Outlier-Resistant Neutrosophic Ratio Estimators based on Generalized M-Estimators

Khaled Ali Abuhasel*

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

INFORMATION

Keywords:
Robust estimation
neutrosophic statistics
generalized M-estimators
engineering uncertainty
environmental monitoring
sensitivity analysis
convergence behavior

DOI: 10.23967/j.rimni.2026.10.76191

Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería

Outlier-Resistant Neutrosophic Ratio Estimators based on Generalized M-Estimators

Khaled Ali Abuhasel*

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

ABSTRACT

Building efficient ratio-type estimators of population parameters, especially the mean and variance, has been a major theme in sampling theory. However, the growing frequency of dirty, inaccurate, and incomplete information still poses a threat to the credibility of the classic estimation processes, especially in environments where outliers are likely to occur. The paper derives a generalized type of neutrosophic robust ratio-type estimator and regression-type estimators that have been developed on the M-estimation platform, including Huber M-estimators and generalized M-estimators (Viz., Mallows-GM, Schweppes-GM, and SIS-GM) formulations, and also incorporating the auxiliary information of the Hodges-Lehmann estimator. The estimators are designed to be asymptotically efficient in clean data models and apply well to contamination and heavy-tailed error distributions. Through a comparative study using real-world data and a Monte Carlo simulation experiment, it is shown that the proposed estimators show better performance, numerical stability, and robustness compared to the current methods in the presence of uncertainty. The simulation results confirm their resilience to contamination and heavy-tailed distributions across varying contamination levels. An application to environmental data involving temperature measurements subject to measurement error and outliers further illustrates the practical relevance of the framework. Collectively, the empirical and simulation evidence support the applicability of the proposed methodology to industrial process analysis and environmental monitoring systems characterized by data imprecision and contamination.

OPEN ACCESS

Received: 16/11/2025

Accepted: 07/01/2026

Published: 20/03/2026

DOI
10.23967/j.rimni.2026.10.76191

Keywords:
Robust estimation
neutrosophic statistics
generalized M-estimators
engineering uncertainty
environmental monitoring
sensitivity analysis
convergence behavior

1 Introduction

In classical statistics, crisp numbers are customarily used to model data, and estimators frequently use auxiliary information to enhance the estimation of population parameters, including mean and variance in finite populations. The precision of the use of the auxiliary information has always been improved, but the relationship between the study (response) variable and the auxiliary variable is critical to the effectiveness of such estimators. In cases where such a relationship is positive, ratio-type estimators are more efficient, and product-type estimators are more appropriate when such a relationship is negative. It is obvious that classical statistics typically assumes data to be

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

precise and deterministic, leading to specific point estimates. However, real-world data are often imprecise, ambiguous, or interval-valued, for example, daily fluctuations in stock prices or variations in city temperatures. In such contexts, fuzzy statistics, introduced by [1], provide greater flexibility by accounting for uncertainty and vagueness. A major limitation of fuzzy statistics, however, lies in its inability to explicitly quantify indeterminacy. To address this, Ref. [2] proposed the framework of neutrosophic statistics, which generalizes both classical and fuzzy methodologies. Building on this, Ref. [3] studied the estimation of the ratio between a crisp and a neutrosophic variable, extending traditional ratio estimation to accommodate indeterminacy, inconsistency, and incompleteness in datasets. Neutrosophic statistics has gained significant attention due to its ability to address indeterminate and inconsistent information. For details, see [4–8].

Motivated by [9–13], the present study proposes generalized robust ratio estimators utilizing generalized M-estimation techniques, namely Mallows GM, Schweppe GM, and SIS GM estimators, instead of OLS and Huber M-estimators as employed by [9]. Furthermore, the proposed approach incorporates auxiliary information of the Hodges–Lehmann estimator, a robust statistic that computes central tendency and location shift using pairwise medians of averages or differences. Since the Huber M-estimator [14] is the least squares efficient in cases of normality, together with being robust against contamination, it uses a loss function that becomes linearly dependent on increasing residuals. This produces the best Gaussian efficiency with a low degree of outlier influence. However, generalized M-estimators extend this principle through adaptive weight and influence functions tailored to specific contamination structures, thereby offering superior performance in regression and location-scale models, especially under heterogeneous or non-Gaussian data. Thus, the current research is based on the existing sound regression principles as it goes beyond an ordinary methodological extension. Current neutrosophic ratio and regression estimators use mostly classical OLS or even simple M-estimation. Instead, this paper builds a coherent generalized M-estimation based neutrosophic regression-ratio framework which incorporates Mallows-GM, Schweppes-GM, and SIS-GM processes in a finite population sample under indeterminacy. This integration allows the contemporaneous control of outliers, leverage points, heavy-tailed errors, and the neutrosophic uncertainty. The newness then consists in the fact that the high-breakdown strong regression operations are entered in a neutrosophic sampling model, explicit expressions of the bias and MSE are given in indeterminacy, and efficiency dominance conditions are established aspects that have never been treated simultaneously in earlier neutrosophic estimation exercises.

Motivation for Neutrosophic Framework: The main models of imprecision are interval-based and fuzzy, do not directly incorporate indeterminacy due to partial, inconsistent or unidentifiable observations. Conversely, neutrosophic statistics models indeterminacy as a distinct element of uncertainty, especially in data used in engineering and other environments with sensors and other sources of indeterminacy such as missing values, sensor error, indeterminacy in reporting etc. in this type of setting, the actual value cannot be singly determined over a given period of observation. The neutrosophic model thus offers a more detailed and interpretable description of uncertainty than fuzzy or interval models and is therefore worth using in the current study. The temperature dataset considered in this study inherently satisfies these conditions, as monthly observations fluctuate within recorded bounds across years, justifying the neutrosophic formulation.

2 Terminology

Consider a neutrosophic random sample n_{+i} and finite population of N units $(A_1, A_2, \ldots, A_N)$ is used to generate a random sample of a given size. Take into consideration the i^{th} sample observation of our data on neutrosophic phenomena, with a form $t_{+i} \in (t_L, t_U)$ and an auxiliary variable of same form $v_{+i} \in (v_L, v_U)$. Due to the interval structure of our neutrosophic data and to address inherent indeterminacy and ambiguity at both the endpoints, a defuzzification strategy was implemented by averaging two boundary levels, thereby yielding a crisp intermediate value that serves as a representative estimate within uncertain interval and also let $\bar{t}_{+i} \in (\bar{t}_L, \bar{t}_U)$ be our research variable and $\bar{v}_{+i} \in (\bar{v}_L, \bar{v}_U)$ be our assisting neutrosophic variable that is correlated with it, where $\bar{t}_M$ and $\bar{v}_M$ are regulated parts or determined parts which are respectively average of two extreme imprecise levels $(\bar{t}_L, \bar{t}_U)$ and $(\bar{v}_L, \bar{v}_U)$. In general, averages of data collected from neutrosophic sources are $\overline{T}$ and $\overline{V}$. $C_{t+i} \in (C_{t+L}, C_{t+U})$ and $C_{v+i} \in (C_{v+L}, C_{v+U})$, respectively, are neutrosophic coefficients of variation of $\bar{t}$ and $\bar{v}$. $\rho_{v+i,t+i}$ is correlation between $\bar{t}$ and $\bar{v}$ neutrosophic variables and $HL = median\left[\left(v_j + v_k\right)/2\right], 1 \leq j \leq k \leq N$ is the Hodges-Lehmann estimator. The diagrammatic representation of procedural framework for implementing proposed techniques within neutrosophic number system is given in Fig. 1.

