The Pharma Innovation Journal 2022; SP-11(9): 1731-1736

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(9): 1731-1736

www.thepharmajournal.com
Received: 19-06-2022
Accepted: 13-07-2022

Archana
Division of Statistics and Computer Science, SKUAST, Jammu, Jammu and Kashmir, India

SEH Rizvi
Division of Statistics and Computer Science, SKUAST, Jammu, Jammu and Kashmir, India

M Iqbal Jeelani Baht
Division of Statistics and Computer Science, SKUAST, Jammu, Jammu and Kashmir, India

P Kumar
FSR, SKUAST, Jammu, Jammu and Kashmir, India

Manpreet Kour
Division of Agronomy, SKUAST, Jammu, Jammu and Kashmir, India

Vijay Khajuria
FSR, SKUAST, Jammu, Jammu and Kashmir, India

Corresponding Author:
Archana
Division of Statistics and Computer Science, SKUAST, Jammu, Jammu and Kashmir, India

Socio economic status and block wise estimation of milk production of Jammu district using small area estimation technique

Archana, SEH Rizvi, M Iqbal Jeelani Baht, P Kumar, Manpreet Kour and Vijay Khajuria

Abstract

The demand for small area statistics is growing day-by-day not only in public but also in private sectors, and small area estimation technique (SAE) is becoming very important in survey sampling due to the thrust of planning process has shifted from macro to micro level. Small Area Estimation technique has been applied for obtaining estimates of per day total milk production at block levels in the Jammu district of J&K state. For this purpose stratified three stage random sampling plan was adopted with blocks constituting the strata, villages as the primary stage units, households possessing livestock as the secondary stage units and cows/buffaloes in milk in the selected households as the third stage units. A cross classified structure with blocks as ‘small areas’ and groups as cows and buffaloes was constituted. The block wise estimates of total milk production of cows and buffaloes per day were estimated through conventional estimators of this study. The average per day milk production (lt.) of all 20 blocks through direct, synthetic and composite small area estimation estimators showed that the total estimate of milk production for Jammu district through direct estimator was found to be 13,62,288 litres and through synthetic and composite estimators were found to be 13,49,910 and 13,54,546. It has been observed that the average agricultural income of all 20 blocks of Jammu district and showed that the agricultural income of block Bishnah was maximum (Rs. 26,95,000), whereas the block Khara Balli had the least (Rs. 13,000), among all the twenty (20) blocks of district Jammu of Jammu and Kashmir State. Also, it was clear that the blocks Arnia and Khour were at par.

It has been observed that the values of mean square error of the composite estimator were smaller as compared to direct and synthetic estimator. The minimum value of composite estimator was found 4265.37 and maximum was 6999522.27. Further the minimum value of absolute relative bias of the composite estimator was 3.89 and maximum was 11.23. In case of milk production it has been observed that the estimate of milk production estimated through composite estimator seems best based on mean square error and absolute relative bias criterion.

In this paper we have empirically investigated the estimate of per day average milk production at block levels of district Jammu of J&K UT through direct and indirect methods of small area estimation using real milk data set for different small domains and the results so obtained are compared in terms of mean square error and absolute relative bias.

Keywords: Small area estimation, direct estimator, synthetic estimator, composite estimator, small domain

1. Introduction

Small area estimation (SAE) is one of the several techniques which involve the estimation of parameters for small sub-population generally used when the sub-population of interest is included in a larger survey. SAE is categorized into two types of estimators direct and indirect estimators. When the estimator uses the values from the sample data only it is called direct estimation. But the major disadvantage of such estimators is that unacceptably a large may result. This happens when the sample size within the domain is small or nil. Direct small area estimation is based on survey deign which includes three estimators called the Horvitz-Thompson (H-T) estimator, Generalized Regression (GREG) estimator and modified direct estimator. On the other hand indirect approaches of small area estimation are further divided into two types statistical and geographical approach. The statistical approach is mainly based on different statistical models and techniques. Whereas, the geographical approach uses techniques, such as micro simulation modeling etc.

