Jour. Ind. Soc. Ag. Statistics
51(2&3), 1998 : 265-282

Crop Yield and Acreage Statistics for Small Areas

B.D. Tikkiwal[1] and G.C. Tikkiwal
J.N.V. University, Jodhpur, India.

SUMMARY

In pursuance of recommendations by a C.S.O. working group [2], for developing suitable methodologies in order to meet the requirements for small area statistics; the two studies, one by Tikkiwal, B.D. on crop yield and the other by Tikkiwal, G.C. on crop acreage, were conducted in Rajasthan. After presenting first the various land-marks in the development of crop yield and acreage statistics in India, the findings of these two studies are discussed. It is brought out in the first study that it is possible to give reliable statistics at district and sub-district levels with direct methods only through an agency alternate to GCES. As regards panchayat samiti/CD-block level and also at lower AAO-circle level, the direct methods can still work in some cases; but, by and large, one may have to resort to Small Area Estimation methods including the SICURE model (Tikkiwal [20]). These are found useful in the second study by Tikkiwal [22]. In his study he uses estimators based on both direct and indirect estimators, referred to in the literature as composite estimators after reviewing the relevant work by Royall [9], Schaible [10] and Singh [12] and then building up a generalised class of composite estimators. The composite estimators considered earlier lie in this generalised class. The findings should be of help in future, in better organisation of current surveys and then in using a better estimation procedure therein, for crop yield and acreage statistics, in India and elsewhere, particularly in other developing countries where similar methodology is used.

Key Words: Simulation-cum-regression (SICURE) model, ANOVA model, Small areas, General crop estimation surveys (GCES), Improvement of crop statistics (ICS) scheme, Timely reporting scheme (TRS), Coefficient of variation (CV), Relative standard error, Absolute relative bias (ARB), Simulated relative standard error (Srse.)

1. Introduction

In the context of recent developments in small area methodology, the first attempt to build up agricultural statistics in India for small areas seems to have been made by Tikkiwal [19], [20] trying Simulation-Cum-Regression (SICURE)

---

1 Institute of Development Research and Statistics, B-45, Dev Nagar, Tonk Road, Jaipur, India.

model in crop estimation for small areas. Earlier attempt due to Panse, Rajgopalan and Pillai [6] to build up crop yield estimates at community development (CD) - block level through the classical double sampling regression method using eye-appraisal of large number of sample fields as auxiliary information did not succeed as the village agency Patwari, responsible for this work, could not do it satisfactorily along with his multifarious responsibilities.

A working group set up by the Central Statistical Organisation (C.S.O.) of Government of India on small area, while considering the need for developing suitable methodologies for meeting the need of data requirements for key characteristics, observed in its report [2], that, from the limited work with regard to the estimation of crop yield for a small area, it appears that the SICURE model is more promising than the various other methods. This is because unlike other methods, the model first artificially enlarges, through simulation, the size of the sample at the small area level and then uses the auxiliary information for building up multi-stage-cum-multiphase regression estimators of the type considered in the literature (Tikkiwal [17], Agarwal and Tikkiwal [1]). The details of how to use SICURE model for estimation of crop yield are given by Tikkiwal [20]. According to these details it in necessary in the beginning to conduct large number of crop cutting experiments in a given district to check on the efficacy of ANOVA-Model.

After presenting the land marks in the development of crop yield and acreage statistics in Sections 2 and 3; the details of the project and its findings are presented in sub-sections 4.1 and 4.2 of the paper. The findings are based on the data on wheat crop, obtained in the field study carried out by M & E Unit, Directorate of Agriculture, Government of Rajasthan in 1995-96 on our behalf. It comes out that the estimates of average yield, through GCES scheme, the only national scheme in India, for those crops, dealt with by above M & E unit also, have been on lower scale of corresponding M & E estimates all these years, partly because of the use of relatively less efficient procedure in GCES scheme. The M & E estimation procedure has further advantage in its being self weighting. Finally, the small area estimation regarding crop acreage statistics at the inspector land revenue circle level of a tehsil is dealt with in subsection 4.3.

2. *Land-Marks in the Development of Crop Yield Statistics*

A brief historical note on the development of methodology for crop estimation surveys for getting necessary yield statistics not only in India, but also elsewhere is presented. Some details of getting yield statistics and its limitations are discussed.

2.1 Historical Note

Before the development of survey methodology for estimation of crop yield in India and elsewhere there were only subjective methods for its estimation resulting in rather unreliable yield statistics. But, now the crop estimation surveys are being conducted in India and other developing countries through survey methodology developed mostly in 1940's (Mahalanobis [4], Sukhatme and Aggarwal [15]). Earlier there was some pioneering work done by Hubback in this area, to which a reference was made by Prof. R.A. Fisher in his memorandum dated 2nd March, 1945 addressed to the then Imperial Council of Agricultural Research, Government of India. It stated:

> "The use of the method of random sampling is theoretically sound. I may mention that its practicability, convenience and economy was demonstrated by an extensive series of crop-cutting experiments on paddy carried out by Hubback (later Sir John Hubback, Governor of Orissa) more than twenty years ago over a greater part of the rice tracts in Bihar and Orissa. So far as I know these were the earliest crop-cutting experiments based on the principle of random sampling anywhere in the world. They influenced greatly the development of my methods at Rothamsted." (From Mahalanobis [4], pp. 269).

