Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

Maejo International Journal of Science and Technology

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

Full Paper

Quantile regression-based mean estimation in circular systematic sampling

Afzal Subhani, Muhammad Irfan* and Maria Javed

Department of Statistics, Government College University Faisalabad, Pakistan

* Corresponding author, e-mail: mirfan@gcuf.edu.pk

Received: 8 August 2025 / Accepted: 29 March 2026 / Published: 31 March 2026

Abstract: The current study addresses a significant gap in sampling theory by introducing novel ratio-type estimators for finite population mean under circular systematic sampling using a quantile regression approach. We derive the theoretical mean squared error expressions for the proposed estimators and conduct comparative analyses against existing estimators. The results clearly demonstrate that our estimators outperform traditional ones in terms of efficiency. To validate our findings, we present numerical illustrations based on a real-life data set. Moreover, a Monte Carlo simulation study based on real-life data set is also included to check the performance of the proposed estimators. Numerical findings endorse the potential of the proposed estimators. The findings may help the researchers to get more estimates that are precise for population mean, which is widely applicable in different fields.

Keywords: mean squared error, ratio-type estimators, circular systematic sampling, outliers, quantile regression

---

INTRODUCTION

In statistical literature various sampling methods have been developed to select truly representative subsets of a population, each chosen based on the nature and structure of the data. Among these, systematic sampling stands out for its simplicity and efficiency. In this method the first unit is selected randomly and the remaining units are chosen according to a fixed, predefined pattern. One of the key advantages of systematic sampling is that the entire sample can be drawn with just a single random start. Beyond its simplicity, which is especially valuable in practical applications, systematic sampling has been shown to yield more efficient estimators over the simple random sampling in many cases, as noted by Cochran [1] and Gautschi [2].

Systematic sampling involves selecting samples at regular intervals from a sampling frame, beginning from a randomly chosen starting point and continuing until the end of the frame is reach.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

Circular systematic sampling (CSS), on the other hand, treats the sampling frame as a loop wrapping around from the end back to the beginning to continue the selection process. This key distinction allows CSS to maintain a consistent sampling interval throughout, potentially reducing bias and improving representativeness, especially in populations with cyclical patterns. To achieve more accurate estimates of the population mean, numerous researchers have suggested different types of estimators by using auxiliary information under different sampling schemes. We briefly review some of the estimators for the estimation of the population mean, which include the use of auxiliary information in the framework of systematic sampling, circular sampling and their modifications.

Sampath and Varalakshmi [3] proposed a new diagonal systematic sampling design and compared their proposed method with traditional CSS. To evaluate the performance of the proposed method, both natural and non-linear populations generated from well-known distributions were considered. Singh and Solanki [4] proposed a novel class of estimators for the population mean in systematic sampling, leveraging auxiliary information to enhance estimation accuracy. Mathematical expressions such as bias and mean squared error (MSE) were derived, particularly under large sample approximations. They also compared their work with that of Shukla [5] and the traditional regression estimator. Subramani and Singh [6] proposed an estimator for determining the sampling interval in CSS, demonstrating its superiority through various methods. They also introduced methods for estimating the population mean in the presence of linear trends, specifically for finite populations. Further evaluating the efficiency of their approach, they compared diagonal systematic sampling methods with single and random starts across multiple populations, aiming to improve the accuracy of their estimates. Verma and Singh [7] suggested exponential-type estimators for estimating the population mean under the CSS scheme. They used both simple and two-phase sampling designs. Their proposed estimators performed better as compared to the usual ratio, product, and regression estimators and the other existing estimators. Riaz et al. [8] examined the problem of estimating the population mean of a study variable by incorporating auxiliary information within a CSS design. Their work considered situations both with and without non-response in the study variable, offering a more comprehensive estimation framework.

Singh and Yadav [9] proposed a generalised family of estimators for the population mean based on auxiliary information in the CSS scheme. Theoretical and empirical evaluations demonstrate that the proposed family of estimators outperforms traditional ratio and regression estimators. Mostafa and Ahmad [10] have presented a review of systematic sampling in their paper and recommended researchers to use this sampling design under different situations. Khan et al. [11] suggested a new estimator through the optimal pairing of units in systematic sampling. Singh et al. [12] introduced a generalised family of estimators for estimating the finite population mean using CSS, exploring various scenarios such as non-response, simple random sampling and two-phase sampling schemes. Their research included critical and empirical evaluations, demonstrating the efficiency and superiority of their proposed estimators as compared to the available estimators in the study.

Singh and Yadav [13] proposed ratio-product ratio-type exponential estimator for estimating the population mean under systematic sampling. Mathematical expressions in terms of bias and MSE were derived. Khan et al. [14] investigated the performance of modified systematic sampling for populations exhibiting autocorrelation. The efficiency of the sampling is evaluated in comparison with linear systematic sampling, CSS and mixed random systematic sampling under various super population models, with particular attention to its suitability for auto-correlated data structures. Pal et al. [15] proposed an efficient class of difference-type estimators for estimating the population mean by using two auxiliary variables in systematic sampling. An empirical study has also been conducted to check the efficiency of the proposed estimators over the competing estimators. Most recently, Irfan et al. [16] introduced a new generalised estimator for estimating the population mean under the CSS framework. The mathematical formulation of the estimator was derived and theoretical comparisons revealed its superiority over the traditional unbiased, ratio, product, and regression estimators.

A summary of different publications that discuss systematic, CSS, diagonal systematic sampling, and modified systematic techniques is presented in Table 1.

Table 1. Summary of different systematic sampling techniques

Publication	**Sampling Method**	**Key Findings**
Sampath and Varalakshmi [3]	Diagonal Systematic	New sampling approach. Real-life and non-linear populations simulated from some distributions are used to check performance of the new approach.
Singh and Solanki [4]	Systematic	Improved performance in population parameters using large sample approximations
Shukla [5]	Systematic	Novel estimator suggested for estimating population mean
Subramani and Singh [6]	Circular Systematic	Efficient estimator proposed to estimate population mean
Verma and Singh [7]	Circular Systematic	Exponential-type estimators suggested for estimation of population mean under simple and two-phase sampling designs
Riaz et al. [8]	Circular Systematic	Situation of non-response considered to estimate population mean
Singh and Yadav [9]	Circular Systematic	Unknown population mean estimated by considering with and without non-response

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

Table 1. (continued).

