer:

Geo referencing usually refers to the method by which locations in the raster and vector GIS files are related to real earth-surface positions.To start geo referencing an unreferenced raster, we must load it using the button you can geo reference an image using already georeferenced image or entering GCPs manually. Here is the procedure for image to image georeferencing. The raster will show up in the main working area of the dialog. Once the raster is loaded, we can start to enter reference points.
Using the Add Point button, add points to the main working area and enter their coordinates
For this procedure you have two options:

a) Click on a point in the raster image and enter the X and Y coordinates manually

b) Click on a point in the raster image and choose the button from map canvas to add the X and Y coordinates with the help of a georeferenced map already loaded in the QGIS map canvas.

c) With the button, you can move the GCPs in both windows, if they are at the wrong place.

Continue entering points. You should have at least 4 points, and the more coordinates you can provide, the better the result will be. There are additional tools on the plugin dialog to zoom and pan the working area in order to locate a relevant set of GCP points.
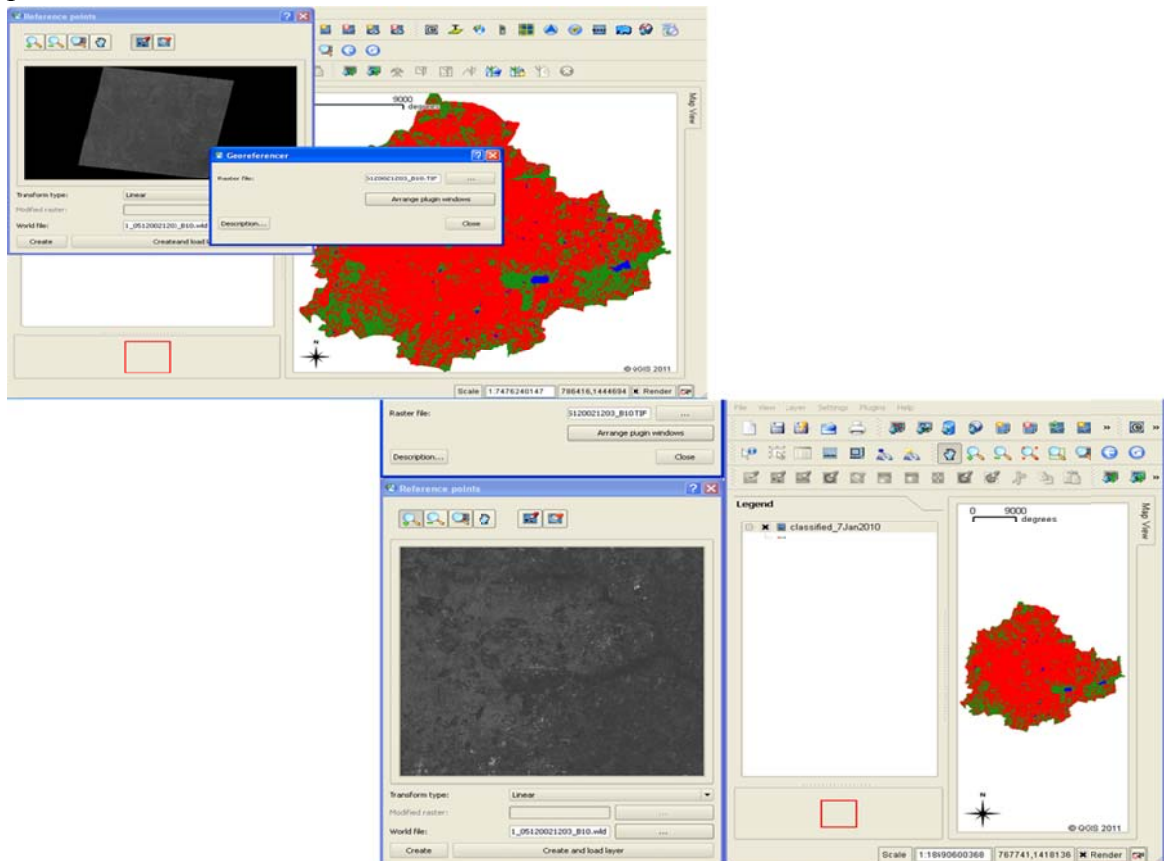
The figure showing the geometrically corrected image of Bangalore overlaid with Bangalore region vector file.
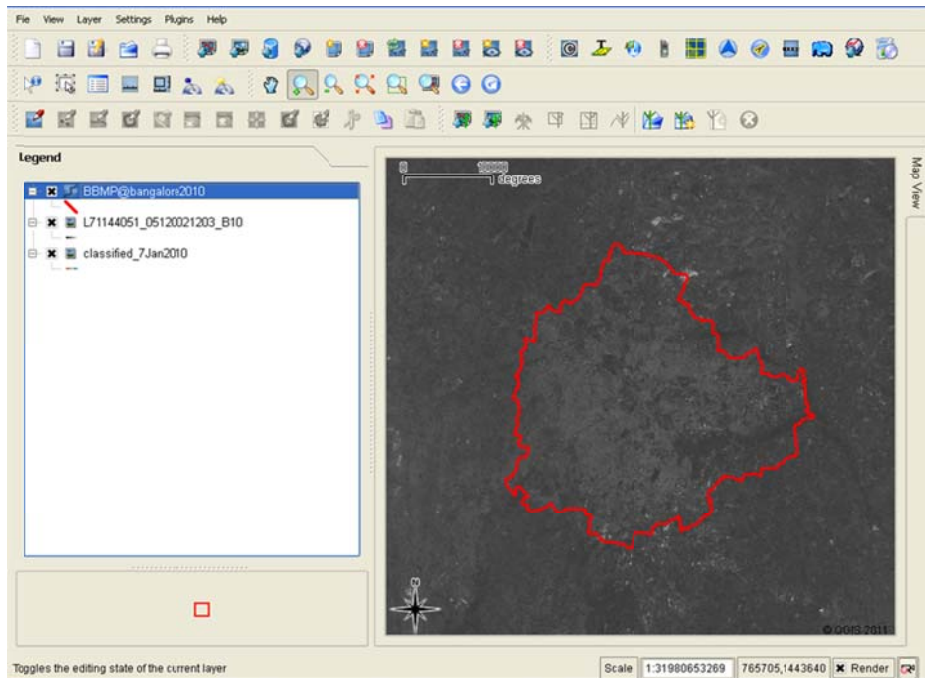
Working with OGC Data

OGC services are increasingly being used to exchange geospatial data between different GIS implementations and data stores. QGIS supports WMS and WFS as data sources. WMS-support is native; WFS and WFS-T is implemented as a plugin. The Open Geospatial Consortium (OGC) is an international organization with more than 300 commercial, governmental, nonprofit and research organisations worldwide. Important OGC specifications are:
– WMS - Web Map Service   – WFS - Web Feature Service   – WCS - Web Coverage Service   – CAT - Web Catalogue Service   – SFS - Simple Features for SQL   GML- Geography Mark-up Language

Digitizing

QGIS supports digitizing from a raster or vector layer. It support editing include shape files, Post GIS, and GRASS data. The attribute values in new table allow real, integer, and string data types. Point, line, and polygon features are supported. Attributes are entered as features are created. Editing Features include →Moving vertices or points, inserting new points where needed, deleting unnecessary points, deleting entire features, cut or copy features from one layer and paste to another, multiple features can be copied/pasted in one operation
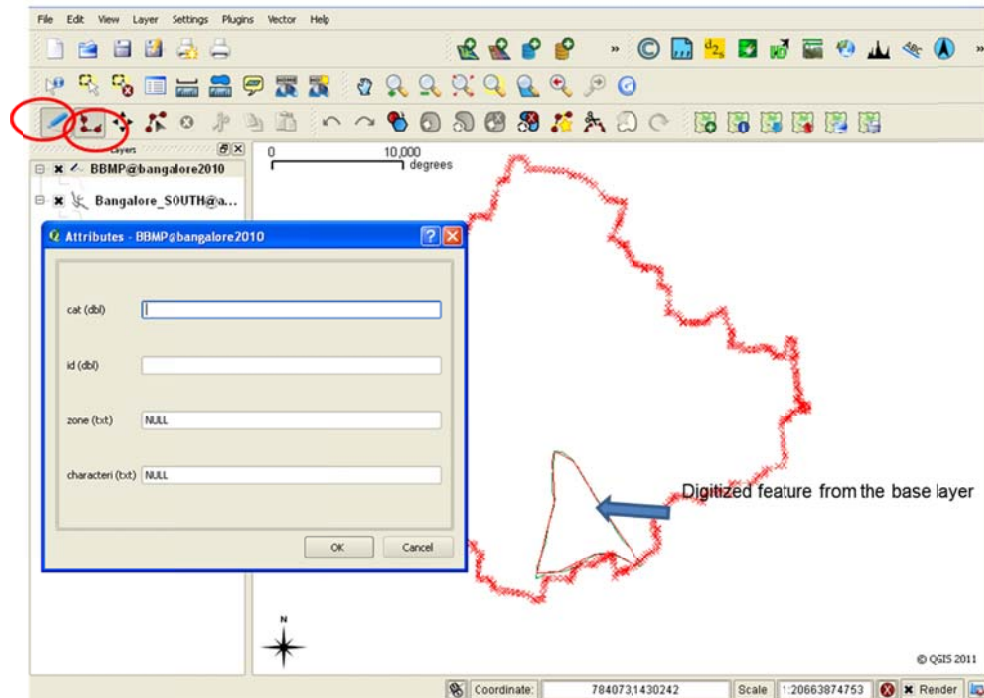
GRASS GIS Integration

The GRASS plugin provides viewing, editing, and geo processing of GRASS data. It provides access to GRASS GIS databases and functionalities. This includes visualization of GRASS raster and vector layers, digitizing vector layers, editing vector attributes, creating new vector layers and analyzing GRASS 2D and 3D data with more than 300 GRASS modules. GRASS data are stored in a directory referred to as GISDBASE. This directory often called grass data, must be created before you start working with the GRASS plugin in QGIS. Within this directory, the GRASS GIS data are organized by projects stored in subdirectories called LOCATION. Each LOCATION is defined by its coordinate system, map projection and geographical boundaries. Each LOCATION can have several MAPSETs. In order to analyze vector and raster layers with GRASS modules, you must import them into a GRASS LOCATION.

Creating a new GRASS LOCATION

Start QGIS and make sure the GRASS plugin is loaded. Click on the Open Mapset icon in GRASS toolbar to open the MAPSET wizard or plugins→Grass→open Mapset(if already created; if not create a new Mapset) .Select an existing GRASS database (GISDBASE) folder grassdata or create one for the new LOCATION using a file manager on your computer then click Next. Create a new MAPSET within an existing LOCATION or to create a new LOCATION altogether by using this wizard to. Click on the radio button Create new location. All the options in Grass tool bar will be highlighted only after opening the Mapset. . This is necessary, because new raster or vector layers created during analysis need

to be written to the currently selected LOCATION and MAPSET. The GRASS toolbox
provides the functionalities to work with Raster/ Vector data.
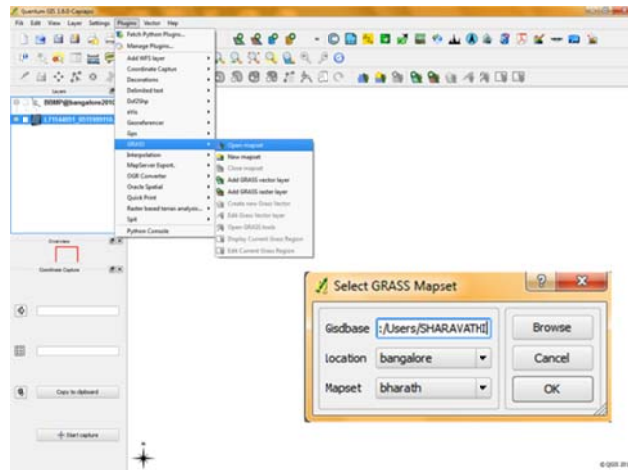
To add Grass raster layer click on Add Grass raster layer icon in the tool bar or
plugins→Grass→Add raster layer a dialogue box will be displayed with name of
Location, Mapset, raster images. Click the name of the raster file to be displayed and
click OK.
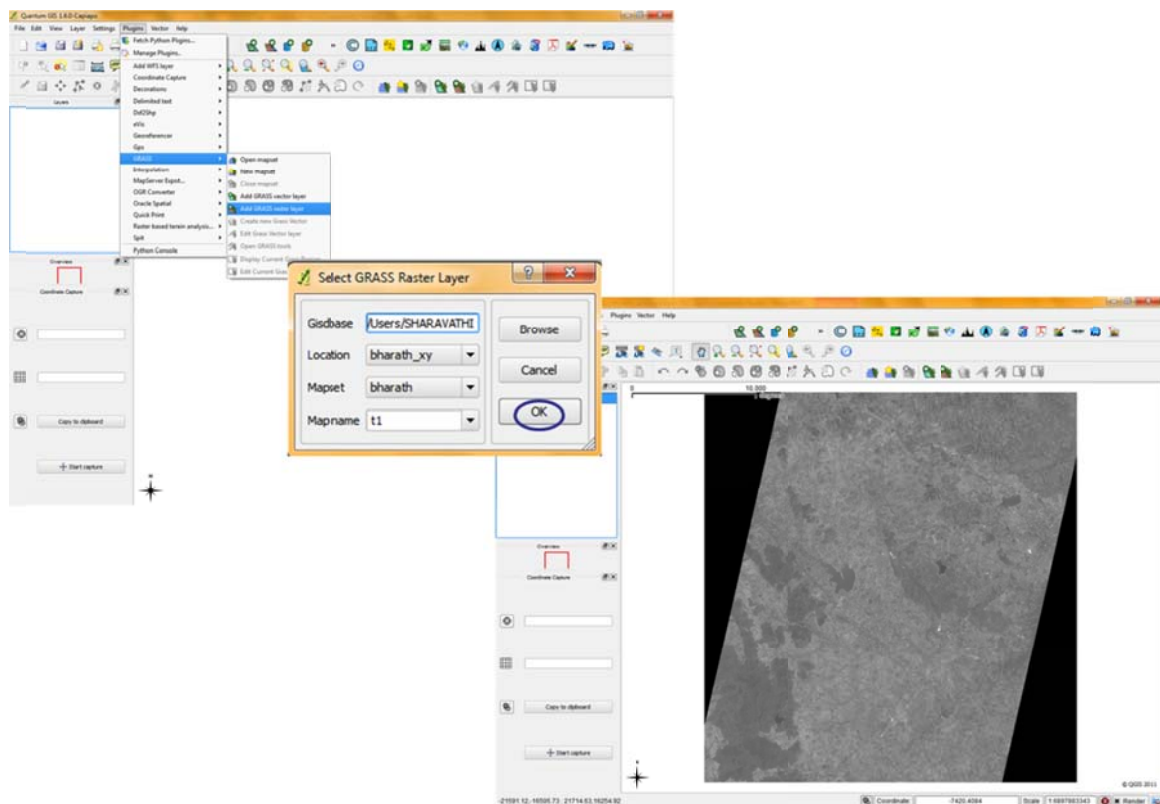
The Open GRASS Tools box provides GRASS module functionalities to work with data inside a selected GRASS LOCATION and MAPSET. Tool box shows the list of commands with description and modules available. If user want to work with specific command, by clicking the command will be executed by showing some options for input the data.
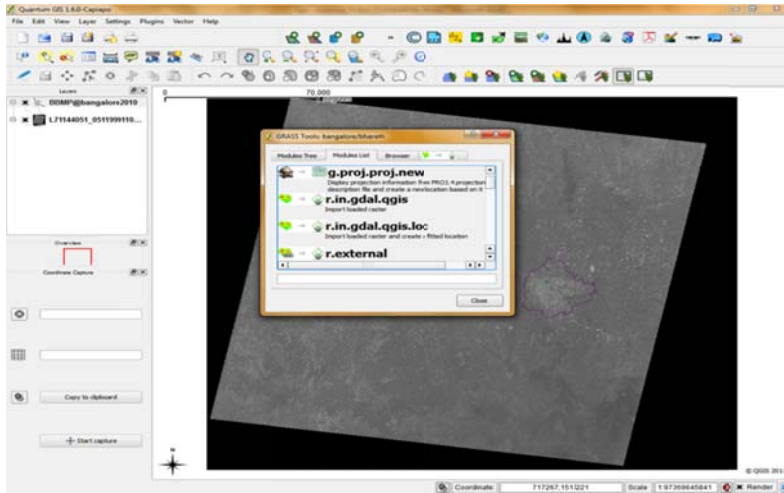
QGIS Plugins

QGIS allows many new features/functions to be easily added to the application with the help of plugins(Similar to Firefox add-ons). The Plugin Manager provides a resource to load or unload plugins. Many of the features in QGIS are actually implemented as either **core** or **external** plugins.

– **Core Plugins** are maintained by the QGIS Development Team and are automatically part of every QGIS distribution. They are written in one of two languages: C++ or Python.

– **External Plugins** are currently all written in Python. They are stored in external repositories and maintained by the individual authors. They can be added to QGIS using the Plugin Installer.
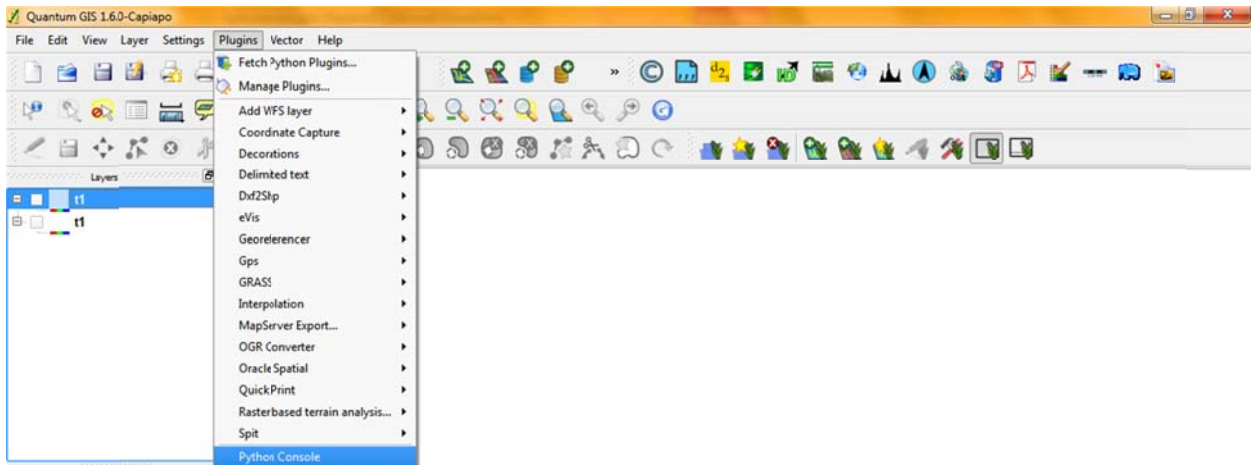
GPS Plugin

GPS, the Global Positioning System, is a satellite-based system allows to find the exact position anywhere in the world. It is used as an aid in navigation. It provides the information of latitude, longitude and elevation by using the signals from the satellites. Waypoints, routes and tracks are the three basic feature types in GPS data. Most receivers also have the capability to store locations (known as waypoints), sequences of locations that make up a planned route and a track log or track of the receivers movement over time. There are different data formats to store GPS data; Qgis supports a standard interchange format (GPS eXchange format), that can contain any number of waypoints, routes and tracks in the same file. To load a GPX file, the plugin should be loaded. To load plugins click on Plugins→ Plugin Manager→GPS Tools.

When this plugin is loaded a button with a small handheld GPS device will show up in the toolbar.
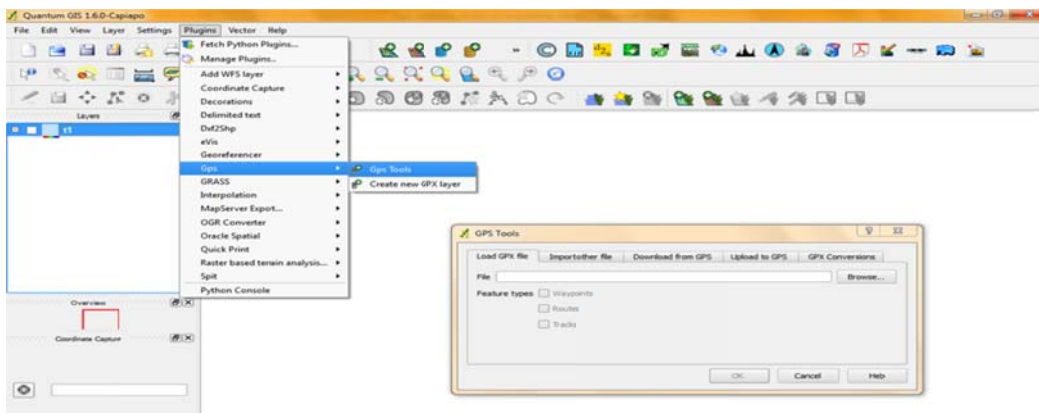
OGR Converter Plugin

This is very useful plugin; which provides the ability to convert vector data from one OGR-supported vector format to another (for example to create KML file from shape file). The plugin requires a few parameters to be specified before running:

– **Source Format/Dataset/Layer**: Enter OGR format and path to the vector file to be converted

– **Target Format/Dataset/Layer**: Enter OGR format and path to the vector output file

Print Composer

The Qgis print composer provides capability of printing maps in different formats. It allows adding elements such as the QGIS map canvas, legend, logo inclusion, north arrow; image formats like PNG/SVG/PDF support, adjustable drawing scale, separate DPI settings, basic shapes, and text labels. Allows to specify the size, group, alignment and position of each element and adjusting the properties. The layout can be printed or exported to image formats; Postscript, PDF and also provides an option to save the layout as template and load it again in another session.

