CHAPTER 1

INTRODUCTION

1.1 CONCEPT OF REMOTE SENSING

Remote sensing is the science and technology of acquiring information about an object, area, or phenomena without being in physical contact with it. Generally, it consists of three parts, the targets either objects or a phenomena in an area; the acquisition of data by certain instruments; and the data analysis by some computing systems. The definition is so broad including the visual scheme of human eyes, sonar sounding of the sea level, ultrasound and x-rays in medical sciences and, laser scanning of atmospheric constituents. The human eyes are a magnificent example of a remote sensing device. We can collect information about our surroundings by measuring the amount and nature of the reflected energy of visible light from some external source (for example the sun or a light bulb) as it reflects off objects in our field of view.

In remote sensing, the electromagnetic radiation (EMR) is used as an information carrier from the target to the sensor. The sun is a source of EMR or energy to illuminate the target of interest. The EMR is considered to span the spectrum of shorter wavelengths (including gamma and X-rays) to the longer wavelengths (including microwaves and broadcast radio waves). Remote sensing process involves an interaction between incident electromagnetic energy and the targets of interest. The energy is reflected, transmitted or emitted by the target of interest and recorded by the sensor. Remote sensors are mounted on platforms such as satellites, space shuttles, aircraft, hot air balloons, and rockets, etc. The sensors include aerial photographic cameras, multi-spectral scanners, RADAR, LIDAR, etc. The complete remote sensing process is shown in Figure 1.1.

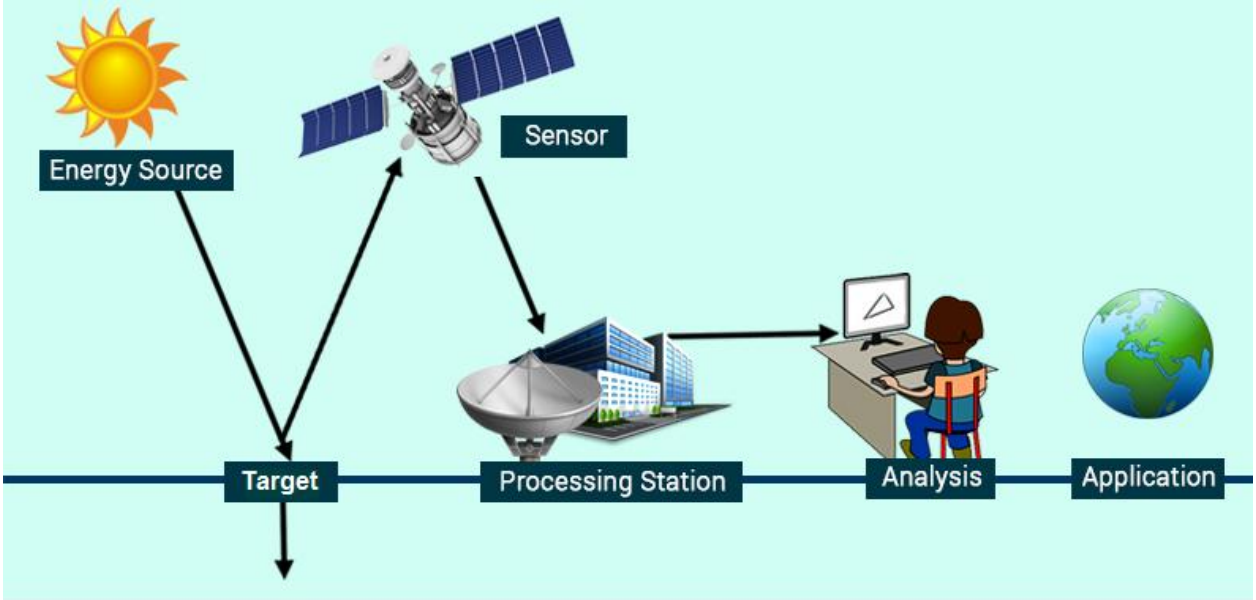

Figure 1.1 Remote sensing process

(**Source:** http://byjus.com/physics/remote-sensing)

1.2 BRIEF HISTORY OF REMOTE SENSING

Remote sensing has been with us for the longer time. In the 1600s, Galileo used optical enhancements to survey celestial bodies. In 1858, Parisian photographer Gaspard Felix Tournachon used his balloon to made photographs over Bievre, France. In later years, kites, messenger pigeons, rockets and unmanned balloons were also used for the acquisition of images. The modern regime of remote sensing arose with the development of aircraft and satellites. The field of remote sensing has experienced some major changes during the period from 1960-2010.

The word remote sensing was first introduced in 1960s and before that, it was generally termed as aerial photography. During 1960s to 1970s, the airborne platforms carrying remote sensing devices were moved to the spaceborne platforms or satellites. It facilitates observations across larger extents of Earth's surface than is possible by traditional ground-based observations at regular interval. During the 1960s, National Aeronautics and Space

Administration (NASA) sponsored several projects to study the application of color infrared and multispectral imaging. The highly successful Landsat program, a series of earth observing satellite missions was launched jointly by NASA and United States Geological Survey in the 1970s. It is the longest-running project for the acquisition of satellite imagery of the Earth. The Lansat-1, originally Earth Resources Technology Satellite (ERTS-1) was launched on 23 July 1972. The most recent, Landsat 8 was launched on 11 February 2013. The Landsat 9, with launch scheduled for late 2020 is in development phase. Many other remote sensing satellites have been flown, including Systeme Probatorie d'Observation de la Terre (SPOT) in 1986, European Radar Satellite in 1991, IKONOS in 1991), OrbView-2 in 1997. The Indian remote sensing (IRS) satellite program was launched in 1988. Subsequently followed by Resourcesat-1 in 2003, Cartosat 2A in 2008, Oceansat-2 in 2009, Resourcesat-2 in 2011, RISAT-1 in 2012, SARAL in 2013 etc. In the mid-1950s, efforts have been made for the development of Synthetic Aperture Radar (SAR) sensors using coherent signals to achieve high-resolution imaging capability from the high-flying aircraft. Some well known SAR sensors are ERS-1/2, RADARSAT-1/2, ALOS-PALSAR, ENVISAT-ASAR, Sentinel-1A etc. In the beginning of 1986 a new type of sensors named as imaging spectrometers were developed including AVIRIS, HySI, HSI, Hyperion etc. The number of spectral bands available had grown from a few to more than 200 in this type of sensing instruments. The potential of remote sensing satellites have been increased considerably over the past two decades. The satellites having spatial resolutions of a few meters or less are now available from commercial vendors. The SAR sensors are now capable of collecting images on demand in various different modes. Satellites are now able to acquire the images of extraterrestrial surfaces in channels having more spectral bands with

better spatial resolution. Moreover, the number of applications has grown as the increasing availability of remote sensing data (Jensen, 2005; Elachi and Van Zyl, 2006; Ulaby et al., 2014).

1.3 TYPES OF REMOTE SENSING

In remote sensing, the sensor measures electromagnetic energy. The sensors can be broadly categorized into two groups: *passive and active*.

1.3.1 Passive remote sensing

In the passive remote sensing, the sun, an ambient energy source is employed for illuminating the target of interest. The sensor detects natural radiation that is either reflected or emitted from the earth's surface. The passive sensor system detects reflected energy only when the sun is illuminating the earth. The reflected energy is unavailable from the sun at night time. The energy that is naturally emitted (such as thermal infrared) can also be detected in day or night time. Passive sensor systems need to deal with the varying illumination condition of the sun, which are significantly affected by the atmospheric conditions. The initial example of passive sensor is film photography. Other passive sensors are available including gamma-ray spectrometer, aerial camera, multispectral scanner, imaging spectrometer, thermal scanner, radiometer etc. (Lillesand and Kiefer, 2002; Elachi and Van Zyl, 2006). The passive remote sensing process is shown in Figure 1.2 (a).

1.3.2 Active remote sensing

Active remote sensing techniques have their own source of energy to illuminate the target of interest. An active sensor emits a controlled beam of energy (EMR) in the direction of the target to be investigated. The sensor then measures the amount of energy that is reflected or backscattered from the target. The key advantages of active sensor systems are

that they can be operated in night time and in any weather conditions. Nevertheless, a large amount of energy is needed to illuminate the targets (Lillesand and Kiefer, 2002; Jensen, 2005). The most familiar active sensor is RADAR, emits energy in the microwave region of the electromagnetic spectrum. The other active sensors are LIDAR, Scatterometer, RADAR altimeters etc. The active remote sensing process is shown in Figure 1.2 (b).

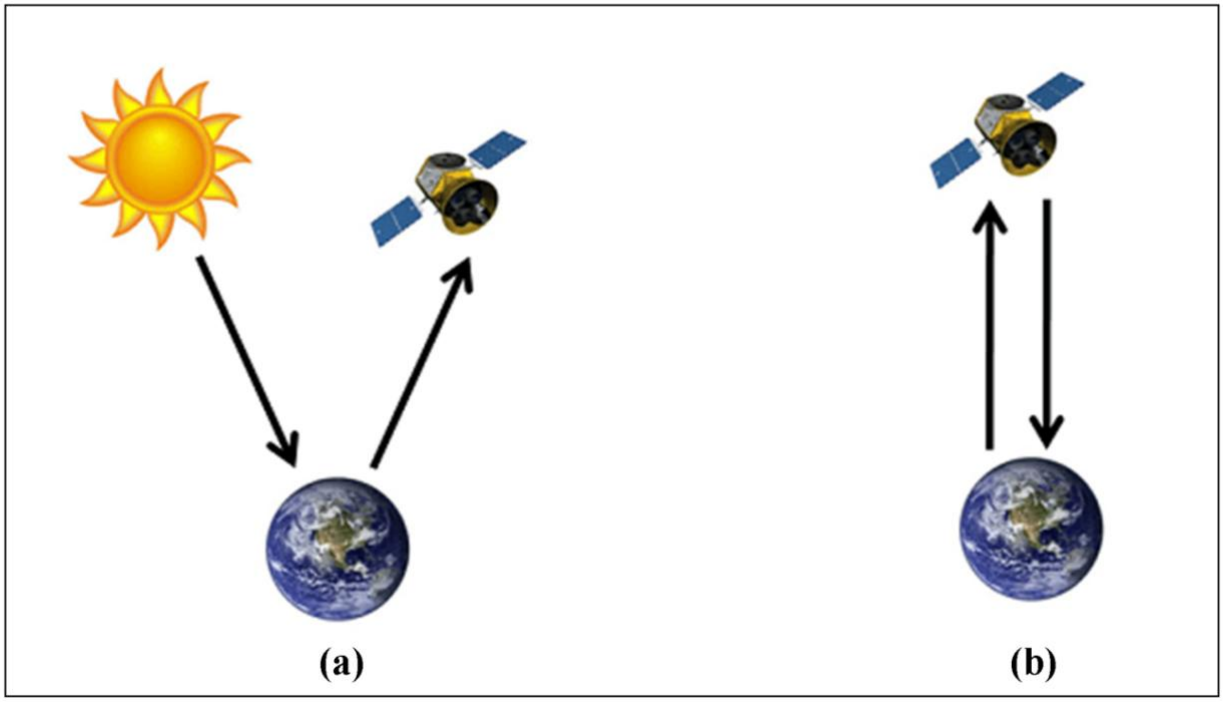

Figure 1.2 Remote sensing processes (a) passive; (b) active
(**Source:** http://www.physics-and-radio-electronics.com/blog/lidar-light-detection-ranging)

1.4 EMR IN REMOTE SENSING

An energy source in the form of EMR is a primary requirement to illuminate the targets of interest. The wide range of wavelength intervals, varying from the shorter wavelengths (gamma and x-rays) to the longer wavelengths (microwaves and radio waves) are usually referred as the electromagnetic spectrum shown in Figure 1.3. Remote sensing operates in several parts of the electromagnetic spectrum. The most widely used regions of the electromagnetic spectrum are visible, infrared and microwaves.

