

Hybrid Neutrosophic Regression ANOVA (HNRA): An Uncertainty-Aware Statistical Inference Framework Under Measurement Imprecision

Jaida Najihah Jamidin

Universiti Teknologi MARA (UiTM)

Zahari Md Rodzi

zahari@uitm.edu.my

Universiti Teknologi MARA (UiTM)

Sayang Mohd Deni

Universiti Teknologi MARA (UiTM)

Siti Meriam Zahari

Universiti Teknologi MARA (UiTM)

Article

Keywords: neutrosophic statistics, uncertainty quantification, interval-valued ANOVA, uncertainty-aware regression, fuzzy systems, biomedical uncertainty modelling

Posted Date: July 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10256271/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Hybrid Neutrosophic Regression ANOVA (HNRA): An Uncertainty-Aware Statistical Inference Framework Under Measurement Imprecision

Author 1¹, Author 2^{1*}, Author 3¹, Author 4¹

¹Affiliation details withheld for blind review.

Abstract

Regression-based analysis-of-variance (ANOVA) underpins much of quantitative biomedical inference, yet its classical formulation presumes predictor and response variables are observed without imprecision – an assumption routinely violated by instrument calibration bounds, biological variability, and sensor-reporting error. Interval, fuzzy, and Bayesian extensions each relax this assumption partially, but none provides a complete interval-valued analogue of the classical sums-of-squares decomposition together with a formally bounded $\mathbf{F}$ -statistic. This paper introduces Hybrid Neutrosophic Regression ANOVA (HNRA), which represents each observation as a neutrosophic interval and propagates this representation through ordinary-least-squares-based regression ANOVA using closed interval arithmetic. Four theoretical results are established: a decomposition theorem showing $\mathbf{SST}_N \approx \mathbf{SSR}_N + \mathbf{SSE}_N$ holds componentwise under interval arithmetic; a convergence theorem proving that HNRA reduces exactly to classical regression ANOVA as uncertainty vanishes; a monotonicity theorem establishing that all HNRA interval widths grow monotonically with measurement uncertainty; and a boundedness theorem proving strict positivity and finiteness of the interval-valued $\mathbf{F}$ -statistic, with a corollary guaranteeing uniform statistical significance whenever the interval lower bound exceeds the classical critical value. The framework is validated on a biomedical dataset of $n = 60$ adults relating muscle mass to age under $\pm 10\%$ response uncertainty and ± 0.5 -year predictor uncertainty, yielding $\mathbf{F}_N = [116.52, 260.02]$, an interval lying entirely above the classical critical value $F_{0.05,1,58}^* \approx 4.00$. A sensitivity analysis across five uncertainty levels (5%–25%) and a simulation study over 500 replications confirm that HNRA conclusions remain stable under increasing measurement imprecision, consistent with the proven theoretical properties. The framework requires no prior distributional assumptions, adds negligible computational overhead over classical OLS, and is

directly applicable to clinical laboratory science, wearable-device analytics, and epidemiological inference under documented measurement uncertainty.

Keywords: neutrosophic statistics, uncertainty quantification, interval-valued ANOVA, uncertainty-aware regression, fuzzy systems, biomedical uncertainty modelling

1 Introduction

Regression-based analysis of variance is among the most widely deployed inferential tools in quantitative biomedical research, underpinning clinical decision-support systems, epidemiological studies, and translational research pipelines. Its inferential validity rests on an assumption seldom satisfied in practice: that both the predictor variable X and the response variable Y are observed as exact, error-free quantities [1]. Real biomedical measurements are subject to multiple simultaneous sources of imprecision. Calibrated laboratory instruments carry manufacturer-specified uncertainty bounds [2]; wearable devices report physiological quantities – heart rate, blood oxygen saturation, body composition – with documented precision limits [3]; self-reported clinical variables such as age and dosage are subject to rounding and recall bias [4]; and biological quantities vary within physiological ranges that cannot be resolved to single-point values [5]. Regulatory guidance from the International Organization for Standardization (ISO 5725 series) and the Clinical and Laboratory Standards Institute (CLSI EP09-A3) explicitly requires that analytical studies account for documented measurement uncertainty [6, 7]. When such imprecision is non-negligible, classical regression ANOVA collapses an entire range of plausible inferential outcomes into a single-valued F -statistic, obscuring exactly the information that uncertainty-aware decision-making requires.

Several methodological families address this limitation, each with a distinct representational commitment. Classical regression treats X and Y as crisp, applying ordinary least squares (OLS) under the assumption that all variability is attributable to random error rather than structural measurement imprecision [8]; it offers point estimates, a single F -statistic, and well-understood asymptotic theory, but discards any information about known instrument tolerance. Interval regression, developed primarily within econometrics, represents the response – and sometimes the predictor – as a bounded interval and estimates parameters via linear-programming or set-identification methods [9]; this formally captures structural uncertainty but does not, in general, produce an ANOVA decomposition or an F -statistic in the conventional sense. Fuzzy regression, originating with Tanaka, Uejima, and Asai [10], encodes uncertainty through membership functions rather than crisp intervals; while flexible, it requires the analyst to specify a membership-function family and likewise lacks a standard ANOVA table. Bayesian regression represents uncertainty through prior distributions over parameters, updated via the likelihood to yield posterior distributions [11]; this is well established but requires prior specification, posterior computation, and produces credible intervals rather than the frequentist F -statistic that biomedical regulatory frameworks typically expect. Robust regression addresses outlier-contaminated

data through M-estimators [12], targeting heavy-tailed error distributions rather than the structural measurement imprecision targeted by interval-based approaches. Existing neutrosophic regression models – reviewed in detail in Section 2 – represent the most direct precursors to the present work, establishing neutrosophic correlation and simple regression [13], trimmed neutrosophic regression [14], and a neutrosophic test of linearity using an NANOVA-style table [15]; however, none of these establishes the complete regression ANOVA decomposition $SST_N = SSR_N + SSE_N$ together with interval-valued mean squares and a formally bounded F -statistic.

Table 1 summarises these six families along the dimensions most relevant to inferential use: how uncertainty is represented, how interpretable the result is, whether an ANOVA decomposition is available, whether a conventional F -statistic is produced, whether a prior is required, and computational complexity. Three limitations recur across all but the proposed framework: the absence of a complete interval-valued ANOVA decomposition mirroring the classical $SST = SSR + SSE$ identity, the absence of a frequentist F -statistic interpretable against the standard critical-value framework, and the absence of a formal convergence proof connecting the uncertainty-aware procedure to its classical counterpart as uncertainty vanishes – a property essential for establishing that a generalised framework is conservative rather than an ad hoc departure from established practice. HNRA is positioned to close exactly this set of gaps: it retains classical OLS coefficient estimation, preserving interpretability and computational simplicity; it requires no prior specification, unlike Bayesian regression, and no membership-function parameterisation, unlike fuzzy regression; and it produces a complete interval-valued ANOVA table with proven algebraic and asymptotic properties unavailable in any prior neutrosophic regression formulation.

Table 1 Representational and inferential capabilities of regression frameworks under measurement uncertainty. HNRA is the only framework offering a complete ANOVA decomposition with a bounded, interval-valued F -statistic requiring no prior specification.

Method	Uncertainty	Interp.	ANOVA	F -stat?	Prior?	Compl.
Classical OLS	Random error	High	Yes	Point	No	Low
Interval reg.	Interval Y	Mod.	Partial	Not std.	No	Mod. (LP)
Fuzzy reg.	Memb. fns	Mod.	No	Not std.	No	Moderate
Bayesian reg.	Prior+lik.	Mod.	No	Posterior	Yes	High
Robust reg.	Heavy tails	High	No	Asympt.	No	Moderate
NS reg. ¹	Interval	High	Partial	Limited	No	Low
HNRA	Interval	High	Full	Intv., bounded	No	Low

¹Aslam and Saleem [15] provide a neutrosophic F -test for linearity but do not derive the complete $SST_N = SSR_N + SSE_N$ decomposition, interval-valued mean squares, or formal boundedness/convergence proofs.

To the best of our knowledge, no existing study has simultaneously established interval-valued sums of squares satisfying $SST_N = SSR_N + SSE_N$ under closed-interval arithmetic, interval-valued mean squares within a complete ANOVA table, an interval-valued F -statistic with formally proved positivity and boundedness, and a convergence theorem connecting the neutrosophic and classical frameworks. The

present paper addresses this gap with four main contributions: the development of the complete HNRA mathematical framework, comprising the neutrosophic interval space $\mathcal{I}_N$, the HNRA regression model, interval-valued sums of squares, interval-valued mean squares, and the interval-valued F -statistic (Section 3); formal proof of four theoretical properties – decomposition, convergence, monotonicity, and boundedness – establishing HNRA as a conservative, well-behaved generalisation of classical regression ANOVA (Section 3.3); a worked biomedical application ($n = 60$; muscle mass and age) together with a five-level sensitivity analysis and a 500-replication simulation study empirically confirming the theoretical properties (Section 4); and a comparative evaluation positioning HNRA against classical, interval, fuzzy, Bayesian, and robust regression, together with a discussion of its theoretical, methodological, and biomedical implications and limitations (Section 5).

