

Javier REIG-MULLOR, PhD

javier.reig@umh.es

Miguel Hernandez University, Elche, Spain

Francisco SALAS-MOLINA, PhD (corresponding author)

frasamo@upv.es

Technical University of Valencia, Alcoi Campus, Spain

An Integrated Approach to Multiple Criteria Evaluation of Short-term Bank Performance

Abstract. *This paper addresses the critical impact of imprecision and vagueness in the evaluation of bank performance, with a particular focus on the information in quarterly reports to enable a more dynamic, timely assessment. We propose a new ranking function for type-2 neutrosophic numbers to address uncertainties in membership functions, relying on the concepts of possibilistic mean and standard deviation. Incorporating the mean and standard deviation into scoring functions provides a more informative basis for ranking than relying solely on the mean, reducing information loss. We illustrate our approach by evaluating the performance of major Spanish banks over 2020-2022, demonstrating its applicability and effectiveness in assessing the evolving dynamics of the banking sector. In this case study, we combine the new ranking functions with three well-known multi-criteria MCDM methods to demonstrate their consistency across methods.*

Keywords: *computational finance, fuzzy sets, bank performance, multiple criteria decision making, seasonality.*

JEL Classification: C44, C61.

Received: 22 May 2026

Revised: 8 September 2026

Accepted: 11 September 2026

1. Introduction

The impact of imprecision and vagueness is especially significant in economics and finance. Banks are fundamental to the efficiency and stability of the economy by facilitating the flow of money, channeling investments, providing essential financial services, and supporting the implementation of economic policies. Although banks play a crucial role in the economy, they can also represent significant risks (credit, liquidity, interest rate, and exchange rate). The inherent risks in banks cannot be eliminated but can be managed and mitigated through effective regulation, prudent supervision, and appropriate risk management practices. At the international level, common standards and principles have been established for banking regulation, led by organizations such as the Committee on Banking Supervision (BIS), the Securities and Exchange Commission (SEC), and the Prudential Regulation Authority (PRA). Most studies evaluating bank solvency have

focused on the information these entities publish annually. However, quarterly reports have the advantage of providing more frequent updates, enabling more regular monitoring, and collecting seasonal features.

To set evaluation criteria, regulatory bodies use the CAMELS (Capital Adequacy, Asset Quality, Management Quality, Earnings, Liquidity, and Sensitivity) rating system, which assesses capital adequacy, asset quality, management, earnings, liquidity, and sensitivity. The CAMELS rating system is based on a set of financial ratios and appears to have significant capacity to detect distress (Wanke et al., 2022). These criteria are based on a set of financial ratios, and several studies have used this methodology, such as Cole and Gunther (1995); DeYoung (1998); Kumar and Ravi (2007); Oshinsky and Olin (2006); Poghosyan and Cihak (2011); Ravisankar et al. (2010), and Reig-Mullor and Brotons-Martinez (2021). In addition, there is a need to include non-financial indicators in evaluating banks' performance. The inclusion of environmental, social, and governance (ESG) indicators is important because sustainability best practices and the integration of the environmental and social dimensions can improve the ability to predict financial distress (Citterio & King, 2023) and create long-term value for all stakeholders (Nizam et al., 2019). According to Nobanee and Ellili (2016), banks' compliance with best practices in sustainability disclosure and the integration of environmental and ecological dimensions in annual reports indicate the banks' commitment to increasing transparency and reducing information asymmetry and costs related to debt financing.

Evaluating bank performance is inherently uncertain. Banks operate in environments shaped by incomplete financial data, subjective managerial assessments, volatile markets, and conflicting expert opinions. In this case, we require specialized tools to identify appropriate solutions to economic and financial problems (Oderanti & De Wilde, 2010). Some recent examples of the application of fuzzy set theory to the assessment of financial performance include Mesgarani et al. (2024) and Sahil, Kaushal, and Lohani (2025). Since the pioneering work of Zadeh (1965), fuzzy set theory has been widely employed across various research fields involving uncertainty (Kimiagari & Keivanpour, 2019; Q.S. Wang et al., 2019). Later, Zadeh (1975) introduced the concept of a type-2 fuzzy set to incorporate uncertainty in the membership function into fuzzy set theory. Neutrosophic logic was introduced by Smarandache (1999) to handle uncertainty, imprecision, and indeterminacy more flexibly.

In this paper, we focus on trapezoidal type-2 neutrosophic sets as a suitable means of evaluating bank performance. Traditional fuzzy systems attempt to model uncertainty, but they still impose structural constraints that limit their expressiveness in real-world financial decision-making. For instance, type-2 fuzzy logic is still tied to a single dimension, namely the degree of truth. By contrast, neutrosophic logic is explicitly designed to handle truth, indeterminacy, and falsity. In addition, bank performance indicators are usually expressed using the maximum, minimum, and modal values. Ranking functions play a crucial role in Type-2 Neutrosophic Number (T2NN) applications, particularly in Multi-Criteria Decision-Making (MCDM)

problems. In essence, score functions act as a bridge, allowing MCDM methods that handle the uncertainty and vagueness inherent in human evaluations to culminate in a clear ranking of alternatives through quantitative calculation.

Summarizing, we present a novel methodology for evaluating bank performance using quarterly financial information published by these entities. We first propose a new score function based on the mean and standard deviation. We extend the CAMELS analysis to integrate ESG indicators. To address the inaccuracy and reliability of the available information, we use a type-2 neutrosophic logic methodology to rank alternatives, which can be easily integrated with multicriteria methods such as MARCOS (Stevic et al., 2020), WASPAS (Zavadskas et al., 2012), and EDAS (Keshavarz et al., 2015). To illustrate our novel approach, we apply the procedure to evaluate the performance of the main Spanish banks during 2020-2022.

In addition to this introduction, Section 2 provides background on fuzzy sets. Section 3 proposes two novel type-2 neutrosophic score functions. Section 4 illustrates our approach through a performance evaluation of major Spanish banks. Finally, Section 5 concludes by highlighting natural extensions of this work.

