Mathematical Sciences (2021) 15:29–40
https://doi.org/10.1007/s40096-020-00360-7

ORIGINAL RESEARCH

Generalized robust-regression-type estimators under different ranked set sampling

Nursel Koyuncu[1] · **Amer Ibrahim Al-Omari**[2]

Received: 27 July 2020 / Accepted: 1 November 2020 / Published online: 17 November 2020

Abstract

In this paper, we have proposed a new generalized robust estimators of population mean under different ranked set sampling. Robust estimators are recently defined by Zaman and Bulut (Commun Stat Theory Methods 48(8):2039–2048, 2019a) and Ali et al. (Commun Stat Theory Methods, 2019. https://doi.org/10.1080/03610926.2019.1645857) under simple random sampling. We have generalized robust-type estimators where Zaman and Bulut (2019a) and Ali et al. (2019) estimators are members of our generalized estimator. We have also extended our results to ranked set and median ranked set sampling designs. The simulation study showed that our proposed robust-type estimator performs better.

Keywords Ratio-type estimators · Regression-type estimators · Robust regression methods · Ranked set sampling · Median ranked set sampling

Mathematics Subject Classification 62D05 · 62F35

Introduction

When the data set does not follow a normal distribution or contains outliers, estimations of parameters are affected badly. To overcome this difficulty and get reliable conclusions about the contaminated data, robust estimators are defined in statistical analysis. This relies on finding proper estimates of the data location and scale ([9, 10, 13]).

Using information of auxiliary variable in the estimates also increases the efficiency. We can list some important studies as follows: Hanif and Shahzad [11] considered the issue of estimating the population variance utilizing trace of kernel matrix in absence of non-response under simple random sampling (SRS) scheme. Shahzad et al. [25] defined a new class of ratio-type estimators for the population mean. Shahzad et al. [26] proposed a new class of exponential-type estimators, based on the known median of study variable.

The authors also studied robust ratio-type estimators when the data contained outliers. Zaman and Bulut [30] have studied the robust estimators in simple random sampling and they also extended their studies to stratified simple random sampling [31]. Recently, Ali et al. [1] generalized Zaman and Bulut [30]'s estimators and they have studied sensitive data case. Subzar et al. [28] adapted the various robust regression techniques to the ratio estimators. Shahzad et al. [27] defined class of regression-type estimators utilizing robust regression tools.

Ranked set sampling (RSS) is an effective design introduced by [20]. The efficiency of RSS depends on the sampling allocation whether balanced or unbalanced. The balanced RSS features an equal allocation of the ranked order statistics. It has been shown theoretically and empirically the variance of the balanced RSS estimator is less than that of the estimator SRS estimator regardless of ranking errors ([3]). In literature, many authors showed that this design performs better compared to the SRS and proposed new RSS designs such as median ranked set (MRSS) by Muttlak [22], double ranked set by Al-Saleh and Al-Kadiri [7], pair ranked set by Muttlak [21], L ranked set by Al-Nasser [2], neoteric ranked set sampling by Zamanzade and Al-Omari [32] and so on.

✉ Nursel Koyuncu
nkoyuncu@hacettepe.edu.tr

Amer Ibrahim Al-Omari
alomari_amer@yahoo.com

[1] Department of Statistics, Hacettepe University, Ankara, Turkey

[2] Department of Mathematics, Faculty of Science, Al al-Bayt University, Mafraq, Jordan

We can also list some important papers to improve estimation under ranked set sampling designs as: Al-Omari [6] studied ratio estimators under MRSS. Koyuncu [17] studied regression-type estimators under different ranked set sampling. Koyuncu [18] has proposed regression-type estimators and more general class of estimators under MRSS. Koyuncu [16] has studied difference-cum-ratio and exponential type estimators under MRSS.

The aim of this study is to propose more general estimators of the population mean using robust statistics under RSS and MRSS.

The article is constructed as follows: In "Robust estimators in SRS" section, we have reviewed the recent robust literature in SRS. In "Proposed class of robust-regression-type-estimators under SRS, RSS and MRSS" section, the new generalized robust estimators under SRS, RSS and MRSS are presented. The mean square error (MSE) is derived up to the first-order of approximation. The theoretical efficiency comparison is given in "Efficiency comparison" section. In "Simulation study" section, a simulation study is conducted using a real data set and summarized our findings in "Conclusion" section.

Robust estimators in SRS

When the information about auxiliary variable is known and the using this in the estimator can results more efficient estimates. Also the normality assumption of data is not hold, we need to use robust statistics. Moving this direction Zaman and Bulut [30] suggested following robust estimators as follows:

$$\bar{y}_{z_1} = \frac{\bar{y} + b_{\text{rob}(zb)}(\bar{X} - \bar{x})}{\bar{x}}\bar{X}, \tag{2.1}$$

$$\bar{y}_{z_2} = \frac{\bar{y} + b_{\text{rob}(zb)}(\bar{X} - \bar{x})}{\bar{x} + C_x}(\bar{X} + C_x), \tag{2.2}$$

$$\bar{y}_{z_3} = \frac{\bar{y} + b_{\text{rob}(zb)}(\bar{X} - \bar{x})}{\bar{x} + \beta_2(x)}(\bar{X} + \beta_2(x)), \tag{2.3}$$

$$\bar{y}_{z_4} = \frac{\bar{y} + b_{\text{rob}(zb)}(\bar{X} - \bar{x})}{\bar{x}\beta_2(x) + C_x}(\bar{X}\beta_2(x) + C_x), \tag{2.4}$$

$$\bar{y}_{z_5} = \frac{\bar{y} + b_{\text{rob}(zb)}(\bar{X} - \bar{x})}{\bar{x}C_x + \beta_2(x)}(\bar{X}C_x + \beta_2(x)) \tag{2.5}$$

where b_i is slope or regression coefficient, calculated from the robust regression methods such as least absolute deviations (LAD), least median of squares method (LMS), least trimmed squares method (LTS), Huber-M, Hample-M, Tukey-M, Huber-MM. When the data contain outliers, these observations cause problems because they may strongly influence the result. Classical methods can be affected by outliers. The aim of using robust statistics in estimators is to describe well the data majority and get reliable estimates. We can summarized these well known robust methods are as follows:

LAD is a method which minimizes the sum of absolute error and is described as

$$\min \sum_{i=1}^{n} | \epsilon_i | .$$

LMS method rather than minimize the sum of the least-squares function, this model minimizes the median of the squared residuals

$$\min \quad \text{median}(\epsilon_i^2)$$

LTS proceeds with OLS after eliminating the most extreme positive or negative residuals. LTS orders the squared residuals from smallest to largest: $(E^2)_{(1)}, (E^2)_{(2)}, \dots, (E^2)_{(n)}$, and then, it calculates b that minimizes the sum of only the smaller half of the residuals

$$\sum_{i=1}^{m}(E^2)_{(i)}$$

where $m = [n/2] + 1$; the square bracket indicates rounding down to the nearest integer.

