https://doi.org/10.56093/JISAS.V78I3.9
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 78(3) 2024 253–266

Ranked Set Sampling for Small Area Estimation using Auxiliary Data: Insights from Crop Production Data

Anoop Kumar[1], Shashi Bhushan[2] and Rohini Pokhrel[3]
[1]*Central University of Haryana, Mahendergarh*
[2]*University of Lucknow, Lucknow*
[3]*Dr. Shakuntala Misra National Rehabilitation University, Lucknow*

Received 19 October 2024; Revised 16 January 2025; Accepted 24 January 2025

SUMMARY

Small area estimation (SAE) is a critical statistical approach used to generate credible estimates for subpopulations or regions with small sample numbers. In agricultural research, reliable crop production estimation at small geographical scales is critical for policy development, resource allocation, and decision-making. Ranked set sampling (RSS), which is recognised for being a cost-effective and accurate data gathering method, is combined with auxiliary data to increase accuracy of the estimates for small regions. This study proposes synthetic ratio type estimators by combining RSS with auxiliary information to improve SAE efficiency and precision, notably in agricultural production estimation. The mean square error (MSE) of the proposed synthetic estimator is obtained to the first order approximation. The approach uses auxiliary information to lower the MSE of the estimators. Comparative investigations with certain well-known adapted SAE estimators reveal that using RSS and auxiliary data considerably improves estimation accuracy for small regions. The theoretical results are supported with a simulation study carried out over an artificially rendered population. The practical benefits of the suggested estimator are demonstrated by an application to crop production data. The findings indicate that this technique is not only more efficient, but also highly applicable in real-world agricultural surveys, making it a useful tool for improving small area estimates in resource-constrained contexts.

Keywords: Small area estimation; Ranked set sampling; Mean square error; Efficiency.

MSC2020:62D05

1. INTRODUCTION

Ranked set sampling (RSS) offers a cost-effective alternative to simple random sampling (SRS), particularly in situations where obtaining precise measurements is expensive or time consuming, but ranking units is relatively easy and inexpensive. McIntyre (1952) initially proposed the concept of RSS in agricultural contexts, proving that it outperformed SRS in terms of mean estimation efficiency. Subsequent research expanded on the findings of McIntyre (1952), analysing the statistical features of RSS. Dell and Clutter (1972) developed a formal theoretical foundation for the use of RSS in mean estimation, proving that RSS produces an unbiased estimate of the population average. Stokes and Sager (1988) used RSS in health research to show that it can efficiently determine the average concentration of a biological marker in a population by using simplified, non-invasive procedures for ranking. Al-Saleh and Samawi (2000) demonstrated that RSS gives more efficient mean estimates than SRS, particularly when the population distribution is skewed or auxiliary information is given to assist ranking. Stokes (1977) conducted a detailed comparison of RSS and SRS for mean estimation, demonstrating that RSS significantly reduces the variance of the estimator, particularly in populations with high variability. Furthermore, Patil *et al.* (1994) investigated the performance of RSS in contrast to stratified and cluster sampling, demonstrating that RSS outperforms both approaches when ranking is possible. Chen *et al.* (2004) expanded these studies by assessing RSS robustness under various ranking errors and

Corresponding author: Shashi Bhushan
E-mail address: bhushan_s@lkouniv.ac.in

departures from ideal circumstances. Zamanzade and Al-Omari (2016) presented a novel RSS technique for calculating population mean and variance. Mahdizadeh and Zamanzade (2022) used a rank-based methodology to estimate the prevalence of breast cancer. In recent decades, various authors have focused on estimating population means using single and multi-auxiliary information under RSS. Samawi and Muttlak (1996) calculated the population mean ratio using RSS. Al-Omari *et al.* (2009) presented a novel ratio estimator of population mean under RSS. Al-Omari and Bouza (2015) proposed ratio estimators of the population mean with missing data in RSS. Bhushan and Kumar (2022a) proposed an optimum class of estimator for RSS, whereas Bhushan and Kumar (2022b) proposed efficient logarithmic estimators using multi-auxiliary information. Rehman and Shabbir (2022) proposed an efficient class of estimators for limited population means in the presence of non-response under RSS. Kocyigit (2023) proposed a new sub-type mean estimator for RSS with dual auxiliary variables. Bhushan and Kumar (2023) investigated the imputation of missing data using multi-auxiliary information via RSS. Punia *et al.* (2024) introduced an enhanced Beluga Whale optimization algorithm for engineering optimization problems. Raj *et al.* (2024) developed a new hybrid pelican-particle swarm optimization algorithm for global optimization issue. Bhushan and Kumar (2024a) proposed various RSS imputation algorithms for correlated measurement errors. Bhushan and Kumar (2024b) examined various unique logarithmic imputation approaches for RSS. Bhushan and Kumar (2024b) inspired Kumar *et al.* (2024a) to develop some unique logarithmic imputation algorithms under RSS that make use of multi-auxiliary information.

Small area estimation (SAE) is a statistical technique used in survey sampling to provide reliable estimates for subpopulations or areas where typical survey methods fail owing to small sample sizes. These small areas might be geographic regions, demographic groupings, or other domains of interest in which direct estimates based simply on survey data are frequently incorrect, resulting in significant variability or large standard errors. To overcome this issue, SAE applies sophisticated approaches such as direct and synthetic estimators. Direct estimators use only data collected from the small area itself, which can result in unreliable estimates when the sample size is small. However, the synthetic estimators borrow strength from adjacent regions by assuming that distinct small areas have comparable features, so, employing data from the larger population or auxiliary information to increase precision. SAE improves the accuracy of estimates by combining these methodologies, making it vital for policymaking, resource allocation, and decision-making in a variety of domains including health, economics, and social sciences.

Several authors developed various design based direct and synthetic estimation methods for the estimation of population mean under SRS using single and multiple auxiliary information. Tikkiwal and Ghiya (2000) suggested a generalized class of synthetic estimators with the application of crop acreage estimation of small areas, whereas Pandey and Tikkiwal (2010) investigated a generalized class of synthetic estimators for small domain under systematic sampling. Rai and Pandey (2013) developed the synthetic estimators for small domains using auxiliary information. Tikkiwal *et al.* (2013) examined the performance of generalized regression estimator for small domains. Ashutosh *et al.* (2024) presented a simulation analysis of non-respondent information for small domain. Kumar *et al.* (2024b) done a small area estimation using some design based direct and synthetic logarithmic estimators, whereas Kumar *et al.* (2024c) designed some enhanced direct and synthetic estimators for domain mean with simulation and applications. Recently, Ahmed *et al.* (2024) developed an indirect estimation of small area parameters under RSS.

