Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

Maejo International
Journal of Science and Technology

e- ISSN 2697-4746
Available online at www.mijst.mju.ac.th

Full Paper

Minimum-covariance-determinant-based mean estimators under systematic sampling

Abdullah Mohammed Alomair [1], Usman Shahzad [2, *], Nadia Hashim Al–Noor [3] and Mohammed Ahmed Alomair [1]

[1] Department of Quantitative Methods, School of Business, King Faisal University, Al-Ahsa 31982, Saudi Arabia
[2] Department of Management Science, College of Business Administration, Hunan University, Changsha 410082, China
[3] Department of Mathematics, College of Science, Mustansiriyah University, Baghdad 10011, Iraq
* Corresponding author, e-mail: usman.stat@yahoo.com

Received: 23 November 2024 / Accepted: 16 April 2025 / Published: 21 April 2025

Abstract: The ordinary least square (OLS) estimates become inappropriate in the presence of outliers and consequently the mean estimators based on the OLS coefficients also become unsuitable. To address this issue, employing robust regression tools and covariance matrices for mean estimation is a well-established practice under a simple random sampling scheme. However, the mean estimation using robust regression tools and covariance matrices under systematic sampling has not been explored yet. To fill this gap, we develop a class of mean estimators under systematic sampling in this article. This study proposes a family of minimum-covariance-determinant-based mean estimators along with their theoretical properties. The Dixon test is used to confirm the presence of extremes in the data. Numerical analyses related to real and simulated data are performed to assess the new proposals. According to the percentage of relative efficiency in the practical study with timber volume data, estimators based on the Huber regression show a percentage relative efficiency (PRE) increase from 226.16 to 241.68. In simulated data scenarios, the PRE of various robust estimators exceeds 450. Hence the proposed estimators provide excellent performance.

Keywords: minimum covariance determinant, mean estimation, robust regression, systematic random sampling

---

INTRODUCTION

The statistical procedure of selecting a sample of components to survey a target population is known as survey sampling. The word ‘survey’ can be used to describe a variety of analytical methods. Various methods of contacting selected sample representatives are the subject of survey data gathering.

The goal of sampling is to minimise the cost or effort needed to sample the entire target population. Although there are numerous methods for obtaining a sample in both probability and non-probability sampling systems, systematic sampling is the focus of this paper.

In applied research, among the many important examples of surveys and sampling procedures are the European Union Labour Force Survey which offers quarterly data on job participation, the Public Service Employee Survey which measures federal government employees' opinions about their leadership, workplace, workforce and reimbursement, and the National Co-morbidity Survey to evaluate anxiety's levels and physical disturbance. The uppermost aim of sampling activities remains to estimate population parameters or trends according to the ongoing study's mandate. One of the most effective approaches to meeting the challenge of obtaining more reliable sample estimates is to use supplementary (additional or auxiliary) information thoughtfully. In survey sampling, mean estimation is one of the major concerns that can be enhanced by using supplementary information.

Kadilar and Cingi [1], Abd-Elfattah et al. [2] and Koyuncu [3] have developed some classes of estimators using supplementary information under a simple random sampling scheme. Kadilar and Cingi [1] pioneered the development of regression-type-ratio estimators based on the ordinary least square (OLS) regression coefficient. However, the OLS estimate becomes inappropriate in the presence of extremes or outliers. To solve this issue, there are some modifications available in the literature such as those of Kadilar et al. [4] who used Huber regression instead of OLS regression coefficient. Abid et al. [5-7] used non-traditional measures of location and dispersion with OLS regression coefficient. Irfan et al. [8] used the median of a study variable and an auxiliary variable. Zaman and Bulut [9] and Zaman [10] extended the work of Kadilar et al. [4] and used least absolute deviation, Huber-M and some other robust regression tools. Bulut and Zaman [11] extended the work by introducing ratio-type mean estimators using minimum covariance determinant (MCD). It is worth mentioning that the algorithm by Rousseeuw and Van Driessen [12] made the MCD computationally practical while Hubert et al. [13] developed even faster algorithms.

Using kernels, the MCD has also been extended to non-elliptical distributions [14] and high-dimensional data sets [15-17]. Shahzad et al. [18] utilised the quantile regression with MCD-based measures of location to introduce a class of quantile-regression-type mean estimators. Bulent [19] combined bootstrapping and MCD robust estimator to offer some benefits for enhanced diagnostics and outlier identification. Alomair and Shahzad [20] used a calibrated MCD to optimise mean estimators in median-ranked set sampling. Anas et al. [21] provided quantile regression coefficients by including the Sarndal idea. Abid et al. [22] used dual auxiliary variable-based exponential-cum-ratio estimators that combined conventional and non-traditional measurements. Zaman and Bulut [23] used robust covariance matrices in mean estimation. Zaman [24] developed robust ratio-type mean estimators by using two tuning parameters. Raymaekers and Rousseeuw [25] proposed a cellwise robust version of the MCD approach by concentrating on single-class multivariate numerical data without a response variable. Further, the works of Hubert and Debruyne [26], Al-Noor and Mohammed [27], Hubert et al. [28], Bhushan and Gupta [29], Subzar et al. [30], Grover and Kaur [31], Koç and Koç [32], Malik et al. [33], Kumar and Siddiqui [34], Bhushan and Kumar [35], and Tian and Qin [36] are among the other important works that interested readers might review.

