

**ASSESSMENT OF CROP AREA AND PRODUCTION
OF MAJOR CROPS OF SEDAM TALUK OF
GULBARGA USING REMOTE SENSING AND GIS**

SHALINI M N

PALB 1161

**DEPARTMENT OF AGRICULTURAL STATISTICS,
APPLIED MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF AGRICULTURAL SCIENCES
GKVK, BENGALURU - 560 065**

2015

**ASSESSMENT OF CROP AREA AND PRODUCTION
OF MAJOR CROPS OF SEDAM TALUK OF
GULBARGA USING REMOTE SENSING AND GIS**

SHALINI M N

PALB 1161

Thesis submitted to the

UNIVERSITY OF AGRICULTURAL SCIENCES, BENGALURU

In partial fulfilment of the requirements

For the award of the degree of

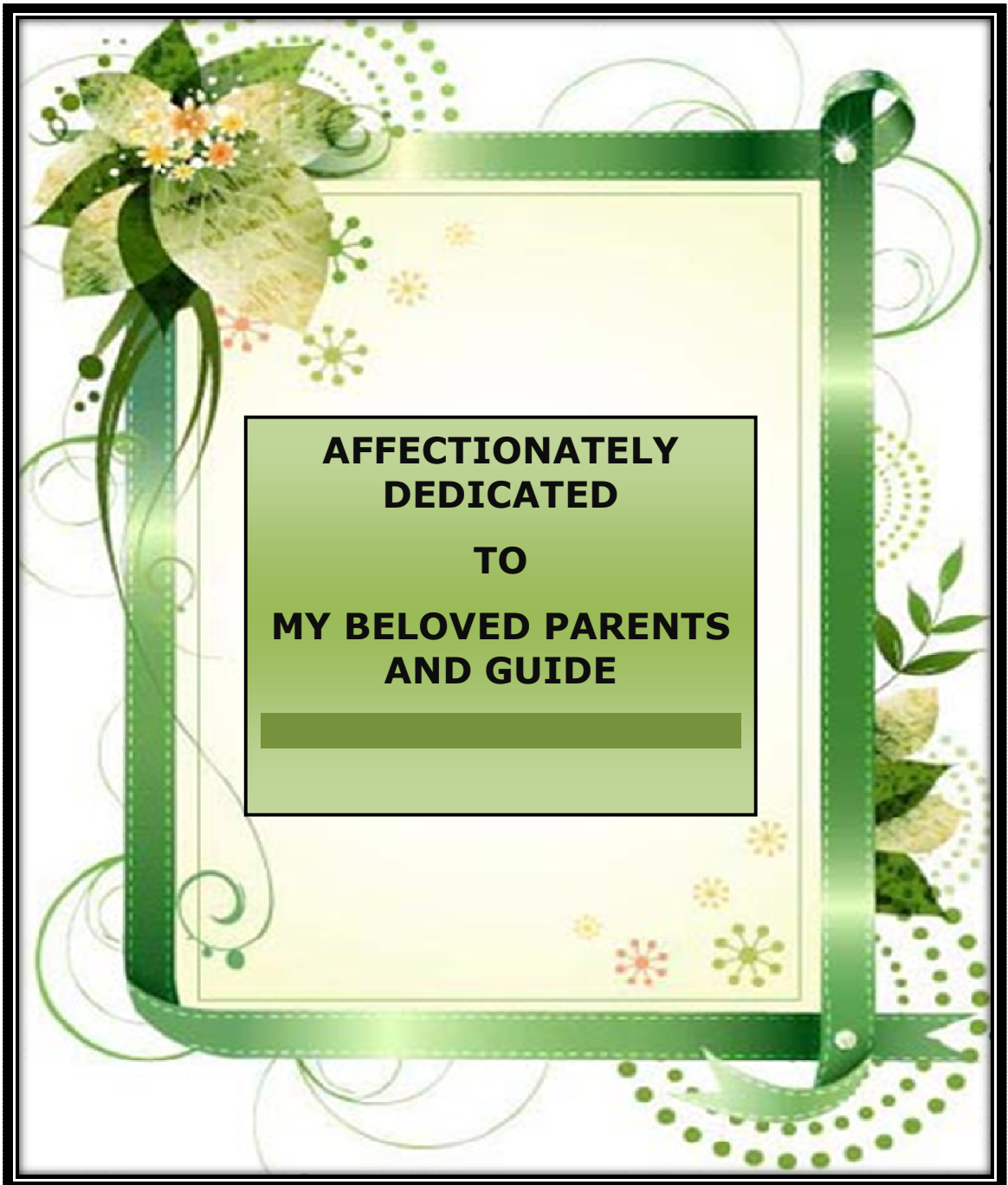
MASTER OF SCIENCE (AGRICULTURE)

in

AGRICULTURAL STATISTICS

BENGALURU

SEPTEMBER, 2015



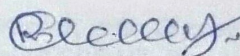
**AFFECTIONATELY
DEDICATED
TO
MY BELOVED PARENTS
AND GUIDE**

**DEPARTMENT OF AGRICULTURAL STATISTICS,
APPLIED MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF AGRICULTURAL SCIENCES
GKVK, BENGALURU-560 065**

CERTIFICATE

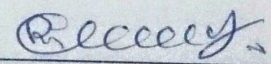
This is to certify that the thesis entitled "ASSESSMENT OF CROP AREA AND PRODUCTION OF MAJOR CROPS OF SEDAM TALUK OF GULBARGA USING REMOTE SENSING AND GIS" submitted by SHALINI, M. N.; I.D NO. PALB 1161, in partial fulfilment of the requirements for the degree of **MASTER OF SCIENCE (AGRICULTURE) IN AGRICULTURAL STATISTICS** to the University of Agricultural Sciences, Bengaluru, is a record of research carried out by her under my guidance and supervision and thesis has not previously formed the basis for the award of any degree, diploma, associate ship, fellowship or other similar titles.

Bengaluru
September, 2015


Mr. R. MUNIRAJAPPA
Associate Professor,
AICSMIP, ZARS, UAS
GKVK, Bengaluru-65
Major Advisor

Approved By

Chairperson:

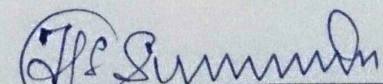


(R. MUNIRAJAPPA)

Members: 1.

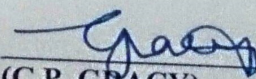
(H. CHANDRASHEKAR)

2.



(H.S. SURENDRA)

3.



(C.P. GRACY)

ACKNOWLEDGEMENT

*It gives me immense pleasure to express my deep sense of gratitude, indebtedness and sincere thanks to **Mr. Munirajappa. R**, Associate professor, AICSMIP, ZARS, UAS, GKVK, Bengaluru and Chairman of my advisory committee for his invaluable guidance, co-operation, encouragement, continuous help and moral support throughout the period of my research. I admit that it has been a great fortune for me to be associated with him during my degree programme.*

*I am grateful and obliged to my advisory committee members, **Dr. Chandrashekar. H**, Professor of Statistics, Co-ordinator, PPMC, GKVK, Bengaluru, **Mr. Surendra. H.S**, Associate Professor, GKVK, UAS, Bengaluru, **Dr. Gracy. C.P**, Professor, Department of Agricultural Marketing Co-operation and Business Management, UAS, Bengaluru, for their constant supervision, invaluable guidance and all the facilities extended during the course of this investigation.*

I wish to express my sincere thanks to all staff of Department of Agricultural Statistics, Applied Mathematics and Computer Sciences, UAS, GKVK, Bengaluru, for their sheer guidance during my period of study.

*I extend my sincere thanks to **Dr. Prabhuraj**, Director, **Mr. Naveen Kumar**, Scientist, ISRO, Bengaluru, **Dr. Ashok Reddy**, Sr. Scientist, KSRSAC, Bengaluru, who helped me, for providing necessary software facilities and digital images of IRS P6 LISS III for my study.*

*There are no words to express my feelings of adoration, love, respect and obligation to my beloved parents **Sri. C. P. Nagaraj** and **Smt. K. C. Padmavathi** who moulded me to what I am now. I sincerely thank them for their love, encouragement and constant inspiration.*

*I would like to express my deep sense of gratitude to my sisters, **Smt. Kavitha** and **Smt. Mamatha**, my brother in laws **Dr. Shankar** and **Mr. Manjunath** and My cousin **Ms. Shilpa** for their encouragement and support throughout my study period.*

*I found no suitable words to express my heartfelt thanks to my friends **Chandana**, **Bindhu**, **Bindhushree**, **Ramya**, **Sowmya**, **Keerthana**, and **Lavanya** for their lovely friendship, encouragement and help.*

Bengaluru

September, 2015

(**Shalini M N**)

ASSESSMENT OF CROP AREA AND PRODUCTION OF MAJOR CROPS OF SEDAM TALUK OF GULBARGA USING REMOTE SENSING AND GIS

SHALINI, M. N

ABSTRACT

Remote sensing is the art and science of making measurements of the earth using sensors on airplanes or satellites. These sensors collect data in the form of images and provide specialized capabilities for manipulating, analysing, and visualizing those satellite images. Remote sensed imagery is integrated within a Geographic Information System (GIS). Several methods exist for remote sensing image classification. They include supervised and unsupervised approaches. In supervised classification, the classifier is trained to identify the classes using training data set where as in an unsupervised classification, the classifier itself develops the spectral classes. For statistical comparisons of different image classification methods test imagery obtained through IRS (Indian Remote Sensing) P6 LISS-III on 22nd December, 2011 for Sedum Taluk of Gulbarga district of Karnataka state were used. Maximum likelihood classification, Mahalanobis distance classification, Minimum distance to means classification under supervised classification approach were performed for the test imagery using ERDAS imagine 9.1 version software. Accuracy assessment was found important to evaluate the final output of remote sensing. Accuracy of the classification of each data set and their classifiers were expressed as an error matrix from which overall accuracy, users accuracy, producers accuracy, mapping accuracy, F-measures, Kappa coefficients and sample variance of Kappa coefficients were estimated. The test of significance of Kappa coefficients were performed using Z-test. Pairwise comparison of Kappa coefficients of different classification methods were performed. Maximum likelihood classification was found to be the best with highest overall accuracy of 93.38 per cent for Sedum Taluk, Gulbarga district of Karnataka.

September, 2015

Department of Agricultural Statistics,
Applied Mathematics and Computer Sciences
UAS, G.K.V.K, Bengaluru-560065

Munirajappa. R
Major Advisor

ದೂರಸಂವೇದಿ ಮತ್ತು ಜಿಐಎಸ್ ಮೂಲಕ ಗುಲ್ಬರ್ಗದ ಸೇಡಂ ತಾಲ್ಲೂಕಿನ ಪ್ರಮುಖ ಬೆಳೆಗಳ
ಪ್ರದೇಶ ಮತ್ತು ಉತ್ಪಾದನೆಯ ಮೌಲ್ಯಮಾಪನ

ಶಾಲಿನಿ, ಎಂ. ಎನ್.

ಸಾರಾಂಶ

ವಿಮಾನಗಳು ಅಥವಾ ಉಪಗ್ರಹಗಳಲ್ಲಿ ಸಂವೇದಕಗಳನ್ನು ಬಳಸಿ ಭೂಮಿಯ ಅಳತೆಯನ್ನು ಮಾಡುವ ಕಲೆ ಹಾಗೂ ವಿಜ್ಞಾನ ವನ್ನು ದೂರ ಸಂವೇದನೆಯೆಂದು ಕರೆಯುತ್ತಾರೆ. ಈ ಸಂವೇದಕಗಳು ಚಿತ್ರಗಳ ರೂಪದಲ್ಲಿ ಮಾಹಿತಿ ಸಂಗ್ರಹಿಸಿ ಮತ್ತು ಆ ಚಿತ್ರಗಳನ್ನು ದೃಶ್ಯೀಕರಿಸುವ ವಿಶೇಷ ಸಾಮರ್ಥ್ಯಗಳನ್ನು ಒದಗಿಸುತ್ತದೆ. ಹಲವು ವಿಧಾನಗಳು ದೂರಸಂವೇದಿ ಚಿತ್ರ ವರ್ಗೀಕರಣದ ಅಸ್ತಿತ್ವದಲ್ಲಿವೆ. ಇವುಗಳಲ್ಲಿ ಮೇಲ್ವಿಚಾರಣೆ ಮತ್ತು ಉಸ್ತುವಾರಿಯಿಲ್ಲದ ವಿಧಾನಗಳು ಸೇರಿವೆ. ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ವಿಧಾನಗಳು ವ್ಯಾಪಕವಾಗಿ ದೂರಸಂವೇದಿ ಅಂಕಿಅಂಶಗಳ ವಿಶ್ಲೇಷಣೆಗೆ ಬಳಸಲಾಗಿದೆ. ವಿವಿಧ ಚಿತ್ರ ವರ್ಗೀಕರಣ ವಿಧಾನಗಳ ಅಂಕಿಅಂಶಗಳ ಹೋಲಿಕೆಗೆ ೨೦೧೧ ರ ಡಿಸೆಂಬರ್ ೨೨ ರಂದು ಐ.ಆರ್.ಎಸ್ ನ ಪಿ-೬ ಲಿಸ್-III ಮೂಲಕ ಪಡೆದ ಕರ್ನಾಟಕ ಗುಲ್ಬರ್ಗ ಜಿಲ್ಲೆಯ ಸೇಡಂ ತಾಲ್ಲೂಕಿನ ಚಿತ್ರಣವನ್ನು ಬಳಸಲಾಗಿದೆ. ಪರಿಕ್ಷೆ ಚಿತ್ರಣಗಳನ್ನು ಗರಿಷ್ಠ ಸಾಧ್ಯತೆ ವರ್ಗೀಕರಣ, ಮಹಲನೊಬಿಸ್ ಅಂತರ ವರ್ಗೀಕರಣ ಮತ್ತು ಕನಿಷ್ಠ ಅಂತರ ವರ್ಗೀಕರಣ, ಮೇಲ್ವಿಚಾರಣೆ ವಿಧಾನ ಅಡಿಯಲ್ಲಿ ಇ. ಆರ್.ಡಿ.ಎ.ಎಸ್ ೯.೧ (ERDAS 9.1) ಆವೃತ್ತಿ ತಂತ್ರಾಂಶವನ್ನು ಬಳಸಿ ವರ್ಗೀಕರಣವನ್ನು ಮಾಡಲಾಗಿದೆ. ದೂರಸಂವೇದಿ ವರ್ಗೀಕರಣ ವಿಧಾನಗಳ ನಿಖರತೆ ಪರಿಕ್ಷಿಸಲು ನಿಖರತೆ ಮೌಲ್ಯಮಾಪನ ಬಹು ಮುಖ್ಯವಾಗಿದೆ. ಪ್ರತಿ ಮಾಹಿತಿ ಸೆಟ್ ವರ್ಗೀಕರಣದ ನಿಖರತೆ ದೋಷ ಮ್ಯಾಟ್ರಿಕ್ಸ್ ಮೂಲಕ ಪರಿಕ್ಷಿಸಲಾಗಿದೆ. ವಿವಿಧ ವರ್ಗೀಕರಣ ವಿಧಾನಗಳ ನಿಖರತೆಯನ್ನು ಪರಿಕ್ಷಿಸಲು ಒಟ್ಟಾರೆ ನಿಖರತೆ, ಬಳಕೆದಾರರ ನಿಖರತೆ, ನಿರ್ಮಾಪಕರ ನಿಖರತೆ, ಮ್ಯಾಪಿಂಗ್ ನಿಖರತೆ, ಎಫ್-ಅಳತೆ, ಕಪ್ಪಾ ಗುಣಾಂಕಗಳನ್ನು ದೋಷ ಮ್ಯಾಟ್ರಿಕ್ಸ್ ಬಳಸಿಕೊಂಡು ಎಣಿಸಲಾಯಿತು. ಕಪ್ಪಾ ಗುಣಾಂಕಗಳ ಪ್ರಾಮುಖ್ಯತೆಯನ್ನು ಜ್ಞೆ ಪರಿಕ್ಷೆ (Z-test) ಮೂಲಕ ನಡೆಸಲಾಯಿತು. ಪ್ರತಿ ಎರಡು ಕಪ್ಪಾ ಗುಣಾಂಕಗಳ ನಡುವೆ ಇರುವ ವ್ಯತ್ಯಾಸದ ಪ್ರಾಮುಖ್ಯತೆಯನ್ನು ಪರಿಕ್ಷಿಸಲಾಯಿತು. ಗರಿಷ್ಠ ಸಾಧ್ಯತೆ ವರ್ಗೀಕರಣ ವಿಧಾನದಲ್ಲಿ ಎಲ್ಲಾ ವರ್ಗೀಕರಣ ವಿಧಾನಗಳಿಗಿಂತ ಹೆಚ್ಚು ನಿಖರತೆ (೯೩.೩೮ ಶೇಖಡ) ದೊರೆತಿದೆ. ಆದ್ದರಿಂದ ಈ ಗರಿಷ್ಠ ಸಾಧ್ಯತೆ ವರ್ಗೀಕರಣ ವನ್ನು ಅತ್ಯುತ್ತಮ ವಿಧಾನ ಎಂದು ಪರಿಗಣಿಸಲಾಗಿದೆ.