Huber-M is based on minimizing another function of outliers instead of error squares. Objective function of M-estimator is given as

$$\min \sum_{i=1}^{n} \rho(e_i)$$

and is asymmetric function of outliers. Huber's function ρ is designed as

$$\rho(e) = \begin{cases} \frac{e^2}{2} & | e | \leq k \\ k | e | - \frac{k^2}{2} & | e | > k \end{cases}$$

The influence function is determined by taking the derivative

$$\varphi(y) = \begin{cases} e & | e | \leq k \\ k\text{sign}(e) & | e | > k \end{cases}$$

where sgn(.) is sign function and represented as

$$sgn(x) = \begin{cases} -1 & e < k \\ 0 & e = k \\ 1 & e > k \end{cases}$$

Hample-M Estimation function is defined as

$$\rho(y)=\begin{cases}\frac{y^2}{2} & 0<\mid y\mid<a\\ a\mid y\mid-\frac{y^2}{2} & a<\mid y\mid\leq b\\ \frac{-a}{2(c-b)}(c-y)^2+\frac{a}{2}(b+c-a) & b<\mid y\mid\leq c\\ \frac{a}{2}(b+c-a) & c<\mid y\mid\end{cases}$$

where $a=1.7$, $b=3.4$ and $c=8.5$.

Tukey M estimation function is given by

$$\rho(y)=\begin{cases}\frac{1}{6}(1-(1-(\frac{y}{k})^2)^3) & \mid y\mid\leq k\\ \frac{1}{6} & \mid y\mid>k\end{cases}$$

where $k=5$ or $k=6$.

Huber MM estimation method is described as follows:

- A starting estimation with high breakdown point (0.5 if possible) is chosen.
- Outliers are calculated as $e_i(T_0)=y_i-T_0x_i, \quad 1\leq i\leq n$

where T_0 is starting estimation. Under $b/a=0.5$ constraints, b is calculated as below

$$b=\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{e_i(\beta)}{s_n}\right)$$

where s_n is M scale estimation and it is calculated as $s_n=s(e(T_0))$. For more information robust estimators kindly see [12, 24, 29])

Ali et al. [1] generalized Zaman and Bulut [30] estimators as

$$\bar{y}_z=\frac{\bar{y}+b_i(\bar{X}-\bar{x})}{(F\bar{x}+G)}(F\bar{X}+G)\quad \text{for}\quad i=1,2,\ldots,7 \tag{2.6}$$

where i shows robust regression methods LAD, LMS, LTS, Huber-M, Hample-M, Tukey-M, Huber-MM, respectively.

$F\neq 0$ and G are any constants, either (0,1) or known characteristics of the population such as, C_x, the coefficients of variation, $\beta_2(x)$, the coefficients of kurtosis from the population having N identifiable units. We can generate many new estimators using suitable variables for b_i, F and G.

The MSE of $\bar{y}_z$ is given as

$$\text{MSE}(\bar{y}_z)=\left(\frac{1-f}{n}\right)\Big[S_y^2+R_{FG}^2S_x^2+2B_iR_{FG}S_x^2+B_i^2S_x^2 - 2R_{FG}S_{xy}-2B_iS_{xy}\Big] \tag{2.7}$$

where B_i robust regression coefficient of population, $R_{FG}=\frac{F\bar{Y}}{F\bar{X}+G}, S_y^2=\frac{\sum_{i=1}^{N}(y_i-\bar{Y})^2}{N-1}, S_x^2=\frac{\sum_{i=1}^{N}(x_i-\bar{X})^2}{N-1}$ are the unbiased variances of Y and X, respectively, $S_{xy}=\frac{\sum_{i=1}^{N}(y_i-\bar{Y})(x_i-\bar{X})}{N-1}$ is the covariance between (X, Y). $\bar{X}, \bar{Y}$ are population means $\bar{x}, \bar{y}$ are the sample means of auxiliary and study variables, respectively.

Then, Ali et al. [1] have proposed classical regression-type estimators using robust b in the equation instead of least-square estimator. They found that their regression-type estimators are more efficient than Zaman and Bulut [30]. Ali et al. [1] estimator and MSE of estimator are given as

$$\bar{y}_a=\bar{y}+b_i(\bar{X}-\bar{x})\quad \text{for}\quad i=1,2,\ldots,7 \tag{2.8}$$

$$\text{MSE}(\bar{y}_a)=\left(\frac{1-f}{n}\right)\left[S_y^2-2B_iS_{xy}+B_i^2S_x^2\right] \tag{2.9}$$

Proposed class of robust-regression-type-estimators under SRS, RSS and MRSS

Ranked-based sampling designs have found many applications in other fields including environmental monitoring [19], clinical trials and genetic quantitative trait loci mappings [8] and medicine ([33, 34]). In this section, RSS and MRSS designs are described and we have introduced a new generalized robust estimators under these designs.

RSS design

The RSS design introduced by McIntyre [20] can be described as follows:

1. Select a simple random sample of size m^2 units from the target finite population and divide them into m samples each of size m.
2. Rank the units within each sample in increasing magnitude by using personal judgment, eye inspection or based on a concomitant variable.
3. Select the ith ranked unit from the ith sample.
4. Repeat steps 1 through 3, k times if needed to obtain a RSS of size $n=mk$.

Let

$$\begin{array}{ll}(X_{11j},Y_{11j}), & (X_{12j},Y_{12j}),\ldots,(X_{1mj},Y_{1mj});\\ (X_{21j},Y_{21j}), & (X_{22j},Y_{22j}),\ldots,(X_{2mj},Y_{2mj});\\ \vdots; & \\ (X_{m1j},Y_{m1j}), & (X_{m2j},Y_{m2j}),\ldots,(X_{mmj},Y_{mmj})\end{array}$$

be m independent bivariate random samples with pdf $f(x, y)$, each of size m in the jth cycle, $(j=1,2,\ldots,k)$. Let

$(X_{i(1:m)j}, Y_{i[1:m]j}), (X_{i(2:m)j}, Y_{i[2:m]j}), \dots,$
$(X_{i(m:m)j}, Y_{i[m:m]j})$

be the order statistics of $X_{i1j}, X_{i2j}, \dots, X_{imj}$ and the judgment order of $Y_{i1j}, Y_{i2j}, \dots, Y_{imj}$ $(i = 1, 2, \dots, m)$, where round parentheses () and square brackets [] indicate that the ranking of X is perfect and ranking of Y has errors. Assume measured units using RSS are

$(X_{1(1:m)j}, Y_{1[1:m]j}), (X_{2(2:m)j}, Y_{2[2:m]j}), \dots, (X_{m(m:m)j}, Y_{m[m:m]j}).$

Then, the RSS estimators of population mean for the study and auxiliary variables can be written as

$$\bar{y}_{[\mathrm{RSS}])} = \frac{1}{mk} \sum_{j=1}^{k} \sum_{i=1}^{m} Y_{i[i:m]j}, \tag{3.1}$$

$$\bar{x}_{(\mathrm{RSS})} = \frac{1}{mk} \sum_{j=1}^{k} \sum_{i=1}^{m} X_{i(i:m)j}. \tag{3.2}$$

MRSS design

The MRSS design can be described as in the following steps:

1. Select m random samples each of size m bivariate units from the population of interest.
2. The units within each sample are ranked by visual inspection or any other cost free method with respect to a variable of interest.
3. If m is odd, select the $((m+1)/2)$ th-smallest ranked unit X together with the associated Y from each set, i.e., the median of each set. If m is even, from the first $m/2$ sets select the $(m/2)$th ranked unit X together with the associated Y and from the other sets select the $((m+2)/2)$ the ranked unit X together with the associated Y.
4. The whole process can be repeated k times if needed to obtain a sample of size $n = mk$ units.