The visible portion of the electromagnetic spectrum extends from 0.4-0.7 µm, is commonly called as light. Another portion of the electromagnetic spectrum of interest is the infrared (IR) region ranging from approximately 0.7-100 µm. It is more than 100 times wider than the visible portion. The IR region can be separated into two types based on their radiation properties the reflected IR, and the emitted or thermal IR. The radiation in the reflected IR region, extending from 0.7-3.0 µm, is very similar to the radiation in the visible region. The thermal IR region, extending from 3.0-100 µm, provides information about the energy emitted from the Earth's surface in the form of heat. The most recent portion of interest of the spectrum is the microwave region, ranging from about 1 mm to 1 m. Microwave region has the longest wavelength used in remote sensing and can provide information about the surface such as roughness, water content, etc.

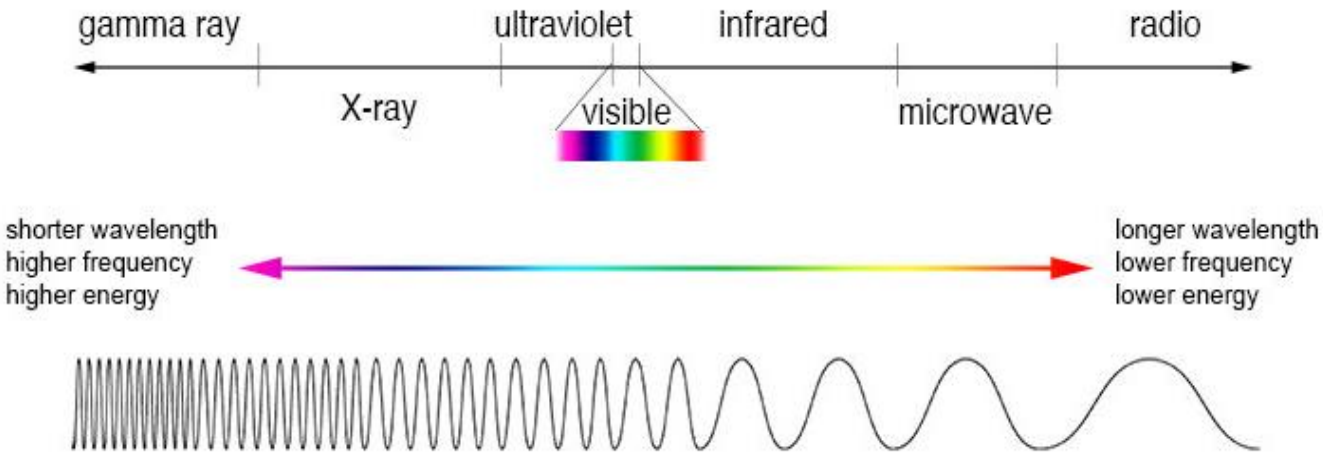

Figure 1.3 Electromagnetic spectrum

(**Source:** https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html)

1.5 INTERACTION OF EMR WITH ATMOSPHERE

The EMR from an energy source travels through some distance in the atmosphere before being recorded by the remote sensor. The EMR before reaching the Earth's surface interacts with the atmosphere in three possible ways: absorption, transmission, and scattering.

The transmitted energy is then either reflected or absorbed by the targets of interest. The incoming EMR is attenuated (scattered and absorbed) by gas and particle constituents of the atmosphere. The atmospheric effects are based on the properties of the radiation such as magnitude and wavelength, the abundance of particles or gases, and also the path length.

1.5.1 Scattering

Atmospheric scattering takes place when the EMR interacts with the particles or gaseous molecules present in the atmosphere and to be redirected from their original path. There are three different types of scattering take place: *Rayleigh scattering*, *Mie scattering*, *and Non-selective scattering*.

1.5.2 Absorption

Absorption is the process in which incident EMR is partly absorbed in a given wavelength by the molecules present in the atmosphere. The most efficient EMR absorbing constituents of the atmosphere are ozone (O_3), water vapour (H_2O), and carbon dioxide (CO_2). The harmful ultraviolet radiation from the Sun is absorbed by O_3. The CO_2 known as the *greenhouse gas*, strongly absorb the far IR portion of the electromagnetic spectrum. Atmospheric absorption causes an effective loss of EMR in contrast to scattering. Although all the wavelengths from the Sun arrive at the top of the atmosphere, because of the atmospheric absorption, only specific wavelengths can pass through the atmosphere. The wavelength regions of the electromagnetic spectrum that are partially or fully transmitted through the atmosphere are known as "*atmospheric windows*" and thus useful in remote sensing (Figure 1.4). The visible and reflected IR portion ranging from 0.4 – 2 μm, is the useful window for optical remote sensing. In the thermal IR portion, two narrow windows about 3 and 5 μm, and relatively wide third window ranging about from 8–14 μm are also

useful for thermal remote sensing. The large atmospheric window at the wavelengths beyond 1 mm is related to the microwave portion of the electromagnetic spectrum.

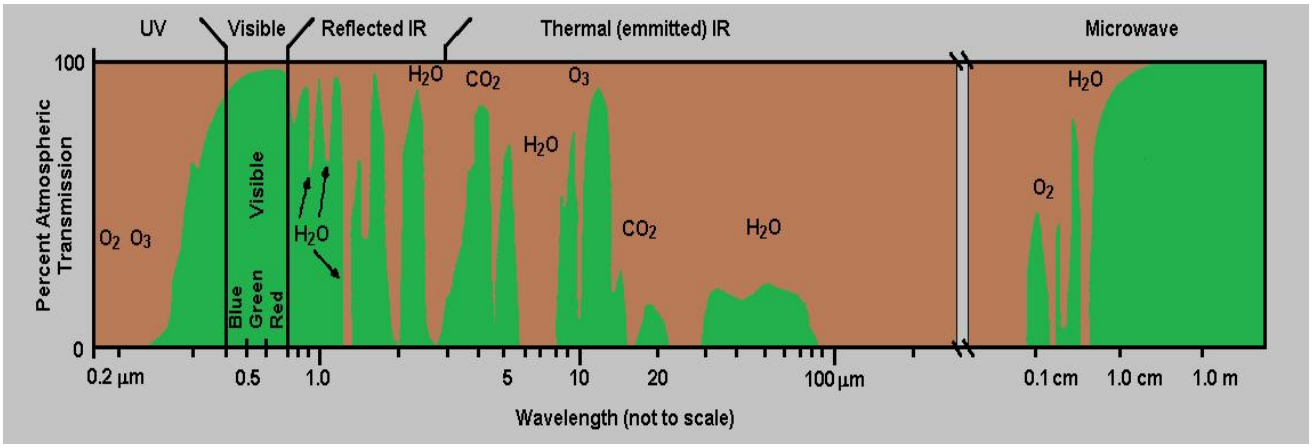

Figure 1.4 Electromagnetic spectrum useful for remote sensing

(**Source:** http://www.oneonta.edu/faculty/baumanpr/geosat2/RS-Introduction/RS Introduction.html)

1.6 INTERACTION OF EMR WITH THE EARTH'S SURFACE

The EMR which is not absorbed or scattered in the atmosphere can arrive at Earth's surface and interact with the land surface features. Three types of interaction can take place when the radiation strikes with the surface. These are: *absorption (A)*; *transmission (T)*; and *reflection (R)*.

Absorption (A) takes place when the EMR (energy) is absorbed by the target. Transmission (T) takes place when the EMR passes through a target. Reflection (R) occurs when the EMR "bounces" off the target and is redirected. Only the reflected radiation is of significant interest in remote sensing because it speaks something about the surface characteristics. There are two types of reflection, showing the way of reflected energy from a target, specular and diffuse.

(a) Specular reflection or mirror-like reflection takes place when a surface is smooth, and almost all the energy is directed away from it in a single direction.

*(*b) Diffuse reflection takes place when the surface is rough and, the energy is reflected uniformly in all the directions.

Most of the Earth's surface features range from perfectly specular or perfectly diffuse reflectors. It depends on the surface roughness in comparison to the wavelength of incoming energy.

1.6.1 Spectral reflectance curves

The surface of the Earth is composed of a variety of target materials. Unfortunately, most of the targets have behaviour between perfectly specular and perfectly diffuse reflectors. The target identification solely based on the reflectance data is a complex task. A spectral reflectance curve can be established by potting reflectance against wavelength. It is sometimes called as the "*signature*". It is a description of the degree to which the energy is either reflected or emitted by the targets on Earth's surface in different portions of the electromagnetic spectrum. A typical spectral reflectance curves for three basic earth's surface features like vegetation, soil and water are shown in Figure 1.5. The feature would appear brighter in an image because of higher reflectance at a specific wavelength. The spectral reflectance of vegetation depends on the properties of the leaves. The reflection from the blue and red wavelengths is comparatively low due to its absorption by chlorophyll content in the leaves but reflects relatively more in green wavelength. The reflectance of vegetation is highest in the NIR region (0.8 - 1.4 μm) in comparison to visible region (0.4 - 0.7 μm). But the amount of reflectance depends on the development and cell structure of the leaves. Generally, the vegetation shows three reflectance valleys at different wavelengths. One is at red wavelength region (0.65 μm) due to high absorptance of radiation by chlorophyll content in the leaves. The other two wavelength regions are at 1.45 - 1.55 μm and 2.10 - 2.20 μm

caused due to high absorptance of radiation by the water content in the leaf tissues. In the visible portion of the spectrum clear water shows low reflectance (< 10%). The water absorbs almost all the incoming radiation at the wavelengths energy longer than 0.75 μm, appears dark in the image. Dry soil shows a relatively flat reflectance response. Its spectral reflectance drops when it is wet due to water absorption only.

