Jour. Ind. Soc. Ag. Statistics
46(2), 1994 : 210–222

Estimation of Crop Yield Using Post-Stratification Based on Satellite Data

R.C. Goyal, Randhir Singh and R.S. Chhikara*
Indian Agricultural Statistics Research Institute, New Delhi
(Received : November, 1990)

Summary

Singh et.al. [7] utilized LANDSAT spectral data and imagery to post-stratify the crop area in terms of vegetation vigour and obtained improved crop yield estimates. However, in this use of spectral imagery for post-stratification, there is some uncertainty involved in the classification of units into different strata. Chhikara and mckeon [3] proposed a set of models that can be used to describe the classification errors for the units that are post-stratified based on their spectral responses. In the present investigation a simulation study has been conducted based on an error model to examine the effect of misclassification of units on the bias and relative efficiency of different post-stratified estimators of crop yield. It is shown that both the bias and efficiency are adversely affected in the presence of misclassification of sample units in the post-strata.

Key words : Spectral classes, Vegetation vigour, Classification errors, Post-stratified estimators, Bias, Relative efficiency.

Introduction

Crop production statistics are of vital importance to India and of course, for that matter, to any other country in the world. These statistics consist of two major components : (i) the area under crop and (ii) the yield per unit area. Under the present system in India, the crop area is estimated through complete enumeration (or sample surveys in the case of permanently settled areas), and crop yields are estimated from yield estimation surveys based on crop cutting experiments. A stratified multi-stage sampling design is used for the yield surveys : The community development blocks in a district form the strata, the villages in a block form the primary sampling units, crop fields within a village form the second stage sampling units, and a plot of specified dimensions is selected from each of the sampled crop fields for crop harvesting. Fuller [4] and Holt and Smith [5] have discussed the use of post-stratification for improving the efficiency of estimators in surveys. Singh et.al. [7] investigated the use of satellite multi spectral data for post-stratification of crop

* University of Houston, Texas - 77058

areas into the average and high vegetation classes and developed an improved yield estimate for the 1986 wheat crop for the district of Sultanpur (U.P.).

For this study the authors used Landsat satellite Thematic Mapper (TM) data acquired for the district on February 23, 1986 to post-stratify the district cropland. This was achieved by transforming the multi spectral data into certain indices which reflect the vegetation vigour and thus are called vegetation indices (Tucker, et.al., [9]). A stratification of the district area was developed using the vegtation indices into three post-strata, namely, non-vegetation, vegetation vigour and high vegetation vigour ; of which only the last two post-strata were utilized. The sample units were assigned to the post-strata based on use of topographic maps and the visual interpretation of satellite imagery. One may refer to Singh, et.al. [7] for further details on the post-strati-fication and the empirical results obtained for this application. This method of assigning sample units to the post-strata may result into some errors of classification. The accuracy of classification depends on several factors such as the ability of the image analyst, quality of the sensors, spectral and spatial resolution, etc. The extent of such misclassification and its effect on the efficiency of the estimator may be of interest. Since theoretically it is not feasible to study the effect of such misclassification on the accuracy and efficiency of a post-stratified estimator, in the present investigation a simulation study has been conducted to investigate some post-stratified estimators of crop yeild. In this study only a single stage stratified sampling design is considered where the average plot yield from the two crop fields from a village represent the average yield for that village; and hence a village is the sampling unit. In simulations, the actual 1986 data set was utilised to estimate the parameters and a mathematical model to generate a large population with the parameters replaced by their estimated values.

2. *Estimator without Using Spectral Data*

For estimating the mean crop yield in a district, consider it to be a population of N units (villages) divided into L strata such that N_i units belong to the ith stratum. Let a stratified sample of size n be selected, where n_i units are selected for the ith stratum with allocation of units being made in proportion to the strata crop areas, and $\sum n_i = n$. Let Y_{ij} denote the crop yield for the jth unit of the ith stratum. An unbiased estimator of the mean crop yield is given by

$$T_1 = \sum_{i=1}^{L} W_i \overline{Y}_i \tag{1}$$

where

$$W_i = \frac{N_i}{N} \text{ and } \overline{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}.$$

The variance of T_1 is given by

$$V(T_1) = \sum_{i=1}^{L} W_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i}\right) S_i^2 \tag{2}$$

where S_i^2 denotes the variance for the i-th stratum.

