Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 77(2) 2023 177–183

Calibration Estimation Approach For Population Ratio under Adaptive Cluster Sampling

Ankur Biswas, Raju Kumar, Deepak Singh and Pradip Basak

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 23 November 2022; Revised 01 March 2023; Accepted 28 March 2023

SUMMARY

Population ratio is one of the most commonly used statistics in official statistics, agriculture and agricultural-related fields. When interest is in the estimation of ratio for rare but highly aggregated geographically distributed population, adaptive cluster sampling (ACS) design is usually used (Dryver and Chang 2007). Under ACS design, neighbouring units are added to the sample if it satisfies a pre-determined criterion. ACS design allow observed values to trigger increased sampling effort during the survey. This intuitively appealing design can have lower variance than conventional designs. In many sampling survey situations, certain auxiliary information is often available and used for increasing the precision of estimator. Calibration approach given by Deville and Särndal (1992) is widely used technique for this purpose. In this article, calibration estimator of population ratio under adaptive sampling has been developed when auxiliary variables are known. The variance and the estimate of variance for these estimators are obtained. The statistical performance of the proposed calibration estimators of population ratio under ACS were evaluated through a simulation study based on real population data with respect to conventional Horvitz Thomson (HT) estimator of population ratio which do not utilize the auxiliary information. The results of the simulation study show that proposed calibration estimators are more efficient than conventional HT estimator of the population mean under ACS with respect to percentage Relative Root Mean Squared Error (%RRMSE).

Keywords: Calibration; Auxiliary information; Rare; Clustered; Population ratio.

1. INTRODUCTION

Survey statisticians often deal with situations in which members with interesting characteristics are sparsely scattered across the geographically distributed population but present in clustered manners. For example, in the estimation of total number of rare birds in an area, number of trees of a rare species, production of non-hybrid minor crops etc. Traditional sampling techniques like simple random sampling, stratified random sampling, PPS sampling, cluster sampling, multiple sampling etc. cannot handle with this situation. Some of the main reason behind the non-applicability of conventional sampling techniques under the situations are sample do not get enough units meeting with specific criteria, difficult sampling and estimation problem, high standard errors etc. In such cases the researchers are interested to search for sampling and estimation methods that go beyond the conventional set of techniques. Thompson (1990, 1991a, 1991b), first, introduced Adaptive Cluster Sampling (ACS) into the literature on survey sampling for sampling geographically rare and hidden clustered population. ACS design allow observed values to trigger increased sampling effort during the survey. ACS designs assign high probabilities to samples that include areas with high density of characteristics under study. This intuitively appealing design can have lower variance than conventional designs. The increase in precision may depend on the spatial distribution of the population, condition determining when to adapt sampling effort etc. In this method, neighbouring units are added to the sample if it satisfies a pre-determined criterion. Also, the neighborhoods do not depend on the data. Technically, the units in the neighborhood don't have to be physically contiguous. For example, the neighborhood may be defined through a social relationship between units, i.e., people. A person's neighborhood could be all people sexually linked

Corresponding author: Raju Kumar
E-mail address: raju.kumar@icar.gov.in

people to the individual with the network units being the subset of people (neighbours) that are HIV positive and edges units being the subset of neighbors that are HIV negative (Dryver and Thompson 1998). Efficiency of ACS design depends on spatial distribution (Brown 2003), but encounter and detection rates tend to be relatively high with only a modest degree of clustering (Smith *et al.* 2004).

Population ratio is one of the most common statistics used in official statistics, agriculture and allied field of agriculture. Often interest is in the estimation of population ratio for rare and clustered geographically distributed population like age or sex ratio of rarest animal in wild life population, the number of bullock available for per acre of holding, productivity of boro rice etc. These statistics are typically expressed in ratio. Again, in many sampling survey situations, certain auxiliary information is often available. The estimators which use auxiliary variables are often more accurate than the standard ones. Calibration approach given by Deville and Särndal (1992) is commonly used in survey sampling to include auxiliary information to increase the precision of the estimators of population parameter. A calibration estimator uses calibrated weights, which are as close as possible, according to a given distance measure, to the original sampling design weights while also respecting a set of constraints, the calibration equations. For every distance measure there is a corresponding set of calibrated weights and a calibration estimator. In this study, an attempt has been made to develop calibration estimator of finite population ratio under adaptive cluster sampling when auxiliary information is available. Brief of ACS design is given in Section 2. The developed methodology under ACS design has been described in Section 3. Under this study, the statistical performance of the proposed calibration estimators of population ratio under ACS are evaluated through a simulation study with respect to conventional Horvitz Thomson (HT) estimator using a real dataset as a study population. Details of the simulation study and discussion on the results are described in Section 4. Concluding remarks are provided in Section 5.

