

Article

A Composite Quality Index and Multi-Criteria Decision-Support Framework for Evaluating Complaint Resolution Performance of Express Delivery Operators

Nikola Knežević , Mladenka Blagojević, Aleksandar Čupić, Dragan Lazarević  and Momčilo Dobrodolac * 

Faculty of Transport and Traffic Engineering, University of Belgrade, 11000 Belgrade, Serbia; n.knezevic@sf.bg.ac.rs (N.K.); m.blagojevic@sf.bg.ac.rs (M.B.); a.cupic@sf.bg.ac.rs (A.Č.); d.lazarevic@sf.bg.ac.rs (D.L.)

* Correspondence: m.dobrodolac@sf.bg.ac.rs

Abstract

The rapid expansion of the express delivery sector, driven by the growth of e-commerce, has increased the importance of effective complaint management as an integral component of service quality and consumer protection. Although operational service performance and complaint handling have been widely investigated, existing studies have generally assessed these dimensions separately, leaving a gap in integrated performance evaluation. This study proposes a comprehensive decision-support framework for evaluating complaint resolution performance of express delivery operators by integrating Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), an Entropy-based Composite Quality Index (CQI), and the MARCOS multi-criteria decision-making method. The proposed framework combines objective operational and complaint-related performance indicators within a unified evaluation model. The CQI is constructed by aggregating three complaint resolution indicators using objectively determined Entropy weights, while the MARCOS method is applied to rank operators according to both operational performance and complaint resolution effectiveness. The applicability of the framework is demonstrated through a case study involving five major express delivery operators in the Republic of Serbia using regulatory operational data. The results reveal meaningful differences in operator performance and show that integrating operational service failures with complaint resolution effectiveness provides a more comprehensive evaluation than evaluating either dimension independently. The proposed framework contributes to the literature by integrating operational service quality evaluation, complaint resolution performance, composite quality measurement, and multi-criteria decision-making into a unified evaluation framework. Furthermore, it provides an objective benchmarking tool that can support regulatory oversight, managerial decision-making, and continuous quality improvement in the express delivery sector.



Academic Editor: Tianwen Zhang

Received: 28 June 2026

Revised: 6 August 2026

Accepted: 9 August 2026

Published: 11 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: express delivery services; complaint resolution; service quality; composite quality index; key performance indicators; entropy weighting; MARCOS method; multi-criteria decision-making; benchmarking; postal services

1. Introduction

The courier, express, and parcel (CEP) sector has experienced substantial transformation over the past decade, driven by the rapid expansion of e-commerce, continuous

digitalisation, and changing consumer expectations. The sustained growth of business-to-consumer (B2C) deliveries has significantly increased the operational complexity of express delivery networks while intensifying competition among service providers. In this environment, express delivery operators are expected not only to ensure rapid and reliable shipment transportation but also to provide flexible, transparent, and customer-oriented services capable of meeting increasingly demanding market requirements. Consequently, service quality has become one of the principal determinants of competitiveness and long-term business sustainability within the postal and express delivery industry.

The role of express delivery has also evolved considerably. Traditionally regarded as the final operational stage of the logistics process, last-mile delivery has become an essential component of the overall customer experience. For many online purchases, the delivery operator represents the only direct physical interaction between the retailer and the customer, meaning that delivery performance strongly influences consumers' perceptions of the entire purchasing process. As a result, service quality is no longer evaluated solely through operational indicators such as delivery speed or shipment reliability, but also through communication quality, shipment visibility, delivery flexibility, accessibility of customer support, and the operator's ability to respond effectively when service failures occur.

Despite continuous technological improvements in logistics operations, service failures remain unavoidable in express delivery systems. Delayed deliveries, damaged or lost shipments, unsuccessful delivery attempts, incorrect deliveries, and inaccurate shipment tracking continue to affect customer experience even within highly optimised distribution networks. Such failures generate additional operational costs, reduce customer confidence, and may negatively influence the reputation of both delivery operators and online retailers. Previous research consistently demonstrates that customers evaluate service providers not only according to the frequency of service failures but also according to the effectiveness with which these failures are resolved. Consequently, complaint management has evolved from a reactive customer-service activity into an essential component of service quality management and customer relationship management.

This evolution has also influenced the regulatory environment. Contemporary postal regulators increasingly recognise complaint-resolution performance as an important indicator of consumer protection, organisational accountability, and compliance with quality-of-service obligations. Transparent complaint procedures, timely complaint resolution, effective compensation mechanisms, and consistent reporting practices have therefore become important regulatory requirements in many postal markets. Beyond protecting individual customer rights, complaint statistics provide valuable information for identifying recurring operational deficiencies, monitoring market performance, and supporting evidence-based regulatory decision-making.

The increasing importance of complaint management has been accompanied by significant developments in performance evaluation. Traditionally, performance evaluation has primarily focused on operational efficiency, using indicators related to delivery time, reliability, accessibility, and network performance. More recently, researchers and regulators have expanded evaluation frameworks by incorporating customer-oriented indicators, complaint-related measures, sustainability objectives, digitalisation, and organisational responsiveness. These developments reflect the growing recognition that service quality represents a multidimensional concept that cannot be adequately assessed using isolated operational or customer-oriented indicators alone.

Although considerable progress has been achieved in developing performance measurement systems for postal and express services, existing evaluation approaches continue to exhibit an important limitation. Operational service quality and complaint-

resolution performance are generally assessed independently despite representing closely related dimensions of the same service process. Operational indicators primarily measure the occurrence of service failures, whereas complaint-related indicators evaluate the effectiveness of organisational responses after failures have occurred. Consequently, current evaluation approaches provide only a partial understanding of overall service quality and offer limited support for identifying how operational deficiencies influence complaint-resolution performance.

This limitation is particularly important from both managerial and regulatory perspectives. Delivery operators require reliable performance evaluation frameworks to identify operational weaknesses, allocate resources effectively, improve complaint-handling procedures, and enhance customer satisfaction. At the same time, regulatory authorities require objective and transparent methodologies that enable consistent benchmarking of operators, facilitate market supervision, and support evidence-based regulatory decisions. Addressing these needs requires evaluation frameworks capable of integrating operational service quality with complaint-resolution performance within a unified analytical approach.

The objective of this study is to develop an integrated decision-support framework for evaluating complaint-resolution performance among express delivery operators. The proposed framework combines Service Failure Indicators (SFI) describing operational service quality, Complaint Resolution Indicators (CRI) reflecting the effectiveness of complaint handling, an entropy-based Composite Quality Index (CQI), and the Measurement of Alternatives and Ranking according to Compromise Solution (MARCOS) method within a unified performance evaluation model. By integrating these complementary components, the proposed framework supports comprehensive benchmarking while providing a more complete evaluation of service quality than approaches that evaluate operational performance and complaint management separately.

The proposed framework is empirically validated through a case study involving major express delivery operators in the Republic of Serbia. The selected case provides an appropriate context for evaluation because complaint-resolution performance is systematically monitored using harmonised indicators reported to the national regulatory authority. Although the empirical analysis focuses on the Serbian express delivery market, the proposed methodological framework is sufficiently general to be applied in other regulatory environments where comparable operational and complaint-related performance indicators are available.

The remainder of this paper is organised as follows. Section 2 provides a critical review of the literature on customer protection in postal and express services, complaint management and service quality evaluation, KPI-based performance evaluation, composite indicators, and multi-criteria decision-making methods. Based on this review, the research gaps, research objective, and scientific contribution of the study are identified. Section 3 describes the proposed methodological framework, including the selection of performance indicators, the development of the Composite Quality Index, entropy-based weighting, and the MARCOS decision-making procedure. Section 4 presents the empirical case study and discusses the obtained results. Finally, Section 5 summarises the principal findings, discusses theoretical and practical implications, outlines the limitations of the study, and identifies directions for future research.

2. Literature Review

The evaluation of complaint-resolution performance in express delivery services draws upon several related research streams, including consumer protection, service quality, complaint management, performance measurement, composite indicators, and multi-

criteria decision-making. Although these fields have generated substantial theoretical and methodological knowledge, they have generally developed independently. As a result, the literature provides detailed insights into individual dimensions of service performance but offers limited support for their integration within a unified operator-evaluation framework.

This section critically reviews the literature relevant to the proposed model. Particular attention is devoted to the role of complaint resolution within consumer protection, the relationship between complaint management and service quality, and the limitations of existing evaluation approaches. The review provides the theoretical basis for integrating operational service failures and complaint-resolution performance within the proposed decision-support framework.

2.1. Customer Protection in Postal and Express Services

Customer protection has become one of the fundamental objectives of contemporary postal regulation, particularly in response to the rapid expansion of e-commerce, market liberalisation, digitalisation, and the continuous growth of parcel and express delivery services [1,2]. Traditionally, customer protection in the postal sector focused on ensuring universal service, affordability, territorial accessibility, delivery reliability, and the availability of complaint procedures. However, the evolution of the express delivery market has considerably broadened this concept to encompass service transparency, shipment visibility, compensation mechanisms, return procedures, personal data protection, digital trust, and the overall quality of customer experience [3,4].

Universal postal service remains the cornerstone of customer protection because it guarantees access to essential postal services under predefined quality standards regardless of geographical location or users' socio-economic characteristics [3]. Beyond ensuring service availability, contemporary regulatory frameworks increasingly recognise that effective customer protection also requires coherent interaction between postal regulation, competition policy, and general consumer-protection legislation [4]. Consequently, regulatory attention has gradually shifted from ensuring access to postal services towards safeguarding the quality of the entire service experience.

In this broader context, service quality has become an integral component of customer protection. Customers no longer evaluate postal and express operators solely on the basis of delivery speed but also through delivery reliability, shipment integrity, information availability, price transparency, accessibility of customer support, and the effectiveness with which service failures are resolved [5–9]. The rapid expansion of e-commerce has further reinforced this relationship, as delivery operators increasingly represent the final point of interaction between online retailers and consumers. Consequently, deficiencies in last-mile delivery performance directly influence customer satisfaction, trust, and future purchasing behaviour [7,10].

Customer protection has therefore expanded beyond preventing service failures to ensuring that, when failures occur, they are managed through transparent, accessible, and effective complaint-resolution procedures. Complaints concerning delayed deliveries, damaged or lost shipments, inaccurate tracking information, compensation claims, and inadequate customer support represent valuable sources of information regarding both operational deficiencies and the effectiveness of organisational responses. Recent regulatory analyses conducted by the European Regulators Group for Postal Services (ERGP) demonstrate that significant differences remain among postal markets with respect to complaint-handling procedures, compensation mechanisms, alternative dispute-resolution systems, and reporting practices, highlighting the continuing importance of regulatory oversight in this area [11,12].

Digital transformation has introduced new opportunities for strengthening consumer protection. Electronic shipment tracking, mobile applications, automated notifications, parcel lockers, and autonomous delivery technologies have improved service accessibility, operational efficiency, and customer convenience [13,14]. At the same time, these innovations have generated new challenges related to personal data protection, transparency, digital inclusion, and consumer trust, requiring regulators and service providers to balance technological innovation with effective protection of customer rights [15,16].

The literature therefore demonstrates that customer protection has evolved from a relatively narrow regulatory concept centered on universal service towards a multidimensional framework encompassing operational quality, transparency, digital trust, complaint management, and customer experience. Nevertheless, existing studies predominantly examine these dimensions separately. In particular, complaint-handling performance is commonly evaluated independently of the operational service failures that generate customer complaints. As a result, current approaches provide only a partial understanding of service quality and offer limited support for comprehensive benchmarking of express delivery operators.

The reviewed studies consistently indicate that complaint handling has become an essential component of customer protection rather than merely an administrative obligation. However, the literature still lacks an integrated perspective that simultaneously considers operational service failures and complaint-resolution performance as complementary dimensions of service quality. This limitation provides the theoretical foundation for the integrated evaluation framework proposed in this study.

2.2. Complaint Management and Service Quality Evaluation

Complaint management has evolved from a reactive customer service function into a strategic component of quality management and organisational learning. In service industries characterised by frequent customer interactions, complaint-handling processes provide valuable information on operational deficiencies, customer expectations, and opportunities for continuous improvement. Rather than representing isolated incidents, customer complaints constitute a feedback mechanism that enables organisations to identify recurring service failures, improve internal processes, and strengthen customer relationships [17–19].

The relationship between complaint management and service quality is particularly important in postal and express delivery services. Unlike many other service industries, service failures in express delivery are often directly observable and measurable through operational events such as shipment damage, loss, delivery delays, service errors, or inaccurate tracking information. These operational failures frequently trigger customer complaints, making complaint management an essential continuation of the service delivery process rather than an independent administrative activity [3–7,20,21]. Consequently, customers evaluate operators not only according to the occurrence of service failures but also according to the effectiveness, fairness, and timeliness with which these failures are resolved [18,20].

The theoretical foundations of complaint management are commonly explained through perceived justice theory and service recovery research. Perceived justice theory suggests that customers evaluate complaint handling through three complementary dimensions: distributive justice, referring to the fairness of the outcome; procedural justice, concerning the accessibility, transparency, consistency, and efficiency of complaint procedures; and interactional justice, reflecting the quality of communication, empathy, and respect demonstrated throughout the resolution process [20]. Service recovery research further indicates that effective complaint handling can restore customer trust,

improve satisfaction, and strengthen long-term customer loyalty, whereas inadequate complaint resolution may amplify dissatisfaction despite otherwise satisfactory operational performance [18].

International quality-management standards reinforce this perspective. ISO 10002 considers complaint handling an integral element of organisational quality management and emphasises visibility, accessibility, responsiveness, objectivity, confidentiality, accountability, and continual improvement as essential principles of effective complaint-management systems [19]. Consequently, complaint handling should not be viewed solely as a regulatory obligation but also as a managerial tool for monitoring organisational performance and supporting evidence-based quality improvement.

The digital transformation of postal services has substantially influenced complaint-management practices. Online complaint portals, mobile applications, automated communication systems, chatbots, and social media platforms have improved complaint accessibility and increased the availability of operational data. Recent studies have also demonstrated the growing application of artificial intelligence, natural language processing, and text analytics for complaint classification, sentiment analysis, and the identification of recurring service failures [17,22–24]. These technologies improve the efficiency of complaint processing and provide valuable information for organisational learning. However, digitalisation alone does not guarantee higher complaint-resolution quality, as procedural fairness, transparency, and customer confidence remain essential determinants of successful complaint management [17,19].

