

(10)

USE OF REMOTE SENSING SATELLITE DATA

IN CROP SURVEYS

RANDHIR SINGH
R.C. GOYAL

FR
96/109



INDIAN AGRICULTURAL STATISTICS RESEARCH INSTITUTE
(ICAR)

LIBRARY AVENUE, NEW DELHI-110012

(2000)

(ICAR AP Cess Fund Project)

CONTENTS

1. Foreward	
2. Preface	
3. Abstract	
	Page
Chapter 1 Introduction	1
1.1 Some Preliminary Concepts on Remote Sensing Technology	2
1.1.1 Satellite and Spectral data	3
1.1.2 Digital Image Processing	6
Chapter-2 Review of Literature	9
2.1 Use of satellite data in crop production estimation	9
2.2 Use of satellite data in crop yield modeling	11
2.3 Orientation of the present study	13
2.4 Objectives of the study	14
2.5 Study area	14
2.6 Extent of data used in the study	15
Chapter 3 Estimation Procedures using Satellite Data	16
3.1 Land use/Land cover statistics	16
3.2 Area statistics	16
3.3 Crop yield estimation using satellite data along with crop yield survey data	17
3.3.1 Digital Image Processing Techniques for Crop Yield Estimation	18
3.3.2 Post stratified Estimation of Crop yield	20
3.4 Crop yield modeling using satellite remote sensing	21
3.4.1 Spectral yield models	21
3.4.2 Integrated yield models using satellite spectral data and farmers eye estimates	23
Chapter 4 Small Area Estimation	24
4.1 Small area estimation of crop yield at Tehsil/Block level	25
4.1.1 Direct Estimator	25
4.1.2 Synthetic Estimator	26
Chapter 5 Results and Discussions	27
5.1 Land use/Land cover statistics	27
5.2 Crop Acreage Estimation	28
5.3 Crop yield estimation	29
5.4 Crop yield modeling	30
5.5 Small area estimation of crop yield at Tehsil/Block level.	31
<i>Tables</i>	32
<i>Summary</i>	38
<i>References</i>	41

FOREWARD

Agriculture is the backbone of Indian economy. Reliable, accurate and timely information on crop statistics are of vital importance for planning agricultural policies. Presently the estimates of crop acreage are generally obtained through complete enumeration whereas the estimates of average yield are obtained through general crop yield estimation surveys based on crop cutting experiments. About 5 lakh crop cutting experiments are conducted annually in the country for different crops. System of preparing crop forecasts/advances estimates is still very subjective and is in the process of being evolved more objectively.

With the advent of remote sensing technology there is great potential in the field of agriculture for improving the agricultural statistics system. Deptt. of Space has taken several studies under the Crop Acreage and Production Estimation (CAPE) project for developing methodology for application of satellite data for crop statistics. The use of remote sensing satellite data along with the crop yield data based on crop cutting experiments can greatly improve the efficiency of crop yield estimators. In an earlier study at IASRI it has been observed that the use of satellite data can significantly improve the efficiency of crop yield estimators and suggested that either number of crop cutting experiments can be considerably reduced for obtaining crop estimates at district level with desired precision resulting in considerable savings every year or small area estimates at lower administrative units like tehsil or block level can be obtained with the existing number of crop cutting experiment in crop yield estimation surveys.

The present study, for which financial support has been provided by the ICAR AP-Cess fund, aims to develop the integrated methodology for agricultural statistics like crop acreage estimation, crop yield estimation and crop yield modeling with the combined use of general crop estimation survey data, satellite spectral data and also the farmers eye estimate of crop yield.

I am happy to note the timely publication of this report as serious efforts currently underway to make efficient use of satellite data and other sources of data for obtaining reliable agricultural statistics. I appreciate the efforts put up by the scientists and other staff members of the Institute in bringing out this report. I am sure the publication of this report is likely to provide crucial information for micro-level planning at lower administrative levels like tehsils or blocks for obtaining reliable crop statistics.

S.D. Sharma

Director

PREFACE

India is predominantly an agrarian economy. The availability of reliable and timely agricultural statistics is hence of paramount importance to the planners, administrators, policy makers and research workers.

India has a long history of organizing various kinds of agricultural statistics. The Indian agricultural statistics systems is quite comprehensive and collects various types of agricultural statistics namely land use statistics, area statistics, yield statistics etc. Directorate of Economics and Statistics in the Department of Agricultural and Cooperation is the nodal official agency for collection, compilation and publication of major agricultural statistics. Crop yield estimation surveys based on crop cutting experiments are conducted for almost all major crops in the country for estimation of crop yield and crop production.

The use of remote sensing satellite data for agricultural statistics like crop area, crop yield and crop yield forecasting is being currently investigated worldwide. The spectral data is manifestation of various factors like soil, weather, agricultural inputs etc. affecting the crop yield and is highly co-related with crop yield and as such can be used with great advantage for developing appropriate methodologies for more reliable agricultural statistics. In the present study remote sensing satellite data in the form of vegetation indices along with crop yield data from crop yield estimation surveys and farmers appraisal of crop yield have been combined together for developing efficient crop yield estimates at **tehsil/block** level and for providing more accurate crop yield forecasting models.

The authors are grateful to the Indian Council of Agricultural Research for providing necessary financial support to undertake the study through which a well equipped Remote Sensing Laboratory has also been established at the Institute. Authors are grateful to Director, NRSA, Hyderabad for providing the satellite data and Director of Agriculture, Additional Director of Agriculture (Statistics) and Jt Director of Agriculture (Statistics), **Haryana**, Chandigarh for providing the crop yield data based on crop cutting experiments and support for completion of the project.

Authors are extremely grateful to the Director, **IASRI** for providing all required facilities and continuous guidance and encouragement during the course of this investigation. Authors are thankful to Dr. A.K. Srivastava, Dr. Anil Rai and Dr. **Tauqueer** Ahmad for consultations and their useful suggestions.

Authors are thankful to **Shri D.C. Dahiya** and **Shri Dinesh Prasad Semwal** for their untiring work in digital analysis of satellite data and preparing the final report and also **Shri C.P. Singh** and **Shri Meer Singh** for their assistance in analyzing the data. Last but not the least authors are thankful to Miss Dolly Das for efficiently typing the manuscript.

AUTHORS

USE OF REMOTE SENSING SATELLITE DATA IN CROP SURVEYS

ABSTRACT

India is predominantly an agrarian **economy**. The availability of reliable and timely agricultural statistics is hence of paramount importance to the planners, administrators, policy makers and research workers. India has a long history of organizing various kinds of agricultural statistics.

General crop yield estimation surveys based on crop cutting experiments are conducted throughout the country for estimating crop yield of all major crops. With the advent of remote sensing technology satellite data has been widely used for obtaining crop statistics. In India several studies have been conducted during the past decade by Dept. of Space under the Crop Acreage and Production Estimation (CAPE) project for various major crop statistics using the satellite spectral data. Singh et al.(1992) and Singh et.al. (2000) used the satellite spectral data along with the survey data on crop yield from general crop yield estimation surveys to develop more efficient post stratified estimators of crop yield at district level and also small area estimators of crop yield at **tehsil** level

In the present study, which has been funded through the ICAR Agricultural Produce- cess fund, an integrated methodology for providing area and yield estimation and yield forecasting models using satellite data, crop yield data from general crop yield estimation surveys based on crop cutting experiments and also the farmers eye appraisal of crop yield has been developed. Also small area estimators of crop yield at Block level have been obtained using satellite spectral data and the crop yield data from general crop estimation surveys.

For crop yield estimation, the satellite data in the form of vegetation indices has been used for stratification of crop area into homogeneous crop growth condition classes like high vegetation, average vegetation, poor vegetation, no vegetation etc. and post-stratified estimators of crop yield have been developed which are more efficient as compared to the usual estimator of crop yield. In case of crop yield forecasting models, the satellite data in the form of vegetation indices along with the farmers eye appraisal of crop yield have been used as explanatory variables to forecast the crop yield.

Two small area estimators namely (i) the Direct estimator and the (ii) the Synthetic estimator of crop yield have also been obtained at block level.

1. INTRODUCTION

Agriculture is the backbone of Indian economy, contributing about 40 percent towards the Gross National Product (GNP) and providing livelihood to about 70 percent of the population. So for a primarily agriculture based country like India, reliable accurate and timely information on types of crops grown and their acreages, crop yield and crop growth conditions are vital components for the planners engaged in formulating and implementing appropriate prices of agricultural commodities, strengthening country's food security and distribution system and import/export policies of these commodities from time to time and in efficient management of natural resources. The availability of accurate and timely data on agricultural production would not only help the planners in formulating development programmes in rural areas but also enable them to take appropriate decisions on policies relating to import/export of these commodities well in advance.

India is one of the few countries which has a well established system of collection of agricultural statistics and detailed statistics of land utilization are continuously available since 1884. The agricultural crop production of principal agricultural crops in the country is usually estimated as a product of area under the crop and the average yield per unit area of the crop. The estimates of the crop acreage at a district level are obtained through complete enumeration whereas the average yield is obtained on the basis of crop cutting experiments conducted on a number of randomly selected fields in a sample of villages in the district. The technique developed during forties has, by and large, been followed for estimating the production of major crops in the country. However the traditional system of estimation of crop production has several problems, viz. lack of timely information, variation in statistical figures, accessibility and quick retrieval of data and heavy burden of work on village level worker (Patwari).

The crop forecasts/advanced estimates of crops are presently developed by Ministry of Agriculture for taking policy decision relating to procurement, marketing, export, import etc. The advance estimates of kharif crops are first prepared in July/August tentatively when behaviour of South West monsoon is clear and reports of coverage of area under crops from the states are available. The advance estimates are reviewed during December/January when estimates of area under kharif crop become available under the Timely Reporting Scheme (TRS) and results of the crop cutting experiments portion from the NSSO (normally 10%) become available. The advance estimates of rabi season are also prepared at the stage. The advance estimates are again reviewed in the month of April based on information obtained from the states giving the final forecast for kharif.

With the advent of Remote Sensing Technology during 1970s, its great potential in the field of agriculture have opened new vistas of improving the agricultural system all over the world. Space borne remotely sensed spectral satellite data has been widely used in the field of agriculture for estimation of area under different major crops like wheat, paddy, groundnut and sugarcane studies have also been made to examine the relationship of crop

growth parameters like leaf area index (LAI) representing crop vigour and the spectral data in the form of several vegetation indices developed from the spectral data of various bands.

Remote sensing satellite data can also be used for improving the crop yield estimation through crop cutting experiments and also for developing models for crop yield using historical data, meteorological data, and remotely sensed satellite data. This may lead to the development of an efficient integrated system for crop statistics like crop acreage estimation, crop yield estimation and crop yield forecasting.

During 1990-93 a study was conducted at the Institute to examine the usefulness of satellite spectral data for stratification of crop area based on vegetation indices for improving crop yield estimation based on yield data from crop cutting experiments under crop yield estimation surveys. The study pertained to wheat crop yield for district Sultanpur UP for 1985-86 and the satellite data was used from the USA satellite Land Sat-4. This study showed that the efficiency of crop yield estimation can be increased considerably by using the satellite data along with the survey data. The results of this study are given in Singh et. al.(1992). Another similar study was undertaken during 1996-98 for improved estimation of wheat crop yield in district Rohtak for 1995-1996 using the IRS 1B - LISS II satellite data for Feb. 17, 1996 and the crop yield data from crop yield estimation surveys for Rabi 1996. The results from this study presented in Singh et. al (2000). also showed that satellite data in the form of vegetation indices greatly improves the efficiency of crop yield estimator.

In the present study integrated methodology for crop statistics like land use statistics, crop acreage, crop yield and crop yield forecasting is attempted using the satellite data along with the crop yield data based on crop cutting experiments from general crop estimation surveys. Also small area estimates of crop yield at Tehsil/block level have been developed..

1.1 Some Preliminary Concepts on Remote Sensing **Technology**

Remote sensing means acquiring information about a phenomenon, an object or surface from a distance and without actually coming into physical contact with it. The technique is based on the characteristic features of earth's surface which exhibit fairly consistent and to a certain extent unique properties of emitting /reflecting/transmitting electromagnetic radiation (EMR). Observations are made on reflected/scattered/emitted energy from the earth in different wavelengths by the sensors of the satellite. The reflectance/emittance of any object at different wavelengths follow a pattern which is characteristic of that object and is known as Spectral signature which leads to the identification of the object.

Electromagnetic radiation forms a very broad spectrum varying from very low frequency to very high frequency or from long wavelength to short wavelength. The electromagnetic energy is available in many forms like visible light, ultra violet rays, radio waves, X-rays and micro waves, etc. All this energy is inherently similar and radiates in accordance with basic wave theory. The entire range of electromagnetic energy from very long wavelength to very short wavelength is known as Electromagnetic Spectrum (EMS) as

given in Fig.1. The reflected/transmitted energy travels in the form of wave at the velocity of light C . The distance from one wave peak to the next peak is defined as **wavelength** and is denoted by λ . The number of peaks passing a fixed point in space per unit time is defined as the wave frequency denoted by f such that $C = \lambda f$.