2. ESTIMATION PROCEDURE UNDER ADAPTIVE CLUSTER SAMPLING

Let, U={1,…,k,…,N} be the finite population under consideration. and Y be the character under study and taking real values as $y_1, y_2, \ldots, y_N$.The objective is to estimate the population mean i.e. $\bar{Y} = \frac{1}{N}\sum_{j=1}^{N} y_j$.

To begin by taking a sample under the ACS design, n_1 units are initially selected from a finite population using simple random sampling without replacement (SRSWOR). If an observed value of a sampling unit meets a predetermind criterion C, additional units in a specified neighbourhood are adaptively added to the sample. The units in their neighbourhoods are then included in the sample if any of these extra units satisfy condition C. This adaptive procedure keeps going until no more units that meet C are found. Finally, we have the sample, in which there are n_1 clusters (not necessarily unique), one for each unit initially chosen. If the selected unit in the initial sample does not satisfy the predefined condition C, then there is no adaptive selection and it is considered as a cluster of size one.

As an example, we used most often cited Smith *et al.* (1995)'s blue-winged teal bird population (Fig. 1). The first part of Fig. 1 represents 200 quadrants of 25 km^2 area each with counts given for quadrats having non-zero bird counts. Pictorial representation of the selection of an adaptive cluster sample from this population is given in second part of Fig. 1. An initial simple random sample of 10 quadrats is selected indicated with $\otimes$. For the predefined condition C i.e. $y_j \geq 1$, all neighbourhood quadrats would be sampled with ACS design denoted with 'o'. Quadrats with 'o' having the teal counts that do not satisfy C are the edge units.

Assume that K distinct networks can be formed from the population. Let, A_i denoted the i^{th} network, m_i denote the number of units in the network, y_i^* is the total of y-values in the i^{th} network i.e. $y_i^* = \sum_{j \in A_i} y_j$.

Thompson (1990) proposed the modified Horvitz-Thompson estimator of the population mean using the initial intersection probability under ACS design as given by

$$\hat{\bar{\mathbf{Y}}}_{HT(AC)} = \frac{1}{N}\sum_{i=1}^{K}\frac{y_i^* J_i}{\pi_i^{'}} = \frac{1}{N}\sum_{i=1}^{k}\frac{y_i^*}{\pi_i^{'}} = \frac{1}{N}\sum_{i=1}^{k} d_i y_i^* \qquad (1)$$

where, J_i takes value 1 if the initial sample intersects the i^{th} network and 0 otherwise, k be the number of distinct networks intersected by the initial sample and $d_i = 1/\pi_i^{'}$, $\pi_i^{'}$ can be interpreted as the probability that the initial sample intersects network A_i, the network

