Advanced Logistic Systems – Theory and Practice, Vol. 19, No. 3 (2025), pp. 72-85.
https://doi.org/10.32971/als.2025.017

PREDICTIVE QUALITY ASSURANCE FOR LOGISTICS PROVIDER SELECTION: AN AI–MCDM FRAMEWORK

GÁBOR NAGY[1] – SZABOLCS SZENTESI[2]

***Abstract**: Quality assurance has evolved from being a competitive advantage to a baseline requirement in logistics services. Traditional multi-criteria decision-making (MCDM) approaches to selecting logistics service providers (LSPs) rely heavily on historical data and expert judgment, which may be unreliable in today's volatile markets. This paper introduces an AI-enhanced MCDM framework that transforms predictive distributions of key quality indicators - such as on-time delivery, lead-time variability, damage rate, CO_2 emissions, and cost - into risk-adjusted preference scores. These scores are then aggregated using established methods (TOPSIS, PROMETHEE, VIKOR) to generate forward-looking and risk-aware rankings. A PRISMA-guided systematic literature review (2019–2025) highlights three major gaps in the field: limited integration of predictive analytics into MCDM, insufficient treatment of tail risks and data drift, and weak attention to explainability. The proposed framework addresses these shortcomings and is illustrated through a case study that demonstrates how predictive, risk-sensitive rankings can enhance robustness and decision quality in LSP selection.*

***Keywords:** logistics service provider selection, predictive analytics, multi-criteria decision-making, quality assurance, machine learning.*

1. INTRODUCTION

The selection of logistics service providers (LSPs) is a strategic decision with long-term implications for service reliability, customer satisfaction, total landed cost, and the overall competitiveness of supply chains. For decades, evaluation frameworks placed their primary emphasis on cost, delivery speed, and reliability, with multi-criteria decision-making (MCDM) methods widely applied to structure these trade-offs [1]. Yet contemporary supply chains have become more complex and exposed to frequent disruptions that necessitate broader perspectives. In particular, resilience, defined as the ability to maintain or quickly restore performance under disruptions, has become a central concern in an era of systemic fragility highlighted by global risk assessments [2].

In parallel, benchmarking of trade logistics has expanded beyond traditional measures; the World Bank's LPI 2023 integrates large-scale shipment-tracking evidence to capture end-to-end speed and reliability, underscoring the centrality of timeliness and network fluidity in competitive performance [3]. At the same time, sustainability has moved to the forefront of logistics decision-making. Organizations are increasingly required to monitor and reduce their environmental footprint, with standardized approaches such as ISO 14083:2023 and the GLEC Framework v3.0 now providing methodologies for calculating greenhouse gas emissions across transport chains [4, 5]. These practices support compliance with the EU Corporate Sustainability Reporting Directive (CSRD), which requires firms to disclose climate-related impacts under the European Sustainability Reporting Standards (ESRS) [6]. Parallel to resilience and sustainability, digital capability has emerged as a decisive dimension in LSP selection. Providers are now expected to demonstrate robust IT integration,

[1] senior lecturer, University of Miskolc, Institute of Logistics, Hungary
gabor.nagy4@uni-miskolc.hu
[2] senior lecturer, University of Miskolc, Institute of Logistics, Hungary
szabolcs.szentesi@uni-miskolc.hu

data quality assurance, and exception handling capacity, enabled through interoperability standards such as GS1 EPCIS 2.0 and the Digital Container Shipping Association (DCSA) APIs [7, 8]. These capabilities have become particularly important in outsourcing and tendering, as they underpin the reliability of real-time visibility and the integrity of logistics event data.

Nevertheless, despite these shifts, many organizations continue to rely on historical key performance indicators (KPIs) aggregated through multi-criteria decision-making (MCDM) approaches. Recent reviews confirm that LSP selection - even when extended to sustainable or hybrid variants - remains largely grounded in backward-looking evidence, which can be unstable in turbulent environments [9]. The literature on supply chain viability and predictive analytics confirms that decisions based on historical snapshots often fail under such conditions, especially when confronted with concept drift in predictive models and KPI distributions [10, 11]. To address this challenge, the present study introduces an integrated framework that combines predictive quality assurance with established MCDM techniques. The framework is designed to produce risk-aware, auditable, and future-oriented evaluations of logistics service providers by embedding distributional forecasts into preference scoring, incorporating Conditional Value-at-Risk (CVaR) penalties to reflect risk aversion, and strengthening explainability through tools such as SHAP values. In doing so, the approach aligns sourcing strategies not only with traditional cost and service objectives but also with the emerging imperatives of resilience, sustainability, and digital governance [12-14].