For odd and even sample sizes, the units measured using MRSS are denoted by MRSSO and MRSSE, respectively. For odd sample size,

$(X_{1(\frac{m+1}{2}:m)j}, Y_{1[\frac{m+1}{2}:m]j}), (X_{2(\frac{m+1}{2}:j)}, Y_{2[\frac{m+1}{2}:j]}), \dots,$
$(X_{m(\frac{m+1}{2}:j)}, Y_{m[\frac{m+1}{2}]j})$

the sample mean estimators using MRSSO are given,

$$\bar{x}_{(\mathrm{MRSSO})} = \frac{1}{mk} \sum_{j=1}^{k} \sum_{i=1}^{m} X_{i(\frac{m+1}{2}:m)j} \tag{3.3}$$

$$\bar{y}_{[\mathrm{MRSSO}]} = \frac{1}{mk} \sum_{j=1}^{k} \sum_{i=1}^{m} Y_{i[\frac{m+1}{2}:m]j}. \tag{3.4}$$

For even sample size,

$(X_{1(\frac{m}{2}:m)j}, Y_{1[\frac{m}{2}:m]j}), (X_{2(\frac{m}{2}:m)j}, Y_{2[\frac{m}{2}:m]j}), \dots,$
$(X_{m(\frac{m}{2}:m)j}, Y_{m[\frac{m}{2}:m]j}),$
$(X_{\frac{m+2}{2}(\frac{m+2}{2}:m)j}, Y_{\frac{m+2}{2}[\frac{m+2}{2}:m]j}),$
$(X_{\frac{m+4}{2}(\frac{m+4}{2}:m)j}, Y_{\frac{m+4}{2}[\frac{m+4}{2}:m]j}), \dots, (X_{m(\frac{m}{2}:m)j}, Y_{m[\frac{m}{2}:m]j})$

the sample means of X and Y using MRSSE are given

$$\bar{x}_{(\mathrm{MRSSE})} = \frac{1}{mk} \sum_{j=1}^{k} \left(\sum_{i=1}^{\frac{m}{2}} X_{i(\frac{m}{2}:m)j} + \sum_{i=\frac{m+2}{2}}^{m} X_{i(\frac{m+2}{2}:m)j} \right) \tag{3.5}$$

and

$$\bar{y}_{[\mathrm{MRSSE}]} = \frac{1}{mk} \sum_{j=1}^{k} \left(\sum_{i=1}^{\frac{m}{2}} Y_{i[\frac{m}{2}:m]j} + \sum_{i=\frac{m+2}{2}}^{m} Y_{i[\frac{m+2}{2}:m]j} \right) \tag{3.6}$$

Proposed class of robust-regression-type-estimators in SRS

In this section, we can generalized Zaman et al. [30] and Ali et al. [1] estimators under SRS as

$$\bar{y}_{N(\mathrm{SRS}))} = [\bar{y} + b_i(\bar{X} - \bar{x})]\left(\frac{F\bar{X} + G}{F\bar{x} + G}\right)^{\alpha} \tag{3.7}$$

To obtain the MSE of the generalized estimators, let us define the following expectations under SRS:

$e_0 = (\bar{y} - \bar{Y})/\bar{Y}, \quad e_1 = (\bar{x} - \bar{X})/\bar{X}.$

We can re-write the $\bar{y}_{N(\mathrm{SRS})}$ using e terms according to first order of approximation as follows:

$$\begin{aligned}(\bar{y}_{N(\mathrm{SRS})} - \bar{Y})^2 &\cong \bar{Y}^2 e_0^2 + B_i^2 \bar{X}^2 e_1^2 + \bar{Y}^2 \alpha^2 R_i^2 e_1^2 \\ &\quad - 2B_i \bar{X}\bar{Y} e_0 e_1 - 2\bar{Y}^2 \alpha R_i e_0 e_1 + 2\alpha B_i \bar{X}\bar{Y} R_i e_1^2\end{aligned} \tag{3.8}$$

The expectations of e terms are

$$E(e_0^2) = \frac{(1-f)}{n}\frac{S_y^2}{\bar{Y}^2}, \quad E(e_1^2) = \frac{(1-f)}{n}\frac{S_x^2}{\bar{X}^2},$$
$$E(e_0e_1) = \frac{(1-f)}{n}\frac{S_{yx}}{\bar{X}\bar{Y}}$$

Substituting these expectations in Eq. 3.8, we can get MSE as given by

$$\begin{aligned}\text{MSE}(\bar{y}_{N(\text{SRS})}) \cong \frac{(1-f)}{n}[S_y^2 + B_i^2S_x^2 + \alpha^2R_{FG}^2S_x^2 \\ - 2B_iS_{yx} - 2\alpha R_{FG}S_{yx} + 2\alpha B_iR_{FG}S_x^2]\end{aligned} \tag{3.9}$$

Putting α=0 in the MSE($\bar{y}_{N(\text{SRS})}$), we can get Ali et al. [1] MSE($\bar{y}_a$) estimator. While for appropriate constants, we can also get MSEs of Zaman and Bulut [30] estimators.

Proposed class of robust-regression-type-estimators in RSS and MRSS

In RSS and MRSS to improve the efficiency of estimators, some authors used information of auxiliary variable in estimators. Koyuncu [18] has proposed regression-type estimators under MRSS. Al-Omari and Bouza [4] studied the ratio estimators of population mean with missing values using RSS. Al-Omari et al. [5] suggested ratio-type estimators of the mean using extreme RSS. Jemain et al. [15] suggested a modified ratio estimator for the population mean using double MRSS.

In this section, we generalized also our proposed estimators in the previous section to new classes based on RSS and MRSS designs as follows

$$\bar{y}_{N(j)} = \left[\bar{y}_{(j)} + b_{i(j)}(\bar{X} - \bar{x}_{(j)})\right]\left(\frac{F\bar{X}+G}{F\bar{x}_{(j)}+G}\right)^{\alpha} \tag{3.10}$$

where (j) represents the sampling design such as SRS, RSS and MRSS.

To obtain the bias and the MSE of suggested class of estimators in Eq. (3.10) under RSS, let us define following notations

To obtain the MSE of the generalized estimators, let us define the following expectations under SRS:

$$e_{0(j)} = (\bar{y}_{(j)} - \bar{Y})/\bar{Y}, \quad e_{1(j)} = (\bar{x}_{(j)} - \bar{X})/\bar{X}$$

$$\begin{aligned}\text{MSE}(\bar{y}_{N(j)}) \cong E\Big[\bar{Y}^2e_{0(j)}^2 + B_i^2\bar{X}^2e_{1(j)}^2 + \alpha^2\psi^2\bar{Y}^2e_{1(j)}^2 \\ - 2B_i\bar{Y}\bar{X}e_{0(j)}e_{1(j)} - 2\alpha\bar{Y}^2\psi e_{0(j)}e_{1(j)} \\ + 2B_i\alpha\psi\bar{Y}\bar{X}e_{1(j)}^2\Big]\end{aligned} \tag{3.11}$$

where $\psi = \frac{F\bar{X}}{F\bar{X}+G}$. One can easily obtain the specific MSE from Eq. 3.11 putting expectation terms belong to design. For example if (j) represents the RSS design we can write following notations:

$$\begin{aligned}
e_{0(\text{RSS})} &= (\bar{y}_{(\text{RSS})} - \bar{Y})/\bar{Y}, \\
e_{1(\text{RSS})} &= (\bar{x}_{(\text{RSS})} - \bar{X})/\bar{X} \\
E(e_{0(\text{RSS})}^2) &= \frac{var(\bar{y}_{\text{RSS}})}{\bar{Y}^2} \\
&= \frac{1}{\bar{Y}^2}\left[\frac{S_y^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{y[i]} - \bar{Y})^2\right], \\
E(e_{1(\text{RSS})}^2) &= \frac{var(\bar{x}_{\text{RSS}})}{\bar{X}^2} \\
&= \frac{1}{\bar{X}^2}\left[\frac{S_x^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{x(i)} - \bar{X})^2\right], \\
E(e_{0(\text{RSS})}e_{1(\text{RSS})}) &= \frac{cov(\bar{x}_{\text{RSS}}, \bar{y}_{\text{RSS}})}{\bar{X}\bar{Y}} \\
&= \frac{1}{\bar{X}\bar{Y}}\left[\frac{S_{xy}}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{y[i]} - \bar{Y})(\mu_{x(i)} - \bar{X})\right].
\end{aligned}$$

$$\begin{aligned}
\text{MSE}(\bar{y}_{N(\text{RSS})}) \cong \Bigg[&\left[\frac{S_y^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{y[i]} - \bar{Y})^2\right] \\
&+B_i^2\left[\frac{S_x^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{x(i)} - \bar{X})^2\right] \\
&+\alpha^2R_{FG}^2\left[\frac{S_x^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{x(i)} - \bar{X})^2\right] \\
&-2B_i\left[\frac{S_{xy}}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{y[i]} - \bar{Y})(\mu_{x(i)} - \bar{X})\right] \\
&-2\alpha R_{FG}\left[\frac{S_{xy}}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{y[i]} - \bar{Y})(\mu_{x(i)} - \bar{X})\right] \\
&+2B_i\alpha R_{FG}\left[\frac{S_x^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{x(i)} - \bar{X})^2\right]\Bigg]
\end{aligned} \tag{3.12}$$

If (j) represents the MRSS design and the sample size n is odd we can write following notations:

$$\begin{aligned}
e_{0(\text{MRSS}_o)} &= \frac{\bar{y}_{\text{MRSS}_o} - \bar{Y}}{\bar{Y}} \quad \text{and} \\
e_{1(\text{MRSS}_o)} &= \frac{\bar{x}_{\text{MRSS}_o} - \bar{X}}{\bar{X}} \\
E(e_{0(\text{MRSS}_o)}^2) &= \frac{1}{n\bar{Y}^2}S^2_{y[\frac{n+1}{2}]}, \\
E(e_{1(\text{MRSS}_o)}^2) &= \frac{1}{n\bar{X}^2}S^2_{x(\frac{n+1}{2})}, \\
E(e_{0(\text{MRSS}_o)}e_{1(\text{MRSS}_o)}) &= \frac{1}{n\bar{X}\bar{Y}}S_{xy(\frac{n+1}{2})}
\end{aligned}$$

Then, we can get the MSE as follows:

$$\begin{aligned}\text{MSE}(\bar{y}_{N(\text{MRSS}_o)})\\ \cong \frac{1}{n}\Bigg[& S^2_{y\left[\frac{n+1}{2}\right]} + B_i^2 S^2_{x\left(\frac{n+1}{2}\right)} + \alpha^2 R_{FG}^2 S^2_{x\left(\frac{n+1}{2}\right)} \\ & -2B_i S_{xy\left(\frac{n+1}{2}\right)} - 2\alpha R_{FG} S_{xy\left(\frac{n+1}{2}\right)} + 2B_i \alpha R_{FG} S^2_{x\left(\frac{n+1}{2}\right)}\Bigg]\end{aligned} \tag{3.13}$$

If (j) represents the MRSS design and the sample size n is even we can write following notations:

$$E(e^2_{0(\text{MRSS}_e)}) = \frac{1}{2n\bar{Y}^2}\left(S^2_{y\left[\frac{n}{2}\right]} + S^2_{y\left[\frac{n+2}{2}\right]}\right),$$

$$E(e^2_{1(\text{MRSS}_e)}) = \frac{1}{2n\bar{X}^2}\left(S^2_{x\left(\frac{n}{2}\right)} + S^2_{x\left(\frac{n+2}{2}\right)}\right) \quad \text{and}$$

$$E(e_{0(\text{MRSS}_e)} e_{1(\text{MRSS}_e)}) = \frac{1}{2n\bar{X}\bar{Y}}\left(S_{yx\left(\frac{n}{2}\right)} + S_{yx\left(\frac{n+2}{2}\right)}\right).$$

After defining these terms, the procedure is very easy. When we put these e terms in Eq. 3.11 we get the MSE is given by

$$\begin{aligned}\text{MSE}(\bar{y}_{N(\text{MRSS}_e)})\\ \cong \Bigg[& \left(S^2_{y\left[\frac{n}{2}\right]} + S^2_{y\left[\frac{n+2}{2}\right]}\right) + B_i^2\left(S^2_{x\left(\frac{n}{2}\right)} + S^2_{x\left(\frac{n+2}{2}\right)}\right) \\ & + \alpha^2 R^2_{FG}\left(S^2_{x\left(\frac{n}{2}\right)} + S^2_{x\left(\frac{n+2}{2}\right)}\right) \\ & -2B_i\left(S_{yx\left(\frac{n}{2}\right)} + S_{yx\left(\frac{n+2}{2}\right)}\right) \\ & -2\alpha R_{FG}\left(S_{yx\left(\frac{n}{2}\right)} + S_{yx\left(\frac{n+2}{2}\right)}\right) \\ & +2B_i \alpha R_{FG}\left(S^2_{x\left(\frac{n}{2}\right)} + S^2_{x\left(\frac{n+2}{2}\right)}\right)\Bigg]\end{aligned} \tag{3.14}$$

Efficiency comparison

In this section, the efficiency comparison between the MSE equations is obtained as below:

(1) Comparison of Ali et al. [1] $\bar{y}_z$ estimator which is general form of Zaman and Bulut [30] with proposed estimators under SRS

$\text{MSE}(\bar{y}_{N(\text{SRS})}) < \text{MSE}(\bar{y}_z)$ if

$$\begin{aligned}&\alpha^2 R^2_{FG} S^2_x + 2\alpha B_i R_{FG} S^2_x - 2\alpha R_{FG} S_{yx} < R^2_{FG} S^2_x \\ &\quad + 2B_i R_{FG} S^2_x - 2R_{FG} S_{xy} \\ &(\alpha^2 - 1) R^2_{FG} S^2_x + 2(\alpha - 1)(B_i R_{FG} S^2_x + R_{FG} S_{yx}) < 0\end{aligned} \tag{4.1}$$

From Eq. (4.1), one can easily see that when the $\alpha = 0$ we get same MSE.

(2) Comparison of Ali et al. [1] $\bar{y}_a$ estimator with proposed estimators under SRS

$\text{MSE}(\bar{y}_{N(\text{SRS})}) < \text{MSE}(\bar{y}_a)$ if

$$\alpha[\alpha R^2_{FG} S^2_x - 2R_{FG} S_{yx} + 2B_i R_{FG} S^2_x] < 0 \tag{4.2}$$

When the conditions Eqs. (4.1–4.2) are satisfied the proposed estimator is more efficient than existing estimators.