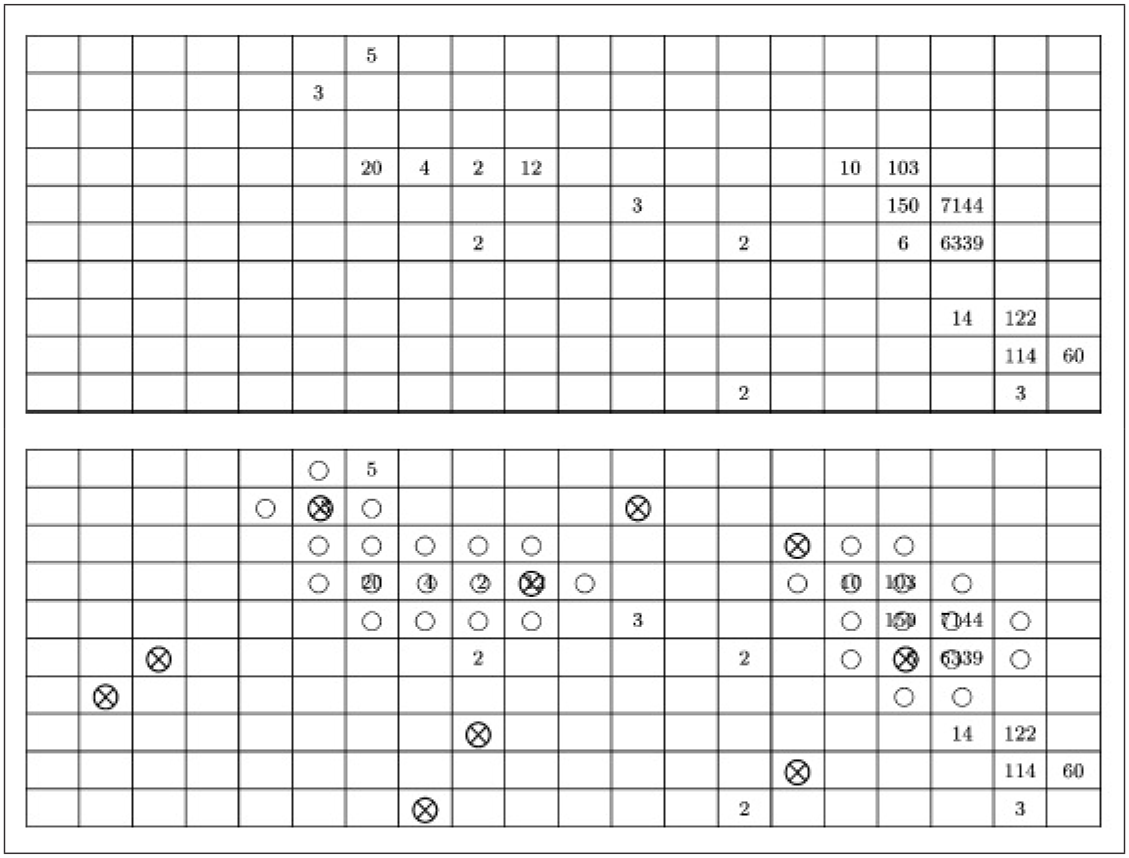

Fig. 1. An adaptive cluster sample (C: $y_i \geq 1$) from the Blue-winged teal population as quoted in Smith *et al.* (1995)

containing unit i. It is same for each unit in the network. π'_{ij} is the joint initial intersection probability.

Thompson (1990) showed that $\hat{\hat{\mathbf{Y}}}_{HT(AC)}$ is unbiased and the sampling variance of the $\hat{\hat{\mathbf{Y}}}_{HT(AC)}$ is given by

$$V\left(\hat{\hat{\mathbf{Y}}}_{HT(AC)}\right) = \frac{1}{N^2}\sum_{i=1}^{K}\sum_{j=1}^{K}\Delta_{ij}\left(d_i y_i^*\right)\left(d_j y_j^*\right) \tag{2}$$

where, $\Delta_{ij} = \pi'_{ij} - \pi'_i \pi'_j$.

3. CALIBRATED ESTIMATORS OF FINITE POPULATION RATIO UNDER ADAPTIVE CLUSTER SAMPLING DESIGN

Let Y and Z are two characteristics under study defined on the population U and taking real values $\{y_1, y_2, \ldots, y_N\}$ and $\{z_1, z_2, \ldots, z_N\}$. Many a times our interest is to estimate the population ratio e.g. production of non-hybrid crop (e.g. boro rice) per acre of holding (Saful, 2005), number of Sahiwal cow per acre of holding in population etc. Let, the parameter of interest is the population ratio of variable Y and Z, i.e. $R = \left(\bar{Y}/\bar{Z}\right) = \left(Y/Z\right) = \sum_{i=1}^{N} y_i \Big/ \sum_{i=1}^{N} z_i$. An adaptive cluster sample is selected using earlier discussed methods and observation on Y and Z variables are taken from all the selected sampling units.