In the literature, there exist no domain mean estimation methods based on bivariate auxiliary information under RSS. Therefore, we develop the fundamental theory and notations for SAE methods under RSS utilizing bivariate auxiliary information. To fill the gap of the literature review, we adapt the synthetic mean, ratio, and power ratio estimators employing bivariate auxiliary information for estimating the domain mean under RSS. To provide efficient estimates of domain mean, we propose the synthetic Searls power ratio estimators employing bivariate auxiliary information under RSS. By incorporating auxiliary data, the proposed method further improves the efficiency of small area estimators, which is especially useful in agricultural contexts where reliable crop production estimates are crucial for decision-making, resource allocation, and policy formulation. Highlighting these scenarios, such as domains with limited resources or sparse data, underscores the practical applicability and significance of our estimators.

1.1 Methodology and notations

The RSS is a statistical technique that enhances estimation efficiency by utilizing auxiliary information to improve the representativeness of a sample. When applied to SAE, the procedure involves the following steps:

i. The population is divided into small areas, often defined by geographic or administrative boundaries. The goal is to produce reliable estimates for each of these areas, despite limited data availability.
ii. From the target population, multiple sets of samples, each of size m, are drawn randomly without measurement. These samples form the basis for ranking.
iii. Within each set, the units are ranked based on a judgment variable or auxiliary information that correlates with the variable of interest. Ranking can be performed visually, through expert opinion, or by using cost-free measurements of the auxiliary variable.
iv. After ranking, a specific unit (e.g., the first smallest unit from first set, second smallest unit from second set, or m^{th} smallest unit from m^{th}) is selected from each set for precise measurement. This process completes a cycle.
v. The cycle is repeated r times to produce $n=mr$ ranked set sample.
vi. The RSS data are integrated into a small area estimator. This estimator combines information from multiple areas and leverages correlations between areas to produce more precise estimates for each small area.
vii. RSS improves efficiency by reducing the variability in estimates due to better representativeness of the sample. The resulting small area estimators are more reliable, especially in cases of limited sample size or scarce direct measurements in some areas.

Let the N identifiable units construct the population $\Lambda=(\Lambda_1,\Lambda_2,\ldots,\Lambda_N)$. Let y_i and (x_i, z_i) be the observed units of the main variable y and auxiliary variables (x,z), respectively. Let the ranking be done on the variable x to estimate the parameter $\bar{Y}$. Let ($y_{[i]}$, $x_{(i)}$, $z_{[i]}$) denote the i^{th} ranked set sample provided $x_{(i)}$ is the i^{th} order statistics in the i^{th} sample for variable x and $y_{[i]}$ and $z_{[i]}$ are the i^{th} judgment order in the i^{th} sample for variables (y,z). The perfect and imperfect ranking of the units are indicated by the parentheses () and [], respectively, utilized in the subscript of x and (y,z). The notations used throughout the paper are defined below:

$\bar{Y}$: population mean utilizing N observations on y;

$\bar{Y}_a$: population mean of domain a utilizing N_a observations on y;

$\bar{X}$: population mean of variable x utilizing N observations;

$\bar{X}_a$: population mean of variable x for domain a utilizing N_a observations;

$\bar{Z}$: population mean of variable z utilizing N observations;

$\bar{Z}_a$: population mean of variable z for domain a utilizing N_a observations;

$\bar{x}$: sample mean utilizing n observations on characteristic x;

$\bar{x}_a$: sample mean utilizing n_a observations on x;

$\bar{z}$: sample mean utilizing n observations on characteristic z;

$\bar{z}_a$: sample mean utilizing n_a observations on z;

$\bar{y}$: sample mean utilizing n observations on y;

$\bar{y}_a$: sample mean of domain a utilizing n_a observations on y;

S_x^2: population mean square of variable x;

$S_{x_a}^2$: population mean square of variable x for the domain a;

S_z^2: population mean square of variable z;

$S_{z_a}^2$: population mean square of variable z for the domain a;

S_y^2: population mean square of variable y;

$S^2_{y_a}$: population mean square of variable y for the domain a;

$C_x = S_x / \bar{X}$: population variation coefficient of variable x;

$C_{x_a} = S_{x_a} / \bar{X}_a$: population variation coefficient of variable x for domain a;

$C_z = S_z / \bar{Z}$: population variation coefficient of variable z;

$C_{z_a} = S_{z_a} / \bar{Z}_a$: population variation coefficient of variable z for domain a;

$C_y = S_y / \bar{Y}$: population variation coefficient of variable y;

$C_{y_a} = S_{y_a} / \bar{Y}_a$: population variation coefficient of variable y for domain a;

ρ_{yx} : correlation coefficient of variables y and x;

ρ_{yz} : correlation coefficient of variables y and z;

ρ_{xz} : correlation coefficient of variables x and z;

$\rho_{y_a x_a}$: correlation coefficient of variables y and x for the domain a;

$\rho_{y_a z_a}$: correlation coefficient of variables y and z for the domain a;

$\rho_{x_a z_a}$: correlation coefficient of variables x and z for the domain a.

The bias and MSE of the synthetic estimators can be determined by assuming the following notations:

$\bar{y}_{[n]} = \bar{Y}\left(1+\varepsilon_0\right)$,

$\bar{x}_{(n)} = \bar{X}\left(1+\varepsilon_1\right)$,

$\bar{z}_{[n]} = \bar{Z}\left(1+\varepsilon_2\right)$,

such that $E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_2) = 0$,

$E\left(\varepsilon_0^2\right) = \gamma C_y^2 - W^2_{y_{[i]}} = \Delta_0$,

$E\left(\varepsilon_1^2\right) = \gamma C_x^2 - W^2_{x_{(i)}} = \Delta_1$,

$E\left(\varepsilon_2^2\right) = \gamma C_z^2 - W^2_{z_{[i]}} = \Delta_2$,

$E(\varepsilon_0\varepsilon_1) = \gamma\, \rho_{yx} C_y C_x - W_{yx_{[i]}} = \Delta_{01}$,

$E(\varepsilon_0\varepsilon_2) = \gamma\, \rho_{yz} C_y C_z - W_{yz_{[i]}} = \Delta_{02}$,

