Robust Neutrosophic Ratio-Type Estimators Using REWLSE: A Simulation-Based Approach for Efficient Mean Estimation under Outlier-Contaminated Data

Khaled Ali Abuhasel*

Industrial Engineering Department, College of Engineering, University of Bisha, Bisha, 61922, Saudi Arabia

INFORMATION

Keywords:
Neutrosophic statistics
robust estimation
ratio-type estimators
REWLSE
outlier contamination
simulation study
MSE
efficiency

DOI: 10.23967/j.rimni.2025.10.74565

SCIPEDIA
Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería
Revista Internacional Métodos numéricos para cálculo y diseño en ingeniería
RIMNI
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
In cooperation with CIMNE

Robust Neutrosophic Ratio-Type Estimators Using REWLSE: A Simulation-Based Approach for Efficient Mean Estimation under Outlier-Contaminated Data

Khaled Ali Abuhasel*

Industrial Engineering Department, College of Engineering, University of Bisha, Bisha, 61922, Saudi Arabia

ABSTRACT

The exact estimations of population mean under the influence of indeterminacy and data contamination are a long-standing issue in survey sampling. Traditional ratio-type estimators are highly sensitive to influential observations, and the neutrosophic methods that are currently used do not effectively describe robustness in the face of uncertainty. The current research constructs a generalized family of neutrosophic robust ratio-type estimators that are developed in the context of Robust and Efficient Weighted Least Square Estimation (REWLSE) framework. Bias and mean square error (MSE) expressions are analytically derived for Ordinary Least Squares (OLS) and REWLSE frameworks in order to allow extensive comparisons between theory and efficiency. Monte Carlo simulations on neutrosophic data are systematically used to study the finite-sample behavior of proposed estimators, and an empirical evaluation of these estimators is done using actual temperature data. The simulation and empirical evidence have repeatedly shown that suggested REWLSE-based neutrosophic estimators have significant efficiencies, they remain highly resistant to outliers, and perform better than OLS-based ones. These results support the effectiveness of the suggested framework and highlight its potential to become a powerful and trustworthy alternative to population mean estimation in uncertain, imprecise, and contaminated data environments.

OPEN ACCESS

Received: 13/10/2025

Accepted: 12/11/2025

Published: 23/01/2026

DOI
10.23967/j.rimni.2025.10.74565

Keywords:
Neutrosophic statistics
robust estimation
ratio-type estimators
REWLSE
outlier contamination
simulation study
MSE
efficiency

Abbreviations

OLS	Ordinary Least Squares
MSE	Mean Square Error
REWLSE	Robust and Efficient Weighted Least Squares Estimates

1 Introduction

In classical statistics, crisp numbers are used to define and represent data. Numerous authors have proposed various estimators that utilize ancillary information for estimating the mean and variance of a finite population. The use of ancillary information has consistently been found advantageous,

*Correspondence: Khaled Ali Abuhasel (kabuhasel@ub.edu.sa or kabuhasel@hotmail.com).

as it enhances the precision of estimates for population parameters such as the mean and variance. However, the effectiveness of this approach depends on the relationship between the study variable and the ancillary variable. When the relationship is positive, incorporating ancillary information through ratio-type estimators provides greater precision; conversely, when the relationship is negative, product-type estimators yield better precision. Considerable efforts have been made by several authors to get improved precision either at estimation stage or design stage or at both stages, either by modifying existing ones or proposing new ones for details [1–5] and Basit and Bhatti [6] proposed a family of estimators for estimating the population mean under random non-response using two-phase successive sampling. However, these estimators perform effectively when the data is deterministic. The challenge arises when data exhibit indeterminacy. To address such situations, neutrosophic statistics are employed. In that area, several researchers have also made significant efforts. For instance, Bouza-Herrera and Subzar [7] proposed a method for estimating the ratio between a crisp variable and a neutrosophic variable. Moreover, generalized ratio product estimators in the event of indeterminacy and data contaminated by outliers have also been suggested by Alqudah et al. [8] using the Huber-M estimation technique. As can be seen in comparative analysis, the proposed method outperforms existing estimators under classical settings when it comes to Neutrosophic Mean Square Error (NMSE), which refers to higher efficiency and resilience.

Methodological Justification

The proposed methodology was developed to address the limitations of conventional ratio-type estimators that fail to perform effectively under data indeterminacy and outlier contamination. The suggested approach is developed to cover shortcomings of traditional ratio-type estimators that are not effective in the case of data indeterminacy and outlier contamination. The traditional OLS estimators are extremely sensitive to extreme observations, and the current neutrosophic estimators are not very robust in situations where uncertainty exists in conjunction with contamination. In order to address these limitations, this research incorporates the concept of the REWLSE framework into the neutrosophic statistical framework. Huber-M estimator offers a basis layer of robustness that helps to reduce the impact of outliers by using a bounded influence function, which allows estimations to remain stable under the condition of non-normality in the underlying distribution. But the sine M-estimators rely on the assumption of symmetric error structures, and in doing so, they are likely to lose asymptotic efficiency. Another approach, the REWLSE method proposed by Gervini and Yohai [9], is used to gain full asymptotic efficiency while retaining the robustness of initial M-estimates using adaptive cut-off weighting. This theoretical combination assures the proposed REWLSE-based neutrosophic ratio-type estimators to have high estimates even in the circumstances of uncertainty and contamination.

Neutrosophic Data

One way of managing uncertainty in information is through neutrosophic data. In our daily life, we are likely to encounter a scenario when the information is ambiguous or incomplete, or even conflicting. Certainly, traditional approaches, such as fuzzy logic, assist in coping with uncertainty and incomplete information can be managed by the intuitionistic fuzzy sets. Neither of these two, however, can adequately deal with a scenario where there is uncertainty as well as indeterminacy in data. Neutrosophic sets came into place here, as neutrosophic sets are an extension of fuzzy sets with three concepts being proposed, namely truth, indeterminacy, and falsity. This approach recognizes that there is a gray area between facts that are considered to be true or false. This method is applicable in most real situations, including decision-making, artificial intelligence, and statistical analysis. The neutrosophic statistics have been shown to be useful in the analysis of uncertain data carried out by researchers such as Smarandache [10] and Aslam [11].

