Ain Shams Engineering Journal 15 (2024) 102836

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: https://www.sciencedirect.com

Full Length Article

Small area estimation using design based direct and synthetic logarithmic estimators

Anoop Kumar [a], Shashi Bhushan [b], Rohini Pokhrel [c], Walid Emam [d,*]

[a] Department of Statistics, Central University of Haryana, Mahendergarh, 123031, India
[b] Department of Statistics, University of Lucknow, Lucknow, 226007, India
[c] Department of Mathematics & Statistics, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, 226017, India
[d] Department of Statistics and Operations Research, Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

ARTICLE INFO

MSC:
62D05

Keywords:
Small area estimation
Direct and synthetic estimators
Simple random sampling

ABSTRACT

In this article, we propose some direct and synthetic logarithmic estimators for estimating the domain mean of small area based on a simple random sampling design. The mean square error expressions of the proposed direct and synthetic estimators are obtained to first order approximation. The efficiency conditions are obtained under which the proposed direct and synthetic estimators dominate their conventional aspirants. The performances of the suggested direct and synthetic logarithmic estimators are examined by a comprehensive simulation study carried out on some artificially drawn symmetric and asymmetric populations. Furthermore, a real data application of the suggested methods is also provided as a case study using the paddy crop acreage data for small domains, where small domains are the revenue inspector circles (RIC) in Mohanlalganj tehsil, Uttar Pradesh, India.

1. Introduction

The role of small area estimation (SAE) is very prominent in the sample survey because of the growing demand for trustworthy small area statistics from different sectors like business, agriculture, economics, health, etc. The goal of sample surveys is to provide relatively accurate direct estimators for the properties of the entire population as well as for a number of subpopulations or domains. A small area is often a subdivision of the population which has small information from the sample surveys. These subsets can refer to a small geographical region, like a municipal council, census division, block, tehsil, gram panchayat, or a demographic group, such as a collection of individuals who share a specific age, sex, or race, inside a larger geographic area, or a combination of the two.

Survey agencies are forced to produce small area estimates from current sample surveys due to the increasing demand for small area decision-making. For providing domain-specific accurate direct estimates for these small areas, small sizes in small areas are frequently very small or perhaps non-existent. Direct estimators use only the information collected from the target small area itself. They are typically based on sample surveys specifically designed to provide estimates for that area. Direct estimators are straightforward to implement and interpret. They consider the real data from the target area, providing estimates that are directly observed.

Direct estimators may suffer from small sample sizes, leading to imprecise estimates, especially for areas with sparse population or rare characteristics. If a reliable direct estimator is not available for a large domain, covering several small domains then a synthetic estimator is used. A synthetic estimator under the assumption of small area has the same characteristic as the large domain [8]. Also, if the sample sizes of the small domain are comparatively small, the synthetic estimator represses the direct estimator, but if the sample size increases, the direct estimator represses the synthetic estimator [14].

Auxiliary variables are instrumental in elevating small area estimates by offering auxiliary information related to the study variable, thereby improving precision and accuracy. By leveraging auxiliary data, researchers can mitigate sampling variability, reduce bias, and model intricate relationships between variables, particularly crucial when direct survey data is scarce for smaller geographic regions. In real-life problems, the abundance of auxiliary variables poses a challenge in selecting the most appropriate ones. To navigate this, it is crucial to follow a systematic approach such as define objective, assess data availability,

* Corresponding author.
E-mail address: wemam.c@ksu.edu.sa (W. Emam).

https://doi.org/10.1016/j.asej.2024.102836
Received 30 November 2023; Received in revised form 29 March 2024; Accepted 20 April 2024
Available online 28 May 2024

consider relevance, evaluate predictive power, account for redundancy, practicality, etc. Following these steps ensures the selection of auxiliary variables that enhance the accuracy of small area estimation in real-world scenarios.

Panse et al. [13] tried to estimate the crop production at the block level utilizing a two phase sampling technique, which was the first stage advancement in SAE for crop-production. Due to physical limitations, this method, however, was not successful. An effort was undertaken to create crop-production estimates at the block level utilizing farmer estimates by Sud et al. [18] within the context of sample design based on the general crop estimating survey technique. These were direct estimates that are based on the usual sample survey methods. Srivastava et al. [17] proposed a synthetic method with an application to crop-estimation at block level using wheat and paddy crops data of Haryana state during 1987-88. Tikkiwal and Ghiya [19] proposed a generalised class of synthetic estimators to estimate crop acreage for small areas, where small domains were RIC in Jodhpur tehsil, Rajasthan. In India, a National Agricultural Insurance Scheme (NAIS) was introduced in 1999-2000 to replace the Comprehensive Crop Insurance Scheme (CCIS), and Gram Panchayat (GP) level was chosen as the area unit in place of blocks. As a result, insurance firms urgently need crop production at the GP level to finalise premiums. Sharma et al. [15] proposed a different strategy for scaling down crop yields from the block level to the GP level by generating correction factors based on the data on crop yields on specific fields obtained through farmer inquiries. Pandey and Tikkiwal [12] proposed a generalized class of synthetic estimators for small areas using systematic sampling. Khare et al. [10] proposed a modified direct regression estimator for the domain mean using auxiliary information when the domain mean of the auxiliary character is known and unknown. Bhushan et al. [5] suggested improved direct estimators for domain mean with some real data applications. The aforementioned studies discussed the direct and synthetic estimators separately.

In this article, our endeavour is different from the aforementioned studies. We propose some direct and synthetic logarithmic estimators together to estimate the paddy crop acreage for small domains using SRS, where small domains are RIC in Mohanlalganj tehsil, Uttar Pradesh, India.

Section 2 provides a brief about the set up of the problem and puts forth the terminologies utilized. In Section 3, the review of the existing literature and prominent works for the direct and synthetic estimators are considered. In Section 4, we propose direct and synthetic logarithmic estimators for estimating the domain mean utilizing auxiliary character and a comparative study is performed theoretically. The performance of the proposed estimators has been evaluated in Section 5 with the help of simulation study. A real data application is also presented in Section 6 using paddy crop acreage data set for small domains. In Section 7, the conclusion of this study is presented.

2. Problem formulation and terminologies

Let us consider that a finite population $\aleph = (\aleph_1, \aleph_2, ..., \aleph_N)$ is divided into non overlapping 'A' small areas, i.e., domains $\aleph_a$ of size N_a for which estimates are needed. The domains might be various and could constitute small areas of a sampled population, such as a district, tehsil, or other state-level subdivision, depending on the situation. Let y and x be the study and auxiliary variables, respectively. A simple random sample $\mathbf{s} = (s_1, s_2, ..., s_n)$ of size n is chosen without replacement such that n_a, $a = 1, 2, ..., A$ units in the sample $\mathbf{s}$ come from the small area 'A'. As a result, $\sum_{a=1}^{A} N_a = N$ and $\sum_{a=1}^{A} n_a = n$. The notations for population and sample means of study variable y and auxiliary variable x are as follows:
$(\bar{Y}, \bar{y})$: the population and sample means of the study characteristic y rely on (N, n) observations, respectively;
$(\bar{Y}_a, \bar{y}_a)$: the population and sample means of characteristic y for a small domain rely on (N_a, n_a) observations, respectively;
$(\bar{X}, \bar{x})$: population and sample means of the auxiliary characteristic x rely on (N, n) observations, respectively;
$(\bar{X}_a, \bar{x}_a)$: the population and sample means of characteristic x for a small domain rely on (N_a, n_a) observations, respectively;
(S_x, S_y): the population standard deviation of variables (x, y), respectively.
(S_{x_a}, S_{y_a}): the population standard deviation of variables (x, y) in domain a, respectively.
(C_x, C_y): the population variation coefficient of variables (x, y), respectively.
(C_{x_a}, C_{y_a}): the population variation coefficient of variables (x, y) in domain a, respectively.

