2024

Computation of Separate Ratio and Regression Estimator Under Neutrosophic Stratified Sampling: An Application to Climate Data

Abhishek Singh

Hemant Kulkarni

Florentin Smarandache
University of New Mexico, smarand@unm.edu

Gajendra K. Vishwakarma

Computation of Separate Ratio and Regression Estimator Under Neutrosophic Stratified Sampling: An Application to Climate Data

Abhishek Singh[1, *], Hemant Kulkarni[1], Florentin Smarandache[2], Gajendra K. Vishwakarma[3]

[1]Department of Mathematics & Statistics,
Dr. Vishwanath Karad MIT World Peace University, Pune-411038
E-mail: abhisuryak@gmail.com, hemant.statistics@gmail.com

[2]Mathematics, Physical and Natural Science Division, University of New Mexico,
Gallup, 705 Gurley Ave., Gallup, NM 87301, USA
E-mail: smarand@unm.edu

[3]Department of Mathematics & Computing,
Indian Institute of Technology (ISM) Dhanbad, Dhanbad-826004, Jharkhand, India
E-mail: vishwagk@rediffmail.com,

Abstract

In this article, we introduce a novel approach by presenting separate ratio and regression estimators in the context of neutrosophic stratified sampling for the very first time, incorporating auxiliary variables. We have conducted a thorough analysis to estimate these newly proposed estimators' bias and mean square error (MSE) up to the first-order approximation. Theoretically using efficiency comparison criteria, our findings demonstrate the superior performance of these estimators compared to traditional unbiased estimators. Also, numerically based on real-life and artificial data, we have shown the supremacy of the neutrosophic stratified sampling over neutrosophic simple random sampling along with the supremacy of our proposed neutrosophic separate stratified estimators over neutrosophic stratified unbiased estimator. Moreover, our research highlights the enhanced reliability of neutrosophic stratified estimators when contrasted with classical stratified estimators.

Keywords: Neutrosophic variables, Neutrosophic stratified sampling, Regression and ratio estimator, Monte-Carlo simulation, Mean square error.

Mathematics Subject Classification: 62D05, 62A86.

1. Introduction

The researchers' primary aim is to pinpoint the most efficient methods for estimating population parameters within sample surveys. This endeavour considers various factors like time, cost management, and the reduction of sampling errors. Achieving these objectives relies on leveraging supplementary information to refine better estimators and minimize sampling errors. This additional data typically comprises secondary traits (auxiliary information) highly correlated with the study variables, usually accessible for each investigated unit. If not readily available, it can often be sourced from preceding surveys or historical records. For instance, to predict the students' current exam scores or performance, the previous year's scores or academic performance can be used as auxiliary information, and to predict current exam scores or performance, they can be considered study characteristics. In another example, to analyze the temperature trend of the current year, past year climate records can be used as auxiliary information, and current year temperature trends can be considered study characteristics. In essence, study and subsidiary characteristics are defined according to the specific context and research objectives to improve the accuracy of population parameter estimations while minimizing sampling errors. Cochran significantly contributed to the advancement of modern sampling theory by introducing the groundbreaking idea of integrating supplementary information. His pioneering work has profoundly impacted the sampling field, leading to the development of two renowned estimation methods. Specifically, Cochran's ratio estimator, presented in [1], and Murthy's product estimator in [2], stand out as well-established techniques for characterizing and harnessing supplementary information in the estimation process. These methods have significantly contributed to enhancing statistical sampling methodologies using auxiliary information.

The above concepts or studies are associated with simple random sampling under classical statistics, which traditionally deals with precise or single-valued data, assuming complete certainty in the data. However, classical statistics may prove inadequate when faced with vague or uncertain data. This uncertainty manifests when data comprises intervals or sets, and addressing such data requires a novel approach known as fuzzy statistics, grounded in fuzzy logic and set theory. The concept of fuzziness in data was first introduced by Zadeh [3] and further developed in [4] as the notion of a fuzzy set. Fuzzy statistics has undergone significant development, giving rise to various specialized branches. These include fuzzy regression analysis, fuzzy probability theory, fuzzy time series analysis and forecasting, estimation of confidence intervals from fuzzy data, fuzzy applications in operations research, hypothesis testing, and addressing issues related to fuzzy arrival/service rates. Also, numerous researchers extensively explored and applied fuzzy logic, which plays a crucial role in decision-making processes, especially in situations characterized by uncertainty or ambiguity, and it has some extensions, such as complex fuzzy logic and intuitionistic fuzzy logic.

Our study specifically focuses on neutrosophic logic and statistics, particularly within the domain of neutrosophic statistics grounded in neutrosophic logic or sets. Neutrosophic logic and statistics come into play when fuzzy or intuitionistic statistics prove inadequate in quantifying the inherent indeterminacy present in fuzzy or vague data. While fuzzy statistics offers a means to handle variable values subject to uncertainty, it lacks the capability to quantify the extent of indeterminacy associated with the data precisely. On the other hand, neutrosophic statistics address this limitation by providing a framework to handle and measure indeterminacy independently in the data that exhibits uncertainty or fuzziness.

The issue of neglecting indeterminacy in fuzzy statistics finds its solution in neutrosophic statistics. This broader framework not only extends fuzzy statistics or classical statistics but also enables the quantification of indeterminacy within uncertain or hazy data. Smarandache initially introduced the concept of neutrosophic sets, logic, and statistics and is further explored in various literature sources, including Smarandache's works in [5], [6], [7], [8], [9], [10], [11], [12]. Neutrosophic statistics stands as a broadening of classical statistics, fuzzy statistics, and intuitionistic statistics (a concept originally conceived by Atanassov [13] and further developed in [14]). The field has witnessed significant research endeavours over the years. In [11], Smarandache introduced the foundation of neutrosophic statistics, laying the groundwork for its application. Subsequent studies have explored various facets of neutrosophic statistics, including the representation of rock joint roughness coefficients using neutrosophic interval statistical numbers by [15], [16]. Furthermore, [17] delved into the realm of neutrosophic probability statistics, and several authors, such as [18], [19], [20], [21], [22], [23], have researched robust single-valued neutrosophic soft aggregation operators in the context of multi-criteria decision making. Additionally, studies have explored classifying trapezoidal bipolar neutrosophic numbers, de-bi-polarization techniques, and their application in solving multiple-criteria group decision-making problems based on cloud services, as presented by [24]. Indeterminacy in data often signifies the presence of neutrosophic data, which can be effectively managed through neutrosophic statistics due to its straightforward methods for handling such indeterminate data. Various research articles have applied neutrosophic data in different contexts, including the analysis of neutrosophic numbers for rock joint roughness coefficients by [15], [16], and the examination of wind speed data [25], [21], [22], [23], [26]. Some more novelty about the neutrosophic concept can be seen through [27], [28], [29], [30], [31], [32], [33], [45] [46], [47], [48], [49].

As neutrosophic statistics replace classical statistics, neutrosophic simple random sampling (NeSRS) replaces classical simple random sampling. A recent and first study utilizing the neutrosophic simple random sampling is given by [34] as neutrosophic ratio-type estimators for population mean. Readers can refer to the article by [34] for a more comprehensive understanding of neutrosophic data. Further, [35], [36], [37] [38], [39], [40] have given different estimation procedures in neutrosophic simple random sampling and neutrosophic ranked set sampling for estimating neutrosophic parameters. All these studies are done under NeSRS for the homogeneous neutrosophic data or population and not for heterogeneous. As if the population is heterogeneous, it is better to use stratified sampling than simple random sampling in which we divide the whole population into some homogeneous strata for better efficacy of the estimators used in estimating population parameters, see [41], [42], [43]. So, inspired by the studies done so far under NeSRS, and like classical stratified sampling, we are giving neutrosophic stratified sampling and estimators under it.

The presentation of this manuscript starts with the introduction section, then followed by the research gap and contribution in section 2; proposed methods of estimation in sections 3, 3.1, and 3.2; efficiency comparisons, simulation study, and empirical study in sections 4, 4.1, and 4.2, respectively. A discussion is given in section 5, followed by concluding remarks with future studies in section 6.

2. Research gap and contribution

In the realm of our sample survey, our exploration has been dedicated solely to well-defined data points that are crisp-type data. We've leaned heavily on conventional sampling techniques to extract precise outcomes using crisp-type data. However, navigating classical methods for undefined values or indeterminacy-type data will pose significant challenges, and dealing with such types of data will be difficult. There is a variety of such data types, such as experimental and equipment-related, water levels, melting points, humidity, wind speed, and day-wise temperature data, that routinely manifest as range values or tend towards ambiguity. The prevalence of interval data in real-world scenarios surpasses the presence of clearly defined or crisp-type data. Procuring information on these ambiguous data types incurs substantial expenses.