2.2 Crop Yield Statistics in India and its Limitations

The overall responsibility for satisfactory conduct of crop estimation surveys in different states and union territories of India was initially entrusted to the Crop Survey Wing of the Statistical Branch of Indian Council of Agricultural Research. Later, in pursuance of a cabinet resolution in June, 1952; this role was entrusted to the Directorate of National Sample Survey (NSS), one of the biggest sample surveys in the world. This arrangement over the years continues. The functions of this central agency as in the case of former one, are

(i) Technical direction and overall coordination of crop estimation in different states,

(ii) Collaboration with the states in training of primary staff for their adherence to the procedural statistical requirements, and

(iii) Central check on the field work to bring lasting improvements in the agricultural statistics.

In addition to performing the above functions, the NSS started the practice of conducting land utilization surveys and crop estimation surveys, as part of its normal rounds, to build up estimates at all India level of its own from the tenth round onward in 1955-56. Subsequently several steps were taken to effect improvements in the system of crop estimation in the states. In view of the improvements so effected, it was decided to drop this practice in 1970-71 and adopt one official estimate of crop production in the country based on the estimates of different states as approved and authorized by the respective state Governments. The entire scheme of getting the production estimates in different states and then in the country as a whole is referred to as General Crop Estimation Surveys (GCES) scheme.

A scheme for Improvement of Crop Statistics (ICS) introduced from 1973-74, provides for supervisory checks through sub-samples on the data collected on both area and yield of various crops. Inspite of this scheme being in operation for so long, some serious problems of non-sampling errors remain as mentioned by Rao [8] in his technical address delivered in Society Conference in 1993 and by Srivastava and Mehta [14]. These errors will be all the more serious in building up crop yield and acreage statistics for small areas.

As regards the demand at present for increase in sample size to meet the small area needs, Rao in his paper (pp. 33) makes the following pertinent observation:

> "I think a ceiling should be imposed on the size of the GCES sample and on further increase should be allowed. In fact even the current scale of experiments (3,50,000 per annum) is excessive having regard to the lack of control over the field operations. I know there is growing demand for large sample sizes to furnish estimates for small areas especially in connection with the crop insurance scheme, but such needs should be met by separate arrangements by the agencies concerned and they should not form part of the GCES."

The solution to the various problems arising in building up reliable crop yield statistics presumably lies in the following:

(i) To reduce the work load on the field agency in temporarily settled states in order to make field operation for GCES more effective. For this, to adopt the practice of carrying out census of all fields (Girdawari) in all villages only once in five years in temporarily settled states, instead of every year and take other administrative measures to improve this agency. For replacing the present yearly census of all fields by a five yearly census, we have to modify the sample design of TRS to bring in the theory of repeat surveys suitably. This step will also reduce subjectivity in the working of this agency.

(ii) To promote the use of SICURE model (Described in Section 1) after necessary trial. This model can provide solution to the non-response problem arising out of the loss of planned crop-cutting experiments in GCES. Without this, all these years we have based yield statistics only on the remaining crop-cutting experiments, thus making our estimates biased and less efficient.

(iii) To strengthen TRS and EARAS schemes (Described in Section 3) in order to meet small area needs and the need of deeper stratification for measuring the impact of irrigation, type of seed etc.

It is to be noted that, while examining the functioning of state Monitoring and Evaluation (M & E) Units in different State Agricultural Departments of the country, Raheja and Jai Krishna [7] make pertinent recommendation to improve and broaden the activities of these units. As regards second phase of the M & E surveys for measuring yield rates of important crops for diagnostic purposes, they recommend suitable integration of these surveys with the GCES surveys. Such an integration should be welcome to effect necessary economy and efficiency. For this, the deeper stratification in (iii) has to take care of the needs of M & E surveys as well.

3. Land-Marks in the Development of Crop Acreage Statistics

In order to obtain yield and production statistics for a particular crop, reliable estimates of its acreage statistics are needed. The responsibility for collection of acreage statistics for different crops lies with the Departments of Land Revenue and Land Records in the country. The Indian territory for this purpose is categorized into (i) Temporarily settled states and Union territories, (ii) Permanently settled states, and (iii) The states and Union territories of the North East Region. These three categories, respectively, cover 88, 11 and 1 percent of the total area in the country. In the first category, there is a permanent agency of village accountants (Patwaris) which is responsible for collection of crop acreage statistics on complete enumeration basis; whereas in the second category, such data were collected through Grid Sampling till 1980-81 and after this, this data is being collected through a scheme known as 'Establishment of an Agency for Reporting Agricultural Statistics (EARAS)'. This scheme makes it possible to conduct area sample surveys on the lines of Timely Reporting Scheme (TRS), a scheme introduced earlier in category (i), in seventies itself for reducing time lag in the availability of crop area statistics. In the third category there is still the old traditional method of giving crop estimates based on personal assessments of village chowkidars, in absence of regular agency.