2. Preliminaries

This section describes basic concepts of neutrosophic fuzzy sets. Since the pioneering work by Zadeh (1965), the theory of fuzzy sets has been developed in several ways. One of them is the proposal of neutrosophic fuzzy sets by Smarandache (1999).

Definition 1. Fuzzy Set. Let X be a non-empty reference set contained in the real numbers. Fuzzy set A is defined by its membership function $\mu_A: X \rightarrow [0,1]$. For any element $x \in X$, the value $\mu_A(x)$ represents the degree of membership of x in the fuzzy set A . Furthermore, a fuzzy set has the following properties:

- Support: $\text{supp}(A) = \{x \in X \mid \mu_A(x) > 0\}$
- Core: $\text{core}(A) = \{x \in X \mid \mu_A(x) = 1\}$
- Height: $h(A) = \sup_{x \in X} \mu_A(x)$. Fuzzy set A is *normal* if $h(A) = 1$
- α -cut for $\alpha \in (0,1]$: $A_\alpha = \{x \in X \mid \mu_A(x) = \alpha\}$.

Definition 2. Fuzzy Number. A fuzzy set A on the real line is called a fuzzy number if: 1) A is normal; 2) for every $\alpha \in (0,1]$, the α -cut A_α is a closed interval; and 3) the support of A is bounded.

Definition 3. Neutrosophic Set. (H. Wang et al., 2010). A neutrosophic set A in a space X is characterized by three membership functions: a truth-membership function $T_A(x)$; an indeterminacy membership function $I_A(x)$, and a falsity membership function $F_A(x)$. Each function maps elements of X to values in the interval $(0^-, 1^+)$.

Definition 4. Single-Valued Neutrosophic Set (SVNS) (H. Wang et al., 2010). A single-valued neutrosophic set is a special case where:

- $T_A(x)$, $I_A(x)$, and $F_A(x)$ all take values in $[0,1]$
- They satisfy: $0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3$.

Definition 5. Single-Valued Triangular Neutrosophic Numbers (Khatter, 2020). An SVNN $\tilde{A} = (a_1, a_2, a_3; w_{\tilde{A}}, y_{\tilde{A}}, u_{\tilde{A}})$ is called a single-valued triangular neutrosophic number (SVTNN) when its truth, indeterminacy, and falsity membership functions are linear and defined piecewise over the interval $[a_1, a_3]$.

- **Truth-membership function $T_{\tilde{A}}(x)$:**

$$T_{\tilde{A}}(x) = \begin{cases} \frac{x - a_1}{a_2 - a_1} w_{\tilde{A}} & \text{if } a_1 \leq x < a_2, \\ w_{\tilde{A}} & \text{if } x = a_2, \\ \frac{a_3 - x}{a_3 - a_2} w_{\tilde{A}} & \text{if } a_2 < x \leq a_3, \\ 0 & \text{otherwise.} \end{cases}$$

- **Indeterminacy-membership function $I_{\tilde{A}}(x)$:**

$$I_{\tilde{A}}(x) = \begin{cases} \frac{a_2 - x + (x - a_1)y_{\tilde{A}}}{a_2 - a_1} & \text{if } a_1 \leq x < a_2, \\ y_{\tilde{A}} & \text{if } x = a_2, \\ \frac{x - a_2 + (a_3 - x)y_{\tilde{A}}}{a_3 - a_2} & \text{if } a_2 < x \leq a_3, \\ 1 & \text{otherwise.} \end{cases}$$

- **Falsity-membership function $F_{\tilde{A}}(x)$:**

$$F_{\tilde{A}}(x) = \begin{cases} \frac{a_2 - x + (x - a_1)u_{\tilde{A}}}{a_2 - a_1} & \text{if } a_1 \leq x < a_2, \\ u_{\tilde{A}} & \text{if } x = a_2, \\ \frac{x - a_2 + (a_3 - x)u_{\tilde{A}}}{a_3 - a_2} & \text{if } a_2 < x \leq a_3, \\ 1 & \text{otherwise.} \end{cases}$$

Definition 6. Trapezoidal Type-2 Neutrosophic Numbers. A trapezoidal type-2 neutrosophic number (TrT2NN) $\tilde{\tilde{A}}$ is defined as $\tilde{\tilde{A}} = ((\tilde{A}_1, \tilde{A}_2, \tilde{A}_3); w_{\tilde{\tilde{A}}}, y_{\tilde{\tilde{A}}}, u_{\tilde{\tilde{A}}})$, where $\tilde{A}_1$, $\tilde{A}_2$, and $\tilde{A}_3$ are SVTNNs, and $w_{\tilde{\tilde{A}}}$, $y_{\tilde{\tilde{A}}}$, and $u_{\tilde{\tilde{A}}}$ are global truth, indeterminacy, and falsity degrees. Then a c can be expressed as:

$$\tilde{\tilde{A}} = \left\{ ((a_{11}, a_{12}, a_{13}, a_{14}); w_{\tilde{A}_1}, y_{\tilde{A}_1}, u_{\tilde{A}_1}), ((a_{21}, a_{22}, a_{23}, a_{24}); w_{\tilde{A}_2}, y_{\tilde{A}_2}, u_{\tilde{A}_2}), \right.$$

$$\left. ((a_{31}, a_{32}, a_{33}, a_{34}); w_{\tilde{A}_3}, y_{\tilde{A}_3}, u_{\tilde{A}_3}); w_{\tilde{\tilde{A}}}, y_{\tilde{\tilde{A}}}, u_{\tilde{\tilde{A}}} \right\}$$

where, $w_{\tilde{A}} = \min(w_{\tilde{A}_1}, w_{\tilde{A}_2}, w_{\tilde{A}_3})$, $y_{\tilde{A}} = \max(y_{\tilde{A}_1}, y_{\tilde{A}_2}, y_{\tilde{A}_3})$, and $u_{\tilde{A}} = \max(u_{\tilde{A}_1}, u_{\tilde{A}_2}, u_{\tilde{A}_3})$.