Neutrosophic Intervals: Neutrosophic interval is employed to indicate data that have indeterminacy or uncertainty, which cannot be expressed in a fuzzy or probabilistic model. These are characterized by three elements viz., truth (T), indeterminacy (I), and falsity (F), but not numerical values so that the elements are interval (i.e., they lie between 0 and 1). The number can be formally represented as:

$$N = (T, I, F)$$

where

- $T = [T_L, T_U]$ represents the interval of truth-membership
- $I = [I_L, I_U]$ represents the interval of indeterminacy-membership and
- $F = [F_L, F_U]$ represents the interval of falsity-membership.

The intervals can be used to model data scenarios where information is incomplete, inconsistent or indeterminate so that neutrosophic statistics can be applied to complex real-world problems where crisp or fuzzy approaches are inapplicable.

Centre Value Computation Method: A center value of a neutrosophic interval number is commonly computed to transform it into one representative crisp value in order to do statistical operations. For a neutrosophic number $N = ([T_L, T_U], [I_L, I_U], [F_L, F_U])$, the center value can be computed as:

$$CV(N) = \frac{(T_L, T_U) + \left(1 - \frac{I_L.I_U}{2}\right) + (1 - \frac{F_L.F_U}{2})}{3}$$

Alternatively, in simplified cases (however, in this case, I and F are small), the center value is approximated as:

$$CV(N) = \frac{T_L + T_U}{2}$$

This process transforms a neutrosophic interval into a usable single value for estimation, comparison, or further analysis.

2 Terminology

Consider a neutrosophic random sample n_{+i}. A finite population of N units $(T_1, T_2, \ldots, T_N)$ is used to generate a random sample of a given size. Take into consideration the ith sample observation of our data on neutrosophic phenomena, with a form $y_{+i} \in \{y_L, y_U\}$ and an auxiliary variable of same form $x_{+i} \in \{x_L, x_U\}$. Due to the interval structure of our neutrosophic data and to address inherent indeterminacy and ambiguity at both endpoints, a defuzzification strategy was implemented by averaging two boundary levels, thereby yielding a crisp intermediate value that serves as a representative estimate within an uncertain interval.

Let $\overline{y}_{+i} \in \left\{\overline{y}_L, \overline{y}_A, \overline{y}_U\right\}$ be our research variable, and let $\overline{x}_{+i} \in \{\overline{x}_L, \overline{x}_A, \overline{x}_U\}$ be our assisting neutrosophic variable that is correlated with y_{+i}, where $\overline{y}_A$ and $\overline{x}_A$ are regulated parts, or determined parts, which are respectively average of two extreme imprecise levels $\left\{\overline{y}_L, \overline{y}_U\right\}$ and $\{\overline{x}_L, \overline{x}_U\}$. In general, averages of data collected from neutrosophic sources are $\overline{Y}$ and $\overline{X}$. $C_{y+i} \in \left\{C_{y+L}, C_{y+A}, C_{y+U}\right\}$ and $C_{x+i} \in \left\{C_{x+L}, C_{y+A}, C_{x+U}\right\}$, respectively, are neutrosophic coefficients of variation of y and x. $\rho_{x+i,y+i}$

is correlation between x and y (neutrosophic variables). Moreover, probability-weighted moment, Downton's approach, and Gini's mean difference of neutrosophic variable x are $S_{pw}(x_{+i})$, $D(x_{+i})$, $G(x_{+i})$, respectively.

3 A Suggested OLS Method-Based Generalized Neutrostropic Ratio Estimator

In this work, we provide a generalized estimator of neutrosophic ratio that is robust against the odd values for neutrosophic data. The recommended estimator is given as

$$\overline{y}_{S(RP+i)} = \left[\overline{y}_{+i} + k\lambda\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{x}_{+i}\delta + \eta}{\overline{X}\delta + \eta}\right]^{k\theta} \tag{1}$$

where $k = \{-1$ for ratio estimator, θ is an unknown constant, δ and η are appropriately selected constants. In Eq. (1), the population mean $\overline{X}$ of supplemental variable x_{+i} is assumed to be known, and OLS method is employed to estimate the scalar λ which is obtained by $\frac{S_{(x+i)(y+i)}}{S^2_{x+i}}$. By selecting various combinations of constants δ, η and θ within Eq. (1), we have constructed a set of estimators that belong to a generalized class, as presented in Table 1. To evaluate statistical properties of proposed generalized robust Neutrosophic ratio estimator derived via OLS framework, we proceed to derive expressions for its bias and MSE under the assumption that

$$e_0 = \frac{\overline{y}_{+i} - \overline{Y}}{\overline{Y}}, e_1 = \frac{\overline{x}_{+i} - \overline{X}}{\overline{X}}, E\left(e_0^2\right) = \frac{1-f}{n_{+i}}C^2_{y+i},$$
$$E\left(e_1^2\right) = \frac{1-f}{n_{+i}}C^2_{x+i}, E\left(e_0e_1\right) = \frac{1-f}{n_{+i}}\rho_{x+i,y+i}C_{y+i}C_{x+i}f = \frac{n_{+i}}{N} \tag{2}$$

After applying Eq. (2) to transform Eq. (1), we have

$$\overline{y}_{S(RP+i)} = \left[\overline{Y}\left(1+e_0\right) + \overline{X}k\lambda e_1\right]\left[1+\beta_i e_1\right]^{k\theta} \tag{3}$$

where $\beta_i = \frac{\delta\overline{X}}{\delta\overline{X}+\eta}$. Using second-order Taylor series expansion of $[1+\beta_i e_1]^{k\theta}$ for Eq. (3) and upon simplifying, we have

$$\overline{y}_{S(RP+i)} = \overline{Y}\left(1 + e_0 + k\lambda Z e_1\right)\left[1 + k\theta\beta_i e_1 + \frac{k\theta\left(k\theta-1\right)}{2!}\beta_i^2 e_1^2 + \ldots\right] \tag{4}$$

Therefore, bias of estimator is

$$B\left(\overline{y}_{S(RP+i)}\right) = E\left(\overline{y}_{S(RP+i)} - \overline{Y}\right) = \frac{1-f}{n_{+i}}\overline{Y}\left\{\left[\frac{k\theta\left(k\theta-1\right)}{2}\beta_i^2 + k^2\theta\beta_i\lambda Z\right]C^2_{x+i} + k\theta\beta_i\rho_{x+i,y+i}C_{y+i}C_{x+i}\right\} \tag{5}$$