Mostafa and Ahmad [10]	Systematic	Detailed review of systematic sampling presented
Khan et al. [11]	Modified Systematic	Performance of modified systematic sampling investigated over linear, circular and mixed random systematic sampling schemes
Singh et al. [12]	Circular Systematic	Different situations e.g. non-response, simple random sampling and two-phase sampling discussed
Singh and Yadav [13]	Systematic	Ratio-product ratio-type exponential estimator for estimation of population mean suggested
Khan et al. [14]	Optimal Systematic	New estimator using optimum matching of units suggested. Newly suggested scheme can be used not only in case when population size is multiple of sample size, but also in other cases when this may not be true
Pal et al. [15]	Systematic	Well-organised class of difference-type estimators developed to estimate population mean utilising bivariate auxiliary information
Irfan et al. [16]	Circular Systematic	Generalised estimator suggested for estimation of unknown population mean by using non-conventional measures of auxiliary variable
This article	Circular Systematic using quantile regression approach	Novel estimators suggested for unknown population mean to fill the gap in sampling theory. Real-life application and Monte Carlo simulation study based on real-life data set are also added to confirm validity of the suggested estimators.

Outliers which are data points that significantly deviate from the overall pattern can skew mean estimates and compromise the accuracy of ordinary least square regression coefficients commonly used to estimate the population mean. The occurrence of outliers can substantially affect ordinary least square estimates, leading to poor performance. In contrast, quantile regression provides a robust alternative, capable of mitigating the influence of outliers and handling non-normal data, thereby offering a more reliable estimation approach. Kadilar and Cingi [17] faced limitations when their method encountered data containing outliers, leading to inefficiency. To overcome this issue, they explored the use of Huber regression, which provides robust estimates of the slope coefficient in non-Gaussian error distributions [18]. By employing a function that behaves quadratic for small residuals and linearly for large residuals, Huber regression effectively neutralises the impact of outliers on population mean estimates, outperforming traditional ordinary least square regression in such scenarios. Many researchers [19-22] have proposed estimators for unknown population parameters when the data sets contain outliers under simple random sampling.

Quantile regression is one of the new techniques which are used to control the effect of outliers. It is strong even in the cases where the data is not Gaussian [23, 24]. Anas et al. [25] suggested an effective family of estimators for estimating finite population mean by applying quantile regression method under missing data case. Koc and Koc [26] proposed another robust class of estimators for estimating population mean based on quantile regression under stratified random sampling. Shahzad et al. [27] first developed quantile regression estimators and later proposed those under systematic sampling to examine the effect of outliers. To compare the theoretical and numerical results, a real-life data set and simulation were also used, which confirmed that the newly proposed estimators perform very well. Shahzad et al. [28] used quantile regression for the estimation of the mean under simple random sampling. The authors used two auxiliary variables and also employed non-conventional measures for the development of newly proposed estimators. For the numerical study, different real-life data sets and simulations were used, which confirmed that their proposed estimators perform better than the existing estimators. Khalid et al. [29] developed some estimators for the estimation of population mean using quantile regression approach. Their findings effectively reduced the impact of large deviations, providing a more resilient estimation method. Alomair and Iftikhar [30] proposed a novel dual-type estimator for the mean based on quantile regression in the presence of outliers. Using asymmetric data, the performance of the proposed estimators was compared with existing estimators, both in the presence and absence of scrambled response. The current study addresses a significant gap in sampling theory by introducing novel estimators for estimating the population mean under CSS using quantile regression approach.

Existing Estimators

Kadilar and Cingi [17] introduced several ratio-type estimators aimed at estimating the population mean within the framework of simple random sampling:

$$\hat{\bar{Y}}_{KC1} = \frac{\bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})}{\bar{x}} \bar{X} \tag{1}$$

$$\hat{\bar{Y}}_{KC2} = \frac{\bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})}{\bar{x} + C_x} (\bar{X} + C_x) \tag{2}$$

$$\hat{\bar{Y}}_{KC3} = \frac{\bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})}{\bar{x} + \beta_{2(x)}} \left(\bar{X} + \beta_{2(x)}\right) \tag{3}$$

$$\hat{\bar{Y}}_{KC4} = \frac{\bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})}{\bar{x}\beta_{2(x)} + C_x} \left(\bar{X}\beta_{2(x)} + C_x\right) \tag{4}$$

$$\hat{\bar{Y}}_{KC5} = \frac{\bar{y} + \hat{\beta}_{(ols)}(\bar{X} - \bar{x})}{\bar{x}C_x + \beta_{2(x)}} \left(\bar{X}C_x + \beta_{2(x)}\right) \tag{5}$$

where C_x represents the population coefficient of variation of the auxiliary variable, $\beta_{2(x)}$ denotes the population coefficient of the kurtosis of the auxiliary variable, $\bar{y}$ is the sample mean of the study

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

variable, $\bar{x}$ is the sample mean of the auxiliary variable, $\bar{X}$ is the known population mean, and $\hat{\beta}_{(ols)}$ indicates the regression coefficient. The MSE expressions of the estimators proposed by Kadilar and Cingi [17] are given as follows:

$$MSE\left(\hat{\bar{Y}}_{KCi}\right) \cong \theta[R_{KCi}^2 S_x^2 + S_y^2(1-\rho_{yx}^2)], i = 1,2,\dots,5 \tag{6}$$

where

$R_{KC1} = R = \frac{\bar{Y}}{\bar{X}}$, $R_{KC2} = \frac{\bar{Y}}{\bar{X}+C_x}$, $R_{KC3} = \frac{\bar{Y}}{\bar{X}+\beta_{2(x)}}$, $R_{KC4} = \frac{\bar{Y}\beta_{2(x)}}{\bar{X}\beta_{2(x)}+C_x}$,

$R_{KC5} = \frac{\bar{Y}C_x}{\bar{X}C_x+\beta_{2(x)}}$ and $\theta = \left(\frac{1}{n}-\frac{1}{N}\right)$.