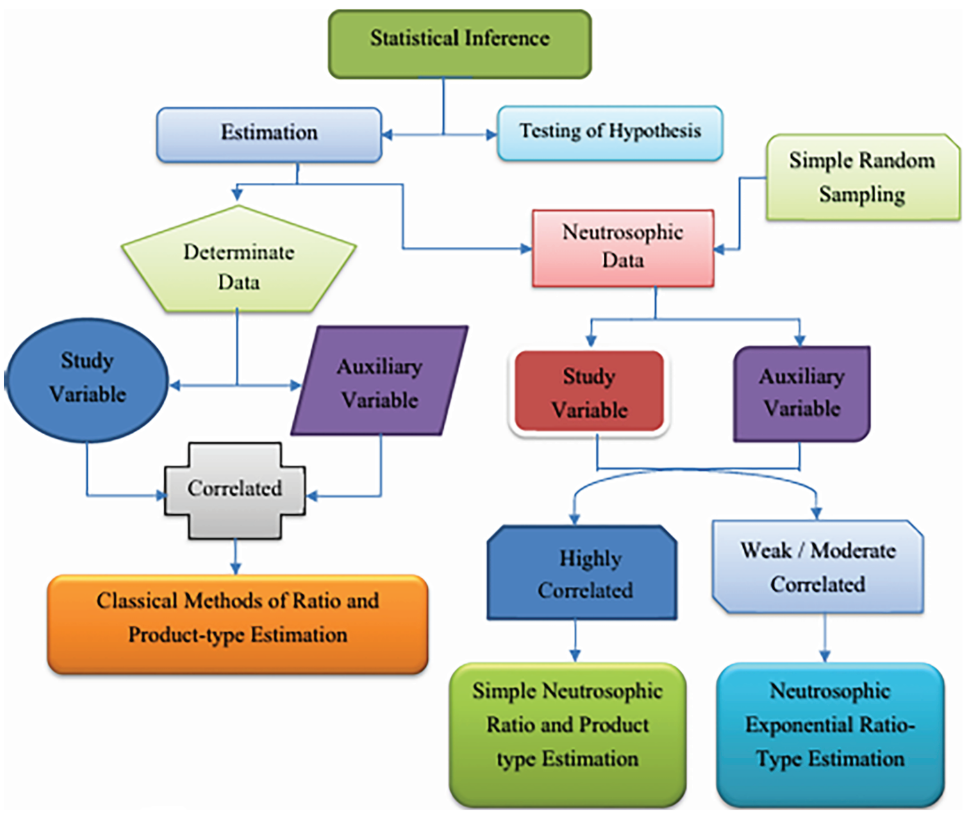

Figure 1: Procedural framework of the proposed neutrosophic estimation techniques

In this study there is no introduction of multiple estimators, but to show a structured robustness spectrum. OLS, Huber-M and GM-type estimates vary essentially in terms of influence control, leverage treatment as well as breakdown behaviour. Introducing these estimators into a single neutrosophic framework will enable practitioners to choose an estimator that fits the level of contamination severity and leverage structure, instead of having one robust alternative.

3 A Proposed OLS Method Based on Generalized Neutrosophic Ratio Estimator Using Hodges Lehmann

A generalized neutrosophic ratio estimator is formulated to achieve robustness against aberrant observations and data irregularities inherent in neutrosophic measurements. The estimator is specified as follows:

$$\bar{t}_{G(Rp+i)} = \left[\bar{t}_{+i} + k\psi\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\bar{v}_{+i}\delta + \eta}{\overline{V}\delta + \eta}\right]^{k\theta} \tag{1}$$

where $k = -1$ for ratio estimator and $k = +1$ for product estimator, $\theta = \rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)$, $\delta = 1$, $\eta = HL$ and $\psi = \psi_{OLS}$, then our proposed estimator takes shape like this

$$\bar{t}_{s(Rp+i)} = \left[\bar{t}_{+i} - \psi_{OLS}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\bar{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{2}$$

where ψ_{OLS} is obtained as $\left[S_{(v+i)(t+i)}/S^2_{v+i}\right]$. In order to evaluate statistical properties of proposed robust neutrosophic ratio estimator via OLS framework, we proceed to derive expressions for its bias and mean square error under standard finite-population regularity assumptions:

(i) bounded second moments of the study and auxiliary variables;

(ii) small error approximation up to terms of order $O\left(n^{-1}\right)$;

(iii) consistency of robust regression coefficients under contamination; and

(iv) differentiability of the underlying loss and influence functions.

These assumptions ensure the validity of the Taylor series expansions used in subsequent derivations.

$$\varphi_0 = \frac{\bar{t}_{+i} - \overline{T}}{\overline{T}}, \phi_1 = \frac{\bar{v}_{+i} - \overline{V}}{\overline{V}}, E\left(\phi_0^2\right) = \frac{1-f}{n_{+i}}C^2_{t+i}, E\left(\varphi_1^2\right) = \frac{1-f}{n_{+i}}C^2_{v+i}, f = \frac{n_{+i}}{N}$$

$$E\left(\varphi_0\varphi_1\right) = \frac{1-f}{n_{+i}}\rho_{v+i,t+i}C_{t+i}C_{v+i} \tag{3}$$

After applying Eq. (3) to transform Eq. (2), we have

$$\bar{t}_{s(Rp+i)} = \left[\overline{T}\left(1+\varphi_0\right) + \overline{V}k\psi_{OLS}\varphi_1\right]\left[1+\zeta_i\varphi_1\right]^{k\theta} \tag{4}$$

where $\zeta_i = \left[\delta\overline{V}/\delta\overline{V} + \eta\right]$, $\delta = 1$, $\eta = HL$.

The expressions are obtained by applying a second–order Taylor series expansion of the estimator $[1+\zeta_i\varphi_1]^{k\theta}$ for Eq. (4) around the true population mean under the small-error approximation framework. Higher-order terms are neglected due to their asymptotically negligible contribution under simple random sampling without replacement, we have

$$\bar{t}_{s(Rp+i)} = \overline{T}\left(1 + \varphi_0 + k\psi_{OLS}Z\varphi_1\right)\left[1 + k\theta\zeta_i\varphi_1 + \frac{k\theta\left(k\theta-1\right)}{2!}\zeta_i^2\varphi_1^2 + \ldots\right] \tag{5}$$

Therefore, bias of estimator is

$$B\left(\bar{t}_{s(Rp+i)}\right) = E\left(\bar{t}_{s(Rp+i)} - \overline{T}\right)$$
$$= \frac{1-f}{n_{+i}}\overline{T}\left\{\left[\frac{k\theta\left(k\theta-1\right)}{2}\zeta_i^2 + k^2\theta\zeta_i\psi_{OLS}Z\right]C^2_{v+i} + k\theta\zeta_i\rho_{v+i,t+i}C_{t+i}C_{v+i}\right\} \tag{6}$$

Taylor series approximation of proposed estimator is also used to generate mean square error statement; Eq. (2) provides this approximation.

$$MSE\left(\bar{t}_{s(Rp+i)}\right) = E\left(\bar{t}_{s(Rp+t)} - \overline{T}\right)^2 = E\left\{\overline{T}\left[\varphi_o + k\left(\theta\zeta_i + \psi_{OLS} Z\varphi_1\right)\right]\right\}^2$$
$$= \frac{1-f}{n_{+i}}\overline{T}\left\{C_{t+i}^2 + 2k\left(\theta\zeta_i + \psi_{OLS}Z\right)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2\left(\theta\zeta_i + \psi_{OLS}Z\right)^2 C_{v+i}^2\right\},$$
$$Z = \frac{\overline{V}}{\overline{T}} = \frac{1}{R} \quad (7)$$

4 Proposed Robust Neutrosophic Ratio Estimators Using Robust Regression Techniques & Incorporating Hodges Lehmann Estimator as Auxiliary Information

In practical applications, data are frequently asymmetric and affected by outliers, rendering classical estimators such as OLS highly sensitive and unreliable. This study undertakes two investigations: first, an OLS-based estimator incorporating Hodges-Lehmann auxiliary information and second, a class of robust estimators employing the same auxiliary information and utilizing robust regression techniques viz., Huber M, Mallows GM, Schweppes GM and SIS GM estimation function. As demonstrated in Fig. 2, OLS exhibits substantial distortion under contamination, whereas GM estimators mitigate the influence of extreme and high-leverage observations through adaptive weighting mechanism as mentioned in Section 1 of this study and thus establishing their clear superiority for statistical inference in contaminated data environments.