Along the lines of Das et al. (2024) to allow the defuzzification (de-neutrosophication) of the two-fold uncertainty in TrT2NN, we propose

Definition 7. The possibilistic ranking function. Given a TrT2NN, denoted by $\tilde{A}_{Tr}$, the possibilistic ranking function $\Re(\tilde{A}_{Tr})$ is defined as:

$$\Re(\tilde{A}_{Tr}) = \left(\frac{(a_{11} + a_{21} + a_{31})}{3} + \frac{(a_{12} + a_{22} + a_{32})}{3} + \frac{(a_{13} + a_{23} + a_{33})}{3} + \frac{(a_{14} + a_{24} + a_{34})}{3} \right) \left(\frac{1}{4} \right) \left(\frac{w_{\tilde{A}} + y_{\tilde{A}} + u_{\tilde{A}}}{3} \right).$$

3. Two New Possibilistic Type-2 Neutrosophic Ranking Functions

The main limitation of considering only the mean for ranking purposes is the loss of information. To solve this problem, we propose to incorporate the standard deviation adding crucial information about the variability, consistency, and risk associated with that performance. To this end, let us first consider the following definitions.

Definition 8. α -level set (Carlsson & Fuller, 2001). The α -level set of a fuzzy number $\tilde{A}$, for $\alpha \in [0, 1]$, is defined as: $[\tilde{A}]_{\alpha} = \{x \in \mathbb{R} | \mu_{\tilde{A}}(x) \geq \alpha\}$.

Let $\tilde{A}$ be a fuzzy number with: $[\tilde{A}]_{\alpha} = [a_L(\alpha), a_R(\alpha)]$, where $a_L(\alpha)$ and $a_R(\alpha)$ are, respectively, the lower (left) and upper (right) bounds of the α -level set $[\tilde{A}]_{\alpha}$. The possibilistic mean and standard deviation are defined as follows:

Definition 9. Possibilistic mean (Gu & Xuan, 2017). The possibilistic mean of $\tilde{A}$ is the level-weighted average of the arithmetic means of all its α -level sets:

$$M(\tilde{A}) = \frac{1}{w^2} \int_0^w \alpha (a_L(\alpha) + a_R(\alpha)) d\alpha.$$

Definition 10. Possibilistic standard deviation (Gu & Xuan, 2017). The possibilistic standard deviation of $\tilde{A}$ is the square root of the expected value of the squared deviations between the arithmetic mean and the endpoints of its α -level sets:

$$\sigma(\tilde{A}) = \left[\frac{1}{2} \int_0^w \alpha (a_R(\alpha) - a_L(\alpha))^2 d\alpha \right]^{1/2}.$$

For a generalized triangular fuzzy number denoted by

$$\tilde{A} = (a_1, a_2, a_3, w),$$

the possibilistic mean and possibilistic standard deviation reduce to the following expressions (Gu & Xuan, 2017):

$$M(\tilde{A}) = \frac{(a_1 + 4a_2 + a_3)}{6}$$

$$\sigma(\tilde{A}) = \frac{w(a_3 - a_1)}{\sqrt{24}}.$$

In the case of SVTNN, Reig-Mullor et al., (2022) proposed the following possibilistic mean and standard deviation definitions.

Definition 11. Given a single-valued triangular neutrosophic number (SVTNN)

$$\tilde{A} = (a_1, a_2, a_3, w_{\tilde{A}}, y_{\tilde{A}}, u_{\tilde{A}}),$$

the possibilistic means of the truth-membership function $T_{\tilde{A}}(x)$, the indeterminacy-membership function $I_{\tilde{A}}(x)$, and the falsity-membership function $F_{\tilde{A}}(x)$ are respectively defined as follows:

$$M_T(\tilde{A}) = \frac{(a_1 + 4a_2 + a_3)w_{\tilde{A}}}{6}$$

$$M_I(\tilde{A}) = \frac{(a_1 + 4a_2 + a_3)(1 - y_{\tilde{A}})}{6}$$

$$M_F(\tilde{A}) = \frac{(a_1 + 4a_2 + a_3)(1 - u_{\tilde{A}})}{6}.$$

Definition 12. Given a single-valued triangular neutrosophic number (SVTNN)

$$\tilde{A} = (a_1, a_2, a_3, w_{\tilde{A}}, y_{\tilde{A}}, u_{\tilde{A}}),$$

the possibilistic standard deviations of the truth-membership function $T_{\tilde{A}}(x)$, the indeterminacy-membership function $I_{\tilde{A}}(x)$, and the falsity-membership function $F_{\tilde{A}}(x)$ are respectively defined as follows:

$$\sigma_T(\tilde{A}) = \frac{(a_3 - a_1)\sqrt{w_{\tilde{A}}}}{\sqrt{24}}$$

$$\sigma_I(\tilde{A}) = \frac{(a_3 - a_1)\sqrt{1 - y_{\tilde{A}}}}{\sqrt{24}}$$

$$\sigma_F(\tilde{A}) = \frac{(a_3 - a_1)\sqrt{1 - u_{\tilde{A}}}}{\sqrt{24}}.$$

Relying on the concepts of possibilistic and standard deviation, Reig-Mullor et al., (2022) finally proposed the following ranking function for SVTNN.

Definition 13. Given a single-valued triangular neutrosophic number (SVTNN)

$$\tilde{A} = (a_1, a_2, a_3, w_{\tilde{A}}, y_{\tilde{A}}, u_{\tilde{A}}),$$

The possibilistic ranking function is defined as:

$$PS(\tilde{A}) = \frac{(a_1 + 4a_2 + a_3)(2 + w_{\tilde{A}} - y_{\tilde{A}} - u_{\tilde{A}})}{18} + \frac{(a_3 - a_1)(\sqrt{w_{\tilde{A}}} + \sqrt{1 - y_{\tilde{A}}} + \sqrt{1 - u_{\tilde{A}}})}{\sqrt{72}}.$$

In what follows, we extend the definitions of Reig-Mullor et al. (2022) and Das et al., (2024) to propose a new possibilistic neutrosophic ranking function.