$E(\varepsilon_1\varepsilon_2) = \gamma\, \rho_{xz} C_x C_z - W_{xz_{[i]}} = \Delta_{12}$,

where $\gamma = 1/n$, $C_y = S_y / \bar{Y}$, $C_x = S_x / \bar{X}$, $C_z = S_z / \bar{Z}$,

$$W^2_{y[i]} = \sum_{i=1}^{m}\left(\mu_{y_{[i]}} - \bar{Y}\right)^2 / m^2 r \bar{Y}^2,\; W^2_{x_{(i)}} = \sum_{i=1}^{m}\left(\mu_{x_{(i)}} - \bar{X}\right)^2 / m^2 r \bar{X}^2,$$

$$W^2_{z_{[i]}} = \sum_{i=1}^{m}\left(\mu_{z_{[i]}} - \bar{Z}\right)^2 / m^2 r \bar{Z}^2,$$

$$W_{yx_{[i]}} = \sum_{i=1}^{m}\left(\mu_{y_{[i]}} - \bar{Y}\right)\left(\mu_{x_{(i)}} - \bar{X}\right) / m^2 r \overline{YX},$$

$$W_{yz_{[i]}} = \sum_{i=1}^{m}\left(\mu_{y_{[i]}} - \bar{Y}\right)\left(\mu_{z_{[i]}} - \bar{Z}\right) / m^2 r \overline{YZ},$$

$$W_{xz_{[i]}} = \sum_{i=1}^{m}\left(\mu_{x_{(i)}} - \bar{X}\right)\left(\mu_{z_{[i]}} - \bar{Z}\right) / m^2 r \bar{X}\overline{Z},\; \mu_{y_{[i]}} = E\left(Y_{[i]}\right),$$

$\mu_{x_{(i)}} = E\left(X_{(i)}\right)$, and $\mu_{z_{[i]}} = E\left(Z_{[i]}\right)$.

In the next part, certain traditional synthetic estimators are adapted and their MSE expressions are calculated. In Section 3, we offer the synthetic Searls power ratio estimators for domain mean under RSS and find out corresponding MSE formulations. In Section 4, we compare the suggested Searls type synthetic estimators to the adapted synthetic mean, ratio, and power ratio estimators. In Section 5, a simulation study is undertaken on an artificially generated symmetric population, while in Section 6, the adapted and suggested synthetic estimators are applied to real data. The study is concluded in Section 7.

2. ADAPTED ESTIMATORS

The choice of synthetic estimators over other indirect estimators, such as composite estimators, was primarily driven by their simplicity and computational efficiency in the context of the study. Synthetic estimators based on pooling information across areas using a model, which is particularly advantageous when dealing with small areas with limited or no direct observations. While composite estimators combine direct and indirect estimates to balance bias and variability, they often require additional modelling complexity and reliable direct estimates, which may not always be available in certain applications, such as those with sparse or imbalanced data. Literature contains no estimator for estimating the domain mean under RSS by utilizing bi-variate auxiliary information. To fill this gap, we have adapted synthetic mean estimator (SME), synthetic ratio estimator (SRE), and synthetic power ratio estimator (SPRE).

2.1 Synthetic mean estimators

A synthetic mean estimator is a technique used in SAE to produce estimates for areas with limited or no direct data (see, Rao and Molina (2015)). When no auxiliary information is available, then we have an obvious choice of the SME for estimating the domain a which is given as

$$t_m = \bar{y}_{[n]}.$$

The bias and MSE of the SME t_m are provided by

$$Bias\left(t_m\right) = \bar{Y} - \bar{Y}_a$$

and $MSE\left(t_m\right) = \left(\bar{Y} - \bar{Y}_a\right)^2 + \bar{Y}^2\Delta_0$.

2.2 Synthetic ratio estimator

A synthetic ratio estimator is an estimation method used in SAE to improve the precision of estimates by using auxiliary information. It combines data from multiple small areas under the assumption that they share a common relationship between the main and the auxiliary variables. Specifically, the estimator computes the ratio of the total of the main variable to the total of the auxiliary variable in a larger domain or combined area and then applies this ratio to the auxiliary total of the target small area. Following Cochran (1977) and utilizing bivariate auxiliary information, the SRE for estimating the mean of domain a under RSS is given as

$$t_r = \bar{y}_{[n]}\left(\frac{\bar{X}_a}{\bar{x}_{(n)}}\right)\left(\frac{\bar{Z}_a}{\bar{z}_{[n]}}\right).$$

The bias and MSE of the SRE t_r are given by

$$Bias\left(t_r\right) = \bar{Y}_a\left(\Delta_1 + \Delta_2 - 2\Delta_{01} - 2\Delta_{02} + 2\Delta_{12}\right)$$

and $MSE\left(t_r\right) = \bar{Y}_a^2\left(\Delta_0 + \Delta_1 + \Delta_2 - 2\Delta_{01} - 2\Delta_{02} + 2\Delta_{12}\right)$.

2.3 Synthetic power ratio estimator

In order to improve the ratio estimator, Khare and Ashutosh (2018) extended the work of Tikkiwal and Ghiya (2000) and introduced power ratio estimator for domain mean estimation under SRS using bivariate auxiliary information. Following Tikkiwal and Ghiya (2000) and Khare and Ashutosh (2018), and utilizing bivariate auxiliary information, the SPRE for estimating the mean of domain a under RSS is given as

$$t_p = \bar{y}_{[n]}\left(\frac{\bar{X}_a}{\bar{x}_{(n)}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{z}_{[n]}}\right)^{\theta}.$$

where β and θ are constants. The bias and minimum MSE of the SPRE t_p are given by

$$Bias\left(t_p\right) = \bar{Y}_a\left\{\frac{\beta(\beta+1)}{2}\Delta_1 + \frac{\theta(\theta+1)}{2}\Delta_2 - \beta\Delta_{01} - \theta\Delta_{02} + \beta\theta\Delta_{12}\right\}$$

and $min.MSE\left(t_p\right) = \bar{Y}_a^2\left(\Delta_0 - \frac{\Delta_{01}^2\Delta_2 + \Delta_{02}^2\Delta_1 - 2\Delta_{01}\Delta_{02}\Delta_{12}}{\Delta_1\Delta_2 - \Delta_{12}^2}\right)$.

3. PROPOSED ESTIMATORS

According to Searls (1964), the efficiency of the estimators can be enhanced by multiplying them with a tuning scalar. The estimator proposed by Searls (1964) indeed relies on a prior knowledge of the population coefficient of variation. RSS fits into this context by using the auxiliary variable for ranking purposes, thereby improving the representativeness of the sample without additional measurement costs. The auxiliary variable facilitates the ranking of units within sets and serves as a source of auxiliary information in the estimation process. Following Searls' suggestion, we multiplied a tuning scalar in the SPRE and proposed the equivalent synthetic Searls power ratio estimator (SSPRE) for the mean of domain 'a' based on bivariate auxiliary information under RSS as

$$t_k = \alpha\bar{y}_{[n]}\left(\frac{\bar{X}_a}{\bar{x}_{(n)}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{z}_{[n]}}\right)^{\theta}.$$

where α , β, and θ are constants.

3.1 Particular cases

When $\alpha = 1$ and $\beta = \theta = 0$, then the SSPRE t_k converts into the synthetic meanestimator t_m .

When α = β = θ = 1, then the SSPRE t_k converts into the synthetic ratio estimator t_r .

When $\alpha = 1$, then the SSPRE t_k converts into the synthetic power ratio estimator t_p .