The synthetic estimation was first introduced in the United Centre for Health Statistics (1968). This technique of indirect estimation was used to calculate State Level disability estimates.

When an estimate is required for a small area we use data from the region outside this small area but belonging to different domains which overlap this small (local) area. This is called as "borrowing strength" from similar regions and such estimators are called "synthetic estimators". Gonzalez (1973) [6] explained an excellent definition of synthetic estimation. "An estimator should be synthetic when a reliable direct estimator for a layer area is used to derive an indirect estimator for a small are belonging to the larger area under the consumption that all small area have the same characteristics as the larger area". Levy (1979) [7] and Rao (2003) [5] overviewed on various synthetic estimation approaches and its applications in small area estimation.
Composite estimation acts as a balancing approach between the synthetic and direct estimators. A weighted sum of these two estimators would be an alternative to choosing one over the other to balance their degree of bias, and this type of estimator is commonly known as a composite estimator. According to Gosh and Rao (1994) [8] composite estimators is a natural way to balance the potential bias of a synthetic estimator against the instability of a direct estimator by choosing an appropriate weight.

2. Direct Estimator
When the estimator uses the values from the sample data only it is called direct estimation. But the major disadvantage of such estimators is that unacceptably a large standard error may result. Direct estimator is the most basic estimator and can only be used when all the areas have been sampled. For the area mean value it is as follows.

$$\hat{\hat{Y}}_{i,Direct}=\frac{\sum_j wij\, yij}{\sum_j wij} \qquad 2.1$$

The weights *wij* have been taken as the inverse of the probability of an individual to be in the sample. Note that since all areas are sampled independently and with replacement, the probability of selecting individual *j* in area *i*is 1/*Ni*, where *Ni* is the number of individuals in area *i*. Thus the weight *wij* may be interpreted as the number of elements in the population represented by the sample element. The choice *wij* satisfies the unbiasedness condition and leads to the well-known Horvitz Thompson (H-T) estimator. If the sample size in region *i* is *ni*, the probability of selecting an individual at least once is $1-\left[1-\frac{1}{Ni}\right]^{ni}$ This is the inclusion probability and we will use weights.

$$w_{ij}{}^{-1}=wi^{-1}=1-\left[1-\frac{1}{Ni}\right]^{ni} \qquad 2.2$$

Direct estimators are generally used when the sample size for each small area is sufficiently large to give reasonably accurate estimates. However, as the sources of data are usually sample surveys designed to give national and regional statistics, sample sizes for the small areas (usually sub domains of the original domains of study) are usually unduly small. Consequently, the associated variances are likely to be unacceptably large since the conditional variances (as can be seen above) are of the order ni^{-1}. Moreover, if information from a national sample is used to make estimates for small areas and there are no sample units in the small area of interest, then obviously direct estimation cannot be used. The variance of the direct estimator, which is also known as design variance, can be estimated to assess the uncertainty about the estimates. This can be used to provide approximate confidence intervals. The design variance of the direct estimator (1) is

$$\mathrm{V}\left[\hat{\hat{Y}}_{iDIRECT}\right]=\left(1-\frac{1}{Ni}\right)\frac{Si^2}{ni}$$

Here, Si^2 is the variance of the sample obtained from area i. The variance can be calculated by

$$\hat{V}\left[\hat{\hat{Y}}_{iDIRECT}\right]=\left(1-\frac{1}{Ni}\right)\frac{\hat{S}i^2}{ni}$$

That is, we substitute the variance of a generic sample$Si^2$by the actual variance of the observed data $\widehat{Si}^2$.