Definition 14. The $\mathfrak{R}_P(\tilde{A}_{Tr})$ possibilistic ranking function. Given a TrT2NN, denoted by $\tilde{A}_{Tr}$, the $\mathfrak{R}_P(\tilde{A}_{Tr})$, the possibilistic ranking function is defined as:

$$\mathfrak{R}_P(\tilde{A}_{Tr}) = \left(\frac{\frac{(a_{11}+a_{21}+a_{31})}{3} + 2 \frac{(a_{12}+a_{22}+a_{32})}{3} + 2 \frac{(a_{13}+a_{23}+a_{33})}{3} + \frac{(a_{14}+a_{24}+a_{34})}{3}}{18} \right) \left(\frac{1}{4} \right) (2w_{\tilde{A}} - y_{\tilde{A}} - u_{\tilde{A}}) + \frac{\left(\frac{(a_{14}+a_{24}+a_{34})}{3} - \frac{(a_{11}+a_{21}+a_{31})}{3} \right) \left(\sqrt{w_{\tilde{A}}} + \sqrt{1-y_{\tilde{A}}} + \sqrt{1-u_{\tilde{A}}} \right)}{\sqrt{72}}.$$

The score function combines the possibilistic mean and standard deviation with the truth, indeterminacy, and falsity memberships so that alternatives with higher expected values and greater neutrosophic reliability receive better evaluations. The possibilistic mean is used to represent the central tendency of a type-2 neutrosophic fuzzy number because it captures the overall magnitude of the information. In contrast, the possibilistic standard deviation captures the extent of uncertainty and imprecision contained in the fuzzy assessment through its support width.

Property 1. Monotonicity with respect to the neutrosophic elements. Let

$$m_1 = \frac{a_{11} + a_{21} + a_{31}}{3}, m_2 = \frac{a_{12} + a_{22} + a_{32}}{3},$$

$$m_3 = \frac{a_{13} + a_{23} + a_{33}}{3}, m_4 = \frac{a_{14} + a_{24} + a_{34}}{3},$$

and define the possibilistic mean μ_P and spread s_P :

$$\mu_P = \frac{m_1 + 2m_2 + 2m_3 + m_4}{6},$$

$$s_P = m_4 - m_1$$

Then, $\mathfrak{R}_P(\tilde{A}_{Tr})$ can be written more compactly as:

$$\mathfrak{R}_P(\tilde{A}_{Tr}) = \frac{\mu_P}{12} (2w_{\tilde{A}} - y_{\tilde{A}} - u_{\tilde{A}}) + \frac{s_P}{\sqrt{72}} \left(\sqrt{w_{\tilde{A}}} + \sqrt{1-y_{\tilde{A}}} + \sqrt{1-u_{\tilde{A}}} \right)$$

The proposed possibilistic function $\mathfrak{R}_P(\tilde{A}_{Tr})$ satisfies the monotonicity property with respect to the neutrosophic components. Since the possibilistic mean and dispersion terms are non-negative, the score increases as the truth-membership degree $w_{\tilde{A}}$ increases and decreases as the indeterminacy and falsity degrees $y_{\tilde{A}}$ and $u_{\tilde{A}}$ increase. This behavior can be verified by examining the sign of the partial derivatives of $\mathfrak{R}_P(\tilde{A}_{Tr})$:

$$\frac{\partial \mathfrak{R}_P}{\partial w_{\tilde{A}}} = \frac{2\mu_P}{12} + \frac{s_P}{2\sqrt{72}\sqrt{w_{\tilde{A}}}} > 0$$

$$\frac{\partial \mathfrak{R}_P}{\partial y_{\tilde{A}}} = -\frac{\mu_P}{12} - \frac{s_P}{2\sqrt{72}\sqrt{1-y_{\tilde{A}}}} < 0$$

$$\frac{\partial \Re_p}{\partial u_{\tilde{A}}} = -\frac{\mu_p}{12} - \frac{s_p}{2\sqrt{72}\sqrt{1-u_{\tilde{A}}}} < 0.$$

Property 2. Dominance. Suppose that two type-2 trapezoidal neutrosophic fuzzy numbers $\tilde{A}_{Tr}$ and $\tilde{B}_{Tr}$ satisfy:

$$\mu_p(\tilde{A}_{Tr}) \geq \mu_p(\tilde{B}_{Tr}),$$

$$s_p(\tilde{A}_{Tr}) \geq s_p(\tilde{B}_{Tr}),$$

$$w_{\tilde{A}} \geq w_{\tilde{B}}, y_{\tilde{A}} \leq y_{\tilde{B}}, u_{\tilde{A}} \leq u_{\tilde{B}}.$$

Then, $2w_{\tilde{A}} - y_{\tilde{A}} - u_{\tilde{A}} \geq 2w_{\tilde{B}} - y_{\tilde{B}} - u_{\tilde{B}}$, and, since the square root function is increasing on $[0, 1]$,

$$\begin{aligned} \sqrt{w_{\tilde{A}}} &\geq \sqrt{w_{\tilde{B}}}, \\ \sqrt{1-y_{\tilde{A}}} &\geq \sqrt{1-y_{\tilde{B}}}, \\ \sqrt{1-u_{\tilde{A}}} &\geq \sqrt{1-u_{\tilde{B}}}. \end{aligned}$$

$$\sqrt{w_{\tilde{A}}} + \sqrt{1-y_{\tilde{A}}} + \sqrt{1-u_{\tilde{A}}} \geq \sqrt{w_{\tilde{B}}} + \sqrt{1-y_{\tilde{B}}} + \sqrt{1-u_{\tilde{B}}}.$$

Combining these inequalities with the assumptions $\mu_p(\tilde{A}_{Tr}) \geq \mu_p(\tilde{B}_{Tr})$ and $s_p(\tilde{A}_{Tr}) \geq s_p(\tilde{B}_{Tr})$, we obtain $\Re_p(\tilde{A}_{Tr}) \geq \Re_p(\tilde{B}_{Tr})$. As a result, the proposed score function satisfies the dominance property, because a trapezoidal neutrosophic fuzzy number with a greater possibilistic mean, no smaller possibilistic standard deviation, a higher degree of truth-membership, and lower degrees of indeterminacy and falsity always receives a score that is at least as large as that of the dominated alternative.