4. Meeting Small Area Needs for Crop Yield and Acreage Statistics

For decentralized planning and other purposes like crop insurance we need reliable agricultural statistics at least at the levels of district and CD - block. But the data thrown up by NSSO, TRS and EARAS are generally reliable at state level and not so at district level. It is demonstrated how to meet these small area needs in Rajasthan, one of the less developed state in India, as a case study.

4.1 An Examination of Reliability of Crop Yield Statistics in Rajasthan for Meeting Small Area Needs

The reliability of crop yield statistics is measured in terms of relative standard error or coefficient of variation in percentage of the estimated average yield per hectare. The theoretical consideration under normality model suggests that this percentage should be any where from 1.2 to 2.6 in order that the estimated average yield does not differ from the true average yield by more than 5 percent with 95 percent confidence. If this percentage is higher, then the error in the estimated average yield proportionally increases and thus the reliability of the estimated average yield decreases. It is felt that for good many practical situations this error upto 5 percent, and upto 10 percent in rare cases, may be tolerated.

The reliability of crop yield statistics in Rajasthan was examined in this light by Tikkiwal [20], for three crops: maize, paddy and wheat for five consecutive years (1981-82 to 1985-86). It was noted that the yield estimates for different crops at state level were satisfactory as expected, but the situation was bad at district level. This position prevails even now.

To meet the small area needs, the number of crop-cutting experiments needs to be increased considerably for yield estimation alone; which is costly and otherwise not desirable. The percentage of crop-cutting experiments correctly conducted during the period 1986-89 in Rajasthan was only between 33 and 45 for both Kharif and Rabi crops. This reflected a highly unsatisfactory state of affairs in its field operations at that time. The situation is reported to have improved considerably since then. Still, far more efficient estimates than required can be obtained through the state agency of M & E Unit in the Directorate of Agriculture, which conducts a much large number of crop-cutting experiments of its own for its M & E surveys every year. This we illustrate from the study conducted in Rajasthan, under our project on Small Areas.

4.2 Rajashthan Study for Crop Yield Statistics

The wheat, a rabi crop sown in Oct.-Nov. of the year and then harvested around March-April of the next year is a major crop in Rajasthan. The M & E Unit carried out the field study, on our behalf, for wheat crop during 1995-96 in Sikar district of the state, a district which is compact and has many on going progressive schemes. The sample design adopted for the study was slightly different from the one followed in the state for these M & E surveys since the beginning of these surveys form Kharif 1980-81. But this change in the sample design makes all the difference in putting the analysis of such surveys in the state on sound footing and further helps in scientific allocation of total sample of clusters (chaks) in a year to different strata, keeping in view the administratively convenience in distributing the work load evenly amongst the concerned field staff.

It may be noted that Rajasthan is the only state (and recently one more state, Karnataka) which independently conducts crop cutting experiments and as a result, the situation in the state is much better than other states in the country where M & E surveys are part of GCES. That, the situation is bad in other states as regards the precision of various estimates is clear from the report by Raheja and Jai Krishna [7]. The various estimates in these states do not reflect the situation properly even at district level whereas such estimates are needed atleast upto the CD-block / Panchayat samiti level. Thus, the selection of Rajasthan for our case study has added significance. This study should provide a good precedence for other states of the country and of other developing countries as well, where similar methodology is followed, in meeting their small area needs.

4.2.1 Sample design in the study

Before presenting the sample design used in the study relevant details are given for the M & E surveys in Rajasthan. A district is divided into certain number of sub-districts. Sikar district, is divided into two sub-districts: Sikar and Srimadhopur. A sub-district consists of certain number of AAO circles, each AAO circle has VEW circles, each VEW circle has chaks and finally a chak consists of certain number of farmers representing their families. Sikar sub-district consists of 9 AAO circles and Srimadhopur of 11 AAO circles. Each AAO circle in Sikar sub-district consists of 7 VEWs and Srimadhopur of 6 VEWs except one circle which consists of 5 VEWs. A mandal in each VEW circle of both the districts consists of 160 farmers. Each of the chaks forming a VEW circle has necessarily some mandal farmers and some non-mandal farmers.