2. SYSTEMATIC LITERATURE REVIEW (PRISMA-GUIDED)

To provide a comprehensive foundation for this study, a systematic literature review (SLR) was conducted. The main objective of the review is to synthesize and critically evaluate the existing body of research on logistics provider selection, quality assurance mechanisms, and the integration of advanced decision-making approaches. Particular attention is given to the intersection of artificial intelligence (AI) and multi-criteria decision-making (MCDM) methods, as these techniques are increasingly recognized as enablers of predictive and sustainable quality management. By identifying prevailing methodologies, emerging trends, and critical research gaps, the review establishes a knowledge base that supports the development of an AI–MCDM framework for predictive quality assurance in logistics provider selection. The review process was designed and conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Applying this structured approach ensures methodological rigor, transparency, and reproducibility throughout the stages of identification, screening, eligibility, and inclusion of studies. Following PRISMA not only enhances the credibility and robustness of the findings but also provides a clear and systematic overview of the research landscape. This evidence-based foundation enables the study to advance both academic understanding and practical applications in predictive quality assurance for logistics networks.

2.1. Protocol and search strategy

The review followed the PRISMA 2020 reporting guidance [15]. Searches were conducted in Scopus, Web of Science, IEEE Xplore, and ScienceDirect, complemented by targeted queries in Google Scholar, covering the period 2019–2025 (final search: 31 August 2025). Search strings combined logistics outsourcing, 3PL/LSP selection, and evaluation terms with multi-criteria decision analysis (MCDA/MCDM) techniques (e.g., AHP, TOPSIS, PROMETHEE, VIKOR, fuzzy, neutrosophic), and predictive/AI/ML terms (e.g., "on-time delivery," "delay prediction," "forecast*"). Inclusion criteria were: (i) English language, (ii) peer-reviewed journal articles or high-quality conference papers, (iii) clear focus on LSP/3PL selection/evaluation or logistics-relevant supplier selection, and (iv) methodological content involving MCDA and/or AI/ML forecasting. Exclusion criteria removed editorials, purely conceptual notes, and studies lacking empirical or algorithmic implementation. Such rigorous scoping is consistent with recent systematic reviews in logistics decision-making research [16].

2.2. Screening results

The reviewed studies can be clustered into four representative streams. The first group encompasses hybrid fuzzy MCDA approaches, where classical multi-criteria methods are extended with fuzzy logic to address linguistic ambiguity and imprecision in LSP/3PL selection [17-19]. A second cluster involves works that integrate machine learning into MCDA frameworks, enabling data-driven ranking and evaluation of suppliers and logistics providers [20-22]. A third stream focuses on predictive modelling of delivery delays and logistics KPIs, which generate valuable insights into risk and disruption probabilities, although their results are rarely propagated into MCDA ranking procedures [23-25]. Finally, a fourth category consists of topical and bibliometric reviews of outsourcing and reverse logistics selection, which synthesize trends and provide broader perspectives on the evolution of logistics decision support [26-28]. Taken together, these four research directions indicate a clear shift of the field toward hybrid, data-driven approaches in supplier and LSP evaluation.

2.3. Thematic synthesis

MCDA core: Classical methods such as AHP/ANP for weight derivation and TOPSIS, PROMETHEE, and VIKOR for ranking/compromise remain dominant. Variants incorporating fuzzy, neutrosophic, or rough set extensions are increasingly used to handle linguistic ambiguity and uncertainty [17, 20]. Criteria trends: Beyond cost, reliability, and lead time, recent studies highlight resilience, sustainability metrics (e.g., CO_2 per parcel), and IT/API integration performance as critical dimensions [18, 21, 23].

Predictive uptake: Applications of ML in logistics focus on forecasting on-time delivery (OTD), delay risks, and disruption probabilities. However, probabilistic predictions are rarely propagated into MCDA ranking models, limiting risk-sensitive decision support [22, 24]. Governance and transparency: Few studies explicitly address explainability (e.g., SHAP, LIME) or rank stability analyses, despite increasing attention to decision accountability in sustainable supply-chain governance [25, 27].

2.4. Research gaps

Based on the synthesis of prior studies, four major research gaps can be identified.

The first concerns distribution-to-preference mapping. Existing studies tend to rely on point estimates of KPIs when constructing decision preferences, even though predictive models frequently yield full probability distributions. The methodological challenge of transforming probabilistic forecasts into MCDA-consistent value functions has rarely been addressed, which constrains the ability of current frameworks to support genuinely risk-aware decision-making under uncertainty [19, 22].

A second gap arises in risk-aware aggregation. Popular aggregation methods such as TOPSIS, PROMETHEE, and VIKOR are typically applied to deterministic or expected values, while the integration of tail-risk measures such as Conditional Value-at-Risk (CVaR) remains uncommon. Moreover, systematic calibration of stakeholder risk aversion within these ranking procedures is still underdeveloped. Consequently, existing models often fail to align ranking outcomes with an organization's actual tolerance for service disruptions or SLA violations [21, 24].

The third gap relates to explainable integration. Although explainable AI techniques such as SHAP and LIME are increasingly applied in logistics forecasting, their outputs are rarely connected to MCDA components such as criteria weighting or ranking justification. The lack of end-to-end frameworks that link model interpretability with structured decision-making weakens both auditability and transparency in LSP evaluations, despite growing institutional and regulatory demands for explainable and accountable AI in supply chains [25, 27].