Theorem 3.1. The bias and minimum MSE of the SSPRE t_k are given up to first order approximation as

$$Bias\left(t_k\right) = \left(\alpha\bar{Y}_a - \bar{Y}_a\right) + \alpha\bar{Y}_a\left\{\frac{\beta(\beta+1)}{2}\Delta_1 + \frac{\theta(\theta+1)}{2}\Delta_2 - \beta\Delta_{01} - \theta\Delta_{02} + \beta\theta\Delta_{12}\right\}$$

and $min.MSE(t_k) = \bar{Y}_a^2\left(1-\frac{Q^2}{P}\right)$.

Proof. Consider the SSPRE t_k as

$$t_k = \alpha \bar{y}_{[n]}\left(\frac{\bar{X}_a}{\bar{x}_{(n)}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{z}_{[n]}}\right)^{\theta}.$$

We represent the SSPRE t_k using the earlier specified notations as

$$t_k = \alpha\bar{Y}(1+\varepsilon_0)\left(\frac{\bar{X}_a}{\bar{X}(1+\varepsilon_1)}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}(1+\varepsilon_2)}\right)^{\theta},$$

$$=\alpha\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}}\right)^{\theta}(1+\varepsilon_0)(1+\varepsilon_1)^{-\beta}(1+\varepsilon_2)^{-\theta},$$

$$=\alpha\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}}\right)^{\theta}(1+\varepsilon_0)\left\{1-\beta\varepsilon_1+\frac{\beta(\beta+1)}{2}\varepsilon_1^2\right\}\left\{1-\theta\varepsilon_2+\frac{\theta(\theta+1)}{2}\varepsilon_2^2\right\},$$

$$=\alpha\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}}\right)^{\theta}\left\{1+\varepsilon_0-\beta\varepsilon_1-\theta\varepsilon_2+\frac{\beta(\beta+1)}{2}\varepsilon_1^2+\frac{\theta(\theta+1)}{2}\varepsilon_2^2-\beta\varepsilon_0\varepsilon_1-\theta\varepsilon_0\varepsilon_2+\beta\theta\varepsilon_1\varepsilon_2\right\}.$$

Subtracting $\bar{Y}_a$ on both sides of the above equation yields:

$$t_k-\bar{Y}_a=\left[\alpha\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}}\right)^{\theta}\left\{\varepsilon_0-\beta\varepsilon_1-\theta\varepsilon_2+\frac{\beta(\beta+1)}{2}\varepsilon_1^2+\frac{\theta(\theta+1)}{2}\varepsilon_2^2-\beta\varepsilon_0\varepsilon_1-\theta\varepsilon_0\varepsilon_2+\beta\theta\varepsilon_1\varepsilon_2\right\}+\left\{\alpha\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\beta}\left(\frac{\bar{Z}_a}{\bar{Z}}\right)^{\theta}-\bar{Y}_a\right\}\right]. \quad (3.1)$$

Assuming synthetic power ratio estimation $\bar{Y}\left(\bar{X}_a/\bar{X}\right)^{\beta}\left(\bar{Z}_a/\bar{Z}\right)^{\theta}\approx\bar{Y}_a$, the equation (3.1) may be expressed as

$$t_k-\bar{Y}_a=\left(\alpha\bar{Y}_a-\bar{Y}_a\right)+\alpha\bar{Y}_a\left\{\varepsilon_0-\beta\varepsilon_1-\theta\varepsilon_2+\frac{\beta(\beta+1)}{2}\varepsilon_1^2+\frac{\theta(\theta+1)}{2}\varepsilon_2^2-\beta\varepsilon_0\varepsilon_1-\theta\varepsilon_0\varepsilon_2+\beta\theta\varepsilon_1\varepsilon_2\right\}. \quad (3.2)$$

We obtain the bias of the proposed SSPRE t_k by considering expectation on both sides of equation (3.2) as

$$Bias(t_k)=\left(\alpha\bar{Y}_a-\bar{Y}_a\right)+\alpha\bar{Y}_a\left\{\frac{\beta(\beta+1)}{2}\Delta_1+\frac{\theta(\theta+1)}{2}\Delta_2-\beta\Delta_{01}-\theta\Delta_{02}+\beta\theta\Delta_{12}\right\}.$$

To derive the MSE of the SSPRE t_k to the first order approximation, we square and take expectation to either side of (3.2) as

$$MSE(t_k)=\bar{Y}_a^2\left[1+\alpha^2\left\{1+\Delta_0+\left(2\beta^2+\beta\right)\Delta_1+\left(2\theta^2+\theta\right)\Delta_2-4\beta\Delta_{01}-4\theta\Delta_{02}+4\beta\theta\Delta_{12}\right\}-2\alpha\left\{1+\frac{\beta(\beta+1)}{2}\Delta_1+\frac{\theta(\theta+1)}{2}\Delta_2-\beta\Delta_{01}-\theta\Delta_{02}+\beta\theta\Delta_{12}\right\}\right],$$

$$=\bar{Y}_a^2\left(1+\alpha^2P-2\alpha Q\right). \quad (3.3)$$

where

$$P=1+\Delta_0+\left(2\beta^2+\beta\right)\Delta_1+\left(2\theta^2+\theta\right)\Delta_2-4\beta\Delta_{01}-4\theta\Delta_{02}+4\beta\theta\Delta_{12}$$

and

$$Q=1+\frac{\beta(\beta+1)}{2}\Delta_1+\frac{\theta(\theta+1)}{2}\Delta_2-\beta\Delta_{01}-\theta\Delta_{02}+\beta\theta\Delta_{12}.$$

Minimizing (3.3) regarding α, we get the optimum value of α as

$$\alpha_{(opt)}=\frac{Q}{P}.$$

Using the aforementioned optimum value of α in (3.3), we obtain the minimum MSE of the SSPRE t_k as

$$min.MSE(t_k)=\bar{Y}_a^2\left(1-\frac{Q^2}{P}\right).$$

4. EFFICIENCY CONDITIONS

This section compares the minimum MSEs of the adapted and proposed synthetic estimators, and the efficiency conditions are given.

