1. Introduction

Systematic and comprehensive compilation of land use statistics is necessary for planned development of agriculture, forests, grasslands, rural settlements, urban spreads, industries and other land based programmes and activities. The need for optimizing land use in an integrated manner has become particularly relevant in recent years as a consequence of compelling and conflicting demands of growing population, increasing land degradation and thus sharply declining man land ratio. The land use planning is a means of optimum utilization of natural resources, specially, the land which is one of the most important resource, based on its related factors like socio-economic condition, soil type, climate, water availability etc. for production as a whole. Naturally, the effective land use planning of particular region needs information related to various input factors of production. The land use statistics not only helps in this planning process but also generates information about other related factors. In India the land use statistics are obtained as per nine fold classification, which is vogue since 1950-51. These statistics are complied and published by Directorate of Economics and Statistics (DES) mostly at district level for most of the states in the country, where cadastral maps exist by complete enumeration through primary reporting agency called Patwari. In permanently settled states of Kerala, Orissa and West Bengal these statistics are based on sample surveys covering 20% of the villages every year. It is general feeling that due to cost and organizational constraints of agencies engaged in collection of data on this aspect, the quality of data is deteriorating and also there is considerable delay in bringing out the results, which hampers the efficient planning process. Further, in view of the micro level planning, there is greater need to provide these statistics at lower levels like tehsils, blocks or at gram panchayat level (as in case of National Crop Insurance Scheme).

The recent advances in the field of geographical techniques, like Geographic Information System (GIS) and Remote Sensing (RS) increased the potential to change substantially the statistical approach to study the geographical realities. The remote sensing techniques to obtain the land utilization statistics gained popularity because of its extensive coverage of geographical area. Although, this technology has vast potential especially in this area, but statistics so obtained are also not free from various kinds of errors. The GIS, is capable of handling geographical data through its geographical coordinates, obtained by different sources like Census, Survey and Remote Sensing. It is one of the important tool having capabilities of integration of different kinds of data and assist in the spatial analysis of data from geographical characters. These recent technologies i.e. GIS & Remote Sensing combined together named as geomatics, can be used as a tool to resolve some of the above issues affecting the land use statistics.

Recently, National Statistical Commission (NSC) reviewed the National Agricultural Statistical System in a great depth and made a important suggestions related to improvement of in the Agricultural Statistical System and quality of data. In relation to crop area statistics, NSC suggested that a 20% sample of villages is large enough to estimate crop area quite
efficiently at the state and district level and hence crop area estimation should be based on 20% of sampled villages.

This study was formulated to explore the various possibilities of obtaining reliable land utilization statistics with minimum time lag by considering these important sources of data collection and processing. The study aimed to (i) examine the quality of data on land use statistics recorded by the village agency (Patwari) by independently observing a sample of villages by specially trained and qualified investigators. (ii) Study also aimed to improve the efficiency of the estimator based on traditional sample surveys by using spatial properties of land use and (iii) spatial models were fitted for correcting the estimates of land use statistics obtained through usual approach of satellite digital data. Further, qualitative aspects of data from revenue records have been studied using survey data as true value. The sampling units in the sample i.e. villages have been selected with the help of GIS by incorporating spatial correlation in the sampling design itself. The spatial models were developed using Remote Sensing data as independent variable and exploiting the properties of spatial relationship of different classes of land for prediction of area under different land use characters. With the help of these spatial models, the area statistics can be easily obtained at lower geographical level under different land use classes in a village at highly reduced cost and with the help of revenue records or field surveys. The study was undertaken in district of Lalitpur in UP as this district has been observed to have considerable area under most of the land use classification categories. The study was initiated with following objectives:

1. To obtain land utilization statistics with the help of survey and Remote Sensing technique.
2. To study the qualitative aspects of land utilization statistics obtained through different sources i.e. Census, Survey and Remote Sensing.
3. To develop model for integration of statistics obtained through different sources.