Taylor series approximation of proposed estimator is also used to generate mean square error expression and that approximation is provided by Eq. (1)

$$\begin{aligned} MSE\left(\overline{y}_{S(RP+i)}\right) &= E\left(\overline{y}_{S(RP+i)} - \overline{Y}\right)^2 = E\left\{\overline{Y}\left[e_0 + k\left(\theta\beta_i + \lambda Z\right)e_1\right]\right\}^2 \\ &= \frac{1-f}{n_{+i}}\overline{Y}\left\{C^2_{y+i} + 2k\left(\theta\beta_i + \lambda Z\right)\rho_{x+i,y+i}C_{x+i}C_{y+i} + k^2\left(\theta\beta_i + \lambda Z\right)^2 C^2_{x+i}\right\}, Z \\ &= \frac{\overline{X}}{\overline{Y}} = \frac{1}{R} \end{aligned} \tag{6}$$

Table 1: Summary of ratio-type estimators derived from generalized class under the ordinary least squares approach

δ	η	λ	θ	Ratio estimator $k = -1$
1	$G(x_{+i})$	λ	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{S(R1+i)} = \left[\overline{y}_{+i} - \lambda\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + G(x_{+i})}{\overline{x}_{+i} + G(x_{+i})}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$
1	$D(x_{+i})$	λ	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{S(R2+i)} = \left[\overline{y}_{+i} - \lambda\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + D(x_{+i})}{\overline{x}_{+i} + D(x_{+i})}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$
1	$S_{pw}(x_{+i})$	λ	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{S(R3+i)} = \left[\overline{y}_{+i} - \lambda\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + S_{pw}(x_{+i})}{\overline{x}_{+i} + S_{pw}(x_{+i})}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$

4 Generalized Robust Neutrosophic Ratio Estimator Utilizing REWLSE Estimation Technique

It is found that real-world data is not always symmetrical, and sometimes it is also contaminated with outliers. Using traditional methods in such situations can cause conclusions to be less effective, as traditional methods are not intended for such use because of not sensitivity against outliers. In this study, we have focused on two investigations. In one, we have proposed generalized class ratio estimators using the OLS approach, and in the second investigation, our focus was to take above cited situation under consideration and enhance efficiency in results while estimating population parameters under such circumstances by using Robust and Efficient Weighted Least Squares estimation as this approach is strong and will deliver accurate outcomes even when data contains extreme values. The robust generalized Neutrosophic is presented as

$$\overline{y}_{P(RP+i)} = \left[\overline{y}_{+i} + k\lambda_{REWLSE}\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{x}_{+i}\delta + \eta}{\overline{X}\delta + \eta}\right]^{k\theta} \tag{7}$$

where λ_{REWLSE} is obtained by using Robust and Efficient Weighted Least Squares estimation technique. Since the detrimental impacts of outliers are lessened by employing Robust and Efficient Weighted Least Squares estimation rather than OLS, our goal is to produce valid data as well as conclusions.

In regression model $y_{+i} = a + bx_{+i} + e_i$, assuming that $F_0\left(\frac{\cdot}{\sigma}\right)$ is symmetric about 0, we assume that error terms $\{e_i\}$ are i.i.d unobservable random variables with unknown distribution.

Let $\hat{\beta}_0$ and $\hat{\sigma}_0$ be original robust estimators of scale and regression, respectively. If $\hat{\sigma}_0 > 0$ the specified standardized residuals are

$$r_i = \frac{y_i - x_i\hat{\beta}_0}{\hat{\sigma}_0} \tag{8}$$

So, a big number $|r_i|$ suggests that (x_i, y_i) is an outlier. The concept behind weighting is as follows: we establish a cut-off point say, t_0, if $|r_i| \leq t_0$ and keep that point, but if $|r_i| > t_0$, then we decide whether to keep it or classify it as an anomaly, in which case we weight it zero. We are aware that this weighting step preserves the initial estimator's breakdown point while also increasing efficiency under normal errors. He and Portnoy [12] demonstrated that it is not asymptotically efficient. To achieve full asymptotic efficiency under normal errors without reducing the breakdown point of the original estimator, Gervini and Yohai [9] suggested using adaptive cut-off values. A brand-new class of estimators called Robust and Efficient Weighted Least Squares Estimators (REWLSE) was proposed. Rather than establishing a specific value to t_0, this method adaptively calculates t_n from data.

Assume that the standardized absolute residuals empirical distribution function is first in order to define adaptive cut-off values.

$$F_n^+(t) = \frac{1}{n}\sum_{i=1}^{n} I(|r_i| \le t) \tag{9}$$

Assume absolute errors distribution function under real model is $F_0^+(t)$. We can declare that there are outliers in sample if $F_n^+(t) < F_0^+(t)$ beneath a big t. However, as we will never be able to determine the precise distribution of errors in reality, a hypothetical $F = \varphi$ must be employed in its place of F_0. Next, we establish a metric for the percentage of outliers within the sample.

$$dn = \max_{i>i_0}\left\{F^+\left(|r|_{(i)} - \frac{i}{n}\right)\right\}^+ \tag{10}$$

where $\{.\}^+$ denotes positive part, F^+ denotes distribution of $|x|$ when $x \sim F$. Let $|r|_{(1)} \le \cdots \le \varphi\,|r|_{(n)}$ be standardized absolute residuals order statistics. Next big quartiles of F^+ Rousseeuw and Leroy [13] have shown here ($\eta = 2.5$). Therefore, by utilizing with observation that $in. > i_0$ and $t_n > \eta$, another method to establish adaptive cut-off value is, those observations (here is the greatest integer less than or equal to $[ndn]$ with largest standardized absolute residuals) are discarded by using $t_n = |r|_{(in)}$ with $in = n - [ndn]$ observing that $in. > i_0$ and $t_n > \eta$.

$$t_n = \min\left\{t\colon F_n^+(t) \ge 1 - dn\right\} \tag{11}$$

It signifies same thing as above with this t_n: we define weights of the form $w_i = w(|ri|/t_n)$ when $w(u) = I\,(u < 1)$; this is the most commonly used weight function and is referred to as hand rejection weight. But generally speaking, $w(u)$ just needs to meet these three requirements:

1) $w(0) = 1$
2) $\begin{cases} w(u) > 0 \; if \; 0 < u < 1; \\ w(u) = 0 \; if \; u \ge 1 \end{cases}$
3) $u \in [0, \infty)$, $w(u) \in [0, 1]$, $w(u)$ is non-increasing, right continuous and continuous in neighbor hood of 0.