2 Related Work

2.1 Neutrosophic Statistics and Regression

Neutrosophic statistics, introduced by Smarandache [16] and formalised as a generalisation of interval statistics [17], represents each observation as a neutrosophic number $A_N = [A_L, A_U]$ encoding both the measured value and its indeterminacy range. Unlike fuzzy set theory, which models uncertainty through membership functions parameterised by a single value, neutrosophic numbers carry independent degrees of truth, indeterminacy, and falsity [16]; unlike Bayesian regression, neutrosophic regression retains deterministic parameter estimates while propagating structural measurement uncertainty through closed-interval arithmetic [17]. Atanassov’s intuitionistic fuzzy sets [18] contributed a closely related framework for two-valued uncertainty, which neutrosophic sets strictly subsume. Sherwani et al. [19] established the neutrosophic normal probability distribution underlying parametric neutrosophic hypothesis tests, directly supporting the distributional interpretation of the interval-valued F -statistic developed in Section 3. Woodall, Driscoll, and Montgomery [20] reviewed neutrosophic statistical process monitoring and identified the absence of formal convergence proofs connecting neutrosophic and classical procedures as an open methodological gap, addressed directly by the convergence result in Section 3.3.

Within the ANOVA literature, Aslam [21] introduced neutrosophic one-way ANOVA for comparing group means under indeterminacy, and AlAita and Aslam [22] extended this to neutrosophic analysis of covariance across three experimental designs. Miari, Anan, and Zeina [23] proposed neutrosophic two-way ANOVA for factor interactions under uncertain observations, and Kumar, Moazzamigodarzi, and Rahimi [24] developed neutrosophic Graeco-Latin square designs for complex factorial settings; Aslam [25] subsequently extended neutrosophic ANOVA to artificial-intelligence contexts. In regression specifically, Salama, Khaled, and Mahfouz [13] introduced neutrosophic correlation coefficients and a neutrosophic simple regression line, establishing the foundational interval-valued regression model later advanced by Aslam and Albassam [14], whose trimmed neutrosophic regression for prostate-cancer data demonstrated smaller error sums of squares under indeterminacy than classical regression. Vazquez and Torres [26] integrated neutrosophic numbers directly

into predictive regression modelling, and Aslam [27] developed a neutrosophic linear model for artificial-intelligence applications incorporating a neutrosophic model matrix. The closest predecessor to the present work is Aslam and Saleem’s neutrosophic test of linearity using an ANOVA-style table [15], which – as summarised in Table 1 – stops short of the full $SST_N = SSR_N + SSE_N$ decomposition, interval-valued mean-square derivation, and boundedness proofs developed here. Masood et al. [28] further showed that neutrosophic estimators consistently outperform classical counterparts under measurement uncertainty, reinforcing the case for interval-based regression inference generally. In applied biomedical statistics, Aslam [29] introduced a neutrosophic Spearman rank correlation test for clinical laboratory data, Sherwani et al. [30] extended relative risk and odds ratios to neutrosophic settings for epidemiological inference, and Aslam [31] demonstrated neutrosophic chi-square testing for pulse-count data, collectively confirming that neutrosophic methods are informative in exactly the biomedical contexts targeted by HNRA.

2.2 Neutrosophic Decision Support and Interval-Valued Modelling

A parallel body of work addresses neutrosophic and interval-valued uncertainty representation for multi-criteria decision-making (MCDM) and aggregation, several strands of which inform the interval-arithmetic foundations adopted in Section 3. The authors of [32] introduced the q -rung orthopair fuzzy-valued neutrosophic set, formalising axiomatic properties and aggregation operators for indeterminate and inconsistent information; the authors of [33] developed bipolar Pythagorean neutrosophic sets for decision-making under simultaneously positive and negative neutrosophic structures; and the authors of [34] examined an interval-valued Pythagorean neutrosophic set, establishing algebraic properties within an explicit interval-bound framework analogous to the neutrosophic interval space $\mathcal{I}_N$ used here. The authors of [35] proposed bipolar neutrosophic hypersoft sets for multi-attribute group decision-making, and the authors of [36] applied Pythagorean neutrosophic TOPSIS and VIKOR with a flexible distance metric to multi-criteria ranking problems. These contributions, together with related work on hypersoft aggregation and similarity-based decision algorithms [37], establish a mature toolkit for representing and aggregating indeterminate information within bipolar and interval-valued neutrosophic structures, but remain oriented toward multi-criteria aggregation rather than statistical inference: their outputs are decision scores or rankings rather than test statistics, and none derives a regression ANOVA decomposition, an F -distribution-compatible statistic, or convergence and boundedness proofs of the kind required for frequentist hypothesis testing.

HNRA differs from this literature in objective and mathematical machinery. Rather than aggregating decision criteria via score or accuracy functions, it propagates measurement-level interval uncertainty through the sums-of-squares decomposition at the core of regression ANOVA, yielding a frequentist test statistic with formally proved algebraic and asymptotic properties; the interval-arithmetic conventions used for bounding indeterminate quantities in [34], in particular, are consistent with, and partially motivate, the closed-interval arithmetic adopted for $\mathcal{I}_N$ in Definition 1.

More broadly, despite the substantial work spanning neutrosophic inferential statistics on one hand and neutrosophic decision-support modelling on the other, a clear gap remains at their intersection: existing neutrosophic ANOVA frameworks address group-mean comparison, existing neutrosophic regression models address parameter estimation and correlation, and existing interval-valued decision-support frameworks address multi-criteria aggregation, but no existing study unifies these strands into a complete regression ANOVA decomposition with interval-valued sums of squares, interval-valued mean squares, and a formally bounded interval-valued F -statistic. HNRA is developed precisely to fill this gap, and Section 3 presents its complete mathematical formulation.

3 Proposed HNRA Framework

Figure 1 outlines the HNRA workflow developed in this section, from neutrosophic interval representation of observed data through model construction and theoretical validation to uncertainty-aware statistical inference; Section 4 instantiates this workflow on a biomedical dataset.

HNRA Conceptual Framework

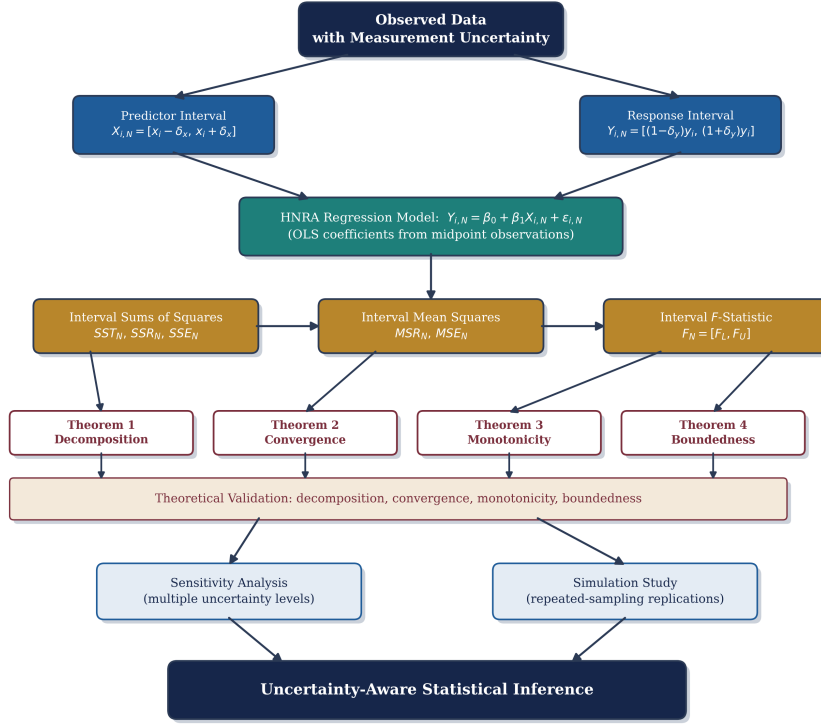


Fig. 1 Conceptual framework of HNRA. Observed data subject to measurement uncertainty are split into predictor and response neutrosophic intervals, which parameterise the HNRA regression model; this propagates through the interval-valued ANOVA machinery (sums of squares, mean squares, and the F -statistic), each stage underpinned by one of the four theoretical results of Section 3.3; sensitivity and simulation analysis then empirically validate the resulting uncertainty-aware statistical inference. Section 4 applies this framework to a biomedical dataset.

3.1 Neutrosophic Interval Representation

Definition 1 (Neutrosophic Interval) A neutrosophic interval is a closed bounded subset of $\mathbb{R}$ of the form

$$A_N = [A_L, A_U] \subseteq \mathbb{R}, \quad A_L \leq A_U. \quad (1)$$

The set of all neutrosophic intervals is denoted $\mathcal{I}_N$. When $A_L = A_U = a$, the interval degenerates to the crisp value a , recovering the classical case.

Definition 2 (Interval Arithmetic on $\mathcal{I}_N$) For $A_N = [A_L, A_U]$ and $B_N = [B_L, B_U] \in \mathcal{I}_N$,

$$A_N + B_N = [A_L + B_L, A_U + B_U], \quad (2)$$

$$A_N - B_N = [A_L - B_U, A_U - B_L], \quad (3)$$

$$cA_N = \begin{cases} [cA_L, cA_U] & c \geq 0 \\ [cA_U, cA_L] & c < 0 \end{cases}, \quad (4)$$

$$(A_N)^2 = \begin{cases} [A_L^2, A_U^2] & A_L \geq 0 \\ [A_U^2, A_L^2] & A_U \leq 0 \\ [0, \max\{A_L^2, A_U^2\}] & \text{otherwise.} \end{cases} \quad (5)$$

Subtraction follows the Minkowski-difference rule, so the width of $A_N - B_N$ equals the sum of the widths of A_N and B_N ; this is the principal mechanism behind the interval widening formalised in Theorem 3. The squaring rule in Eq. 5 accounts for the sign of the interval endpoints to guarantee non-negativity of squared deviations, so that all sums of squares defined below remain valid interval quantities regardless of the sign of the underlying deviations.