2. GIS and Remote Sensing Applications

The application of remote sensing satellite digital data for estimation of crop area was initiated in USA under Corn Blight Watch Experiment (CBWE) in 1971. Crop Identification and Technology Assessment for Remote Sensing (CITARS) experiment was started in 1973 to quantify the Crop Identification Performance (CIP) followed by Large Area Crop Inventory Experiment (LACIE) during 1974-1977 for forecasting of wheat production of major wheat grains regions. A major programme for research and development named as Agriculture and Resources Inventory Survey through Aerospace Remote Sensing (AGRISTARS) was taken up in 1988. Number of methodological studies were carried out in Africa, Europe, Argentina, Australia, Brazil, Canada, Japan etc.. Currently, major programs are under way in Africa under Global Information and Early Warning System (GIEWS) and in Europe Under Monitoring Agriculture Through Remote Sensing (MARS). A large number of small level studies have been carried out for acreage estimation of major crops. Some of such studies are by Lepoutre (1991), Meyer (1991), Fang (1998) etc.

In India, Indian Council of Agricultural Research (ICAR) and Indian space Research Organization (ISRO) jointly conducted the first multi-spectral air born study for identification of root-wilt disease in coconut in 1969.

The country level studies related to applications of remote sensing technologies were initiated after launch of IRS-IA satellite. Crop Acreage and Production Estimation (CAPE) was one of the important projects in this direction for estimation of crop area under wheat, rice, cotton, ground nut, sorghum& mustered. Apart from these national level projects, numbers of small
studies have been carried out to develop methodologies for application of satellite data in various fields of agricultural and rural development by Department of Space. Some of these studies are by Dadhwal et al. (1985, 1991), etc. The details of GIS and remote sensing techniques are provided by Burrough (1986), Curan (1985), Goodchild, (1987). Several methodological studies related to estimation of crop area and production have been carried out at Indian Agricultural Statistics Research Institute (IASRI), New Delhi. Singh et al. (1992) used satellite data for stratification of crop area for the general crop estimation surveys and obtained more precise estimator of crop yield. Singh et al. (1999) also developed small area estimator of crop yield. Singh et al. (2002) used satellite data and the farmers eye estimate for developing a reliable crop yield model.

3. Data Collection

In this study, data from three different sources namely revenue records based on complete enumeration of villages; sample survey and remote sensing have been used as per the objectives of the study. In order to plan the field study of district, the census data of 1991 has been obtained from the Office of Registrar General of India, New Delhi. This data was in digital format and it has been utilized for planning of the field work and selection of sample. A sample of size 20 villages has been selected from the district following proportional representation of the villages of each tehsils. The copy of revenue records collected by the respective patwaris of the selected villages for all three seasons in the district has been taken from the tehsil head quarters. This record pertains to survey number wise land use statistics under different classes for respective villages. In order to check the quality of the data collected by the patwaris in the selected villages, field staff of the institute was deputed to completely enumerate all survey numbers of the selected villages using categories of the area falling under different land use classes as per the standard nine fold classification. The copy of cadastral maps of the selected villages have been obtained from tehsils head quarter and provided to field staff to assist in data collection. The data collected by two independent sources on same set of 20 villages were used for qualitative checking of revenue records. The total numbers of survey numbers covered in each season under this study were 36320. Further, satellite remote sensing data from Indian Remote Sensing Satellite IRS-1C, LISS-III sensor as well as WiFs for all three seasons has been procured from NRSA, Hyderabad.

Further, training sites belonging to different land use classes were identified in all three seasons with the help ground survey carried out in the entire district. In identification of these training sites, Survey of India toposheets of 1:50,000 and 1: 2,50,000 have been extensively utilized.

4. Spatial Stratified Sampling Technique

In present study a stratified spatial sampling technique has been suggested for selection of areal sampling units (villages) in space from the district (spatial region), which is a function of combined spatial correlation coefficient for different land use variables and geographical area (i.e. size) of the villages. Tehsils (lower administrative units in a district) in a districts forms the strata and villages in a tehsil form the sampling units. The neighbors are defined on the basis of optimum distances, from centroid of each village. In this study census data for 1991 census for district Lalitpur has been used for sample selection. There are 754 villages in the district and three land use variables which are available in the census data with respect to land use namely (i) Un-irrigated area (U) (ii) Culturable waste (W) and (iii) Area not available for cultivation (C) have been used for sample selection.
Let \( \mathbf{X}'=(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N) \) a vector of a land use variable deviated from its mean. The corresponding lagged variable as defined by Cliff and Ord (1981) is given by \( \mathbf{X}''=\mathbf{W}' \mathbf{X} \)

where

\[
\mathbf{W}' = \begin{bmatrix}
0 & w_{12} & \cdots & \cdots & w_{1N} \\
w_{21} & 0 & \cdots & \cdots & w_{2N} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
w_{N1} & w_{N2} & \cdots & \cdots & 0
\end{bmatrix}
\]