Lemma 4.1. The suggested SSPRE t_k dominates the SME t_m, if

$$MSE(t_k)<MSE(t_m)$$

$$\Rightarrow \bar{Y}_a^2\left(1-\frac{Q^2}{P}\right) < \left(\bar{Y}-\bar{Y}_a\right)^2 + \bar{Y}^2\Delta_0$$

$$\Rightarrow \frac{Q^2}{P} > 1 - \frac{\left(\bar{Y}-\bar{Y}_a\right)^2}{\bar{Y}_a^2} - \frac{\bar{Y}^2}{\bar{Y}_a^2}\Delta_0 .$$

Lemma 4.2. The suggested SSPRE t_k dominates the SME t_r, if

$$MSE(t_k) < MSE(t_r)$$

$$\Rightarrow \bar{Y}_a^2\left(1-\frac{Q^2}{P}\right) < \bar{Y}_a^2\left(\Delta_0+\Delta_1+\Delta_2-2\Delta_{01}-2\Delta_{02}+2\Delta_{12}\right)$$

$$\Rightarrow \frac{Q^2}{P} > 1-\Delta_0-\Delta_1-\Delta_2+2\Delta_{01}+2\Delta_{02}-2\Delta_{12}.$$

Lemma 4.3. The suggested SSPRE t_k dominates the SME t_p, if

$$MSE(t_k) < MSE(t_p)$$

$$\Rightarrow \bar{Y}_a^2\left(1-\frac{Q^2}{P}\right) < \bar{Y}_a^2\left(\Delta_0 - \frac{\Delta_{01}^2\Delta_2+\Delta_{02}^2\Delta_1-2\Delta_{01}\Delta_{02}\Delta_{12}}{\Delta_1\Delta_2-\Delta_{12}^2}\right)$$

$$\Rightarrow \frac{Q^2}{P} > 1-\Delta_0+\frac{\Delta_{01}^2\Delta_2+\Delta_{02}^2\Delta_1-2\Delta_{01}\Delta_{02}\Delta_{12}}{\Delta_1\Delta_2-\Delta_{12}^2}.$$

These inequalities show that when the quality measure Q^2/P surpasses a certain threshold, SSPRE t_k outperforms SME t_m, SRE t_r, and SPRE t_p in terms of MSE reduction. As a result, this confirms the claim that SSPRE t_k is a more effective estimator under the provided parameters, increasing total estimation reliability. In the coming section, the inequalities derived under the above lemmas will be verified through a comprehensive simulation study.

5. SIMULATION STUDY

In this section, a simulation study is conducted to compare the accuracy of the proposed synthetic estimator with the adapted synthetic estimators. The simulation generates a symmetric population of size N = 14,000, with means $\bar{Y}$ = 16, $\bar{X}$ = 21, $\bar{Z}$ = 27, and standard deviations S_y = 12, S_x = 15, S_z = 18, while exploring various combinations of correlation coefficients ρ_{yx}, ρ_{yz}, and ρ_{xz} to assess the performance under different correlation structures.

The aforementioned population consists of seven equal domains, each containing 2000 units. From each domain, a ranked set sample of size $n=15$ is selected using the RSS scheme with a set size 3 and 5 cycles. To evaluate the performance of the adapted and suggested synthetic estimators, the MSE and percent relative efficiency (PRE) are computed through 5,000 simulation iterations. The detailed steps of the simulation process are outlined in the following points, providing clarity on the methodology employed for this analysis.

- Select a ranked set sample of size $n=15$ from each domain with the set size 3 and the number of cycles 5.
- Using the above ranked set sample, figure out the needed descriptive statistics.
- The MSE and PRE of the adapted and suggested synthetic estimators have been calculated over 5,000 iterations using the following equations.

$$MSE(t_*) = \frac{1}{5,000}\sum_{i=1}^{5,000}\left(t_* - \bar{Y}_a\right)^2,$$

$$PRE(t_*) = \frac{MSE(t_m)}{MSE(t_*)}\times 100.$$

where $t_* = t_m, t_r, t_p, \text{and}\, t_k$.

- Tables 1-3 contain the findings of the adapted and proposed synthetic estimators for different correlation combinations.

5.1 Discussion of simulation findings

Discussing simulation findings entails analysing the output data to identify patterns, confirm models, and draw conclusions about the phenomena being simulated. Therefore, it is essential to interpret the simulation results of the synthetic estimators compiled in Tables 1-3 for hypothetically generated normal population.

The findings of Table 1 indicate that the SSPRE t_k outperforms SME t_m, SRE t_r, and SPRE t_p in terms of lesser MSE and greater PRE for fixed values of $\rho_{yz}=0.75$, $\rho_{xz}=0.55$, and passably chosen values of $\rho_{yx}=0.2, 0.4, 0.6, 0.8$ throughout each domain. This shows that the SSPRE t_k is more efficient and trustworthy at capturing the relationship between variables, particularly when there are fixed correlations between variables (y, z), (x, z), and moderate to

Table 1. MSE and PRE of synthetic estimators for fixed values of $\rho_{yz} = 0.75$, $\rho_{xz} = 0.55$, and varying values of ρ_{yx} = (0.2, 0.4, 0.6, 0.8)

Domains	Estimators	ρ_{yx}							
		0.2		0.4		0.6		0.8	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	t_m	16.69	100.00	17.35	100.00	16.78	100.00	16.92	100.00
	t_r	15.40	108.37	12.67	136.92	8.37	200.40	4.86	348.03
	t_p	2.82	590.93	4.17	415.88	3.00	559.38	0.52	3272.28
	t_k	2.79	597.23	4.09	423.71	2.95	568.54	0.51	3309.88
2	t_m	16.45	100.00	17.04	100.00	16.44	100.00	16.59	100.00
	t_r	15.37	107.01	11.63	146.47	8.62	190.64	5.13	323.58
	t_p	2.84	579.99	3.83	444.40	3.09	531.95	0.55	3032.39
	t_k	2.81	585.28	3.76	452.76	3.04	540.67	0.54	3067.99
3	t_m	16.89	100.00	16.82	100.00	16.90	100.00	16.89	100.00
	t_r	15.31	110.35	11.81	142.49	8.64	195.52	5.13	329.14
	t_p	2.81	601.03	3.89	432.80	3.09	547.21	0.54	3122.15
	t_k	2.78	607.81	3.82	440.96	3.04	556.19	0.53	3161.31
4	t_m	16.63	100.00	16.91	100.00	16.33	100.00	16.32	100.00
	t_r	15.51	107.25	11.45	147.71	8.53	191.44	5.07	322.12
	t_p	2.84	585.91	3.78	447.68	3.06	534.53	0.53	3068.87
	t_k	2.80	593.46	3.71	456.09	3.01	543.29	0.52	3109.73
5	t_m	16.90	100.00	17.19	100.00	16.91	100.00	17.00	100.00
	t_r	15.83	106.78	11.88	144.67	8.66	195.36	5.08	334.55
	t_p	2.90	582.24	3.92	438.89	3.10	545.48	0.54	3137.87
	t_k	2.87	587.94	3.84	447.15	3.05	554.41	0.54	3173.36
6	t_m	16.77	100.00	17.21	100.00	16.82	100.00	16.91	100.00
	t_r	14.87	112.80	12.10	142.26	8.26	203.75	4.89	345.72
	t_p	2.73	614.73	3.99	431.04	2.96	568.69	0.52	3265.83
	t_k	2.70	622.05	3.92	439.14	2.91	578.01	0.51	3304.71
7	t_m	16.55	100.00	17.27	100.00	16.55	100.00	16.62	100.00
	t_r	15.68	105.56	11.72	147.27	8.73	189.48	5.16	322.03
	t_p	2.88	573.68	3.87	446.69	3.14	527.44	0.55	3015.22
	t_k	2.85	579.74	3.79	455.08	3.09	536.09	0.55	3049.68