Consequently, applying traditional (classical) methods for computing precise values of unknown parameters becomes a high-cost, high-risk endeavour when confronted with indeterminate or vague data. So, the neutrosophic estimation method takes over the old classical method for the neutrosophic (vague or indeterminate) data types. Notably, [34] addressed this gap in sampling theory by introducing neutrosophic ratio-type estimators for population means in simple random sampling. Inspired by their innovative work and aiming to bridge the gap between traditional and neutrosophic statistical approaches, further, [35], [36], [37] [38], [39], [40] have given different estimation procedures in neutrosophic simple random sampling and neutrosophic ranked set sampling for estimating neutrosophic parameters.

All these studies are done under NeSRS for the homogeneous neutrosophic data or population only and not for heterogeneous. As if the population is heterogeneous, it is better to use stratified sampling than simple random sampling, in which we divide the whole population into some homogeneous strata for better efficacy (precision) of the estimators used in estimating population parameters. When the population has distinct subgroups with varying characteristics, stratified sampling allows you to capture these differences more accurately. For instance, if you're studying a city's population with different income brackets and know income influences spending habits, stratified sampling ensures each income group is represented proportionally in the sample. This increases the precision of estimates within each subgroup compared to simple random sampling, where certain groups might be underrepresented or overrepresented by chance. Still, overall precision will be increased compared to simple random sampling. In another example, suppose a household survey is conducted in different regions according to language. If a Hindi-speaking surveyor visits a Tamil-speaking region, then both the surveyor and respondent will definitely face difficulty, which will lead to non-sampling errors, and consequently will cost time and money. This problem can be removed by stratification, that is, by applying stratified sampling. In which, we divide the whole population into homogeneous strata in language-wise regions and apply the language-wise surveyor in the respective regions for separate estimation. The combined results obtained according to stratified sampling will be better (in terms of precision, time, and money) than a simple random method. For more detail, see [41], [42], [43].

So, inspired by the studies done so far under NeSRS, and like classical stratified sampling, we estimate the population mean under neutrosophic stratified sampling for the first time. For example, a company conducts a job satisfaction survey among employees, separately stratified into office-based and remote workers, with a proportional allocation based on workforce composition. Upon review, it's found that a few employees have hybrid work arrangements, blurring the distinction between the two strata. To accommodate this ambiguity, the survey shifts to stratified random neutrosophic sampling. If we use classical statistics or the classical stratified random method here, it will cause information loss. In addressing this issue, we employ neutrosophic statistics to obtain interval-type results while minimizing the Mean Squared Errors (MSEs). In this manuscript, we first give neutrosophic stratified random sampling, then ratio and regression estimators under it. Neutrosophic random sampling can be useful in medical research or diagnostics; there might be cases where symptoms or test results are ambiguous or contradictory, leading to uncertainty in classification. Neutrosophic stratified sampling could be applied to ensure that samples are representative across various subsets of patients with uncertain or overlapping symptoms for more accurate diagnostic models or treatment assessments. It can be useful in dealing with risk assessment in various domains such as finance, environmental studies, or cybersecurity; there could be scenarios where the available data is imprecise or conflicting. Neutrosophic stratified sampling might assist in creating samples that account for this uncertainty, aiding in better risk evaluation and decision-making processes. Also, it can be useful in market research, consumer behaviour, opinion polls, social experiments, and so on in many fields where it can be useful.

The neutrosophic observations can be represented in various formats, with the neutrosophic numbers potentially encompassing an unknown interval [a, b]. Here, we are illustrating neutrosophic values as $Z_N = Z_L + Z_U I_N$ with $I_N \in \left[I_L, I_U\right]$, the symbol 'N' is used to represent a neutrosophic number. Consequently, our neutrosophic observations will fall within an interval $Z_N \in [a,b]$, where 'a' and 'b' denote the lower and upper values of the neutrosophic data.

3. Proposed neutrosophic stratified sampling and estimators under it

Neutrosophic stratified sampling is a statistical sampling technique that combines stratified sampling and neutrosophic set theory elements. It is used to collect data from a population to ensure that different subgroups or strata within the population are represented adequately while also accounting for uncertainty, vagueness, and indeterminacy in the data. For examples:

(i) Two distinct strata, consisting of men and women, are considered within the city of Gallup, New Mexico. However, in light of the population's demographic breakdown, with women constituting 51% and men 49%, a random sample consisting of 51 women and 49 men is initially drawn. Subsequently, it is discovered that one individual identified as a man and two individuals identified as women are, in fact, transgender. This revelation introduces an element of indeterminacy involving three individuals. As a result, the sampling method shifts to stratified random neutrosophic sampling, acknowledging the inherent uncertainty in the gender identities of these individuals and reflecting the complexities of the population composition more accurately.

(ii) A university surveys student satisfaction with two strata: undergraduate and graduate students. Given that 70% are undergraduates and 30% are graduate students, an initial sample comprises 70 undergraduates and 30 graduate students. Later, it is discovered that some students are pursuing dual degrees, making their educational status indeterminate. Therefore, the survey adopts stratified random neutrosophic sampling to account for these ambiguous cases.

(iii) A market research study stratifies respondents into two groups: those who prefer Product A and Product B, with an equal allocation of 50 respondents to each group. After collecting data, it becomes apparent that a few respondents are uncertain or neutral in their preferences. Consequently, the study shifts to stratified random neutrosophic sampling to accommodate this indeterminacy.

(iv) A healthcare provider assesses service utilization among residents in District X and District Y, with 60% of respondents from District X and 40% from District Y. Subsequently, it's revealed that some respondents have dual residency, making it unclear which district they primarily belong to. The study adopts stratified random neutrosophic sampling to address this uncertainty.

These examples demonstrate how the transition from conventional stratified sampling to neutrosophic stratified random sampling can occur when dealing with situations where certain individuals or cases exhibit indeterminate characteristics or affiliations, requiring a more nuanced sampling approach.

3.1 Proposed neutrosophic stratified sampling method

As a classical stratified sampling method, this sampling method also consists of dividing the whole neutrosophic heterogeneous population into neutrosophic homogeneous subgroups/subpopulations such that units within each subgroup are homogeneous and between subgroups/subpopulations are heterogeneous concerning characteristics under study or study variables. Such subgroups/subpopulations are known as strata, and each subgroup is a stratum. Then, we apply the neutrosophic simple random sampling method to each stratum to obtain neutrosophic separate stratified samples. Let $N_N = N_L + N_U I_N, I_N \in [I_L, I_U]$ be the finite neutrosophic heterogeneous population and divided into non-overlapping neutrosophic homogeneous L strata of each size of $N_{hN} = N_{hL} + N_{hU} I_{hN}, I_{hN} \in [I_{hL}, I_{hU}]$ such that $\sum_{h=1}^{L} N_{hN} = N_N$, $h = 1, 2, ..., L.$ Then, we apply neutrosophic simple random sampling method to each stratum and draw separate stratified samples $n_{hN} = n_{hL} + n_{hU} I_{hN}, I_{hN} \in [I_{hL}, I_{hU}]$ from each stratum population of size N_{hN} such that we have total separate stratified sample $n_N = n_L + n_U I_N, I_N \in [I_L, I_U]$, $n_N = \sum_{h=1}^{L} n_{hN}$, $h = 1, 2, ..., L,$ see in below Fig. 1 placed in this section. Also, let Y_N be the neutrosophic study characteristics and Y_{hjN} be the population value of the study character Y_N and y_{hjN} be the sample value of the j^{th} unit $(j = 1, 2, ..., N_h)$ in the h^{th} stratum. Then,

Population means of h^{th} stratum = $\bar{Y}_{hN} = \bar{Y}_{hL} + \bar{Y}_{hU} I_{hN} = \dfrac{1}{N_{hN}} \sum_{j=1}^{N_{hN}} Y_{hjN}$, $I_{hN} \in [I_{hL}, I_{hU}]$.

Population mean = $\bar{Y}_N = \bar{Y}_N + \bar{Y}_N I_N = \dfrac{1}{N_{hN}} \sum_{h=1}^{L} \sum_{j=1}^{N_{hN}} Y_{hjN}, I_N \in [I_L, I_U]$

$$= \sum_{h=1}^{L} w_{hN} \bar{Y}_{hN},\ w_{hN} = \frac{N_{hN}}{N_N} = w_{hL} + w_{hU} I_{hN}.$$

Population mean square of h^{th} stratum = $S_{hN}^2 = S_{hL}^2 + S_{hU}^2 I_{hN} = \dfrac{1}{N_{hN} - 1} \sum_{j=1}^{N_{hN}} \left(Y_{hjN} - \bar{Y}_{hN}\right)^2.$

Population mean square = $S_N^2 = S_L^2 + S_U^2 I_N = \frac{1}{N_N - 1}\sum_{h=1}^{L}\sum_{j=1}^{N_N}\left(Y_{hjN} - \bar{Y}_N\right)^2, I_N \in [I_L, I_U]$.