At the time of launching this study, the sample design in use in Rajasthan's M & E surveys was stratified-multi-stage, the strata being the sub-district in the concerned district. The first stage units were the chaks in the sub-district, the second-stage units were the farmers and then ultimate units of sampling were plots of specified size in one of the fields of each selected farmer. The simple random sampling scheme was adopted at all stages of sampling; though some varying probabilities sampling scheme with replacement at the first-stage could have been tried profitably (Sukhatme and Narain [16], and Tikkiwal [18]). For each crop 50 chaks were taken in the sample; in turn, from each selected chak only one mandal farmer and one non-mandal farmer; and finally one plot for harvest from each type of farmer was selected. The size of this plot and the mode of its selection was the same as in GCES (Sec. 2.2) scheme.

The above sample design has been prevalent in Rajasthan from Kharif 1980-81 itself and is in conformity with the guidelines of the World Bank manual [23], strictly for monitoring surveys but easily extendable to M & E surveys as well. However, in order to have some even distribution of workload at AAO level, 50 chaks were selected from a given sub-district in a manner that a certain minimum number of chaks fell in each of AAO circles of the sub-district and that no two chaks fell in the same VEW circle. But this control over random selection invalidated the formula used for estimating the precision of the estimates of yield rate in these surveys. Weather this invalidation meant over estimation or under estimation of the precision was not immediately clear, as it meant technically arbitrary allocation of chaks to different AAO circles in a given sub-district.

In order to overcome this difficulty in our study, AAO circles in each of the two sub-districts are taken as strata, in place of sub-districts, with VEW cricles and chaks, as first-stage and second-stage units, the remaining stages of sampling remaining the same. However, this creates another difficulty, viz., it is no more possible to measure the exact precision of the various estimators of the average yield considered in the sub-section 4.2.2, as it is not possible to measure the variability between chaks within VEWs and between farmers within chaks with this sample design in which only one chak is selected within a selected VEW and then only one farmer within a given selected chak. The respective formulae for various estimators used in the study give over estimates of their precision. However, this difficulty is overcome in some other way.

4.2.2 The main findings of the study

Let $\overline{X}_{i.}$ denote the sample mean yield of a particular crop in i^{th} AAO-circle, for different i, of a particular sub-district. Let $\overline{X}_{st} = \sum_{i=1}^{S} W_i \overline{X}_{i.}$ denote the weighted mean yield of the crop in the sub-district, where S denotes the number of AAO circles in the sub-district and W_i is the weight associated with the mean yield for the crop in i^{th} AAO-circle. Let M_i and m_i denote respectively the total number of VEWs and the selected number of VEWs in i^{th} AAO-circle. Let A_i denote the area sown for the the crop in i^{th} AAO-circle and A the area sown for the crop in the sub district. Thus $A = \sum_{i=1}^{S} A_i$.

By taking W_i equal to m_i/n, A_i/A, M_i/N, and $1/S$, where $N = \sum_{i=1}^{S} M_i$ and $n = \sum_{i=1}^{S} m_i$ we get four different estimators of the mean yield of the crop in the sub-district. The first estimator here corresponds to the estimator used in M & E surveys in Rajasthan. The second estimator corresponds to the one used in GCES (NSSO Report [5]). It may be noted that if the total number n of VEWs selected from the sub-district is distributed over different AAO circles of the sub-district in proportion to the areas sown for the crop in different AAO circles, then the first estimator reduces to the second one. Table 1 and Table 2 give various estimators, their estimates, estimated variances $V(\overline{X}_{st})$, estimated gains in efficiency due to stratification and their estimated coefficient of variations (CVs) for both the sub-districts is Sikar district. In these tables $V(\overline{X})$ denotes the estimate of the variance of the estimator, when there is no stratification at sub-district level.

It may be noted from the tables that estimator used in M & E surveys has throughout maximum precision amongst all the four estimators considered here. But the estimator used in GCES behave differently. In case of Srimadhopur sub-district, it is second best for both Mandal and Non-mandal farmers, provided we ignore its CVs. But in case of Sikar district, it is poorest of the four estimators for both Mandal and Non-mandal farmers. Further, it under-estimates the average yield of the crop in both the sub-districts for Mandal farmers as well as Non-mandal farmers. Though, this under-estimation lies only between 4.2 and 0.23 percent of the corresponding M & E estimate, there is additional advantage in using M & E estimator. It makes the modified sample design

Table 1. Results regarding the wheat-crop-yield in the year 1995-96 for Mandal and Non-Mandal farmers in Sikar sub-district, Rajasthan