To obtain the properties of the direct estimators, we take the notations as follows:
$\bar{y}_a = \bar{Y}_a(1+\epsilon_1)$, $\bar{x}_a = \bar{X}_a(1+\epsilon_2)$ such that $E(\epsilon_1) = 0$, $E(\epsilon_2) = 0$, where $|\epsilon_i| < 1; i = 1, 2$,
$E(\epsilon_1^2) = f_a C_{y_a}^2$, $E(\epsilon_2^2) = f_a C_{x_a}^2$, and $E(\epsilon_1\epsilon_2) = f_a C_{y_a x_a}$, where $f_a = (N_a - n_a)/N_a n_a$, $S_{x_a}^2 = (N_a - 1)^{-1}\sum_{i=1}^{N_a}(X_{ai} - \bar{X}_a)^2$, $S_{y_a}^2 = (N_a - 1)^{-1}\sum_{i=1}^{N_a}(Y_{ai} - \bar{Y}_a)^2$, $C_{x_a} = S_{x_a}/\bar{X}_a$, $C_{y_a} = S_{y_a}/\bar{Y}_a$, $S_{x_a y_a} = (N_a - 1)^{-1}\sum_{i=1}^{N_a}(X_{ai} - \bar{X}_a)(Y_{ai} - \bar{Y}_a)$, and $C_{x_a y_a} = \rho_{y_a x_a} C_{y_a} C_{x_a}$, respectively, such that X_{ai} and Y_{ai}, $a = 1, 2, ..., A$ and $i = 1, 2, ..., N_a$, denote the i^{th} observation of small domain a of the population for the characteristic x and y, respectively.

Similarly, to obtain the properties of the synthetic estimators, we assume that $\bar{y} = \bar{Y}(1+\epsilon_3)$, $\bar{x} = \bar{X}(1+\epsilon_4)$, such that $E(\epsilon_3) = 0$, $E(\epsilon_4) = 0$, where $|\epsilon_i| < 1; i = 3, 4$, $E(\epsilon_3^2) = fC_y^2$, $E(\epsilon_4^2) = fC_x^2$, and $E(\epsilon_3\epsilon_4) = fC_{yx}$, where $f = (N-n)/Nn$, $S_x{}^2 = (N-1)^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$, $S_y{}^2 = (N-1)^{-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$, $C_x = S_x/\bar{X}$, $C_y = S_y/\bar{Y}$, $S_{yx} = (N-1)^{-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})$, and $C_{yx} = \rho_{yx} C_y C_x = S_{yx}/\bar{Y}\bar{X}$.

3. Existing estimators for domain mean

The direct and synthetic usual mean estimators are given, respectively, by

$$t_{m,a}^d = \bar{y}_a$$

$$t_{m,a}^s = \bar{y}$$

where the superscripts "d" and "s" in the estimators stand for "direct" and "synthetic" throughout the article.
The MSE of the direct and synthetic usual mean estimator are given, respectively, by

$$MSE(t_{m,a}^d) = f_a \bar{Y}_a^2 C_{y_a}^2$$

$$MSE(t_{m,a}^s) = (\bar{Y} - \bar{Y}_a)^2 + f\bar{Y}^2 C_y^2$$

The direct and synthetic ratio estimators are given, respectively, by

$$t_{r,a}^d = \bar{y}_a\left(\frac{\bar{X}_a}{\bar{x}_a}\right)$$

$$t_{r,a}^s = \bar{y}\left(\frac{\bar{X}_a}{\bar{x}}\right)$$

The MSE of the direct and synthetic estimators are given, respectively, by

$$MSE(t_{r,a}^d) = \bar{Y}_a^2 f_a (C_{y_a}^2 + C_{x_a}^2 - 2C_{y_a x_a})$$

$$MSE(t_{r,a}^s) = \left(\frac{\bar{Y}}{\bar{X}}\bar{X}_a - \bar{Y}_a\right)^2 + \frac{\bar{Y}}{\bar{X}}\bar{X}_a f\left\{\frac{\bar{Y}}{\bar{X}}\bar{X}_a\left(3C_x^2 + C_y^2 - 4C_{yx}\right) - 2\bar{Y}_a(C_x^2 - C_{yx})\right\}$$

Under the synthetic assumption $\bar{Y}/\bar{Y}_a = \bar{X}/\bar{X}_a$, the MSE of the synthetic estimator $t_{r,a}^s$ are reduced to

$$MSE(t^s_{r,a}) = \bar{Y}_a^2 f(C_x^2 + C_y^2 - 2C_{yx})$$

The direct and synthetic generalized estimators suggested by Tikkiwal and Ghiya [19] are given by

$$t^d_{tg,a} = \bar{y}_a \left(\frac{\bar{x}_a}{\bar{X}_a}\right)^{\alpha}$$

$$t^s_{tg,a} = \bar{y} \left(\frac{\bar{x}}{\bar{X}_a}\right)^{\alpha_a}$$

The minimum MSE of the direct estimator at the optimum value of $\alpha_{(opt)} = -C_{y_a x_a}/C^2_{x_a}$ and the minimum MSE of the synthetic estimator under the synthetic assumption $(\bar{Y}_a/\bar{Y}) = (\bar{X}/\bar{X}_a)^{\alpha_a}$, and at the optimum value of $\alpha_{a(opt)} = -C_{yx}/C_x^2$ as

$$minMSE(t^d_{s,a}) = \bar{Y}_a^2 f_a \left(C^2_{y_a} - \frac{C^2_{y_a x_a}}{C^2_{x_a}}\right)$$

$$minMSE(t^s_{tg,a}) = \bar{Y}_a^2 f \left(C_y^2 - \frac{C^2_{yx}}{C^2_x}\right)$$

Following Bahl and Tuteja [1], the ratio exponential type direct and synthetic estimators are given, respectively, by

$$t^d_{bt,a} = \bar{y}_a \exp\left(\frac{\bar{X}_a - \bar{x}_a}{\bar{X}_a + \bar{x}_a}\right)$$

$$t^s_{bt,a} = \bar{y} \exp\left(\frac{\bar{X}_a - \bar{x}}{\bar{X}_a + \bar{x}}\right)$$

The MSE of the above estimators are given, respectively, by

$$MSE(t^d_{bt,a}) = \bar{Y}_a^2 f_a \left(C^2_{y_a} + \frac{C^2_{x_a}}{4} - C_{y_a x_a}\right)$$

$$MSE(t^s_{bt,a}) = \left\{ \begin{array}{l} (K_{1a}\bar{Y} - \bar{Y}_a)^2 + \bar{Y}^2 f(K^2_{1a}C^2_y + K^2_{2a}C^2_x - 2K_{1a}K_{2a}C_{yx}) \\ +2(\bar{Y}K_{1a} - \bar{Y}_a) f \bar{Y} \left(K_{3a}C^2_x - K_{2a}C_{yx}\right) \end{array} \right\}$$

where $K_{1a} = 1 + A_a + A_a^2/2 + A_a^3/6$, $K_{2a} = \left(2A_a + 2A_a^2 + A_a^3 \bar{X}\bar{X}_a\right) / \left(\bar{X}_a^2 - \bar{X}^2\right)$
$K_{3a} = \left[2\{(\bar{X}_a - \bar{X})A_a + (2\bar{X}_a - \bar{X})A_a^2 + A_a^3\}\bar{X}^2\bar{X}_a/(\bar{X}_a^2 - \bar{X}^2)^2\right]$, and $A_a = \left(\bar{X} - \bar{X}_a\right) / \left(\bar{X} + \bar{X}_a\right)$.

4. Proposed direct and synthetic estimators

The objectives of this article are:

(i). to obtain efficient estimates of the domain mean $\bar{Y}_a$ via both direct and synthetic estimation approach,
(ii). to effectively utilize the information on the auxiliary variable to obtain efficient estimates.

Under the abovementioned objectives, we propose the following direct and synthetic logarithmic estimators for $\bar{Y}_a$ in SRS as

$$t^d_{bk,a} = \alpha_1 \bar{y}_a \left\{1 + \beta_1 \log\left(\frac{\bar{x}_a}{\bar{X}_a}\right)\right\}$$

$$t^s_{bk,a} = \alpha_2 \bar{y} \left\{1 + \beta_2 \log\left(\frac{\bar{x}}{\bar{X}_a}\right)\right\}$$

where $\alpha_j,\ j = 1,2$ and β_j are suitably chosen constants. The mathematical expressions of the MSE and the minimum MSE for the proposed direct and synthetic estimators are given in the following theorems.