Sample mean of h^{th} stratum = $\bar{y}_{hN} = \bar{y}_{hL} + \bar{y}_{hU} I_{hN} = \frac{1}{n_{hN}}\sum_{j=1}^{n_{hN}} y_{hjN}, I_{hN} \in [I_{hL}, I_{hU}]$.

Sample mean = $\bar{y}_{stN} = \bar{y}_{stL} + \bar{y}_{stU} I_N = \sum_{h=1}^{L} w_{hN}\bar{y}_{hN}, I_N \in [I_L, I_U]$ where $w_{hN} = \frac{N_{hN}}{N_N}$.

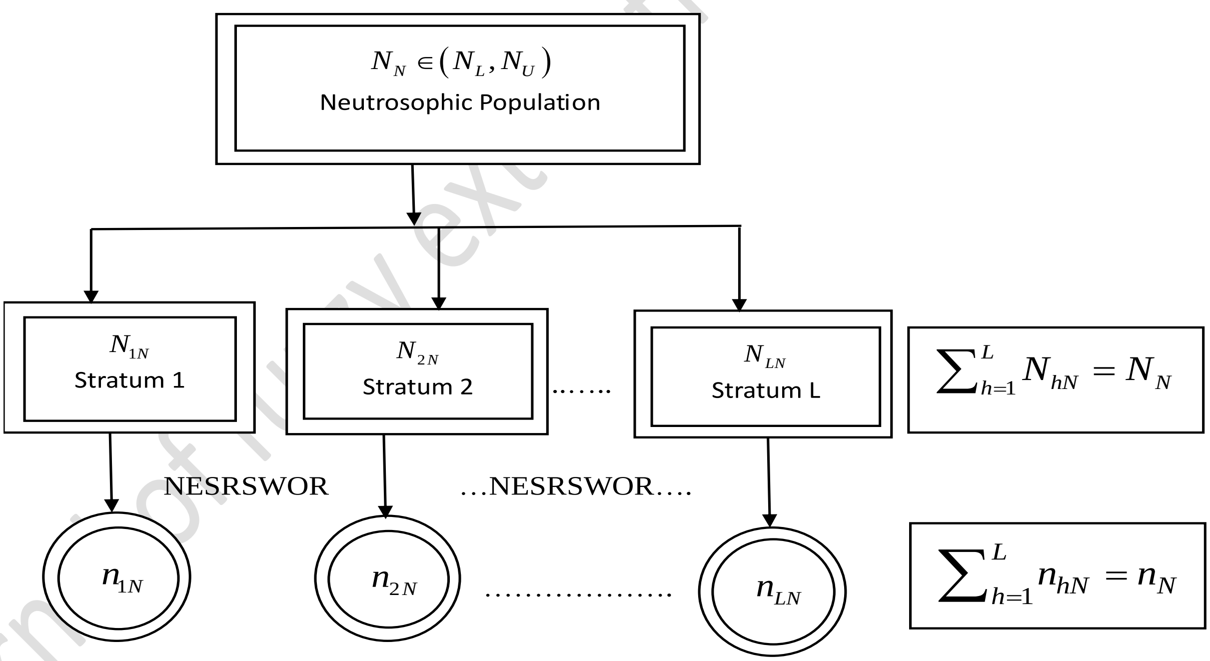
Like in classical stratified sampling, neutrosophic unbiased estimator and its variance is given by as:

$$\bar{y}_{stN} = \sum_{h=1}^{L} w_{hN}\bar{y}_{hN}, w_{hN} = w_{hL} + w_{hU} I_{hN}, I_{hN} \in [I_{hL}, I_{hU}] \tag{1}$$

$$V\left(\bar{y}_{stN}\right) = \sum_{h=1}^{L} w_{hN}^2 V\left(\bar{y}_{hN}\right) = \sum_{h=1}^{L} w_{hN}^2\left(\frac{1}{n_{hN}} - \frac{1}{N_{hN}}\right) S_{hN}^2 \tag{2}$$

Below here in Fig. 1, we are given a neutrosophic stratified sampling structure:

Fig 1:



NESRSWOR: Neutrosophic Simple Random Sampling without Replacement

Every neutrosophic value in this section is represented like this $Z_N = [Z_L + Z_U I_N]$ with $I_N \in [I_L, I_U]$ and stratified neutrosophic value is represented as $Z_{hN} = [Z_{hL} + Z_{hU} I_{hN}]$ with

$I_{hN} \in [I_{hL}, I_{hU}]$. The symbol 'N' used to represent a neutrosophic number and 'h' for each stratum. Consequently, our neutrosophic value will fall within an interval that is $Z_N \in [a, b]$ and stratified neutrosophic value as $Z_{hN} \in [a_h, b_h]$.

3.2 Proposed separate ratio and regression estimator under neutrosophic stratified sampling

To estimate the neutrosophic separate stratified population mean using neutrosophic auxiliary characteristics. We examine a finite, diverse population with a total size of N_N, and we partition it into L homogeneous and non-overlapping strata, with each stratum having a size of $N_{hN}, h = 1, 2, ..., L, \sum_{h=1}^{L} N_{hN} = N_N$. Let (y_{hjN}, x_{hjN}) represent the paired values associated with the study character y_N and the auxiliary character x_N, respectively of the j^{th} unit $(j = 1, 2, ..., N_{hN})$ in the h^{th} stratum. Also, let (y_{hjN}, x_{hjN}) be the pair of values on (y, x) drawn from the h^{th} stratum $(j = 1, 2, ..., n_{hN}; h = 1, 2, ..., L)$. Let $\bar{y}_{hN} = \frac{1}{n_{hN}} \sum_{j=1}^{n_{hN}} y_{hjN} \in [\bar{y}_{hL}, \bar{y}_{hU}]$ and $\bar{x}_{hN} = \frac{1}{n_{hN}} \sum_{j=1}^{n_{hN}} x_{hjN} \in [\bar{x}_{hL}, \bar{x}_{hU}]$ be the stratum sample means, $\bar{y}_{stN} = \sum_{h=1}^{L} w_{hN} \bar{y}_{n_{hN}} \in [\bar{y}_{stL}, \bar{y}_{stU}]$ and $\bar{x}_{stN} = \sum_{h=1}^{L} w_{hN} \bar{x}_{n_{hN}} \in [\bar{x}_{stL}, \bar{x}_{stU}]$ be the total sample means of the strata for the neutrosophic study and ancillary variables are represented, respectively. Also, $\bar{Y}_{hN} = \frac{1}{N_{hN}} \sum_{j=1}^{N_{hN}} Y_{hjN} \in [\bar{Y}_{hL}, \bar{Y}_{hU}]$ and $\bar{X}_{hN} = \frac{1}{N_{hN}} \sum_{j=1}^{N_{hN}} X_{hjN} \in [\bar{X}_{hL}, \bar{X}_{hU}]$ be the stratum population means, $\bar{Y}_N = \sum_{h=1}^{L} w_{hN} \bar{Y}_{hN} \in [\bar{Y}_L, \bar{Y}_U]$ and $\bar{X}_N = \sum_{h=1}^{L} w_{hN} \bar{X}_{hN}$ $\in [\bar{X}_L, \bar{X}_U]$ be the population means of the neutrosophic study variable $Y_N \in [Y_L, Y_U]$ and ancillary $X_N \in [X_L, X_U]$ characteristics, respectively. The stratum correlation coefficient between both neutrosophic study and subsidiary variables is $\rho_{yxhN} \in [\rho_{yxhL}, \rho_{yxhU}]$, $C_{xhN} \in [C_{xhL}, C_{xhU}]$ and $C_{yhN} \in [C_{yhL}, C_{yhU}]$ be the stratum coefficient of variation of neutrosophic variables Y_N and X_N. The parameter $\beta_{2(x)hN} \in [\beta_{2(x)hL}, \beta_{2(x)hU}]$ is the neutrosophic stratum coefficient of kurtosis of a neutrosophic subsidiary variable X_N. Let the neutrosophic stratum mean error terms are $e_{h0N} \in [e_{h0L}, e_{h0U}]$ and $e_{h1N} \in [e_{h1L}, e_{h1U}]$, where $e_{h0N} = (\bar{y}_{hN} - \bar{Y}_{hN}) / \bar{Y}_{hN}$ and $e_{h1N} = (\bar{x}_{hN} - \bar{X}_{hN}) / \bar{X}_{hN}$ are such