4. Case Study

To illustrate our new ranking functions, we next propose combining them with three well-known multi-criteria MCDM techniques. First, we consider MARCOS (Stevic et al., 2020), which relies on testing the alternatives' reference values in relation to ideal values and on an application methodology that is comprehensive, reasonable, and rational. Second, we use WASPAS (Zavadskas et al., 2012), that combines a weighted sum model (WSM) and a weighted product model (WPM) using a parameter, typically restricted to the interval between 0 and 1. Finally, we implement EDAS (Keshavarz et al., 2015) using average solutions to appraise the alternatives.

4.1 Measurement Framework

As mentioned in the introduction, the CAMELS rating system is an acronym commonly used in the banking industry to assess the overall health of a financial

institution. The set of criteria and subcriteria used based on the CAMELS-ESG rating system is summarized in Table 1. Each letter in CAMELS represents a different component. C for Capital Adequacy. This point assesses the bank's capital in relation to its risk profile. Adequate capital is crucial for absorbing losses. A for Asset Quality. This point focuses on the quality of the bank's loans and other assets. It assesses the risk level in the bank's portfolio. M for Management Quality. This point evaluates the quality of the bank's management, including its strategies, governance, and risk management practices. E for Earnings. This point assesses the bank's profitability and the sustainability of its earnings over time. L for Liquidity. This point looks at the bank's ability to meet its short-term obligations without causing significant harm. S for Sensitivity. This point evaluates the bank's exposure to various market risks, including interest rate and foreign exchange risk.

Table 1. CAMELS-ESG Rating System Criteria Indicators

Criteria	Subcriteria	Indicator description
(C)	C1	Core TIER 1 ratio (BIS)
	C2	TIER1 ratio (BIS)
	C3	Regulatory Total Capital ratio (BIS)
	C4	Regulatory Leverage ratio (BIS)
	C5	Equity to assets ratio
(A)	A1	Non-performance loans to gross loan
	A2	Loan loss reserve to non-performance loan
	A3	Loan loss provision to net interest revenue
(M)	M1	Net interest revenue to average assets ratio
	M2	Other operational income to assets ratio
	M3	Non-interest expenses to assets ratio
	M4	Operating leverage
(E)	E1	Return on asset (ROA)
	E2	Return on equity (ROE)
	E3	Return on tangible equity (ROTE)
	E4	Return on risk-weighted assets (RORWA)
	E5	Efficiency ratio
(L)	L1	Net loan to deposits
	L2	Net loan to asset
	L3	Liquid assets to deposits
	L4	Regulatory Liquidity Coverage Ratio (BIS)
	L5	Regulatory Net Stable Funding Ratio (BIS)
(S)	S1	Risk-weighted assets to assets ratio
	S2	Rating
	S3	Probability of Default (EIKON)
(ESG)	ESG1	Environmental
	ESG2	Social
	ESG3	Governance

Source: Own elaboration from the CAMELS rating system (Wanke et al., 2022).

4.2 Ranking of Spanish Banks

Next, we will evaluate the performance of the six main Spanish commercial banks in the Ibex-35 index from 2020 to 2022, as shown in Table 2. This evaluation will consider the financial and sustainability criteria outlined in the expanded analysis. The CAMELS-ESG rating system is used to quantify the model's performance, as it includes non-financial criteria such as sustainability.

Table 2. Largest commercial banks in Spain in terms of net interest income 2022
(Figures in millions of EUR)

Name of Bank	Commercial Name	Net Interest Income (EUR mn)
Banco Santander S.A.	SANTANDER	38,619
Banco Bilbao Vizcaya Argentaria S.A.	BBVA	19,152
CaixaBank S.A.	CAIXABANK	6,916
Banco Sabadell S.A.	SABADELL	3,800
Bankinter S.A.	BANKINTER	1,537
Unicaja Banco S.A.	UNICAJA	1,058

Source: Bank of Spain (www.bde.es).

Although geographically concentrated in Spain, the selected sample holds substantial systemic relevance within the European Union (EU). The six analyzed commercial banks account for over 80% of the domestic market share and are all categorized as Significant Institutions, directly supervised by the European Central Bank (ECB) through the Single Supervisory Mechanism (SSM). Furthermore, the sample includes systemically important global banks (G-SIBs), such as Banco Santander, as well as major European institutions (such as BBVA and CaixaBank) that rank among the largest banks in the Eurozone by total assets. Thus, while the empirical study serves primarily as an illustrative application of the proposed neutrosophic multi-criteria framework, the sample reflects structural dynamics, regulatory standards, and financial performance patterns that are representative of major banking groups operating within the European Monetary Union.

We assume neutrality across the set of criteria in Table 1, so the weight assigned to each subcriterion is the same ($1/28 = 0.0357$). To this end, we consider the period ranging from 2020 to 2022. As a result, the set of alternatives indexed by s is the set of banks, and the subcriteria indexed by q are those summarized in Table 1.