All of the above-mentioned works were done using a simple random sampling scheme or median-ranked set sampling. However, no significant development is considered under systematic sampling with MCD. Therefore, by exploring this gap and drawing inspiration from the work of Shahzad et al. [37, 38], we are introducing the MCD-based mean estimators under systematic sampling.

The systematic sampling technique is a probability sampling technique where elements are taken from a population at equal intervals from a random point. This strategy is used when the population structures are even or inhabitants are evenly distributed to ensure even coverage without clustering residents. It is useful for testing product quality control examination or sampling the workforce. Systematic sampling is more efficient and cost-effective compared to simple random sampling when the sampling frame is easily accessible. It is particularly convenient for studies involving periodic monitoring or spatial analysis where balance within the population is critical.

In the next sections we briefly present the mean estimates from MCD-based estimators by Bulut and Zaman [11] in the context of systematic sampling. In addition, we propose MCD-based regression-type estimators under systematic sampling. Efficiency comparisons are provided based on practical and quantified simulation studies.

METHODS

Adapted Estimators in Systematic Sampling

Bulut and Zaman [11] developed MC-based mean estimators under simple random sampling. We adapt their class of estimators under systematic sampling as given below:

$$T_{uj} = \frac{\bar{y}_s + b_{rob}(\bar{X} - \bar{x}_s)}{F_j \bar{x}_s + G_j}\left(F_j \bar{X} + G_j\right) \text{ for } j = 1,2,\ldots,35 \tag{1}$$

All the 35 family members of Bulut and Zaman [11] are provided in Table 1, where

b_{h-rrc} = Huber regression coefficient for $j = 1, \ldots, 5$,
$b_{lad-rrc}$ = Least absolute deviation (LAD) regression coefficient for $j = 6, \ldots, 10$,
$b_{lms-rrc}$ = Least median of square (LMS) regression coefficient for $j = 11, \ldots, 15$,
$b_{lts-rrc}$ = Least trimmed square (LTS) regression coefficient for $j = 16, \ldots, 20$,
$b_{hpl-rrc}$ = Hample regression coefficient for $j = 21, \ldots, 25$,
$b_{tky-rrc}$ = Tukey regression coefficient for $j = 26, \ldots, 30$,
$b_{hmm-rrc}$ = Huber method of moments (HMM) regression coefficient for $j = 31, \ldots, 35$.

The mean square error (MSE) of T_{uj} family is given below:

$$\text{MSE}(T_{uj}) = \frac{1-f}{n}\left[\vartheta_y S_y^2 + \left(K_{uj} + B_{lad-rrc}\right)^2 \vartheta_x S_x^2 - 2\left(K_{uj} + B_{lad-rrc}\right)\vartheta_x t^* S_{xy}\right] \text{ for } j = 1, \ldots, 5 \tag{2}$$

where $\vartheta_x = 1 + (n-1)r_x$, $\vartheta_y = 1 + (n-1)r_y$, $t^* = \sqrt{\frac{\vartheta_y}{\vartheta_x}}$. Further, $\left(S_x^2, S_y^2\right)$ are the unbiased variances of (X, Y), S_{xy} represents covariance, and $\left(r_x, r_y\right)$ are the intra-class correlations of (X, Y). Note that the intra-class correlation $\frac{S_B^2}{S_B^2 + S_W^2}$ can be calculated using within-group and between-group variances (S_B^2, S_W^2) of any variable. Also, $G_c = C_x$ and $G_b = \beta_{2x}$ are the coefficients of variation and

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

kurtosis, and $K_{uj} = \frac{F_j\bar{Y}}{F_j\bar{X}+G_j}$ for $j = 1,2,\dots,35$. It is worth mentioning that all characteristics are calculated through MCD.

Table 1. Adapted estimators from Bulut and Zaman [11]

Estimator	b_{rob}	F_{j1}	G_{j1}
T_{u1}	b_{h-rrc}	1	0
T_{u2}	b_{h-rrc}	1	G_c
T_{u3}	b_{h-rrc}	1	G_b
T_{u4}	b_{h-rrc}	G_b	G_c
T_{u5}	b_{h-rrc}	G_c	G_b
T_{u6}	$b_{lad-rrc}$	1	0
T_{u7}	$b_{lad-rrc}$	1	G_c
T_{u8}	$b_{lad-rrc}$	1	G_b
T_{u9}	$b_{lad-rrc}$	G_b	G_c
T_{u10}	$b_{lad-rrc}$	G_c	G_b
T_{u11}	$b_{lms-rrc}$	1	0
T_{u12}	$b_{lms-rrc}$	1	G_c
T_{u13}	$b_{lms-rrc}$	1	G_b
T_{u14}	$b_{lms-rrc}$	G_b	G_c
T_{u15}	$b_{lms-rrc}$	G_c	G_b
T_{u16}	$b_{lts-rrc}$	1	0
T_{u17}	$b_{lts-rrc}$	1	G_c
T_{u18}	$b_{lts-rrc}$	1	G_b
T_{u19}	$b_{lts-rrc}$	G_b	G_c
T_{u20}	$b_{lts-rrc}$	G_c	G_b
T_{u21}	$b_{hpl-rrc}$	1	0
T_{u22}	$b_{hpl-rrc}$	1	G_c
T_{u23}	$b_{hpl-rrc}$	1	G_b
T_{u24}	$b_{hpl-rrc}$	G_b	G_c
T_{u25}	$b_{hpl-rrc}$	G_c	G_b
T_{u26}	$b_{tky-rrc}$	1	0
T_{u27}	$b_{tky-rrc}$	1	G_c
T_{u28}	$b_{tky-rrc}$	1	G_b
T_{u29}	$b_{tky-rrc}$	G_b	G_c
T_{u30}	$b_{tky-rrc}$	G_c	G_b
T_{u31}	$b_{hmm-rrc}$	1	0
T_{u32}	$b_{hmm-rrc}$	1	G_c
T_{u33}	$b_{hmm-rrc}$	1	G_b
T_{u34}	$b_{hmm-rrc}$	G_b	G_c
T_{u35}	$b_{hmm-rrc}$	G_c	G_b