3. Synthetic Estimator
The term "synthetic estimates" was first used by the U.S. National Centre for Health Statistics (1968). This technique of indirect estimation was used to calculate State Level disability estimates. When an estimate is required for a small area we use data from the region outside this small area but belonging to different domains which overlap this small (local) area. This is called as "borrowing strength" from similar regions and such estimators are called "synthetic estimators". Gonzalez (1973) [6] explained an excellent definition of synthetic estimation as follows:
"An estimator should be synthetic when a reliable direct estimator for a layer area is used to derive an indirect estimator for a small are belonging to the larger area under the consumption that all small area have the same characteristics as the larger area".
Suppose the population is partitioned into large domains for which reliable direct estimators $Y'_{.g}$, of the totals, $Y_{.g}$, can be calculated from the survey data: the small areas, i, may cut across g so that $Y_{.g}=\sum_i Y_{ig}$, where, Y_{ig} is the total for cell(i, g.) we assume that auxiliary information in the form of total X_{ig}, is also available. A synthetic estimator of small area total $Y_i=\sum_g Y_{ig}$ is then given by $\hat{Y}_i^S=\sum_g\left(X_{ig}/X_{.g}\right)\hat{Y}'_{.g}$, where, $X_{.g}=\sum_i X_{ig}$ (Ghangurde and Singh, 1977).
The synthetic method of estimation is by far one of the most widely used small area estimation methods due to its ease of calculation.

4. Composite Estimator
The synthetic estimators outperform the simple direct estimators, when small area samples are relatively small; however, when small area sample sizes are large, the direct estimators outperform the synthetic estimators. Thus it was concluded that a weighted sum of these two estimators would be better than choosing one over the other.
Composite estimation acts as a balancing approach between the synthetic and direct estimators. A weighted sum of these two estimators would be an alternative to choosing one over the other to balance their degree of bias, and this type of estimator is commonly known as a composite estimator. According to Gosh and Rao (1994) [8] composite estimators is a natural way to balance the potential bias of a synthetic estimator against the instability of a direct estimator by choosing an appropriate weight. The weights are defined so that if the sample size is "large" the direct estimate is given more weight than the synthetic one and when the sample is not reliable, the synthetic estimate will be given more weight.

Thus a natural way to balance the potential bias of a synthetic estimator, say $\hat{Y}_{iSYNTH}$ against the instability of a direct estimator, say, $\hat{Y}_{iDIRECT}$ is to take the weighted average of $\hat{Y}_{iDIRECT}$ and $\hat{Y}_{iSYNTH}$. Such composite estimator of small area total may be written as:

$$\hat{\hat{Y}}_{iCOMP} = wi\hat{\hat{Y}}_{iDIRECT} + (1 - wi)\hat{\hat{Y}}_{iSYNTH} \quad 3.1$$

For a suitably chosen weight wi (0≤wi≤1) which controls the shrinkage of the two estimators. That is, depending on how large is the sample in the small area it will give more weight to the direct estimate (if the sample is large) or to the synthetic estimate (if information is needed from other areas). The design MSE of the composite estimator is given by

$$MSEp(\hat{\hat{Y}}_{iCOMP}) = øi^2MSEp(\hat{Y}_{iDIRECT}) + (1 - øi)^2MSEp(\hat{Y}_{iSYNTH}) + 2øi(1 - øi)Ep(\hat{Y}_{iDIRECT} - Yi)(\hat{Y}_{iSYNTH} - Yi) \quad 3.2$$

By minimizing (3.2) with respect to $øi$, we get the optimal weight $øi$ as

$$øi^* = MSEp(\hat{Y}_{iSYNTH}) / [MSEp(\hat{Y}_{iDIRECT}) + MSEp(\hat{Y}_{iSYNTH})]$$

The approximate optimal weight $øi^*$depends only the ratio of the MSEs

$øi^*$=1/ (1+F_i)

Where

$F_i = øi^* \frac{MSEp(\hat{Y}_{iDIRECT})}{MSEp(\hat{Y}_{iDIRECT})}$

It is easy to show $\hat{Y}_{i,COMP}$ is better than either component estimator in terms of MSE when max $(0,2øi - 1) \leq øi \leq min\,(2\,øi^*, 1)$. The latter reduces to the whole range $0 \leq øi \leq 1$. When F_i =1, and it becomes narrower as F_i deviates from 1. The optimal weight $øi^*$will be close to zero or one when one of the component estimators has a much larger MSE than the other that is when Fi is either large or small. In this case estimator with large MSE adds little information and therefore it better to use the component estimator with small MSE in preference to the composite estimator.