3. *Post-stratified Estimators Using Spectral Data*

The stratification in crop yield surveys is based on geographical and political boundaries and hence may not be very effective in term of the efficiency of the estimator. Singh et.al. [7] showed that the use of spectral data for post-stratification of the crop area based on vegetation vigour led to more efficent estimates. Therefore some post-stratified estimators of crop yield are presented which make use of regular yield data obtained from the yield estimation surveys alongwith remotely sensed satellite spectral data used for post-stratification of the crop area.

Consider a population of N units divided into L strata and a stratified sample of size n selected such that n_i units are selected from the i-th stratum. Further, let the crop area for the population of N units (villages) be post-stratified into L′ post-strata based on vegetation indices derived from satellite spectral data. Let N_{ih} and n_{ih} denote the number of population units and sample units, respectively, falling in the overlap between the h-th post-stratum and i-th initial stratum.

If each n_{ih} $(i = 1 \ldots L, h = 1 \ldots, L')$ is non-zero, an unbiased post-stratified estimator of the mean crop yield is given by

$$T_2 = \sum_{h=1}^{L'} \sum_{i=1}^{L} W_{ih} \overline{Y}_{ih} \tag{3}$$

where $$W_{ih} = \frac{N_{ih}}{N} \text{ and } \overline{Y}_{ih} = \frac{1}{n_{ih}} \sum_{k=1}^{n_{ih}} Y_{ihk}$$

The conditional variance of T_2, given that n_{ih} is fixed, is easily seen to be as following :

$$V(T_2 \mid n_{ih}) = \sum_h \sum_i W_{ih}^2 \left(\frac{1}{n_{ih}} - \frac{1}{N_{ih}}\right) S_{ih}^2$$

where S_{ih}^2 is the variance for the overlap area between the hth post-stratum and the ith initial stratum. The unconditional variance of T_2 is then approximately given by

$$V(T_2) = \sum_h \sum_i W_{ih}^2 \left[\frac{1}{nW_{ih}} + \frac{1 - W_{ih}}{n^2 W_{ih}^2} - \frac{1}{N_{ih}}\right] S_{ih}^2 \qquad (4)$$

which follows by using the approximate result due to Stephen [8],

$$E\left(\frac{1}{n_{ih}}\right) = \frac{1}{nW_{ih}} + \frac{1 - W_{ih}}{n^2 W_{ih}^2}.$$

(Ignoring terms of order $\frac{1}{n^r}$. $r > 2$)

But in situations when some of the n_{ih} are zero, the estimator given in (3) is no longer unbiased. In this case an alternative estimator, based on first pooling the sampled units from across the post-strata and then utilizing suitable weights for the initial strata, may be more appropriate as defined by

$$T_3 = \sum_{i=1}^{L} W_i \frac{\sum_{h=1}^{L_i^1} W_{ih} \overline{Y}_{ih}}{\sum_{h=1}^{L_i^1} W_{ih}} \qquad (5)$$

where, L_i^1 is the number of post-strata corresponding to sample units from the ith inital stratum. (Those post-strata which do not contain any unit from the i-th initial strata will be excluded from this count).The revised weight for the i-th strata in this situation will be

$$W_i = \sum_{h=1}^{L_i^1} W_{ih}$$

Similarly, another estimator can be developed, first by pooling the sample units over the initial strata and then making use of suitable weights for the post-strata, as defined by

$$T_4 = \sum_{h}^{L'} W_{.h} \frac{\sum_{i=1}^{L_h} W_{ih} \overline{Y}_{ih}}{\sum_{i=1}^{L_h} W_{ih}} \tag{6}$$

where L_h is the number of initial strata from which sample units have fallen into the h-th post-stratum. (Excluded from this count are those strata from which no unit fell into the h-th post-stratum). The post-stratum weight is

$$W_{.h} = \sum_{i=1}^{L_h} W_{ih}$$

Clearly the estemators given in (3), (5) and (6) take into consideration simultaneously the effect of initial strata based on geographical boundaries and the post-strata based on vegetation vigor. Hence these estimators can be expected to be better than the estimator given in (1).