Existing studies evaluate complaint-management performance using a broad range of indicators, including complaint rates, response time, complaint resolution time, resolution success rate, compliance with regulatory deadlines, escalation frequency, compensation mechanisms, and customer satisfaction following complaint resolution [18,19,25,26]. These indicators capture different aspects of complaint handling and provide valuable insights into organisational performance. Nevertheless, they are frequently analysed independently, making it difficult to achieve a comprehensive evaluation of complaint-resolution effectiveness.

An additional limitation concerns the interpretation of complaint outcomes. Complaint resolution is often represented as a binary outcome (resolved versus unresolved), although actual complaint handling may involve partial settlements, negotiated solutions, procedural closure without customer acceptance, or technical resolution without restoring customer confidence. Likewise, the use of average complaint-resolution time may conceal substantial differences among complaint categories, since complaints related to shipment loss, damage, delayed delivery, or compensation claims differ considerably in terms of investigation complexity and required processing time. These methodological limitations should therefore be recognised when developing performance evaluation frameworks based on objective administrative data.

Despite these limitations, objective complaint-resolution indicators remain particularly valuable for regulatory benchmarking because they are routinely collected, consistently reported, and directly comparable across operators. In contrast, customer satisfaction, perceived fairness, and trust are usually measured through surveys that are not always available or collected using comparable methodologies. Consequently, administrative complaint indicators continue to provide the most appropriate basis for comparative regulatory evaluation, provided that their interpretive limitations are appropriately acknowledged.

2.3. KPI-Based Performance Evaluation

KPI-based performance evaluation has become an essential component of performance management in postal and express delivery services due to declining letter volumes, the rapid growth of parcel traffic, expanding e-commerce, digital transformation, and increasing sustainability requirements. While traditional performance measurement focused primarily on universal service obligations, delivery standards, network accessibility, and affordability, contemporary postal and express services are increasingly recognised as complex logistics systems that require the simultaneous evaluation of operational, customer, environmental, financial, and digital performance dimensions [27–31].

A substantial body of research has focused on operational service quality and delivery reliability. Indicators such as delivery time, on-time performance, delivery frequency, shipment integrity, accessibility, complaint frequency, and service continuity remain among the most widely applied measures in both regulatory practice and operational management. At the international level, the Universal Postal Union (UPU) evaluates postal development through dimensions including reliability, reach, relevance, and resilience, while its statistical framework provides numerous indicators supporting international benchmarking. These developments demonstrate that postal performance is increasingly understood as a multidimensional concept rather than a single measure of delivery efficiency.

Recent research has further expanded KPI systems by incorporating indicators related to last-mile delivery performance. Last-mile operations directly influence operational costs, customer satisfaction, environmental performance, and service competitiveness. Consequently, recent studies increasingly employ indicators describing delivery density, route productivity, vehicle utilisation, failed delivery rates, operational efficiency, and real-time delivery performance, while artificial intelligence, IoT technologies, and digital monitoring systems provide additional opportunities for improving operational control and decision-making.

Customer-oriented performance evaluation has also gained increasing attention. The quality of last-mile delivery significantly influences customer satisfaction, trust, and future purchasing intentions, particularly within e-commerce environments, where the delivery operator represents the final point of interaction between retailers and consumers [9]. Consequently, contemporary KPI systems increasingly incorporate customer-focused measures alongside operational indicators, including service accessibility, communication quality, shipment visibility, flexibility, and complaint-related performance. These developments reflect the transition from operational efficiency towards customer-centred performance evaluation.

Complaint-related KPIs have become particularly important because they assess organisational responsiveness once service failures occur. Commonly applied indicators include complaint frequency, complaint rate, complaint-resolution time, complaint-resolution success rate, compliance with regulatory deadlines, compensation rates, and post-complaint customer satisfaction [18,19,25,26]. Unlike operational KPIs, which primarily describe the occurrence of service failures, complaint-related indicators evaluate the effectiveness of service recovery and the operator's ability to restore customer confidence.

In parallel, sustainability has emerged as an additional dimension of KPI-based evaluation. Postal and express operators increasingly monitor environmental indicators such as CO₂ emissions, energy consumption, empty mileage, alternative-fuel vehicle utilisation, and delivery consolidation. Similarly, digital transformation has encouraged the use of business intelligence systems, real-time monitoring platforms, simulation models, and advanced analytical tools for continuous KPI monitoring and operational decision support. These developments illustrate that modern KPI systems extend beyond operational

efficiency and increasingly support strategic management, sustainability evaluation, and regulatory supervision.

Despite these significant developments, the literature also reveals an important limitation. Existing KPI frameworks predominantly evaluate operational performance, customer satisfaction, environmental sustainability, or complaint management as separate dimensions of service quality. Operational indicators primarily measure the occurrence of service failures, whereas complaint-related indicators evaluate organisational responses after failures have occurred. Consequently, the relationship between operational deficiencies and complaint-resolution performance is rarely examined within a single analytical framework.

This separation has important practical implications. Operators demonstrating excellent operational performance may nevertheless exhibit ineffective complaint-handling procedures, while operators experiencing relatively frequent service failures may partially compensate through efficient complaint resolution and effective customer communication. Evaluating these dimensions independently therefore provides only a partial representation of overall service quality, limiting the usefulness of KPI systems for comprehensive benchmarking and regulatory evaluation.

From a methodological perspective, an effective KPI framework should satisfy several fundamental requirements. Selected indicators should be relevant to organisational objectives, objectively measurable, consistently available across operators, sufficiently sensitive to performance differences, and suitable for comparative benchmarking. These requirements are particularly important in regulatory environments, where harmonised administrative data provide an objective basis for objective performance evaluation. Consequently, objective operational and complaint-related indicators remain particularly appropriate for comparative evaluation because they are routinely collected and consistently reported by postal operators.

2.4. Composite Indicators in Service Quality Evaluation

Composite indicators have become an important analytical tool for evaluating complex systems characterised by multiple performance dimensions. Unlike individual indicators, which describe only specific aspects of performance, composite indicators synthesise several complementary variables into a single measure, thereby facilitating benchmarking, comparative analysis, performance monitoring, and decision-making. Their application has expanded considerably in recent years, particularly in public administration, healthcare, transport, logistics, education, sustainability evaluation, and digital governance, where multidimensional phenomena cannot be adequately captured by a single performance measure [32–37].

The methodological foundations of composite indicators are largely based on the OECD/JRC framework, which defines a systematic procedure consisting of several interrelated stages, including the establishment of a theoretical framework, indicator selection, data preprocessing, normalisation, weighting, aggregation, uncertainty analysis, and validation [33,34]. This structured approach ensures that composite indicators represent coherent analytical constructs rather than simple mathematical combinations of variables. Among these stages, the selection of indicators, determination of weighting coefficients, and aggregation procedures have been consistently recognised as the most influential methodological decisions because different choices may substantially affect the resulting rankings and their interpretation.

In service quality evaluation, composite indicators offer a significant advantage by integrating multiple dimensions that collectively determine organisational performance. Individual indicators frequently describe only one aspect of service quality—for example,

delivery reliability, complaint frequency, waiting time, customer satisfaction, or operational efficiency—whereas service quality itself is inherently multidimensional. Composite indicators therefore enable a more comprehensive evaluation by simultaneously considering several complementary performance dimensions while reducing the complexity of interpreting large numbers of individual indicators [5,36,37].

The application of composite indicators has expanded beyond traditional service-quality evaluation to encompass transport and logistics. Recent studies have employed composite measures to evaluate transport sustainability, logistics efficiency, urban quality of life, governance performance, and digital public services [35–39]. These studies demonstrate that composite indicators provide effective decision-support tools whenever multiple criteria must be synthesised into an interpretable performance measure. Their value lies not only in ranking alternatives but also in supporting strategic planning, benchmarking, policy evaluation, and performance monitoring across heterogeneous systems.

Despite these advantages, composite indicators are not without methodological limitations. One of the most frequently discussed issues concerns the selection of indicators. The validity of a composite indicator depends fundamentally on whether the selected variables adequately represent the underlying construct being measured. Consequently, indicator selection should be guided by theoretical justification, empirical evidence, data availability, and practical relevance rather than solely by statistical considerations.

A second methodological challenge concerns the assignment of indicator weights. Previous studies have proposed a wide range of weighting approaches, including equal weighting, expert judgement, Analytic Hierarchy Process (AHP), entropy weighting, Data Envelopment Analysis (DEA), and other objective or subjective weighting techniques. Each approach reflects different assumptions regarding the relative importance of indicators. Objective methods such as entropy weighting derive weights directly from data variability and reduce the influence of subjective judgement, whereas expert-based methods incorporate stakeholder preferences but may introduce evaluator bias. Consequently, the choice of weighting methodology should be consistent with the objective of the evaluation framework and the characteristics of the available data.

Aggregation represents a further important methodological issue. Composite indicators should preserve meaningful differences among alternatives while avoiding excessive compensation effects, where poor performance on one dimension is completely offset by excellent performance on another. For this reason, transparency regarding normalisation, weighting, and aggregation procedures has become a fundamental requirement for ensuring the credibility, robustness, and interpretability of composite measures [34–37].

Although composite indicators have been widely applied in transport, logistics, healthcare, sustainability evaluation, and public administration, their use for evaluating complaint-resolution performance in the express delivery sector remains relatively limited. Existing studies predominantly employ composite indicators to evaluate overall logistics performance, transport sustainability, or service quality, whereas complaint-resolution performance is generally assessed using individual indicators such as complaint rate, resolution time, or compliance with regulatory deadlines. Consequently, current approaches provide only fragmented information regarding complaint-handling effectiveness and offer limited support for comprehensive benchmarking of express delivery operators.

The framework proposed in this study adopts the established methodological principles of composite indicator construction rather than introducing a fundamentally new methodology. Its contribution lies in applying these principles to the evaluation of complaint-resolution performance through the development of a Composite Quality In-

dex (CQI) that integrates objective Complaint Resolution Indicators (CRI) into a single synthetic measure. By relying on routinely collected regulatory indicators and objective entropy-based weighting, the proposed CQI provides a transparent, objective, and reproducible basis for comparative evaluation while remaining fully compatible with existing regulatory reporting systems. This distinction is important because the originality of the proposed framework derives from the integration and application of existing methodological principles to a previously underexplored evaluation problem, rather than from proposing an entirely new composite indicator methodology. This clarification directly addresses the concern that the study's novelty lies primarily in application rather than in methodological invention.

2.5. Multi-Criteria Decision-Making Models in Transport and Logistics

Multi-criteria decision-making (MCDM) methods have become an important component of decision-support systems in transport and logistics because decision-makers are frequently required to evaluate alternatives using multiple, often conflicting, criteria. Problems such as transport planning, supplier selection, logistics network design, warehouse location, sustainable mobility, city logistics, and service-quality evaluation simultaneously involve economic, operational, environmental, technical, and customer-oriented objectives that cannot be adequately represented by a single performance indicator [40–50]. Consequently, MCDM methods provide structured analytical procedures for integrating heterogeneous evaluation criteria into transparent and reproducible decision-making processes.

Among the most frequently applied MCDM approaches are the Analytic Hierarchy Process (AHP), TOPSIS, ELECTRE, PROMETHEE, and VIKOR [40,51–54]. These methods have been successfully employed to a wide range of transport and logistics problems, including infrastructure prioritisation, supplier evaluation, route selection, sustainable transport planning, logistics performance evaluation, and city logistics. Their widespread adoption demonstrates the suitability of MCDM techniques for analysing complex decision problems characterised by multiple performance criteria and competing stakeholder objectives.

Recent developments indicate a gradual transition from applying MCDM methods to purely economic optimisation problems towards supporting multidimensional performance evaluation. Contemporary studies increasingly integrate MCDM techniques with composite indicators, sustainability evaluation, business intelligence, simulation models, and objective weighting methods to enhance the robustness, transparency, and interpretability of decision-support frameworks [48–50]. These developments are particularly relevant for service industries, where operational efficiency, customer satisfaction, environmental performance, and service quality must be evaluated simultaneously rather than independently.

Despite their methodological diversity, the effectiveness of MCDM methods depends primarily on the quality of the evaluation framework rather than on the mathematical characteristics of a particular ranking algorithm. The selection of relevant criteria, their theoretical justification, data quality, weighting methodology, and the definition of decision alternatives generally have a greater influence on the credibility of evaluation results than the choice of a specific MCDM technique. Consequently, MCDM methods should be regarded as analytical decision-support tools whose performance depends on the adequacy of the underlying performance model rather than as independent sources of methodological novelty.

This consideration is particularly important in postal and express delivery services. The evaluation of operators requires the simultaneous consideration of operational service

failures, complaint-resolution effectiveness, and overall service quality, involving indicators with different measurement scales, directions of optimisation, and relative importance. Such multidimensional characteristics make this problem well suited to MCDM-based analysis, provided that the underlying evaluation framework appropriately captures the relationships among the considered performance dimensions.

Among recently developed MCDM methods, the MARCOS (Measurement of Alternatives and Ranking according to Compromise Solution) method has attracted considerable attention in transport, logistics, supply chain management, sustainability evaluation, and performance evaluation because it simultaneously evaluates alternatives with respect to the ideal and anti-ideal solutions while determining their corresponding utility functions [45,55]. Compared with classical distance-based approaches, MARCOS provides stable rankings, supports simultaneous treatment of benefit and cost criteria, and offers straightforward interpretation of the relative utility of evaluated alternatives. Recent applications further demonstrate its suitability for evaluating logistics systems characterised by heterogeneous criteria and complex performance structures [45,48–50].

The selection of the MARCOS method in this study was motivated by the characteristics of the proposed evaluation problem rather than by any assumption of methodological superiority over TOPSIS or other MCDM methods. The proposed framework simultaneously evaluates Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), and the Composite Quality Index (CQI), which include both benefit and cost criteria with objectively determined weights. Under these conditions, MARCOS provides a transparent procedure for comparing operators with respect to both ideal and anti-ideal reference solutions while preserving the interpretability of the obtained rankings. Accordingly, the methodological contribution of this study lies not in introducing a new MCDM technique but in integrating an established MCDM method within a novel framework for evaluating complaint-resolution performance in express delivery services. This distinction is important because the research contribution arises from the proposed evaluation framework rather than from the choice of the ranking algorithm itself.