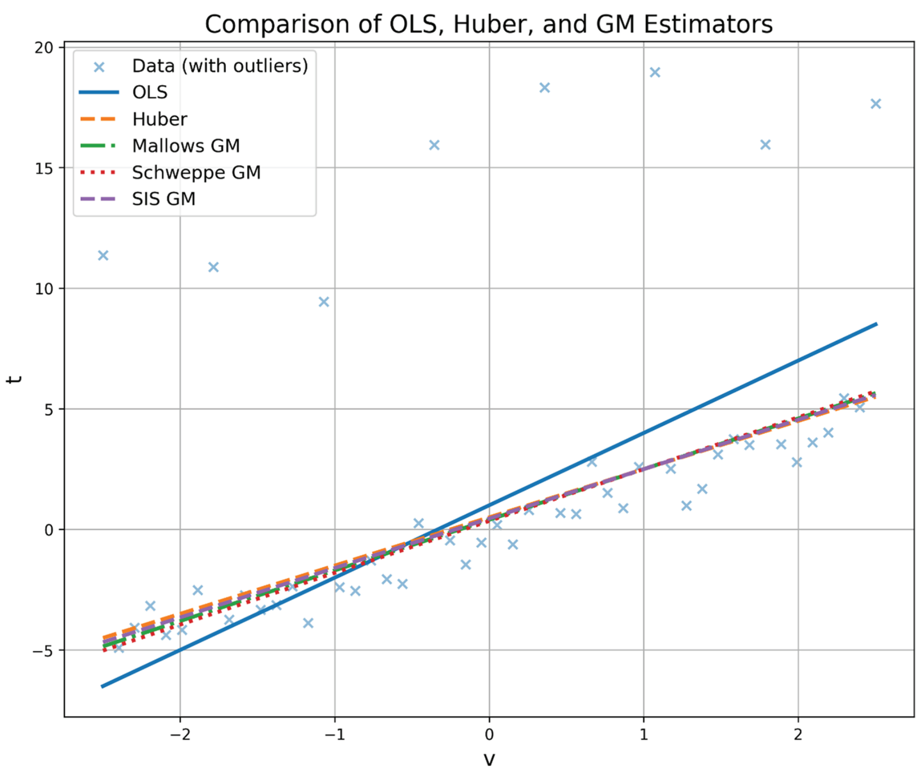

Figure 2: Comparison of OLS, Huber M and GM Estimators under data contamination, highlighting sensitivity and robustness

Therefore, in the present study, we first compare the OLS estimator with other robust regression techniques and subsequently conduct a comparative assessment of the Huber M-estimation with the aforementioned Generalized M robust regression methods. The proposed estimator is defined as follows:

$$\bar{t}_{p(Rp+i)} = \left[\bar{t}_{+i} - \psi_{rob}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\overline{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{8}$$

where ψ_{rob} is obtained using the Huber-M, Mallows-GM, Schweppes-GM and SIS-GM estimation function. The Huber M-estimation has also been employed by [9] to obtain precise estimates. While traditional M-estimator rely solely on residuals to minimize a loss function under specific distribution assumptions, GM-estimators extend this framework by incorporating robust weighting schemes that account heteroscedasticity and leverage points. Consequently, GM-estimators achieve superior robustness and efficiency in the presence of outliers and model deviations. To derive the analytical expressions for bias and mean square error associated with the proposed class of estimators, Eq. (3) is invoked, yielding the following theoretical formulations:

$$\bar{t}_{p(Rp+i)} = \left[\overline{T}\left(1 + \varphi_0\right) + \overline{V}k\psi_{rob}\varphi_1\right]\left[1 + \zeta_i\varphi_1\right]^{k\theta} \tag{9}$$

where $\zeta_i = \left[\delta\overline{V}/\left(\delta\overline{V} + \eta\right)\right]$, $\delta = 1$, $\eta = HL$. Following algebraic simplification of Eq. (9), a second-order $\left[1 + \zeta_i\varphi_1\right]^{k\theta}$, taylor series expansion is employed to approximate the estimator, yielding the subsequent analytical expression

$$\bar{t}_{p(Rp+i)} = \overline{T}\left(1 + \varphi_0 + k\psi_{rob}Z\varphi_1\right)\left[1 + k\theta\zeta_i\varphi_1 + \frac{k\theta\left(k\theta - 1\right)}{2!}\zeta_i^2\varphi_1^2 + \ldots\right] \tag{10}$$

As a result, the bias associated with the proposed estimator is derived as follows:

$$\begin{aligned} B\left(\bar{t}_{p(Rp+i)}\right) &= E\left(\bar{t}_{p(Rp+t)} - \overline{T}\right) \\ &= \frac{1-f}{n_{+i}}\overline{T}\left\{\left[\frac{k\theta\left(k\theta - 1\right)}{2}\zeta_i^2 + k^2\theta\zeta_i\psi_{rob}Z\right]C_{v+i}^2 + k\theta\zeta_i\rho_{v+i,t+i}C_{t+i}C_{v+i}\right\} \end{aligned} \tag{11}$$

Furthermore, Taylor series approximation of the proposed estimator yields the corresponding mean square error expression, as presented in Eq. (8).

$$\begin{aligned} MSE\left(\bar{t}_{p(Rp+i)}\right) &= E\left(\bar{t}_{p(Rp+t)} - \overline{T}\right)^2 = E\left\{\overline{T}\left[\varphi_o + k\left(\theta\zeta_i + \psi_{rob}Z\varphi_1\right)\right]\right\}^2 \\ &= \frac{1-f}{n_{+i}}\overline{T}\left\{C_{t+i}^2 + 2k\left(\theta\zeta_i + \psi_{rob}Z\right)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2\left(\theta\zeta_i + \psi_{rob}Z\right)^2 C_{v+i}^2\right\}, \\ Z &= \frac{\overline{V}}{\overline{T}} = \frac{1}{R} \end{aligned} \tag{12}$$

A series of robust neutrosophic ratio-type estimators are systematically derived from the proposed generalized class by appropriately substituting ψ_{rob} in Eq. (8) with the Huber M, Mallows GM, Schweppes GM and SIS GM estimation functions. Consequently, the following estimators emerge in the specified forms:

$$\bar{t}_{p1(Rp+i)} = \left[\bar{t}_{+i} - \psi_{Huber}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\overline{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{13}$$

$$\bar{t}_{p2(Rp+i)} = \left[\bar{t}_{+i} - \psi_{Mallows\,GM}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\bar{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{14}$$

$$\bar{t}_{p3(Rp+i)} = \left[\bar{t}_{+i} - \psi_{Schweppes\ \ GM}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\bar{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{15}$$

$$\bar{t}_{p4(Rp+i)} = \left[\bar{t}_{+i} - \psi_{SIS\ GM}\left(\bar{v}_{+i} - \overline{V}\right)\right]\left[\frac{\bar{v}_{+i} + HL}{\overline{V} + HL}\right]^{-\rho_{v+i,t+i}\left(C_{t+i}/C_{v+i}\right)} \tag{16}$$

In this section, we further propose a class of neutrosophic robust regression estimators that exhibit consistent and unequivocal efficiency across diverse sampling scenarios. These estimators are described to be of structural simplicity since they only use the regression element and deliberately do not use the ratio-type adjustments used in their adjusted counterparts. The suggested estimator type will be based on the following definitions:

$$\bar{t}_{p5(Rp+i)} = \left[\bar{t}_{+i} + \psi_{OLSr}\left(\overline{V} - \bar{v}_{+i}\right)\right] \tag{17}$$

$$\bar{t}_{p6(Rp+i)} = \left[\bar{t}_{+i} + \psi_{Huber}\left(\overline{V} - \bar{v}_{+i}\right)\right] \tag{18}$$

$$\bar{t}_{p7(Rp+i)} = \left[\bar{t}_{+i} + \psi_{Mallows\,GM}\left(\overline{V} - \bar{v}_{+i}\right)\right] \tag{19}$$

$$\bar{t}_{p8(Rp+i)} = \left[\bar{t}_{+i} + \psi_{Schweppes\,GM}\left(\overline{V} - \bar{v}_{+i}\right)\right] \tag{20}$$

$$\bar{t}_{p9(Rp+i)} = \left[\bar{t}_{+i} + \psi_{SIS\ GM}\left(\overline{V} - \bar{v}_{+i}\right)\right] \tag{21}$$

