Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 66(3) 2012 391-400

On Small Area Estimation Techniques – An Application in Agriculture

B.V.S. Sisodia and Anupam Singh
Narendra Deva University of Agriculture & Technology, Kumarganj, Faizabad, U.P.

Received 18 June 2010; Revised 11 May 2012; Accepted 11 July 2012

SUMMARY

Small area estimation approach development by Sisodia and Singh (2001) is revisited. This approach does not require any additional survey or conducting extra CCEs for crop-production estimate at block level. The district level data on crop-production and related auxiliary variables are exploited through regression models to obtain reliable estimate of crop production at block level. A new scaled estimator of block estimate is also proposed. Some alternative procedures to obtain weights for the auxiliary variables based on partial correlation and standardized regression coefficients are also suggested. An empirical study for wheat production in Barabanki district of State Uttar Pradesh (India) suggests that among the scaled estimators, $\tilde{Y}_q^{(2)}$ is the best one and the best choice of weights (w_j) be based on partitioning of sum of squares due to regression. Some limitations of the study are also highlighted.

Keywords: Small area statistics, Scaled estimators, Regression model, Crop cutting experiments.

1. INTRODUCTION

Reliable estimates of various parameters at micro level, *i.e.* at Block/Panchayat level are in great demand in the recent times by the administrators and policy planners for policy formulation in the context of de-centralized planning process at smaller level in India and developing countries. In view of such demand the thrust of research efforts has also shifted to development of precise estimators for small area. An offshoot of this development is that various small area estimation (SAE) techniques are being proposed by different research workers for their implementation. One of the SAE technique involves micro level information related to study character for scaling down estimates available at the higher level to lower level.

In the present paper we describe SAE approach as an alternative methodology for estimating crop production at Block level. This approach does not require any additional survey or conducting extra crop cutting experiment (CCEs) for producing the crop production estimate at Block level. In fact, the District level estimates of crop production are developed on the basis of scientifically designed CCEs conducted under the scheme of General Crop Estimation Surveys (GCES). Main theme is to use statistical models to link the variable of interest (crop production) with auxiliary information available from various secondary sources at district level. These statistical models are then used to obtain reliable estimate of crop-production at Block level. The remaining sections of the paper are organized as follows. Section 2 highlights some relevant references and describes the existing methodology for estimating the crop-production at Block level based on regression models. In section 3, we propose a new small area estimator based on a regression model and its

*Corresponding author : B.V.S. Sisodia
E-mail address : bvssisodia@gmail.com

properties are described. Section 4 describes the empirical results using wheat production data of Barabanki district of the State of Uttar Pradesh, India, to examine the relative efficiency of various estimators. Concluding remarks and further avenues of research are presented in section 5.

2. SMALL AREA ESTIMATION OF CROP-PRODUCTION

The efforts have been made by several research workers in the past [Panse *et al*. (1966), Singh (1968), Stasny *et al*. (1991), Srivastava *et al*. (1999), Singh and Goel (2000), Sisodia and Singh (2001), Sharma *et al*. (2004) etc.] to develop methodology for estimation of crop-production at Block/Panchayat level. We describe a scale down approach using multiple regression model due to Sisodia and Singh (2001) to obtain the Block level estimate from the District level crop-production statistics. We first assume the availability of auxiliary variables that are related to the crop-production at both District and Block level. We then postulate a regression model between the crop-production and auxiliary variables at District level as follows

$$Y_i = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \varepsilon_i \qquad (1)$$

where Y_i is the crop-production in the i^{th} year, (i = 1, 2, 3, ..., n), X_{ij} is the value of j^{th} auxiliary variable, (j = 1, 2, 3, ..., p), in the i^{th} year, $\beta' = (\beta_0, \beta_1, \beta_2, ..., \beta_p)$ is vector of unknown parameters and ε_i is error term. It is assumed that ε_i follows normal distribution with mean 0 and variance σ^2. Let the fitted model be denoted as,

$$\hat{Y}_i = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j X_{ij} \qquad (2)$$

where $\hat{\beta}_j$'s are the least square estimate of $\hat{\beta}_j$'s (j = 0, 1, 2, ... p).