ಕೃಷಿ ಸಂಖ್ಯಾಶಾಸ್ತ್ರ ವಿಭಾಗ,

ಕೃ.ವಿ.ವಿ., ಗಾ.ಕೃ.ವಿ.ಕೆ, ಬೆಂಗಳೂರು- ೬೫

ಆರ್.ಮುನೀರಾಜಪ್ಪ

(ಪ್ರಮುಖ ಸಲಹೆಗಾರರು)

CONTENTS

CHAPTER	TITLE	PAGE No.
I	INTRODUCTION	1-9
II	REVIEW OF LITERATURE	10-19
III	MATERIAL AND METHODS	20-36
IV	RESULTS	37-50
V	DISCUSSION	51-56
VI	SUMMARY	57-58
VII	REFERENCES	59-62

LIST OF TABLES

TABLE No.	TITLE	PAGE No.
1.1	Electromagnetic regions, their wavelengths and Salient features.	4
3.1	Error Matrix built using two data sets, classification and reference data.	30
3.2	Error matrix for Precision and Recall.	32
4.1	Class Means Estimated from Training site	39
4.2	Descriptive statistics of training site for class jowar.	39
4.3	Descriptive statistics of training site for class red gram.	39
4.4	Descriptive statistics of training site for class other crops.	39
4.5	Descriptive statistics of training site for class fallow.	40
4.6	Descriptive statistics of training site for class water body.	40
4.7	Variance-covariance matrix estimated from training site for land cover class jowar.	40
4.8	Variance-covariance matrix estimated from training site for land cover class red gram.	40
4.9	Variance-covariance matrix estimated from training site for land cover class other crops.	41
4.10	Variance-covariance matrix estimated from training site for land cover class fallow.	41
4.11	Variance-covariance matrix estimated from training site for land cover class Water body.	41
4.12	Area estimates estimated using different classification methods and actual area estimates as per DOA.	42
4.13	Production estimates estimated using different classification methods and Actual production estimates for the study area.	42

TABLE No.	TITLE	PAGE No.
4.14	Absolute deviation values of area estimates obtained by classifiers from actual estimates of DOA.	43
4.15	Absolute deviation of production estimates obtained by classifiers from actual estimates of DOA.	43
4.16	Classification results obtained with Maximum Likelihood classification.	44
4.17	Classification results obtained with Mahalanobis distance classification.	45
4.18	Classification results obtained with Minimum distance classification.	45
4.19	User`s and Producer`s accuracy of three classification methods.	46
4.20	Test of significance of Kappa coefficients for different classification methods.	46
4.21	The Pair wise comparison between Kappa coefficients	46
4.22	The overall accuracy and mapping of three methods.	47
4.23	Recall, Precision and F-measures for classes under Maximum likelihood classification.	48
4.24	Recall, Precision and F-measures for classes under Mahalanobis distance classification.	48
4.25	Recall, Precision and F-measures for classes under Minimum distance classification.	49
4.26	The overall F-measure for all three classifications.	49
4.27	Deviation values of area estimates of classifiers from the actual estimates of DOA and test of significance of Friedman two way analysis of Variance using area estimates.	50

LIST OF FIGURES

FIGURE No.	TITLE	BETWEEN PAGES
1	Satellite Chronology	1-2
2	Mechanism of remote sensing data acquisition.	2
3	Wave length Reflectance of water, soil and vegetation.	3
4	Procedure for Image classification.	7

LIST OF PLATES

PLATE No.	TITLE	BETWEEN PAGES
1	Location map of study area in Karnataka state	21-22
2	IRS-P6 LISS III Satellite image of Sedam Taluk, Gulbarga District.	21-22
3	Supervised classification map of Sedam Taluk, Gulbarga District using Maximum likelihood algorithm	41-42
4	Supervised classification map of Sedam Taluk, Gulbarga District using Mahalanobis distance algorithm.	41-42
5	Supervised classification map of Sedam Taluk, Gulbarga District using Minimum distance algorithm.	41-42



INTRODUCTION

I INTRODUCTION

India is primarily an agriculture based country. Agriculture is the backbone of Indian economy and our country development depends on agriculture.

The Department of Agriculture and Cooperation under the Ministry of Agriculture is the nodal organisation responsible for the development of the agriculture sector in India. Under it, several other bodies such as the National Dairy Development Board (NDDB) work for the development of the other allied agricultural sectors. The agriculture production information is very important for planning and allocation of resources to different sectors of agriculture.

Increasing agricultural production is one of the most important national requirements. Fulfilling this requirement would entail meticulous planning based on elaborate information on crop acreage, yield etc. Primarily agriculture based country like India must be aware of reliable, accurate, crop yield, crop growth conditions and timely information on types of crops grown with their acreages. Systematic collection of this information in the country has begun since 1884. The successful launching of the earth resource technology satellite named as **Landsat-1** on July 23, 1972 provided a new tool for gathering information on crops.

The Landsat program is the longest running enterprise for acquisition of satellite imagery of Earth. Followed by **Landsat-1** many satellites were launched by NASA with Landsat series in upcoming years and currently active satellite is **Landsat-8** on February 11, 2013. World-wide coverage provided by Landsat series of satellites has been important factor in the promotion of Remote Sensing Technology for crop inventory.

Landsat uses remote sensing technique for acquisition of information about an object or phenomenon without making physical contact with the object. This process involves making observations using sensors mounted on the satellites, which are at a considerable distance from the earth surface and recording the observations on images of photographic films and videotapes or digital data on magnetic tapes.

Satellite chronology

Instrument ↕	Picture ↕	Launched ↕	Terminated ↕	Duration ↕	Notes ↕
Landsat 1		July 23, 1972	January 6, 1978	2 years, 11 months and 15 days	Originally named Earth Resources Technology Satellite 1.
Landsat 2		January 22, 1975	February 25, 1982	2 years, 10 months and 17 days	Nearly identical copy of Landsat 1
Landsat 3		March 5, 1978	March 31, 1983	5 years and 26 days	Nearly identical copy of Landsat 1 and Landsat 2
Landsat 4		July 16, 1982	December 14, 1993	11 years, 4 months and 28 days	
Landsat 5		March 1, 1984	June 5, 2013	29 years, 3 months and 4 days	Nearly identical copy of Landsat 4. Longest Earth-observing satellite mission in history.
Landsat 6		October 5, 1993	October 5, 1993	0 days	Failed to reach orbit.
Landsat 7		April 15, 1999	Still active	15 years, 7 months and 28 days	Operating with scan line corrector disabled since May 2003.
Landsat 8		February 11, 2013	Still active	1 year, 10 months and 2 days	Originally named Landsat Data Continuity Mission from launch until May 30, 2013, when NASA operations were turned over to USGS.

Fig. 1: Satellite Chronology

When electromagnetic radiation falls upon a surface some of its energy is absorbed, some of its energy is transmitted through the surface and the rest is reflected. Surfaces also naturally emit radiation, mostly in the form of heat. Such reflected and emitted electromagnetic radiation is recorded either on the photographic film or digital sensor.

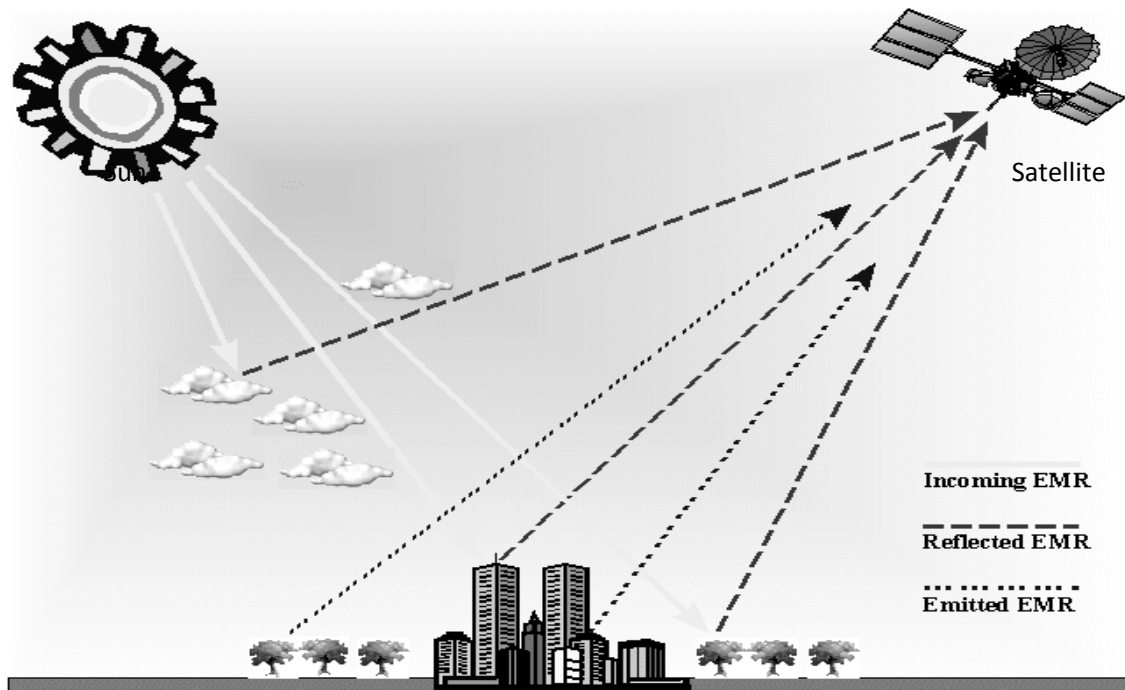


Fig. 2: Mechanism of remote sensing data acquisition.

Since the intensity and wavelengths of this radiation is a function of the surface in question, each surface is described as processing a characteristic “spectral signature”. If an instrument can identify and distinguish between different spectral signatures, then it will be possible to map the extent of surfaces using remote sensing. Remote sensing is widely used as a tool in many parts of the world for the management of the resources.

Crop identification and discrimination is based upon the fact that each crop has a unique spectral signature. Typical spectral reflectance of a crop shows absorption due to pigments in the visible region (0.4 to 0.7 m), high reflectance in the near infrared region because of internal cellular structure of the leaves and absorption at 1.45, 1.95 and 2.6 m spectral bands due to water content.

Spectral response of a crop canopy is influenced by:

- The Leaf – Area Index (LAI) and percentage ground cover
- Growth stages
- Differences in cultural practices
- Stress conditions and
- Canopy architecture.

Background soil and water are also important factors influencing the spectral response. Each crop has its own architecture and growing period thus enabling discrimination through Remote Sensing data.

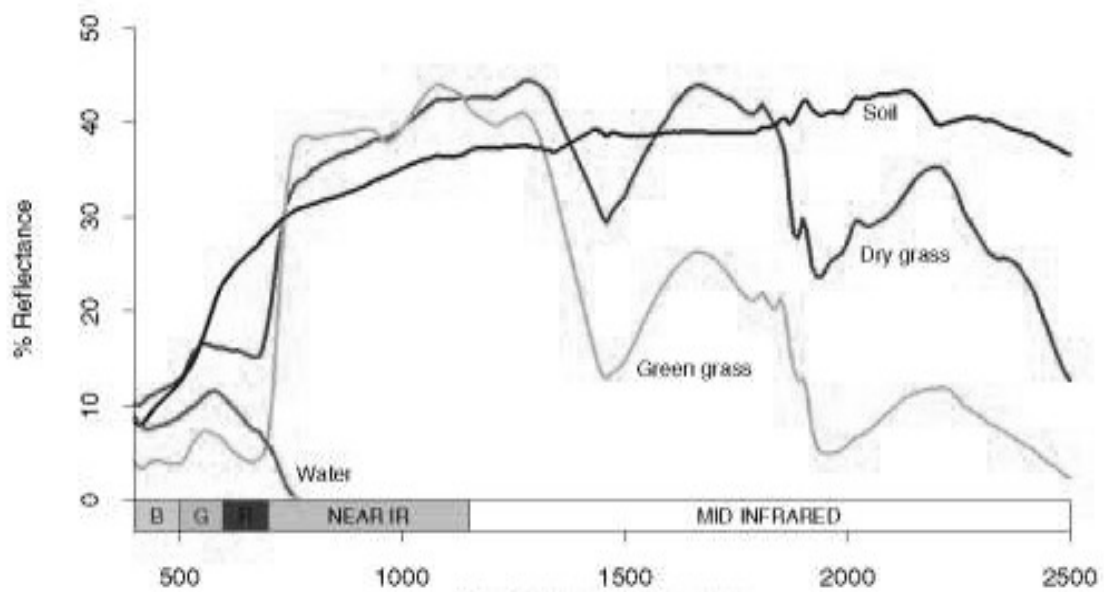


Fig. 3: Wave length Reflectance of water, soil and vegetation.

Source: Muhammad Aqeel Ashraf, Mohd. JamilMaah and Ismail Yusoff (2011). Introduction to Remote Sensing of Biomass, Biomass and Remote Sensing of Biomass, Dr. Islam Atazadeh (Edⁿ.), p: 262.

Satellite sensors record the intensity of electromagnetic radiation (sunlight) reflected from the earth at different wavelengths. Energy that is not reflected by an object is absorbed. Each object has its own unique 'spectrum' as shown in the diagram.

Table 1.1: Electromagnetic regions, their wavelengths and Salient features

Spectral regions	Wave length	Salient features
Gamma-ray region	< 0.03 nm	Incoming radiation completely absorbed by the upper atmosphere and not available for remote sensing.
X-ray region	0.03 - 30 nm	Completely absorbed by the atmosphere. Not available for remote sensing.
Ultraviolet region	0.3 - 0.4 μm	Incoming wavelengths of less than 0.3 μm will be completely absorbed by ozone layer in the upper atmosphere.
Photographic UV band	0.3 - 0.4 μm	Transmitted through the atmosphere. Detectable with film and photo detectors, but atmospheric scattering is severe.
Visible region	0.4 - 0.7 μm	Imaged with film and photo-detectors. Includes reflected energy peak of earth at 0.5 μm .
Infrared region	0.7 - 100 μm	Interaction with matter varies with the wavelength. Atmospheric transmission windows are separated by absorption bands.
Reflected IR band	0.7 - 3.0 μm	Reflected solar radiation that contains no information about thermal properties of materials. The interval from 0.7 - 0.9 μm is detectable with film and is called photographic IR band.
Thermal IR band	3 - 5 μm , 8 - 14 μm	Principal atmospheric windows in the thermal region. Images at these wavelengths are acquired by optical-mechanical scanners and special vidicon systems but not by film.
Microwave region	0.1 - 100 cm	Longer wavelengths that can penetrate clouds, fog and rain. Images may be acquired in the active or passive mode.
Radar	0.1 - 100 cm	Active form of microwave remote sensing. Radar images are acquired at various wavelength bands.
Radio	> 100 cm	Longest-wavelength portion of electromagnetic spectrum.

Remote sensing relies on the fact that particular features of the landscape such as bush, crop, salt-affected land and water reflect light differently in different wavelengths. Grass looks green because it reflects green light and absorbs other visible wavelengths. This can be seen as a peak in the green band in the reflectance spectrum for green grass above. The spectrum also shows that grass reflects even more strongly in the infrared part of the spectrum. While this can't be detected by the human eye, it can be detected by an infrared sensor. Instruments mounted on satellites detect and record the energy that has been reflected. The detectors are sensitive to particular ranges of wavelengths called 'bands'. The satellite systems are characterised by the bands at which they measure the reflected energy.

The Landsat satellite which provides the data has bands at the blue, green and red wavelengths in the visible part of the spectrum, three bands are near and mid infrared part of the spectrum and one band in the thermal infrared part of the spectrum. The satellite detectors measure the intensity of the reflected energy and record it.

Geographic information system (GIS) is a computer based information system used to digitally represent and analyse the geographic features present on the earth surface and the events that take place on it. The meaning to represent digitally is to convert analog into a digital form.

“Every object present on the earth can be geo-referenced” is the fundamental key of associating any database to GIS. Here, term ‘database’ is a collection of information about things and their relationship to each other and ‘geo-referencing’ refers to the location of a layer or coverage in space defined by the co-ordinate referencing system. Evolution of GIS has transformed and revolutionized the ways in which planners, engineers, managers and scientists conduct the database management and analysis. A GIS is a system of hardware, software, data, people, organizations and institutional arrangements of collecting, storing, analysing and disseminating information about areas of the earth. It is also defined as an information system designed to work with the data referenced by spatial / geographical co-ordinates. In other words GIS is both a database

system with specific capabilities for spatially referenced data as well as a set of operations for working with the data.

Remote Sensing is detecting and measuring of electromagnetic energy (usually photons) emanating from distant objects made of various materials, so that we can identify and categorize these object by class or type, substance and spatial distribution. Image Classification has the overall objective to automatically categorize all pixels in an image into classes or themes. The Spectral pattern or signature of surface materials belonging to a class or theme determines an assignment to a class.