Proposed Robust Regression-Type Estimators

Taking motivation from Bulut and Zaman [11], we define the following class of MCD-based regression-type estimators under systematic random sampling:

$$T_{vi} = \bar{y}_s + b_{rob}(\bar{X} - \bar{x}_s) \text{ for } i = 1,2,\ldots,7$$

where $(\bar{x}_s, \bar{y}_s)$ represents the systematic sample means of (X, Y). The derivation of the MSE of the T_{vi} family using Taylor-series is given below:

$$\text{MSE}(T_{vi}) = V(\bar{y}_s) - 2B_{rob}Cov(\bar{x}_s, \bar{y}_s) + B_{rob}^2 V(\bar{x}_s) \text{ for } i = 1,\ldots,7$$

Putting the values of $V(\bar{y}_s), V(\bar{x}_s)$ and $Cov(\bar{x}_s, \bar{y}_s)$, we get the finalised MSE expression as given below :

$$\text{MSE}(T_{vi}) = \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{rob}\vartheta_x t^* S_{xy} + B_{rob}^2 \vartheta_x S_x^2\right] \text{ for } i = 1,\ldots,7.$$

Now utilising 7 diverse robust regression tools and MCD estimation, we get the proposed class containing 7 members, each with their own MSE as follows:

$$T_{vi} = \begin{cases} [\bar{y}_s + b_{h-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 1, b_{rob} = b_{h-rrc} \\ [\bar{y}_s + b_{lad-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 2, b_{rob} = b_{lad-rrc} \\ [\bar{y}_s + b_{lms-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 3, b_{rob} = b_{lms-rrc} \\ [\bar{y}_s + b_{lts-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 4, b_{rob} = b_{lts-rrc} \\ [\bar{y}_s + b_{hpl-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 5, b_{rob} = b_{hpl-rrc} \\ [\bar{y}_s + b_{tky-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 6, b_{rob} = b_{tky-rrc} \\ [\bar{y}_s + b_{hmm-rrc}(\bar{X} - \bar{x}_s)] & \text{for } i = 7, b_{rob} = b_{hmm-rrc} \end{cases} . \quad (3)$$

The corresponding MSEs of T_{vi} members are given by

$$\text{MSE}(T_{vi}) = \begin{cases} \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{h-rrc}\vartheta_x t^* S_{xy} + B_{h-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 1, B_{rob} = B_{h-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{lad-rrc}\vartheta_x t^* S_{xy} + B_{lad-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 2, B_{rob} = B_{lad-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{lms-rrc}\vartheta_x t^* S_{xy} + B_{lms-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 3, B_{rob} = B_{lms-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{lts-rrc}\vartheta_x t^* S_{xy} + B_{lts-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 4, B_{rob} = B_{lts-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{hpl-rrc}\vartheta_x t^* S_{xy} + B_{hpl-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 5, B_{rob} = B_{hpl-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{tky-rrc}\vartheta_x t^* S_{xy} + B_{tky-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 6, B_{rob} = B_{tky-rrc} \\ \frac{1-f}{n}\left[\vartheta_y S_y^2 - 2B_{hmm-rrc}\vartheta_x t^* S_{xy} + B_{hmm-rrc}^2 \vartheta_x S_x^2\right] & \text{for } i = 7, B_{rob} = B_{hmm-rrc} \end{cases} . \quad (4)$$

It is worth noting that the proposed estimators are simple in nature compared to adapted estimators as the latter are based on both ratio and regression while the former are only based on the regression part.

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

Efficiency Comparisons

This section includes numerical analyses of both real and simulated data to evaluate the performance of the proposals. Efficiency comparisons are analysed using the percentage relative efficiency (PRE) criteria.

Population 1: practical study

For numerical illustration, the tree data are taken from Murthy [39], where X = strip length and Y = volume of timber. The size of the population is $N = 176$. Note that Murthy [39] also provided some values of intra-class correlation, i.e. $r_x = r_y = r_\omega$ for instance, with respect to different sample sizes as follows:

$r_\omega = -0.1510$ when $n = 4$,
$r_\omega = -0.1106$ when $n = 8$,
$r_\omega = -0.0522$ when $n = 16$,
$r_\omega = -0.0435$ when $n = 22$.

In light of all the aforementioned values, the performance of the adapted and proposed estimators is compared. For the presence of outliers in (X, Y) individually, Dixon test (DT) is used [40, 41], whose results are:

For X, DT $= 111.35$, P-value$= 2.2e - 16$,
For Y, DT $= 63.02$, P-value$= 1.998e - 15$.

