https://doi.org/10.56093/JISAS.V79I3.3
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 79(3) 2025 221–237

Some Imputation Methods for the Estimation of Domain Mean with an Application to Crop Production Data

Anoop Kumar[1], Shashi Bhushan[2] and Rohini Pokhrel[3]

[1]Central University of Haryana,
[2]University of Lucknow, Lucknow
[3]Dr. Shakuntala Misra National Rehabilitation University, Lucknow

Received 04 April, 2025; Revised 01 December 2025; Accepted 09 December 2025

SUMMARY

Missing data is an unavoidable yet pestering issue in most of the surveys, and if it is not tackled carefully during the survey, the results may exhibit significant biases in the survey findings. This research proposes mean, ratio, product, and some general classes of imputation methods to solve the problem of missing data at domain level and proposes a domain mean estimation methods utilizing simple random sampling. The properties of the suggested imputation methods are determined up to first order approximation, followed by a simulation analysis utilising hypothetically created data sets. Further, the simulation study is extended with a real data application. Additionally, some appropriate recommendations for practical applications have been provided to survey experts.

Keywords: Missing data; Imputation; Efficiency.

MSC2020: 62D10, 62D05.

1. INTRODUCTION

In recent years, the focus of planning has changed from a macro to a micro level. Administrators and policy makers are looking for accurate estimates of many factors at the micro level. The focus of researchers has also turned to the construction of precise estimators for small areas (domains) in response to contemporary demands. As a result of this development, the researchers are proposing several small area estimation (SAE) methods for use. The SAE problems are frequently characterized by the fact that large scale sample surveys, which are intended to generate accurate estimates at the national or state level, typically do not furnish estimates with high accuracy at lower levels, such as district, tehsil/sub-division/county, block, revenue inspector circle (RIC), Gram Panchayat (GP), and Gram. This is due to the fact that, in general, at the lower level, the sample sizes are too small to furnish accurate estimates utilizing conventional estimators. As a result, it was found necessary to come up with alternative estimators employing the information previously gathered from large-scale surveys to produce small area statistics. In the literature on SAE, the conventional design-based domain estimators and alternative estimators are also referred to as direct and synthetic estimators, respectively. The direct estimators consider only the domain information, whereas the synthetic estimators are based on the techniques that increase the efficient sample size by: (i). simulating sufficient data through adequate analysis of the available data under adequate modelling, (ii). by employing data from other domains or time periods with the help of models that assume similarities across domain and/or time periods. The sole approach that is currently known to belong to category (i) is the SICURE modelling developed by Tikkiwal (1993), whereas other techniques such as synthetic,

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

extended regression, and composite belong to category (ii). Out of these methods, the synthetic estimators are utilized for SAE due to their flexibility, wide applicability to general sampling design and potential to provide accuracy in the estimation. The synthetic estimator, on the other hand, may be seriously design biased if the implicit model assumption of similarities throughout domain violates. With the use of auxiliary data, a number of direct and synthetic estimators have been presented to estimate the domain mean.

In SAE, an early advancement for the estimation of crop yield, Panse *et al.* (1966) tried to estimate the crop yields at Block level utilizing a double sampling technique. Due to physical limitations, this method could not be used, and at that time it could not be developed further.

Srivastava *et al.* (1999) 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. Singh and Goel (2000) used remote sensing data to estimate wheat crop production at the Tehsil level (a level higher than the Block), utilizing the synthetic method of estimation. Tikkiwal and Ghiya (2000) proposed a generalized class of synthetic estimators to estimate crop acreage for small domains, 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 GP level was chosen as the area unit level instead of Blocks. This made it necessary for the insurance firms to get the agricultural production at the GP level as soon as possible in order to finalize the premium, make an insurance claim, etc. Sisodia and Singh (2001) developed estimators for crop-production at block level using crop-production and other related information at district level. Sharma *et al.* (2004) proposed a different strategy for scaling down crop production from the Block level to the GP level, based on the development of correction factors using the data on crop yields on particular fields obtained through farmer's evaluation. Pandey and Tikkiwal (2010) proposed a generalized class of synthetic estimators for small areas under systematic sampling. Using the methodology of Sisodia and Singh (2001), Sisodia and Chandra (2012) made additional attempts to enhance the estimates for block level crop production. Tikkiwal *et al.* (2013) examined the performance of generalized regression estimators for small domains. Sharma and Singh (2016) investigated the use of discriminant function analysis to estimate agricultural production at the block level using auxiliary variables related to crop output. Further, Khare and Ashutosh (2017) proposed a synthetic estimator for domain mean by employing auxiliary information and conducted a simulation study on Swedish municipalities data. Khare *et al.* (2018) proposed a modified direct regression estimator for domain mean using auxiliary information when domain mean of auxiliary character is known and unknown. Ashutosh (2020) suggested estimating the domain mean using a traditional synthetic estimator with bivariate auxiliary information.

The issue of missing data is persistent in sample surveys and necessitates quick action to prevent the validity of any conclusions drawn from such data. The properties such as unbiasedness and efficiency of the estimators might both be compromised by the missing data. Imputation of missing data is the preferred and most often used method for dealing with missing data. Rubin (1976) proposed three fundamental conceptions in his landmark work: missing at random (MAR), observed at random (OAR), and parameter distribution (PD). A discrimination between missing at random (MAR) and missing completely at random (MCAR) was provided by Heitjan and Basu (1996). Many renowned writers have addressed the issue of missing data, and different imputation approaches have been used to fill in the gaps. The accessibility of adequate supplementary information is critical for the creation of effective imputations schemes. Numerous prominent researchers, including Rueda *et al.* (2005), Toutenberg and Srivastava (1998), Toutenberg *et al.* (2008), Singh and Horn (2000), Singh and Deo (2003) and Ahmed *et al.* (2006), Singh (2009), Prasad (2017), Bhushan and Pandey (2016, 2018, 2020, 2021), and Bhushan *et al.* (2022) have studied in this field and developed imputations and the corresponding estimators for missing data utilizing auxiliary information. In this study, we use the MCAR model to impute missing data altogether with the following objectives:

i) to propose some conventional mean, ratio, product, and general class of imputation methods for estimating the population mean;

ii) while imputing the missing observations, do not modify the original responses.

The following paper is organised into many sections. In Section 3, the methodology and notations used in this paper are defined. Also, some conventional imputation methods are proposed with the expressions of the bias and mean square error. In Section 4, the comparison of the proposed imputation methods. In Section 5, a broad simulation study is performed by using some hypothetical populations and the interpretation of the simulation results is provided. In Section 6, a real data application is provided. The conclusion of this article is provided in Section 7.

2. METHODOLOGY AND NOTATIONS USED

Consider a specified population $\mathrm{K} = 1, 2, \ldots, N$ of the size N from which a simple random sample S of the size n is drawn without replacement. In order to estimate the mean of domain a, we use the information collected in the sample. Further, let r_a and r be the amount of responding units from chosen n_a and n units and let R_a and R be the set of responding units in the domain a and population K, respectively. Also, R_a^c and R^c symbolize the set of non-responding units in the domain a and total population, respectively. For the unit $i\epsilon R$, the quantity y_i is obtained, but for the units $i\epsilon R^c$, the y quantities are missing and imputed values must be obtained to finalize the formation of sample data set. The imputation is accomplished by utilizing the additional auxiliary information, X, so X_i, the value of X for unit i, is available and positive for all $i\epsilon s$ such that the data $\boldsymbol{X_s} = \{X_i; i\epsilon s\}$ are available.

To establish the bias and mean square error (MSE) of the consequent direct estimators of the proposed direct imputation methods, we take the following notations: $\bar{y}_{r,a} = \bar{Y}_a(1+e_{0a})$, $\bar{x}_{r,a} = \bar{X}_a(1+e_{1a})$, and $\bar{x}_{n,a} = \bar{X}_a(1+e_{2a})$, e's are error terms such that $E(e_{ka}) = 0, k = 0,1,2$ and $E(e_{0a}^2) = f_{r_a} C_{y_a}^2$, $E(e_{1a}^2) = f_{r_a} C_{x_a}^2$, $E(e_{2a}^2) = f_{n_a} C_{x_a}^2$, $E(e_{0a}e_{1a}) = f_{r_a}\rho_{y_a x_a} C_{y_a} C_{x_a}$, $E(e_{0a}e_{2a}) = f_{n_a}\rho_{y_a x_a} C_{y_a} C_{x_a}$, $E(e_{1a}e_{2a}) = f_{n_a} C_{x_a}^2$, where $f_{r_a} = \left(\frac{1}{r_a} - \frac{1}{N_a}\right)$ and $f_{n_a} = \left(\frac{1}{n_a} - \frac{1}{N_a}\right)$. Here, $\bar{y}_{r,a}$ and $\bar{x}_{r,a}$ are the sample means of responding units for the study variable y and auxiliary variable x in the domain a, while $\bar{x}_{n,a}$ is the sample mean of the auxiliary variable x in the domain a. Also, $\bar{Y}_a$ and $\bar{X}_a$ are the population means of study and auxiliary variables in the domain a. Further, C_{y_a} and C_{x_a} are the coefficient of variation of study and auxiliary variables in the domain a, respectively, $\rho_{y_a x_a}$ is the correlation coefficient between study and auxiliary variables in the domain a. Moreover, N_a is the size of domain a.