Estimator of the mean yield $\overline{X}_{st}$	Results for Mandal Farmers				Results for Non-Mandal Farmers			
	Value of $\overline{X}_{st}$	$\hat{V}(\overline{X}_{st})$	Estimated gain in efficiency due to stratification in percentage	Estimated coefficient of variation in percentage	Value of $\overline{X}_{st}$	$\hat{V}(\overline{X}_{st})$	Estimated gain in efficiency due to stratification in percentage	Estimated coefficient of variation in percentage
$\overline{X}_{st} = \frac{1}{n}\sum_{i=1}^{S} m_i X_{i.}$	7·817	0.0121	60.33	1.407	7.058	0.0115	97.39	1.519
$\overline{X}_{st} = \frac{1}{A}\sum_{i=1}^{S} A_i X_{i.}$	7.799	0.0157	23.57	1.607	6.765	0.0149	52.35	1.804
$\overline{X}_{st} = \frac{1}{N}\sum_{i=1}^{S} M_i X_{i.}$	7.869	0.0129	50.39	1.443	7.123	0.0120	89.17	1.538
$\overline{X}_{st} = \frac{1}{S}\sum_{i=1}^{S} X_{i.}$	7.869	0.0129	50.39	1.443	7.123	0.0120	89.17	1.538
	$\hat{V}(\overline{X}) = 0{\cdot}0194$				$\hat{V}(\overline{X}) = 0{\cdot}0227$			

Table 2. Results regarding the wheat-crop-yield in the year 1995-96 for Mandal and Non-Mandal farmers in Srimadhopur sub-district, Rajasthan

Estimator of the mean yield $\overline{X}_{st}$	Results for Mandal Farmers				Results for Non-Mandal Farmers			
	Value of $\overline{X}_{st}$	$\hat{V}(\overline{X}_{st})$	Estimated gain in efficiency due to stratification in percentage	Estimated coefficient of variation in percentage	Value of $\overline{X}_{st}$	$\hat{V}(\overline{X}_{st})$	Estimated gain in efficiency due to stratification in percentage	Estimated coefficient of variation in percentage
$\overline{X}_{st} = \frac{1}{n} \sum_{i=1}^{S} m_i X_{i.}$	9.078	0.0264	62.50	1.790	8.638	0.0180	50.00	1.553
$\overline{X}_{st} = \frac{1}{A} \sum_{i=1}^{S} A_i X_{i.}$	9.035	0.0308	39.29	1.942	8.555	0.0194	39.18	1.628
$\overline{X}_{st} = \frac{1}{N} \sum_{i=1}^{S} M_i X_{i.}$	9.640	0.0334	28.44	1.896	9.458	0.0216	25.00	1.554
$\overline{X}_{st} = \frac{1}{S} \sum_{i=1}^{S} X_{i.}$	9.650	0.0323	32.82	1.862	8.596	0.0209	29.19	1.682
	$\hat{V}(\overline{X}) = 0.0429$				$\hat{V}(\overline{X}) = 0.0270$			

used in the study self-weighting, which, in turn, makes tabulation and subsequent analysis of data easy.

It has been observed all these years that M & E estimators of average yield for different crops are higher than those obtained through GCES in Rajasthan. The above discussion explains why it happens this way.

It may be noted here that each of the four estimators gives estimates of sub-district yield with precision much more than required. We can easily build up estimates also for Panchayat samities in Sikar district and also of this district itself. The precision of these estimates should also be much more than required, particularly for the district. However, from Table 3 it is clear that it is not invariably so at AAO level. Considering that CV normally should not exceed 5 percent limit; the precision is as required in nineteen cases out of 40 in all. To bring down the CVs in the rest, we need to resort to small area estimation methods of the kind presented in the following sub-section. For this, we need to organize such field studies in series over some years.

It is observed in sub-section 4.2.1 that with the design used in the study it is not possible to measure the exact precision of the various estimators of the average yield for wheat crop, as it is not possible to measure the variability

Table 3. The coefficient of variations (CVs), in percentage, for the two types of farmers AAO-circle wise in the two sub-district of Sikar, Rajasthan

AAO	Srimadhopur Sub-district		AAO	Sikar Sub-district	
	CV for Mandal farmers	CV for Non-Mandal farmers		CV for Mandal farmers	CV for Non-Mandal farmers
1	0	0	1	0	0
2	5.284	4.978	2	5.796	6.276
3	10.570	6.384	3	5.221	5.387
4	3.611	5.176	4	1.083	4.297
5	3.346	3.561	5	3.493	2.123
6	5.453	6.994	6	6.483	3.750
7	6.270	5.480	7	6.346	6.306
8*	17.505	14.580	8	2.885	4.364
9	3.860	4.522	9	4.213	5.091
10	5.904	6.773			
11	5.422	4.541			

*In this circle No 8 of Srimadhopur, the CVs for both types of farmers based on three crop-cutting experiments in each of the type appear to be out-liers. Therefore, the data from this circle have not been included in the analysis of data presented in the previous two tables. But, even when we include this data in the analysis, the results regarding the relative performance of the two estimators, one used in GCES and the other in Rajasthan M & E unit remain unaffected.

between chaks within VEWs and between farmers within chaks. In future field studies for research suggested above we have to modify sample design further to take care of this limitation as well. However, it may be noted that above observations hold good inspite of this limitation; particularly for the first two estimators, one used in M & E surveys and the other in GCES, where there is so much difference between their estimated gains in efficiency due to stratification. As regards the exact precision of these estimators, their CVs should not exceed more than 5 percent for any sub-district in view of what follows.