Table 2. MSE and PRE of synthetic estimators for fixed values of $\rho_{yx} = 0.65$, $\rho_{xz} = 0.50$, and varying values of ρ_{yz} = (0.2, 0.4, 0.6, 0.8)

Domains	Estimators	ρ_{yz}							
		0.2		0.4		0.6		0.8	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	t_m	16.71	100.00	16.73	100.00	16.76	100.00	16.84	100.00
	t_r	15.30	109.18	12.20	137.10	9.09	184.47	5.92	284.35
	t_p	2.52	664.34	3.37	496.38	3.17	528.66	1.90	887.38
	t_k	2.49	671.43	3.32	504.24	3.12	537.76	1.88	896.90
2	t_m	17.10	100.00	16.88	100.00	16.62	100.00	16.46	100.00
	t_r	16.29	105.00	12.89	130.92	9.49	175.12	6.15	267.45
	t_p	2.68	638.89	3.55	474.77	3.31	502.55	1.97	834.38
	t_k	2.65	645.51	3.50	482.28	3.25	511.22	1.95	843.36
3	t_m	17.00	100.00	16.98	100.00	16.94	100.00	16.89	100.00
	t_r	16.40	103.71	12.97	130.93	9.54	177.65	6.16	274.13
	t_p	2.68	634.68	3.56	476.40	3.31	511.49	1.97	858.18
	t_k	2.65	641.84	3.51	483.99	3.26	520.32	1.95	867.43
4	t_m	16.51	100.00	16.37	100.00	16.29	100.00	16.34	100.00
	t_r	15.77	104.68	12.52	130.70	9.29	175.35	6.10	267.76
	t_p	2.57	643.40	3.43	476.69	3.23	504.49	1.96	835.51
	t_k	2.54	650.59	3.38	484.29	3.17	513.19	1.93	844.49
5	t_m	17.14	100.00	17.06	100.00	16.97	100.00	16.94	100.00
	t_r	15.67	109.38	12.54	136.07	9.37	181.00	6.16	274.82
	t_p	2.58	664.55	3.46	492.73	3.27	519.29	1.97	857.78
	t_k	2.55	671.57	3.41	500.53	3.21	528.23	1.95	867.00
6	t_m	17.12	100.00	17.03	100.00	16.91	100.00	16.83	100.00
	t_r	15.73	108.84	12.38	137.54	9.08	186.22	5.88	286.33
	t_p	2.58	663.08	3.41	499.40	3.16	535.03	1.88	893.50
	t_k	2.55	670.10	3.36	507.30	3.11	544.24	1.86	903.12
7	t_m	16.69	100.00	16.64	100.00	16.58	100.00	16.57	100.00
	t_r	16.22	102.89	12.91	128.89	9.56	173.42	6.22	266.39
	t_p	2.66	626.87	3.56	467.46	3.34	497.03	2.00	828.55
	t_k	2.63	633.34	3.50	474.84	3.28	505.60	1.98	837.47

Table 3. MSE and PRE of synthetic estimators for fixed values of $\rho_{yx} = 0.62$, $\rho_{yz} = 0.48$, and varying values of ρ_{xz} = (0.2, 0.4, 0.6, 0.8)

Domains	Estimators	ρ_{xz}							
		0.2		0.4		0.6		0.8	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	t_m	16.78	100.00	16.72	100.00	16.73	100.00	16.83	100.00
	t_r	6.84	245.34	9.95	168.11	13.03	128.42	14.50	116.03
	t_p	3.08	544.17	3.73	448.43	3.78	442.97	2.82	597.58
	t_k	3.02	555.47	3.65	457.64	3.70	451.68	2.31	727.62
2	t_m	16.70	100.00	16.71	100.00	16.84	100.00	17.03	100.00
	t_r	7.18	232.51	10.42	160.40	13.74	122.57	15.48	110.00
	t_p	3.23	516.57	3.90	428.62	3.98	423.12	2.81	605.98
	t_k	3.17	527.30	3.82	437.42	3.91	431.34	2.75	619.40
3	t_m	16.95	100.00	16.96	100.00	16.98	100.00	16.97	100.00
	t_r	7.22	234.80	10.48	161.87	13.83	122.79	15.54	109.18
	t_p	3.24	523.25	3.91	433.78	4.00	424.66	2.96	572.60
	t_k	3.17	534.13	3.83	442.68	3.92	433.18	2.61	650.87
4	t_m	16.30	100.00	16.31	100.00	16.35	100.00	16.42	100.00
	t_r	7.02	232.17	10.15	160.66	13.36	122.43	15.06	109.04
	t_p	3.15	516.99	3.79	430.50	3.86	424.21	2.93	559.65
	t_k	3.09	527.72	3.71	439.33	3.78	432.79	2.48	661.51
5	t_m	17.01	100.00	17.00	100.00	17.05	100.00	17.13	100.00
	t_r	7.06	240.82	10.23	166.15	13.39	127.34	14.96	114.51
	t_p	3.18	534.80	3.83	443.71	3.88	439.02	2.60	659.15
	t_k	3.12	545.91	3.75	452.82	3.81	447.54	2.40	714.98
6	t_m	16.96	100.00	16.95	100.00	17.01	100.00	17.11	100.00
	t_r	6.88	246.42	9.99	169.75	13.19	129.00	14.87	115.05
	t_p	3.09	548.34	3.73	454.11	3.81	446.13	2.83	604.05
	t_k	3.03	559.74	3.66	463.43	3.74	454.86	2.52	677.46
7	t_m	16.60	100.00	16.60	100.00	16.63	100.00	16.68	100.00
	t_r	7.23	229.70	10.48	158.46	13.77	120.79	15.44	108.06
	t_p	3.25	510.13	3.92	423.19	3.98	417.72	2.85	585.75
	t_k	3.19	520.74	3.84	431.88	3.91	425.79	2.34	713.31

high correlations between variables (y, x). The constantly reduced MSE across different ρ_{yx} values emphasise the stability of SSPRE in minimizing prediction errors, while its superior PRE reveals a stronger explanatory power than the other estimators. Thus, the SSPRE is the best option, especially in circumstances where the correlations between the auxiliary variables and the study variable are moderate to high, delivering consistent and reliable estimates across various correlation structures.