Following Montegomery and Peck (1982), sum of squares due to regression, *i.e.* $SS_R(\beta_1, \beta_2, ..., \beta_p|\beta_0)$ is decomposed to define weight for the j^{th} auxiliary variable on the basis of its relative contribution in the model as follows.

$$w_j = \frac{\text{SS due to } j^{\text{th}} \text{ auxiliary variable}}{SS_R\ (\beta_1, \beta_2, ..., \beta_p | \beta_0)} \qquad (3)$$

Let $\bar{Y}$ be the average yield of a crop at district level in a given year based on crop cutting experiment and A be the corresponding area under that crop during the same year. Let $\hat{Y}$ be the estimated crop-production obtained from the fitted model (2) during the same year. Thus, the estimated average yield of that crop based on fitted model (2) in the same year is defined as $\hat{\bar{Y}} = \hat{Y} / A$.

Using weights in (3), an estimator of crop-production Y_q of block q for given year is constructed as follows

$$\hat{Y}_q = \left[\sum_{j=1}^{p} w_j x_j(q)\right] \hat{\bar{Y}};\ q = 1,\ 2,\ ...,\ Q \qquad (4)$$

where Q is total number of blocks in a given district and X_j is the value of j^{th} auxiliary variable at q^{th} block level in that year. The natural choice could be $\bar{Y}$ in place of $\hat{\bar{Y}}$ in (4). However, since standard error of $\bar{Y}$ is not being reported in statistical bulletins of State Govt./Union Govt., $\hat{\bar{Y}}$ is used in place of $\bar{Y}$ as standard error of $\hat{\bar{Y}}$ can be obtained by the fitted model (2), which is required to find out standard error $\hat{Y}_q$. The estimator $\hat{Y}_q$ is, in fact, conditionally an unbiased estimator of Y_q for given $\sum_{j=1}^{p} w_j X_j$ with respect to the block q since under model (1) expected value of $\hat{\bar{Y}}$ is $\bar{Y}$. A procedure to find out stable value of W_j is also suggested in the paper. The variance of $\hat{Y}_q$ is derived as follows:

$$\mathrm{V}(\hat{Y}_q) = \mathrm{V}\left[\left(\sum_{j=1}^{p} W_j X_j(q)\right)\hat{\bar{Y}}\right]$$

$$= \delta_q^2 \mathrm{V}\left(\frac{\hat{Y}}{A}\right), \quad \delta_q = \sum_{j=1}^{p} W_j X_j(q),$$

being constant for block q

$$= \left(\delta_q / A\right)^2 \mathrm{V}\left(\hat{Y}\right) \qquad (5)$$

where A is constant because it is area under the crop in a given year. The variance of $\hat{Y}$ can easily be computed from the fitted model (2), which is equal to $\hat{\sigma}^2$, the estimated error variance. It is obvious that, in general, $\sum_{q=1}^{Q} \hat{Y}_q \neq Y$, where Y is the actual crop production reported at district level through CCEs in a given year. Thus, a scaled estimator of Y_q proposed by Sisodia and Singh (2001) is

$$\tilde{Y}_q = a_q \hat{Y}_q \tag{6}$$

where a_q are constant such that $\sum_{q=1}^{Q} \tilde{Y}_q = \sum_{q=1}^{Q} a_q \hat{Y}_q = Y$.

Two alternative choices of a_q suggested by them are given by

(i) $a = Y / \sum_{q=1}^{Q} \hat{Y}_q$; if $a_q = a$; for all q = 1, 2, .., Q (7)

and

(ii) $a_q = 1 + \left(Y - \sum_{q=1}^{Q} \hat{Y}_q \right) / Q\hat{Y}_q$ (8)

by minimizing sum of squares of differences $(\tilde{Y}_q - \hat{Y}_q)$ subject to the condition that $\sum_{q=1}^{Q} a_q \hat{Y}_q = Y$.

Using the choice of a_q in (7) and (8), two improved scaled estimators of Y_q, are given by

$$\tilde{Y}_q^{(1)} = \hat{Y}_q \left(\frac{Y}{\sum_{q=1}^{Q} \hat{Y}_q} \right) \tag{9}$$

$$\tilde{Y}_q^{(2)} = \hat{Y}_q + \left(Y - \sum_{q=1}^{Q} \hat{Y}_q \right) / Q \tag{10}$$

with their variances

$$\mathrm{V}(\tilde{Y}_q^{(1)}) = a^2 (\delta_q / A)^2 \mathrm{V}(\hat{Y}) \tag{11}$$

and $$\mathrm{V}(\tilde{Y}_q^{(2)}) = \frac{Q-1}{Q} (\delta_q / A)^2 \mathrm{V}(\hat{Y}) \tag{12}$$

Remarks

(i) The scaled estimator $\tilde{Y}_q^{(1)}$ is not an unbiased estimator as claimed by Sisodia and Singh (2001). Thus, the expression in (11) is not correct.