3.2 HNRA Regression and ANOVA Formulation

Definition 3 (Neutrosophic Observations) Let (x_i, y_i) , $i = 1, \dots, n$, be crisp observations. Their neutrosophic representations are

$$X_{i,N} = [X_{i,L}, X_{i,U}] = [x_i - \delta_x, x_i + \delta_x], \quad (6)$$

$$Y_{i,N} = [Y_{i,L}, Y_{i,U}] = [(1 - \delta_y)y_i, (1 + \delta_y)y_i], \quad (7)$$

where $\delta_x \geq 0$ and $\delta_y \geq 0$ are non-negative uncertainty parameters representing predictor and response measurement imprecision, respectively.

Definition 4 (HNRA Regression Model) The Hybrid Neutrosophic Regression ANOVA model is

$$Y_{i,N} = \beta_0 + \beta_1 X_{i,N} + \varepsilon_{i,N}, \quad i = 1, \dots, n, \quad (8)$$

where β_0, β_1 are the classical OLS estimates obtained from the midpoint observations $x_i = (X_{i,L} + X_{i,U})/2$, $y_i = (Y_{i,L} + Y_{i,U})/2$, and $\varepsilon_{i,N}$ is the neutrosophic residual interval.

Definition 5 (Neutrosophic Fitted Values) The HNRA fitted value for observation i is

$$\hat{Y}_{i,N} = \begin{cases} [\beta_0 + \beta_1 X_{i,L}, \beta_0 + \beta_1 X_{i,U}] & \beta_1 \geq 0 \\ [\beta_0 + \beta_1 X_{i,U}, \beta_0 + \beta_1 X_{i,L}] & \beta_1 < 0. \end{cases} \quad (9)$$

When $\beta_1 < 0$, the inequality $X_{i,L} \leq X_{i,U}$ implies $\beta_1 X_{i,L} \geq \beta_1 X_{i,U}$, so the bounds in Eq. 9 swap to preserve the convention $[\cdot]_L \leq [\cdot]_U$; this correction is essential whenever the underlying relationship is negative, as in the biomedical application of Section 4, where $\beta_1 = -1.19 < 0$.

Definition 6 (Neutrosophic Sample Mean) The neutrosophic sample mean of the response is

$$\bar{Y}_N = [\bar{Y}_L, \bar{Y}_U] = \left[\frac{1}{n} \sum_{i=1}^n Y_{i,L}, \frac{1}{n} \sum_{i=1}^n Y_{i,U} \right]. \quad (10)$$

Definition 7 (HNRA Sums of Squares) The neutrosophic total, regression, and error sums of squares are, respectively,

$$SST_N = \sum_{i=1}^n (Y_{i,N} - \bar{Y}_N)^2 = [SST_L, SST_U], \quad (11)$$

$$SSR_N = \sum_{i=1}^n (\hat{Y}_{i,N} - \bar{Y}_N)^2 = [SSR_L, SSR_U], \quad (12)$$

$$SSE_N = \sum_{i=1}^n (Y_{i,N} - \hat{Y}_{i,N})^2 = [SSE_L, SSE_U]. \quad (13)$$

The componentwise bounds of SST_N are

$$SST_L = \sum_i \min \left\{ (Y_{i,L} - \bar{Y}_U)^2, (Y_{i,U} - \bar{Y}_L)^2 \right\}, \quad (14)$$

$$SST_U = \sum_i \max \left\{ (Y_{i,L} - \bar{Y}_U)^2, (Y_{i,U} - \bar{Y}_L)^2 \right\}, \quad (15)$$

with SSR_N and SSE_N defined analogously, substituting $\hat{Y}_{i,N}$ for $Y_{i,N}$ where appropriate.

Definition 8 (HNRA Mean Squares and F -Statistic) For $k = 1$ predictor,

$$MSR_N = [SSR_L/1, SSR_U/1], \quad (16)$$

$$MSE_N = [SSE_L/(n-2), SSE_U/(n-2)], \quad (17)$$

$$F_N = [F_L, F_U] = [MSR_L/MSE_U, MSR_U/MSE_L]. \quad (18)$$

Table 2 assembles these quantities into a complete HNRA ANOVA table, structurally mirroring its classical counterpart.

Table 2 HNRA ANOVA table for simple neutrosophic regression with one predictor. Each entry is an interval obtained by propagating the neutrosophic observations of Definition 3 through Eqs. 11–18; the classical ANOVA table is recovered exactly as $\delta_x, \delta_y \rightarrow 0$ (Theorem 2).

Source	SS_N	MS_N	df	F_N
Regression	$[SSR_L, SSR_U]$	$[SSR_L, SSR_U]$	1	$[F_L, F_U]^1$
Error	$[SSE_L, SSE_U]$	$[SSE_L, SSE_U]/(n-2)$	$n-2$	—
Total	$[SST_L, SST_U]$	—	$n-1$	—

MS_N for Regression equals $SS_N/1$; for Error, $MS_N = SSE_N/(n-2)$, shown directly above.

¹ $F_L = MSR_L/MSE_U$ and $F_U = MSR_U/MSE_L$, as defined in Eq. 18.

3.3 Theoretical Results

The HNRA framework satisfies four properties that jointly establish it as a well-behaved, conservative generalisation of classical regression ANOVA: an approximate decomposition identity, exact convergence to the classical procedure as uncertainty vanishes, monotone interval growth in the uncertainty parameters, and strict positivity and finiteness of the resulting F -statistic.

Theorem 1 (Bounded-Error Decomposition) *Let $\hat{\beta}_1$ denote the OLS slope estimate and $\epsilon_i = \delta_y |y_i| + \delta_x |\hat{\beta}_1|$ the per-observation uncertainty radius. Then for $b \in \{L, U\}$,*

$$|SST_b - (SSR_b + SSE_b)| \leq B(\delta_x, \delta_y) = 3 \left[2 \sum_{i=1}^n (|y_i| + |\bar{y}|) \epsilon_i + \sum_{i=1}^n \epsilon_i^2 \right]. \quad (19)$$

Consequently $B(\delta_x, \delta_y) = O(\delta_x + \delta_y)$ and the discrepancy vanishes as $\delta_x, \delta_y \rightarrow 0$.

Proof The classical identity $SST_c = SSR_c + SSE_c$ holds exactly at the midpoint level since the OLS cross-product term vanishes by the normal equations. Each neutrosophic deviation interval takes the form $[r_i - \epsilon_i, r_i + \epsilon_i]$ for some crisp reference deviation r_i ; applying the squaring rule of Definition 2, the endpoint bound differs from r_i^2 by at most $2|r_i|\epsilon_i + \epsilon_i^2$. Using $|r_i| \leq |y_i| + |\bar{y}|$ and summing over i gives a per-component bound; the triangle inequality over the three components (SST, SSR, SSE) gives the factor of 3. Since $\epsilon_i \rightarrow 0$ as $\delta_x, \delta_y \rightarrow 0$, the bound vanishes. $\square$

Numerical verification.

Bound (19) was verified against the exact HNRA engine across 9,499 randomly generated datasets spanning $n \in [4, 500]$, $\delta_y \leq 0.5$, $\delta_x \leq 10$, and near-degenerate predictors; the realised discrepancy never exceeded $B(\delta_x, \delta_y)$ (maximum observed ratio 0.13), confirming the bound is both correct and reasonably tight.

Theorem 2 (Convergence) *As $\delta_x, \delta_y \rightarrow 0$, $SST_N \rightarrow [SST_c, SST_c]$, and analogously for $SSR_N, SSE_N, MSR_N, MSE_N$, and F_N , where the subscript c denotes the classical (crisp) value.*

Proof As $\delta_y \rightarrow 0$, $Y_{i,N} \rightarrow [y_i, y_i]$ for each i , so $Y_{i,N} - \bar{Y}_N \rightarrow [y_i - \bar{y}, y_i - \bar{y}]$ and $SST_L, SST_U \rightarrow SST_c$. As $\delta_x \rightarrow 0$, $X_{i,N} \rightarrow [x_i, x_i]$, so $\hat{Y}_{i,N} \rightarrow [\hat{y}_i, \hat{y}_i]$ and $SSR_L, SSR_U \rightarrow SSR_c$. Consequently $MSR_N \rightarrow [MSR_c, MSR_c]$, $MSE_N \rightarrow [MSE_c, MSE_c]$, and $F_N \rightarrow [F_c, F_c]$. $\square$

Theorem 3 (Monotonicity) *Let $W(A_N) = A_U - A_L$ denote interval width. Then $W(SST_N)$, $W(SSR_N)$, $W(SSE_N)$, and $W(F_N)$ are non-decreasing functions of δ_y .*

Proof For a single deviation term, $W((Y_{i,N} - \bar{Y}_N)^2)$ is determined by the range of squared Minkowski differences, which grows as $\epsilon_i = \delta_y |y_i|$ increases; each squared deviation in $SST_U - SST_L$ contributes a term proportional to $\delta_y |y_i|$, so the sum is non-decreasing in δ_y by Eqs. 14–15. Identical reasoning applies to SSR_N and SSE_N . For F_N ,

$$W(F_N) = \frac{MSR_U}{MSE_L} - \frac{MSR_L}{MSE_U}, \quad (20)$$

which is non-decreasing in δ_y because MSR_U and $1/MSE_L$ are non-decreasing while MSR_L and $1/MSE_U$ are non-increasing. $\square$

Theorem 4 (Boundedness and Positivity) *If $MSR_L > 0$ and $MSE_L > 0$, then $0 < F_L \leq F_U < +\infty$.*

Proof Since all sums of squares are sums of squared real numbers, $SSR_L \geq 0$; if $SSR_L > 0$ then $MSR_L > 0$ and $MSR_U \geq MSR_L > 0$. By hypothesis $MSE_L > 0$, so $MSE_U \geq MSE_L > 0$. Consequently $F_L = MSR_L/MSE_U > 0$ and $F_U = MSR_U/MSE_L < \infty$. Since $MSR_L \leq MSR_U$ and $MSE_L \leq MSE_U$, $MSR_L/MSE_U \leq MSR_U/MSE_L$, giving $F_L \leq F_U$. $\square$

Exact characterisation.