An estimator of population ratio using modified Horvitz-Thompson (HT) estimator which does not make use of auxiliary information (Dryver and Chao, 2007) is given by

$$\hat{R}_{HT(AC)} = \sum_{i=1}^{k} d_i y_i^* \Big/ \sum_{i=1}^{k} d_i z_i^* \tag{3}$$

where, $y_i^* = \sum_{j \in A_i} y_j$, $z_i^* = \sum_{j \in A_i} z_j$ and $d_i = 1/\pi'_i$ are respective design weights.

Let, U and V be available auxiliary variables of the study variable Y and Z respectively with real values as $\{u_1, u_2, \ldots, u_N\}$ and $\{v_1, v_2, \ldots, v_N\}$. Let, assume that population total of auxiliary variable U and V are available i.e. $U = \sum_{i=1}^{N} u_i$ and $V = \sum_{i=1}^{N} v_i$ are assumed to be known. Using the well-known Calibration Approach (Deville and Särndal, 1992), attempt has been made to improve the Horvitz-Thompson estimator based estimator of the population ratio under adaptive cluster sampling i.e. $\hat{R}_{HT(AC)}$.

The proposed calibration estimator is given by

$$\hat{R}_{CAL(AC)} = \sum_{i=1}^{k} w_{1i} y_i^* \Big/ \sum_{i=1}^{k} w_{2i} z_i^* \tag{4}$$

where, w_{1i} and w_{2i} are the respective calibrated weights for variable Y and Z.

The new weights w_{1i} and w_{2i} are chosen as close as possible to the design weight d_i subjected to calibration constraints. For this purpose, we minimize the chi-square type distance function given by

$$L_1 = \sum_{i=1}^{k} \frac{(w_{1i} - d_i)^2}{d_i q_i} \text{ and } L_2 = \sum_{i=1}^{k} \frac{(w_{2i} - d_i)^2}{d_i q_i} \tag{5}$$

subject to constraints $N\bar{U} = \sum_{i=1}^{k} w_{1i} u_i^*$ and $N\bar{V} = \sum_{i=1}^{k} w_{2i} v_i^*$, where q_i are suitably chosen constants.

We minimize following function using the method of Lagrange multiplier as

$$\varphi(w_{1i}, w_{2i}, \lambda_1, \lambda_2) = \frac{1}{2}\sum_{i=1}^{k} \frac{(w_{1i} - d_i)^2}{d_i q_i} + \frac{1}{2}\sum_{i=1}^{k} \frac{(w_{2i} - d_i)^2}{d_i q_i} - \lambda_1\left(U - \sum_{i=1}^{k} w_{1i} u_i^*\right) - \lambda_2\left(V - \sum_{i=1}^{k} w_{2i} v_i^*\right) \tag{6}$$

By minimization of objective function $\varphi(w_{1i}, w_{2i}, \lambda_1, \lambda_2)$ using the Lagrange multiplier technique, the new set of calibration weights are obtained as

$$w_{1i} = d_i + \frac{d_i q_i u_i^*}{\sum_{i=1}^{n} d_i q_i u_i^{*2}}\left(U - \sum_{i=1}^{k} d_i u_i^*\right) \text{ and}$$

$$w_{2i} = d_i + \frac{d_i q_i v_i^*}{\sum_{i=1}^{n} d_i q_i v_i^{*2}}\left(V - \sum_{i=1}^{k} d_i v_i^*\right). \tag{7}$$

The proposed calibration estimator based on the revised weights is given by

$$\hat{R}_{CAL(AC)} = \frac{\sum_{i=1}^{k} w_{1i} y_i^*}{\sum_{i=1}^{k} w_{2i} z_i^*} = \frac{\sum_{i=1}^{k} d_i y_i^* + \frac{\sum_{i=1}^{k} d_i q_i u_i^* y_i^*}{\sum_{i=1}^{k} d_i q_i u_i^{*2}}\left(U - \sum_{i=1}^{k} d_i u_i^*\right)}{\sum_{i=1}^{k} d_i z_i^* + \frac{\sum_{i=1}^{k} d_i q_i v_i^* z_i^*}{\sum_{i=1}^{k} d_i q_i v_i^{*2}}\left(V - \sum_{i=1}^{k} d_i v_i^*\right)}. \tag{8}$$