The DT's significant results emphasise the presence of extremes in the data. As a result, the data are appropriate for handling the estimators under consideration.

Population 2: simulation study

In the simulation study the variable Xj follows a Gamma distribution with shape parameter of 2.7 and scale parameter of 3.9, i.e. $X_j \sim$ G$(2.7, 3.9)$. Further, Y_j is defined as $Y_j = a + RX_j + bX_j^p$ with $a = 6$, $R = 2.2$, $p = 1.7$, and b following a standard normal distribution with $N = 1000$. The systematic sampling was replicated 1000 times with $n = 100$.

The empirical MSE values of T_{uj} and T_{vi} are investigated as MSE$(t_i) = \frac{\sum_{i=1}^{K}(t_i - \bar{t})}{K}$, where t_i representing proposed and existing estimators. The PRE is computed as

$$PRE(t_i) = \frac{V(\bar{y}_s)}{MSE(t_i)} \times 100.$$

RESULTS AND DISCUSSION

The results of the PRE values related to population 1 are provided in Tables 2-5, and PRE values related to population 2 are provided in Table 6.

With different sample sizes and various values of intra-class correlation, the proposed estimators provide excellent performance, where their performance can be arranged respectively for populations 1 and 2 in the following manner:

$$T_{v3}\text{: LTS} > T_{v4}\text{: LMS} > T_{v5}\text{: Hample} > T_{v6}\text{: Tukey} > T_{v1}\text{: Huber} > T_{v7}\text{: HMM} > T_{v2}\text{: LAD}$$

$$T_{v6}\text{: Tukey} > T_{v3}\text{: LTS} > T_{v7}\text{: HMM} > T_{v2}\text{: LAD} > T_{v4}\text{: LMS} > T_{v1}\text{: Huber} > T_{v5}\text{: Hample.}$$

Thus, for real data, among the proposed estimators, the estimators T_{v3} and T_{v4} based on employing LTS and LMS record the highest efficiency. On the other hand, for simulated data, the estimators T_{v6} and T_{v3} based on employing Tukey and LTS record the highest efficiency. Further, in comparison to other robust regression estimators, the proposed estimators perform best or nearly so overall when examined on real and simulated data. Given these outcomes, the proposed estimators can be regarded as highly robust and effective estimators. These results are also provided in Figures 1-5.

Table 2. PRE of adapted and proposed estimators with $n = 4$ and $r_\omega = -0.1510$ in Population 1

Estimator	Adapted					Proposed
Huber	T_{u1} 226.1582	T_{u2} 241.6759	T_{u3} 419.8329	T_{u4} 521.1829	T_{u5} 520.5255	T_{v1} 461.8194
LAD	T_{u6} 231.2658	T_{u7} 247.1671	T_{u8} 427.3614	T_{u9} 518.9625	T_{u10} 522.4009	T_{v2} 454.9534
LTS	T_{u11} 214.6931	T_{u12} 229.3321	T_{u13} 401.9831	T_{u14} 524.3695	T_{u15} 514.1828	T_{v3} 477.0048
LMS	T_{u16} 217.6949	T_{u17} 232.5662	T_{u18} 406.7810	T_{u19} 523.7974	T_{u20} 516.1412	T_{v4} 473.0761
Hample	T_{u21} 222.5908	T_{u22} 237.8377	T_{u23} 414.4180	T_{u24} 522.4563	T_{u25} 518.8771	T_{v5} 466.5895
Tukey	T_{u26} 225.8080	T_{u27} 241.2993	T_{u28} 419.3071	T_{u29} 521.3184	T_{u30} 520.3762	T_{v6} 462.2888
HMM	T_{u31} 226.4114	T_{u32} 241.9483	T_{u33} 420.2124	T_{u34} 521.0836	T_{u35} 520.6317	T_{v7} 461.4798

Table 3. PRE of adapted and proposed estimators with $n = 8$ and $r_\omega = -0.1106$ in Population 1

Estimator	Adapted					Proposed
Huber	T_{u1} 250.6176	T_{u2} 269.3937	T_{u3} 499.0314	T_{u4} 627.5682	T_{u5} 633.5100	T_{v1} 534.5242
LAD	T_{u6} 256.8328	T_{u7} 276.1473	T_{u8} 509.2763	T_{u9} 623.4399	T_{u10} 635.4197	T_{v2} 524.6076
LTS	T_{u11} 237.1506	T_{u12} 254.7440	T_{u13} 475.5721	T_{u14} 634.1993	T_{u15} 626.3378	T_{v3} 556.2317
LMS	T_{u16} 237.9507	T_{u17} 255.6149	T_{u18} 477.0127	T_{u19} 633.9061	T_{u20} 626.8873	T_{v4} 554.9404
Hample	T_{u21} 246.2905	T_{u22} 264.6889	T_{u23} 491.6788	T_{u24} 630.0701	T_{u25} 631.6744	T_{v5} 541.4801
Tukey	T_{u26} 250.1924	T_{u27} 268.9315	T_{u28} 498.3168	T_{u29} 627.8283	T_{u30} 633.3484	T_{v6} 535.2062
HMM	T_{u31} 250.9253	T_{u32} 269.7281	T_{u33} 499.5474	T_{u34} 627.3782	T_{u35} 633.6244	T_{v7} 534.0310

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

Table 4. PRE of adapted and proposed estimators with $n = 16$ and $r_\omega = -0.0522$ in Population 1