Icon	Purpose	Icon	Purpose
	Load from template		Save as template
	Export to an image format		Export as PDF
	Export print composition to SVG		Print or export as Postscript
	Zoom to full extend		Zoom in
	Zoom out		Refresh view
	Add new map from QGIS map canvas		Add Image to print composition
	Add label to print composition		Add new legend to print composition
	Add new scalebar to print composition		Add basic shape to print composition
	Add arrow to print composition		Add attribute table to print composition
	Select/Move item in print composition		Move content within an item
	Group items of print composition		Ungroup items of print composition
	Raise selected items		Lower selected items
	Move selected items to top		Move selected items to bottom
	Align selected items left		Align selected items right
	Align selected items center		Align selected items center vertical
	Align selected items top		Align selected items bottom

Help (Useful websites for Quantum GIS references)

http://lists.osgeo.org/mailman/listinfo/qgis-user

https://lists.fossgis.de/mailman/listinfo/fossgis-talk-liste

http://lists.osgeo.org/mailman/listinfo/qgis-developer

http://lists.osgeo.org/mailman/listinfo/qgis-commit

http://lists.osgeo.org/mailman/listinfo/qgis-trac

http://lists.osgeo.org/mailman/listinfo/qgis-community-team

http://lists.osgeo.org/mailman/listinfo/qgis-edu

Geographic Resources Decision Support System – an Open Source GIS

* Centre for Ecological Sciences, Indian Institute of Science
+ Centre for Sustainable Technologies, Indian Institute of Science
# Salim Ali Centre for Ornithology and Natural History, Dehradun

T. V. Ramachandra (*,+), Uttam Kumar (*), S.N. Prasad (#) &Vaishnav B(*)
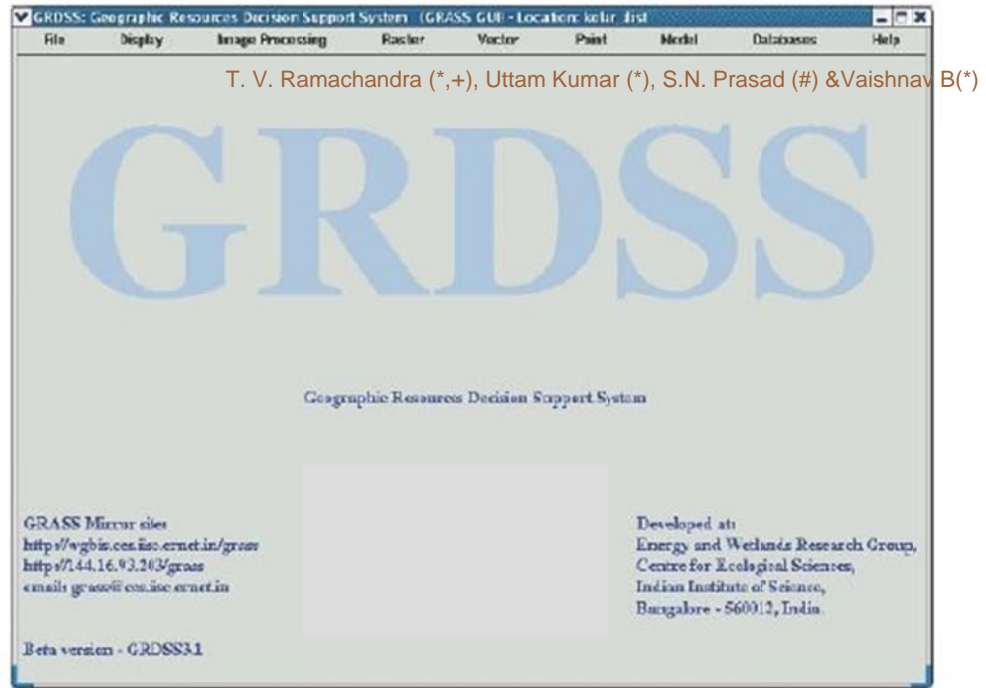
Figure 1: GRDSS

The process of developing the GUI also involved undertaking the task of identifying the commands and implementation of the interfacing code. The layout of the interface had to be altered to suit the end user's perspective. Figure 2 shows the structural diagram of the GRDSS graphical user interface.

Fully structured and matured open source spatial and temporal analysis technology seems to be the official carrier of the future for planning of the natural resources especially in the developing nations.
This technology has gained enormous momentum because of technical superiority, affordability and ability to join expertise from all sections of the society. Sustainable development of a region depends on the integrated planning approaches adopted in decision making which requires timely and accurate spatial data. With the increased developmental programmes, the need for appropriate decision support system has increased in order to analyse and visualise the decisions associated with spatial and temporal aspects of natural resources. In this regard Geographic Information System (GIS) along with remote sensing data support the applications that involve spatial and temporal analysis on digital thematic maps and the remotely sensed images. Open source GIS would help in wide scale applications involving decisions at various hierarchical levels (for example from village panchayat to planning commission) on economic viability, social acceptance apart from technical feasibility. GRASS (Geographic Resources Analysis Support System, http://wgbis.ces.iisc.ernet.in/grass) is an open source GIS that works on Linux platform (freeware), but most of the applications are in command line argument, necessitating a user friendly and cost effective graphical user interface (GUI). Keeping these aspects in mind, Geographic Resources Decision Support System (GRDSS) has been developed with functionality such as raster, topological vector, image processing, statistical analysis, geographical analysis, graphics production, etc. This operates through a GUI developed in Tcltk (Tool command language / Tool kit) under Linux as well as with a shell in X-Windows. GRDSS include options such as Import /Export of different data formats, Display, Digital Image processing, Map editing, Raster Analysis, Vector Analysis, Point Analysis, Spatial Query, which are required for regional planning such as watershed Analysis, Landscape Analysis etc. This is customised to Indian context with an option to extract individual band from the IRS (Indian Remote Sensing Satellites) data, which is in BIL (Band Interleaved by Lines) format. The integration of PostgreSQL (a freeware) in GRDSS aids as an efficient

Address for
Correspondence :
Dr. T.V. Ramachandra
Energy & Wetlands
Research Group
Centre for Ecological
Sciences
Indian Institute of
Science,
Bangalore – 560 012,
INDIA

Tel :
91-80-23600985/22932506/
22933099

Fax :
91-80-23601428/23600085/
23600683[CES-TVR]

E-mail :
cestvr@ces.iisc.ernet.in
cestvr@hamsadhvani.sers.
iisc.ernet.in
energy@ces.iisc.ernet.in

URL :
http://ces.iisc.ernet.in
/energy/index.htm

database management system.

Introduction

GIS (Geographic Information System) is a system of hardware, software and procedures to facilitate the management, manipulation and analyses of spatial-temporal data. Its application is wide ranging from micro level to macro level planning. However this boundless capabilities are limited by ones ability tovisualise its implications. It has become an important component in regional planning with a pivotal role in major decisions related to earth resources. This includes urban sprawl analysis involving growth and migration of population, location of industries, employment activities, basic infrastructure (water, electricity, etc.), transportation (railway and road networks), drainage system (river basin quality, etc.), town planning, etc.

With the growing incidence of ecological and environmental impacts, GIS has been used to analyse the impact of human activities on the ecosystems. This is possible due to the options to display and analyse changes in the geo-referenced temporal and spatial data. This helps to solve complex problems regarding planning and management of resources and also biophysical modeling. The data input includes spatial, statistical and thematic data derived from a combination of existing maps, aerial photographs, interpretation of remotely sensed images and secondary data from the government agencies (census data, land use data, etc.). Digital image processing technique aids in restoration and rectification of data, segmentation of data and also in visualisation. The integration of spatial and temporal technology help in addressing challenges faced by the environment. The integration of GIS with remote sensing data offers enhanced capability for inventorying, mapping, monitoring and modeling to understand many environmental processes. This integration also helps in using remotely sensed image to update the GIS and also maps. The GIS thematic data and attributes are used to guide image classification. GRASS, a GIS based open source software distributed under the GNU GPL (General Public License) integrates these two technologies resulting in numerous advantages.

GRASS was originally developed by the U.S. Army Construction Engineering Research Laboratories, a branch of the US Army Corp of Engineers, as a tool for land management and environmental planning by the military. GRASS is a raster/vector GIS, image processing system, and graphics production system (http://wgbis.ces.iisc.ernet.in/grass). GRASS contains over 350 programs and tools to render maps and images on monitor and paper; manipulate raster, vector, and sites data; process multi spectral image data; and create, manage, and store spatial data. GRASS has the options to interface with printers, plotters, digitizers, and databases to develop new data as well as manage existing data. GRASS uses both intuitive windows interface as well as command line syntax for ease of operations and is supported by developers and users worldwide. Numerous sub modules existed in GRASS but most of them are in command line arguments. The command line syntax for GRASS was cumbersome and time consuming for individuals, who were not exposed to computer programming and did not possess any programming skills. In order to overcome this difficulty, it was the necessity for an economically viable, technically superior and a more flexible spatial and temporal analyses tools that led to the development of a user friendly graphical user interface (GRDSS) for GRASS GIS along with the database component as depicted in figure 1.
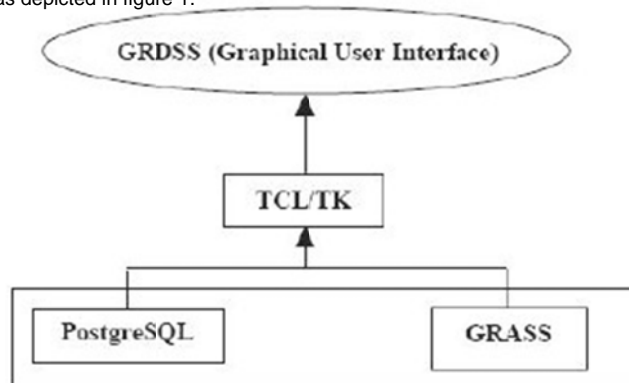
Figure 2: GRDSS - GUI structural diagram

Methodology

The current development in GRDSS in the allied technologies of remote sensing and GIS is to build an integrated spatial decision support system (SDSS). Recent advances in the development of graphical user interface based programming have enabled rapid prototyping, testing, and application development. Various GIS packages such as MapInfo (http://mapinfo.com), Idrisi (http://www.idrisi.clarku.edu), Geomedia (http:// w ww.intergraph.com) etc were comparatively studied as per the set criteria listed in table 1.

Sl. No.	Characteristics	MAPINFO	IDRISI	GEOMEDIA	GRDSS
1.	Menus and user interface	Y	Y	Y	Y
2.	Command line syntax	N	N	N	Y
3.	Data Import/Export	Y	Y	Y	Y
4.	IRS LISS–MSS Data extraction (Data in BIL and BSQ format to individual spectral bands)	N	Y	N	Y
5.	Map Registration & projection	Y	Y	Y	Y
6.	Digitization (vector data)	Y	Y	Y	Y
7.	Coordinate transformation	N	Y	Y	Y
8.	Buffer generation	Y	Y	Y	Y
9.	Digital Image analysis	N	Y	Y	Y
10.	Map Algebra	N	Y	Y	Y
11.	Terrain Modelling (DTM)	Y	Y	Y	Y
12.	Distance calculation	Y	Y	Y	Y
13.	Database support	Y	Y	Y	Y
14.	Error Explanation	Y	Y	Y	Y
15.	Help	Y	Y	Y	Y
16.	Online manuals and tutorials	Y	Y	Y	Y
17.	Proprietary software	Y	Y	Y	N
18.	Open source	N	N	N	Y
19.	Free software	N	N	N	Y
20.	Mailing List	N	N	Y	Y
21.	Digital Elevation Model (DEM)	Y	Y	Y	Y
22.	Fly through and animation	N	Y	Y	Y
23.	Fusion	N	N	N	Y

Y = features present
N = features not present

Table 1: Comparative analyses of GIS softwares

The limitations of the commercial GIS softwares were also considered while implementing the GRDSS to offer optimal performance. The following feature characterises GRDSS development.

i. System requirements - The GRDSS GUI sources can be compiled on all PC's running LINUX. It is a full system with X Window graphical user interface. Minimum PC can be 486 (or better), with "LINUX". The Linux- system with a three-button mouse, 32 MB RAM (64 MB recommended) and a hard disk space of 250MB (beside DOS/Windows partition). GRDSS binaries require only 45-50 MB and to compile additional 150 MB disk space is required (to store the sources). The GRDSS source code can be compiled using a GNU "C" compiler.

ii. User friendly interface and menu organisation - The user interface is designed to suit to the users in India and those who are not much exposed to programming. The interface is customized and simplified comparable to proprietary softwares in the market. The menu is organised sequentially in a user-friendly environment that satisfies both specific and common users.

iii. Functional capabilities - GRDSS is equipped with the functional capabilities such as data import/export in different formats (including the IRS data format), map analysis, data modeling, buffer generation, spatial and non spatial statistical calculation, digital image processing (spatial, spectral and temporal data analysis), raster, vector and point analysis and query, map calculator, terrain modeling, erosion modeling, report generation and can also handle spatial and non-spatial database (both spatial data and attribute information), etc.

iv. Help facilities – GRDSS extends all the help facilities present in GRASS. All the commands with their functions and options are present in help to get started with the GRDSS. There is a mailing list and mailing list archive where users can post their doubts and can discuss through e-mails worldwide. The help facilities andthe documents available on the GRASS official website provide enough information to start with the software.

v. Error explanation and manual support - GRDSS is equipped with error explanation, which automatically renders error messages through the mail facility in Linux. Beside many sample data sets are available on the GRASS official website, including online tutorials and manuals.
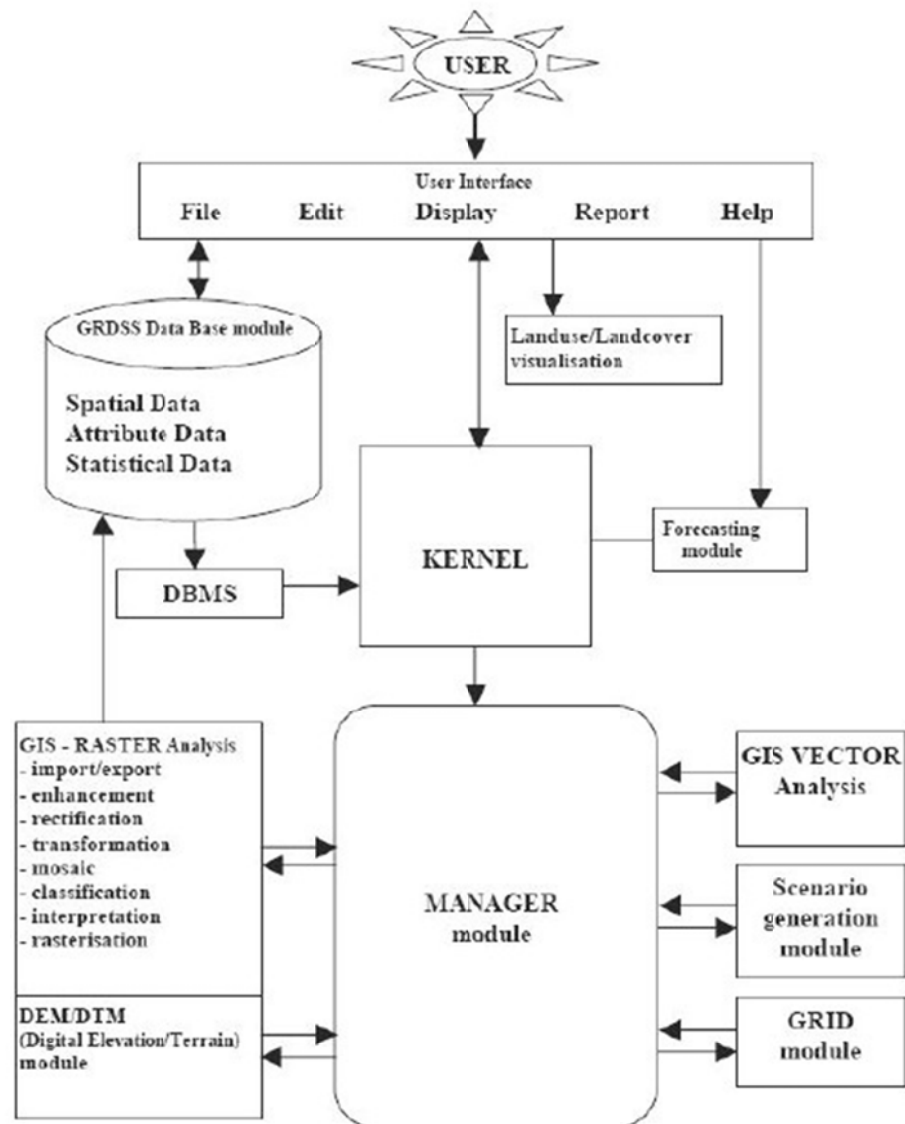
Figure 3: The GRDSS design and conceptual diagram.

Overview and description of GRDSS

GRDSS supports all basic features required by any GIS package.

  i. It supports a wide range of data Import/Export that are also supported by other standard GIS including IRS (LISS3 and PAN) data extraction and DTEM/DTED (Digital Terrain Elevation Model / Digital Terrain Elevation Data) extraction.

ii. It has capabilities for displaying the raster/vector images. GRDSS can further interact with commercial plotters and printers and is capable of color management.

iii. GRDSS has the image processing capabilities (rectification, enhancement, transformation, classification, data analysis, etc.).

iv. GRDSS has the capability of manipulating raster, vector and point data. This includes digitization, topology creation, map overlay, arithmetic operations, DEM generation, decision support, distance, area and statistical report, etc.

v. It can perform wetland analysis, terrain modeling with slope, aspect, relief, profile, cost between two locations to analyze terrain data, runoff-modeling, etc.

vi. Thematic data preparation, histogram generation, and charts of different formats.

vii. Database (PostgreSQL) management system.

viii. Online tutorials with help manuals and sample datasets. Figure 4 depicts the GRDSS hierarchical menu arrangement.
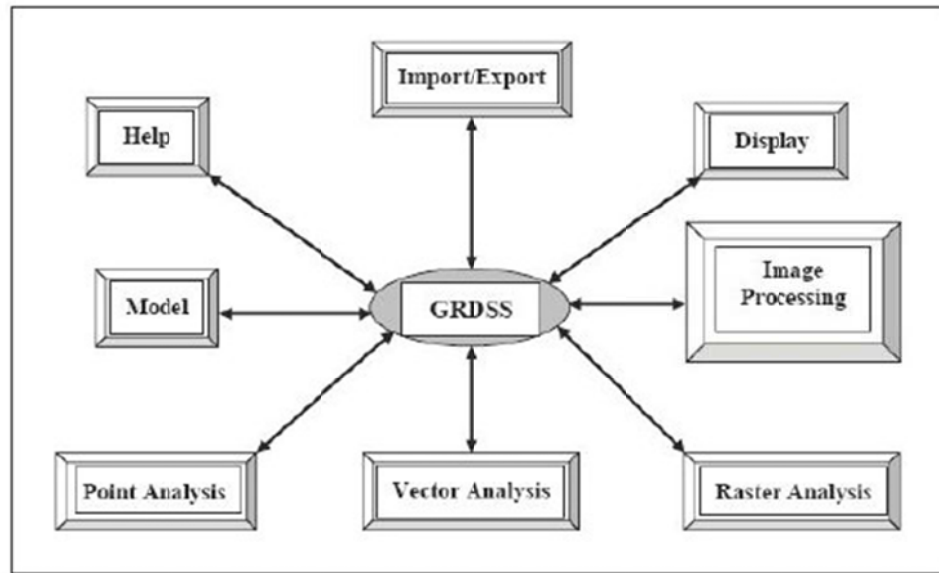
Figure 4: GRDSS hierarchical menu arrangement.Table 2: Functions of GRDSS modules

Import/Export (supported data formats)	Point (site) data formats
Raster formats – IRS Data format (.bil – band interleaved by lines), Binary, Bsq (band sequential format), tiff (Tagged Image File Format), geotiff, img (image file), jpeg (joint photographic expert group), ppm (Portable Pixel Map), png (Portable Network Graphics), pbm (Portable Bit Map), pgm (Portable Gray Map), utm (Universal Transverse Mercator) sunraster, arcinfo, Esri, Erdas, ascii, tang, gridatb, Landsat tape formats, Idrisi etc. **Vector formats –** Autocad (dxf-Drawing eXchange Format), Atlas, shp (shape file), ascii, Esri, Arcinfo (dlg-Digital Line Graphs), Mapinfo (mif-MapInfo Interchange File), Sdts (Spatia Data Transfer Standard), Moss, xyz etc. **Site formats –** Ascii, Shp (shape file), Dbf (data base file), Esri, etc. **DEM/DTED extraction module.**	• Basic information attributes. • Quadrat counts, univariate statistics. • Buffer generation • Reprojection • Triangulation • Interpolation • Rainfall data preparation • Conversion tools – point to raster, vector, grid 3d etc. **Image Processing** • Segmentation • Rectification and restoration • Georegistration (geocorrection) • Transformation • Image enhancement • False color composite • Signature development • Clustering • Classification (land use analysis) • Vegetation indices (land cover analysis) **Models** • Wetland • Watershed • Landscape • Wildfire • Erosion tools

Vector format	Database (PostgreSQL)
• Digitize • Basic information about layers attributes, statistical calculation. • Report • Distance operators • Length, Perimeter, areas etc. • Reprojection • Grid generation • Conversion tool – vector to raster, vector to point, areas to line etc.	• Import /Export attribute data from different formats. • Select database/select table • Display table • Modify (add/delete) values • Query table • List table, list columns • Query raster, vector, sites data • Display raster, vector, sites data
Raster Analysis	Display
• Image information, report, area, category range, and statistical calculations. • Map algebra, regression analyses. • Distance operators, buffer creation. • Coordinate conversion tools • Topographic variables • Digital elevation model • Decision support • Grid generation, reclass, rescale, recode, color management. • Subimage extraction • Change resolution • Digitise • Conversion tools raster to vector lines, vector areas, point etc. • Cropping • Mosaic	• Display raster, vector, sites, 3d map. • Color management. • Display HIS, RGB, north arrow, frame, legend, map title. • Generate 3d view. • Animation. • Fly through. • Draping multiple raster, vector, site data. • Zoom, pan. • Generate histogram, pie chart, bar chart. • Select font size, symbols, appearance, style, width, and orientation. • Colors, line and polygon fill. • Placing text strings, map title, labels on raster, vector and point data.