2.6. Research Gap, Research Objective and Scientific Contribution

The literature reviewed in the preceding sections demonstrates substantial progress in customer protection, complaint management, KPI-based performance evaluation, composite indicators, and multi-criteria decision-making methods. Collectively, these research streams provide valuable theoretical foundations and methodological tools for evaluating service quality in transport and logistics. Nevertheless, they have largely evolved independently, resulting in fragmented approaches that explain individual aspects of service quality but provide limited support for comprehensive evaluation of complaint-resolution performance in express delivery services.

Based on the reviewed literature, four principal research gaps can be identified.

G1. Integration of operational service quality and complaint-resolution performance. Existing studies predominantly examine operational service quality and complaint management as separate research domains. Operational performance indicators primarily quantify service failures, whereas complaint-management studies focus on organisational responses after failures have occurred. Consequently, integrated frameworks simultaneously evaluating both dimensions remain limited within the postal and express delivery sector.

G2. Comprehensive KPI-based evaluation of complaint resolution. Although numerous KPI systems have been developed for postal and logistics services, existing approaches primarily emphasise operational performance and provide limited support for systematically evaluating complaint-resolution effectiveness as a strategic component of service

quality. As a result, complaint-related indicators are commonly analysed individually rather than as complementary elements of a broader performance evaluation framework.

G3. Application of composite indicators to complaint-resolution performance. Composite indicators have found widespread application in healthcare, sustainability evaluation, transport, and public administration. However, their application to complaint-resolution performance in express delivery services remains relatively limited. Existing studies rarely integrate multiple complaint-resolution dimensions into a single synthetic measure suitable for benchmarking postal operators.

G4. Decision-support frameworks integrating operational failures and complaint management. Although MCDM methods are extensively used in transport and logistics, they are generally applied within established evaluation frameworks rather than to integrated models linking operational service failures with complaint-resolution performance. Consequently, existing decision-support approaches provide limited support for simultaneously identifying the operational sources of customer complaints and evaluating the effectiveness of organisational responses within a unified benchmarking framework. This is the unresolved scientific problem addressed by the present study. This formulation directly addresses the reviewer's observation that the absence of a particular analytical method does not, by itself, constitute a research gap.

2.6.1. Research Objective

To address the identified research gaps, this study develops an integrated decision-support framework for evaluating complaint-resolution performance among express delivery operators. The proposed framework combines Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), an Entropy-weighted Composite Quality Index (CQI), and the MARCOS multi-criteria decision-making method within a unified analytical model. Rather than evaluating operational performance and complaint management independently, the framework explicitly integrates these complementary dimensions to support comprehensive benchmarking and evidence-based decision-making.

2.6.2. Scientific Contribution

The scientific contribution of this study lies in the integration of established theoretical and methodological concepts into a coherent evaluation framework specifically designed for complaint-resolution performance in express delivery services. The contribution should not be interpreted as the introduction of entirely new individual indicators or a novel MCDM algorithm. Instead, it derives from the conceptual integration of operational service failures, complaint-resolution performance, objective composite indicator construction, and multi-criteria decision-making within a single analytical framework applicable to regulatory benchmarking and managerial performance evaluation. This clarification also addresses the reviewer's request to distinguish between the originality of the framework and the originality of its individual components.

More specifically, this study makes four principal contributions:

It develops an integrated conceptual framework linking operational service failures and complaint-resolution performance, thereby connecting two research streams that have predominantly been investigated independently.

It proposes a structured KPI framework that integrates Service Failure Indicators (SFI) and Complaint Resolution Indicators (CRI) into a unified performance evaluation framework.

It develops a Composite Quality Index (CQI) that synthesises multiple complaint-resolution indicators into a single objective performance measure using entropy-based weighting.

It demonstrates the applicability of the proposed framework through the integration of the CQI within a MARCOS-based decision-support model for benchmarking express delivery operators.

Overall, the proposed framework contributes to the literature on service quality evaluation, complaint management, logistics performance evaluation, and customer protection by providing an integrated analytical framework for evaluating complaint-resolution performance. Furthermore, it offers a practical decision-support tool for regulators, postal operators, and other stakeholders responsible for monitoring, benchmarking, and improving service quality and complaint-resolution performance.

3. Development of the KPI Framework

3.1. Conceptual Framework

Efficient complaint resolution represents one of the key elements of customer protection and service quality management in the express delivery sector. Although operators regularly monitor various performance indicators, neither regulatory practice nor the existing literature provides a unified framework for comprehensively assessing complaint resolution performance and objectively comparing operators. Therefore, this study develops a KPI-based framework designed to evaluate the quality and effectiveness of complaint-handling processes among express delivery operators.

Although the primary objective of this study is to evaluate complaint resolution performance, the proposed framework assumes that complaint handling cannot be interpreted independently of the operational conditions that generate customer complaints. From a service-quality perspective, complaints represent the consequence of operational service failures, while complaint handling represents the organisational response to those failures. Accordingly, operational performance and complaint resolution are considered complementary stages of the same service process. Evaluating complaint-related indicators independently of the operational context may therefore result in an incomplete evaluation of overall operator performance. For this reason, the proposed framework integrates operational Service Failure Indicators (SFI) with Complaint Resolution Indicators (CRI), enabling complaint resolution performance to be interpreted in relation to the frequency of service failures that precede it.

The proposed framework is based on two complementary performance dimensions. The first dimension focuses on service quality and includes indicators that capture the most common causes of customer complaints. The second dimension evaluates complaint resolution performance and reflects the operator's ability to respond effectively to customer requests and resolve service-related problems. Accordingly, the indicators are organised into two main groups:

- Service Failure Indicators (SFI)
- Complaint Resolution Indicators (CRI)

Service Failure Indicators represent operational service quality measures and quantify the frequency of the most common service failures occurring during the collection, transportation, sorting, and delivery of shipments. Complaint Resolution Indicators assess the effectiveness of complaint-handling processes in terms of resolution success, timeliness, and compliance with regulatory requirements.

The proposed KPI framework does not treat Service Failure Indicators as direct measures of complaint-resolution performance. Instead, SFI provide the operational context within which complaint-resolution performance is evaluated. This distinction is important because operators experiencing similar complaint-resolution outcomes may differ substantially in the frequency of service failures that generate complaints. Simultaneously evaluating both dimensions therefore provides a more balanced and compre-

hensive evaluation of overall service performance than considering complaint-related indicators alone.

The proposed KPI framework enables the simultaneous evaluation of service quality and complaint resolution performance, thereby providing a foundation for the development of a Composite Quality Index (CQI) and the subsequent multi-criteria evaluation and ranking of express delivery operators.

3.2. Service Failure Indicators

Service Failure Indicators (SFI) represent the primary group of indicators designed to assess the quality of express delivery services through the identification and quantification of the most common operational failures occurring during the acceptance, transportation, sorting, and delivery of shipments. In the literature, shipment damage, loss, service errors, and delivery delays are widely recognized as key indicators of logistics system reliability and service quality, as they directly affect customer satisfaction, complaint frequency, and overall perceptions of service quality [5,6,46].

For the purposes of this study, four Service Failure Indicators were defined to capture the most significant categories of irregularities in the provision of express delivery services. To ensure comparability among operators of different sizes and market shares, all indicators were expressed per 1000 processed shipments. Since higher values indicate poorer operator performance, all indicators within this group were classified as cost criteria. The selected indicators cover the most common categories of service failures in express delivery operations, including shipment damage, loss, service errors, and delivery delays, which are widely recognized in the literature as major causes of customer complaints and dissatisfaction [47].

Although the primary objective of this study is to evaluate complaint-resolution performance, operational service failures represent the events that initiate the complaint process. From a service-quality perspective, complaint handling cannot be interpreted independently of the operational conditions that generate customer complaints. Consequently, the proposed framework incorporates Service Failure Indicators (SFI) not as direct measures of complaint-resolution performance but as contextual indicators that enable complaint outcomes to be interpreted in relation to the frequency of operational failures. This integration provides a more comprehensive evaluation of operator performance than analysing complaint-related indicators alone by establishing the operational context for the subsequent evaluation of Complaint Resolution Indicators (CRI).

Since express delivery operators differ considerably in the annual number of processed shipments, all Service Failure Indicators are expressed per 1000 processed shipments. This normalization eliminates the influence of operator size and enables meaningful comparison based on relative rather than absolute frequencies of operational failures. The selected denominator is consistent with comparative regulatory benchmarking, where operators with substantially different traffic volumes are evaluated using harmonized relative performance measures [56].

Damage Rate (SFI₁) represents the frequency of shipments that have suffered physical damage during the service process, affecting the contents, functionality, or market value of the shipment. Damage may occur during handling, sorting, transportation, or delivery operations and is considered one of the most common causes of customer complaints in express delivery services.

The Service Failure Indicators (SFI) proposed in this study were developed by the authors as a standardized set of key performance indicators for evaluating operational service failures in the express delivery sector. Although their mathematical formulation is based on conventional relative-frequency calculations commonly used in performance

measurement, the proposed indicators represent a novel regulatory-oriented KPI framework specifically designed to quantify the frequency of the most relevant service failure events. Accordingly, the methodological contribution of this study lies in the development of the SFI indicator set and its integration into the proposed evaluation framework rather than in the mathematical form of the individual equations.

The proposed SFI are defined as follows:

$$SFI_i = \frac{N_i}{N_T} \times 1000, \quad (1)$$

where

SFI —value of the i -th Service Failure Indicator per 1000 processed shipments;

N_i —number of operational service failures corresponding to indicator i ;

N_T —total number of processed shipments.

The same calculation procedure was applied to all four Service Failure Indicators, with N_i representing the corresponding category of operational service failure (damage, loss, service error, or delay). Thus, the numerator differs depending on the observed category of service failure, whereas the denominator remains identical for all indicators, ensuring methodological consistency and comparability across operators.

Damage Rate (SFI_1) is therefore obtained by substituting the number of damaged shipments (N_D) into the general expression where

SFI_1 —Damage Rate;

N_D —number of damaged shipments;

N_T —total number of processed shipments.

The unit of measurement is the number of damaged shipments per 1000 processed shipments (‰). Lower values of this indicator indicate a higher level of shipment protection and greater reliability of the logistics process.

Lost Shipment Rate (SFI_2) is obtained by substituting the number of lost shipments (N_L) into the general expression.

Where

SFI_2 —Lost Shipment Rate;

N_L —number of lost shipments;

N_T —total number of processed shipments.

The unit of measurement is expressed as the number of lost shipments per 1000 processed shipments (‰). Lower values of this indicator indicate a higher level of reliability and efficiency of the logistics network.

Service Error Rate (SFI_3) is obtained by substituting the number of service errors (N_E) into the general expression.

Where

SFI_3 —Service Error Rate;

N_E —number of service errors;

N_T —total number of processed shipments.

The unit of measurement is expressed as the number of service errors per 1000 processed shipments (‰). Lower values of this indicator indicate higher operational accuracy, better process control, and improved service quality.

This indicator provides insight into the accuracy of operational processes and the overall quality of service execution. Lower values indicate a higher degree of operational accuracy, improved process control, and better service performance.

Delay Rate (SFI_4) is obtained by substituting the number of delayed shipments (N_{DL}) into the general expression.

Where

SFI_4 —Delay Rate;

N_{DL} —number of delayed shipments;

N_T —total number of processed shipments.

The unit of measurement is expressed as the number of delayed shipments per 1000 processed shipments (‰). Lower values of this indicator indicate a higher level of service reliability, better compliance with delivery standards, and greater customer satisfaction.

In accordance with the national quality monitoring methodology, a shipment is classified as delayed when it is delivered after the prescribed delivery standard. Accordingly, the Delay Rate reflects the frequency of delayed shipments rather than the duration or severity of individual delays. Although delay duration may provide additional information regarding service performance, such data are not included in the standardized regulatory reporting system used in this study. Therefore, the proposed indicator is intended to support comparative regulatory evaluation based on harmonized performance measures that are consistently available for all operators.

Based on the defined indicators, all Service Failure Indicators (SFI) are treated as cost criteria, as an increase in their values reflects a higher frequency of operational failures and consequently lower service quality performance.

Within the proposed multi-criteria decision-making framework, the term cost criterion indicates that lower values correspond to better performance and should therefore be minimized during the evaluation process. This classification reflects the optimisation direction of the indicators rather than their economic or financial cost.

An overview of the proposed indicators, their units of measurement, and criterion classifications is presented in Table 1.

Table 1. Service Failure Indicators and Criterion Classification.

Symbol	Indicator	Unit of Measurement ¹	Criterion Type
SFI_1	Damage Rate	‰	Cost
SFI_2	Lost Shipment Rate	‰	Cost
SFI_3	Service Error Rate	‰	Cost
SFI_4	Delay Rate	‰	Cost

¹ All indicators are expressed as the number of occurrences per 1000 processed shipments (‰).

The proposed indicators provide a systematic framework for assessing the operational performance of express delivery operators by quantifying the most common service failures.

Although these indicators do not directly evaluate complaint-resolution performance, they establish the operational context in which customer complaints arise. In the following subsection, they are complemented by Complaint Resolution Indicators (CRI), enabling complaint handling to be evaluated in relation to the operational service failures that precede it and thereby providing the basis for the integrated KPI framework proposed in this study.

4. Development of the Composite Quality Index (CQI)

4.1. Conceptual Background

Complaint resolution performance is widely recognized as one of the most important indicators of customer service quality, reflecting an operator's ability to effectively address and resolve problems arising during the provision of express delivery services [57]. In the logistics and express delivery sector, the quality of customer experience and the effectiveness of complaint resolution are regarded as key dimensions of overall service quality, as they directly influence customer satisfaction, trust, and service performance [6,58].

Although individual indicators enable the monitoring of different aspects of the complaint-handling process, their separate analysis often limits a comprehensive evaluation of operator performance and objective benchmarking across different service providers [7,52].

In the existing literature and regulatory practice, complaint resolution performance is most commonly assessed using individual indicators that measure resolution success, compliance with prescribed deadlines, and the average complaint resolution time [2,12]. However, this approach does not enable a comprehensive evaluation of complaint-handling performance, nor does it provide a unified framework for benchmarking service providers and supporting decision-making [33,34].

To overcome these limitations, this study develops a Composite Quality Index (CQI) as a synthetic indicator for assessing the overall effectiveness of complaint resolution among express delivery operators. The proposed index integrates multiple complaint-resolution performance indicators into a single performance measure, enabling easier comparison of operators, identification of weaknesses in complaint-handling processes, and support for service quality improvement. The construction of composite indicators is a widely accepted approach for evaluating complex systems and measuring service performance [34,35].