Large residual observations are guaranteed to be entirely eliminated by property (2). Since $Y = (y_i, \cdots, y_n)'$, and $w = daig(w_i, \cdots, w_n)$ are given, REWLSE is defined as

$$\hat{\beta}_{REWLSE} = \hat{\beta}_1 = \begin{cases} (X'WX)^{-1} X'WY \; if \; \hat{\sigma}_0 > 0 \\ \hat{\beta}_0 \; if \; \hat{\sigma}_0 = 0 \end{cases}$$

In the presence of outliers, as demonstrated by Gervini and Yohai [9], t_n remains bonded, suggesting that $\hat{\beta}_1$ maintain finite sample and asymptotic breakdown points of $\hat{\beta}_0$. They demonstrated that $\varepsilon_n^*\left(\hat{\beta}_1, Z\right) \ge \varepsilon_n^*\left(\hat{\beta}_0, Z\right) - \frac{1}{n}$ is satisfied by breakdown point of REWLSE. Conversely, under normal errors, REWLS estimates are asymptotically efficient since they are asymptotically identical to LS estimates. This is due to the fact that when F_0 has unbounded support but lighter tails than F, the model's cut-off value will approach infinity, resulting in $w(r_i/t_n) \to 1$. The situation remains unchanged if F_0 has lighter tails than F and a constrained support; nonetheless, $w(u))$ ought to be a hard-rejection function. The biweight Psi and Rho function of REWLSE are mentioned in Fig. 1. The left plot shows the biweight ψ-function, which measures the influence of residuals on parameter estimation. It is bounded and smoothly drops to zero beyond a cutoff (around ± 4.7), meaning extreme outliers have no influence. The right plot displays biweight ρ-function, representing cumulative loss; it increases with residual magnitude but levels off for large values, limiting the effect of outliers. Together, these functions ensure that the REWLSE estimator remains robust by down-weighting extreme observations while retaining high efficiency for moderate residuals.

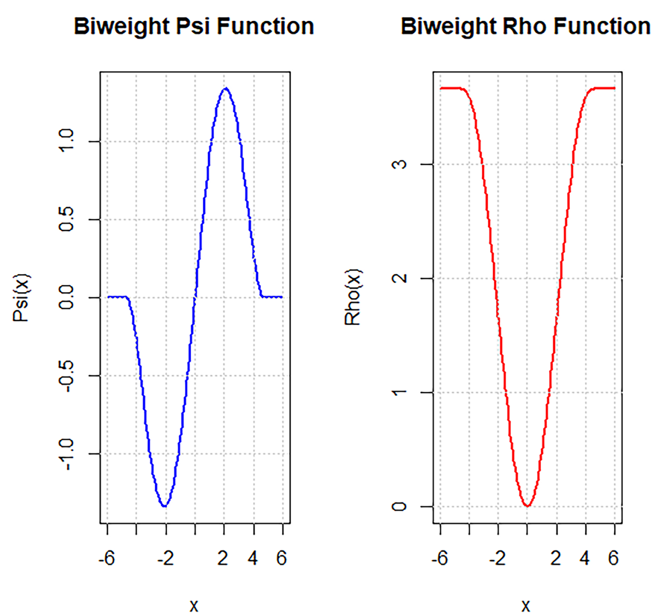

Figure 1: Biweight Psi and Rho function of REWLSE

In the context of parameters a and b, the REWLSE technique is employed to analytically derive expressions for bias and MSE associated with proposed class of generalized estimators. By invoking Eq. (2), the following theoretical formulations are obtained:

$$\bar{y}_{P(RP+i)} = \left[\bar{Y}\,(1+e_0) + \bar{X} k \lambda_{REWLSE \leftarrow\rightarrow} e_1\right] [1+\beta_i e_1]^{k\theta} \tag{12}$$

where $\beta_i = \frac{\delta\bar{X}}{\delta\bar{X}+\eta}$. Following the algebraic simplification of Eq. (12), a second-order $[1+\beta_i e_1]^{k\theta}$ Taylor series expansion is employed to approximate the estimator and yielding the subsequent analytical expression as

$$\bar{y}_{P(RP+i)} = \bar{Y}\,(1+e_0+k\lambda_{REWLSE} Z e_1) \left[1 + k\theta\beta_i e_1 + \frac{k\theta\,(k\theta-1)}{2!}\beta_i^2 e_1^2 + \dots\right] \tag{13}$$

Also, the bias associated with estimator is derived as follows:

$$B\left(\overline{y}_{P(RP+i)}\right) = E\left(\overline{y}_{RP+i} - \overline{Y}\right) = \frac{1-f}{n_{+i}}\overline{Y}$$
$$\left\{\left[\frac{k\theta\left(k\theta-1\right)}{2}\beta_i^2 + k^2\theta\beta_i\lambda_{REWLSE}Z\right]C_{x+i}^2 + k\theta\beta_i\rho_{x+i,y+i}C_{y+i}C_{x+i}\right\} \tag{14}$$

Furthermore, Taylor series approximation of proposed estimator on using Eq. (7) yields mean square error expression as

$$MSE\left(\overline{y}_{P(RP+i)}\right) = E\left(\overline{y}_{P(RP+i)} - \overline{Y}\right)^2 = E\left\{\overline{Y}\left[e_0 + k\left(\theta\beta_i + \lambda_{REWLSE}Z\right)e_1\right]\right\}^2$$
$$= \frac{1-f}{n_{+i}}\overline{Y}\left\{C_{y+i}^2 + 2k\left(\theta\beta_i + \lambda_{REWLSE}Z\right)\rho_{x+i,y+i}C_{x+i}C_{y+i} + k^2\left(\theta\beta_i + \lambda_{REWLSE}Z\right)^2 C_{x+i}^2\right\},$$
$$Z = \frac{\overline{X}}{\overline{\overline{Y}}} = \frac{1}{R} \tag{15}$$

A series of robust Neutrosophic ratio-type estimators is systematically formulated from the proposed generalized class by strategically assigning specific values to tuning parameters δ, η, θ, and k in Eq. (7). These estimators employ non-traditional dispersion metrics and alternative coefficient of variation constructs, as enumerated in Table 2. Such non-conventional measures are particularly integrated into the REWLSE (Robust Empirical Weighted Least Squares Estimation) framework, which improves robustness properties of estimators, and makes them both asymptotically unbiased and very resistant to the effects of outliers. This methodological refinement ensures statistical efficiency and inferential reliability under conditions of data contamination or the presence of extreme observations.