Theorem 4.1. *The MSE and minimum MSE of the suggested direct estimator $t^d_{bk,a}$ up to 1^{st} order approximation are, respectively, given below.*

$$MSE(t^d_{bk,a}) = \bar{Y}_a^2 \left[\begin{array}{l} 1 + \alpha_1^2 \left\{1 + f_a C^2_{y_a} + \beta_1(\beta_1 - 1) f_a C^2_{x_a} + 4\beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \\ -2\alpha_1 \left(1 + \beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a} - \frac{\beta_1}{2} f_a C^2_{x_a}\right) \end{array} \right] \tag{1}$$

$$minMSE(t^d_{bk,a}) = \bar{Y}_a^2 \left(1 - \frac{M_1^2}{L_1}\right) \tag{2}$$

Proof. To determine the MSE and minimum MSE of the suggested direct estimator $t^d_{bk,a}$, the notations defined in Section 2 are used to express the estimator $t^d_{bk,a}$ as

$$t^d_{bk,a} = \alpha_1 \bar{y}_a \left\{1 + \beta_1 \log\left(\frac{\bar{x}_a}{\bar{X}_a}\right)\right\}$$

$$= \alpha_1 \bar{Y}_a (1 + \epsilon_1) \left[1 + \beta_1 \log\left\{\frac{\bar{X}_a(1 + \epsilon_2)}{\bar{X}_a}\right\}\right]$$

$$= \alpha_1 \bar{Y}_a (1 + \epsilon_1) \left\{1 + \beta_1 \log\left(1 + \epsilon_2\right)\right\}$$

Using Taylor series expansion, multiplying and neglecting the error terms having power more than two, we get

$$t^d_{bk,a} \approx \alpha_1 \bar{Y}_a (1 + \epsilon_1) \left\{1 + \beta_1 \left(\epsilon_2 - \frac{\epsilon_2^2}{2} + \ldots\right)\right\}$$

After simplifying and subtracting $\bar{Y}_a$ on both sides of the above equation, we get

$$t^d_{bk,a} - \bar{Y}_a \approx \bar{Y}_a \left\{\alpha_1 \left(1 + \epsilon_1 + \beta_1\epsilon_2 - \frac{\beta_1}{2}\epsilon_2^2 + \beta_1\epsilon_1\epsilon_2\right) - 1\right\} \tag{3}$$

Squaring and taking expectation on both sides of (3), we get

$$MSE(t^d_{bk,a}) \approx \bar{Y}_a^2 \left[\begin{array}{l} 1 + \alpha_1^2 \left\{1 + f_a C^2_{y_a} + \beta_1(\beta_1 - 1) f_a C^2_{x_a} + 4\beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \\ -2\alpha_1 \left(1 + \beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a} - \frac{\beta_1}{2} f_a C^2_{x_a}\right) \end{array} \right]$$

$$\approx \bar{Y}_a^2 \left(1 + \alpha_1^2 L_1 - 2\alpha_1 M_1\right) \tag{4}$$

where $L_1 = 1 + f_a C^2_{y_a} + \beta_1(\beta_1 - 1) f_a C^2_{x_a} + 4\beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a}$ and $M_1 = 1 + \beta_1 f_a \rho_{y_a x_a} C_{y_a} C_{x_a} - \frac{\beta_1}{2} f_a C^2_{x_a}$.
Differentiating (4) partially w.r.t. parameter α_1 and equating to zero, we get

$$\alpha_{1(opt)} = \frac{M_1}{L_1}$$

The minimum MSE of the suggested direct estimator $t^d_{bk,a}$ is determined by using the value of $\alpha_{1(opt)}$ in (4) as

$$minMSE(t^d_{bk,a}) \approx \bar{Y}_a^2 \left(1 - \frac{M_1^2}{L_1}\right)$$

It is important to note that the simultaneous minimization of α_1 and β_1 is not possible. The optimum value of β_1 can be obtained by putting $\alpha_1 = 1$ in the estimator $t^d_{bk,a}$ and minimizing the MSE expression of the estimator regarding β_1. The optimum value of $\beta_{1(opt)}$ is tabulated below.

$$\beta_{1(opt)} = -\rho_{y_a x_a} \left(\frac{C_{y_a}}{C_{x_a}}\right) \quad \square$$

Theorem 4.2. *The MSE and minimum MSE of the suggested synthetic estimators $t^s_{bk,a}$ up to 1^{st} order approximation are, respectively, given below.*

$$MSE(t^s_{bk,a}) = \bar{Y}_a^2 \left[\begin{array}{l} 1 + \alpha_2^2 \left\{ 1 + fC_y^2 + \beta_2^2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right)^2 fC_x^2 \right. \\ \left. -\beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) fC_x^2 + 4\beta_2 f\rho_{yx}C_yC_x \right\} \\ -2\alpha_2 \left\{ 1 - \beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) f \left(\frac{C_x^2}{2} - \rho_{yx}C_yC_x\right) \right\} \end{array} \right]$$

$$minMSE(t^s_{bk,a}) = \bar{Y}_a^2 \left(1 - \frac{M_2^2}{L_2}\right) \quad (5)$$

Proof. To determine the MSE and minimum MSE of the suggested synthetic estimator $t^s_{bk,a}$, the notations defined in Section 2 are used to express the estimator $t^s_{bk,a}$ as

$$\begin{aligned} t^s_{bk,a} &= \alpha_2 \bar{y} \left\{ 1 + \beta_2 \log\left(\frac{\bar{x}}{\bar{X}_a}\right) \right\} \\ &= \alpha_2 \bar{Y}(1+\epsilon_3) \left\{ 1 + \beta_2 \log\left(\frac{\bar{X}(1+\epsilon_4)}{\bar{X}_a}\right) \right\} \\ &= \alpha_2 \bar{Y}(1+\epsilon_3) \left\{ 1 + \beta_2 \log\left(\frac{\bar{X}}{\bar{X}_a}\right) + \beta_2 \log(1+\epsilon_4) \right\} \end{aligned}$$

Using Taylor series expansion, multiplying and neglecting the error terms having power more than two, we get

$$t^s_{bk,a} \approx \alpha_2 \bar{Y}(1+\epsilon_3) \left\{ 1 + \beta_2 B_a + \beta_2 \left(\epsilon_4 - \frac{\epsilon_4^2}{2}\right) \right\}$$

where $B_a = \log(\bar{X}/\bar{X}_a)$. Subtracting $\bar{Y}_a$ on both sides to the above equation, we get

$$t^s_{bk,a} - \bar{Y}_a \approx \alpha_2 \bar{Y} \left(1 + \beta_2 B_a + \epsilon_3 + \beta_2\epsilon_4 + \beta_2 B_a \epsilon_3 - \frac{\beta_2}{2}\epsilon_4^2 + \beta_2\epsilon_3\epsilon_4\right) - \bar{Y}_a \quad (6)$$

Squaring and taking expectation on both side to (6) and using the synthetic assumption $\bar{Y}_a = \bar{Y}(1 + \beta_2 B_a)$, we get

$$\begin{aligned} MSE(t^s_{bk,a}) &\approx \bar{Y}_a^2 \left[\begin{array}{l} 1 + \alpha_2^2 \left\{ 1 + fC_y^2 + \beta_2^2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right)^2 fC_x^2 \right. \\ \left. -\beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) fC_x^2 + 4\beta_2 f\rho_{yx}C_yC_x \right\} \\ -2\alpha_2 \left\{ 1 - \beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) f \left(\frac{C_x^2}{2} - \rho_{yx}C_yC_x\right) \right\} \end{array} \right] \\ &\approx \bar{Y}_a^2 \left(1 + \alpha_2^2 L_2 - 2\alpha_2 M_2\right) \end{aligned} \quad (7)$$

where

$$L_2 = 1 + fC_y^2 + \beta_2^2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right)^2 fC_x^2 - \beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) fC_x^2 + 4\beta_2 f\rho_{yx}C_yC_x,$$

$$M_2 = 1 - \beta_2 \left(\frac{\bar{Y}}{\bar{Y}_a}\right) f \left(\frac{C_x^2}{2} - \rho_{yx}C_yC_x\right).$$

By differentiating (7) partially w.r.t. β_2 and equating to zero, we get

$$\alpha_{2(opt)} = \frac{M_2}{L_2}$$

The minimum MSE is determined by using the value of $\alpha_{2(opt)}$ in (7) as

$$minMSE(t^s_{bk,a}) \approx \bar{Y}_a^2 \left(1 - \frac{M_2^2}{L_2}\right)$$

It is important to note that the simultaneous minimization of α_2 and β_2 is not possible. The optimum value of β_2 can be obtained by putting $\alpha_2 = 1$ in the estimator $t^s_{bk,a}$ and minimizing the MSE expression of the estimator regarding β_2. The optimum value of $\beta_{2(opt)}$ is tabulated below.

$$\beta_{2(opt)} = -\rho_{yx} \frac{\bar{Y}_a C_y}{\bar{Y} C_x} \quad \square$$

Corollary 4.1. *The suggested synthetic estimator $t^s_{bk,a}$ dominates the suggested direct estimator $t^d_{bk,a}$ iff*

$$\frac{M_2^2}{L_2} > \frac{M_1^2}{L_1} \quad (8)$$

and contrarily. Otherwise, both are equally efficient if the equality holds in (8).