The Complaint Resolution Indicators selected for the construction of the CQI were chosen because they represent standardized regulatory performance measures that are systematically collected for all observed operators and directly reflect the key dimensions of complaint-handling performance. Their selection was based on three criteria: (i) regulatory relevance, (ii) data availability and comparability, and (iii) suitability for objective benchmarking across operators. Although additional dimensions, such as customer satisfaction, procedural transparency, accessibility of complaint channels, or perceived procedural fairness, may also influence the quality of complaint handling, these dimensions are not consistently available within the regulatory reporting system used in this study and therefore were not included in the proposed Composite Quality Index.

For the construction of the Composite Quality Index (CQI), three key Complaint Resolution Indicators (CRI) were selected to represent the most important dimensions of complaint-handling performance:

- CRI₁—Successful Complaint Resolution Rate;
- CRI₂—Resolution within Regulatory Deadline;
- CRI₃—Average Resolution Time.

Together, these three indicators capture the principal regulatory dimensions of complaint-resolution performance. CRI₁ reflects the effectiveness of complaint resolution, CRI₂ measures compliance with prescribed regulatory deadlines, while CRI₃ represents the operational efficiency of the complaint-handling process. Their integration into a single composite indicator enables a balanced and comprehensive evaluation of complaint-resolution performance while reducing reliance on any individual performance measure.

The first two indicators are classified as benefit criteria, since higher values indicate better complaint-handling performance. In contrast, Average Resolution Time (CRI₃) is treated as a cost criterion, as shorter resolution times reflect greater operational efficiency. Therefore, before integrating these indicators into the Composite Quality Index (CQI), the data must be appropriately normalized to ensure comparability and consistency among indicators with different measurement scales and optimization directions.

4.2. Complaint Resolution Indicators

Complaint Resolution Indicators (CRI) represent the second group of Key Performance Indicators (KPIs) included in the proposed evaluation framework and are intended to assess the effectiveness and quality of complaint-handling processes among express delivery operators. Unlike Service Failure Indicators, which quantify operational deficiencies occurring during service provision, Complaint Resolution Indicators evaluate the operator's response after a complaint has been submitted. Consequently, they measure the effectiveness, timeliness, and efficiency of complaint resolution, thereby complementing the operational perspective provided by the Service Failure Indicators.

Successful Complaint Resolution Rate (CRI_1) represents the percentage of complaints successfully resolved relative to the total number of submitted complaints. It measures an operator's ability to address customer complaints effectively and provide satisfactory resolutions. A high complaint resolution rate contributes directly to customer satisfaction, strengthens customer trust, and enhances long-term customer loyalty, making it one of the key indicators of complaint-handling performance [58].

Although successful complaint resolution is represented as a binary outcome (resolved or unresolved), this indicator was selected because it corresponds to the standardized regulatory reporting methodology applied in the observed market and enables objective comparison among operators. Consequently, CRI_1 reflects regulatory complaint-resolution performance rather than customers' subjective perception of satisfaction following the complaint-handling process.

The Successful Complaint Resolution Rate is calculated using the following expression:

$$CRI_1 = \frac{NSR}{NC} \quad (2)$$

where

CRI_1 —Successful Complaint Resolution Rate;

NSR —number of successfully resolved complaints;

NC —total number of submitted complaints.

The indicator is expressed as a dimensionless ratio ranging from 0 to 1. Higher values indicate better complaint-resolution performance and greater effectiveness in addressing customer complaints. Therefore, CRI_1 is classified as a benefit criterion.

Resolution within Regulatory Deadline (CRI_2) represents the proportion of complaints resolved within the prescribed regulatory deadline relative to the total number of resolved complaints. Timely complaint resolution is an essential component of customer protection and one of the key indicators of customer service quality, as it reflects an operator's ability to comply with regulatory requirements and provide prompt responses to customer complaints [2,12].

Compliance with regulatory deadlines is treated as an objective indicator of procedural efficiency because the prescribed complaint-resolution deadlines are identical for all operators included in the analysis, thereby ensuring methodological consistency and comparability of results.

The Resolution within Regulatory Deadline is calculated using the following expression:

$$CRI_2 = \frac{NRD}{NR} \quad (3)$$

where

CRI_2 —Resolution within Regulatory Deadline;

NRD —number of complaints resolved within the regulatory deadline;

NR —total number of successfully resolved complaints.

This indicator measures the timeliness of the complaint-handling process and the degree of an operator's compliance with regulatory requirements. It is expressed as a dimensionless ratio ranging from 0 to 1, where higher values indicate better regulatory compliance and more efficient complaint handling. Therefore, CRI_2 is classified as a benefit criterion.

Average Resolution Time (CRI_3) represents the average number of days required to process and resolve customer complaints. The time required to handle complaints is widely recognized as a key determinant of complaint-handling efficiency, customer satisfaction, and perceived service quality, as shorter resolution times contribute to a more positive customer experience and greater confidence in the service provider [51,54].

The Average Resolution Time represents an aggregate indicator of complaint-handling efficiency at the operator level. Although individual complaint categories may require different processing times depending on their complexity, the regulatory reporting system provides only aggregate information on complaint-resolution time. Consequently, the proposed indicator reflects the overall operational efficiency of complaint handling rather than the duration associated with individual complaint categories. This limitation should be considered when interpreting the results.

The Average Resolution Time is calculated using the following expression:

$$CRI_3 = \frac{\sum_{r=1}^{NR} T_r}{NR} \quad (4)$$

where

CRI_3 —Average Resolution Time;

T_r —resolution time of complaint r (days);

NR —total number of successfully resolved complaints.

The indicator is expressed in days. Lower values indicate a more efficient complaint-handling process and faster resolution of customer complaints. Therefore, CRI_3 is classified as a cost criterion.

The Average Complaint Resolution Time represents the mean time required to resolve customer complaints reported within the standardized regulatory monitoring system. Since the available regulatory data are aggregated at the operator level, they do not distinguish complaint resolution times according to individual complaint categories or levels of complexity. Consequently, the proposed indicator reflects the overall operational efficiency of complaint handling rather than the resolution performance for specific complaint types. This approach is consistent with the primary objective of the proposed framework, namely to support comparative regulatory evaluation using harmonized performance indicators that are uniformly available for all operators.

The three Complaint Resolution Indicators represent complementary dimensions of complaint-handling performance and therefore should not be interpreted independently. CRI_1 reflects the effectiveness of complaint resolution, CRI_2 evaluates compliance with prescribed regulatory requirements, while CRI_3 measures the operational efficiency of the complaint-handling process. Their combined interpretation provides a more comprehensive evaluation of complaint-resolution performance than any individual indicator alone and provides the conceptual foundation for the Composite Quality Index (CQI) developed in the following sections.

4.3. Entropy-Based Weight Determination

To avoid subjectivity in determining the importance of individual indicators, the Entropy method was employed to derive objective criterion weights based on the information content and variability of the observed data [59,60].

The Entropy weighting method was selected because it enables the objective determination of criterion weights using only the information contained in the observed dataset, without requiring subjective judgments from experts. Since the proposed evaluation framework is based on standardized regulatory performance indicators collected under identical reporting conditions for all observed operators, an objective weighting approach was considered more appropriate than expert-based methods, such as the Analytic Hierarchy Process (AHP). This choice improves the transparency, reproducibility, and comparability of the proposed evaluation framework, making it particularly suitable for benchmarking studies based on regulatory data.

The Entropy weighting method is based on the assumption that criteria exhibiting greater information dispersion contain more useful information and, consequently, should be assigned higher weights in the evaluation process [59,60].

It should be emphasized that Entropy weights reflect the discriminatory power of individual indicators within the analysed dataset rather than their intrinsic managerial or strategic importance. Consequently, indicators exhibiting greater variability among operators receive higher weights because they provide more information for distinguishing performance differences. This characteristic makes the Entropy method particularly appropriate for objective performance evaluation while avoiding potential bias associated with subjective weighting procedures. At the same time, it should be acknowledged that different weighting methods may produce different weight distributions. Therefore, the selected approach should be interpreted as an objective data-driven weighting procedure rather than the only possible weighting strategy.

The first step involves normalizing the decision matrix $X = [x_{ij}]_{m \times n}$, where m denotes the number of express delivery operators and n represents the number of evaluation criteria. The normalized values are calculated using the following expression:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (5)$$

where

- x_{ij} —value of criterion j for operator i ;
- p_{ij} —normalized value of criterion j for operator i ;
- m —number of operators.

In this way, normalization is performed independently for each criterion, such that the sum of the normalized values for each criterion is equal to one:

$$\sum_{i=1}^m p_{ij} = 1, \quad j = 1, 2, \dots, n \quad (6)$$

This condition provides the basis for calculating the entropy values of the evaluation criteria. After normalization, the entropy of each criterion is calculated using the following expression:

$$e_j = -k \sum_{i=1}^m p_{ij} \ln(p_{ij}) \quad (7)$$

where

- e_j —entropy value of criterion j ;
- p_{ij} —normalized value of criterion j for operator i ;
- $k = 1/\ln(m)$ normalization constant ensuring that the entropy values are bounded within the interval $[0, 1]$;
- m —number of operators.

The degree of divergence of each criterion is calculated using the following expression:

$$d_j = 1 - e_j \quad (8)$$

where

d_j —degree of divergence of criterion j .

The objective weights of the evaluation criteria are calculated using the following expression:

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad (9)$$

where

w_j —objective weight of criterion j ;

d_j —degree of divergence of criterion j ;

n —number of criteria.

The following condition holds:

$$\sum_{j=1}^n w_j = 1 \quad (10)$$

which ensures that the sum of the criterion weights is equal to one.

4.4. Calculation of the Composite Quality Index

After determining the criterion weights, the Composite Quality Index (CQI) is constructed. By integrating multiple complaint resolution indicators into a single composite indicator, the CQI provides a comprehensive evaluation of complaint-handling performance. The use of composite indicators is a widely accepted approach for evaluating complex systems and measuring service performance [33,34].

The primary purpose of the Composite Quality Index is to synthesize the multidimensional nature of complaint-resolution performance into a single performance measure suitable for comparative evaluation. Rather than replacing the individual Complaint Resolution Indicators, the CQI preserves their information through entropy-based aggregation while providing a simplified representation of overall complaint-handling performance. Such an approach facilitates objective benchmarking of operators and supports the subsequent integration of complaint-resolution performance with operational service quality indicators within a unified evaluation framework.

Since CRI3 is classified as a cost criterion, it must be transformed into a benefit criterion prior to its integration into the Composite Quality Index (CQI). The transformation is performed using the following expression:

$$NCR3 = 1 - \frac{CRI3}{CRI3_{max}} \quad (11)$$

where

$NCR3$ —normalized value of the Average Resolution Time indicator;

$CRI3_{max}$ —maximum value of the Average Resolution Time indicator in the observed sample.

After transforming the cost criterion, the Composite Quality Index (CQI) is calculated as follows:

$$CQI = w_1 CRI1 + w_2 CRI2 + w_3 NCR3 \quad (12)$$

where

CQI —Composite Quality Index;

w_1 —Entropy weight of CRI1;

w_2 —Entropy weight of CRI2;

w_3 —Entropy weight of NCRI3;

CRI1—Successful Complaint Resolution Rate;

CRI2—Resolution within Regulatory Deadline;

NCRI3—normalized value of the Average Resolution Time indicator.

Higher CQI values indicate more effective and higher-quality complaint resolution performance. Similar approaches integrating multiple indicators into a single composite index have been successfully applied in studies on logistics, transport, and service quality evaluation.

The proposed CQI integrates three key dimensions of the complaint-handling process: complaint resolution success, timeliness of complaint resolution, and operational efficiency. By combining these dimensions into a single composite indicator, the CQI provides a comprehensive evaluation of complaint-handling performance and facilitates objective benchmarking of express delivery operators.

It should be emphasized that the Composite Quality Index is not intended to be used together with the individual Complaint Resolution Indicators in the subsequent multi-criteria evaluation. Instead, the CQI replaces the individual CRI by representing complaint-resolution performance as a single aggregated criterion. Consequently, the final decision matrix used in the MARCOS method combines four Service Failure Indicators (SFI₁–SFI₄) with one Composite Quality Index (CQI), thereby avoiding double counting of complaint-related information while enabling the simultaneous evaluation of operational service quality and complaint-resolution performance within a unified decision-support framework.

In the subsequent stage of the proposed evaluation framework, the CQI is incorporated into the MARCOS multi-criteria decision-making method to support the evaluation and ranking of express delivery operators.

5. Research Methodology

5.1. Research Framework

The aim of this study is to develop an integrated framework for evaluating the performance of express delivery operators based on service quality and complaint resolution performance indicators. The proposed methodological approach combines Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), objective criterion weighting using the Entropy method, and the MARCOS multi-criteria decision-making method for the comprehensive evaluation and ranking of express delivery operators.

The proposed research framework consists of four interrelated steps. In the first step, the Key Performance Indicators (KPIs) are defined and organized into two groups: Service Failure Indicators (SFI) and Complaint Resolution Indicators (CRI). Service Failure Indicators quantify the most significant operational failures occurring during the provision of express delivery services, including shipment damage, loss, service errors, and delivery delays. Complaint Resolution Indicators evaluate the effectiveness of the complaint-handling process in terms of complaint resolution success, compliance with regulatory deadlines, and average complaint resolution time.

The proposed framework follows a sequential evaluation procedure. First, operational service quality is described through the Service Failure Indicators, while complaint-handling performance is evaluated using the Complaint Resolution Indicators. Subsequently, the three Complaint Resolution Indicators are objectively aggregated into the Composite Quality Index (CQI) using the Entropy weighting method. This sequential approach enables complaint-resolution performance to be represented by a single synthetic

criterion before its integration with operational service quality indicators within the final multi-criteria evaluation framework.

In the second step, the Composite Quality Index (CQI) is developed by integrating the three key Complaint Resolution Indicators into a single synthetic measure of complaint-handling performance. To eliminate subjectivity in determining the relative importance of individual criteria, criterion weights are determined using the Entropy weighting method, which objectively quantifies the information contribution of each criterion [59,60].