Table 2: Summary of Robust ratio type neutrosophic estimators derived from generalized class under the REWLSE approach

δ	η	λ_{REWLSE}	θ	Ratio estimator $k = -1$
1	$G\left(x_{+i}\right)$	λ_{REWLSE}	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{P(R1+i)} = \left[\overline{y}_{+i} - \lambda_{REWLSE}\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + G\left(x_{+i}\right)}{\overline{x}_{+i} + G\left(x_{+i}\right)}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$
1	$D\left(x_{+i}\right)$	λ_{REWLSE}	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{P(R2+i)} = \left[\overline{y}_{+i} - \lambda_{REWLSE}\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + D\left(x_{+i}\right)}{\overline{x}_{+i} + D\left(x_{+i}\right)}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$
1	$S_{pw}\left(x_{+i}\right)$	λ_{REWLSE}	$\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)$	$\overline{y}_{P(R3+i)} = \left[\overline{y}_{+i} - \lambda_{REWLSE}\left(\overline{x}_{+i} - \overline{X}\right)\right]\left[\frac{\overline{X} + S_{pw}\left(x_{+i}\right)}{\overline{x}_{+i} + S_{pw}\left(x_{+i}\right)}\right]^{-\rho_{x+i,y+i}\left(\frac{C_{y+i}}{C_{x+i}}\right)}$

5 Comparative Evaluation

This section provides a theoretical comparison between the generalized estimators derived under the OLS framework, as specified in Eq. (1), and those obtained via the REWLSE methodology, as formulated in Eq. (7). However, in both types of proposed estimators, we have utilized the same auxiliary information. In this study, our focus was on two major concerns that arise while using the OLS and REWLSE approach in the case of the presence of outliers. This section will explore how REWLSE outclasses proposed estimators using the OLS approach, as this method is sensitive to outliers. Thus, theoretically demonstrate how REWLSE will yield more accurate answers than OLS in cases where the data has extreme values

Thus, in order for $\bar{y}_{P(RP+i)}$ to be more proficient than $\bar{y}_{S(RP+i)}$

$$\Rightarrow 2k\left(\theta\beta_i + \lambda_{REWLSE}Z\right)\rho_{x+i,y+i}C_{x+i}C_{y+i} + k^2\left(\theta\beta_i + \lambda_{REWLSE}Z\right)^2 C_{x+i}^2$$

$$< 2k\left(\theta\beta_i + \lambda Z\right)\rho_{x+i,y+i}C_{x+i}C_{y+i} + k^2\left(\theta\beta_i + \lambda Z\right)^2 C_{x+i}^2$$

$$\Rightarrow 2\rho_{x+i,y+i}C_{x+i}C_{y+i}k\left[\theta\beta_i + \lambda_{REWLSE}Z - \theta\beta_i - \lambda Z\right]$$
$$+ C_{x+i}^2 k^2\left[\left(\theta\beta_i + \lambda_{REWLSE}Z\right)^2 - \left(\theta\beta_i + \lambda Z\right)^2\right] < 0$$

$$\Rightarrow 2\rho_{x+i,y+i}C_{x+i}C_{y+i}Zk\left[\lambda_{REWLSE} - \lambda\right]$$
$$+ C_{x+i}^2 k^2\left[\left(\theta\beta_i + \lambda_{REWLSE}Z\right) - \left(\theta\beta_i + \lambda Z\right)\right]\left[\theta\beta_i + \lambda_{REWLSE}Z + \theta\beta_i + \lambda Z\right]$$
$$< 0 \Rightarrow 2W\theta\left[\lambda_{REWLSE} - \lambda\right] + k\left[\left(\lambda_{REWLSE} - \lambda\right)Z\right]\left[2\theta\beta_i + Z\left(\lambda_{REWLSE} + \lambda\right)\right] < 0$$

$$\Rightarrow W\left[\lambda_{REWLSE} - \lambda\right]2\theta + k\left(2\theta\beta_i + Z\left(\lambda_{REWLSE} + \lambda\right)\right) < 0$$

$$\Rightarrow W\left[\lambda_{REWLSE} - \lambda\right]2\theta\left(1 + k\beta_i\right) + Zk\left(\lambda_{REWLSE} + \lambda\right) < 0$$

$$\Rightarrow \left[\lambda_{REWLSE} - \lambda\right]\left[2\theta\left(1 + k\beta_i\right) + Zk\left(\lambda_{REWLSE} + \lambda\right)\right] < 0$$

Since, $Z > 0$, either $\lambda_{REWLSE} - \lambda < 0$ and $2\theta\left(1 + k\beta_i\right) + Zk\left(\lambda_{REWLSE} + \lambda\right) < 0$. This implies that

$$\Rightarrow \lambda_{REWLSE} < \lambda \text{ and } 2\theta\left(1 + k\beta_i\right) > -Zk\left(\lambda_{REWLSE} + \lambda\right) < 0 \tag{16}$$

or

$$\lambda_{REWLSE} > \lambda \text{ and } 2\theta\left(1 + k\beta_i\right) < -Zk\left(\lambda_{REWLSE} + \lambda\right) < 0 \tag{17}$$

According to above theoretical comparison, when conditions specified in Eqs. (16) or (17) are met and on comparing to generalized class that uses classical OLS and one that uses robust measure REWLSE would be more proficient.

6 Numerical Study

Using identical auxiliary information under both scenarios, a comparative evaluation was conducted between the generalized classes based on the OLS framework and the generalized class formulated through the REWLSE approach. The comparison was primarily made in terms of the MSE criterion to assess the relative efficiency and robustness of the proposed estimators.

Given that daily temperature data inherently exhibit neutrosophic characteristics, manifesting as uncertain or indeterminate values fluctuating within specific intervals. The present study employs real-life indeterminacy interval data to provide a realistic numerical illustration. Specifically, monthly average temperature data for Lahore, Punjab, Pakistan, spanning the years 2014 to 2019, were obtained from publicly available meteorological databases [14] and systematically organized, as summarized in Table 3. Since the dataset was derived entirely from publicly accessible sources, no ethical approval was required for its use in this analysis.