Proof. To obtain (8), we compare the minimum MSE of the suggested direct and synthetic estimators given in (2) and (5), respectively. $\square$

Additionally, we will compare the minimum MSEs of the suggested and currently available direct and synthetic estimators under the following lemmas:

Lemma 4.1. *The proposed direct estimator $t^d_{bk,a}$ dominates the direct mean per unit estimator $t^d_{m,a}$, if*

$$MSE(t^d_{bk,a}) < MSE(t^d_{m,a}) \implies \frac{M_1^2}{L_1} > 1 - f_a C_{y_a}^2$$

Lemma 4.2. *The proposed synthetic estimator $t^s_{bk,a}$ dominates the synthetic mean per unit estimator $t^s_{m,a}$, if*

$$MSE(t^s_{bk,a}) < MSE(t^s_{m,a}) \implies \frac{M_2^2}{L_2} > 1 - \left(1 - \frac{\bar{Y}^2}{\bar{Y}_a^2}\right) + f\left(\frac{\bar{Y}}{\bar{Y}_a}\right)^2 C_y^2$$

Lemma 4.3. *The proposed direct estimator $t^d_{bk,a}$ dominates the direct ratio estimator $t^d_{r,a}$, if*

$$MSE(t^d_{bk,a}) < MSE(t^d_{r,a}) \implies \frac{M_1^2}{L_1} > 1 - f_a(C_{y_a}^2 + C_{x_a}^2 - 2C_{y_a x_a})$$

Lemma 4.4. *The proposed synthetic estimator $t^s_{bk,a}$ dominates the synthetic ratio estimator $t^s_{r,a}$, if*

$$MSE(t^s_{bk,a}) < MSE(t^s_{r,a}) \implies \frac{M_2^2}{L_2} > 1 - f(C_y^2 + C_x^2 - 2C_{yx})$$

Lemma 4.5. *The proposed direct estimator $t^d_{bk,a}$ dominates the Tikkiwal and Ghiya (2000) direct estimator $t^d_{tg,a}$, if*

$$MSE(t^d_{bk,a}) < MSE(t^d_{tg,a}) \implies \frac{M_1^2}{L_1} > 1 - f_a\left(C_{y_a}^2 - \frac{C_{y_a x_a}^2}{C_{x_a}^2}\right)$$

Lemma 4.6. *The proposed synthetic estimator $t^s_{bk,a}$ dominates Tikkiwal and Ghiya [19] synthetic estimator $t^s_{tg,a}$, if*

$$MSE(t^s_{bk,a}) < MSE(t^s_{tg,a}) \implies \frac{M_2^2}{L_2} > 1 - f\left(C_y^2 - \frac{C_{yx}^2}{C_x^2}\right)$$

Lemma 4.7. *The proposed direct estimator $t^d_{bk,a}$ dominates Bahl and Tuteja [1] type direct exponential ratio estimator $t^d_{bt,a}$, if*

$$MSE(t^d_{bk,a}) < MSE(t^d_{bt,a}) \implies \frac{M_1^2}{L_1} > 1 - f_a\left(C_{y_a}^2 + \frac{C_{x_a}^2}{4} - C_{y_a x_a}\right)$$

Lemma 4.8. *The proposed synthetic estimator $t^s_{bk,a}$ dominates Bahl and Tuteja [1] type synthetic exponential ratio estimator $t^s_{bt,a}$, if*

$$MSE(t^s_{bk,a}) < MSE(t^s_{bt,a}) \implies \frac{M_2^2}{L_2} > 1 - \frac{1}{\bar{Y}_a^2}\left\{\begin{array}{l}\bar{Y}^2 f(K_{1a}^2 C_y^2 + K_{2a}^2 C_x^2 - 2K_{1a}K_{2a}C_{yx}) \\ +2(\bar{Y}K_{1a} - \bar{Y}_a) f \bar{Y}\left(K_{3a}C_x^2 - K_{2a}C_{yx}\right) \\ +(K_{1a}\bar{Y} - \bar{Y}_a)^2\end{array}\right\}$$

The proposed direct and synthetic estimators outperform their conventional counterparts under the efficiency conditions reported under Lemma 4.1 to Lemma 4.8.

5. Simulation study

To validate the theoretical results, following Singh and Horn [16] and motivated by Bhushan and Kumar [3,4], we conduct a simulation study. In the process of simulation study, we hypothetically draw some symmetric and asymmetric populations using the following models:

$$y = 8.7 + \sqrt{(1-\rho_{xy}^2)}\, y^* + \rho_{xy}\left(\frac{S_y}{S_x}\right) x^*$$

$$x = 4.1 + x^*$$

where x^* and y^* are independent variables for the corresponding distributions. Using the models mentioned above, we generated the populations shown below:

(1). A Normal population of size $N = 12000$ utilizing $x^* \sim N(19,47)$ and $y^* \sim N(13,27)$ with varying correlation coefficients $\rho_{xy} = 0.1, 0.3, 0.5, 0.7, 0.9$.
(2). An exponential population of size $N = 12000$ utilizing $x^* \sim Exp(0.05)$ and $y^* \sim Exp(0.09)$ with varying correlation coefficients $\rho_{xy} = 0.1, 0.3, 0.5, 0.7, 0.9$.

The above populations are divided into 6 equal domains of size 2000. A sample of size 50 is being randomly selected from every domain and the descriptive statistics are computed. Performing 20,000 iterations, we have computed MSE and percent relative efficiency (PRE) of the proposed logarithmic type direct and synthetic estimators with respect to the direct and synthetic mean estimators, respectively, by using the following formulae:

$$MSE(t^*) = \frac{1}{20,000}\sum_{s=1}^{20,000}(t^* - \bar{Y}_a)^2$$

$$MSE(t^{**}) = \frac{1}{20,000}\sum_{s=1}^{20,000}(t^{**} - \bar{Y}_a)^2$$

$$PRE^d = \frac{MSE(t^d_{m,a})}{MSE(t^*)} \times 100$$

$$PRE^s = \frac{MSE(t^s_{m,a})}{MSE(t^{**})} \times 100$$

where $t^* = t^d_{m,a}, t^d_{r,a}, t^d_{tg,a}, t^d_{bk,a}$, and $t^{**} = t^s_{m,a}, t^s_{r,a}, t^s_{tg,a}, t^s_{bt,a}, t^s_{bk,a}$. The simulation results of the direct estimators for normal and exponential populations are reported in Tables 1-2, respectively, whereas the simulation results of the synthetic estimators for normal and exponential populations are reported in Tables 3-4, respectively.

5.1. Discussion of simulation results

Following a thorough examination of the outcomes of the simulation study, we outline the discussion of the results in point-wise fashion.

(i). From the results of Table 1 based on the artificially generated normal population, the proposed logarithmic type direct estimator $t^d_{bk,a}$ dominates direct mean per unit estimator $t^d_{m,a}$, direct ratio estimator $t^d_{bk,r}$, direct exponential ratio estimator $t^d_{bt,a}$ envisaged on the lines of Bahl and Tuteja [1] estimator, and direct type Tikkiwal and Ghiya [19] estimator $t^d_{tg,a}$ for each value of correlation coefficient in each domain by minimum MSE and maximum PRE. This fact can easily be observed from Fig. 1.
(ii). The similar tendency as seen from the outcomes of Table 1 and Fig. 1 may also be seen from the results of Table 2 and Fig. 2 based on the artificially generated exponential population.
(iii). From the results of Table 3 consisting of an artificially generated normal population, the proposed logarithmic type synthetic estimator $t^s_{bk,a}$ surpasses the synthetic mean per unit estimator $t^s_{m,a}$, synthetic ratio estimator $t^s_{bk,r}$, synthetic exponential ratio estimator $t^s_{bt,a}$ envisaged on the lines of Bahl and Tuteja [1] estimator, synthetic regression estimator $t^s_{reg,a}$, and Tikkiwal and Ghiya [19] synthetic estimator $t^s_{tg,a}$ for each value of correlation coefficient in each domain by minimum MSE and maximum PRE. This fact can easily be observed from Fig. 3.
(iv). The similar tendency as observed from the outcomes of Table 3 and Fig. 3 may also be observed from the results of Table 4 and Fig. 4 based on the artificially generated exponential population.
(v). From the results of Tables 1-2, the MSE and PRE of the proposed direct estimator $t^d_{bk,a}$ decrease and increase, respectively, as the correlation coefficient increases. The similar tendency can be observed from the results of the proposed synthetic estimator $t^s_{bk,a}$ reported in Tables 3-4.
(vi). From the results of Tables 1-4, the suggested synthetic estimator $t^s_{bk,a}$ represses the suggested direct estimator $t^d_{bk,a}$ in all domains.