Estimator			Adapted			Proposed
Huber	T_{u1} 266.7416	T_{u2} 287.2540	T_{u3} 550.3302	T_{u4} 666.4366	T_{u5} 684.9603	T_{v1} 556.0156
LAD	T_{u6} 273.5493	T_{u7} 294.6704	T_{u8} 561.5533	T_{u9} 660.4106	T_{u10} 685.4877	T_{v2} 544.7254
LTS	T_{u11} 247.6978	T_{u12} 266.4809	T_{u13} 516.4603	T_{u14} 679.6654	T_{u15} 678.1090	T_{v3} 588.4671
LMS	T_{u16} 255.5326	T_{u17} 275.0313	T_{u18} 530.8252	T_{u19} 674.9641	T_{u20} 681.9524	T_{v4} 575.0023
Hample	T_{u21} 620.0060	T_{u22} 282.0918	T_{u23} 542.2432	T_{u24} 670.2710	T_{u25} 684.0290	T_{v5} 563.9847
Tukey	T_{u26} 266.2761	T_{u27} 286.7467	T_{u28} 549.5453	T_{u29} 666.8273	T_{u30} 684.8899	T_{v6} 556.7950
HMM	T_{u31} 267.0785	T_{u32} 287.6211	T_{u33} 550.8967	T_{u34} 666.1521	T_{u35} 685.0084	T_{v7} 555.4521

Table 5. PRE of adapted and proposed estimators with $n = 22$ and $r_\omega = -0.0435$ in Population 1

Estimator			Adapted			Proposed
Huber	T_{u1} 247.1293	T_{u2} 265.1170	T_{u3} 487.0173	T_{u4} 593.1248	T_{u5} 602.0028	T_{v1} 510.7594
LAD	T_{u6} 253.0916	T_{u7} 271.5699	T_{u8} 496.2643	T_{u9} 589.0617	T_{u10} 603.1537	T_{v2} 501.7501
LTS	T_{u11} 231.1322	T_{u12} 247.7784	T_{u13} 460.4040	T_{u14} 601.1056	T_{u15} 595.0309	T_{v3} 535.0574
LMS	T_{u16} 237.2821	T_{u17} 254.4481	T_{u18} 470.9394	T_{u19} 598.6002	T_{u20} 598.4194	T_{v4} 525.7297
Hample	T_{u21} 242.9736	T_{u22} 260.6162	T_{u23} 480.3531	T_{u24} 595.6374	T_{u25} 600.7559	T_{v5} 517.0712
Tukey	T_{u26} 246.7211	T_{u27} 264.6750	T_{u28} 486.3706	T_{u29} 593.3838	T_{u30} 601.8969	T_{v6} 511.3785
HMM	T_{u31} 247.4246	T_{u32} 265.4368	T_{u33} 487.4842	T_{u34} 592.9358	T_{u35} 602.0772	T_{v7} 510.3116

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

Table 6. PRE of adapted and proposed estimators through simulation study of Population 2

Estimator	Adapted					Proposed
Huber	T_{u1} 104.146052	T_{u2} 115.864207	T_{u3} 454.023148	T_{u4} 446.128689	T_{u5} 453.923680	T_{v1} 450.577311
LAD	T_{u6} 104.132586	T_{u7} 115.764958	T_{u8} 453.459726	T_{u9} 443.971893	T_{u10} 454.279508	T_{v2} 451.854133
LTS	T_{u11} 104.117622	T_{u12} 115.655183	T_{u13} 452.519469	T_{u14} 441.290052	T_{u15} 454.362501	T_{v3} 452.975709
LMS	T_{u16} 104.016751	T_{u17} 114.929177	T_{u18} 437.617667	T_{u19} 415.967362	T_{u20} 446.066805	T_{v4} 451.754292
Hample	T_{u21} 104.170478	T_{u22} 116.045355	T_{u23} 454.368573	T_{u24} 449.409010	T_{u25} 452.616125	T_{v5} 447.641910
Tukey	T_{u26} 104.11241	T_{u27} 115.61707	T_{u28} 452.11405	T_{u29} 440.28650	T_{u30} 454.31313	T_{v6} 453.29108
HMM	T_{u31} 104.125055	T_{u32} 115.709643	T_{u33} 453.027957	T_{u34} 442.659365	T_{u35} 454.362750	T_{v7} 452.458326

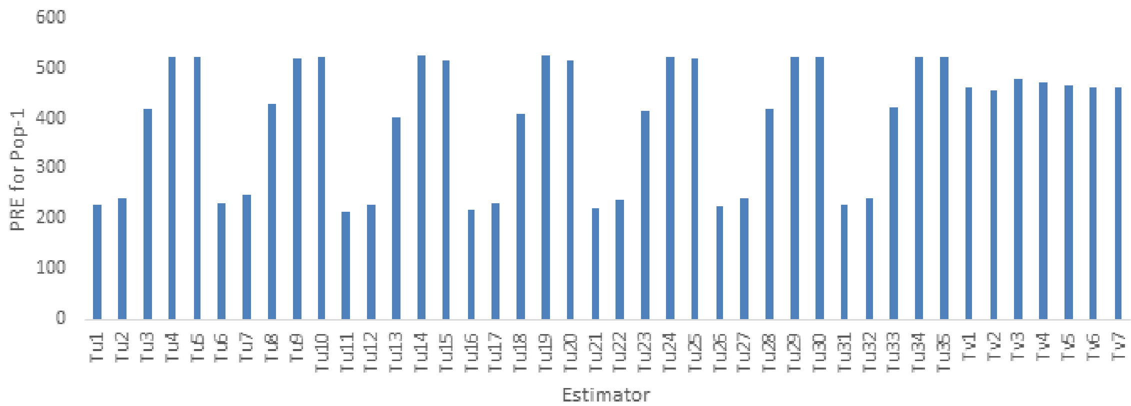

Figure 1. PRE of adapted and proposed estimators with $n = 4$ and $r_\omega = -0.1510$ in Population 1