Ground truth data collection has an important bearing on most of the Remote Sensing application. It involves the gathering of information on land cover types, their spatial extent, condition and geographical coordinates of the location. In crop related applications, ground truth data collection is necessarily done in synchronous or near-synchronous with the time of satellite data acquisition in order to maintain the correspondence between remote sensing and field observations.

Training site is an identified feature or area on an image established as a comparison standard for the automatic classification of features on the image. Training area is a small sample of homogeneous areas selected by the image analyst prior to classification. Each area is determined from maps, ground data, or other information.

Remote sensed data are usually digitally image data. Therefore data processing in remote sensing is treated as Digital Image Processing and it is concerned with below basic operations:

- **Image restoration:** It is concerned with the correction and calibration of images in order to achieve appropriate representation of the earth surface.
- **Image enhancement:** It is predominantly concerned with the modification of images to optimize their appearance to the visual system. Visual analysis is a key element in digital image processing and the effects of these techniques can be dramatic.
- **Image classification:** Computer-assisted interpretation of images thus it is an operation that is vital to GIS.

- **Image transformation:** Derivation of new imagery as a result of some mathematical treatment of the raw image bands.

As Image classification plays a vital role to GIS, we would be focusing more on this topic for the present study.

Image classification refers to computer-assisted interpretation of remotely sensed images. Although some procedures are able to incorporate information about such image characteristics as texture and context, the majority of image classification is based solely on the detection of the spectral signatures (spectral response patterns) of land cover classes. The success with which this can be done will depend on two things:

- The presence of distinctive signatures for the land cover classes of interest in the band set being used and
- The ability to reliably distinguish these signatures from other spectral response patterns that may be present.

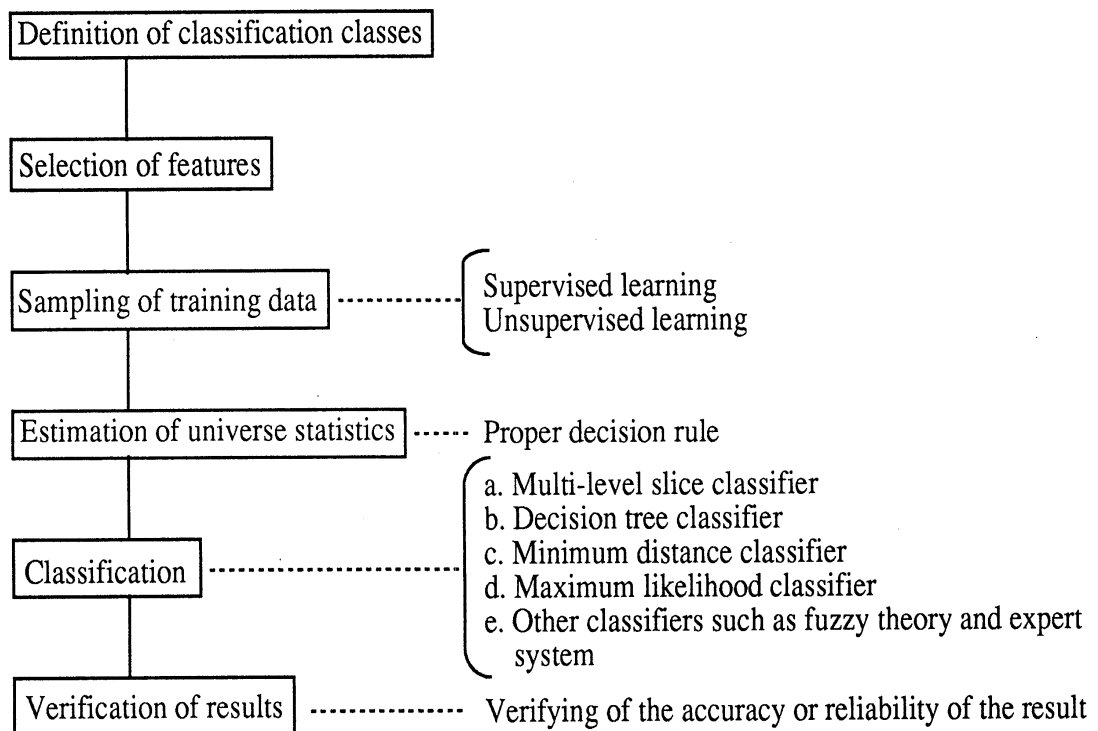


Fig. 4: Procedure for Image classification

An important step in digitally image processing is the process of land cover classification. In this process the image pixels are assigned to different land cover classes based on the spectral measurements of each pixel. There are two general approaches to image classification: supervised and unsupervised. They differ in how the classification is performed. In the case of supervised classification, the software system delineates specific land cover types based on statistical characterization data drawn from known examples in the image known as training sites. With unsupervised classification, clustering software is used to uncover the commonly occurring land cover types with the analyst providing interpretation of those cover types at a later stage. Other methods and algorithms also available such as decision tree approach, knowledge based classifier, fuzzy classification, neural networks etc., and as such there is no general method of classification. Different methods are found to be suitable in different corresponding situations.

Satellite Remote Sensing technology has played a vital role in the agriculture sector by providing relevant information for strategic decisions and development activities. Reliable and timely estimates of crop acreage and production provided a base for the formulation of marketing strategies and price stabilisation.

The high spatial resolution of LISS-III (Linear Imaging Self Scanner) data enabled the identification of crops which are grown under the multiple crop situations. Crop acreage estimation, crop condition assessment, crop yield forecasting, extent of crop damage due to pests and disease, identification of potential areas suitable for agriculture development etc., are some of the applications of remote sensing in the field of agriculture.

The scope of the present study is land use or land cover classification using digital image classification methods, their comparisons and accuracy assessment.

Objectives of the present study are

1. Acreage and production estimation of major crops of sedum Taluk of Gulbarga district.
2. To compare the estimates of remote sensing with the estimates of Department of Agriculture.
3. Comparison of classification methods and their accuracy assessment.



REVIEW OF LITERATURE

II REVIEW OF LITERATURE

The recent advances in the field of geographical techniques like Geographical Information System (GIS) and Remote Sensing increased the potential to change substantially the statistical approach to study the geographical realities. The remote sensing technology to obtain land utilization statistics gained popularity because of its extensive coverage of geographical area. Remote sensing has enabled mapping, studying, monitoring and management of various resources like agriculture, forestry, geology, water, ocean etc. It has further enabled monitoring of environment and thereby helping in conservation. Image classification and accuracy assessment of the classified images are important aspects of remote sensing image analysis. Several research works carried out using different methodologies and algorithms with their accuracy are reviewed in this section.

Sharma and Garg (1985) analysed Landsat Computer Compatible Tape (CCT) of March 8, 1977 on Multispectral Interactive Data Analysis System (MIDAS) for digital mapping of Land use classes at 1:25,000 scale in a portion of Northern Himalayan region (INDIA). Parallelepiped and minimum distance classification methods were employed under supervised classification approach. It was found that both the methods produce the Land use maps of nearly same accuracy. However, Parallelepiped method is relatively cheaper as it took nearly half of the computer time than minimum distance classification method. The overall Land use classification is made and mapping accuracy for Parallelepiped and minimum distance classification were also computed which was about 88 and 72 per cent respectively.

Roy and Das (1991) studied mapping of forest types in two districts, Karbi Angling and North Cachar of Assam using Landsat MSS digital data. Supervised maximum likelihood classification has been performed using training sets collected during field work. Overall classification accuracy of 96 per cent has been achieved at 95 per cent confidence interval.

Russell (1991) has reviewed the factors and techniques to be considered when using the accuracy of classification of remotely sensed data. Three classification approaches like supervised approach, unsupervised approach and modified approach were used. Error matrices for these three approaches were constructed for assessing the accuracy. The overall accuracy of 84, 78 and 88 per cent was achieved for the above three approaches respectively. Kappa coefficients were calculated for all the three approaches. Kappa coefficient for supervised approach, unsupervised approach and modified approach was found to be 0.77, 0.69 and 0.85 respectively.

Sugumaran *et al.* (1994) have made a study to delineate the waste lands at micro level in Matur Taluk of Kheda district. IRS LISS 11 digital data of 1991 has been used for the study which was analysed on VAX 11/780 image processing system. The digital data was classified following supervised classification algorithm. Three types of waste lands namely water logged, pasture/grazing land and salt affected lands could be identified and mapped. The overall accuracy estimated for all the different categories of waste lands was 87 per cent at 90 per cent confidence level.

Steffen Bock (1996) analysed a geocoded 7-channel Landsat TM image acquired on 16 April 1992 of the city of Brazil. Two classifiers were used: The neural network approach such as the ATL (Adoptive threshold learning) neural network model and maximum-likelihood classifier. The overall classification accuracy for two methods were obtained using error matrix. It was found that the neural network approach exceeds the overall classification accuracy achieved by the maximum-likelihood method by 14 per cent.

Sharma and Sarkar (1998) studied two methods of contextual classification techniques - one for low-resolution data and for high-resolution data. A new method was proposed by combining these two methods. The new method was compared with the Gaussian maximum-likelihood classification and the two methods of contextual classification for low-resolution and high-resolution data. Classification algorithms are compared using normalized classification accuracies and the Kappa coefficient. The

results showed that contextual classifiers produced significantly better results than the GML classification and the new method was found to be superior to other methods.

Krishna Prasad *et al.* (1999) have studied multi temporal data sets from coarse resolution sensors of IRS to classify various forest types using their phenological attributes reflected in temporal NDVI profiles using supervised classification with minimum distance classifier algorithm.

Sudhakar *et al.* (1999) used IRS-IB LISS data of Dec 1994 for land use/land cover classification with special emphasis on forest type mapping of Jaldapara, wild life sanctuary. Three classification techniques namely maximum likelihood, contextual and neural network classifications were applied. The classification accuracy showed that the neural network achieved maximum accuracy of 95 per cent, maximum likelihood algorithm with 91.06 per cent and contextual classifier with 87.42 per cent. It was concluded that neural network classifier worked better in homogeneous forest lands where as the maximum likelihood was best in both the conditions.

Nancy Thomas *et al.* (2003) have conducted a study on mapping and classification of the built-up contiguous portion of the City of Scottsdale South of the Central Arizona Project (CAP) canal. They made a study to assess the accuracy of three different methods for extracting land-cover/land-use information from high-resolution imagery of urban environments: combined supervised/ unsupervised spectral classification, raster-based spatial modelling and image segmentation classification using classification tree analysis. The overall accuracy of the fuzzy error matrix from the combined supervised/unsupervised spectral classification was only 58 per cent. The spatial modelling techniques used in this project achieved the highest overall accuracy of 81 per cent.

Dengsheng Lu *et al.* (2004) studied mapping and classification of The Brazilian Amazon basin. Four distinctly different classifiers were used to analyse multispectral data. They made a comparison of Minimum-Distance Classifier (MDC), Maximum-Likelihood Classifier (MLC), Extraction and Classification of Homogeneous Objects (ECHO) and

Decision-Tree Classifier based on Linear Spectral Mixture Analysis (DTC-LSMA). Each of the classifiers used both Landsat Thematic Mapper data and identical field-based training sample datasets in a western Brazilian Amazon study area. Seven land-cover classes such as mature forest, advanced secondary succession, initial secondary succession, pasture lands, agricultural lands, bare lands, and water were classified. Classification results indicate that the DTC-LSMA and ECHO classifiers were more accurate than were the MDC and MLC. The overall accuracy of the DTCLSMA approach was 86 per cent with a 0.82 Kappa coefficient and ECHO had an accuracy of 83 per cent with a 0.79 Kappa coefficient. The accuracy of the other classifiers ranged from 77 to 80 per cent with Kappa coefficients from 0.72 to 0.75.

Dwivedi *et al.* (2004) have studied to evaluate the potential of the Gaussian maximum likelihood classifier, Mahalanobis minimum distance classifier, minimum distance classifier, and artificial neural network classifier in deriving information on land use/land cover over part of Ethiopia using various band combinations of Landsat TM data. The values of Kappa coefficient were used to compare the performance of the classifiers, two at a time, by means of Z-statistics. The results demonstrated the superiority of artificial neural network classifier over per-pixel Gaussian maximum likelihood classifier, Mahalanobis minimum distance classifier in deriving information on the land use/land cover. The study also revealed data set with first three principal components significantly improved the classification accuracy.

Elnazir Ramadan *et al.* (2004) have studied the use of remote sensing technology to investigate the quantitative estimates of urban growth of Shaoxing city in china. Multi date Landsat thematic mapper images for 1984, 1997 and 2000 were obtained. Image rectification was performed and supervised maximum likelihood algorithm was used to classify the images. Confusion matrix was generated and an overall accuracy of 92, 84 and 88 per cent were obtained for 2000, 1997 and 1984 respectively.

Kelley *et al.* (2004) studied multiple semi-automated methods for discriminating Land Use/Land Cover (LULC) classification in the absence of field data. The study site comprises a heterogeneous mixture of physiographic landscape elements in a remote area

of eastern Mexico. Classification of modified Tasseled Cap transformations as well as a hybrid unsupervised-supervised classification technique was employed using a Landsat 7ETM+ scene; the panchromatic band was reserved for accuracy assessment. Using a random stratified sampling scheme, the overall accuracies ranged from 61.55 to 83.3 per cent and Khat accuracies varied from 56 to 81.1 per cent. The classifications derived from modified Tasseled Cap transformations did not appropriately discriminate landscape components possibly owing to landscape heterogeneity. Classification results were improved through the use of landscape stratification in the attribution phase in the hybrid classification procedure recommended for complex environments.

Donald and Robert (2005) developed land cover classification for 5 research watersheds in Garland and Saline counties in Arkansas using 2002 LANDSAT 7 Enhanced Thematic Mapper Plus (ETM+) satellite imagery. The supervised classification was based upon 146 training areas identified from reference data and then applied to the imagery using the maximum likelihood classification algorithm. The unsupervised classification used an Iterative Self- Organizing Data Analysis Techniques (ISODATA) algorithm to classify the imagery into 300 spectral classes which then were identified from reference data. Data from 171 field locations were used to assess the accuracy of the final classifications using an error matrix. The supervised classification had an overall accuracy of 74.85 per cent compared to 40.94 per cent for the unsupervised classification.

Shamsudheen *et al.* (2005) have studied land use/land cover mapping for Kumta Taluk of Uttar Kannada district of Karnataka. The IRS ID LISS 111 image was used. Signatures for training sets were collected by ground truth to perform supervised maximum likelihood classification. The accuracy of the classification was assessed using stratified sampling method. An overall accuracy of 75 per cent was obtained.

Sharma and Bren (2005) studied the supervised classification using maximum algorithm on three dates of IRS LISS 111 satellite data to know the effect of seasonal spectral variation on land cover classification for the study area falling in the Solan district of Himachal Pradesh. It was found that the summer dataset was better with

overall classification accuracy of 76 per cent as compared to winter and spring dataset with classification accuracy of 49 and 46 per cent respectively.

Fei Yuan *et al.* (2005) developed a methodology to map and monitor land cover change using multi temporal Landsat Thematic Mapper (TM) data in the seven-county Twin Cities Metropolitan Area of Minnesota for 1986, 1991, 1998 and 2002. Error matrix was constructed; the overall seven-class classification accuracies averaged 94 per cent for the four years. The overall accuracy of land cover change maps generated from post-classification change detection methods and evaluated using several approaches ranged from 80 to 90 per cent. The maps showed that between 1986 and 2002 the amount of urban or developed land increased from 23.7 to 32.8 per cent of the total area, while rural cover types of agriculture, forest and wetland decreased from 69.6 to 60.5 per cent. The results quantify the land cover change patterns in the metropolitan area and demonstrate the potential of multi temporal Landsat data to provide an accurate economical means to map and analyze changes in land cover over time that can be used as inputs to land management and policy decisions.

Yagoub and kolan (2006) attempted to evaluate and quantify Abu Dhabi coastal zone land use /land cover changes from 1972 to 2000 using multi temporal Landsat satellite data and digital change detection techniques. Supervised classification was performed by minimum distance statistical processing in ENVI software. Accuracy assessment was done by generating confusion matrix and accuracy of 88 per cent was obtained.

Al-Ahmadi and Hames (2008) applied four different classification techniques unsupervised (ISODATA) and supervised (Maximum likelihood, Mahalanobis Distance and Minimum Distance) in three sub-catchments in Saudi Arabia for the classification of the raw TM5 images. The developed maps are then visually compared with each other and accuracy assessments utilizing ground-truths are undertaken. Error matrices produced to evaluate the classification methods show that the best overall classification accuracy method was the Maximum likelihood for all the three sub-catchments; with an average accuracy of about 80 per cent. The second best overall classification accuracy method

was Mahalanobis distance; with an average accuracy of 74 per cent and the worst overall classification accuracy method was minimum distance with an average accuracy of 67 per cent.