(ii) Although $\tilde{Y}_q^{(2)}$ is unbiased estimator but its variance expression in (12) is not correct as it has been derived assuming that the $\mathrm{V}(\hat{Y}_q)$ is same for all q, q = 1, 2, ... Q, which is, in fact, not true.

Therefore, the bias and mean square error (MSE) of $\tilde{Y}_q^{(1)}$ and correct version of the variance of $\tilde{Y}_q^{(2)}$ will be derived in section 3 of the paper.

It is obvious that the value of w_j depends upon the fitted model (1) based on a set of values of Y and X_{ij}'s. Any variation in the set of Y and X_{ij}'s would cause change in w_j. Therefore, the value of w_j can not be unique. Thus, the estimate $\hat{Y}_q$ and its standard error would be subject to change depending on the set of Y and X_{ij}'s that has been chosen by an investigator. Therefore, it will be worthwhile to find out the stable value of w_j in order to find out the stable estimates $\hat{Y}_q$. An iterative technique is suggested to find out the stable value of w_j in this paper. In this technique, the model (1) is fitted with data set, initially on n points of time and w_j are calculated. The process is continued with the data set on (n + 1), (n + 2), (n + 3) ... points of time unless the value of w_j is stabilized. These stabilized values of w_j are considered as an approximately constant, although they are subject to the little error as they results from the fitted model (2), in investigating the properties of the estimators. These stabilized values of w_j are used in the estimator $\hat{Y}_q$ and subsequently in scaled estimators in order to find out stable estimates of Y_q in the empirical study presented in section 4.

Two more choices of w_j are also proposed in this paper. Firstly, the w_j is defined on the basis of partial correlation coefficient between Y and auxiliary variables and secondly, the w_j is defined on the basis of standardized regression coefficient of auxiliary variables. The stable values of w_j on the basis of partial correlation coefficients and standard regression coefficients are also obtained in the empirical study using the aforesaid iterative technique.

3. PROPOSED NEW SCALED ESTIMATOR

Another choice of a_q is obtained by minimizing sum of squares of relative differences $(\tilde{Y}_q - \hat{Y}_q)/\hat{Y}_q$ subject to the condition $\sum_{q=1}^{Q} a_q \hat{Y}_q = Y$, which is given by

$$a_q = 1 + \frac{Y - \sum_{q=1}^{Q} \hat{Y}_q}{\sum_{q=1}^{Q} \hat{Y}_q^2} \hat{Y}_q \qquad (13)$$

A new proposed scaled estimator is then given by

$$\tilde{Y}_q^{(3)} = \left[1 + \frac{Y - \sum_{q=1}^{Q} \hat{Y}_q}{\sum_{q=1}^{Q} \hat{Y}_q^2} \hat{Y}_q\right] \hat{Y}_q = \hat{Y}_q + \frac{\hat{Y}_q^2}{\sum_{q=1}^{Q} \hat{Y}_q^2}\left(Y - \sum_{q=1}^{Q} \hat{Y}_q\right) \qquad (14)$$

Obviously, the scaled estimator $\tilde{Y}_q^{(3)}$ is obtained by adjusting the original estimates $\hat{Y}_q$ by adding a factor which is a proportion of the difference between district total Y and the sum of the original estimates of block. The proportion here is based on the squared values of the original estimates $\hat{Y}_q$. Note that $\tilde{Y}_q^{(3)}$ is not an unbiased estimator of Y_q as it is evident from the expression (14) that $E(\tilde{Y}_q^{(3)}) \neq Y_q$.

3.1 Bias and Mean Square Error (MSE) and Variance of Estimators

First, we derive the bias and MSE of $\tilde{Y}_q^{(1)}$. Expressing $\tilde{Y}_q^{(1)}$ in the form

$$\tilde{Y}_q^{(1)} = c\hat{Y}_q, \quad \text{where} \quad c = \left(Y / \sum_{q=1}^{Q} \hat{Y}_q\right) \qquad (15)$$

We get the expected value of (15) as $E\left[\tilde{Y}_q^{(1)}\right] = cY_q$

The bias in $\tilde{Y}_q^{(1)}$ is, therefore, given by

$$\text{Bias}\left[\tilde{Y}_q^{(1)}\right] = \frac{\left(Y - \sum_{q=1}^{Q} \hat{Y}_q\right) Y_q}{\sum_{q=1}^{Q} \hat{Y}_q} \qquad (16)$$