To enhance the efficiency of the CAMELS model, four new indicators are incorporated alongside those already established by various authors. These additions include Operating Leverage (M4), which measures the extent to which a bank can augment operating income through revenue increases. Operating leverage is calculated as the contribution margin divided by net operating income. The Regulatory Liquidity Coverage (L4) Ratio assesses the ratio of high-quality liquid assets (HQLA) to a bank's net cash outflows expected within a 30-day period during a liquidity stress scenario. This metric is used to evaluate liquidity risks and illustrates the extent to which a company's liquid assets can cover its immediate cash

obligations. Basel III agreements mandate that this ratio exceed 100%. The Regulatory Net Stable Funding Ratio (L5) is the ratio of available stable funding to required stable funding, calculated using the liquidity characteristics of assets over a one-year horizon. This ratio demonstrates a company's ability to finance long-term assets with minimal reliance on stable liabilities. Basel III agreements also stipulate this ratio to exceed 100%. Finally, the Probability of Default indicator (S3), developed by Refinitiv Eikon, computes the probability of default. This indicator derives from 25 financial ratios organized into five distinct components: profitability, leverage, coverage, liquidity, and growth and stability.

The precision and reliability of the information gathered in each of the 28 subcriteria vary. In addition, there is the issue of the information's origin: some data come from the financial statements of banks subject to regulatory supervision, while other data, such as the subcriterion Rating, come from private entities (credit rating agencies). Data for subcriteria such as Probability of Default and ESG originate from the Refinitiv Eikon database, in both cases without any supervision, verification, or validation.

The diversity in the data underlying each indicator creates uncertainty about the model's validity when all indicators are managed uniformly. To overcome this drawback, a model based on a neutrosophic environment is proposed, enabling the individualized management of the quality of the different data that constitute each indicator. The proposed model aims to quantify the reliability and imprecision of the information collected for each of the 28 subcriteria, reflecting the information quality under the supervision of subcriterion data. Without loss of generality and for illustrative purposes, we consider the following quantification:

- Supervision: $(w_{\tilde{r}_q}, y_{\tilde{r}_q}, u_{\tilde{r}_q}) = (0.9, 0.1, 0.1)$
- Not supervision: $(w_{\tilde{r}_q}, y_{\tilde{r}_q}, u_{\tilde{r}_q}) = (0.7, 0.2, 0.3)$

Additionally, managing quarterly data independently within each fiscal year is challenging. To address this point, we propose using TrT2NN. Our model considers data from three periods: 2020, 2021, and 2022. Table 3 displays alternative rankings using the TrT2NN ranking function $\Re(\tilde{A}_{Tr})$ in Definition 7, and the ranking obtained by applying the proposed ranking function $\Re_P(\tilde{A}_{Tr})$ in Definition 14.

Based on the results in Table 3, we derive the following conclusions. First, for both ranking functions, there is a notable consistency in the top-ranked banks. By computing the Spearman's correlation coefficient for any pair of sets of ranks, we find that all values are above 0.99. More precisely, BBVA consistently ranks first across all three MCDM methods and ranking functions. This fact suggests that BBVA is the most efficient according to these assessments. While consistently ranked third with the $\Re(\tilde{A}_{Tr})$ ranking function, the rank of Santander drops to fifth with the $\Re_P(\tilde{A}_{Tr})$ ranking function across all methods. This indicates that the choice of ranking function significantly impacts Santander's perceived efficiency.

Caixabank, Sabadell, Bankinter, or Unicaja present mixed results. However, for most banks, the rankings within each ranking function are largely consistent across the EADS, MARCOS, and WASPAS methods. This suggests that while the ranking function can influence the ranking, the chosen MCDM method often leads to similar relative efficiencies.

Table 3. Benchmarking of Spanish banks using two different ranking functions

Bank	$\mathfrak{R}(\tilde{A}_{Tr})$			$\mathfrak{R}_p(\tilde{A}_{Tr})$		
	EADS	MARCOS	WASPAS (0.5)	EADS	MARCOS	WASPAS (0.5)
SANTANDER	3	3	3	5	5	5
BBVA	1	1	1	1	1	1
CAIXABANK	2	2	2	3	2	2
SABADELL	6	5	5	6	4	4
BANKINTER	5	6	6	4	6	6
UNICAJA	4	4	4	2	3	3

Source: Authors' own creation.

The most striking implication is that the perceived efficiency of certain banks varies with the ranking function used. For instance, Santander consistently ranks third with the $\mathfrak{R}(\tilde{A}_{Tr})$ function but drops to fifth with $\mathfrak{R}_p(\tilde{A}_{Tr})$. The divergent results require discussion and justification for selecting one ranking function over another. Decision-makers must clearly articulate why one ranking function is more appropriate than the other for a given context. This often depends on the specific objectives of the ranking and the nature of available data. However, in the context a comprehensive evaluation of banks, where stability, consistency, and risk are paramount, a ranking function $\mathfrak{R}_p(\tilde{A}_{Tr})$ that incorporates both the mean and standard deviation provides a far richer and more informative basis for ranking than one relying solely on the mean as in the case of $\mathfrak{R}(\tilde{A}_{Tr})$.

Table 4 presents the findings of a sensitivity analysis assessing how variations in the weights assigned to seven core categories (Capital, Asset Quality, Management, Earnings, Liquidity, Sensitivity to Market Risk, and ESG) influence the overall ranking of Spanish banks using the MARCOS method and ranking function $\mathfrak{R}_p(\tilde{A}_{Tr})$. The stability of the rankings was evaluated by computing Spearman's rank correlation coefficient between each of the 15 alternative weighting schemes (SET 1 to SET 15) and a baseline set (SET*), which assigns an equal weight to every subcriterion (1/28). Overall, the results demonstrate a high level of robustness across most scenarios, with Spearman correlation values remaining predominantly strong (0.714).

SET 2 (C) and SET 5 (E) both achieved a perfect correlation (1.000), indicating that high weighting towards financial soundness and profitability produces rankings identical to the SET*. SET 8 represents the main divergence (0.429). Elevating the ESG category weight to 0.40 significantly alters the final rankings, showing that non-

financial ESG criteria prioritize different performance dimensions compared to traditional CAMELS metrics. When category weights are moderately skewed (0.25) weight in SETs 9–15, rankings remain stable, consistently yielding correlations between 0.771 and 0.829.