The proposed estimators need relatively a smaller auxiliary data, which leads to the simplified structural form not considering any ratio-type adjustment but only the regression-type adjustment. The estimator of each of them is built with the help of strong regression coefficients namely OLS (ψ_{OLS}), Huber regression coefficient (ψ_{Huber}), Mallows GM regression coefficient ($\psi_{Mallows\,GM}$), Schweppes GM regression coefficient $\left(\psi_{Schweppess\,GM}\right)$ and SIS GM regression coefficient ($\psi_{SIS\ GM}$). The formulations that have been used to represent these estimators have been stated in the previous section. Based on these definitions and simple straight forward algebraic manipulations to avoid unnecessary or tedious computations, we obtain the formulas to the mean squared error (MSE) of the suggested set of estimators, to the order of n^{-1} in the following form:

$$MSE\left(\bar{t}_{p5(Rp+i)}\right) = \left[Var\left(\bar{t}_{+i}\right) - 2\psi_{OLS}Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) + \psi^2_{OLSr}Var\left(\bar{v}_{+i}\right)\right] \tag{22}$$

$$MSE\left(\bar{t}_{p6(Rp+i)}\right) = \left[Var\left(\bar{t}_{+i}\right) - 2\psi_{Huber}Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) + \psi^2_{Huber}Var\left(\bar{v}_{+i}\right)\right] \tag{23}$$

$$MSE\left(\bar{t}_{p7(Rp+i)}\right) = \left[Var\left(\bar{t}_{+i}\right) - 2\psi_{Mallows\,GM}Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) + \psi^2_{Mallows\,GM}Var\left(\bar{v}_{+i}\right)\right] \tag{24}$$

$$MSE\left(\bar{t}_{p8(Rp+i)}\right) = \left[Var\left(\bar{t}_{+i}\right) - 2\psi_{Schweppes\,GM}Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) + \psi^2_{Schweppes\,GM}Var\left(\bar{v}_{+i}\right)\right] \tag{25}$$

$$MSE\left(\bar{t}_{p9(Rp+i)}\right) = \left[Var\left(\bar{t}_{+i}\right) - 2\psi_{SIS\ GM}Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) + \psi^2_{SIS\ GM}Var\left(\bar{v}_{+i}\right)\right] \tag{26}$$

So on substituting $Var\left(\bar{t}_{+i}\right) = \upsilon S^2_{t+i} Cov\left(\bar{t}_{+i}, \bar{v}_{+i}\right) = \upsilon \rho_{t+i,,v+i} S_{t+i} S_{v+i}$, $Var\left(\bar{v}_{+i}\right) = \upsilon S^2_{v+i}$ to get final expressions of MSE as follows:

$$MSE\left(\bar{t}_{p5(Rp+i)}\right) = \upsilon\left[\overline{T}^2 C^2_{t+i} - 2\psi_{OLS}\overline{T}\,\overline{V}\rho_{t+i,v+i}C_{t+i}C_{v+i} + \psi^2_{OLS}\overline{V}^2 C^2_{v+i}\right] \tag{27}$$

$$MSE\left(\bar{t}_{p6(Rp+i)}\right) = \upsilon\left[\overline{T}^2 C^2_{t+i} - 2\psi_{Huber}\overline{T}\,\overline{V}\rho_{t+i,v+i}C_{t+i}C_{v+i} + \psi^2_{Huber}\overline{V}^2 C^2_{v+i}\right] \tag{28}$$

$$MSE\left(\bar{t}_{p7(Rp+i)}\right) = \upsilon\left[\overline{T}^2 C^2_{t+i} - 2\psi_{Mallows\,GM}\overline{T}\,\overline{V}\rho_{t+i,v+i}C_{t+i}C_{v+i} + \psi^2_{Mallows\,GM}\overline{V}^2 C^2_{v+i}\right] \tag{29}$$

$$MSE\left(\bar{t}_{p8(Rp+i)}\right) = \upsilon\left[\overline{T}^2 C^2_{t+i} - 2\psi_{Schweppes\,GM}\overline{T}\,\overline{V}\rho_{t+i,v+i}C_{t+i}C_{v+i} + \psi^2_{Schweppes\,GM}\overline{V}^2 C^2_{v+i}\right] \tag{30}$$

$$MSE\left(\bar{t}_{p9(Rp+i)}\right) = \upsilon\left[\overline{T}^2 C^2_{t+i} - 2\psi_{SIS\,GM}\overline{T}\,\overline{V}\rho_{t+i,v+i}C_{t+i}C_{v+i} + \psi^2_{SIS\,GM}\overline{V}^2 C^2_{v+i}\right] \tag{31}$$

where $\upsilon = \left(n^{-1}_{+i} - N^{-1}\right)$.

5 Comparative Evaluation

The comparative efficiency analysis is structured by directly contrasting the MSE expressions of competing estimators under identical auxiliary information. Efficiency dominance is established by analyzing the sign of the difference between the corresponding MSE functions. The inequalities derived in this section provide sufficient conditions under which robust and GM-based estimators outperform classical OLS and Huber-M estimators, thereby offering a mathematically transparent basis for comparison.

(a) Comparison between estimators using OLS and Robust regression techniques incorporating the same ancillary information

Thus, in order for $\bar{t}_{p(Rp+i)}$ to be more efficient than $\bar{t}_{s(Rp+i)}$, we have

$$\Rightarrow 2k\left(\theta\zeta_i + \psi_{rob}Z\right)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2\left(\theta\zeta_i + \psi_{rob}Z\right)^2 C^2_{v+i} <$$
$$2k\left(\theta\zeta_i + \psi_{OLS}Z\right)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2\left(\theta\zeta_i + \psi_{OLS}Z\right)^2 C^2_{v+i}$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}k\left[\theta\zeta_i + \psi_{rob}Z - \theta\zeta_i - \psi_{OLS}Z\right] + C^2_{v+i}k^2\left[\left(\theta\zeta_i + \psi_{rob}Z\right)^2 - \left(\theta\zeta_i + \psi_{OLS}Z\right)^2\right] < 0$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}Zk\left[\psi_{rob} - \psi_{OLS}\right] + C^2_{v+i}k^2\left[\left(\theta\zeta_i + \psi_{rob}Z\right) - \left(\theta\zeta_i + \psi_{OLS}Z\right)\right]$$
$$\left[\theta\zeta_i + \psi_{rob}Z + \theta\zeta_i + \psi_{OLS}Z\right] < 0$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}Zk\left[\psi_{rob} - \psi_{OLS}\right] + C^2_{v+i}k^2\left[\left(\theta\zeta_i + \psi_{rob}Z\right) - \left(\theta\zeta_i + \psi_{OLS}Z\right)\right]$$
$$\left[\theta\zeta_i + \psi_{rob}Z + \theta\zeta_i + \psi_{OLS}Z\right] < 0$$

$\Rightarrow 2W\theta\left[\psi_{rob} - \psi_{OLS}\right] + k\left[\left(\psi_{rob} - \psi_{OLS}\right)Z\right]\left[2\theta\zeta_i + Z\left(\psi_{rob} + \psi_{OLS}\right)\right] < 0$, where $W = \rho_{v+i,t+i}C_{t+i}C_{v+i}Zk$

$$\Rightarrow W\left[\psi_{rob} - \psi_{OLS}\right]2\theta + k\left(2\theta\zeta_i + Z\left(\psi_{rob} + \psi_{OLS}\right)\right) < 0$$

$$\Rightarrow W\left[\psi_{rob} - \psi_{OLS}\right]2\theta\left(1 + k\zeta_i\right) + Zk\left(\psi_{rob} + \psi_{OLS}\right) < 0$$

$$\Rightarrow \left[\psi_{rob} - \psi_{OLS}\right]\left[2\theta\left(1 + k\zeta_i\right)\right] + Zk\left(\psi_{rob} + \psi_{OLS}\right) < 0$$

Since $Z > 0$, either $(\psi_{rob} - \psi_{OLS}) < 0$ and $[2\theta(1 + k\zeta_i)] + Zk(\psi_{rob} + \psi_{OLS}) < 0$. This implies that

$$\Rightarrow \psi_{rob} < \psi_{OLS} \text{ and } [2\theta(1 + k\zeta_i)] > -Zk(\psi_{rob} + \psi_{OLS}) < 0 \quad (32)$$

$$\text{or } \psi_{rob} > \psi_{OLS} \text{ and } [2\theta(1 + k\zeta_i)] < -Zk(\psi_{rob} + \psi_{OLS}) < 0 \quad (33)$$

Based on the above theoretical comparison, when the conditions specified in Eqs. (32) or (33) are satisfied, the generalized class of estimators employing robust regression techniques demonstrates superior profiency compared to the class based on classical OLS.