Finally, there is a gap in ensuring robustness under drift. Very few studies explicitly stress-test the stability of rankings under dynamic conditions such as concept drift, demand volatility, or shifting forecast horizons. Techniques like bootstrap resampling, sliding time windows, and horizon sensitivity analysis are largely absent from the literature. This omission is critical, as logistics networks operate in environments where data distributions evolve rapidly, making ranking robustness a prerequisite for reliable and resilient sourcing decisions [23, 28].

3. Model and methods: AI-enhanced MCDM for predictive QA

3.1. Workflow

The proposed framework integrates predictive analytics with multi-criteria decision-making (MCDM) to achieve predictive quality assurance in LSP selection. Fig. 1 outlines the six stages.

This modular design follows contemporary AI-MCDM hybrid frameworks in supply chain decision-making [29].

3.2. Criteria and weights

Representative criteria are selected to capture economic, operational, environmental, and digital performance dimensions:

- Cost per parcel (↓),
- On-time delivery (OTD) rate (↑),
- Lead-time variance (↓),
- Damage rate (↓),
- Flexibility (↑),
- CO_2 emissions per parcel (↓),
- IT/API uptime (↑),
- Exception handling quality (↑).

Weights w_i are derived via AHP/ANP or fuzzy AHP to account for linguistic uncertainty, subject to $\sum_i w_i = 1$. The use of fuzzy pairwise comparisons for logistics evaluation is well established in the literature [30, 31].

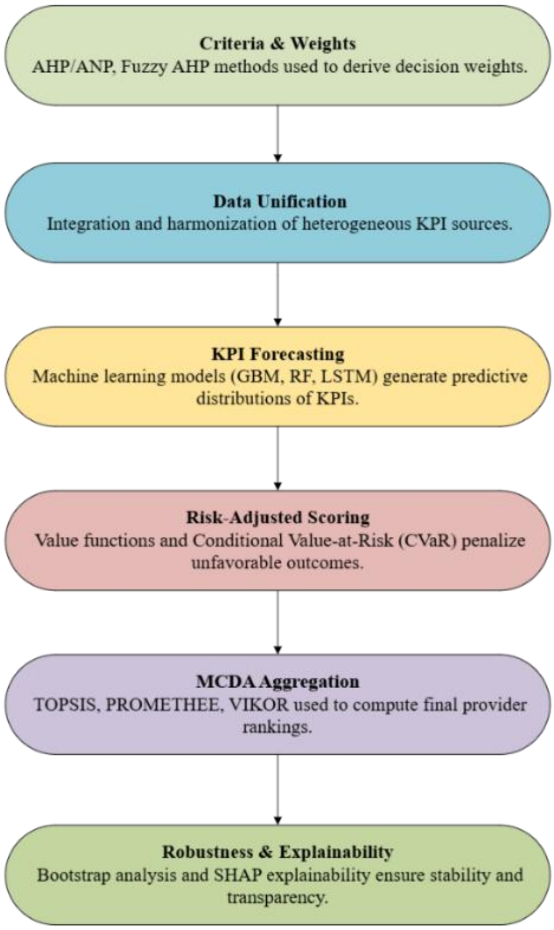

Figure 1. Workflow of the Proposed Predictive QA–MCDM Framework

3.3. KPI forecasting

For each LSP j and criterion i, predictive distributions $\widehat{Y_{ij,t+h}}$ are generated using ensemble forecasters (gradient boosted trees, random forests) for mixed tabular inputs, or sequence models (LSTM, Temporal CNN) for time-series signals. Forecast validation employs rolling-origin evaluation and nested cross-validation, optimizing MAE/RMSE for central tendency and pinball loss for quantiles.

Exogenous drivers (seasonality, promotions, weather, holidays, network changes) are incorporated as covariates, reflecting best practices in predictive logistics modelling [32-34]. Predictive distributions rather than point estimates are emphasized, aligning with calls for distributional forecasting in supply chains [35].

3.4. Value functions and risk adjustment

Forecasted KPI distributions are mapped to [0, 1] preference scores via monotone value functions $g_i(\cdot)$, with lower and upper bounds (L_i, U_i) derived from policy targets or historical percentiles.

- For benefit criteria (1):

$$g_i(y) = \min\left\{1, \max\left\{0, \frac{y - L_i}{U_i - L_i}\right\}\right\}. \tag{1}$$

- For cost criteria (2):

$$g_i(y) = 1 - \min\left\{1, \max\left\{0, \frac{y - L_i}{U_i - L_i}\right\}\right\}. \tag{2}$$

To incorporate risk sensitivity, unfavourable distribution tails are penalized via Conditional Value-at-Risk (CVaR) at level α (3):

$$S_j^{\text{risk}} = \sum_i w_i \ E!\left[g_i!\left(\widehat{Y_{ij}}\right)\right] - \lambda R_j(\alpha) \tag{3}$$

where $R_j(\alpha)$ aggregates per-criterion CVaR penalties and $\lambda \geq 0$ captures stakeholder risk aversion. Tail-risk integration in MCDM has recently been advocated for logistics and financial decision systems [36, 37].