Table 1: General features of Jammu district

1.	Area	2336 sq km
2.	Population	1529958
3.	Population density	650 person per sq km
4.	Blocks	20
5.	Tehsil	04
6.	Panchayats	244
7.	Livestock population	78.908 lakhs
8.	Principal crops	Wheat, Rice, Maize
9.	Literacy rate	83.45%
10.	Sex Ratio	880 Female per 1000 male.
11.	Sub-divisions	7
12.	Urban local bodies	8
13.	Assembly Constituencies	11

Source: Official website/http://jammu.nic.in

Table 2: List of Sub-divisions of Jammu district

S. No.	Names
1	Jammu South
2	Jammu North
3	R.S.Pura
4	Marh
5	Akhnoor
6	Chowki Choura
7	Khour

For obtaining the estimates of per day total milk production of cows as well as buffaloes at block levels in the Jammu district of J&K state, all blocks of the Jammu district were considered and then the number of villages were taken in proportion to the total number of villages in a block so as to take 80 villages in total from the entire district. The villages from each block were selected randomly from the list of villages of the respective block, and then complete enumeration of selected villages was done. All the twenty blocks, total numbers of villages in each block and randomly selected villages as per detail are given in the table 3.

Table 3: List of blocks of Jammu district and randomly selected villages from each block

S. No	Name of the Block	Total no. of Villages	Name of randomly selected villages
1.	Akhnoor	34	Sungal, Dasgal, Manda
2.	Arnia	30	Karyal Khurd, Adlehar, Chak Fateh Khan
3.	Bhalwal	35	Lower Kote, Thather, Gurha Brahamana
4.	Bishnah	96	Chumbian Brahamana, Sarore, Chackbana, Shahpur, Fatwal, Pasgal, Kheri, Makhanpur, Kotlichakran
5.	Bhalwal Brahamana	35	Gurah Jagir, Naziachack, Garh
6.	Chowki Choura	27	Dhana Danu, Chowki
7.	Dansal	69	Jajjarkotli, Churta, Slay, Palli, Suketer, Jandrah, Sandrote
8.	Khour	49	Platan, Dhar, Chack Malal, Sainth, Pahariwala
9.	Khara Balli	15	Kharah
10.	Marh	105	Halqa Marh, Klasey Chack, Pandorian, Taru Chack, Saharan, Dhatriyal, Zaffrachack, Flora Nagbani, Shama Chack, Sui
11.	Mandal Phallian	53	Khandwal, Lalyal, Mangu Chack, Nandwal, Murad Pur
12.	Miran Sahib	40	Tutre, Ban Sultan, Langotia, Kharian
13.	Mathwar	20	Keri, Rabta
14.	Maira Mandrian	26	Maira-Mandrian, Rehani
15.	Nagrota	37	Jagti Nala Kamini, Sitni, Kandoli Nagrota, Dok Bajree
16.	Pargwal	16	Belajamana, Sajwal
17.	R.S. Pura	50	Puropana, Abdullian, Jassore, Kana chack, Badyal Brahamana
18.	Satwari	24	Chatha Pind, Raipur
19.	Samwan	11	Kachrial
20.	Suchetgarh	62	Biaspur, Kotli Mirdian, Sindhey, Chak Baja, Chakroi, Sai Khurad

Table 4: Average agricultural income, Average total income, Total number of livestock and Average area under fodder cultivation of blocks of Jammu district