Following Deville and Särndal (1992) and Särndal *et al.* (1992), the asymptotic variance of proposed calibration estimator by Taylor's series linearization technique is obtained as

$$AV\left(\hat{R}_{CAL(AC)}\right) = \frac{1}{Z^2}\sum_{i=1}^{K}\sum_{j=1}^{K} \Delta_{ij}\left(d_i E_{iyzuv}\right)\left(d_j E_{jyzuv}\right) \tag{9}$$

where,

$E_{iyzuv} = E_{iyu} - RE_{izv}$, $E_{iyu} = y_i^* - B_{yu} u_i^*$, $E_{izv} = z_i^* - B_{zv} v_i^*$,

$B_{yu} = \sum_{i=1}^{K} q_i u_i^* y_i^* \Big/ \sum_{i=1}^{K} q_i u_i^{*2}$, $B_{zv} = \sum_{i=1}^{K} q_i v_i^* z_i^* \Big/ \sum_{i=1}^{K} q_i v_i^{*2}$,

$\Delta_{ij} = \pi'_{ij} - \pi'_i \pi'_j$.

Following Särndal *et al.* (1992), the expression of estimator of variance is obtained as

$$\hat{V}\left(\hat{R}_{CAL(AC)}\right) = \frac{1}{\hat{Z}^2_{CAL(AC)}}\sum_{i=1}^{k}\sum_{j=1}^{k} \frac{\Delta_{ij}}{\pi'_{ij}}\left(d_i e_{iyzuv}\right)\left(d_j e_{jyzuv}\right) \tag{10}$$

where,

$e_{iyzuv} = e_{iyu} - \hat{R}_{CAL(AC)} e_{izv}$, $e_{iyu} = y_i^* - \hat{B}_{yu} u_i^*$,

$e_{izv} = z_i^* - \hat{B}_{zv} v_i^*$, $\hat{B}_{yu} = \sum_{i=1}^{k} w_{1i} q_i u_i^* y_i^* \Big/ \sum_{i=1}^{k} w_{1i} q_i u_i^{*2}$,

$\hat{B}_{zv} = \sum_{i=1}^{k} w_{2i} q_i v_i^* z_i^* \Big/ \sum_{i=1}^{k} w_{2i} q_i v_i^{*2}$ and $\hat{Z}^2_{CAL(AC)} = \sum_{i=1}^{k} w_{2i} z_i^*$.

4. SIMULATION STUDY

In this study, using a real dataset, a simulation study was carried out in order to study the statistical performance of the proposed calibration estimators under adaptive cluster sampling design framework. The blue-winged teal bird population given in the often cited Smith *et al.* (1995) has been utilized for the simulation study. Generally, wintering waterfowl (e.g. blue-winged teal bird) are counted from aircraft to monitor density and test ecological hypotheses. Because the distribution of wintering waterfowl populations is spatially clustered, precise estimation of density is difficult. Ideally, spatially clustered populations are sampled by allocating lower sampling effort where density is low, and higher effort where density is high and variable. However, blue-winged teal birds are highly mobile and move readily in response to changes in environmental conditions. Thus, information on areas with high and low densities of waterfowl is not always available before a survey is implemented, excluding stratification except at regional scales. Thompson (1990, 1991a, 1991b) developed adaptive cluster sampling designs that allow observed values to trigger increased sampling effort during the survey. When applied to waterfowl surveys, adaptive cluster sampling designs assign high probabilities to samples that include areas with high waterfowl density. These intuitively appealing designs can have lower variance than conventional designs.

Description of the target population as per Smith *et al.* (1995) is given here. Efforts were made to count every individual duck of the three waterfowl species (Ring-necked ducks, Blue-winged teal bird and green-winged teal) in a 5000 km^2 area of central Florida by making systematic flights over the entire study region. The areas of available habitat was computed from a land classification by the Florida Game and Fresh Water Fish Commission (Wickham and Kautz, 1989). The study area of 5000 km^2 area can be divided in 200 quadrats of 25 km^2 each. Thus, the population size is considered to be N=200. Fig. 2 shows the blue-winged teal bird population given in Smith *et al.* (1995). It shows number of blue-winged teal bird observed in 25 km^2 quadrats. Under the simulation study, this variable has been considered as character of interest *Y*. Zero counts of the bird are also given in the dataset. It can be observed that the population mean of this blue-winged teal population is $\bar{Y}$ =70.60 i.e. approximately 71 birds per quadrats.