that $E\left(e_{h1N}\right)=0,$ $E\left(e_{h1N}\right)=0,$ $E\left(e_{h0N}^{2}\right)=f_{hN}C_{yyhN},$ $E\left(e_{h1N}^{2}\right)=f_{hN}C_{xxhN}$ and $E\left(e_{h0N}e_{h1N}\right)=f_{hN}C_{yxhN},$
where,
$f_{hN}=\left(\frac{1}{n_{hN}}-\frac{1}{N_{hN}}\right), S_{xxhN}\in\left[S_{xxhL},S_{xxhU}\right]; S_{yyhN}\in\left[S_{yyhL},S_{yyhU}\right]; S_{xyhN}\in\left[S_{xyhL},S_{xyhU}\right].\ f_{hN}\in\left[f_{hL},f_{hU}\right]$ $; e_{h0N}^{2}\in\left[e_{h0L}^{2},e_{h0U}^{2}\right];\ e_{h1N}^{2}\in\left[e_{h1L}^{2},e_{h1U}^{2}\right];\ e_{h0N}e_{h1N}\in\left[e_{h0L}e_{h1L},e_{h0U}e_{h1U}\right];\ C_{yyhN}\in\left[C_{yyhL},C_{yyhU}\right];$ $C_{xxhN}\in\left[C_{xxhL},C_{xxhU}\right];\ C_{yxhN}\in\left[C_{yxhL},C_{yxhU}\right]=\rho_{yxhN}C_{yhN}C_{xhN};\ S_{xhN}\in\left[S_{xhL},S_{xhU}\right];$ $S_{yhN}\in\left[S_{yhL},S_{yhU}\right];\ \rho_{yxhN}\in\left[\rho_{yxhL},\rho_{yxhU}\right].$

To ascertain population means amidst uncertainty, we define established separate stratified ratios and regression estimators into neutrosophic separate stratified estimators. So, inspired by [1], [34], [35], [43]. We propose a neutrosophic separate stratified ratio and regression estimator for population means by utilizing neutrosophic subsidiary information. The neutrosophic separate stratified ratio and regression estimator are given, respectively

$$\bar{y}_{RhN}=\bar{y}_{hN}\left(\frac{\bar{X}_{hN}}{\bar{x}_{hN}}\right)=\left(\bar{y}_{hL}\left(\frac{\bar{X}_{hL}}{\bar{x}_{hL}}\right)\right)+\left(\bar{y}_{hU}\left(\frac{\bar{X}_{hU}}{\bar{x}_{hU}}\right)\right)I_{hN}, I_{hN}\in\left[I_{hL},I_{hU}\right] \tag{3}$$

$$\begin{aligned}\bar{y}_{reghN}=\bar{y}_{hN}+b_{hN}\left(\bar{X}_{hN}-\bar{x}_{hN}\right)&=\left(\bar{y}_{hL}+b_{hL}\left(\bar{X}_{hL}-\bar{x}_{hL}\right)\right)\\&\quad+\left(\bar{y}_{hU}+b_{hU}\left(\bar{X}_{hU}-\bar{x}_{hU}\right)\right)I_{hN}, I_{hN}\in\left[I_{hL},I_{hU}\right]\end{aligned} \tag{4}$$

So, total neutrosophic separate stratified ratio and regression estimators are given as

$$\bar{y}_{RN}=\sum_{h=1}^{L}w_{hN}\bar{y}_{RhN}=\sum_{h=1}^{L}w_{hN}\left(\bar{y}_{hN}\frac{\bar{X}_{hN}}{\bar{x}_{hN}}\right) \tag{5}$$

$$\bar{y}_{regN}=\sum_{h=1}^{L}w_{hN}\bar{y}_{reghN}=\sum_{h=1}^{L}w_{hN}\left(\bar{y}_{hN}+b_{hN}\left(\bar{X}_{hN}-\bar{x}_{hN}\right)\right) \tag{6}$$

To calculate the *Bias* and *MSE*, we extend the neutrosophic separate stratified ratio estimator from Eq. (5) by breaking it down into neutrosophic separate stratified mean error terms as follows:

$$\begin{aligned}\bar{y}_{RN}&=\sum_{h=1}^{L}w_{hN}\bar{y}_{RhN}=\sum_{h=1}^{L}w_{hN}\bar{Y}_{hN}\left(1+e_{h0N}\right)\left(\frac{\bar{X}_{hN}}{\bar{X}_{hN}\left(1+e_{h1N}\right)}\right);\\&\approx\sum_{h=1}^{L}w_{hN}\bar{Y}_{hN}\left(1+e_{h0N}\right)\left[1+e_{h1N}\right]^{-1}\end{aligned}$$

$$\bar{y}_{RN} \approx \sum_{h=1}^{L} w_{hN}\bar{Y}_{hN}\left(1+e_{h0N}-e_{h1N}+e_{h1N}^{2}-e_{h1N}e_{h0N}\right)$$

$$\left(\bar{y}_{RN}-\bar{Y}_{N}\right) \approx \sum_{h=1}^{L} w_{hN}\bar{Y}_{hN}\left(e_{h0N}-e_{h1N}+e_{h1N}^{2}-e_{h1N}e_{h0N}\right) \tag{7}$$

By taking expectations on both sides of Eq. (7), we get the *Bias* of neutrosophic separate stratified ratio estimator as

$$Bias\left(\bar{y}_{RN}\right)=\sum_{h=1}^{L} w_{hN}\bar{Y}_{hN} f_{hN}\left(C_{xxhN}-C_{yxhN}\right) \tag{8}$$

By squaring and taking expectations on both sides of Eq. (7), with a degree no greater than 2, we determine the Mean Squared Error (*MSE*) of the neutrosophic separate stratified ratio estimator as:

$$MSE\left(\bar{y}_{RN}\right)=\sum_{h=1}^{L} w_{hN}^{2}\bar{Y}_{hN}^{2} f_{hN}\left(C_{yyhN}+C_{xxhN}-2C_{yxhN}\right) \tag{9}$$

To obtain *Bias* and *MSE*, we opened the neutrosophic separate stratified estimator from Eq. (6) by breaking it down into the first-order approximation of neutrosophic separate stratified mean error terms as:

$$\bar{y}_{regN}=\sum_{h=1}^{L} w_{hN}\bar{y}_{reghN}=\sum_{h=1}^{L} w_{hN}\left(\bar{Y}_{hN}\left(1+e_{h0N}\right)+b_{hN}\left(\bar{X}_{hN}-\bar{X}_{hN}\left(1+e_{h1N}\right)\right)\right);$$

$$\approx \sum_{h=1}^{L} w_{hN}\bar{Y}_{hN}\left(1+e_{h0N}\right)-b_{hN}\bar{X}_{hN}e_{h1N}$$

$$\left(\bar{y}_{regN}-\bar{Y}_{N_N}\right) \approx \sum_{h=1}^{L} w_{hN}\left(\bar{Y}_{hN}e_{h0N}-b_{hN}\bar{X}_{hN}e_{h1N}\right) \tag{10}$$

By taking expectations on both sides of Eq. (10), we get the *Bias* of neutrosophic separate stratified regression estimator as

$$Bias\left(\bar{y}_{regN}\right)=0 \tag{11}$$

By squaring and taking expectations on both sides of Eq. (10), with a degree no greater than 2, we determine the Mean Squared Error (*MSE*) of the neutrosophic separate stratified ratio estimator as:

$$MSE\left(\bar{y}_{regN}\right)=\sum_{h=1}^{L} w_{hN}^{2} f_{hN}\left(\bar{Y}_{hN}^{2}C_{yyhN}+b_{hN}^{2}\bar{X}_{hN}^{2}C_{xxhN}-2b_{hN}\bar{Y}_{hN}\bar{X}_{hN}C_{yxhN}\right) \tag{12}$$

where the constant term b_{hN} is such that minimized the *MSE* of $\bar{y}_{regN}$ and for this, we differentiate the *MSE* in Eq. (12) with respect to b_{hN} and by equating it to 0, we get

$$\frac{\partial\left(MSE\left(\bar{y}_{regN}\right)\right)}{\partial b_{hN}}=0 \quad \Rightarrow \quad b_{hN}=\frac{\bar{Y}_{hN}}{\bar{X}_{hN}}\left(\frac{C_{YXhN}}{C_{xxhN}}\right) \tag{13}$$

Replacing b_{hN} from Eq. (13) into Eq. (12), we get the *MSE* of $\bar{y}_{regN}$ as

$$MSE\left(\bar{y}_{regN}\right)=\sum_{h=1}^{L} w_{hN}^{2} f_{hN} \bar{Y}_{hN}^{2} C_{yyhN}\left(1-\rho_{YXhN}^{2}\right) \tag{14}$$