6. Real data application- a case study

The state of Uttar Pradesh similar to the most of the other Indian states is divided into a several districts for revenue collection and other administrative duties. Additionally, every district is divided into several tehsils, and each tehsil is further subdivided into several revenue inspector circles (RICs). Each RIC is made up of many villages. In this study, we take RICs as small domains.

It is observed that the cultivated area under any particular crop decreases or increases every year. Thus, for application, we take the crop acreage estimation issue for the RICs of Mohanlalganj tehsil of Uttar Pradesh. Six RICs of Mohanlalganj tehsil are taken as small domains. The paddy crop acreage (in hectares) for the agricultural season 2018-19 is taken as study variable, whereas the paddy crop acreage for the agricultural season 2017-18 is taken as auxiliary variable. Various information of the RICs of Mohanlalganj tehsil are reported in Table 5, whereas the parameters of each domain are reported in Table 6 for ready reference.
Using the domain parameters given in Table 6, we have computed the MSE and PRE of the proposed direct and synthetic logarithmic estimators with respect to the direct and synthetic mean estimators, respectively, by using the following formulae:

$$PRE^d = \frac{MSE(t^d_{m,a})}{MSE(t^*)} \times 100$$

$$PRE^s = \frac{MSE(t^s_{m,a})}{MSE(t^{**})} \times 100$$

The results of the numerical study for the direct and synthetic estimators are shown in Tables 7-8, respectively, which show the dominance of the proposed direct and synthetic estimators over their conventional counterparts. Further, the suggested direct estimator represses the suggested synthetic estimator in RICs, Mohanlalganj, Khujauli, and Amethi, whereas the suggested synthetic estimator represses the suggested direct estimator in RICs Nagram, Gosaiganj, and Behrauli.

7. Conclusion

In this paper, we have proposed direct and synthetic logarithmic estimators for the domain mean under SRS. The MSE of the proposed

Table 1
Results of direct estimators for normal population.

Estimators Domains	ρ_{xy}	$t^d_{m,a}$		$t^d_{r,a}$		$t^d_{bt,a}$		$t^d_{tg,a}$		$\mathbf{t^d_{bk,a}}$	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	0.1	43.12	100.00	56.16	76.78	44.71	96.43	41.59	103.66	**38.06**	**113.29**
	0.3	43.12	100.00	43.14	99.95	39.08	110.33	38.01	113.45	**35.20**	**122.50**
	0.5	43.12	100.00	32.64	132.10	34.33	125.61	31.19	138.26	**29.75**	**144.94**
	0.7	43.12	100.00	23.23	185.61	29.84	144.50	21.16	203.78	**20.94**	**205.94**
	0.9	43.12	100.00	13.31	323.83	24.69	174.63	7.91	545.32	**6.94**	**620.86**
2	0.1	43.13	100.00	56.46	76.41	44.81	96.27	41.63	103.61	**38.13**	**113.14**
	0.3	43.13	100.00	43.32	99.56	39.14	110.21	38.06	113.32	**35.28**	**122.27**
	0.5	43.13	100.00	32.75	131.72	34.36	125.54	31.26	138.00	**29.83**	**144.58**
	0.7	43.13	100.00	23.28	185.30	29.85	144.50	21.23	203.22	**20.92**	**206.15**
	0.9	43.13	100.00	13.31	324.02	24.69	174.73	7.94	543.30	**7.15**	**603.07**
3	0.1	43.12	100.00	56.25	76.66	44.73	96.41	41.62	103.62	**38.13**	**113.09**
	0.3	43.12	100.00	43.19	99.84	39.08	110.34	38.04	113.36	**35.27**	**122.26**
	0.5	43.12	100.00	32.66	132.03	34.32	125.66	31.23	138.10	**29.81**	**144.65**
	0.7	43.12	100.00	23.23	185.66	29.82	144.61	21.20	203.39	**20.97**	**205.62**
	0.9	43.12	100.00	13.29	324.57	24.66	174.84	7.93	543.72	**7.09**	**608.18**
4	0.1	43.15	100.00	56.34	76.59	44.75	96.41	41.62	103.66	**38.14**	**113.13**
	0.3	43.15	100.00	43.20	99.88	39.08	110.42	38.01	113.53	**35.25**	**122.42**
	0.5	43.15	100.00	32.63	132.24	34.30	125.80	31.16	138.47	**29.73**	**145.15**
	0.7	43.15	100.00	23.18	186.14	29.80	144.81	21.13	204.17	**20.91**	**206.36**
	0.9	43.15	100.00	13.25	325.69	24.65	175.07	7.90	546.16	**6.69**	**645.11**
5	0.1	43.14	100.00	56.34	76.57	44.77	96.36	41.63	103.63	**38.12**	**113.16**
	0.3	43.14	100.00	43.24	99.78	39.11	110.32	38.03	113.44	**35.25**	**122.38**
	0.5	43.14	100.00	32.68	132.01	34.33	125.65	31.20	138.28	**29.78**	**144.85**
	0.7	43.14	100.00	23.23	185.71	29.83	144.61	21.17	203.76	**20.95**	**205.88**
	0.9	43.14	100.00	13.29	324.62	24.68	174.83	7.92	544.89	**6.79**	**635.08**
6	0.1	43.10	100.00	56.53	76.25	44.81	96.19	41.62	103.56	**38.12**	**113.09**
	0.3	43.10	100.00	43.35	99.44	39.13	110.16	38.06	113.26	**35.27**	**122.22**
	0.5	43.10	100.00	32.75	131.63	34.34	125.52	31.26	137.91	**29.85**	**144.42**
	0.7	43.10	100.00	23.27	185.24	29.83	144.49	21.23	203.06	**20.98**	**205.43**
	0.9	43.14	100.00	13.29	324.62	24.68	174.83	7.92	544.89	**6.79**	**635.08**

Table 2
Results of direct estimators for exponential population.

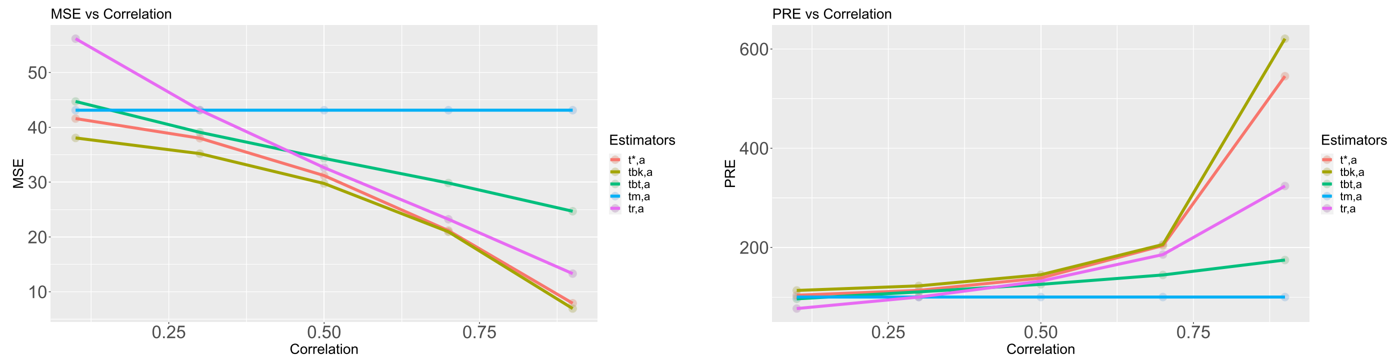
Estimators Domains	ρ_{xy}	$t^d_{m,a}$		$t^d_{r,a}$		$t^d_{bt,a}$		$t^d_{tg,a}$		$\mathbf{t^d_{bk,a}}$	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	0.1	7.75	100.00	10.07	76.99	8.08	95.99	7.51	103.18	**7.42**	**104.55**
	0.3	7.75	100.00	7.77	99.84	7.05	109.95	6.86	113.09	**6.78**	**114.38**
	0.5	7.75	100.00	5.86	132.34	6.17	125.65	5.60	138.48	**5.56**	**139.46**
	0.7	7.75	100.00	4.13	187.83	5.33	145.36	3.82	203.23	**3.81**	**203.72**
	0.9	7.75	100.00	2.31	336.34	4.38	176.93	1.47	527.65	**1.45**	**533.97**
2	0.1	7.72	100.00	10.02	77.07	8.04	96.04	7.48	103.23	**7.38**	**104.58**
	0.3	7.72	100.00	7.73	99.94	7.02	110.01	6.82	113.17	**6.75**	**114.46**
	0.5	7.72	100.00	5.83	132.46	6.14	125.71	5.57	138.62	**5.53**	**139.59**
	0.7	7.72	100.00	4.11	187.97	5.31	145.41	3.80	203.44	**3.79**	**203.92**
	0.9	7.72	100.00	2.29	336.44	4.36	176.96	1.46	527.86	**1.45**	**534.19**
3	0.1	7.69	100.00	10.00	76.92	8.01	95.97	7.45	103.20	**7.36**	**104.56**
	0.3	7.69	100.00	7.71	99.71	7.00	109.91	6.80	113.10	**6.72**	**114.39**
	0.5	7.69	100.00	5.82	132.09	6.12	125.58	5.56	138.37	**5.52**	**139.35**
	0.7	7.69	100.00	4.10	187.40	5.29	145.27	3.79	202.69	**3.79**	**203.18**
	0.9	7.69	100.00	2.29	335.45	4.35	176.81	1.47	524.49	**1.45**	**530.69**
4	0.1	7.7	100.00	10.01	76.89	8.02	95.98	7.46	103.20	**7.36**	**104.56**
	0.3	7.7	100.00	7.73	99.64	7.01	109.90	6.81	113.05	**6.73**	**114.34**
	0.5	7.7	100.00	5.83	131.98	6.13	125.57	5.57	138.22	**5.53**	**139.20**
	0.7	7.7	100.00	4.11	187.22	5.30	145.24	3.81	202.32	**3.80**	**202.81**
	0.9	7.7	100.00	2.30	335.14	4.36	176.79	1.47	522.84	**1.46**	**528.98**
5	0.1	7.75	100.00	10.07	76.97	8.08	95.97	7.51	103.25	**7.41**	**104.61**
	0.3	7.75	100.00	7.76	99.85	7.05	109.95	6.85	113.18	**6.77**	**114.47**
	0.5	7.75	100.00	5.86	132.37	6.17	125.66	5.59	138.61	**5.55**	**139.58**
	0.7	7.75	100.00	4.13	187.91	5.33	145.38	3.81	203.45	**3.80**	**203.94**
	0.9	7.75	100.00	2.30	336.54	4.38	176.97	1.47	528.16	**1.45**	**534.54**
6	0.1	7.74	100.00	10.05	77.03	8.06	96.01	7.50	103.23	**7.40**	**104.59**
	0.3	7.74	100.00	7.75	99.90	7.04	109.98	6.84	113.16	**6.76**	**114.45**
	0.5	7.74	100.00	5.85	132.41	6.16	125.68	5.59	138.61	**5.55**	**139.58**
	0.7	7.74	100.00	4.12	187.92	5.33	145.39	3.80	203.49	**3.80**	**203.97**
	0.9	7.74	100.00	2.30	336.45	4.38	176.95	1.47	528.42	**1.45**	**534.80**