From the first two tables of this sub-section, it is noted that there is considerable gain in efficiency due to stratification for all the estimators in the two sub-districts. This is going to remain so even when one takes into account the variability between chaks within VEWs and the variability between farmers within chaks, as the contribution to the overall variance of an estimator due to further stages of sampling after the first is generally much less. Thus, the assumption that the sample of chaks is randomly drawn directly from the sub-district itself, as was done in M & E surveys all along, will give an upper limit to the CV of an estimator. Therefore, through the formula used in M & E surveys, the upper limits of respective CVs are obtained as given in Table 4.

Table 4. The upper limits of the coefficient of variations in percentages for wheat crop in 1995-96 for the two types of farmers in the two sub-districts of Sikar, Rajasthan

Sub-district	Type of Farmer	
	Mandal	Non-Mandal
Sikar	3.34	1.12
Srimadhopur	4.31	4.18

4.3 Rajasthan Study for Crop Acreage Statistics

Some empirical studies by Schaible *et al* [11] conclude that when sample sizes for small domains are relatively small; the synthetic or indirect estimator out-performs the direct estimator. However, for relatively large sample sizes reverse happens. In absence of precise idea about the nature of sample size for a given situation, it becomes necessary to use composite estimators, which are weighted sums of the two types of estimators.

Defined below a generalized class of composite estimators, due to Tikkiwal [22] having earlier composite estimators as special cases. Seven estimators, direct and indirect are given special cases of the generalized class and their relative performance considered in our Crop Acreage Study carried out in a tehsil, only slightly bigger, on an average, than a CD-block or panchayat samiti.

Let n_a and N_a denote the sample units and population units in the small domain a. Let $\overline{Y}_a$, $\overline{X}_a$ and $\overline{Y}$, $\overline{X}$ denote respectively the sample means based on n_a and $n = \sum_a n_a$ observations, selected from N_a and $N = \sum_a N_a$ observations, for characteristics y and x. Correspondingly, let $\bar{y}_a$, $\bar{x}_a$ and $\bar{y}$, $\bar{x}$ denote respectively the population means based on N_a and N observations. Now, we define the generalized class of composite estimators for estimating the population mean $\bar{y}_a$ of the domain 'a' as follows:

$$\overline{Y}_{c,a} = w_a \, \overline{Y}_a (\overline{X}_a/\bar{x}_a)^{\beta_1} + (1 - w_a) \, \overline{Y} (\overline{X}/\bar{x}_a)^{\beta_2} \tag{4.3.1}$$

where β_1, β_2 and weights w_a are suitably chosen constants. The first component of the composite estimator is a direct estimator ($\overline{Y}_{d,a}$) due to Srivastava [13] and the second component is a generalized synthetic estimator ($\overline{Y}_{syn,a}$) due to Ghiya [3]. The best way to choose the values of β_1, β_2 and w_a is in which MSE ($\overline{Y}_{c,a}$) is minimum. This gives approximately

$$\beta_1 = -\frac{C_{xa\,ya}}{C_{xa}^2} \tag{4.3.2}$$

and

$$\beta_2 = \frac{\bar{y}(C_x^2 - 4C_{xy}) - 2\bar{y}_a \{ (C_x^2/2) - C_{xy} \}}{(4\bar{y}\, C_x^2 - 2\bar{y}_a C_x^2)} \tag{4.3.3}$$

in usual notations. The weight w_a is given by

$$w_a = \frac{\text{MSE}(\overline{Y}_{syn,a})}{\text{MSE}(\overline{Y}_{d,a}) + \text{MSE}(\overline{Y}_{syn,a})} \tag{4.3.4}$$

We assess the relative performance, empirically through a simulation study, of this generalized composite estimator $\overline{Y}_{c,a}$ with other estimators, direct and indirect, belonging to the general class. Given below is the list of all these estimators of population total T_a of small domain 'a', as this total has a meaning for crop acreage.

Direct Estimators

(1) Direct ratio estimator $t_{1,a} = N_a (\bar{Y}_a / \bar{X}_a) \bar{x}_a$

(2) Direct general estimator $t_{2,a} = N_a \bar{Y}_a (\bar{X}_a / \bar{x}_a)^{\beta_1}$

Indirect Estimators

(3) Ratio synthetic estimator $t_{3,a} = N_a (\bar{Y} / \bar{X}) \bar{x}_a$

(4) Generalized synthetic estimator $t_{4,a} = N_a \bar{Y} (\bar{X} / \bar{x}_a)^{\beta_2}$

(5) Generalized composite estimator $t_{5,a} = N_a \bar{Y}_{c,a}$

(6) Composite estimator $t_{6,a} = N_a [w_a \bar{Y}_a + (1 - w_a)(\bar{Y} / \bar{X}) \bar{x}_a]$