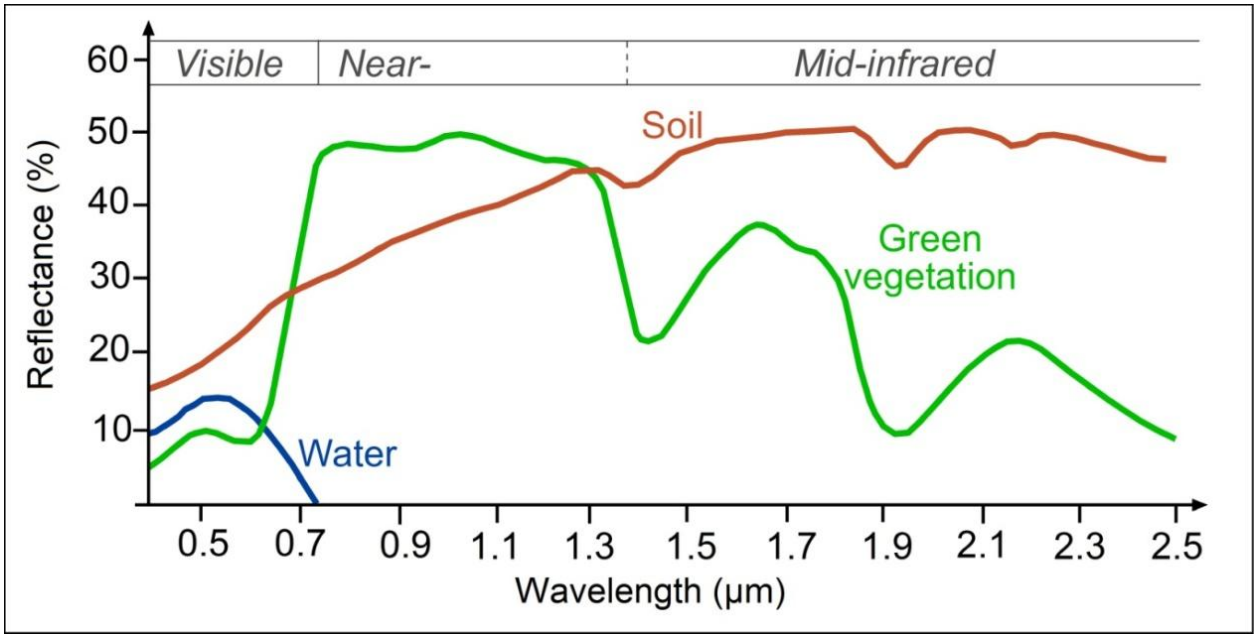

Figure 1.5 Spectral reflectance curves for soil, vegetation and water
(**Source:** http://www.seos-project.eu/modules/classification/classification-c01-p05.html)

1.7 CHARACTERISTICS OF REMOTE SENSING IMAGE

Remote sensing image is more than a picture. It detects electromagnetic energy either photographically or electronically, and stores in a standard grid format (rows and columns). Each image is composed of small, equal-sized and shaped areas, called "*pixels*", an abbreviation of picture elements. The detected energies are stored as discrete digital number values or *"DN values"*. The quality of an image mainly depends on the characteristics of the sensor-platform system such as spatial, spectral, radiometric and temporal resolutions.

1.7.1 Spatial resolution

Spatial resolution refers to the smallest unit-area that can be detected by the sensor. It determines how finely an image can record the spatial detail of the real world and therefore the number of pixels utilized in the construction of an image. The image with higher spatial resolution is composed of a substantial amount of pixels than those of lower spatial resolution. The higher spatial resolution leads to the more details available on the image that can be observed. It is stated in meters (m). Spatial resolution of some common satellite sensors are as IKONOS (1 m), LISS-IV (5.8 m), LISS-III (23.5 m), Sentinel-1A (20 m), RISAT-1 (25 m), Landsat 8-OLI (30 m), AWiFS (56 m).

1.7.2 Spectral resolution

Spectral resolution refers to the ability of a sensor to define fine wavelength intervals or to resolve the energy received in a spectral wavelength ranges to characterize different components of the earth's surface. The finer the spectral resolution, the narrower the wavelength ranges for a specific band. Most of the remote sensing sensors are multi-spectral, that record energy over discrete wavelength ranges at different spectral resolutions. For example, LISS-IV sensor has three spectral bands, B2 (green): 0.52 - 0.59 μm, B3 (red): 0.62 - 0.68 μm and B4 (NIR): 0.77 - 0.86 μm.

1.7.3 Radiometric resolution

Radiometric resolution refers to the ability of a sensor to discriminate very slight differences in electromagnetic energy. It is measured through the number of the grey-scale values. The higher the radiometric resolution of a sensor is, the more sensitive it is in detecting the slight differences in reflected or emitted energy. The number of grey-scale values is defined by the number of bits or binary numbers. An 8 bit depiction has 256 grey-

scale values, while a 16 bit depiction has 65536 grey-scale values. Radiometric resolutions of some common satellite sensors are LISS-IV (10 bits), LISS-III (10 bits), Landsat 8-OLI (12 bits), RISAT-1 (16 bits) and Sentinel-1A (16 bits).

1.7.4 Temporal resolution

Temporal resolution is defined as the time interval between two successive image acquisitions over the same area on the earth. The revisit time of a satellite sensor is usually described in days. The temporal resolution is high when the revisit time is low and vice-versa. The real temporal resolution is a function of a variety of factors, including the satellite/sensor capabilities, the swath overlap, and latitude. Temporal resolutions of some common satellite sensors are LISS-IV (24 days), Landsat 8-OLI (16 days), RISAT-1 (25 days) and Sentinel-1A (12 days).

1.8 IMAGE RESTORATION

A remote sensing image acquired by a sensor either on board spaceborne or airborne platforms at a great distance from the Earth's surface, will have various distortions including geometric and radiometric. So, it is needed to correct distorted or degraded image data to construct a more authentic representation of the original scene through several image processing steps.

1.8.1 Geometric distortions and its correction

It is an inherent distortion in remote sensing which may rise during the accurate representation of three-dimensional surface of the Earth as a two-dimensional image. Geometric distortions in the image rise due to moving Earth at the time of image acquisition by orbiting satellite. Thus the image geometry is in constant flux. Consequently, there are two types of geometric distortions that can degrade the quality of an image.

(a) Internal distortions: These geometric distortions are introduced by the geometry of remote sensing system itself or in combination with earth rotation or curvature. Most of the sensing systems are based on lens that produces radial distortions, tangential distortions, projection distortions, scale errors etc. The scanning systems generate scan skew, mirror scan velocity error, platform velocity etc. These are systematic or predictable distortions and may be corrected through the analysis of sensor characteristics and ephemeris data.

(b) External distortions: These geometric distortions are generally introduced by phenomena that vary in nature through space and time. These are also known as non-systematic or random distortions. The most important sources that can cause external distortion in an image are random variation of platforms (aircraft or spacecraft) and properties of the ground targets at the time of data collection. These sources include the changes in altitude and attitude (roll, yaw and pitch) of the platforms. Earth curvature and irregular relief displacements of terrain also cause another type of external distortions.

An image preprocessing is necessary to correct either internal or external distortions so that the pixels are in their proper planimetric map locations. However, a lot of unknown sources also contribute to distortions in a remote sensing image. The simplest method to associate an image to a map projection system is to solve geometric corrections with reference data or ground control points (GCPs) under selected geometric transformation model. The process of an image transformation from pixel-based grid system into a georeferenced grid system is performed using a series of GCPs in combination with resampling algorithms. There is need of enough number of GCPs depending on the selected model. Generally, 5-10 GCPs are selected and sufficient for a transformation model. This procedure is also called as *georeferencing*. It is essential to perform image registration that is

a process to compare the images collected from multi-source or multi-temporal data in a GIS environment. Another approach known as geocoding is used, when all the image data is required to be stored in the same map projection. There are two types of image registration, "*image to map*" *and* "*image to image*".

1.8.2 Radiometric distortions and its correction

Radiometric distortions are resulted due to various elements including sensor's sensitivity, sun angle, topography and atmospheric constituents. In optical sensor data, the defects may be in the form of vignetting effect, scan line dropout, line striping and random noise or spike. The sun angle and topography create radiometric distortions in the form of sun glitter, shade, and shadow. The radiation coming from the sun as well as the reflected and/or emitted from earth's surface must pass through the atmosphere before it reaches to the sensor. The atmospheric constituents (such as particulate matter, moisture content and turbulence) substantially affect the incoming plus the reflected and/or emitted radiation in specific wavelength regions of the electromagnetic spectrum (Kaufman, 1989). The atmospheric defects are introduced by direct sunlight, skylight, path radiance, scattering and their reflections incident to sensors. It affects the quality of imagery and leads to information loss. The atmospheric corrections are not easy because the spatial distribution of the atmosphere are not known and are usually not homogeneous.

The correction of sensor induced distortions involves all the operations that aimed to correct apparent errors and noise in an image. In the restoration process of scan line dropout, the first step is to calculate the average DN-value per scan line for the entire image. Then the average DN-value is compared with this image average. A scan line is identified as defective if it deviates from a chosen threshold value. The next step is to replace the defective lines by

calculating average DN values for the corresponding pixels in the preceding and succeeding scan lines. The most popular procedure to correct the line striping error is the histogram matching. In this process, separate histograms are constructed corresponding to each detector unit and matched. The gains and offsets for other detector units are adjusted, and new DN-values are computed and assigned by choosing one of the histograms as standard. On the other hand, random noises occur due to errors in data transmission and temporary disturbance. Therefore, it is required to apply a more sophisticated restoration method, such as digital filtering to correct random noises in an image.

1.8.2.1 Atmospheric correction

Atmospheric correction is the process of the removal of distortions in the degree of electromagnetic radiation detected by the sensor to produce surface reflectance. It is a critical preprocessing step for the quantitative analysis of the spectral properties of the data. It becomes significant, when the adjacent or overlapping scenes acquired under differing atmospheric conditions are going to be used in combination with each other, such as change detection or land use and land cover mapping over a large area using multiple satellite images. It can enhance the interpretability and usefulness of an image. Several parameters are needed to accurately apply atmospheric correction, including the amount of water vapor, distribution of aerosols, and scene visibility at the time of image acquisition. There are number of atmospheric correction techniques available in many software packages. One of the most common techniques is dark object subtraction (DOS) for ameliorating atmospheric effects when there is the unavailability of atmospheric conditions. Other atmospheric correction techniques are given as follows:

- ATmospheric REMoval (ATREM)

- QUick Atmospheric Correction (QUAC)
- Atmospheric CORection Now (ACORN)
- Atmospheric and Topographic Correction (ATCOR)
- Second Simulation of the Satellite Signal in the Solar Spectrum (6S)
- MODerate spectral resolution atmospheric TRANsmittance (MODTRAN)
- Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH)

1.9 REMOTE SENSING APPLICATIONS

- Agriculture
- Forestry
- Water resources
- Geosciences
- Land use/land cover
- Monitoring of atmospheric constituents
- Biodiversity conservation
- Disaster management

1.10 WHAT ARE LAND USE, LAND COVER, AND LAND CHANGE?

Landscapes provide the natural resources and living space for humankind. Physically, it is composed of a montage of heterogeneous land use and land cover (LULC) categories that are produced by various anthropogenic activities and environmental processes (Alberti, 2008). LULC are two separate terminologies which are frequently used vice versa (Dimyati et al., 1996). Land use is defined as the way in which the land and its resources have been exploited by humans and their habitat, usually with special emphasis on economic activities.