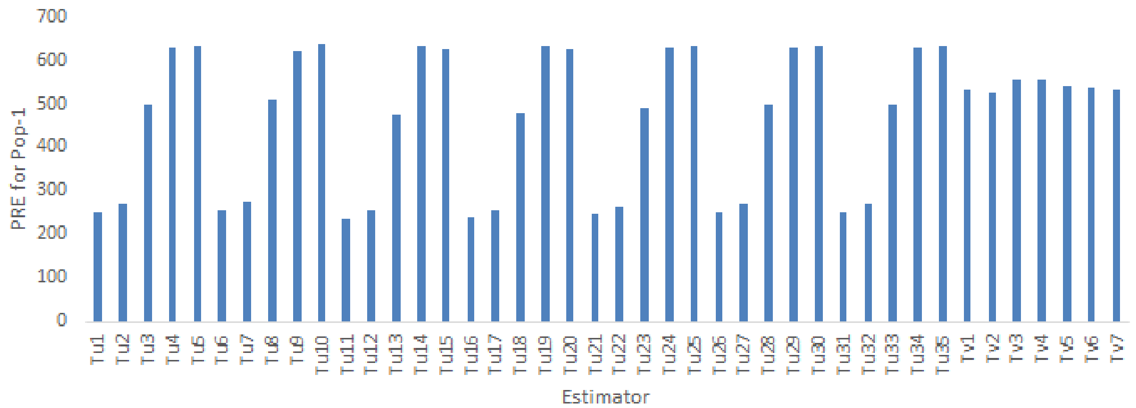

Figure 2. PRE of adapted and proposed estimators with $n = 8$ and $r_\omega = -0.1106$ in Population 1

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

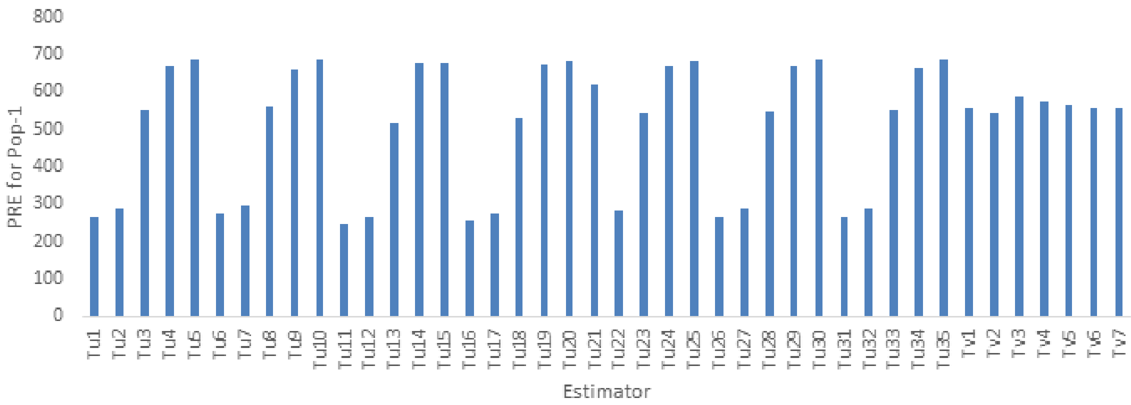

Figure 3. PRE of adapted and proposed estimators with $n = 16$ and $r_{\omega} = -0.0522$ in Population 1

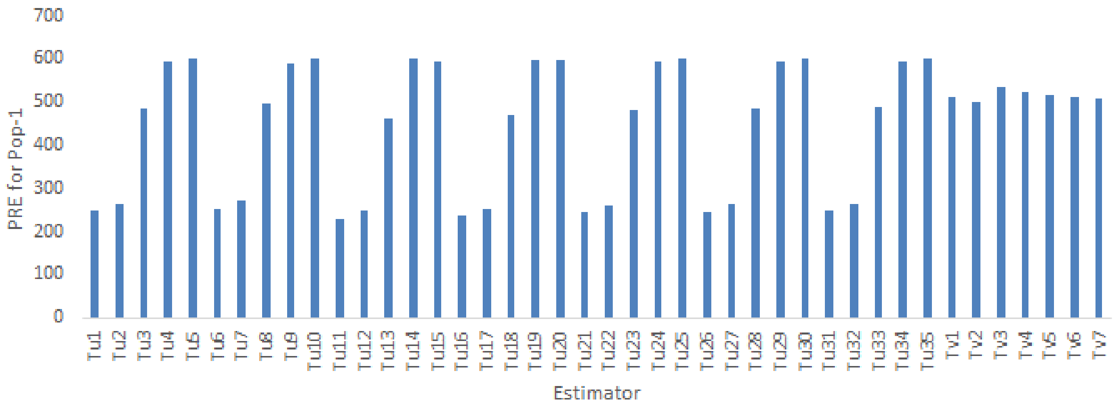

Figure 4. PRE of adapted and proposed estimators with $n = 22$ and $r_{\omega} = -0.0435$ in Population 1