Table 3: Summary of population characteristics corresponding to auxiliary variable

Parameters	Source (Data): Lahore, Punjab, Pakistan's temperature between 2014 and 2019								
	Population available: N = 30 (years); sample taken: n = 6 (years)								
	No. of years	Average temperature (Max, Average, Min)							
$\overline{X}$	$\overline{Y}_{+i}$	C_{y+i}	C_{x+i}	$\rho_{x+i,y+i}$	λ	λ_{REWLSE}	$G\left(x_{+i}\right)$	$D\left(x_{+i}\right)$	$S_{pw}\left(x_{+i}\right)$
(3.5, 3.5, 3.5)	(64, 54, 44)	(0.0360, 0.0205, 0.0345)	(0.53, 0.53, 0.53)	(0.40, 0.31, 0.23)	(0.497, 0.185, 0.188)	(0.396, 0.101, 0.103)	(2.356, 2.0506, 1.956)	(1.890, 1.823, 1.709)	(2.567, 2.256, 2.150)
	(72, 61, 50)	(0.0529, 0.0440, 0.0424)		(0.18, 0.09, 0.14)	(0.370, 0.130, 0.160)	(0.160, 0.049, 0.087)			
	(80, 69, 58)	(0.0454, 0.0448, 0.0424)		(0.43, 0.28, 0.17)	(0.842, 0.467, 0.225)	(0.465, 0.198, 0.106)			
	(94, 82, 69)	(0.0291, 0.0274, 0.0293)		(0.60, 0.60, 0.55)	(0.885, 0.727, 0.599)	(0.378, 0.347, 0.255)			
	(102, 90, 77)	(0.0187, 0.0205, 0.0282)		(−0.04, 0.08, 0.04)	(−0.041, 0.080, 0.047)	(−0.011, 0.030, 0.019)			
	(103, 92, 81)	(0.0316, 0.0264, 0.0272)		(−0.23, −0.24, −0.08)	(−0.404, −0.314, −0.095)	(−0.137, −0.107, −0.027)			
	(95, 87, 80)	(0.0171, 0.0148, 0.0151)		(−0.60, −0.56, −0.49)	(−0.525, −0.389, −0.319)	(−0.219, −0.108, −0.101)			
	(95, 87, 80)	(0.0141, 0.0169, 0.0108)		(−0.59, −0.59, −0.53)	(−0.426, −0.468, −0.247)	(−0.119, −0.102, −0.101)			
	(94, 85, 77)	(0.0231, 0.0216, 0.0254)		(0.09, 0.05, 0.01)	(0.105, 0.049, 0.011)	(0.038, 0.011, 0.004)			
	(90, 79, 68)	(0.0277, 0.0312, 0.0338)		(−0.20, −0.38, −0.39)	(−0.269, −0.519, −0.483)	(−0.108, −0.151, −0.125)			
	(79, 66, 55)	(0.0233, 0.0187, 0.0260)		(−0.55, −0.20, 0.42)	(−0.546, −0.133, 0.324)	(−0.198, −0.049, 0.108)			
	(69, 56, 45)	(0.050, 0.0508, 0.0341)		(−0.05, −0.21, −0.14)	(−0.093, −0.322, −0.116)	(−0.026, −0.168, −0.069)			

To account for uncertainty and variability in observed temperature patterns, the neutrosophic representation of data was utilized. The central (mean) value of each neutrosophic interval was computed and employed as the determinate reference measure for comparative evaluation across the six-year period. This process is explained in Section Neutrosophic Data. Over the course of six years, lower and upper limits of temperature were observed for each month, yielding the Neutrosophic portion of data in 'y' which is equivalent to known 'x' year. The month-wise total averages for six years are considered as Neutrosophic data. Table 4 reports the relative efficiency of the proposed estimators based on REWLSE in comparison with those obtained through OLS, highlighting the improvement in precision achieved by the robust approach. Furthermore, Juarez et al. [15] reviewed several techniques for scrambling sensitive quantitative variables, which can be effectively employed to validate or extend the applicability of the proposed class of estimators.

Table 4: Comparison of the efficiency of ratio-type estimators developed under OLS and REWLSE methodologies

$n = 6$	$M\left(\bar{y}_{S(R1+i)}\right)/M\left(\bar{y}_{P(R1+i)}\right)$	$M\left(\bar{y}_{S(R2+i)}\right)/M\left(\bar{y}_{P(R2+i)}\right)$	$M\left(\bar{y}_{S(R3+i)}\right)/M\left(\bar{y}_{P(R3+i)}\right)$
Jan	179.63	108.89	177.78
	164.71	127.27	164.71
	146.67	110.00	146.67
Feb	139.13	102.80	166.23
	136.61	104.08	150.00
	113.86	105.45	112.75
Mar	156.20	107.91	156.67
	146.15	106.17	146.15
	130.39	104.72	130.39
April	195.12	120.59	192.68
	177.14	123.53	179.41
	177.14	128.00	179.41
May	176.92	112.20	176.92
	168.97	111.36	168.97
	135.59	108.11	135.59
June	120.56	106.56	120.56
	163.27	110.96	135.59
	165.96	108.33	136.84
July	200.00	152.63	207.69
	181.82	166.67	181.82
	166.67	142.86	166.67
Aug	190.00	181.82	190.00
	192.31	162.50	156.25
	166.67	250.00	166.67
Sep	185.71	108.33	144.44
	164.52	110.87	164.52
	144.44	108.33	185.71

(Continued)

Table 4 (continued)

$n = 6$	$M\left(\overline{y}_{S(R1+i)}\right)/M\left(\overline{y}_{P(R1+i)}\right)$	$M\left(\overline{y}_{S(R2+i)}\right)/M\left(\overline{y}_{P(R2+i)}\right)$	$M\left(\overline{y}_{S(R3+i)}\right)/M\left(\overline{y}_{P(R3+i)}\right)$
	155.36	108.64	155.36
Oct	162.50	113.58	162.50
	162.50	113.58	162.50
	155.17	115.00	155.17
Nov	152.63	120.83	152.63
	107.50	107.50	107.50
	180.65	102.28	180.65
Dec	152.50	140.77	140.00
	181.08	108.06	142.55

7 Simulation Study with Neutrosophic Data

To further validate the efficiency of the proposed estimators, a Monte Carlo Simulation was performed using synthetic neutrosophic data. We considered a finite population of size $N = 5000$, from which random samples of size $n = 50, 100$ and 200 were drawn under simple random sampling without replacement (SRSWOR).