Let $U = (U_1, U_2, \dots, U_N)$ denote a finite population of size N. The values of the study and auxiliary variables for the j^{th} unit in the i^{th} sample are denoted by y_{ij} and x_{ij} , where $i = 1,2,\dots,k$ and $j = 1,2,\dots,n$. In this framework, k samples are selected, each consisting of n units. From the population, N circular systematic samples of size n are then obtained with $1 \le r \le N$. The sample means under CSS scheme are: $\bar{y}_{css} = n^{-1}\sum_{j=0}^{n-1} y_{r+jk}$ and $\bar{x}_{css} = n^{-1}\sum_{j=0}^{n-1} x_{r+jk}$, and the intra-class correlation coefficients ρ_y and ρ_x between pairs of units within the CSS are for y and x variables respectively. The variance and covariance of $\bar{y}_{css}$ and $\bar{x}_{css}$ are as follows:

$$Var(\bar{y}_{css}) = \left(\frac{N-1}{N}\right)[1+(n-1)\rho_y]\frac{S_y^2}{n} = \tilde{S}_y^2 = \bar{Y}^2\tilde{C}_y^2 \tag{7}$$

$$Var(\bar{x}_{css}) = \left(\frac{N-1}{N}\right)[1+(n-1)\rho_x]\frac{S_x^2}{n} = \tilde{S}_x^2 = \bar{X}^2\tilde{C}_x^2 \tag{8}$$

$$Cov(\bar{y}_{css}, \bar{x}_{css}) = \left(\frac{N-1}{N}\right)[1+(n-1)\rho_y]^{1/2}[1+(n-1)\rho_x]^{1/2}\frac{S_{yx}}{n} = \tilde{S}_{yx} = \bar{Y}\,\bar{X}\tilde{C}_{yx} \tag{9}$$

where

$$S_y^2 = \frac{1}{n(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}(y_{ij}-\bar{Y})^2$$

$$S_x^2 = \frac{1}{n(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}(x_{ij}-\bar{X})^2$$

$$S_{yx} = \frac{1}{n(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}(y_{ij}-\bar{Y})(x_{ij}-\bar{X})$$

$$\rho_y = \frac{\sum_{i=1}^{N}\sum_{j\neq j=1}^{n}(y_{ij}-\bar{Y})(y_{ij}-\bar{Y})}{kn(n-1)S_y^2}$$

$$\rho_x = \frac{\sum_{i=1}^{N}\sum_{j\neq j=1}^{n}(x_{ij}-\bar{X})(x_{ij}-\bar{X})}{kn(n-1)S_x^2}.$$

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

SUGGESTED ESTIMATORS BASED ON QUANTILE REGRESSION

Inspired by the work of Kadilar and Cingi [17], this study introduces five innovative ratio-type estimators, developed within the framework of quantile regression under CSS. The formulations of the estimators are presented below:

$$\hat{\hat{Y}}_{P1} = \frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}}\bar{X} \quad (10)$$

$$\hat{\hat{Y}}_{P2} = \frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}+C_x}(\bar{X}+C_x) \quad (11)$$

$$\hat{\hat{Y}}_{P3} = \frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}+\beta_{2(x)}}\left(\bar{X}+\beta_{2(x)}\right) \quad (12)$$

$$\hat{\hat{Y}}_{P4} = \frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}\beta_{2(x)}+C_x}\left(\bar{X}\beta_{2(x)}+C_x\right) \quad (13)$$

$$\hat{\hat{Y}}_{P5} = \frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}C_x+\beta_{2(x)}}\left(\bar{X}C_x+\beta_{2(x)}\right) \quad (14)$$

where $\hat{\beta}_{(qi)}$ is the quantile regression coefficient.

The following relative error terms, along with their expectations, are used to drive the expressions for the bias and MSE of the proposed estimators:

$$\zeta_0 = \frac{\bar{y}_{css}-\bar{Y}}{\bar{Y}} \text{ and } \zeta_1 = \frac{\bar{x}_{css}-\bar{X}}{\bar{X}} \quad (15)$$

such that

$$E(\zeta_0) = E(\zeta_1) = 0$$
$$E(\zeta_0^2) = \tilde{C}_y^2, \qquad E(\zeta_1^2) = \tilde{C}_x^2 \qquad \text{and} \qquad E(\zeta_0\zeta_1) = \tilde{C}_{yx}.$$

Expressing the estimator $\hat{\hat{Y}}_{Pi}, i = 1,2,3,4,5$ in terms of relative error terms, we have

$$\hat{\hat{Y}}_{Pi} = [\bar{Y}(1+\zeta_0) - \hat{\beta}_{(qi)}\bar{X}\zeta_1](1+\alpha_i\zeta_1)^{-1} \quad (16)$$

$$\hat{\hat{Y}}_{Pi} = [\bar{Y}+\bar{Y}\zeta_0 - \hat{\beta}_{(qi)}\bar{X}\zeta_1](1-\alpha_i\zeta_1+\alpha_i^2\zeta_1^2)$$

$$\hat{\hat{Y}}_{Pi} = \bar{Y} - \alpha_i\bar{Y}\zeta_1 + \alpha_i^2\bar{Y}\zeta_1^2 + \bar{Y}\zeta_0 - \alpha_i\bar{Y}\zeta_0\zeta_1 - \hat{\beta}_{(qi)}\bar{X}\zeta_1 + \alpha_i\hat{\beta}_{(qi)}\bar{X}\zeta_1^2. \quad (17)$$

Subtracting $\bar{Y}$ on both sides of Eq. (17), we get

$$\left(\hat{\hat{Y}}_{Pi} - \bar{Y}\right) = -\alpha_i\bar{Y}\zeta_1 + \alpha_i^2\bar{Y}\zeta_1^2 + \bar{Y}\zeta_0 - \alpha_i\bar{Y}\zeta_0\zeta_1 - \hat{\beta}_{(qi)}\bar{X}\zeta_1 + \alpha_i\hat{\beta}_{(qi)}\bar{X}\zeta_1^2. \quad (18)$$

Taking expectation on both side of Eq. (18), we get Bias$\left(\hat{\hat{Y}}_{Pi}\right)$:

$$Bias\left(\hat{\hat{Y}}_{Pi}\right) \cong \alpha_i^2\bar{Y}\tilde{C}_x^2 - \alpha_i\bar{Y}\tilde{C}_{yx} + \alpha_i\hat{\beta}_{(qi)}\bar{X}\tilde{C}_x^2. \quad (19)$$