3.5. Aggregation and ranking

Risk-adjusted score matrices are aggregated with three canonical MCDA methods:
- TOPSIS: ranking based on relative closeness to the ideal solution,
- PROMETHEE: net preference flows via pairwise outranking,
- VIKOR: compromise solution balancing group utility and individual regret.

This method-agnostic design ensures compatibility with different organizational decision cultures and risk attitudes [38,3 9].

3.6. Robustness and explainability

To ensure transparency and resilience of the framework:
- Bootstrap resampling of input windows provides rank stability intervals.
- Sensitivity analysis on weights w_i and risk-aversion parameter λ highlights decision leverage points.
- Explainability is achieved through SHAP values linking predictive features (e.g., route length, depot congestion) to forecasted KPIs, documented in model cards for auditability.

This responds directly to the growing demand for explainable and auditable AI in supply chain decision systems [40-42].

4. Illustrative case

To demonstrate the applicability of the proposed framework, we consider a parcel logistics outsourcing scenario involving five competing LSPs operating across a rolling 12-month observation window. The evaluated KPIs comprise:

- On-time delivery (OTD) measured at D+1 and D+2 service levels,
- Damage rate,
- Lead-time variance,
- CO_2 emissions per parcel, and
- Cost per parcel.

Forecasting of KPIs is performed using gradient boosting models with quantile loss, enabling distributional prediction of service quality. Naïve seasonal benchmarks serve as baselines to validate predictive skill, following accepted forecasting evaluation practices in logistics analytics [43, 44]. Prediction horizons span 1-3 months, reflecting tactical sourcing decisions.

Weights (w_i) for the criteria are elicited from three senior logistics managers using the Analytic Hierarchy Process (AHP), ensuring managerial alignment and traceability of preference structures [45]. Risk parameters are set at $\alpha=0.9$ for Conditional Value-at-Risk (CVaR) estimation and λ calibrated against service-level agreement (SLA) tolerances, enabling explicit mapping of organizational risk appetite to source decisions [46].

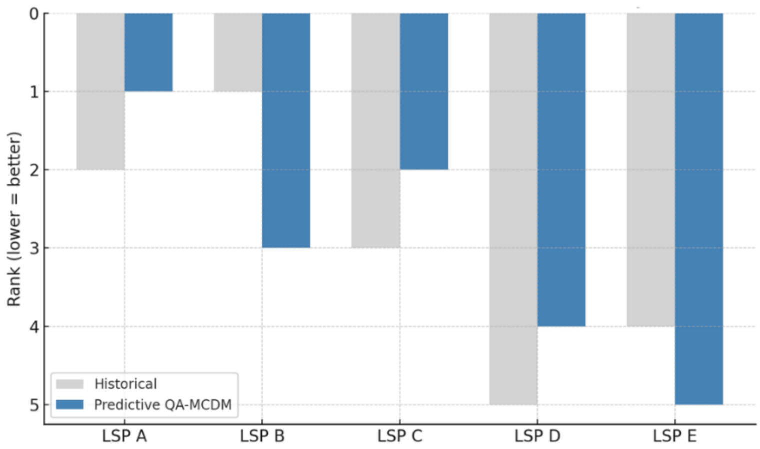

Figure 2. LSP Ranking Shifts: Historical vs Predictive QA-MCDM

Results (illustrative). Fig. 2 presents comparative rankings of LSPs under the proposed predictive QA-MCDM framework versus a purely historical KPI-based ranking. Shifts are observed in two providers: one improves due to predicted OTD gains, while other declines due to forecasted CO_2 intensity deterioration. Fig. 3 displays a tornado sensitivity diagram, showing OTD and cost per parcel as the most influential drivers of ranking volatility. Such visualization supports managerial sense-making in procurement negotiations [47].

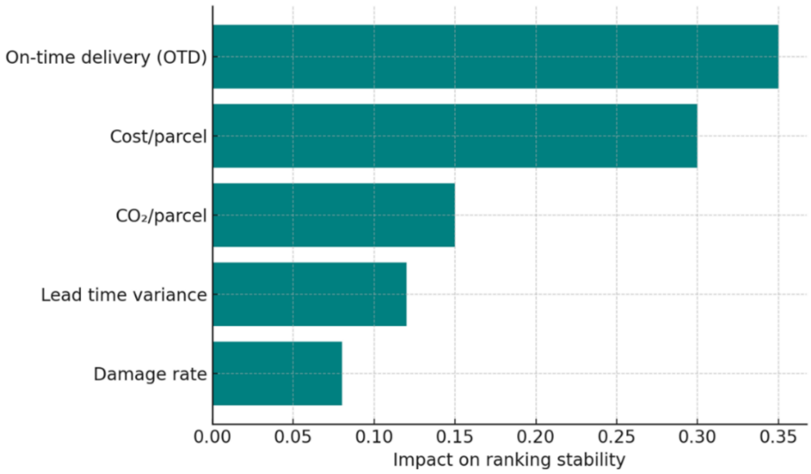

Figure 3. Tornado Sensitivity Analysis of Criteria

5. MANAGERIAL IMPLICATIONS

The proposed AI-enhanced MCDM framework carries substantial managerial implications for procurement, risk management, and sustainability governance in logistics outsourcing. Four key domains are particularly relevant.