Further, it was stated by Skole (1994) that the land use itself is the human utilization of a land cover type, the way by which human activities appropriates the results of net primary production (NPP) as determined by a complex of socio-economic factors. Land cover is defined by the observed bio-physical attributes of earth's surface and immediate subsurface. It describes the physical state of the land surface including forest, vegetation, water bodies and built up areas (Meyer and Turner, 1996). The LULC pattern of a region is the result of natural and socio-economic factors and their utilization by human beings in space and time. The land change science investigates the changes in LULC which are termed as the alteration of Earth's environmental system by various anthropogenic activities. The main goal of land change science is to quantify the changes as well as the identification of drivers and its impacts.

1.11 NEED OF LULC INFORMATION AND LAND CHANGE ANALYSIS

LULC is the most remarkable attributes of landscapes on the earth's surface, and so it has become a significant component of global environmental change studies (Dickinson, 1995). Information on LULC is a fundamental requirement for the selection, planning, and execution of land use schemes to meet the increasing demands for basic human needs and welfare. Timely and accurate LULC information is also vital to scientists, planners, and decision-makers in forming policies regarding economic, demographic and environmental concerns at different observational scales (Mishra and Rai, 2016). In land change science, LULC are frequently studied in combination with each other, particularly in remote sensing based studies, because satellite images and aerial photography can identify land cover categories. However, inferring land use often need more knowledge of the study area. Therefore some agreement is usually made between identifying the variable of interest (land

use) and the related substitute (land cover). The earth's surface, all over its history and existence, has undergone changes and alterations at different space and time scales (Mariwah et al., 2017).

In the past several centuries, various anthropogenic activities have altered the Earth's environment by changing the LULC (Hurtt et al., 2006; Liu and Tian, 2010). LULC changes (LULCC) are major driving forces for various environmental concerns of human populations today, including climate change, biodiversity loss, deforestation, increased natural disasters and the pollution of water, soils, and air from local to the global scales. (Mas et al., 2004; Feddema et al., 2005; Jain and Yang, 2005; Tian et al., 2012; Tao et al., 2013). These changes also assist in understanding the complex interactions between human activities and global environmental changes (Dickinson, 1995; Zhu and Woodcock, 2014). Since 1850, LULCC has only contributed to around 35% of human-induced carbon dioxide (CO_2) emissions across the globe (Houghton et al., 2012). The rapid development and growing human population mainly since the start of the industrial revolution has intensified various anthropogenic activities which have led to spontaneous and uncontrolled impacts on LULC (Szilassi et al., 2006; Petropoulos et al., 2013). According to Lambin et al. (2001), in recent decades the rate, magnitude and spatial reach of direct and indirect alterations of the earth's surface by human's interferences are unprecedented. It is posited that LULCC is the most important indicator of human induced alteration of Earth's surface. Furthermore, when LULCC aggregated globally, becomes insidious leading to significantly influence key aspects of the functioning of the Earth system (Lambin et al., 2001). Nevertheless, according to Otukei and Blaschke (2010) the lack of LULC information and its changing pattern can be attributed to: (1) weak government support for planning agencies and decision-makers; (2)

expensive software and hardware; (3) inadequate budget allotment for data procurement; and (4) resistance to change in developing countries like India.

Therefore, monitoring and quantifying the LULCC and its negative consequences while sustaining the management and development of natural resources has turned out as a major priority of researchers and policy-makers around the world. Despite the significant role played by LULCC in environmental monitoring and understanding, it is limited by the scarcity of accurate and timely information. These changes also occur at multiple spatial and temporal scales that may vastly differ among regions (Tian et al., 2014). Such information, particularly in the form of thematic maps, is not normally readily available whereas, the speed and flexibility of the produced information and analysis is obviously an important aspect. A substantial quantity of the data about the Earth's surface is obligatory for useful analysis, monitoring and quantifying the ongoing LULCC in an area.

1.12 NEED OF LAND CHANGE MODELING

Modern technologies offer the acquisition and analysis of data from space-borne platforms, with linkages to GPS data, GIS layers and functions, and emerging modeling capabilities (Franklin, 2001). The land change models are the tools to support the analysis of the causes and consequences of LULCC to better understand the functioning of the landscape system. They are valuable in disentangling the sophisticated suite of socio-economic and biophysical forces that influence the rate and spatial pattern of LULCC and for estimating its impacts.

Furthermore, the models can also support the investigation of future LULCC under different geospatial scenarios (Verburg et al., 2004). Recently the modeling of spatio-temporal landscape dynamics is gaining a lot of recognition in solving the problems

occurring due to the alteration and conversion of LULC (Lambin et al., 2001). Nevertheless, the future prediction is directly connected to the changes observed from the past to the current also (Bhatta, 2010). The study of current situations and model applications both are spatial in nature; thus the solutions also need spatial approach. Consequently, it is essential to apply spatially explicit models to simulate and predict the future LULCC scenarios.

In the past, all over the world, information on LULCC was produced primarily through conventional ground survey methods which were not only time consuming, expensive, and tedious but also demand lots of human resources (Chaudhary et al., 2008). These methods are somewhat ineffective and impractical for monitoring the changes over shorter periods and for real time applications. Recent advances in remote sensing, digital image processing and geographic information system (GIS), in contrast to conventional techniques, are appeared as an impressive way for LULC classification, mapping, and change analysis.

1.13 POTENTIAL OF REMOTE SENSING IN LULC AND LAND CHANGE STUDIES

Remote sensing technologies have emerged as one of the fascinating subjects over the past three decades. Earth observation through spaceborne sensors is an attractive pathway, providing valuable information for describing the extent of landscape types, performing spatio-temporal LULCC analysis quantitatively at various scales (Kamusoko and Aniya, 2007; Srivastava et al., 2014; Petropoulos et al., 2015; Lamine et al., 2017). During the last three decades, remote sensing technologies and methods have made tremendous development, including a suite of sensors operating at a broad range of imaging scales with potential interest and value to planners, land managers and decision-makers (Clark et al., 2004). These sensors include aerial cameras, multispectral scanners, and imaging radars. Nowadays, the images with high spatial resolution to medium spatial resolution such as

QuickBird, IKONOS, LISS-IV, LISS-III, Landsat 8-OLI, AWiFS, etc. are offering new opportunities for mapping and monitoring of LULC. Optical remote sensing is a viable approach for the classification, mapping and monitoring of LULC and its changing pattern widely at both regional and global levels. But the use of optical sensors is restricted mostly due to frequent cloud cover and adverse weather conditions. The advent of SAR, with its all weather and all-time imaging capabilities, contributed the wealth of information to LULC and land change studies.

The advantages of space-borne remote sensing are given as:

1. Remote sensing technologies facilitate observations across larger extents of Earth's surface than is possible by traditional ground based observations.
2. Remote sensing images provide useful information for detecting land features at various wavelengths which are not visible to the human eyes.
3. It provides cost-effective data of inaccessible regions regularly.
4. It provides rapid acquisition of images with varying spatial and temporal resolution.
5. The integrated approach of remote sensing, geospatial analysis, and modeling together with the interdisciplinary assortment of natural and socio-economic scientific methods required to study the causes and consequences of LULCC over time.
6. It is easy to manipulate and fast processing using computers.

1.14 RESEARCH QUESTIONS

Pertaining to the domain of the present research and focus of the study on mining LULC information and changing landscape pattern analysis, the following research questions are to be answered:

1. How the textural features are useful in improving LULC classification accuracy over a heterogeneous landscape?
2. How the multi-source remote sensing images with varying spatial resolutions affect the selection of textural features?
3. How dual-polarimetric SAR data at C-band improve the accuracy of rice crop mapping using machine learning technique?
4. How a geographically weighted method is useful for producing and visualizing the spatial variation in accuracy of remote sensing image classification?
5. How the integration of remote sensing, GIS and land change models, together with the assortment of socio-economic factors and environmental variables are helpful to assess the changes in landscape and predict its future scenario?

1.15 RESEARCH OBJECTIVES

The work presented in this thesis mainly focuses on the establishment of a robust classification approach to improve LULC classification accuracy using ancillary information in a heterogeneous landscape. The study was also performed to investigate LULCC and predict the future scenario. More specifically the research objectives of this thesis are given as follows:

1. Comprehensive evaluation of textural features in improving LULC classification accuracy using multi-sensor and multi-resolution remote sensing images.
2. Potential of dual-polarimetric SAR data at C-band for rice crop mapping using decision tree approach.
3. Analysis of spatial variation in accuracy measures of remote sensing image classification.

4. Assessment of LULC changes over a decade (2000-2014) using multi-temporal remote sensing data.
5. Modeling and prediction of spatio-temporal LULC dynamics in Varanasi district of Uttar Pradesh, India using geospatial approach.

1.16 STUDY AREA

The study area for this research work is Varanasi district of Uttar Pradesh, a part of northern India. It is located along the left crescent bank of the holy river Ganga. It is considered as one of the oldest living places of the world. It is well known for being a hotspot of heritage, education, cultural activities and biodiversity for over many years. Varanasi, the holy city of India, is very well known for its fine quality silks, ghats, and temples etc. It is the most popular pilgrimage point for the Hindus and known as the religious capital of India. Literally, the meaning of Varanasi is the land between river Varuna and Assi while in recent times the city has expanded beyond this limit. It is also known by different names such as Kashi, Benares, and Banaras. Varanasi district is surrounded by river Gomati and Jaunpur district in the north, by Sant Ravidas Nagar Bhadohi in the west and Chandauli district in the east. The land is very fertile and rich in agriculture due to its location in the middle Ganga plain. The soil types found in Varanasi district are alluvial and calcareous. The major crops grown in Varanasi district are rice, corn, pigeon pea, wheat, pea, gram, mustard, black gram, green gram, barley, lentil, linseed, sorghum and millet etc. The location map of the study area is shown in Figure 1.6.

The other characteristics of the study area are given as follows:

Geographical location **:** 25°10′30″ N to 25°35′15″ N latitude

82°40′50″ E to 83°12′18″ E longitude

Projection system	**:**	UTM (zone 44, N), WGS 84
Geographical area	**:**	1532.91 km^2
Physiographic location	**:**	Middle Ganga plain
Climate Zone	**:**	Humid subtropical climate with large variations between summer and winter temperatures
Elevation range	**:**	0 to 98 meters

The rapid population growth is one of the most intimidating issues in India. Varanasi district is one of the fastest growing urban areas in India over the past few decades. It has a total population of 3,676,841 people with 1,921,857 males and 1,754,984 females. While in 2001, Varanasi had a population of 3,138,671 people with 1,649,187 males and 1,489,484 females (Census of India, 2011). One of the most significant reasons for population growth in the Varanasi is the large-scale rural to urban relocation and rapid urbanization. In the year 2001, the population density was 2,045 persons per km^2 while in the year 2011 it increased up to 2,395 persons per km^2. The average literacy rate in the year 2001 was 66.12% and increased to 75.60% in the year 2011.