To establish the bias and MSE of the consequent synthetic estimators of the proposed synthetic imputation methods, we take the following notations: $\bar{y}_r = \bar{Y}(1+\varepsilon_0)$, $\bar{x}_r = \bar{X}(1+\varepsilon_1)$, $\bar{x}_n = \bar{X}(1+\varepsilon_2)$, the $\varepsilon's$ are error terms such that $E(\varepsilon_k) = 0$, $k = 0,1,2$ and $E(\varepsilon_0^2) = f_r C_y^2$, $E(\varepsilon_1^2) = f_r C_x^2$, $E(\epsilon_2^2) = f_n C_x^2$, $E(\varepsilon_0\varepsilon_1) = f_r \rho_{yx} C_y C_x$, $E(\varepsilon_0\varepsilon_2) = f_n \rho_{yx} C_y C_x$, $E(\varepsilon_1\varepsilon_2) = f_n C_x^2$, where $f_r = \left(\frac{1}{r} - \frac{1}{N}\right)$ and $f_n = \left(\frac{1}{n} - \frac{1}{N}\right)$. Here, $\bar{y}_r$ and $\bar{x}_r$ are the sample means of responding units for the study variable y and auxiliary variable x, while $\bar{x}_n$ is the sample mean of the auxiliary variable x. Also, $\bar{Y}$ and $\bar{X}$ are the population means of study and auxiliary variables, respectively. Further, C_y and C_x are the coefficient of variation of study and auxiliary variables, respectively, ρ_{yx} is the correlation coefficient between study and auxiliary variables.

3. PROPOSED IMPUTATION METHODS

In this section, we propose some conventional imputation methods for the estimation of mean of domain a using SRS.

3.1 Mean imputation method

We propose mean imputation of domain mean extending the results of Lee *et al.* (1994), for unit value imputation, when the i^{th} sample unit in domain a is absent and requires imputation, the technique of imputation is expressed for direct and synthetic methods as

$$y_{.i_m}^{d} = \begin{cases} y_i \, for \quad i \in R_a \\ \bar{y}_{r,a} \;\; for \quad i \in R_a^c \end{cases}$$

and

$$y_{.i_m}^{s} = \begin{cases} y_i \; for \quad i \in R \\ \bar{y}_r \, for \quad i \in R^c \end{cases}$$

The consequent estimator for mean of domain a are

$$\beth_m^d = \bar{y}_{r,a}.$$

and $\beth_m^s = \bar{y}_r$

The superscript "d" in the imputation method $y_{.i_m}^{d}$ and the corresponding estimator $\beth_m^d$ stands for "direct", while the superscript "s" in the imputation method $y_{.i_m}^{s}$ and the corresponding estimator $\beth_m^s$ stands for "synthetic". The bias and MSE of the consequent direct and synthetic mean estimators are

$$Bias\left(\beth_m^d\right) = 0. \tag{3.1}$$

$$Bias\left(\beth_m^s\right) = \bar{Y} - \bar{Y}_a \tag{3.2}$$

$$MSE\left(\beth_m^d\right) = \bar{Y}_a^2 f_{r_a} C_{y_a}^2. \tag{3.3}$$

$$MSE\left(\beth_m^s\right) = \left(\bar{Y} - \bar{Y}_a\right)^2 + \bar{Y}^2 f_r C_y^2. \tag{3.4}$$

The imputation approaches are distinguished into two schemes when additional auxiliary information is taken into account.

Scheme I: When $\bar{X}_a$ is known and $\bar{x}_{n,a}$ is used.

Scheme II: When $\bar{X}_a$ is known and $\bar{x}_{r,a}$ is used.

Further, in this section, we propose some conventional direct and synthetic imputation methods for the estimation of domain mean in presence of missing data under the above-mentioned schemes.

3.2 Direct and synthetic general class of imputation methods

The proposed direct and synthetic general class of imputation methods under scheme I are given below:

$$y_{.i_1}^{d} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{X}_a}{\bar{x}_{n,a}}\right)^{\theta_1} - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

$$y_{.i_3}^{s} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n - r}\left[n \bar{y}_r\left(\dfrac{\bar{X}_a}{\bar{x}_n}\right)^{\theta_3} - r \bar{y}_r\right] for \;\; i \in R^c \end{cases}$$

The proposed direct and synthetic general class of imputation methods under scheme II are given below:

$$y_{.i_2}^{d} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{X}_a}{\bar{x}_{r,a}}\right)^{\theta_2} - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

$$y_{.i_4}^{s} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n - r}\left[n \bar{y}_r\left(\dfrac{\bar{X}_a}{\bar{x}_r}\right)^{\theta_4} - r \bar{y}_r\right] for \, i \in R^c \end{cases}$$

The resulting estimators are calculated under the schemes described above as

$$\beth_1^d = \bar{y}_{r,a}\left(\frac{\bar{X}_a}{\bar{x}_{n,a}}\right)^{\theta_1}$$

$$\beth_2^d = \bar{y}_{r,a}\left(\frac{\bar{X}_a}{\bar{x}_{r,a}}\right)^{\theta_2}$$

$$\beth_3^s = \bar{y}_r\left(\frac{\bar{X}_a}{\bar{x}_n}\right)^{\theta_3}$$

$$\beth_4^s = \bar{y}_r\left(\frac{\bar{X}_a}{\bar{x}_r}\right)^{\theta_4}$$

where θ_1, θ_2, θ_3, and θ_4 are suitably chosen scalars to optimize the MSE of the corresponding estimators. The optimum values of $\theta_j ; j = 1,2,3,4$ are determined laterin Theorem 3.1 and Theorem 3.2.

3.3 Special cases

(i). When $\theta_j = 1; j = 1,2;$ then under schemes I and II, the proposed direct imputation methods $y^d_{.i_j}$ and the corresponding resultant direct estimators $\beth^d_j$ deform into the classical direct ratio imputation methods $y^d_{.i_{r_j}}$ and the corresponding resultant classical direct ratio estimators $\beth^d_{r_j}$, respectively.

Scheme I

$$y^d_{.i_{r_1}} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{X}_a}{\bar{x}_{n,a}}\right) - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

Scheme II

$$y^d_{.i_{r_2}} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{X}_a}{\bar{x}_{r,a}}\right) - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

The consequent direct ratio estimators under above schemes are

$$\beth^d_{r_1} = \bar{y}_{r,a}\left(\frac{\bar{X}_a}{\bar{x}_{n,a}}\right)$$

$$\beth^d_{r_2} = \bar{y}_{r,a}\left(\frac{\bar{X}_a}{\bar{x}_{r,a}}\right)$$

(ii) When $\theta_j = 1, \; j = 3,4;$ then under schemes I and II, the proposed synthetic imputation methods $y^s_{.i_j}$ and the corresponding resultant synthetic estimators $\beth^s_j$ deform into the classical synthetic ratio imputation methods $y^s_{.i_{r_j}}$ and the corresponding resultant classical synthetic ratio estimators $\beth^s_{r_j}$, respectively.

Scheme I

$$y^s_{.i_{r_3}} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n - r}\left[n\bar{y}_r\left(\dfrac{\bar{X}_a}{\bar{x}_n}\right) - r\bar{y}_r\right] for \, i \in R^c \end{cases}.$$

Scheme II

$$y^s_{.i_{r_4}} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n - r}\left[n\bar{y}_r\left(\dfrac{\bar{X}_a}{\bar{x}_r}\right) - r\bar{y}_r\right] for \, i \in R^c \end{cases}.$$

The consequent synthetic ratio estimators under above schemes are

$$\beth^s_{r_4} = \bar{y}_r\left(\frac{\bar{X}_a}{\bar{x}_r}\right)$$

(iii) When $\theta_j = -1, \; j = 1,2;$ then under schemes I and II, the proposed direct imputation methods $y^d_{.i_j}$ and the corresponding resultant direct estimators $\beth^d_j$ deform into the classical direct product imputation methods $y^d_{.i_{p_j}}$ and the corresponding resultant classical direct product estimators $\beth^d_{p_j}$, respectively.

Scheme I

$$y^d_{.i_{p_1}} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{x}_{n,a}}{\bar{X}_a}\right) - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

Sheme II

$$y^d_{.i_{p_1}} = \begin{cases} y_i \, for \, i \in R_a \\ \dfrac{1}{n_a - r_a}\left[n_a \bar{y}_{r,a}\left(\dfrac{\bar{x}_{r,a}}{\bar{X}_a}\right) - r_a \bar{y}_{r,a}\right] for \, i \in R_a^c \end{cases}$$

The consequent estimators under above schemes are

$$\beth^d_{p_1} = \bar{y}_{r,a}\left(\frac{\bar{x}_{n,a}}{\bar{X}_a}\right)$$

$$\beth^d_{p_2} = \bar{y}_{r,a}\left(\frac{\bar{x}_{r,a}}{\bar{X}_a}\right)$$

(iv) When $\theta_j = -1, \; j = 3,4;$ then under schemes I and II, the proposed synthetic imputation methods $y^s_{.i_j}$ and the corresponding resultant synthetic estimators $\beth^s_j$ deform into the classical synthetic product imputation methods $y^s_{.i_{p_j}}$ and the corresponding resultant classical synthetic product estimators $\sqsupset_{p_j}^{s}$, . respectively.

Scheme I

$$y_{.i_{p3}}^{s} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n-r}\left[n\bar{y}_r\left(\dfrac{\bar{x}_n}{\bar{X}_a}\right) - r\bar{y}_r\right] for \, i \in R^c \end{cases}$$

Scheme II

$$y_{.i_{p4}}^{s} = \begin{cases} y_i \, for \, i \in R \\ \dfrac{1}{n-r}\left[n\bar{y}_r\left(\dfrac{\bar{x}_r}{\bar{X}_a}\right) - r\bar{y}_r\right] for \, i \in R^c \end{cases}$$