Simulation study

In this section, we used a real data set which is used by Jemain et al. [14] to illustrate the efficiency of proposed estimators over existing ones. These data consist of height and the diameter at breast height of 399 trees. We have used Height (H) in feet as study variable and Diameter (D) as auxiliary variable. The summary statistics of data is given in Table 1. The scatter plot of the data set is given in Fig. 1. From the scatter plot, we can say that the data set contains outliers. In the simulation study, we have selected 10000 samples with different sample sizes ($n = 3, 4, 5, 7, 8$) under SRS, RSS and MRSS designs using R Software version 3.1.1 [23] We have calculated the MSE of Zaman and Bulut [30], Ali et al. [1] and the proposed estimators under SRS which is given by Table 2 based on linear, Huber-M, LMS, LTS, LAD, S and MM robust estimators. In this table, we

Table 1 Summary of tree data

$N = 399$	$\bar{Y} = 52.33835$	$\bar{X} = 21.08897$
$\rho = 0.8762623$	$\sigma^2_y = 3262.817$	$\sigma^2_x = 329.6326$
$\sigma_{xy} = 908.7525$	$q_1 = 6.25$	$q_3 = 34$
$C_x = 0.8609138$	$\beta_2(x) = 0.5612123$	

Table 2 MSE of estimators under SRS

Type of estimators	n	Robust betas $b_{\text{rob}(zb)}$						
		Linear	Huber	LMS	LTS	LAD	S	MM
Zaman and Bulut [30]	3	3316.163	2717.506	1692.179	1992.434	2554.539	1839.175	1794.209
$\bar{y}_{z_1} = \frac{\bar{y}+b_{\text{rob}(zb)}(\bar{X}-\bar{x})}{\bar{x}}\bar{X}$	4	1822.856	1497.67	942.3292	1104.674	1409.246	1021.772	997.4632
	5	1210.12	997.247	634.5172	740.4148	939.4121	686.3196	670.4645
	7	707.1964	585.1299	377.6735	438.1449	551.999	407.2423	398.1896
	8	556.9105	463.5243	304.698	351.0141	438.1709	327.3478	320.4139
Zaman and Bulut [30]	3	2518.164	2052.397	1261.701	1492.024	1926.034	1374.302	1339.824
$\bar{y}_{z_2} = \frac{\bar{y}+b_{\text{rob}(zb)}(\bar{X}-\bar{x})}{\bar{x}+C_x}(\bar{X}+C_x)$	4	1454.437	1189.396	740.8671	871.272	1117.576	804.5877	785.0696
	5	1000.186	820.9085	518.2033	606.091	772.3704	561.1328	547.9798
	7	607.2781	500.6415	321.0498	373.1111	471.7985	346.4689	338.6785
	8	485.7631	403.0742	263.7009	304.1234	380.7016	283.44	277.3909
Proposed	3	946.4828	754.1013	445.8923	532.4438	703.0277	487.7866	474.868
$\bar{y}_{N(\text{SRS})_1} = \frac{\bar{y}+b_{\text{rob}(zb)}(\bar{X}-\bar{x})}{\bar{x}+q_1}(\bar{X}+q_1)$	4	628.2487	504.5441	307.2425	362.4848	471.7566	333.9598	325.7164
	5	479.707	387.8118	241.6403	282.4927	363.4794	261.3881	255.2928
	7	325.2891	264.5348	168.1446	195.0376	248.463	181.1382	177.1263
	8	276.4936	226.9415	148.1996	170.1922	213.8256	158.8287	155.5476
Proposed	3	289.2675	241.0532	189.9096**	199.4164	229.8427	193.8569	192.4968
$\bar{y}_{N(\text{SRS})_2} = \frac{\bar{y}+b_{\text{rob}(zb)}(\bar{X}-\bar{x})}{\bar{x}+q_3}(\bar{X}+q_3)$	4	217.3862	182.9273	146.8066**	153.3996	174.9415	149.5193	148.5783
	5	179.7786	152.4561	124.266**	129.2835	146.1516	126.304	125.5902
	7	130.2347	110.5632	90.36465**	93.93151	106.03	91.8074	91.30054
	8	119.6148	102.5568	84.82652***	88.02005	98.61282	86.13167	85.67665
Ali et al. [1]	3	256.5864	262.9883	321.3053	295.9877	267.615	307.8455	311.7328
$\bar{z}_a = \bar{y} + b_i(\bar{X}-\bar{x})$	4	183.3275	188.2805	231.4627	212.779	191.7421	221.5353	224.4035
	5	148.5035	153.3822	189.8203	174.263	156.4206	181.5723	183.9589
	7	101.9058	105.5672	132.1744	120.8439	107.8024	126.1699	127.9079
	8	95.01786	98.08421	121.0242	111.2284	99.99623	115.8307	117.3335

**Shows most efficient estimators according to sample size

***Shows most efficient estimators according to all cases

can also see the efficiency of different robust betas in the same type estimator. The same procedure is also extended for the RSS and MRSS estimators as given Tables 3 and 4. We can summarized the simulation study as follows:

- From Table 2, we can say that $\bar{y}_{N(\text{SRS})_2}$ proposed estimator which used third quantile of auxiliary variable and LMS robust beta is the best for all sample sizes.
- From Table 3, it can be seen that $\bar{y}_{N(\text{RSS})_4}$ proposed estimator which used third quantile of auxiliary variable and LMS robust beta is the best for all sample sizes under RSS design.
- From Table 4, under MRSS design $\bar{y}_{N(\text{MRSS})_4}$ is the best using third quantile and LMS robust beta when the sample size $n = 3, 4, 5$. When the sample size is increasing, proposed $\bar{y}_{N(\text{MRSS})_4}$ with S robust beta is perform better.
- PRE of estimators over different sampling designs is given in Tables 5 and 6. From these tables, we can see that ranked set sampling designs have better performance over SRS for all simulation cases.