Table 2: Functions of GRDSS modules
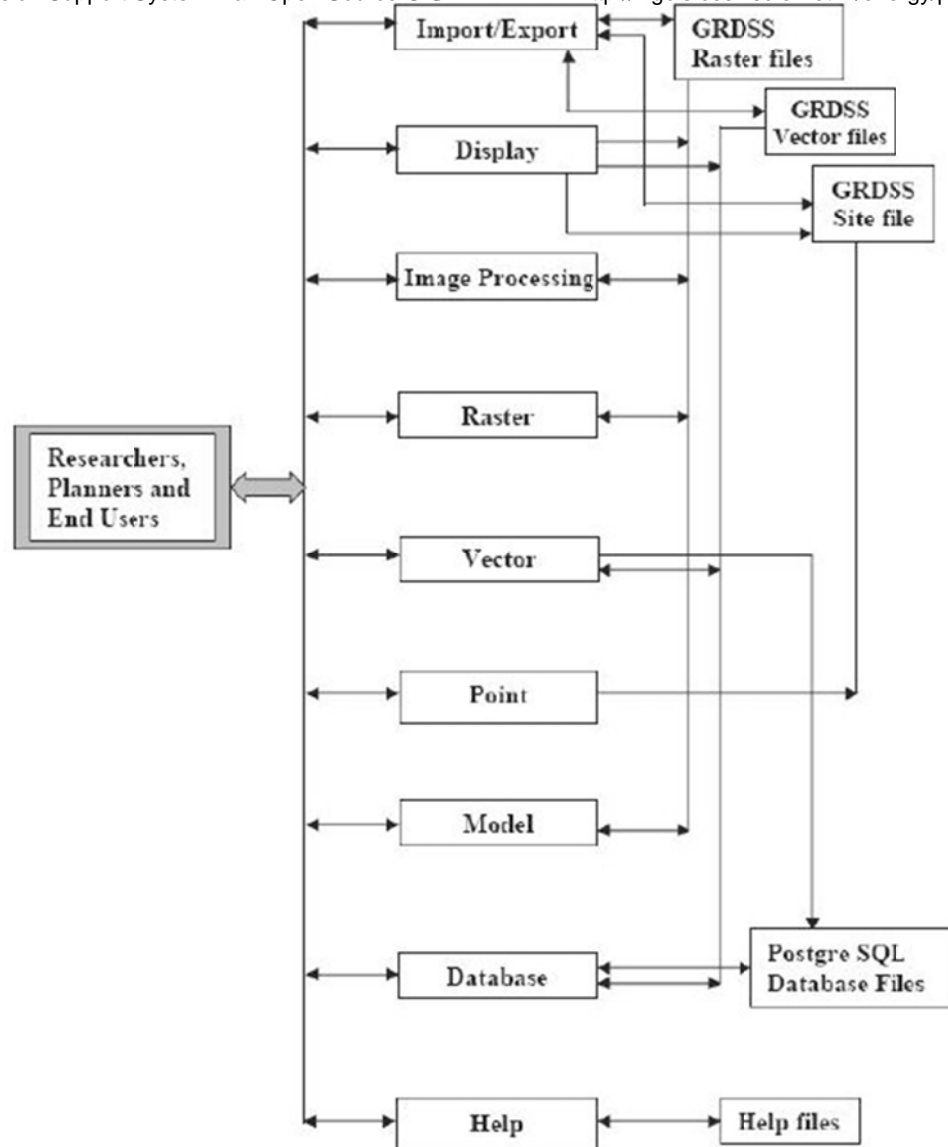
Figure 5: GRDSS data flow diagram.

Database connectivity

GRDSS uses PostgreSQL (a freeware) as a database for data storage, data retrieval, query and manipulation of both spatial and attribute information. It is the most advanced open source database server used primarily for data storage and retrieval, and complex data analysis. PostgreSQL is a sophisticated object-relational database management system (ORDBMS). An ORDBMS is an extension of the more traditional relational database management systems (RDBMS). For the advanced developer, PostgreSQL even supports extensibility of its data types and procedural languages. PostgreSQL has comprehensive SQL support, referential integrity, MVCC or Multi- Version Concurrency Control to avoid unnecessary locking that increases the reliability of the database by logging changes before they are written to the database.

Applications

Land cover analysis: This has been done by computing both slope based and distance based vegetation indices as the district had varying extent of vegetation covers. The result of land cover analysis is listed in table 3.

GRDSS provide a support for spatial data as well as attribute data by integrating GIS, image processing, GPS (global positioning system) and other techniques.

Kolar	Cutoff point	Area in hectares		Area in percentage	
		Non-vegetation	Vegetation	Non-vegetation	Vegetation
RATIO	1	444659	379184	53.98	46.03
NDVI	0	418783	401654	51.05	48.96
RVI	1	82824.8	68955.7	54.56	45.43
NRVI	0.99	401647	369692	52.08	47.93
TVI	0.71	444099	379183.2	53.94	46.06
CTVI	0.707	418203	379183.2	52.46	47.55
TTVI	0.73	364415.1	305628	54.39	45.61
PVI	26.00	446960	376808	54.26	45.74
PVI1	7.00	453558	369512	55.10	44.89
PVI2	104	435853	387416	52.95	47.05
PVI3	3850	448812	375025	54.47	45.53
DVI	45	452326	371438	54.91	45.09
AVI	1	468967	354879	56.92	43.08
SAVI	0.02	72405.7	55919.4	56.42	43.58
TSAVI1	0.0199	409483	390422	51.19	48.81
TSAVI2	0.024	453360	370482	55.03	44.97
MSAVI1	0.17	451548	372294	54.81	45.19
MSAVI2	0	444626	355103	55.60	44.41
WDVI	42.01	450160	373627	54.65	45.35

Table 3: Vegetation Indices

Land use analysis: GRDSS through its capability of analyzing the statistical, raster and vector data with options for map overly, map algebra and other operators aid in analyzing the landuse pattern. Both supervised and unsupervised classification approaches were tried to identify landuse categories in the district using Gaussian maximum likelihood classifier (GMLC). The level of accuracy in GMLC was 94.67% compared to unsupervised classifier (78.07%). The composition of land use categories (agriculture, forest, plantation, built-up and wasteland) are listed in the table 4.

Categories	Supervised		Unsupervised	
	Area (in ha)	Area (%)	Area (in ha)	Area (%)
Agriculture	233519	28.34	222416	27.00
Builtup	131468	15.96	70970	8.62
Forest	68300	8.29	85295	10.35
Plantation	70276	8.53	84716	10.28
Waste land	320284	38.88	360450	43.75

Accuracy Estimation   Table 4: Land use details of Kolar district

Accuracy estimation in terms of producer's accuracy, user's accuracy, overall accuracy and Kappa coefficient were calculated after generating confusion matrix for supervised classification (table 5) and unsupervised classification (table 6).

Table 5: Error Matrix Resulting from Classifying Training Set Pixels.

Classification data	Agriculture	Builtup	Forest	Plantation	Waste land	Row Total
Agriculture	42	1	0	0	0	43
Built up	0	16	0	0	0	16
Forest	0	0	20	3	0	23
Plantation	0	0	2	34	0	36
Waste land	2	0	0	0	30	32
Column Total	44	17	22	37	30	150

Classification data	Agriculture	Builtup	Forest	Plantation	Waste land	Row Total
Agriculture	33	0	0	0	1	34
Built up	0	17	0	0	4	21
Forest	0	0	38	13	0	51
Plantation	0	0	0	28	0	28
Waste land	8	8	5	2	30	53
Column Total	41	25	43	43	35	187

Table 6: Error Matrix for Unsupervised Classification.

The producer's accuracy, user's accuracy corresponding to the various categories and overall accuracy results obtained are summarized in table 7.

Category	Supervised Classification			Unsupervised Classification		
	Producer's Accuracy	User's accuracy (%)	Overall accuracy (%)	Producer's accuracy (%)	User's accuracy (%)	Overall accuracy (%)
Agriculture	95.45	97.67		80.49	97.06	
Builtup	94.11	100.00		68.00	80.95	
Forest	90.90	86.96	94.67	88.37	74.51	78.07
Plantation	91.89	94.44		65.16	100.00	
Waste land	100.00	93.75		85.71	56.60	

$\hat{k}$

Table 7: Producer's accuracy, user's accuracy and overall accuracy.

A value was computed (0.931577) which is as an indication that an observed classification is 93 percent better than one resulting from a chance.

GRDSS was also used to study the watershed status in Kolar district along with land use and land cover analyses. The digital elevation model generated for the district is depicted in figure 6. Draping a land use data on the DEM enabled to assess the status of the catchment of major water bodies. The temporal analysis of Kolar district provided a picture of land use changes. The analysis indicates the unplanned developmental activities without proper planning have lead to the conversion of large scale productive land to waste land (43.75%).

Watershed mismanagement and conversion of waterbodies for various anthropogenic activities (agriculture, buildings, etc.) were the prime reasons for the depletion of the water table along with increasing ecological problems such as salinization, runoff, etc.

agriculture
builtup
forest
plantation
waste land
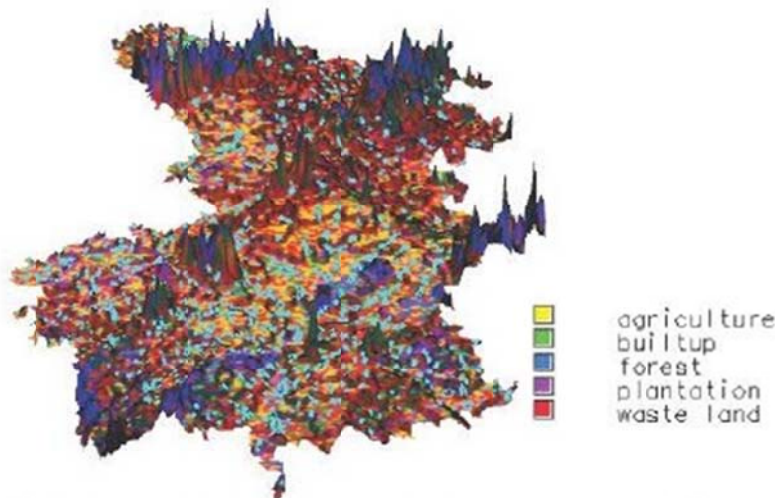
Figure 6: 3d-view of the Kolar district with landuse map draped on the image.

The land cover and land use analyses showed NRVI, CTVI, PVI1 results are comparable to GMLC in terms of percentage area. The assessment of the watershed and holistic approaches in landuse planning considering the actual ground condition could help in ensuring the sustainable development.

GRDSS a open source GIS with capabilities such as raster analysis, vector analysis, image processing, map algebra and other functionalities is comparable to any commercial software available in the market for various applications including natural resource management. An economical (free ware) with user friendly GUI's, GRDSS is hoped to penetrate in all sections of decision making and contribute to the sustainable development of India.

Summary

References :

1.GRASS official website :http://grass.itc.it
2.GRASS mirror site in India :http://wgbis.ces.iisc.ernet.in/grass/welcome.html
3.GRASS mailing list and mailing list archive :http://wgbis.ces.iisc.ernet.in/grass/support.html
4.GRASS software download page :http://wgbis.ces.iisc.ernet.in/grass/download.html
5.Requirements to compile GRASS GIS :http://wgbis.ces.iisc.ernet.in/grass/grass5/source
/REQUIREMENTS.html
6. Mapinfo :http://www.mapinfo.com/location/integration
7. Idrisi :http://www.idrisi.clarku.edu
8. Geomedia :http://www.intergraph.com/dynamicdefault.asp

E-mail | Sahyadri| ENVIS | Energy | GRASS | CES | IISc | E-mail

The Geospatial Resource Portal

Engaging Web for better administration

Home > **GIS Development** > **The Global Geospatial Magazine** > **2008** > **December**

Home | About Us | Site Map

Dr S Narendra Prasad
Senior Principal Scientist,
Salim Ali
Centre for Ornithology and
Natural
History, Deccan Regional
Station,
Hyderabad and Secretary,
OSGeo
snarendra.prasad@gmail.com

Santosh Gaikwad
IT Officer , Salim Ali Centre
for
Ornithology and Natural
History
Deccan Regional Station ,
Hyderabad
santosh@osgeo.in

Dr KS Rajan
Associate Professor,
International
Institute of Information
Technology
Hyderabad and Treasurer,
OSGeo
rajan@iiit.ac.in

India has been witnessing tremendous growth, the like of which has not been observed since Independence. The rapid pace of development has impacted practically every aspect of environment and the common man.

It has, therefore, become imperative to put in place effective governance systems with major emphasis on efficiency, accessibility and transparency and at the same time being able to understand the problems and issues faced by the society. Whether it is delivery of proper health care to the citizens, access to safe drinking water, social infrastructure like schools and other quality of life (QoL) concerns of the residents, the knowledge and understanding of geography or 'location' plays an important role in making the right decisions by the respective agencies or departments of the government machinery.

The governance systems, therefore, would require accurate and timely information and data, which is location specific in nature.

A WEB GIS FOR RAJAHMUNDRY

GIS is an indispensable tool to provide a spatial data infrastructure (SDI) for implementing e-governance. A GIS has been developed for the Rajahmundry Parliamentary Constituency, East Godavari district, Andhra Pradesh, which focuses on both the urban and rural regions, integrates data from multiple sources - remote sensing imagery, GPS surveys and field studies - and brings everything together on to an Open Source GIS platform for the development of a Spatial Decision Support System for civil and public administration from a desktop to Web enabled GIS. A Web GIS has been created both in English as well as in the local language, Telugu.

The work is unique in that multiple stakeholders were involved at various stages of development of the GIS. First, Member of Parliament V Aruna Kumar readily saw the possibilities and potential of developing such a system for the benefit of common man and funded the project through the MPLADS scheme.

The collector of East Godavari district, in turn, appreciated the merits of implementing such a project and facilitated its administration and implementation. The OSGeo India chapter endorsed the project and the Salim Ali Centre for Ornithology and Natural History executed the project. M.Sc students of Adi Kavi Nannaya University, Rajahmundry, were trained to carry out the project work in part fulfillment of their course requirements.
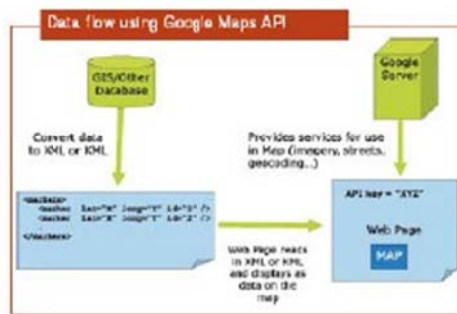
A PIGGYBACK RIDE ON GOOGLE MAPS APIS

Google Maps provides a highly responsive, intuitive mapping interface with embedded, detailed street and aerial imagery data. Google Maps provides not only the map, satellite image or a hybrid of both but also a range of operations on the map including zooming, panning, information pop-ups and overlays. Google Maps API provides an interface into these operations through JavaScript objects. The GIS application for Rajahmundry parliamentary constituency has been developed using Google Maps API.

A spatial database was developed for it in PostgreSQL/ PostGIS database. The advan- GIS DEVELOPMENT 50 Open Source Engaging Web for better administration DECEMB E R 200 8 Data flow using Google Maps API tage of creating spatial database for Rajahmundry was having all the data in a central database including well defined privileges which makes it possible to extend the more standard SQL queries with spatial queries.

DATASETS USED FOR RAJAHMUNDRY PARLIAMENTARY CONSTITUENCY
The following shape files were used while developing the spatial database for Rajahmundry Parliamentary
Constituency.
Point data
Bank and ATM centres, cemeteries, bus stations, police stations, places of worship, cinema halls, clubs,
commercial complexes, community centres, customer care centres, e-Seva centres, educational centres ,
electric sub-stations.
Polygon data
Rajahmundry parliamentary constituency boundary, slums in Rajahmundry parliamentary constituency

Figure 1
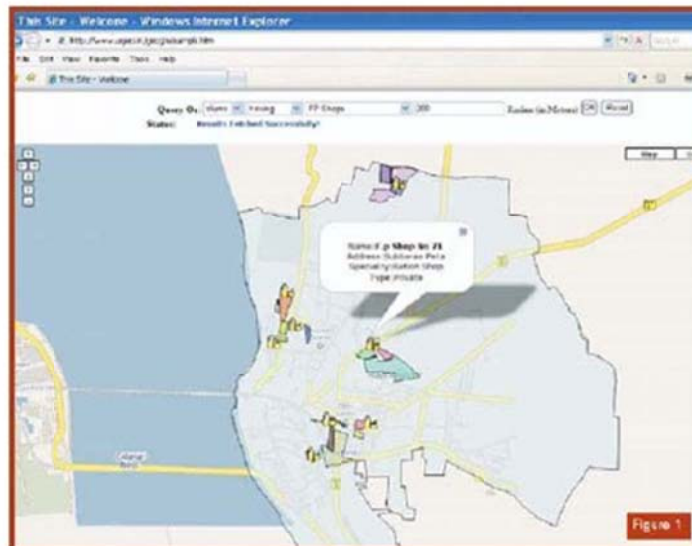
.
Steps involved in developing the Web-GIS application
The following are prerequisites that ensured the development of Web-GIS application for Rajahmundry
parliamentary constituency.
• An Apache Web server running PHP and PostgreSQL/PostGIS
• Spatial data in PostGIS database for Rajahmundry.
• Populating the spatial data into Post- GIS database
• Outputting XML with PHP
• Generating HTML page for map visualisation
The URL for Rajahmundry parliamentary constituency Web-GIS application is http://www.osgeo.in/google
/sample. htm. The interface which was developed for Rajahmundry Parliamentary Constituency Web-GIS
application allows users to query against spatial data available in the PostgreSQL/Post- GIS database.
FEATURES
• Intuitive user interface for querying the spatial data
General public is familiar with Google Maps interface and its basic navigation. In this, we added a simple query
tool. In Figure 1, the results have been showed by the query slums having fair price (FP) shops within the
radius of 100 meters. A user can click on the FP shops marker icons which will pop up an info window to show
the information about the particular FP shop. User can even
Advantages
• Free (except for developer time)
• Quick development time - depending on complexity of application
• End product is light weight application, client side scripting
• Intuitive user interface - general public already familiar with Google Maps interface and basic navigation
• Fast, good response time
• Google provides solid background services (satellite data, roads, traffic data, street view, geocoding)
• Effective for displaying selected GIS data - not every mapping application requires multiple, complex map
layers
• Great on-line resources for learning, multitude of samples and tutorials
• Best for focussed applications
• Best for small municipalities, companies (not huge datasets)
Limitations
• Limited functionality compared to some commercial products
• Difficult to overlay more complex GIS data
• Difficult to overlay multiple GIS layers

• Cannot read directly from GIS database, must convert to other formats (XML, KML) (but parts of process can be automated with scripting)

Figure 1

Figure 3

Figure 4

get the information about slums by clicking on slums polygon. The interface allows the user to query slums not having FP shops within the radius of 100 meters to know which slums don't have the FP shops nearby. Figure 2 shows the result for the query slums not having FP shops within the radius of 100 meters.

• Light weight Web-GIS application

This application gets loaded faster on the browser and even the querying time is less. Results too get displayed faster.

• Google provides background data such as satellite data, roads etc.

• This application is useful for common public, policy makers, decision makers, government officials etc.

A simple Web- GIS application showing point of interests in the Rajahmundry parliamentary constituency was developed to get the information about point of interests such as parks and gardens, Banks and ATM centres, Police stations etc (Figure 3). Following the URL for it http:// www.osgeo.in/google. overlay. htm. In this application, point of interests can be overlaid by checking the respective point of interest check boxes.

The Rajahmundry parliamentary constituency Web-GIS is also available in the regional language Telugu (Figure

4). Following the URL for Telugu version of Rajahmundry parliamentary constituency Web-GIS application http:// www.osgeo.in/google/sample_ te.htm http://www.osgeo.in/ google /overlay_te.htm.

The Indic IME for Telugu software from www.bhashaindia. com was used to type the data in Telugu. The PostgreSQL database was encoded to UTF-8 and then the data was loaded into PostgreSQL database.

THE WAY FORWARD

We increasingly see the use of Open Source tools by a large number of stakeholders in virtually all thematic areas of concern. Combined with the power of the use of Indic languages as a preferred medium, the Web GIS will emerge as one of the most potent tools for societal benefit.

ACKNOWLEDGEMENTS

It is a pleasure to thank Vundavalli Aruna Kumar, MP for making this small scale experiment a success, the Open Source geospatial and particularly Dr PS Roy, V Ravi Kumar, Dr Hanumantha Rao, Ramamurthy, Sinha, Aneel Kumar, Dr Sahu for their active involvement and help.