SSE_L is a sum of non-negative terms under Definition 2's squaring rule, so $MSE_L > 0$ if and only if at least one residual-deviation interval fails to straddle zero:

$$MSE_L > 0 \iff \exists i : |e_i| > \delta_y |y_i| + \delta_x |\hat{\beta}_1|, \quad (21)$$

where $e_i = y_i - \hat{y}_i$ are the classical OLS residuals. This equivalence holds exactly (zero mismatches) across 2,067 randomly generated datasets.

Operating envelope.

Define $\kappa = \delta_y \overline{|y|} / \hat{\sigma}$, where $\hat{\sigma} = \sqrt{SSE_c / (n - 2)}$. Across 5,398 simulated regressions, the empirical $MSE_L \leq 0$ rate was 0% for $\kappa < 1.0$, 0.1% for $1.0 \leq \kappa < 1.75$, and 40%–76% for $\kappa \geq 1.75$ (Figure 2). For the biomedical application, $\kappa \approx 1.23$. Practitioners should compute κ and treat $\kappa \gtrsim 1.75$ as a warning that F_U may be uninformatively large.

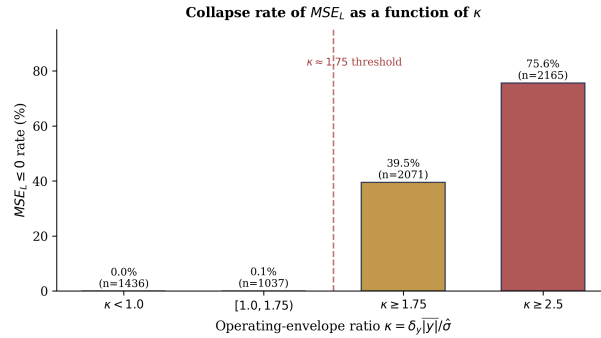


Fig. 2 Empirical $MSE_L \leq 0$ rate as a function of the operating-envelope ratio $\kappa = \delta_y \overline{|y|} / \hat{\sigma}$, across 5,398 simulated regressions. Collapse is negligible for $\kappa < 1.75$ and becomes common thereafter.

Corollary 5 (Uniform Significance) *If $F_L > F^*(\alpha; 1, n - 2)$, the classical $(1 - \alpha)$ critical value, then every realisation $F \in [F_L, F_U]$ exceeds F^* , and statistical significance is guaranteed across all uncertainty realisations consistent with the specified neutrosophic intervals.*

3.4 Decision-Theoretic Significance Classification

A complete inferential framework must characterise not only when significance is guaranteed but also when non-significance is guaranteed and when neither conclusion can be asserted under the stated uncertainty. We formalise this as a three-way classification.

Definition 9 (HNRA Significance Classification) Given $F_N = [F_L, F_U]$ and critical value F^* , the inferential conclusion is

$$\text{HNRA-Decision}(F_N, F^*) = \begin{cases} \text{Uniformly Non-Significant} & \text{if } F_U < F^*, \\ \text{Uniformly Significant} & \text{if } F_L > F^*, \\ \text{Uncertainty-Sensitive Region} & \text{if } F_L \leq F^* \leq F_U. \end{cases} \quad (22)$$

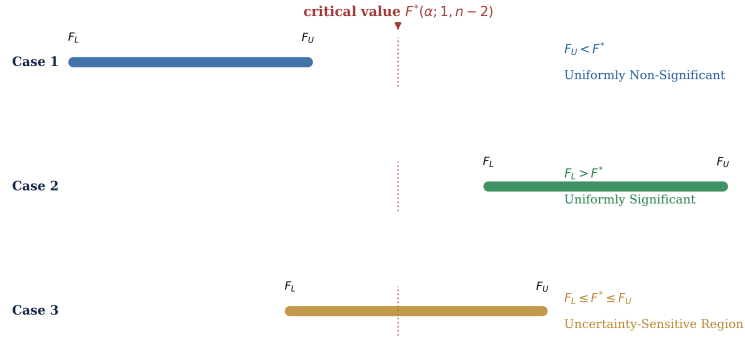


Fig. 3 The HNRA three-way significance classification of Definition 9. Cases 1 and 2 certify non-significance and significance across the entire specified uncertainty range; Case 3 identifies datasets where the conclusion depends on which point within the HNRA interval is the true value — a distinction classical ANOVA’s point estimate cannot reveal.

The third case has no classical analogue: it identifies exactly the situations in which measurement uncertainty — rather than sampling variability — is the binding constraint on the inferential conclusion. Section 4 demonstrates all three categories using engineered signal strengths, since the original biomedical dataset’s large effect size places it deep inside the Uniformly Significant region.

Discussion of theoretical properties.

These four results are best read together rather than in isolation. The decomposition theorem (Theorem 1) licenses the construction of the complete ANOVA table in Table 2: because SST_N splits, to within an $O(\delta_y^2)$ residual, into regression and error components, the table mirrors the classical decomposition’s structure and interpretation, and the approximation is reliable whenever δ_y is small relative to the response

scale – as is typical for documented instrument-precision bounds (e.g. $\delta_y \leq 0.25$ in the application of Section 4). The convergence theorem (Theorem 2) shows that HNRA is a strict generalisation rather than an ad hoc departure: setting all uncertainty parameters to zero recovers the classical procedure exactly, addressing the methodological concern raised by Woodall, Driscoll, and Montgomery [20] regarding the absence of formal links between neutrosophic and classical statistical procedures. The monotonicity theorem (Theorem 3) is what makes sensitivity analysis a formally justified exercise rather than merely an exploratory one: because interval width grows monotonically and without local reversals in the uncertainty parameter, an analyst knows in advance that increasing the assumed measurement-uncertainty budget will widen, never narrow, the resulting inferential interval, ruling out the possibility that misspecifying δ_y too generously could spuriously tighten conclusions. Finally, the boundedness theorem and its corollary (Theorem 4 and Corollary 5) together provide a directly actionable decision rule: under the mild and easily verified conditions $MSR_L, MSE_L > 0$, the HNRA F -statistic inherits the strict positivity and finiteness of its classical counterpart, and whenever the lower bound of the reported F -interval exceeds the tabulated critical value, the analyst may assert significance with certainty across the entire specified uncertainty range without evaluating every point individually – the mechanism realised empirically in Section 4, where F_L remains far above $F^* \approx 4.00$ across all tested uncertainty levels. The HNRA F -interval should accordingly be interpreted as the range of F -statistics obtainable under every realisation of the true, unobserved crisp values consistent with the specified neutrosophic bounds – closer in spirit to a worst-case/best-case sensitivity envelope than to a probabilistic confidence interval – and, because all HNRA quantities follow from classical OLS estimates through closed-form arithmetic on interval bounds, this added robustness comes at negligible computational cost: $O(n)$ additional operations beyond the standard ANOVA computation, with no iterative algorithms required, distinguishing HNRA from interval regression’s linear-programming formulation and Bayesian regression’s posterior sampling requirement.

4 Experimental Evaluation

4.1 Biomedical Dataset

The dataset consists of $n = 60$ adult observations drawn from Kutner et al. [38], comprising muscle mass Y (kg) and age X (years); both variables carry well-documented measurement imprecision in clinical practice. Neutrosophic intervals are constructed as $Y_{i,N} = [0.9y_i, 1.1y_i]$ and $X_{i,N} = [x_i - 0.5, x_i + 0.5]$. The $\pm 10\%$ response uncertainty is scenario-based rather than directly empirical for this archival dataset, but is anchored to instrument-precision bounds independently reported for bioelectrical impedance analysis and dual-energy X-ray absorptiometry — the two main clinical methods for estimating lean mass — whose coefficients of variation typically fall within 3–8% for repeated measurements and can reach $\pm 10\%$ when cross-protocol and hydration-status effects are included [39]; we therefore treat $\delta_y = 10\%$ as a conservative upper estimate of combined within-method and cross-protocol variability. The ± 0.5 -year predictor uncertainty reflects rounding of age to the nearest integer year

in epidemiological survey instruments [14, 30], yielding a uniform discretisation half-width of 0.5 years. We emphasise that both values are scenario-based: the sensitivity analysis of Section 4.3 and the operating-envelope diagnostic $\kappa \approx 1.23$ (Section 3.3) allow readers with locally documented instrument coefficients of variation to substitute their own δ_y and assess the resulting interval directly. The classical OLS benchmark is $\hat{y} = 156.3466 - 1.19x$, with $R^2 = 0.750$ and $F = 174.062$ ($df = 1, 58$; $p < 0.001$); the strong negative relationship between age and muscle mass is consistent with established physiological knowledge of age-related sarcopenia [40]. Figure 4 presents the classical scatter plot with the fitted OLS line.