Table 3 MSE of estimators under RSS

Type of estimator	n	Robust betas						
		Linear	Huber	LMS	LTS	LAD	S	MM
Proposed estimators under RSS	3	1201.504	1001.184	660.5236	759.8591	946.801	709.1006	694.2293
$\bar{y}_{N(\mathrm{RSS})_1} = \frac{\bar{y}_{(\mathrm{RSS})} + b_{\mathrm{rob}(zb)(\mathrm{RSS})}(\bar{X} - \bar{x}_{(\mathrm{RSS})})}{\bar{x}_{(\mathrm{RSS})}} \bar{X}$	4	558.0791	475.4833	336.1066	376.5585	453.1262	355.8637	349.81
	5	351.5762	305.2557	227.2356	249.8543	292.7264	238.2795	234.8949
	7	200.7543	179.3209	143.0137	153.5759	173.5108	148.1756	146.5947
	8	163.9655	148.3553	121.7979	129.5439	144.1168	125.5861	124.4264
	3	1005.839	836.6798	551.8172	634.3906	790.9273	592.1334	579.7772
$\bar{y}_{N(\mathrm{RSS})_2} = \frac{\bar{y}_{(\mathrm{RSS})} + b_{\mathrm{rob}(zb)(\mathrm{RSS})}(\bar{X} - \bar{x}_{(\mathrm{RSS})})}{\bar{x}_{(\mathrm{RSS})} + C_{x(\mathrm{RSS})}} (\bar{X} + C_{x(\mathrm{RSS})})$	4	497.8215	424.4606	301.8816	337.2438	404.6772	319.1249	313.8354
	5	323.7731	281.498	210.9677	231.2954	270.1041	220.8774	217.8369
	7	190.4861	170.3872	136.6323	146.4006	164.9566	141.3995	139.938
	8	157.2872	142.4989	117.5384	124.7838	138.4956	121.0772	119.9929
	3	514.8465	428.6654	292.6133	330.4445	405.9088	310.8745	305.2324
$\bar{y}_{N(\mathrm{RSS})_3} = \frac{\bar{y}_{(\mathrm{RSS})} + b_{\mathrm{rob}(zb)(\mathrm{RSS})}(\bar{X} - \bar{x}_{(\mathrm{RSS})})}{\bar{x}_{(\mathrm{RSS})} + q_{1(\mathrm{RSS})}} (\bar{X} + q_{1(\mathrm{RSS})})$	4	318.4984	274.5989	205.8316	224.8522	263.0396	214.9991	212.1636
	5	231.7789	203.9525	160.3374	172.406	196.6239	166.1549	164.3557
	7	152.1012	137.4716	114.2346	120.7227	133.5999	117.3701	116.4021
	8	131.2887	120.0253	101.9496	107.0314	117.0332	104.4102	103.6516
	3	242.8171	218.6456	195.5329**	199.1191	213.1793	196.8773	196.3773
$\bar{y}_{N(\mathrm{RSS})_4} = \frac{\bar{y}_{(\mathrm{RSS})} + b_{\mathrm{rob}(zb)(\mathrm{RSS})}(\bar{X} - \bar{x}_{(\mathrm{RSS})})}{\bar{x}_{(\mathrm{RSS})} + q_{3(\mathrm{RSS})}} (\bar{X} + q_{3(\mathrm{RSS})})$	4	192.2831	177.3924	162.7855**	165.1668	174.0026	163.7063	163.3717
	5	154.4981	143.9715	133.3354**	135.1636	141.5563	134.0642	133.8054
	7	112.2813	105.8765	98.74063**	100.1632	104.3665	99.35083	99.14556
	8	101.9688	96.46909	89.83767***	91.29443	95.14176	90.48815	90.27563
	3	239.0086	244.0064	278.7986	264.0452	246.9646	270.9858	273.2483
$\bar{y}_{N(\mathrm{RSS})_5} = \bar{y}_{(\mathrm{RSS})} + b_{\mathrm{rob}(zb)(\mathrm{RSS})}(\bar{X} - \bar{x}_{(\mathrm{RSS})})$	4	190.3964	192.2727	211.3533	203.0046	193.7498	206.9092	208.1916
	5	149.3587	150.4655	163.2738	157.6241	151.4315	160.2625	161.1306
	7	105.9015	105.9851	111.982	109.2073	106.3644	110.4919	110.9193
	8	95.7731	95.13028	98.42656	96.69609	95.22311	97.48048	97.74852

**Shows most efficient estimators according to sample size

***Shows most efficient estimators according to all cases

- When we compare the efficiency according to sample size, we can say that for all sample sizes the proposed estimator is more efficient than existing estimators. Especially for small sample sizes, efficiency over existing estimators is quite high than large sample sizes.

Conclusion

When the data contain outliers or not hold normality assumption to get more reliable results on estimates, we need to consider robust statistics. In this study, moving

Table 4 MSE of estimators under MRSS

Type of estimator	n	Robust betas						
		Linear	Huber	LMS	LTS	LAD	S	MM
Proposed estimators under MRSS	3	1889.569	1552.346	976.5284	1144.846	1460.654	1058.892	1033.689
$\bar{y}_{N(\mathrm{MRSS})_1} = \frac{\bar{y}_{(\mathrm{MRSS})}+b_{\mathrm{rob}(zb)(\mathrm{MRSS})}(\bar{X}-\bar{x}_{(\mathrm{MRSS})})}{\bar{x}_{(\mathrm{MRSS})}}\bar{X}$	4	670.2312	548.5602	342.8301	402.6134	515.6008	372.0385	363.0909
	5	788.6304	645.8916	405.3352	475.0983	607.2737	439.4011	428.9615
	7	513.9472	416.4074	254.2308	300.873	390.1524	276.9556	269.9805
	8	344.692	274.714	159.6635	192.5192	255.9571	175.6407	170.73
	3	1547.773	1264.804	785.9423	925.165	1188.126	853.9715	833.1336
$\bar{y}_{N(\mathrm{MRSS})_2} = \frac{\bar{y}_{(\mathrm{MRSS})}+b_{\mathrm{rob}(zb)(\mathrm{MRSS})}(\bar{X}-\bar{x}_{(\mathrm{MRSS})})}{\bar{x}_{(\mathrm{MRSS})}+C_{x(\mathrm{MRSS})}}(\bar{X}+C_{x(\mathrm{MRSS})})$	4	578.2221	470.5606	290.2786	342.3567	441.5034	315.682	307.8913
	5	679.2632	553.2658	343.1026	403.6661	519.31	372.626	363.5675
	7	449.5154	361.7994	217.654	258.8061	338.2921	237.6639	231.5134
	8	302.5978	239.1196	136.0931	165.2725	222.1864	150.2504	145.8921
	3	685.2468	547.0278	326.4918	388.2553	510.3881	356.3652	347.1484
$\bar{y}_{N(\mathrm{MRSS})_3} = \frac{\bar{y}_{(\mathrm{MRSS})}+b_{\mathrm{rob}(zb)(\mathrm{MRSS})}(\bar{X}-\bar{x}_{(\mathrm{MRSS})})}{\bar{x}_{(\mathrm{MRSS})}+q_{1(\mathrm{MRSS})}}(\bar{X}+q_{1(\mathrm{MRSS})})$	4	296.9665	235.2678	138.4248	165.2477	219.0099	151.3579	147.3589
	5	345.7691	274.8707	164.5515	194.924	256.2474	179.1711	174.6451
	7	236.1753	184.152	105.3714	126.6455	170.6188	115.5544	112.3894
	8	157.8372	119.1884	61.99551	77.17739	109.2157	69.22565	66.97027
	3	202.0913	167.0964	130.6803**	137.2514	159.0026	133.3683	132.4318
$\bar{y}_{N(\mathrm{MRSS})_4} = \frac{\bar{y}_{(\mathrm{MRSS})}+b_{\mathrm{rob}(zb)(\mathrm{MRSS})}(\bar{X}-\bar{x}_{(\mathrm{MRSS})})}{\bar{x}_{(\mathrm{MRSS})}+q_{3(\mathrm{MRSS})}}(\bar{X}+q_{3(\mathrm{MRSS})})$	4	90.44566	73.34087	57.46462**	59.77861	69.50191	58.29567	57.97628
	5	106.6025	88.66478	73.81949**	75.40392	84.74875	74.23809	74.03294
	7	67.12364	55.03597	48.90096	48.1625	52.63271	48.21179**	48.35186
	8	38.16683	29.86154	28.33111	26.57027	28.3738	27.18355***	27.47294
	3	140.113	145.7237	184.2986	167.9619	149.0152	175.6492	178.1544
$\bar{y}_{N(\mathrm{MRSS})_5} = \bar{y}_{(\mathrm{MRSS})} + b_{\mathrm{rob}(zb)(\mathrm{MRSS})}(\bar{X} - \bar{x}_{(\mathrm{MRSS})})$	4	57.0171	61.28342	84.75624	75.06935	63.42936	79.65024	81.13363
	5	71.93104	77.79206	105.7019	94.41887	80.476	99.77613	101.502
	7	47.11555	54.83533	84.32635	72.85911	57.92768	78.34642	80.09664
	8	28.6256	36.59961	64.60855	53.90899	39.64437	59.04758	60.67895

**Shows most efficient estimators according to sample size

***Shows most efficient estimators according to all cases

this direction we have considered robust techniques for estimation of population mean. First, we have examined newly proposed robust estimators in SRS and tried to define more general class of estimators in SRS which newly proposed estimators are member of our class. Then, we have extended our theoretical findings to different sampling designs such as RSS and MRSS. A general form of estimators of population mean and MSE formula are obtained. Theoretical MSEs of the robust family are also given for each designs. To see the performance of