The consequent estimators under above schemes are

$$\sqsupset_{p_3}^{s} = \bar{y}_r\left(\frac{\bar{x}_n}{\bar{X}_a}\right)$$

$$\sqsupset_{p_4}^{s} = \bar{y}_r\left(\frac{\bar{x}_r}{\bar{X}_a}\right)$$

Torem 3.1. *The bias, MSE, and minimum MSE of the consequent direct estimators* $\sqsupset_j^d, j = 1,2$ *of the proposed imputation methods* $\sqsupset_{.i_j}^{d}$ *under schemes I and II are given by*

$$Bias\left(\sqsupset_1^d\right) = \bar{Y}_a\left\{\frac{\theta_1(\theta_1+1)}{2} f_{n_a} C_{x_a}^2 - \theta_1 f_{n_a} \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.5}$$

$$Bias\left(\sqsupset_2^d\right) = \bar{Y}_a\left\{\frac{\theta_2(\theta_2+1)}{2} f_{r_a} C_{x_a}^2 - \theta_2 f_{r_a} \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.6}$$

$$MSE\left(\sqsupset_1^d\right) = \bar{Y}_a^2\left\{f_{r_a} C_{y_a}^2 + \theta_1^2 f_{n_a} C_{x_a}^2 - 2\theta_1 f_{n_a} \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.7}$$

$$MSE\left(\sqsupset_2^d\right) = \bar{Y}_a^2 f_{r_a}\left\{C_{y_a}^2 + \theta_2^2 C_{x_a}^2 - 2\theta_2 \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.8}$$

$$minMSE\left(\sqsupset_1^d\right) = \bar{Y}_a^2 C_{y_a}^2\left\{f_{r_a} - f_{n_a}\rho_{y_a x_a}^2\right\}. \tag{3.9}$$

$$minMSE\left(\sqsupset_2^d\right) = \bar{Y}_a^2 C_{y_a}^2 f_{r_a}\left\{1 - \rho_{y_a x_a}^2\right\} \tag{3.10}$$

Proof. Consider the proposed consequent direct estimator $\sqsupset_1^d$ as

$$\sqsupset_1^d = \bar{y}_{r,a}\left(\frac{\bar{X}_a}{\bar{x}_{n,a}}\right)^{\theta_1}$$

We can express the above estimator using the notations established in the previous section as

$$\sqsupset_1^d = \bar{Y}_a(1+e_{0a})\left(\frac{\bar{X}_a}{\bar{X}_a(1+e_{2a})}\right)^{\theta_1}.$$

$$= \bar{Y}_a(1+e_{0a})(1+e_{2a})^{-\theta_1}$$

$$= \bar{Y}_a(1+e_{0a})\left\{1 - \theta_1 e_{2a} + \frac{\theta_1(\theta_1+1)}{2} e_{2a}^2\right\}$$

$$= \bar{Y}_a\left\{1 + e_{0a} - \theta_1 e_{2a} - \theta_1 e_{0a} e_{2a} + \frac{\theta_1(\theta_1+1)}{2} e_{2a}^2\right\}$$

Subtracting $\bar{Y}_a$ on both sides to the above expression, we get

$$\sqsupset_1^d - \bar{Y}_a = \bar{Y}_a\left\{e_{0a} - \theta_1 e_{2a} - \theta_1 e_{0a} e_{2a} + \frac{\theta_1(\theta_1+1)}{2} e_{2a}^2\right\} \tag{3.11}$$

Taking expectation both sides to (3.11), we get bias of the estimator $\sqsupset_1^d$ to the first order expression as

$$Bias\left(\sqsupset_1^d\right) = \bar{Y}_a f_{n_a}\left\{\frac{\theta_1(\theta_1+1)}{2} C_{x_a}^2 - \theta_1 \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.12}$$

Squaring and taking expectation both sides to (3.11), we get MSE of the estimator $\sqsupset_1^d$ to the first order approximation as

$$MSE\left(\sqsupset_1^d\right) = \bar{Y}_a^2\left\{f_{r_a} C_{y_a}^2 + \theta_1^2 f_{n_a} C_{x_a}^2 - 2\theta_1 f_{n_a} \rho_{y_a x_a} C_{y_a} C_{x_a}\right\} \tag{3.13}$$

Partially differentiating (3.13) regarding θ_1 and equating to zero, we get optimum value of θ_1 as

$$\theta_{1(opt)} = \rho_{y_a x_a}\frac{C_{y_a}}{C_{x_a}} \tag{14}$$

Putting the optimum value of θ_1 from the above expression to (3.13), we get minimum MSE of the estimator $\sqsupset_1^d$ as

$$min.MSE\left(\beth_1^d\right)=\bar{Y}_a^2C_{y_a}^2\left\{f_{r_a}-f_{n_a}\rho_{y_ax_a}^2\right\}$$

Similarly, the first order approximated expressions of bias, MSE, and minimum MSE of the proposed direct estimator $\beth_2^d$ can be obtained.

Theorem 3.2. *The bias, MSE, and minimum MSE of the consequent synthetic estimators $\beth_j^s, j=3,4$ of the proposed synthetic imputation methods $y_{.i_j}^s$ under schemes I and II are given by*

$$Bias\left(\beth_3^s\right)=\bar{Y}_a\left\{\frac{\theta_3\left(\theta_3+1\right)}{2}f_nC_x^2-\theta_3f_n\rho_{yx}C_yC_x\right\} \tag{3.15}$$

$$Bias\left(\beth_4^s\right)=\bar{Y}_af_r\left\{\frac{\theta_4\left(\theta_4+1\right)}{2}C_x^2-\theta_4\rho_{yx}C_yC_x\right\} \tag{3.16}$$

$$MSE\left(\beth_3^s\right)=\bar{Y}_a^2\left\{f_rC_y^2+\theta_3^2f_nC_x^2-2\theta_3f_n\rho_{yx}C_yC_x\right\} \tag{3.17}$$

$$MSE\left(\beth_4^s\right)=\bar{Y}_a^2f_r\left\{C_y^2+\theta_4^2C_x^2-2\theta_4\rho_{yx}C_yC_x\right\} \tag{3.18}$$

$$minMSE\left(\beth_3^s\right)=\bar{Y}_a^2C_y^2\left\{f_r-f_n\rho_{yx}^2\right\} \tag{3.19}$$

$$minMSE\left(\beth_4^s\right)=\bar{Y}_a^2C_y^2f_r\left\{1-\rho_{yx}^2\right\} \tag{3.20}$$

Proof. Consider the proposed consequent synthetic estimator $\beth_3^s$ as

$$\beth_3^s=\bar{y}_r\left(\frac{\bar{X}_a}{\bar{x}_n}\right)^{\theta_3}$$

Using the notations defined in earlier section, we express the above estimator as

$$\beth_3^s=\bar{Y}\left(1+\varepsilon_0\right)\left(\frac{\bar{X}_a}{\bar{X}\left(1+\varepsilon_2\right)}\right)^{\theta_3}$$

$$=\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}\left(1+\varepsilon_0\right)\left(1+\varepsilon_2\right)^{-\theta_3}$$

$$=\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}\left(1+\varepsilon_0\right)\left(1-\theta_3\varepsilon_2+\frac{\theta_3\left(\theta_3+1\right)}{2}\varepsilon_2^2\right)$$

$$=\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}\left(1+\varepsilon_0-\theta_3\varepsilon_2-\theta_3\varepsilon_0\varepsilon_2+\frac{\theta_3\left(\theta_3+1\right)}{2}\varepsilon_2^2\right)$$

Subtracting $\bar{Y}_a$ on both sides to the above expression, we get

$$\beth_3^s-\bar{Y}_a=\left\{\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}-\bar{Y}_a\right\}+ \bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}\left(\varepsilon_0-\theta_3\varepsilon_2-\theta_3\varepsilon_0\varepsilon_2+\frac{\theta_3\left(\theta_3+1\right)}{2}\varepsilon_2^2\right) \tag{3.21}$$

Under the assumption for generalized synthetic estimation $\bar{Y}_a=\bar{Y}\left(\frac{\bar{X}_a}{\bar{X}}\right)^{\theta_3}$, the above expression becomes

$$\beth_3^s-\bar{Y}_a=\bar{Y}_a\left(\varepsilon_0-\theta_3\varepsilon_2-\theta_3\varepsilon_0\varepsilon_2+\frac{\theta_3\left(\theta_3+1\right)}{2}\varepsilon_2^2\right) \tag{3.22}$$

Taking expectation both sides to (3.22), we get bias of the synthetic estimator $\beth_3^s$ to the first order approximation as

$$Bias\left(\beth_3^s\right)=\bar{Y}_af_n\left\{\frac{\theta_3\left(\theta_3+1\right)}{2}C_x^2-\theta_3\rho_{yx}C_yC_x\right\} \tag{3.23}$$

Squaring and taking expectation both sides to (3.22), we get MSE of the synthetic estimator $\beth_3^s$ to the first order approximation as

$$MSE\left(\beth_3^s\right)=\bar{Y}_a^2\left\{f_rC_y^2+\theta_3^2f_nC_x^2-2\theta_3f_n\rho_{yx}C_yC_x\right\} \tag{3.24}$$

Partially differentiating (3.24) regarding θ_3 and equating to zero, we get the optimum value of θ_3 as

$$\theta_{3(opt)}=\rho_{yx}\frac{C_y}{C_x}$$

Putting the optimum value of θ_3 from the above expression to (3.24), we get minimum MSE of the estimator $\beth_3^s$ as

$$min.MSE\left(\beth_3^s\right)=\bar{Y}_a^2C_y^2\left\{f_r-f_n\rho_{yx}^2\right\} \tag{3.25}$$

Similarly, the first order approximated expressions of bias, MSE, and minimum MSE of the proposed synthetic estimator $\sqsupset_4^s$ can be obtained.

4. ANALYTICAL STUDY

In the present section, we compare the minimum MSE of the proposed direct and synthetic imputation methods with the corresponding minimum MSE of the existing direct and synthetic imputation methods under schemes I and II.