Table 3
Results of synthetic estimators for normal population.

Estimators Domains	ρ_{xy}	$t^s_{m,a}$		$t^s_{r,a}$		$t^s_{bt,a}$		$t^s_{tg,a}$		$\mathbf{t^s_{bk,a}}$	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	0.1	39.89	100.00	3.97	1005.53	47.42	84.11	3.07	1300.01	**3.05**	**1306.89**
	0.3	39.89	100.00	3.09	1291.80	50.89	78.37	2.82	1414.3	**2.81**	**1420.84**
	0.5	39.89	100.00	2.35	1700.33	54.29	73.48	2.32	1716.02	**2.32**	**1720.77**
	0.7	39.89	100.00	1.66	2402.74	58.00	68.78	1.58	2523.55	**1.58**	**2524.66**
	0.9	39.89	100.00	0.93	4302.86	63.20	63.11	0.59	6773.75	**0.59**	**6798.54**
2	0.1	39.58	100.00	3.99	993.11	46.98	84.24	3.08	1284.44	**3.07**	**1291.24**
	0.3	39.58	100.00	3.10	1276.03	50.45	78.45	2.83	1397.36	**2.82**	**1403.82**
	0.5	39.58	100.00	2.36	1679.88	53.84	73.52	2.33	1695.46	**2.33**	**1700.16**
	0.7	39.58	100.00	1.67	2374.31	57.53	68.80	1.59	2493.33	**1.59**	**2494.43**
	0.9	39.58	100.00	0.93	4252.64	62.67	63.15	0.59	6692.62	**0.59**	**6717.09**
3	0.1	39.31	100.00	3.99	986.06	46.64	84.29	3.08	1275.11	**3.07**	**1281.86**
	0.3	39.31	100.00	3.10	1266.85	50.04	78.55	2.83	1387.21	**2.82**	**1393.62**
	0.5	39.31	100.00	2.36	1667.65	53.38	73.64	2.34	1683.15	**2.33**	**1687.81**
	0.7	39.31	100.00	1.67	2356.95	57.02	68.94	1.59	2475.22	**1.59**	**2476.31**
	0.9	39.31	100.00	0.93	4221.98	62.13	63.27	0.59	6644.01	**0.59**	**6668.30**
4	0.1	39.66	100.00	4.00	992.62	46.98	84.43	3.09	1283.88	**3.07**	**1290.67**
	0.3	39.66	100.00	3.11	1275.39	50.45	78.62	2.84	1396.74	**2.83**	**1403.20**
	0.5	39.66	100.00	2.36	1679.06	53.85	73.66	2.34	1694.72	**2.33**	**1699.41**
	0.7	39.66	100.00	1.67	2373.34	57.55	68.92	1.59	2492.23	**1.59**	**2493.33**
	0.9	39.66	100.00	0.93	4251.71	62.72	63.24	0.59	6689.67	**0.59**	**6714.12**
5	0.1	39.58	100.00	3.98	994.61	47.14	83.97	3.08	1285.39	**3.06**	**1292.19**
	0.3	39.58	100.00	3.10	1277.55	50.53	78.33	2.83	1398.39	**2.82**	**1404.86**
	0.5	39.58	100.00	2.35	1681.21	53.86	73.49	2.33	1696.71	**2.33**	**1701.41**
	0.7	39.58	100.00	1.67	2375.11	57.51	68.83	1.59	2495.16	**1.59**	**2496.26**
	0.9	39.58	100.00	0.93	4251.97	62.63	63.20	0.59	6697.54	**0.59**	**6722.09**
6	0.1	39.74	100.00	3.99	996.71	46.84	84.85	3.08	1288.71	**3.07**	**1295.53**
	0.3	39.74	100.00	3.10	1280.47	50.33	78.96	2.83	1402.00	**2.82**	**1408.49**
	0.5	39.74	100.00	2.36	1685.46	53.76	73.92	2.34	1701.10	**2.33**	**1705.81**
	0.7	39.74	100.00	1.67	2381.86	57.52	69.10	1.59	2501.61	**1.59**	**2502.71**
	0.9	39.74	100.00	0.93	4265.81	62.78	63.30	0.59	6714.85	**0.59**	**6739.43**

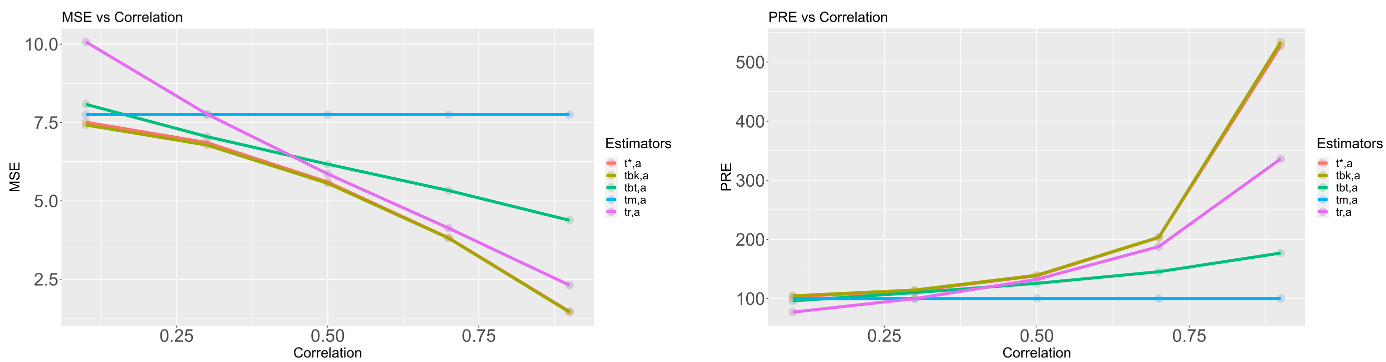
Table 4
Results of synthetic estimators for exponential population.