In remote sensing it is common to categorize the electromagnetic waves by their location on EMS and the common unit to measure the wave length along the spectrum is micro meter (μm) which is equal to 10^{-6}m . The EMS of great interest to us is the optical wavelength which extends from 0.3 μm to 15 μm . The region between 0.38 μm to 3.0 μm is usually referred to as the reflective part of the spectrum. Energy sensed in these wavelengths is primarily the radiation originating from the Sun and reflected by the objects on the earth. The reflective part of the spectrum is divided into visible band (0.38 μm to 0.72 μm) since human eye responds to the radiation of these wavelengths and the reflective infra-red band (0.72 μm to 3.0 μm) which may be further divided into near infra-red (0.72 μm to 1.30 μm) and middle infra-red(1.30 μm to 3.0 μm). In visible band approximate range of blue colour is from 0.38 μm to 0.5 μm , green from 0.5 μm to 0.6 μm and red from 0.6 μm to 0.72 μm . The ultra violet energy extends to just shorter wavelengths than the visible wave length while the wave length between 7-15 μm are termed as thermal and microwaves. The wavelength region between 3.0 μm to 7.0 μm is not usually attributed any special terms since the atmospheric effects complicate interpretation of the radiation data in this region and in fact limit the usefulness of these wavelengths.

1.1.1 Satellites and spectral data :

The satellites can be divided into two major categories depending upon their altitude and the orbits in which they are moving.

i) Geostationary Satellite

An equatorial west to east satellite orbiting the earth at an altitude of about 36,000 Km, the altitude at which it makes one revolution in 24 hours, synchronous with the earth's rotation is called a Geostationary satellite. These satellites cover the same place and give continuous near hemispheric coverage over the same area day and night. These satellites are mainly used for communication and weather monitoring viz. GOES, METEOSAT, INTELSAT, INSAT, etc.

ii) Sun Synchronous Satellite

A satellite orbiting in the plane which is near to polar, at an altitude between 700 to 900 Kms. such that satellite passes over all the places on earth having the same latitude twice in each orbit at the same local sun-time is known as Sun synchronous. Through these satellites, the entire globe is covered on regular basis and giving repetitive coverage on periodic basis. All the remote sensing resources satellites may be grouped into this category. A few of these satellite are LANDSAT, SPOT, IRS-IA, 1B, 1C, 1D, NOAA, SEASAT etc.

The orbit of these satellites are Sun-synchronous so that the spectral data for a given point on earth are collected at the same time of the day each time the satellite passes.

Sensors and bands

Different object **reflect/emit/scatter** different amount of energy in different wavelength bands of EMS and these typical characteristics are the source of information to describe the nature and condition of the objects. The multi-spectral scanners in remote sensing satellite provides potential tools to differentiate, identify and map various ground cover types such as vegetation, water, **habitation**, snow covered areas and base soils, etc. The signal which are reflected by different objects are received by the sensor detectors for different bands. The sensor is a device that gathers energy (EMR) and converts it into signals and presents it in the form of digital data for obtaining information about the objects. These sensors operate in visible, near infra-red, thermal and microwave regions of the EMS and these regions are given the name '**spectral bands**'. Bands of spectral sensors employed on different satellites correspondence to spectral regions defined for different purposes. LANDSAT, 1, 2 and 3 series of satellites had multi-spectral scanners(MSS) having 4 bands while LANDSAT-4 and 5 included Thematic Mapper' (TM) with 7 spectral bands. Indian Remote Sensing Satellite IRS-1A launched on March 17, 1988 and IRS-1B launched on August 29, 1991 had two sensors LISS-I (linear Imaging self scanner) which had spatial resolution of 72.5 meters and LISS-II which had the spatial resolution of 36.25 meters. Both LISS-I and LISS-II operate in four bands in visible and near infra-red regions. IRS-1C launched on Dec.28, 1995 and IRS-1D lunched in 1997 have three different sensors LISS-III, Panchromatic (PAN) and Wide Field Sensor (WiFS). LISS-III sensor has a spatial resolution of around 23.50 meters. PAN has a very high spatial resolution of about 5.8 meters along with stereo capability while the spatial resolution of WiFS is about 188 meters. The salient features of various remote sensing resources Satellites are tabulated in Table 1.1

Satellite Resolutions

The information acquired through remote sensing depends upon the satellite resolutions which measure of ability of an optical system to distinguish between the signals. The different types of resolutions are discussed below:

i) **Spectral Resolution**: This refers to the location of the spectral bands in the electromagnetic spectrum (EMS). It is a measure of both the discreteness of the band width and the sensitivity of the sensor to clearly distinguish between grey levels.

ii) **Spatial Resolution** : It is the ability of the sensors to measure the spectral properties of the smallest target i.e. the minimum distance between two objects that a sensor can record distinctly, It is also called ground resolution element (GRE) or the picture element (pixel).

iii) **Temporal Resolution**: Obtaining spatial and spectral data at certain time intervals is called the temporal resolution

Table 1.1 : Characteristics of different satellites

SATELLITE	LANDSAT 1,2,3 (MSS)	LANDSAT 4,5 (TM)	IRS- 1A,1B	IRS-1C,1D	SPOT	AVHRR (NOAA)
CHARACTERISTIC						
Lunch date	1972,1975,1978	1982,1984	1988,1991	1995,1997	1984	1984
Linear Resolution	80 mtrs.	30 mtrs.	LIIS-72 mtrs.	LISS-III 23.5 mtrs. PAN5.8mtrs WiFS 188 mtrs.	20 mtrs Multi Spectral	1.1 Kms Nadir
Orbit Repeat Period	18 days	16 days	22 days	25 days	26 days	1 day
Mean Altitude	919 kms	705 kms	904 kms	821 kms	822 kms	833 kms
Swath Width (Nadir)	185 kms	185 kms	148 kms	141 kms	2x62 kms	3000 kms
No. of bands	4	7	4	4-LISS-III 1-PAN 2-WiFS	Pointable Across Track	5
Spectral bands(μm)	0.5-0.6 0.6-0.7 0.7-0.8 0.8-1.1	0.45-0.52 0.52-0.62 0.63-0.69 0.76-0.90 1.55-1.75 10.4-12.5 2.08-2.35	0.44-0.50 0.52-0.58 0.62-0.69 0.77-0.86	0.52-0.59 0.62-0.68 0.77-0.86 1.55-1.70 PAN 0.50-0.75 WiFS 0.62-0.68 0.77-0.86	MSS 0.50-0.59 0.61-0.69 0.79-0.90	0.58-0.68 0.725-1.10 3.55-3.93 10.30- 11.30 11.50- 12.50
Sponsoring Country	U.S.A.	U.S.A.	INDIA	INDIA	FRANCE	U.S.A.

1.1.2 Digital **Image Processing**

The digital nature of remotely sensed data supporting quantitative and statistical analysis of spectral measurements led to rapid advancement in the field of digital image processing. Digital image processing encompasses the operations such as noise removal, information extraction, and image data manipulation and management. Digital image consists of discrete picture elements called pixels. Associated with each pixel is a digital number (DN) that depicts the average radiance of pixel area. The various important steps involved in image processing including image correction image enhancement and information extraction are described below.

a) Image Enhancement:

Image enhancement techniques improve the quality of an image as perceived by the human. These techniques are useful because many satellite images when examined on a colour display give inadequate information for image interpretation. There exist a wide variety of techniques for improving the image quality. The contrast stretch, density slicing, edge enhancement and spatial filtering are the most commonly used techniques. Image enhancement techniques are applied after the image is corrected for geometric and radiometric distortions i.e. after the image is rectified. Image enhancement methods are applied separately to each band of the multi spectral image.

b) Contrast Enhancement

Contrast generally refers to the difference in luminance or grey level values in a image and is a very important characteristic. Contrast ratio can be defined as the ratio of the maximum intensity to the minimum intensity of an image. Contrast ratio has a strong bearing on the resolving power and detectability of an image. Larger this ratio more easy to interpret the image. Contrast enhancement techniques expand the range of brightness values in the image so that the image can be efficiently displayed in a manner described by the analyst. The DN values in a scene are literally pulled further apart that is expanded over a greater range. The effect is to increase the visual contrast between two areas of different uniform densities. This enables the analyst to discriminate easily between areas initially having a small difference in density. Contrast enhancement can be effected by linear or non linear transformation. Linear Contrast stretching is the simplest contrast stretch algorithm. The grey values in the original image and the modified image follow a linear relationship in this algorithm. The image display and recording devices typically operate over a range of 256 grey levels (The maximum number represented in 8 bit computer recording) The sensor data in a single image rarely extends over the entire range. So the intent of contrast stretching is to expand the narrow range of grey level brightness values present in an input image over a wider range of grey values. The resulting output image could be more easily distinguished as the light tonal areas would appear darker. The linear stretch is applied to each pixel value in the image using the algorithm

$$y = \frac{X - X_{\min}}{X_{\max} - X_{\min}} * 255$$

Where

y is the stretched digital gray value assigned to pixel in the output image.

x is the original gray value of pixel in the input image.

X_{\min} - is the minimum gray value present in the input image .

X_{\max} - is the maximum gray value present in the input image.

For spatial analysis, a specific feature may be analyzed in greater radiometric details by assigning the display range exclusively to a particular range of image values. For example if agricultural features could be enhanced by stretching the small range to full range on the stretched display, minute tonal variations in the agricultural range would be greatly exaggerated. In short the contrast stretch displays permit the image interpreter to evaluate the radiometric details in a better way as compared to the original image.

There exists some non-linear contrast enhancements also where the input and output data values follows a non linear transformation.

c) Density slicing

Digital images have high radiometric resolutions. Images in same wavelength band contain 256 distinct gray levels. But a human eye interpreter can reliably detect and consistently differentiate between 20 to 25 shades of gray level only. However human eye is more sensitive to colour than different shades between black and white. Density slicing is a technique that converts the continuous gray tone of an image into a series of density intervals or slices, each corresponding to a specified digital range. Each slice is displayed in a separate colour. This technique is applied on each band separately.

d) False Colour Composite (FCC):

The number of bands varies from 4 to 7 in different satellites and the digital data is received on the ground station for all the bands. The number of channels/filters provided for processing the digital data are restricted to only three which means that the digital data can be processed for a maximum of three bands simultaneously taking suitable combination of different bands. False Colour Composites (FCC's) are generated by assigning blue, green and red filters to three different bands and the resultant output is a colour picture image called false colour Composite (FCC). In case of IRS series of the satellites generally band 4 (Near infra red) data is passed through red channel, band 3 (red) data is passed through green channel and band 2 (green) data is passed through blue channel for preparing the FCC's. The red shades in this image corresponds to vegetation which includes agriculture crops as well as

forest region. With a prior information about geographical location besides built in knowledge and experience and with the help of 'ground truths' the forests and the crops can be distinguished by the tone and texture of the red shades on the FCC's. The higher red tones could address to terrain cultivation features and dark red colour tone may be classified into deep forest areas. In general blue shade is associated with water bodies, dark blue colour refers to deep water, very light bluish shades corresponds to human settlement regions and light yellow to whitish shades represents the fallow/barren lands and white spots corresponds to snow covered areas and clouds. The clouds are separable from snow covered areas as the clouds are associated with shadows. The colour reference as mentioned above is however, only representative as the photo processing variation may cause changes in the colour tones.

2. REVIEW OF LITERATURE

The Board of Agriculture in India recommended as early as 1919, an objective method of conducting crop cutting experiments for estimation of crop yield through a random selection of villages, fields and plots. But the credit of visualizing the problem of crop estimation from an entirely new angle of carrying out the first yield estimation survey in 1923-25 on the principles of random sampling goes to Hubback (1927). The concept of randomization however was lacking in the surveys conducted by Hubback as the sampling at a given center was limited to only those specified fields where the harvesting was in progress on the day of the visit of the investigator. Sampling methods similar to those evolved by Hubback were adopted by Mahalanobis (1945) where crop cutting experiments conducted for the yield estimation of wheat and gram in two districts, Shahbad and Monghyr, as part of Bihar crop survey in 1943-44 were presented. Mahalanobis (1946) presented the investigations on Bihar and Bengal surveys.