(7) Composite estimator $t_{7,a} = w_a t_{1,a} + (1 - w_a) t_{3,a}$

In simulation study, the problem of crop acreage estimation is undertaken for all seven Inspector Land Revenue Circles (ILRCs) of Jodhpur Tehsil of Rajasthan. These ILRCs are small domains from TRS point of view. The crop under consideration is Bajra (Indian Corn or Millet) for agriculture season 1993-94 and crop acreage for 1992-93 is taken as auxiliary characteristic x. Before simulation, first examine the synthetic assumptions of ratio synthetic estimator $t_{3,a}$ and generalized synthetic estimator $t_{4,a}$ with respect to the above small domains. These assumptions respectively are as follow:

$$\bar{y}_a (\bar{x}_a)^{\beta_2} \cong \bar{y} (\bar{x})^{\beta_2} \qquad (4.3.5)$$

for $\beta_2 = -1$ for estimator $t_{3,a}$ and for optimum value of β_2 given in (4.3.3) for estimator $t_{4,a}$. From this examination, it may be that both the assumptions closely meet in ILRCs 3, 5 and 7 and deviate moderately in ILRCs 4 and 6; whereas they deviate considerably in ILRCs 1 and 2.

Now taking villages as sampling units for simulation purpose and otherwise, five hundred independent simple random samples for each size of 25, 50, 63, 76 and 88 are selected from the population of 252 villages of Jodhpur Tehsil. Then to assess the relative performance of the estimators under consideration, their Absolute Relative Bias (ARB) and Simulated relative standard error (Srse) or simply coefficient of variation are calculated for each of ILRCs as follows:

$$\text{ARB}(t_{k,a}) = \frac{\left| \frac{1}{500} \sum_{s=1}^{500} t^s_{k,a} - T_a \right|}{T_a} \times 100 \qquad (4.3.6)$$

and

$$\text{Srse}\,(t_{k,a}) = \frac{\sqrt{\text{ASE}\,(t_{k,a})}}{E\,(t_{k,a})} \times 100 \tag{4.3.7}$$

where

$$\text{ASE}\,(t_{k,a}) = \frac{1}{500} \sum_{s=1}^{500} (t_{k,a}^{s} - T_a)^2$$

for k = 1, ..., 7 and a = 1, ..., 7

Results of ARB and Srse are presented in Table 5 only for n = 50. (A sample of 20 percent villages, as presently done in TRS); as the findings from other tables are similar.

Table 5. Simulated relative standard errors and absolute relative biases, in percentage, of various estimators for n = 50 in different ILRCs

ILRC / Estimator	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$t_{1,a}$	37.37 (0.21)	17.46 (2.28)	8.51 (0.76)	16.29 (0.13)	12.73 (2.41)	12.28 (0.32)	15.29 (2.78)
$t_{2,a}$	18.55 (0.96)	18.32 (1.50)	6.56 (0.12)	15.43 (0.18)	11.27 (1.12)	13.68 (0.54)	11.34 (0.61)
$t_{3,a}$	19.11 (17.90)	20.67 (19.50)	5.71 (0.72)	10.11 (8.66)	5.71 (0.05)	12.14 (11.03)	5.85 (1.02)
$t_{4,a}$	40.54 (39.67)	21.00 (19.84)	5.96 (0.11)	10.17 (8.68)	5.95 (0.18)	8.43 (4.06)	7.16 (3.97)
$t_{5,a}$	16.87 (5.70)	13.80 (9.29)	4.41 (0.10)	8.49 (6.21)	6.17 (0.88)	6.29 (2.82)	6.05 (3.09)
$t_{6,a}$	17.02 (0.91)	17.14 (1.20)	5.07 (0.11)	9.95 (0.17)	8.03 (1.09)	11.42 (0.51)	5.61 (0.59)
$t_{7,a}$	16.48 (8.40)	13.48 (10.20)	4.78 (0.50)	10.15 (6.30)	5.01 (0.38)	8.14 (4.60)	5.45 (1.20)

Note: The figure shown in the parentheses are the absolute relative biases in percentage.

For assessing relative performance of the various estimators, we adopt the thumb rule that at ILRC level, an estimator should not have Srse more than 10 percent and bias more than 5 percent. We note from Table 5 that none of the estimators satisfy the thumb rule in circles 1 and 2. This is happening because, in these circles, there is considerable deviation from the synthetic assumptions, as observed earlier. In ILRC 4, where the synthetic assumptions deviate moderately, $t_{6,a}$ alone satisfy the thumb rule. But, in ILRC 6, where too deviations are moderate; $t_{5,a}$ is best. $t_{5,a}$ is best in ILRC 4 too provided we restrict ourselves only to Srse. In circle 3, where the synthetic assumptions closely meet, $t_{5,a}$ is best here also. However, in ILRC 5 and 7, where too the synthetic assumptions closely meet, $t_{7,a}$ alone is competitive with others.