Here, \( \mathbf{W}' \) is a weight matrix, which measures the spatial proximity of locations in arrangement of areal units over a space. The elements \( w_{ij} \) may take binary values i.e. 0 or 1 i.e. \( w_{ij}=1 \) if units i and j are neighbors and 0 otherwise. The other options may be providing weights based on universal distance, lengths of common boundary, inverse of distance etc. The use of binary numbers is more straightforward and computationally convenient. The off diagonal elements must be chosen such that sum of the diagonal elements are equal to one. Obviously, this matrix is a symmetric matrix. The spatial correlation coefficient based on above weight matrix may be defined as

\[
\rho = \left( \mathbf{X}' \mathbf{X}' \right)^{\frac{1}{2}} \left( \mathbf{X}'' \mathbf{X}'' \right)^{\frac{1}{2}}\left( \mathbf{X}'' \mathbf{X}' \right) \]  \quad \ldots \quad (1)

In the present study spatial correlations for different land use variables i.e. unirrigated area \( (\rho_u) \), culturable waste \( (\rho_w) \), and area not available for cultivation \( (\rho_c) \) were calculated.

Let their variances be denoted by \( \sigma_u^2 \), \( \sigma_w^2 \) and \( \sigma_c^2 \) respectively. Since, number of land use variables needs to be estimated through same sample, so it is desirable to estimate a combined spatial correlation coefficient for selection of the sample. A reasonable value of combined correlation coefficient can be obtained as

\[
\rho = \frac{\sigma_u^2 \rho_u + \sigma_w^2 \rho_w + \sigma_c^2 \rho_c}{\sigma_u^2 + \sigma_w^2 + \sigma_c^2} \]  \quad \ldots \quad (2)

This combined spatial correlation was used for defining neighborhoods, so that this sampling design is reasonably efficient for all land use classes in the region. In this study distance based neighbors were defined by associating information at the centroid of each village and obtaining optimum distance from the village under consideration. The centroid is the point with coordinates that are the mean values of coordinates of all points in the configuration (polygon). The distances between any two areal units (villages) were taken to be Euclidean distance between the centroids of areal units (villages).

5. Optimum Distance for Defining Neighbors

Once distance between any two villages was computed, the next step was to decide areal extent of neighborhood. A circle of radius r is considered from the centroid of the village for which the neighbors were to be defined. All those villages whose centroids fall within this radius were identified as first lag neighbors. Villages whose centroids were falling were identified as second lag neighbors. Similarly, third, fourth lag neighbors were identified.
The optimum value of $r$ was obtained based on data from census 1991 on 754 villages of the district. To obtain the optimum value of $r$, the spatial correlation $\rho$ for different land use variables along with their variances were calculated for different distances (in degree decimals) $h$. It was observed that an increase in spatial correlation occurs with increase in values of $h$ up to certain point and after that spatial correlation started decreasing. The variance was also minimized at this peak value of $\rho$ for a particular value of $h$, which was considered as the optimum distance (or radius) for defining neighbors. The optimum value of distance i.e. $r_o$ was defined as the weighted average of optimum distances obtained for the three land use variables as given below

$$
 r_o = \frac{\sigma^2_u h_u + \sigma^2_w h_w + \sigma^2_c h_c}{\sigma^2_u + \sigma^2_w + \sigma^2_c} \quad \ldots (3)
$$

After obtaining value of combined spatial correlation coefficient $\rho$ and defining criterion for neighbors using optimum distances, a stratified sample of size $n$ villages from a population of size $N$ villages were selected by using proportional allocation following a spatial sampling procedure, which incorporates spatial correlation coefficient and size measure of areal units with respect to neighbors of already selected units in the sample. The spatial sampling procedure is given in the following section.