The third step involves constructing the final decision matrix using four Service Failure Indicators (SFI) and the Composite Quality Index (CQI). Since the CQI already incorporates the information contained in the three Complaint Resolution Indicators through entropy-based aggregation, the individual CRI are not included separately in the decision matrix. This approach avoids double counting of complaint-related information while preserving the multidimensional representation of complaint-resolution performance.

The resulting decision matrix serves as the basis for the multi-criteria evaluation of express delivery operators using the MARCOS (Measurement Alternatives and Ranking according to COMpromise Solution) method. MARCOS enables the evaluation of alternatives relative to both the ideal and anti-ideal solutions and has been widely applied in recent years in studies on transport, logistics, and performance evaluation systems [45,48–50].

In the final step, the express delivery operators are ranked based on the calculated utility function values. The obtained results enable the identification of the relative strengths and weaknesses of individual operators and facilitate the benchmarking of performance within the express delivery sector.

The overall research framework is illustrated in Figure 1.

Overall, the proposed research framework enables an integrated evaluation of service quality and complaint resolution performance by combining Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), objective Entropy-based criterion weighting, and MARCOS multi-criteria decision-making within a unified decision-support framework. Consequently, it provides an objective basis for evaluating and benchmarking the performance of express delivery operators while supporting informed decision-making by managers, regulators, and other stakeholders.

The MARCOS (Measurement Alternatives and Ranking according to Compromise Solution) multi-criteria decision-making method was employed to evaluate and rank express delivery operators in this study. MARCOS is a relatively recent MCDM approach that evaluates the utility of alternatives relative to both the ideal and anti-ideal solutions [45,55]. Owing to its simplicity, result stability, ability to simultaneously handle both benefit and cost criteria, and utility-based evaluation relative to the ideal and anti-ideal solutions, the MARCOS method has been widely applied in recent years in studies on transport, logistics, supply chain management, and performance evaluation [48–50].

In the proposed evaluation framework, the MARCOS method is applied after the construction of the Composite Quality Index (CQI). Consequently, the final decision matrix consists of four Service Failure Indicators (SFI₁–SFI₄) and one Composite Quality Index (CQI). The individual Complaint Resolution Indicators (CRI₁–CRI₃) are first integrated into the CQI through the Entropy weighting procedure described in Section 4 and are therefore not included separately in the MARCOS evaluation. This sequential procedure avoids double counting of complaint-related information while preserving the multidimensional representation of complaint-resolution performance.

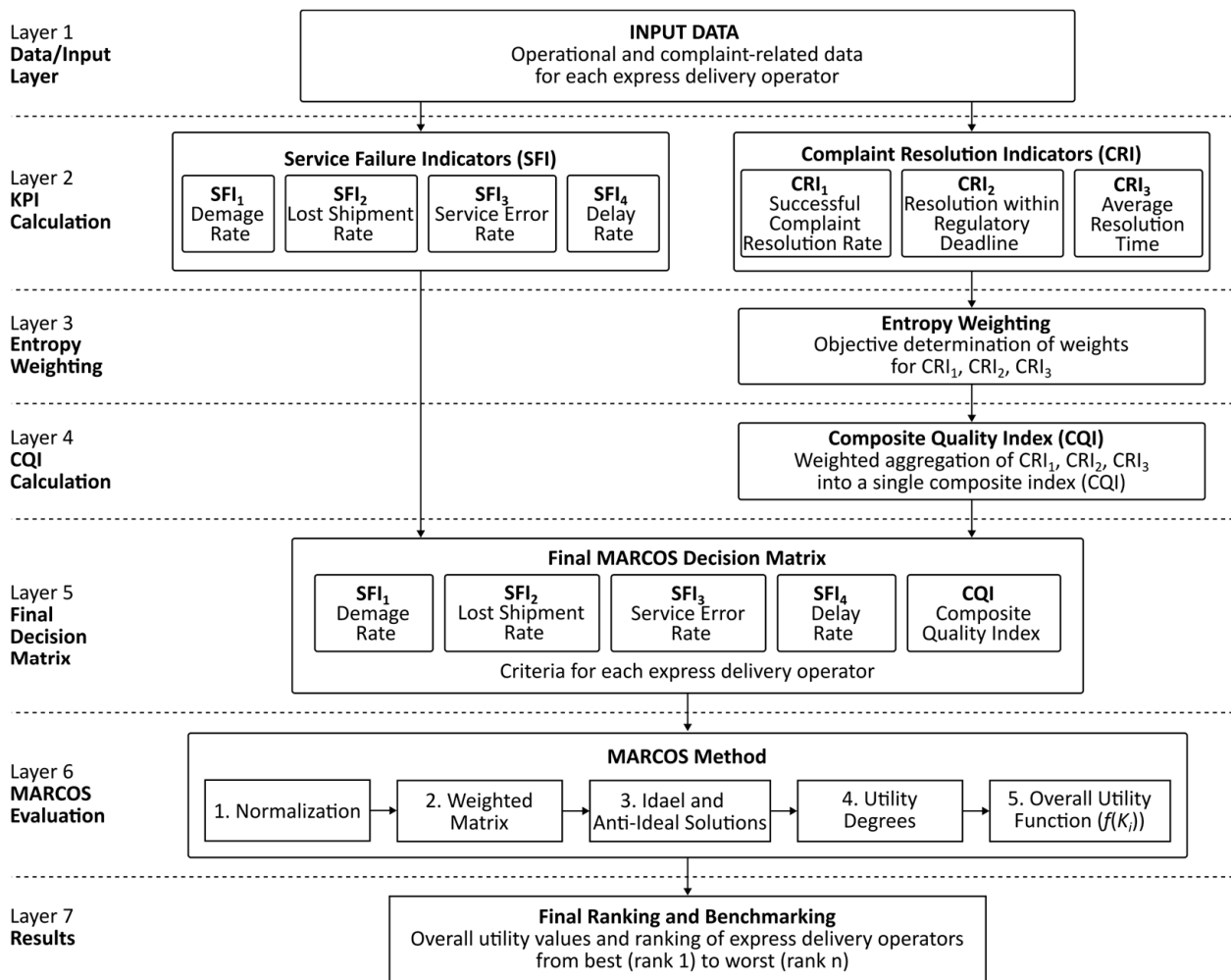


Figure 1. Sequential structure of the proposed SFI-CRI-CQI-MARCOS evaluation framework.

5.2. MARCOS Method

Although other well-established MCDM methods, such as TOPSIS, could also be applied to the proposed KPI framework, the selection of the MARCOS method was motivated by the characteristics of the present evaluation problem rather than by any assumption of methodological superiority. MARCOS was selected because it evaluates alternatives relative to both the ideal and anti-ideal solutions while providing utility functions that facilitate the interpretation of operator performance. The principal scientific contribution of this study lies in the development of the integrated SFI-CRI-CQI evaluation framework, whereas MARCOS serves as an appropriate decision-support tool for ranking express delivery operators within the proposed methodology.

The objective of the MARCOS method is to evaluate the relative performance of each operator by comparing it with the ideal and anti-ideal solutions constructed from the observed set of operators.

The application of the MARCOS method consists of the following steps.

Step 1: Formation of the Initial Decision Matrix. The initial decision matrix is defined as follows:

$$X = [x_{ij}]_{m \times n} \quad (13)$$

where

x_{ij} —value of criterion j for operator i ;

m —number of operators;

n —number of evaluation criteria.

The rows of the decision matrix represent the express delivery operators (alternatives), whereas the columns represent the evaluation criteria.

Step 2: Determination of Ideal and Anti-Ideal Solutions. After constructing the initial decision matrix, the ideal (AI) and anti-ideal (AAI) solutions are determined, and the extended initial decision matrix is constructed.

$$X = \begin{matrix} & \begin{matrix} C_1 & C_2 & \cdots & C_n \end{matrix} \\ \begin{matrix} AAI \\ A_1 \\ A_2 \\ \cdots \\ A_m \\ AI \end{matrix} & \begin{bmatrix} x_{aa1} & x_{aa2} & \cdots & x_{aan} \\ x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \\ x_{ai1} & x_{ai2} & \cdots & x_{ain} \end{bmatrix} \end{matrix} \quad (14)$$

The anti-ideal solution (AAI) represents a hypothetical alternative characterized by the least favorable values of all considered criteria, while the ideal solution (AI) represents a hypothetical alternative with the most favorable criterion values. According to the type of criterion (benefit or cost), the anti-ideal and ideal solutions are determined using Equations (15) and (16), respectively.

$$AAI = \min_i x_{ij} \text{ if } j \in B \quad \text{and} \quad \max_i x_{ij} \text{ if } j \in C \quad (15)$$

$$AI = \max_i x_{ij} \text{ if } j \in B \quad \text{and} \quad \min_i x_{ij} \text{ if } j \in C \quad (16)$$

where

B —the set of benefit criteria;

C —the set of cost criteria.

Step 3: Normalization of the Extended Decision Matrix. The elements of the extended decision matrix are normalized using different expressions for benefit and cost criteria.

For cost criteria:

$$n_{ij} = \frac{x_{ai}}{x_{ij}} \text{ if } j \in C \quad (17)$$

For benefit criteria:

$$n_{ij} = \frac{x_{ij}}{x_{ai}} \text{ if } j \in B \quad (18)$$

where elements x_{ij} and x_{ai} denote the elements of the extended decision matrix X .

Step 4: Construction of the Weighted Decision Matrix. Following the normalization procedure, the weighted decision matrix $V = [v_{ij}]_{m \times n}$ is constructed. In the general MARCOS procedure, criterion weights may be determined using an appropriate weighting method depending on the characteristics of the evaluation problem. In the proposed framework, however, equal weights are assigned to all evaluation criteria. This decision was made because the Composite Quality Index already incorporates entropy-based weighting of the Complaint Resolution Indicators. Consequently, assigning additional differential weights within the MARCOS procedure would result in repeated weighting of complaint-related information. The use of equal criterion weights therefore ensures methodological consistency while allowing the CQI to represent complaint-resolution performance as a single aggregated criterion within the final evaluation model.

The weighted values are calculated as follows:

$$v_{ij} = n_{ij} \times w_j \quad (19)$$

where

v_{ij} —weighted value of criterion j for operator i ;

n_{ij} —normalized value of criterion j for operator i ;

w_j —weight of criterion j obtained using the Entropy method.

The resulting weighted decision matrix represents the basis for determining the utility degrees of the alternatives and the subsequent ranking of express delivery operators using the MARCOS method.

Step 5: Determination of Utility Degrees. The weighted sum of the normalized criteria for each alternative is calculated as follows:

$$S_i = \sum_{j=1}^n v_{ij} \quad (20)$$

Based on the obtained weighted values, the utility degrees relative to the anti-ideal and ideal solutions are determined using the following expressions:

$$K_i^- = \frac{S_i}{S_{AAI}} \quad (21)$$

$$K_i^+ = \frac{S_i}{S_{AI}} \quad (22)$$

where

S_i —total weighted value of alternative i ;

S_{AAI} —total weighted value of the anti-ideal solution;

S_{AI} —total weighted value of the ideal solution.

Step 6: Determination of the Utility Function. After determining the utility degrees with respect to the ideal and anti-ideal solutions, the corresponding utility functions are calculated. These functions quantify the relative position of each alternative with respect to the ideal and anti-ideal reference solutions and provide the basis for the final evaluation of operator performance.

The utility function with respect to the anti-ideal solution is calculated as:

$$f(K_i^-) = \frac{K_i^+}{K_i^+ + K_i^-} \quad (23)$$

The utility function with respect to the ideal solution is calculated as:

$$f(K_i^+) = \frac{K_i^-}{K_i^+ + K_i^-} \quad (24)$$

where

$f(K_i^-)$ —utility function of alternative i with respect to the anti-ideal solution;

$f(K_i^+)$ —utility function of alternative i with respect to the ideal solution;

K_i^- —utility degree of alternative i relative to the anti-ideal solution;

K_i^+ —utility degree of alternative i relative to the ideal solution.

The final utility function is determined based on the relationship between each alternative and the ideal and anti-ideal solutions, as follows:

$$f(K_i) = \frac{K_i^+ + K_i^-}{\left[1 + \left(\frac{1-f(K_i^+)}{f(K_i^+)}\right) + \left(\frac{1-f(K_i^-)}{f(K_i^-)}\right)\right]} \quad (25)$$

Step 7: Ranking of Alternatives. The final ranking of express delivery operators is based on the overall utility function values, $(f(K_i))$, calculated in the previous step.

The utility function provides a comprehensive measure of alternative performance by simultaneously considering the relative position of each alternative with respect to the ideal and anti-ideal solutions.

The alternatives are ranked in descending order according to their utility function values. Higher values of $(f(K_i))$ indicate better overall performance, reflecting greater proximity to the ideal solution and a larger distance from the anti-ideal solution. Accordingly, the alternative with the highest utility function value is assigned the first rank, while the remaining alternatives are ranked in descending order of their utility values.

Within the proposed framework, the MARCOS method represents the final stage of the evaluation process. Rather than evaluating complaint-resolution indicators independently, the method integrates the Composite Quality Index with the Service Failure Indicators into a unified decision-making model. This sequential structure enables simultaneous evaluation of operational service quality and complaint-resolution performance while ensuring methodological consistency, preserving the complementary roles of SFI and CRI, and avoiding the duplication of complaint-related information.

The resulting ranking provides a robust basis for comparative performance evaluation, benchmarking, and evidence-based decision-making by identifying opportunities to improve service quality and complaint-management performance within the express delivery sector.

6. Case Study and Results

6.1. Case Study Description

To demonstrate the applicability of the proposed evaluation framework, an empirical case study was conducted using data collected from major express delivery operators operating in the Serbian market. The selection of operators was based on their market share in the express delivery segment.

According to the national methodology for monitoring the quality of express delivery services, quality evaluation is conducted for postal operators whose market share exceeds 5% of the total express delivery volume in the Republic of Serbia. Based on this criterion and the 2024 market data, five operators meeting the prescribed threshold were included in the analysis [56]. These operators collectively account for the majority of the national express delivery market and therefore provide a representative basis for comparative performance evaluation and regulatory benchmarking.

In accordance with the same methodology, the annual sample consists of 2000 test shipments. The sample is equally distributed among the selected operators, with 400 shipments assigned to each operator. Such a sampling design ensures comparability of results and provides a balanced basis for assessing service quality and complaint-handling performance across operators.