MSE of $\tilde{Y}_q^{(1)}$ is obtained as

$$\begin{aligned} \text{MSE}\left[\tilde{Y}_q^{(1)}\right] &= \left[\text{Bias}\left(\tilde{Y}_q^{(1)}\right)\right]^2 + V\left(\tilde{Y}_q^{(1)}\right) \\ &= (c-1)^2 Y_q^2 + c^2 V\left(\hat{Y}_q\right) \end{aligned}$$

Substituting the value of c in above expression, we get

$$\text{MSE}\left[\tilde{Y}_q^{(1)}\right] = \frac{\left(Y - \sum_{q=1}^{Q} \hat{Y}_q\right)^2 Y_q^2}{\left(\sum_{q=1}^{Q} \hat{Y}_q\right)^2} + \frac{Y^2}{\left(\sum_{q=1}^{Q} \hat{Y}_q\right)^2} V\left(\hat{Y}_q\right) \qquad (17)$$

The correct expression for the variance of $\tilde{Y}_q^{(2)}$ is derived as follows,

$$\begin{aligned} V\left[\tilde{Y}_q^{(2)}\right] &= V\left[\hat{Y}_q + \left(Y - \sum_{q=1}^{Q} \hat{Y}_q\right)/Q\right] \\ &= V\left(\hat{Y}_q\right) + \frac{1}{Q^2}\sum_{q=1}^{Q} V\left(\hat{Y}_q\right) - 2\frac{V\left(\hat{Y}_q\right)}{Q}, \\ &\qquad \text{as } \text{Cov}\left(\hat{Y}_q, \hat{Y}_{q'}\right) = 0 \\ &= \frac{Q-2}{Q} V\left(\hat{Y}_q\right) + \frac{1}{Q^2}\sum_{q=1}^{Q} V\left(\hat{Y}_q\right) \end{aligned} \qquad (18)$$

We now derive the bias of $\tilde{Y}_q^{(3)}$ by taking expectation of (14) and we get

$$E\left[\tilde{Y}_q^{(3)}\right] = Y_q + YE\left(\frac{\hat{Y}_q^2}{\sum_{q=1}^{Q} \hat{Y}_q^2}\right) - E\left(\frac{\hat{Y}_q^2 \sum_{q=1}^{Q} \hat{Y}_q}{\sum_{q=1}^{Q} \hat{Y}_q^2}\right)$$

So, bias in $\tilde{Y}_q^{(3)}$ is given by

$$\begin{aligned} \text{Bias}\left[\tilde{Y}_q^{(3)}\right] &= E\left[\tilde{Y}_q^{(3)}\right] - Y_q \\ &= YE\left(\frac{\hat{Y}_q^2}{\sum_{q=1}^{Q} \hat{Y}_q^2}\right) - E\left(\frac{\hat{Y}_q^2 \sum_{q=1}^{Q} \hat{Y}_q}{\sum_{q=1}^{Q} \hat{Y}_q^2}\right) \end{aligned} \qquad (19)$$

Evaluation of (19) is not so tractable. We also note from (19) that if $\sum_{q=1}^{Q}\hat{Y}_q = Y$, then bias is exactly zero. However, in case $\sum_{q=1}^{Q}\hat{Y}_q \neq Y$, and if the difference $\left(\sum_{q=1}^{Q}\hat{Y}_q - Y\right)$ is expected to be nominal, then RHS of (19) is generally expected to be very close to zero. Hence, $\tilde{Y}_q^{(3)}$ could be considered as an almost unbiased estimator. If bias is negligible, then the variance of $\tilde{Y}_q^{(3)}$ is given by

$$\mathrm{V}\left[\tilde{Y}_q^{(3)}\right] = a_q^2 \mathrm{V}\left(\hat{Y}_q\right) \tag{20}$$

where a_q is given by the expression (13). Although a_q is based on the estimates $\hat{Y}_q$ for the q^{th} block, for the time being it can be considered as a constant for a given block. Thus, an approximate MSE of $\tilde{Y}_q^{(3)}$ is obtained as

$$\mathrm{MSE}\left[\tilde{Y}_q^{(3)}\right] \cong \left[1 + \frac{Y - \sum_{q=1}^{Q}\hat{Y}_q}{\sum_{q=1}^{Q}\hat{Y}_q^2}\hat{Y}_q\right]^2 \mathrm{V}\left(\hat{Y}_q\right) \tag{21}$$

4. EMPIRICAL INVESTIGATION

The time series data on production of wheat, area under wheat, irrigated area under wheat, gross cropped area and fertilizer consumption (*N*, *P*, *K*) in kg/hectare pertaining to the period 1980-81 to 2003-04 for Barabanki district were taken from the Bulletin of Agricultural Statistics, published by Directorate of Agricultural Statistics and Crop Insurance, Govt. of Uttar Pradesh, Lucknow (India). Since most of the area under wheat are irrigated, the area under wheat is not included in the model. It may be noted that the fertilizer consumption (*N*, *P*, *K*) is not being reported crop wise, hence the total fertilizer consumption in a year was apportioned for the wheat crop. Approximately 40% of the total fertilizer consumption in a year is considered to have been used for wheat crop.