Table 4. Sensitivity analysis results. Spearman correlation (r_s) with baseline SET* for $\mathfrak{R}_p$ (MARCOS)

SET	WEIGHTING SCHEME	r_s
SET 1	C=0.143, A=0.143, M=0.143, E=0.143, L=0.143, S=0.143, ESG=0.143	0.771
SET 2	C=0.4, A=0.1, M=0.1, E=0.1, L=0.1, S=0.1, ESG=0.1	1.000
SET 3	C=0.1, A=0.4, M=0.1, E=0.1, L=0.1, S=0.1, ESG=0.1	0.829
SET 4	C=0.1, A=0.1, M=0.4, E=0.1, L=0.1, S=0.1, ESG=0.1	0.714
SET 5	C=0.1, A=0.1, M=0.1, E=0.4, L=0.1, S=0.1, ESG=0.1	1.000
SET 6	C=0.1, A=0.1, M=0.1, E=0.1, L=0.4, S=0.1, ESG=0.1	0.829
SET 7	C=0.1, A=0.1, M=0.1, E=0.1, L=0.1, S=0.4, ESG=0.1	0.771
SET 8	C=0.1, A=0.1, M=0.1, E=0.1, L=0.1, S=0.1, ESG=0.4	0.429
SET 9	C=0.25, A=0.125, M=0.125, E=0.125, L=0.125, S=0.125, ESG=0.125	0.771
SET 10	C=0.125, A=0.25, M=0.125, E=0.125, L=0.125, S=0.125, ESG=0.125	0.829
SET 11	C=0.125, A=0.125, M=0.25, E=0.125, L=0.125, S=0.125, ESG=0.125	0.829
SET 12	C=0.125, A=0.125, M=0.125, E=0.25, L=0.125, S=0.125, ESG=0.125	0.771
SET 13	C=0.125, A=0.125, M=0.125, E=0.125, L=0.25, S=0.125, ESG=0.125	0.771
SET 14	C=0.125, A=0.125, M=0.125, E=0.125, L=0.125, S=0.25, ESG=0.125	0.771
SET 15	C=0.125, A=0.125, M=0.125, E=0.125, L=0.125, S=0.125, ESG=0.25	0.771

Source: Authors' own creation.

5. Conclusions

In conclusion, this paper introduces two novel ranking functions for evaluating bank performance, specifically addressing the challenges of imprecision and vagueness inherent to economic and financial analyses. The study provides a more nuanced and accurate assessment of bank performance by focusing on quarterly financial information, incorporating ESG indicators, and using two novel ranking functions adapted to type-2 neutrosophic fuzzy set theory.

The novel approach captures short-term fluctuations and seasonal peculiarities, enhancing the evaluation framework by integrating environmental, social, and governance dimensions. The adoption of type-2 neutrosophic logic methodology further acknowledges and accommodates the inherent uncertainties and inaccuracies in available information. Because of information loss in the defuzzification process, we consider that incorporating both the mean and standard deviation into scoring functions provides a far richer and more informative basis for ranking than one that relies solely on the mean. This statement is based on the fact that more information will be preserved, thereby improving the quality of the ranking results.

Future extensions of this work could explore additional dimensions for evaluating banks, such as incorporating advanced machine learning techniques for predictive analytics and expanding the sample to include a broader range of international banks, thereby addressing limitations in sample size and external validity. Additionally, the methodology could be adapted to assess the impact of

external factors, such as economic downturns or global events, on bank performance. Furthermore, ongoing refinements to incorporate ESG indicators and explore additional factors that enhance the robustness of the evaluation process could enhance the proposed methodology's overall effectiveness and applicability.

References

- [1] Abdel-Basset, M., Saleh, M., Gamal, A., Smarandache, F. (2019), *An approach of TOPSIS technique for developing supplier selection with group decision making under type-2 neutrosophic number*. *Applied Soft Computing*, 77, 438-452.
- [2] Betz, F., Oprüca, S., Peltonen, T.A., Sarlin, P. (2014), *Predicting distress in European banks*. *Journal of Banking and Finance*, 45, 225-241.
- [3] Carlsson, C., Fullér, R. (2001), *On possibilistic mean value and variance of fuzzy numbers*. *Fuzzy Sets and Systems*, 122, 315-326.
- [4] Citterio, A., King, T. (2023), *The role of environmental, social, and governance (ESG) in predicting bank financial distress*. *Finance Research Letters*, 51, 103411.
- [5] Cole, R.A., Gunther, J.W. (1995), *Separating the likelihood and timing of bank failure*. *Journal of Banking and Finance*, 19, 1073-1089.
- [6] Das, S.K., Vincent, F.Y., Roy, S.K., Weber, G.W. (2024), *Location–allocation problem for green efficient two-stage vehicle-based logistics system: A type-2 neutrosophic multi-objective modeling approach*. *Expert Systems with Applications*, 238, 122174.
- [7] DeYoung, R. (1998), *Management quality and X-inefficiency in national banks*. *Journal of Financial Service Research*, 13, 5-22.
- [8] Doumpos, M., Zopounidis, C. (2010), *A multicriteria decision support system for bank rating*. *Decision Support Systems*, 50, 55-63.
- [9] Gu, Q., Xuan, Z. (2017), *A new approach for ranking fuzzy numbers based on possibility theory*. *Journal of Computational and Applied Mathematics*, 309, 674-682.
- [10] Karaaslan, F., Hunu, F. (2020), *Type-2 single-valued neutrosophic sets and their applications in multi-criteria group decision making based on TOPSIS method*. *Journal of Ambient Intelligence and Humanized Computing*, 11, 4113-4132.
- [11] Keshavarz Ghorabae, M., Zavadskas, E.K., Olfat, L., Turskis, Z. (2015), *Multi-criteria inventory classification using EDAS*. *Informatica*, 26, 435-451.
- [12] Khatter, K. (2020), *Neutrosophic linear programming using possibilistic mean*. *Soft Computing*, 24, 16847-16867.
- [13] Kimiagari, S., Keivanpour, S. (2019), *An interactive risk visualisation tool for large-scale and complex engineering and construction projects under uncertainty and interdependence*. *International Journal of Production Research*, 57, 6827-6855.
- [14] Kumar, P.R., Ravi, V. (2007), *Bankruptcy prediction in banks and firms via statistical and intelligent techniques – a review*. *European Journal of Operational Research*, 180, 1-28.
- [15] Liu, Q., Wu, C., Lou, L. (2018), *Evaluation research on commercial bank counterparty credit risk management based on new intuitionistic fuzzy method*. *Soft Computing*, 22, 5363-5375.