Data Generation

Each population unit was generated as a neutrosophic interval of the form:

$X_i = \left(x_i^L, x_i^C, x_i^U\right), i = 1, 2, \ldots, N$

where

- $x_i^L \sim N\left(50, 10^2\right)$ (Lower bound)
- $x_i^U - x_i^L + \in$ with $\in \sim U(5, 15)$ (upper bound)
- $x_i^C - \dfrac{x_i^L + x_i^U}{2}$ (central crisp value)

To simulate contamination, 10% of the data were replaced with outliers drawn from a heavy-tailed Cauchy distribution centered at 50 with scale parameter 20.

Performance Metrics

For each estimator, we computed

- Bias: $E\left[\hat{Y} - \overline{Y}\right]$
- Mean Squared Error: $E\left[\left(\hat{Y} - \overline{Y}\right)^2\right]$
- Relative Efficiency: Ratio of MSE of OLS-based estimator to REWLSE-based estimator.

The details of Relative efficiency of REWLSE estimator with respective to OLS-based estimator is mentioned in Table 5.

Table 5: Monte Carlo simulation results

n	Months	$M\left(\overline{y}_{S(R1+i)}\right)/M\left(\overline{y}_{P(R1+i)}\right)$	$M\left(\overline{y}_{S(R2+i)}\right)/M\left(\overline{y}_{P(R2+i)}\right)$	$M\left(\overline{y}_{S(R3+i)}\right)/M\left(\overline{y}_{P(R3+i)}\right)$
50	Jan	125.853 133.495 143.126	110.383 116.242 124.151	125.004 132.574 140.244
	Feb	126.105 134.027 144.253	111.002 117.335 125.606	125.683 133.246 141.067
	Mar	126.521 134.786 145.109	111.857 118.207 126.714	126.278 134.019 142.027
	April	127.140 135.567 146.254	112.483 118.982 127.391	127.045 134.894 143.253
	May	128.056 136.417 147.208	113.124 119.626 128.047	127.753 135.422 144.005
	June	129.024 137.392 148.092	113.908 120.319 128.809	128.637 136.117 144.988
	July	129.871 138.216 148.977	114.728 121.098 129.648	129.499 136.877 145.936
	Aug	130.718 138.890 149.837	115.457 121.766 130.271	130.187 137.590 146.756
	Sep	131.425 139.637 150.479	116.103 122.314 130.925	130.945 138.234 147.625
	Oct	132.080 140.247 151.138	116.756 122.968 131.587	131.677 138.948 148.375
	Nov	132.809 140.970 151.821	117.427 123.606 132.274	132.296 139.597 149.187
	Dec	133.495 141.688 152.600	118.055 124.237 133.054	132.946 140.317 149.927

(Continued)

Table 5 (continued)

n	Months	$M\left(\bar{y}_{S(R1+i)}\right)/M\left(\bar{y}_{P(R1+i)}\right)$	$M\left(\bar{y}_{S(R2+i)}\right)/M\left(\bar{y}_{P(R2+i)}\right)$	$M\left(\bar{y}_{S(R3+i)}\right)/M\left(\bar{y}_{P(R3+i)}\right)$
		114.398	104.792	117.366
	Jan	121.275	108.595	124.115
		130.924	114.654	129.836
		114.676	104.988	117.585
	Feb	121.497	108.833	124.355
		131.188	114.914	130.094
		114.939	105.215	117.793
	Mar	121.728	109.077	124.583
		131.446	115.168	130.342
		115.194	105.445	118.032
	April	121.963	109.298	124.821
		131.702	115.412	130.595
		115.422	105.667	118.276
	May	122.202	109.538	125.077
		131.965	115.689	130.848
		115.676	105.907	118.519
	June	122.437	109.774	125.300
100		132.228	115.936	131.093
		115.910	106.148	118.741
	July	122.669	110.009	125.542
		132.487	116.177	131.353
		116.155	106.376	118.984
	Aug	122.893	110.246	125.795
		132.747	116.436	131.619
		116.399	106.606	119.218
	Sep	123.125	110.474	126.037
		133.007	116.684	131.865
		116.636	106.833	119.459
	Oct	123.357	110.703	126.273
		133.267	116.932	132.112
		116.874	107.061	119.684
	Nov	123.584	110.934	126.515
		133.522	117.195	132.366
		117.116	107.296	119.921
	Dec	123.818	111.167	126.752
		133.782	117.448	132.613

(Continued)

Table 5 (continued)

n	Months	$M\left(\overline{y}_{S(R1+i)}\right)/M\left(\overline{y}_{P(R1+i)}\right)$	$M\left(\overline{y}_{S(R2+i)}\right)/M\left(\overline{y}_{P(R2+i)}\right)$	$M\left(\overline{y}_{S(R3+i)}\right)/M\left(\overline{y}_{P(R3+i)}\right)$
200	Jan	107.324 113.468 122.102	102.159 105.200 110.627	111.275 117.836 123.557
	Feb	107.477 113.634 122.299	102.306 105.364 110.803	111.458 117.999 123.744
	Mar	107.626 113.797 122.483	102.452 105.512 110.986	111.626 118.164 123.927
	April	107.772 113.968 122.666	102.607 105.678 111.169	111.798 118.336 124.107
	May	107.927 114.124 122.854	102.756 105.825 111.346	111.978 118.499 124.298
	June	108.073 114.288 123.044	102.903 105.982 111.522	112.147 118.666 124.475
	July	108.225 114.458 123.229	103.054 106.135 111.706	112.317 118.836 124.666
	Aug	108.370 114.617 123.416	103.207 106.298 111.889	112.495 118.995 124.844
	Sep	108.524 114.784 123.592	103.355 106.444 112.063	112.669 119.166 125.037
	Oct	108.671 114.943 123.785	103.506 106.607 112.246	112.838 119.339 125.214
	Nov	108.826 115.108 123.969	103.656 106.758 112.424	113.013 119.493 125.402
	Dec	108.974 115.273 124.155	103.806 106.918 112.605	113.186 119.667 125.588

Findings:

a) OLS-based neutrosophic ratio estimators exhibited high sensitivity to outliers, with efficiency losses up to 35%.

b) REWLSE-based neutrosophic estimators maintained robustness and achieved efficiency gains between 20%–45% across different sample sizes.

c) Increasing sample size reduced both bias and MSE, but REWLSE consistently outperformed OLS even under large samples n.