First, proactive quality assurance emerges as a major benefit. Traditional LSP selection practices often rely on historical averages of KPIs, which fail to anticipate future deterioration in performance. By embedding predictive models into the evaluation pipeline, procurement managers can base their decisions on anticipated service trajectories rather than backward-looking data. This allows them to proactively exclude providers at risk of declining performance in terms of on-time delivery, damage rates, or CO_2 emissions, thereby protecting customer satisfaction and ensuring continuity of operations. Such anticipatory assurance is consistent with the growing emphasis on supply chain viability, where the focus shifts from reactive recovery to proactive adaptation [48].

Second, the framework supports risk-sensitive contracting. The incorporation of risk aversion parameters (λ, α) into the ranking mechanism enables decision-makers to align sourcing outcomes with their organization's risk appetite. Firms with strict service-level agreements may choose higher λ values to penalize tail-risk events such as extreme delays or excessive variability in lead times. By contrast, organizations that prioritize cost efficiency may adopt more risk-neutral parameter settings. This approach provides a quantifiable, contract-aligned mechanism for embedding risk management policies into supplier selection, moving beyond qualitative assessments of reliability [49].

A third implication concerns transparent governance. By explicitly documenting decision criteria, weighting schemes, and forecasting assumptions, the framework strengthens auditability and compliance in procurement. Such transparency enhances internal governance - through board-level oversight and cross-functional alignment across procurement, operations, and sustainability departments - while also meeting external accountability requirements. Regulatory frameworks such as ISO 9001:2015 (quality management systems) and the EU Corporate Sustainability Reporting Directive (CSRD) explicitly call for auditable and traceable decision processes. Aligning LSP selection with such compliance-oriented governance models reduces exposure to both regulatory penalties and reputational risks [50].

Finally, the framework promotes sustainability alignment. Forecasted CO_2 intensity per parcel provides managers with the ability to integrate carbon-aware considerations into tendering and contracting. By accounting not only for historical but also for projected emissions, procurement decisions can prioritize LSPs whose anticipated environmental performance matches corporate climate targets and disclosure requirements under the European Sustainability Reporting Standards (ESRS). In doing so, procurement becomes directly linked to the firm's ESG agenda, supporting both regulatory compliance and competitive differentiation. For organizations competing in carbon-sensitive markets, such capability functions as a strategic lever for green branding and stakeholder trust [51].

6. LIMITATIONS AND FUTURE WORK

Despite the methodological contributions of this study, several limitations remain that provide fertile ground for future research.

A first limitation relates to data quality and drift. The predictive fidelity of the framework is inherently dependent on the stability and accuracy of input data. Concept drift - where the statistical properties of service processes evolve over time - poses a major threat to model validity. Pandemic shocks altering demand distributions, geopolitical disruptions reshaping routing patterns, or regulatory changes tightening emission limits can all render previously trained models obsolete. Addressing this challenge requires continuous monitoring, re-calibration, and potentially automated drift detection mechanisms [52]. Without such safeguards, the predictive QA component risks degradation, undermining the credibility of supplier rankings.

A second limitation involves the challenge of cold-start LSPs. The framework presumes sufficient historical observations to train reliable KPI forecasters. However, new or niche logistics service providers, often central to innovation-driven markets, lack such records and face the "cold-start" problem. Tackling this issue may require advanced transfer learning techniques that borrow knowledge from related providers or regions, or hierarchical Bayesian models that share statistical strength across similar network nodes [53]. Future work should explicitly test these approaches to ensure that innovative but data-scarce providers are not systematically disadvantaged in sourcing decisions.

Third, the scope of optimization in the current framework is limited by its sequential pipeline-forecast, risk adjustment, and MCDA aggregation. While this modular structure is transparent, it may fail to capture cross-criterion trade-offs holistically. A promising alternative is to embed forecasting outputs directly into a multi-objective optimization (MOO) problem where cost, service quality, and sustainability objectives are jointly optimized under uncertainty. Recent advances in evolutionary multi-objective algorithms and robust stochastic optimization highlight potential directions for such integrated formulations [54].

A further limitation is the reliance on static weighting. At present, the framework applies weights derived via AHP/ANP that capture managerial preferences at a given point in time. Yet criterion salience is not fixed: sustainability metrics may gain prominence under new regulatory regimes, while cost may dominate during recessionary cycles. Incorporating uncertainty-aware dynamic weighting schemes that adapt to predictive distributions or external triggers could enhance resilience and adaptivity. Research on time-varying MCDA models offers valuable building blocks for such developments [55].