Small Areas	Agril. Income (Rs.)	Total income (Rs)	No. of Livestock	Area (ha)
Nagrota	15,000	19,65,000	10,957	0.95
Mathwar	3,40,000	13,40,000	1,915	10.20
Maira Mandrian	3,31,000	31,00,000	4,521	6.25
Akhnoor	1,20,000	21,10,000	11,404	1.85
Marh	25,49,000	80,37,000	34,099	47.95
Satwari	5,40,000	10,80,000	10,855	9.10
Balwal	3,41,000	21,40,000	6790	8.20
Miran sahib	75,000	13,65,000	12,495	2.11
Mandal Phallian	10,63,000	35,29,000	10,856	17.55
Bishnah	26,95,000	80,65,000	24,822	35.27
Suchetgarh	11,55,000	27,79,000	17,666	20.62
Chowki Choura	85,000	6,75,000	6,262	4.85
Arnia	9,10,000	25,75,000	7,982	11.10
R.S. Pura	12,90,000	40,95,000	16,984	19.90
Khour	9,11,000	37,78,000	13,860	14.20
Balwal Brahamana	83,000	67,7,000	9,749	2.20
Dansal	2,45,000	22,46,000	17,980	11.70
Khara Balli	13,000	1,43,000	4,620	0.60
Samwan	1,00,000	6,10,000	2,464	3.30
Pargwal	2,05,000	11,55,000	4,928	3.45

Table 4 represents the Average agricultural income, Average total income, Total number of livestock and Average area under fodder cultivation of blocks of Jammu district. From the values of agricultural income it was noticed that the minimum value of average agricultural income was Rs. 13,000 of block Khara Balli and its maximum value was found Rs. 26, 95,000 of Marh block. Also, minimum total income was observed for block Khara Balli with Rs. 1, 43,000. Further it was observed that the block Dansal had the maximum number of livestock population 17,980 (cows and buffaloes) whereas block Mathwar was noticed with minimum livestock population numbered 1,915. The average area under fodder cultivation was found minimum for block Khara Balli with 0.6 ha. And it was maximum for block Marh with 47.95 ha.