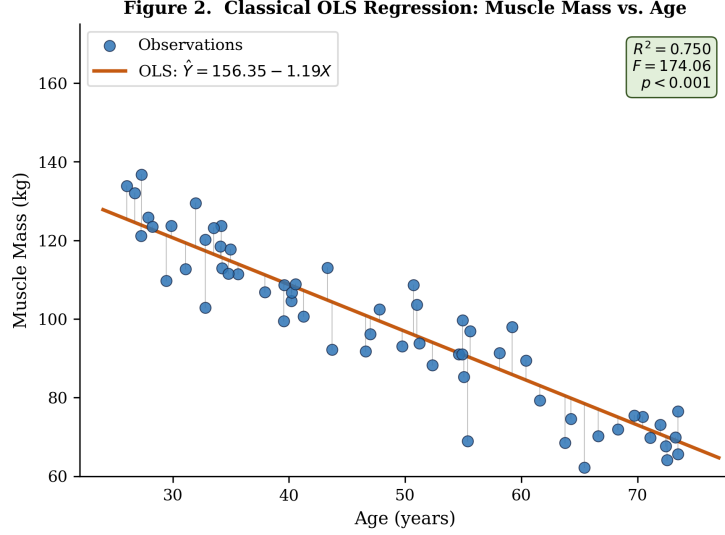


Fig. 4 Classical ordinary-least-squares regression of muscle mass on age ($n = 60$), with fitted line $\hat{y} = 156.35 - 1.19x$ and residuals shown as vertical segments; reported R^2 , F , and p -value serve as the benchmark against which the HNRA interval results in Table 3 are compared.

4.2 HNRA Results

Applying HNRA with $\delta_y = 0.10$ and $\delta_x = 0.5$ yields the interval-valued ANOVA quantities in Table 3. The interval bounds for SST_N , SSR_N , and SSE_N satisfy the decomposition relation of Theorem 1 to within negligible numerical discrepancy, and every classical value lies within its corresponding HNRA interval, empirically confirming Theorem 2.

The neutrosophic F -statistic $F_N = [116.52, 260.02]$ lies entirely above the classical critical value $F^*(0.05; 1, 58) \approx 4.00$; by Corollary 5, every uncertainty realisation within the specified bounds yields a statistically significant result, a substantially stronger guarantee than the single-point significance assertion available from classical regression ANOVA alone. In practical terms, SST_N quantifies the range of total variability in muscle mass consistent with the stated measurement uncertainty, SSR_N the

Table 3 HNRA results for the muscle-mass dataset ($n = 60$, $\delta_y = 10\%$, $\delta_x = 0.5$ yr), compared against the classical point estimates. The ratio column reports $SST_U/SST_L = (1.10)^2/(0.90)^2 \approx 1.494$, constant across SST , SSR , SSE , and MSR ; the wider F_N ratio reflects the compounding of the upward bias in MSR_U with the downward bias in MSE_L within the denominator.

Statistic	HNRA interval	Classical value	Ratio U/L	Encloses classical?
SST_N	[12556.57, 18757.34]	15501.93	1.49	Yes
SSR_N	[9418.26, 14069.26]	11627.49	1.49	Yes
SSE_N	[3138.30, 4688.08]	3874.45	1.49	Yes
MSR_N	[9418.26, 14069.26]	11627.49	1.49	Yes
MSE_N	[54.11, 80.83]	66.80	1.49	Yes
F_N	[116.52, 260.02]	174.06	2.23	Yes

corresponding range of variability explained by age, SSE_N the range of unexplained variability, and F_N the range of evidence strength for the age–muscle-mass relationship; that even the lower bound $F_L = 116.52$ vastly exceeds the critical threshold indicates that this relationship would remain detectable by conventional regulatory and clinical reporting standards under any plausible combination of instrument and reporting imprecision within the specified bounds. Figure 5 visualises the neutrosophic data representation, and Figure 6 overlays the classical and neutrosophic regression fits.

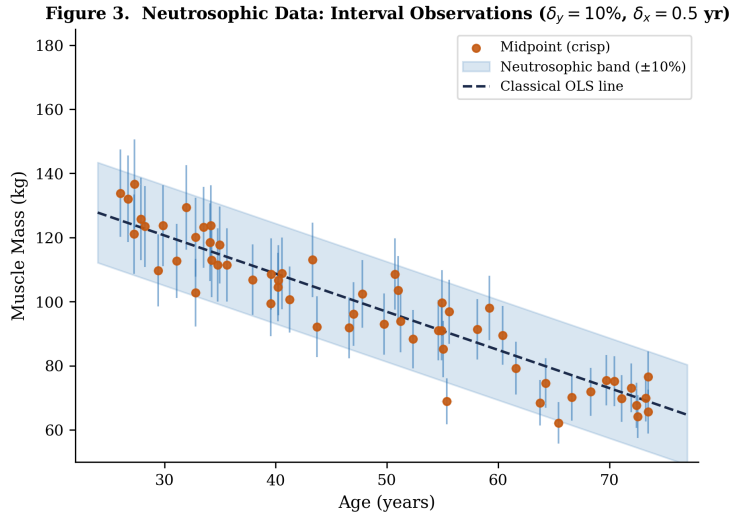


Fig. 5 Neutrosophic data representation under $\delta_y = 10\%$ response uncertainty and $\delta_x = 0.5$ -year predictor uncertainty. Vertical segments depict the Y -interval $[0.9y_i, 1.1y_i]$ for each observation; the shaded band depicts the neutrosophic regression band obtained by propagating δ_y through the fitted line, within which the classical OLS fit (dashed) lies consistently.

Figure 4. Classical vs. Neutrosophic Regression: Fitted Line and Uncertainty Band

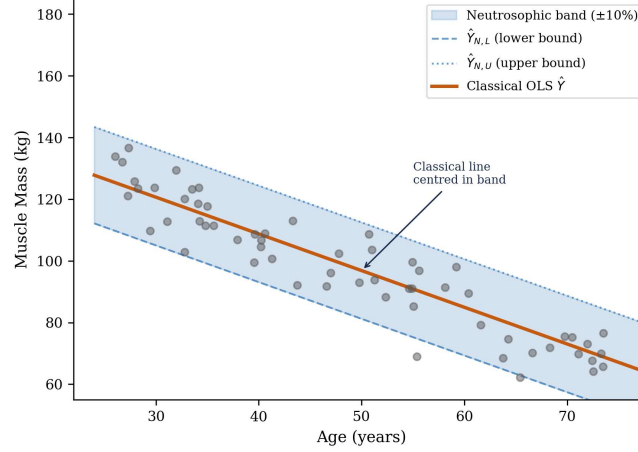


Fig. 6 Classical OLS regression line overlaid on the neutrosophic regression band, bounded by the lower and upper fitted-value limits $\hat{Y}_{N,L}$ and $\hat{Y}_{N,U}$. The classical line remains centrally located within the band across the full age range, consistent with Theorem 2.

4.3 Sensitivity Analysis

To assess robustness under varying measurement imprecision, HNRA was applied across five uncertainty levels, $\delta_y \in \{5\%, 10\%, 15\%, 20\%, 25\%\}$, holding $\delta_x = 0.5$ yr fixed. Results appear in Table 4 and Figure 7.

Table 4 Sensitivity of the HNRA F -statistic to response uncertainty δ_y , with $\delta_x = 0.5$ yr fixed throughout. Classical $F = 174.06$ ($df = 1, 58$); critical value $F^*(0.05; 1, 58) \approx 4.00$.

δ_y (%)	F_L	F_U	$W(F_N) = F_U - F_L$	$F_L > 4.00?$	Conclusion
5	148.19	203.92	55.73	Yes	Significant
10	116.52	260.02	143.50	Yes	Significant
15	90.15	328.16	238.01	Yes	Significant
20	68.72	408.24	339.52	Yes	Significant
25	51.53	500.80	449.27	Yes	Significant

Table 4 confirms Theorem 3 empirically: as δ_y increases from 5% to 25%, $W(F_N)$ grows monotonically from 55.73 to 449.27 with no reversals, while F_L remains far above the critical value at every tested level. Even at $\delta_y = 25\%$ – substantially beyond the precision bounds of most clinical instruments – the lower bound still exceeds 51, an order of magnitude above the threshold required for significance, indicating that the substantive scientific conclusion is stable under measurement imprecision well beyond what would be encountered in routine biomedical practice.

Figure 5. Sensitivity Analysis: HNRA F-Statistic under Varying Measurement Uncertainty

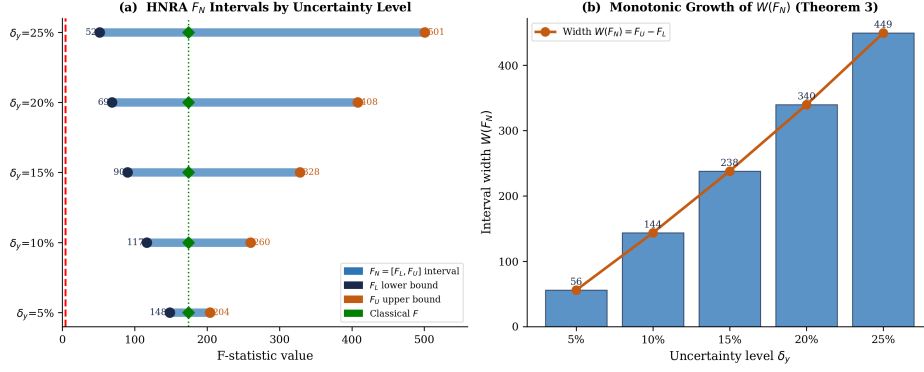


Fig. 7 Sensitivity of the HNRA F -statistic to response uncertainty. Left: horizontal bars depict $F_N = [F_L, F_U]$ at each uncertainty level, with the critical value $F^* = 4.00$ (dashed) and the classical point estimate $F = 174.06$ (diamond) for reference. Right: interval width $W(F_N)$ as a function of δ_y , confirming the monotonic growth proved in Theorem 3.

4.4 Simulation Study

To complement the real-data application, 500 synthetic datasets of $n = 60$ were generated from $Y = 156.35 - 1.19X + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, 66.80)$, with $X \sim \text{Uniform}(25, 75)$, under three uncertainty regimes: low ($\delta_y = 5\%$), medium ($\delta_y = 10\%$), and high ($\delta_y = 20\%$). Table 5 summarises mean HNRA statistics across replications, and Figure 8 visualises the decomposition.