From the above analysis and otherwise, the use of composite estimators $t_{5,a}$ and $t_{7,a}$ at ILRC level and upto the level of district in TRS is recommended, when there is no considerable deviations from synthetic assumptions. When this condition is not satisfied one should look for other types of estimators such as those through SICURE Model (Tikkiwal [20]) and assess their relative performance through studies of the kind, in series, over some years for crop acreage estimation. We have already seen, in Sec. 4.2, a similar need for organization of further studies, in series, to develop small area methods for crop yield estimation too.

REFERENCES

[1] Agarwal, C.L. and Tikkiwal, B.D., 1980. Two-stage sampling on successive occasions. *Sankhya, C*, 31-44.

[2] C.S.O., 1990. Report of the working group on small area development programme statistics. *Department of Statistics, Ministry of Planning, Govt. of India, New Delhi.*

[3] Ghiya, A., 1997. On direct and indirect estimation in survey sampling. A Ph.D. thesis submitted to J.N.V. University, Jodhpur, India under the supervision of Dr. G.C. Tikkiwal.

[4] Mahalanobis, P.C., 1946. Sample surveys of crop yields in India. *Sankhya*, 269-280.

[5] NSSO Report, 1993-94. Consolidated results of crop estimation survey on principal crops. *National Sample Survey Organization, Department of Statistics, Ministry of Planning and Programme Implementation, Govt. of India.*

[6] Panse, V.G., Rajgopalan, M. and Pillai, S.S., 1966. Estimation of crop yield for small domains. *Biometrics*, 374-384.

[7] Raheja, S.K. and Jaya Krishna, 1991. A report on special study on functioning of state monitoring and evaluation units. *Sponsored by the Directorate of Extension, Ministry of Agriculture, Govt. of India.*

[8] Rao, V.R., 1993. Technical address: Statistics for rural development-An agenda for nineties. *J. Ind. Soc. Agric. Statist.*, **XLV (1)**, 1-47.

[9] Royall, R.A., 1977. Statistical theory of small area estimates - use of prediction models. *Unpublished report prepared under contract from the National Center for Health Statistics.*

[10] Schaible, W.L., 1978. Choosing weights for composite estimators for small area statistics. Proceedings of the survey Research Methods Section, *J. Amer. Statist. Assoc.*, Washington, D.C., 741-746.

[11] Schaible, W.L., Brock, D.B., Casady, R.J. and Schnack, G.A., 1977. An empirical comparison of the simple inflation, synthetic and composite estimators for small area statistics. *Proceedings of American Statistical Association, Social Statistics Section*, 1017-1021.

[12] Singh, D., 1993. Some studies in small area statistics. Ph.D. thesis, Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi.

[13] Srivastava, S.K., 1967. An estimator using auxiliary information in sample surveys. *Calcutta Statist. Assoc. Bull.*, 121-132.

[14] Srivastava, S.S. and Mehta, R., 1996. Qualitative issues in agricultural statistics. Souvenir, 50th Annual Conference of *Ind. Soc. Agril. Statist.*, 49-55.

[15] Sukhatme, P.V. and Aggarwal, O.P., 1946-47, 47-48. Report on the crop cutting survey on wheat by the random sampling method in Delhi. *Indian Council of Agricultural Research, New Delhi.*

[16] Sukhatme, P.V. and Narain, R.D., 1952. Sampling with replacement, *J. Ind. Soc. Agril. Statist.*, **4**, 42-49.

[17] Tikkiwal, B.D., 1965a. The theory of two-stage sampling on successive occasions. *J. Ind. Statist. Assoc.*, 125-136.

[18] Tikkiwal, B.D., 1965b. Some aspects of sampling with varying probabilities from a finite population. *University of Rajasthan Studies (Science Series)*, 31-52.

[19] Tikkiwal, B.D., 1991. Modelling through survey data for small domains. *The Key Note address at the symposium on modelling held at Kurukshetra University*, March 7-9, 1991.

[20] Tikkiwal, M.D., 1993. Modelling through survey data for small domains. *Proceedings International Scientific Conference on Small Area Statistics and Survey Design* (held in September, 1992 at Warsaw, Poland).

[21] Tikkiwal, B.D. and Tikkiwal, G.C., 1998. Small area estimation in India: Crop yield and acreage statistics. *Invited paper in the International Conference; "Agricultural Statistics-2000" at Washington, D.C.* (Invited paper hand-out - Day **2**, 127-142).

[22] Tikkiwal, G.C., 1997. On estimation of crop area statistics for small domains. An invited paper, presented in person before the 2nd Annual Conference of Information Theory and Applications and National Symposium in Statistics with Focus on Application in Agriculture and Allied Sciences, held at CCS Haryana Agricultural University, Hissar from February 7-9, 1997. Paper is under publication in the proceedings.

[23] World Bank Manual, 1981. The monitoring and evaluation of training and visit extension in India. *A manual of instructions*; Slade, R.H. and G. Feder.