4.1 Direct imputation methods

Lemma 4.1. (i). T*he proposed direct imputation method* $y_{.i_1}^d$ *dominates the mean imputation method* $y_{.i_m}^d$ *under scheme I, if*

$$MSE\left(\sqsupset_1^d\right) < MSE\left(\hat{u}_m^d\right) \Rightarrow \rho_{y_a x_a} > 0$$

(ii). The proposed direct imputation method $y_{.i_2}^d$ dominates the mean imputation method $y_{.i_m}^d$ under scheme II, if

$$MSE\left(\sqsupset_2^d\right) < MSE\left(\sqsupset_m^d\right) \Rightarrow \rho_{y_a x_a} > 1$$

Lemma 4.2. The proposed direct imputation methods $y_{.i_j}^d\ ; j = 1;2$ dominate the direct ratio imputation methods $y_{.i_{r_j}}$ under schemes I and II, if

$$MSE\left(\sqsupset_j^d\right) < MSE\left(\sqsupset_{r_j}^d\right) \Rightarrow$$

$$\rho_{y_a x_a} > \frac{1}{C_{y_a}}\sqrt{\left(2\rho_{y_a x_a} C_{y_a} C_{x_a} - C_{x_a}^2\right)}$$

Lemma 4.3. The proposed direct imputation methods $y_{.i_j}^d, j = 1,2$ dominate the direct product imputation methods $y_{.i_{p_j}}^d$ under schemes I and II, if

$$MSE\left(\sqsupset_j^d\right) < MSE\left(\sqsupset_{r_j}^d\right) \Rightarrow$$

$$\rho_{y_a x_a} > -\frac{1}{C_{y_a}}\sqrt{\left(2\rho_{y_a x_a} C_{y_a} C_{x_a} + C_{x_a}^2\right)}$$

4.2 Synthetic imputation methods

Lemma 4.4. (i). The proposed synthetic imputation method $y_{.i_1}^s$ dominates the synthetic mean imputation method $y_{.i_m}^s$ under scheme I, if

$$MSE\left(\sqsupset_1^s\right) < MSE\left(\sqsupset_m^s\right) \Rightarrow \rho_{yx} > 0$$

(ii) The proposed synthetic imputation method $y_{.i_2}^s$ dominates the synthetic mean imputation method $y_{.i_{r_j}}^s$ under scheme II, if

$$MSE\left(\sqsupset_2^s\right) < MSE\left(\sqsupset_m^s\right) \Rightarrow \rho_{yx} > 1$$

Lemma 4.5. The proposed synthetic imputation methods $y_{.i_j}^s\ ; j = 1;2$ dominate the synthetic ratio imputation methods $y_{.i_{r_j}}^s$ under schemes I and II, if

$$MSE\left(\sqsupset_j^s\right) < MSE\left(\sqsupset_{r_j}^s\right) \Rightarrow \rho_{yx} > \frac{1}{C_y}\sqrt{\left(2\rho_{yx} C_y C_x - C_x^2\right)}$$

Lemma 4.6. The proposed synthetic imputation methods $y_{.i_j}^s, j = 1,2$ dominate the synthetic product imputation methods $y_{.i_{p_j}}^s$ under schemes I and II, if

$$MSE\left(\sqsupset_j^s\right) < MSE\left(\sqsupset_{r_j}^s\right) \Rightarrow \rho_{yx} > -\frac{1}{C_y}\sqrt{\left(2\rho_{yx} C_y C_x + C_x^2\right)}$$

The suggested direct and synthetic imputation methods demonstrate dominance over the current direct and synthetic imputation methods if the aforementioned lemmas are satisfied. How successful are these conditions in practice? Only an empirical study can provide a solution to this. Additionally, it is a source of relief that these lemmas are easily satisfied in practise when conducting an empirical study using simulated and real data.

5. SIMULATION STUDY

The efficiency of the proposed direct and synthetic imputation methods is evaluated in relation to the existing direct and synthetic imputation methods by using a simulation study designed on the lines of Singh and Horn (1998). In the simulation process, some symmetrical and asymmetrical populations are generated using the following models:

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

$$x = 4.1 + x^*$$

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

1. A normal population of size $N = 12000$ using $x^* \sim N(10,57)$ and $y^* \sim N(17,20)$ with varying correlation coefficients $\rho_{yx} = -0.9, -0.5, -0.1, 0.1, 0.5, 0.9$.
2. An exponential population of size $N = 12000$ using $x^* \sim Exp(0.017)$ and $y^* \sim Exp(0.097)$ with varying correlation coefficients $\rho_{yx} = -0.9, -0.5, -0.1, 0.1, 0.5, 0.9$.

The populations are divided into 6 equal domains of size 2000. We have drawn a random sample of size $(n_1, n_2, n_3, n_4, n_5, n_6) = (200, 250, 300, 350, 500, 400)$ from the respective domains with different response rates $(r_1, r_2, r_3, r_4, r_5, r_6) = (170, 210, 270, 315, 490, 280)$ from the respective samples. The imputation technique is considered and the bias and MSE of the consequent estimators are computed by using 20,000 iterations. The simulation study is devised using the following steps.

(i) For direct estimators, quantify a sample of size n_a randomly from the population of size N_a, while for synthetic estimators, quantify a sample of size n randomly from the population of size N.

(ii) In case of direct estimator, drop randomly $(n_a - r_a)$ units from sample every time, while in case of synthetic estimators, drop randomly $(n - r)$ units from sample every time.

(iii) Impute extracted units utilizing proposed imputation methods studied for quantified samples.

(iv) Tabulate appropriate statistic.

(v) Iterated the prior steps (i) to (iv) 20,000 times.

The simulated bias and MSE of the proposed and existing general class of direct and synthetic estimators are calculated by using the following formulae:

$$Bias\left(\beth_*^d\right) = \frac{1}{20,000}\sum_{i=1}^{20,000}\left(\beth_*^{d(i)} - \bar{Y}_a\right)$$

$$Bias\left(\beth_*^s\right) = \frac{1}{20,000}\sum_{i=1}^{20,000}\left(\beth_*^{s(i)} - \bar{Y}_a\right)$$

$$MSE\left(\beth_*^d\right) = \frac{1}{20,000}\sum_{i=1}^{20,000}\left(\beth_*^{d(i)} - \bar{Y}_a\right)^2$$

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

where $\beth_*^d = \beth_m^d, \beth_{r_j}^d, \beth_{p_j}^d, \beth_j^d, \; j = 1,2$ and $\beth_*^s = \beth_m^s, \beth_{r_j}^s, \beth_{p_j}^s, \beth_j^s, \; j = 3,4$. The simulation results of the consequent direct and synthetic estimators for normal and exponential populations are reported in Table 1 to Table 4.

5.1 Interpretation of simulation results

We interpret the simulation results disclosed from Table 1 to Table 4 in following points.

1. The outcomes drawn from normal population for the consequent direct estimators are reported in Table 1. These outcomes show that:
   a) the MSE of the consequent direct ratio estimator $\beth_{r_1}^d$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to -0.1, while the MSE of the consequent direct ratio estimator $\beth_{r_1}^d$ under scheme I decreases with the successive increase in the values of correlation coefficient ρ_{yx} from 0.1 to 0.9. This tendency in the MSE values of $\beth_{r_1}^d$ can be also observed from scheme II for the direct ratio estimator $\beth_{r_2}^d$. This is opposite to the behaviour of the consequent direct product estimator.
   b) the MSE of the consequent direct product estimator $\beth_{p_1}^d$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to 0.9. This tendency in the MSE values of $\beth_{p_1}^d$ can be also observed from scheme II. This is opposite to the behaviour of the consequent direct ratio estimator.

c) the MSE of the consequent direct general class of estimator $\beth_1^d$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to -0.1, while the MSE of the consequent direct general class of estimator $\beth_1^d$ under scheme I decreases with the successive increase in the values of correlation coefficient ρ_{yx} from 0.1 to 0.9. This tendency in the MSE values of $\beth_1^d$ can be also observed from scheme II for the direct general class of estimator $\beth_2^d$.

d) the consequent direct general class of estimators $\beth_j^d, j = 1,2$ perform better than the proposed direct mean estimators $\beth_m^d$, direct ratio estimators $\beth_{r_j}^d$, and direct product estimator $\beth_{p_j}^d$ under schemes I and II.

2. The similar tendency as observed from the results of Table 1 can also be observed from the results of Table 2 obtained from exponential population.

3. The outcomes drawn from normal population for the consequent synthetic estimators are reported in Table 3. These outcomes show that:

a) the MSE of the consequent synthetic ratio estimator $\beth_{r_3}^s$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to -0.1, while the MSE of the consequent synthetic ratio estimator $\beth_{r_3}^s$ under scheme I decreases with the successive increase in the values of correlation coefficient ρ_{yx} from 0.1 to 0.9. This tendency in the MSE values of $\beth_{r_3}^s$ can be also observed from scheme II for the synthetic ratio estimator $\beth_{r_4}^s$. This is opposite to the behaviour of the consequent synthetic product estimator.

b) the MSE of the consequent synthetic product estimator $\beth_{p_3}^s$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to 0.9. This tendency in the MSE values of $\beth_{p_3}^s$ can be also observed from scheme II for synthetic product estimator $\beth_{p_4}^s$. This is opposite to the behaviour of the consequent synthetic ratio estimator.

c) the MSE of the consequent synthetic general class of estimator $\beth_3^s$ under scheme I increases with the successive decrease in the values of correlation coefficient ρ_{yx} from -0.9 to -0.1, while the MSE of the consequent synthetic general class of estimator $\beth_3^s$ under scheme I decreases with the successive increase in the values of correlation coefficient ρ_{yx} from 0.1 to 0.9. This tendency in the MSE values of $\beth_3^s$ can be also observed from scheme II for the synthetic general class of estimator $\beth_4^s$.

d) the consequent synthetic general class of estimators $\beth_j^s, j = 3,4;$ perform better than the corresponding synthetic mean estimators $\beth_m^s$, synthetic ratio estimators $\beth_{r_j}^s$, and synthetic product estimators $\beth_{p_j}^s$ under schemes I and II.

4. The similar tendency as observed from the results of Table 3 obtained from normal population for synthetic estimators can also be observed from the results of Table 4 obtained from exponential population for synthetic estimators.

5. The simulation results reported in Tables 1-4 for both normal and exponential populations, varying correlation levels, and both schemes, show that the direct estimators consistently achieve lower MSE than the synthetic estimators. This indicates that under the present simulation conditions with moderate domain sample sizes, the direct estimators are more efficient. Although synthetic estimators are theoretically preferable in situations with very small domain samples. Consequently, the synthetic estimators do not gain an advantage, and the direct imputation-based estimators clearly dominate in terms of empirical MSE.