(b) Comparison between estimators using Huber M and Generalized M Robust regression techniques incorporating the same ancillary information

Thus, in order for $\bar{t}_{p_i(Rp+i)}$ to be more efficient than $\bar{t}_{p1(Rp+i)}$, $pi = p2, p3, p4$ we have

$$\Rightarrow 2k(\theta\zeta_i + \psi_{rob\,GM}Z)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2(\theta\zeta_i + \psi_{rob\,GM}Z)^2 C_{v+i}^2 <$$
$$2k(\theta\zeta_i + \psi_{Huber}Z)\rho_{v+i,t+i}C_{t+i}C_{v+i} + k^2(\theta\zeta_i + \psi_{Huber}Z)^2 C_{v+i}^2$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}k[\theta\zeta_i + \psi_{rob\,GM}Z - \theta\zeta_i - \psi_{Huber}Z] + C_{v+i}^2k^2$$
$$\left[(\theta\zeta_i + \psi_{rob\,GM}Z)^2 - (\theta\zeta_i + \psi_{Huber}Z)^2\right] < 0$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}Zk[\psi_{rob\,GM} - \psi_{Huber}] + C_{v+i}^2k^2[(\theta\zeta_i + \psi_{rob\,GM}Z) - (\theta\zeta_i + \psi_{Huber}Z)]$$
$$[\theta\zeta_i + \psi_{rob\,GM}Z + \theta\zeta_i + \psi_{Huber}Z] < 0$$

$$\Rightarrow 2\rho_{v+i,t+i}C_{t+i}C_{v+i}Zk[\psi_{rob\,GM} - \psi_{Huber}] + C_{v+i}^2k^2[(\theta\zeta_i + \psi_{rob\,GM}Z) - (\theta\zeta_i + \psi_{Huber}Z)]$$
$$[\theta\zeta_i + \psi_{rob\,GM}Z + \theta\zeta_i + \psi_{Huber}Z] < 0$$

$$\Rightarrow 2W\theta[\psi_{rob\,GM} - \psi_{Huber}] + k[(\psi_{rob\,GM} - \psi_{Huber})Z][2\theta\zeta_i + Z(\psi_{rob\,GM} + \psi_{Huber})] < 0,$$

where $W = \rho_{v+i,t+i}C_{t+i}C_{v+i}Zk$.

$$\Rightarrow W[\psi_{rob\,GM} - \psi_{Huber}]2\theta + k(2\theta\zeta_i + Z(\psi_{rob\,GM} + \psi_{Huber})) < 0$$

$$\Rightarrow W[\psi_{rob\,GM} - \psi_{Huber}]2\theta(1 + k\zeta_i) + Zk(\psi_{rob\,GM} + \psi_{Huber}) < 0$$

$$\Rightarrow [\psi_{rob\,GM} - \psi_{Huber}][2\theta(1 + k\zeta_i)] + Zk(\psi_{rob\,GM} + \psi_{Huber}) < 0$$

Since $Z > 0$, either $(\psi_{rob\,GM} - \psi_{Huber}) < 0$ and $[2\theta(1 + k\zeta_i)] + Zk(\psi_{rob\,GM} + \psi_{Huber}) < 0$. This implies that

$$\Rightarrow \psi_{rob\,GM} < \psi_{Huber} \text{ and } [2\theta(1 + k\zeta_i)] > -Zk(\psi_{rob\,GM} + \psi_{Huber}) < 0 \quad (34)$$

$$\text{or } \psi_{rob\,GM} > \psi_{Huber} \text{ and } [2\theta(1 + k\zeta_i)] < -Zk(\psi_{rob\,GM} + \psi_{Huber}) < 0 \quad (35)$$

When conditions specified in Eqs. (34) or (35) are satisfied, it is concluded that estimators based on Generalized M techniques outperform those using the Huber M in terms of efficiency, which is consistent with the findings of [15].

6 Numerical Study

For numerical analysis, we have considered real-life indeterminate interval data for temperature, recognizing that daily temperature exhibits neutrosophic characteristics fluctuating within an interval and containing vague or uncertain values. One of the primary reasons for modeling temperature as neutrosophic data is its inherent variability: the reported daily temperature may correspond to minimum, maximum, or any intermediate value recorded during that day. We compiled six years of monthly temperature data (2014–2019) for Lahore, Punjab, Pakistan, obtained from publicly available weather websites [16], as presented in Table 1. This data is freely accessible online, and therefore, no ethical approval was required. For each month, we recorded average of lowest and highest temperatures for each of six years. The variable V represents time, coded from 1 to 6, corresponding to number of years. The neutrosophic average values of lower and upper temperature bounds for each month were calculated. These form neutrosophic data T, associated with known year V. Hence, monthly total averages over six-year period are treated as neutrosophic data. Temperature is thus represented as neutrosophic data in the form (Temp, $t_{+i} \in [t_L, t_U]$), where t_L and t_U denote lower and upper limits of monthly temperature interval, respectively. The average of six-year time variable is expressed as $v_{+i} \in [v_L, v_U]$, which remains constant across all temperature intervals. Additionally, HL denotes Hodges-Lehmann estimator. Table 2 reports the mean square error (MSE) values of the proposed neutrosophic ratio estimators, while the corresponding percent relative efficiency (PRE) of the neutrosophic ratio estimators is presented in Table 3. Similarly, Table 4 summarizes the MSE of the neutrosophic regression estimator and its associated PRE are reported in Table 5.

Although the illustrative temperature dataset consists of six annual observations per month, it is used solely as a demonstrative case study to highlight estimator behavior under neutrosophic intervals. The primary validation of estimator performance is therefore supported by extensive Monte Carlo simulations with 10,000 replications, ensuring that the empirical conclusions do not rely on the small sample illustration alone.

The suggested framework is placed as a general inference procedure that can be used in uncertain and contaminated measurements that are typically found in engineering and environmental monitoring. As much as the current study provides a methodological basis, additional validation using large domain specific datasets would reinforce application based conclusions.

7 Simulation and Sensitivity Analysis

In addition to MSE and bias, robustness was assessed indirectly through contamination sensitivity and stability of estimator's performance across increasing contamination levels. While formal breakdown point and influence function derivations are beyond the scope of this study, the observed stability of GM-type estimators under up to 20% contamination is consistent with their known high-breakdown and bounded-influence properties documented in robust regression literature.

A thorough Monte Carlo simulation with 10,000 replications was carried out under various contamination levels (0%, 5%, 10%, and 20%) using a gross-error model defined as

$$T_i = (1 - \varepsilon)\, T_i^{(0)} + \varepsilon T_i^{(c)}$$

where $T_i^{(0)}$ follows the assumed baseline distribution and $T_i^{(c)}$ is drawn from a heavy-tailed distribution. The contamination proportion ε varies across simulation scenarios.

Table 1: Auxiliary variable characteristics of population

Parameters	Source (Data): Lahore, Punjab, Pakistan's temperature between 2014 and 2019									
	Population available: N = 30 (years); sample taken: n = 6 (years)									
	No. of Year V	Average temperature (Max, Min) T								
$\overline{V}$	$\overline{T}_{+i}$	C_{t+i}	C_{v+i}	$\rho_{t+i,v+i}$	ψ_{OLS}	ψ_{Huber}	$\psi_{Mallows\ GM}$	$\psi_{Schweppes\ GM}$	$\psi_{SIS\ GM}$	HL
(3.5, 3.5)	(64, 44)	(0.0360, 0.0345)	(0.53, 0.53)	(0.40, 0.23)	(0.497, 0.188)	(0.416, 0.113)	(0.403, 0.107)	(0.397, 0.101)	(0.380, 0.090)	(2.567, 1.956)
	(72, 50)	(0.0529, 0.0424)		(0.18, 0.14)	(0.370, 0.160)	(0.260, 0.107)	(0.215, 0.101)	(0.208, 0.096)	(0.201, 0.083)	
	(80, 58)	(0.0454, 0.0424)		(0.43, 0.17)	(0.842, 0.225)	(0.665, 0.156)	(0.624, 0.143)	(0.614, 0.133)	(0.605, 0.125)	
	(94, 69)	(0.0291, 0.0293)		(0.60, 0.55)	(0.885, 0.599)	(0.678, 0.357)	(0.645, 0.324)	(0.632, 0.313)	(0.624, 0.302)	
	(102, 77)	(0.0187, 0.0282)		(−0.04, 0.04)	(−0.041, 0.047)	(−0.021, 0.029)	(−0.019, 0.027)	(−0.013, 0.026)	(−0.009, 0.017)	
	(103, 81)	(0.0316, 0.0272)		(−0.23, −0.08)	(−0.404, −0.095)	(−0.234, −0.047)	(−0.227, −0.043)	(−0.218, −0.036)	(−0.211, −0.025)	
	(95, 80)	(0.0171, 0.0151)		(−0.60, −0.49)	(−0.525, −0.319)	(−0.313, −0.167)	(−0.302, −0.157)	(−0.295, −0.142)	(−0.281, −0.135)	
	(95, 80)	(0.0141, 0.0108)		(−0.59, −0.53)	(−0.426, −0.247)	(−0.219, −0.116)	(−0.207, −0.104)	(−0.201, −0.100)	(−0.191, −0.092)	
	(94, 77)	(0.0231, 0.0254)		(0.09, 0.01)	(0.105, 0.011)	(0.068, 0.007)	(0.058, 0.006)	(0.052, 0.004)	(0.043, 0.003)	
	(90, 68)	(0.0277, 0.0338)		(−0.20, −0.39)	(−0.269, −0.483)	(−0.158, −0.250)	(−0.151, −0.237)	(−0.142, −0.228)	(−0.134, −0.215)	
	(79, 55)	(0.0233, 0.0260)		(−0.55, −0.42)	(−0.546, −0.133)	(−0.298, −0.122)	(−0.289, −0.115)	(−0.277, −0.111)	(−0.266, −0.105)	
	(69, 45)	(0.050, 0.0341)		(−0.05, −0.14)	(−0.093, −0.116)	(−0.046, −0.069)	(−0.042, −0.061)	(−0.038, −0.055)	(−0.025, −0.046)	