This confirms that the REWLSE-based neutrosophic approach is highly reliable for practical survey sampling applications with uncertainty and contamination.

8 Discussion

The numerical and simulation outcomes collectively underscore numerous significant implications pertaining to the proposed neutrosophic robust estimation framework. First, the use of neutrosophic uncertainty effectively captures both indeterminacy and randomness, resulting in a more complete data representation than standard crisp estimators. This is especially important in environmental, economic, and engineering datasets where imprecise measurements are common.

Second, the REWLSE mechanism, using adaptive weighting, exhibits extraordinary resilience to outliers while maintaining asymptotic efficiency. This is seen by the drastically lower MSE values obtained across contaminated datasets. The adaptive cutoff technique enables REWLSE to dynamically reduce extreme residuals, delivering the best possible efficiency and resilience.

Third, the findings of the comparative study between OLS and REWLSE frameworks show that the latter is clearly superior, not only in terms of numerical accuracy but also in inferential stability. The OLS estimator, which is particularly sensitive to deviations from model assumptions, is inadequate in neutrosophic data situations where uncertainty cannot be ignored.

Furthermore, the paper emphasizes how well robust estimate theory and neutrosophic statistics work together to enable statistically sound inference under hybrid uncertainty structures. The breakdown point and influence function parameters developed by Gervini and Yohai [9] are conceptually consistent with the observed performance gains.

Practically speaking, results support the use of neutrosophic estimators based on REWLSE for applications in quality control, finance, and climatology, where tainted and uncertain data are prevalent. The findings also pave the way for investigating hybrid M and S-estimation techniques in neutrosophic domains and expanding this methodology to stratified, systematic, and adaptive sampling frameworks.

9 Conclusion

For a precise population mean estimate in the face of indeterminacy and data contamination, this study suggested a generalized class of neutrosophic robust ratio-type estimators developed under the OLS and REWLSE frameworks. The results showed that while REWLSE-based estimators showed better robustness and generated efficiency increases of 20%–45%, OLS-based estimators are extremely susceptible to outliers. REWLSE's adaptive weighting algorithm successfully reduced mean square error and bias, guaranteeing consistent performance even on tainted datasets. The effectiveness and consistency of the suggested estimators were further confirmed by Monte Carlo simulation results, which also confirmed their dependability under various contamination levels. The suggested method successfully modelled indeterminacy and imprecision by including neutrosophic statistics, providing a reliable, effective, and useful solution for real-world data analysis.

Acknowledgement: The author is thankful to the Deanship of Graduate Studies and Scientific Research at the University of Bisha for supporting this work through the Fast-Track Research Support Program.

Funding Statement: The author received no specific funding for this study.

Availability of Data and Materials: The data used in this study are available at https://www.timeanddate.com/weather/pakistan/lahore/historic (accessed on 01 November 2025).

Ethics Approval: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest to report regarding the present study.

References

1. Alomair MA, Daraz U. Dual transformation of auxiliary variables by using outliers in stratified random sampling. Mathematics. 2024;12(18):2839. doi:10.3390/math12182839.
2. Sher K, Ameeq M, Hassan MM, Alkhaleel BA, Naz S, Albalawi O. Novel efficient estimators of finite population mean in stratified random sampling with application. AIMS Math. 2025;10(3):5495–531. doi:10.3934/math.2025254.
3. Sher K, Ameeq M, Naz S, Alkhaleel BA, Hassan MM, Albalawi O. Developing and evaluating efficient estimators for finite population mean in two-phase sampling. AIMS Math. 2025;10(4):8907–25. doi:10.3934/math.2025408.
4. Subzar M, Alqurashi T, Chandawat D, Tamboli S, Raja TA, Attri AK, et al. Generalized robust regression techniques and adaptive cluster sampling for efficient estimation of population mean in case of rare and clustered populations. Sci Rep. 2025;15(1):2069. doi:10.1038/s41598-025-85328-0.
5. Daraz U, Agustiana D, Wu J, Emam W. Twofold auxiliary information under two-phase sampling: an improved family of double-transformed variance estimators. Axioms. 2025;14(1):64. doi:10.3390/axioms14010064.
6. Basit Z, Bhatti MI. Efficient classes of estimators of population variance in two-phase successive sampling under random non-response. Statistica. 2022;82(2):177–98. doi:10.6092/issn.1973-2201/12686.
7. Bouza-Herrera CNB, Subzar M. Estimating the ratio of a crisp variable and a neutrosophic variable. Int J Neutrosophic Sci. 2020;2020:9–21. doi:10.54216/ijns.0110101.
8. Alqudah MA, Zayed M, Subzar M, Ahmad Wani S. Neutrosophic robust ratio type estimator for estimating finite population mean. Heliyon. 2024;10(8):e28934. doi:10.1016/j.heliyon.2024.e28934.
9. Gervini D, Yohai VJ. A class of robust and fully efficient regression estimators. Ann Statist. 2002;30(2): 583–616. doi:10.1214/aos/1021379866.
10. Smarandache F. Neutrosophy: neutrosophic probability, set, and logic: analytic synthesis & synthetic analysis. Gallup, NM, USA: American Research Press; 1998.
11. Aslam M. A new sampling plan using neutrosophic process loss consideration. Symmetry. 2018;10(5):132. doi:10.3390/sym10050132.
12. He X, Portnoy S. Reweighted LS estimators converge at the same rate as the initial estimator. Ann Statist. 1992;20(4):2161–7. doi:10.1214/aos/1176348910.
13. Rousseeuw PJ, Leroy A. Robust regression and outlier detection. Hoboken, NJ, USA: John Wiley & Sons Inc.; 1987.
14. Custom Weather. Historic weather data for Lahore [Internet]. [cited 2025 Nov 1]. Available from: https://www.timeanddate.com/weather/pakistan/lahore/historic.
15. Juarez PO, Bouza CN, Mareno AS, Vallejo JM. Procedures for scrambling sensitive quantitative variables: an updated review. Rev Investig Operacional. 2023;44(3):438–50.