Finally, the framework requires cross-industry validation. While the illustrative case in this study focuses on parcel logistics, broader validation is essential to assess generalizability. Application to cold-chain logistics - with its temperature-control risks - bulk commodities, where cost-volume trade-offs dominate, and maritime logistics, characterized by global route volatility, may reveal new dynamics and constraints. Cross-industry empirical studies could stress-test the robustness of the framework, highlight sector-specific extensions, and improve its scalability as a universal procurement decision tool [56].

7. SUMMARY

This paper has presented a modular framework that integrates predictive analytics with multi-criteria decision-making (MCDM) in order to support risk-aware, auditable, and sustainability-aligned logistics service provider (LSP) selection. Whereas conventional approaches predominantly rely on historical KPIs, the proposed methodology emphasizes the anticipation of future service trajectories by incorporating probabilistic KPI forecasts into evaluation and ranking. This predictive orientation represents a significant step forward for quality assurance in logistics outsourcing, since it enables decision-makers to identify potential risks and opportunities before they materialize.

The study contributes conceptually by combining predictive quality assurance (PQA) with structured MCDM techniques, thus bridging two research traditions that have so far developed largely in parallel. Methodologically, the framework operationalizes predictive QA through a six-stage workflow that unifies criteria weighting, data consolidation, KPI forecasting, risk-adjusted preference scoring, aggregation, and robustness analysis. By embedding Conditional Value-at-Risk (CVaR) and explainable AI (e.g., SHAP) within the MCDM pipeline, the model ensures both sensitivity to tail-risk events and interpretability of results. Empirically, the illustrative case of five parcel LSPs demonstrated that predictive quality assurance can substantially alter rankings compared to historical evaluations: anticipated improvements in on-time delivery or forecasted deterioration in CO_2 intensity shifted contracting preferences in ways that would not have been visible using retrospective metrics alone.

Furthermore, sensitivity analysis revealed that on-time delivery and cost exerted the strongest influence on rank stability, providing procurement managers with clear insights into which levers shape sourcing outcomes most decisively. Beyond its methodological and empirical contributions, the framework also advances governance practice by aligning sourcing decisions with regulatory and compliance mandates, including ISO 9001 for quality management and the EU CSRD/ESRS for sustainability disclosure. The explicit articulation of criteria, weights, and forecasts strengthens auditability and transparency, thereby enhancing both internal accountability and external legitimacy. In doing so, the framework positions itself at the intersection of predictive analytics, risk management, and sustainable supply-chain governance, addressing critical gaps such as distribution-to-preference mapping, tail-risk integration, and robustness under concept drift.

Looking ahead, further research should focus on embedding multi-objective optimization directly on predictive distributions to capture trade-offs among cost, service, and sustainability objectives under uncertainty. The incorporation of dynamic weighting schemes that adapt to changing contexts and the extension of empirical validation beyond parcel logistics to sectors such as cold-chain, bulk, or maritime transport would enhance the model's generalizability and resilience. Taken together, the findings demonstrate that predictive quality assurance, when integrated with MCDM, has the potential to redefine LSP selection by aligning anticipatory analytics with structured decision-making, thus embedding resilience, transparency, and sustainability into the strategic core of logistics outsourcing.

References

[1] Abbasi, S., Sıcakyüz, Ç., Santibanez Gonzalez, E. D. & Ghasemi, P. (2024). A systematic literature review of logistics services outsourcing. *Heliyon,* **10**(9), e33374, https://doi.org/10.1016/j.heliyon.2024.e33374

[2] World Economic Forum. (2024). *The Global Risks Report 2024* (19th ed.). World Economic Forum. Retrieved from https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

[3] World Bank. (2023). *Connecting to Compete 2023: Trade logistics in an uncertain global economy (Logistics Performance Index)*. World Bank. Retrieved from https://lpi.worldbank.org/sites/default/files/2023-04/LPI_2023_report.pdf

[4] International Organization for Standardization. (2023). *ISO 14083:2023 — Greenhouse gases — Quantification and reporting of GHG emissions arising from transport chain operations*. ISO. Retrieved from https://www.iso.org/standard/78864.html

[5] Smart Freight Centre. (2023). *Global Logistics Emissions Council (GLEC) Framework v3.0*. Smart Freight Centre. Retrieved from https://www.smartfreightcentre.org

[6] European Commission. (2023). *Corporate sustainability reporting (CSRD) and European Sustainability Reporting Standards (ESRS)*. European Commission. Retrieved from https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en

[7] GS1. (2022). *EPCIS & CBV 2.0 Standards for Supply Chain Event Data*. GS1. Retrieved from https://www.gs1.org/standards/epcis

[8] Digital Container Shipping Association. (2024). *Track & Trace Standard (v2.2)*. DCSA. Retrieved from https://dcsa.org/standards

[9] Karakoç, Ö., Memiş, S., & Sennaroglu, B. (2023). A review of sustainable supplier selection with decision-making methods from 2018 to 2022. *Sustainability, 16*(1), 125. https://doi.org/10.3390/su16010125

[10] Ivanov, D. & Dolgui, A. (2023). Toward supply chain viability theory: From lessons learned through COVID-19 pandemic to viable ecosystems. *International Journal of Production Research,* **61**(8), 2402–2415. https://doi.org/10.1080/00207543.2023.2177049

[11] Hinder, F., Vaquet, V. & Hammer, B. (2024). One or two things we know about concept drift - A survey on monitoring in evolving environments: Part A. *Frontiers in Artificial Intelligence,* **7**, 1330257. https://doi.org/10.3389/frai.2024.1330257

[12] International Organization for Standardization. (2016). *ISO 9001:2015 and risk-based thinking - Implementation guidance (ISO/TC 176/SC 2)*. ISO.