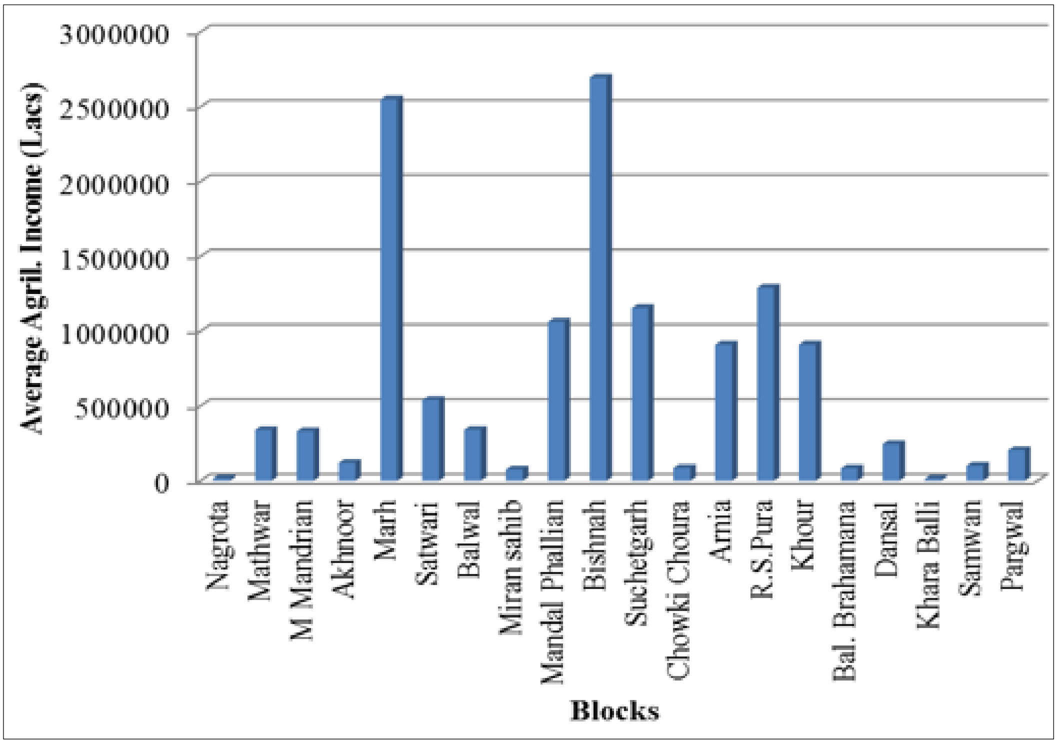

Fig 1: Avg. agricultural income of blocks of Jammu district

Fig. 1 represents the average agricultural income of all 20 blocks of Jammu district and showed that the agricultural income of block Bishnah was maximum (Rs. 26,95,000), whereas the block Khara Balli had the least (Rs. 13,000), among all the twenty (20) blocks of district Jammu of Jammu and Kashmir State. Also, it was clear that the blocks Arnia and Khour were at par.