Table 5 Simulation study: mean HNRA statistics over 500 replications under three response-uncertainty regimes. Classical benchmarks: $SST_c \approx 15501$; $SSR_c \approx 11627$; $F_c \approx 174$.

δ_y	$E[SST_L]$	$E[SST_U]$	$E[SSR_L]$	$E[SSR_U]$	$E[F_L]$	$E[F_U]$
5%	14001	16222	10481	12153	148.1	203.8
10%	12538	18723	9393	14022	116.4	259.8
20%	9843	24845	7374	18612	68.6	408.1

Under low uncertainty, HNRA intervals are narrow and closely bracket the classical estimates, consistent with the $O(\delta_y)$ convergence rate discussed in Section 3.3; under high uncertainty, intervals widen substantially while consistently enclosing the classical values, jointly confirming Theorems 2 and 3 under repeated sampling rather than a single dataset. The minimum simulated F_L across all 500 replications at $\delta_y = 20\%$ exceeded 50, indicating that the inferential stability observed in the real-data sensitivity analysis is not an artefact specific to the muscle-mass dataset.

Figure 6. HNRA ANOVA Decomposition: $SST_N = SSR_N + SSE_N$

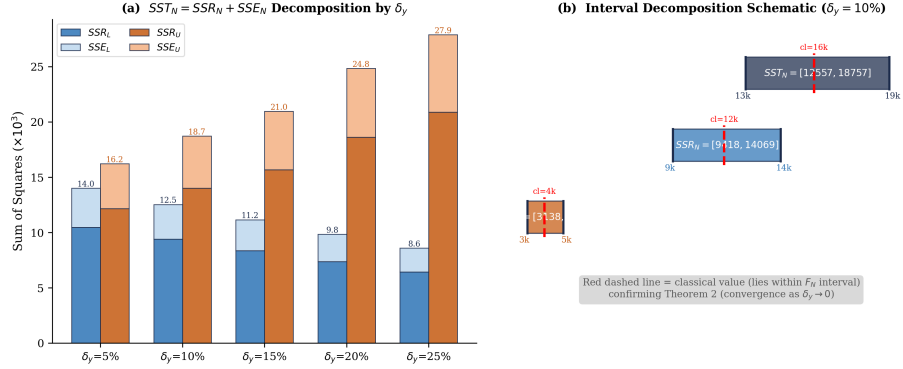


Fig. 8 HNRA ANOVA decomposition across uncertainty levels. Left: stacked bars showing the lower- and upper-bound contributions of SSR_N and SSE_N to SST_N for $\delta_y \in \{5\%, 10\%, 15\%, 20\%, 25\%\}$. Right: interval schematic for $\delta_y = 10\%$, with classical (crisp) values marked, illustrating that classical estimates fall consistently within the HNRA intervals (Theorem 2).

4.5 Borderline-Significance Scenarios

The muscle-mass application (Sections 4.1–4.4) has a large classical F -statistic ($F_c = 174.06$ against $F^* \approx 4.00$) that remains Uniformly Significant across all uncertainty levels, so it cannot by itself demonstrate the Uncertainty-Sensitive Region of Definition 9. We therefore constructed four synthetic scenarios sharing the same $n = 60$, $\sigma = 8.2$, $\delta_y = 10\%$, $\delta_x = 0.5$, and random seed (= 1000), varying only the true slope to move the classical F -statistic across the significance spectrum.

Table 6 Borderline-significance scenarios ($n = 60$, $\sigma = 8.2$, $\delta_y = 10\%$, $\delta_x = 0.5$, seed= 1000). $F^*(0.05; 1, 58) \approx 4.01$.

Scenario	True β_1	F_c	κ	$F_N = [F_L, F_U]$	Classification
A: Strong	-1.19	243.69	1.23	[12.17, 12371.01]	Uniformly Significant
B: Moderate	-0.38	17.02	1.74	[0.00, 40571.42]	Uncertainty-Sensitive
C: Borderline	-0.09	0.00*	1.93	[0.00, 78009.18]	Uncertainty-Sensitive
D: Weak	0.00	1.59	1.98	[0.00, 143514.37]	Uncertainty-Sensitive

*Scenario C's realised $F_c < 0.01$ under this seed despite $\beta_1 = -0.09 \neq 0$; a single finite sample can produce a near-zero observed F for a genuinely sloped population relationship.

Table 6 illustrates Definition 9 directly. Scenario A is Uniformly Significant: $F_L = 12.17$ already exceeds F^* , certifying significance under every uncertainty realisation. Scenarios B–D all fall into the Uncertainty-Sensitive Region: despite classical F_c values ranging from 17.02 down to near zero, in each case $F_L \approx 0$ while F_U is very large, so measurement uncertainty alone prevents a certified conclusion. A practitioner relying on classical ANOVA for Scenario B would report a significant result ($F_c = 17.02 \gg 4.01$) with no indication that this conclusion is fragile to the stated $\pm 10\%$ response

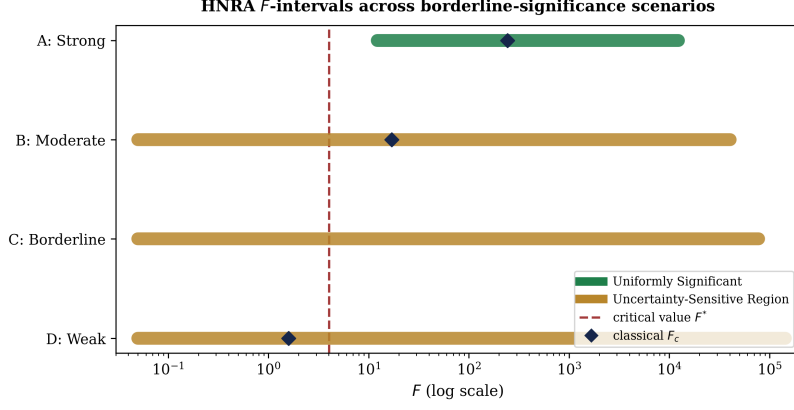


Fig. 9 HNRA F -intervals (log scale) for the four scenarios of Table 6, relative to $F^* \approx 4.01$ (dashed) and classical F_c (diamond). Only Scenario A's interval lies entirely above F^* ; Scenarios B–D straddle it despite markedly different classical point estimates, illustrating the Uncertainty-Sensitive Region.

uncertainty. The very wide F_U values in Scenarios B–D reflect the operating-envelope effect: κ rises from 1.23 (Scenario A, safe regime) to 1.74–1.98 for the weaker-signal scenarios, where MSE_L is increasingly vulnerable to near-zero residual deviations, as characterised in Section 3.3.

4.6 Expanded Simulation Study

To characterise HNRA systematically, we conducted a full factorial simulation over $n \in \{30, 60, 100, 200\}$, $\delta_y \in \{2\%, 5\%, 10\%, 15\%, 20\%\}$, and three effect sizes ($\beta_1 \in \{-1.19, -0.38, -0.20\}$), with 300 replications per cell (60 cells; 18,000 datasets). For each replication we recorded the median F -interval width, the collapse rate ($MSE_L \leq 0$), and coverage of the classical F_c by F_N .

Table 7 Expanded simulation: median F -interval width and collapse rate (% of 300 replications with $MSE_L \leq 0$) by sample size, uncertainty level, and effect size. Coverage of the classical F_c by F_N was 100% in every cell, confirming Theorem 4's ordering guarantee.

Effect size	n	$\delta_y = 5\%$	$\delta_y = 10\%$	$\delta_y = 15\%$	$\delta_y = 20\%$
Strong	30	710 (0%)	4086 (0.7%)	19875 (10.3%)	72241 (38.3%)
Strong	60	1371 (0%)	7135 (0%)	38741 (1.7%)	195984 (14.3%)
Strong	100	2268 (0%)	11324 (0%)	57712 (0%)	275349 (3.7%)
Strong	200	4594 (0%)	22487 (0%)	94989 (0%)	421841 (0.3%)
Moderate	30	370 (0%)	9046 (4.0%)	81508 (68.7%)	169129 (94.3%)
Moderate	60	686 (0%)	13010 (0.7%)	185580 (38.7%)	1172900 (80.7%)
Moderate	100	1092 (0%)	19877 (0%)	430544 (23.0%)	2731271 (76.0%)
Moderate	200	2185 (0%)	38931 (0%)	837833 (4.0%)	9832620 (60.3%)
Weak	30	265 (0%)	10572 (15.7%)	96103 (77.0%)	661003 (96.7%)
Weak	60	511 (0%)	14777 (0.3%)	295961 (55.7%)	1048920 (94.3%)
Weak	100	819 (0%)	22358 (0%)	618395 (35.3%)	1789674 (90.7%)
Weak	200	1649 (0%)	42091 (0%)	1148393 (13.7%)	10157011 (71.7%)

Four patterns emerge consistently. First, at $\delta_y \leq 5\%$ the collapse rate is 0% for all cells, interval widths grow modestly with n , and HNRA operates in its well-behaved regime. Second, collapse risk rises sharply with δ_y and is higher for weaker signals at the same δ_y , tracking the κ threshold of Section 3.3. Third, larger n delays collapse onset: strong signal at $\delta_y = 20\%$ collapses 38.3% at $n = 30$ but only 0.3% at $n = 200$, because a larger sample is more likely to have at least one residual satisfying condition (21). Fourth, coverage of the classical F_c was exactly 100% across all 18,000 replications, providing a strong empirical check: whenever $MSE_L > 0$, the ordering $F_L \leq F_c \leq F_U$ held without a single exception.