Mohd Hasmadi and Kamaruzaman (2008) made a study to assess classification accuracy of classified forest map on Landsat TM data from difference number of reference data (200 and 388 reference data). This comparison was made through observation (200 reference data), interpretation and observation approaches (388 reference data). Five land cover classes namely primary forest, logged over forest, water bodies, bare land and agricultural crop/mixed horticultural has been identified by the differences in spectral wavelength. Result showed that an overall accuracy from 200 reference data was 83.5 per cent (Kappa value 0.7502; Kappa variance 0.0028), which was considered acceptable or good for optical data. However, when 200 reference data was increased to 388 in the confusion matrix, the accuracy slightly improved from 83.5 to 89.17 per cent, with Kappa coefficient increased from 0.75 to 0.80, respectively. The accuracy in the classification suggested that this strategy for the selection of training area, interpretation approaches and number of reference data used were importance to perform better classification result.

Krishna Bahadur (2009) made a study to explore and evaluate the use of modified band and ancillary data in Landsat and IRS image classification and to produce a land use land cover map of the Galaudu watershed of Nepal. Classification of land uses were explored using supervised and unsupervised classification for 12 feature sets containing the Landsat MSS, TM and IRS original bands, ratios, normalized difference vegetation index, principal components and a digital elevation model. Overall the supervised classification method produced higher accuracy than the unsupervised approach. The result from the combination of bands ration 4/3, 5/4 and 5/7 ranked the highest in terms of accuracy (82.86 %) while the combination of bands 2, 3 and 4 ranked the lowest (45.29 %).

Mohd Hasmadi *et al.* (2009) studied mapping and classification of the Ayer Hitam Forest Reserve area. It was found that the unsupervised classification was more

“noisy” compared to the supervised classification. However the problem has been overcome by conducting a majority filter where isolated pixels were grouped into the closest value of spectral or digital number. SPOT 5 satellite image data has been used and is capable to map and classify the Ayer Hitam Forest Reserve and its surrounding area with five classes namely vegetation, urban, water body, grassland and barren land for both classification techniques. The overall accuracy for supervised classification was higher than the unsupervised classification where the accuracy for supervised was 90.28 per cent while the unsupervised produced 80.56 per cent. Thus, the supervised classification appeared more accurate than the unsupervised classification.

Dengsheng Lu *et al.* (2010) made an approach to examine the suitable methods for land-use/land-cover classification in a urban landscape, Lucas do Rio Verde in Mato Grosso State, Brazil using quick bird image. Traditional per-pixel spectral-based supervised classification like Maximum Likelihood Classifier, incorporation of textural images, multispectral image like ECHO, spectral-spatial classifier and segmentation-based classification were used. The accuracy assessment was conducted for each classification result. Comparing the MLC-based classification result the ECHO improved overall land-cover classification accuracy by 6 per cent. Comparing the classification accuracies between using segmentation-based mean-spectral images and using per-pixel spectral-based multispectral images indicated that segmentation-based method significantly improved classification performance for all land-covers. The results show that use of spatial information during the image classification procedure either through the integrated use of textural and spectral images or through the use of segmentation-based classification method can significantly improve land-cover classification performance.

Zhengrong Li *et al.* (2010) studied on the empirical comparison of seven machine learning algorithms in texture classification with application to vegetation management in power line corridors. Aiming at classifying tree species in power line corridors object-based method was employed. Individual tree crowns were segmented as the basic classification units and three classic texture features were extracted as the input to the classification algorithms. Several widely used performance metrics like overall accuracy,

precision/recall, F-measure, ROC analysis and computational cost were used to evaluate the classification algorithms. The experimental results demonstrated that the classification performance depends on the performance matrix, the characteristics of datasets and the feature used.

Offer and Arnon (2011) analysed a Landsat-5 TM image of the northern Negev and classified in several ways using ERDAS IMAGINE 2010. Several established methods for land-use classification from RS data like ISODATA unsupervised classification, Maximum likelihood supervised classification were compared. A new approach of hybrid classification was used by combining signatures from both supervised and unsupervised training data which gave significantly more accurate results than each approach separately. In addition, ancillary land-use data were used to update and improve the RS classification accuracy within a GIS framework. It was also found that updating this product using ancillary data and GIS techniques can improve the product accuracy by up to 10 per cent. The final product's overall accuracy was 81 per cent. The McNemer's test confirmed that the accuracy improvement by the hybrid classification was statistically significant in comparison to the supervised classification, but not significant in comparison to the unsupervised classification.

Renuka Devi and Santhosh Baboo (2011) worked on a new method for classifying the areas in a remotely sensed image under the category of supervised classification techniques. This classification technique describes how to classify the geographical areas in given image under supervised classification techniques conventions. The new method RGB&L algorithm was performed and its performance was analysed with other classification algorithm. The usual supervised classifications like Parallelepiped Classification, Minimum Distance to Means Classification and Maximum Likelihood Classification were performed. The overall accuracy of classification was computed from the error matrix. The new method gave 83 per cent of overall accuracy than the other methods.

Asmala and Shaun (2012) made detail analyses of Maximum likelihood classification for tropical land covers in Malaysia recorded from Landsat 5 TM satellite.

ML classified the classes that exist in the study area with a good agreement with the reference map. ML classified the study area into 11 classes with accuracy 97 per cent ($\kappa=0.97$). The classification was carefully examined using visual analysis, classification accuracy, band correlation and decision boundary. The results indicated that the separation between mean of the classes in the decision space was the main factor that lead to the high classification accuracy of ML.

Manisha *et al.* (2012) conducted a study on mapping and classification of Pune Municipal Corporation area. Both maximum likelihood classifier and minimum distance classifier were used. Per-pixel maximum likelihood classification of Landsat-TM data offered the most satisfactory results. From the results, it has been confirmed that supervised classification using maximum likelihood method is most accurate as compared to minimum distance method. Maximum likelihood supervised classification method yielded higher accuracy of 0.91 per cent whereas accuracy of minimum distance method was found to be 0.87 per cent.



MATERIAL AND METHODS

III MATERIALS AND METHODS

The methodology followed to achieve the objectives indicated in chapter I is discussed in this chapter.

An important part of remote sensing is land cover classification. The main objective of image classification is to automatically categorize all pixels in a digital image into land cover classes or theme. Image classification can be defined as the process of reducing an image to information classes. The categorization of image pixels is based on their digital numbers/grey values in one or more spectral bands. Information classes are the categories of interest to the data users. For example land use/land cover classes like residential, industrial, agricultural, forest, water bodies etc., and each land cover class may have its own characteristic spectral response in different spectral bands of the electromagnetic spectrum. These spectral classes are the distinct groups with similar digital numbers (Spectral response curves). Similar features have similar spectral responses. The spectral response of a certain feature is unique when compared to other features of interest. By quantifying the spectral response of a known feature, we can use this information to find all occurrences (Instances) of that feature.

The methods of classification techniques adopted in the present study are Maximum likelihood classification, Minimum distance classification and Mahalanobis distance classification.

3.1 Material

3.1.1 Description of the study area

3.1.2 Details of image data

3.1.3 Details of land cover classes considered

3.1.1 Description of the study area

Gulbarga also known as Kalaburgi is a district of Karnataka state in India. It is totally 623 km from the north of Bengaluru which is state capital of Karnataka.

Gulbarga is situated in Deccan Plateau located at 17.33°N 76.83°E and the general elevation ranges from 300 to 750 mts above mean sea level. Two main rivers, Krishna and Bhima, flow in this district. Black soil is predominant type and the Upper Krishna Project is major irrigation venture in this district. Bajra, tur, sugarcane, groundnut, sunflower, sesamum, castor bean, black gram, jowar, wheat, cotton, ragi, Bengal gram, and linseed are very commonly grown in this district.

The weather prevailing in Gulbarga consists of three main seasons. The summer which spans from late February to mid-June. It is followed by the south west monsoon which spans from the late June to late September heavy rainfall may go up to 750 mm. It is then followed by dry winter weather until mid-January.

Sedum Taluk of Gulbarga district of Karnataka is considered as the study area for the present study which lies at 17°10'59" N 77°16'59" E (Plate 1).

3.1.2 Details of image data

IRS (Indian Remote Sensing Satellite)-P6 LISS-III (Linear Imaging Self Scanner) imageries of 22nd December 2011 is undertaken for the present study. The geometrically corrected imageries are obtained from National Remote Sensing Agency, Department of Space, Government of India, Hyderabad. The topographical map of the study area is laid on this image to extract the digital image of the study area (Plate 2). The spatial resolution of the images is 23.5 m and the imageries were recorded in three spectral bands.

Among these only the first two namely green and red are in small range of electromagnetic spectrum. Further the third one i.e. blue band is useful in identification of green vegetation based on crops.

Ground truth data collected during field visits in the selected study area and the toposheets are used to accomplish the task of selection of training sites for each category for training the classifier in supervised classification. However, a part of the data was used as test sites for assessing classification with accuracy.

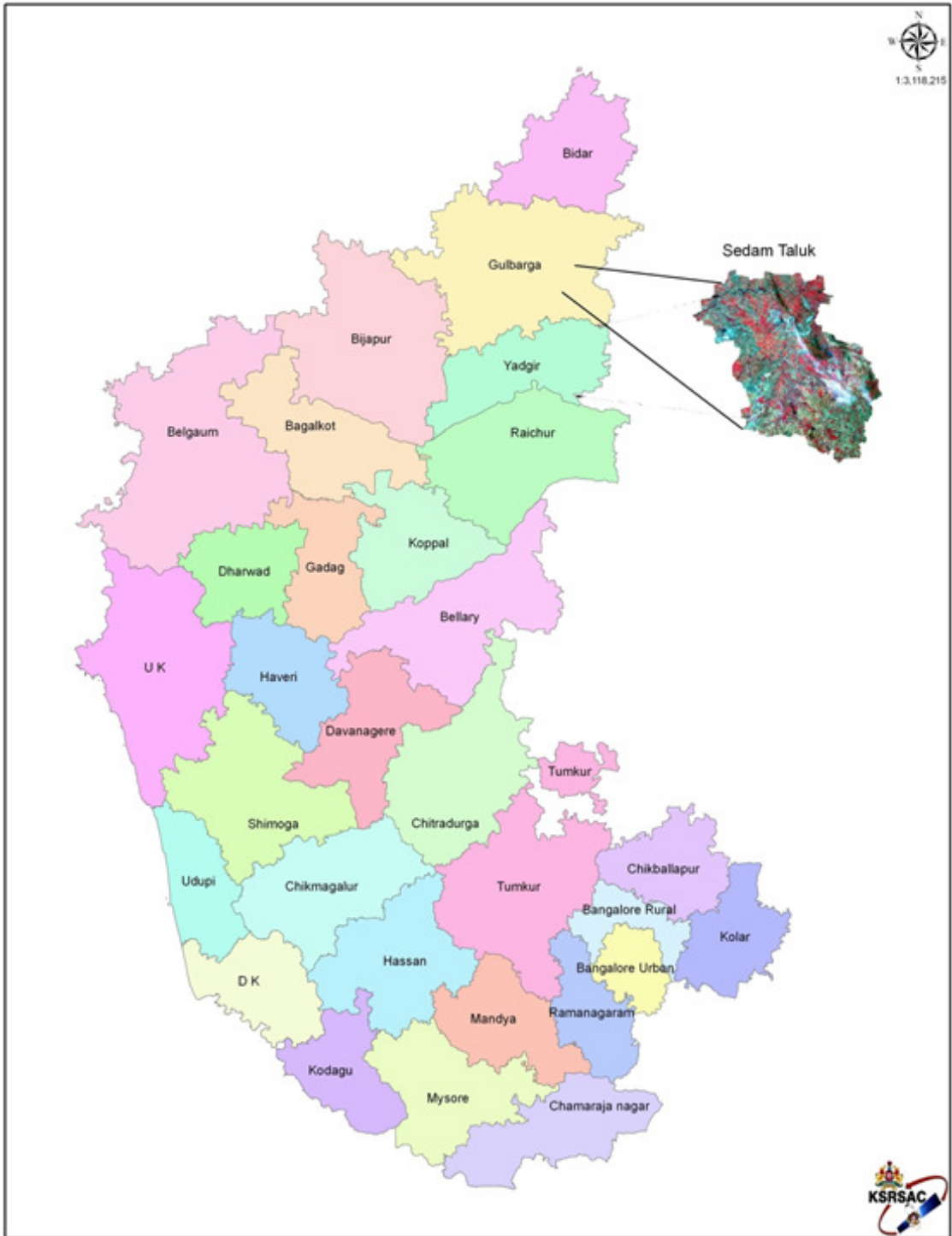


Plate 1: Location map of study area in Karnataka state

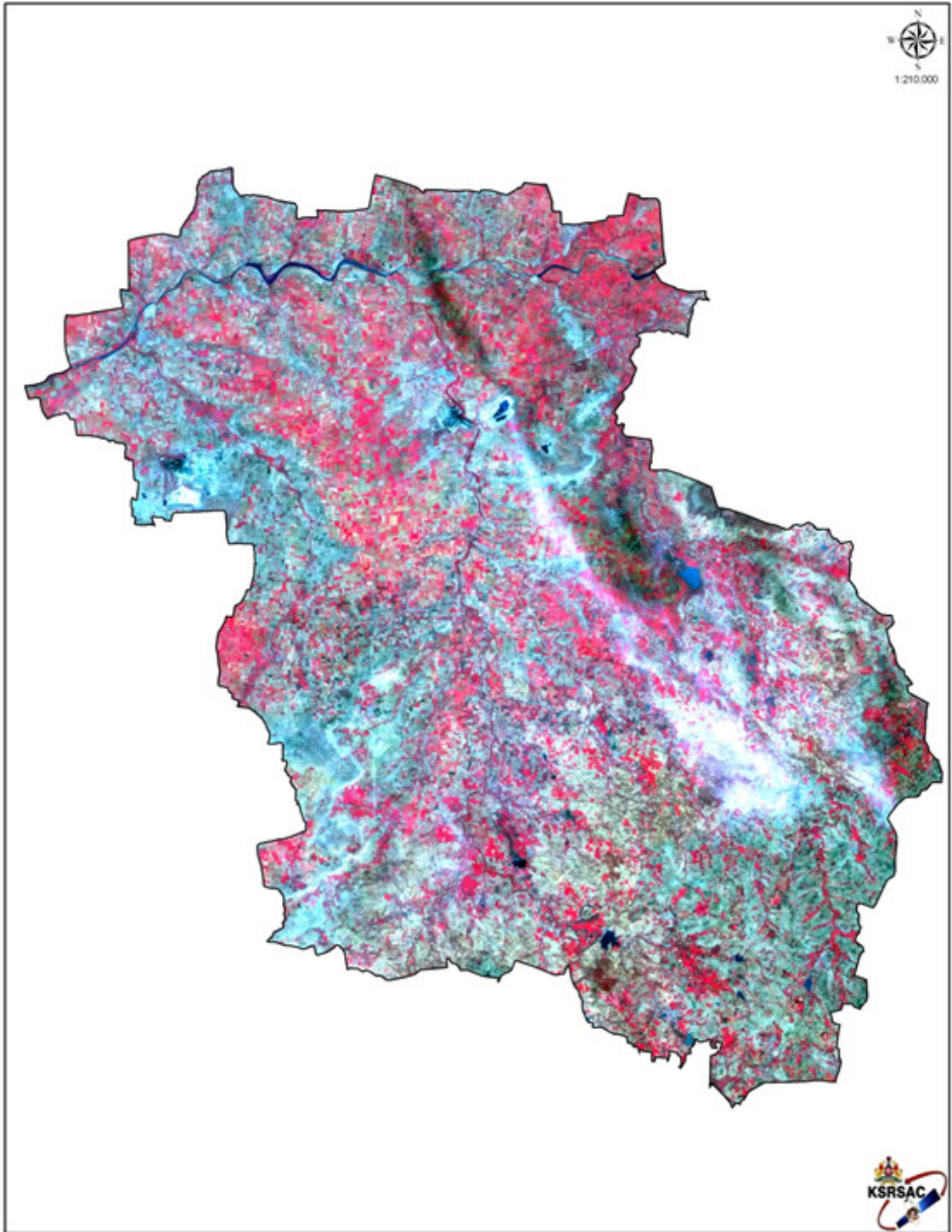


Plate 2: IRS-P6 LISS III Satellite image of Sedam Taluk, Gulbarga District dated 22nd Dec 2011

3.1.3 Details of land cover classes considered

The categories of interest were carefully selected and defined to successfully and effectively perform digital image classification. In the present study land use/land cover classification system is adopted with six different categories for the study area as follows.

Land use/land cover classification of Sedum Taluk of Gulbarga district.

- Jowar
- Red gram
- Other crops (Paddy, Bajra, Maize, Wheat etc.),
- Fallow land and
- Water bodies.