The block wise data on the auxiliary variables in Barabanki district were also taken for the year 1996-97 from District Statistical Bulletin, published by Directorate of Economics and Statistics, Govt. of Uttar Pradesh, India. The block wise actual data of wheat production based on CCEs in Barabanki district for the year 1996-97 were also available in District Statistical Bulletins as a part of pilot studies which were used for the validation of the results.

We consider the following model at district level

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i \tag{22}$$

where, Y_i is the wheat production, X_{i1} is the irrigated area under wheat crop, X_{i2} is the fertilizer consumption (kg/ha) for wheat crop and X_{i3} is the relative area under wheat crop as percent to gross cropped area of Barabanki district, Uttar Pradesh (India) in the i^{th} year. Here β_j 's (j = 0, 1, 2, 3) are unknown parameters and ε_i is the random error component distributed normally with mean 0 and variance σ^2. The model (22) was fitted with the time series data as mentioned above.

4.1 Results Based on w_j Obtained from Relative Contribution of Individual Auxiliary Variable to Total Sum of Square due to Regression

The estimates of regression coefficients, their standard errors and value of coefficient of determinations (R^2) etc. are presented in Table 1.

Table 1. The estimate of regression coefficients, their standard error and R^2

Variable	Regression Coefficient	Standard error	R^2 (%)
X_1	41.615135**	12.91245	91.25**
X_2	10710.849*	1591.633	
X_3	–77598.92	43723.51	
Constant	–1277636		

*$p < 0.05$, **$p < 0.01$

Standard error of estimates $\hat{Y}$ = 233861.24

The irrigated area of wheat and fertilizer consumption have shown positive and significant effect on wheat production in the district. The value of coefficient of determination (R^2) was found to be quite high, *i.e.* 91.25%, which is indicative of the fact that these variables included in the model are quite sufficient to explain the variability in the data of wheat production at district level.

In order to find out the relative contribution of individual auxiliary variable, the first auxiliary variable namely irrigated area under wheat (X_1) was included in the model followed by fertilizer consumption in kg/ha (X_2) and percent relative area under wheat to the gross cropped area (X_3) depending upon the order of magnitude of correlation coefficients between auxiliary variable and crop production (y). On the basis of their relative contribution, the stable values of weights w_j as defined in previous section were calculated through iteration technique. In this technique, the model was fitted initially with 15 years data starting from 1980-81 to 1995-96. The process is continued with data increasing year by year and the stable value of w_j were found at 19th year starting from 1980-81 to 2000-01, which is presented in Table 2.

Table 2. Contribution of individual variable towards sum of squares due to regression and the stable value of weights (w_j)

S.No.	Variables	Contribution of variable	w_j
1.	X_1	$SSR(\beta_1 \mid \beta_0) = 4.242 \times 10^{12}$	0.66
2.	X_2	$SSR(\beta_2 \mid \beta_0, \beta_1) = 2.028 \times 10^{12}$	0.32
3.	X_3	$SSR(\beta_3 \mid \beta_0, \beta_1, \beta_2) = 0.164 \times 10^{12}$	0.02
	Total	$SSR(\beta_1, \beta_2, \beta_3 \mid \beta_0) = 6.434 \times 10^{12}$	1.00

Table 3. Block estimates of wheat production based on different estimators, standard error and their overall average error during the year 1996-97