**Sample Selection:** Let the population is (a district) divided in $L$ strata (Tehsil). Let $N_l$ be the number of sampling units (Village) in the $l$-th strata. Let $n_l$ denote the sample size of villages to be selected from the $l$ th strata. Following steps were followed for selecting sampling units from a stratum:

**Step 1: Selection of first unit from a stratum**

The first unit in each stratum was selected with probability proportional to size measure i.e. geographical area of the village $X_g$. Let $x_{gil}$ denote the geographical area of the $i$th village in the $l$th stratum and $X_{gl}$ denotes total geographical area of the $l$th stratum. Then the probability of selecting $i$th unit (village) of $l$th stratum at first draw in the sample is $P_{i(l)} = X_{gil} / X_{gl}$.

**Step 2: Selection of second unit in the sample from $l$th stratum**

The second unit in the $l$th stratum was selected from remaining $N_l - 1$ unit by following the steps given below:

1. Select a random number from 1 to $N_1$ (say $i$)
2. Select another random number from 1 to $M_i$ (say $j$), where $M_i$ is the maximum value of the auxiliary character in the stratum i.e. geographical area of the village.
3. Define weight $U_{i_2(l)} = \left(1 - \rho^{d_{i_1,i_2}}\right)$ where, $i_2(l)$ denotes the $i$th unit from stratum 1 selected in the second draw of the sample and $d_{i_1,i_2}$ denotes the distance between $1^{st}$ and $2^{nd}$ units which takes value 1 if the $1^{st}$ and $2^{nd}$ units selected in the sample are $1^{st}$ lag neighbors, takes value 2 if $1^{st}$ and $2^{nd}$ units selected are $2^{nd}$ lag neighbors etc. It takes value $n_l$ if $1^{st}$ and $2^{nd}$ units selected are $n_l$-th lag neighbors.
4. Select the unit $i$ if $j \leq U_{i_2(l)}X_{gil}$.

5. Reject unit $i$ and repeat the above process if $j > U_{i_2(l)}X_{gil}$.

The probability of selecting $i^{th}$ unit of the population in the $2^{nd}$ draw is given by

$$P_{(2)il} = \frac{U_{i_2(l)}X_{gil}}{\sum_{i \in \Omega_l} U_{i_2(l)}X_{gil}}$$

Here, $\Omega_l$ denotes the set of all units belonging to stratum $l$ and $s_i^*$ denotes the set of earlier selected units in the sample. It can be seen that sum of these probabilities over $N_l-1$ units of the population becomes one.

**Step 3: Selection of subsequent Units**

Following the above procedure define

$$U_{i_3(l)} = \left(1 - \rho^{d_{(l)13}}\right) \left(1 - \rho^{d_{(l)23}}\right)$$

$$U_{i_4(l)} = \left(1 - \rho^{d_{(l)14}}\right) \left(1 - \rho^{d_{(l)24}}\right) \left(1 - \rho^{d_{(l)34}}\right)$$

$$U_{i_n(l)} = \left(1 - \rho^{d_{(l)1n}}\right) \left(1 - \rho^{d_{(l)2n}}\right) \cdots \left(1 - \rho^{d_{(l)mn}}\right)$$

It can be easily seen that the probability of selecting $i^{th}$ unit in $n_1$-th draw is given by

$$P_{i_{n_1}(l)} = \frac{U_{i_{n_1}(l)}X_{gil}}{\sum_{i \in \Omega_l} U_{i_{n_1}(l)}X_{gil}}.$$ 

It may be noted that above sampling procedure is equivalent to probability proportional to size without replacement and the procedure reduces to simple random sampling with out replacement (SRSWOR) sampling design if $\rho = 0$ i.e. the areal unit are spatially independent and size measure is not taken in to consideration.