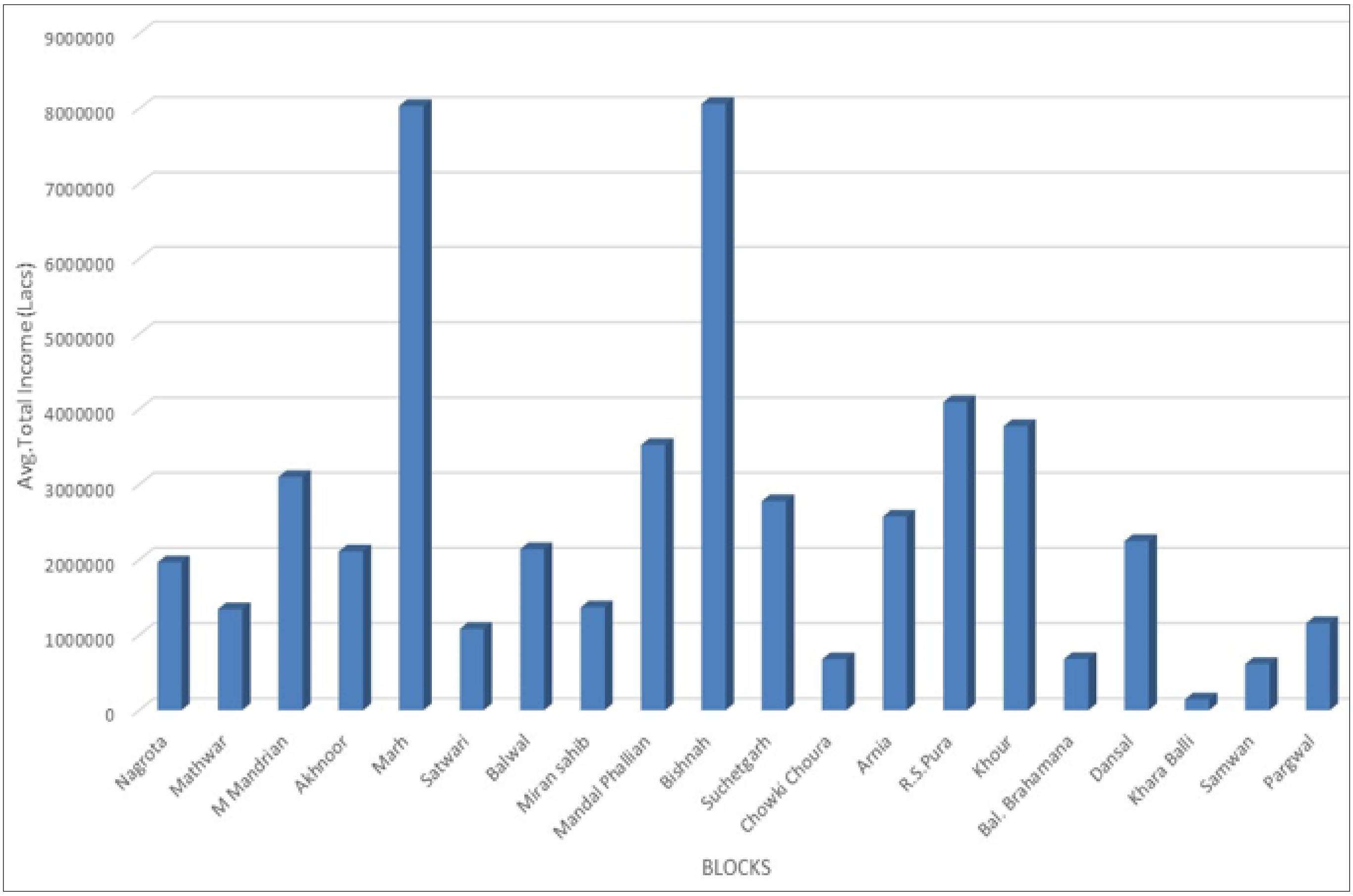

Fig 2: Avg. total income of blocks of Jammu district

Fig. 2 represents that the total income of all 20 blocks of district Jammu and it was noted that the block Bishnah showed maximum total Income, whereas the block Khara Balli had the least, among all the twenty (20) blocks of Jammu district.

In this section, we make an empirical comparison between direct, synthetic and composite estimators. The performance of estimators is examined from the accuracy of the point estimates standpoints. This is considered through the mean squared error.

5. Methodology

For obtaining the estimates of milk production of district Jammu of Jammu and Kashmir UT at block levels though small area estimation techniques which are already existing, we considered all the 20 blocks of district Jammu of Jammu and Kashmir UT and Stratified three stage random sampling plan was adopted with blocks constituting the strata, villages as the primary stage units, households possessing livestock as the secondary stage units and cows / buffaloes in milk in the selected households as the third stage units. The data was collected through personal interview with the three randomly selected households and from each household two animals in milk (2 cows/ 2 buffaloes or 1 cow + 1 buffalo) were selected randomly for recording data on milk, using pre-structured schedule in the light of study. Thus, out of 80 villages, 240 households and a total number of 480 milch animals (consisting of cows and buffaloes) were selected from the entire blocks of Jammu district. Based on the collected information the empirical investigation were undertaken in order to have small area estimates of milk production at block levels in Jammu district.

Table 5. Revealed the average per day milk production (lt.) of all 20 blocks through direct, synthetic and composite small area estimation techniques. From the estimates presented above through different estimators it has been observed that the block Samwan yields minimum milk production 12,810 litres and block Marh yields maximum of 1,59,383 litres. The milk production is minimum in Samwan block and is maximum in Marh block through all estimators.

The empirical estimates of direct, synthetic and composite estimators based on our sample study revealed that the Composite estimates are close to the actual values (13,54,546, litres) as compared to Synthetic and Direct estimates.

Table 5: Average per day milk production (lt.) of 20 small areas (Blocks) through Direct, Synthetic and Composite small area estimation techniques

Small Area	Direct Estimate	Synthetic Estimate	Composite Estimate
Nagrota	61704.32	61117.7	61337.44
Mathwar	11019.78	10917.25	10955.65
M Mandrian	29072.71	28830.66	28921.33
Akhnoor	84181.98	83571.43	83800.13
Marh	248164.4	246338.8	247022.7
Satwari	97739.61	97158.45	97376.14
Balwal	54107.53	53744	53880.17
Miran sahib	87723.17	87054.21	87304.79
Mandal Phallian	75104.68	74523.47	74741.18
Bishnah	147967.6	146638.6	147136.4
Suchetgarh	123577.3	122631.5	122985.8
Chowki Choura	13847.88	13512.62	13638.21
Arnia	37854.36	37427.01	37587.09
R.S. Pura	108469	107559.7	107900.3
Khour	69426.07	68684.02	68961.98
Bal. Brahamana	19903.14	19381.19	19576.7
Dansal	41879.2	40916.57	41277.15
Khara Balli	9985.74	9738.39	9831.044
Samwan	14412.71	14280.79	14330.21
Pargwal	26147.09	25883.25	25982.08
	13,62,288	13,49,910	13,54,546