Spatio-temporal landscape modelling for natural hazard vulnerability analysis in select watersheds of Central Western Ghats

T. V. Ramachandra, *Senior Member, IEEE*, AninditaDasgupta,Uttam Kumar, *Student Member, IEEE,* Bharath H Aithal, *Student Member, IEEE,* P. G. Diwakar and N. V. Joshi

Abstract—**Natural hazards such as landslides are triggered by numerous factors such as ground movements, rock falls, slope failure, debris flows, slope instability, etc. Changes in slope stability happen due to human intervention, anthropogenic activities, change in soil structure, loss or absence of vegetation (changes in land cover), etc. Loss of vegetation happens when the forest is fragmented due to anthropogenic activities. Hence land cover mapping with forest fragmentation can provide vital information for visualising the regions that require immediate attention from slope stability aspects**. The main objective of this paper is to understand the rate of change in forest landscape from 1973 to 2004 through multi-sensor remote sensing data analysis. The forest fragmentation index presented here is based on temporal land use information and forest fragmentation model, in which the forest pixels are classified as patch, transitional, edge, perforated, and interior, that give a measure of forest continuity. The analysis carried out for five prominent watersheds of Uttara Kannada district– Aganashini, Bedthi, Kali, Sharavathi and Venkatpura revealed that interior forest is continuously decreasing while patch, transitional, edge and perforated forest show increasing trend. The effect of forest fragmentation on landslide occurrence was visualised by overlaying the landslide occurrence points on classified image and forest fragmentation map. The increasing patch and transitional forest on hill slopes are the areas prone to landslides, evident from the field verification, indicating that deforestation is a major triggering factor for landslides. This emphasises the need for immediate conservation measures for sustainable management of the landscape. Quantifying and describing land use - land cover change and fragmentation is crucial for assessing the effect of land management policies and environmental protection decisions.

Index Terms—**Forest fragmentation, landslide, SVM, MLC**

INTRODUCTION

Natural hazards such as landslides involving small to large ground movements are mainly triggered due to unstable slopes with scanty green cover. Unstable sloped are induced due to the removal of vegetation cover, rock falls, deep failure of slopes, shallow debris flows, ground water pressure, erosion, soil nutrients, soil structure, etc. The actual landslide often requires a trigger before being released and in most cases, a change in land cover (LC) due to the loss of vegetation is a primary factor that builds up specific sub-surface conditions for landslide to occur. The loss in forest cover due to LC change has increased rapidly in recent times due to increasing mankind needs. Forest cover is reduced to almost half of the ecologically desired amount [1] and about 72 percent of India's forests have lost their viability for regeneration, with forest grazing being one of the most important causes [2]. Forest fragmentation apart from affecting the biodiversity and ecology of the region has a significant influence in the movement of soil (silt) and debris in undulating terrains with high intensity rainfall. Forest fragmentation analysis spatially aids in visualising the regions that require immediate attention to minimise natural calamities such as landslides. Spatial fragmentation map depicts the type and extent of fragmentation derived from land use (LU) data which are obtained from multi-source, multi-sensor, multi-temporal, multi-frequency or multi-polarization remote sensing (RS) data. The objectives of this paper are

i.) Classification of multi-temporal RS data using Maximum Likelihood classifier to obtain LU map.
ii.) Multi-temporal forest fragmentation analysis for five watersheds in Uttara Kannada to characterise the type and extent of fragmentation or loss of vegetation cover.
iii.) Visualising the consequences of fragmentation for landslide susceptibility.

Methods

Maximum Likelihood classifier (MLC) – Supervised classification of the image was performed using MLC. MLC has become popular and widespread in RS because of its robustness [3–6]. It quantitatively evaluates both the variance and covariance of the category spectral response

pattern [7] assuming the distribution of data points to be Gaussian [8] which is described by the mean vector and the covariance matrix. The statistical probability of a given pixel value being a member of a particular class is computed and the pixel is assigned to the most likely class (highest probability value). If the training data pertaining to different classes contain n samples and the samples in each class are i.i.d. (independent and identically distributed) random variables and further if we assume that the spectral classes for an image is represented by ω_n, n=1,..., N, where N is the total number of classes, then probability density $p(\omega_n|x)$ gives the likelihood that the pixel x belongs to class ω_n where x is a column vector of the observed digital number of the pixels. It describes the pixel as a point in multispectral space (M-dimensional space, where M is the number of spectral bands). The maximum likelihood (ML) parameters are estimated from representative i.i.d. samples. Classification is performed according to

$$x \in \omega_n \text{ if } p(\omega_n \mid x) > p(\omega_j \mid x) \ \forall \ j \neq n \tag{1}$$

i.e., the pixel x belongs to class $\omega_n$if $p(\omega_n|x)$ is largest. The ML decision rule is based on a normalised estimate of the probability density function (p.d.f.) of each class. The discriminant function, $g_{ln}(x)$ for ω_n in MLC is expressed as

$$g_{ln}(\mathbf{x}) = p(x|\omega_n)p(\omega_n) \tag{2}$$

where $p(\omega_n)$ is the prior probability of ω_n, $p(x|\omega_n)$ is the p.d.f. for pixel vector x conditioned on ω_n[6]. Pixel vector x is assigned to the class for which $g_{ln}(x)$ is greatest. In an operational context, the logarithm form of (2) is used, and after the constants are eliminated, the discriminant function for ω_n is stated as

$$G_{ln}(\mathbf{x}) = (\mathbf{x} - \mu_n)^T \Sigma_n^{-1}(\mathbf{x} - \mu_n) + \ln | \Sigma_n | -2\ln P(\omega_n) \tag{3}$$

where Σ_n is the variance-covariance matrix of ω_n, μ_n is the mean vector of ω_n. A pixel is assigned to the class with the lowest $G_{ln}(\mathbf{x})$ [6, 9-10].For each LU class (agricultural land, human settlement / residential / commercial areas, roads, forest, plantation, waste land / open land and water bodies) training samples were collected representing approximately 10% of the study area. With these 10% known pixel labels from training data, the aim was to assign labels to all the remaining pixels in the image.

Forest fragmentation: Forest fragmentation is the process whereby a large, continuous area of forest is both reduced in area and divided into two or more fragments. The decline in the size of the forest and the increasing isolation between the two remnant patches of the forest has been the major cause of declining biodiversity [11-15]. The primary concern is direct loss of forest area, and all disturbed forests are subject to "edge effects" of one kind or another. Forest fragmentation is of additional concern, insofar as the edge effect is mitigated by the residual spatial pattern [16-19].

LU map indicate only the location and type of forest, and further analysis is needed to quantify the forest fragmentation. Total extent of forest and its occurrence as adjacent pixels, fixed-area windows surrounding each forest pixel is used for calculating type of fragmentation. The result is stored at the location of the centre pixel. Thus, a pixel value in the derived map refers to between-pixel fragmentation around the corresponding forest location. As an example [20], if Pf is the proportion of pixels in the window that are forested and Pff is the proportion of all adjacent (cardinal directions only) pixel pairs that include at least one forest pixel, for which both pixels are forested. Pff estimates the conditional probability that, given a pixel of forest, its neighbour is also forest. The six fragmentation model that identifies six fragmentation categories are: (1) interior, for which Pf = 1.0; (2), patch, Pf < 0.4; (3) transtitional, 0.4 < Pf < 0.6; (4) edge, Pf > 0.6 and Pf-Pff> 0; (5) perforated, Pf

> 0.6 and Pf-Pff< 0, and (6) undetermined, Pf > 0.6 and Pf = Pff. When Pff is larger than Pf, the implication is that forest is clumped; the probability that an immediate neighbour is also forest is greater than the average probability of forest within the window. Conversely, when Pff is smaller than Pf, the implication is that whatever is non-forest is clumped. The difference (Pf-Pff) characterises a gradient from forest clumping (edge) to non-forest clumping (perforated). When Pff = Pf, the model cannot distinguish forest or non-forest clumping. The case of Pf = 1 (interior) represents a completely forested window for which Pff must be 1.

Support Vector Machine (SVM): SVM are supervised learning algorithms based on statistical learning theory, which are considered to be heuristic algorithms [21]. SVM map input vectors to a higher dimensional space where a maximal separating hyper plane is constructed. Two parallel hyper planes are constructed on each side of the hyper plane that separates the data. The separating hyper plane maximises the distance between the two parallel hyper planes. An assumption is made that the larger the margin or distance between these parallel hyper planes, the better the generalisation error of the classifier will be. The model produced by support vector classification only depends on a subset of the training data, because the cost function for building the model does not take into account training points that lie beyond the margin [21]. When it is not possible to define the hyper plane by linear equations, the data can be mapped into a higher dimensional space through some nonlinear mapping functions.

A free and open source software – openModeller [22] was used for predicting the probable landslide areas using SVM. openModeller (http://openmodeller.sourceforge.net/) includes facilities for reading landslide occurrence and environmental data, selection of environmental layers on which the model should be based, creating a fundamental niche model and projecting the model into an environmental scenario as shown in Fig. 1.

Study area and Data

The Uttara Kannada district lies 74°9' to 75°10' E longitude and 13°55' to 15°31' N latitude, extending over an area of 10,291 km^2 in the mid-western part of Karnataka state (Fig. 2). It accounts for 5.37 % of the total area of the state with a population above 1.2 million [23]. This region has gentle undulating hills, rising steeply from a narrow coastal strip bordering the Arabian sea to a plateau at an altitude of 500 m with occasional hills rising above 600–860m.

Landslide occurrence points + Environmental Layers (Land Cover, DEM, Aspect, Slope, Rainfall) = Probability Distribution
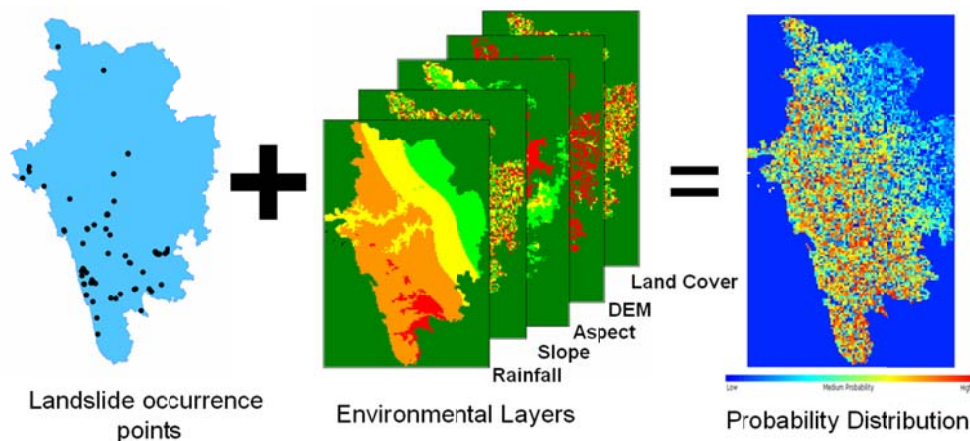
Fig. 2.Uttara Kannada district, Karnataka, India.

This district, with 11 taluks, can be broadly categorised into three distinct regions — coastal lands (Karwar, Ankola, Kumta, Honnavar and Bhatkaltaluks), mostly forested Sahyadrian interior (Supa, Yellapur, Sirsi and Siddapurtaluks) and the eastern margin where the table land begins (Haliyal, Yellapur and Mundgodtaluks). Climatic conditions range from arid to humid due to physiographic conditions ranging from plains, mountains to coast.

Survey of India (SOI) toposheets of 1:50000 and 1:250000 scales were used to generate base layers – district and taluk boundaries, water bodies, drainage network, etc. Field data were collected with a handheld GPS. RS data used in the study were Landsat MSS (1973, spatial resolution – 79m), Landsat TM (1989, spatial resolution – 30m), Landsat ETM+ (2000, spatial resolution – 30m) [downloaded from Global Land Cover Facility, http://www.landcover.org] and LISS-III Multi-spectral (2004, spatial resolution – 23.5m) procured from NRSC, Hyderabad, India. Google Earth data (http://earth.google.com) served in pre and post classification process and validation of the results.

Environmental data such as precipitation of wettest month were downloaded from WorldClim – Global Climate Data [http://www.worldclim.org/bioclim]. Other environmental layers (Aspect, DEM, Flow accumulation, Flow direction, Slope, Compound Topographic Index) used for modelling landslide were obtained from USGS Earth Resources Observation and Science (EROS) Center based Hydro1Kdatabase   [http://eros.usgs.gov/#/Find_Data/ Products_and_Data_Available/gtopo30/hydro/asia]. The global LC change maps were obtained from Global   Land   Cover   Facility,   Land   Cover   Change [http://glcf.umiacs.umd.edu/services/landcoverchange/landcover.shtml; http://www.landcover.org/services/landcover change/landcover.shtml]. The spatial resolution of all

the data were 1 km. 125 landslide occurrence points of low, medium and high intensity were recorded using handheld GPS from the field and published reports.

Results and Discussion

RS data were geometrically corrected on a pixel by pixel basis and the images were resampled to a common resolution of 30m with a dimension of 7562 x 6790. LC mapping (Fig. 3) was done using normalised difference vegetation index (NDVI) given as

$$\frac{NIR\text{-}Red}{NIR\text{+}Red}$$

(4)

NDVI values range from -1 to +1; increasing values from 0 indicate presence of vegetation (agriculture, forest and plantation) and negative values indicate absence of greenery (builtup, sand, fallow, water) as given in Table I.
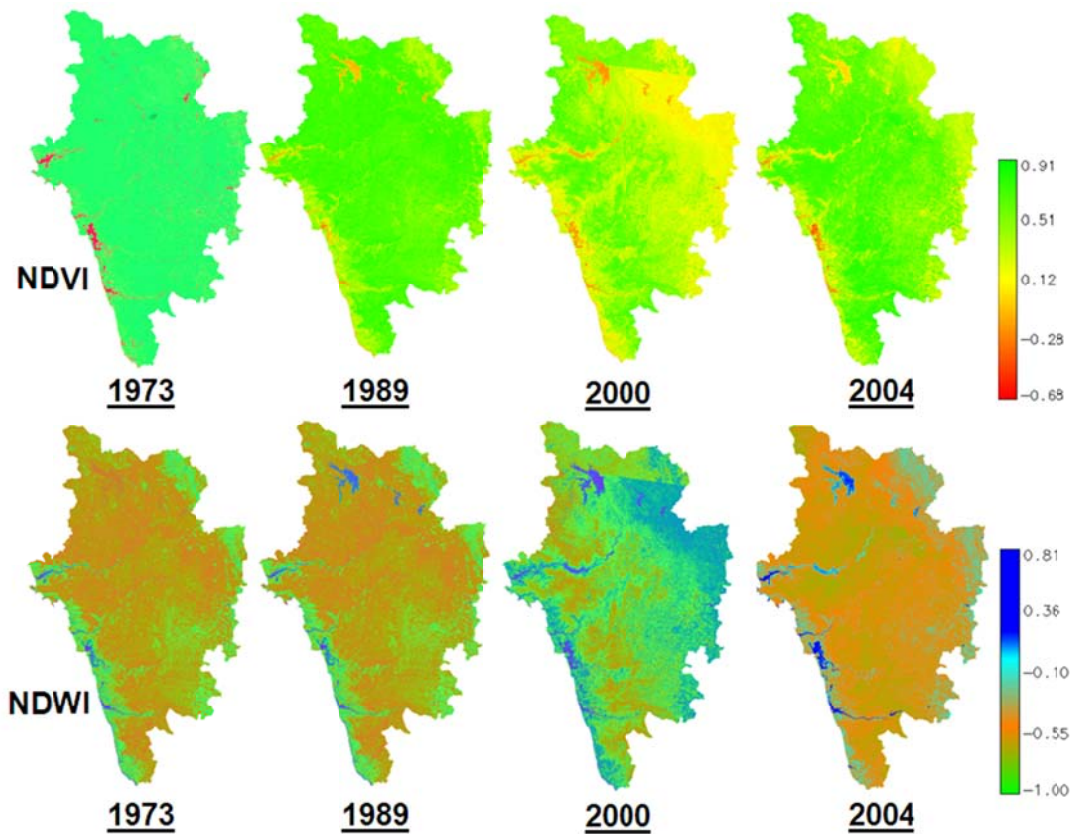
Fig. 3.NDVI and NDWI for Uttara Kannada.

TABLE I: NDVI AND WDVI OF UTTARA KANNADA.

INDEX	NDVI		WDVI	
	VEGETATION (%)	NON-VEGETATION (%)	WATER (%)	NON-WATER (%)
1973	97.82	2.18	1.75	98.25
1989	96.13	3.87	2.20	97.80

| 2000 | 94.33 | 5.67 | 2.80 | 97.20 |
| 2004 | 94.00 | 6.00 | 2.50 | 97.50 |

Mapping of water bodies (Fig. 3) was done using normalised difference water index (NDWI) [24] given as

$$\frac{Green-NIR}{Green+NIR}$$

(5)

NDWI values above 0 indicate presence of water bodies and values below 0 indicate other classes.

Training pixels were collected from the false colour composite of the respective bands (for time period 1973, 1989, and 2000) since historical data were unavailable. For 2004 data, training data uniformly distributed over the study area collected with pre calibrated GPS were used. The class spectral characteristics for six LU categories (agriculture, builtup / settlement, forest (evergreen, semi-evergreen, deciduous), plantation, waste land / fallow / sand, and water bodies / streams) using RS data were obtained to assess their inter-class separability. and the images were classified using MLC. Temporal classified images (into 6 LC classes – agriculture, built-up, forest, plantation, waste land and water bodies) are shown in Fig. 4 and the statistics are given in Table II. This was validated with the representative field data (covering ~ 10% of the study area) and also using Google Earth image.Producer's, user's, overall accuracy and Kappa values computed are listed in Table III.

Further, five watersheds were delineated from the classified images of four time period as shown in Fig. 5 and the statistics are listed in Table IV. Forest fragmentation model was used to obtain fragmentation indices as shown in Fig. 6 and the statistics are presented in Table V. Forest was categorised into patch, transitional, edge, perforated, and interior types using programs in GRASS GIS (http://wgbis.ces.iisc.ernet.in/foss).

Fig. 4. Temporal classified image of Uttara Kannada district.

TABLE II: LU DETAILS OF UTTARA KANNADA DISTRICT

YEAR	1973		1989		2000		2004	
	AREA		AREA		AREA		AREA	
CLASS	(HA)	(%)	(Ha)	(%)	(Ha)	(%)	(Ha)	(%)

AGRICULTURE	25794	2.51	57465	5.59	97749	9.51	115464	11.23
BUILTUP	4490	0.44	8671	0.84	13674	1.33	27577	2.68
FOREST	897420	87.29	845665	82.25	749877	72.94	627455	61.02
PLANTATION	82439	8.02	85190	8.29	122160	11.88	184340	17.93
WASTELANDLAND	-	-	8483	0.83	15884	1.54	23456	2.28
WATER bodies	17995	1.75	22666	2.20	28796	2.80	49924	4.86
TOTAL	1028151	100	1028151	100	1028151	100	1028151	100
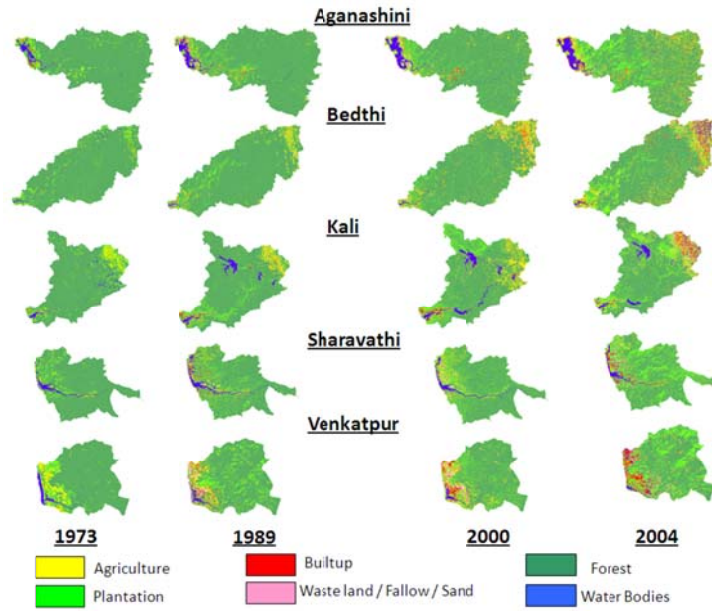
Fig. 5. Temporal LU image of the five rive basins of Uttara Kannada district.

TABLE III: ACCURACY ASSESSMENT OF THE CLASSIFIED IMAGES

YEAR	CATEGORY	USER'S ACCURACY	PRODUCER'S ACCURACY	OVERALL ACCURACY	KAPPA
1973	AGRICULTURE	74.12	77.97	76.89	71.92
	BUILTUP	68.45	65.42		
	FOREST	87.23	85.22		
	PLANTATION	81.27	77.23		
	WASTELAND	85.62	81.21		
	WATER BODIEE	78.50	74.29		
1989	AGRICULTURE	70.43	80.39	75.78	69.82
	BUILTUP	78.78	81.11		
	FOREST	73.96	73.39		
	PLANTATION	79.62	77.19		
	WASTELAND	80.21	77.78		
	WATER BODIEE	31.02	64.82		
2000	AGRICULTURE	81.16	80.47	83.21	79.21
	BUILTUP	84.29	87.50		
	FOREST	87.15	82.54		
	PLANTATION	86.51	87.23		
	WASTELAND	85.22	81.79		
	WATER BODIEE	86.40	97.56		
2004	AGRICULTURE	85.21	84.54	87.83	83.11
	BUILTUP	86.47	83.11		
	FOREST	94.73	96.20		
	PLANTATION	92.27	91.73		
	WASTELAND	88.49	87.88		
	WATER BODIEE	83.13	81.33		

It is to be noted that Supa dam was constructed in late 1970s. Hence this water body is absent in the 1973 Landsat multi-spectral classified image and prominent in the classified images of 1989, 2000 and 2004 (Fig. 4). Also, there was no waste land in 1973 as per the statistics in Table II. All the rivers are perennial since the catchment areas are mainly composed of interior forest. However, eventually there are drastic land cover changes consequent to unplanned developmental activities. In addition, the dams of major rivers in the district have inundated large vegetation areas. These areas have silt deposit at the river beds and have been classified as sand / waste land.