0	0	0	0	0	0	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	**20**	**4**	**2**	**12**	0	0	0	0	0	**10**	**103**	0	0	0
0	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	**150**	**7144**	**1**	0
0	0	0	0	0	0	0	0	**2**	0	0	0	0	**2**	0	0	**6**	**6339**	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**14**	**122**	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**114**	**60**
0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0	**3**	0

Fig. 2. Blue-winged teal bird population from Smith *et al.* (1995)

In environmental and ecological studies, in addition to total number (density) of a particular species of bird in certain area, researchers often study the ratio of major bird species to assess the resource availability under different habitats for proper management and conservation of bird species. In such scenario, ratio of bird species may act as the parameter of interest considering number of each birds (say ring necked duck and blue winged teal) as variable *X* and *Y*. It has been observed that the teals shared herbaceous wetlands with ring-necked ducks, teal were not found in great numbers in open water. Since, habitat acts as an important driving force to the numbers of specific bird species, thus its area can be chosen as respective auxiliary information. Here, areas of open water and wetlands can be considered as auxiliary information for ring necked duck and blue winged teal count.

In this study, we have proposed calibration estimator for estimation of finite population ratio under adaptive cluster sampling design using the available auxiliary information when the members bearing a characteristic of interest are sparsely scattered in a geographically distributed population. Under the simulation study, an auxiliary variable that are highly correlated with study variable *Y* were generated. Based on levels of correlation, two population sets i.e. Set 1 and Set 2 have been generated. Here, it has been assumed that the auxiliary variable *X* follows Normal distribution considering:

Mean of X ($\bar{X}$) = 30, Standard deviation of $X(\sigma_X)$ = 10 and Correlation coefficient between *X* and the study variate *Y* i.e. ρ_{XY} = 0.9 (Set 1) and 0.65 (Set 1).

It was found that the population ratio $R = \bar{Y}/\bar{X}$ were 2.35 and 2.31 in population Set 1 and 2 respectively. Further, for the purpose of empirical evaluation of proposed calibration estimators of population ratio under adaptive cluster sampling, couple of auxiliary variable *U* and *V* which are correlated with the variable *Y* and *X* respectively have been generated. Here, it has been assumed that *U* and *V* follow Normal distribution considering:

Variable	Mean	Standard deviation	Correlation coefficient
U	$\bar{U}$ = 35	σ_U = 10	ρ_{YU} = 0.85 (Set 1) and 0.6 (Set 2)
V	$\bar{V}$ = 40	σ_V = 10	ρ_{XV} = 0.8 (Set 1) and 0.6 (Set 2)

Under the simulation study, there is need to select samples using adaptive cluster sampling design from the above described study population. An initial sample of size n_1 is taken using Simple Random Sampling (SRS) from the population of N=200 quadrats. If the observed *y* value of a sampled unit satisfies a condition C, i.e. $y_i > 0$, then the rest of the unit's neighbourhood is added to the sample. If any other units in that neighbourhood satisfy the condition C, then their neighbourhoods are also added to the sample. The process is continued until a cluster of units is obtained that contains a network and edge units. The final sample then consists of n_1 (not necessarily distinct) clusters, one for each unit selected in the initial sample.

In this simulation study, initial samples of different sizes has been taken to obtain adaptive cluster samples. It ranges from 5% to 20% of N. Different sizes of the initial samples are considered as n_1 = 10, 15, 20, 25, 30, 35 and 40. We used Monte Carlo simulation to draw samples from the enumerated populations. From both the study populations, we have selected a total of 5000 independent samples of above mentioned sizes using adaptive cluster sampling design.