What if all W_{ih} are not known ? This may happen in the case of post-stratification based on vegetation vigour using satellite data since it would require a complete enumeration of units to obtain each W_{ih} and this may not be feasible. However, the post-strata weights in terms of total areas for the post-stratified classes may be easily obtained by aggregating all pixels falling in each post-stratum. On the other hand, if the intital sample allocation is made in proportion to the crop area, N_{ih} may be estimated as

$$N_{ih} = (n_{ih}/n_h)\, N_h ,$$

and an alternative estimator of the mean crop yield can be obtained from (6) with N_{ih} replceed by $\hat{N}_{ih}$. The form of this estimator is

$$T_5 \doteq \sum_{h=1}^{L'} W_h \overline{Y}_h \tag{7}$$

where $\overline{Y}_h$ is sample mean based on all sample units falling into the h-th post-stratum and W_h is the h-th post-stratum weight. This estimator is best when post-strata weights W_h are known. Hence this estimator may prove quite efficient in the case of post-stratification of crop area based on spectral data where

W_h is available based on either the estimated N_{ih} or the total area for the post-strata as discussed above.

The estimators T_3, T_4 and T_5 are based on complex survey design and hence the expression for their Mean Square Errors will be quite complicated and will not be of comparable form. Therefore the efficiencies of these estimators are examined empirically through a simulation study.

4. *Errors of Classification*

The vegetation vigor is measured in terms of certain vegetation indices obtained by transforming the satellite multi spectral data. In the past, one or more crop growth variables have been defined and ultilized as vegetation indices (Barnett and Thompson, [1]. Based on the computed values of these indices, a color composite imagery is prepared for the area of interest. The analyst interprets the image and manually delineates the vegetation vigour classes which can subsequently be used as post-stata (Singh et.al [7]).

For a post-stratification of the crop area into vegetation vigour classes using satellite data, the sample units are first identified on the FCC imagery and then each unit is classified or assigned to a post-stratum. This process involves a manual procedure to identify unit of the imagery for the purpose of its assignment to one of the vegetation vigour classes.

Hence some of the sample units may get misclassified, i.e. a unit may be classified belonging to a class other than its actual class (Chhikara, [2]).

From a classification of all the n sample units into the two post-strata (say, PS1 and PS2) supose one has the following :

n_1 = actual number of units in PS1

n_2 = actual number of units in PS2

n_{11} = number of units correctly classified to PS1

n_{12} = number of units classified from PS1 to PS2

n_{21} = number of units classified from PS2 to PS1

n_{22} = number of units correctly classified to PS2

Next, define

$$\theta = E\left(\frac{n_{12}}{n_1}\right),$$

the expected proportion of units misclassified from PS1 to PS2, and

$$\varphi = E\left(\frac{n_{21}}{n_2}\right),$$

the expected proportion of units misclassified from PS2 to PS1.

Chhikara and Mckeon [3] provide a set of models that can be used to describe the misclassification of units. In the present context, the most suitable model is their exponential model. Under this model the chance of misclassification for an individual unit increases as the variable value deviates further away from the mean of its true class in the direction of the mean of the other class. Following chhikara and Mckeon [3], let

$$g_{12}(t) = 1 - \exp[-k_1(t - \mu_1^2)/2\sigma_1^2], \tag{8}$$

a monotone increasing function, and

$$g_{21}(t) = 1 - \exp\left[-k_2 \frac{(t-\mu_2)^2}{2\sigma_2^2}\right] \tag{9}$$

a monotone decreasing function with t, the variable value to lie on the axis joining the means μ_1 and μ_2 of the two post-strata PS1 and PS2. $\mu_1 < t < \mu_2$ and $K_1, K_2 > 0$. Let $h_i(t)dt$, $i = 1, 2$, denote the probability that an individual unit belongs to PSi. In the present study $h_i(t)$ is taken as the normal pdf,

$$h_i(t) = \frac{1}{2\pi\sigma_i} \exp\left[\frac{-(t-\mu_i)^2}{2\sigma_i^2}\right], i = 1,2. \tag{10}$$

Accordingly, the expected error rates are

$$\theta = \int_{-\infty}^{\infty} g_{12}(t) h_1(t)\, dt, \quad \text{and}$$

$$\varphi = \int_{-\infty}^{\infty} g_{21}(t) h_2(t)\, dt. \tag{11}$$

It is easily seen that when some of the sample units are misclassified in terms of their post-strata, this may introduce certain amount of bias and also increase in the variance of a post-stratified estimator. Since it is not possible to examine these effects theoretically, we now consider simulations to study these effects.