Table 2: Mean Square error of proposed Neutrosophic ratio estimators using OLS, Huber M & generalized M

$n = 6$	$MSE\left(\bar{t}_{s(Rp+i)}\right)$	$MSE\left(\bar{t}_{p1(Rp+i)}\right)$	$MSE\left(\bar{t}_{p2(Rp+i)}\right)$	$MSE\left(\bar{t}_{p3(Rp+i)}\right)$	$MSE\left(\bar{t}_{p4(Rp+i)}\right)$
Jan	0.009879357 0.006764878	0.009593328 0.006635020	0.009556182 0.006629699	0.009539855 0.006625129	0.009496398 0.006618700
Feb	0.026285120 0.011846830	0.026062740 0.011772770	0.026016220 0.011767630	0.026011300 0.011763860	0.026007010 0.011756190
March	0.019273900 0.011860140	0.018467350 0.011707110	0.018331780 0.011688060	0.018301640 0.011675520	0.018275490 0.011666810
April	0.008065292 0.006489940	0.007242574 0.005643202	0.007150072 0.005588081	0.007116551 0.005572925	0.007096742 0.005559379
May	0.004750203 0.008156840	0.004747835 0.008152290	0.004747796 0.008152022	0.004747895 0.008151907	0.004748141 0.008151401
June	0.014373180 0.007960167	0.013915910 0.007940082	0.013971660 0.007939587	0.013886130 0.007939155	0.013873820 0.007939600
July	0.004911308 0.002088241	0.004098842 0.001864027	0.004068532 0.001858567	0.004049853 0.001852528	0.004013914 0.001850592
August	0.001933296 0.001038754	0.001648901 0.000899012	0.001645107 0.000896053	0.001643731 0.000895434	0.001642211 0.000894745
Sep	0.006651561 0.006623291	0.006636418 0.006623042	0.006634620 0.006623010	0.006634009 0.006622981	0.006633752 0.006622984
Oct	0.009723542 0.009430058	0.009630350 0.008822410	0.009628341 0.008810084	0.009626431 0.008802887	0.009625369 0.008794420
Nov	0.007289144 0.004085250	0.006140053 0.004083153	0.006111785 0.004082869	0.006075558 0.004083075	0.006043820 0.004083883
Dec	0.022961640 0.006896673	0.022942790 0.006847813	0.022942550 0.006843983	0.022942510 0.006841967	0.022943870 0.006840318

Table 3: Percent relative efficiency of proposed Neutrosophic ratio estimators using generalized M with the estimators using OLS and Huber M

Month	RE of OLS v/s Huber	RE of OLS v/s Mallows GM	RE of OLS v/s Schweppes GM	RE of OLS v/s SIS GM
Jan	102.98154 101.95716	103.38184 102.03899	103.55878 102.10938	104.03268 102.20856
Feb	100.85325 100.62908	101.03359 100.67303	101.05270 100.70530	101.06937 100.77100
March	104.36744 101.30715	105.13927 101.47227	105.31242 101.58126	105.46311 101.65709
April	111.35947 115.00457	112.80015 116.13898	113.33147 116.45482	113.64781 116.73858
May	100.04988 100.05581	100.05070 100.05910	100.04861 100.06051	100.04343 100.06672
June	103.28595 100.25296	102.87382 100.25921	103.50746 100.26466	103.59930 100.25904
July	119.82184 112.02847	120.71450 112.35759	121.27127 112.72386	122.35708 112.84178
August	117.24755 115.54395	117.51795 115.92551	117.61633 116.00565	117.72519 116.09498
Sep	100.22818 100.00376	100.25534 100.00424	100.26458 100.00468	100.26846 100.00464
Oct	100.96769 106.88755	100.98876 107.03710	101.00880 107.12461	101.01994 107.22774
Nov	118.71468 100.05136	119.26375 100.05832	119.97489 100.05327	120.60492 100.03347
Dec	100.08216 100.71351	100.08321 100.76987	100.08338 100.79957	100.07745 100.82387

Table 4: Mean Square error of proposed estimators Neutrosophic regression estimators using OLS, Huber M & generalized M

$n = 6$	$MSE\left(\bar{t}_{s(Rp+i)}\right)$	$MSE\left(\bar{t}_{p1(Rp+i)}\right)$	$MSE\left(\bar{t}_{p2(Rp+i)}\right)$	$MSE\left(\bar{t}_{p3(Rp+i)}\right)$	$MSE\left(\bar{t}_{p4(Rp+i)}\right)$
Jan	1.0474300 0.3558393	0.9765913 0.3325345	0.9657827 0.3308930	0.9608463 0.3292846	0.9470394 0.3264216
Feb	2.1220210 0.6343308	2.0529380 0.6200605	2.0278760 0.6186074	2.0241450 0.6174217	2.0204580 0.6144462
March	2.7339360 0.6624271	2.4748780 0.6380706	2.4189700 0.6339706	2.4055680 0.6309223	2.3935850 0.6285497
April	2.0749850 1.0388010	1.7585660 0.7996068	1.7117550 0.7711520	1.6935880 0.7618891	1.6824860 0.7527371
May	0.4873254 0.6315386	0.4859832 0.6301380	0.4858691 0.6300007	0.4855490 0.6299334	0.4853539 0.6293693
June	1.5019650 0.6594707	1.4459220 0.6521603	1.4273750 0.6516465	1.4420130 0.6507827	1.4403760 0.6495161
July	0.5035114 0.3345773	0.4118036 0.2561982	0.4081703 0.2517847	0.4059160 0.2453365	0.4015423 0.2423980
August	0.4889136 0.1834280	0.3467680 0.1319495	0.3397331 0.1280211	0.3362652 0.1267409	0.3305587 0.1242247
Sep	0.6437115 0.5100549	0.6371995 0.5099832	0.6356551 0.5099675	0.6347725 0.5099390	0.6335105 0.5099261
Oct	1.0877120 1.0252950	1.0361680 0.8436684	1.0332960 0.8350016	1.0296700 0.8290923	1.0265100 0.8206879
Nov	0.6157228 0.3202029	0.5072972 0.3156490	0.5044234 0.3128089	0.5007071 0.3112062	0.4974167 0.3088296
Dec	1.5985040 0.3323765	1.5914980 0.3233948	1.5909950 0.3220678	1.5905070 0.3211111	1.5890230 0.3197380

Table 5: Percent relative efficiency of proposed Neutrosophic regression estimators using generalized M with the estimators using OLS and Huber M

Month	RE of OLS v/s Huber	RE of OLS v/s Mallows GM	RE of OLS v/s Schweppes GM	RE of OLS v/s SIS GM
Jan	107.25367 107.00824	108.45400 107.53908	109.01119 108.06436	110.60047 109.01218
Feb	103.36508 102.30144	104.64254 102.54174	104.83542 102.73866	105.02673 103.23618
March	110.46751 103.81721	113.02067 104.48862	113.65033 104.99345	114.21930 105.38977
April	117.99301 129.91398	121.21974 134.70768	122.52006 136.34543	123.32852 138.00316
May	100.27618 100.22227	100.29973 100.24411	100.36585 100.25482	100.40620 100.34468
June	103.87594 101.12095	105.22568 101.20068	104.15752 101.33501	104.27590 101.53262
July	122.26979 130.59315	123.35817 132.88230	124.04325 136.37486	125.39436 138.02808
August	140.99156 139.01379	143.91109 143.27951	145.39524 144.72676	147.90523 147.65824
Sep	101.02197 100.01406	101.26742 100.01714	101.40822 100.02273	101.61023 100.02526
Oct	104.97448 121.52820	105.26625 122.78959	105.63695 123.66476	105.96214 124.93117
Nov	121.37319 101.44271	122.06468 102.36374	122.97065 102.89091	123.78410 103.68271
Dec	100.44021 102.77732	100.47197 103.20079	100.50280 103.50826	100.59666 103.95277

To assess the robustness of the estimators, bias and mean square error (MSE) were computed for each scenario. The results highlight the superior robustness of the Generalized M (GM) estimators in the presence of outliers and heavy-tailed distributions with Schweppes-GM and SIS-GM in particular, exhibiting nearly constant MSE values even as contamination levels increase is reported in Table 6.