A random sampling method for estimation of crop yield where main emphasis was given on the selection of plot for harvesting (i.e.the ultimate sampling unit) by a strict process of randomization in place of subjective selection by investigator was introduced by Sukhatme and Panse (1951). This technique has been adopted by the ministry of Agriculture and has by and large been followed by the Central and State Departments of Agriculture for estimating the production of major crops in the country. About 5 lakh crop cutting experiments are conducted annually under this scheme known as General Crop Estimation Surveys.(GCES). A stratified multistage random sampling design is adopted in these surveys where the blocks constituted the strata, the villages selected randomly formed the primary sampling unit, the fields selected from each village formed the second stage unit and the plot within the field formed the ultimate stage of sampling. A sample of villages is selected from different strata in proportion to the area under crop. From each selected village, two fields are selected randomly and from each field, a plot measuring 10m x 5m is selected for recording the yield by actual harvesting the crop.

2.1 Use of satellite data in crop production estimation

Global Scenario

The era of high quality observations of the earth surface through space began in 1972 with the launch of ERTS-1, (Later named as LANDSAT) in USA with the Multispectral Scanners (MSS) at 80-metre spatial resolution. Since 1982 Landsat Thematic Mapper (T.M.) has provided 30 meter spatial resolution data which are of better quality than the data received from MSS.

The use of space borne remote sensing data for large area crop survey was explored in USA under Corn Blight Watch Experiment (CBWE) in 1971, under Crop Identification Technology Assessment for Remote Sensing (CITARS) in 1973 followed by an attempt to forecast wheat crop production for major growing regions of the world under Large Area

Crop Inventory Experiment (LACIE) during 1974-1977. Later a six year programme of research and development named Agriculture and Resource Inventory Survey Through Aerospace Remote Sensing (AGRISTARS) was taken up in 1988. Since then large scale methodology **development-cum-demonstration** studies for crop statistics have been carried out in Africa and Europe as well as in a number of other countries (Argentina, Australia, **Brazil**, Canada, Japan etc). Currently major programmes are underway in Africa under Global Information and Early Warning System (GIEWS) and in Europe under Monitoring Agriculture through Remote Sensing (MARS). The MARS project has developed rapid crop survey procedure for Crop Growth and Monitoring System (CGMS) which employs crop simulation models, agro-meteorological models, and real time data for crop forecasting and assessment. USDA Statistical Reporting Service (SRS) has integrated Landsat data in domestic crop estimation programme. USDA makes use of the Landsat data for stratification (based on visual interpretation) and to classify the digital data into crop types and regress SRS ground-collected data results from the area sampling segments (0.7 sq. miles) on the classified Landsat data for each crop type. A direct expansion estimator is used to expand the data to state, regional or national level (Hanuschak et al. 1982).

The use of spectral data has also been investigated for obtaining crop yield estimation and many empirical studies have been conducted to evaluate the feasibility of the application. Several vegetation indices have been developed and shown to be very well correlated with the agronomic variables and hence the crop yield. Tucker *et.al.*(1980) and Rudroft and Batista (1991) are among many other studies related with the relationship of spectral data and crop yield.

Experience in India

India entered the space age by launching of the Indian Remote Sensing satellite (IRS-1A) on March 17, 1988 which has opened new vistas in data collection techniques for agriculture and natural resources in terms of speed and quality. IRS-1A provided comparable data to those obtained from LANDSAT, as the sensor LISS-I system is well comparable with MSS while LISS-II was comparable with TM as far as resolutions were concerned. As a follow up and towards ensuring continuity of data availability to the users the IRS-1B Satellite identical with the IRS-1A was launched in August 1991. Subsequently IRS-1C was launched in 1995 and IRS-1D in 1997. Data obtained from these satellites opened vast area of research and applications and several organizations have been engaged in development of new and improved methodologies for the applications of satellite data.

Systematic multi-crop and large scale investigation on use of remote sensing data for crop statistics has been undertaken under Crop Acreage and Production Estimation (CAPE) project by Deptt. of Space at Space Application Center (SAC) Ahmedabad since 1986, This country-wide project funded by the Ministry of Agriculture and being executed jointly by SAC, State Remote Sensing Centers, State Department of Agriculture and Agricultural Universities. The project aims at **pre-harvest** production forecasts at regional (district /group of districts level) with pre specified goals. It covers major crops like wheat, rice, rabi sorghum, cotton, mustard, groundnut in their major growing areas. First attempt in the

country towards the use of Satellite digital data in crop acreage estimation was made in the Karnal district of Haryana using Landsat MSS data (Dadhwal and Parihar 1985). Problems involved in operationalizing of remote sensing techniques in India have been pointed out by Sahai and Ajai (1988). A concerted effort has been made to develop methodology for applicability over large areas at Space Applications Centre. The crops being studied include wheat, rice, sorghum and groundnut. The work has so far emphasized the use of single-date data and supervised Maximum Likelihood (MXL) classification approach.

Wheat acreage estimates using administrative-boundary-overlaying approach, single-date Landsat MSS digital data and supervised maximum likelihood classification approach were made for 1983-84 Rabi wheat in parts of Karnal district (Dadhwal and Parihar 1985) and for 1984-85 wheat season in Patiala tehsil (Kalubarme and Mahey 1986). As these results were considered encouraging, large area studies using sample segment approach were taken up in Haryana, Punjab and selected districts in Western Uttar Pradesh. While the basic methodology has remained the same, in the above 3 states, there are some procedural "differences especially in stratification. In Haryana in 1985-86 originally historical wheat acreages were used for stratifying into 3 strata (Anon. 1986a). When the results were compared with two-step stratification in 1986-87 advantages of using 5 agrophysical strata and second-stage of stratification based on potential agricultural area became apparent (Dadhwal et al. 1987)). In case of Punjab, in 1986-87 each district was considered as a stratum (Sridhar et al 1987) whereas with a known agro-climatic map lower coefficient of variation was attained in 1987-88 results (Sridhar et al. 1989). In case of Uttar Pradesh, 1986-87 satellite data was used to stratify study area into vegetation density classes and sample segment of 5x5 km area were used.

Use of satellite data along with survey data of crop yield from the general crop estimation surveys based on crop cutting experiments (GCES) for obtaining improved estimators of crop yield has been undertaken at IASRI since 1990 Singh et al (1992). Singh and Goyal (1993) and Singh et al (2000) have presented improved post-stratified estimator of crop yield using satellite spectral data in the form of vegetation indices for stratification of crop area into homogeneous crop growth condition classes like very good crop, average crop, poor crop etc.

2.2 Use of Satellite **data in crop yield modeling.**

Forecasting of crop production is one of the most important aspect of agricultural statistics system. Yield forecasts at present are based on eye estimates and the final crop production estimates based on objective crop cutting surveys become available long after the harvests. This as such calls for the necessity of objectives methods for **pre-harvest** forecast of crop yields.

The main factors affecting crop yield are inputs and weather. Use of these factors forms one class of models for forecasting crop yields. The other approach uses plant vigour measured through plant characters. It is assumed that plant characters are integrated affects

of all the factors affecting; yield. Yet another approach is measurement of crop vigour through remotely sensed data. These approaches are being tried by various organizations.

The approach using weather parameters is normally based on time series data. The major work in this regard has been attempted at **IMD** (Sarker, 1977, Sarwade, 1988). Their studies involve identification of significant weather parameters in different periods and utilizing these parameters in the regression model along with trend. At **IASRI**, studies have been carried out at district level using weekly weather parameters. Various composite weather variables were derived as weighted accumulations of weekly weather parameters/interactions up to the time of forecast and were used as regressors in the model along with trend. **Principal** components of weather variables were also tried for developing the model (Agarwal et al.1986; Jain et al. 1980). The problem associated with meteorological model is assumption of same weather prevailing in a larger area as observatories are sparsely located. These models also require long series of data which are not available for most of the locations.

The other approach using plant characters collected at farmers' fields has been attempted through **pilot** studies at **IASRI**, New Delhi. The data have been collected at different periodic intervals through suitable sampling design for 3 to 4 years. Mainly two types of models, between year and within year models have been used. Between year models are based on historic data and involve an assumption that present year is a part of the composite population of the previous years. These models utilize the plant characters at some suitable phenological stage of crop growth either as such or their suitable transformations through multiple regression technique (Sardana et al.1972; Jha et al. 1981, Singh et al. 1988). Models were also developed using plant characters data of two or more periods through growth indices/principal components (Jain et al. 1984, 1985). Agarwal, Jain and Jha (1986) studied models based on crop weather relationship for Rice. Box and Jenkins (1976) used time series models for forecasting where the variation in yield during different years is explained using historical data through trend analysis and presented the well known technique of auto regressive integrated moving averages **ARIMA**.

In case of crop yield modeling using satellite data, several studies have been undertaken to establish relationship between spectral parameters through vegetation indices and the crop yield. Sridhar et. al.(1994) presented wheat production forecasting for a predominantly un-irrigated region in Madhya Pradesh. Singh and Ibrahim (1996) examined the use of multi date satellite spectral data for crop yield modeling using Markov Chain Model. Saha (1999) used satellite data and **GIS** for a developing several crop yield models.

Recently with a view to collect, collate and assimilate large data from different sources, a National Crop Forecasting Centre (**NCFC**) has been set up under the Ministry of Agriculture during 1998. Deptt. of Space have also recently launched a project - Forecasting Agricultural output using Space, Agro-meteorology and Land based observations (**FASAL**) envisaging advance reliable assessment of crop acreage and production using remote sensing techniques and also other data. Very recently National Wheat production forecast for (1998-99) using multi date WiFs and meteorological data have been developed under this project.

A study on "Evaluation of crop cut method and farmers reports for estimating crop production" Verma et al (1988) was undertaken at Longacre Agricultural Development Centre UK. This study was carried out in 5 countries in Africa during 1987 with the objective of comparing crop estimates based on crop cut methods with estimates obtained by asking farmers directly to state their production. The results of the study showed that farmers eye estimates are remarkably close to actual production figures in all the countries and they also show considerably small variance compared to the estimates based on crop cutting experiments. After the publication of this report considerable interest is again focused on using farmers estimates which are much cheaper to obtain and easier to conduct.

In the present study, therefore an effort is made to use the farmers eye estimate more objectively as a auxiliary variable along with the spectral indices to improve the efficiency of crop yield models for forecasting crop yield. Farmers estimates were obtained for the same fields in which crop cutting experiments were conducted .

- Note: Because the study has been conducted for developing crop yield forecasting model using the past data of yield based on crop cutting experiments and the satellite data for the year 1997-98, the farmers estimates were obtained only after the crop had been harvested. However for operational use of the model based on spectral data and farmers eye estimate the optimum time for recording the spectral data is Feb. March when the crop has maximum vigour and hence highest correlation with yield. The farmers estimate may also be obtained at the same time for timely availability of crop yield forecast before actual harvest of the crop.

2.3 Orientation of the present study

At the Indian Agricultural Statistics Research Institute (IASRI) New Delhi several studies have been taken up since 1990 for using satellite data along with the survey data of crop yield based on crop cutting experiments for crop yield estimation. Singh et. al. (1992) Singh et. al.(2000) have shown that use of spectral data for stratification based on vegetation indices of crop area in general crop yield estimation surveys can greatly improve the efficiency of crop yield estimator.

In our country the field sizes are fairly small (some times less than an acre). It is therefore not feasible to correlate spectral data with the grain yield at field level instead it is more appropriate to utilize the spectral data as auxiliary information. In this regard a study on 'Use of Remote Sensing Technology in crop yield estimation surveys' in Sultanpur District of U.P. was conducted at the Institute during 1990-93 where the LANDSAT (TM) satellite spectral data dated 23rd Feb., 1986 was utilized to stratify the crop area into different vegetation vigor classes based on the vegetation indices NDVI and RVI and the post stratified estimators of crop yield based on plot yield as obtained from crop cutting experiment were developed. The study revealed that the post stratified estimators using satellite data along with survey data are more efficient as compared to the usual crop yield estimator. It was also observed that the use of NDVI for post stratification of crop area has provided more efficient estimator as compared to RVI.

Another similar study was undertaken during 1996-98 for Rohtak district of Haryana which is major wheat growing area and IRS-1B, LISS-II satellite spectral data have been used to develop improved estimation of wheat crop yield. This study also provided similar results and thus confirmed that spectral data in the form of vegetation indices is suitable as a stratification variable in general crop estimation surveys. From these studies it could also be observed that for obtaining crop yield estimates at district level, the use of satellite data can lead to a significant reduction in number of crop cutting experiments without affecting the efficiency. Alternatively the crop yield estimates can be developed at small administrative units level (say Tehsil/Block) using appropriate small area estimators.

The present study was planned to develop an integrated methodology for crop acreage and crop yield estimation and crop yield modeling using the survey data on crop yield from general crop yield estimation surveys along with the satellite spectral data and farmers personal eye estimates of crop yield. Further small area estimates of crop yield at Tehsil level could also be developed as the post strata based on vegetation vigour are quite homogeneous and cut across small areas.