Table 1. Bias and MSE of direct estimators for hypothetically drawn normal population

Domain	ρ_{yx}	Strategy-I								Strategy-II					
		$\sqsupset_m^d$		$\sqsupset_{r_1}^d$		$\sqsupset_{p_1}^d$		$\sqsupset_1^d$		$\sqsupset_{r_2}^d$		$\sqsupset_{p_2}^d$		$\sqsupset_2^d$	
		Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	-0.9	0.000	2.136	1.282	24.407	-0.294	8.464	-0.326	0.689	-1.533	28.774	-0.351	9.705	-0.390	0.406
	-0.5	0.000	2.138	1.729	44.931	-0.163	30.429	-0.287	1.691	2.068	53.322	-0.195	35.976	-0.344	1.603
	-0.1	0.000	2.168	1.829	50.541	-0.033	47.145	-0.241	2.150	2.187	60.026	-0.039	55.963	-0.289	2.147
	0.1	0.000	2.185	1.803	49.227	0.033	52.734	-0.213	2.167	2.157	58.451	0.039	62.646	-0.255	2.163
	0.5	0.000	2.211	1.602	38.888	0.166	56.083	-0.147	1.749	1.916	46.079	0.199	66.646	-0.176	1.659
	0.9	0.000	2.203	1.053	16.737	0.298	40.838	-0.056	0.711	1.260	19.587	0.357	48.414	-0.067	0.419
2	-0.9	0.000	1.719	2.577	48.776	-0.266	33.34	-0.462	0.576	3.138	59.02	-0.324	40.224	-0.563	0.327
	-0.5	0.000	1.740	3.484	89.534	-0.148	75.837	-0.499	1.383	4.242	108.645	-0.181	91.967	-0.608	1.305
	-0.1	0.000	1.705	3.682	100.208	-0.029	97.102	-0.472	1.691	4.484	121.651	-0.036	117.869	-0.574	1.688
	0.1	0.000	1.679	3.632	97.486	0.029	100.635	-0.443	1.666	4.422	118.342	0.035	122.176	-0.540	1.663
	0.5	0.000	1.630	3.23	77.092	0.144	91.998	-0.356	1.296	3.933	93.519	0.175	111.67	-0.434	1.223
	0.9	0.000	1.620	2.122	33.242	0.258	53.561	-0.196	0.542	2.584	40.126	0.314	64.867	-0.238	0.308
3	-0.9	0.000	1.328	1.059	16.208	-0.203	4.99	-0.219	0.377	1.289	18.154	-0.230	5.469	-0.247	0.252
	-0.5	0.000	1.352	1.419	29.539	-0.114	19.452	-0.193	1.053	1.728	33.223	-0.129	21.818	-0.218	1.014
	-0.1	0.000	1.325	1.495	32.929	-0.023	30.625	-0.16	1.313	1.820	37.060	-0.025	34.455	-0.181	1.312
	0.1	0.000	1.304	1.473	32.000	0.022	34.344	-0.141	1.292	1.794	36.013	0.025	38.663	-0.159	1.291
	0.5	0.000	1.260	1.312	25.324	0.110	36.49	-0.097	0.981	1.597	28.469	0.124	41.095	-0.110	0.945
	0.9	0.000	1.244	0.869	11.015	0.197	26.501	-0.037	0.353	1.058	12.292	0.223	29.803	-0.041	0.236
4	-0.9	0.000	1.064	1.164	14.694	-0.175	4.881	-0.191	0.304	1.418	16.529	-0.199	5.395	-0.217	0.202
	-0.5	0.000	1.059	1.563	26.767	-0.097	18.146	-0.173	0.826	1.903	30.229	-0.110	20.447	-0.197	0.794
	-0.1	0.000	1.062	1.651	29.993	-0.019	28.011	-0.146	1.052	2.010	33.889	-0.022	31.641	-0.166	1.051
	0.1	0.000	1.064	1.628	29.198	0.019	31.228	-0.130	1.055	1.982	32.987	0.022	35.291	-0.147	1.054
	0.5	0.000	1.071	1.448	23.100	0.098	32.936	-0.090	0.835	1.763	26.067	0.111	37.227	-0.102	0.803
	0.9	0.000	1.075	0.956	10.004	0.176	23.665	-0.034	0.308	1.164	11.206	0.200	26.708	-0.039	0.204
5	-0.9	0.000	0.651	0.488	9.224	-0.114	2.762	-0.122	0.138	0.501	9.457	-0.117	2.819	-0.126	0.124
	-0.5	0.000	0.642	0.657	16.984	-0.063	11.282	-0.109	0.486	0.675	17.429	-0.065	11.572	-0.112	0.482
	-0.1	0.000	0.623	0.694	19.046	-0.012	17.753	-0.092	0.617	0.713	19.547	-0.013	18.219	-0.094	0.617
	0.1	0.000	0.615	0.685	18.553	0.012	19.868	-0.081	0.609	0.704	19.041	0.013	20.392	-0.083	0.609
	0.5	0.000	0.605	0.611	14.709	0.061	21.009	-0.056	0.458	0.627	15.093	0.063	21.564	-0.058	0.454
	0.9	0.000	0.618	0.402	6.309	0.111	15.101	-0.021	0.131	0.413	6.464	0.114	15.495	-0.021	0.117
6	-0.9	0.000	1.235	0.494	9.817	-0.128	2.838	-0.134	0.584	0.758	14.414	-0.197	3.696	-0.206	0.235
	-0.5	0.000	1.232	0.668	17.942	-0.071	11.556	-0.115	1.031	1.026	26.893	-0.109	17.086	-0.176	0.924
	-0.1	0.000	1.225	0.707	20.151	-0.014	18.667	-0.094	1.217	1.085	30.290	-0.022	28.011	-0.144	1.213
	0.1	0.000	1.222	0.697	19.641	0.014	21.166	-0.082	1.214	1.070	29.508	0.022	31.85	-0.126	1.210
	0.5	0.000	1.219	0.620	15.597	0.071	23.023	-0.055	1.021	0.952	23.300	0.109	34.704	-0.084	0.914
	0.9	0.000	1.224	0.407	6.852	0.128	17.290	-0.019	0.578	0.625	9.867	0.196	25.896	-0.029	0.232

Table 2. Bias and MSE of direct estimators for hypothetically drawn exponential population

Domain	ρ_{yx}	Strategy-I								Strategy-II					
		$\beth_{m}^{s}$		$\beth_{r_1}^{s}$		$\beth_{p_1}^{s}$		$\beth_{1}^{s}$		$\beth_{r_2}^{s}$		$\beth_{p_2}^{s}$		$\beth_{2}^{s}$	
		Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	-0.9	0.000	0.572	0.055	1.018	-0.038	0.300	-0.063	0.185	0.066	1.105	-0.045	0.247	-0.075	0.109
	-0.5	0.000	0.593	0.071	1.018	- 0.021	0.689	-0.021	0.469	0.085	2.080	-0.025	0.707	-0.026	0.445
	-0.1	0.000	0.603	0.074	1.018	- 0.004	1.767	-0.011	0.598	0.089	2.385	-0.005	1.996	-0.013	0.597
	0.1	0.000	0.603	0.073	2.056	0.004	2.417	-0.008	0.598	0.087	2.341	0.005	2.773	-0.009	0.597
	0.5	0.000	0.594	0.066	1.675	0.021	3.709	-0.003	0.470	0.079	1.887	0.025	4.320	-0.004	0.446
	0.9	0.000	0.573	0.049	0.834	0.038	4.387	0.000	0.185	0.058	0.885	0.045	5.134	0.000	0.109
2	-0.9	0.000	0.445	0.044	0.802	- 0.030	0.230	-0.048	0.149	0.054	0.879	-0.036	0.183	-0.059	0.084
	-0.5	0.000	0.432	0.056	1.392	- 0.016	0.529	- 0.016	0.343	0.068	1.600	-0.020	0.550	-0.020	0.324
	-0.1	0.000	0.430	0.059	1.594	- 0.003	1.352	-0.009	0.427	0.072	1.847	-0.004	1.553	-0.011	0.426
	0.1	0.000	0.432	0.059	1.581	0.003	1.850	-0.006	0.428	0.072	1.831	0.004	2.158	-0.008	0.428
	0.5	0.000	0.440	0.054	1.321	0.016	2.862	-0.003	0.350	0.066	1.512	0.020	3.389	-0.003	0.330
	0.9	0.000	0.452	0.040	0.674	0.030	3.458	0.000	0.151	0.048	0.722	0.037	4.113	0.000	0.086
3	-0.9	0.000	0.350	0.044	0.633	- 0.024	0.175	-0.041	0.099	0.054	0.670	-0.028	0.152	-0.046	0.067
	-0.5	0.000	0.345	0.056	1.136	- 0.013	0.414	-0.014	0.268	0.069	1.240	-0.015	0.423	-0.015	0.258
	-0.1	0.000	0.344	0.059	1.306	- 0.003	1.102	-0.007	0.341	0.072	1.432	-0.003	1.201	-0.008	0.340
	0.1	0.000	0.345	0.059	1.293	0.003	1.519	-0.005	0.342	0.072	1.416	0.003	1.673	-0.006	0.341
	0.5	0.000	0.349	0.054	1.069	0.014	2.365	-0.002	0.271	0.066	1.163	0.015	2.629	-0.002	0.261
	0.9	0.000	0.354	0.039	0.527	0.025	2.852	0.000	0.100	0.048	0.549	0.028	3.179	0.000	0.067
4	-0.9	0.000	0.272	0.043	0.527	- 0.020	0.129	-0.031	0.078	0.053	0.541	-0.022	0.110	-0.035	0.052
	-0.5	0.000	0.267	0.055	0.904	- 0.011	0.331	-0.011	0.208	0.067	0.990	-0.012	0.340	-0.012	0.200
	-0.1	0.000	0.267	0.058	1.039	- 0.002	0.878	-0.006	0.265	0.071	1.143	-0.022	0.960	-0.007	0.265
	0.1	0.000	0.269	0.058	1.029	0.002	1.208	-0.004	0.267	0.070	1.143	0.002	1.334	-0.005	0.266
	0.5	0.000	0.273	0.053	0.852	0.011	1.875	-0.002	0.213	0.064	0.930	0.012	2.091	-0.002	0.205
	0.9	0.000	0.277	0.039	0.423	0.020	2.256	-0.000	0.079	0.047	0.442	0.022	2.523	-0.000	0.053
5	-0.9	0.000	0.163	0.039	0.310	- 0.013	0.071	-0.021	0.034	0.019	0.314	-0.013	0.069	-0.021	0.031
	-0.5	0.000	0.164	0.024	0.576	- 0.007	0.199	-0.007	0.124	0.024	0.587	-0.007	0.200	-0.007	0.123
	-0.1	0.000	0.166	0.025	0.664	- 0.001	0.557	-0.004	0.164	0.025	0.678	-0.001	0.568	-0.004	0.164
	0.1	0.000	0.167	0.025	0.655	0.001	0.775	-0.003	0.165	0.025	0.669	0.001	0.791	-0.003	0.165
	0.5	0.000	0.167	0.022	0.534	0.007	1.211	-0.001	0.127	0.023	0.544	0.007	1.240	-0.001	0.126
	0.9	0.000	0.166	0.016	0.253	0.013	1.451	-0.000	0.035	0.017	0.255	0.013	1.486	0.000	0.032
6	-0.9	0.000	0.341	0.016	0.540	-0.017	0.216	-0.029	0.161	0.039	0.647	-0.027	0.149	-0.045	0.065
	-0.5	0.000	0.344	0.032	0.906	-0.010	0.387	-0.010	0.288	0.050	1.207	-0.015	0.410	-0.015	0.258
	-0.1	0.000	0.345	0.034	1.024	-0.002	0.877	-0.005	0.343	0.052	1.388	-0.003	1.162	-0.008	0.342
	0.1	0.000	0.346	0.034	1.011	0.002	1.174	-0.004	0.343	0.052	1.367	0.003	1.618	-0.005	0.342
	0.5	0.000	0.345	0.031	0.845	0.010	1.773	-0.001	0.288	0.047	1.113	0.015	2.538	-0.002	0.258
	0.9	0.000	0.341	0.022	0.460	0.017	2.107	-0.000	0.161	0.035	0.524	0.027	3.053	0.000	0.065