The analysis was performed using standardized regulatory quality indicators reported by express delivery operators in accordance with the national quality monitoring methodology. These indicators are aggregated at the operator level and are not disaggregated according to shipment categories (e.g., documents, parcels, fragile goods, or other commodity groups). Consequently, the proposed evaluation framework assesses the overall service quality and complaint-resolution performance of express delivery operators based on standardized regulatory data rather than shipment-type-specific service quality. This approach is consistent with the primary objective of the proposed framework, namely to support comparative regulatory evaluation and benchmarking of express delivery operators using harmonized performance indicators that are uniformly available for all operators.

The analysis was conducted over a one-year observation period and was based on operational and complaint-handling data collected through the monitoring process. The dataset included information on service failures, complaint resolution outcomes, compliance with regulatory deadlines, and complaint processing times. These data were subsequently used to calculate the Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), and the Composite Quality Index (CQI) for each operator.

The selected case study provides an appropriate environment for demonstrating the applicability of the proposed evaluation framework under real regulatory conditions. Since the analysis includes the largest express delivery operators monitored under a standardized national quality evaluation methodology, the obtained results provide a representative basis for validating the practical applicability of the proposed evaluation framework. Although the empirical evaluation is limited to the Serbian express delivery market and a single observation period, the methodological procedure itself is sufficiently general to be adapted to other postal markets where comparable operational and complaint-related data are available, subject to the applicable regulatory framework.

To ensure confidentiality and avoid potential commercial implications, the operators are anonymized and identified as O1, O2, O3, O4, and O5 throughout the study. The anonymization procedure does not affect the validity of the analysis, as the objective of the research is to evaluate the applicability of the proposed methodology rather than to assess the market position of individual operators.

The empirical analysis was carried out in several stages. First, the performance indicators were calculated for each operator. Subsequently, criterion weights were determined using the Entropy method, and the Composite Quality Index (CQI) was calculated. Finally, the MARCOS method was applied to evaluate and rank the operators according to their overall performance.

The proposed framework enables a comprehensive evaluation of service quality and complaint resolution performance in the express delivery sector. Its practical application is demonstrated in the following sections through the calculation of Entropy criterion weights, the construction of the Composite Quality Index (CQI), the development of the MARCOS decision matrix, and the final evaluation and ranking of express delivery operators.

6.2. Decision Matrix Construction

The construction of the decision matrix represents the first step in the empirical evaluation of express delivery operators. In accordance with the proposed evaluation framework, the analysis was based on four Service Failure Indicators (SFI) and three Complaint Resolution Indicators (CRI).

The selected indicators provide a comprehensive representation of operator performance by simultaneously considering operational service quality and complaint resolution efficiency. Service Failure Indicators capture the most significant operational deficiencies occurring during the delivery process, while Complaint Resolution Indicators assess the effectiveness and timeliness of complaint handling procedures.

According to their characteristics, the Service Failure Indicators (SFI₁–SFI₄) and Average Resolution Time (CRI₃) were treated as cost criteria, since lower values indicate better performance. In contrast, Successful Complaint Resolution Rate (CRI₁) and Resolution within Regulatory Deadline (CRI₂) were considered benefit criteria because higher values reflect more efficient complaint management. The indicators included in the empirical analysis are presented in Table 2.

The selected indicators represent two complementary dimensions of operator performance. While the Service Failure Indicators assess operational reliability during service provision, the Complaint Resolution Indicators evaluate the effectiveness and efficiency of

complaint handling. Their combined use enables a more comprehensive evaluation of operator performance than evaluating either operational service quality or complaint-handling performance independently.

Table 2. Evaluation Indicators.

Criterion	Description	Unit	Criterion Type
SFI ₁	Damage Rate	%	Cost
SFI ₂	Lost Shipment Rate	%	Cost
SFI ₃	Service Error Rate	%	Cost
SFI ₄	Delay Rate	%	Cost
CRI1	Successful Complaint Resolution Rate	–	Benefit
CRI2	Resolution within Regulatory Deadline	–	Benefit
CRI3	Average Resolution Time	Days	Cost

Based on the collected data, the values of the Service Failure Indicators and Complaint Resolution Indicators were calculated for the five observed operators. The resulting indicator values are presented in Table 3.

Table 3. Service Failure and Complaint Resolution Indicators for the Observed Operators.

Operator	SFI ₁	SFI ₂	SFI ₃	SFI ₄	CRI1	CRI2	CRI3
O1	12.5	5.0	20.0	30.0	0.8889	0.9259	12.4
O2	15.0	2.5	22.5	30.0	0.8571	0.8929	12.2
O3	20.0	2.5	27.5	27.5	0.9355	0.8667	10.7
O4	17.5	5.0	25.0	32.5	0.8750	0.8667	12.8
O5	12.5	7.5	20.0	35.0	0.8000	1.0000	13.3

The values presented in Table 2 constitute the basis for the calculation of the Composite Quality Index (CQI). The entropy-based weighting procedure described in Section 4 was applied to the Complaint Resolution Indicators (CRI1–CRI3), and the obtained weights were subsequently used to calculate the CQI values for each operator.

At this stage of the evaluation procedure, the Complaint Resolution Indicators are analysed individually in order to determine their objective entropy weights and construct the Composite Quality Index. Consequently, Table 3 should be interpreted as an intermediate decision matrix used exclusively for the development of the CQI rather than for the final multi-criteria evaluation.

The resulting CQI values were then combined with the Service Failure Indicators to construct the final decision matrix used in the MARCOS evaluation and ranking procedure.

Accordingly, the three individual Complaint Resolution Indicators are replaced by the Composite Quality Index in the final MARCOS decision matrix. This sequential procedure preserves the multidimensional representation of complaint-resolution performance while avoiding double counting of complaint-related information during the final multi-criteria evaluation.

6.3. Entropy Weight Calculation and CQI Determination

To develop the Composite Quality Index (CQI), the Entropy weighting procedure described in Section 4 was applied to the Complaint Resolution Indicators (CRI1–CRI3). The procedure included data normalization, calculation of entropy values, determination of divergence measures, and computation of criterion weights using Equations (5)–(10). The obtained entropy values, divergence values, and final criterion weights are presented in Table 4.

Table 4. Entropy Weights of Complaint Resolution Indicators.

Criterion	Entropy (e_j)	Divergence (d_j)	Weight (w_j)
<i>CRI1</i>	0.999204	0.000796	0.2398
<i>CRI2</i>	0.999089	0.000911	0.2745
<i>CRI3</i>	0.998389	0.001611	0.4857
Total	–	–	1.0000

The results indicate that *CRI3* (Average Resolution Time) has the highest weight (0.4857), suggesting that this indicator plays the most important role in distinguishing the complaint resolution performance of the observed operators. *CRI2* ranks second with a weight of 0.2745, while *CRI1* has the lowest weight (0.2398). Consequently, the efficiency and timeliness of complaint handling have the greatest influence on the overall evaluation of complaint resolution performance.

It should be noted that the obtained Entropy weights reflect the discriminatory power of the analysed indicators within the observed dataset rather than their intrinsic managerial importance. The higher weight assigned to *CRI3* indicates that this indicator exhibits the greatest variability among the analysed operators and therefore contributes most to differentiating their complaint-resolution performance. This interpretation is fully consistent with the objective nature of the Entropy weighting method described in Section 4.3.

Following the determination of criterion weights, the Composite Quality Index (CQI) was calculated for each operator using Equations (11) and (12). Since *CRI3* represents a cost criterion, it was first transformed into a benefit-type criterion through the normalization procedure described in Section 4. The resulting normalized values were then combined with the entropy-based weights to obtain the final CQI scores. The calculated CQI values for the observed operators are presented in Table 5.

Table 5. Composite Quality Index (CQI) Values.

Operator	CQI
O1	0.5006
O2	0.4923
O3	0.5578
O4	0.4667
O5	0.4663

The results reveal noticeable differences in complaint resolution performance among the analyzed operators.

Operator O3 achieved the highest CQI value (0.5578), indicating the most effective overall complaint resolution performance. This result can be attributed to its strong performance across all complaint resolution indicators, particularly its high successful complaint resolution rate and the shortest average complaint resolution time among the observed operators. Operator O1 ranked second with a CQI value of 0.5006, followed closely by O2 (0.4923).

The lowest CQI values were recorded for operators O4 (0.4667) and O5 (0.4663). Although O5 achieved the highest value of *CRI2*, indicating full compliance with regulatory deadlines, its relatively low successful complaint resolution rate and the longest average complaint resolution time negatively affected its overall CQI score.

Overall, the results suggest that efficient complaint handling is influenced not only by compliance with regulatory deadlines but also by the ability to successfully resolve complaints and minimize complaint processing time. Consequently, operators demonstrating

balanced performance across all complaint resolution indicators tend to achieve higher CQI values.

The Composite Quality Index should therefore be interpreted as a synthetic measure of complaint-resolution performance rather than as a replacement for the individual Complaint Resolution Indicators. Its primary role is to integrate the complementary dimensions of complaint handling into a single performance measure that can subsequently be incorporated into the final multi-criteria evaluation together with the Service Failure Indicators. This aggregation simplifies the evaluation process while preserving the information contained in the underlying complaint-resolution indicators through the Entropy weighting procedure.

The CQI values presented in Table 5 summarize the overall effectiveness of complaint resolution processes and are subsequently incorporated into the final decision matrix used for MARCOS-based evaluation and ranking.

6.4. MARCOS Evaluation and Ranking

Following the calculation of the Composite Quality Index (CQI), the final decision matrix was constructed and used as the input for the MARCOS evaluation procedure. In accordance with the proposed framework, the matrix included four Service Failure Indicators (SFI₁–SFI₄) and the Composite Quality Index (CQI), thereby integrating both operational service quality and complaint resolution performance into a single evaluation model.

The use of the Composite Quality Index in the final decision matrix represents the final stage of the sequential evaluation framework proposed in this study. Instead of including the three Complaint Resolution Indicators as separate evaluation criteria, they are first objectively aggregated into a single composite indicator using the Entropy weighting procedure. Consequently, the MARCOS evaluation simultaneously considers operational service quality and complaint-resolution performance while avoiding double counting of complaint-related information.

The Service Failure Indicators were treated as cost criteria because lower values indicate better service performance, whereas the Composite Quality Index was considered a benefit criterion, since higher values reflect superior complaint resolution performance. The final decision matrix used in the MARCOS evaluation is presented in Table 6.

Table 6. Final Decision Matrix.

Operator	SFI ₁	SFI ₂	SFI ₃	SFI ₄	CQI
O1	12.5	5.0	20.0	30.0	0.5006
O2	15.0	2.5	22.5	30.0	0.4923
O3	20.0	2.5	27.5	27.5	0.5578
O4	17.5	5.0	25.0	32.5	0.4667
O5	12.5	7.5	20.0	35.0	0.4663

The MARCOS procedure was subsequently applied following the steps described in Section 5.2. First, the anti-ideal and ideal solutions were determined for each criterion. The resulting extended decision matrix is presented in Table 7.

Since the Composite Quality Index already incorporates entropy-based weighting of complaint resolution indicators, an additional weighting procedure was not applied within the MARCOS evaluation. To ensure methodological consistency and avoid double weighting of complaint-related information, equal weights were assigned to all evaluation criteria. The criterion weights used in the MARCOS analysis are presented in Table 7.

Table 7. Criterion Weights Used in the MARCOS Evaluation.

Criterion	Weight
SFI ₁	0.20
SFI ₂	0.20
SFI ₃	0.20
SFI ₄	0.20
CQI	0.20
Total	1.00

The assignment of equal weights within the MARCOS procedure should not be interpreted as assuming equal managerial importance of all evaluation criteria. Rather, it reflects the sequential structure of the proposed methodology. Since the relative importance of complaint-resolution performance has already been incorporated into the Composite Quality Index through the Entropy weighting procedure, introducing additional differential weights at this stage would disproportionately increase the influence of complaint-related information. Therefore, equal criterion weights were adopted to ensure methodological consistency throughout the final multi-criteria evaluation.

The first step of the MARCOS procedure involves the determination of the anti-ideal (AAI) and ideal (AI) solutions for each criterion. The anti-ideal solution represents the least favorable performance, whereas the ideal solution corresponds to the most favorable performance among the observed operators. Based on these values, the initial decision matrix is extended by adding the anti-ideal and ideal alternatives.

For cost criteria (SFI₁–SFI₄), the anti-ideal solution was defined as the maximum observed value, while the ideal solution corresponded to the minimum observed value. For the benefit criterion (CQI), the anti-ideal and ideal solutions were determined using the minimum and maximum values, respectively. The resulting extended decision matrix is presented in Table 8.

Table 8. Extended Decision Matrix.

Alternative	SFI ₁	SFI ₂	SFI ₃	SFI ₄	CQI
AAI	20.0	7.5	27.5	35.0	0.4663
O1	12.5	5.0	20.0	30.0	0.5006
O2	15.0	2.5	22.5	30.0	0.4923
O3	20.0	2.5	27.5	27.5	0.5578
O4	17.5	5.0	25.0	32.5	0.4667
O5	12.5	7.5	20.0	35.0	0.4663
AI	12.5	2.5	20.0	27.5	0.5578

Following the construction of the extended decision matrix, the normalization procedure was performed according to the criterion type using Equations (15) and (16). For cost criteria (SFI₁–SFI₄), normalization was carried out by dividing the ideal value by the observed value. For the benefit criterion (CQI), normalization was performed by dividing the observed value by the ideal value. The resulting normalized matrix is presented in Table 9.

The normalized values provide a comparable scale for all criteria and represent the basis for constructing the weighted decision matrix in the next step.

After normalization, the weighted normalized matrix was constructed by multiplying the normalized criterion values by their corresponding weights. Since equal weights were assigned to all criteria, each normalized value was multiplied by 0.20. The resulting weighted normalized matrix is presented in Table 10.

Table 9. Normalized Matrix.

Alternative	SFI ₁	SFI ₂	SFI ₃	SFI ₄	CQI
AAI	0.6250	0.3333	0.7273	0.7857	0.8361
O1	1.0000	0.5000	1.0000	0.9167	0.8975
O2	0.8333	1.0000	0.8889	0.9167	0.8825
O3	0.6250	1.0000	0.7273	1.0000	1.0000
O4	0.7143	0.5000	0.8000	0.8462	0.8367
O5	1.0000	0.3333	1.0000	0.7857	0.8361
AI	1.0000	1.0000	1.0000	1.0000	1.0000

Table 10. Weighted Normalized Matrix.