S.N.	Block	Actual Prod. (qt) Y	Block Estimates				% Standard Error			
			$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$	$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$
1.	Dewan	262653	287459	283684	284240	283104	3.67	3.86	3.59	3.67
2.	Harakh	257177	271468	267903	268249	267584	3.82	4.02	3.74	3.82
3.	Barabanki	180724	225148	222191	221929	222476	3.65	3.80	3.65	3.65
4.	Masauli	182515	201456	198810	198237	199317	4.13	4.31	4.14	4.13
5.	Dariyabad	258617	235698	232603	232479	232770	4.27	4.51	4.20	4.27
6.	Banikodar	289730	278453	274796	275234	274366	3.81	4.05	3.73	3.81
7.	Pure Dalai	212968	236587	233480	233368	233637	2.88	3.12	2.94	2.88
8.	Mavai	217740	245838	242610	242619	242653	3.52	3.71	3.50	3.52
9.	Fatehpur	238394	257824	254438	254605	254321	3.94	4.13	3.87	3.94
10.	Nindura	265905	275486	271868	272267	271486	4.90	5.06	4.74	4.90
11.	Ramnagar	228755	201458	198812	198239	199319	4.45	4.70	4.42	4.45
12.	Suratganj	313909	285467	281718	282248	281172	4.17	4.42	4.06	4.17
13.	Haidargarh	241654	265483	261997	262264	261768	4.62	4.77	4.49	4.62
14.	Siddhaur	226005	245682	242456	242463	242501	4.24	4.41	4.16	4.24
15.	Trivediganj	209903	186594	184144	183375	184759	5.10	5.32	5.06	5.10
16.	Sirauli	233011	201658	199010	198439	199515	3.79	4.09	3.82	3.79
17.	Rudauli	292860	265487	262001	262268	261772	4.41	4.65	4.30	4.41
	Total	4112520	4167246	4112522	4112520	4112520				
Overall Average Error (E_i)								3245	3219	3315

Using these weights and block wise data on X_1, X_2, and X_3 the block estimates of wheat production based on four estimators, their percent standard error and an overall average error (E_i) were computed for the year 1996-97 and are presented in Table (3).

The percent standard error in the estimates was calculated as

$$\%\text{ Standard error} = \frac{\sqrt{\text{Variance/ MSE of estimator}}}{\text{Estimate of } Y_q} \times 100$$

The overall average error in $\tilde{Y}_q^{(i)}$ as compared to $\hat{Y}_q$ was calculated as

$$E_i = \sqrt{\frac{\sum_{q=1}^{Q}\left(\hat{Y}_q - \tilde{Y}_q^{(i)}\right)^2}{Q}} \quad \text{for } i = 1, 2, 3$$

The results presented in Table (3) show that the block estimates obtained from four different estimators are subject to maximum of almost 5% standard error. The percent standard error for block estimates varied between 2.88 to 5.10% in case $\hat{Y}_q$ followed by 3.12 to 5.32% for $\tilde{Y}_q^{(1)}$, 2.94 to 5.06% for $\tilde{Y}_q^{(2)}$ and 2.88 to 5.10% for $\tilde{Y}_q^{(3)}$. It shows that the range of percent standard error for block estimates is smaller for $\tilde{Y}_q^{(2)}$ as compared to other estimators. An overall average error has also been found to be 3219 in case of $\tilde{Y}_q^{(2)}$ which is smaller as compared to other estimators. Therefore, it can be concluded that the $\tilde{Y}_q^{(2)}$ is the best scaled improved estimator. It may also be observed from Table (3) that the block estimates obtained from the estimators $\tilde{Y}_q^{(1)}$, $\tilde{Y}_q^{(2)}$ and $\tilde{Y}_q^{(3)}$ have been almost at par with the actual value of the block production except in few blocks.

4.2 Results Based on w_j Obtained from Partial Correlation Coefficients

The stable weights w_j using iteration technique were also worked out on the basis of partial correlation coefficient between Y and X_1, X_2, X_3 and are presented in Table (4). The weights w_j were calculated as

$$w_j = \frac{\text{Partial correlation coefficient between } Y \text{ and } X_j}{\text{Sum of partial correlation coefficients between } Y \,\&\, X_1, Y \,\&\, X_2 \text{ and } Y \,\&\, X_3}$$

Table 4. Partial correlation coefficient between Y and X_1, X_2, X_3 and the value of weights (w_j)

S.N.	Variables	Partial Correlation Coefficient with Y	w_j
1.	X_1	0.605	0.57
2.	X_2	0.846	0.79
3.	X_3	–0.386	–0.36
	Total	1.065	1.00

Using these weights and block wise data on X_1, X_2, and X_3, the block estimates of wheat production based on four estimators, their percent standard error and an overall average error (E_i) were computed for the year 1996-97 and are presented in Table (5).

The magnitudes of percent standard error of block estimates have been found to be less to the order of maximum 3.48% for $\tilde{Y}_q^{(2)}$ as compared to other estimators (Table 5). Overall average error has also been minimum for $\tilde{Y}_q^{(2)}$ as compared to other estimators. Thus, $\tilde{Y}_q^{(2)}$ is preferred over other scaled improved estimators. It can also be seen from Table 5 that the block estimates obtained from the estimators $\tilde{Y}_q^{(1)}$, $\tilde{Y}_q^{(2)}$ and $\tilde{Y}_q^{(3)}$ have been found to be at par with the actual value of the block production except in few blocks.