Now, again, to show *Bias* and *MSE*, we opened the neutrosophic stratified unbiased estimator from Eq. (1) in neutrosophic separate stratified mean error terms as

$$\bar{y}_{stN}=\sum_{h=1}^{L} w_{hN} \bar{y}_{hN}=\sum_{h=1}^{L} w_{hN} \bar{Y}_{hN}\left(1+e_{h0N}\right) \tag{15}$$

To the 1st order approximation, the Bias and MSE of the neutrosophic separate stratified unbiased estimator are given respectively by Eq. (16) and (17)

$$Bias\left(\bar{y}_{stN}\right)=0 \tag{16}$$

$$MSE\left(\bar{y}_{stN}\right)=\sum_{h=1}^{L} w_{hN}^{2} f_{hN} \bar{Y}_{hN}^{2} C_{yyhN} \tag{17}$$

Further, for the comparison with NeSRS and by [34], we are giving an unbiased estimator under NeSRS corresponding to a neutrosophic separate stratified unbiased estimator as

$$\bar{y}_{N}=\frac{1}{n_{N}} \sum_{i=1}^{n_{N}} y_{iN} \tag{18}$$

To the 1st order approximation, the Bias and MSE of the neutrosophic unbiased estimator under NeSRS are given respectively by Eq. (19) and (20)

$$Bias\left(\bar{y}_{N}\right)=0 \tag{19}$$

$$MSE\left(\bar{y}_{N}\right)=Var\left(\bar{y}_{N}\right)=f_{N} S_{N}^{2},\ f_{N}=\left(\frac{1}{n_{N}}-\frac{1}{N_{N}}\right),\ S_{N}^{2}=\frac{1}{N_{N}-1} \sum_{h=1}^{L} \sum_{j=1}^{N_{N}}\left(Y_{hjN}-\bar{Y}_{N}\right)^{2}. \tag{20}$$

$$\bar{y}_{RN} \in\left[\bar{y}_{RL}, \bar{y}_{RU}\right];\ \bar{y}_{regN} \in\left[\bar{y}_{regL}, \bar{y}_{regU}\right];\ \bar{y}_{stN} \in\left[\bar{y}_{stL}, \bar{y}_{stU}\right];\ \bar{y}_{N} \in\left[\bar{y}_{L}, \bar{y}_{U}\right];\ f_{N} \in\left[f_{L}, f_{U}\right]$$

$$Bias\left(\bar{y}_{RN}\right) \in\left[Bias\left(\bar{y}_{RL}\right), Bias\left(\bar{y}_{RU}\right)\right];\ MSE\left(\bar{y}_{RN}\right) \in\left[MSE\left(\bar{y}_{RL}\right), MSE\left(\bar{y}_{RU}\right)\right];$$

$$Bias\left(\bar{y}_{regN}\right) \in\left[Bias\left(\bar{y}_{regL}\right), Bias\left(\bar{y}_{regU}\right)\right];\ MSE\left(\bar{y}_{regN}\right) \in\left[MSE\left(\bar{y}_{regL}\right), MSE\left(\bar{y}_{regU}\right)\right];$$

$$Bias\left(\bar{y}_{stN}\right) \in\left[Bias\left(\bar{y}_{stL}\right), Bias\left(\bar{y}_{stU}\right)\right];\ MSE\left(\bar{y}_{stN}\right) \in\left[MSE\left(\bar{y}_{stL}\right), MSE\left(\bar{y}_{stU}\right)\right];$$

$$Bias\left(\bar{y}_{N}\right) \in\left[Bias\left(\bar{y}_{L}\right), Bias\left(\bar{y}_{U}\right)\right];\ MSE\left(\bar{y}_{N}\right) \in\left[MSE\left(\bar{y}_{L}\right), MSE\left(\bar{y}_{U}\right)\right];\ \ b_{N} \in\left[b_{L}, b_{U}\right].$$

Every neutrosophic value in this section is represented like this $Z_{N}=\left[Z_{L}+Z_{U} I_{N}\right]$ with $I_{N} \in\left[I_{L}, I_{U}\right]$ and stratified neutrosophic value is represented as $Z_{hN}=\left[Z_{hL}+Z_{hU} I_{hN}\right]$ with $I_{hN} \in\left[I_{hL}, I_{hU}\right]$. The symbol 'N' used to represent a neutrosophic number and 'h' for each stratum. Consequently, our neutrosophic value will fall within an interval that is $Z_{N} \in[a, b]$ and stratified neutrosophic value as $Z_{hN} \in\left[a_{h}, b_{h}\right]$.

4. Efficiency Comparisons

Efficiency comparisons of the neutrosophic separate stratified ratio and regression estimators over neutrosophic separate stratified unbiased estimator

(i) $\bar{y}_{RN}$ will be efficient to $\bar{y}_{stN}$, iff

$MSE(\bar{y}_{RN}) < MSE(\bar{y}_{stN})$ i.e. $\left(MSE(\bar{y}_{stN}) - MSE(\bar{y}_{RN})\right) > 0$

from Eq. (9) and (17), we have

$$\left[\sum_{h=1}^{L} w_{hN}\bar{Y}_{hN}^{2} f_{hN} - \sum_{h=1}^{L} w_{hN}\bar{Y}_{hN}^{2} f_{hN}\left(C_{yyhN} + C_{xxhN} - 2C_{yxhN}\right)\right] > 0 \text{ i.e.}$$

$$C_{yxhN} > \frac{C_{xxhN}}{2} \tag{18}$$

(ii) The estimator $\bar{y}_{regN}$ will be efficient to $\bar{y}_{rsswe}$, iff

$MSE(\bar{y}_{regN}) < MSE(\bar{y}_{stN})$ i.e. $\left(MSE(\bar{y}_{stN}) - MSE(\bar{y}_{regN})\right) > 0$

from Eq. (20) and (46), we have

$$\left[\sum_{h=1}^{L} w_{hN} f_{hN}\bar{Y}_{hN}^{2} C_{yyhN} - \sum_{h=1}^{L} w_{hN} f_{hN}\bar{Y}_{hN}^{2} C_{yyhN}\left(1-\rho_{YXhN}^{2}\right)\right] > 0 \text{ i.e.}$$

$$\sum_{h=1}^{L} w_{hN} f_{hN}\bar{Y}_{hN}^{2} C_{yyhN}\rho_{YXhN}^{2} > 0 \tag{19}$$

which is always true.

4.1 Empirical study

To numerically elucidate the characteristics of neutrosophic stratified estimators, we have compiled real-life indeterminate climate data of the USA state. We have taken two states, Alabama and Georgia, as strata, and then November month is taken for the data from both states. There are many variables, but we are considering Dew Point Temperature vs Relative Humidity variables only here. The Dew Point Temperature variable is taken as the auxiliary variable $X_{hN} = \in\left[X_{hL}, X_{hU}\right]$, and the Relative Humidity variable as the study variable $Y_{hN} \in\left[Y_{hL}, Y_{hU}\right]$. Along with indeterminate data, classical data is also taken by averaging lower and upper values of the indeterminate data. The parameter descriptions are given in Table 1 below, and the data is available in Tables 6 and 7 of Appendix A. Also, one can visit for the data on this link: https://mrcc.purdue.edu.

Table 1: Description of real data parameters for estimating means in the context of neutrosophic stratified random sampling.

Parameters	Neutrosophic	Classical	Parameters	Neutrosophic	Classical
		1st Stratum			
N_{1N}	[19, 19]	19	S_{y1N}	[13.55, 23.30]	12.12
n_{1N}	[6, 6]	6	C_{x1N}	[0.2295, 0.5549]	0.5549
$\bar{X}_{1N}$	[19.58, 61.95]	19.58	C_{y1N}	[0.1405, 0.8261]	0.1944
$\bar{Y}_{1N}$	[28.21, 96.47]	62.34	$\beta_{2(x)1N}$	[6.4058, 6.0892]	6.0892
S_{x1N}	[10.86, 14.22]	10.86	ρ_{yx1N}	[0.946, 0.941]	0.889
		2nd Stratum			
N_{2N}	[22, 22]	22	S_{y2N}	[11.78, 21.88]	9.457
n_{2N}	[7, 7]	7	C_{x2N}	[0.1393, 0.5794]	0.1622
$\bar{X}_{2N}$	[22.55, 62.23]	42.39	C_{y2N}	[0.1255, 0.6887]	0.1506
$\bar{Y}_{2N}$	[31.77, 93.86]	62.82	$\beta_{2(x)2N}$	[2.5435, 8.5923]	3.2652
S_{x2N}	[8.67, 13.06]	6.873	ρ_{yx2N}	[0.8854, 0.9481]	0.8588

Further, we have taken $n_{1N} = 6$ samples from the 1st strata of size 19 and $n_{2N} = 7$ from the strata of size 22 for both neutrosophic and classical using simple random form the both strata. The calculated results of the *MSEs* and *REs* obtained using the proposed neutrosophic separate stratified estimators along with classical separate stratified estimators are presented in Tables 5 and 6.