Table 5 Percent relative efficiency (PRE) of proposed estimators under RSS over SRS

SRS $\bar{y}_{z1}$ Zaman and Bulut estimator [30]							RSS $\bar{y}_{N(\mathrm{RSS})_1}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	122.03	195.97	166.44	129.81	180.31	184.83	276.00	331.22	502.05*	436.42	350.25	467.66	477.68
100	121.71	193.44	165.01	129.35	178.40	182.75	326.63	383.37	542.34*	484.08	402.28	512.23	521.10
100	121.35	190.72	163.44	128.82	176.32	180.49	344.20	396.43	532.54*	484.33	413.40	507.86	515.18
100	120.86	187.25	161.41	128.12	173.65	177.60	352.27	394.37	494.50*	460.49	407.58	477.27	482.42
100	120.15	182.77	158.66	127.10	170.13	173.81	339.65	375.39	457.24*	429.90	386.43	443.45	447.58
SRS $\bar{y}_{z2}$ Zaman and Bulut estimator [30]							RSS $\bar{y}_{N(\mathrm{RSS})_2}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	122.69	199.58	168.78	130.74	183.23	187.95	250.35	300.97	456.34*	396.94	318.38	425.27	434.33
100	122.28	196.32	166.93	130.14	180.77	185.26	292.16	342.66	481.79*	431.27	359.41	455.76	463.44
100	121.84	193.01	165.02	129.50	178.24	182.52	308.92	355.31	474.09*	432.43	370.30	452.82	459.14
100	121.30	189.15	162.76	128.72	175.28	179.31	318.80	356.41	444.46*	414.81	368.14	429.48	433.96
100	120.51	184.21	159.73	127.60	171.38	175.12	308.84	340.89	413.28*	389.28	350.74	401.20	404.83
SRS $\bar{y}_{N(\mathrm{SRS})_1}$ proposed estimator							RSS $\bar{y}_{N(\mathrm{RSS})_3}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	125.51	212.27	177.76	134.63	194.04	199.31	183.84	220.80	323.46*	286.43	233.18	304.46	310.09
100	124.52	204.48	173.32	133.17	188.12	192.88	197.25	228.79	305.22*	279.41	238.84	292.21	296.12
100	123.70	198.52	169.81	131.98	183.52	187.90	206.97	235.21	299.19*	278.24	243.97	288.71	291.87
100	122.97	193.46	166.78	130.92	179.58	183.65	213.86	236.62	284.76*	269.45	243.48	277.15	279.45
100	121.83	186.57	162.46	129.31	174.08	177.75	210.60	230.36	271.21*	258.33	236.25	264.81	266.75
SRS $\bar{y}_{N(\mathrm{SRS})_2}$ proposed estimator							RSS $\bar{y}_{N(\mathrm{RSS})_3}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	120.00	152.32	145.06	125.85	149.22	150.27	119.13	132.30	147.94*	145.27	135.69	146.93	147.30
100	118.84	148.08	141.71	124.26	145.39	146.31	113.06	122.55	133.54*	131.62	124.93	132.79	133.06
100	117.92	144.67	139.06	123.01	142.34	143.15	116.36	124.87	134.83*	133.01	127.00	134.10	134.36
100	117.79	144.12	138.65	122.83	141.86	142.64	115.99	123.01	131.90*	130.02	124.79	131.09	131.36
100	116.63	141.01	135.89	121.30	138.87	139.61	117.31	123.99	133.15*	131.02	125.72	132.19	132.50

*Shows the most efficient estimator

our estimators, we have conducted a simulation study using a real data set. We have compared the existing estimators with our proposed estimators and concluded that our proposed estimators perform better than recently proposed Zaman and Bulut [30] and Ali et al. [1] estimators. As a future work, these robust estimators can be introduced under new RSS designs and new robust methods also can be proposed.

Table 6 PRE of proposed estimators under MRSS over SRS

SRS $\bar{y}_{z1}$ Zaman and Bulut estimator [30]							MRSS $\bar{y}_{N(\text{MRSS})_1}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	122.03	195.97	166.44	129.81	180.31	184.83	175.50	213.62	339.59*	289.66	227.03	313.17	320.81
100	121.71	193.44	165.01	129.35	178.40	182.75	271.97	332.30	531.71*	452.76	353.54	489.96	502.04
100	121.35	190.72	163.44	128.82	176.32	180.49	153.45	187.36	298.55*	254.71	199.27	275.40	282.10
100	120.86	187.25	161.41	128.12	173.65	177.60	137.60	169.83	278.17*	235.05	181.26	255.35	261.94
100	120.15	182.77	158.66	127.10	170.13	173.81	161.57	202.72	348.80*	289.28	217.58	317.07	326.19
SRS $\bar{y}_{z2}$ Zaman and Bulut estimator [30]							MRSS $\bar{y}_{N(\text{MRSS})_2}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	122.69	199.58	168.78	130.74	183.23	187.95	162.70	199.10	320.40*	272.19	211.94	294.88	302.25
100	122.28	196.32	166.93	130.14	180.77	185.26	251.54	309.09	501.05*	424.83	329.43	460.73	472.39
100	121.84	193.01	165.02	129.50	178.24	182.52	147.25	180.78	291.51*	247.78	192.60	268.42	275.10
100	121.30	189.15	162.76	128.72	175.28	179.31	135.10	167.85	279.01*	234.65	179.51	255.52	262.31
100	120.51	184.21	159.73	127.60	171.38	175.12	160.53	203.15	356.93*	293.92	218.63	323.30	332.96
SRS $\bar{y}_{N(\text{SRS})_1}$ proposed estimator							MRSS $\bar{y}_{N(\text{MRSS})_3}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	125.51	212.27	177.76	134.63	194.04	199.31	138.12	173.02	289.89*	243.78	185.44	265.59	272.65
100	124.52	204.48	173.32	133.17	188.12	192.88	211.56	267.04	453.86*	380.19	286.86	415.07	426.34
100	123.70	198.52	169.81	131.98	183.52	187.90	138.74	174.52	291.52*	246.10	187.20	267.74	274.68
100	122.97	193.46	166.78	130.92	179.58	183.65	137.73	176.64	308.71*	256.85	190.65	281.50	289.43
100	121.83	186.57	162.46	129.31	174.08	177.75	175.18	231.98	445.99*	358.26	253.16	399.41	412.86
SRS $\bar{y}_{N(\text{SRS})_2}$ proposed estimator							MRSS $\bar{y}_{N(\text{MRSS})_4}$ proposed estimator						
Linear	Huber	LMS	LTS	LAD	S	MM	Linear	Huber	LMS	LTS	LAD	S	MM
100	120.00	152.32	145.06	125.85	149.22	150.27	143.14	173.11	221.36*	210.76	181.93	216.89	218.43
100	118.84	148.08	141.71	124.26	145.39	146.31	240.35	296.41	378.30*	363.65	312.78	372.90	374.96
100	117.92	144.67	139.06	123.01	142.34	143.15	168.64	202.76	243.54*	238.42	212.13	242.16	242.84
100	117.79	144.12	138.65	122.83	141.86	142.64	194.02	236.64	266.32	270.41	247.44	270.13*	269.35
100	116.63	141.01	135.89	121.30	138.87	139.61	313.40	400.56	422.20	450.18	421.57	440.03*	435.39