5. *A Simulation Study*

To study the effect of misclassification on the bias and relative efficiency for the different post-stratified estimators of crop yield an empirical study based on simulated data has been undertaken. A discrete population of size N is generated randomly, where associated with each unit is a parametric value generated as the average yield of the unit. For the present study the size of the simulated population is taken to be 2490 (which is equal to the number of villages in the district of Sultanpur.) The wheat yields for the units are generated using the normal distribution with the parmatric values estimated from the real wheat crop yields obtained from crop cutting expriments in the district. The estimated parametric values were as : mean = 20.12 (quintal/hectare) and standard deviation = 5.91. For more details on generating normal variate values one may refer to Marsagalia and Maclaren [6].

Based on the observed yield data, the district of Sultanpur can be stratified into 6 homogeneous strata. Therefore, the simulated population is also divided into six initial strata of sizes $N_1, N_2 \ldots N_6$ to represent the blocks in the district. For the post-stratifications the yield simulated for the units are arranged in ascending order and the units are divided into two groups of specified sizes, say N_p and N_q (such that $N_p + N_q = N$) representing the two post-strata of average yield and high yield, PS1 and PS2, repectively. The basic idea is that a high vegetation vigour leads to a higher crop yield and a low to moderate vegetation vigour would result in an average yield for the crop.

The classification of sample units into the two post-strata PS1 and PS2 are based on the exponential model assumed for the errors of classification as discussed in Section 4. Suppose $\hat{\mu}_1$ and $\hat{\mu}_2$ are the average vegetation indices obtained for PS1 and PS2 respectively, and $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are their corresponding standard deviations. Then the following method is used for the classification of sample units : For an individual unit to belong to PS1, let

$$\begin{aligned} g_{12}(t) &= 0, t \leq \hat{\mu}_1 \\ &= 1 - \exp[-k_1 (t - \hat{\mu}_1)^2 / 2\hat{\sigma}_1^2], \; t > \hat{\mu}_1 \end{aligned}$$

and for the unit to belong to PS2, let

$$\begin{aligned} g_{21}(t) &= 0, t \geq \hat{\mu}_2 \\ &= 1 - \exp[-k_2 (t - \hat{\mu}_2)^2 / 2\hat{\sigma}_2^2], \; t < \hat{\mu}_2 \end{aligned}$$

where k_1 and k_2 are easily obtained for the specified values of θ and φ. After replacing μ_1 by $\hat{\mu}_1$ and σ_1 by $\hat{\sigma}_1$ in (10), it follows from equation (11) that

$$\theta = \int_{\mu_1}^{\infty} [1 - \exp(-k_1 \frac{(t-\hat{\mu}_1)^2}{2\hat{\sigma}_1^2})] \left[\frac{1}{\sqrt{2\pi}\sigma_1} \exp \frac{-(t-\hat{\mu}_1)^2}{2\hat{\sigma}_1^2} \right] dt$$

$$= \frac{1}{2} - \frac{1}{\sqrt{2\pi}\sigma_1} \int_{\mu_1}^{\infty} \exp [\frac{-(k_1+1)(t-\hat{\mu}_1)^2}{2\sigma_1^2}] \, dt$$

$$= \frac{1}{2} - \frac{1}{2} (\sqrt{k_1 + 1})^{-1}$$

Hence $k_1 = (1 - 2\theta)^{-2} - 1.$

Similarly, $k_2 = (1 - 2\varphi)^{-2} - 1.$

For sample units satisfying $t \leq \hat{\mu}_1$, or $t \geq \hat{\mu}_2$, the allocation is obvious. To classify the units satisfying the condition $\hat{\mu}_1 < t < \hat{\mu}_2$, we generate a uniform random variable u. For $t \in$ PS1, classify t to PS1 if $u > g_{12}(t)$, Otherwise to PSI. Similarly for $t \in$ PS2, classify t to PS1 if $u \leq g_{21}(t)$, otherwise to PS2.