4.2 Monte-Carlo Simulation

Again, to numerically elucidate the characteristics of neutrosophic estimators, we conducted a Monte-Carlo simulation utilizing the methodology proposed by [35], [44]. This simulation was specifically conducted within the framework of neutrosophic analysis. Our population is separately stratified into two distinct categories, denoted as h=1 and h=2. It is important to note that within this context, neutrosophic stratified random variables (NSRV) naturally adhere to a neutrosophic separate stratified normal distribution (NSND), i.e.,

$\left(X_{hN}, Y_{hN}\right) \sim NN\left[\left(\mu_{xhN}, \sigma_{xhN}^2\right), \left(\mu_{yhN}, \sigma_{yhN}^2\right)\right]$, $X_{hN} \in \left[X_{hL}, X_{hU}\right]$; $Y_{hN} \in \left[Y_{hL}, Y_{hU}\right]$; $\mu_{xhN} \in \left[\mu_{xhL}, \mu_{xhU}\right]$; $\mu_{yhN} \in \left[\mu_{yhL}, \mu_{yhU}\right]$; $\sigma_{xhN}^2 \in \left[\sigma_{xhL}^2, \sigma_{xhU}^2\right]$; $\sigma_{xhN}^2 \in \left[\sigma_{xhL}^2, \sigma_{xhU}^2\right]$. The neutrosophic data is generated from a 4-variate multivariate stratified normal distribution with means $\left(\mu_{xhL},\ \mu_{yhL}, \mu_{xhU},\ \mu_{yhU}\right)$ and covariance matrix

$$\begin{pmatrix} \sigma_{xhL}^2 & \rho_{yxhL}\sigma_{xhL}\sigma_{yhL} & 0 & 0 \\ \rho_{yxhL}\sigma_{xhL}\sigma_{yhL} & \sigma_{yhL}^2 & 0 & 0 \\ 0 & 0 & \sigma_{xhU}^2 & \rho_{yxhU}\sigma_{xhU}\sigma_{xhU} \\ 0 & 0 & \rho_{yxhU}\sigma_{xhU}\sigma_{yhU} & \sigma_{yhU}^2 \end{pmatrix}.$$

The essential parameters for simulating neutrosophic separate stratified data are detailed in the following Table 2.

Table 2: Description of simulated data parameters for estimating means in the context of neutrosophic stratified random sampling.

Parameters	Neutrosophic	Classical	Parameters	Neutrosophic	Classical
			1st Stratum		
N_{1N}	[80, 80]	80	S_{y1N}	[9.34, 12.72]	9.88
n_{1N}	[25, 25]	25	C_{x1N}	[0.2501, 0.2652]	0.2343
$\bar{X}_{1N}$	[39.16, 50.35]	45.46	C_{y1N}	[0.2502, 0.2386]	0.2163
$\bar{Y}_{1N}$	[39.13, 50.85]	45.70	$\beta_{2(x)1N}$	[2.5189, 3.2019]	2.7548
S_{x1N}	[9.79, 13.35]	10.65	ρ_{yx1N}	[(0.5, 0.7, 0.9), (0.5, 0.7, 0.9]	0.5, 0.7, 0.9
			2nd Stratum		
N_{2N}	[120, 120]	120	S_{y2N}	[4.61, 5.56]	5.82
n_{2N}	[40, 40]	40	C_{x2N}	[0.1811, 0.2362]	0.2341
$\bar{X}_{2N}$	[19.54, 30.31]	25.25	C_{y2N}	[0.1837, 0.2324]	0.2339
$\bar{Y}_{2N}$	[19.84, 30.28]	24.86	$\beta_{2(x)2N}$	[2.6657, 3.2219]	3.3548
S_{x2N}	[4.62, 5.48]	5.91	ρ_{yx2N}	[(0.5, 0.7, 0.9), (0.5, 0.7, 0.9]	0.5, 0.7, 0.9

We conducted a study in which we divided a total population of size 200 into two neutrosophic strata of sizes 80 and 120. Further, neutrosophic stratum samples of sizes 25 and 40, respectively, are drawn using the neutrosophic simple random sampling method. So, we have a total of 65 neutrosophic separate stratified samples, and we utilize these samples to calculate the mean square errors (*MSEs*) and relative efficiencies (*REs*) for the proposed neutrosophic ratio and regression estimators. This entire process of obtaining MSEs and REs for the separate neutrosophic stratified estimators via the neutrosophic stratified sampling method was repeated 7000 times. The same process is also done for classical stratified sampling along with neutrosophic stratified sampling. The results of the *MSEs* and *REs* obtained using the proposed neutrosophic separate stratified estimators along with classical separate stratified estimators are presented in Tables 3 and 4.

Table 3: *MSEs* of the proposed estimators under neutrosophic stratified sampling.

Estimators	*MSEs*					
	$\rho_{yx1N}=[.5,.5]$ $\rho_{yx2N}=[.9,.9]$	$\rho_{yx1}=.5$ $\rho_{yx2}=.9$	$\rho_{yx1N}=[.7,.7]$ $\rho_{yx2N}=[.9,.9]$	$\rho_{yx1}=.7$ $\rho_{yx2}=.9$	$\rho_{yx1N}=[.5,.5]$ $\rho_{yx2N}=[.7,.7]$	$\rho_{yx1}=.5$ $\rho_{yx2}=.7$
$\bar{y}_N$	[16.1, 19.8]	17.2	[17.2, 21.6]	20.6	[18.1, 22.7]	19.1
$\bar{y}_{stN}$	[1.04, 1.85]	1.46	[1.27, 2.30]	1.53	[1.13, 2.05]	1.41
T_{RN}	[0.84, 1.27]	1.19	[0.62, 0.71]	0.79	[1.06, 1.57]	1.36
T_{regN}	**[0.56, 0.93]**	**0.92**	**[0.563, 0.68]**	**0.73**	**[0.74, 1.25]**	**1.01**

Table 4: *REs* of the proposed estimators over neutrosophic stratified unbiased.

Estimators	*MSEs*					
	$\rho_{yx1N}=[.5,.5]$ $\rho_{yx2N}=[.9,.9]$	$\rho_{yx1}=.5$ $\rho_{yx2}=.9$	$\rho_{yx1N}=[.7,.7]$ $\rho_{yx2N}=[.9,.9]$	$\rho_{yx1}=.7$ $\rho_{yx2}=.9$	$\rho_{yx1N}=[.5,.5]$ $\rho_{yx2N}=[.7,.7]$	$\rho_{yx1}=.5$ $\rho_{yx2}=.7$
$\bar{y}_{stN}$	[1.00, 1.00]	1.00	[1.00, 1.00]	1.00	[1.00, 1.00]	1.00
T_{RN}	[1.24, 1.45]	1.23	[2.04, 3.23]	1.95	[1.07, 1.30]	1.03
T_{regN}	**[1.85, 1.97]**	**1.59**	**[2.26, 3.37]**	**2.10**	**[1.53, 1.64]**	**1.39**

Table 5: *MSEs* of the proposed estimators under neutrosophic stratified sampling.

E	*MSEs*

	Neutrosophic	Classical
$\bar{y}_N$	[8.26, 26.2]	5.94
$\bar{y}_{stN}$	[3.61, 11.02]	2.57
T_{RN}	[2.03, 2.07]	1.01
T_{regN}	**[0.58, 1.01]**	**0.86**

Table 6: *REs* of the proposed estimators over neutrosophic stratified unbiased.

Estimators	*MSEs*	
	Neutrosophic	Classical
$\bar{y}_{stN}$	[1.00, 1.00]	1.00
T_{RN}	[1.75, 5.42]	2.53
T_{regN}	**[3.58, 18.9]**	**2.98**

5. Results and Discussion

We have derived the mathematical expressions for the proposed neutrosophic separate stratified ratio and regression estimators, limited to the first-order approximation. To gain a deeper understanding of the properties of these neutrosophic separate stratified estimators, we conducted a real-life neutrosophic data-based empirical study and a Monte-Carlo simulation study using artificial neutrosophic data, specifically targeting the correlation coefficients [(0.5, 0.7, 0.9), (0.5, 0.7, 0.9)]. We then computed the Mean Squared Errors (MSEs) and Relative Efficiencies (REs) for the proposed neutrosophic separate stratified ratio and regression estimators. These numerical results are presented in Tables 3, 4, 5, and 6.