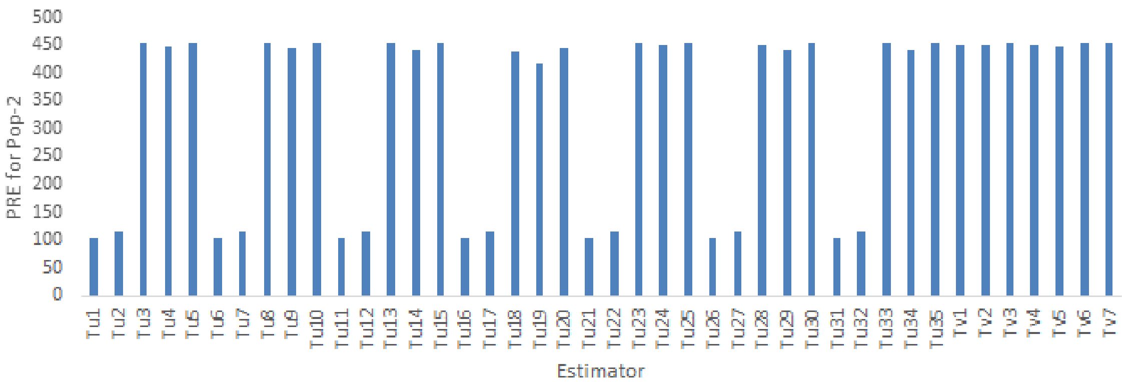

Figure 5. PRE of adapted and proposed estimators through simulation study of Population 2

CONCLUSIONS

In this study we have comprehensively evaluated MCD-based estimators across diverse scenarios and populations. The numerical results present PRE for each estimator under different sample sizes and intra-class correlations. The findings reveal variations in accuracy and efficiency across distinct conditions. Notably, the proposed estimators demonstrate superior performance compared to most existing ones. These results contribute valuable insights into the selection of appropriate estimators based on specific conditions, thereby advancing the understanding of estimation methodologies and facilitating informed decision-making in practical applications. In future studies the work can be extended in light of cluster sampling.

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

ACKNOWLEDGEMENTS

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Grant No. KFU250458).

REFERENCES

1. C. Kadilar and H. Cingi, "Ratio estimators in simple random sampling", *Appl. Math. Comput.*, **2004**, *151*, 893-902.
2. A. M. Abd-Elfattah, E. A. El-Sherpieny, S. M. Mohamed and O. F. Abdou, "Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute", *Appl. Math. Comput.*, **2010**, *215*, 4198-4202.
3. N. Koyuncu, "Efficient estimators of population mean using auxiliary attributes", *Appl. Math. Comput.*, **2012**, *218*, 10900-10905.
4. C. Kadilar, M. Candan and H. Cingi, "Ratio estimators using robust regression", *Hac. J. Math. Stat.*, **2007**, *36*, 181-188.
5. M. Abid, R. A. K. Sherwani, N. Abbas and M. A. Raza, "Some improved modified ratio estimators based on decile mean of an auxiliary variable", *Pak. J. Stat. Oper. Res.*, **2016**, *12*, 787-797.
6. M. Abid, N. Abbas, H. Z. Nazir and Z. Lin, "Enhancing the mean ratio estimators for estimating population mean using non-conventional location parameters", *Revista Colomb. Estad.*, **2016**, *39*, 63-79.
7. M. Abid, N. Abbas, R. A. K. Sherwani, and H. Z. Nazir, "Improved ratio estimators for the population mean using non-conventional measures of dispersion", *Pak. J. Stat. Oper. Res.*, **2016**, *12*, 353-367.
8. M. Irfan, M. Javed, M. Abid and Z. Lin, "Improved ratio type estimators of population mean based on median of a study variable and an auxiliary variable", *Hac. J. Math. Stat.*, **2018**, *47*, 659-673.
9. T. Zaman and H. Bulut, "Modified ratio estimators using robust regression methods", *Commun. Stat. Theory Meth.*, **2019**, *48*, 2039-2048.
10. T. Zaman, " An efficient exponential estimator of the mean under stratified random sampling", *Math. Popul. Stud.*, **2021**, *28*, 104-121.
11. H. Bulut and T. Zaman, "An improved class of robust ratio estimators by using the minimum covariance determinant estimation", *Commun. Stat. Simul. Comput.*, **2022**, *51*, 2457-2463.
12. P. J. Rousseeuw and K. van Driessen, "A fast algorithm for the minimum covariance determinant estimator", *Technometr.*, **1999**, *41*, 212-223.
13. M. Hubert, P. J. Rousseeuw and T. Verdonck, "A deterministic algorithm for robust location and scatter", *J. Comput. Graph. Stat.*, **2012**, *21*, 618-637.
14. J. Schreurs, I. Vranckx, M. Hubert, J. A. K. Suykens and P. J. Rousseeuw, "Outlier detection in non-elliptical data by kernel MRCD", *Stat. Comput.*, **2021**, *31*, Art.no.66.
15. K. Boudt, P. J. Rousseeuw, S. Vanduffel and T. Verdonck, "The minimum regularized covariance determinant estimator", *Stat. Comput.*, **2020**, *30*, 113-128.
16. J. Kalina and J. Tichavský, "The minimum weighted covariance determinant estimator for high-dimensional data", *Adv. Data Anal. Classif.*, **2022**, *16*, 977-999.