Table 3. Bias and MSE of synthetic estimators for hypothetically drawn normal population

Domain	ρ_{yx}	Strategy-I								Strategy-II					
		$\beth_{m}^{s}$		$\beth_{r_1}^{s}$		$\beth_{p_1}^{s}$		$\beth_{3}^{s}$		$\beth_{r_2}^{s}$		$\beth_{p_2}^{s}$		$\beth_{4}^{s}$	
		Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	-0.9	-2.119	21.867	0.041	3.693	-0.034	1.253	-0.062	0.702	0.048	3.962	-0.040	1.075	-0.073	0.423
	-0.5	-2.119	21.867	0.014	2.737	-0.012	1.901	-0.020	1.755	0.017	2.831	-0.014	1.842	-0.023	1.669
	-0.1	-2.119	21.867	0.005	2.366	-0.002	2.222	-0.002	2.207	0.006	2.391	-0.002	2.222	-0.002	2.203
	0.1	-2.119	21.867	0.001	2.220	0.002	2.359	0.000	2.207	0.002	2.220	0.002	2.384	0.000	2.203
	0.5	-2.119	21.867	-0.005	1.938	0.010	2.648	-0.009	1.755	-0.006	1.885	0.012	2.726	-0.010	1.669
	0.9	-2.119	21.867	-0.014	1.509	0.023	3.169	-0.032	0.702	-0.017	1.378	0.028	3.341	-0.038	0.423
2	-0.9	0.415	13.416	0.041	2.670	-0.029	0.903	-0.052	0.507	0.048	2.864	-0.034	0.774	-0.062	0.305
	-0.5	0.415	13.416	0.014	1.977	-0.010	1.372	-0.017	1.267	0.017	2.045	-0.012	1.329	-0.020	1.205
	-0.1	0.415	13.416	0.005	1.708	-0.002	1.604	-0.002	1.593	0.006	1.727	-0.002	1.604	-0.002	1.590
	0.1	0.415	13.416	0.001	1.603	0.002	1.704	0.000	1.593	0.002	1.602	0.002	1.721	0.000	1.590
	0.5	0.415	13.416	-0.005	1.398	0.008	1.913	-0.007	1.267	-0.006	1.360	0.010	1.969	-0.009	1.205
	0.9	0.415	13.416	-0.014	1.089	0.020	2.289	-0.027	0.507	-0.017	0.994	0.023	2.414	-0.032	0.305
3	-0.9	-0.464	10.643	0.043	2.956	-0.031	1.002	-0.056	0.562	0.051	3.171	-0.036	0.860	-0.066	0.338
	-0.5	-0.464	10.643	0.015	2.190	-0.011	1.521	-0.018	1.405	0.018	2.265	-0.012	1.474	-0.021	1.336
	-0.1	-0.464	10.643	0.005	1.893	-0.002	1.778	-0.002	1.766	0.006	1.914	-0.002	1.778	-0.002	1.763
	0.1	-0.464	10.643	0.001	1.777	0.002	1.888	0.000	1.766	0.002	1.776	0.002	1.908	0.000	1.763
	0.5	-0.464	10.643	-0.006	1.550	0.009	2.120	-0.008	1.405	-0.007	1.508	0.011	2.182	-0.009	1.336
	0.9	-0.464	10.643	-0.015	1.207	0.021	2.536	-0.029	0.562	-0.018	1.102	0.025	2.674	-0.034	0.338
4	-0.9	0.372	9.012	0.041	2.631	-0.029	0.892	-0.053	0.500	0.048	2.823	-0.034	0.765	-0.062	0.301
	-0.5	0.372	1.585	0.014	1.949	-0.010	1.354	-0.017	1.250	0.017	2.016	-0.012	1.311	-0.020	1.189
	-0.1	0.372	1.585	0.005	1.685	-0.002	1.583	-0.002	1.571	0.006	1.703	-0.002	1.582	-0.002	1.569
	0.1	0.372	1.585	-0.001	1.581	0.002	1.680	0.000	1.571	0.002	1.581	0.002	1.698	0.000	1.569
	0.5	0.372	1.585	-0.005	1.380	0.009	1.886	-0.007	1.250	-0.006	1.342	0.010	1.942	-0.009	1.189
	0.9	0.372	1.585	0.014	1.074	0.020	2.257	-0.027	0.500	-0.017	0.981	0.024	2.380	-0.032	0.301
5	-0.9	0.374	5.960	0.000	2.597	-0.029	0.883	-0.053	0.494	0.000	2.785	-0.034	0.757	-0.062	0.298
	-0.5	0.374	5.960	0.000	1.925	-0.010	1.338	-0.017	1.235	0.000	1.991	-0.012	1.296	-0.020	1.174
	-0.1	0.374	5.960	0.000	1.664	-0.002	1.564	-0.002	1.553	0.000	1.682	-0.012	1.563	-0.002	1.550
	0.1	0.374	5.960	0.000	1.562	0.002	1.660	0.000	1.553	0.000	1.562	0.012	1.677	0.000	1.550
	0.5	0.374	5.960	0.000	1.364	0.009	1.863	-0.007	1.235	0.000	1.327	0.010	1.918	-0.009	1.174
	0.9	0.374	5.960	0.000	1.062	0.020	2.229	-0.027	0.494	0.000	0.970	0.024	2.350	-0.032	0.298
6	-0.9	0.356	9.423	0.041	2.639	-0.029	0.888	-0.052	0.500	0.050	2.830	-0.030	0.760	-0.060	0.300
	-0.5	0.356	9.423	0.014	1.951	-0.010	1.353	-0.017	1.250	0.020	2.020	-0.010	1.310	-0.020	1.190
	-0.1	0.356	9.423	0.005	1.685	-0.002	1.583	-0.002	1.571	0.010	1.700	0.000	1.580	0.000	1.570
	0.1	0.356	9.423	0.001	1.581	0.002	1.681	0.000	1.571	0.000	1.580	0.000	1.700	0.000	1.570
	0.5	0.356	9.423	-0.005	1.379	0.008	1.887	-0.007	1.250	-0.010	1.340	0.010	1.940	-0.010	1.190
	0.9	0.356	9.423	-0.014	1.073	0.020	2.259	-0.027	0.500	-0.020	0.980	0.020	2.380	-0.030	0.300

Table 4. Bias and MSE of synthetic estimators for hypothetically drawn exponential population