In Table 3, the *MSEs* of the proposed neutrosophic separate stratified ratio and regression estimators, neutrosophic stratified unbiased and neutrosophic unbiased estimators, along with classical separate stratified estimators, are given through Monte-Carlo simulation on artificial neutrosophic data for the combination values of the correlation coefficient. The highlighted bold font shows the least *MSE* of the proposed neutrosophic stratified estimators. We see the *MSEs* of the proposed neutrosophic separate regression estimator for different sets of correlation coefficients are least along with the classical stratified separate regression estimator but better than the classical separate regression estimator. Also, we see that the unbiased estimator under NeSRS is greater than the unbiased estimator under neutrosophic stratified sampling, proving that neutrosophic stratified is better than NeSRS.

Again, in Table 4, we see that the *REs* of the proposed separate regression estimator are higher than other estimators along with the classical separate regression estimator but better than the classical separate regression estimator.

Also, we see the *MSE* and *RE* of the neutrosophic stratified separate regression estimator, along with the classical stratified separate regression estimator, are the lowest and highest, respectively, for the correlation coefficient value $\rho_{yx1N} = [.7,.7], \rho_{yx2N} = [.9,.9]$. So, the neutrosophic stratified estimators are better than classical stratified estimators when the 'uncertain or vague' heterogeneous 'data or population' exist.

In Table 5, similar to Table 3, the *MSEs* of the proposed neutrosophic separate stratified ratio and regression estimators, neutrosophic stratified unbiased and neutrosophic unbiased estimators, along with classical separate stratified estimators are given through real-life neutrosophic climate data. The highlighted bold font shows the least *MSE* of the proposed neutrosophic stratified estimators. We see the *MSEs* of the proposed neutrosophic separate regression estimator are least along with the classical stratified separate regression estimator but better than the classical separate regression estimator. Also, we see that the unbiased estimator under NeSRS is greater than the unbiased estimator under neutrosophic stratified sampling, proving that neutrosophic stratified is better than NeSRS.

Again, in Table 6, similar to Table 4, we see that the *RE* of the proposed separate regression estimator is higher than other estimators along with the classical separate regression estimator but better than the classical separate regression estimator.

6. Conclusion

Motivated by the work of [34] and [35] and with the help of neutrosophic subsidiary variables in our research, we have introduced novel concepts in dealing with neutrosophic heterogeneous data or populations. Specifically, we have put forth the concept of neutrosophic stratified sampling and then separate ratio and regression estimators under this sampling method. The *Biases* and *MSEs* of the proposed estimators are derived up to the 1st-order approximation. The neutrosophic stratified sampling and proposed estimators are examined theoretically and numerically very well through efficiency criteria. Numerical comparisons show that the neutrosophic stratified sampling is better than neutrosophic simple random sampling, and the neutrosophic separate regression estimator is better than the neutrosophic separate ratio and unbiased estimators. Also, comparisons with classical estimators and neutrosophic estimators show that neutrosophic estimators or neutrosophic methods are better and more reliable than classical estimation methods when there is indeterminacy in data.

These findings open the door for further investigations in various directions, such as modifying estimation procedures, exploring alternative estimators, or experimenting with different sampling methodologies. The current study can also be done further using combined stratified estimators instead of separate ones. Motivated by the neutrosophic ranked set sampling method given in [35], a neutrosophic stratified ranked set sampling can be defined, and further neutrosophic stratified median ranked set, extreme ranked set, unbalanced ranked set, etc., can be defined.

Acknowledgements: Authors are heartily thankful to Editors and anonymous learned reviewers for their valuable comments, which have substantially improved the original manuscript to its present form.

Funding: The authors declare that there is no funding for this research.

Conflict of interest: The authors declare no conflict of interest.

References

[1] Cochran, G. (1940). Some properties of estimators based on sampling scheme with varying probabilities. *Aust. J. Stat*. *17*, 22–28.

[2] Murthy, M. N. (1964). Product method of estimation. *Sankhyā: The Indian Journal of Statistics, Series A*, 69-74. https://www.jstor.org/stable/25049308

[3] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8* (3), 338-353.

[4] Zadeh, L. A. (1996). Fuzzy sets. In Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh (pp. 394-432). https://doi.org/10.1142/2895

[5] Smarandache, F. (1998). Neutrosophy: neutrosophic probability, set, and logic: analytic synthesis & synthetic analysis. *ProQuest Information & Learning, Ann Arbor, MI, USA*, *105*, 118–123.

[6] Smarandache, F. (1999). A unifying field in Logics: Neutrosophic Logic. In Philosophy (pp. 1-141). *American Research Press*.

[7] Smarandache, F. (2001). A unifying field in logics: Neutrosophic logic, neutrosophic set, neutrosophic probability, and statistics. https://arxiv.org/abs/math/0101228

[8] Smarandache, F. (2005). Neutrosophic set a generalization of the intuitionistic fuzzy set. *International journal of pure and applied mathematics*, *24*(3), 287.

[9] Smarandache, F. (2010). Neutrosophic logic-a generalization of the intuitionistic fuzzy logic. Multispace & multi structure. *Neutrosophic transdisciplinarity* (100 collected papers of science), 4, 396. http://dx.doi.org/10.2139/ssrn.2721587

[10] Smarandache, F. (2013). Introduction to neutrosophic measure, neutrosophic integral, and neutrosophic probability. *Infinite Study*. https://www.google.co.in/books/edition/Introduction_to_Neutrosophic_Measure_Neu/NSjdBAAAQBAJ?hl=en&gbpv=0

[11] Smarandache, F. (2014). Introduction to neutrosophic statistics: *Infinite Study*. https://doi.org/10.48550/arXiv.1406.2000

[12] Smarandache, F. (2019). Neutrosophic set is a generalization of intuitionistic fuzzy set, inconsistent intuitionistic fuzzy set (picture fuzzy set, ternary fuzzy set), pythagorean fuzzy set, spherical fuzzy set, and q-rung orthopair fuzzy set, while neutrosophication is a generalization of regret theory, grey system theory, and three-ways decision (revisited). *Journal of New Theory*, *29*, 1-31.

[13] Atanassov K. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets Syst*, *20*(1), 87–96.

[14] Atanassov K. (1999). New operations defined over the intuitionistic fuzzy sets. *Fuzzy Sets Syst*, *61*(2):137– 142.

[15] Chen, J., Ye, J., & Du, S. (2017). Scale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics. *Symmetry*, *9*(10), 208.

[16] Chen, J., Ye, J., Du, S., & Yong, R. (2017) Expressions of rock joint roughness coefficient using neutrosophic interval statistical numbers. *Symmetry*, 9(7), 123.

[17] Alhabib, R., Ranna, M. M., Farah, H., & Salama, A. A. (2018). Some neutrosophic probability distributions. *Neutrosophic Sets and Systems*, *22*, 30-38.

[18] Chen, C. T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy sets and systems*, *114*(1), 1-9.

[19] Jana, C., & Pal, M. (2019). A robust single-valued neutrosophic soft aggregation operators in multi-criteria decision making. *Symmetry*, 11(1), 110.

[20] Nabeeh, N. A., Smarandache, F., Abdel-Basset, M., El-Ghareeb, H. A., & Aboelfetouh, A. (2019). An integrated neutrosophic-topsis approach and its application to personnel selection: A new trend in brain processing and analysis. *IEEE Access*, 7, 29734-29744.

[21] Aslam M (2019). Neutrosophic analysis of variance: application to university students. *Complex & Intelligent System, 5*, 403–407. https://doi.org/10.1007/s40747-019-0107-2

[22] Aslam, M., Arif, O. H., & Sherwani, R. A. K. (2020). New diagnosis test under the neutrosophic statistics: an application to diabetic patients. *BioMed Research International*, *2020*, 1-7. https://doi.org/10.1155/2020/2086185

[23] Aslam, M., & Algarni, A. (2020). Analyzing the Solar Energy Data Using a New Anderson-Darling Test under Indeterminacy. *International Journal of Photoenergy*, *2020*, 1-6. https://doi.org/10.1155/2020/6662389

[24] Chakraborty, A., Mondal, S. P., Alam, S., & Dey, A. (2021). Classification of trapezoidal bipolar neutrosophic number, de-bipolarization technique and its execution in cloud service-based MCGDM problem. *Complex & Intelligent Systems*, *7*(1), 145-162.