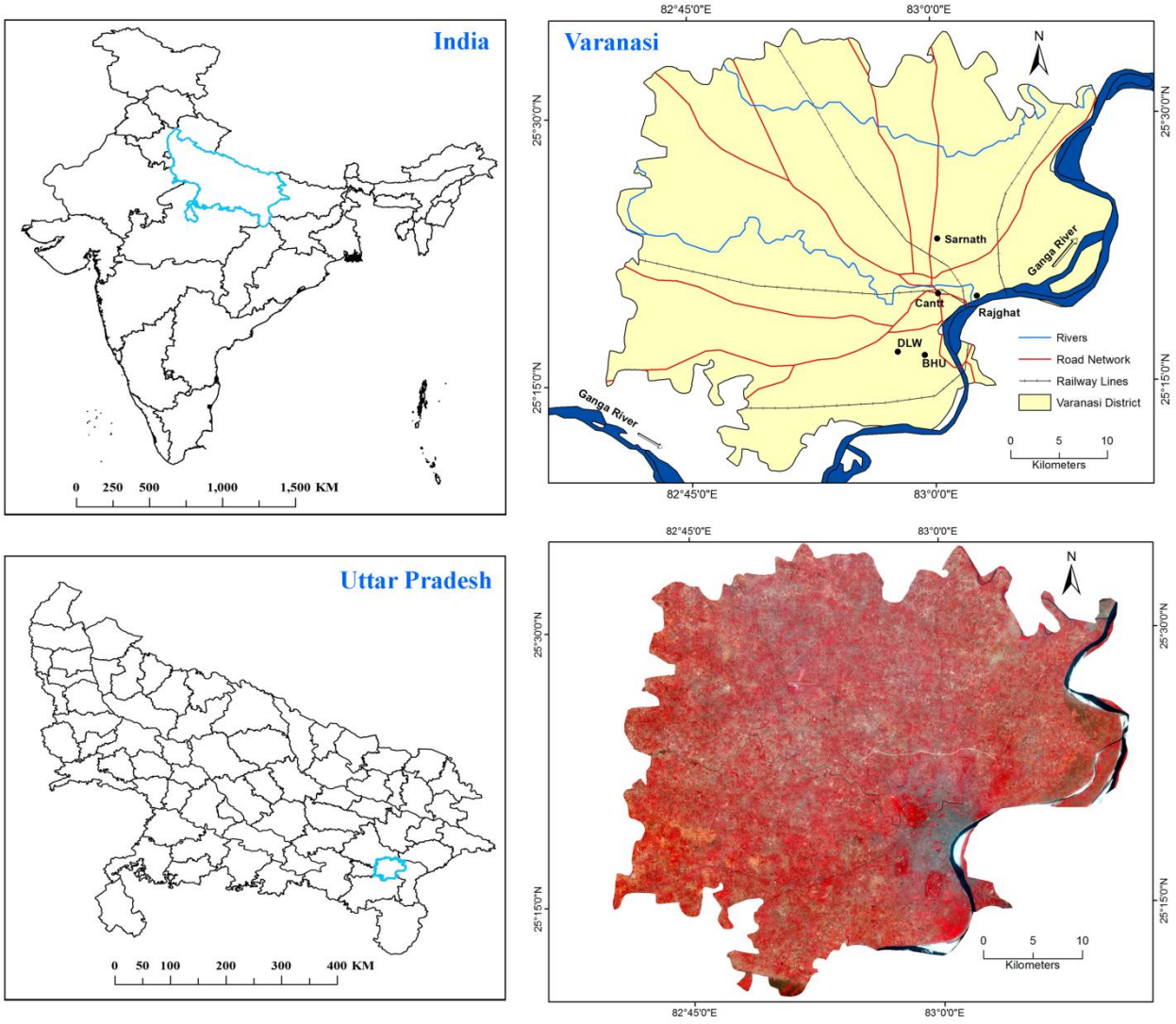

Figure 1.6 Geographical location of the study area as viewed on Landsat 8-OLI image in FCC

1.17 DATA USED

Several types of remote sensing images were collected during 2013-2015 for the analysis and classification of the heterogeneous landscape of Varanasi district of Uttar Pradesh, India. The images were collected from the optical and microwave sensors depending upon their availability and quality for the study area. The optical remote sensing images from the multispectral sensors such as Linear Imaging Self-Scanner-III (LISS-III),

Linear Imaging Self-Scanner-IV (LISS-IV), Landsat-Thematic Mapper (TM)/Enhanced Thematic Mapper Plus (ETM+)/Operational Land Imager (OLI) images and Advanced Wide Field Sensor (AWiFS) were collected. Microwave remote sensing images from SAR sensors such as Radar Imaging Satellite-1 (RISAT-1) and Sentinel-1A were collected. Reference data were also collected in situ during the field survey.

1.17.1 Collection of ground truth data

For this study, ground truth data were collected in years 2013 and 2015. During the field campaigns, various locations in the study area were visited to collect the ground truth data of different land surface features on the day of satellite image acquisition. These data will be helpful in the categorization of various land surface features. The data collected during the field visit were aggregated into seven generalized LULC classes according to the landscape of study area. The precise geolocation of the samples containing various LULC classes were captured with the help of a handheld GPS receiver (Trimble Juno 3B) with positional accuracy up to 5 meter in X, Y coordinates. The seven generalized LULC classes of interest for the study area are the agricultural land, dense vegetation, sparse vegetation, built up, fallow land, water bodies and sand.

1.17.2 Collection of remote sensing data

The multispectral remote sensing images from LISS-IV, LISS-III, and Landsat TM/ETM+/OLI sensors were collected and used for comparative analysis of LULC classification and land change analysis of a heterogeneous landscape. The LISS-IV and LISS-III images were obtained from National Remote Sensing Centre (NRSC) Hyderabad, whereas TM/ETM+/OLI images were downloaded from United States Geological Survey (USGS) website (http://glovis.usgs.gov). For this study, polarized SAR data from RISAT-1

were obtained from National Remote Sensing Centre (NRSC) Hyderabad and Sentinel-1A satellite were obtained from the European Space Agency (ESA). Both the datasets are available at C-band with dual polarization. The major characteristics of the above mentioned datasets are given in Table 1.1. The multi-temporal images from Landsat series Satellites (TM/ETM+/OLI) were also obtained during the period of 1988-2015 from the official website of USGS website. The multi-temporal remote sensing images are listed in Table 1.2.

1.17.3 Auxiliary data

In addition to remote sensing images, some auxiliary data were also used to support the study. The Google Earth images were also used as ground truth information to verify the LULC classification results. The digital elevation model (DEM) by SRTM with 90 m spatial resolution was downloaded from (http://srtm.csi.cgiar.org/) and used to generate slope and aspect. The vector layers of roads and railway network were gathered from the toposheet and Google Earth images.

Table 1.1 Major characteristics of multi-source remote sensing data

Satellite/sensor	Date of acquisition	Major characteristics	Data sources
Resourcesat 2-LISS-IV	06 April 2013	Spatial resolution: 5.8 m B2-green: 0.52-0.59 µm B3-red: 0.62-0.68 µm B4-NIR: 0.77-0.86 µm Radiometric resolution: 10-bit Swath: 70 Km	NRSC
Resourcesat 2-LISS-III	13 March 2013	Spatial resolution: 23.5 m B2-green: 0.52-0.59 µm B3-red: 0.62-0.68 µm B4-NIR: 0.77-0.86 µm B5-SWIR: 1.55-1.70 µm Radiometric resolution: 10-bit Swath: 140 Km	NRSC
Landsat 8-OLI	15 April 2013 06 September 2013	Spatial resolution: 30 m B1-Coastal aerosol: 0.43-0.45 µm B2-Blue: 0.45-0.51 µm B3-Green: 0.53-0.59 µm B4-Red: 0.64-0.67 µm B 5-NIR: 0.85-0.88 µm B 6-SWIR 1: 1.57-1.65 µm B 7-SWIR 2: 2.11-2.29 µm B 8-Panchromatic: 0.50-0.68 µm B 9-Cirrus: 1.36-1.38 µm Radiometric resolution: 16-bit Swath: 185 Km Product type: L1T	USGS
Sentinel-1A	29 April 2015	Spatial Resolution: 5x20 m Band: C (5.40 GHz) Level: L1-GRDH Imaging mode: IW Polarizations: VV & VH Unsigned 16-bit Swath: 250 Km	ESA
RISAT-1	09 August 2013 03 September 2013 01 April 2015	Spatial Resolution: 25 m Band: C (5.35 GHz) Level: L1-slant range, ground range Imaging mode: MRS Polarizations: HH & HV Unsigned 16-bit Swath: 115 Km	NRSC

Table 1.2 Multi-temporal remote sensing data

Satellite-sensor	Path/row	Date of acquisition	Spectral resolution (μm)	Spatial resolution (m)	Temporal resolution (days)
Landsat 5-TM	142/42 142/43	04 November 1988 11 May 2005 23 June 2009	B1-blue: 0.45-0.52 B2-green: 0.52-0.60 B3-red: 0.63-0.69 B4-NIR: 0.76-0.90 B5-SWIR: 1.55-1.75 B6-TIR: 10.40-12.50 B7-MIR: 2.08-2.35	30 30 30 30 30 120 30	16
Landsat 7-ETM+	142/42 142/43	03 April 2000 31 October 2001	B1-blue: 0.45-0.52 B2-green: 0.52-0.60 B3-red: 0.63-0.69 B4-NIR: 0.77-0.90 B5-SWIR 1: 1.55-1.75 B6-TIR: 10.40-12.50 B7-SWIR 2: 2.09-2.35 B8-PAN: 0.52-0.90	30 30 30 30 30 60 30 15	16
Landsat 8-OLI	142/42 142/43	05 June 2014 15 November 2015	B1- coastal aerosol: 0.43-0.45 B2-blue: 0.45-0.51 B3-green: 0.53-0.59 B4-red: 0.64-0.67 B5-NIR: 0.85-0.88 B6-SWIR 1: 1.57-1.65 B7-SWIR 2: 2.11-2.29 B8-PAN: 0.50-0.68 B9-cirrus: 1.36-1.38	30 30 30 30 30 30 30 15 30	16

1.18 METHODOLOGY

In this thesis, multi-source remote sensing images were investigated for achieving improved LULC information and change analysis. All the subsequent pre-processing, interpretation, texture analysis, LULC classification, accuracy assessment were performed using ENVI-SARscape (v 5.1) image processing software. Also, to model LULCC using three hybrid models which are Stochastic-Markov chain (ST-MC), Cellular automata-

Markov chain (CA-MC), and Multi-layer perceptron-Markov chain (MLP-MC), IDRISI Selva software was employed as well as to predict the future scenarios.

1.18.1 Preprocessing of remote sensing images

The images were first imported into ENVI environment. The atmospheric correction of a single date image is not required for classification purpose. Therefore atmospheric correction of single date images from LISS-IV, LISS-III and Landsat 8-OLI sensors was not performed. However, the collected multi-temporal Landsat images were atmospherically corrected using QUick Atmospheric Correction (QUAC) module available in ENVI software. All the images were spatially referenced in the Universal Transverse Mercator (UTM) projection system (Zone 44, North) with datum World Geodetic System (WGS) 1984.