**6. Estimation of Different Land Use Classes**

All distinct fields in a village was having distinct identity (survey number), which were given in a cadastral map of the village. All the fields in the selected villages were completely enumerated for the land use and area under different land use classes. Land use statistics for important classes were obtained for each selected village. Let $y_{a(l)i}$, where $(a = 1,2,\ldots,9)$
denotes the area under \( a^{th} \) land use category from \( i^{th} \) village of \( l^{th} \) stratum. Now sample mean of \( l^{th} \) stratum for \( a^{th} \) category i.e. \( \bar{y}_{a(l)} \) was obtained using following estimator

\[
\bar{y}_{a(l)} = \frac{\sum_{i=1}^{n_l} d_{a(i)}}{n_l}
\]

where

\[
d_{a1(l)} = \frac{y_{a1(l)}}{N_l p_{(l)il}}
\]

\[
d_{a2(l)} = \frac{1}{N_l} \left[ y_{a1(l)} + \frac{y_{a2(l)}}{p_{(2)il}} \right]
\]

\[
\vdots
\]

\[
d_{an(l)} = \frac{1}{N_l} \left[ y_{a1(l)} + y_{a2(l)} + \ldots + \frac{y_{an(l)}}{p_{(n)il}} \right]
\]

Then the estimator for area under \( a^{th} \) land use category in a district was obtained by

\[
\hat{Y}_a = N \sum_{l=1}^{L} W_l \bar{y}_{a(l)}
\]

where \( W_l \) is the stratum weight for \( l^{th} \) stratum given by \( N_l / N \).

This estimator is similar to the estimator given by Des Raj (1956). Hence, it can be easily shown that it is unbiased.

An estimate of variance on the similar lines was obtained as :

\[
V(\hat{Y}_a) = N^2 \sum W_l^2 V(\bar{y}_{a(l)})
\]

where \( V(\bar{y}_{a(l)}) = \frac{1}{n_l(n_l - 1)} \sum_{i=1}^{n_l} (d_{a(i)} - \bar{y}_{a(l)})^2 \)

The land use statistics through remote sensing suffer from several deficiencies especially the cloud cover, location and classification errors etc. At the same time the remote sensing satellite data can be available for the entire region at a very low cost and result of land use statistics can be obtained at the earliest (all most on line). Therefore in the present study an attempt has been made to improve the estimates of land use statistics using remote sensing digital data and patwari land records data of selected villages through spatial models. Since, the digital data is available for all the villages of the district, these models can be used to predict the land use statistics for all non-sampled villages so that the land use statistic for the district as a whole can be obtained using the above developed relationship. Here, it is proposed to apply spatial model by incorporating the spatial structure of the land use variables. These models may provide better prediction as compared to traditional prediction models as the variable under study is spatial in nature.
7. **Spatial Regression Model**

Let, $Y_a$ denotes the vector of observations pertaining to the values of $a^{th}$ class of land use through a suitable model is developed using this data of villages selected in the sample. Similarly $X_a$ is corresponding vector of observation obtained through patwari records.

Further as NSC has recommended to do away with the complete girdavali and use only data for 20% villages for accarage estimation, this methodology of spatial modeling can be used with greater advantage to improve the land use statistics using the patwari land records of this 20% sample of villages and the satellite data of the whole district of all the villages.

Let us consider the following simple regression model

$$ Y = X\beta + e $$

where $e \sim N(0, \sigma^2 I)$

$$ Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_9 \end{bmatrix}, \quad Y_a = \begin{bmatrix} Y_{1,a} \\ Y_{2,a} \\ \vdots \\ Y_{20,a} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_9 \end{bmatrix}, \quad (a = 1,2,3,\ldots,9) $$

$$ X = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \text{and} \quad I \text{ is identity matrix.} $$

In this model errors $e$ are assumed to follow normal distribution with mean vector $0$ and variance-covariance matrix $\sigma^2 I$. So ordinary least square estimator of $\beta$ i.e. $\hat{\beta}_{ols}$ is identical to its maximum likelihood estimator. It is well known that in case of spatial data the structure of variance – covariance matrix of error $e$ i.e. $\sigma^2 I$ will not have simplified structure due to spatial dependence of observations. Hence, structure of variance –covariance matrix here may be represented by $\Sigma$. The elements of $\Sigma$ may be obtained by following two different approaches.