Two separate cases were considered for the population units to be subivided into the two post-strata. These corresponded to the two proportions of P = 0.75 and 0.50 for the average vegetation post-stratum and consequently the proprtions of P = 0.25 and 0.50, respectively, for the high vegetation post-stratum. For the misclassification rates we considered θ = 0.00, 0.10, 0.20 and φ = 0.00, 0.10 and 0.20 where θ is the misclassification rate for the units in the average vegetation class but classified into the high vegetation class, and φ is the misclassification rate for the units in the high vegetation class but classified to the average vegetation class. For the sample size, two cases of n = 90 and 30 were considered. Moreover, to estimate the mean crop yield in the district, stratified random samples of sizes 90 and 30 units each were generated repeatedly 500 times as well as 1000 times, and these samples were each time post-stratified into two post-strata. However, it was observed that there is no significant difference in the results based on 500 samples and 1000 samples and hence the results are presented only for the case of 500 repeated samples.

For each simulated sample, five estimates of the mean yield using estimators given in (1), (3), (5), (6) and (7) were computed, and then the average was computed from the 500 estimates obtained corresponding to each estimator. The observed relative bias and variance from repeated samples are computed as follows :

For the i-th estimator, say T_1,

$$\text{Relative bias} = \frac{\overline{T}_i - \overline{Y}}{\overline{Y}} \times 100$$

where $\bar{T}_i = \frac{1}{k} \sum_j T_{ij}$, and

$$\text{Variance} = \frac{1}{k-1} \sum_j (T_{ij} - \bar{T}_i)^2. \tag{12}$$

Here k = 500 and T_{ij} denotes the j-th estimate corresponding to i-th estimator, j = 1, 2, . . . 500.

6. *Results and Discussions*

Tables 1–2 give the results obtained for relative bias, relative efficiency based on variance and the relative efficency based on mean square error of different estimators of average wheat yield for various misclassification rates and post-strata proportions. From these results it can be seen that when $\varphi = O$ and $\theta = O$ i.e. there is no misclassification the estimator T_5 is most efficient and estimators T_2, T_3, T_4 also perform far better as compared to T_1. When $\varphi = O$ and $\theta \neq O$ then in most of the cases the estimator tend to understimate the average yield and this tendency to understimate becomes more pronounced with an increase in the value of θ. For $\theta = O$ and $\varphi \neq O$ then the average yield is generally over estimated for N = 90 and the over estimation is increased with an increase in the value of φ. When sample size is reduced to 30 the direction in the similarity of results still holds, though generally estimator T_2 showed higher under estimation as the misclassification rate is increased.

A perusal of the relative efficiencies of the estimators compared to T_1 based on variance (i.e. ignoring the bias) showed that in general T_5 is most efficient followed by T_4, T_3 and T_2 respectively and when the sample size is large (i.e. n = 90) the estimators T_3, T_4 and T_5 are fairly competitive. Also in most of the cases the efficiency of an estimator decreases with an increase in misclassification rates θ and φ. The decrease is relatively more in case of .75 than for p = .50.

But in case of relative efficiency based on mean square error the efficiency of most of the estimators T_2, T_3, T_4 and T_5 is considerably reduced indicating that the bias in the new estimator also is not reduced in case of misclassification.

From these results it has also been observed that the efficiency of the post-stratified estimators except T_2 is not affected much by reducing the sample size which suggests that there is a scope to compensate the additional cost incurred from the use of satellite data for post-stratification by reducing the size of the sample of crop cutting experiments.

Table 1. Relative bias B, Relative efficiency based on variance E_1 and relative efficiency based on MSEE_2 of various yield estimators for different rates of misclassification based on 500 samples of sizes 90 and 30 for P = 0.75.