Taking square on both sides of Eq. (18), we get

$$\left(\hat{\hat{Y}}_{Pi} - \bar{Y}\right)^2 = \alpha_i^2\bar{Y}^2\zeta_1^2 + \bar{Y}^2\zeta_0^2 + \hat{\beta}_{(qi)}^2\bar{X}^2\zeta_1^2 - 2\alpha_i\bar{Y}^2\zeta_0\zeta_1 + 2\alpha_i\hat{\beta}_{(qi)}\bar{X}\bar{Y}\zeta_1^2 - 2\hat{\beta}_{(qi)}\bar{X}\bar{Y}\zeta_0\zeta_1. \quad (20)$$

Taking expectation on both side of the Eq. (20), we get $MSE\left(\hat{\hat{Y}}_{Pi}\right)$:

$$MSE\left(\hat{\hat{Y}}_{Pi}\right) \cong (\alpha_i\bar{Y}+\hat{\beta}_{(qi)}\bar{X})^2\tilde{C}_x^2 + \bar{Y}^2\tilde{C}_y^2 - 2\bar{Y}\left(\alpha_i\bar{Y}+\hat{\beta}_{(qi)}\bar{X}\right)\tilde{C}_{yx} \quad (21)$$

where

$\alpha_1 = 1, \alpha_2 = \frac{\bar{X}}{\bar{X}+C_x}, \alpha_3 = \frac{\bar{X}}{\bar{X}+\beta_{2(x)}}, \alpha_4 = \frac{\bar{X}\beta_{2(x)}}{\bar{X}\beta_{2(x)}+C_x}$ and $\alpha_5 = \frac{\bar{X}C_x}{\bar{X}C_x+\beta_{2(x)}}$.

We also propose several other novel estimators by combining the two estimators already introduced in Eq (10) to (14), as follows:

$$\hat{\bar{Y}}_{P6} = t\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}}\bar{X} + (1-t)\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}+C_x}(\bar{X}+C_x) \quad (22)$$

$$\hat{\bar{Y}}_{P7} = t\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}}\bar{X} + (1-t)\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}+\beta_{2(x)}}(\bar{X}+\beta_{2(x)}) \quad (23)$$

$$\hat{\bar{Y}}_{P8} = t\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}}\bar{X} + (1-t)\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\beta_{2(x)}\bar{x}_{css}+C_x}(\beta_{2(x)}\bar{X}+C_x) \quad (24)$$

$$\hat{\bar{Y}}_{P9} = t\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{\bar{x}_{css}}\bar{X} + (1-t)\frac{\bar{y}_{css}+\hat{\beta}_{(qi)}(\bar{X}-\bar{x}_{css})}{C_x\bar{x}_{css}+\beta_{2(x)}}(C_x\bar{X}+\beta_{2(x)}). \quad (25)$$

Expressing estimator $\hat{\bar{Y}}_{Pj}, j = 6,7,8,9$ in terms of relative error terms, we have

$$\hat{\bar{Y}}_{Pj} = t_j[\bar{Y}(1+\zeta_0) - \hat{\beta}_{(qi)}\bar{X}\zeta_1](1+\zeta_1)^{-1} + (1-t_j)[\bar{Y}(1+\zeta_0) - \hat{\beta}_{(qi)}\bar{X}\zeta_1](1+\alpha_i\zeta_1)^{-1}.$$

After simplification and subtracting $\bar{Y}$ on both sides, we get

$$\left(\hat{\bar{Y}}_{Pj} - \bar{Y}\right) = -t_j\bar{Y}\zeta_1 + t_j\bar{Y}\zeta_1^2 - t_j\bar{Y}\zeta_0\zeta_1 + t_j\hat{\beta}_{(qi)}\bar{X}\zeta_1^2 + \bar{Y}\zeta_0 - \hat{\beta}_{(qi)}\bar{X}\zeta_1 - \alpha_i\bar{Y}\zeta_1 - \alpha_i\bar{Y}\zeta_0\zeta_1 + \alpha_i\hat{\beta}_{(qi)}\bar{X}\zeta_1^2 + t_j\alpha_i\bar{Y}\zeta_1 + t_j\alpha_i\bar{Y}\zeta_0\zeta_1 - t_j\alpha_i\hat{\beta}_{(qi)}\bar{X}\zeta_1^2 + \alpha_i^2\bar{Y}\zeta_1^2 - t_j\alpha_i^2\bar{Y}\zeta_1^2. \quad (26)$$

Taking expectation and square on both side of Eq.(26), we get the MSE of the estimator $\hat{\bar{Y}}_{Pj}, j = 6,7,8,9$ up to the first order of approximation in Eq. (27):

$$MSE\left(\hat{\bar{Y}}_{Pj}\right) \cong t_j^2\bar{Y}^2\tilde{C}_x^2(1-\alpha_i)^2 + 2t_j\bar{X}\bar{Y}(\hat{\beta}_{(qi)}\tilde{C}_x^2 - R\tilde{C}_{yx} + \alpha_iR\tilde{C}_x^2 + \alpha_iR\tilde{C}_{yx} - \alpha_i\hat{\beta}_{(qi)}\tilde{C}_x^2 - \alpha_i^2R\tilde{C}_x^2) + \hat{\beta}_{(qi)}^2\bar{X}^2\tilde{C}_x^2 + \bar{Y}^2\tilde{C}_y^2 + \alpha_i^2\bar{Y}^2\tilde{C}_x^2 - 2\hat{\beta}_{(qi)}\bar{X}\bar{Y}\tilde{C}_{yx} - 2\alpha_i\bar{Y}^2\tilde{C}_{yx} + 2\alpha_i\hat{\beta}_{(qi)}\bar{X}\bar{Y}\tilde{C}_x^2. \quad (27)$$

where $i = 2,3,4,5$ and $j = 6,7,8,9$;

$t_6 = \frac{R_{KC2}}{R_{KC2}-R}, t_7 = \frac{R_{KC3}}{R_{KC3}-R}$ $t_8 = \frac{R_{KC4}}{R_{KC4}-R}$ and $t_9 = \frac{R_{KC5}}{R_{KC5}-R}$ are optimal values.