3.2 METHODS OF IMAGE CLASSIFICATION

Image classification is an effective process of classifying a remotely sensed image into pixel classes. A common example is land cover classes. It is the process of dividing the image into different regions with some similarity and labelling the regions using supplementary ground truth information. In the present study the technique of supervised classification is used for image classification. Classification is done using ERDAS (Earth Resource Data Analysis System) imagine 9.1 software at the Karnataka State Remote Sensing Application Centre, Department of IT and BT, M. S. Building, Government of Karnataka, Dr.Ambedkar veeidhi, Bengaluru.

Supervised Classification

In supervised classification, selection of training sites is based on known a priori through a combination of fieldwork, map analysis and personal experience. The spectral characteristics of these sites are used to train the classification algorithm for eventual land-cover mapping of the remainder of the image. Supervised classification can be defined normally as the process of samples of known identity to classify pixels of unknown identity. Samples of known identity are those pixels located within training areas. Pixels located within these areas called the training samples/units used to monitor

the classification algorithm in assigning specific spectral values to appropriate informational class.

Typically, supervised classification comprising of three step process:

- 1) Delineate representative areas called training fields for each cover class.
- 2) Use training polygons to develop descriptive statistics for each cover class.
- 3) Use the representative statistics for each cover class to predict the cover class of every pixel in the image.

Various supervised classification algorithms may be used to assign an unknown pixel to one of 'm' possible classes. The choice of a particular classifier or decision rule mainly depends on the nature of the input data and the expected desired output. Parametric classification algorithms assume that the observed measurement vectors X_c obtained for each class in each spectral band during the training phase of the supervised classification are Gaussian; which is normally distributed in nature. Nonparametric classification algorithms make no such assumption in its application.

The following are the important steps in supervised classification of remotely sensed data.

3.2.1 Choosing an appropriate classification algorithm.

3.2.2 Training site selection and statistic extraction.

3.2.1 Choosing an appropriate classification algorithm

The image classification process involves translating the pixel values in a satellite image into useful and meaningful categories. In the case of land cover classification these categories comprise different types of land cover defined by the classification scheme which is implemented. A classification algorithm is used to assign individual pixels or groups of pixels to one of the specified and valid categories. The primary difference between classification algorithms is to determine how an individual pixel is assigned to a land cover category.

The difference between the different types of supervised statistical classification algorithms is how they determine similarity between pixels.

In the present study three algorithms were used for supervised classification namely

- a) Maximum likelihood algorithm,
- b) Mahalanobis distance algorithm, and
- c) Minimum distance to means algorithm.

A digital image in 'p' spectral bands is represented by p-matrices of order M x N, where 'M' represents the number of rows and 'N' represents the number of columns in the image. At each pixel location (i, j) there is a p-dimensional measurement vector, $X_{ij} = (X_{ij1}, X_{ij2}, X_{ij3}, \dots, X_{ijp})$ representing the digital values in the p-spectral bands. In an 8 bit image, the possible values of X_{ijk} vary from 0 to 255. It is assumed that there 'm' land cover classes represented by w_1, w_2, \dots, w_m in the underlying image and the purpose of classification is to assign each pixel to its class based on the observations.

3.2.1.1 Maximum Likelihood Classification Algorithm

The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes and that the input bands consider following normal distributions. The probabilities of a pixel belonging to each of a predefined set of 'm' classes were calculated and the pixel is then assigned to the class for which the probability is the highest.

In this algorithm each pixel is modelled to have multivariate normal distribution. Let $\mu_1, \mu_2, \dots, \mu_m$ and $\Sigma_1, \Sigma_2, \dots, \Sigma_m$ denotes the population mean vectors and population variance matrices for 'm' classes respectively.

The observation vector X_r at pixel 'r' when it belongs to class 'c' is distributed as a multivariate normal distribution with mean μ_c and covariance matrix Σ_c .

$$\text{Then, } P_{rc} = \left[\frac{1}{2\pi} \right]^{p/2} * |\Sigma|^{-1/2} * \exp \left\{ \frac{-1}{2} (X_r - \mu_c)' \Sigma_c^{-1} (X_r - \mu_c) \right\}$$

Gives the likelihood of pixel 'r' belonging to class 'c'

Further taking logarithm, the expression converts to

$$\ln P_{rc} = \frac{p}{2} \ln \left[\frac{1}{2\pi} \right] - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} (X_r - \mu_c)' \Sigma_c^{-1} (X_r - \mu_c)$$

Ignoring $\frac{p}{2} \ln \left[\frac{1}{2\pi} \right]$ which is a constant,

The maximum likelihood algorithm assigns pixel r to class c if and only if

$$\ln P_{rc} \geq \ln P_{rq}, \text{ For all } q=1, 2, \dots, m \text{ classes, } q \neq c.$$

Since the class mean vectors μ_c and covariance matrix Σ_c are unknown, the sample estimates are obtained from the training set.

Let $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ be the sample mean vectors and V_1, V_2, \dots, V_m , be the sample variance-covariance matrices estimated from the training data for 'm' classes respectively.

The pixel assignment is made based on the estimated value of p_{rc}

$$\ln \hat{p}_{rc} = \left[-0.5 \det(V_c) \right] - \left[0.5 (X_r - \bar{X}_c)' (V_c)^{-1} (X_r - \bar{X}_c) \right]$$

The pixel is assigned to that class for which it has the highest similarity of being a member. The decision rule is given as assign pixel 'r' to the class 'c' if and only if,

$$\ln \hat{P}_{rc} \geq \ln \hat{P}_{rq},$$

For all $q = 1, 2, \dots, m$ classes, $q \neq c$.

The classification of the complete image is performed on a pixel by pixel basis.

Every pixel is assigned to one of the mutually exclusive classes based on the likelihood as described above and no pixel remains unclassified.

3.2.1.2 Mahalanobis Distance Classification Algorithm

In Mahalanobis distance algorithm each pixel follows multivariate normal distribution. Let $\mu_1, \mu_2, \dots, \mu_m$ and $\Sigma_1, \Sigma_2, \dots, \Sigma_m$ denote the population mean and population variance covariance matrices for m classes respectively.

The observation vector X_r at pixel 'r' when it belongs to class 'c' a multivariate normal distribution with mean μ_c and covariance matrix Σ_c .

Then,

$$D_{rc}^2 = \sqrt{(X_r - \mu_c)' \Sigma_c^{-1} (X_r - \mu_c)}$$

Gives the Mahalanobis distance of pixel 'r' belonging to class 'c'.

Since the class mean vectors μ_c covariance matrix Σ_c are unknown, the sample estimates are obtained from the training set.

Let $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ be the sample mean vectors and V_1, V_2, \dots, V_m be the sample variance covariance matrices estimated from the training data for m classes respectively.

The classification is performed for the whole data set on pixel by pixel basis.

The pixel assignment is made based on the value of d_{rc}

$$D_{rc}^2 = \sqrt{(X_r - \bar{X}_c)' V_c^{-1} (X_r - \bar{X}_c)}$$

Where X_r is the observation vector for an unclassified pixel 'r'.

3.2.1.3 Minimum Distance to means Classification Algorithm

In this technique the spectral response vectors of each class are modelled to have mean vectors. The mean vectors of the classes are estimated from training sets of each class. The Euclidean distance of a pixel from the mean vector of each class is computed and the pixel is assigned to a class for which this distance is shortest.

Symbolically let w_1, w_2, \dots, w_m indicate the m land cover classes in the image with unknown mean vectors $\mu_1, \mu_2, \dots, \mu_m$. let $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ represent the sample mean vectors of the m classes estimated from the training set. $\bar{X}_c = (\bar{x}_{c1}, \bar{x}_{c2}, \dots, \bar{x}_{cp})^T$ Where mean is calculated over all pixels in the training set of class c , for $c = 1, 2, \dots, m$ classes and $k = 1, 2, \dots, s$ denotes the digital value of r^{th} pixel.

$$X = (X_{r1}, X_{r2}, \dots, X_{rk}, \dots, X_{rp})^T.$$

Let D_{rc} denote the Euclidean distance between pixel 'r' and the class 'c'.

$$\text{then, } d_{rc} = \sqrt{\sum_{k=1}^p (x_{rk} - \bar{x}_{rk})^2}, \text{ for all } c = 1, 2, \dots, m$$

The minimum distance to means classifier assigns pixels 'r' to the class 'c' if

$$d_{rc} \leq d_{rq}, \text{ for all } q = 1, 2, \dots, m \text{ classes, } q \neq c.$$

3.2.2 Training Site Selection and Statistics Extraction

Set of homogeneous pixels selected from the imageries whose identity is completely known on the basis of ground truth are called training sites. Representative training sites are selected for each of the land cover classes. Each site is composed of many pixels. The general rule is that if training data are being extracted from 'p' bands, the minimum number of pixels in a class should be $(p+1)$. This condition will allow the inverse of covariance matrix for each class to be calculated.

Let n_c denote the number of pixels in the training site of class 'c' and X_{ic} denote the observation vector corresponding to pixel i in class 'c', then,

$$X_{ic} = \begin{bmatrix} x_{ic1} \\ x_{ic2} \\ \cdot \\ \cdot \\ x_{icp} \end{bmatrix}$$

The digital values for each pixel in each band in training site is analyzed statistically to yield mean measurement vector, \bar{X}_c

$$\bar{X}_c = \begin{bmatrix} \bar{x}_{c1} \\ \bar{x}_{c2} \\ \cdot \\ \cdot \\ \bar{x}_{cp} \end{bmatrix}$$

Where, $\bar{x}_c = \frac{\sum_{i=1}^{n_c} x_{ick}}{n_c}$, for $c = 1, 2, \dots, p$ spectral bands. The measurement

vector is analyzed to yield the covariance matrix for each class V_c , where V_{ck1} is the covariance between the bands 'k' and '1' in class 'c'.

$$V_c = \begin{bmatrix} v_{c11} & v_{c12} & \dots & v_{c1p} \\ v_{c21} & v_{c22} & \dots & v_{c2p} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ v_{cp1} & v_{cp2} & \dots & v_{cpp} \end{bmatrix}$$

$$\text{Where, } V_{ck1} = \frac{\sum_{i=1}^{n_c} (x_{ick} - \bar{x}_{ck})(x_{icl} - \bar{x}_{cl})}{n_c - 1}$$

For $k = 1, 2, \dots, p$ and $l = 1, 2, \dots, p$ spectral bands.

3.3 Classification Accuracy Assessment

Accuracy assessment is an important part of any classification. Accuracy assessment tries to quantify how good a classifier is. The accuracy of a classification is usually assessed by comparing the classification with some reference data that is believed to accurately reflect the true land-cover. Sources of reference data include among other things like ground truth, higher resolution satellite images and maps derived from aerial photo interpretation. Classification accuracy is estimated using testing data i.e., the spatial data consisting of pixels for which the correct classification is known but not used in generating training statistics.

3.3.1 Error Matrix

The accuracy assessment reflects the difference between the classification and the reference data. Consequently, if the reference data is highly inaccurate, assessment might indicate that the classification is poor, while it really is a good classification. It is better to get fewer, but more accurate reference data. Quantitative methods to assess classification accuracy involve an error matrix or confusion matrix built from two data sets, classification and reference data. Comparison between the classification obtained by the method under consideration and the accurate classification using test data is made, a count of number of pixels correctly classified and misclassified are recorded for each class in an error matrix.

An error matrix is a square array of numbers in rows and columns which express the number of a sample units (pixels) assigned to a particular category relative to the actual category as verified by test data set.

In the Table 3.1, n_{ij} is the number of pixels classified as category 'i' which are found to belong to category 'j' in the reference data, N is the test sample size (pixels) used, m is the number of classes, n_{ij} and n_j are marginal totals of corresponding rows and columns. The main diagonal element represents the correct classification and off diagonal elements represents the misclassifications.

Table 3.1: Error Matrix built using two data sets, classification and reference data

Classification categories	Reference categories							
	1	2	j	M	Total
1	n ₁₁	n ₁₂	n _{1j}	n _{1m}	n _{1.}
2	n ₂₁	n ₂₂	n _{2j}	n _{2m}	n _{2.}
3
.
.
I	n _{i1}	n _{im}	n _{i.}
.
M	n _{m1}	n _{m2}	n _{mj}	n _{mm}	n _{m.}
Total	n _{.1}	n _{.2}	n _{.j}	n _{.m}	N

Classification accuracy is assessed using overall accuracy and Kappa coefficient. Overall accuracy is relatively simple and intuitive measure of agreement. It uses only the main diagonal elements of the error matrix. It does not take into account the proportion of agreement between data sets that is due to chance alone and as such it tends to overestimate classification accuracy.

Overall accuracy is obtained as the ratio of total correct classifications for all the classes to the total test sample units. The user and producer accuracy are two widely used measures of class accuracy. The producer's accuracy refers to the probability that a certain land-cover of an area on the ground is classified as such, while the user's accuracy refers to the probability that a pixel labelled as a certain land-cover class in the map is really this class. The user and producer accuracy for any given class typically are not the same. Producer's accuracy is the ratio of correct classifications to the actual data. User's accuracy is the ratio of the correct classification to the total number of pixels classified in specific classes.

$$\text{Overall accuracy} = \frac{\sum_{i=1}^m n_{ii}}{N}$$

$$\text{User's accuracy} = \frac{n_{ii}}{n_i}$$

$$\text{Producer's accuracy} = \frac{n_{ii}}{n_i}$$

Where, n_{ii} = correctly classified pixels.

n_i = Marginal totals of classification categories.

n_i = Marginal totals of reference categories.

N = Total number of pixels.

3.3.2 Mapping accuracy (MA)

The mapping accuracy of a particular class i , M_i should take into account the commission of wrong observations and omission of correct observation. Thus M_i is defined as:

$$M_i = n_{ii} / (n_{i.} + n_{.i} - n_{ii})$$

The overall mapping accuracy, MA is defined as the weighted average of M_i 's where weights are taken as proportion of observations in the class.

$$MA = (\sum n_i M_i) / N$$

3.3.3 F-Measures

It is worth noting that the overall accuracy does not distinguish between types of errors the classifier makes. For example, two classifiers may obtain the same accuracy but they may behave quite differently on each category. If one classifier obtains 100 per cent accuracy on one category but only 41 per cent on the other category, while another classifier generate 70 per cent for each category, it is hard to claim that the first classifier is better. Therefore, overall accuracy may not be used blindly as the evaluation method for classifiers on a dataset. Precision and Recall can avoid the problem encountered by Accuracy. Precision can be seen as a measure of exactness or fidelity, whereas Recall is a measure of completeness.

Their definitions are:

$$\text{Precision} = TP / (TP+FP)$$

$$\text{Recall} = TP / P$$

Where, TP = True Positive count

FP = False Positive count

P = True positive + False negative.

Usually, Precision and Recall scores are discussed jointly and a single measure can be derived by combing both measures called ‘F’ Measure.

F-measure is the weighted harmonic mean of precision and recall. We use the ‘F’ measure in which the precision and Recall are evenly weighted.

$$F = 2 \left[\frac{\text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \right]$$

Thus, higher the value of ‘F’ measure, the better classifier’s performance.

Table 3.2: Error matrix for Precision and Recall

Classification categories	Reference categories								Recall
	1	2	j	...	M	Total	
1	n ₁₁	n ₁₂	n _{1j}	...	n _{1m}	n _{1.}	n ₁₁ /n _{1.}
2	n ₂₁	n ₂₂	n _{2j}	...	n _{2m}	n _{2.}	n ₂₂ /n _{2.}
3
.
.
i	n _{i1}	n _{im}	n _{i.}	n _{im} /n _{i.}
.
M	n _{m1}	n _{m2}	n _{mj}	...	n _{mm}	n _{m.}	n _{m1} /n _{m.}
Total	n _{.1}	n _{.2}	n _{.j}	...	n _{.m}	N	
Precision	n ₁₁ /n _{.1}	n ₂₂ /n _{.2}	n _{mj} /n _{.j}	...	n _{mm} /n _{.m}		

3.3.3.1 Overall F-measure:

The F-measure for each class over the entire data is based on the cluster that best describes each class. The overall F-measure of a classification is calculated by using the weighted sum of individual F-measure for all the classes.

$$\text{Overall F-measure} = \sum F_i / N,$$

Where, F_i = Individual F-measure.

N = Number of land use classes.

3.3.4 Kappa coefficients (K)

The Kappa index of agreement for categorical data was developed by Cohen and associates in the context of psychology and psychiatric diagnosis. Kappa was subsequently adopted by the remote sensing community as a useful measure of classification accuracy. It attempts to control the chance agreement by incorporating all marginal distributions of the error matrix. The basic idea behind Kappa is that some of the apparent classification accuracy given by naive measures could be due to chance. Kappa statistics gives the proportion of agreement after the chance agreement is removed and it may be calculated as,

$$\hat{K} = \frac{P_0 - P_c}{1 - P_c}$$

$$\text{Where, } P_0 = \frac{\sum_{i=1}^m n_{ii}}{N}, \quad P_c = \frac{\sum_{i=j=1}^m n_{i.} \cdot n_{.j}}{N}$$

An approximation of variance of Kappa statistics is given by

$$\hat{\sigma}_k^2 = \frac{P_0(1 - P_c)}{N(1 - P_c)}$$

Using the estimated Kappa and its variance, we can compute a confidence interval for Kappa. We can also compare two classifications to see if one is significantly better than the other.