- [16] Maghyereh, A.I., Awartani, B. (2012), *Financial integration of GCC banking markets: A non-parametric bootstrap DEA estimation approach*. *Research in International Business and Finance*, 26, 181-195.
- [17] Mesgarani, H., Aghdam, Y.E., Beiranvand, A., Gómez-Aguilar, J. (2024), *A novel approach to fuzzy based efficiency assessment of a financial system*. *Computational Economics*, 63, 1609-1626.
- [18] Nizam, E., Ng, A., Dewandaru, G., Nagayev, R., Nkoba, M.A. (2019), *The impact of social and environmental sustainability on financial performance: A global analysis of the banking sector*. *Journal of Multinational Financial Management*, 49, 35-53.
- [19] Nobanee, H., Ellili, N. (2016), *Corporate sustainability disclosure in annual reports: Evidence from UAE banks*. *Renewable and Sustainable Energy Reviews*, 55, 1336-1341.
- [20] Oderanti, F.O., De Wilde, P. (2010), *Dynamics of business games with management of fuzzy rules for decision making*. *International Journal of Production Economics*, 128, 96-109.
- [21] Oshinsky, R., Olin, V. (2006), *Troubled banks: Why don't they all fail?* *FDIC Banking Review*, 18, 1-51.
- [22] Poghosyan, C., Cihak, M. (2011), *Distress in European banks: An analysis based on a new data set*. *Journal of Financial Service Research*, 40, 163-184.
- [23] Ravisankar, P., Ravi, V., Bose, I. (2010), *Failure prediction of dotcom companies using neural network-genetic programming hybrids*. *Information Sciences*, 180, 1257-1267.
- [24] Reig-Mullor, J., Brotons-Martinez, J.M. (2021), *The evaluation performance for commercial banks by intuitionistic fuzzy numbers: the case of Spain*. *Soft Computing*, 25, 9061-9075.
- [25] Reig-Mullor, J., Garcia-Bernabeu, A., Pla-Santamaria, D., Vercher-Ferrandiz, M. (2022), *Evaluating ESG corporate performance using a new neutrosophic AHP-TOPSIS approach*. *Technological and Economic Development of Economy*, 28, 1242-1266.
- [26] Sahil, M.A., Kaushal, M., Lohani, Q.D. (2025), *A novel Pythagorean approach based sine-shaped fuzzy DEA model: Assessment of Indian public sector banks*. *Computational Economics*, 65, 1373-1395.
- [27] Seçme, N.Y., Bayraktaroğlu, A., Kahraman, C. (2009), *Fuzzy performance evaluation in Turkish banking sector using analytic hierarchy process and TOPSIS*. *Expert Systems with Applications*, 36, 11699-11709.
- [28] Simić, V., Milovanović, B., Pantelić, S., Pamučar, D., Tirkolaee, E.B. (2023), *Sustainable route selection of petroleum transportation using a type-2 neutrosophic ITARA-EDAS model*. *Information Sciences*, 622, 732-754.
- [29] Smarandache, F. (1999), *A unifying field in logics: Neutrosophic logic*. *American Research Press*, 1-141.
- [30] Stević, Ž., Pamučar, D., Puška, A., Chatterjee, P. (2020), *Sustainable supplier selection in healthcare industries using MARCOS*. *Computers & Industrial Engineering*, 140, 106231.
- [31] Wang, H., Smarandache, F., Zhang, Y., Sunderraman, R. (2010), *Single valued neutrosophic sets*. *Multispace and Multistructure*, 4, 410-413.

- [32] Wang, Q.S., Zhou, J.Z., Huang, K.D., Dai, L., Zha, G., Chen, L., Qin, H. (2019), *Risk assessment and decision-making based on mean-CVaR-entropy for flood control operation of large-scale reservoirs*. *Water*, 11(4).
- [33] Wanke, P., Azad, M.A.K., Barros, C.P., Hadi-Vencheh, A. (2016), *Predicting performance in ASEAN banks: An integrated fuzzy MCDM-neural network approach*. *Expert Systems*, 33, 213-229.
- [34] Wanke, P., Azad, M.A.K., Yazdi, A.K., Birau, F.R., Spulbar, C.M. (2022), *Revisiting CAMELS rating system and the performance of ASEAN banks: A comprehensive MCDM/Z-numbers approach*. *IEEE Access*, 10, 54098-54109.
- [35] Zadeh, L.A. (1965), *Information and control*. *Fuzzy Sets*, 8, 338-353.
- [36] Zadeh, L.A. (1975), *The concept of a linguistic variable and its application to approximate reasoning*. *Information Sciences*, 8, 199-249.
- [37] Zavadskas, E.K., Turskis, Z., Antucheviciene, J., Zakarevicius, A. (2012), *Optimization of weighted aggregated sum product assessment*. *Elektronika ir Elektrotechnika*, 122, 3-6.
- [38] Zhao, H., Sinha, A.P., Ge, W. (2009), *Effects of feature construction on classification performance: An empirical study in bank failure prediction*. *Expert Systems with Applications*, 36, 2633-2644.