[25] Almaraashi, M. (2017). Short-term prediction of solar energy in Saudi Arabia using automated-design fuzzy logic systems. *PloS one*, *12*(8), e0182429. https://doi.org/10.1371/journal.pone.0182429

[26] Aslam, M. (2020). Monitoring the road traffic crashes using NEWMA chart and repetitive sampling. *International journal of injury control and safety promotion*, *28*(1), 39-45. https://doi.org/10.1080/17457300.2020.1835990

[27] Kumar, R., Edalatpanah, S. A., Jha, S., & Singh, R. (2019). A novel approach to solve gaussian valued neutrosophic shortest path problems. *International Journal of Engineering and Advanced Technology*, *8*(3), 347–353.

[28] Veeramani, C., Edalatpanah, S. A., & Sharanya, S. (2021). Solving the multiobjective fractional transportation problem through the neutrosophic goal programming approach. *Discrete Dynamics in Nature and Society*, *2021*, 1-17. https://doi.org/10.1155/2021/7308042

[29] Mohanta, K. K., & Sharanappa, D. S. (2023). Neutrosophic data envelopment analysis: a comprehensive review and current trends. *Optimality*, *1*(1), 10–22.

[30] Mohanta, K. K., & Toragay, O. (2023). Enhanced performance evaluation through neutrosophic data envelopment analysis leveraging pentagonal neutrosophic numbers. *J. Oper. Strateg Anal*, 1(2), 70-80.

[31] Hosseinzadeh, E., & Tayyebi, J. (2023). A compromise solution for the neutrosophic multi-objective linear programming problem and its application in transportation problem. *Journal of applied research on industrial engineering*, *10*(1), 1-10.

[32] Edalatpanah, S. A., Abdolmaleki, E., Khalifa, H. A. E. W., & Das, S. K. (2023). A novel computational method for neutrosophic uncertainty related quadratic fractional programming problems. *Neutrosophic Sets and Systems*, *58*(1), 38.

[33] Edalatpanah, S. A., Hassani, F. S., Smarandache, F., Sorourkhah, A., Pamucar, D., & Cui, B. (2024). A hybrid time series forecasting method based on neutrosophic logic with applications in financial issues. *Engineering Applications of Artificial Intelligence*, *129*, 107531. https://doi.org/10.1016/j.engappai.2023.107531

[34] Tahir, Z., Khan, H., Aslam, M., Shabbir, J., Mahmood, Y., & Smarandache, F. (2021). Neutrosophic ratio-type estimators for estimating the population mean. *Complex & Intelligent Systems*, *7*(6), 2991-3001.

[35] Vishwakarma, G. K., & Singh, A. (2022). Generalized estimator for computation of population mean under neutrosophic ranked set technique: An application to solar energy data. *Computational and Applied Mathematics*, *41*(4), 144. https://doi.org/10.1007/s40314-022-01820-7

[36] Singh, A., Aslam, M., Vishwakarma, G. K., Dhital, A., & Patrascu, I. (2023). Neutrosophic regression cum ratio estimators for the population mean: an application in medical science. *In Cognitive Intelligence with Neutrosophic Statistics in Bioinformatics* (pp. 313-333). Academic Press.

[37] Singh, R., Smarandache, F., & Mishra, R. (2023). Generalized robust-type neutrosophic ratio estimators of pharmaceutical daily stock prices. In *Cognitive Intelligence with Neutrosophic Statistics in Bioinformatics* (pp. 417-429). Academic Press.

[38] Alomair, A. M., & Shahzad, U. (2023). Neutrosophic Mean Estimation of Sensitive and Non-Sensitive Variables with Robust Hartley–Ross-Type Estimators. *Axioms*, *12*(6), 578. https://doi.org/10.3390/axioms12060578

[39] Yadav, S. K., & Smarandache, F. (2023). Generalized Neutrosophic Sampling Strategy for Elevated estimation of Population Mean. *Neutrosophic Sets and Systems*, *53*(1), 14. https://digitalrepository.unm.edu/nss_journal/vol53/iss1/14/

[40] Tahir, Z., Khan, H., Alamri, F. S., & Aslam, M. (2023). Neutrosophic ratio-type exponential estimators for estimation of population mean. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-25. https://content.iospress.com/articles/journal-of-intelligent-and-fuzzy-systems/ifs223539

[41] Iliyasu, R., & Etikan, I. (2021). Comparison of quota sampling and stratified random sampling. *Biom. Biostat. Int. J. Rev*, *10*(1), 24-27.

[42] Vishwakarma, G. K., Singh, A., & Singh, N. (2020). Calibration under measurement errors. *Journal of King Saud University-Science*, *32*(7), 2950-2961.

[43] Kadilar, C., & Cingi, H. (2003). Ratio estimators in stratified random sampling. *Biometrical Journal*, *45*, 218-225. https://doi.org/10.1002/bimj.200390007

[44] Vishwakarma, G., & Singh, A. (2022). Computing the effect of measurement errors on ranked set sampling estimators of the population mean. *Concurrency and Computation: Practice and Experience*, 34(27), e7333. https://doi.org/10.1002/cpe.7333

[45] Wang, H., Smarandache, F., Sunderraman, R., & Zhang, Y. Q. (2005) interval neutrosophic sets and

logic: theory and applications in computing: Theory and applications in computing (Vol. 5). *Infinite Study*. https://doi.org/10.48550/arXiv.cs/0505014
[46] Hosseinzadeh, E., & Tayyebi, J. (2023). A compromise solution for the neutrosophic multi-objective linear programming problem and its application in transportation problem. *Journal of applied research on industrial engineering*, *10*(1), 1-10.
[47] Zhang, K., Xie, Y., Noorkhah, S. A., Imeni, M., & Das, S. K. (2023). Neutrosophic management evaluation of insurance companies by a hybrid TODIM-BSC method: a case study in private insurance companies. *Management decision*, *61*(2), 363-381.
[48] Edalatpanah, S. A. (2019). A nonlinear approach for neutrosophic linear programming. *Journal of applied research on industrial engineering*, *6*(4), 367-373.
[49] Das, S. K., & Dash, J. K. (2020). Modified solution for neutrosophic linear programming problems with mixed constraints. Infinite Study. https://doi.org/10.22105/riej.2020.224198.1127

Appendix A

The data in Table 6 is taken as 1st strata, and the data in Table 7 is taken as 2nd strata.

Table 6: The climate data is taken for November month from the Alabama state of USA

Station Name	Min Dew Point Temp (F)	Max Dew Point Temp (F)	Min RH (%)	Max RH (%)
ANNISTON METRO AP	10	66	17	100
MOBILE DWTN AP	12	69	12	94
BIRMINGHAM AP	14	64	16	94
DECATUR PRYOR FLD	15	62	18	100
DOTHAN RGNL AP	17	69	20	100
ALABASTER SHELBY CO AP	16	66	18	100
EVERGREEN MIDDLETON FLD	18	68	22	100
MONTGOMERY AP	19	67	19	100
MOBILE RGNL AP	15	70	14	100
MUSCLE SHOALS RGNL AP	13	61	14	100
MAXWELL AFB	50	64	94	100
CAIRNS FLD FT RUCKER	48	64	82	100
CRAIG AFB	14	68	17	108
TUSCALOOSA MUNI AP	15	64	16	100
TROY MUNI AP	18	66	18	94
FAIRHOPE 3 NE	25	27	41	60
GADSDEN 19 N	17	20	52	63
HUNTSVILLE INTL AP JONES FIELD	16	62	18	100
ANNISTON METRO AP	20	80	28	120

Table 7: The climate data is taken for November month from Georgia state of USA

Station Name	Min Dew Point Temp (F)	Max Dew Point Temp (F)	Min RH (%)	Max RH (%)
ALBANY SW GA RGNL AP	19	68	22	100
AUGUSTA BUSH FLD AP	13	66	16	100
ATHENS BEN EPPS AP	11	64	16	100
ALMA BACON CO AP	21	67	21	100
ATLANTA HARTSFIELD INTL AP	14	64	22	94
COLUMBUS METRO AP	16	66	18	100
AUGUSTA DANIEL FLD AP	13	65	16	100
PEACHTREE CITY FALCON FLD	10	65	17	100
ATLANTA FULTON CO AP	14	64	22	100
GAINESVILLE GILMER AP	12	61	19	100
FT BENNING LAWSON FLD	23	66	28	100
MACON MIDDLE GA RGNL AP	16	67	19	100
MARIETTA DOBBINS AFB	36	47	59	87
ATLANTA PEACHTREE AP	10	64	20	100
ROME R B RUSSELL AP	13	63	19	94
SAVANNAH INTL AP	22	67	15	100
BRUNSWICK MALCOLM MCKINNON AP	47	56	59	73
VALDOSTA MOODY AFB	50	66	88	94
VALDOSTA RGNL AP	24	71	26	100
WARNER ROBINS' AFB	43	66	72	94
BRUNSWICK 23 S	46	55	67	74
NEWTON 8 W	23	31	38	55