3.3.5 Test of significance of K

The estimate of Kappa coefficient gives a measure that indicates that the confusion matrix is significantly different from a random result. The percentage agreement estimate (P_0) follows Binomial distribution. Where as P_c is the percentage of misclassification. When the sample size is large (>100) it can be approximated by normal distribution. Hence the distribution of the Kappa coefficient is also normal.

Test of significance can be performed to determine whether the Kappa coefficient is significantly different from zero or not by using Z test and is given by,

$$H_0 : K = 0$$

$$H_1 : K \neq 0$$

Test statistics,

$$Z = \frac{\hat{K}}{\sqrt{\hat{\sigma}_k^2}}$$

Follows normal distribution with zero mean and unit variance.

Decision: if observed value of $|Z| > 1.96$ null hypothesis is rejected at 0.05 level of significance indicating that the confusion matrix is significantly different from a random result.

3.3.6 Test of significance difference between two Kappa coefficients

The test of significance between two Kappa coefficients can be used to determine whether the corresponding classification accuracies differ significantly or not. This is done by testing null hypothesis (H_0): $K_1 = K_2$ against alternate hypothesis (H_1): $K_1 \neq K_2$.

Where K_1 and K_2 are the Kappa coefficients obtained from two different classification methods.

Since the statistic \hat{K} is to follow a normal distribution the difference $(\hat{K}_1 - \hat{K}_2)$ also follows normal distribution. Accordingly the test statistic,

$$Z = \frac{\hat{K}_1 - \hat{K}_2}{\sqrt{\hat{\sigma}_{K1}^2 + \hat{\sigma}_{K2}^2}}$$

follows a standard normal distribution.

The hypothesis H_0 is rejected if $|Z| > 1.96$, leading to the conclusion that the corresponding classification methods differ significantly with respect to their classification accuracy.

The classified images were used to estimate the area under different land cover classes. This is done by aggregation.

$$\text{Area under class } c = (\text{No. of pixels in class } c) \times (\text{Area of a pixel})$$

3.4 Validation of schemes of classifiers

In the present study, validation of classifiers is achieved through the non-parametric tests namely Friedman two way analysis of variance and chi square for goodness of fit.

3.4.1 Friedman two way analysis of variance

In this method the data arranged in two way table having 'N' rows and 'M' columns. The rows represent the various subjects and the columns represent the various conditions. If the scores of subjects serving under all the conditions are under study, then each row gives the scores of one subject under the K conditions.

The data of the test are ranked and the scores in each row are ranked separately. The Friedman determines whether it is likely that the different columns of ranks came from the same population.

In the present study, we have taken three classifiers as column features and six land use patterns (classes) as row features. We have taken the absolute deviation of area estimates of three classifiers from the observed or ground truth value, and then given ranks for deviated value within the each class.

Test of significance can be performed to determine whether classifiers are differing significantly using chi-square test.

$$H_0 : M_1 = M_2 = M_3$$

$$H_1 : M_1 \neq M_2 \neq M_3$$

Test statistic,

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum_{j=1}^k (R_j)^2 - 3N(k+1)$$

Where, N= number of rows

k = number of columns

R_j = sum of ranks in jth column

Where, χ_r^2 is distributed approximately as chi-square with (k-1) df.

Note: if k or N exceeds table values then one can make large sample approximation

χ^2 with (k-1) df.

We may determine the probability of occurrence under H₀ of χ_r^2 by referring to table N which gives the exact probability associated with obtained values.

Decision: If observed value of $\chi_r^2 >$ table of χ^2 with (k-1) df, the null hypothesis is rejected at 0.05 or 0.01 level of significance.

Further, based on the magnitude of ranks we can compare the particular classifier with the ground truth, i.e., higher the magnitude, classifier is more nearer to the ground truth.



RESULTS

IV RESULTS

The most common uses of satellite images are mapping land cover via image classification and detection of land cover change. In the present study land use/land cover classifications for the study area is performed by several methods. Assessment of classification accuracy is performed for all the classified images. The results of the classification by different methods and their accuracy assessment are presented in this chapter.

The IRS-P6 imageries of the study area were used for the study. The basic sequence of steps for supervised classification as mentioned in chapter III under section 3.2 is followed. Once the categories of interest are defined, sample of homogeneous pixels are selected as training sites for each category by drawing polygons on the false colour composite imageries. These training sites are used to produce statistical descriptors for each one and cover category. These include mean values and variance-covariance matrix associated with each class. The statistics obtained for the training site of each class for the study area is presented in the tables.

In order to achieve the objectives specified in chapter I classification of test image of Sedam Taluk, Gulbarga district obtained by remote sensing through the satellite IRS-P6 LISS-III on 22nd December, 2011 are done. The results obtained by applying the methods are described in the chapter III are discussed below.

4.1 Maximum likelihood classification

The mean vectors and variance covariance matrices are estimated from training site under each class is presented in the Tables 4.2 to 4.11 for the study area. Each pixel is classified into one of the land cover classes as described in chapter III under section 3.2.1.1. The resulting classified images are presented in Plate-3. Sets of color coding are given to distinguish the corresponding land cover categories. The error matrix for the data sets of Maximum likelihood method under the study is presented in Table 4.16.

4.2 Mahalanobis distance classification

The mean vectors and variance–covariance matrices are estimated from the training set of each class are presented in the Tables 4.2 to 4.11 for the study area. Each pixel is classified into one of the land cover classes as described in chapter III under section 3.2.1.2. The resulting classified images are presented in Plate-4. Different colour coding is given to distinguish the different land cover categories. The error matrix for two data sets of Mahalanobis distance method under the study is presented in Table 4.17.

4.3 Minimum distance to means classification

In this technique the different land cover classes are characterized by the mean vectors of each class. These class means are estimated from the training sets of each class and are presented in Table 4.2 to 4.11. Each pixel is then classified into one of the corresponding classes as described in the third chapter under section 3.2.1.3. The resulting classified image is presented in Plate-5. The error matrix for two data sets of Minimum distance method is presented in Table 4.18.

Classification accuracy is estimated using testing data i.e., the data consisting of pixels for which the correct classification is known. The comparison between the classification obtained by the method under consideration and the correct classification is made and count of correctly classified and misclassified is recorded for each class in an error matrix. The error matrices for all three methods are presented in the tables.

The area for the land cover classes estimated using different classification methods are presented in Table 4.12 and the production estimated using different classification methods for the land cover classes Jowar, Red gram and other crops are presented in the Table 4.13. The Area and Production estimates collected from DOA (Directorate of Agriculture, Govt of Karnataka) are also presented in the Table 4.12 and 4.13 respectively. The comparison of the estimates collected from Directorate of Agriculture and remote sensing are made.

Table 4.1: Class means estimated from training site

Class	Red	Blue	Green	Infrared
Jowar	91.40	76.44	218.44	160.12
Redgram	84.67	57.25	256.83	148.58
Other Crops	84.00	68.00	177.00	146.44
Fallow	101.84	106.89	169.26	204.84
Water body	100.25	81.00	83.03	107.63

Table 4.2: Descriptive statistics of training site for class Jowar

Band	Min	Max	Mean	SD
Red	89.00	93.00	91.40	1.16
Blue	74.00	80.00	76.44	1.64
Green	211.00	229.00	218.44	6.36
Infrared	156.00	164.00	160.12	2.09

Table 4.3: Descriptive statistics of training site for class Red gram

Band	Min	Max	Mean	SD
Red	83.00	85.00	84.70	0.70
Blue	56.00	58.00	57.30	0.60
Green	247.00	261.00	256.80	3.70
Infrared	148.00	150.00	148.60	0.80

Table 4.4: Descriptive statistics of training site for class other crops

Band	Min	Max	Mean	SD
Red	83.00	85.00	84.00	0.87
Blue	66.00	70.00	68.00	1.50
Green	174.00	184.00	177.00	3.16
Infrared	145.00	147.00	146.44	0.73

Table 4.5: Descriptive statistics of training site for class Fallow

Band	Min	Max	Mean	SD
Red	83.00	85.00	84.70	0.70
Blue	56.00	58.00	57.30	0.60
Green	247.00	261.00	256.80	3.70
Infrared	148.00	150.00	148.60	0.80

Table 4.6: Descriptive statistics of training site for class Water body

Band	Min	Max	Mean	SD
Red	98.00	104.00	100.20	1.40
Blue	79.00	85.00	81.00	1.50
Green	81.00	88.00	83.00	1.90
Infrared	106.00	110.00	107.60	1.00

Table 4.7: Variance-covariance matrix estimated from training site for land cover class Jowar

Band	Red	Blue	Green	Infrared
Red	1.30	0.90	3.00	1.20
Blue	0.90	2.70	-1.90	1.50
Green	3.00	-1.90	40.50	5.70
Infrared	1.20	1.50	5.70	4.40

Table 4.8: Variance-covariance matrix estimated from training site for land cover class Red gram

Band	Red	Blue	Green	Infrared
Red	0.40	0.10	1.90	0.20
Blue	0.10	0.40	1.00	0.30
Green	1.90	1.00	13.60	1.60
Infrared	0.20	0.30	1.60	0.60

Table 4.9: Variance-covariance matrix estimated from training site for land cover class Other crops

Band	Red	Blue	Green	Infrared
Red	0.80	-0.50	0.40	0.10
Blue	-0.50	2.30	-0.80	-0.40
Green	0.40	-0.80	10.00	0.00
Infrared	0.10	-0.40	0.00	0.50

Table 4.10: Variance-covariance matrix estimated from training site for land cover class Fallow

Band	Red	Blue	Green	Infrared
Red	0.40	0.10	1.90	0.20
Blue	0.10	0.40	1.00	0.30
Green	1.90	1.00	13.60	1.60
Infrared	2.00	0.30	1.60	0.60

Table 4.11: Variance-covariance matrix estimated from training site for land cover class Water body

Band	Red	Blue	Green	Infrared
Red	1.90	1.50	2.20	0.70
Blue	1.50	2.20	2.20	0.80
Green	2.20	2.20	3.70	1.00
Infrared	0.70	0.80	1.00	1.00

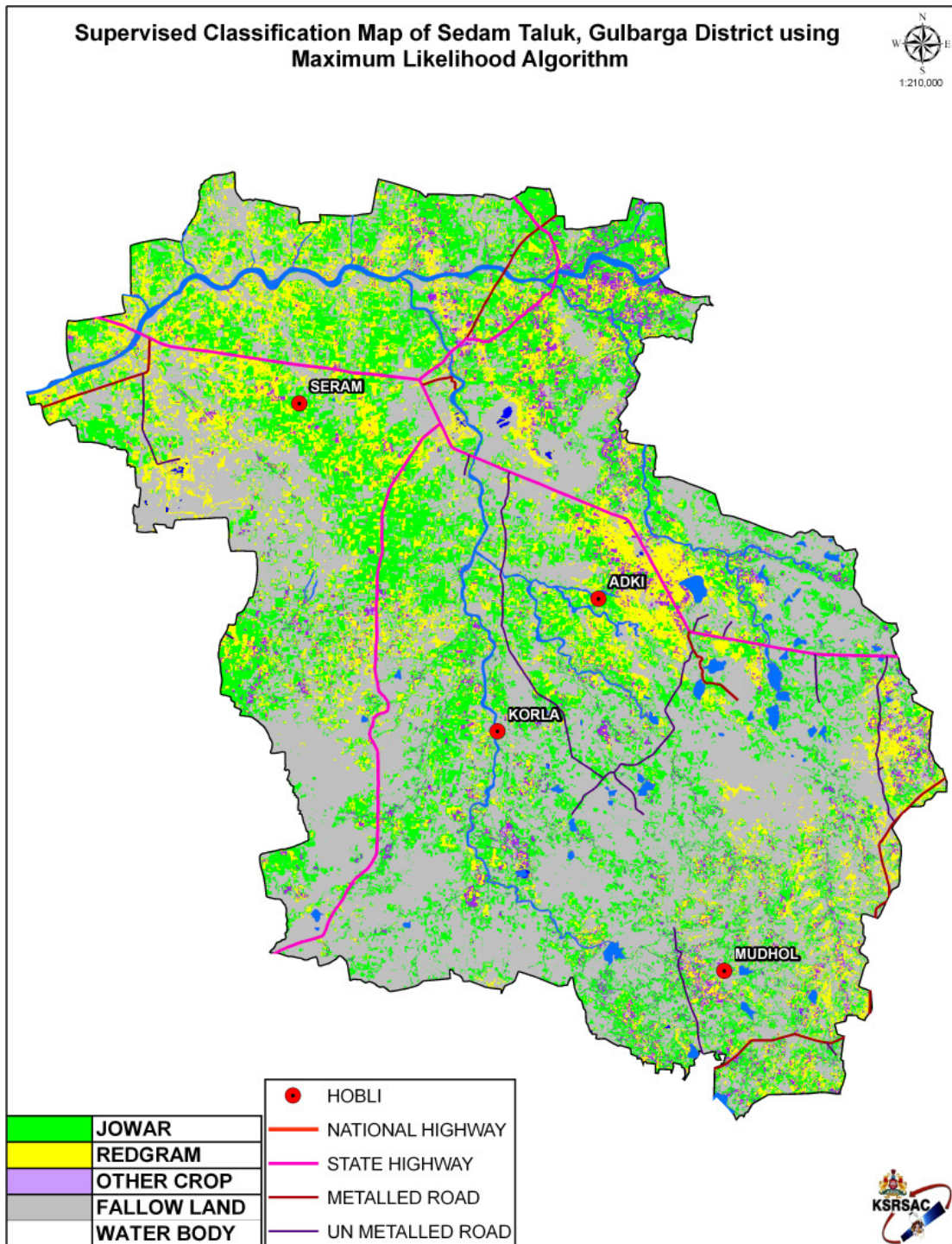


Plate 3: Supervised classification map of Sedam Taluk, Gulbarga District using Maximum likelihood algorithm

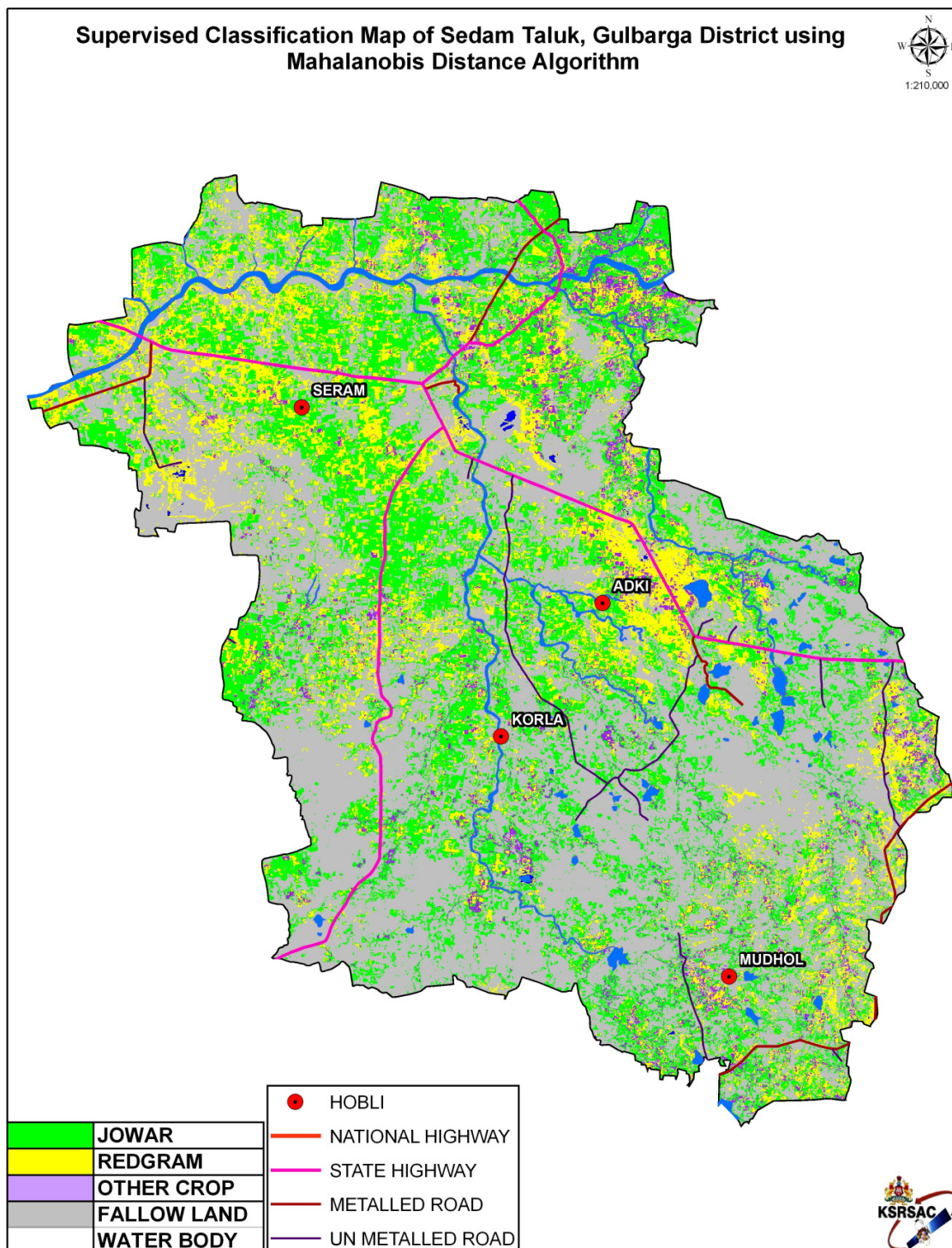


Plate 4: Supervised classification map of Sedam Taluk, Gulbarga District using Mahalanobis distance algorithm