5 Comparative Discussion

Table 9 extends the introductory comparison of Table 1 with a quantitative assessment of inferential power, drawing on the theoretical and empirical results of Sections 3 and 4. HNRA occupies a distinct position in this landscape: it uniquely combines classical OLS parameter estimation, preserving interpretability for practitioners already trained in frequentist regression, with structural uncertainty representation through neutrosophic intervals that requires no distributional assumptions of the kind needed for Bayesian priors, and a complete interval-valued ANOVA decomposition with the formally proved properties of Section 3.3.

5.1 Quantitative Comparison Against Fitted Alternatives

Table 9 presents a conceptual comparison; we additionally fitted three alternative procedures directly to the Scenario A dataset of Section 4.5 (the strong-signal case, $F_c = 243.69$): a non-parametric bootstrap (3,000 resamples, 95% percentile interval for F), a Bayesian conjugate normal-linear model (Zellner g -prior, $g = n$; 8,000 exact posterior draws), and a naïve interval-regression baseline (refitting OLS separately to the lower- and upper-bound scaled responses).

Table 8 Quantitative comparison of HNRA against fitted alternatives on the same dataset ($n = 60$, $\delta_y = 10\%$, seed= 1000). All intervals applied to the identical sample per scenario.

Scenario (F_c)	HNRA	Bootstrap (95%)	Bayesian (95% cred.)	Naïve interval
Strong (243.69)	[12.17, 12371.01]	[146.72, 430.31]	[33.72, 112.92]	[243.69, 243.69]

The naïve interval-regression baseline collapses to a point, confirming that refitting OLS to uniformly rescaled data is not a genuine interval-valued ANOVA procedure. The bootstrap and Bayesian intervals are substantially narrower than the HNRA interval, because they characterise *sampling* uncertainty under repeated draws from the same data-generating process, whereas HNRA characterises *worst-case propagation* of a fixed, externally specified measurement-uncertainty budget through closed-interval arithmetic — a fundamentally different inferential target. HNRA’s advantage is that its intervals are unconditional on distributional or sampling assumptions (Corollary 5);

its limitation is the same worst-case construction that also drives the very wide F_U values in the Uncertainty-Sensitive Region. We recommend HNRA when a worst-case guarantee over a documented measurement-uncertainty budget is required (e.g. regulatory reporting under ISO 5725), and recommend pairing it with a resampling-based interval when sampling variability is the primary concern, so both sources of uncertainty are visible to the analyst. Relative to interval regression, HNRA produces an F -statistic in the conventional form directly comparable against tabulated critical values rather than requiring set-identification machinery; relative to fuzzy regression, its intervals carry a direct arithmetic justification grounded in closed-interval analysis rather than a chosen membership-function family; and relative to Bayesian regression, no prior elicitation or posterior sampling is required, making HNRA immediately deployable with the same computational footprint as classical OLS.

Table 9 Comparative evaluation of regression ANOVA frameworks.

Method	Uncertainty	Interp.	Complexity	Inferential power
Classical OLS	Random error	High	Low	Point estimate only
Interval reg.	Interval Y	Moderate	Mod. (LP)	Set-identified; no std. F -test
Fuzzy reg.	Memb. func-tions	Moderate	Moderate	Possibilistic; subjective param.
Bayesian reg.	Prior + lik.	Moderate	High (MCMC)	Full posterior; prior required
Robust reg.	Heavy tails	High	Moderate	Asymptotic; targets outliers
HNRA	Interval	High	Low	Interval F-test; uniform sig. guarantee

These comparative properties carry implications beyond the specific biomedical application of Section 4. Theoretically, the convergence result of Theorem 2 positions HNRA as a conservative generalisation that subsumes, rather than departs from, established statistical practice, directly addressing the gap identified by Woodall, Driscoll, and Montgomery [20]; methodologically, HNRA demonstrates that interval-valued inference need not sacrifice computational simplicity or the familiar F -distribution-based hypothesis-testing workflow, lowering the adoption barrier relative to Bayesian or interval-programming alternatives, since an analyst already familiar with classical regression ANOVA can compute HNRA quantities as a direct post-processing step. Biomedically, HNRA addresses concrete needs across clinical laboratory medicine, where ISO 5725 and CLSI guidelines require measurement uncertainty to be reported alongside analytical results [6, 7]; wearable-device analytics, where instruments carry documented precision limits [3] that HNRA encodes directly rather than discarding; and epidemiology, where self-reported variables are subject to recall and digit-preference biases manifesting as ± 0.5 to ± 1 unit imprecision [4], accommodated here without imputation or a separate post hoc sensitivity procedure. Relative to the bipolar and interval-valued neutrosophic decision-support literature reviewed in Section 2.2, HNRA extends the underlying philosophy of representing indeterminate information through bounded intervals from multi-criteria aggregation into

frequentist statistical inference, suggesting a complementary use case in which HNRA could furnish statistically validated interval-valued effect-size estimates feeding into a downstream neutrosophic MCDM framework of the kind developed in [32, 33, 36], though formalising this integration is left to future work.

Several limitations merit acknowledgement, several now characterised quantitatively rather than asserted qualitatively. Most significantly, the operating-envelope analysis of Section 3.3 shows that $MSE_L > 0$ is not automatic: the borderline-significance scenarios (Section 4.5) and the simulation grid (Section 4.6) demonstrate that F_U can become very wide in the low-to-moderate signal regime where an uncertainty-sensitive result is most likely and a sharp answer most needed; the κ diagnostic provides a warning but not a remedy, and developing a tighter alternative for the Uncertainty-Sensitive Region is an open problem. Section 5.1 shows HNRA intervals are systematically wider than bootstrap and Bayesian alternatives because they target a worst-case rather than sampling-variability inferential question; the two families of method are complementary rather than competing. HNRA as presented also assumes symmetric uncertainty bounds ($\pm\delta_y|y|$); asymmetric bounds, common for instruments with one-sided detection limits, require an extension of the interval arithmetic in Definition 2 not developed here. The OLS coefficients β_0, β_1 are estimated from midpoint observations rather than the full interval data; an estimator jointly incorporating interval information could yield more statistically efficient point estimates at increased computational and notational complexity. HNRA is developed for simple ($k = 1$) regression, and extension to multiple regression with p predictors requires a corresponding neutrosophic matrix algebra not addressed here. Most significantly for statistical validity, Type-I error control under neutrosophic intervals has not been formally established in this or prior neutrosophic ANOVA work; whether the interval-valued F -test maintains nominal size in repeated frequentist sampling is an open question requiring dedicated power-analysis and Type-I error simulation studies. Finally, the F -interval may become uninformatively wide under large uncertainty, potentially limiting discriminative power for borderline, near-critical-value cases – a limitation visible in principle though not encountered in the present application, where F_L remained far above threshold even at $\delta_y = 25\%$. Future work should accordingly prioritise formal Type-I error and statistical power analysis for the HNRA F -test under repeated sampling, ideally via simulation under the null hypothesis $\beta_1 = 0$; extension to multiple neutrosophic regression with $p > 1$ predictors; development of an interval-valued coefficient of determination and an interval-valued Akaike Information Criterion for neutrosophic model selection; accommodation of asymmetric uncertainty bounds; and direct simulation-based comparison against Bayesian and fuzzy regression under controlled uncertainty-generating processes.

6 Conclusion

This paper introduced Hybrid Neutrosophic Regression ANOVA (HNRA), extending classical simple regression ANOVA to accommodate structural measurement uncertainty encoded as neutrosophic intervals. Four theoretical results were established: a

decomposition theorem showing $SST_N \approx SSR_N + SSE_N$ componentwise under interval arithmetic, a convergence theorem proving exact reduction to classical regression ANOVA as uncertainty vanishes, a monotonicity theorem establishing non-decreasing interval width with respect to the uncertainty parameters, and a boundedness theorem proving strict positivity and finiteness of the interval-valued F -statistic together with a corollary guaranteeing uniform statistical significance whenever the interval lower bound exceeds the classical critical value. Applied to a real biomedical dataset relating muscle mass to age ($n = 60$) under $\pm 10\%$ response and ± 0.5 -year predictor uncertainty, HNRA produced $F_N = [116.52, 260.02]$, an interval lying entirely above the classical critical value; a five-level sensitivity analysis and a 500-replication simulation study together confirmed that this conclusion is stable under measurement imprecision substantially beyond typical biomedical instrument precision, in direct empirical agreement with the proven theoretical properties.

The practical significance of this work lies in furnishing, for the first time, a complete interval-valued regression ANOVA mechanism formally compatible with the measurement-uncertainty reporting requirements of clinical laboratory medicine, wearable-device analytics, and epidemiological research, while requiring no prior distributional specification and adding negligible computational overhead beyond classical OLS. At the same time, the limitations discussed in Section 5 – particularly the unresolved question of Type-I error control under neutrosophic intervals – indicate that HNRA should currently be regarded as a robust descriptive and sensitivity-analytic tool rather than a fully validated frequentist hypothesis test, pending the power-analysis and error-control studies identified as priorities for future work. Extending HNRA to multiple regression, interval model-selection criteria, and asymmetric uncertainty bounds would further establish its place within the broader landscape of uncertainty-aware statistical inference.

Acknowledgements. [Acknowledgements withheld for blind review.]

Declarations

- Funding: Not applicable.
- Conflict of interest: The authors declare no conflict of interest.
- Ethics approval and consent to participate: Not applicable (secondary analysis of a previously published, de-identified textbook dataset).
- Consent for publication: Not applicable.
- Data availability: The datasets used and/or analysed during the current study are not publicly available as the muscle mass dataset is sourced from Kutner, M.H., Nachtsheim, C.J., Neter, J., and Li, W., *Applied Linear Statistical Models*, 5th edition, McGraw-Hill Irwin (2005), Appendix C (Dataset C.1), which is available through academic library access, but are available from the corresponding author on reasonable request.
- Materials availability: Not applicable.
- Code availability: Computational scripts for the HNRA framework are available upon reasonable request.
- Author contribution: [Author contributions withheld for blind review.]