APPLICATION

To estimate the average of the number of teachers in primary and secondary schools in Turkey, the data set reported by Javed et al. [31] was utilised. In this study the number of teachers in both primary and secondary schools was treated as the study variable while the number of students in these schools was used as the auxiliary variable. Important descriptive measures are reported in Table 2. Firstly, we calculated MSE of the Kadilar and Cingi [17] ratio-type estimators for the population mean under simple random sampling. Table 3. presents the MSE values. Figure 1 presents the scatter plot and box plot of the data set to examine the possible presence of outliers. From the figure, it is evident that some potential outliers are present in the data. A histogram was also constructed to evaluate the normality of the distribution, which indicates that the data do not follow a normal pattern. Under such circumstances, the efficiency of estimators based on traditional ordinary least square coefficients for estimating the population mean under CSS may be

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

questionable [32]. Therefore, the performance of the proposed estimators under CSS was assessed using quantile regression.

Table 2. Important parameters related to the study and auxiliary variables

$N = 923, n = 180, C_x = 1.814, C_y = 1.718, \bar{Y} = 436.436, \bar{X} = 11440.50, \rho_{yx} = 0.954,$
$\rho_x = -0.0031, \rho_y = -0.0025, \beta_{2(x)} = 18.566, t_6 = -6135.868, t_7 = 616.199, t_8 = -113920,$
$t_9 = -1148.9, \mathrm{R}_{KC1} = 0.0381, \mathrm{R}_{KC2} = 0.0382, \mathrm{R}_{KC3} = 0.0380, \mathrm{R}_{KC4} = 0.0381.$

Table 3. MSE of $\hat{\hat{Y}}_{KCi}$ based on real-life application under simple random sampling

Estimator	$\hat{\hat{Y}}_{KC1}$	$\hat{\hat{Y}}_{KC2}$	$\hat{\hat{Y}}_{KC3}$	$\hat{\hat{Y}}_{KC4}$	$\hat{\hat{Y}}_{KC5}$
MSE	3185.975	3185.01	3176.387	3185.923	3180.827

Table 4 presents MSE corresponding to all the proposed estimators, specifically $\hat{\hat{Y}}_{Pi}, i = 1,2,3,4,5$ and $\hat{\hat{Y}}_{Pj}, j = 6,7,8,9$. To evaluate the performance of the estimators, different quantile levels, i.e. $q_{10th} = .10, q_{20th} = .20, q_{30th} = .30, q_{40th} = .40, q_{50th} = .50, q_{60th} = .60, q_{70th} = .70, q_{80th} = .80$ and $q_{90th} = .90$, are considered. The results indicate that as the value of q_{th} increases, and the MSE values also show an increasing trend. Nevertheless, the estimators based on the CSS sampling scheme consistently outperform those derived from the simple random sampling scheme. This reduction underscores the utility and robustness of quantile regression in enhancing the precision of population mean estimation. Empirical results based on real-life data set indicate that the suggested estimators $\hat{\hat{Y}}_{Pi}$ consistently outperform the traditional estimators, i.e. $\hat{\hat{Y}}_{KCi}$, in terms of the lower MSE. Furthermore, among all the proposed estimators, $\hat{\hat{Y}}_{Pj}$ demonstrates superior performance by yielding the lowest MSE when compared to $\hat{\hat{Y}}_{Pi}$. Notably, all the estimators within the $\hat{\hat{Y}}_{Pj}$ class exhibit identical minimum MSE values, emphasising their reliability and consistency in estimation accuracy.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