Forest and plantation were considered as a single class - forest and all other classes were considered as non-forest as the extent of LC is a decisive factor in landslides. While the forest fragmentation map produced

valuable information, it also helped to visualise the state of forest for tracking the trends and to identify the areas where forest restoration might prove appropriate to reduce the impact of forest fragmentation. Forest fragmentation also depends on the scale of analysis (window size) and various consequences of increasing the window size are reported in [25]. The measurements are also sensitive to pixel size.Nepstad et al., (1999a, 1999b)[26 and 27]reported higher fragmentation when using finer grain maps over a fixed extent (window size) of tropical rain forest. Finer grain maps identify more non-forest area where forest cover is dominant but not exclusive.

The criterion for interior forest is more difficult to satisfy over larger areas. Although knowledge of the feasible parameter space is not critical, there are geometric constraints [28]. For example, it is not possible to obtain a low value of Pff when Pf is large. Percolation theory applies strictly to maps resulting from random processes; hence, the critical values of Pf (0.4 and 0.6) are only approximate and may vary with actual pattern. As a practical matter, when Pf > 0.6, non-forest types generally appeared as "islands" on a forest background, and when Pf < 0.4, forests appeared as "islands" on a non-forest background. Fig. 6 shows the temporal change in forest patch type, revealing that patch, transitional and edge forest are increasing due to deforestation and interior forest is decreasing. The statistics in Table V shows the decrease in forest patch types from 1973 to 2004. Forest fragmentation is a vital indicator that is accountable for environmental changes. With the expansion of human settlement there is higher risk of forest degradation disturbing the biodiversity, affecting the water quality, endangering wildlife survival, habitat protection, etc.
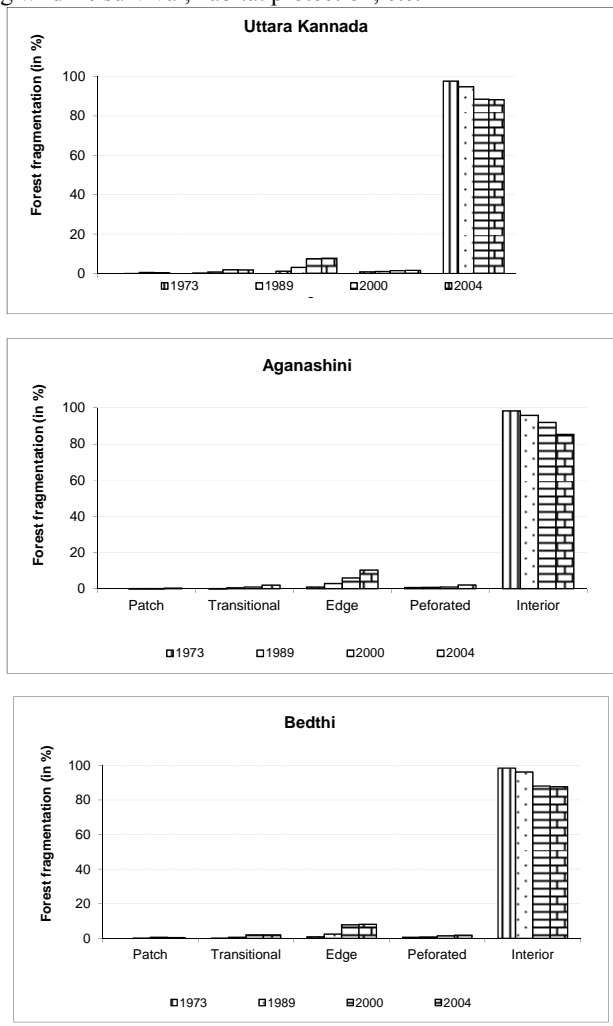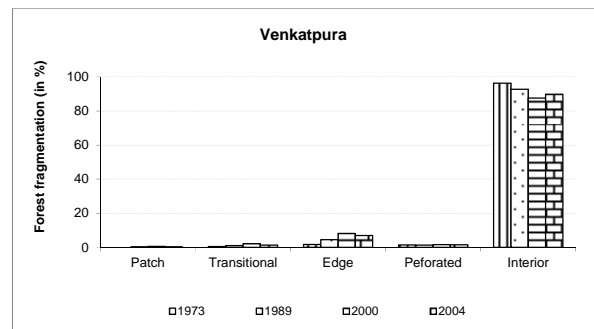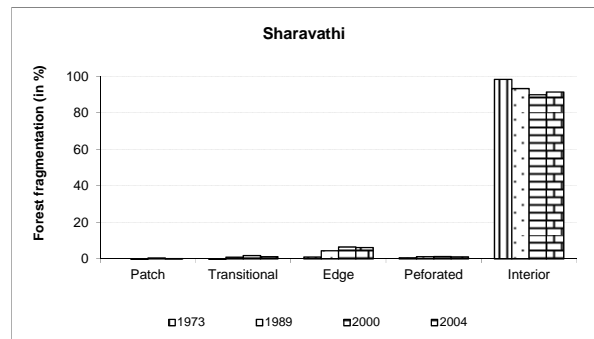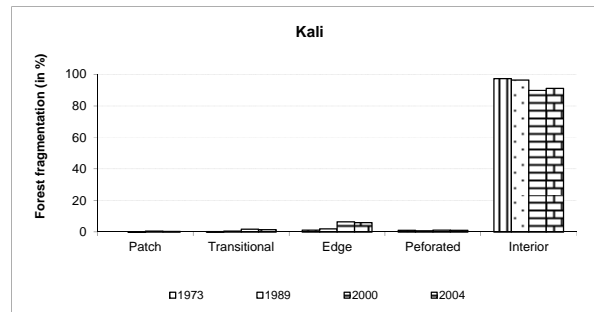
Fig. 6.Temporal Forest fragmentation change of the five river basins of Uttara Kannada district.
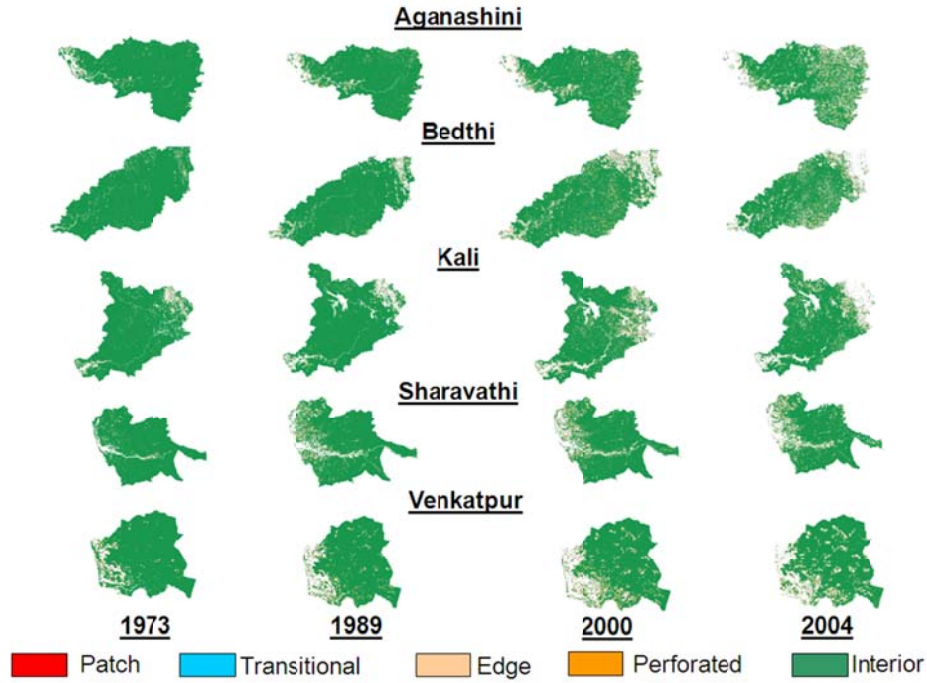
Fig. 6.Forest fragmentation in five river basins in Uttara Kannada district.

TABLE IV: LU DETAILS OF THE FIVE RIVER BASINS (IN HA AND %)

YEAR	CATEGORY	AGANASHINI		BEDTHI		KALI		SHARAVATHI		VENKATPURA	
1973	AGRICULTURE	1886.6	1.4	3616.42	1.3	14294	3.8	1302	1.6	1355.2	5.9
	BUILTUP	627.6	0.5	799.83	0.3	1318.7	0.4	103.976	0.1	74.69	0.3
	FOREST	123500	90	251363	92	320174	86	71536	88	18282	81
	PLANTATION	7930.9	5.8	16315.9	6	31593	8.4	2900	7.3	2289	10
	WASTELAND	-	-	-	-	-	-	-	-	-	-
	WATER BODIES	2662.5	2	1556.7	0.6	6696.8	1.8	2329	2.9	615.9	2.7
1989	AGRICULTURE	4466	3.3	12550.9	4.6	19667	5.3	4110	5.1	1529.7	6.8
	BUILTUP	883.9	0.7	856.9	0.3	2886.6	0.8	784	1.0	404.4	1.8
	FOREST	118724	87	239894	87	303758	81	63686	78	16725	74
	PLANTATION	8217.7	6.0	17783	6.5	34152	9.1	8807	11	3165.8	14
	WASTELAND	784.9	0.6	1360.8	0.5	1696.5	0.5	729	1	508	2.3
	WATER BODIES	3531.6	2.6	1205.8	0.4	11916	3.2	3055	3.8	284	1.3
2000	AGRICULTURE	6179.3	4.5	29413	11	33421	8.9	4781	5.9	1881.3	8.3
	BUILTUP	969.7	0.7	3838.4	1.4	4364	1.2	414.8	0.5	717.6	3.2
	FOREST	11130	81	218541	80	265508	71	59379	73	15363	68
	PLANTATION	12004.6	8.8	16379	6	50542	14	12936	16	3519.7	16
	WASTELAND	1571.8	1.1	3978	1.5	2788.6	0.8	1258	1.6	822.65	3.7
	WATER BODIES	4580.5	3.4	1501.8	0.6	17453	4.7	2404	3	312.9	1.4
2004	AGRICULTURE	13028	9.5	36414	13	36575	9.8	3778.9	4.6	1489.9	6.6
	BUILTUP	2706.5	1.9	5221	1.9	7364.2	1.9	1723.2	2.1	959.98	4.3
	FOREST	84127.8	61	181674	66	232097	62	47020	58	13048	58
	PLANTATION	27836.2	20	35315	13	67093	18	23472	29	5896	26
	WASTELAND	1847.4	1.4	5411	2	10332	2.8	1720.5	2.1	381.97	1.7
	WATER BODIES	7064.4	5.2	9603	3.5	20549	5.5	3484.7	4.3	836.26	3.7

TABLE V: FORESTFRGMENTATION DETAILS OF THE FIVE RIVER BASINS (IN HA)

YEAR	CATEGORY	AGANASHINI	BEDTHI	KALI	SHARAVATHI	VENKATPURA
1973	PATCH	0.59	0.16	1.49	-	0.11
	TRANSITIONAL	189.17	322.04	1050.52	94.99	85.40
	EDGE	1177.55	2398.25	4442.65	729.81	347.26
	PERFORATED	787.33	1478.46	3588.78	457.62	293.29
	INTERIOR	127895.85	261524.81	337778.54	74929.99	19506.20
	PATCH	83.69	217.57	333.17	94.67	38.33
	TRANSITIONAL	631.03	1443.65	1995.02	670.22	215.23
	EDGE	3638.08	6198.87	6753.60	3155.90	889.215
	PERFORATED	857.74	1907.86	2526.22	839.13	269.26
	INTERIOR	120188.05	246239.30	322328.53	66064.81	18178.51
2000	PATCH	213.07	1326.61	1637.77	198.690	95.52
	TRANSITIONAL	1201.70	4622.81	5654.96	919.972	390.37
	EDGE	7263.70	18204.16	20319.45	4262.67	1507.74
	PERFORATED	1169.60	3506.29	3977.67	770.327	302.99
	INTERIOR	112006.30	205359.76	280116.92	64838.71	16249.80
2004	PATCH	363.82	947.32	1117.11	144.83	40.424
	TRANSITIONAL	2106.73	4207.65	4414.24	781.60	243.40
	EDGE	11098.47	17250.44	17408.67	4189.86	1274.87
	PERFORATED	2199.95	3942.96	3120.97	717.20	275.79
	INTERIOR	92213.32	186500.95	267026.02	62076.267	16337.97

Precipitations of wettest month along with the seven other layers (as mentioned in section III) were used to predict landslides using SVM as shown in Fig. 7. The landslide occurrence points were overlaid on the probability map to validate the prediction as shown in Fig. 8.
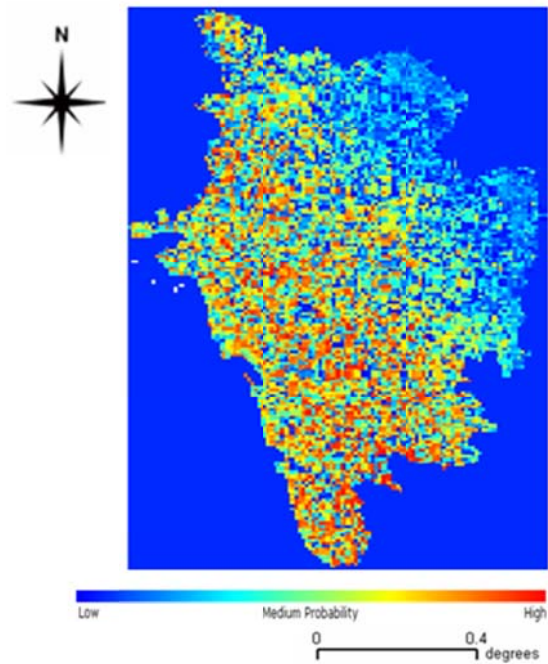
Fig. 7.Probability distribution of the landslide prone areas.
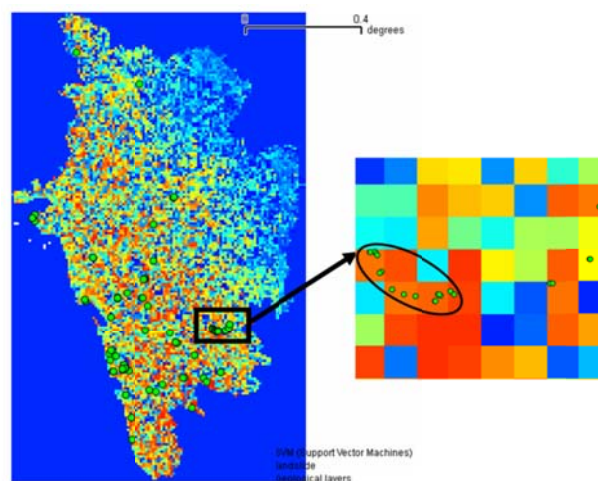
Fig. 8.Validation of the probability distribution of the landslide prone areas by overlaying landslide occurrence points.

The SVM map was 96% accurate with respect to the ground and Kappa values 0.8733 and 0.9083 respectively [29]. Most of the predicted landslide areas as probable landslide prone zones (indicated in

red in Fig. 7) reside on terrain that is highly undulating with steep slopes and are frequently exposed to landslides induced by rainfall. Maximum number of landslide points occurring in the undulating terrain, collected from the ground fell in fragmented areas - patch, transitional and edge forest. Fig. 9 shows the landslide occurrence points obtained from ground using handheld GPS overlaid on the classified image of the district (A) and forest fragmentation map (B) based on 2004 RS data. This map indicates that hill slopes with undulating terrain and less vegetation cover (patch, transitional and edge forest) are more susceptible to landslides compared to north-eastern part of the district which has relatively flat terrain with large area utilised under agricultural practices.

The present study makes an effort to quantify forest cover by using fragmentation index for measuring the disturbances due to human impact giving a complete picture of the change that has occurred. The five prominent watersheds of Uttara Kannada district were assessed for quantifying forest fragmentation caused due to anthropogenic disturbances. Forest fragmentation model is useful to visualize state of forest fragmentation for an area and identifying areas where forest restoration might prove appropriate. If the vegetative areas have been cleared, the water retaining capacity of the soil has decreased, triggering landslides in those areas. Hence, there is a immediate need to restore those vegetative by forestation to ensure that the soil is retained on the hill slopes and do not activate any downward movement of the hill tops.

Natural disasters have drastically increased over the last decades. National, state and local government including NGOs are concerned with the loss of human life and damage to property caused by natural disasters. The trend of increasing incidences of landslides occurrence is expected to continue in the next decades due to urbanisation, continued anthropogenic activities, deforestation in the name of development and increased regional precipitation in landslide-prone areas due to changing climatic patterns [30].

Conclusion

Landslides occur when masses of rock, earth or debris move down a slope. Mudslides, debris flows or mudflows, are common type of fast-moving landslides that tend to flow in channels. These are caused by disturbances in the natural stability of a slope, which are triggered by high intensity rains. The primary criteria that influence landslides are precipitation intensity, slope, soil type, elevation, vegetation cover and LC type.

In this paper, LU analyses along with forest fragmentation were carried out for five perennial watersheds of Uttara Kannada district. SVM was used to generate probable landslide map. By overlaying the fragmentation map on the landslide probable map, past and future occurrences of landslides can be visualised. It is also evident that degradation of forest and LC change is an important factor that is not only responsible for triggering landslides, but also a major contributor to global warming, climate change, natural resource depletion and consequent detrimental effect on our environment.
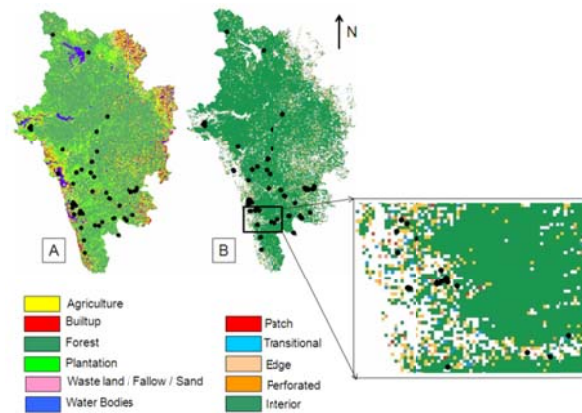
Fig. 9. Landslide susceptibility points overlaid on the classified image of 2004 (A) and forest fragmentation map (B). Black points represent the occurrence of landslide points obtained using handheld GPS.

VI. Acknowledgment

REFERENCES

[1] G. B. Maithani, "*Non-forestry technology for forest based wasteland development,*" The Indian Forester, vol. 116, pp. 931-937, 1990.

[2] S. N. Rai and A. Saxena, "*The extent of forest fire, grazing and regeneration status in inventoried forest areas in India,*" The Indian Forester, vol. 123, pp. 689-702, 1997.

[3] A. H. Strahler, "*The use of prior probabilities in maximum likelihood classification of remotely sensed data,*" Remote Sensing of Environment, vol. 10, pp. 135-163, 1980.

[4] C. Conese and F. Maselli, "*Use of error matrices to improve area estimates with maximum likelihood classification procedures,*" Remote Sensing of Environment, vol. 40, pp. 113-124, 1992.

[5] L. Ediriwickrema and S. Khorram, "*Hierarchical maximum-likelihood classification for improved accuracies,*" IEEE Transactions on Geoscience and Remote Sensing, vol. 35, pp. 810-816, 1997.

[6] M. Zheng, Q. Cai and Z. Wang, "*In: Effect of prior probabilities on maximum likelihood classifier,*" Geoscience and Remote Sensing Symposium, 2005, IGARS'05, Proceedings 2005 IEEE International, vol. 6, pp. 3753-3756, 2005.

[7] T. M. Lillesand and R. W. Kiefer, "*Remote Sensing and Image Interpretation,*" Fourth Edition, John Wiley and Sons, 2002, ISBN 9971-51-427-3.

[8] U. Bayarsaikhan, B. Boldgiv, K-R. Kim, K-A. Park and D. Lee, "*Change detection and classification of land cover at HustaiNational Park in Mongoli*a," International Journal of Applied Earth Observation and Geoinformation, vol. 11, pp. 273-280, 2009.

[9] A. R. John and J. Xiuping, "*Remote Sensing Digital image Analysis: An Introduction,*" Springer-Verlag Inc., New York, 1999.

[10] R. O. Duda, P. E. Hart and D, G. Stork, "*Pattern classification,*" New York, A Wiley-Interscience Publication, Second Edition, 2000, ISBN 9814-12-602-0.

[11] R. P. Tucker, and J. F. Richards, "*Global deforestation and the nineteenth century world economy,*" Duke University Press, Durham, North Carolina, USA, 1983.

[12] B. L. Turner, W. C. Clark, R. W. Kates, J. F. Richards, J. T. Mathews, and W. B. Meyer, *"The Earth as transformed by human action,"* Cambridge University Press, Cambridge, UK, 1990.

[13] W. B. Meyer, and B. L. Turner, *"Changes in land use and land cover: a global perspective,"* Cambridge University Press, Cambridge, UK, 1994.

[14] R. A. Houghton, *"The worldwide extent of land-use change,"* BioScience, vol. 44, pp. 305-313, 1994.

[15] J. D. Hurd, E. H. Wilson, S. G. Lammey, and D. L. Civco, *"Characterization of Forest Fragmentation and Urban Sprawl using time sequential Landsat Imagery,"* In ASPRS 2001 Annual Convention, St. Louis, MO, 2001.

[16] R. T. T. Forman, and M. Godron, *"Landscape Ecology,"* John Wiley, New York, 1986.