Domain	ρ_{yx}	Strategy-I								Strategy-II					
		$\sqsupset_{m}^{s}$		$\sqsupset_{r_1}^{s}$		$\sqsupset_{p_1}^{s}$		$\sqsupset_{3}^{s}$		$\sqsupset_{r_2}^{s}$		$\sqsupset_{p_2}^{s}$		$\sqsupset_{4}^{s}$	
		Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	-0.9	-0.262	18.335	0.158	14.927	-0.045	3.188	-0.046	0.543	0.187	17.347	-0.053	3.457	-0.054	0.327
	-0.5	-0.262	18.335	0.023	3.803	-0.009	1.486	-0.009	1.358	0.027	4.184	-0.011	1.443	-0.011	1.291
	-0.1	-0.262	18.335	0.008	2.353	-0.001	2.024	-0.002	1.707	0.010	2.469	-0.002	2.079	-0.002	1.704
	0.1	-0.262	18.335	0.005	1.953	0.001	2.251	0.000	1.707	0.006	1.996	0.001	2.348	0.000	1.704
	0.5	-0.262	18.335	0.000	1.361	0.005	2.678	-0.001	1.358	-0.001	1.295	0.006	2.853	-0.001	1.291
	0.9	-0.262	18.335	-0.005	0.826	0.009	3.230	-0.005	0.543	-0.005	0.662	0.011	3.506	-0.006	0.327
2	-0.9	1.084	15.732	0.155	14.308	-0.044	3.056	-0.045	0.521	0.183	16.627	-0.052	3.314	-0.053	0.314
	-0.5	1.084	15.732	0.023	3.644	-0.009	1.424	-0.009	1.302	0.027	4.010	-0.010	1.383	-0.010	1.238
	-0.1	1.084	15.732	0.008	2.255	-0.001	1.940	-0.002	1.636	0.010	2.366	-0.001	1.993	-0.002	1.634
	0.1	1.084	15.732	0.005	1.872	0.001	2.157	0.000	1.636	0.005	1.913	0.001	2.250	0.000	1.634
	0.5	1.084	15.732	0.000	1.305	0.005	2.567	-0.001	1.302	-0.001	1.242	0.006	2.735	-0.001	1.238
	0.9	1.084	15.732	-0.004	0.792	0.009	3.096	-0.005	0.521	-0.005	0.635	0.011	3.361	-0.006	0.314
3	-0.9	-0.674	11.960	0.159	15.090	-0.045	3.222	-0.046	0.549	0.188	17.537	-0.054	3.494	-0.054	0.331
	-0.5	-0.674	11.960	0.023	3.845	-0.009	1.503	-0.009	1.373	0.028	4.231	-0.011	1.459	-0.011	1.306
	-0.1	-0.674	11.960	0.009	2.380	-0.001	2.046	-0.002	1.726	0.010	2.497	-0.002	2.102	-0.002	1.724
	0.1	-0.674	11.960	0.005	1.975	0.001	2.276	0.000	1.726	0.006	2.018	0.001	2.374	0.000	1.724
	0.5	-0.674	11.960	0.000	1.377	0.005	2.708	-0.001	1.373	-0.001	1.310	0.006	2.886	-0.001	1.306
	0.9	-0.674	11.960	-0.005	0.836	0.009	3.266	-0.005	0.549	-0.005	0.670	0.011	3.546	-0.006	0.331
4	-0.9	1.099	10.549	0.154	14.268	-0.044	3.046	-0.045	0.519	0.183	16.581	-0.052	3.302	-0.053	0.313
	-0.5	1.099	10.549	0.023	3.637	-0.009	1.421	-0.009	1.299	0.027	4.001	-0.010	1.380	-0.010	1.235
	-0.1	1.099	10.549	0.008	2.251	-0.001	1.936	-0.002	1.633	0.010	2.361	-0.001	1.988	-0.002	1.630
	0.1	1.099	10.549	0.005	1.868	0.001	2.153	0.000	1.633	0.005	1.909	0.001	2.246	0.000	1.630
	0.5	1.099	10.549	0.000	1.302	0.005	2.561	-0.001	1.299	-0.001	1.239	0.006	2.729	-0.001	1.235
	0.9	1.099	10.549	-0.004	0.790	0.009	3.089	-0.005	0.519	-0.005	0.634	0.011	3.353	-0.006	0.313
5	-0.9	-0.467	6.342	0.000	14.961	-0.045	3.192	-0.046	0.545	0.000	17.387	-0.054	3.461	-0.054	0.328
	-0.5	-0.467	6.342	0.000	3.815	-0.009	1.491	-0.009	1.363	0.000	4.197	-0.011	1.448	-0.011	1.296
	-0.1	-0.467	6.342	0.000	2.361	-0.001	2.030	-0.002	1.713	0.000	2.477	-0.002	2.086	-0.002	1.710
	0.1	-0.467	6.342	0.000	1.960	0.001	2.258	0.000	1.713	0.000	2.003	0.001	2.356	0.000	1.710
	0.5	-0.467	6.342	0.000	1.366	0.005	2.687	-0.001	1.363	0.000	1.300	0.006	2.863	-0.001	1.296
	0.9	-0.467	6.342	0.000	0.829	0.009	3.241	-0.005	0.545	0.000	0.665	0.011	3.518	-0.006	0.328
6	-0.9	-0.420	9.880	0.158	14.982	-0.045	3.202	-0.046	0.545	0.187	17.412	-0.054	3.472	-0.054	0.328
	-0.5	-0.420	9.880	0.023	3.815	-0.009	1.491	-0.009	1.362	0.027	4.197	-0.011	1.448	-0.011	1.295
	-0.1	-0.420	9.880	0.008	2.361	-0.001	2.030	-0.002	1.713	0.010	2.477	-0.002	2.086	-0.002	1.710
	0.1	-0.420	9.880	0.005	1.960	0.001	2.258	0.000	1.713	0.006	2.002	0.001	2.355	0.000	1.710
	0.5	-0.420	9.880	0.000	1.366	0.005	2.687	-0.001	1.362	-0.001	1.300	0.006	2.863	-0.001	1.295
	0.9	-0.420	9.880	-0.005	0.829	0.009	3.240	-0.005	0.545	-0.005	0.664	0.011	3.518	-0.006	0.328

6. REAL DATA APPLICATION

In terms of administrative subdivisions, the state of Uttar Pradesh is similar to the majority of the other Indian states which are divided into a several districts for revenue collection and other administrative duties. Additionally, each district is divided into several tehsils, and each tehsil is further subdivided into several blocks. In this study, blocks are considered as small domains. Since the production of any crop depends upon area under production. Therefore, for real data application, we consider the crop production estimation problem for different blocks of Agra district of Uttar Pradesh. We consider 15 blocks of Agra district as small domains. The Bajra crop production (in tonnes) for the agricultural season 2022-2023 is considered as study variable y, whereas the area under production (in hectare) of Bajra crop for the agricultural season 2022-2023 is considered as auxiliary variable x. Various information regarding the blocks of Agra district is reported in Table 5, whereas the parameters of each domain are reported in Table 6 for ready reference.

Table 5. Total production and area under Bajra crop in Blocks of Agra district for agricultural season 2022-23

Sr. No.	Blocks of Agra district (a)	Numbers Of villages in Blocks (N_a)	Totals production (in tonne) under the Bajra crop in 2022-23 (Y_a)	Total area (in hectare) under the Bajra crop in 2022-23 (X_a)
1	Achhnera	50	13509	5972
2	Akola	38	11472	4865
3	Baah	50	22975	9689
4	Bichpuri	22	6626	22799
5	Baroli aheer	51	20094	8315
6	Eatmadpur	46	21604	9007
7	Fatehabad	70	39904	16361
8	Jagner	28	7680	3968
9	Jaitpur Kalan	44	17675	7866
10	Khandoli	41	15674	6822
11	Khera Garh	36	14527	7696
12	Pinahat	35	21950	9630
13	Sainya	44	15183	7697
14	Seekri	50	14589	6344
15	Shamshabad	58	25917	10720
Total		663	269379	117751

From the domain sizes (N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8, N_9, N_{10}, N_{11}, N_{12}, N_{13}, N_{14}, N_{15}) = (50, 38, 50, 22, 51, 46, 70, 28, 44, 41, 36, 35, 44, 50, 58) mentioned in Table 5, we have selected samples (n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8, n_9, n_{10}, n_{11}, n_{12}, n_{13}, n_{14}, n_{15}) = (10, 8, 10, 4, 10, 9, 14, 6, 9, 8, 7, 7, 9, 10, 12), respectively. Out of these selected samples, the responding units are (r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9, r_{10}, r_{11}, r_{12}, r_{13}, r_{14}, r_{15}) =(8, 6, 7, 3, 7, 7, 11, 5, 7, 6, 6, 6, 7, 8, 9), respectively. Using the domain parameters given in Table 6, we have computed the bias and MSE of the proposed direct and synthetic methods with respect to the direct and synthetic mean imputation methods, respectively.

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	50	270.18	119.44	260.43	109.33	0.996
2	38	301.89	128.03	217.10	88.79	0.992
3	50	459.50	193.78	265.07	113.58	0.997
4	22	301.18	127.23	467.46	203.32	0.998
5	51	394.00	163.04	330.73	133.91	0.991
6	46	469.65	195.80	450.43	188.92	0.997
7	70	570.06	233.73	420.02	173.15	0.993
8	28	274.28	141.71	283.64	146.56	0.989
9	44	401.70	178.77	253.80	114.88	0.990
10	41	382.29	166.39	336.17	150.32	0.992
11	36	403.53	213.78	370.22	192.73	0.983
12	35	627.14	275.14	484.95	209.78	0.996
13	44	345.06	174.93	228.79	113.68	0.975
14	50	291.78	126.88	226.34	94.99	0.995
15	58	446.84	184.83	304.09	123.45	0.991

The results based on the real data for direct and synthetic estimators are reported in Tables 7 and 8, respectively, which show the dominance of the proposed general class of direct and synthetic imputation methods over the proposed mean, ratio and product imputation methods. Further, Tables 7 and 8 show that both bias and MSE are higher for direct estimators compared to synthetic estimators, except for the traditional mean estimator. Thus, the synthetic estimators outperform the direct estimators, except the mean estimator.

7. CONCLUSIONS

In the current article, we have proposed mean, ratio, product and general class of imputation methods for the estimation of domain mean in case of missing data using simple random sampling. The equations of biases and MSE are determined till first order approximation. The efficiency conditions are obtained by comparing the MSE expressions. Further, to examine the performance of the proposed imputation methods, an extensive simulation is performed using hypothetically rendered normal and exponential populations. It has been observed from the simulation results that the proposed general class of imputation methods dominate the mean, ratio, and product imputation methods for the different amount of correlation coefficient in each domain. The applicability of proposed imputation methods is shown through a real data application based on Bajra crop production in Agra district of Uttar Pradesh, India. As a result, survey practitioners may be urged to use the proposed imputation approaches wherever missing data is noticed.