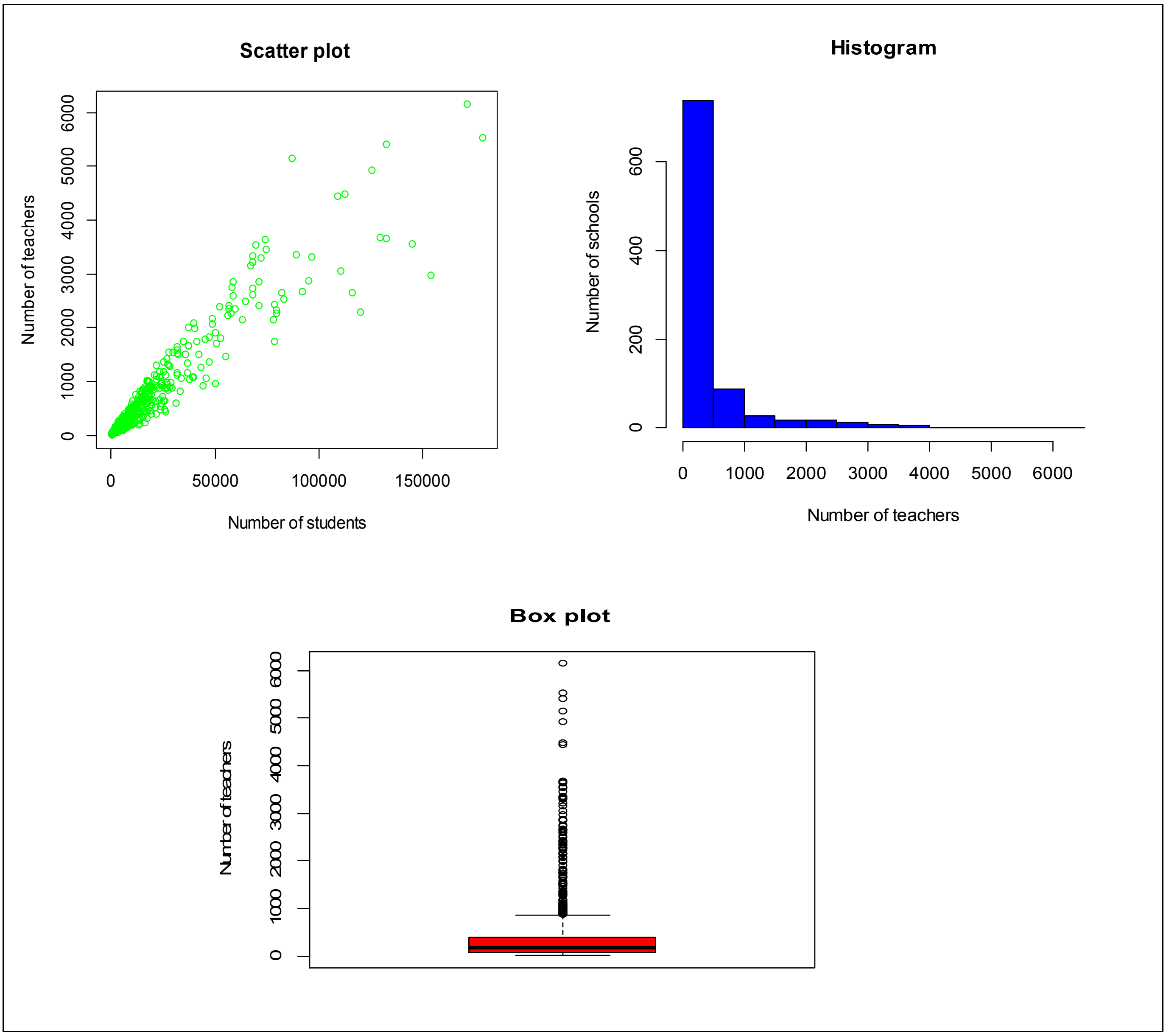

Figure 1. Scatter plot, box-plot and histogram of real-life application

Table 4. MSE of proposed estimators based on real-life application

Estimator	$q_{.10}$	$q_{.20}$	$q_{.30}$	$q_{.40}$	$q_{.50}$	$q_{.60}$	$q_{.70}$	$q_{.80}$	$q_{.90}$
$\hat{\bar{Y}}_{P1}$	732.24	1000.55	1167.62	1416.23	1711.75	1885.08	2088.95	2501.79	2798.66
$\hat{\bar{Y}}_{P2}$	731.93	1001.17	1167.20	1415.76	1711.24	1884.54	2088.37	2501.16	2797.99
$\hat{\bar{Y}}_{P3}$	729.14	997.78	1163.51	1411.63	1706.65	1879.70	2083.25	2495.52	2792.01
$\hat{\bar{Y}}_{P4}$	732.23	1001.52	1167.59	1416.20	1711.72	1885.05	2088.92	2501.76	2798.62
$\hat{\bar{Y}}_{P5}$	730.58	999.52	1165.40	1413.76	1709.01	1882.20	2085.89	2498.42	2795.09
$\hat{\bar{Y}}_{P6}$	406.97	271.48	219.84	173.57	154.24	156.74	170.28	226.281	285.65
$\hat{\bar{Y}}_{P7}$	406.97	271.48	219.84	173.58	154.24	156.74	170.28	226.281	285.65
$\hat{\bar{Y}}_{P8}$	406.97	271.48	219.84	173.58	154.24	156.74	170.28	226.281	285.65
$\hat{\bar{Y}}_{P9}$	406.97	271.48	219.84	173.58	154.24	156.74	170.28	226.281	285.65

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

SIMULATION STUDY

The numerical findings provide compelling evidence that the proposed estimators denoted as $\hat{\hat{Y}}_{Pi}$ and $\hat{\hat{Y}}_{Pj}$ demonstrate superior efficiency over other estimators under study. This enhanced performance is reflected in their consistently lower MSE values establishing them as more reliable and precise alternatives for estimating population mean. To further validate these results, a comprehensive Monte Carlo simulation study was carried out using R software. The objective of this simulation was to rigorously assess the comparative performance of the proposed estimators $\hat{\hat{Y}}_{Pi}$ and $\hat{\hat{Y}}_{Pj}$ against existing estimators $\hat{\hat{Y}}_{KCi}$. The simulation draws upon a real-life data set for 923 districts in Turkey in 2007, while the corresponding summary statistics are provided in Table 2. In order to evaluate robustness across different conditions, the simulation considers a range of sample sizes, i.e. $n = 190, 200, 210$ and 220. This variation helps to assess the estimators' stability and effectiveness across practical scenarios. Equation (27) is employed to compute the MSE values, serving as the basis for identifying the estimator with the minimum error. The results consistently show that the proposed estimators yield the lowest MSEs across all sample sizes, reinforcing their statistical efficiency and applicability in real-world data settings.

$$MSE\left(\hat{\hat{Y}}_l\right) = \frac{\sum_{l=1}^{20000}\left(\hat{\hat{Y}}_l - \bar{Y}\right)^2}{20000} \text{ where } l = \hat{\hat{Y}}_{Pi}, \hat{\hat{Y}}_{Pj} \text{ and } \hat{\hat{Y}}_{KCi} \qquad (27)$$

Procedure for computing minimum MSE of the estimators:

1. Generate N circular systematic samples, each containing n observations, corresponding to various values of r satisfying $1 \leq r \leq N$;
2. For each selected sample, compute the minimum MSE for every estimator;
3. Repeat the above two steps 20,000 times to obtain stable results;
4. This process produces 20,000 minimum MSE values for each estimator, namely $\hat{\hat{Y}}_{KCi}, \hat{\hat{Y}}_{Pi}$ and $\hat{\hat{Y}}_{Pj}$;
5. Finally, compute the mean of these 20,000 MSE values to determine the overall minimum MSE for each estimator.

Tables 5 and 6 present the minimum MSE values across varying sample sizes. Consistent with real data results, the simulation study confirms that: **(i)** $\hat{\hat{Y}}_{Pi}$ consistently outperforms $\hat{\hat{Y}}_{KCi}$ in terms of lower MSE; **(ii)** the proposed estimators $\hat{\hat{Y}}_{Pj}$ demonstrate the best overall performance among all existing and $\hat{\hat{Y}}_{Pi}$ estimators; and **(iii)** MSE values decrease as sample size increases, highlighting improved estimation precision with larger samples.