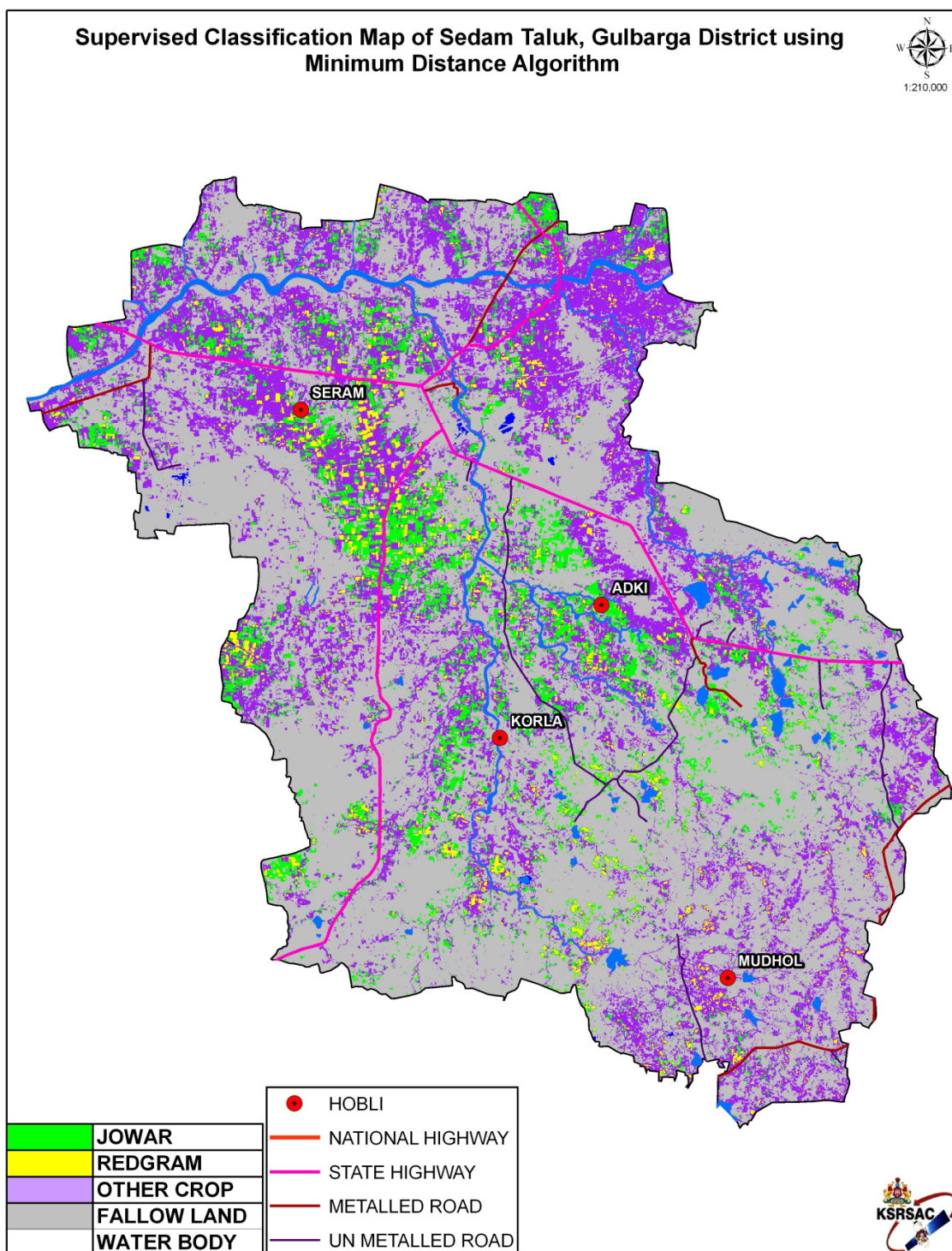


Plate 5: Supervised classification map of Sedam Taluk, Gulbarga district using Minimum distance algorithm

Table 4.12: Area estimates estimated using different classification methods and actual area estimates as per DOA

Class	Estimates of Area (ha)			
	Maximum likelihood	Mahalanobis distance	Minimum distance	Actual estimates of Area (ha)
Jowar	22,131	21,939	2,1743	22,232
Red gram	46,272	45,967	45,642	46,370
Other crops	7,914	7,339	7,593.00	8,443
Fallow	1,05,054	1,06,132	1,06,401	1,04,318
Water body	515	509	507	523
Total	1,81,886	1,81,886	1,81,886	1,81,886

DOA: Department of Agriculture

The estimates of area under each land cover class obtained from different classification methods for the study area presented in the above table indicate that the estimates obtained using Maximum likelihood method is more nearer to the actual estimates compared to other two methods for all the classes.

Table 4.13: Production estimates using different classification methods and Actual production estimates for the study area

Class	Production (Lakh kg) estimates			
	Maximum likelihood	Mahalanobis distance	Minimum distance	Actual estimates (DOA)
Jowar	229.27	227.29	225.26	259.46
Red gram	221.64	220.18	218.62	150.85
Other Crops	118.56	109.95	118.56	126.48

The estimates of production under each land cover class obtained from different classification methods for the study area are presented in above table reveals that the production estimates for class jowar and other crops obtained under maximum likelihood method are more nearer to the actual estimates. Whereas, the estimates obtained under minimum distance method for class red gram is nearer to the actual estimate.

Comparison of estimates of remote sensing methods and DOA estimates using absolute deviation values

Table 4.14: Absolute deviation values of area estimates obtained by classification methods from actual estimates of DOA

Classes	Maximum Likelihood	R _j	Mahalanobis Distance	R _j	Minimum Distance	R _j
Jowar	101	3	293	2	489	1
Red gram	98	3	403	2	728	1
Other crops	529	3	1104	1	850	2
Fallow	736	3	1814	2	2083	1
Water body	8	3	14	2	16	1
Total (R_j)		15		9		6

R_j: Ranks given based on the magnitude of deviation value.

The area under all the classes estimated using maximum likelihood method are found to be more close to the actual estimate of the classes.

Table 4.15: Absolute deviation values of production estimates obtained by classification methods from actual estimates of DOA

Classes	Maximum Likelihood	R _j	Mahalanobis Distance	R _j	Minimum Distance	R _j
Jowar	30.19	3	32.17	2	34.2	1
Red gram	70.79	1	69.33	2	67.77	3
Other crops	7.91	3	16.53	1	7.92	2
Total (R_j)		7		5		6

R_j: Ranks given based on the magnitude of deviation value.

The production estimates under the class jowar and other crops estimated using Maximum likelihood method are found to be more nearer to the actual estimates. Whereas, the production estimates under class red gram obtained using minimum distance method is nearer to the actual estimate of red gram.

4.4 Classification accuracy analysis

Accuracy assessment of classification is an essential component of the image classification process. The error matrix is accepted as a standard descriptive tool for classification accuracy assessment. Testing samples of 210 pixels were used for accuracy assessment. Error matrix is generated for classified imageries using the reference data.

The error matrices for different classification methods considered in the study are presented in Tables 4.16 to 4.18 and 4.19 gives the comparison of accuracy of individual classes obtained from different methods. Test of significance is performed for Kappa coefficients of each method as described in section 3.4. The results are presented in Table 4.20. The pair wise comparisons between Kappa coefficients for different methods are also done and presented in Table 4.21.

The overall accuracy and the mapping accuracy are also calculated as described in chapter III using error matrix for all three methods and the results are presented in Table 4.22.

Table 4.16: Classification results obtained with Maximum Likelihood classification

Categories	Reference Categories					
	1	2	3	4	5	Total
Jowar	38	0	1	2	0	41
Red gram	0	27	1	1	0	29
Other crop	1	1	47	0	1	50
Fallow land	1	2	1	56	1	61
Water body	0	0	0	1	28	29
Total	40	30	50	60	30	210

Reference categories are same as classification categories.

Table 4.17: Classification results obtained with Mahalanobis Distance classification

Categories	Reference Categories					
	1	2	3	4	5	Total
Jowar	35	2	3	1	0	41
Redgram	1	24	1	2	1	29
Other crop	2	1	45	2	2	52
Fallow land	2	3	1	54	1	61
Water body	0	0	0	1	26	27
Total	40	30	50	60	29	210

Reference categories are same as classification categories.

Table 4.18: Classification results obtained with Minimum Distance classification

Categories	Reference Categories					
	1	2	3	4	5	Total
Jowar	33	3	4	3	0	43
Redgram	1	21	3	4	1	30
Other crop	2	1	42	4	2	51
Fallow land	4	5	1	47	1	58
Water body	0	0	0	2	26	28
Total	40	30	50	60	30	210

Reference categories are same as classification categories.

The results of Table 4.19 depicts that the user's and producer's accuracy for all the classes under maximum likelihood method is found to be highest compared to other two methods. Whereas, it is less in case of minimum distance method for all the classes.

Table 4.19: User's and Producer's accuracy of three classification methods

Class	Accuracy (%)					
	Maximum Likelihood		Mahalanobis Distance		Minimum Distance	
	User	Producer	User	Producer	User	Producer
Jowar	92.68	95.00	85.37	87.50	76.74	82.5
Red gram	93.10	90.00	82.76	80.00	70.00	70.00
Other crops	94.00	94.00	86.54	90.00	82.35	84.00
Fallow land	91.80	93.33	88.52	90.00	81.03	78.33
Water body	96.55	93.33	96.30	89.66	92.85	86.66

4.5 Kappa coefficients

Table 4.20: Test of significance of Kappa coefficients for different classification methods

Classification method	Kappa coefficient (K)	Variance of Kappa (K)	Z-Value	Table value
Maximum likelihood	0.9149	0.0004	41.6543	>2.58
Mahalanobis distance	0.8419	0.0008	29.0197	
Minimum distance	0.7512	0.0012	21.554	

**Significant at 1% level.

Test of significance for each classification is performed using the Kappa coefficient. All three classifiers are found to be significant at 1 per cent level.

Table 4.21: The pair wise comparison between Kappa coefficients

Difference between two Kappa coefficients		Z-value
k1-k2	0.0729	2.0054*
k1-k3	0.1637	3.9736**
k2-k3	0.0907	2.0006*

* Significant at 5 % level, **Significant at 1 % level.

The pairwise comparison is made between the classifiers using the Kappa coefficients calculated for each classification method to determine whether the classification accuracies differ significantly or not.

All the three classification methods differ significantly with respect to their classification accuracy at 5 per cent level of significance ($p > 0.05$). Whereas, only the maximum likelihood method and minimum distance method differ significantly from each other at 1 per cent level of significance ($p < 0.01$).

4.6 Overall and Mapping Accuracy

Table 4.22: The overall accuracy and mapping accuracy of three methods

Classification method	Overall accuracy (%)	Mapping accuracy (%)	Difference
Maximum likelihood	93.38	87.51	5.87
Mahalanobis distance	87.66	78.42	9.24
Minimum distance	80.44	67.61	12.83

The overall accuracy and mapping accuracy were calculated for all three classification methods using an error matrix. Maximum likelihood method is found to have highest overall (93.38%) and mapping accuracy (87.51%) among all the methods. There is a marginal gap between the overall and mapping accuracy under the maximum likelihood method indicating superior classification compared to other methods. Whereas, there is a larger difference under minimum distance method.

4.7 F-measures

Recall, precision and individual F-measures used to obtain overall F-measure is calculated as described in section 3.3.3 of chapter III and results for each classification is presented in Tables 4.23, 4.24 and 4.25. The overall F-measure for all three classification methods are presented in the Table 4.26.

Maximum likelihood classification

Table 4.23: Recall, Precision and F-measures for classes under Maximum likelihood classification

Classes	Recall	Precision	Individual F-measure
Jowar	0.92	0.95	0.93
Red gram	0.93	0.90	0.91
Other crop	0.94	0.94	0.94
Fallow land	0.91	0.93	0.92
Water body	0.96	0.93	0.94
Total	4.66	4.65	4.64

The individual F-measure calculated using Recall and Precision under maximum likelihood classification for each class are presented in the above table indicates, that the classes other crop and water body is found to have a highest F-measure of 0.94. Whereas, the class red gram has a least F-measure of 0.91 compared to other classes.

Mahalanobis distance classification

Table 4.24: Recall, Precision and F-measures for classes under Mahalanobis distance classification

Classes	Recall	Precession	Individual F-measure
Jowar	0.85	0.87	0.86
Red gram	0.83	0.80	0.81
Other crop	0.87	0.90	0.88
Fallow land	0.89	0.90	0.89
Water body	0.96	0.90	0.93
Total	4.40	4.37	4.37

The individual F-measure calculated using Recall and Precision under Mahalanobis distance classification for each class are presented in the above table indicates that the class water body has highest F-measure of 0.93. Whereas, the class red gram has a least F-measure of 0.81 compared to other classes.

Minimum distance classification

Table 4.25: Recall, Precision and F-measures for classes under Minimum distance classification

Classes	Recall	Precision	Individual F-measure
Jowar	0.76	0.82	0.79
Red gram	0.70	0.70	0.70
Other crop	0.83	0.84	0.83
Fallow land	0.81	0.78	0.80
Water body	0.93	0.87	0.89
Total	4.03	4.01	4.01

The individual F-measure calculated using Recall and Precision under minimum distance classification for each class are presented in the above table indicates that the class water body has highest F-measure of 0.89 whereas, the class red gram has a least F-measure of 0.70 compared to other classes.

The overall F-measure was calculated as described in the chapter III of section 3.3.3.1.

Table 4.26: The overall F-measure for all three classifications

Classification method	Total F-measure
Maximum likelihood	0.928
Mahalanobis distance	0.874
Minimum distance	0.802

The overall F-measure calculated for all the three methods. The maximum likelihood method (0.928) has the highest F-measure among all the methods followed by Mahalanobis distance method (0.874) and minimum distance method (0.802) is found to have lesser F-measure. Higher the value of F-measure, better the classifier's performance. It can be concluded that Maximum likelihood method's performance is better when compared to other two methods.

Validation of schemes of classifiers

4.8 Friedman two way analysis of variance

In the present study we have taken the three classifiers as column features and five land use classes as row features. In this technique, we have taken the absolute deviation between four classifiers from the observed or ground truth value and ranked based on the magnitude of deviated value within the each class for area estimates and using Friedman two analysis of variance the significance is tested which are presented in Table 4.27.

Table 4.27: Deviation values of area estimates of classifiers from the actual estimates of DOA and test of significance of Friedman two way analysis of Variance using area estimates

Classes	Maximum Likelihood	R _j	Mahalanobis Distance	R _j	Minimum Distance	R _j	χ_r^2
Jowar	101	3	293	2	489	1	8.40*
Red gram	98	3	403	2	728	1	
Other crops	529	3	1104	1	850	2	
Fallow	736	3	1814	2	2083	1	
Water body	8	3	14	2	16	1	
Total R_j		15		9		6	

* Significance at 5 per cent level.

R_j: Ranks given based on the magnitude of deviation value.

It is found that the coefficient of Friedman two way analysis of variance (χ_r^2) is significant at five per cent level which is presented in the Table 4.27 indicating that all the three classifiers differ significantly and they have not emerged from the same population. Based on the magnitude of rank of the classifier we are comparing the classifier which is nearer to the ground truth .i.e. higher the magnitude of the classifier more nearer to the ground truth.



DISCUSSION

V DISCUSSION

The results presented in chapter 4 are discussed in this section. The land cover classes considered were jowar, red gram, other crops, fallow and water body. The area and production statistics for land cover classes were estimated using different classification methods such as Maximum likelihood, Mahalanobis distance and Minimum distance classification.

The spectral characteristics of pixel digital numbers within each of the land cover types can be used to generate multivariate statistics parameters for each of the training sites. As the supervised classification methods are based on the statistical concepts, this classification is also termed as per-point or per-pixel classification.

The estimation of statistical properties, in particular the mean and variance of each spectral band and covariance of all pairs of spectral bands for training sites are made. The class means were estimated from training sites for all the land use classes and are presented in Table 4.1.

The descriptive statistics like mean and standard deviation for all the land use classes are presented in Table 4.2 to 4.6. The variance-covariance matrix estimated from training sites for all the land use classes are presented in Table 4.7 to 4.11.