[13] International Organization for Standardization. (2018). *ISO 19011:2018 — Guidelines for auditing management systems*. ISO. Retrieved from https://www.iso.org/standard/70017.html

[14] International Organization for Standardization. (2022). *ISO 8000-1:2022 — Data quality — Part 1: Overview*. ISO. Retrieved from https://www.iso.org/standard/74374.html

[15] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D. & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ,* **372**, n71. Retrieved from https://doi.org/10.1136/bmj.n71

[16] Jahani, H., Jain, R. & Ivanov, D. (2023). Data science and big data analytics: A systematic review of methodologies used in the supply chain and logistics research. *Annals of Operations Research*. Advance online publication. https://doi.org/10.1007/s10479-023-05390-7

[17] Qu, G., Zhang, Z., Qu, W. & Xu, Z. (2020). Green Supplier Selection Based on Green Practices Evaluated Using Fuzzy Approaches of TOPSIS and ELECTRE with a Case Study in a Chinese Internet Company. *International Journal of Environmental Research and Public Health*, **17**(9), 3268. https://doi.org/10.3390/ijerph17093268

[18] Caristi, G., Boffardi, R., Ciliberto, C., Arbolino, R. & Ioppolo, G. (2022). Multicriteria Approach for Supplier Selection: Evidence from a Case Study in the Fashion Industry. *Sustainability*, **14**(13), 8038. https://doi.org/10.3390/su14138038

[19] Aydın, S., Yörükoğlu, M. & Kabak, M. (2021). Fourth party logistics firm assessment using a novel neutrosophic MCDM. *Journal of Intelligent & Fuzzy Systems,* **42**(1), 529–539. https://doi.org/10.3233/JIFS-219209

[20] Sudha, S., Deepak, F. X., Edwin, N. & Martin, N. (2023). *Integrated Machine Learning Algorithms and MCDM Techniques in Optimal Ranking of Battery Electric Vehicles*. In *E3S Web of Conferences* (Vol. 405). https://doi.org/10.1051/e3sconf/202340502005

[21] Agrawal, S., Murthy, P., Kumar, R., Jain, S. & Agarwal, R. (2023). Data-Driven Decision Making in Supply Chain Management. *Innovative Research Thoughts*, **9**(5), 265–281. https://doi.org/10.36676/irt.v9.i5.1487

[22] Aljohani, A., Shafique, M. & Banerjee, A. (2023). Predictive Analytics and Machine Learning for Real-Time Risk Management in Supply Chains. *Sustainability*, **15**(20), 15088. https://doi.org/10.3390/su152015088

[23] Ivanov, D. & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research*, **58**(10), 2904–2915. https://doi.org/10.1080/00207543.2020.1750727

[24] Lu, J., Liu, A., Dong, F., Gu, F., Gama, J. & Zhang, G. (2020). Learning under Concept Drift: A Review. *IEEE Transactions on Knowledge and Data Engineering*, **31**(12), 2346–2363. https://doi.org/10.1109/TKDE.2018.2876857

[25] Ribeiro, M. T., Singh, S. & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

[26] Selviaridis, K. & Spring, M. (2007). Third party logistics: A literature review and research agenda. *International Journal of Logistics Management,* **18**(1), 125–150. https://doi.org/10.1108/09574090710748207

[27] Büyüközkan, G. & Göçer, F. (2018). Digital supply chain: Literature review and a proposed framework for future research. *Computers in Industry,* **97**, 157–177. https://doi.org/10.1016/j.compind.2018.02.010

[28] Govindan, K., Mina, H. & Alavi, B. (2020). A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: A case study of coronavirus disease 2019 (COVID-19). *Transportation Research Part E: Logistics and Transportation Review*, **138**, 101967. https://doi.org/10.1016/j.tre.2020.101967

[29] Stević, Ž., Pamučar, D., Puška, A. & Chatterjee, P. (2020). Sustainable supplier selection in healthcare supply chains using a new MCDM method: Measurement of alternatives and ranking according to compromise solution (MARCOS). *Computers & Industrial Engineering*, **140**, 106231. https://doi.org/10.1016/j.cie.2019.106231

[30] Rezaei, J. (2015). Best–worst multi-criteria decision-making method. *Omega,* **53**, 49–57. https://doi.org/10.1016/j.omega.2014.11.009