While the proposed methods show clear advantages, certain limitations merit acknowledgement. The study is based on the MCAR mechanism, therefore, deviations from this assumption may influence the robustness of the estimators. Moreover, the analysis is restricted to SRS, whereas many real surveys employ complex designs. These limitations open several

Table 7. Bias and MSE of direct estimators for real population

mains	$\beth_m^d$		Strategy I						Strategy II					
			$\beth_{r_1}^d$		$\beth_{p_1}^d$		$\beth_1^d$		$\beth_{r_2}^d$		$\beth_{p_2}^d$		$\beth_2^d$	
	Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	0.000	7121.622	-0.887	1749.024	18.998	22280.340	-2.729	1737.275	-1.165	70.087	24.935	27017.440	-3.582	54.667
2	0.000	6615.244	-0.421	2034.614	14.752	19848.43	-2.008	2030.875	-0.599	100.57	20.980	25435.77	-2.856	95.253
3	0.000	8632.135	0.234	3044.785	12.394	25825.15	-1.464	3042.784	0.360	51.562	19.034	35035.700	-2.248	48.490
4	0.000	62907.370	4.588	18285.700	152.743	202299.400	-17.439	18245.400	6.457	106.495	214.972	259088.700	-24.544	49.789
5	0.000	13480.900	-0.279	4842.518	21.646	38956.49	-2.812	4841.084	-0.428	237.388	33.185	52537.550	-4.311	235.189
6	0.000	24573.710	0.343	6544.924	38.731	79304.980	-4.714	6543.507	0.465	140.876	52.488	98746.210	-6.389	138.957
7	0.000	13517.520	0.217	3573.633	17.661	43843.670	-2.128	3572.134	0.291	183.670	23.681	54182.130	-2.853	181.660
8	0.000	13217.320	0.395	2897.096	38.021	44611.640	-4.606	2895.981	0.496	270.130	47.699	52602.930	-5.779	268.731
9	0.000	7738.423	0.389	2164.993	14.273	25091.340	-1.643	2157.271	0.529	158.255	19.399	31329.080	-2.234	152.612
10	0.000	16079.450	1.068	4894.586	30.325	51267.060	-3.408	4880.699	1.510	262.471	42.884	65839.700	-4.820	242.832
11	0.000	19036.790	-0.044	3769.700	37.790	64766.550	-4.740	3769.679	-0.053	610.992	45.608	74227.88	-5.721	610.967
12	0.000	32476.400	-0.452	5779.133	42.118	111435.200	-5.437	5776.062	-0.546	217.202	50.893	127884.900	-6.569	213.492
13	0.000	6288.304	0.059	1885.061	12.820	19579.99	-1.580	1884.967	0.081	303.489	17.425	24354.150	-2.148	303.36
14	0.000	5379.404	-0.406	1324.907	13.490	17068.84	-1.845	1321.228	-0.533	57.877	17.705	20721.79	-2.421	53.048
15	0.000	8680.483	-0.139	2668.37	13.315	26467.91	-1.717	2667.712	-0.198	141.54	18.912	33943.79	-2.439	140.606

Table 8. Bias and MSE of synthetic estimators for real population

Domains	$\beth_m^s$		Strategy I						Strategy II					
			$\beth_{r_3}^s$		$\beth_{p_3}^s$		$\beth_3^s$		$\beth_{r_4}^s$		$\beth_{p_4}^s$		$\beth_4^s$.	
	Bias	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**	**Bias**	**MSE**
1	136.123	19544.030	0.003	139.456	1.147	1379.469	-0.142	139.454	0.004	11.836	1.621	1763.742	-0.201	11.833
2	104.408	11915.630	0.004	174.117	1.282	1722.332	-0.159	174.114	0.005	14.778	1.811	2202.113	-0.225	14.774
3	-53.197	3844.414	0.005	403.368	1.951	3990.035	-0.242	403.362	0.008	34.236	2.757	5101.52	-0.342	34.227
4	105.121	12065.010	0.004	173.296	1.279	1714.207	-0.159	173.293	0.005	14.709	1.807	2191.725	-0.224	14.705
5	12.303	1165.878	0.005	296.567	1.673	2933.581	-0.207	296.562	0.007	25.172	2.364	3750.775	-0.293	25.164
6	-63.349	5027.607	0.006	421.389	1.995	4168.294	-0.247	421.382	0.008	35.766	2.818	5329.436	-0.349	35.756
7	-163.745	27829.880	0.007	620.823	2.421	6141.048	-0.300	620.812	0.009	52.693	3.420	7851.730	-0.424	52.678
8	132.017	18443.120	0.003	143.727	1.165	1421.713	-0.144	143.724	0.005	12.199	1.646	1817.753	-0.204	12.196
9	4.599	1035.658	0.005	308.279	1.706	3049.434	-0.211	308.274	0.007	26.166	2.410	3898.899	-0.299	26.158
10	24.010	1591.014	0.004	279.205	1.624	2761.835	-0.201	279.200	0.006	23.698	2.294	3531.185	-0.284	23.691
11	2.775	1022.213	0.005	311.084	1.714	3077.178	-0.212	311.079	0.007	26.404	2.421	3934.372	-0.300	26.396
12	-220.840	49784.68	0.007	751.387	2.663	7432.564	-0.330	751.375	0.010	63.775	3.763	9503.017	-0.466	63.757
13	61.235	4764.234	0.004	227.479	1.465	2250.171	-0.182	227.475	0.006	19.308	2.070	2876.99	-0.257	19.302
14	114.523	14130.070	0.003	162.645	1.239	1608.854	-0.154	162.643	0.005	13.805	1.751	2057.025	-0.217	13.801
15	-40.542	2658.137	0.005	381.456	1.898	3773.281	-0.235	381.449	0.007	32.377	2.681	4824.386	-0.332	32.367

promising avenues for future research. Extending the proposed imputation methods to alternative missing-data mechanisms such as MAR and MNAR, and to more complex sampling designs, would significantly broaden their applicability.

Disclosure statement

No potential conflict of interest was reported by the authors.

REFERENCES

Ahmed, M.S., Al-Titi, O., Al-Rawi, Z. and Abu-Dayyeh, W.(2006). Estimation of a population mean using different using different imputation methods. Statistics in Transition, **7**(6), 1247-1264.

Ashutosh (2020). Estimation of domain mean using conventional synthetic estimator with two auxiliary characters. Annals of Data Science, **10**(1), 153-166.

Bhushan, S. and Pandey, A.P. (2016). Optimal imputation of missing data for estimation of population mean. Journal of Statistics and Management System, **19**(6),755-769.

Bhushan, S. and Pandey, A.P. (2018). Optimality of ratio type estimation methods for population mean in presence of missing data. Communication in Statistics- Theory and Methods, **47**(11), 2576-2589.

Bhushan, S. and Pandey, A.P. (2020). Optimal imputation of missing data using multi auxiliary information. Computational Statistics, **36**(1), 449-477.

Bhushan, S. and Pandey, A.P. (2021). Optimality of ratio-type imputation methods for estimation of population mean using higher order moment of an auxiliary variable. Journal of Statistical Theory and Practice, 15, 48.

Bhushan, S., Kumar, A., Pandey, A.P. and Singh, S. (2022). Estimation of population mean in presence of missing data under simple random sampling. Communications in Statistics- Simulation and Computation, **52**(12), 6048-6069.

Heitjan, D.F. and Basu S. (1996). Distinguishing 'Missing at Random' and 'Missing Completely at Random'. The American Statistician, **50**, 207-213.

Khare, B.B. and Ashutosh (2017). Synthetic estimator for domain mean using auxiliary character. International Journal of Statistics and Economics, **18**(4), 32-44.

Khare,B.B., Ashutosh and Khare, S.(2018). Modified direct regression estimator for domain mean using auxiliary character. International Journal of Applied Mathematics and Statistics, **57**(6), 10-20.

Lee, H., Rancourt, E. and Sarndal, C.E.(1994). Experiments with variance estimation from survey data with imputed values. Journal of Official Statistics, **10**, 231-243.

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.

Panse, V.G., Rajagopalan, M. and Pillai, S.S.(1966). Estimation of crop yields for small areas. Biometrics, 374-384.

Prasad, S.(2017). A study on new methods of ratio exponential type imputation in sample surveys. Hecettepe J Math Stat., **47**(5), 1281-1301.

Rubin, R.B.(1976). Inference and missing data. Biometrika, **63**(3), 581-592.

Rueda, M., Gonzalez, S. and Arcos, A. (2005). Indirect methods of imputation of missing data based on available units. Appl. Math. Computat., **164**, 249-261.

Sharma, S.D., Srivastava, A.K. and Sud, U.C.(2004). Small area crop estimation for crop yield estimators at Gram Panchayat level. Journal of Indian Society of Agricultural Statistics, **57,** 26-37.

Sharma, M.K. and Sisodia, B.V.S.(2016). Estimation of crop production for smaller geographical area: An application of discriminant function analysis. Journal of Statistical Theory and Practice, **10**, 444-455.

Singh, H.P. and Horn, S. (1998). An alternative estimator for multi-character surveys. Metrika, **48**, 99-107.

Singh, R. and Goel, R.C. (2000). Use of remote sensing satellite data in crop survey. Technical report, Indian Agricultural Statistics Research Institute, New Delhi.

Sisodia, B.V.S. and Singh, A. (2001). On small area estimation-An empirical study. Journal of Indian Society Agricultural Statistics, **54**(3), 303-316.

Sisodia, B.V.S. and Chandra, H. (2012). Estimation of crop production at smaller geographical level in India. Journal of the Indian Society of Agricultural Statistics, **66**(2), 313-319.

Srivastava, A.K., Ahuja, D.L., Bhatia, D.K. and Mathur, D.C.(1999). Small area estimation technique in agriculture. Souvenir in Advances in Agricultural Statistics and Computer Application, Indian Agricultural Statistics Research Institute, New Delhi, 32-45

Singh, S. and Deo, B. (2003). Imputation by power transformation. Statistical Papers, **44**, 555-579.

Singh, S. (2009). A new method of imputation in survey sampling: A Journal of Theoretical and Applied Statistics, **43**(5), 499-511.

Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. Metrika, **51**, 267-276.

Tikkiwal, B.D. (1993). Modeling through survey data for small domains. Proceedings of International Scientific Conference on small Area Statistics and Survey Design (An invited paper), held in September 1992 at Warsaw, Poland.

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.

Toutenburg, H. and Srivastava, V.K. (1998). Estmation of ratio of population means in survey sampling when some observations are missing. Metrika, **48**, 177-187.

Toutenburg, H. and Srivastava, V.K. and Shalabh (2008). Amputation versus imputation of missing values through ratio method in sample surveys. Statistical Papers, 49, 237-247.