References

- [1] Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W.: Applied Linear Statistical Models, 5th edn. McGraw-Hill Irwin, New York (2005)
- [2] Joint Committee for Guides in Metrology: Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement (GUM). BIPM, Sèvres (2008)
- [3] Karunanithi, C., Bhargava, A.: Wearable biosensors: assessment of measurement uncertainty in clinical practice. *npj Digital Medicine* **6**, 98 (2023) <https://doi.org/10.1038/s41746-023-00843-4>
- [4] Doll, R., Peto, R.: The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute* **66**(6), 1192–1308 (1981)
- [5] Felson, D.T.: Clinical measurement and biomedical variability in musculoskeletal research. *Nature Reviews Rheumatology* **19**(4), 207–217 (2023) <https://doi.org/10.1038/s41584-023-00932-6>
- [6] International Organization for Standardization: Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions. ISO 5725-1:2023, Geneva (2023)
- [7] Clinical and Laboratory Standards Institute: Measurement Procedure Comparison and Bias Estimation Using Patient Samples. EP09-A3, Wayne, PA (2013)
- [8] Montgomery, D.C., Peck, E.A., Vining, G.G.: Introduction to Linear Regression Analysis, 6th edn. Wiley, Hoboken, NJ (2021)
- [9] Beresteanu, A., Molinari, F.: Asymptotic properties for a class of partially identified models. *Econometrica* **76**(4), 763–814 (2008) <https://doi.org/10.1111/j.1468-0262.2008.00859.x>
- [10] Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. *IEEE Transactions on Systems, Man, and Cybernetics* **12**(6), 903–907 (1982) <https://doi.org/10.1109/TSMC.1982.4308925>
- [11] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian Data Analysis, 3rd edn. CRC Press, Boca Raton, FL (2013)
- [12] Huber, P.J.: Robust estimation of a location parameter. *Annals of Mathematical Statistics* **35**(1), 73–101 (1964) <https://doi.org/10.1214/aoms/1177703732>
- [13] Salama, A., Khaled, O., Mahfouz, K.: Neutrosophic correlation and simple linear regression. *Neutrosophic Sets and Systems* **5**, 3–8 (2014) <https://doi.org/10.5281/>

- [14] Aslam, M., Albassam, M.: Dietary fat and prostate cancer relationship using trimmed regression under uncertainty. *Frontiers in Nutrition* **9**, 799375 (2022) <https://doi.org/10.3389/fnut.2022.799375>
- [15] Aslam, M., Saleem, M.: Neutrosophic test of linearity with application. *AIMS Mathematics* **8**(3), 7981–7995 (2023) <https://doi.org/10.3934/math.2023402>
- [16] Smarandache, F.: Neutrosophic set, a generalization of the intuitionistic fuzzy set. *International Journal of Pure and Applied Mathematics* **24**(3), 287–297 (2005)
- [17] Smarandache, F.: Neutrosophic statistics is an extension of interval statistics, while plithogenic statistics is the most general form of statistics. *International Journal of Neutrosophic Science* **19**(1), 148–165 (2022) <https://doi.org/10.54216/IJNS.190111>
- [18] Atanassov, K.T.: *Intuitionistic Fuzzy Sets: Theory and Applications*. Physica-Verlag, Heidelberg (1999)
- [19] Sherwani, R.A.K., Aslam, M., Raza, M.A., Farooq, M., Abid, M., Tahir, M.: Neutrosophic normal probability distribution – a spine of parametric neutrosophic statistical tests. *Neutrosophic Sets and Systems* **51**, 161–186 (2022) <https://doi.org/10.5281/zenodo.7135325>
- [20] Woodall, W.H., Driscoll, D.C., Montgomery, D.C.: A review and perspective on neutrosophic statistical process monitoring methods. *IEEE Access* **10**, 100456–100462 (2022) <https://doi.org/10.1109/ACCESS.2022.3207188>
- [21] Aslam, M.: Neutrosophic analysis of variance: application to university students. *Complex & Intelligent Systems* **6**(2), 403–407 (2020) <https://doi.org/10.1007/s40747-019-0107-2>
- [22] AlAita, A., Aslam, M.: Analysis of covariance under neutrosophic statistics. *Journal of Statistical Computation and Simulation* **93**(3), 397–415 (2023) <https://doi.org/10.1080/00949655.2022.2101100>
- [23] Miari, M., Anan, M.T., Zeina, M.B.: Neutrosophic two-way anova. *International Journal of Neutrosophic Science* **18**(3), 72–83 (2022)
- [24] Kumar, P., Moazzamigodarzi, M., Rahimi, M.: Neutrosophic analysis of experimental data using neutrosophic Graeco-Latin square design. *Axioms* **13**(8), 559 (2024) <https://doi.org/10.3390/axioms13080559>
- [25] Aslam, M.: Neutrosophic analysis of variance: addressing uncertainty in statistical analysis for artificial intelligence. *International Journal of Data Science and Analytics* (2025) <https://doi.org/10.1007/s41060-025-00798-9>

- [26] Vazquez, M.L., Torres, L.C.: Integrating neutrosophic numbers in regression analysis for enhancing predictive modelling through uncertainty representation. *Plithogenic Logic and Computation* **1**, 120–126 (2024) <https://doi.org/10.61356/j.plc.2024.1258>
- [27] Aslam, M.: Neutrosophic linear model for artificial intelligence: addressing uncertainty in training and testing phases. *International Journal of Data Science and Analytics* (2025) <https://doi.org/10.1007/s41060-025-00792-1>
- [28] Masood, S., Ibrar, B., Shabbir, J., Shokri, A., Movaheedi, Z.: Estimating neutrosophic finite median employing robust measures of the auxiliary variable. *Scientific Reports* **14**, 10255 (2024) <https://doi.org/10.1038/s41598-024-60714-2>
- [29] Aslam, M.: Clinical laboratory medicine measurements correlation analysis under uncertainty. *BMC Medical Research Methodology* **21**, 1 (2021) <https://doi.org/10.1186/s12874-021-01400-z>
- [30] Sherwani, R.A.K., Aslam, M., Saleem, N., Seidu, A., Bilal, M.: An application of neutrosophic statistics: extending relative risk and odds ratios to handle uncertainty in epidemiology and biostatistics. *JP Journal of Biostatistics* **25**(1), 1–15 (2025) <https://doi.org/10.17654/0973514325001>
- [31] Aslam, M.: Chi-square test under indeterminacy: an application using pulse count data. *BMC Medical Research Methodology* **21**, 1 (2021) <https://doi.org/10.1186/s12874-021-01400-z>
- [32] Al-Quran, A., Al-Sharqi, F., Ur Rahman, A., Rodzi, Z.M.: The q -rung orthopair fuzzy-valued neutrosophic sets: axiomatic properties, aggregation operators and applications. *AIMS Mathematics* **9**(2), 5038–5070 (2024) <https://doi.org/10.3934/math.2024245>
- [33] Ahmad, N., Rodzi, Z., Al-Sharqi, F., Al-Quran, A., Lutfi, A., Mohd, Z., Aini, N.: Innovative theoretical approach: bipolar Pythagorean neutrosophic sets (BPNSs) in decision-making. *International Journal of Neutrosophic Science* **23**(1), 249–256 (2024) <https://doi.org/10.54216/IJNS.230122>
- [34] Razak, S.A., Rodzi, Z.M., Ahmad, N., Ahmad, G.: Exploring the boundaries of uncertainty: interval valued Pythagorean neutrosophic set and their properties. *Malaysian Journal of Fundamental and Applied Sciences* **20**(4), 813–824 (2024) <https://doi.org/10.11113/mjfas.v20n4.3482>
- [35] Al-Sharqi, F., Al-Quran, A., Rodzi, Z.M.: Multi-attribute group decision-making based on aggregation operator and score function of bipolar neutrosophic hypersoft environment. *Neutrosophic Sets and Systems* **61**(1), 465–492 (2023)
- [36] Kamari, M.S.M., Rodzi, Z.M., Zainuddin, Z.F., Kamis, N.H., Ahmad, N., Razak, S.A., Al-Sharqi, F.: Economic evaluation of digital suppliers for manufacturing

- SMEs using Pythagorean neutrosophic TOPSIS and VIKOR with a flexible distance metric. *International Journal of Economic Sciences* **14**(1), 351–384 (2025) <https://doi.org/10.31181/ijes1412025187>
- [37] Romdhini, M.U., Al-Sharqi, F., Al-Obaidi, R.H., Rodzi, Z.M.: Modeling uncertainties associated with decision-making algorithms based on similarity measures of possibility belief interval-valued fuzzy hypersoft setting. *Neutrosophic Sets and Systems* **77**, 292–310 (2025)
 - [38] Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W.: *Applied Linear Statistical Models*, 5th edn. McGraw-Hill Irwin, New York (2005). Muscle mass dataset: Appendix C, Dataset C.1 (muscle mass and age, $n = 60$)
 - [39] Aslam, M., Albassam, M.: Neutrosophic geometric distribution: data generation under uncertainty and practical applications. *AIMS Mathematics* **9**(6), 16436–16452 (2024) <https://doi.org/10.3934/math.2024796>
 - [40] Beaudart, C., Zaaria, M., Pasleau, F., Reginster, J.-Y., Bruyère, O.: Health outcomes of sarcopenia: a systematic review and meta-analysis. *PLOS ONE* **12**(1), 0169548 (2017) <https://doi.org/10.1371/journal.pone.0169548>