[17] M. G. Turner, *"Landscape ecology: the effect of pattern on process,"* Annual Review of Ecology and Systematics, vol. 20, pp. 171-197, 89

[18] S. A. Levin, *"The problem of pattern and scale in ecology,"* Ecology, vol. 73, pp. 1943-1967, 1992.

[19] T. V. Ramachandra, U. Kumar, P. G. Diwakar and N. V. Joshi, *"Land Cover Assessment using À Trous Wavelet fusion and K-Nearest Neighbour classification,"* In proceedings of the 25th Annual In-House Symposium on Space Science and Technology, ISRO IISc Space Technology Cell, Indian Institute of Science, Bangalore, India, 29-30 January, 2009.

[20] K., J. Riitters, R. O' Neill. Wickham, B. Jones, and E. Smith, *"Global-scale patterns of forest fragmentation,"* Conservation Ecology, vol. 4(2-3), 2000.

[21] T. Kavzoglu, I. Colkesen, *"A kernel functions analysis for support vector machines for land cover classification,"* International Journal of Applied Earth Observation and Geoinformation, vol 11, pp. 352-359, 2009.

[22] M. E. S. Munoz, R. Giovanni, M. F. Siqueira, T. Sutton, P. Brewer, R. S. Pereira, D. A. L. Canhos, and V. P. Canhos, *"openModeller- a generic approach to species' potential distribution modeling,"Geoinformatica.* DOI: 10.1007/s10707-009-0090-7.

[23] T. V. Ramachandra and A. V. Nagarathna, *"Energetics in paddy cultivation in Uttara Kannada district,"* Energy Conservation and Management, vol. 42(2), pp. 131-155, 2001.

[24] S. K. McFeeters, *"The use of normalized difference water index (NDWI) in the delineation of open water features,"*International Journal of Remote Sensing, vol. 17(7), pp. 1425–1432, 1996.

[25] K. J. Riitters, R. O' Neill. Wickham, B. Jones, and E. Smith, *"Global-scale patterns of forest fragmentation,"* Conservation Ecology, vol. 4(2-3), 2000.

[26] D. C. Nepstad, P. Lefebvre, and E. A. Davidson, *"Positive feedbacks in the fire dynamic of closed canopy tropical forests,"*Science, vol. 284, pp. 1832-1835, 1999a.

[27] D. C. Nepstad, A. Verissimo, A. Alencar, C. Nobre, E. Lima, P. Lefebvre, P. Schlesinger, C. Potter, P. Moutinho, E. Mendoza, M. Cochrane, and V. Brooks, *"Large-scale impoverishment of Amazonian forests by logging and fire,"* Nature, vol. 98, pp. 505-508, 1999b.

[28] R. V. O'Neill, C. T. Hunsaker, S. P. Timmins, B. L. Jackson, K. B. Jones, K. H. Ritters, and J. D. Wickham, *"Scale problems in reporting landscape pattern at the regional scale,"* Landscape Ecology, vol. 11, pp. 169-180, 1996.

[29] T. V. Ramachandra, U. Kumar, Bharath H. Aithal, P. G. Diwakar and N. V. Joshi, *"Landslide Susceptible Locations in Western Ghats: Prediction through open Modeller,"* In proceedings of the 26th Annual In-House Symposium on Space Science and Technology, ISRO-IISc Space Technology Cell, Indian Institute of Science, Bangalore, India, 28-29 January, 2010.

[30] T. V. Ramachandra, M. D. Subashchandran and AnantHegdeAshisar, *"Landslides at Karwar, October 2009: Causes and Remedial Measures,"* ENVIS Technical Report: 32,Environmental Information System, Centre for Ecological Sciences, Bangalore, 2009.

Land Surface Temperature with Land Cover Dynamics: Multi-Resolution, Spatio-Temporal Data Analysis of Greater Bangalore

Ramachandra, T. V.,[1] **and Uttam, K.,**[2]

[1]Energy and Wetlands Research Group, Centre for Ecological Sciences & Centre for Infrastructure, Sustainable Transportation and Urban Planning [CiSTUP], Indian Institute of Science, Bangalore – 560 012 India, E-mail: cestvr@ces.iisc.ernet.in, energy@ces.iisc.ernet.in

[2]Centre for Sustainable Technologies, Department of Management Studies, Indian Institute of Science

Abstract

Bangalore is experiencing unprecedented urbanisation in recent times due to concentrated developmental activities with impetus on IT (Information Technology) and BT (Biotechnology) sectors. The concentrated developmental activities has resulted in the increase in population and consequent pressure on infrastructure, natural resources, ultimately giving rise to a plethora of serious challenges such as urban flooding, climate change, etc. One of the perceived impact at local levels is the increase in sensible heat flux from the land surface to the atmosphere, which is also referred as heat island effect. In this communication, we report the changes in land surface temperature (LST) with respect to land cover changes during 1973 to 2007. A novel technique combining the information from sub-pixel class proportions with information from classified image (using signatures of the respective classes collected from the ground) has been used to achieve more reliable classification. The analysis showed positive correlation with the increase in paved surfaces and LST. 466% increase in paved surfaces (buildings, roads, etc.) has lead to the increase in LST by about 2 °C during the last 2 decades, confirming urban heat island phenomenon. LSTs' were relatively lower (~ 4 to 7 °C) at land uses such as vegetation (parks/forests) and water bodies which act as heat sinks.

1. Introduction

Many cities in developing countries are now undergoing rapid urbanisation evident from the increase in urban population from 13% (220 million) in 1900, to 29% (732 million) in 1950, to 49% (3.2 billion) in 2005 and is projected to rise to 60% (4.9 billion) by 2030 (WUP, 2005). Accurate and timely information in land use (LU) and LU changes is crucial for long-term economic development planning and also for short-term land management. Increase in paved land covers (LC) consequent to the concentrated human activities often leads to increased land surface temperatures (LST). Enhanced LST in certain urban pockets compared to its immediate surroundings consequent to the increase in paved surfaces is known as urban heat island (UHI) phenomenon (Landsbeg, 1981). Specifically, surface and atmospheric temperatures are increased by anthropogenic heat discharge due to energy consumption, increased land surface coverage by artificial materials having high heat capacities and conductivities, and the associated decreases in vegetation and water impervious surfaces, which reduce surface temperature through evapotranspiration (Kato and Yamaguchi, 2005). Temperatures have been monitored through space borne remote sensing (RS) sensors, which measure top of the atmosphere (TOA) radiances in the Thermal Infrared (TIR) region. TOA radiance is the net result of emitted radiance from the Earth's surface, upwelling radiance from the atmosphere, and downwelling radiance from the sky. Brightness temperatures (also known as blackbody temperatures) can be derived from the TOA radiance (Dash et al., 2002). These brightness temperatures are further corrected with spectral emissivity values prior to the computation of LST to account for the roughness properties of the land surface, the amount and nature of vegetation cover, and the thermal properties and moisture content of the soil (Friedl, 2002). However, lack of knowledge of emissivity can introduce an error ranging from 0.2 to 1.2 K for mid-latitude summers and from 0.8 to 1.4 K for the winter conditions for an emissivity of 0.98 and at the ground height of 0 km (Dash et al., 2002). Two approaches have been developed to recover LST from multispectral TIR imagery (Schmugge et al., 1998). The first approach utilises a radiative transfer equation to correct the at-sensor radiance to surface radiance, followed by an emissivity model to separate the surface radiance

LISS-III data of 1999 and 2006 (23.5 m) and MODIS data of 2002 and 2007 (250 m to 500 m spatial resolution) using supervised pattern classifiers based on Gaussian maximum likelihood (GML) estimation followed by a Bayesian statistical approach. This technique quantifies the tradeoffs between various classification decisions using probability and costs that accompany such decisions (Duda et al., 2000). It makes assumptions that the decision problem is posed in probabilistic terms, and that all of the relevant probability values are known with a number of design samples or training data collected from field that are particular representatives of the patterns to be classified. The mean and covariance are computed using maximum likelihood estimation with the best estimates that maximises the probability of the pixels falling into one of the classes. LU analysis considering temporal data (1973, 1992, 1999, 2000, 2002, 2006 and 2007) was done using the open source programs (i.gensig, i.class and i.maxlik) of Geographic Resources Analysis Support System (http://wgbis.ces.iisc.ernet.in/grass).

2.3.2 Change detection

LU change detection is performed by change/no-change recognition followed by boundary delineation on images of two different time periods. Pixels which show significant changes are checked and validated on the ground and the boundaries of the changed patches are category wise delineated. This is supplemented with visual interpretation and online digitisation. Many LU change detection techniques have been developed, but no single algorithm is suitable for all cases (Lu et al., 2004), as the implementation of change detection analysis is dependent on the data itself (Zhang and Zhang, 2007). Bi-temporal multispectral images have been analysed to understand LU dynamics through:

- **Principal Component Analysis** (PCA) – PCA has been an effective tool for change detection (Fung and LeDrew, 1987 and Michener and Houhoulis, 1997). Major components of the time two (t_2) data are subtracted from the corresponding components of the time one (t_1) data to obtain differences related to changes in LU. This method provides a better result than simple image differencing when the radiometric differences between the two images are too large due to different imaging circumstances and cannot be effectively dealt with by the radiometric normalisation process (Zhang and Zhang, 2007). Landsat MSS (1973) and IRS LISS-III (2006) scenes having different radiometry were used with PCA to understand the

- **Tasseled Cap** (TC) or Kauth-Thomas (KT) transformation – Here, multi-temporal TM and ETM+ data are transformed into the brightness-greenness-wetness space, then changed areas are generated by differencing the brightness (ΔB) and greenness (ΔG) values (Kauth and Thomas, 1976). Changes in brightness (ΔB) are associated with most LU changes, especially constructed related changes. TC results can be physically interpreted as its coefficients are predetermined and independent of each image scene, while PCA coefficients are not. However, for the purpose of simply detecting change/no-change areas, PCA is better than TC in many cases, although the physical interpretation is difficult. TC transformations were performed on Landsat TM (1992) and ETM+ (2000) data.

- **Image Differencing** (ID) – ID is effective for identifying LU changes from visible and NIR band pairs acquired in similar circumstances (imaging conditions) using same sensor over two different time periods (Macleod and Congalton, 1998). IRS LISS-III (1999 and 2006) data used to visualise the LU changes through ID.

2.3.3 Derivation of LST from Landsat TM and Landsat ETM+

LST were computed (Weng et al., 2004) from TIR bands (Landsat TM and ETM+). Emissivity correction for the specified LC is carried out using surface emissivities as per Synder et al., (1998); Stathopoplou et al., (2007) and Landsat 7 science data user's handbook (2008). The emmissivity corrected land surface temperature (Ts) are computed as per Artis and Carnahan (1982)

$$T_s = \frac{T_B}{1 + (\lambda \times T_B / \rho)\ln \varepsilon}$$

Equation 1

Where, λ is the wavelength of emitted radiance for which the peak response and the average of the limiting wavelouths ($\lambda = 11.5$ μm) (Markham and Barker, 1985) were used, $\rho = h \times c/\sigma$ (1.438 x 10^{-2} mK), σ = Stefan Bolzmann's constant (5.67 x 10^{-8} Wm^{-2}K^{-4} = 1.38 x 10^{23} J/K), h = Planck's constant (6.626 x 10^{-34} Jsec), c = velocity of light (2.998 x 10^{8} m/sec), and ε is spectral emissivity.

2.3.4 LST from MODIS

MODIS LST/Emissivity 16-bit unsigned integer data with 1 km spatial resolution are multiplied by a scale factor of 0.02 (http://lpdaac.usgs.gov/modis-/dataproducts.asp#mod11) and are converted to degree Celsius.

2.3.5 NDVI from Landsat TM, ETM+ and MODIS

NDVI was computed using visible Red (0.63 – 0.69 μm) and NIR (0.76 – 0.90 μm) bands of Landsat TM (1992)/ETM+ (2000) and Red (0.62 – 0.68 microns) and NIR (0.77 – 0.86 microns) bands of LISS-III data of 2006.

2.3.6 Estimation of abundance maps

Linear unmixing method was adopted for solving the mixed pixel problem as the spectral radiance measured by a sensor consists of the mixture of radiances reflected in proportion to the sub-pixel area covered (Kumar et al., 2008). Endmembers corresponding to pure pixels are given by the reference spectra of each of the individual pure materials. Spectra corresponding to sub-pixel areas in a pixel are assumed to be linearly independent, and the target pixel spectra are a combination of these spectra (which is proportional to respective LC in a pixel). The spectrum measured by a sensor is a linear combination of the spectra, therefore,

$$y_i = \sum_{j=1}^{n} (a_{ij} x_j) + e_i$$

Equation 2

where n = the number of distinct LU classes; y_i = Spectral reflectance of respective pixels in a band; a_{ij} = Spectral reflectance of the jth component in the pixel for i^{th} spectral band; x_j = Proportion value of the jth component in the pixel; j = 1, 2, 3 ... n (number of land classes assumed); i = 1, 2, 3 ... m (number of multispectral bands) and e_i = error term for the ith spectral band. The error term (e_i) is due to the assumption made that the response of each pixel in any spectral wavelength is a linear combination of the proportional responses of each component. Assuming that the error term is 0, equation 2 can be written as:

$$AX = Y$$

Equation 3

where A is a m x n matrix (a_{11}, a_{12}, ..., a_{mn}), X is a n x 1 vector (x_1, x_2, ..., x_n) and Y is a n x 1 vector (y_1, y_2, ..., y_n) written as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Equation 4

$$X_{constrained} = (A^T A)^{-1} A^T Y + \frac{(1 - 1^T ((A^T A)^{-1} A^T Y))}{1^T (A^T A)^{-1}} (A^T A)^{-1} 1$$

Equation 5 gives the Constrained Least Squares (CLS) estimate of the abundance expressed in terms of matrix A, X and Y. The reflected spectrum of a pure feature is called a reference or endmember spectrum. Endmembers are extracted using scatter plot or through automatic endmember extraction techniques (Kumar et al., 2008).

2.3.7 LC derivation from abundance values along with training polygons using Bayesian classification

Abundances of each category (pixel wise) are used as a prior probability of the class. In the Baye's classifier for the multispectral data, the posterior probability of the class given the observation is computed by multiplying the prior probability of the class with the conditional probability P(x|k), where x denotes the multispectral observation vector and k any class. The class label assigned to the pixel is:

$$l = \arg \min_k \frac{P(k|i,j) P(x|k)}{P(x)}$$

Equation 6

3. Results

Temporal LU details are displayed in figure 2 and class statistics are listed in table 1. The classified images of 1973, 1992, 1999, 2000, 2002, 2006 and 2007 showed an overall accuracy of 72%, 75%, 71%, 77%, 60%, 73% and 55%. Accuracy assessment was performed which showed higher accuracy for high resolution data (~ 70-75% for Landsat and IRS LISS-III) and decreasing accuracy with coarse spatial resolution (~ 55-60% for MODIS). Figure 3 (a) – (f) depicts the LU change based on differencing techniques of PCA and TC. The disappearance of water bodies from 1973 to 2006 is given in Figure 4. 55% decline (from 207 to 93) in the number of water bodies and 61% decline in the spatial extent (of water bodies) is noticed from the temporal analysis. Validation was done considering training data and Google Earth image, covering approximately 15% of the study area. Then, pixels corresponding to urban category were extracted for further analysis. Figure 5 shows the LST and NDVI of Greater Bangalore in 1992, 2000 and 2007. The minimum (min) and maximum (max) temperature was 12 °C and 21 °C with a mean of 16.5±2.5 from Landsat TM (1992, winter). Similarly MODIS data of 2000 (summer) show the min, max and mean temperature of 20.23, 28.29 and 23.71±1.26 °C respectively. Corresponding values for 2007 (summer) are 23.79, 34.29 with a mean of 28.86±1.60 °C. LC wise NDVI and LST are listed in table 2.
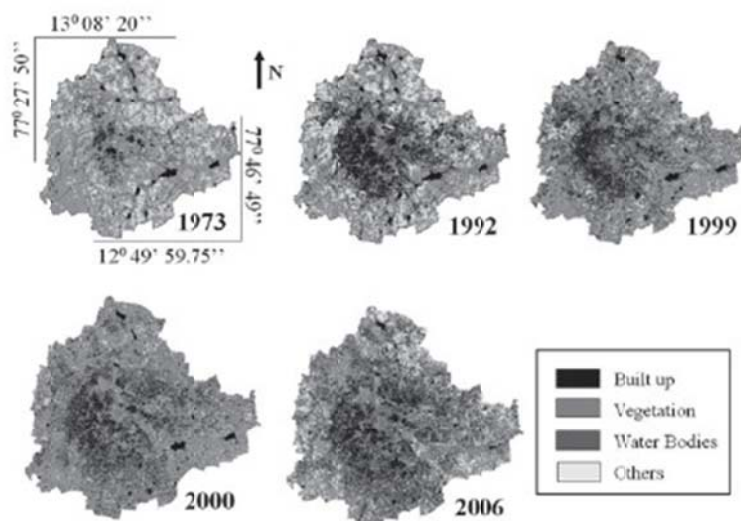
Figure 2: Greater Bangalore in 1973, 1992, 1999, 2000 and 2006
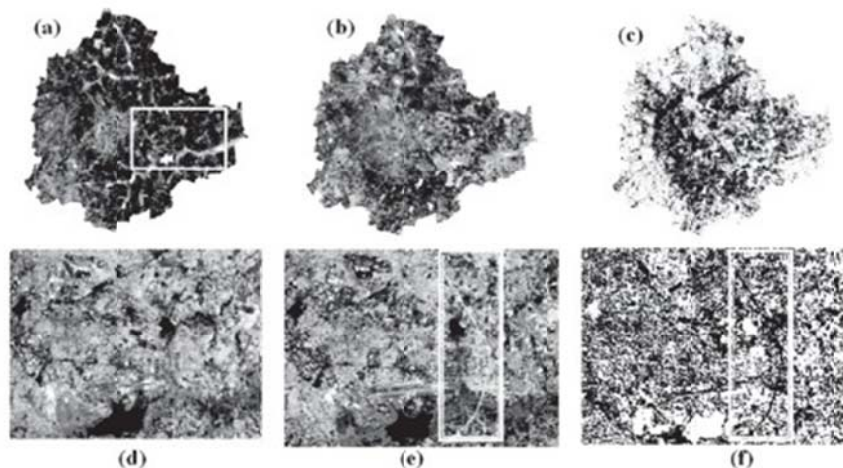
Figure 3: PC1 of Landsat MSS-1973 (a), PC1 of IRS LISS-III-2006 (b) and the change in LU is highlighted in (c) using PCA differencing methods. The highlighted box in (a) and (b) are further enlarged in (d) to (f). Brightness values from TC transformation on Landsat TM-1992 (d) and ETM+2000 (e). The changes are highlighted in (f) by differencing the brightness values emphasizing new built up

Table 1: Greater Bangalore land cover statistics

Class → Year ↓		Built up	Vegetation	Water Bodies	Others
1973	Ha	5448	46639	2324	13903
	%	7.97	68.27	3.40	20.35
1992	Ha	18650	31579	1790	16303
	%	27.30	46.22	2.60	23.86
1999	Ha	23532	31421	1574	11794
	%	34.44	45.99	2.30	17.26
2000	Ha	24163	31272	1542	11346
	%	35.37	45.77	2.26	16.61
2002	Ha	26992	28959	1218	11153
	%	39.51	42.39	1.80	16.32
2006	Ha	29535	19696	1073	18017
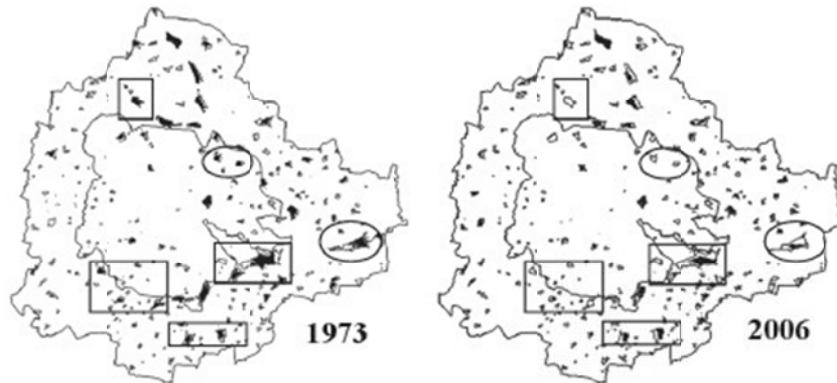
Figure 4: Temporal changes in water bodies from 1973 (using Landsat MSS) to 2006 (using IRS LISS-III) highlighted in rectangular boxes and circles.
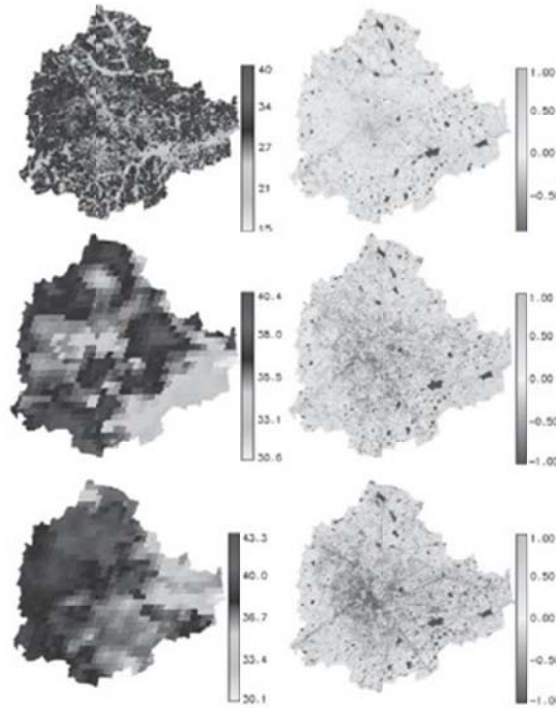
Figure 5: LST from Landsat TM (1992), MODIS (2000 and 2007), NDVI from Landsat TM (1992), Landsat ETM+ (2000) and IRS LISS-III (2006)

Table 2: NDVI and LST (°C) for respective land uses

Land cover	1992 (TM)		2000 (MODIS)		2007 (MODIS)	
	Mean LST (SD)	Mean NDVI (SD)	Mean LST (SD)	Mean NDVI (SD)	Mean LST (SD)	Mean NDVI (SD)
Builtup	19.03 (1.47)	-0.162 (0.096)	26.57 (1.25)	-0.614 (0.359)	31.24 (2.21)	-0.607 (0.261)
Vegetation	15.51 (1.05)	0.467 (0.201)	22.21 (1.49)	0.626 (0.27)	25.29 (0.44)	0.348 (0.42)
Water bodies	12.82 (0.62)	-0.954 (0.055)	21.27 (1.03)	-0.881 (0.045)	24.00 (0.27)	-0. 81 (0.27)
Open ground	17.66	-0.106	24.72	0.016	28.85	-0.007

The relationship between LST and NDVI were investigated for each LC type through the Pearson's correlation coefficient (CC) at a pixel level, which are listed in table 3. It is apparent that values tend to negatively correlate with NDVI for all LC types. NDVI values ranges from -0.05 to -0.6 (built up) and 0.15 to 0.6 (vegetation). Temporal increase in temperature with the increase in the number of urban pixels is noticed during 1992 to 2007 (63%) and 'r' confirms this relationship for the respective years. The decrease in vegetation is reflected by the respective increase in temperature. Further analysis is done by considering vegetation abundance. Landsat ETM+ (band 1, 2, 3, 4, 5 and 7) were unmixed to get the abundance maps of 5 classes (1) dense urban (commercial/industrial/residential), (2) mixed urban (urban with vegetation and open ground), (3) vegetation, (4) open ground and (5) water bodies. We considered only dense urban, mixed urban and vegetation abundance for further analysis as shown in figure 6. The min and max temperature from ETM+ data was 13.49 and 26.32 °C with a mean of 21.75±2.3. These abundance images were further analysed to see their contribution to the UHI by separating the pixels that contains 0-20%, 20-40%, 40-60%, 60-80% and 80-100% of the commercial/industrial/residential (dense urban), mixed urban and vegetation. Table 4 gives the mean and standard deviation (SD) of the LST for various LU. Application of decision based unmixing approach, systematically exploited the information from both the sources (sub-pixel class proportions and classified image based on training data collected from the ground) for achieving more reliable classification, shown in figure 7. Table 5 lists LC wise LST, NDVI and correlation coefficient. Relationship of population density with LST (Landsat ETM+) is evident in figure 8, which corroborate that the increase in LST is due to urbanisation and consequent increase in population.