The findings in Table 2 and Table 3 for various correlation coefficient combinations provide a similar interpretation. In both Tables, the SSPRE t_k continues to outperform in terms of minimizing MSE and maximizing PRE, independent of the various correlation structures. These constant trends across several scenarios demonstrate the stability and flexibility of SSPRE t_k when used to various correlation settings. The results also show that the estimator can retain its efficiency and forecast accuracy even when the inter-relationships between the variables change, extending its value to a wider variety of applications. This suggests that SSPRE t_k may be securely implemented in a variety of data contexts where correlations between auxiliary variables and study variable display varying strengths and patterns, giving it an adaptable option for reliable estimations.

6. REAL DATA APPLICATION

Uttar Pradesh (UP), like the majority of Indian states, is administratively split into districts, largely for tax collection and efficient handling of governmental functions. Each district is further divided into tehsils, which function as intermediate administrative divisions. These tehsils are subsequently split into blocks, which are the smallest administrative divisions with separate administration and development responsibilities. In the current study, blocks are viewed as small domains that reflect localised locations where government plans, resources and policies are implemented and monitored.

In this section, we estimate the paddy crop production for the 2022-2023 agricultural season over nine unique blocks of Kanpur district in Uttar Pradesh, treating each block as a discrete domain. The primary variable y depends on total production (in tonnes) of paddy during 2022-2023, while the auxiliary variables x and z correspond to the total production of paddy during 2021-2022 and 2020-2021, respectively. The data on the production of paddy for these agricultural seasons are obtained from the agricultural department, Directorate of Agricultural Statistics and Crop Insurance, Government of Uttar Pradesh, India. The production of paddy is estimated from crop-cutting experiments (CCEs) under general crop estimation survey (GCES). Such surveys are conducted twice a year to cover different types of crops. The CCE is a method used by governments and agricultural bodies to estimate the yield of a crop or region. The CCEs are an integral part of the implementation of Pradhan Mantri Fasal Bima Yojana. It helps in analysing the overall yield of any village and helps us in estimating the yield of any region.

Table 4. Total production of paddy crop in blocks of Kanpur district of Uttar Pradesh for agricultural seasons 2020-21, 2021-22, and 2022-23

S. No.	**Blocks of Kanpur District (a)**	**Number of villages in Blocks (N_a)**	**Total production (in tonne) of paddy in 2022-23 (Y_a)**	**Total production (in tonne) of paddy in 2021-22 (X_a)**	**Total production (in tonne) of paddy 2020-21 (Z_a)**
1	Bidhnu	44	13041	13654	15118
2	Bilhaur	24	7044	6966	8612
3	Chaubepur	25	8220	7919	9075
4	Ghatampur	14	3124	2952	3141
5	Kakwan	22	10637	10595	11601
6	Kalyanpur	16	5632	5454	5594
7	Patara	27	6144	6132	6345
8	Sarsaul	27	4298	4053	4377
9	Shivrajpur	49	17137	17397	18719

Table 4 displays the total production of paddy in different years on the blocks, while, Table 5 summarises statistical parameters for each block. A ranked set sample of size $n=15$ is selected from each domain with set size 3 and number of cycles 5. The domain parameters reported in Table 5 are taken to compute the MSE and PRE of the proposed synthetic estimators. The PRE of the synthetic estimators is computed by employing the following formula:

$$PRE = \frac{MSE(t_m)}{MSE(t_*)} \times 100$$

Table 6 presents the MSE and PRE of synthetic estimators based on real data. These findings clearly illustrate the superiority of the SSPRE estimator t_k,

Table 5. Features of the population for different domains

Domains	N_a	$\bar{Y}_a$	$\bar{X}_a$	$\bar{Z}_a$	S_{y_a}	S_{x_a}	S_{z_a}	$\rho_{y_a x_a}$	$\rho_{y_a z_a}$	$\rho_{x_a z_a}$
1	44	296.39	310.32	343.59	295.22	337.12	393.92	0.98	0.96	0.98
2	24	293.50	290.25	358.83	298.33	309.29	411.78	0.96	0.96	0.98
3	25	328.80	316.76	363.00	242.59	249.85	263.98	0.97	0.97	0.95
4	14	223.14	210.86	224.36	195.34	190.57	196.56	0.98	0.98	0.99
5	22	483.50	481.59	527.32	346.62	351.30	391.26	0.97	0.97	0.94
6	16	352.00	340.88	349.62	220.49	198.19	206.20	0.98	0.97	0.98
7	27	227.56	227.11	235.00	165.46	166.31	164.53	0.98	0.98	0.99
8	27	162.89	150.11	162.11	109.40	104.14	115.45	0.98	0.98	0.98
9	49	349.73	355.04	382.02	261.11	257.47	264.96	0.97	0.98	0.99

Table 6. Synthetic estimators' MSE and PRE for a simulated normal population

Domains	Estimators	MSE	PRE
1	t_m	4449.19	100.00
	t_r	3645.20	122.06
	t_p	119.09	3736.05
	t_k	116.10	3832.33
2	t_m	4501.13	100.00
	t_r	3574.55	125.92
	t_p	116.78	3854.36
	t_k	113.85	3953.70
3	t_m	5010.19	100.00
	t_r	4486.09	111.68
	t_p	146.56	3418.52
	t_k	142.88	3506.62
4	t_m	10920.24	100.00
	t_r	2066.19	528.52
	t_p	67.50	16177.60
	t_k	65.81	16594.53
5	t_m	36634.11	100.00
	t_r	9700.58	377.65
	t_p	316.92	11559.54
	t_k	308.95	11857.45

Domains	Estimators	MSE	PRE
6	t_m	6701.95	100.00
	t_r	5141.50	130.35
	t_p	167.97	3989.92
	t_k	163.75	4092.74
7	t_m	10226.65	100.00
	t_r	2148.72	475.94
	t_p	70.20	14568.22
	t_k	68.43	14943.66
8	t_m	24287.42	100.00
	t_r	1101.00	2205.94
	t_p	35.97	67522.16
	t_k	35.07	69262.32
9	t_m	6489.34	100.00
	t_r	5075.54	127.86
	t_p	165.82	3913.55
	t_k	161.65	4014.41

over the other estimators such as the SME t_m, SRE t_r, and SPRE t_p across all domains. The SSPRE t_k consistently shows reduced MSE and higher PRE, suggesting better accuracy and precision in estimating population parameters. This implies that the SSPRE t_k gives more accurate findings by minimising MSE and maximizing PRE than the adapted synthetic estimators.