θ	Estimator	φ	0.0		0.10		0.20	
			90	30	90	30	90	30
0.0	T_1	B	−0.02	−0.12	−0.11	0.06	0.05	0.32
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−0.66	−7.66	0.34	−7.86	0.97	−7.24
		E_1	1.34	0.63	1.21	0.55	0.95	0.52
		E_2	1.26	0.29	1.20	0.25	0.86	0.27
	T_3	B	−0.18	−2.69	0.93	−1.79	1.83	−1.03
		E_1	1.97	1.42	1.91	1.35	1.69	1.31
		E_2	1.96	1.07	1.62	1.18	1.07	1.24
	T_4	B	0.09	−0.08	1.27	1.41	2.33	0.37
		E_1	2.09	2.01	2.06	1.85	1.93	1.94
		E_2	2.09	2.01	1.51	1.64	0.93	1.42
	T_5	B	0.19	0.00	1.37	1.54	2.40	2.49
		E_1	2.35	2.28	2.21	2.06	2.07	2.18
		E_2	2.33	2.27	1.52	1.76	0.93	1.50
0.10	T_1	B	−0.18	0.07	−0.23	−0.30	−0.77	0.32
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−3.81	−8.98	−2.50	−7.91	−1.79	−7.50
		E_1	1.84	0.82	1.58	0.66	1.15	0.70
		E_2	0.47	0.26	0.84	0.28	0.86	0.30
	T_3	B	−3.70	−4.84	−2.31	−3.72	−1.36	−2.81
		E_1	2.01	1.52	1.95	1.39	1.54	1.38
		E_2	0.50	0.71	1.01	0.85	1.22	1.01
	T_4	B	−3.64	−3.48	−2.21	−2.12	−1.13	−0.73
		E_1	2.01	2.00	2.08	1.18	1.60	1.67
		E_2	0.51	1.12	1.09	1.42	1.35	1.63
	T_5	B	−3.53	−3.28	−2.26	−2.08	−1.05	−0.67
		E_1	2.10	2.36	2.28	2.04	1.78	1.93
		E_2	0.55	1.29	1.13	1.57	1.51	1.88
0.20	T_1	B	−0.29	−0.10	0.01	0.18	−0.23	−0.37
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−7.79	−10.90	−6.22	−9.96	−4.84	−8.74
		E_1	1.77	1.02	1.81	0.79	1.49	0.91
		E_2	0.15	0.20	0.22	0.23	0.34	0.27
	T_3	B	−7.76	−7.63	−6.18	−6.32	−4.76	−5.31
		E_1	1.81	1.51	1.84	1.30	1.59	1.39
		E_2	0.15	0.39	0.22	0.49	0.35	0.60
	T_4	B	−7.74	−7.59	−6.16	−6.01	−4.71	−4.46
		E_1	1.81	1.90	1.83	1.54	1.62	1.48
		E_2	0.15	0.41	0.22	0.56	0.36	0.74
	T_5	B	−7.70	−7.44	−6.16	−5.90	−4.56	−4.42
		E_1	1.94	2.30	1.96	1.81	1.72	1.87
		E_2	0.15	0.44	0.22	0.60	0.38	0.84

Table 2. Relative bias B, Relative efficiency based on variance E_1 and relative efficiency based on MSE E_2 of various yield estimators for different rates of misclassification based on 500 samples of sizes 90 and 30 for P = 0.50.