Furthermore, sensitivity analysis with respect to the correlation coefficient ($\rho = 0.3$–0.9) confirmed theoretical expectations, indicating that estimator efficiency increases monotonically with stronger association between study and auxiliary variables. These results further demonstrate the adaptability of GM-type estimators to varying data structures and quality levels.

Table 6: Simulation study results

Contamination level (%)	OLS MSE	Huber M	Mallows GM	Schweppes GM	SIS GM
0	0.01034	0.00982	0.00961	0.00953	0.00941
5	0.01295	0.01012	0.00984	0.00962	0.00953
10	0.01613	0.01124	0.01005	0.00984	0.00964
20	0.02895	0.01566	0.01186	0.01095	0.01045

Although formal derivations of influence functions and breakdown points have been known to be well defined in the GM-type estimators in the robust regression literature, this work measures the robustness in terms of contamination-based sensitivity analysis and stability of MSE as the level of contamination increases. The numerical stability observed up to 20 percent contamination is in line with the bounded-influence and high-breakdown of GM estimators, which is of practical application as robustness evidence in the neutrosophic setting.

8 Convergence Analysis of the Proposed GM-Type Neutrosophic Estimators

Algorithmic Framework: The proposed GM-type neutrosophic regression and ratio estimator is calculated by an iteratively reweighted least squares (IRLS) algorithm.

Let $\psi^{(t)} = \left(\psi_0^{(t)}, \psi_1^{(t)}\right)'$ denote the regression parameter set in the t iteration. The update in each iteration is defined as $\psi^{(t+1)} = \arg\min_{\psi} \sum_{i \subset s} \omega_i^t \rho \left(\frac{t_i - \psi_0 - \psi_1 v_i}{\hat{\sigma}}\right)$, where $\rho(.)$ is the loss function of Huber-M, Mallows-GM, Schweppe-GM and SIS-GM estimators, $\hat{\sigma}$ is a robust scale estimator and ω_i^t are iteration-dependent weights that consider both magnitude and leverage effects of the residual values.

The updated weight is in the form of

$$\omega_i^t = \frac{\varphi\left(r_i^t\right)}{r_i^{(t)}} . \omega\left(h_i\right)$$

where $r_i^{(t)}$ is the standardized residuals, $\varphi(.)$ is the influence function and $\omega(h_i)$ is a leverage-based weight peculiar to GM-type estimators.

The iterative process is set off with the OLS estimation that provides a consistent starting point with finite population sampling.

Convergence Theoretical Reasoning: Convergence to the estimates of the proposed estimation algorithm is classical in its results of M and GM-estimation in convex loss functions. Specifically:

1. **Monotonicity:** As each iteration is designed to minimize the non-increasing objective function: $Q\left(\psi^{(t+1)}\right) \leq Q\left(\psi^{(t)}\right)$,
   where $Q(.)$ denotes the weighted robust loss. This makes it numerical stability.
2. **Bounded Influence:** In case of GM-type estimators, the influence function $\varphi(.)$ is bounded. This means that big residuals will not destabilize the update sequence, which ensures contamination resistance.

3. **Convexity**: The loss function associated with Huber-M and GM-type estimators are convex (or locally convex) in the neighborhood of the optimum. Therefore, the IRLS algorithm converges to a stationary point of the objective function.

The iterative estimation process of the proposed GM-type estimators reached a convergence in a short period (4 to 6 iterations) at a given (10^{-5}) tolerance parameter update threshold. Convergence plots help to verify numerical stability as well as decreasing trend of error with increasing iterations. The proposed methods are suitable to large-scale engineering datasets and can be used in data-driven systems and real-time monitoring applications with their low computational cost and rapid convergence because of their approximate computational complexity. Convergence of the GM estimators with the iterations is reported in Table 7. The estimator's convergence behaviour and the relative efficiency at the various contamination levels are given by Figs. 3 and 4, respectively.

Table 7: Convergence of GM estimators

Iteration	OLS	Huber M	Mallows GM	Schweppes GM	SIS GM
1	0.02456	0.02034	0.01913	0.01873	0.01843
2	0.02075	0.01654	0.01527	0.01471	0.01432
3	0.01746	0.01383	0.01278	0.01234	0.01204
4	0.01562	0.01226	0.01135	0.01096	0.01065
5	0.01483	0.01138	0.01058	0.01017	0.00986
6	0.01467	0.01119	0.01039	0.01008	0.00977

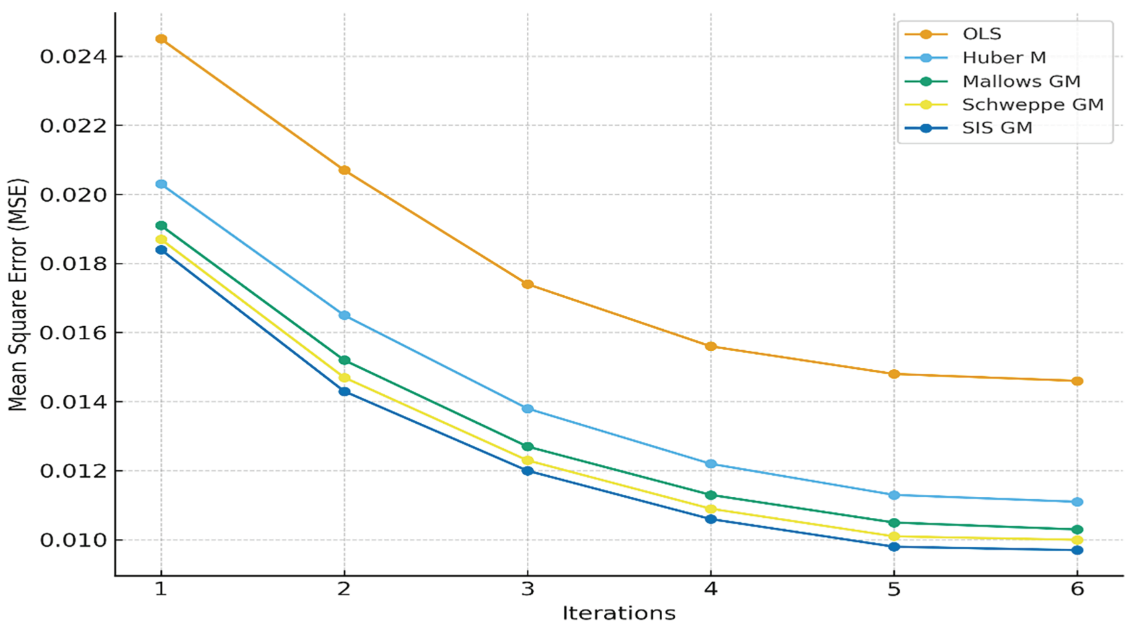

Figure 3: Convergence behavior of the estimators

In this study the convergence analysis is primarily algorithmic and empirical in nature. Formal convergence rates and closed –form guarantees for IRLS-based generalized M-estimators under neutrosophic sampling are difficult to establish due to the presence of interval-valued observations and indeterminacy. Nevertheless, convergence is supported under standard sufficient conditions from robust estimation theory, including bounded influence functions, local convexity of the loss function and monotonic decrease of the objective function across iterations. Empirical convergence results consistently demonstrate rapid stabilization within a small number of iterations, indicating numerical reliability for practical implementation.