8. **Spatial Weighted Regression Model**

This approach is based on obtaining optimum weights depending on the location of the spatial aerial units. Hence, in this approach the location specific regression coefficient can be obtained. This technique provides non-parametric estimates of regression coefficients. It also leads to localized ordinary least square estimation and may be called spatial weighted regression. This technique was suggested by Brunsdon, Furtheringham & Chariton(1998).
In this case for a given location “i” the circle of inclusion was drawn using the optimum radius \( r_0 \) obtained for defining the neighbors. The weights \( \alpha_{ik} \) for \( k \)–th villages were obtained using the procedure as given below:

\[
\alpha_{ik} = \begin{cases} 
\exp(-d_{ik}/s \times r_0), & \text{if } d_{ik} < s \times r_0 \\
0, & \text{otherwise}
\end{cases}
\]  

… (6)

Where, \( s \) denotes the order of neighborhood and takes values 1,2, 3,… and \( d_{ik} \) is the distance between i-th and k-th location. This is also called Kernel function or Kernel and denoted by the letter K such that \( \alpha_{ik} = K(d_{ik}) \). The above Kernel has all the desirable feature of a Kernel function such as

(i) \( K(0) = 1 \)

(ii) \( \lim_{d \to \infty} \{K(d)\} = 0 \), and \( K \) is a monotone decreasing function for positive real value of \( d \).

Now the regression coefficient \( \beta_{s} \) of the regression equation can be estimated using the structure of \( \Sigma \) where

\[
\Sigma = \begin{bmatrix}
V_1 & 0_2 & - & - & - & 0_9 \\
0_1 & V_2 & - & - & - & 0_9 \\
0_1 & 0_2 & V_3 & - & - & 0_9 \\
0_1 & 0_2 & 0_3 & V_4 & - & 0_9 \\
0_1 & 0_2 & - & - & - & V_9
\end{bmatrix}_{20}, \text{ where}
\]

\[
V_a = \sigma_a^2 \begin{bmatrix}
1 & \alpha_{a12} & - & - & \alpha_{a1n} \\
\alpha_{a21} & 1 & - & - & \alpha_{a2n} \\
\alpha_{a31} & \alpha_{a32} & 1 & - & \alpha_{a3n} \\
- & - & - & 1 & - \\
\alpha_{an1} & \alpha_{an2} & - & - & 1
\end{bmatrix} \quad a = 1,2,3,...,9
\]

Thus using this structure of \( \Sigma \), location specific equations were obtained separately. \( \hat{Y}_a \), area under different land use classes can be obtained using the estimated regression coefficients.

\[
\hat{\beta} = (X'\Sigma X)^{-1} X \Sigma Y
\]

The estimate of variance for the above estimator of regression coefficient is

\[
\hat{\text{V}}(\hat{\beta}) = (X'\Sigma^{-1} X)^{-1}
\]

9. Spatial Variogram Regression Models

In the spatial variogram approach the variance-covariance matrix is generated based on spatial dependence instead of distances of neighborhoods. This reflects, in a physically
realistic fashion, the possible spatial dependence of the errors. The simplest way of denoting the spatial dependence of errors is the use of authorized variogram function. In this approach, elements of $V_a$ i.e. $(\alpha_{ij})$ are obtained by using following equations:

$$\alpha_{aij} = 1 - r(h_{ij}) \quad a, i, j = 1, 2, 3, \ldots, 9$$

where $h_{ij}$ is the lag separating spatial units $i$ and $j$ in space. This also preserves the symmetric nature of variance-covariance matrix.

The examples of authorized variogram function are the exponential and spherical variograms. The exponential variogram function is defined as

$$r(h) = 1 - e^{-(h/r_0)}$$

where $r_0$ is optimum distance parameter and $h$ is the lag distance.

The spherical variogram function is defined as

$$r(h) = \begin{cases} \frac{3h}{2r_0} - \frac{h}{2} \left( \frac{h}{r_0} \right)^2, & h \leq r_0 \\ 1, & h > r_0 \end{cases}$$