The RISAT-1 and Sentinel-1A images were preprocessed using ENVI-SARscape (version 5.1) and SNAP (version 3.0.0) software respectively. Both the images were first imported into corresponding software platforms. The speckles in a SAR image generally refer to a noises produced by coherent phenomena leading to degradation in the image quality. It makes complexity for the examination of various features of interest and classification process. Therefore, different speckle filtering methods were employed to reduce the speckles in SAR images. After speckle reduction, the geometric correction was performed, and the DN values were converted into backscattering coefficient (σ°).

1.18.2 Texture analysis

The texture of an image describes the pattern of the arrangement of intensities in space. The image consists of different land features such as agricultural land, vegetation, built up area, water bodies etc. can be characterized based on their distinct textural features (Kandaswamy et al., 2005). Texture analysis of remote sensing images can be performed

efficiently using Gray Level Co-occurrence Matrix (GLCM) method. GLCM is the two-dimensional matrix of joint probabilities $P_{d,r}(i,j)$, between pairs of pixels, separated by a distance *d*, in a given direction *r*. It calculates how frequently a pixel with gray level value i occurs horizontally adjacent to a pixel with the value j. Every element (i,j) in GLCM identifies the number of times that the pixel with value *i* occurs horizontally adjacent to a pixel with the value *j*. It is based on the repeated occurrence of some gray level pattern in the texture. The texture measured, and the related formulae are listed in Table 1.3.

Table 1.3 Texture measures used in the study

No.	Texture measures	Formulae
1	Mean (MEA)	$\mathrm{MEA} = \sum_{\mathrm{i,j}=0}^{\mathrm{N}-1} \mathrm{i}\left(\mathrm{P_{i,j}}\right)$
2	Variance (VAR)	$\mathrm{VAR} = \sum_{i,j=0}^{N-1} P_{i,j}(i-ME)^2$
3	Homogeneity (HOM)	$\mathrm{HOM} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$
4	Contrast (CON)	$\mathrm{CON} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$
5	Dissimilarity (DIS)	$\mathrm{DIS} = \sum_{i,j=0}^{N-1} P_{i,j}\lvert i-j\rvert$
6	Entropy (ENT)	$\mathrm{ENT} = \sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right)$
7	Second Moment (SM)	$\mathrm{SM} = \sum_{i,j=0}^{N-1} P_{i,j}^2$
8	Correlation (COR)	$\mathrm{COR} = \sum_{\mathrm{i,j}=0}^{\mathrm{N}-1} \mathrm{P_{i,j}}\left[\frac{(\mathrm{i}-\mathrm{MEA_i})(\mathrm{j}-\mathrm{MEA_j})}{\sqrt{(\mathrm{VAR_i})(\mathrm{VAR_j})}}\right]$

here $P_{i,j} = {V_{i,j}}/{\sum_{i,j=0}^{N-1} V_{ij}}$, where $V_{i,j}$ is the value in the cell i,j (row i and column j) of the moving window and N is the number of rows or columns.

1.18.3 LULC classification scheme and accuracy assessment

The multi-source remote sensing images were used to perform LULC classification. The visual interpretation and image analysis techniques were used to identify and delineate different LULC classes on the images. In this study, four machine learning classification techniques such as Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF) and Support Vector Machine (SVM) were used to perform LULC classification. A conventional statistical method Maximum Likelihood Classifier (MLC) was used to compare the performances of machine leaning classification techniques. The detailed description of these classification techniques will be discussed in respective chapters. However, the reasons for selecting them are given below.

MLC was selected because it is most commonly used and available in many software packages. ANN is based on artificial intelligence pattern recognition systems to determine class membership. It is selected because of its independence from the normal distribution of training data and non-parametric nature. It provides better results when non-spectral data are incorporated into the classification. It was found to provide more accurate LULC classification results when compared to MLC. DT is selected because of its ability to handle both numeric and non-numeric datasets on different measurement scales. It was found to provide reasonably good classification accuracy. The performance of RF and SVM algorithms is comparable to each other and has gained a lot of attention because of numerous advanced image-handling capabilities. Both have potentials to control over-training and

unbalanced data-sets. SVM requires less computational time and training data in comparison to other algorithms.

The classification results were assessed to test the validity and reliability of the classified LULC results. The assessment of classification results was performed using accuracy measures based on confusion matrix or error matrix approach (Congalton and Green 1999). The classification results were assessed by computing the accuracy measures such as User's Accuracy (UA), Producer's Accuracy (PA), Overall Accuracy (OA), and Kappa coefficient (Kc) using equations given as follows:

$$UA = \frac{n_{ii}}{n_{irow}} \tag{1.1}$$

$$PA = \frac{n_{ii}}{n_{icol}} \tag{1.2}$$

$$OA = \frac{1}{N}\sum_{i=1}^{r} n_{ii} \tag{1.3}$$

$$K_c = N\sum_{i=1}^{r} n_{ii} - \sum_{i=1}^{r} \frac{n_{icol}\, n_{irow}}{N^2} - \sum_{i=1}^{r} n_{icol}\, n_{irow} \tag{1.4}$$

where n_{ii} is the number of pixels correctly classified in a category; N is the total number of pixels in the confusion matrix; r is the number of rows; and n_{icol} and n_{irow} are the column (reference data) and row (predicted classes) total, respectively.

Some other accuracy measures such as F-measure, Hellden's Mean Accuracy (MAH), Short's Mapping Accuracy (MAS), Individual Classification Success Index (ICSI), and Classification Success Index (CSI) were also used for more comprehensive assessment of classification results (Liu et al., 2007). The F-measure, MAH, MAS and ICSI are class-level accuracy measures, whereas CSI is map-level accuracy measure and given as follows:

$$F - \text{measure} = \frac{2 \times UA_i \times PA_i}{UA_i + PA_i} \quad (1.5)$$

$$MAH = \frac{2n_{ii}}{(n_{icol} + n_{irow})} \quad (1.6)$$

$$MAS = \frac{n_{ii}}{(n_{irow} + n_{icol} - n_{ii})} \quad (1.7)$$

$$ICSI = UA_i + PA_i - 100 \quad (1.8)$$

$$CSI = \frac{1}{m}\sum_{i=1}^{m}(UA_i + PA_i) - 100 \quad (1.9)$$

where UA_i and PA_i are the user's and producer's accuracies of individual LULC class and m is the total number of classes.

The Receiver Operating Characteristic (ROC) curve analysis was also performed for visualizing the performance of classifiers.

1.18.4 Land change analysis, modelling and prediction

The post-classification method was employed to analyze the changes in LULC in Varanasi district of Uttar Pradesh, India. The multi-temporal Landsat images were used to produce LULC maps of different years using digital image classification techniques. The LULC maps of years 1988, 2000, 2001, 2005, 2009, 2014, and 2015 were used to investigate the changes occurred in the study area. Modeling of landscape dynamics was performed using spatially explicit hybrid models such as ST-MC, CA-MC and MLP-MC using LULC maps of different years within GIS environment. The comparison of models and prediction of the future LULC scenario was also carried out for Varanasi district of Uttar Pradesh, India.

The elaborated details of the data used, methodology, and procedure adopted under various studies are provided in the respective chapters.

1.19 MOTIVATION OF STUDY

Land use, land cover, and their interactions are longstanding interests for researchers around the world. The studies on geospatial distribution of LULC through digital image classification methods are a prerequisite for analyzing various environmental and socio-economic concerns. The changes in LULC, partly driven by rapid urbanization, are vital for the better understanding of landscape dynamic during a known period along with sustainable development and management of an area. It remarkably affects the natural environment, ecosystem and climate from local to global level. Therefore, mapping and monitoring of LULC while sustaining the production of essential natural resources has become a major priority for researchers and policymakers in India and around the world. A considerable quantity of research has been done in improving the accuracy of LULC classification and the analysis and modeling of landscape dynamics using remote sensing data and methods. However, the following motivated this study which primarily has an eventual aim to enrich the scientific knowledge and references of this particular subject.

Over the last few decades, several studies have been performed for LULC classification and mapping using well known remote sensing datasets. Nevertheless, LULC classification of complex landscapes is still a challenging task due to several reasons. A major problem is the frequent cloud cover, which may be somewhat repress through the use of SAR imagery. The presence of spectral confusion caused by land surface features is the main concern which makes it hard to acquire LULC classification results with high accuracy. Other constraints are the scarcity of previous LULC maps, and of ancillary data sources that may be used to improve the classification results.

Different sensor data may have various capabilities for LULC classification. For optical sensor data, the radiometric and temporal properties are constant, but the spectral and spatial properties are the most important features to be further explored. It is not always possible to acquire better classification results with improved spatial resolution images using pure spectral properties. High spatial resolution implies high within class spectral variability corresponding to land surface features, which decreases their spectral separability and leads to the lower accuracy of classification results. SAR data have unique features with dual or full polarization options (HH, HV, VV, and VH). Different polarizations have different sensitivities and backscattering response for the same ground object.

It is possible to produce new variables from multispectral bands using different image processing techniques, like vegetation indices, image transforms. There are various supervised and unsupervised classification techniques for extracting LULC information from remote sensing images. However, these techniques are based on individual pixels without incorporation of spatial relationships among them. The classification accuracy is likely to improve by incorporating the spatial information or features for spatially complex and spectrally mixed land surface features. The selection of adequate input variables has the significant role in acquiring desired classification results with high accuracy.

Spatial features reflect the relationships between one central pixel and its surroundings. Texture analysis of images supply valuable spatial information and can be investigated through its relationship with spectral information. Textural features represent a wealth of information regarding the spatial relation of pixels and can be extracted from the image itself. It can be incorporated as an ancillary data during the classification process.

Although a variety of image classification algorithms have been proposed, the classification accuracy is not ideal. A fundamental issue is still the selection of the suitable algorithm to address the problems related to LULC classification of heterogeneous landscapes. The use of machine learning techniques has gained a lot of attention in recent years. Therefore, it is equally important to select an appropriate classification algorithm for successful and accurate classification of landscape features, particularly in an urban environment. The most common approach of expressing the accuracy of remote sensing image classifications is confusion or error matrix. However, it has been criticized for not providing any information about the spatial distribution of errors. So, there is need of methods to analyze the spatial heterogeneity of error and to visualize the spatial variation in the measures of classification accuracy and uncertainty.

Varanasi district experienced rapid population growth and urban expansion in last few decades. So, understanding the past changes in LULC and predicting the future scenarios are vital to preserve natural resources and promote socio-economic welfare. The outcomes may be very valuable for the recently launched smart city initiative by Government of India. It could also be employed in other geographical locations of India and around the world.

Remote sensing data acquired by LISS-IV, LISS-III, Landsat 8-OLI, RISAT-1 and Sentinel-1A sensors have been used in this research work. RISAT-1 is the first Indian space-borne hybrid polarimetric SAR system launched on 26 April 2012 by Indian Space Research Organization (ISRO). Sentinel-1A was launched on 03 April 2014 by European Space Agency (ESA). The capabilities of all the above mentioned datasets have not been explored much for LULC classification based on texture analysis in a complex landscape. The

potential of multi-temporal remote sensing images from Landsat series satellites has not been investigated much for LULCC analysis for Varanasi district of Uttar Pradesh, India.