Further, from each of the 5000 independent samples, using the observations from the variables Y, X, U and V, estimates of the proposed calibration estimators $\left(\hat{R}_{CAL(AC)}\right)$ of population ratio under adaptive cluster sampling as well as Horvitz-Thompson estimator based estimator of the population ratio under adaptive cluster sampling i.e. $\hat{R}_{HT(AC)}$ have been calculated. The formula as given in Equation 6 (putting $q_i = 1$) and 1 respectively are used to obtain the estimates of population ratio. Values of these estimators were averaged over 5000 replicate samples which can be interpreted as expected values, i.e., $E\left(\hat{R}_D\right) = \frac{1}{S}\sum_{i=1}^{S}\hat{R}_{D,i}$, where S denotes the number of sample replications (i.e., S = 5000), D denotes the design: a combination of initial sample size (n_1) and the estimation procedure i.e. Calibration estimators or Horvitz-Thompson estimator and $\hat{R}_{D,i}$ denotes the value of the estimator of population ratio (R) obtained by D^{th} design at i^{th} sample iteration.

Developed calibration estimators as well as usual Horvitz-Thompson estimators of population ratio under adaptive cluster sampling were evaluated on the basis of two measures viz. percentage Relative Bias (% RB) and percentage Relative Root Mean Squared Error (% RRMSE) as given by

$$\%RB\left(\hat{R}_D\right) = \frac{1}{S}\sum_{i=1}^{S}\left(\frac{\hat{R}_{D,i} - R}{R}\right)\times 100 \text{ and}$$

$$\%RRMSE\left(\hat{R}_D\right) = \sqrt{\frac{1}{S}\sum_{i=1}^{S}\left(\frac{\hat{R}_{D,i} - R}{R}\right)^2}\times 100 .$$

Simulation study was implemented based on developed SAS macros for empirical evaluation of the proposed calibration estimators.

5. RESULTS AND DISCUSSION

The following tables contain the results obtained for each of the cases considered under the simulation study. Table 1 and 2 present the % RB and % RRMSE of both the estimators of population ratio $\left(R = \bar{Y}/\bar{X}\right)$ under adaptive cluster sampling for different initial sample sizes (n_1) from population Set 1 and 2 respectively. The proposed calibration estimator i.e. $\hat{R}_{CAL(AC)}$ of population ratio (R) is compared with usual Horvitz-Thompson based estimator ($\hat{R}_{HT(AC)}$) estimator under adaptive cluster sampling considering $q_i = 1$.

Table 1. % RB and % RRMSE of both the estimators of population ratio ($R = \bar{Y}/\bar{X}$) for different initial sample sizes (n_1) under adaptive cluster sampling from population Set 1

n_1	% RB		% RRMSE	
	$\hat{R}_{HT(AC)}$	$\hat{R}_{CAL(AC)}$	$\hat{R}_{HT(AC)}$	$\hat{R}_{CAL(AC)}$
10	-4.60	-7.93	138.42	131.20
15	-3.23	-6.02	109.02	103.35
20	-3.43	-5.17	89.94	86.03
25	-0.55	-2.84	75.58	73.19
30	-1.00	-2.15	65.24	63.70
35	-0.83	-1.38	56.47	55.09
40	-0.63	-1.19	49.04	48.41

Table 2. % RB and % RRMSE of both the estimators of population ratio ($R = \bar{Y}/\bar{X}$) for different initial sample sizes (n_1) under adaptive cluster sampling from population Set 2

n_1	% RB		% RRMSE	
	$\hat{R}_{HT(AC)}$	$\hat{R}_{CAL(AC)}$	$\hat{R}_{HT(AC)}$	$\hat{R}_{CAL(AC)}$
10	-6.11	-8.75	141.13	136.71
15	-0.52	-2.23	110.66	115.74
20	-6.53	-5.77	91.31	95.08
25	-1.28	-2.53	76.72	81.53
30	3.00	1.38	64.34	70.31
35	-2.52	-2.06	58.16	63.44
40	-2.03	-2.27	50.41	55.50

From Table 1, it is visible that the proposed calibration estimator ($\hat{R}_{CAL(AC)}$) of the population ratio ($R = \bar{Y}/\bar{X}$) is producing slightly higher amount of % RB than Horvitz-Thompson based estimator ($\hat{R}_{HT(AC)}$) of the population ratio from population Set 1.