Estimators Domains	ρ_{xy}	$t^s_{m,a}$		$t^s_{r,a}$		$t^s_{bt,a}$		$t^s_{tg,a}$		$\mathbf{t^s_{bk,a}}$	
		MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	0.1	7.82	100.00	1.67	467.02	8.99	86.95	1.27	615.60	**1.27**	**616.97**
	0.3	7.82	100.00	1.29	605.91	9.63	81.23	1.17	669.72	**1.17**	**671.03**
	0.5	7.82	100.00	0.97	805.44	10.24	76.42	0.96	812.59	**0.96**	**813.56**
	0.7	7.82	100.00	0.68	1147.41	10.88	71.89	0.65	1194.99	**0.65**	**1195.22**
	0.9	7.82	100.00	0.38	2052.04	11.74	66.63	0.24	3207.60	**0.24**	**3212.39**
2	0.1	7.90	100.00	1.67	473.03	9.00	87.80	1.27	623.54	**1.26**	**624.93**
	0.3	7.90	100.00	1.29	613.69	9.65	81.92	1.16	678.35	**1.16**	**679.68**
	0.5	7.90	100.00	0.97	815.81	10.27	76.97	0.96	823.07	**0.96**	**824.04**
	0.7	7.90	100.00	0.68	1162.26	10.93	72.31	0.65	1210.40	**0.65**	**1210.63**
	0.9	7.90	100.00	0.38	2078.88	11.82	66.87	0.24	3248.96	**0.24**	**3253.80**
3	0.1	7.96	100.00	1.67	477.27	9.09	87.48	1.26	629.20	**1.26**	**630.60**
	0.3	7.96	100.00	1.28	619.23	9.75	81.61	1.16	684.52	**1.16**	**685.86**
	0.5	7.96	100.00	0.97	823.21	10.37	76.68	0.96	830.55	**0.96**	**831.53**
	0.7	7.96	100.00	0.68	1172.83	11.04	72.05	0.65	1221.39	**0.65**	**1221.63**
	0.9	7.96	100.00	0.38	2097.70	11.94	66.65	0.24	3278.47	**0.24**	**3283.36**
4	0.1	7.93	100.00	1.67	476.15	9.08	87.30	1.26	627.65	**1.26**	**629.05**
	0.3	7.93	100.00	1.28	617.75	9.73	81.49	1.16	682.83	**1.16**	**684.17**
	0.5	7.93	100.00	0.97	821.20	10.35	76.60	0.96	828.50	**0.96**	**829.48**
	0.7	7.93	100.00	0.68	1169.90	11.01	72.00	0.65	1218.38	**0.65**	**1218.62**
	0.9	7.93	100.00	0.38	2092.50	11.90	66.63	0.24	3270.40	**0.24**	**3275.28**
5	0.1	7.95	100.00	1.68	474.74	9.06	87.79	1.27	625.74	**1.27**	**627.13**
	0.3	7.95	100.00	1.29	615.90	9.71	81.87	1.17	680.75	**1.17**	**682.08**
	0.5	7.95	100.00	0.97	818.71	10.34	76.90	0.96	825.97	**0.96**	**826.95**
	0.7	7.95	100.00	0.68	1166.31	11.01	72.22	0.65	1214.67	**0.65**	**1214.90**
	0.9	7.95	100.00	0.38	2085.89	11.91	66.77	0.24	3260.42	**0.24**	**3265.28**
6	0.1	7.81	100.00	1.67	467.14	8.94	87.37	1.27	615.79	**1.27**	**617.16**
	0.3	7.81	100.00	1.29	606.07	9.58	81.55	1.17	669.92	**1.16**	**671.23**
	0.5	7.81	100.00	0.97	805.66	10.19	76.67	0.96	812.84	**0.96**	**813.80**
	0.7	7.81	100.00	0.68	1147.79	10.84	72.08	0.65	1195.35	**0.65**	**1195.58**
	0.9	7.81	100.00	0.38	2052.92	11.71	66.75	0.24	3208.57	**0.24**	**3213.35**



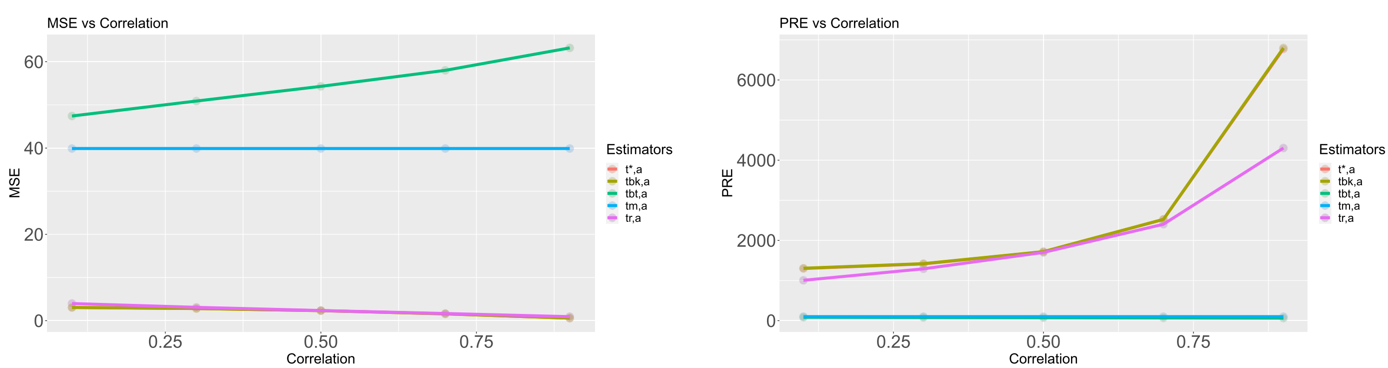
(a) **(b)**

Fig. 1. Graph of (**a**) MSE, (**b**) PRE of direct estimators for normal population.



(c) **(d)**

Fig. 2. Graph of (**c**) MSE, (**d**) PRE of direct estimators for exponential population.



(e) **(f)**

Fig. 3. Graph of (**e**) MSE, (**f**) PRE of synthetic estimators for normal population.

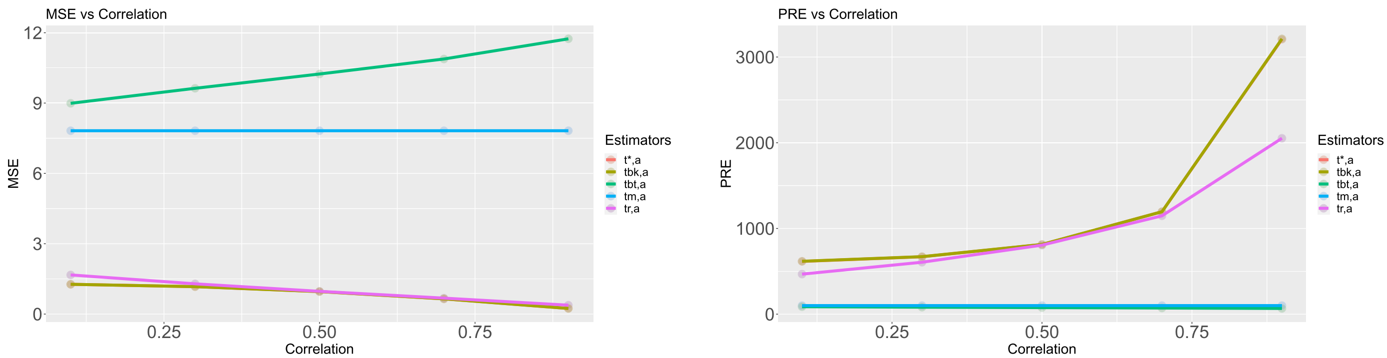
Table 5
Total areas (Irrigated and Unirrigated) under Paddy crop in RICs of Mohanlalganj tehsil for agricultural season 2017-18 and 2018-19.

S. No.	RIC of Mohanlalganj tehsil	Number of villages in RIC	Total area (Irr.+U. Irr.) under the Paddy crop in 2017-18	Total area (Irr.+U. Irr.) under the Paddy crop in 2018-19
1	Mohanlalganj	26	2324	2280
2	Nagram	25	4483	4620
3	Khujauli	32	3740	3708
4	Gosaiganj	34	2887	2944
5	Amethi	29	2957	3220
6	Behrauli	36	4945	4494

Table 6
Population parameters for different domains.