DENSE URBAN   MIXED URBAN

VEGETATION   LST from ETM+
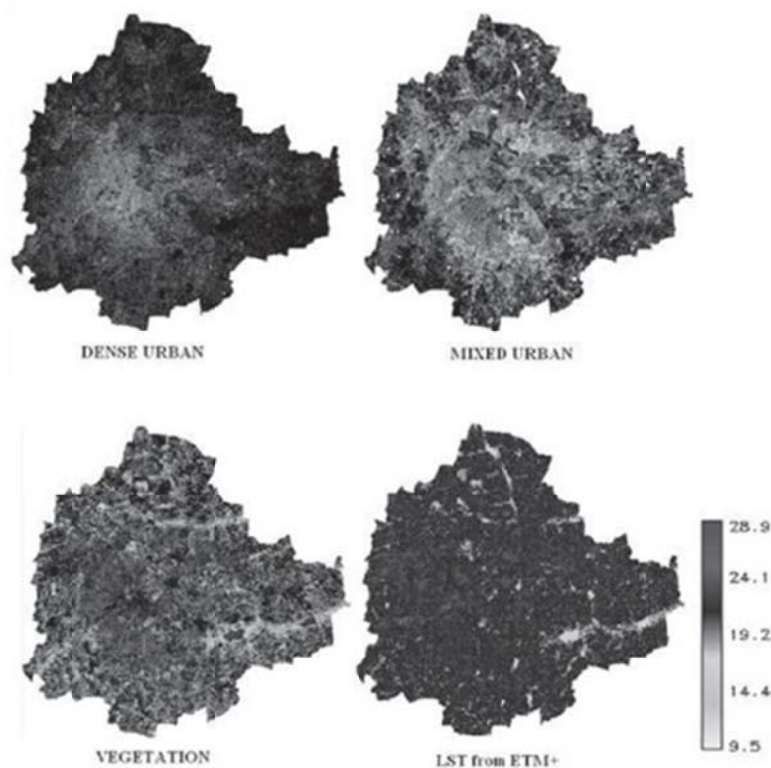
Figure 6: Abundance maps and LST obtained from Landsat ETM+ data (2000)

Table 3: Correlation coefficients between LST and NDVI by LC type (significant at 0.05 level)

LU	1992	2000	2007
Builtup	-0.7188	-0.7745	-0.7900
Vegetation	-0.8720	-0.6211	-0.6071
Open ground	-0.6017	-0.5827	-0.6004

Table 4: Mean LST for different LC classes for various abundances

Class → Abundance ↓	Mean Temperature ± SD of dense urban	Mean Temperature± SD of mixed urban	Mean Temperature± SD of vegetation
0-20%	21.99±2.37	21.57±2.36	17.91±2.19
20-40%	22.06±2.15	21.58±2.36	17.39±1.37
40-60%	22.27±2.00	21.67±2.41	17.22±0.89
60-80%	22.33±2.22	22.28±2.02	17.13±0.85
80-100%	22.47±1.96	22.37±2.17	17.12±0.91

DENSE URBAN
MIXED URBAN
VEGETATION
OPEN GROUND
WATER BODIES
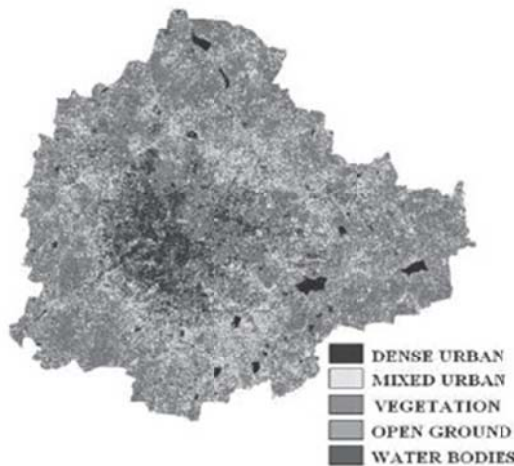
Figure 7: Classified image obtained from combining unmixed images and classified image using spectral signatures from ground as input to Baye's classifier from 6 MSS bands of Landsat ETM+

$y = 18.964x^{0.0796}$
$r = 0.572$
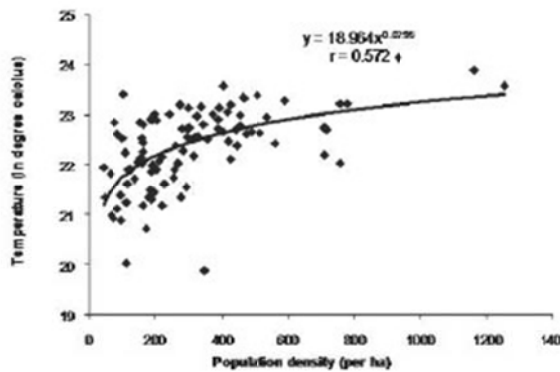
Figure 8: LST with ward wise population density

Table 5: LST, NDVI and correlation coefficient for different LC classes

Landuse	LST Mean ± SD	NDVI Mean ±SD	Correlation coefficient between LST and NDVI
Dense builtup	23.09± 1.16	-0.2904± 0.395	-0.7771
Mixed builtup	22.14± 1.06	-0.138± 0.539	-0.6834
Vegetation	19.27± 1.59	0.3969± 0.404	-0.8500
Open Ground	22.40± 1.97	-0.0193± 0.164	-0.6319
Water Bodies	19.57± 1.72	-0.301± 0.47	0.2319

Table 6: MMU sizes for different RS data sources used

Data source	MODIS	Landsat MSS	Landsat TM/ETM+	IRS LISS-III
MMU (ha)	6.25	0.62	0.09	0.055

4. Discussion

The analysis showed that there has been a 466% increase in built up area from 1973 to 2007 as evident from temporal analysis leading to a sharp decline of 61% area in water bodies. LU changes were more prominent in the city during the last 2 decades consequent to rapid urbanisation accompanied with urban sprawl with IT (Information Technology) and BT (Biotechnology)

different regions. Identification of the changed patches or specific region that underwent changes (from one LU to other) is not easily achievable in a highly dynamic and large urbanising environment considering the varying range of spatial resolutions (23.5 to 250 m) across time (1973 to 2007). This requires the development and maintenance of database for each LU across a large time-scale

A newly constructed flyover or road (for example see a new road in Figure 3 (e) running from south to north highlighted in a rectangular box in the second half of the image), was detected as a broken line in figure 3 (f) due to shadow or getting merged with other classes (mixed pixels), etc. This was obtained by differencing TC transformed 92 and 2000 brightness values and the changed patches were identified using Region Growing Method (RGM) along with thresholding. The exact delineation of the boundary for every recognised change patch also depends on the minimum mapping unit (MMU) and its proper selection (Zhang and Zhang, 2007). Large MMU changes result in significant differences in the accuracy estimates of LC classification. They make change detection more accurate but often miss real and small LU changes. A small MMU present detailed LU changes but also increase the uncertainty of the RS derived change product. MMU for various pixel sizes are listed in table 6. The minimum built up size in residential areas is ~ 0.011 ha to 0.037 ha and often extend up to 1 ha (in case of multi-storied buildings) while industrial buildings can spread from 50 to 150 ha approximately. One way to delineate such features is to perform interactive on-screen editing along with visual interpretation in a synergistic way. MODIS and Landsat MSS present significant challenge due to mixed pixel problems. Landsat TM/ETM+ and IRS LISS-III are adequate for detecting pixel patch changes, yet most of the objects fall on boundaries of two adjacent pixels and they are smaller than the MMU size (form mixed pixels). Hence, these data tend to overestimate the LU classes. The area derived from RS data is closer to the area on the ground when patch size increases. However, area measurements from high spatial resolution were more accurate. The correlation (r) between NDVI and temperature of 1992 (based on TM data) was 0.88. Similarly for 2000, r was 0.72 (MODIS 2000) and 0.65 (MODIS 2007) respectively. Analysis of LST with NDVI suggests that the extent of LC with vegetation plays a significant role in the regional LST. Although many factors can contribute to the variations of LST, the spatial arrangement and aerial extent of various LU types is a fundamental one. Within the city, if we consider the LU nature, then we see that temperature falls from densely builtup area to medium builtup area and again from medium builtup to vegetative land. The temperature close to vegetation and wetlands were lower by 4 and 5-7 °C respectively (see table 2). This highlights that LU characteristics play a significant role in maintaining the ambient temperature and also in the regional

bodies are negatively correlated with temperature suggesting that these LU aid as heat sinks and hence maintains salubrious regional climate. Accuracy of the unmixing decision based approach considering information from sub-pixel class proportions and classified map obtained using training data was 85.46%. This approach is apt for fusing the information obtained from multi-sensors (such as MODIS and IRS). MODIS based sub-pixel information obtained from one of the unmixing techniques can be fused with IRS LISS-III MSS classified information. The endmembers for MODIS data unmixing can be extracted from the image itself (Kumar et al., 2008), while IRS LISS-III MSS classification is done with the training data. This technique would be helpful in optimising the benefits of higher spectral and spatial resolutions of multi-sensors. Certain areas in the city with sparsely located building having parks and lakes in the surroundings have lower air temperature compared to completely urbanised regions. UHI is evident with the enhanced surface temperatures in urbanised landscapes compared to their surroundings (non-urbanised). Temperature profile with respect to human population density is depicted in figure 8, suggesting that dense urban areas have higher temperatures. When the population of a region exceeds its carrying capacity, it exerts pressure on the local natural resources (land, water, etc.). Correlations among LST and NDVI are the result of unique signatures of the biophysical parameters due to interactions between thermal and vegetation dynamics in each LU.

5. Conclusion

This study analysed the role of increased urbanisation through spatial change analysis on LST, which is positively correlated with urbanising landscapes and negatively correlated with vegetation and water bodies. Abundance maps with pixels having 80-100% dense builtup and mixed builtup class proportions show increase in urban temperature by an average of 2 °C. LST were comparatively lower in areas having parks, healthy vegetation and lakes that aid as heat sinks. Although the algorithms adopted here seems to work satisfactorily, the classification and change detection methods in use were not very suitable for handling multi-sensor, multi-resolution condition as the classification accuracies obtained from MODIS data were low. These require advanced image fusion and LU classification algorithms or machine learning techniques to be adopted to achieve higher classification accuracies. The unmixing decision based approach adopted in this paper is an attempt

accuracies per class. This could be automated to obtain LU classes and bring out changes across multi-temporal data routinely. Methods of change detection from multi-resolution images integrating spectral, structural and textural features to generate changed patches and change attribute is also desirable and a challenging area of research.

References

Artis, D. A., and Carnahan, W. H., 1982, Survey of Emissivity Variability in Thermography of Urban Areas. *Remote Sensing of Environment*, 12, 13-329.

Balling, R. C., and Brazell, S. W., 1988, High Resolution Surface Temperature Patterns in a Complex Urban Terrain. *Photogrammetric Engineering and Remote Sensing*, 54, 289-1293.

Carnahan, W. H., and Larson, R. C., 1990, An analysis of an urban heat sink. *Remote Sensing of Environment*, 33, 65-71.

Dash, P., Gottsche, F. M., Olesen, F. S., and Fischer, H., 2002, Land Surface Temperature and Emissivity Estimation from Passive Sensor Data: Theory and Practice-Current Trends. *International Journal of Remote Sensing*, 23(13), 2563–2594.

Duda, R. O., Hart, P. E., and Stork, D. G., 2000, *Pattern Classification*, A Wiley-Interscience Publication, Second Edition, ISBN 9814-12-602-0.

Forghani, A., Cechet, B., and Nadimpalli, K., 2005, Object-Based Classification of Multi-Sensor Optical Imagery to Generate Terrain Surface Roughness Information for Input to wind Risk Simulation. *Proceedings of the Geosciences and Remote Sensing Symposium*, 2007, IGARSS '07, 23-27 July, 2007, Barcelona, Spain, IEEE International Publication, 3090-3095.

Friedl, M. A., 2002, Forward and Inverse Modeling of Land Surface Energy Balance using Surface Temperature Measurements. *Remote Sensing of Environment*, 79, 344-354.

Fung, T., and LeDrew, E., 1987, Application of Principal Component Analysis for Change Detection. *Photogrammetric Engineering and Remote Sensing*, 53(12), 1649-1658.

Gallo, K. P., and Owen, T. W., 1998, Assessment of Urban Heat Island: A Multi-Sensor Perspective for the Dallas-Ft. Worth, USA Region. *Geocarto International*, 13, 35– 41.

Gallo, K. P., McNab, A. L., Karl, T. R., Brown, J. F., Hood, J. J., and Tarpley, J. D., 1993, The use of NOAA AVHRR Data for Assessment of the Urban Heat Island Effect. *Journal of Applied*

Kato, S., and Yamaguchi, Y., 2005, Analysis of Urban Heat-Island Effect using ASTER and ETM+ Data: Separation of Anthropogenic Heat Discharge and Natural Heat Radiation from Sensible Heat Flux. *Remote Sensing of Environment*, 99, 44–54.

Kauth, R. J., and Thomas, G. S., 1976, The Tasseled Cap-A Graphic Description of The Spectral-Temporal Development of Agricultural Crops as seen by LANDSAT. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Purdue University of West Lafayette, Indiana, 4B-41-4B – 51.

Kidder, S. Q., and Wu, H. T., 1987, A Multispectral Study of the St. Louis Area Under Snow-Covered Conditions using NOAA-7 AVHRR Data. *Remote Sensing of Environment*, 22, 159–172.

Kumar, U., Kerle, N., and Ramachandra T. V., 2008, Contrained Linear Spectral Unmixing Technique for Regional Land Cover Mapping using MODIS Data. *In: Innovations and Advanced Techniques in Systems, Computing Sciences and Software Engineering*, Edited by Khaled Elleithy. Berlin: Springer, ISBN: 978-1-4020-8734-9, 87 - 95.

Landsat 7 Science Data User's Handbook, Landsat Project Science Office, Goddard Space Flight Center, 2002, URL: http://ltwww.gsfc.nasa.gov-/IAS/handbook/handbook_toc.html (last accessed: 25 March, 2008).

Landsberg, H. E., 1981, *The Urban Climate*, (New York: Academic Press).

Li, F., Jackson, T. J., Kustas, W., Schmugge, T. J., French, A. N., Cosh, M. L., and Bindlish, R., 2004, Deriving land surface temperature from Landsat 5 and 7 during SMEX02/SMACEX. *Remote Sensing of Environment*, 92, 521 – 534.

Lu, D., Mausel, P., Brondizio, E., and Moran, E., 2004, Change Detection Techniques. *International Journal of Remote Sensing*, 25(12), 2365-2407.

Macleod, R. D., and Congalton, R. G., 1998, A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing*, 64(3), 207-216.

Markham, B. L., and Barker, J. K., 1985, Spectral Characteristics of the LANDSAT Thematic Mapper Sensors. *International Journal of Remote Sensing*, 6, 697-716.

Michener, W. K., and Houhoulis, P. F., 1997, Detection of Vegetation Associated with Extensive Flooding in a Forest Ecosystem.

Photogrammetric Engineering and Remote Sensing, 63(12), 1363-1374.

Nikolakopoulos, K. G., Vaiopoulos, D. A., and Skianis, G. A., 2003, Use of Multitemporal Remote Sensing Thermal Data for the Creation of Temperature Profile of Alfios River Basin, *Proceedings of the Geoscience and Remote Sensing Symposium, 2003, IGARSS '03*, 21-25 July 2003. IEEE International Publication, 4:2389-2391, ISBN: 0-7803-7929-2.

Ramachandra, T. V., and Kumar, U., 2008, Wetlands of Greater Bangalore, India: Automatic Delineation through Pattern Classifiers. *Electronic Green Journal*, 26, WebURL: http://egj.lib.uidaho.edu/index.php/egj/article/view/3171

Schmugge, T., Hook, S. J., and Coll, C., 1998, Recovering Surface Temperature and Emissivity from Thermal Infrared Multispectral Data. *Remote Sensing of Environment*, 65, 121– 131.

Snyder, W. C., Wan, Z., Zhang, Y., and Feng, Y. Z., 1998, Classification Based Emissivity for Land Surface Temperature Measurement from Space. *International Journal of Remote Sensing*, 19, 2753– 2774.

Stathopoplou, M., Cartalis, C., and Petrakis, M., 2007, Integrating CORINE Land Cover Data and landsat TM for Surface Emissivity Definitions: An Application for the Urban Areas of Athens, Greece. *International Journal of Remote Sensing*, 28, 3291-3304.

Stathopoulou, M., and Cartalis, C., 2007, Daytime Urban Heat Island from Landsat ETM+ and Corine land Cover Data: An application to Major Cities in Greece. *Solar Energy*, 81, 358-368.

Tanaka, Y., Shibata, S., and Gotoh, K., 2005, Appearance Characteristic Analysis of Heat Island Phenomenon by Using Satellite Remote Sensing and GIS. *Proceedings of the Geosciences and Remote Sensing Symposium, 2005, IGARSS '05*, 25-29 July 2005, IEEE International Publication, 3:1855-1858, ISBN: 0-7803-9050-4.

Weng, Q., Lu, D., and Schubring, J., 2004, Estimation of Land Surface Temperature – Vegetation Abundances Relationship for Urban Heat Island Studies. *Remote Sensing of Environment*, 89, 467-483.

WUP (World Urbanization Prospects): The 2005 Revision, Population Division, Department of Economic and Social Affairs, UN.

Zhang, J., and Zhang, Y., 2007, Remote Sensing Research Issues of the National Land Use Change Program of China. *Photogrammetry & Remote Sensing*, 62(6), 461-472.

Hybrid Bayesian Classifier for Improved Classification Accuracy

Uttam Kumar, S. Kumar Raja, C. Mukhopadhyay, and T. V. Ramachandra, *Senior Member, IEEE*

Abstract— **Widely used Bayesian classifier is based on the assumption of equal prior probabilities for all the classes. However, inclusion of equal prior probabilities may not guarantee high classification accuracy for the individual classes. Here, we propose a novel technique - *Hybrid Bayesian Classifier* (HBC), where the class prior probabilities are determined by unmixing a supplement low spatial-high spectral resolution multispectral (MS) data that are assigned to every pixel in a high spatial-low spectral resolution MS data in Bayesian classification. This is demonstrated with two separate experiments – firstly, class abundances are estimated per pixel by unmixing MODIS data to be used as prior probabilities while posterior probabilities are determined from the training data obtained from ground. These have been used for classifying the IRS LISS-III MS data through Bayesian classifier. In the second experiment, abundances obtained by unmixing Landsat ETM+ are used as priors and posterior are determined from the ground data to classify IKONOS MS images through Bayesian classifier. The results indicated that HBC systematically exploited the information from two image sources improving the overall accuracy of LISS-III MS classification by 6% and IKONOS MS classification by 9%. Inclusion of prior probabilities increased the average producer's and user's accuracy by 5.5% and 6.5% in case of LISS-III MS with 6 classes and 12.5% and 5.4% in IKONOS MS for the 5 classes considered.**

Index Terms— **prior probability, unmixing, Bayesian classifier**

INTRODUCTION

IMAGE classification using conventional Bayesian classifier is a supervised method based on the prior probabilities [1]. Prior probability resolves confusions among classes that are not well separable[2]and is therefore effective in improving classification accuracy. Bayesian classifier is generally used with the assumption that prior probabilities are equal, asreliable prior probabilities are not easily available. However, with the use of equal prior probabilities, the performance of the classifier is not optimal [3] which is evident from the misclassified pixels primarily due to spectral confusion between classes, apart from the increased computing and sampling requirements[4]. The theoretical analysis of the effect of prior probability has been discussed in detail by Z. Mingguo et al., (2009) [5]. Earlier works used prior probabilities based on previous year crop statistics[6], geographical data [7],elevation data[8] and spatial characteristics specified through a Markov random field model at reference resolution [9], improving the overall accuracy and Kappa coefficient. Therefore, it is desirable to obtain reliable prior probabilities for each spectral class and use them to classify the pixels that are likely to misclassify[1], even though they are difficult to determine within the same time period and for the same spectral classes.