2.4 Objectives of the study

The basic goal of the proposed study is to develop suitable methodology for application of Remote Sensing Technology for planning of crop surveys, for estimation of crop acreage, estimation of crop yield and crop yield forecasting. To achieve this the following specific objectives are set forth.

1. To develop sampling methodology (involving planning of the surveys and method of estimation) for estimation of crop acreage and crop yield based on the combined use of satellite data and ground survey data on crop yield based on crop cutting experiments.
2. To develop small area estimators of crop yield at Tehsil/Block level.
3. To develop suitable crop yield models using satellite spectral data and farmers personal eye estimates of crop yield.

2.5 Study area

The study was conducted for district Rohtak of Haryana State which is one of the major wheat growing areas having an acreage of more than 66 percent under wheat crop during Rabi season Geographically the district lies between $76^{\circ}15'$ to $77^{\circ}00'$ east longitude and $28^{\circ}20'$ to $29^{\circ}05'$ north latitude. The district is bounded by districts of Jind and Panipat on Northern side, by Hisar district on North-Western side, by Bhiwani district on Western side, by Rewari and Gurgaon districts on Southern side and by Sonapat district and National capital Territory of Delhi on the Eastern side. The total area of the district is 3911 sq. Kms. The district headquarter is located at Rohtak. At present the district consists of four tehsils, namely (i) Rohtak (ii) Mehum (iii) Bahadurgarh and (iv) Jhajjar. The land of the district is generally plane. Three railway lines, one Northern Railway main line from Delhi to Jind another from Rewari to Charkhi Dadri and the third from Rohtak to Gohana and one National

High way No. 10 from Delhi to Hissar passes through the district .The Western Yamuna canal and Bohar and Bhalaut branches of Western Yamuna main canal pass through the district. All these features provide great aid in identification of villages selected for crop cutting experiments and interpretation of satellite spectral data used in the present study.

2.6 Extent of data used in the study

(a) General crop yield estimation survey data

In the present study the yield data for the Rabi season for the years 1995-96 and 1997-98 from general crop estimation surveys based on crop cutting experiments for wheat crop for district Rohtak. Haryana has been used.

b) Satellite data

The satellite data in the study has been used for 1995-96 from IRS-1B, LISS-II of path 30 and Row 47 of 17th February, 1996. The total area of Rohtak district is covered in one sub scene B₂ of 30-47. For 1997-98 IRS-1D data of sensor LISS-III of path 95 and row 51 for Feb. 4th, 1998 has been used. The digital Image processing was carried out in the Remote Sensing Laboratory which has been set up at the Institute under this project using PENTIUM pro hardware system and Digital image processing software ERMAPPER .

A Global Positioning System (GPS) was also used to identify the exact locations of the plots selected for crop cutting experiment for wheat crop in terms of their latitudes/longitudes and also the locations of ground control points(GCP's) which were later used to rectify the raw digital spectral data.

(c) Farmers yield appraisal data

Verma et al (1988) report on evaluation of crop cut methods and farmers reports for estimation of crop production suggested that farmer's eye estimate of crop yield is quite close to the actual yield. Since farmers eye estimate are subjective but can be obtained at a much smaller cost, it was considered prudent to examine the use of farmers estimate as auxiliary variable for improving the efficiency of crop yield models based on satellite spectral data. The data has been collected for the years 1995-96 and 1997-98 for wheat crop yield from the same farmers where fields have been selected for crop cutting experiments in general crop estimation surveys. The data should be collected for eye estimate of yield for only the same fields at the time of maximum crop growth stage where satellite data has highest correlation with yield.

3. ESTIMATION PROCEDURES USING SATELLITE DATA

3.1 Land use/land cover statistics

Land use statistics have been available in India continuously since 1884. Over the years, its geographical coverage has changed, scope expanded and quality improved. At present, land-use statistics are being collected according to the following nine-fold classification:

- (i) Forests
- (ii) Area under non-agricultural uses
- (Hi) Barren and unculturable land
- (iv) Permanent pastures and other grazing lands
- (v) Miscellaneous tree crops and groves not included in sown area
- (vi) Culturable waste
- (vii) Fallow lands other than current fallows
- (viii) Current fallows
- (ix) Net area sown

With the availability of satellite data comprehensive land use/land cover classification system amenable to remote sensing application has been developed by NRSA, Department of Space for countrywide unified application and in order to demonstrate application and feasibility of digital technique of analysis executed a project in 42 districts of the country and prepared a Manual of nation wide land use/land cover mapping using digital techniques(1990) which gave first level classification into 6 classes and second level classification into 22 classes

In order to develop appropriate first level land use statistics for district Rohtak in the two years 1996, 1998 for Rabi crop season, satellite data from IRS-1B and IRS-1D was analyzed using unsupervised clustering technique. This process enables us to separate all land surface areas into distinct and consistent groups. The central mathematical concept in clustering is the computation of central tendency or mean value. The main assumptions of this method are that the Euclidean distance separating the n-points in a p-dimensional space are proportional to the dissimilarities between the objects and that no object belong simultaneously to two clusters.

3.2 Area statistics

India has a long history of developing various kinds of agricultural statistics and the Directorate of Economics in the Ministry of Agriculture is the nodal agency for collection compilation and publication of major agricultural statistics.

Area statistics in respect of cadastrally surveyed known as temporarily settled states, which account for about 86% of the reporting area, are built up as a part of the land records

maintained by the land record agencies at village level (by village Patwaris), In other states (known as permanently settled states) of Kerala, West Bengal and Orissa the estimation of area under different crops is based on a sample of 20% villages.

There is presently a thinking to adopt a survey based on 20% sample approach for area statistics instead of complete enumeration even in the temporarily settled states also. This may lead to reduction in Patwaris work load to a great extent and thus may lead to improved statistic because of reduction of many of the non sampling errors associated with the present complete enumeration approach.

Crop acreage estimation using satellite data

One of the earliest applications of remote sensing for crop acreage has been reported in LACJE and AGRISTARS experiments conducted in the US using land sat data. The first systematic attempt in India directed towards crop inventory through remote sensing technique was carried out under a joint ISRO - ICAR experimental project named Agricultural Resources Inventory and Survey Experiment (ARISE) during 1974-75 and inventory and acreage under various crops were estimated. Further concerted efforts for developing suitable methodology for using Satellite data for crop acreage and crop production and developing acreage estimates for various major crops in the country have been made under the crop Acreage and Production Estimation (CAPE) project of Deptt of Space. With the satellite data predominantly digital classification techniques are used, the most important being the supervised maximum likelihood approach. Under this approach representative training sites of known class are selected. The spectral data of these training sites is used to develop appropriate statistics like mean vector and the variance covariance matrix. Using appropriate classification algorithm each unknown pixel is assigned to any one of the number of classes.

In the present study a Global Positioning System (GPS) was used to identify the selected wheat crop plots for crop cutting experiments in the general crop estimation surveys which are used as training sites. The location of these plots (in terms of longitudes and latitudes) were identified on the spectral imagery and the corresponding spectral signatures were used to develop the appropriate statistics. The supervised maximum likelihood classification algorithm was used to classify all wheat pixels and thus to obtain the crop acreage under wheat.

3.3 Crop yield estimation using satellite data along with crop yield survey data

Crop yield estimation surveys based on crop cutting experiments are conducted throughout the country for obtaining precise estimates of average yield for all major crops. The estimation procedure of estimating average yield and crop production based on data collected from crop cutting experiments under general crop yield estimation surveys are described in Sukhatme and Panse (1951)

The factors like different soil types, agricultural inputs, adoption of improved technology, etc. affect the crop yield and hence cause a lot of variability in the yield even

within a stratum. Since the spectral reflectance is a manifestation of all factors affecting the crop put together, hence a stratification of crop area on the basis of crop vigour as reflected by the spectral data is expected to result in a greater efficiency of the crop yield estimation. Singh *et. al* (1992), Singh and Goyal (1993) and Singh *et.al.*(2000) showed that the stratification based on NDVI improved the efficiency of crop yield estimation considerably.

In the present study **also** it is proposed to use the spectral data for post-stratified estimator of crop yield when the post-stratification of the crop area is done on the basis of vegetation vigour using Spectral Vegetation Indices NDVI and RVI as observed through spectral **reflectance**. For developing a post stratified estimator we have to identify the different sampling units (crop cutting site) belonging to different vegetation strata for which we have to identify sampling units on imageries. But in practice, it is not possible to identify the plots selected for the crop cutting experiments in the selected villages on the topographic maps vis a vis on the satellite data based imagery. However now with the availability the Global Positioning System (GPS) it is possible to obtain the locations of these plots in terms of longitudes and latitudes and then identify them on the spectral imagery.

A topographic map is the best tool to supply ground truth information for visual interpretation and identification of various features on satellite imageries. From these maps locations of villages along with related features like continuous roads, canals railway tracks etc. can be easily identified on FCC's. Survey of India topographical maps of Rohtak district on 1:50,000 scale were used to identify the location of villages selected for the crop yield estimation surveys.

3.3.1. Digital image **processing techniques for crop yield estimation**

The technique of digital image processing needs a computer system with appropriate hardware and software for processing pixel data. The image processing for the present study was undertaken at the recently established Remote Sensing Laboratory equipped with digital Image Processing System at the Institute based on Pentium server and the ER-MAPPER Software. It involves the following steps.

- (a) Generation of district boundary mask, FCC and identification of villages,
 - (b) Identification of crop plots selected for crop cutting experiments using a GPS.
 - (c) Generation of vegetation indices, RVI and NDVI
 - (d) Stratification of imageries using the density slicing technique.
- (a) **Generation of District Boundary Mask, FCC and identification of villages**

District boundary mask was generated with the help of topographic maps of scale 1:250,000 using a **Digitizer** for digitizing the district boundary and then overlaying the digitized map over the satellite image to extract all pixels belonging to the study district. After that a FCC of the district was generated using the band-2 (Green band) band-3 (Red

band) and band-4 (Near infrared band). The cover picture represents the FCC of district Rohtak for Rabi season of 1996 based on IRS-1B LISS-II dated Feb. 17, 1996. Further the topographic maps of scale 1:50,000 containing the identified locations of the villages selected in the survey were used as base-material for identification of villages on the FCC. All the village locations in different segments of the FCC's were identified, by comparing the features seen on FCC imageries to the identifiable features like surrounding roads, canals or water logged area etc. available on topographic maps. The co-ordinates (scan line and column number of each sampled village were recorded to be used to identify the location of these villages on the Normalized Difference Vegetation Index (NDVI) imagery and the Ratio Vegetation Index (RVI) imagery.

b) Identification of crop plots selected for crop cutting experiments using GPS

A Global Positioning System (GPS) was used to locate the crop plots selected for crop cutting experiments. The GPS was taken to the plots and locations of the plots (longitude and latitude) were recorded. These locations were identified in the selected villages already earmarked on the FCC,s. However, GPS has the limitations that recordings given by a GPS have some deviation from actual location varying from 10 to 50 meters. But in the present case where crop plot is selected from a large field and also vast continuous areas are generally under the same crop, this may not affect the result much.

c) Generation of Vegetation Indices RVI and NDVI

Spectral response characteristics of healthy vegetation, can easily be characterised in the different parts of the electromagnetic spectrum. To further enhance the discrimination between different spectral vegetation classes, computation of different vegetation indices using infrared and red band data in the electromagnetic spectrum, for describing the crop growth conditions, are commonly used. Two most commonly used vegetation indices are:

(i) The Normalized Difference Vegetation Index (NDVI) defined as

$$NDVI = \frac{IR - R}{IR + R}, \text{ and}$$

(ii) The Ratio Vegetation Indices (RVI) defined as

$$RVI = \frac{IR}{R}$$

Where IR and R refer to radiance in infrared (band-4) and red (band-3) bands of the satellite. These two indices have been used in the present study to generate the index images for post-stratification of the study area on the basis of vegetation vigour.

d) Stratification of imageries using density slicing technique

The concept of density slicing was used to divide the RVI and NDVI imageries into different vegetation classes. The RVI and NDVI grey level values were linearly stretched over the total range (0-255) of grey level values and were divided into 3 classes named as

- (i) Non-vegetation class
- (ii) Average Vegetation class, and
- (iii) High Vegetation class

Assigning different colours to different class range values, the stratified imageries were generated and area falling under different strata could be obtained which have been used as the strata weights. The magenta colour was assigned to high-vegetation class, green to average vegetation class and white was assigned to non-vegetation class. Figure 2 and 3 represent the stratified imagery based on RVI and NDVI respectively for the Rabi 1995-96 based on IRS-1B LISS-II data of dated Feb. 17, 1996.

3.3.2 Post stratified Estimator of crop yield

In case of yield estimation surveys, the original stratification is only based on geographical considerations and may not be much effective in terms of making the strata more homogeneous. As such in the present study for simplification of results the original stratification has not been taken into consideration.