θ	Estimator	φ	0.0		0.10		0.20	
			90	30	90	30	90	30
0.0	T_1	B	−0.02	0.00	0.00	−0.11	−0.06	−0.30
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−0.14	−3.88	2.06	−3.01	5.14	−1.70
		E_1	2.61	0.95	1.89	0.76	1.69	0.54
		E_2	2.63	0.64	0.99	0.62	0.31	0.51
	T_3	B	−0.13	−0.05	2.09	1.23	5.22	3.15
		E_1	2.65	1.83	2.04	1.63	1.94	1.45
		E_2	2.67	1.83	1.02	1.51	0.31	0.98
	T_4	B	−0.13	−0.03	2.11	1.95	5.27	4.86
		E_1	2.67	2.64	2.07	2.13	1.98	2.10
		E_2	2.18	2.64	1.01	1.68	0.31	0.79
	T_5	B	−0.04	0.12	2.16	2.19	5.38	4.97
		E_1	2.96	3.25	2.32	2.63	2.19	2.58
		E_2	2.99	3.24	1.04	1.86	0.30	0.85
0.10	T_1	B	−0.02	0.17	0.11	−0.10	0.10	−0.59
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−2.44	−5.66	0.59	−3.78	2.91	−2.80
		E_1	2.26	1.13	2.13	0.73	1.84	0.56
		E_2	0.45	0.55	1.97	0.53	0.74	0.49
	T_3	B	−2.41	−1.55	0.60	0.13	2.98	1.47
		E_1	2.31	1.59	2.13	1.52	2.07	1.34
		E_2	0.97	1.43	1.97	1.52	0.75	1.23
	T_4	B	−2.42	−2.12	0.60	0.48	3.01	2.58
		E_1	2.34	2.26	2.13	1.92	2.10	1.60
		E_2	0.97	1.75	1.97	1.88	0.75	1.18
	T_5	B	−2.34	−2.07	0.71	0.53	3.05	2.62
		E_1	2.65	2.71	2.44	2.21	2.35	2.00
		E_2	1.06	2.03	2.16	2.16	0.76	1.36
0.20	T_1	B	−0.25	0.39	0.09	0.53	−0.12	−0.39
		E_1	1.00	1.00	1.00	1.00	1.00	1.00
		E_2	1.00	1.00	1.00	1.00	1.00	1.00
	T_2	B	−5.03	−9.02	−2.47	−5.97	−0.04	−4.12
		E_1	2.22	1.11	2.15	0.87	1.89	0.85
		E_2	0.34	0.29	0.94	0.42	1.89	0.58
	T_3	B	−4.91	−3.10	−2.43	−1.37	−0.02	−0.22
		E_1	2.28	1.53	2.16	1.39	1.93	1.51
		E_2	0.36	1.04	0.96	1.29	1.93	1.51
	T_4	B	−4.96	−4.87	−2.45	−2.14	−0.03	−0.22
		E_1	2.33	2.32	2.18	1.78	1.93	1.79
		E_2	0.36	0.84	0.96	1.40	1.93	1.79
	T_5	B	−4.91	−4.78	−2.34	−2.25	−0.05	−0.05
		E_1	2.63	2.67	2.43	2.13	2.06	2.11
		E_2	0.37	0.90	1.06	1.56	2.06	2.12

ACKNOWLEDGEMENT

The authors are thankful to the referee for his valuable suggestions.

REFERENCES

[1] Barnett, T.L and Thompson, D.R., 1982. The Use of Large Area Spectral Data in Wheat Yield Estimation, *Remote Sensing of Environment*, **13**, 509–518.

[2] Chhikara, R.S., 1986. Error Analysis of Crop Acreage Estimation Using Satellite Data, *Technometrics*, **28**, 73–80.

[3] Chhikara, R.S. and Mckeon, J., 1984. Linear Discriminant Analysis with Misallocation in Training Samples, *J. Amer. Statist. Assoc.*, **79**, 899–906.

[4] Fuller, W.A., 1966. Estimation Employing Post-Strata, *J. Amer, Statist Assoc.*, **61**, 1172–1183.

[5] Holt, D. and Smith, T.M.F., 1979. Post-Stratification, *J. Roy. Statist. Soc. (A)*, **142**, 33–46.

[6] Marsagalia, G. and Maclaren, M.D., 1964. A fast procedure for generating normal random variables. *Communi. ACM VII*, 4–10.

[7] Singh, R., Goyal, R.C., Saha, S.K. and Chhikara, R.S., 1992. Use of Satellite Spectral Data in Yield Estimation Surveys. *International J. Remote Sensing.* **13**(14), 2583–2592

[8] Stephan, F.F., 1945. The Expected Value and Variance of the Reciprocal and Other Negative Powers of a Positive Bernoullian Variate, *Ann. Math. Stat.*, **16**, 50–61.

[9] Tucker, C.J., Holben, B.N., Elgin, J.H., and McMurtey–III, 1980. Relationship of Spectral Data to Grain yield Variation, *Photogram. Eng. Remote. Sens.*, **46**, 657–666.