Table 5. MSE of $\hat{\hat{Y}}_{KCi}$ based on simulation study under simple random sampling

Estimator	$\hat{\hat{Y}}_{KC1}$	$\hat{\hat{Y}}_{KC2}$	$\hat{\hat{Y}}_{KC3}$	$\hat{\hat{Y}}_{KC4}$	$\hat{\hat{Y}}_{KC5}$
$n = 190$	2977.669	2976.767	2968.708	2977.621	2972.857
$n = 200$	2790.194	2789.349	2781.797	2790.148	2785.685
$n = 210$	2620.573	2619.779	2612.686	2620.530	2616.339
$n = 220$	2466.373	2465.626	2458.950	2466.332	2462.387

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

Table 6. MSE of proposed estimators based on simulation study

Estimator	$q_{.10}$	$q_{.20}$	$q_{.30}$	$q_{.40}$	$q_{.50}$	$q_{.60}$	$q_{.70}$	$q_{.80}$	$q_{.90}$
					$n = 190$				
$\hat{\hat{Y}}_{P1}$	622.903	854.276	997.370	1211.980	1467.566	1617.617	1794.256	2152.341	2410.078
$\hat{\hat{Y}}_{P2}$	622.635	853.950	997.013	1211.580	1467.544	1617.147	1793.760	2151.793	2409.497
$\hat{\hat{Y}}_{P3}$	620.238	851.035	993.819	1208.009	1467.099	1612.955	1789.324	2146.900	2404.300
$\hat{\hat{Y}}_{P4}$	622.888	854.259	997.351	1211.950	1467.520	1617.591	1794.229	2152.311	2410.047
$\hat{\hat{Y}}_{P5}$	621.472	852.536	995.463	1209.848	1465.172	1615.114	1791.608	2149.420	2406.976
$\hat{\hat{Y}}_{P6}$	378.129	253.917	205.756	161.323	140.294	140.300	149.763	194.373	243.482
$\hat{\hat{Y}}_{P7}$	378.129	253.917	205.756	161.323	140.294	140.300	149.763	194.373	243.482
$\hat{\hat{Y}}_{P8}$	378.129	253.917	205.756	161.323	140.294	140.300	149.763	194.373	243.482
$\hat{\hat{Y}}_{P9}$	378.129	253.917	205.756	161.323	140.294	140.300	149.763	194.373	243.482
					$n = 200$				
$\hat{\hat{Y}}_{P1}$	525.530	722.852	845.309	1029.237	1249.040	1378.213	1530.392	1839.265	2061.832
$\hat{\hat{Y}}_{P2}$	525.302	722.573	845.002	1029.033	1248.657	1377.909	1529.964	1838.793	2061.330
$\hat{\hat{Y}}_{P3}$	523.264	720.082	842.267	1025.967	1245.238	1374.198	1526.141	1834.569	2056.840
$\hat{\hat{Y}}_{P4}$	525.518	722.837	845.292	1029.358	1249.019	1378.191	1530.369	1839.240	2061.805
$\hat{\hat{Y}}_{P5}$	524.313	721.365	843.676	1027.546	1246.990	1376.057	1528.109	1836.744	2059.152
$\hat{\hat{Y}}_{P6}$	352.543	238.563	193.583	150.868	128.384	126.179	132.002	166.438	206.359
$\hat{\hat{Y}}_{P7}$	362.543	238.563	193.583	150.868	128.384	126.179	132.002	166.438	206.359
$\hat{\hat{Y}}_{P8}$	352.543	238.563	193.583	150.868	128.384	126.179	132.002	166.438	206.359
$\hat{\hat{Y}}_{P9}$	352.543	238.563	193.583	150.868	128.384	126.179	132.002	166.438	206.359

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

Table 6. (continued).

Estimator	$q_{.10}$	$q_{.20}$	$q_{.30}$	$q_{.40}$	$q_{.50}$	$q_{.60}$	$q_{.70}$	$q_{.80}$	$q_{.90}$
					$n = 210$				
$\hat{\hat{Y}}_{P1}$	438.630	605.240	709.075	865.579	1052.835	1163.139	1293.233	1557.665	1748.468
$\hat{\hat{Y}}_{P2}$	438.438	605.003	708.815	865.287	1052.509	1162.794	1292.867	1557.226	1748.037
$\hat{\hat{Y}}_{P3}$	436.724	602.893	706.493	862.677	1049.591	1159.710	1289.597	1553.642	1744.186
$\hat{\hat{Y}}_{P4}$	438.620	605.227	709.061	865.563	1052.818	1163.121	1293.215	1557.644	1748.445
$\hat{\hat{Y}}_{P5}$	437.606	603.979	707.689	864.021	1051.093	1161.298	1291.280	1555.505	1746.169
$\hat{\hat{Y}}_{P6}$	329.841	225.214	183.163	142.072	118.346	114.180	116.754	142.071	173.734
$\hat{\hat{Y}}_{P7}$	329.841	225.214	183.163	142.072	118.346	114.180	116.754	142.071	173.734
$\hat{\hat{Y}}_{P8}$	329.841	225.214	183.163	142.072	118.346	114.180	116.754	142.071	173.734
$\hat{\hat{Y}}_{P9}$	329.841	225.214	183.163	142.072	118.346	114.180	116.754	142.061	173.734
					$n = 220$				
$\hat{\hat{Y}}_{P1}$	361.044	499.647	586.813	718.341	876.223	969.421	1079.411	1303.621	1465.612
$\hat{\hat{Y}}_{P2}$	360.885	497.883	586.597	718.095	875.948	969.129	1079.180	1303.278	1465.246
$\hat{\hat{Y}}_{P3}$	359.463	499.834	584.647	715.898	873.485	966.521	1076.411	1300.207	1461.975
$\hat{\hat{Y}}_{P4}$	361.035	499.642	586.801	718.328	876.208	969.406	1079.473	1303.603	1465.593
$\hat{\hat{Y}}_{P5}$	360.195	498.791	584.647	717.029	874.752	967.864	1077.837	1301.788	1463.660
$\hat{\hat{Y}}_{P6}$	309.729	213.717	174.389	134.856	110.088	104.189	103.861	120.987	145.209
$\hat{\hat{Y}}_{P7}$	309.729	213.717	174.389	134.856	110.088	104.189	103.861	120.987	145.209
$\hat{\hat{Y}}_{P8}$	309.729	213.717	174.389	134.856	110.088	104.189	103.861	120.987	145.209
$\hat{\hat{Y}}_{P9}$	309.729	213.717	174.389	134.856	110.088	104.189	103.861	120.987	145.209

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

CONCLUSIONS

This study proposes novel ratio-type estimators for estimating population mean under CSS based on quantile regression approach. The findings show that the proposed estimators consistently perform better than the existing ones in terms of lower MSE. The findings also support the robustness and efficiency of the proposed estimators under CSS, particularly in data sets containing outliers. These estimators are highly recommended for use in future studies conducted under similar conditions. Potential directions for future research include:

- Estimating population median using both conventional and non-conventional auxiliary variables under CSS based on quantile regression approach,
- Estimating population mean using auxiliary information in the presence of non-sampling errors under CSS using quantile regression,
- Estimating population mean, median and variance using partial auxiliary information under various sampling schemes.