Table 6: Mean Square Error (MSE) of Estimators of Variance components for 20 blocks of district Jammu

Direct	Synthetic	Composite
4838559.31	4821201.31	4792087.31
265508.34	248150.34	219036.34
2079864.59	2062506.59	2033392.59
1702894.17	1685536.17	1656422.17
1227027.97	1209669.97	1180555.97
7045994.27	7028636.27	6999522.27
1517409.59	1500051.59	1470937.59
2638766.58	2621408.58	2592294.58
4414800.02	4397442.02	4368328.02
1184204.46	1166846.46	1137732.46
3804688.73	3787330.73	3758216.73
120978.78	103620.78	74506.78
1715864.25	1698506.25	1669392.25
1432566.10	1415208.10	1386094.10
1206372.75	1189014.75	1159900.75
22638.16	5280.16	4265.37
545417.91	528059.91	498945.91
57279.99	39921.99	17089.35
4460546.00	4443188.00	4414074.00
2201002.34	2183644.34	2154530.34

From the table 6, it has been observed that the values of mean squared error of the composite estimator were smaller as compared to direct and synthetic estimator. The minimum value of mean squared error in case of composite estimator was found 4265.37 and maximum was 6999522.27 whereas the minimum values of mean squared error for direct and synthetic, estimator were 22638.16, 5280.16 and their maximum values were 7045994.27 and 7028636.27 respectively. On the basis mean squared error it was observed that the composite estimator performed better than direct and synthetic estimator.

6. Conclusion

In the present investigation the average agricultural income of all 20 blocks of Jammu district has been analysed and showed that the agricultural income of block Bishnah was maximum (Rs. 26,95,000), whereas the block Khara Balli had the least (Rs. 13,000), among all the twenty (20) blocks of district Jammu of Jammu and Kashmir UT. Also, it was clear that the blocks Arnia and Khour were at par. Also, it was noted that the block Bishnah showed maximum total Income, whereas the block Khara Balli had the least, among all the twenty (20) blocks of Jammu district.

The estimate of per day average milk production at block levels of district Jammu of J&K UT through direct and indirect methods of small area estimation using real milk data set for different small domains. It is concluded that composite estimator performed better than the direct and synthetic estimators.

7. References

1. Archana, Rizvi SEH, Sharma M, Kand Bhat MIJ. A Generalized Class of Synthetic Estimator with Application to Estimation of Milk Production for Small Domains. International research journal of Agricultural Economics and Statistics. 2019;10(1):115-121.
2. Gonzales ME. Use and Evaluation of Synthetic Estimators, Proceedings of the Social Statistics Section of the American Statistical Association; c1973. p.33-36.
3. Ghosh M, Rao JNK. Small area estimation: An appraisal. Statistical Science. 1994;6(1):55-93.
4. Levy PS. The Use of Mortality Data in Evaluating Synthetic Estimates, Proceedings of the Social Statistics Section, American Statistical Association; c1971. p.328-331.
5. Rao JNK. Small Area Estimation. New York: Wiley; c2003.
6. Gonzalez-Barcena D, Kastin AJ, Schalch DS, Torres-Zamora M, Perez-Pasten E, Kato AL, *et al*. Responses to thyrotropin-releasing hormone in patients with renal failure and after infusion in normal men. The Journal of Clinical Endocrinology & Metabolism. 1973 Jan 1;36(1):117-20.
7. Levy MR. Watching TV news as para-social interaction. Journal of Broadcasting & Electronic Media. 1979 Jan 1;23(1):69-80.
8. Ghosh M, Rao J. Small area estimation: an appraisal. Statistical science. 1994 Feb;9(1):55-76.