Domains	N_a	$\bar{Y}_a$	$\bar{X}_a$	S_{y_a}	S_{x_a}	$\rho_{y_a x_a}$
1	26	87.69	89.38	63.78	64.33	0.989
2	25	184.804	179.32	125.67	122.00	0.966
3	32	115.88	116.87	74.35	73.18	0.980
4	34	86.59	84.91	70.76	68.10	0.971
5	29	111.03	101.97	75.35	74.53	0.980
6	36	124.83	137.36	90.57	103.72	0.783

estimators is obtained to the first order approximation. Under some efficiency conditions, the proposed direct and synthetic estimators outperform the conventional direct and synthetic estimators. The theoretical



(g) (h)

Fig. 4. Graph of (**g**) MSE, (**h**) PRE of synthetic estimators for exponential population.

Table 7
Results of direct estimators for real population.

Estimators Domains	$t^d_{m,a}$		$t^d_{r,a}$		$t^d_{bt,a}$		$t^d_{tg,a}$		$\mathbf{t^d_{bk,a}}$	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	657.12	100.00	13.83	4751.59	174.64	376.28	13.83	4751.59	**0.41**	**159925.50**
2	2526.77	100.00	171.49	1473.45	716.78	352.52	168.49	1499.65	**139.21**	**1815.05**
3	748.64	100.00	30.00	2495.60	211.11	354.62	29.99	2496.59	**21.39**	**3499.30**
4	568.02	100.00	32.98	1722.35	163.74	346.90	32.91	1725.79	**25.25**	**2249.82**
5	750.52	100.00	37.37	2008.31	176.26	425.81	30.24	2481.91	**22.86**	**3283.81**
6	944.07	100.00	427.66	220.75	430.23	219.43	365.03	258.63	**363.16**	**259.96**

Table 8
Results of synthetic estimators for real population.

Estimators Domains	$t^s_{m,a}$		$t^s_{r,a}$		$t^s_{bt,a}$		$t^s_{tg,a}$ S		$\mathbf{t^s_{bk,a}}$	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	1026.66	100.00	14.36	7148.66	249.34	411.75	13.64	7524.80	**13.56**	**7571.66**
2	4794.44	100.00	63.78	7517.17	1772.37	270.51	60.59	7912.69	**60.22**	**7961.97**
3	177.66	100.00	25.08	708.48	55.95	317.52	23.82	745.76	**23.68**	**750.40**
4	1092.26	100.00	14.00	7800.59	208.81	523.08	13.30	8211.03	**13.22**	**8262.16**
5	210.49	100.00	23.02	914.20	53.37	394.43	21.87	962.30	**21.74**	**968.29**
6	240.51	100.00	29.10	826.41	67.88	354.34	27.65	869.89	**27.48**	**875.31**

results are supported with a simulation study using artificially generated populations and the results are reported in Tables 1-4. The results of the simulation study show that the proposed direct and synthetic estimators dominate their conventional counterparts existing till date in each domain of the simulated populations. Also, from the results of Tables 1-4, the proposed synthetic estimators are found to be better than the proposed direct estimators in all domains. Further, from the simulation results of Tables 1-4, the MSE and PRE for the suggested direct and synthetic estimators $t^d_{bk,a}$ and $t^s_{bk,a}$ reduces and increases, respectively, as the correlation coefficient increases. Moreover, a case study is also presented by using a real data of crop acreage of paddy in Mohanlalganj tehsil of Uttar Pradesh. The numerical results based on real data are reported in Tables 7-8 which demonstrate the ascendancy of the proposed estimators over their counterparts in each domain. Therefore, for the estimation of small domain means, survey practitioners are strongly encouraged to use the proposed direct and synthetic logarithmic estimators.

The proposed direct and synthetic estimators may be examined in the presence of measurement errors using SRS (see, [11]), under stratified sampling (see, [9,6]) under ranked set sampling (see, [2]), under stratified ranked set sampling (see, [7]), etc.

CRediT authorship contribution statement

Anoop Kumar: Conceptualization, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Shashi Bhushan:** Supervision, Writing – review & editing. **Rohini Pokhrel:** Data curation, Formal analysis, Software. **Walid Emam:** Funding acquisition, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Data availability

The article includes all data utilised for this investigation.

Acknowledgements

The study was funded by Researchers Supporting Project number (RSPD2024R749), King Saud University, Riyadh, Saudi Arabia.

References

[1] Bahl S, Tuteja RK. Ratio and product type exponential estimators. J Inf Optim Sci 1991;12(1):159–64.
[2] Bhushan S, Kumar A. On optimal classes of estimators under ranked set sampling. Commun Stat, Theory Methods 2022;51(8):2610–39.

[3] Bhushan S, Kumar A. Imputation of missing data using multi auxiliary information under ranked set sampling. Commun Stat, Simul Comput 2023. https://doi.org/10.1080/03610918.2023.2288796.

[4] Bhushan S, Kumar A. Novel logarithmic imputation methods under ranked set sampling. Measurement 2024. https://doi.org/10.1080/15366367.2023.2247692.

[5] Bhushan S, Kumar A, Pokhrel R. Improved direct estimators for domain mean with an application to municipalities data. Int J Agric Stat Sci 2023;19(2):835–41.

[6] Bhushan S, Kumar A, Singh S. Some efficient classes of estimators under stratified sampling. Commun Stat, Theory Methods 2023;52(6):1767–96.

[7] Bhushan S, Kumar A, Hussam E, Mustafa MS, Zakarya M, Alharbi WR. On stratified ranked set sampling for the quest of an optimal class of estimators. Alex Eng J 2024;86:79–97.

[8] Gonzalez ME. Use and evaluation of synthetic estimates. In: Proceedings of the social statistics section, American statistical association; 1973. p. 33–6.

[9] Iftikhar S, Khalil A, Ali A. A novel and improved logarithmic ratio-product type estimator of mean in stratified random sampling. Math Probl Eng 2022;2022:1–20. https://doi.org/10.1155/2022/6179625.

[10] Khare BB, Ashutosh, Khare S. Comparative study of synthetic estimators with ratio synthetic estimator for domain mean in survey sampling using auxiliary character. Int J Appl Math Stat 2018;57(3):152–61.

[11] Kumar A, Bhushan S, Shukla S, Emam W, Tashkandy Y, Gupta R. Impact of correlated measurement errors on some efficient classes of estimators. J Math 2023;2023:1–27.

[12] Pandey KK, Tikkiwal GC. Generalized class of synthetic estimators for small area under systematic sampling design. Stat Transit 2010;11(1):75–89.

[13] Panse VG, Rajagopalan M, Pillai SS. Estimation of crop yields for small areas. Biometrics 1966;374(384).

[14] Schaible WL, Brock DB, Casady RJ, Schnack GA. An empirical comparison of the simple synthetic and composite estimators for small area statistics. In: Proceedings of amer statist assoc, social statistics section; 1977. p. 1017–21.

[15] Sharma SD, Srivastava AK, Sud UC. Small area crop estimation for crop yield estimates at Gram Panchayat level. J Indian Soc Agric Stat 2004;57:26–37.

[16] Singh HP, Horn S. An alternative estimator for multi-character surveys. Metrika 1998;48:99–107.

[17] Srivastava AK, Ahuja DL, Bhatia DK, Mathur DC. Small area estimation techniques in agriculture. In: Souvenir in advances in agricultural statistics and computer application. New Delhi: Indian Agricultural Statistics Research Institute; 1999. p. 32–45.

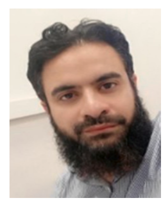
[18] Sud UC, Chandra H, Gupta VK. Calibration-based product estimator in single and two phase sampling. J Stat Theory Pract 2014;8(1):1–11.

[19] Tikkiwal GC, Ghiya A. A generalized class of synthetic estimators with application to crop acreage estimation for small domains. Biom J 2000;42:865–76.

Dr. Anoop Kumar completed Ph.D. in Applied Statistics from the Department of Mathematics and Statistics, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, India. He is currently working as an Assistant Professor in the Department of Statistics, Central University of Haryana, Mahendergarh, India. He also qualified UGC NET twice in population studies. His research area includes sampling survey, missing data, measurement errors, small area estimation, etc.

Prof. Shashi Bhushan received his Ph.D. degree in statistics from Lucknow University, Lucknow, India. He is currently working as a Professor in the Department of Statistics, University of Lucknow, Lucknow, India. His research interests include Sample survey, missing data, non-response, measurement errors, small area estimation, etc.

Mr. Rohini Pokhrel is research scholar in the Department of Mathematics and Statistics, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, Lucknow, India. He qualified UGC NET in population studies. His research interest is in sampling theory, small area estimation, missing data, etc.



Dr. Walid Emam is currently working as Associate Professor in the Department of Statistics and Operation Research, Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.