Alternative	SFI ₁	SFI ₂	SFI ₃	SFI ₄	CQI
AAI	0.1250	0.0667	0.1455	0.1571	0.1672
O1	0.2000	0.1000	0.2000	0.1833	0.1795
O2	0.1667	0.2000	0.1778	0.1833	0.1765
O3	0.1250	0.2000	0.1455	0.2000	0.2000
O4	0.1429	0.1000	0.1600	0.1692	0.1673
O5	0.2000	0.0667	0.2000	0.1571	0.1672
AI	0.2000	0.2000	0.2000	0.2000	0.2000

The weighted normalized matrix represents the basis for calculating the utility degrees of the alternatives relative to the anti-ideal and ideal solutions. These values are subsequently used to determine the overall utility functions and the final ranking of the observed operators.

Following the construction of the weighted normalized matrix, the utility values were calculated according to the MARCOS procedure. First, the sums of weighted normalized values (S_i) were determined using Equation (20). The utility degrees relative to the anti-ideal and ideal solutions (K_i^- and K_i^+) were then calculated using Equations (21) and (22). Based on these values, the utility functions $f(K_i^-)$, $f(K_i^+)$, and the overall utility function $f(K_i)$ were obtained using Equations (23)–(25). The resulting utility values and the final ranking of the observed operators are presented in Table 11.

Table 11. Utility Degrees and Ranking Results.

Operator	S_i	K_i^-	K_i^+	$f(K_i^-)$	$f(K_i^+)$	$f(K_i)$	Rank
O1	0.8628	1.3043	0.8628	0.3981	0.6019	0.6829	3
O2	0.9043	1.3670	0.9043	0.3981	0.6019	0.7158	1
O3	0.8705	1.3159	0.8705	0.3981	0.6019	0.6890	2
O4	0.7394	1.1178	0.7394	0.3981	0.6019	0.5853	5
O5	0.7910	1.1958	0.7910	0.3981	0.6019	0.6261	4

As shown in Table 11, the values of $f(K_i^-)$, and $f(K_i^+)$ are identical for all operators because they are constants derived solely from the anti-ideal (AAI) and ideal (AI) reference solutions in accordance with the MARCOS methodology. Since these values depend only on the reference solutions rather than on the individual alternatives, they remain constant across all operators. Consequently, the overall utility function $f(K_i)$, is the only value that distinguishes operator performance and determines the final ranking of express delivery operators.

The results show that operator O2 achieved the highest overall utility value ($f(K_i) = 0.7158$), and therefore ranked first. Operator O3 ranked second, followed by

O1. Operators O5 and O4 achieved lower utility values, with O4 ranked last among the observed operators.

The obtained ranking confirms the importance of evaluating operational service quality and complaint-resolution performance within a unified decision-making framework. Although operator O3 achieved the highest Composite Quality Index, reflecting superior complaint-handling performance, its comparatively weaker results for several Service Failure Indicators reduced its overall ranking. Conversely, operator O2 achieved a more balanced performance across both operational and complaint-related dimensions, which ultimately resulted in the highest overall utility value. This finding demonstrates that superior complaint handling alone is insufficient to achieve the highest overall performance if it is not accompanied by consistently high operational service quality.

6.5. Sensitivity Analysis

To examine the robustness of the obtained ranking, a sensitivity analysis was conducted using five weighting scenarios: one reference scenario and four alternative scenarios. The reference scenario retained equal weights of 0.20 for all criteria. In the alternative scenarios, greater importance was assigned to complaint-resolution performance, overall operational performance, lost-shipment performance, and delivery-delay performance, respectively. In each scenario, the criterion weights were normalized so that their sum remained equal to one (Table 12).

Table 12. Definition of weighting scenarios used in the sensitivity analysis.

Scenario	Criterion Weights (SFI1, SFI2, SFI3, SFI4, CQI)
S0—Equal-weights reference scenario	(0.20, 0.20, 0.20, 0.20, 0.20)
S1—CQI-priority scenario	(0.15, 0.15, 0.15, 0.15, 0.40)
S2—Operational-performance-priority scenario	(0.225, 0.225, 0.225, 0.225, 0.10)
S3—Lost-shipment-priority scenario	(0.15, 0.40, 0.15, 0.15, 0.15)
S4—Delivery-delay-priority scenario	(0.15, 0.15, 0.15, 0.40, 0.15)

The sensitivity analysis indicates that the ranking is generally stable across the considered weighting scenarios (Table 13). Operator O2 retains the first position in four of the five scenarios, whereas O3 becomes the highest-ranked operator only in the CQI-priority scenario. Operator O1 remains in the second or third position across all scenarios, while O4 and O5 generally occupy the two lowest positions. The observed rank reversals are limited to operators with relatively similar performance profiles and occur when greater importance is assigned to specific criteria. These results confirm the overall robustness of the proposed evaluation framework while demonstrating its expected responsiveness to different regulatory or managerial priorities.

Table 13. Overall utility values and operator rankings under alternative weighting scenarios.

Operator	S0 $f(K_i)$	Rank	S1 $f(K_i)$	Rank	S2 $f(K_i)$	Rank	S3 $f(K_i)$	Rank	S4 $f(K_i)$	Rank
O1	0.6829	3	0.6747	3	0.6871	2	0.6368	3	0.6828	3
O2	0.7158	1	0.6959	2	0.7259	1	0.7655	1	0.7070	1
O3	0.6890	2	0.6990	1	0.6837	3	0.7446	2	0.7035	2
O4	0.5853	5	0.5913	5	0.5820	5	0.5604	4	0.5970	5
O5	0.6261	4	0.6212	4	0.6285	4	0.5580	5	0.6153	4

The sensitivity analysis also provides an initial validation of the proposed evaluation framework. The stability of the obtained rankings under alternative weighting scenarios indicates that the framework is not overly sensitive to moderate changes in criterion

importance and is capable of producing consistent evaluation results. Furthermore, the framework is constructed using standardized regulatory performance indicators, an objective Entropy weighting procedure, and the well-established MARCOS decision-making method, which together contribute to its transparency, reproducibility, and practical applicability. Although additional validation using datasets from other regulatory environments would further strengthen the generalizability of the proposed framework, the present results demonstrate its capability to provide meaningful and robust comparative evaluation of express delivery operators.

6.6. Discussion of Results

The results obtained using the proposed evaluation framework demonstrate that the integrated evaluation of operational service quality and complaint-resolution performance provides a more comprehensive understanding of express delivery operator performance than evaluating either dimension independently. While the Service Failure Indicators quantify the reliability of the operational delivery process, the Composite Quality Index summarizes the effectiveness of complaint handling into a single synthetic performance measure. Their integration within the MARCOS framework therefore enables a balanced evaluation of operator performance by considering both service execution and post-service customer support.

An important finding of the empirical analysis is that superior complaint-resolution performance alone does not necessarily lead to the highest overall ranking. Although operator O3 achieved the highest Composite Quality Index, reflecting the most effective complaint-handling performance, its weaker results for several Service Failure Indicators reduced its overall performance within the MARCOS evaluation. Conversely, operator O2 achieved the highest overall ranking by maintaining consistently good performance across both operational service quality and complaint-resolution dimensions. This result confirms that balanced performance across multiple quality dimensions is more important than excellence in a single performance area.

These findings are consistent with previous studies emphasizing that service quality in postal and express delivery systems is inherently multidimensional and cannot be adequately represented by a single performance indicator. Earlier research has shown that operational reliability, complaint handling, and customer-oriented quality management jointly influence overall service quality. The proposed framework extends existing approaches by integrating operational service failures and complaint-resolution performance into a unified evaluation model, thereby providing a more comprehensive basis for evaluating express delivery operator performance than approaches relying exclusively on operational indicators or customer complaint statistics.

From a regulatory perspective, the proposed framework provides a transparent benchmarking tool that enables supervisory authorities to identify operators requiring performance improvements in specific quality dimensions. Instead of evaluating operational failures and complaint handling separately, regulators may use the proposed evaluation framework to support evidence-based monitoring, quality control, and regulatory decision-making. Similarly, express delivery operators may use the obtained results to identify weaknesses in operational processes or complaint management and to prioritize improvement measures according to their relative impact on overall performance.

The proposed methodology also has broader applicability beyond the analysed case study. Although the empirical evaluation was conducted using data from the Serbian express delivery market, the methodological framework relies on standardized performance indicators, objective weighting through the Entropy method, and the MARCOS multi-criteria decision-making approach. Consequently, the framework can be adapted to other

postal markets where comparable operational and complaint-related data are available, subject to the characteristics of the respective regulatory environment.

Nevertheless, several limitations of the present study should be acknowledged. First, the empirical analysis was conducted using data from one national postal market and one observation period, which limits the generalizability of the obtained results. Second, the Composite Quality Index was constructed using only those Complaint Resolution Indicators that are systematically collected within the national regulatory reporting system. Other dimensions of complaint handling, such as customer satisfaction, perceived fairness of complaint resolution, or accessibility of complaint channels, were not included because comparable data were unavailable. Future research may therefore extend the proposed framework by incorporating additional quality dimensions, alternative weighting approaches, and longitudinal datasets covering multiple regulatory environments.

Overall, the empirical results confirm that the proposed evaluation framework provides an objective and practically applicable approach for assessing the performance of express delivery operators. By integrating operational service quality and complaint-resolution performance within a unified decision-support framework, the methodology offers a valuable tool for comparative performance evaluation, benchmarking, and quality improvement in the postal and express delivery sector.

7. Conclusions

This study proposed an integrated evaluation framework for assessing the performance of express delivery operators by simultaneously considering operational service quality and complaint-resolution performance. Unlike conventional approaches that evaluate these dimensions independently, the proposed framework combines Service Failure Indicators (SFI), Complaint Resolution Indicators (CRI), the Entropy weighting method, the Composite Quality Index (CQI), and the MARCOS multi-criteria decision-making method into a unified decision-support framework.

The proposed methodology follows a sequential evaluation procedure. First, operational performance is assessed through Service Failure Indicators, while complaint-resolution performance is evaluated using three standardized Complaint Resolution Indicators. Subsequently, the Entropy method is applied to determine the objective weights of the complaint-resolution indicators, which are aggregated into the Composite Quality Index. Finally, the CQI is integrated with the Service Failure Indicators within the MARCOS framework to provide a comprehensive evaluation and ranking of express delivery operators. This sequential structure enables the integration of complementary quality dimensions while avoiding duplication of complaint-related information in the final evaluation.

The empirical case study conducted using data from the Serbian express delivery market demonstrated the applicability of the proposed framework under real regulatory conditions. The obtained results showed that the highest overall performance is achieved not necessarily by operators excelling in a single quality dimension, but by those maintaining balanced performance across both operational service quality and complaint-resolution processes. The findings confirm that joint evaluation of these dimensions provides a more comprehensive understanding of operator performance than analysing them independently.

The proposed framework makes several contributions to the existing literature. First, it introduces an integrated methodology that combines operational service quality and complaint-resolution performance within a single evaluation framework. Second, it develops a Composite Quality Index that objectively summarizes complaint-handling performance using Entropy-based weighting. Third, it demonstrates how the aggregated complaint-resolution indicator can be incorporated into a multi-criteria decision-making

model without duplicating complaint-related information. Finally, the study extends the application of the MARCOS method to the evaluation of express delivery operators by integrating operational and customer-oriented quality dimensions into a unified evaluation framework.

From a practical perspective, the proposed framework provides a transparent and objective tool for benchmarking operator performance and identifying areas requiring improvement. Regulatory authorities may use the framework to support evidence-based monitoring of service quality and complaint handling, while express delivery operators may employ the obtained results to identify operational weaknesses, improve complaint-management processes, and prioritize quality improvement initiatives.

Nevertheless, several limitations should be acknowledged. The empirical evaluation was conducted using data from a single national postal market and one observation period. Furthermore, the Composite Quality Index was constructed using only those Complaint Resolution Indicators that are systematically collected within the national regulatory reporting system. Consequently, other potentially relevant dimensions of complaint handling, such as customer satisfaction, perceived fairness of complaint resolution, or accessibility of complaint channels, were not included because comparable data were unavailable.

One limitation of the present study is that the regulatory dataset does not distinguish complaints according to shipment categories or commodity types. Consequently, differences in complaint characteristics associated with fragile goods, electronic devices, documents, or other shipment categories cannot be analysed separately. Future research could address this limitation by incorporating category-specific performance indicators and complaint-type-specific resolution times whenever such disaggregated regulatory data become available, thereby enabling a more detailed evaluation of complaint-resolution performance while preserving the regulatory benchmarking capability of the proposed framework. Future research may also extend the proposed framework by applying it to different postal markets and regulatory environments, analysing longitudinal datasets, incorporating additional complaint-related quality indicators, and comparing alternative objective and hybrid weighting methods. Such studies would further examine the robustness, transferability, and practical applicability of the proposed evaluation framework. Although the proposed framework has been demonstrated using real operational data from express delivery operators, future research should focus on comprehensive methodological validation through sensitivity analysis, comparative application of alternative MCDM methods, and empirical testing in different regulatory and operational environments.

Overall, the results demonstrate that integrating operational service quality and complaint-resolution performance within a unified evaluation framework provides a more comprehensive basis for assessing the performance of express delivery operators than treating these dimensions separately. The proposed methodology therefore represents a useful decision-support tool for researchers, regulators, and postal operators seeking to improve service quality through objective, transparent, and reproducible performance evaluation.