REFERENCES

1. W. G. Cochran, "Relative accuracy of systematic and stratified random samples for a certain class of populations", *Ann. Stat.*, **1946**, *17*, 164-177.
2. W. Gautschi, "Some remarks on systematic sampling", *Ann. Stat.*, **1957**, *28*, 385-394.
3. S. Sampath and V. Varalakshmi, "Diagonal circular systematic sampling", *Model Assist. Stat.*, **2008**, *3*, 345-352.
4. H. P. Singh and R. S. Solanki, "An efficient class of estimators for the population mean using auxiliary information in systematic sampling", *J. Stat. Theory Pract.*, **2012**, *6*, 274-285.
5. N. D. Shukla, "Systematic sampling and product method of estimation", Proceedings of All India Seminar on Demography and Statistics, **1971**, Varanasi, India.
6. J. Subramani and S. Singh, "Estimation of population mean in the presence of linear trend", *Commun. Stat.-Theory Meth.*, **2014**, *43*, 3095-3116.
7. H. K. Verma and R. Singh, "A family of efficient estimator in circular systematic sampling", *J. Adv. Comput.*, **2014**, *3*, 56-68.
8. S. Riaz, G. Diana and J. Shabbir, "Modified classes of estimators in circular systematic sampling", *Hacet J. Math. Stat.*, **2017**, *46*, 743-765.
9. H. P. Singh and A. Yadav, "A two-parameter ratio-product-ratio type exponential estimator for finite population mean in sample surveys", *Pak. J. Stat. Oper. Res.*, **2018**, *14*, 215-232.
10. S. A. Mostafa and I. A. Ahmad, "Recent developments in systematic sampling: A review", *J. Stat. Theory Pract.*, **2018**, *12*, 290-310.
11. Z. Khan, J. Shabbir, S. Gupta and A. Shamim, "An optimal systematic sampling scheme", *J. Stat. Comput. Simul.*, **2020**, *90*, 2023-2036.
12. H. P. Singh, A. Yadav and S. M. Gorey, "A generalized family of estimators for estimating finite population mean in circular systematic sampling (CSS)", *Commun. Stat.-Theory Meth.*, **2021**, *50*, 631-662.
13. H. P. Singh and A. Yadav, "Estimation of finite population mean through a two parameter ratio-product-ratio-type exponential estimator in systematic sampling", *J. Stat. Appl. Prob.*, **2023**, *12*, 277-293.

14. Z. Khan, L. A. Al-Essa, F. S. Al-Duais, J. Shabbir and S. Gupta, "An application of modified systematic sampling in auto-correlated populations", *Kuwait J. Sci.*, **2025**, *52*, Art.no.100404.
15. S. Pal, S. A. Mahmud and H. P. Singh, "An efficient estimation of finite population mean through difference estimator in systematic sampling", *Afr. Mat.*, **2025**, *36*, Art.no.14.
16. M. Irfan, F. Zunaira, M. Javed, S. C. Shongwe, S. H. Bhatti and M. A. Hussain, "Advancements in population mean estimation for circular systematic sampling", *Contemp. Math.*, **2025**, *6*, 5053-5072.
17. C. Kadilar and H. Cingi, "Ratio estimators in simple random sampling", *Appl. Math. Comput.*, **2004**, *151*, 893-902.
18. C. Kadilar and H. Cingi, "Ratio estimators using robust regression", *Hacet J. Math. Stat.*, **2007**, *36*, 181-188.
19. T. Zaman, "Improvement of modified ratio estimators using robust regression methods", *Appl. Math. Comput.*, **2019**, *348*, 627-631.
20. T. Zaman, U. Shazad and K. V. Yadav, "An efficient Hartley-Ross type estimators of nonsensitive and sensitive variables using robust regression methods in sample surveys", *J. Comput. Appl. Math.*, **2024**, *440*, Art.no.115645.
21. T. Zaman, H. Bulut and K. S. Yadav, "Robust ratio-type estimators for finite population mean in simple random sampling: A simulation study", *Concurr. Comput. Pract. Exp.*, **2022**, *34*, Art.no.e7273.
22. N. Ali, I. Ahmad, M. Hanif and U. Shahzad, "Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information", *Commun. Stat.-Theory Meth.*, **2021**, *50*, 979-992.
23. U. Shahzad, N. H. Al-Noor, M. Hanif and M. M. Anas, "Minimum covariance determinant-based quantile robust regression-type estimators for mean parameter", *Math. Prob. Eng.*, **2021**, https://doi.org/10.1155/2021/5255839.
24. U. Shahzad, M. Hanif, I. Sajjad and M. M. Anas, "Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information", *Sci. Iran*, **2022**, *29*, 1705-1715.
25. M. M. Anas, Z. Huang, U. Shahzad, T. Zaman and S. Shahzadi, "Compromised imputation based mean estimators using robust quantile regression", *Commun. Stat.-Theory Meth.*, **2022**, *53*, 1700-1715.
26. T. Koc and H. Koc, "A new class of quantile regression ratio-type estimators for finite population mean in stratified random sampling", *Axioms*, **2023**, *12*, Art.no.713.
27. 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.
28. U. Shahzad, I. Ahmad, I. M. Almanjahie, N. H. Al- Noor and M. Hanif, "Mean estimation using robust quantile regression with two auxiliary variables", *Sci. Iran*, **2023**, *30*, 1245-1254.
29. M. Khalid, H. Khan and J. Shabbir, "Exponentially quantile regression-ratio-type estimators for robust mean estimation", *Commun. Stat.-Theory Meth.*, **2023**, *53*, 7069-7086.
30. A. M. Alomair and S. Iftikhar "Dual-type quantile regression approach for mean estimation incorporating sampled and non-sampled data", *Heliyon*, **2024**, *10*, Art.no.e31034.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 67-82

31. M. Javed, M. Irfan and T. Pang, "Utilizing bivariate auxiliary information for enhanced estimation of population mean under simple and stratified random sampling schemes", *Natl. Sci. Found.*, **2019**, *47*, 199-211.
32. N. Buasiri and T. Mayureesawan, "The effect of sampling methods for linear regression estimation of population means of dependent variables", *Thai Stat.*, **2013**, *11*, 31-43.