7. CONCLUSIONS

This work proposed the synthetic ratio estimator for small area estimation under RSS using bivariate auxiliary data. The incorporation of RSS into the SAE framework proven to be a helpful strategy, increasing the precision of estimates in small areas where standard direct sampling approaches frequently fail owing to insufficient sample numbers. The use of auxiliary information in combination with RSS enabled more effective data utilisation, resulting in higher estimation accuracy. An extensive simulation study has been conducted to evaluate the theoretical implications using artificially generated normal population in various settings. The simulation study examined different levels of variability and correlation structures on the effectiveness of the estimators. The simulation study consistently demonstrated that the suggested synthetic estimator t_k outperformed the SME t_m, the SRE t_r, and the SPRE t_p. across various situations. Furthermore, the real-life applicability of the adapted and suggested estimators has been demonstrated using real crop production data, which revealed comparable performance benefits. The real data analysis revealed that the suggested synthetic estimator t_k consistently generated more precise and trustworthy estimates than the adapted synthetic estimators, making it more robust for practical usage in small area estimation issues, notably in agricultural and environmental investigations.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the anonymous reviewers for providing suggestions and comments which led to the significant improvement in the manuscript.

REFERENCES

Al-Omari, A.I. and Bouza, C. (2015). Ratio estimators of the population mean with missing values using ranked set sampling. Environmetrics, **26**(2), 67-76.

Al-Saleh, M.F. and Samawi, H. (2000). On the efficiency of Monte Carlo methods using steady state ranked simulated samples. Communications in Statistics-Simulation and Computation, **29**, 941-954.

Ahmed, S., Albalawi, O. and Shabbir, J. (2024). On indirect estimation of small area parameters under ranked set sampling. Alexandria Engineering Journal, **107**, 690-697.

Al-Omari, A.I., Jemain, A.A. and Ibrahim, K. (2009). New ratio estimators of the mean using simple random sampling and ranked set sampling methods. Investigacin Operacional, **30**(2), 97-108.

Ashutosh, A., Stefan, M., Rai, P.K., Emam, W., Iftikhar, S. and Anas, M.M. (2024). Simulation analysis of non-respondent information in context of small domain. Heliyon, **10**(6), e26897.

Bhushan, S. and Kumar, A. (2022a). On the quest of optimal class of estimators using ranked set sampling. Scientia Iranica. doi: 10.24200/sci.2022.58063.5547

Bhushan, S. and Kumar, A. (2022b). New efficient logarithmic estimators using multi auxiliary information under ranked set sampling. Concurrency and Computation: Practice and Experience, **34**(27), e7337.

Bhushan, S. and Kumar, A. (2023). Imputation of missing data using multi auxiliary in formation under ranked set sampling. Communications in Statistics-Simulation and Computation, 1-22.

Bhushan, S. and Kumar, A. (2024a). Ranked set sampling imputation methods in presence of correlated measurement errors. Communications in Statistics-Theory and Methods, 1-22.

Bhushan, S. and Kumar, A. (2024b). Novel logarithmic imputation methods under ranked set sampling. Measurement: Interdisciplinary Research and Perspectives, **22**(3), 235-257.

Chen, Z., Bai, Z. and Sinha, B. (2004). Ranked set sampling: Theory and Applications. Springer Verlag. New York.

Cochran, W. G. (1977). Sampling techniques. Johan Wiley & Sons Inc.

Dell, T.R. and Clutter, J.L. (1972). Ranked set sampling theory with order statistics background. Biometrics, **28**, 545-555.

Khare, B.B. and Ashutosh (2018). Simulation study of the generalized synthetic estimator for domain mean in the sample survey. International Journal of Tomography & Statistics, **31**(3), 87-100.

Kocyigit, E.G. (2023). A novel sub-type mean estimator for ranked set sampling with dual auxiliary variables. Journal of New Theory, **44**, 79-86.

Kumar, A., Bhushan, S., Emam, W., Tashkandy, Y. and Khan, M.J.S. (2024a). Novel logarithmic imputation procedures using multi auxiliary information under ranked set sampling. Scientific Reports, **14**(1), 18027.

Kumar, A., Bhushan, S., Pokhrel, R. and Emam, W. (2024b). Small area estimation using design based direct and synthetic logarithmic estimators. Ain Shams Engineering Journal, 102836.

Kumar, A., Bhushan, S., Pokhrel, R., Emam, W., Tashkandy, Y. and Khan, M.J.S. (2024c). Enhanced direct and synthetic estimators for domain mean with simulation and applications. Heliyon, 10(14), e33839.https://doi.org/10.1016/j.heliyon.2024.e33839

McIntyre, G.A. (1952). A method of unbiased selective sampling using ranked set. Aust. J. Agr. Res. **3**, 385-390.

Mahdizadeh, M. and Zamanzade, E. (2022). Using a rank-based design in estimating prevalence of breast cancer. Soft Computing, **26**(7), 3161-3170.

Pandey, K.K. and Tikkiwal, G.C. (2010). Generalized class of synthetic estimators for small area under systematic sampling design. Statistics in Transition- new series, Poland, **11**(1), 75-89.

Patil, G.P, Sinha, A.K. and Taillie, C. (1994). Ranked set sampling for multiple characteristics. International Journal of Ecology and Environmental Sciences, **20**, 357-373.

Punia, P., Raj, A. and Kumar, P. (2024). An enhanced Beluga Whale optimization algorithm for engineering optimization problems. Journal of Systems Science and Systems Engineering, 1-38.

Rai, P.K. and Pandey, K.K. (2013). Synthetic estimators using auxiliary information in small domains. Statistics in Transition new series, **14**(1), 31-44.

Raj, A., Punia, P. and Kumar, P. (2024). A novel hybrid pelican-particle swarm optimization algorithm (HPPSO) for global optimization problem. International Journal of System Assurance Engineering and Management, 1-16.

Rao, J. N. K. and Molina, I. (2015). Small area estimation. John Wiley & Sons.

Rehman, S.A. and Shabbir, J. (2022). An efficient class of estimators for finite population mean in the presence of non-response under ranked set sampling (RSS). Plos one, **17**(12), e0277232.

Samawi, H.M. and Muttlak, H.A.(1996).Estimation of ratio using ranked set sampling. Biom. J., **38**, 753-764.

Searls, D.T. (1964). The utilization of a known coefficient of variation in the estimation procedure. Journal of the American Statistical Association, **59**(308), 1225-1226.

Stokes, S.L. (1977). Ranked set sampling with concomitant variables. Communications in Statistics-Theory and Methods, **A**(6), 1207-1211.

Stokes, S.L. and Sager, T.W. (1988). Characterization of a ranked-set sample with ap plication to estimating distribution functions. Journal of the American Statistical Association, **83**, 374-381.

Tikkiwal G.C. and Ghiya. A. (2000). A generalized class of synthetic estimators with application to crop acreage estimation for small domains. Biometrical Journal, **42**, 865-876.

Tikkiwal, G.C., Rai, P.K. and Ghiya, A. (2013). On the performance of generalized regression estimator for small domains. Communications in Statistics-Simulation and Computation, **42**(4), 891-909.

Zamanzade, E. and Al-Omari, A.I. (2016). New ranked set sampling for estimating the population mean and variance. Hacettepe Journal of Mathematics and Statistics, **45**(6), 1891-1905.