To obtain the post-stratified estimator, let us assume that n villages selected in the sample have been post-stratified into L' strata such that n'_k villages fall in the k-th post-stratum. Let Y_{kij} denote the yield for the j-th field in the i-th village of the k-th post stratum.

The sample mean for the k-th post stratum can be defined as

$$\bar{y}_k = \frac{\sum_{i=1}^{n'_k} \sum_{j=1}^{m_k} Y_{kij}}{m_k} \quad \dots\dots(3.1)$$

where m_k is total number of field experiments falling in the k-th post-strata.

Now the post stratified estimator of district average yield can be given by

$$\bar{y} = \frac{\sum_{k=1}^{L'} A_k \bar{y}_k}{\sum_{k=1}^{L'} W_k} = \sum_{k=1}^{L'} W_k \bar{y}_k \quad \dots\dots(3.2)$$

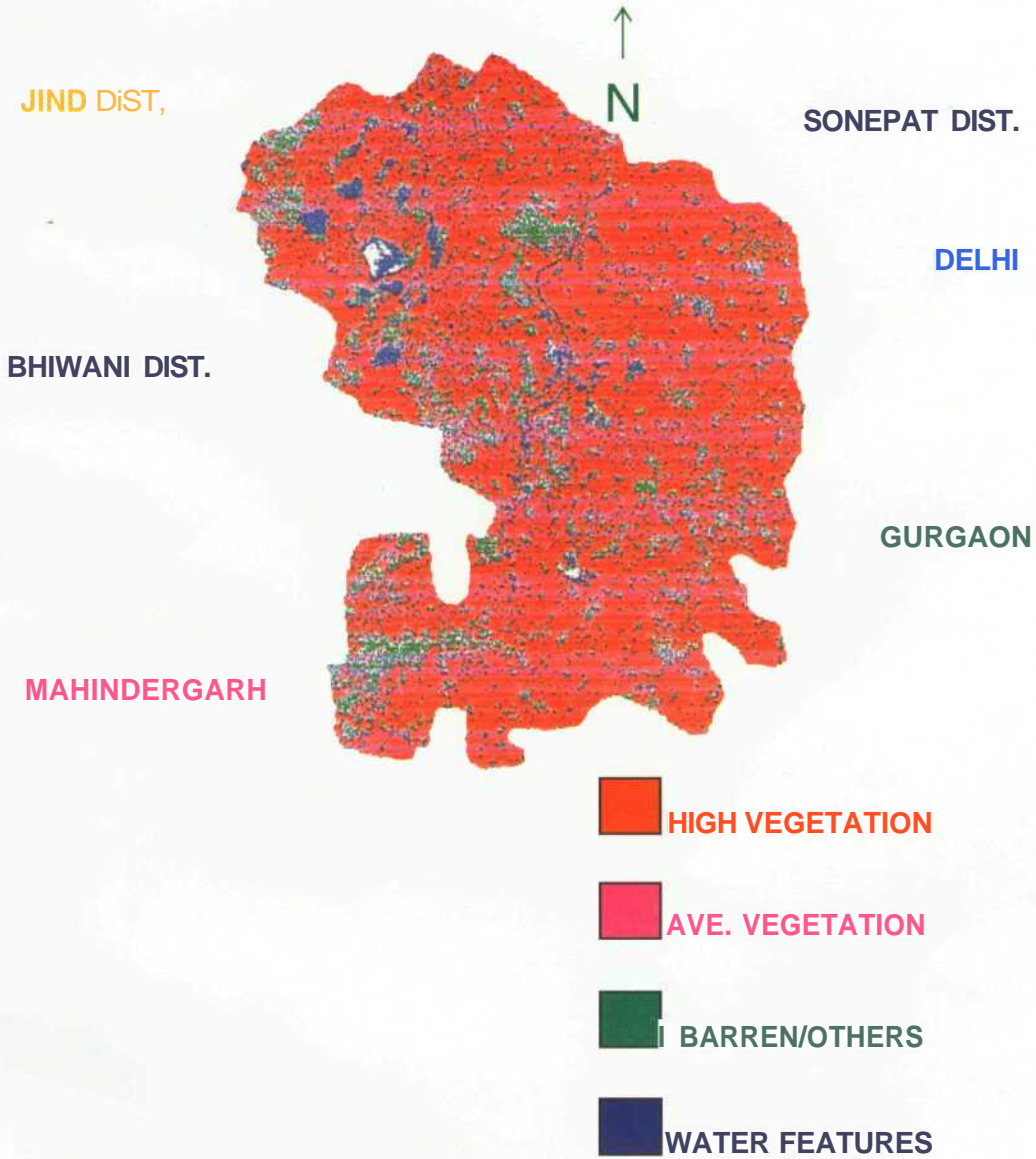
Where A_k denote the area under crop in the k-th post-stratum. $W_k = \frac{A_k}{\sum A_k}$

Ignoring the pre stratification and also ignoring the contribution to sampling error due to post stratification the Variance and the Estimator of variance of y can be obtained easily given by

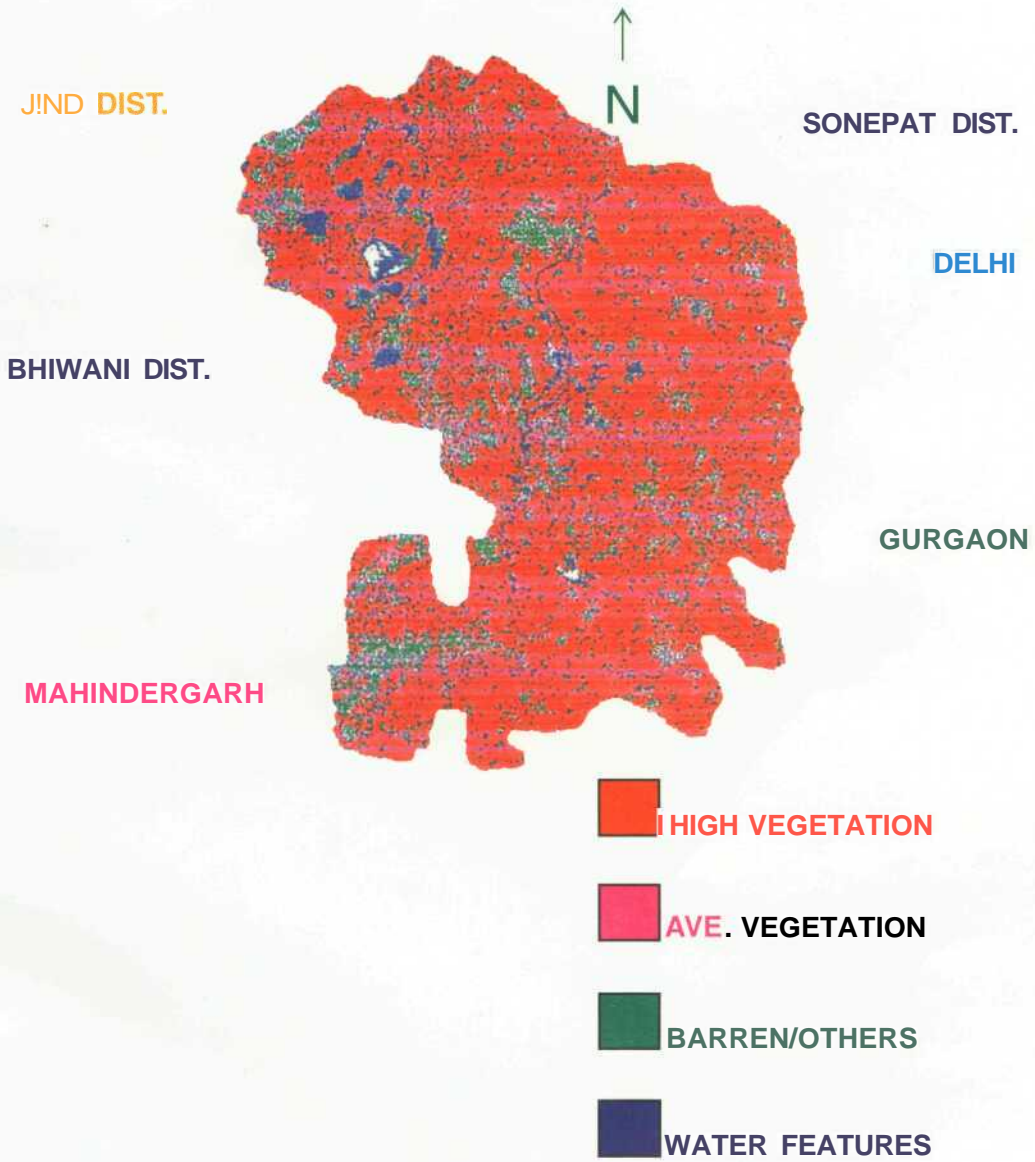
$$V(\bar{y}) = \sum W_k^2 V(\bar{y}_k) \quad \dots\dots(3.3)$$

$$\text{and } \hat{V}(\bar{y}) = \sum W_k^2 \hat{V}(\bar{y}_k) \quad \dots\dots(3.4)$$

RVI [4/3] IRS 1B, 17 FEB, 1996
ROHTAK DISTRICT



NDVI [4-3/4+3] IRS 1B, 17 FEB, 1996 ROHTAK DISTRICT



3.4 Crop yield modeling using satellite remote sensing

Crop yield is influenced by a large number of factors related to soil, weather, agro climatic factors, management practices etc. Satellite data is integrated manifestation of effects of all these factors on the crop growth and hence can provide immense potential for use in crop yield modeling.

Several approaches in crop yield modeling using satellite remote sensing data have recently been developed like spectral yield models using spectral vegetation indices or spectral growth profile, meteorological yield models using meteorological data pertaining to some significant crop phenological stages in the form of some indices. Agriculture and climate are closely inter linked in the sense that crop growth development and production are greatly affected by variation in agro-meteorological parameters during crop growth period. In this modeling approach remote sensing derived SVI is coupled with meteorological indices and multiple regression model is developed. A large number of meteorological indices like Growing Degraded Day (GDD) mean Temperature (T_{mean}), Rainfall index (RI), Crop Water Stress Index (CWSI) etc are being used in agromet spectral yield models. However the difficulty and delay in availability of weather parameters make this approach less attractive. Studies have also been conducted for developing integrated or combined models incorporating parameters from diverse sources or combining two or more independent forecasting models.

3.4.1 Spectral yield model

These are empirical models which directly relates the crop yield to the multi-spectral satellite data or derived parameters in the form of spectral vegetation indices (SVI). In this procedure SVI at the time of maximum vegetation growth stage of the crop is related to final crop yield through regression techniques and pre harvest crop yield is forecasted. In India district level yields of major crops like wheat, paddy, sorghum etc. have been developed under crop acreage and production estimation (CAPE) project undertaken by National Remote Sensing Agency (NRSA), Hyderabad, Deptt. of Space. However these models could explain about 60% variation in yield and hence are not very efficient.

In the present situation of post-stratification based on vegetation indices the regression coefficient of y (yield) on the spectral response parameter (x) for the h-th post stratum may be of interest and this may lead to improvement in efficiency of the regression model.

For this we define the regression model as

$$y_{hi} = \beta_0 + \beta_h x_{hi} + e_{hi} \quad (6.2.1)$$

where

y_{hi} and x_{hi} denote the yield and the spectral response of the i^{th} unit in the h^{th} stratum and m_h is number of sampling units in the h -th stratum, $h = 1, 2, \dots, L$, $i = 1, 2, \dots, m_h$

$$\begin{aligned} E(e_{hi}) &= 0 \\ V(e_{hi}) &= \sigma_h^2 \\ \text{Cov}(e_{hi}, e_{hj}) &= 0 \text{ for } i \neq j \text{ and} \end{aligned}$$

Here β_h may be estimated separately for each stratum. Under the given assumption, the Ordinary Least Square (OLS) estimator of the regression coefficient f_{hi} for the h^{th} stratum may be given by $\hat{\beta}_h$ as

$$\hat{\beta}_h \approx \frac{\sum_{i=1}^{m_h} (x_{hi} - \bar{x}_h)(y_{hi} - \bar{y}_h)}{\sum_{i=1}^{m_h} (x_{hi} - \bar{x}_h)^2}$$

with $\bar{x}_h = \sum_{i=1}^{m_h} x_{hi} / m_h$, $\bar{y}_h = \sum_{i=1}^{m_h} y_{hi} / m_h$, and

The variance of $\hat{\beta}_h$ is given by

$$V(\hat{\beta}_h) = \frac{\sigma_h^2}{\sum_{i=1}^{m_h} (x_{hi} - \bar{x}_h)^2} \quad \text{and}$$

An the unbiased estimator of σ_h^2 is given by

$$s_h^2 = \frac{\sum_{i=1}^{m_h} (y_{hi} - \hat{\beta}_0 - \hat{\beta}_h x_{hi})^2}{(m_h - 2)}$$

The fitted regression equation can be used to predict the value of yield corresponding to a chosen value x'_{hi} of the spectral response as given by

$$\hat{y}_{hi} = \hat{\beta}_0 + \hat{\beta}_h x'_{hi}$$

The variance of the predicted value is given by

$$V(\hat{y}_h) = V(\bar{y}_h) + (x'_h - \bar{x}_h)^2 V(\hat{\beta}_h)$$

$$= \frac{\sigma_h^2}{m_h} + \frac{(x'_h - \bar{x}_h)^2 \sigma_h^2}{\sum_{i=1}^{m_h} (x'_{hi} - \bar{x}_h)^2}$$

The regression coefficient estimator for the population may be defined as

$$\hat{\beta} = \sum_{h=1}^L w_h \hat{\beta}_h$$

and the variance of the estimator of the regression coefficient is given by

$$V(\hat{\beta}) = \sum_{h=1}^L w_h^2 V(\hat{\beta}_h) = \sigma^2$$

Where w_h is the weight of the h^{th} stratum.