[31] Alves de Araújo, F., Mendes dos Reis, J. G., Terra da Silva, M. & Aktas, E. (2022). A Fuzzy Analytic Hierarchy Process Model to Evaluate Logistics Service Expectations and Delivery

Methods in Last-Mile Delivery in Brazil. *Sustainability*, **14**(10), 5753. https://doi.org/10.3390/su14105753
[32] Makridakis, S., Spiliotis, E. & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting,* **38**(4), 1346–1364. https://doi.org/10.1016/j.ijforecast.2021.11.013
[33] Bandara, K., Bergmeir, C. & Hewamalage, H. (2020). Forecasting across large time series databases using recurrent neural networks. *Applied Soft Computing,* **103**, 107872. https://doi.org/10.1016/j.eswa.2019.112896
[34] Seyedan, M. & Tian, X. (2020). *Predictive big data analytics for supply chain demand forecasting*. *Journal of Big Data*, **7**, 66. https://doi.org/10.1186/s40537-020-00329-2
[35] Klein, N., Smith, N. S. & Nott, D. J. (2023). Deep distributional time series models and the probabilistic forecasting framework. *Journal of Applied Econometrics*. **38**(4), 493-511, https://doi.org/10.1002/jae.2959
[36] Rockafellar, R. T. & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk,* **2**(3), 21–42. https://doi.org/10.21314/JOR.2000.038
[37] Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research,* **253**(1), 1–13. https://doi.org/10.1016/j.ejor.2015.12.023
[38] Behzadian, M., Kazemzadeh, R. B., Albadvi, A. & Aghdasi, M. (2010). PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research,* **200**(1), 198–215. https://doi.org/10.1016/j.ejor.2009.01.021
[39] Opricovic, S. & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research,* **156**(2), 445–455. https://doi.org/10.1016/S0377-2217(03)00020-1
[40] Lundberg, S. M. & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems (NeurIPS 2017)* (pp. 4765–4774). Retrieved from https://arxiv.org/abs/1705.07874
[41] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D. & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287596
[42] Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K. & Müller, K.-R. (Eds.). (2019). *Explainable AI: Interpreting, explaining and visualizing deep learning*. Springer. https://doi.org/10.1007/978-3-030-28954-6
[43] Hyndman, R. J. & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. Retrieved from https://otexts.com/fpp3/
[44] Long, X., Bui, Q., Oktavian, G., Schmidt, D. F., Bergmeir, C., Godahewa, R., Lee, S. P., Zhao, K. & Condylis, P. (2023). Scalable Probabilistic Forecasting in Retail with Gradient Boosted Trees: A Practitioner's Approach. *arXiv:2311.00993*. Retrieved from https://arxiv.org/abs/2311.00993
[45] Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences,* **1**(1), 83–98. https://doi.org/10.1504/IJSSCI.2008.017590
[46] Liu, C. & Zhu, W. (2023). Newsvendor conditional value-at-risk minimisation: a feature-based approach under adaptive data selection. *European Journal of Operational Research (accepted manuscript / preprint)*, *arXiv:2209.11121*. Retrieved from https://arxiv.org/abs/2209.11121
[47] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S. (2008). *Global sensitivity analysis: The primer*. Wiley. https://doi.org/10.1002/9780470725184
[48] Ivanov, D. (2021). Supply chain viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation strategies. *International Journal of Production Research,* **59**(12), 3535–3552. https://doi.org/10.1080/00207543.2021.1890852

[49] Chopra, S. & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. *MIT Sloan Management Review,* **46**(1), 53–61. Retrieved from https://sloanreview.mit.edu/article/managing-risk-to-avoid-supplychain-breakdown/

[50] International Organization for Standardization. (2015). *ISO 9001:2015 - Quality management systems - Requirements*. ISO. Retrieved from https://www.iso.org/standard/62085.html

[51] European Commission. (2023). *Commission adopts European Sustainability Reporting Standards (ESRS)*. European Commission. Retrieved from https://finance.ec.europa.eu/news/commission-adopts-european-sustainability-reporting-standards-2023-07-31_en

[52] Peng, G., Selvik, J. T., Abrahamsen, E. B., Bang, K. E. & Markeset, T. (2024). Integrating structure time series forecasting and multicriteria decision analysis for adaptive operational risk assessment: An empirical study using real-time data. *International Journal of System Assurance Engineering and Management*, **15**, 3162–3181. https://doi.org/10.1007/s13198-024-02322-x

[53] Pan, S. J. & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering,* **22**(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

[54] Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. (2002). A fast elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation,* **6**(2), 182–197. https://doi.org/10.1109/4235.996017

[55] Campello, B. S. C., Duarte, L. T. & Romano, J. M. T. (2022). Dealing with multi-criteria decision analysis in time-evolving approach using a probabilistic prediction method. *Engineering Applications of Artificial Intelligence,* **116**, 105462. https://doi.org/10.1016/j.engappai.2022.105462

[56] Christopher, M. & Peck, H. (2004). Building the resilient supply chain. *The International Journal of Logistics Management,* **15**(2), 1–14. https://doi.org/10.1108/09574090410700275