Absolute value of the % RB decreases with the increase in initial sample sizes (n_1). Table 1 also dictates that the proposed calibration estimator ($\hat{R}_{CAL(AC)}$) of the population ratio is giving lesser % RRMSE than Horvitz-Thompson based estimator ($\hat{R}_{HT(AC)}$). % RRMSE decreases with the increase in initial sample sizes (n_1) which shows the consistency of the proposed estimator. Similar trends in % RB and % RRMSE can also be seen in Table 2 for population Set 2 with lesser correlations among the variables. Although, the % RRMSE of both the estimators were found to be higher in this case in comparison to population Set 1 with higher correlations among the variables which makes the estimators less efficient. Thus, the proposed calibration estimator of the population ratio is more efficient than Horvitz-Thompson based estimator of the population ratio for different levels of correlations among the study and auxiliary variables.

6. CONCLUSIONS

Adaptive Cluster Sampling (ACS) is a survey sampling procedure suitable for sampling geographically rare and hidden clustered population. ACS design allow observed values to trigger increased sampling effort during the survey. Under this study, calibration estimator of finite population ratio has been developed under adaptive cluster sampling when auxiliary information is available for both the variables associated to the population ratio. Expressions of approximate variance and estimator of variance of the proposed estimator have also been derived. Statistical performance of the proposed calibration estimator of population ratio under ACS are evaluated through a simulation study with respect to conventional Horvitz-Thomson estimator using a real dataset as a study population. The results of the simulation study show that the proposed calibration estimator ($\hat{R}_{CAL(AC)}$) of the population ratio ($R = \bar{Y}/\bar{X}$) is more efficient than Horvitz-Thompson based estimator ($\hat{R}_{HT(AC)}$) of the population ratio on the basis of % RRMSE.

ACKNOWLEDGEMENTS

The authors are grateful to the editors and the reviewers for their constructive suggestions/comments that helped to improve the quality of the manuscript.

REFERENCES

Brown, J.A. (2003). Designing an efficient adaptive cluster sample. *Environmental and Ecological Statistics*, **10**, 95-105.

Chao C.T., Lin, F.M. and Chiang, T.C. (2008). Improved ratio estimators in adaptive cluster sampling. *Section on Survey Research Methods - JSM 2008*.

Chutiman, N. (2013). Adaptive cluster sampling using auxiliary variable. *J. Math. Stat.,* **9(3),** 249-255.

Chutiman, N. and Chiangpradit, M. (2014). Ratio estimator in adaptive cluster sampling without replacement of networks. *Journal of Probability and Statistics*, DOI: http://dx.doi.org/10.1155/2014/726398.

Deville, J.C. and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, **87**, 376-382.

Dryver A.L. and Chao C.T. (2007). Ratio estimators in adaptive cluster sampling. *Environmetrics*, **18**, 607-620.

Dryver, A.L. and Thompson, S.K. (1998). Improving unbiased estimators in adaptive cluster sampling. In *ASA Proceedings of the Section on Survey Research Methods*, 727-731.

Särndal, C.E., Swensson, B. and Wretman, J. (1992). *Model-Assisted Survey Sampling*. Springer-Verlag.

Särndal, C.E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, **33**, 99-119.

Smith, D.R., Conroy, M.J. and Brakhage, D.H. (1995) Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics*, **51**, 777-788.

Smith, D.R., Villella, R.F. and Lemarié, D.P. (2003). Application of adaptive cluster sampling to low-density populations of freshwater mussels. *Environmental and Ecological Statistics*, **10**, 7-15.

Smith, D.R., Brown, J.A., and Lo, N.C.H. (2004). *Application of adaptive cluster sampling to biological populations*. Pages 77-122 In Thompson, W.L. editor. Sampling rare or elusive species: Concepts, designs, and techniques for estimating population parameters. Island Press, Covelo, CA.

Thompson, S.K. (1990). Adaptive Cluster Sampling. *Journal of the American Statistical Association*, **85**, 1050-1059.

Thompson, S.K. (1991a). Adaptive cluster sampling: Designs with primary and secondary units. *Biometrics*, **47**, 1103-1115.

Thompson, S.K. (1991b). Stratified adaptive cluster sampling. *Biometrika*, **78**, 389-397.

Thompson, S.K. (1992). *Sampling*, John Wiley and Sons, New York.

Thompson, S.K. and Seber, G.A.F. (1996). *Adaptive Sampling*, John Wiley and Sons, New York.