3.4.2 Integrated yield model using spectral data and farmers eye estimate of crop yield

Most of the crop yield models developed so far could not be adopted in practice either because of delay in the availability of data on different variables to be used in the model or the high cost in collecting the data and in analysing the results.

For any operational yield model to be successful for adoption it is necessary that data should be available much before the harvest of the crop and it should be cost effective. Spectral data in the form of vegetation indices have proved to be very useful variable for explaining variability of the crop yield which can be early available for use in yield forecasting models. In a recent study for 'evaluation of crop cutting methods and farmers reports for estimating crop production' undertaken at Longacre Agricultural Development Centre UK, it has been shown that farmers eye estimates are remarkably close to actual production figures. But, eye estimates being subjective and amenable to several non-sampling errors, it is desirable that these estimates are not used directly for estimation of crop yield. However, this information can be used as auxiliary variable along with the spectral vegetation indices to improve the efficiency of the crop yield models. An earlier such attempt on using eye appraisal of crop yield of a large number of sample fields as auxiliary information had been made by Panse, Rajgopalan and Pillai (1966).

In the present study, therefore suitable models using spectral vegetation indices in the form of NDVI. and farmers eye estimate as explanatory variables in the regression model have been developed for improved crop yield forecasting models. Both these variables can be easily obtained at the time of maximum growth of crop and can prove very effective for developing suitable yield forecasting models.

4. SMALL AREA ESTIMATION

Issue of small area estimation has gained importance in view of growing needs of micro level planning. The advances in computer facilities have provided convenient tools for many theoretical developments for providing small area estimates. The small area estimation techniques make use of information from other available sources and borrow strength from related or similar areas through explicit and implicit models that connect this small area via supplementary data.

Most of the small area estimation techniques in the early stages were developed in the context of demographic studies. Purcell and Kish (1979) categorised these areas under the general heading of Symptomatic Accounting Technique (SAT). Gonzales (1973) described a small area estimation technique well known as synthetic estimator. In this method an unbiased estimate is obtained from a sample survey for a larger area and this estimate is used to derive estimates for sub areas having the same characteristics as the larger area. Sarndal (1984) interprets the word synthetic in two different senses. In the first interpretation synthesizing is commonly used for combining different parts to get the whole and thus the synthetic estimate is a combination of sub estimates. The other interpretation of synthetic estimator is the artificial nature of these estimates. The estimates developed for the larger area are scaled down to smaller areas on the basis of certain model assumptions which assume that the relation for the study character as well as for the auxiliary characters between the larger area and the small area remain the same. Considering the need for developing suitable methodologies for meeting the need of data requirements for key characteristics recently a working group of small area has been set up by the Central Statistical Organization(CSO) of the Government of India.

4.1 Small Area Estimation of Crop Yield at Tehsil/Block level

General crop estimation surveys have been designed to obtain crop yield estimates at the district level. However with increasing emphasis on micro level planning, the estimates at lower administrative unit level like tehsil or block level are needed. Since it is not possible and desirable to further increase the number of crop cutting experiments, it is desirable to make use of the satellite data and the small area statistics techniques to develop reliable crop yield estimates at tehsil and block level.

Consider the population (a district) consisting of T small areas (Tehsil/Block). Let the district area be divided into V post strata representing crop condition like very good crop, average crop, poor crop, no crop etc. based on the vegetation indices derived from the satellite spectral data. The crop within these post strata is homogeneous in respect of the character under study (the crop yield) and the boundaries of these post strata cut across the small areas. Hence it can be easily assumed that the units within a small area belonging to particular post strata will have the same characteristics as the units belonging to that particular post strata irrespective of the small area.

In order to develop crop yield estimates at tehsil level from general crop yield estimation surveys based on crop cutting experiments we propose two estimators namely

- (i) The Direct estimator and
- (ii) The Synthetic estimator

These estimators make use of available information on crop yield and also the information of crop acreage for all the post strata which overlap the tehsil. It has been seen earlier from the results that stratification based on NDVI provides more efficient estimates of crop yield for the district as whole. Therefore, for small area estimation only NDVI has been used to develop post strata.

4.1.1 Direct Estimator

Let y_{tvi} and x_{tvi} denote the crop yield and the crop acreage for the i -th plot in the v -th post strata of the t -th small area (tehsil). Let \bar{y}_{tv} and \bar{x}_{tv} denote estimators of the character under study (crop yield) and the auxiliary character (crop acreage) from the t -th small area (tehsil) and v -th post strata (based on vegetation vigour). Further let S_{tv} denote the set. of sample observations belonging to the t -th tehsil in the v -th post strata. If all S_{tv} 's are non empty then an unbiased post stratified estimator known as the direct estimator for crop yield may be obtained as

$$\bar{y}_{dt} = \sum_v w_{tv} \bar{y}_{tv} \quad \dots\dots(4.1)$$

where, $y_{tv} = \frac{1}{n_{tv}} \sum_1^{n_{tv}} y_{tvi}$ is the average yield for tv -th cell

$$W_{tv} = \frac{X_{tv} X_{00}}{X_{t0} X_{0v}}$$

X_{tv} = the crop acreage for tv -th cell

$X_{t0} = \frac{\sum_v X_{tv}}$ is the crop area for the t -th tehsil

$X_{0v} = \sum_t x_{tv}$ in the crop area for the v -th post stratum, and

$x_{00} = \sum_t \sum_v x_{tv}$, total crop acreage in the district.

The approximate estimate of variance of y_{dt} can be written as

$$\hat{V}(\bar{y}_{dt}) = \sum_v W_{tv}^2 \hat{V}(\bar{y}_{tv}) \quad \dots(4.2)$$

where

$$\hat{V}(\bar{y}_{tv}) = \frac{s_{tv}^2}{n_{tv}} \text{ and } s_{tv}^2 = \frac{1}{n_{tv} - 1} \sum_I^{n_{tv}} (y_{tvi} - \bar{y}_{tv})^2$$

4.1.2. Synthetic Estimator

The direct estimator is based on only the number of crop cutting experiments belonging to tv - th cell i.e. t -th tehsil in the post strata which is quite small and hence the estimator will not be quite efficient.

To improve the efficiency of the direct estimator a synthetic estimator is proposed which make use of the information from the whole sample given by y_{st}

$$\bar{y}_{st} = \sum_v w'_v \bar{y}_{ov} \text{ where } V_{..} = \frac{X}{X_{to}} \quad \dots(4.3)$$

and y_{ov} is the average crop yield for the v^t post-stratum given by

$$\bar{y}_{ov} = \sum_t w''_t \bar{y}_{tv} \text{ where } w''_t = \frac{X_{..}}{X_{ov}}$$

The estimator of variance of y_{st} can be approximately written as

$$\begin{aligned} \hat{V}(\bar{y}_{st}) &= \sum_v w'^2_v \hat{V}(\bar{y}_{ov}) \\ &= \sum_v w'^2_v \sum_t W_{tv}^2 \hat{V}(\bar{y}_{tv}) \end{aligned} \quad \dots(4.4)$$

since, sample in each tehsil has been selected independently,

5. RESULTS AND DISCUSSION

District Rohtak in Haryana State was taken up for the study of various crop statistics for wheat crop. The district is one of the major wheat growing areas having area of acreage more than 66% under wheat. The main findings of the study are divided into the following five sections.

- a) Land use/Land cover statistics
- b) Crop Acreage Estimation
- c) Crop yield Estimation
- d) Crop yield Forecasting and
- e) Small area estimation of crop yield at Block level

5.1 Land use/Land cover statistics

First level land cover statistics for district Rohtak, Haryana for the Rabi season for 1995-96 and 1997-98 have been developed using the satellite spectral data of Feb., 17, 1996 from IRS-1B LISS-II and of 4th February, 1998 from IRS-1D LISS-III. The results are given in table 5.1.1.

Table 5.1.1 Land use classification of district Rohtak based on satellite data for the Rabi season for 1995-1996 and 1997-1998.

S.No	LAND USE CLASSIFICATION	Area in square kilometers based on	
		IRS-1B, LISS-II February, 1996	IRS-1D, LISS-III February, 1998
1.	AGRICULTURAL LAND	2471.943	2088.198
	i) High veg.	1240.109	1199.159
	ii) Av veg	1231.834	889.039
2.	BUILT UP LAND	513.213	516.997
3.	WATER BODIES	116.830	290.621
4.	SALINE SOILS/WASTE LAND	134.027	-
5.	FALLOW WET LAND	-	267.481
6.	OTHERS	675.427	748.143
	TOTAL	3911.440	3911.440

By comparing the results of land use statistics for 1996 and 1998 it could be seen that area under crops was around 400 square kilometers less in rabi 1998 compared to rabi 1996 of which 267 km area were classified as fallow wet land during 1998 while there was none in 1996. Also water bodies and saline soils/waste lands could be differentiated during 1996 but during 1998 all saline soils or waste lands also being under water or wet could be classified along with water bodies.

Thus the major land use classification for district Rohtak could be achieved by using the supervised classification technique. The study also highlighted that extent of floods could be obtained using satellite data for two separate years of the same season and using the digital analysis technique which can prove a great help in planning suitable flood relief measures and suitable preventive measures.

5.2 Crop Acreage Estimation:

Haryana is a temporary settled state and entire area has been cadastrally surveyed. The estimates of crop acreage at district level are obtained through complete enumeration of all the villages in the district by the revenue agencies of the state. The information of wheat acreage is collected during the growth period of the crop and supplied at the district H.Q. well in advance before the crop is actually harvested.

In the present study an attempt has been made to obtain the District and Tehsil level estimates of area under wheat crop using the IRS-1D, LISS-III satellite data for 1998. The estimates of area under wheat using remote sensing satellite data and the actual area under wheat as reported by the state government are given in the table 5.2.1.

Table 5.2.1 Crop acreage under wheat at Tehsil level as obtained from satellite data for district Rohtak for Rabi 1997-1998 in (oo ha)

S.No.	Tehsil	Estimate of area under wheat using satellite data.	Area under wheat as reported by State Govt.
1.	Rohtak	609.25	553.00
2.	Meham	282.99	185.00
3.	Bahadurgarh	193.00	284.00
4.	Jhajjar	638.09	615.00
	Overall for the district	1613.33	1637.00

It is seen from here that the deviation in the estimate of acreage estimate at District level is about (1.5%). But Tehsil level deviations are quite large.

5.3 Crop yield estimation

An earlier study (Singh et.al. (1999)) was undertaken in Rohtak district of Haryana to study wheat crop yield estimation for the year 1995-96 using crop yield data obtained from general crop estimation surveys for 1995-96 and IRS-1B, LISS—II satellite spectral data for 17th February, 1996. The results obtained from this study showed considerable gain in efficiency of the post-stratified estimator over the usual estimator of crop yield at district level. The study also gave small area estimators of crop yield at tehsil level which also had standard error of the order of less than 5 per cent. The results of this study for Rabi 1995-96 are also presented here for the sake of continuity in tables 5.3.1 to 5.3.3.

Keeping in view the higher resolution of IRS-1D, LISS-III the present study was taken up to confirm the efficiency of the post stratified estimator of crop yield at district level and to extend the methodology of crop yield estimation to Block level. The crop yield data for Rabi 1997-98 was obtained from the general crop yield estimation surveys based on crop cutting experiment and the satellite spectral data of IRS-1D, LISS-III was obtained for 4th February, 1998.