In this context, a new classification method - *Hybrid Bayesian classifier* (HBC)based on linear unmixing and Bayesian classifier is proposed to assign a class label to each pixel in a *high spatial-low spectral resolution* (HS-LSR) multispectral (MS) data. The prior probabilities of the different classes in the Bayesian classifier to classify each pixel in the HS-LSR data such as IRS LISS-III MS and IKONOS MS are obtained from abundance estimates by unmixing *low spatial-high spectral resolution* (LS-HSR) MS data such as MODIS and Landsat ETM+ respectively. The terms HS-LSR and LS-HSR have been used in a relative sense here, depending upon the spatial resolution of the images. For example, in the first experiment, MODIS bands are referred as LS-HSR and IRS LISS-III MS bands are HS-LSR data, while in the second experiment, Landsat ETM+ bands are LS-HSR and IKONOS MS bands are HS-LSR data.

The reason for selecting MODIS and Landsat ETM+ images as LS-HSR supplement data is because of their economic viability and high temporal resolution that enables their procurement for any part of the globe, throughout the year corresponding to any HS-LSR data. However, the technique in general, can be applied on any other LS-HSR (such as hyperspectral bands) and HS-LSR (such as multispectral bands) data classification. The novelty of this approach lies in the fact that low spatial-high spectral and low spectral-high spatial resolution MS data are combined to improve the classification results which can be thought of as a fusion process, in the sense, that information from two
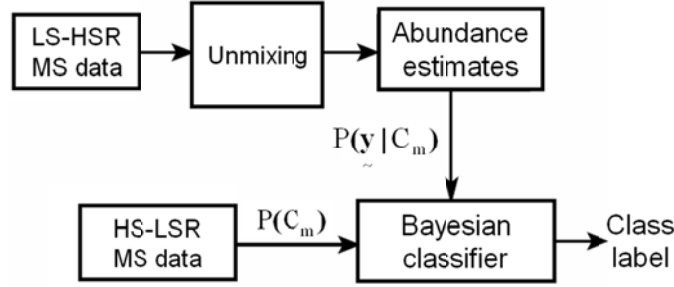
Fig. 1.Hybrid Bayesian Classifier.

different sources (sensors) are combined to arrive at a improved classified image by systematically exploiting the relevant information from both the sources as shown in Fig. 1. Zhukov et al., (1999) [10] attempted a similar multisensormultiresolution image fusion based on low spatial resolution (LSR) unmixing and high spatial resolution (HSR) classification. However, the end product was not a classified image, but a set of MS fused images equivalent to the number of LSR bands at the HSR image dimension. In the following section, linear unmixing using Orthogonal Subspace Projection (OSP) and Bayesian classifier are reviewed. The proposed HBC is discussed in Section III while section IV demonstrates the experimental results and validation followed by conclusion in Section V.

Review of Methods

Linear Unmixing:*With K spectral bands and M classes ($C_1,..,C_M$), each pixel has an associated K-dimensional pixel vector $y = (y_1,...,y_K)'$ whose components are the gray values corresponding to the K spectral bands. Let $E = [e_1,...,e_M]$, where, for $m = 1,...,M$, e_m is a Kx1 column vector representing the endmember spectral signature of the mth target material. For a pixel, let α_m denote the fraction of the mth target material signature, and $\alpha = (\alpha_1,...,\alpha_M)'$ denote the M-dimensional abundance column vector. The linear mixture model for y is given by*

$$y = E\alpha + \eta \qquad (1)$$

where, $\eta = (\eta_1,...,\eta_K)'$, are i.i.d. $N(0,\sigma^2)$ [11]. Equation (1) represents a standard signal detection model where $E\alpha$ is a desired signal vector to be detected. Since, OSP detects one signal (target) at a time, we divide a set of M targets into desired (C_m) and undesired ($C_1,...,C_{m-1},C_{m+1},...C_M$) targets. A logical approach is to eliminate the effects of undesired targets that are considered as "interferers" to C_m before the detection of C_m takes place. Now, in order to find α_m, the desired target material is e_m. The term $E\alpha$ in (1) can be rewritten to separate the desired spectral signature e_m:

$$y = e_m\alpha_m + Rr + \eta \qquad (2)$$

where $r = (\alpha_1,...\alpha_{m-1},\alpha_{m+1},...,\alpha_M)$ and $R = [e_1,...,e_{m-1},e_{m+1},...,e_M]$. Thus the interfering signatures in R can be removed by the operator,

$$P = (I - R(R^T R)^{-1} R^T)$$ (3)

which is used to project y into a space orthogonal to the space spanned by the interfering spectral signatures, where I is the KxK identity matrix [12]. Operating on y with P, and noting that $PR = 0$,

$$Py = Pe_m\alpha_m + P\eta.$$ (4)

After maximizing the SNR, an optimal estimate of α_m is

$$\hat{\alpha}_m = \frac{y^T P^T Py}{e_m^T P^T Pe_m}$$ (5)

Note that $1 \geq \alpha_m \geq 0$ and thus $(\alpha_1,...,\alpha_M)$ can be taken as proportional probabilities of the M classes i.e. $P(C_m) \propto \alpha_m$.

Bayesian classifier: ***Associated with any pixel, there is an observation $\underset{\sim}{y}$. With M classes $(C_1,...,C_M)$, Bayesian classifier calculates the posterior probability of each class conditioned on y [13]:***

$$P(C_m|\underset{\sim}{y})=\frac{P(\underset{\sim}{y}|C_m)P(C_m)}{P(\underset{\sim}{y})}$$ (6)

In (6), since $P(\underset{\sim}{y})$ is constant for all classes, only $P(\underset{\sim}{y}|C_m)P(C_m)$ is considered. $P(\underset{\sim}{y}|C_m)$ is computed assuming class conditional independence, so, $P(\underset{\sim}{y}|C_m)$ is given by $P(\underset{\sim}{y}|C_m)=\prod_{k=1}^{K} P(y_k|C_m)$ (7)

New Hybrid Bayesian Classifier (HBC)

HBC uses the abundance of each class obtained from LS-HSR data by linear unmixing as prior probability while classifying the HS-LSR data using Bayesian classifier of the same geographical area and time frame. That is, given the observation vector $\underset{\sim}{y}$ for a pixel, it is classified to fall in class *l* if

$l = \underset{m}{Arg\ Max}\ P(\underset{\sim}{y}|C_m)P(C_m)$, where $P(\underset{\sim}{y}|C_m)$ is as in (7) calculated using the HS-LSR data and $P(C_m) \propto \alpha_m$ calculated using the LS-HSR data. The assumptions in this method are (i) if there are r HS-LSR pixels contained in one LS-HSR pixel, i.e. resolution ratio is (r:1), the prior probabilities for all the r HS-LSR pixels are equal corresponding to the same LS-HSR pixel, and (ii) the two data types have a common origin or upper left corner, i.e. the edges of the r x r HS-LSR pixels overlaps exactly with the corresponding LS-HSR pixel. The limitations are (i) (K-1) should be ≥ M in LS-HSR data and (ii) M in HS-LSR should be ≤ M in LS-HSR data.

Experimental Results

Two separate experiments were carried out. In the first case, LISS-III MS data (3 bands of 23.5m x 23.5m spatial resolution resampled to 25m, acquired on December 25, 2002) of 5320 x 5460 size and MODIS 8-day composite data (7 bands of 250m x 250m, acquired from 19-26 December, 2002) of 532 x 546 dimensionwere co-registered with known ground control points (RMSE - 0.11). Training data were collected from the ground representing approximately 10% of the study area covering the entire spectral gradient of the classes. Separate test data were collected for validation. LISS-III MS classified image using conventional Bayesian classifier is shown in Fig. 2 (a). Assuming that there are six (fixed) number of representative endmembers (pure pixels), the entire image was modeled in terms of those spectral components, extracted using N-FINDR algorithm [14] from MODIS images. In the absence of pure pixels, alternative algorithms [15] can be used for endmember extraction, which is a limitation of N-FINDR. It may be noted that some objects (for example buildings with concrete roofs, tiled roofs, asphalt, etc.) exhibit high degrees of spectral heterogeneity representing variable endmembers. This intra-class spectral variation with variable endmembers can be addressed through techniques discussed in [16-18], and is beyond the scope of the current study.

Abundance values were estimated for each pixel through unmixing and used as prior probabilities in HBC to classify LISS-III MS data (Fig. 2 (b)). Table I is the class statistics and Table II indicates the producer's and user's accuracies. The overall accuracy and Kappa for HBC (93.54%, 0.91) is higher than Bayesian classifier (87.55%, 0.85). For any particular class, if the reference data has more pixels with correct label, the producer's accuracy is higher and if the pixels with the incorrect label in classification result is less, its user's accuracy is higher [5]. Bayesian classifier wrongly classified many pixels belonging to waste / barren / fallow as builtup. Forest class was over-estimated and plantation was under-estimated by Bayesian classifier. The only minority class in the study area is water bodies. Classified image using conventional Bayesian classifier had 1.08% (8854 ha) of water bodies in the study area. After the ground visit, we found that there are not many water bodies and the extents of most individuals were $< 2000m^2$. Therefore, only a few water bodies that had spatial extent ≥ 62500 m^2 could be used as endmembers. The minimum detected water class was 5% using unmixing of LS-HSR data. It may have happened that the prior probability was unlikely for this class while classifying the LISS-III MS data using HBC. The classified image obtained from HBC showed 0.81% (66.20 ha) of water bodies. A few pixels were wrongly classified using HBC, therefore, the producer's accuracy decreased from 90.91 (in conventional Bayesian classifier) to 88.18% (in HBC). However, given the same set of training pixels for classification, the user's accuracy has increased from 88.89 to 97%. Producer's accuracy increased for agriculture (2.6%), builtup (4.3%), forest (7%), plantation (11.5%) and waste land (10.6%) and user's accuracy increased for agriculture (8%),
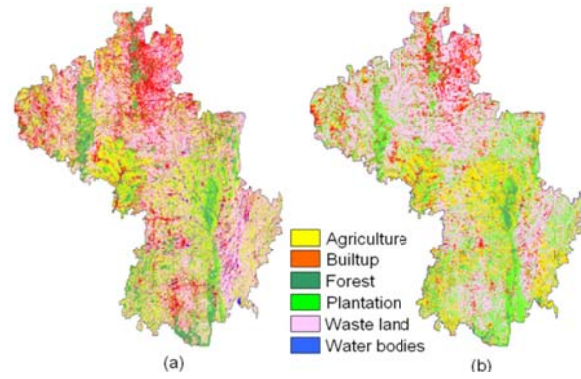
Fig. 2. LISS-III MS classified images: (a) Bayesian classifier, (b) HBC.

TABLE I
CLASS STATISTICS FROM BAYESIAN AND HBC FOR LISS-III DATA

Classifiers →	Bayesian classifier		HFC	
Class ↓	Ha	%	Ha	%
Agriculture	155, 451	19.04	142931	17.51
Builtup	139, 759	17.12	79280	9.71
Forest	93, 241	11.42	55721	6.83
Plantation	89,493	10.96	176132	21.58
Waste land	329, 473	40.36	355587	43.56
Water bodies	8854	1.08	66.20	0.81

TABLE II
ACCURACY ASSESSMENT FOR LISS-III DATA

Classifiers →	Bayesian classifier		HBC			
Class ↓	PA*	UA*	PA*		UA*	
Agriculture	87.54	87.47	90.15	↑	95.56	↑

Builtup	85.11	81.68	89.39	↑	98.33	↑
Forest	85.71	88.73	92.61	↑	96.36	↑
Plantation	84.44	91.73	95.95	↑	91.03	↓
Waste land	88.03	90.37	98.67	↑	89.66	↓
Water bodies	90.91	88.89	88.18	↓	97.00	↑
Average	86.96	88.15	92.49	↑	94.66	↑

* PA – Producer's Accuracy; UA – User's Accuracy.

builtup (16.6%), forest (7.6%) and water bodies (8%) in HBC output. On the other hand, producer's accuracy decreased (2.7%) for water bodies and user's accuracy decreased (~0.7%) for plantation and waste land classes in agreement with the similar observations reported in [5]. HBC was intended to improve classification accuracies by correctly classifying pixels which were likely to be misclassified by Bayesian classifier. Therefore a cross comparison of the two classified images located the pixels that were assigned different class labels at the same location. These wrongly classified pixels when validated with ground data revealed a 6% (~1742832 pixels) improvement in classification by HBC.

In the second experiment, IKONOS MS data (4 bands of 4m, acquired on November 24, 2004) of 700 x 700 size and Landsat ETM+ data (6 bands excluding Thermal and Panchromatic, of 30m, acquired on November 22, 2004) of 100 x 100 dimension were co-registered (RMSE - 0.09). Landsat pixels were resampled to 28m so that 49 IKONOS pixels would fit in 1 Landsat pixel. The scenes correspond to Bangalore city, India near the central business district having race course, bus stand, railway lines, parks, builtup with concrete roofs, asbestos roof, blue plastic roofs, coal tarred roads with flyovers and a few open areas (play ground, walk ways, vacant land, etc.). Class proportions from Landsat image pixels were used as prior probabilities to classify the IKONOS MS images using HBC (Fig. 3 (b)). The overall accuracy and Kappa for HBC (89.53%, 0.87) is higher than Bayesian classifier (80.46%, 0.69). Producer's accuracy increased by 17% for concrete, asbestos, blue plastic roof and open area and decreased by 7% for vegetation in HBC (Table IV). User's accuracy increased by ~5.5% for all the classes in HBC. Bayesian classifier wrongly classified and overestimated many pixels belonging to open area as asbestos roof (Table III). Vegetation was overestimated by 9% using Bayesian classifier and open area was underestimated by ~15%. A cross comparison of the two classified images showed that 95733 pixels (0.33% of the study area) were differently classified by the two classifiers. Validation of these pixels showed an improvement of 9% by HBC, higher than accuracies reported in [6, 7 and 8].
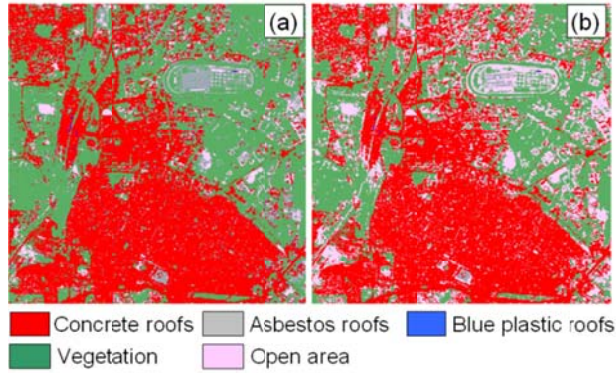
Fig. 3. IKONOS classified images: (a) Bayesian classifier, (b) HBC.

Conclusion

The major contribution of this technique lies in the fact that abundance estimates from LS-HSR data were utilized as prior probabilities to classify HS-LSR data using a Bayesian classifier improving the overall accuracy by 6% and 9% with IRS LISS-III MS and IKONOS MS data respectively, as compared to conventional Bayesian classifier, demonstrating the robustness of the approach.

REFERENCES

[31] J. Ediriwickrema and S. Khorram, "Hierarchical Maximum-Likelihood Classification for Improved Accuracies," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 810-816, 1997.

[32] D. K. McIver and M. A. Friedl, "Using prior probabilities in decision-tree classification of remotely sensed data," *Remote Sens. Environ.*, vol. 81, pp. 253-261, 2002.

[33] C. Lee and D. A. Landgrebe, "Fast likelihood classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 4,pp. 509–517, 1991.

[34] L. Pedroni, "Improved classification of Landsat Thematic Mapper data using modified prior probabilities in large and complex landscapes," *Int. J. Remote Sens.*, vol.24, no. 1, pp. 91-113, 2003.

[35] Z. Mingguo, C. Qiangguo and Q. Mingzhou, "The Effect of Prior Probabilities in the Maximum Likelihood Classification on Individual Classes: A Theoretical Reasoning and Empirical Testing," *PhotogrammEng. Remote Sens.*, vol. 75, no. 9, pp. 1109-1116, 2009.

TABLE III
CLASS STATISTICS FOR IKONOS DATA

Classifiers →	Bayesian classifier		HFC	
Class ↓	Ha	%	Ha	%
Concrete roofs	346.67	44.34	356.18	45.56
Asbestos roofs	47.99	7.41	6.79	0.87
Blue plastic roof	5.83	0.75	0.21	0.03
Vegetation	329.72	42.18	260.01	33.26
Open area	41.60	5.32	158.58	20.28

TABLE IV
ACCURACY ASSESSMENT FOR IKONOS DATA

Classifiers →	Bayesian classifier	HFC

Class ↓	PA	UA	PA		UA	
Concrete roofs	69.99	84.01	76.49	↑	93.89	↑
Asbestos roofs	84.77	87.77	91.89	↑	94.46	↑
Vegetation	94.21	87.55	87.24	↓	89.13	↑
Blue plastic roof	84.33	81.17	97.00	↑	85.60	↑
Open area	51.49	69.49	95.00	↑	74.22	↑
Average	76.96	81.99	89.52	↑	87.46	↑

[36] L. L. F. Janssen and H. Middelkoop, "Knowledge-based crop classification of a Landsat Thematic Mapper image," *Int. J. Remote Sensing*, vol. 13, pp. 2827–2837, 1992.

[37] H. Cetin, T. A. Warner and D. W. Levandowski, "Data classification, visualization, and enhancement using n-dimensional probability density functions (NPDF): AVIRIS, TIMS, TM, and geophysical applications," *Photogramm. Eng. Remote Sens.*, vol. 59, pp. 1755–1764, 1993.

[38] A. H. Strahler, "The use of prior probabilities in maximum likelihood classification of remotely sensed data," *Remote Sens. Environ.*, vol. 10, pp. 135–163, 1980.

[39] G. Storvik, R. Fjortoft, and A. H. S. Solberg, "A Bayesian Approach to Classification of Multiresolution Remote Sensing Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 539–547, 2005.

[40] B. Zhukov, D. Oertel, F. Lanzl, and G. Reinhackel, "Unmixing-based MultisensorMultiresolution Image Fusion" *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1212–1226, 1999.

[41] U. Kumar, N. Kerle, and T. V. Ramachandra, "Constrained linear spectral unmixing technique for regional land cover mapping using MODIS data," *In: Innovations and advanced techniques in systems, computing sciences and software engineering/ed by KhaledElleithy*, Berlin: Springer, pp. 87-95, 2008.

[42] C-I. Chang, "Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502–518, 2005.

[43] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, San Francisco, CA, USA, 2003.

[44] M. E. Winter, "N-Findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," *In Proceedings of the SPIE: Imaging Spectrometry*, vol. 3753, pp. 266-275, 1999.

[45] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A Quantitative and Comparative Analysis of Endmember Extraction Algorithms FromHyperspectral Data" IEEE Trans. Geosci. Remote Sens., vol. 42, no. 3, pp. 650–663, 2004.

[46] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember Bundles: A New Approach to Incorporating Endmember Variability into Spectral Mixture Analysis" *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 1083–1094, 2000.

[47] C. Song, "Spectral mixture analysis for sub-pixle vegetation fractions in the urban environment: How to incorporate endmember variability?" *Remote Sens.* Environment, vol. 95, pp. 248–263, 2005.

[48] G. M. Foody, and H. T. X. Doan, "Variability in Soft Classification Prediction and its Implications for Sub-pixel Scale Change Detection and Super Resolution Mapping" *Photogrammetric Engineering & Remote Sensing*, vol. 73, no. 8, pp. 923–933, 2007.

FOSS @ IISC

http://wgbis.ces.iiscernet.in/foss
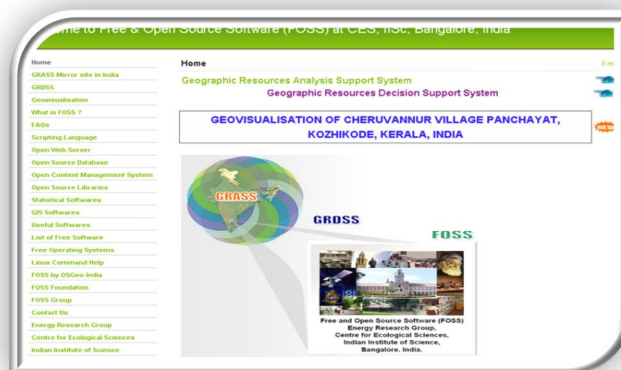
Dr. T. V. Ramachandra,

Geoinformatics lab,

Centre for *Infrastructure*, Sustainable Transportation and Urban Planning (C*i*STUP)

Indian Institute of Science, Bangalore 560012

Tel: 91-080- 22933099 / 22932786 / 2293 2506 / 23600985 (Extn: 215/232) E mail: cestvr@ces.iisc.ernet.in ,

grass@ces.iisc.ernet.in,