Author Contributions: Conceptualization, N.K., M.B. and M.D.; methodology, N.K., M.B. and A.Č.; software, D.L.; validation, N.K., M.B. and D.L.; formal analysis, N.K., M.B. and A.Č.; investigation, N.K., M.B. and D.L.; resources, N.K., M.B. and M.D.; data curation, N.K., M.B. and M.D.; writing—original draft preparation, N.K., M.B., A.Č., D.L. and M.D.; writing—review and editing, N.K., M.B., A.Č., D.L. and M.D.; visualization, N.K.; supervision, N.K. and M.D.; project administration, N.K. and M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. UNCTAD. *E-Commerce and Digital Economy Programme: Year in Review 2023*; United Nations Conference on Trade and Development: Geneva, Switzerland, 2023. Available online: <https://unctad.org/publication/e-commerce-and-digital-economy-programme-year-review-2023> (accessed on 20 June 2026).
2. Universal Postal Union (UPU). *State of the Postal Sector 2023: A Hyper-Collaborative Path to Postal Development*; UPU Publishing: Bern, Switzerland, 2023. Available online: <https://www.upu.int/en/publications/2ipd/the-state-of-the-postal-sector-2023> (accessed on 20 June 2026).
3. Pateiro-Rodríguez, C.; Martín-Bermúdez, F.; Pateiro-López, C.; Escourido-Calvo, M. Universal Service Regulation and Network Effects in Services of General Economic Interest in the European Union. *Economies* **2024**, *12*, 280. [CrossRef]
4. Hearn, J. *The Confluence of Competition Law, Consumer Protection, and Sector-Specific Regulation in the Postal, Delivery, and Related Sectors*; Springer: Cham, Switzerland, 2023.
5. Parasuraman, A.; Zeithaml, V.A.; Berry, L.L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *J. Retail.* **1988**, *64*, 12–40.
6. Thai, V. Logistics service quality: Conceptual model and empirical evidence. *Int. J. Logist. Res. Appl.* **2013**, *16*, 114–131. [CrossRef]
7. Vakulenko, Y.; Shams, P.; Hellström, D.; Hjort, K. Online retail experience and customer satisfaction: The mediating role of last-mile delivery. *Int. Rev. Retail Distrib. Consum. Res.* **2019**, *29*, 306–320. [CrossRef]
8. Blagojević, M.; Knežević, N.; Marković, D. Assessment of Customer Satisfaction with Postal Services—A Statistical Approach. *Zagreb Int. Rev. Econ. Bus.* **2024**, *27*, 43–74. [CrossRef]
9. James, N.; Kumar, A.K.P.; Nathan, R.J. Logistic Service Improvement Parameters for Postal Service Providers Using Analytical Hierarchy Process and Quality Function Deployment. *Economies* **2025**, *13*, 120. [CrossRef]
10. Aljohani, K. The Role of Last-Mile Delivery Quality and Satisfaction in Online Retail Experience: An Empirical Analysis. *Sustainability* **2024**, *16*, 4743. [CrossRef]
11. European Regulators Group for Postal Services (ERGP). *Report on the Contractual Situation of Consumers of Postal Services*; ERGP: Brussels, Belgium, 2021.
12. European Regulators Group for Postal Services (ERGP). *Report on QoS, Consumer Protection and Complaint Handling—An Analysis of Trends*; ERGP Publishing: Brussels, Belgium, 2024. Available online: <https://ec.europa.eu/docsroom/documents/57555/attachments/1/translations/en/renditions/native> (accessed on 15 June 2026).
13. Tsai, J.-F.; Ngo, H.-N.; Che, Z.-H.; Lin, M.-H. Sustainable Solutions in E-Commerce: An Examination of Customer Acceptance of Automated Delivery Stations in Vietnam. *Sustainability* **2024**, *16*, 4570. [CrossRef]
14. Chen, Y.; Ryu, M.H. Sustainable Logistics: Exploring the Determinants of Consumer Attitudes and Intention to Use Toward Autonomous Delivery Services. *Sustainability* **2025**, *17*, 3290. [CrossRef]
15. Hipólito, F.; Dias, Á.; Pereira, L. Influence of Consumer Trust, Return Policy, and Risk Perception on Satisfaction with the Online Shopping Experience. *Systems* **2025**, *13*, 158. [CrossRef]
16. Morić, Z.; Dakić, V.; Djekić, D.; Regvart, D. Protection of Personal Data in the Context of E-Commerce. *J. Cybersecur. Priv.* **2024**, *4*, 731–761. [CrossRef]
17. Üstündağlı Erten, E. Complaint Management through the E-State Portal: Is Digitalization Actually Beneficial? *Proceedings* **2024**, *101*, 1. [CrossRef]
18. Adzhigalieva, Z.; Hurriyati, R.; Hendrayati, H. The Influence of Complaint Handling and Service Recovery on Customer Satisfaction, Customer Loyalty, and Customer Retention. In *Proceedings of the 6th Global Conference on Business, Management, and Entrepreneurship (GCBME 2021)*, Bandung, Indonesia, 18 August 2021; pp. 281–285. [CrossRef]
19. *ISO 10002:2018; Customer Satisfaction—Guidelines for Complaints Handling in Organizations*. International Organization for Standardization (ISO): Geneva, Switzerland, 2018.
20. Tax, S.S.; Brown, S.W.; Chandrashekar, M. Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing. *J. Mark.* **1998**, *62*, 60–76. [CrossRef]
21. Arabelen, G.; Kaya, H.T. Assessment of Logistics Service Quality Dimensions: A Qualitative Approach. *J. Shipp. Trade* **2021**, *6*, 14. [CrossRef]

22. Blandón Andrade, J.C.; Castaño Toro, A.; Morales Ríos, A.; Orozco Ospina, D. Computational Methods for Information Processing from Natural Language Complaint Processes—A Systematic Review. *Computers* **2025**, *14*, 28. [\[CrossRef\]](#)
23. Cai, H.; Shao, X.; Zhou, P.; Li, H. Multi-Label Classification of Complaint Texts: Civil Aviation Service Quality Case Study. *Electronics* **2025**, *14*, 434. [\[CrossRef\]](#)
24. Ho, C.-I.; Liu, Y.; Chen, M.-C. Assessment of Customers' Evaluations of Service Quality in Live-Streaming Commerce: Conceptualizing and Testing a Multidimensional and Hierarchical Model. *Information* **2024**, *15*, 510. [\[CrossRef\]](#)
25. Vicente, H.; Fernandes, A.; Neves, J.; Figueiredo, M. Enhancing Service Quality—A Customer Opinion Assessment in Water Laboratories through Artificial Neural Networks. *Appl. Sci.* **2024**, *14*, 7626. [\[CrossRef\]](#)
26. Aulin, V.; Liashuk, O.; Mironov, D.; Staliński, P.; Rutkowski, M.; Lysenko, S. Management-Oriented Assessment of Transport Service Quality Using Logistics Monitoring System and Harrington's Desirability Function in Support of SDG 9. *Sustainability* **2025**, *17*, 7837. [\[CrossRef\]](#)
27. Shuaibu, A.S.; Mahmoud, A.S.; Sheltami, T.R. A Review of Last-Mile Delivery Optimization: Strategies, Technologies, Drone Integration, and Future Trends. *Drones* **2025**, *9*, 158. [\[CrossRef\]](#)
28. Jovanović, B.; Stević, Ž.; Mitrović Simić, J.; Stupar, A.; Kopic, M. Assessment of Efficiency of Last-Mile Delivery Zones: A Novel IRN OWCM-IRN AROMAN Model. *Mathematics* **2025**, *13*, 2845. [\[CrossRef\]](#)
29. Muharemović, E.; Kosovac, A.; Begović, M.; Tadić, S.; Krstić, M. Cost Modeling for Pickup and Delivery Outsourcing in CEP Operations: A Multidimensional Approach. *Logistics* **2025**, *9*, 96. [\[CrossRef\]](#)
30. Čižiūnienė, K.; Draugelytė, G.; Sokolovskij, E.; Matijošius, J. Improving Small Parcel Delivery Efficiency and Sustainability: A Study of Lithuanian Private Delivery Company. *Sustainability* **2025**, *17*, 1838. [\[CrossRef\]](#)
31. Ferreira, J.C.; Esperança, M. Enhancing Sustainable Last-Mile Delivery: The Impact of Electric Vehicles and AI Optimization on Urban Logistics. *World Electr. Veh. J.* **2025**, *16*, 242. [\[CrossRef\]](#)
32. Picozzi, P.; Nocco, U.; Pezzillo, A.; De Cosmo, A.; Cimolin, V. The Use of Business Intelligence Software to Monitor Key Performance Indicators (KPIs) for the Evaluation of a Computerized Maintenance Management System (CMMS). *Electronics* **2024**, *13*, 2286. [\[CrossRef\]](#)
33. Nardo, M.; Saisana, M.; Saltelli, A.; Tarantola, S. *Tools for Composite Indicators Building*; EUR 21682 EN; European Commission, Joint Research Centre: Ispra, Italy, 2005.
34. OECD/JRC. *Handbook on Constructing Composite Indicators: Methodology and User Guide*; OECD Publishing: Paris, France, 2008. [\[CrossRef\]](#)
35. Greco, S.; Ishizaka, A.; Tasiou, M.; Torrisi, G. On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Soc. Indic. Res.* **2019**, *141*, 61–94. [\[CrossRef\]](#)
36. Kara, P.; Valentin, J.B.; Mainz, J.; Johnsen, S.P. Composite Measures of Quality of Health Care: Evidence Mapping of Methodology and Reporting. *PLoS ONE* **2022**, *17*, e0268320. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Barclay, M.; Dixon-Woods, M.; Lyratzopoulos, G. The Problem with Composite Indicators. *BMJ Qual. Saf.* **2019**, *28*, 338–344. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Ayadi, H.; Benaissa, M.; Hamani, N.; Kermad, L. Assessing the Sustainability of Transport Systems through Indexes: A State-of-the-Art Review. *Sustainability* **2024**, *16*, 1455. [\[CrossRef\]](#)
39. Zhiyenbayev, M.; Kunicina, N.; Mansurova, M.; Patlins, A.; Caiko, J.; Beliaev, V.; Grants, R.; Bisenieks, M.; Shyntore, G. Development of Aggregated Sustainability Indicators for Quality of Life Evaluations in Urban Areas of the Republic of Kazakhstan. *Sustainability* **2024**, *16*, 9483. [\[CrossRef\]](#)
40. Hwang, C.L.; Yoon, K. *Multiple Attribute Decision Making: Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 1981.
41. Yannis, G.; Kopsacheili, A.; Dragomanovits, A.; Petraki, V. State-of-the-Art Review on Multi-Criteria Decision-Making in the Transport Sector. *J. Traffic Transp. Eng.* **2020**, *7*, 413–431. [\[CrossRef\]](#)
42. Paul, A.; Shukla, N.; Paul, S.K.; Trianni, A. Sustainable Supply Chain Management and Multi-Criteria Decision-Making Methods: A Systematic Review. *Sustainability* **2021**, *13*, 7104. [\[CrossRef\]](#)
43. Broniewicz, E.; Ogrodnik, K. Application Potential of MCDM/MCDA Methods in Transport—Literature Review and Case Study. *Sustainability* **2025**, *17*, 7671. [\[CrossRef\]](#)
44. Veličković, M.; Stojanović, Đ.; Ilin, V.; Mirčetić, D. A Combined Multi-Criteria Decision-Making and Social Cost-Benefit Analysis Approach for Evaluating Sustainable City Logistics Initiatives. *Sustainability* **2025**, *17*, 884. [\[CrossRef\]](#)
45. Stević, Ž.; Pamučar, D.; Puška, A.; Chatterjee, P. Sustainable Supplier Selection in Healthcare Industries Using a New MCDM Method: Measurement of Alternatives and Ranking according to Compromise Solution (MARCOS). *Comput. Ind. Eng.* **2020**, *140*, 106231. [\[CrossRef\]](#)
46. Bienstock, C.C.; Mentzer, J.T.; Bird, M.M. Measuring Physical Distribution Service Quality. *J. Acad. Mark. Sci.* **1997**, *25*, 31–44. [\[CrossRef\]](#)
47. Čižiūnienė, K.; Šiugždinytė, A.; Matijošius, J. Analysis of Parcel Delivery Issues at 'State Parcel Company': An Examination of Customer Complaints and Interrelationships. *Logistics* **2025**, *9*, 16. [\[CrossRef\]](#)

48. Stanković, M.; Stević, Ž.; Das, D.K.; Subotić, M.; Pamučar, D. A New Fuzzy MARCOS Method for Road Traffic Risk Analysis. *Mathematics* **2020**, *8*, 457. [CrossRef]
49. Stević, Ž.; Brković, N. A Novel Integrated FUCOM–MARCOS Model for Evaluation of Human Resources in a Transport Company. *Logistics* **2020**, *4*, 4. [CrossRef]
50. Ulutaş, A.; Karabasevic, D.; Popovic, G.; Stanujkic, D.; Nguyen, P.T.; Karaköy, Ç. Development of a Novel Integrated CCSD–ITARA–MARCOS Decision-Making Approach for Stackers Selection in a Logistics System. *Mathematics* **2020**, *8*, 1672. [CrossRef]
51. Roy, B. The Outranking Approach and the Foundations of ELECTRE Methods. *Theory Decis.* **1991**, *31*, 49–73. [CrossRef]
52. Brans, J.P.; Vincke, P. Note—A Preference Ranking Organisation Method (The PROMETHEE Method for Multiple Criteria Decision-Making). *Manag. Sci.* **1985**, *31*, 647–656. [CrossRef]
53. Opricovic, S.; Tzeng, G.-H. Compromise Solution by MCDM Methods: A Comparative Analysis of VIKOR and TOPSIS. *Eur. J. Oper. Res.* **2004**, *156*, 445–455. [CrossRef]
54. Universal Postal Union (UPU). *State of the Postal Sector 2025: Postal Power Reimagined—Scale, Service and Survival*; UPU Publishing: Bern, Switzerland, 2025. Available online: https://www.upu.int/UPU/media/upu/publications/01-StateofthePostalSector2025_EN.pdf (accessed on 10 June 2026).
55. Johnston, R. Linking Complaint Management to Profit. *Int. J. Serv. Ind. Manag.* **2001**, *12*, 60–69. [CrossRef]
56. RATEL. *Overview of the Market of Telecommunications and Postal Services in the Republic of Serbia in 2024*; RATEL Publishing: Belgrade, Serbia, 2025. Available online: <https://www.ratel.rs/storage/upload/2026/01/RATEL-SRB-Pregled-trzista-2024-IZVESTAJ.pdf> (accessed on 23 June 2026).
57. Homburg, C.; Fürst, A. How Organizational Complaint Handling Drives Customer Loyalty: An Analysis of the Mechanistic and the Organic Approach. *J. Mark.* **2005**, *69*, 95–114. [CrossRef]
58. Michel, S.; Bowen, D.; Johnston, R. Why Service Recovery Fails: Tensions Among Customer, Employee, and Process Perspectives. *J. Serv. Manag.* **2009**, *20*, 253–273. [CrossRef]
59. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656. [CrossRef]
60. Wang, T.C.; Lee, H.D. Developing a Fuzzy TOPSIS Approach Based on Subjective Weights and Objective Weights. *Expert Syst. Appl.* **2009**, *36*, 8980–8985. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.