For digital analysis of satellite data, first district boundary mask was generated using a Digitizer and toposheets of the district. The digitized map was overlaid over the satellite image to extract all the pixels belonging to the district. FCC was generated using band 1 (green), band 2 (red) and band 3 (near infrared) and all the 72 villages selected for crop cutting experiments were identified on the FCC's. Normalized Difference Vegetation Index (NDVI) and Ratio Vegetation Index (RVI) were generated from the spectral data using band 2 and band 3 of the LISS-III sensor. The total range of RVI and NDVI values were divided into three classes using the density slicing technique to form three post strata based on the vegetation vigour namely

- (i) Strata having high vegetation growth implying very good crop yield.
- (ii) Strata having average vegetation growth implying average crop yield and
- (iii) Strata consisting of non-vegetation area.

The three classes were assigned different colours and the corresponding stratified RVI and NDVI imageries were generated. The range of grey level values and the corresponding strata weights in terms of area in the above three strata based on NDVI and RVI are given in table 5.3.4.

For estimation of crop yield at district level following three different estimates have been developed.

1. Usual estimate of crop yield based on crop cutting experiments without using the satellite data.
2. Post-stratified estimate of crop yield using RVI for stratification
3. Post-stratified estimate of crop yield using NDVI for stratification

The relative efficiency of post-stratified estimators using satellite data has been obtained as compared to the usual estimator of crop yield. The results are given in table 5.3.5.

From here it is seen that the post-stratified estimators are highly efficient compared to the usual estimator. Further post-stratified estimator based on NDVI is better compared to post-stratified estimator based on RVI. The relative efficiency of the two estimators compared to the usual estimator are 101% and 66% respectively. As expected the results are quite similar (even far better) to the results obtained for Rabi 1995-96 by Singh et.al.(1999)

5.4. Crop yield modeling

It has been shown in several studies that crop yield is having high correlation with spectral vegetation indices RVI and NDVI. But when used for crop yield forecasting spectral indices alone have been able to explain only about 50% to 60% variability in crop yield and hence it is important to include some other variables to improve the yield models. The use of agro-climatic variables and biometrical variables has been examined by various workers. However all these models have their own limitations and have not been used operationally.

Recently in a study at Longacre Research centre it has been seen that farmers eye estimate of crop yield are quite close to the actual yield but the enquiry method being subject to several limitations cannot be used directly for forecasting crop yield. Therefore in the present study the farmers eye estimate of crop yield corresponding to the crop plots selected for crop cutting experiment has been used as an auxiliary variable along with the vegetation indices for improving the crop yield forecasting models.

The yield data pertains to wheat crop yield data for district Rohtak for the year 1995-96 based on crop cutting experiments. Spectral data in the form of vegetation indices RVI and NDVI has been obtained from IRS 1B-LISS II data dated Feb., 1996 for the region. The farmer's eye estimate is obtained from the selected farmers for the fields in which crop cutting experiments were conducted.

The usual linear regression based models were developed with the crop yield (y) as the dependent variable and three independent variables, namely RVI (x_1), NDVI (x_2) and the farmer's eye estimate of crop yield of the corresponding plot (x_3). The model has been developed using the data for the Rabi 1995-96.

This model has been used to forecast the crop yield for Rabi 1997-98 using the independent variables for 1997-98. The results and predicted value of crop yield using different independent variables independently as well as together are given in table 5.4.1.

From this table it is seen that R^2 value is 0.45 and 0.54 respectively when only RVI and NDVI alone are used in the model and it increases to 0.59 when both these variables are used. However the R^2 value is 0.86 when only farmers eye estimate is used as the

explanatory variable and the R^2 value increases to around 0.90 when it is used along with RVI or along with NDVI or along with both RVI and NDVI together. The deviation of the predicted yield from the actual yield is very low. In almost all the cases it is less than 2%. The standard error of the predicted value is also small in all the cases and as expected it is smallest when all the three variables are used together but it is not much different in the case when only farmers eye estimate and NDVI is used.

* The results suggest that a reliable and timely forecast may be obtained using NDVI from the satellite spectral data along with the farmers eye estimate as the two explanatory variables. These both variables can be obtained at the time of maximum vigour of the crop and hence objective reliable forecast may be made available about 6-8 weeks before actual harvest of the crop.

5.5 Small area estimation of crop yield at Tehsii and Block level

Singh et. al. (1999) and Singh et. al. (2000) presented the methodology for small area estimation of crop yield at Tehsii level using the yield data from general crop estimation surveys and satellite data in the form of vegetation indices for stratification of crop area in to homogeneous crop growth condition strata so that appropriate small area statistics can be used.

The crop area had been post-stratified into two homogeneous crop growth conditions based on NDVI namely (i) high crop growth and (ii) average crop growth condition. These strata cut across the small area and hence provide useful information for developing small area estimates. Two small area estimators of crop yield at Tehsii level had been developed namely

1. A Direct estimate of crop yield and
2. A Synthetic estimate of crop yield

The results of this study for small area estimates at tehsil level using Direct estimator and Synthetic estimator for district Rohtak for Rabi 1995-96 are given in Table 5.3.3.

In the present study also the small area estimates have first been developed at Tehsii level for Rabi 1997-98 and the results are presented in Table 5.5.1.

Further the study has been extended to obtain the small area estimates at Block level and since the selection of villages and the corresponding crop cutting sites is made independently at block level there is no problem in extending the study to obtain Block level estimates. Again the two estimators i.e. the Direct estimator and the Synthetic estimator have been obtained for Block level yield estimates for Rabi 1997-98. The block level results are given in Table 5.5.2.

The results show that synthetic estimator is better than the Direct estimator both at Tehsii as well as block level (as expected). The standard error of the synthetic estimator at Tehsii level varies from 5% to 8%. The standard error at block level also is less than 10% in most of the blocks

Table : 5.3.1 Distribution of grey values and area of different strata based on RVI and NDVI for District Rohtak and for different Tehsils in the district for Rabi 1995-96.

Sl.No.	Stratum/Ve g classes	NDVI			RVI		
		Range of grey values	No.of villages selected	Area Sq.Km.	Range of grey values	No.of villages selected	Area Sq.Km.
District Rohtak (Complete)							
1.	Non-Vegetation	0-167	-	.934.986	0.67	-	851.412
2.	Av.-vegetation	168-217	39	1263.129	68-187	42	1360.522
3.	High-vegetation	218-255	36	1240.109	188-255	33	1226.468
Tehsil Level Estimates							
Tehsil Maham							
I	Non-vegetation	0-161	-	127.562	0-67	-	120.562
2.	Av.-vegetation	162-220	4	102.327	68-190	5	140.200
3.	High-vegetation	221-255	5	211.115	191-255	4	100.242
Tehsil Jhajjar							
1.	Non-vegetation	0-171	-	245.788	0-61	-	210.700
2.	Av.-vegetation	172-219	17	680.714	62-185	17	620.000
3.	High-vegetation	220-255	11	439.093	186-255	11	474.825
Tehsil Bahadurgarh							
1.	Non-vegetation	0-157	-	199.189	0-64	-	170.150
2.	v Vegetation	158-215	6	156.828	65-177	5	160.120
3.	High Vegetation	216-255	4	100.615	178-255	5	126.362
Tehsil Rohtak							
1.	Non-vegetation	0-165	-	362.447	0-70	-	350.00
2.	Av.-vegetation	166-208	12	323.260	71-180	15	380.202
3.	High-vegetation	209-255	16	489.286	181-255	13	444.791

Table 5.3.2 : Wheat crop yield Estimation using stratification based on NDVI and RVI for district Rohtak for Rabi 1995-96.

	Ave. Yield (Qtls./Hect.)	S.E.	%SE	RE
Usual Estr	35.92	0.7594	2.11	1.00
Post Str. Estr, based on NDVI	33.66	0.5324	1.58	1.42
Post Str Estr based on RVI	34.05	0.5753	1.69	1.28

Table S.3.3 : Tehsil level wheat crop yield estimation using stratification based on NDVI for district Rohtak for Rabi 1995-96.

Tehsil	Total No. Villages	No. of villages selected	Estimator	Average Yield (Qtls/Hect.)	S.E.	%S.E.
Rohtak	115	28	Direct	37.14	0.8732	2.35
			Synthetic	34.34	0.5568	1.59
Meham	26	9	Direct	39.18	1.3466	3.44
			Synthetic	34.78	0.9216	2.65
B. Garh	55	10	Direct	30.52	2.0310	6.65
			Synthetic	32.18	1.2374	3.85
Jhajjar	190	28	Direct	33.56	0.8824	2.63
			Synthetic	33.22	0.6692	2.01

Table 5.3.4 Distribution of grey values and area of different strata based on RVI and NDVI for District Rohtak for Rabi 1997-98.

SI. No	Stratum/ Veg. classes	NDVI			RVI		
		Range of grey values	No.of villages selected	Area Sq.Km.	Range of grey values	No.of villages selected	Area Sq.Km.
District Rohtak (Complete)							
1.	Non Vegetation	0-167	-	1525.38	0-78	-	1440.07
2.	Av.-vegetation	168-214	30	861.62	79-186	31	908.64
3.	High-vegetation	215-255	42	1121.85	187-255	41	1160.14
Tehsil Level Estimates							
Tehsil Maham							
1.	Non-vegetation	0-165	-	135.11	0-80	-	125.00
2.	Av.-vegetation	166-215	4	100.23	81-195	5	121.95
3.	High-vegetation	216-255	5	196.86	196-255	4	185.25
Tehsil Jhajjar							
1.	Non-vegetation	0-162	-	578.20	0-70	-	560.00
2.	Av.-vegetation	163-217	11	345.47	71-180	10	352.79
3.	High-vegetation	218-255	15	426.06	181-255	16	436.94
Tehsil Bahadurgarh							
1.	Non-vegetation	0-154	-	217.07	0-67	-	200.05
2.	Av. Vegetation	155-210	5	114.00	68-177	4	131.95
3.	High Vegetation	211-255	5	150.93	178-255	6	150.00
Tehsil Rohtak							
1.	Non-vegetation	0-162	-	595.00	0-85	-	555.02
2.	Av.-vegetation	163-217	10	301.92	86-192	12	300.85
3.	High-vegetation	218-255	17	348.00	193-255	15	389.05

Table 5.3.5 : Wheat crop yield estimation using stratification based on NDVI and RVI, for district Rohtak for Rabi 1997-98.

	Average yield (Qtls/Hect.)	S.E.	% S.E.	R.E.
Usual Estimator	36.17	3.5610	9.48	1.00
Post Stratified based on NDVI	36.36	1.7124	4.71	2.01
Post Stratified based on RVI	36.77	2.0985	5.90	1.66

Table 5.4.1 **Wheat crop yield forecasting model using RVI (x_1), NDVI (x_2) and the farmers eye Estimate (x_3) as independent variables for district Rohtak for forecasting crop yield for Rabi 1997-98. (using the model based on data for Rabi 1995-96).**

	R^2	a	b	Predicted value(Q/hac.) \hat{y}	%S.E.	Percentage of Deviation
$y = a + b_1x_1$	0.451596 (3.102028)	3.3445 (0.5724)	4.251948 (1.998869)	35.86 (4.8568)	13.5434	0.1653
$y = a + b_1x_2$	0.543511 (2.830157)	-6.18267 (2.721178)	44.87417 (5.024239)	35.86 (4.7004)	13.1071	0.1652
$y = a + b_1x_3$	0.867496 (1.124314)	2.036448 (1.52888)	0.216212 (0.021125)	34.85 (2.1200)	6.0828	3.0619
$y = a + b_1x_1 + b_2x_2$	0.59259 (2.03613)	-3.798801 (8.212242)	$b_1 = -2.530032$ (5.094414) $b_2 = 55.993241$ (46.451635)	35.22 (5.2418)	14.8829	1.9872
$y = a + b_1x_1 + b_2x_3$	0.90009 (1.00829)	0.785252 (1.479713)	$b_1 = 1.146923$ (0.510314) $b_2 = 0.179795$ (0.024910)	34.86 (1.8352)	5.2647	3.0424
$y = a + b_1x_2 + b_2x_3$	0.90345 (0.99122)	-1.049277 (1.865161)	$b_1 = 11.219523$ (1.679861) $b_2 = 0.175972$ (0.025054)	34.86 (1.8046)	5.1771	3.0430
$y = a + b_1x_1 + b_2x_2 + b_3x_3$	0.90406 (1.02274)	-2.144346 (4.132285)	$b_1 = -0.770271$ (2.572191) $b_2 = 18.257945$ (23.994514) $b_3 = 0.175182$ (0.025984)	34.86 (1.7992)	5.1613	3.0433

Actual crop yield for Rabi 1997-98=35.92 (Q/hac)
(Figures in braces give the corresponding standard error)

Table: 5.5.1 Tehsil level wheat crop yield estimation using stratification based on NDVI for district Rohtak for Rabi 1997-98.