The variance–covariance structure $\sum$ can be estimated based on these two variogram functions spatially. The Gaussian variogram function was also fitted but the fit was poor so it has not be considered further. Using the data of selected villages in these models, the land use statistics of the non-selected villages can be obtained. Again the estimated variance-covariance for the regression coefficient is given by

$$\hat{V} = \left( X' \sum^{-1} X \right)^{-1}$$

10. Results and Discussions

This study was undertaken in district of Lalitpur in UP due to the fact that this district has been observed to have considerable area under most of the land use classification categories. It has been observed in this study that quality of revenue records in the study area i.e. Lalitpur district is quite reliable for most of the usual nine fold classified land use classes. The statistics of land use classes were restricted to five broader classes, which can be identified by using single time digital data of Remote Sensing out of the above nine-fold classification can be easily obtained using RS. These statistics of land use classes obtained through RS could be used as auxiliary information in spatial /non-spatial models to get reliable statistics of different classes. The above models can be used to predict the statistics related to these classes for non-surveyed area/villages of the districts. Hence, it is possible to develop reliable land use statistics at any smaller level i.e. panchayat/block/tehsils using above models.

To take into account the spatial dependence of the neighboring units, the classical sampling technique approach is being modified such that the probabilities of selecting neighboring units, once a particular unit is selected in the sample, becomes less as compared to distant areal units.

The best fitted spatial model for each class of land use was found to be different, depending on the spatial distribution of the land use class patches of land in the district. The prediction of area under different land use categories covered under nine fold classification based on satellite data using spatial model seems to be quite satisfactory. In case of land use under forest class and Barren and unculturable land class, accuracy of prediction seems to be poor.
The best-fitted model for each class may be different in different regions depending on the distribution pattern of the land use categories. The comparison of the estimates of area under different classes of land use obtained from field surveys, spatial models and remote sensing has been also presented. These estimates were obtained for only those classes which were identifiable in single year and single date IRS-IC satellite data from LISS-III sensor. It may be noted that in case of remote sensing, one year data, only five classes can be identified out of nine classes of traditional nine fold classification of land use. The land under forest (Class-1), Barren and unculturable land (Class-2) and land put to non-agricultural use (Class-3) are identified clearly using digital data. The area under different crops (Class-5) is mixed with the area under permanent pastures and other grazing lands (Class-4) and area under net area sown (Class-9). Another new class is formed as fellow land, which is a combination of area under cultivable waste (Class-6), area under current fellow (Class-7) and area under fellow land other than current fellow (Class-8). The variance and % standard errors for different land use classes were also calculated only for field survey data by incorporating sampling design for selection of the sample. The salient achievements of this analysis are as follows:

1. The data quality of revenue records were found to be quite satisfactory at least for nine fold classification of land use classes as compared to field survey data in the study district. Hence, to obtain land use statistics falling under each of the classes, the data of revenue records could be used even for sampled villages instead of completely enumerating these villages.

2. The prediction of area under different land use categories covered under nine-fold classification based on satellite data using spatial model seems to be quite satisfactory. In case of land use land 2 accuracy of prediction seems to be poor. The best-fitted model for each class may be different in different regions depending on the distribution pattern of the land use categories. It may be noted that these results holds true only for the study district i.e. Lalitpur. This analysis for development of the spatial models has to be performed for each district/region separately as the spatial pattern of land use classes for other district/region may be different.

3. The estimates of area under different land use classes obtained through remote sensing have been corrected through spatial models. The overall correction factor is around 4% of the total geographical area with respect to actual geographical area of the district.

4. The spatial models developed in this project may be used for estimation of area under different land use statistics even at village level. Hence, these models are quite useful for obtaining these land use statistics at small area levels, such as village panchayats/blocks/tehsils etc..

5. It is observed that the techniques developed under this project may be useful for implementation of National Statistical Commission recommendations in which it has been observed that these statistics should be obtained only on the basis of 20% sampled villages. It may be noted that by remote sensing technology and spatial models, the sample size of the villages may be further reduced, which may leads to considerable saving of national resources.

The total of estimated area for different land use classes using spatial model is around 14% away from actual geographical area where as the total area of all the classes obtained through remote sensing is almost 18% higher than actual geographical area of the district. It clearly indicates that the overall correction of area estimates obtained through remote sensing is around 4% by using spatial models. It may be noted that the total estimates of geographical area obtained through spatial survey is closer to actual
geographical area of the district, it does not imply that the estimates of area under different land use classes estimated through survey are also close to actual area under the same class in the district. Further, it is not possible to estimate area at small area level like village panchayats/blocks/tehsils using spatial sampling technique whereas spatial models will provide area estimates at any of these small area levels.

References