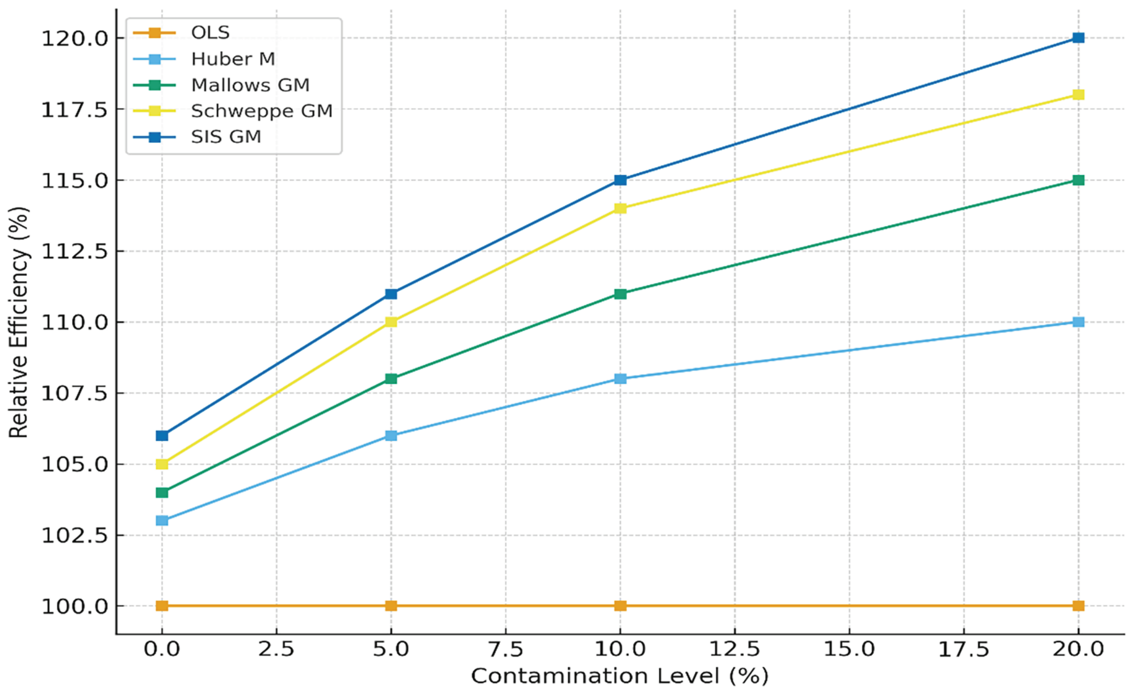

Figure 4: Relative efficiency of the estimators at different contamination levels

9 Results & Discussion

Our recommendation in this work is a new type of neutrosophic robust ratio estimators to estimate the population mean by three regression estimation techniques OLS (Ordinary Least Square), Huber-M and generalized M (GM) estimators. The proposed estimators use the Hodges-Lehmann estimator as a supporting information to increase the accuracy in an array of data. The empirical evidence suggests that Huber M-estimator has always been superior to the classical OLS estimator because it has an adaptive weighting mechanism which minimizes the effect of extreme values yet remains efficient even in the presence of contamination. Therefore, the Huber M-estimator is an efficient even in the presence of contamination. Therefore, the Huber M-estimator is an efficient alternative of OLS in estimating ratios. In addition, the neutrosophic estimators that use GM estimators namely Mallows, Schweppes and SIS estimators are more robust and efficient than the OLS estimators and the Huber M-estimators. These results are consistent with those provided by [15] who also accentuated the benefit of GM-type estimators in polluted data environment. The performance metrics of higher quality of GM-based procedures result in the mean squared error (MSE) and relative efficiency (RE) in Tables 2 and 3, respectively.

Also, five regression-type estimators were suggested where OLS, Huber M, Mallows GM, Schweppes GM and SIS GM were extended analysis estimators. Comparative analysis in similar circumstances proved the preceding observation: GM-type estimators greatly exceed the performance of OLS and also both OLS and Huber M-estimator are significantly less effective than GM-type estimators in terms of their ability to estimate a population mean estimation when using complex-data conditions and robust methods. Tables 4 and 5 show GM-type estimators are much more effective and powerful in various data conditions and thus should be employed with adaptive sampling to accurately estimate a population mean estimation using complex-data conditions. Generally, the suggested neutrosophic and regression type estimators have much more benefits to robustness and accuracy in that they are too applicable in real world where data is uncertain or contaminated.

10 Conclusion

The study presents a single structure of Neutrosophic robust ratio estimators, which involve the integration of OLS, Huber M, and generalized M-type regression models with the Hodges-Lehmann auxiliary information. Theoretical derivations and empirical studies have repeatedly shown that OLS is exposed to contamination and that Huber M and GM-type estimators are stronger and more efficient. Specifically, Mallows-GM, Schweppes-GM, and SIS-GM models are found to have consistent performance and lower mean squared error in the presence of indeterminacy and outlier contamination. Simulation and sensitivity analysis are further provided to support the numerical stability and convergence behavior. The suggested framework generalizes the robust finite-population estimation in the neutrosophic paradigm and offers a convenient inference tool in data settings that can be characterized by uncertainty and error of measurement. The current research concentrates on Huber-M and generalized M-type estimators and does not specifically address better high-breakdown estimators like MM-, S-, and τ-estimators. It is also known that these estimators provide greater strength when faced with high contamination and complicated leverage design. Consequently, the findings of this paper are to be viewed in the light of moderate levels of contamination and normal data patterns. Although the present results already show obvious robustness benefits compared to classical methods, another set of research including the use of MM-, S-, and τ-estimators would be helpful in order to define whether more benefits in efficiency or robustness might be obtained in more serious conditions of the data. The extension is hence recommended as a future research direction.

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

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

Availability of Data and Materials: The data that support the findings of this study are openly available in Custom Weather at https://www.timeanddate.com/weather/pakistan/lahore/historic.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zadeh LA. Fuzzy sets. Inf Control. 1965;8(3):338–53. doi:10.1016/S0019-9958(65)90241-X.
2. Smarandache F. Neutrosophy: neutrosophic probability, set, and logic: analytic synthesis & synthetic analysis. Rehoboth, NM, USA: American Research Press; 1998.

3. Bouza-Herrera CN, Subzar M. Estimating the ratio of a crisp variable and a neutrosophic variable. Int J Neutrosophic Sci. 2020;10(2):95–104. doi:10.54216/ijns.0110101.
4. Vishwakarma GK, Singh A. Generalized estimator for computation of population mean under neutrosophic ranked set technique: an application to solar energy data. Comput Appl Math. 2022;41(4):144. doi:10.1007/s40314-022-01820-7.
5. Alomair AM, Shahzad U. Neutrosophic mean estimation of sensitive and non-sensitive variables with robust Hartley-Ross-type estimators. Axioms. 2023;12(6):578. doi:10.3390/axioms12060578.
6. Singh A, Aslam M, Vishwakarma GK, Dhital A, Patrascu I. Neutrosophic regression cum ratio estimators for the population mean: an application in medical science. In: Smarandache F, Aslam M, editors. Cognitive intelligence with neutrosophic statistics in bioinformatics. Cambridge, MA, USA: Academic Press; 2023. p. 313–33. doi:10.1016/B978-0-323-99456-9.00018-0.
7. Yadav SK, Smarandache F. Generalized neutrosophic sampling strategy for elevated estimation of population mean [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://ssrn.com/abstract=4652035.
8. Singh R, Kumari A, Smarandache F, Tiwari SN. Construction of almost unbiased estimator for population mean using neutrosophic information. Neutrosophic Sets Syst. 2025;76:449–63. doi:10.61356/j.nois.2025.7571.
9. 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.
10. Singh P, Gupta S. Combining two auxiliary variables for elevated estimation of finite population mean under neutrosophic framework. Neutrosophic Sets Syst. 2025;76:275–87. doi:10.61356/j.nswa.2025.25443.
11. Kumar A, Kumar V. Estimating population mean of climate data via neutrosophic imputation in simple random sampling. Qual Quant. 2025;2025:1–30. doi:10.1007/s11135-025-02503-1.
12. Rehman A, Ahmed SH, Ahmad I. Optimized nonlinear robust controller along with model-parameter estimation for blood glucose regulation in type-1 diabetes. ISA Trans. 2025;160:163–74. doi:10.1016/j.isatra.2025.02.030.
13. Rehman A, Ahmed SH, Ghias R, Ahmad I. Reinforcement learning based sliding mode control for optimal chemotherapy drug in cancerous tumor. Biomed Signal Process Control. 2025;103:107485. doi:10.1016/j.bspc.2024.107485.
14. Huber PJ. Robust estimation of a location parameter. Ann Math Stat. 1964;35(1):73–101. doi:10.1214/aoms/1177703732.
15. Subzar M, Al-Omari AI, Alanzi ARA. Robust regression methods for estimating finite population mean based on SRSWOR in the presence of outliers. Comput Mater Contin. 2020;65(1):125–38. doi:10.32604/cmc.2020.010230.
16. CustomWeather. Historic weather data for Lahore, Pakistan [Internet]. 2020 [cited 2025 Jan 1]. Available from: https://www.timeanddate.com/weather/pakistan/lahore/historic.