*Shows the most efficient estimator

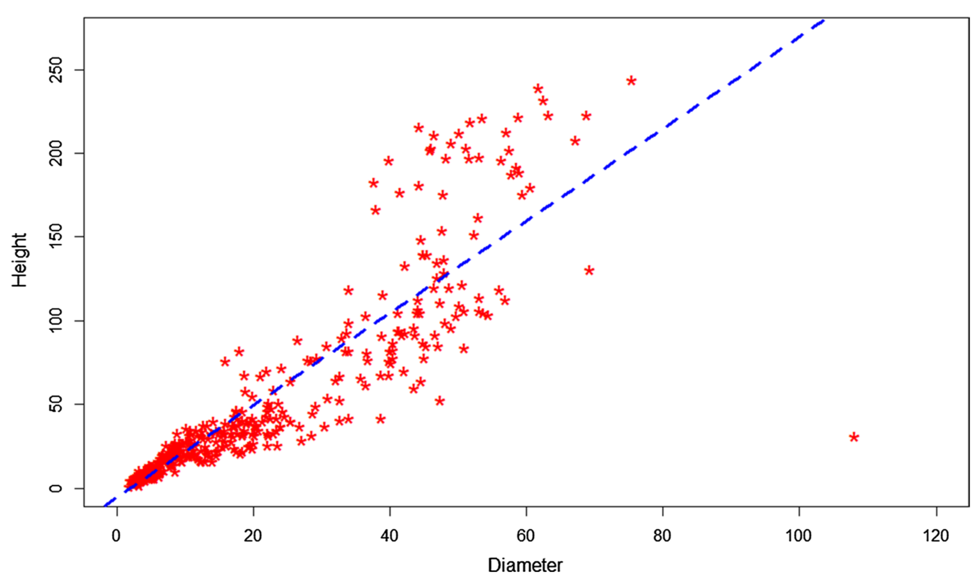

Fig. 1 Scatter plot of the tree data

References

1. Ali, N., Ahmad, I., Hanif, M., Shahzad, U.: Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information. Commun. Stat. Theory Methods (2019). https://doi.org/10.1080/03610926.2019.1645857
2. Al-Nasser, A.D.: L ranked set sampling: a generalization procedure for robust visual sampling. Commun. Stat. Simul. Comput. **36**(1), 33–43 (2007)
3. Al-Omari, A.I., Bouza, C.N.: Review of ranked set sampling: modifications and applications. Investigación Operacional **35**(3), 215–235 (2014)
4. Al-Omari, A.I., Bouza, C.N.: Ratio estimators of population mean with missing values using ranked set sampling. Environmetrics **26**(2), 67–76 (2015)
5. Al-Omari, A.I., Jaber, K., Ibrahim, A.: Modified ratio-type estimators of the mean using extreme ranked set sampling. J. Math. Stat. **4**(3), 150 (2008)
6. Al-Omari, A.I.: Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling. Stat. Probab. Lett. **82**, 1883–1890 (2012)
7. Al-Saleh, M.F., Al-Kadiri, M.A.: Double-ranked set sampling. Stat. Probab. Lett. **48**(2), 205–212 (2000)
8. Chen, Z.: Ranked set sampling: its essence and some new applications. Environ. Ecol. Stat. **14**(4), 355–363 (2007)
9. Collins, J.R.: Robust estimation of a location parameter in the presence of asymmetry. Ann. Stat. **4**, 68–85 (1976)
10. Daszykowski, M., Kaczmarek, K., Vander Heyden, Y., Walczak, B.: Robust statistics in data analysis–a review: basic concepts. Chemometr. Intell. Lab. Syst. **85**(2), 203–219 (2007)
11. Hanif, M., Shahzad, U.: Estimation of population variance using kernel matrix. J. Stat. Manag. Syst. **22**(3), 563–586 (2019)
12. Huber, P.J.: Robust regression: asymptotics, conjectures and Monte Carlo. Ann. Stat. (1973). https://doi.org/10.1214/aos/1176342503
13. Huber, P.J.: Robust Estimation of a Location Parameter. Breakthroughs in Statistics, pp. 492–518. Springer, New York (1992)
14. Jemain, A.A., Al-Omari, A.I., Ibrahim, K.: Balanced groups ranked set sampling for estimating the population median. J. Appl. Stat. Sci. **17**(1), 39–46 (2008a)
15. Jemain, A.A., Al-Omari, A.I., Ibrahim, K.: Modified ratio estimator for the population mean using double median ranked set sampling. Pak. J. Stat. **24**(3), 217–226 (2008b)
16. Koyuncu, N.: New difference-cum-ratio and exponential type estimators in median ranked set sampling. Hacet. J. Math. Stat. **45**(1), 207–225 (2016)
17. Koyuncu, N.: Regression estimators in ranked set, median ranked set and neoteric ranked set sampling. Pak. J. Stat. Oper. Res. **14**(1), 89–94 (2018)
18. Koyuncu, N.: A class of estimators in median ranked set sampling. İstatistikçiler Dergisi: İstatistik ve Aktüerya **12**(2), 58–71 (2019)
19. Kvam, P.H.: Ranked set sampling based on binary water quality data with covariates. J. Agric., Biol., Environ. Stat. **8**(3), 271 (2003)
20. McIntyre, G.A.: A method for unbiased selective sampling using ranked sets. Aust. J. Agric. Res. **3**, 385–390 (1952)
21. Muttlak, H.A.: Pair rank set sampling. Biom. J. **38**, 879–885 (1996)
22. Muttlak, H.A.: Median ranked set sampling. J. Appl. Stat. Sci. **6**, 245–255 (1997)
23. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2020)
24. Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. Wiley Series in Probability and Mathematical Statistics. Wiley, New York (1987)
25. Shahzad, U., Perri, P.F., Hanif, M.: A new class of ratio-type estimators for improving mean estimation of nonsensitive and sensitive variables by using supplementary information. Commun. Stat. Simul. Comput. **48**(9), 2566–2585 (2019)
26. Shahzad, U., Al-Noor, N.H., Hanif, M., Sajjad, I.: An exponential family of median based estimators for mean estimation with simple random sampling scheme. Commun. Stat. Theory Methods 1–10 (2020)
27. Shahzad, U., Al-Noor, N.H., Hanif, M., Sajjad, I., Muhammad Anas, M.: Imputation based mean estimators in case of missing data utilizing robust regression and variance-covariance matrices. Commun. Stat. Simul. Comput. 1–20 (2020)
28. Subzar, M., Bouza, C.N., Al-Omari, A.I.: Utilization of different robust regression techniques for estimation of finite population mean in srswor in case of presence of outliers through ratio method of estimation. Investigación Operacional **40**(5), 600–609 (2019)
29. Tukey, J.W.: Exploratory Data Analysis. Addison-Wesley, Boston (1977)
30. Zaman, T., Bulut, H.: Modified ratio estimators using robust regression methods. Commun. Stat. Theory Methods **48**(8), 2039–2048 (2019a)
31. Zaman, T., Bulut, H.: Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling. Commun. Stat. Theory Methods (2019b). https://doi.org/10.1080/03610926.2019.1588324
32. Zamanzade, E., Al-Omari, A.I.: New ranked set sampling for estimating the population mean and variance. Hacet. J. Math. Stat. **45**(6), 1891–1905 (2016)
33. Zamanzade, E., Mahdizadeh, M.: A more efficient proportion estimator in ranked set sampling. Stat. Probab. Lett. **129**, 28–33 (2017)
34. Zamanzade, E., Wang, X.: Estimation of population proportion for judgment post-stratification. Comput. Stat. Data Anal. **112**, 257–269 (2017)