The outcomes are of substantial scientific and practical values to the broader remote sensing community and research studies on LULC and land change analysis. The classification approach suggested in this study is expected to provide an accurate and rapid mapping of land surface features in a complex landscape. This research work may also be useful for sustainable development and management of natural resources at different observational scales.

1.20 REVIEW OF LITERATURE

Thematic mapping describes the pattern and spatial distribution of LULC, is one of the core areas of remote sensing data applications (Foody, 2002; Qi et al., 2012; Jia et al., 2014; Idol et al., 2015; Zakeri et al., 2017). Regularly updated LULC information plays a crucial role in various socio-economic and environmental activities from regional to the global scales (Lu and Weng, 2007; Jensen, 2005). Remote sensing images, with larger geographical area overage and high temporal frequency, provide a unique opportunity for deriving LULC information through the process of image interpretation and classification. The interpretation of remotely sensed images utilizes many techniques including pattern recognition, artificial intelligence, computer vision, image processing and statistical analysis. The progress towards automated analysis of remotely sensed images is encouraged with increasing data volumes as well as the high cost of traditional ground surveying methods. The new generation of spaceborne remote sensing instruments is offering spatial and spectral resolution images, leading to the potential application of remote sensing products and further emphasizing the requirement for more automated analysis. Remote sensing image classification is the most

commonly used process in deriving spatially distributed LULC maps (Petropoulos et al., 2012; Jia et al., 2014; Hütt et al., 2016; Sonobe et al., 2017). Classification is the process of performing a conversion from the numerical spectral measurements into a set of meaningful classes or thematic information, which can describe a landscape. The process typically involves the use of the reflectance and radiance values of each pixel to assign it to some LULC categories of earth's surface. It is assumed that the information of a pixel originates solely from within its footprint. Spectral reflectance characteristics of a pixel, generally their response in different wavelength intervals are employed for the identification of various LULC categories (Richards, 1993). A set of samples for each pixel are analyzed to provide a class or label that associated with the pixel of a particular land surface feature.

The remote sensing data sets from Landsat series satellites having multispectral observations are the most common sources for LULC classification and mapping. The launch of space-borne satellites like IKONOS, QuickBird, SPOT, LISS-IV, LISS-III, AWiFS, Landsat 8-OLI, etc. are offering more opportunities in LULC mapping and change analysis studies. However, LULC classification remains a challenging task because of several reasons. The challenges come to rise mainly due to the heterogeneity of landscapes. One of the major problems is the difficulty in attaining cloud-free multispectral images, which may be overcome through the use of SAR imaging system. Some of the most common SAR systems are RADARSAT-1/2, ENVISAT ASAR, ALOS PALSAR, RISAT-1, Sentinel-1A, etc. Classification of remote sensing image is a comprehensive process that needs careful consideration of several aspects, such as user's requirements, the complexity of landscapes, the extent of the study region, classification system, choice of remote sensing variables and related classification algorithms, as well as the analyst's expertise (Lu and Weng, 2007). It is

essential to select an appropriate classification algorithm for deriving reliable LULC information efficiently from satellite remote sensing images. Many algorithms have been developed for digital image classification over the past decades and reviewed by Lu and Weng (2007). Most of the classification algorithms are based on information of image pixel, with which each pixel is assigned a distinct LULC class. Supervised classification algorithms are based on employing training samples collected directly from the image to be classified, and they try to group the spectrally similar pixels using various statistical approaches The classification algorithms based on pixel information including statistics-based (e.g. maximum likelihood, minimum distance), non-statistics-based (e.g. artificial neural network, decision tree, support vector machine, random forest, k-nearest neighbour) and from pixel-based to subpixel-based and object-oriented algorithms have been explored and various degrees of success are reported (Lu and Weng, 2007; Marpu et al., 2012; Li et al., 2012).

The conventional pixel-based classifier, like maximum likelihood, needs to have normal Gaussian distribution for representative pixel samples (Mather, 1999). This is not apparently true for remotely sensed images. It is also not able to handle mixed-pixel problems efficiently in complex landscapes (Jensen, 2005). Most favorable alternatives to the statistics-based classifiers are the advance classification algorithms based on machine learning theory that do not rely to the statistical nature of the data. These advance classification algorithms include ANN, DT, RF, SVM etc.

Although many neural networks have been developed, the most broadly used is the backpropagation ANN algorithm by the remote sensing community (Foody, 2004; Mass and Flores, 2008; Dixon and Candade, 2008; Szantoi et al., 2015). Several studies have been carried out on the use of back propagation ANN, and many guidelines have been reported by

research regarding fitting of ANN parameters for LULC classification (Kavzoglu and Mather, 2003; Srivastava et al., 2012; Kumar et al., 2015). ANN, with its non-parametric approach, facilitates to avoid some of the problems of MLC (Huang et al., 2002). It is capable of handling multitudes of data and can classify LULC categories that are not linearly separable in the original spectral space. However, it is noted that accuracy using ANN is rarely increases beyond 80% (Atkinson and Tate, 2000). According to Kavzoglu and Mather (2003), accurate description of the procedures that transform input data into output classes can be complicated because of the combined use of multiple nonlinear activation functions at different layers. Therefore it is frequently referred as a "black box" technique (Qiu and Jensen, 2004). The classification accuracy may also be affected by the required amount of training data which increases exponentially with the increased dimensionality of the input data and number of hidden nodes (Foody and Arora, 1997; Dixon and Candade, 2008).

DT, a non-parametric approach that can train rapidly with the ability to handle data sets from numeric and non-numeric sources on different measurement scales. DT classifiers have been used successfully for LULC classification of digital images (Pal and Mather, 2003; Kandrika and Roy, 2008; Punia et al., 2011; Aimaiti, 2016). DT classifiers are appropriate for remote sensing classification problems due to its flexibility, intuitive simplicity and computational efficiency, leading to increased acceptance (Quiland, 1993; Pal and Mather, 2003). In several studies, the performance of DT is compared with other supervised classifiers and found to provide reasonably good classification accuracy (Pal and Mather, 2003; Szantoi et al., 2015; Zeyada et al., 2016). In some studies, it is also reported that DT performed better in comparison to SVM for classifying various land surface features (Otukei and Blaschke, 2010; Li et al., 2015). In spite of providing acceptably good accuracy, DT

classifiers are not suggested for high-dimensional data sets. The selection of an appropriate pruning method also affects the improvement in classification accuracy positively, whereas the use of attribute selection measures was not found to be significant (Pal and Mather, 2003).

RF classifier is a robust and ensemble machine learning method proposed by Breiman (2001). RF is a combination of decision trees used as the base classifiers. It works on large data-sets efficiently and does not require normally distributed training data samples (Rodriguez-Galiano et al., 2012). Among the advanced classification algorithms, RF and SVM have received the considerable amount of attention due to several superior image-handling abilities. They can synthesize regression or classify functions based on discrete or continuous data sets, obtuse to over-training, and capable of controlling unbalanced data-sets (Breiman 2001). In several studies, RF is used for LULC classification using common remotely sensed data sets (Pal 2005; Gislason et al., 2006; Rodriguez-Galiano et al., 2012; Puissant et al., 2014; Adam et al., 2014). The advantage of RF classifier is that it needs only two parameters to be set while the SVM require several user-defined parameters. The accuracy achieved by RF classifier is also comparable to that obtained by SVM (Pal 2005; Adam et al., 2014). However, RF is bias-prone in conditions where the number of cases is unequally distributed among the classes of interest (Puissant et al., 2014).

SVM based on machine learning technique, is one of the excellent additions to the existing catalogue of image classification algorithms (Vapnik, 1998). SVM can be applied successfully to the problems of image classification with high input dimensionality and small training datasets (Roli and Fumera, 2000; Melgani and Bruzzone, 2004; Pal and Mather, 2005; Hosseini and Ghassemian, 2011). Gong et al. (2013) applied SVM to create global

land cover data set using Landsat TM and ETM+ satellite images. The classified products from SVM were compared with other classification algorithms such as MLC, ANN, DT and found that it outperformed the others (Huang et al., 2002; Pal and Mather, 2005; Dixon and Candade, 2008; Szuster et al., 2011; Noi and Kappas, 2018). SVM in comparison to ANN provides a probabilistic guarantee on how well the classifier will generalize on unseen data (Perkins et al., 2001). It operates on structural risk minimization (SRM) principle, which minimizes the error on unseen data (Dixon and Candade, 2008). SVM also tends to be less prone to the overfitting problems than the other techniques because it relies on the number of support vectors rather than the dimensionality of the transformed space (Duda et al., 2002). It is easy to achieve acceptable thematic accuracy by SVM at less computational or training time and sampling cost (Lizarazo, 2008; Dixon and Candade, 2008). However, the performance of SVM is affected by the kernel parameters, selection of training samples, and input variables (Huang et al., 2002; Dixon and Candade, 2008).

These pixel-based classification algorithms, when applied to complex areas, however, are with restrictions, as the object size may be much smaller than the pixel size. A pixel may not have only a single LULC type, but a mixture of several LULC categories. The spectral information of the LULC categories being classified is the most important in information extraction from remote sensing images. Nevertheless, it is hard to acquire classification results with high accuracy due to the spectral confusion caused by land surface features (Rongqun and Daolin, 2011). Traditional spectral classification algorithms employed on a pixel-by-pixel basis ignores the spatial information, which is inherited in remote sensing images (Li et al., 2014). This subject becomes critical with the availability of high spatial resolution remote sensing images. The higher the spatial resolutions, images are likely to

comprise higher spectral variability within the class. As a consequence, less satisfactory results have been achieved using spectral classifiers (Myint et al., 2011). Therefore, in addition to the development of advanced classification algorithms, a fundamental issue is the adequate selection of input variables. Texture analysis of images is one of the conventional methods to extract spatial information. Texture refers to the frequency of change and arrangement of image tones. The perception of texture is supposed to play an essential part in human visualization for classification. Texture is created by aggregating unit features that may be too small to be distinguished individually on an image. It is often employed to quantify the variability of pixels in a neighborhood (Jensen, 2005). The first use of texture analysis was reported by Fasler (1980) using SEASAT L-band HH polarization data for urban applications. Several methods have been proposed to extract spatial information from an image (Haralick, 1973; Haralick, 1979; Gool et al., 1985: Marceau et al., 1990). The GLCM proposed by Haralick (1973) is the most popular of all the methods. In addition to the GLCM, other methods for extracting statistical texture properties including grey level difference histogram (GLDH), grey level run length matrices (GLRLM), neighbouring grey level dependency matrix (NGLDM), sum and difference histograms (SADH), power spectrum (PS), simple statistical transformations (SST), and texture spectrum (TS) are available. Other texture measures such as Gabor wavelets and wavelet packets are also popular in remote sensing image analysis (Manjunath and Ma, 1996; Pun and Lee, 2003).