Table 5. Block estimates of wheat production based on different estimators, their standard error and overall average error during the year 1996-97

S.N.	Block	Actual Prod. (qt) Y	Block Estimates				% Standard Error			
			$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$	$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$
1.	Dewan	262653	144617	254713	249180	257473	5.70	5.71	3.20	5.70
2.	Harakh	257177	142639	251230	247203	252431	5.70	5.71	3.19	5.70
3.	Barabanki	180724	114975	202506	219539	186310	5.70	5.71	2.94	5.70
4.	Masauli	182515	116033	204368	220596	188685	5.70	5.71	2.95	5.70
5.	Dariyabad	258617	137869	242828	242432	240438	5.70	5.71	3.15	5.70
6.	Banikodar	289730	145307	255929	249871	259244	5.70	5.71	3.21	5.70
7.	Pure Dalai	212968	94438	166334	199002	142565	5.70	5.71	2.72	5.70
8.	Mavai	217740	118624	208932	223188	194558	5.70	5.71	2.98	5.70
9.	Fatehpur	238394	140061	246689	244624	245918	5.70	5.71	3.17	5.70
10.	Nindura	265905	183919	323935	288482	366451	5.70	5.71	3.48	5.70
11.	Ramnagar	228755	125503	221048	230067	210499	5.70	5.71	3.04	5.70
12.	Suratganj	313909	162811	286759	267375	305851	5.70	5.71	3.34	5.70
13.	Haidargarh	241654	167639	295261	272202	319287	5.70	5.71	3.38	5.70
14.	Siddhaur	226005	143639	252992	248203	254975	5.70	5.71	3.20	5.70
15.	Trivediganj	209903	131816	232168	236380	225579	5.70	5.71	3.10	5.70
16.	Sirauli	233011	105870	186469	210434	166354	5.70	5.71	2.85	5.70
17.	Rudauli	292860	159178	280359	263741	295904	5.70	5.71	3.32	5.70
	Total	4112520	2334937	4112520	4112520	4112520				
Overall Average Error (E_i)								105974	104563	109943

4.3 Results Based on w_j Obtained from Standardized Regression Coefficients

The stable weights w_j using iteration techniques were calculated on the basis of standardized regression coefficients of auxiliary variables X_1, X_2 and X_3. The weights w_j's have been computed by

$$w_j = \frac{\text{Values of standardized regression coefficient of the regressor variable } X_j}{\text{Sum of the values of standardized regression coefficients of all regressor variables}}$$

and are presented in Table 6.

Table 6. The stable values of standardized regression coefficients of auxiliary variables X_1, X_2, X_3 and the weights (w_j)

S.N.	Variables	Standardized Regression Coefficient	w_j
1.	X_1	0.297	0.33
2.	X_2	0.739	0.83
3.	X_3	–0.142	–0.16
	Total	0.894	1.00

Using these weights and block wise data on X_1, X_2, and X_3, the block estimates of wheat production based on four estimators, their percent standard error and the overall average error (E_i) have been computed for the year 1996-97 and are presented in Table 7.

The perusal of results in Table 7 reveal that the block estimates based on $\tilde{Y}_q^{(2)}$ have been subject to the maximum of 2% standard error as compared to other estimators, where others accounted for about 5.50% standard error. An overall average error is also minimum in case of $\tilde{Y}_q^{(2)}$. Hence, $\tilde{Y}_q^{(2)}$ is again the best-scaled improved estimator. It may also be noted from Table 7 that the block estimates obtained from the estimators, $\tilde{Y}_q^{(1)}$, $\tilde{Y}_q^{(2)}$ and $\tilde{Y}_q^{(3)}$ have been found to be very close to the actual value of block production except for few blocks.

Table 7. Block estimates of wheat production based on different estimators, their standard error and overall average error during the year 1996-97