5.1 Maximum likelihood classification

The area and production estimates for all the classes under maximum likelihood method are estimated along with the actual estimates collected from department of agriculture are presented in Table 4.12 and 4.13 respectively. The estimates obtained under this method are found to be nearer to the actual estimates for all the classes compared to other two methods.

The classification accuracy analysis is made using error matrix are discussed under section 5.4. Kappa coefficients, overall and mapping accuracy and F-measures are calculated and discussed further in this chapter under section 5.5, 5.6 and 5.7 respectively.

The Friedman two way analysis of variance is used to validate the maximum likelihood method and is discussed in section 5.8.

5.2 Mahalanobis distance classification

The area and production estimates for all the classes under Mahalanobis distance method are estimated along with the actual estimates collected from department of agriculture are presented in Table 4.12 and 4.13 respectively. The estimates obtained under this method are found nearer to the actual estimates next to maximum likelihood method. The classification accuracy analysis is made using error matrix are discussed in section 5.4. Kappa coefficients, overall and mapping accuracy and F-measures are calculated and discussed further in this chapter under section 5.5, 5.6 and 5.7 respectively. The Friedman two way analysis of variance is used to validate the Mahalanobis distance method and is discussed in section 5.8.

5.3 Minimum distance to means classification

The area and production estimates for all the classes under minimum distance method are estimated along with the actual estimates collected from department of agriculture are presented in Table 4.12 and 4.13 respectively.

The classification accuracy analysis is made using error matrix are discussed in section 5.4. Kappa coefficients, overall and mapping accuracy and F-measures are calculated and discussed further in this chapter under section 5.5, 5.6 and 5.7 respectively. The Friedman two way analysis of variance is used to validate the minimum distance method and is discussed in section 5.8.

The area statistics for land cover classes were estimated using different classification methods such as Maximum likelihood, Mahalanobis distance and Minimum distance classification and actual estimates of area collected from Department of Agriculture (DOA) are presented in Table 4.12. The classes estimated under maximum likelihood method are found to be more nearer to the actual estimates compared to other two methods. The production estimates using all three classifiers are also presented along with the actual production estimates for all land use classes in Table 4.13. the production

estimates under classes jowar and other crops obtained using Maximum likelihood method are found to be nearer to the actual estimates whereas the production estimate under class red gram obtained using Minimum distance classification is found close to the actual value.

The comparison of the area and production estimates obtained from classifiers with DOA estimates are made by taking their absolute deviation values from the DOA estimates and ranking has been done based on the magnitude of total deviation of each classifier which are presented in Tables 4.14 and 4.15 respectively . It is found that maximum likelihood estimate is nearer to the ground truth value. Higher the magnitude more nearer to the ground truth value.

5.4 Classification accuracy analysis:

The error matrix is generated using the actual and reference data for Maximum likelihood, Mahalanobis and Minimum distance method which are presented in Tables 4.16, 4.17 and 4.18 respectively. The error matrix is used to measure the accuracy of the classification methods.

The user's accuracy and producer's accuracy for individual classes ranged between 91.80 to 96.55 per cent and 90.00 to 95.00 per cent respectively for different classes using maximum likelihood algorithm which is presented in Table 4.19. The maximum likelihood method found to have the highest user's and producer's accuracy.

The user's accuracy and producer's accuracy for individual classes ranged between 82.76 to 96.30 per cent and 80.00 to 90.00 per cent respectively for different classes using Mahalanobis distance classifier, which is presented in Table 4.19. The next best to maximum likelihood method is the Mahalanobis method in terms of accuracy estimated.

The user's accuracy and producer's accuracy for individual classes ranged between 72.00 to 92.85 per cent and 70.00 to 86.66 per cent respectively for different classes using minimum distance to means classifier, which is presented in Table 4.19.

Among all the three methods the minimum distance classification method is found have less user's and producer's accuracy.

5.5 Kappa coefficients

The validity of classification accuracy of all the methods of classification algorithms are assessed using Kappa statistics computed from error matrix for each classification.

The Kappa coefficients for all the classifications were highly significant, which implies that all the classifiers produced classification significantly different from random result. All three classifiers are significant at 1 per cent level. The results are presented in Table 4.20.

The pair wise comparison between Kappa coefficients showed high significant difference between maximum likelihood and Mahalanobis distance classification, Maximum distance classification and Minimum distance to means classification, Mahalanobis distance classification and minimum distance classification at 5 per cent of level of significance. Whereas, the classification accuracy between maximum likelihood and minimum distance method do not differ significantly at 1 per cent level of significance. The results are presented in Table 4.21.

5.6 Overall and Mapping Accuracy

Overall classification accuracy evaluated from the error matrix and compared with methods, the highest classification accuracy was obtained with Maximum likelihood classification with 93.38 per cent for the study area, followed by Mahalanobis distance with overall accuracy of 87.66 per cent. Next to this was minimum distance classification with overall accuracy of 80.44 per cent which is presented in the Table 4.22.

The mapping accuracies for all three methods are obtained using error matrix which is represented in Table 4.22 and it is found that Maximum likelihood method has highest overall mapping accuracy of 87.51 per cent, next highest overall mapping accuracy was 78.42 per cent found with Mahalanobis distance classification and the least

overall mapping accuracy of 67.61 was found with minimum distance to means classification.

In this study the maximum likelihood method gave the best result with the highest overall accuracy of 93.38 per cent and overall mapping accuracy of 87.51 per cent compared to the other two methods.

5.7 F-measures

F-measures for all the classes under each classification method were calculated using error matrix. Recall, precision and individual F-measure were calculated by using the error matrix under each classification.

In Maximum likelihood method the recall, precision and individual F-measure for classes were found to be from 91 to 96 per cent, 90 to 95 per cent and 91 to 94 per cent respectively. For Mahalanobis the recall, precision and individual F-measure for classes was found to be 83 to 96 per cent, 80 to 90 per cent and 81 to 93 per cent respectively. Whereas for minimum distance to means classification it was found that the recall for individual classes ranges from 70 to 93, precision ranges from 70 to 87 per cent and individual F-measure ranges from 70 to 89 per cent which are presented in Table 4.23, 4.24 and 4.25 respectively.

The overall F-measure for all three methods was calculated and was found to be 0.928 0.874 and 0.802 in case of maximum likelihood, Mahalanobis distance and minimum distance to means classification respectively and the results are presented in Table 4.26.

Maximum likelihood method was found to have highest overall F-measure compared to other two methods. The next best is Mahalanobis distance classification method. Higher the value of F-measure better is the classifier's performance. Hence, the maximum likelihood method has performed better than the other two methods.

5.8 Friedman two way analysis of variance

The classifiers were subjected to non-parametric Friedman's two way analysis of variance (χ_r^2) for the purpose of validation. The results obtained using area estimates were presented in the Table 4.27. The results show that there is significant difference between the classifiers among the different classes for each data set.

Based on the magnitude of rank of classifiers, it is observed that rank of maximum likelihood classifier is high which indicates that it is most nearer to the ground truth observation and it is observed that the rank of Mahalanobis distance is next nearer to the ground truth observation. Rank of minimum distance to means classifier is next nearer to the ground truth observation after maximum likelihood and Mahalanobis distance classification.

Finally, it can be concluded that the classifier maximum likelihood method is more appropriate and best method for estimation compared to other methods. Mahalanobis distance method is next best to the maximum likelihood classification followed by minimum distance to means classifier.



SUMMARY

VI SUMMARY

Remote sensing technology has been proved as a useful application for natural resource evaluation and management. It has wide applications in various fields because of its inherent advantages. Satellite Remote Sensing technology has played a vital role in the agriculture sector by providing relevant information for strategic decisions and development activities. One of the important steps in remote sensing digital image analysis is land cover/ land use classification. The spectral response of a particular land cover class deviates from its ideal response due to the presence of noise. Application of statistical techniques in remote sensing image classification in order to partition the noisy image into its constituent classes is of great importance. With the objective of statistical comparison of remote sensing image classification methods with respect to their accuracy, test image obtained through IRS-P6 LISS-III on 22nd December, 2011 for Sedum Taluk of Gulbarga district was used for classification.

In the present study an attempt has been made to compare the classifications performed using Maximum likelihood classification, Mahalanobis distance classification and Minimum distance to means classification. The difference in classification results produced by different methods may not be apparent as such, quantitative methods are required to assess and compare the classification methods. Therefore classification results are evaluated using overall accuracy, overall mapping accuracy and Kappa statistics through the analysis of error matrix.

Supervised classification like Maximum likelihood classification, Mahalanobis distance classification and Minimum distance to means classification were performed for the test image. Supervised classification can be defined normally as the process of samples of known identity to classify pixels of unknown identity. Supervised classification requires interaction of the analyst to determine known classes and training regions.

Maximum likelihood classification is found to be the best with higher overall accuracy, overall mapping accuracy and overall F-measure of 93.38 per cent, 87.51 and

92.08 per cent respectively. The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. It is a parametric method of classification which relies on second order statistics of Gaussian probability model for each class. The maximum likelihood assumes that each pixel is independent of the other pixel and it has equal probability of being classified into any of land cover classes.

Mahalanobis distance classification is found to be the next best classification after maximum likelihood classification with overall accuracy of 87.66 per cent, mapping accuracy of 78.42 per cent and overall F-measure found to be 87.04 per cent. Mahalanobis distance considers mean vectors and population variance covariance matrices for each class, which is next best method of classification after maximum likelihood.

Minimum distance to means algorithm makes use of mean vectors of training site to assign an unknown pixel to the category it belongs using Euclidean distance. In the present study minimum distance classification resulted in overall accuracy, overall mapping accuracy, and overall F-measure of 80.44 per cent, 67.61 per cent and 80.02 per cent respectively.

A non-parametric test of Friedman two way analysis of variance tests were used for the validation purpose using area estimates. In the Friedman two analysis of variance it was found that the classifier maximum likelihood is very nearer to the ground truth observation based on the magnitude of ranks.

Future lines of work

The maximum likelihood classification is found be superior method among other classifiers. Other methods and algorithms available such as decision tree approach, knowledge based classifier, fuzzy classification, neural networks etc., can also be applied. Incorporation of these may improve the classification results. As such there is no general method of classification available. Estimation of land areas and production estimates can be achieved through classification procedures. The study can be extended to large area say districts and can be compared.

VII REFERENCES

- AL-AHMADI, F. S. AND HAMES, A. S., 2008, Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia. *JKAU Earth Sci*, **20**(1): 67-191.
- ASMALA AHMAD AND SHAUN QUEGAN, 2012, Analysis of Maximum Likelihood Classification on Multispectral Data. *Applied Mathematical Sciences*, **6**(129): 6425 – 6436.
- DENGSHENG LU, PAUL MAUSEL, MATEUS BATISTELLA AND EMILIO MORAN, 2004, Comparison of Land-Cover Classification Methods in the Brazilian Amazon Basin. *Photogrammetric engineering & remote sensing*, **70**(6): 723-731.
- DENGSHENG LU, SCOTT HETRICK AND EMILIO MORAN, 2010, Land Cover Classification in a Complex Urban-Rural Landscape with QuickBird Imagery. *Photogrammetric Engineering & Remote Sensing*, **76**(10): 1159–1168.
- DONALD I. M. ENDERLE AND ROBERT C. WEIH, JR, 2005, Integrating Supervised and Unsupervised Classification Methods to Develop a More Accurate Land Cover Classification. *Journal of the Arkansas Academy of Science*, **59**: 65-73.
- DWIVEDI, R. S., KANDRIKA, S. AND RAMANA, K. V., 2004, Comparison of classifiers of remote-sensing data for land use/land cover mapping. *Current Science*, **86**(2): 328-334.
- ELNAZIR RAMADAN, FENG XUE-ZHI AND CHENG ZHENG, 2004, Satellite remote sensing for urban growth assessment in Shaoxing city, Zhejiang province. *Journal of Zhejiang University Science*, **5**(9): 1095-1101.

- FEI YUAN, KALI E. SAWAYA, BRIAN C. LOEFFELHOLZ, N BAUER, 2005, Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment* **98**: 317 – 328
- KELLEY A. CREWS-MEYER, PAUL F. HUDSON AND RENÉ R. COLDITZ, 2004, landscape complexity and remote classification in eastern coastal mexico: applications of landsat-7 etm+ data. *geocarto International*, **19**(1): 45-56.
- KRISHNA BAHADUR, K.C., 2009, Improving Landsat and IRS Image Classification: Evaluation of Unsupervised and Supervised Classification through Band Ratios and DEM in a Mountainous Landscape in Nepal. *Remote Sensing*, **1**: 1257-1272.
- KRISHNA PRASAD, V., YOGESH KANT AND BADRINATH K. V. S., 1999, Vegetation discrimination using IRS-P3 WiFS temporal data set-A case study from Rampa forests, Eastern Ghats, A.P. *Journal of Indian Society of Remote Sensing*, **27**(3): 147-154.
- MANISHA B. PATIL, CHITRA G. DESAI AND BHAVANA N. UMRIKAR, 2012, Image classification Tool for Land Use / Land Cover Analysis: A Comparative Study of Maximum Likelihood and Minimum Distance Method. *International Journal of Geology, Earth and Environmental Sciences*, **2**(3): 189-196.
- MOHD HASMADI ISMAIL AND KAMARUZAMAN JUSOFF, 2008, Satellite Data Classification Accuracy Assessment Based from Reference Dataset. *International Journal of Computer and Information Engineering*, **2**: 386-392.
- MOHD HASMADI, I., PAKHRIAZAD H. Z. AND SHAHRIN M. F., 2009, Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. *Malaysian Journal of Society and Space*, **1**: 1-10.

- NANCY THOMAS, CHAD HENDRIX, AND RUSSELL G. CONGALTON A., 2003
Comparison of Urban Mapping Methods Using High-Resolution Digital Imagery.
Photogrammetric Engineering & Remote Sensing **69**(9): 963–972
- OFFER ROZENSTEIN AND ARNON KARNIELI, 2011, Comparison of methods for
land-use classification incorporating remote sensing and GIS inputs. *Applied
Geography*, **31**: 533-544.
- RENUKA DEVI, M. AND SANTHOSH BABOO, S., 2011, Land use and Land Cover
Classification using RGB&L Based Supervised Classification Algorithm.
International Journal of Computer Science & Engineering Technology, **2**(10):
167-180.
- ROY, P. S. AND DAS, K. K., 1991, Forest cover and use mapping in Kabri Anglong and
North Cachar hilly districts of Assam using Landsat MSS data. *Journal of Indian
Society of Remote Sensing*, **19**(2): 113-123
- RUSSELL G CONGALTON, 1991, A Review of assessing the accuracy of
classifications of remotely sensed data. *Department of Forestry and Resource
Management, University of California, Berkely*. **37**:35-46.
- SHAMSUDHEEN, M., DASOG, G. S. AND TEJASWINI, B., 2005, Land use/land
cover mapping in the coastal area of North Karnataka using remote sensing
data. *Journal of Indian Society of Remote Sensing*, **33**(1): 253-257.
- SHARMA AND SARKAR, 1998, Important contextual classification techniques used in
remote sensing. *Journal of Indian Society of Remote Sensing*, **43**(2): 145-178.
- SHARMA, D. P. AND BREN LEON, 2005, Effect of seasonal spectral variations on land
cover classification. *Journal of Indian Society of Remote Sensing*, **33**(2): 203-209.
- SHARMA, K. P. AND GARG, P. K. 1985, Digital Landuse Classification In Himalayan
Region From Remotely Sensed Data. *Civil Engineering Department, University
of Roorkee*, pp: 555-562.

- STEFFEN BOCK, 1996, A Region-Based Approach to land-use classification of remotely-sensed image data using Artificial Neural Networks. *International archives of photogrammetry and remote sensing*, **31**: 71-76
- SUDHAKAR, S., SRIDEVI, G., RAMANA, I. V., VENKATESHWARAA RAO, V., AND RAHA, A. K., 1999, Techniques of classification for land use/land cover with the special reference to forest type mapping in Jaldapara wild life sanctuary, Jalpaiguri district, West Bengal – A case study. *Journal of Indian Society of Remote Sensing*, **27**(4): 217-224.
- SUGUMARAN, R., SANDHYA, G., RAO K. S., JADHAV, R. N. AND KIMOTHI, M. M., 1994, Potential of satellite data in delineation of waste lands and correlation with ground information. *Journal of Indian Society of Remote Sensing*, **22**(2): 217-224.
- YAGOUB, M. M. AND KOLAN, G. R.M., 2006, Monitoring coastal zone land use and land cover changes of Abu Dhabi using remote sensing. *Journal of Indian Society of Remote Sensing*, **34**(1): 57-68.
- ZHENGRONG LI, YUEE LIU, ROSS HAYWARD, AND RODNEY WALKER, 2010 Empirical comparison of machine learning algorithms for image texture classification with application to vegetation management in power line corridors. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (Part A), ISPRS, Vienna, Austria*, pp. 128-133.