Sl.	Tehsil	Total no. of villages	No. of villages selected	Estimator	Average yield (Qts/Hect.)	S.E.	% S.E.
1.	Rohtak	115	27	Direct	33.47	3.1925	9.20
				Synthetic	31.87	1.7805	5.58
2.	Mehan	26	9	Direct	40.32	3.2570	8.07
				Synthetic	40.32	3.2570	8.07
3.	Bahadurgarh	55	10	Direct	34.92	2.4785	7.10
				Synthetic	34.92	2.4785	7.10
4.	Jhajjar	190	26	Direct	38.36	3.1218	8.14
				Synthetic	38.05	1.5542	4.06

Table : 5.5.2 Block level wheat crop yield estimation using stratification based on NDVI for district Rohtak for Rabi 1997-98.

S.No.	Block	Total no.villages	No. of villages selected	Estimator	Ave. yield (Qtls./Hect)	S.E.	% S.E.
1	Rohtak	34	7	Direct	30.96	3.0733	9.93
				Synthetic	31.74	1.7569	5.53
2	Lakhan Majra	17	7	Direct	35.35	2.5785	7.29
				Synthetic	32.11	1.7461	5.44
3	Kalanaur	41	6	Direct	34.02	3.8645	11.36
				Synthetic	32.02	1.7376	5.43
4	Sampla	23	7	Direct	32.69	2.0441	6.06
				Synthetic	32.11	1.6314	5.08
5	Meham	26	9	Direct	40.32	3.2570	8.07
				Synthetic	40.32	3.2570	8.07
6	Bahadurgarh	55	10	Direct	34.92	2.4785	7.10
				Synthetic	34.92	2.4785	7.10
7	Matenhahi	41	6	Direct	40.56	3.5513	8.75
				Synthetic	39.31	1.5020	3.82
8	Sahlawas	43	7	Direct	38.99	1.7857	4.76
				Synthetic	38.17	1.5319	4.01
9	Beri	33	7	Direct	36.25	4.2946	11.85
				Synthetic	38.28	1.5220	3.96
10	Jhajjar	73	6	Direct	36.89	2.4965	6.77
				Synthetic	38.39	1.5132	3.94

SUMMARY

With the advent of Remote Sensing Technology during 1970s, its great potential in the field of agriculture have opened new vistas of improving the agricultural statistics system all over the world. Space borne remotely sensed spectral satellite data has been widely used in the field of agriculture for estimation of area under different major crops like **wheat**, paddy, groundnut and sugarcane.

Studies on use of satellite data for improving the estimator of crop yield obtained from general crop estimation surveys have been taken up at I.A.S.R.I., New Delhi since 1990. During 1990-93 a study was conducted at the Institute to examine the usefulness of satellite spectral data for stratification of crop area based on vegetation indices for improving crop yield estimation based on yield data from crop cutting experiments under general crop estimation surveys. The study pertained to wheat crop yield for district Sultanpur UP for Rabi 1985-86 and the **satellite** data was used from the USA satellite Land Sat-4. This study showed that the efficiency of crop yield estimators can be increased considerably by using the satellite data along with the survey data. Another similar study was undertaken during 1996-98 for improved estimation of wheat crop yield in district Rohtak, Haryana for Rabi 1995-1996 using the IRS 1B - LISS II satellite data for Feb. 17, 1996 and the crop yield data from crop yield estimation surveys for Rabi 1996. The results from this study also showed that satellite data in the form of vegetation indices greatly improves the efficiency of crop yield estimator.

In the present study improved post stratified estimators of crop yield at district level have been obtained for wheat crop for district Rohtak for Rabi 1997-98. The satellite data has been taken from IRS 1D -LISS III for Feb. 4, 1998 and crop yield data has been taken from GCES for Rabi 1997-98 for the district. The spectral data has been used in the form of vegetation indices NDVI and RVI for stratifying the crop area into homogeneous crop growth areas and corresponding **post-stratified** estimators of crop yield have been developed at district level. In this study Global Positioning System (GPS) has been used to identify the location of **crop** cutting fields sites on the satellite imageries. The results are extremely good showing that post-stratified estimators are considerably more efficient compared to the usual estimator of crop yield. The post-stratified estimator based on NDVI is more than 101% more efficient and the post stratified estimator based on RVI about 66% more efficient compared to the usual estimator.

Forecasting of crop production is one of the most important aspect of agricultural statistics systems. Yield forecasts at present are based on eye estimates and are quite subjective. The final crop production **estimates** based on objective crop cutting surveys become available long after the harvests. This as such calls for the necessity of developing some objective methods for **pre-harvest** forecast of crop yield and several efforts are being made continuously in this direction.

Verma et al (1988) presented the results of a study on "Evaluation of crop cut method and farmers reports for estimating crop production" undertaken at Longacre Agricultural Development Centre UK". This study was carried out in 5 countries in Africa during 1987 with the objective of comparing crop estimates based on crop cut methods with estimates obtained by asking farmers directly to state their production. The results of the study showed that farmer's eye estimates are remarkably close to actual production figures in all the countries and they also show considerably small variance compared to the estimates based on crop cutting experiments. After the publication of this report considerable interest is again focused on using farmers estimates which are much cheaper to obtain and easier to conduct. Since farmers eye estimates are subjective but can be obtained at a much smaller cost, it was considered prudent to examine the use of farmers eye estimate of crop yield as auxiliary variable for improving the efficiency of crop yield models based on satellite spectral data.

In the present study crop yield forecasting models have been developed using satellite data in the form of vegetation indices NDVI and RVI and the farmers eye estimate of crop yield for the corresponding plots. For developing the model wheat crop yield data for Rabi 1995-96 from GCES and also the farmers eye estimate of crop yield for the corresponding plots for district Rohtak and corresponding satellite spectral data for Feb. 17, 1996 from IRS 16 LISS II have been used. For testing the model the respective data has been taken for 1997-98. The results show that the predicted yield is very close to the actual yield in almost all the models, However the most efficient model is achieved when the satellite data in the form of NDVI along with the farmers eye estimate of crop yield are used as independent variables. In this case the value of R^2 is 0.90 with a standard error of 1.02 and the predicted value is very close the actual value with a standard error of approximately 5%.

Issue of small area estimation has also gained importance in view of growing needs of micro level planning. The advances in computer facilities have provided convenient tools for many theoretical developments for providing small area estimates. The small area estimation techniques make use of information from other available sources and borrow strength from related or similar areas through explicit and implicit models that connect this small area via supplementary data.

Singh et. al. (1999) developed methodology for small area estimation of crop yield. Using satellite spectral data and the crop acreage data. The population (a district) consisting of T small areas (Tehsil/Block) was divided into V post strata representing crop condition like very good crop, average crop, poor crop, no crop etc. based on the vegetation indices derived from the satellite spectral data. The crop within these post strata is homogeneous in respect of the character under study (the crop yield) and the boundaries of these post strata cut across the small areas. Hence it can be easily assumed that the units within a small area belonging to particular post strata will have the same characteristics as the units belonging to that particular post strata irrespective of the small area. Therefore the crop acreage for small area, which is known, may be used to apportion the crop yield for the small area from the yield estimates of post-strata, Two small area estimators (i) the Direct Estimator and (ii) the Synthetic Estimator were developed at tehsil level. These estimators make use of available information on crop yield and also the information of crop acreage for all the small areas.

In the present study also the same two small area estimators namely (i) the Direct estimator and (ii) the Synthetic estimator for wheat crop yield at block level for Rabi 1997-98 for district Rohtak have been developed. The results show that Synthetic estimator is more efficient as compared to the Direct estimator (as expected). The standard error of all the estimates of block level is less than 10%.

Thus the results of the present study provide an integrated approach for crop acreage estimation, crop yield estimation at district level and small area estimates of crop yield at block level and crop yield forecasting model using crop yield data from GCES, and farmers eye estimates of yield of corresponding plots and the corresponding satellite data.

REFERENCES

1. Aggarwal, Ranjana, Jain R.C., and Jha, M.P. (1986). Models for studying rice crop - weather relationship, *Mausam*, 37(1)pp.67-70.
2. Box, G.E.P. and Jenkins, G.M., (1976), *Time series analysis: Forecasting and control*, Holden-Day, San Francisco, 575 p.
3. Dadhwal V.K., and Parihar, J.S., (1985), Estimation of 1983-84 wheat acreage of Karnal district (Haryana) using Landsat MSS digital data, Technical note, IRS-UP/SAC/CPF/TN/09, Space Application Centre, Ahmedabad.
4. Gonzales M.E. (1973) Use and evaluation of synthetic estimates. *Proceedings of the American Statistical Association, Social Statistics Section* 33-36.
5. Hubback, L.A. (1927). Sampling for rice yield in Bihar and Orissa, *Pusa Agric. Res Inst. Bull.No.* 166.
6. Jain R.C., Aggarwal, Ranjana and Jha, M.P. (1980). Effects of climatic variables on rice yields and its forecast. *Mausam*, 31(4), 591-596.
7. Jha, M.P., Jani, R.C. and Singh, D.(1981). Pre-harvest forecasting of sugarcane yield. *Indian J. Agric. Sci.* 51(11), 757-61.
8. Mahalanobis, P.C. (1945). A report on the Bihar crop survey, 1943-44. *Sankhya*, 7, 29-118
9. Mahalanobis, P.C. (1946). Recent Experiments in Statistical Sampling in the Indian Statistical Institute, *J. Roy, Statist. Soc.* 190,325-370.
10. Manual of nationwide land use/land cover mapping using digital techniques part-II (1999), RRSSC Nagpur, Deptt. of space.
11. Panse, V.G. Rajgopalan, M and Pillai, S.S. (1966). "Estimation of crop yield for small domain" *Biometrics* pp 374-384.
12. Parihar, J.S., Panigrahi, S., Dadhwal, V.K., Bhatt, H.P., Dass, N.K., Gosh B.K. and Behra, D. (1987). Rice acreage estimation in South Orissa using stratified random sampling approach. *Scientific Note, IRS-UP/SAC/CPR/SN/09/87*, 24.
13. Purcell S., and Kish, N.J. (1979) Estimation of small domains. *Biometrics* 35 : 365-384
14. Saha, S.K. (1999), Cop yield modeling using satellite remote sensing and GIS-current status and future prospects. Processing "Geoinformatics - Beyond 2000, an international conference on Geoinformatics for natural resources assessment, monitoring and management" held at IIRS Dehradun during 9-11 March, 1999.
15. Sahai Baldev and Navalgund, R R.,(1988). Indian Remote Sensing Satellite utilization programe for Agriculture, *Remote Sensing in Agriculture* (ed. Baldev sahai), Ind. Soc. Of Rem. Sen., Ahmedabad Chapter, 41.
16. Sarndal C.E. (1984) Design consistent versus model dependent estimation for small domains. *JASA* 79, 624-631.

17. Singh, R., Goyal, R.C. Saha, S.K. and Chhikara, R.S. (1992). Use of Satellite Spectral data in Crop Yield Estimation Surveys. *Int. J. Remote Sensing*, Vol.13.No.14, 2583-2592.
18. Singh, Randhir and Goyal, R.C, (1993) Use of remote sensing technology in crop yield estimation surveys. Project Report, IASRI, New Delhi.
19. Singh. Randhir and Ibrahim. AEI (1996) Use of spectral data in Markov Chain model for crop yield forecasting. *Jr.Indian Socy. of Remote Sensing* V.24 No.3, 145-152
20. Singh, Randhir, Goyal, R.C., Pandey, L.M., and Saha, S.K., (1999), Use of remote sensing technology in crop yield estimation surveys - II, Project Report IASRI, New Delhi.
21. Singh R, Semwal Dinesh P., Rai, Anil and Chhikara, Raj S, (2000) small area estimation of crop yield using remote sensing satellite data. Accepted for publication) *International Journal of Remote Sensing*
22. Sridhar, V.N., Dadhwal, V.K. Choudhary, K.N. Sharma, R., Bairagi, G.D. and Sharma A.K., (1994), Wheat production forecasting for a predominantly unirrigated region in Madhya Pradesh (India), *Int. J. Rem. Sens.*, 15(6):1307-1316.
23. Sukhatme, P.V. and Panse, V.G. (1951). Crop surveys in India - II. *Journal of Indian Society of Agricultural Statistics* 2. 95-168.
24. Tucker. C.J., Holban, B.N., Elgin, J.H., and Mc Murtrey, J.E. (1980). Relationship of spectral data to grain yield variations. *Photogramm. Engg. Remote Sensing*, 46, 657-662
25. Verma Vijay, Marchant, Tun and Scott.C. (1988). "Evaluation of crop cut methods and farmers reports for estimating crop production" Longacre Agricultural Development Centre Ltd., London, U.K.
26. Forecasting Agricultural Output using space, Agrometeorology and Land based observation (FASAL) (1999) -- Scientific Note, Space Application Centre. Ahmedabad