S.N.	Block	Actual Prod. (qt) Y	Block Estimates				% Standard Error			
			$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$	$\hat{Y}_q$	$\tilde{Y}_q^{(1)}$	$\tilde{Y}_q^{(2)}$	$\tilde{Y}_q^{(3)}$
1.	Dewan	262653	86152	253962	245994	258025	5.54	5.55	1.88	5.54
2.	Harakh	257177	85169	251065	245012	253143	5.54	5.55	1.86	5.54
3.	Barabanki	180724	69872	205972	229715	182926	5.54	5.55	1.65	5.54
4.	Masauli	182515	70185	206894	230027	184254	5.54	5.55	1.66	5.54
5.	Dariyabad	258617	82129	242103	241972	238326	5.54	5.55	1.82	5.54
6.	Banikodar	289730	86317	254449	246159	258850	5.54	5.55	1.88	5.54
7.	Pure Dalai	212968	56951	167881	216793	132056	5.54	5.55	1.46	5.54
8.	Mavai	217740	70626	208192	230468	186130	5.54	5.55	1.67	5.54
9.	Fatehpur	238394	83742	246856	243584	246131	5.54	5.55	1.85	5.54
10.	Nindura	265905	108766	320625	268609	382712	5.54	5.55	2.15	5.54
11.	Ramnagar	228755	76253	224780	236095	210896	5.54	5.55	1.75	5.54
12.	Suratganj	313909	96579	284698	256421	312571	5.54	5.55	2.01	5.54
13.	Haidargarh	241654	99476	293238	259318	328622	5.54	5.55	2.04	5.54
14.	Siddhaur	226005	86027	253592	245869	257400	5.54	5.55	1.88	5.54
15.	Trivediganj	209903	79255	233631	239098	224711	5.54	5.55	1.79	5.54
16.	Sirauli	233011	63686	187735	223528	157606	5.54	5.55	1.56	5.54
17.	Rudauli	292860	93915	276846	263757	298158	5.54	5.55	1.97	5.54
	Total	4112520	1395101	4112519	4112418	4112520				
Overall Average Error (E_i)								161833	159842	167413

5. CONCLUSION

Among all the scaled improved estimators, $\tilde{Y}_q^{(2)}$ has been found most precise estimator at block level. The overall average error has been found to be minimum when the choice of w_j is based on $\tilde{Y}_q^{(2)}$ partition sum of square due to regression. It may also be noted that the percent standard error of block estimates based on $\tilde{Y}_q^{(2)}$ have been found minimum up to the order of 2% when w_j are based on standardized regression coefficients. However, on the other hand, estimates based on initial estimator $\hat{Y}_q$ using weights derived from standardized regression coefficients have been found to be far away from the actual block estimates based on crop-cutting experiments. Therefore, it is recommended that the weight w_j be computed on the basis of portioning sum of square due to regression, and scaled estimator $\tilde{Y}_q^{(2)}$ be used for estimation of Block crop production Y_q in practice for practical application.

It may, however, be noted from the Table 3, 5 and 7 that some block estimates of crop production obtained from the estimators $\tilde{Y}_q^{(1)}$, $\tilde{Y}_q^{(2)}$ and $\tilde{Y}_q^{(3)}$ are far away from their actual block production. This is probably because of the local effects of the blocks. It is natural that all the blocks may not be uniform in terms of soil type, land type, ecological environment etc. If these factors could be indexed and taken into account while developing estimators for block estimates, one may get more precise block estimates, which may be the subject of further investigation.

ACKNOWLEDGEMENT

Authors are highly thankful to the referee for valuable comments and suggestions which have brought out considerable improvement in the text of the paper.

REFERENCES

Srivastava, A.K., Ahuja, D.L., Bhatia, D.K. and Mathur, D.C. (1999). Small area estimation techniques in agriculture. Souvenir in Advances in Agricultural Statistics and Computer Application, Indian Agricultural Statistics Research Institute, New Delhi, 32-45.

Singh, D. (1968). Double sampling and its application in agriculture, contributions in Statistics and Agricultural Sciences. Dr. V.G. Panse lecture, Indian Society of Agricultural Statistics, New Delhi.

Montgomery, D.C. and Peck, E.A. (1982). *Introduction to Linear Regression Analysis.* John Wiley & Sons, Canada.

Stasny, E.A, Goel, P.K. and Rumsey, D.J. (1991). *County estimates of wheat production*. Technical Report 458, Department of Statistics, The Ohio State University, Columbus, Ohio.

Singh, R. and Goel, R.C. (2000). *Use of remote sensing satellite data in crop surveys*. Technical Report, Indian Agricultural Statistics Research Institute, New Delhi.

Sharma, S.D., Srivastava, A.K. and Sud, U.C. (2004). Small area crop estimation for crop yield estimates at Gram Panchayat level. *J. Ind. Soc. Agril. Statist.*, **57,** 26-37.

Sisodia, B.V.S. and Singh, A. (2001). On small area estimation – An empirical study. *J. Ind. Soc. Agril. Statist.,* **54(3),** 303-316.

Panse, V.G., Rajagopalan, M. and Pillai, S. (1966). Estimation of crop yields for small areas. *Biometrics*, **66,** 374-388.