Textural features are the potential source of ancillary information and their benefits in improving LULC classification accuracy have been reported in several studies (Gong et al., 1992; Herold et al., 2004; J. Paneque-Gálvez et al., 2013; Wang and Zhang, 2014; Zakeri et al., 2017). The GLCM method can be used to extract textural images that can be incorporated

into data bands during the classification process (Gong et al., 1992). With all the fourteen texture measures developed by Haralick et al., (1973), eight of them, including mean, variance, contrast, entropy, homogeneity, dissimilarity, angular second moment and correlation have been extensively applied and achieved reasonably satisfactory results (Pacifici et al., 2009). Although the textural images have been widely explored and implemented in earlier studies, but how different sensor data with different spatial resolutions affect the choice of textural images and how it influences LULC classification accuracy are inadequately understood. The identification of suitable textural images is a challenging task due to several reasons such as image bands, moving window sizes, quantization levels etc. (Chen et al., 2004). There is the unavailability of general guidelines for the selection of an optimal texture because of different spatial patterns and compositions of the LULC types under investigation (Lu et al., 2014). The texture analysis of images becomes essential in extracting spatial information to improve LULC classification accuracy for heterogeneous landscapes. The challenges still remain on how to efficiently employ textural information in detailed and improved LULC mapping.

Assessing the accuracy of regional to global scale thematic maps is vital to support most of the mapping projects and scientific applications around the world (Foody, 2002). In remote sensing, the confusion matrix and its associated measures, such as OA, PA and UA, are the conventional approach for assessing the mapping accuracy (Congalton and Green, 1999). However, there are some restrictions associated with the conventional confusion matrix method. The error summaries provided by confusion matrix is aspatial in nature. It does not contribute information regarding the spatial distribution of errors, and a global estimation of overall accuracy may be inappropriate for local sub-regions (McGuire and

Fisher, 2001; Foody, 2005; Comber et al., 2012). Some methods have been developed to overcome the problem of estimating and describing the spatial distribution of errors in landscape mapping accuracy. A phenomenon differs across a landscape due to spatial non-stationarity. So, this spatial variability restricts to use any conventional global regression technique which leads to erroneous conclusions in spatial analysis and generates spatially autocorrelated residuals (LeSage and Pace, 2001). Some local regression techniques have been proposed to overcome the challenges caused by spatial non-stationarity. One of the best alternatives to spatial regression is the Geographically Weighted Regression (GWR), a statistical local regression technique that explicitly deals with spatial non-stationarity (Brunsdon et al., 1996; Fotheringham et al., 2002). GWR creates spatially varying data for the relationships among variables compared to conventional regression techniques. Several studies have demonstrated the utility of GWR for different applications (Brunsdon et al., 2001; Foody, 2005; Comber et al., 2012; Tsutsumida and Comber, 2015, Maimaitijianga et al., 2015).

An improved and accurate LULC classification is of critical significance in identifying the changes on earth's surface. It is required to determine the number of discrete LULC classes that are considered to be the most appropriate depiction of the landscape of study area. Land cover is observed directly while the interpretation of images with excellent field knowledge is needed to determine land use from remotely sensed images (Anderson et al. 1976). Remote sensing images are stored in the digital format and then processed by the computers to generate images for interpretation purposes.

In spite of the several studies, researchers have never ceased of dipping in this field of land change analysis. The earth is undergoing a crucial alteration, is continuously threatening

the humankind, therefore, LULCC studies represent an urgent issue around the world. Satellite remote sensing has emerged as an essential tool in land change science (Ellis and Pontius 2011; Feng et al., 2017). It can provide spatially explicit and continuous information on LULCC, which can assist in understanding the direction and quantity of changes (Elmore et al., 2000; Rogan and Chen, 2004; Buenemann et al., 2011; Feng et al., 2017). There are different methods of approaching the utilization of satellite images for determining LULCC in heterogeneous landscapes. Yuan et al., (1998) categorized the change detection methods into pre-classification and post-classification. The pre-classification methods apply a range of algorithms, including image differencing and image ratioing, to single or multiple spectral bands, directly to multi -date satellite images to generate ‘‘change’’ vs. ‘‘no-change’’ maps. These methods provide only the location but not the nature of changes (Ridd and Liu, 1998; Yuan et al., 1998). On the other hand, post-classification methods employ separate classified maps of images acquired at different times to generate difference maps providing ‘‘from - to’’ change information (Jensen, 2005). Several techniques of LULCC detection are available and have been reviewed by Lu et al., (2004). In many studies, images from space-borne remote sensing technology have been explored in providing accurate and timely geospatial information describing the changes in LULC (Yuan et al., 2005; Otukei and Blaschke, 2010; Kibert et al., 2016; Lamine et al. 2017). Several studies have been performed on LULCC using remote sensing images in India (Singh et al. 2014; Misra et al. 2015; Rawat and Kumar 2015). Singh and Jeganathan (2016) used time-series satellite data to determine the spatio-temporal pattern of forest changes in Palamu District of Jharkhand, India. Mishra and Rai (2016) carried out a study to analyze, simulate, and predict the future scenario of changes in LULC for Patna district of Bihar, India.

Growing understanding regarding the importance of sustainable development is stimulating the improvement in current methods to understand better the landscape evolution, which is the outcome of complex interactions between physical, biological and social forces in time and space (Turner, 1987; Avtar et al., 2017). Modelling of landscape dynamics begin with the method of examining what, where and to what extents landscape change has occurred, and also, to understand how and why the changes happened (Weng, 2002, Yang and Lo, 2002). Modeling and analysis of LULCC would not be possible without a significant amount of remote sensing data which is one of the most widely used data sources used in GIS. A fundamental problem is a complexity of finding an effectual model that can incorporate both, spatial and temporal information into predicting future scenarios (Tang et al., 2007). Such spatially explicit model is crucial for the analysis, understanding, demonstration and modelling of landscape dynamics. Remote sensing data and techniques in combination with GIS have contributed to support the development of robust models. Consequently, various modelling methods have particular focus because of the complexity of LULC changing systems and their interactions with the underlying driving forces (Zhu et al., 2016). In recent years, many simulation and prediction models have been used within a GIS environment to determine realistic future scenarios of LULC patterns. These models include markov chain (MC) (Kumar et al., 2014; Fathizad et al., 2015), artificial neural network (ANN) model (Maithani, 2015; Shooshtari and Gholamalifard, 2015; Losiri et al., 2016), cellular automata (CA) model (Mitsova et al., 2011; Al-sharif and Pradhan, 2014; Rimal et al., 2017), logistic regression (LR) model (Kumar et al., 2014; Shooshtari and Gholamalifard, 2015), GEOMOD (Paudel and Yuan 2012; Y. Sakieh and Salmanmahiny, 2016), SLEUTH model (Hua et al., 2014), conversion of land use and its effects (CLUE) model (Zhu et al.,

2010; Zheng et al., 2015). Musa et al., (2016) reviewed various models based on their spatial, temporal and complexity in selection. The problems associated with an individual model must be overcome by integrating them to work as complementary to each other. Therefore, in recent years several hybrid models have been developed to improve the demonstration of landscape dynamics and processes (Kamusoko et al., 2009; Thapa and Murayama, 2012; Bozkaya et al., 2015; Mishra and Rai, 2016). The performance of land change modelling methods is different for different study area due to varied environmental conditions and landscapes (Arsanjani et al., 2011). The comparison of models and prediction of the future scenario of LULC using best result providing model are gaining the considerable amount of attention in remote sensing community (Mas et al., 2014; Mozumder et al., 2016).

In the context of above discussion, this PhD work contributes to the research studies on LULC classification and change analysis in Varanasi district of Uttar Pradesh, India. Firstly, it endeavors to establish a robust classification approach based on ancillary data sources to map accurately all the broad LULC classes considered in a heterogeneous landscape of Varanasi. Secondly, it attempts to investigate the changes in LULC and predict the future landscape scenario based on multi-temporal earth observation datasets integrated with land change models.

1.21 ORGANIZATION OF THESIS

The research work described in this thesis covers the period from July 2013 to December 2017. New ideas were developed obviously during the study period, and the research is still progressing in areas such as the establishment of more robust texture-based classification approach, integration of optical and SAR images, development of new prediction models by incorporating more socio-economic, physical and climatic variables.

The organization of thesis provides cursory glance and a brief introduction of all the chapters and the studies carried out. The thesis has been organized as follows:

Chapter 1 pertains to the introduction of LULC classification and change analysis, research background, research objectives, data used and methodology. The literature review is also provided to develop insights about the previous studies carried out by various researchers in the domain of LULC classification, land change analysis and modeling.

Chapter 2 pertains to the evaluation of dual-polarimetric RISAT-1 data for LULC classification based on textural images. The GLCM approach was used to extract the textural images. Inter-class separability analysis was performed to identify the potential textural images. The best combination of textural images was identified using the standard deviation of preferred textural images and correlation coefficients between them. Also, the classification results of MLC, ANN, RF and SVM were compared to recognize suitable method.

Chapter 3 pertains to the performance evaluation of textural images in improving LULC classification accuracy in a complex landscape using multi-source remote sensing images with varying spatial resolution. Textural images were extracted from Landsat 8-OLI, RISAT-1, LISS-III, Sentinel-1A and LISS-IV images with spatial resolution of 30, 25, 23.5, 20 and 5.8 m respectively using GLCM method. It was examined that how the images from different sensors with varying spatial resolutions influence the selection of textural images. A supervised SVM classifier is then employed to carry out LULC classification based on different combinations of textural images from different sensor data. The significance of the incorporation of textural features to improve classification accuracy was compared with the individual datasets.

Chapter 4 pertains to explore the feasibility of dual-polarimetric RISAT-1 data at C-band for rice crop mapping. The composite images namely ratio (HH/HV) and difference (HH−HV) were produced to enhance the backscatter difference between rice and non-rice categories. The separability index was calculated for each class pair. The DT classifier (QUEST) is then employed to map rice crop areas. Results were further compared with Landsat-8 OLI optical sensor-derived rice crop map.

Chapter 5 pertains to describe a geographically weighted method combined with logistic regression for producing and visualizing spatially distributed accuracy measures. A kernel-based procedure characterizes the data and weighting schemes that are used to calculate the varying accuracies at each location across the landscape.

Chapter 6 pertains to evaluate three classification techniques namely MLC, ANN and SVM using multi-temporal Landsat series images. Assessment of spatio-temporal LULCC occurred in a fast growing Varanasi district of Uttar Pradesh, India over a period of 14 years (2001–2014) based on best recognized classification technique. A paired samples t-test was also carried out to determine the statistical significance of changes in LULC between different time periods.

Chapter 7 pertains to evaluate and compare three MC-based hybrid land change models namely ST-MC, CA-MC and MLP-MC to predict future landscape scenario in Varanasi district of Uttar Pradesh, India. LULC information was extracted from Landsat TM/ETM+/OLI images using SVM classifier for years 1988, 2001 and 2015 respectively. The validation of models is carried out using kappa index statistics. The prediction of future LULC scenario for years 2030 and 2050 were performed based on the better result providing MLP-MC hybrid model.

Chapter 8 presents the overall conclusions drawn from this research work and scope for the future work.