17. R. Muthukrishnan and S. S. Nair, "Robust Mahalanobis depth based on minimum regularized covariance determinant estimator for high-dimensional data", *Reliab. Theory Appl.*, **2023**, *18*, 446-452.
18. U. Shahzad, N. H. Al-Noor, N. Afshan, D. A. Alilah, M. Hanif and M.M. Anas, "Minimum covariance determinant-based quantile robust regression-type estimators for mean parameter", *Math. Probl. Eng.*, **2021**, *2021*, Art.no.5255839.
19. B. Tutmez, "Minimum covariance determinant-based bootstrapping for appraising air passenger arrival data", *J. Eng. Manag. Compet.*, **2022**, *12*, 176-185.
20. A. M. Alomair and U. Shahzad, "Optimizing mean estimators with calibrated minimum covariance determinant in median ranked set sampling", *Symmetry*, **2023**, *15*, Art.no.1581.
21. M. M. Anas, Z. Huang, U. Shahzad and S. Iftikhar, "A new family of robust quantile-regression-based mean estimators using Sarndal approach", *Commun. Stat. Simul. Comput.*, **2024**, doi: 10.1080/03610918.2024.2359504.
22. M. Abid, M. Sun, W. Latif and T. Nawaz, "A novel robust class of estimators for estimation of finite population mean: A simulation study", *Pak. J. Stat. Oper. Res.*, **2024**, *20*, 417-444.
23. T. Zaman and H. Bulut, "Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling", *Commun. Stat. Theory Meth.*, **2020**, *49*, 3407-3420.
24. T. Zaman, "Improvement of modified ratio estimators using robust regression methods", *Appl. Math. Comput.*, **2019**, *348*, 627-631.
25. J. Raymaekers and P. J. Rousseeuw, "The cellwise minimum covariance determinant estimator", *J. Am. Stat. Assoc.*, **2024**, *119*, 2610-2621.
26. M. Hubert and M. Debruyne, "Minimum covariance determinant", *WIREs Comput. Stat.*, **2010**, *2*, 36-43.
27. N. H. Al-Noor and A. A. Mohammad, "Model of robust regression with parametric and non-parametric methods", *Math. Theory Model.*, **2013**, *3*, 27-39.
28. M. Hubert, M. Debruyne and P. J. Rousseeuw, "Minimum covariance determinant and extensions", *WIREs Comput. Stat.*, **2018**, *10*, Art.no.e1421.
29. S. Bhushan and R. Gupta, "An improved log-type family of estimators using attribute", *J. Stat. Manag. Syst.*, **2020**, *23*, 593-602.
30. M. Subzar, A. I. Al-Omari and A. R. A. Alanzi, "The robust regression methods for estimating of finite population mean based on SRSWOR in case of outliers", *CMC-Comput. Mater. Contin.*, **2020**, *65*, 125-138.
31. L. K. Grover and A. Kaur, "An improved regression type estimator of population mean with two auxiliary variables and its variant using robust regression method", *J. Comput. Appl. Math.*, **2021**, *382*, Art.no.113072.
32. T. Koç and H. Koç, "A new class of quantile regression ratio-type estimators for finite population mean in stratified random sampling", *Axioms*, **2023**, *12*, Art.no.713.
33. S. Malik, Kanika and Atul, "An improvement in regression estimator through exponential estimator using two auxiliary variables," *J. Modern Appl. Stat. Meth.*, **2023**, *22*, doi: 10.56802/Jmasm.V22.i1.1.
34. A. Kumar and A. S. Siddiqui, "Enhanced estimation of population mean using simple random sampling", *Res. Stat.*, **2024**, *2*, Art.no. 2335949.
35. S. Bhushan and A. Kumar, "Improved estimation of population mean in simple random sampling using attribute", *Thai. Statistician*, **2024**, *22*, 374-389.

Maejo Int. J. Sci. Technol. **2025**, *19*(01), 67-79

36. W. Tian and Z. Qin, "The minimum covariance determinant estimator for interval-valued data", *Stat. Comput.*, **2024**, *34*, Art.no.80.
37. U. Shahzad, M. Hanif, N. Koyuncu and A. V. G. Luengo, "Estimation of population mean under systematic random sampling in absence and presence non-response", *J. Statisticians: Stat. Actuarial Sci.*, **2017**, *1*, 23-39.
38. U. Shahzad, I. Ahmad, N. H. Al-Noor, M. Hanif and I. M. Almanjahie, "Robust estimation of the population mean using quantile regression under systematic sampling", *Math. Popul. Stud.*, **2023**, *30*, 195-207.
39. M. N. Murthy, "Sampling Theory and Methods", 2nd Edn., Statistical Publishing Society, Calcutta, **1967**, pp.133-182.
40. W. J. Dixon, "Analysis of extreme values", *Ann. Math. Stat.*, **1950**, *21*, 488-506.
41. W. J. Dixon, "Ratios involving extreme values", *Ann. Math. Stat.*, **1951**, *22*, 68-78.