mathematics

MDPI

Article

Benchmarking Analysis of Railway Infrastructure Managers: A Hybrid Principal Component Analysis (PCA), Grey Best–Worst Method (G-BWM), and Assurance Region Data Envelopment Analysis (AR-DEA) Model

Snežana Tadić [1], Aida Kalem [2], Mladen Krstić [1,3,*], Nermin Čabrić [2], Adisa Medić [2] and Miloš Veljović [1]

[1] Logistics Department, Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia; s.tadic@sf.bg.ac.rs (S.T.); m.veljovic@sf.bg.ac.rs (M.V.)
[2] Faculty of Traffic and Communications, University of Sarajevo, Zmaja od Bosne 8, 71000 Sarajevo, Bosnia and Herzegovina; aida.kalem@fsk.unsa.ba (A.K.); nermin.cabric@fsk.unsa.ba (N.Č.); adisa.medic@fsk.unsa.ba (A.M.)
[3] Department of Economic Sciences, University of Salento, Via Monteroni Snc, 73100 Lecce, Italy
* Correspondence: m.krstic@sf.bg.ac.rs or mladen.krstic@unisalento.it

Abstract: Benchmarking railway infrastructure managers (RIMs) has become a crucial tool in the context of European transport market liberalization, facilitating efficiency improvements and strategic decision-making. RIMs face challenges in increasing capacity, optimizing operations, and ensuring competitive, safe, and economically sustainable services. To address these challenges, this study proposes a hybrid benchmarking model that integrates Principal Component Analysis (PCA) to identify key performance indicators (KPIs) and reduce data dimensionality, the Grey Best–Worst Method (G-BWM) to determine KPI weight coefficients based on expert evaluations, and Assurance Region Data Envelopment Analysis (AR-DEA) to assess the relative efficiency of RIMs while incorporating real-world constraints. The research findings confirm that RIM8 is the most efficient unit, driven by high electrification levels, strong accident prevention measures, and optimal use of infrastructure. In contrast, RIM2 and RIM4 record the lowest efficiency scores, primarily due to poor safety performance, high infrastructure-related delays, and suboptimal resource utilization. By introducing weight constraints through AR-DEA, the model ensures that efficiency assessments reflect actual operational conditions, rather than relying on unrestricted weight allocations. The main contribution of this study lies in developing a systematic and objective framework for evaluating RIM efficiency, ensuring consistency and reliability in performance measurement. The practical implications extend to policy development and operational decision-making, providing insights for infrastructure managers, regulatory bodies, and policymakers to optimize resource allocation, enhance infrastructure resilience, and improve railway sector sustainability. The results highlight key efficiency factors and offer guidance for targeted improvements, reinforcing benchmarking as a valuable tool for long-term railway infrastructure management and investment planning. By offering a quantitatively grounded efficiency assessment, this model contributes to the competitiveness and sustainability of railway networks across Europe.

Keywords: benchmarking; railway infrastructure managers; KPIs; performance; PCA; G-BWM; AR-DEA

MSC: 03B52; 90B50

check for updates

Academic Editor: Snezhana Gocheva-Ilieva

Received: 14 February 2025
Revised: 27 February 2025
Accepted: 28 February 2025
Published: 1 March 2025

Citation: Tadić, S.; Kalem, A.; Krstić, M.; Čabrić, N.; Medić, A.; Veljović, M. Benchmarking Analysis of Railway Infrastructure Managers: A Hybrid Principal Component Analysis (PCA), Grey Best–Worst Method (G-BWM), and Assurance Region Data Envelopment Analysis (AR-DEA) Model. *Mathematics* **2025**, *13*, 830. https://doi.org/10.3390/math13050830

Copyright: © 2025 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) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Benchmarking, although defined in various ways, is widely accepted as a structured process aimed at improving standards through the adoption of best practices. Its primary purpose is the implementation of proven solutions in dynamic environments, serving as a tool for comparison and evaluation of all aspects of operations, products, or services against leading industry examples [1]. Traditionally, the focus of benchmarking has been on performance analysis in the later stages of the process, where retrospective evaluation of results supports planned organizational changes. However, to fully realize the potential of this method, its application needs to be expanded to anticipate unexpected changes and integrate it into the early stages of the process [2]. In addition to enabling organizations to identify key processes that require improvement, benchmarking provides an opportunity to adopt applicable solutions from industry best practices [3,4]. Modern approaches increasingly emphasize proactive and preventive indicators that extend beyond financial metrics. These indicators enable the prediction of future outcomes and active influence over them, ensuring business stability, more efficient utilization of knowledge resources, and the strengthening of an organization's competitive advantage [5,6]. Benchmarking has become an integral part of various sectors, including aviation, road, maritime, and railway transportation (performance analysis of airports and terminals [7], efficiency assessment of container terminals and ports [8], road safety evaluation [9], and assessment of service quality, safety, and punctuality of railway operators [10]). Also, it is widely applied in industries such as manufacturing, healthcare services, finance, education, and the public sector [11]. Specifically, in the research of railway infrastructure managers (RIMs), which are the subject of this study, the benchmarking framework has been used for measuring railways' safety on a national [12] and cross-continental level [13], environmental and economic impacts on high-speed rails [14], the success of railway construction megaprojects [15], railway power management systems [16], etc.

European railway infrastructure is facing increasing operational demands, with rail networks expected to accommodate a 25% growth in freight transport and a 40% rise in passenger traffic by 2050 [17]. However, inefficiencies in infrastructure management remain a major concern, with studies indicating that outdated systems and operational bottlenecks can reduce network capacity by up to 30% [18]. This underscores the critical need for robust benchmarking tools to enhance performance and sustainability in railway infrastructure management. Effective railway infrastructure management is crucial for enhancing transportation systems, facilitating sustainable economic development, improving connectivity for people and goods, and reducing environmental impacts. As an environmentally friendly mode of transport, rail transport plays a significant role in achieving sustainable development goals. However, the complexity of railway infrastructure management, coupled with challenges such as outdated infrastructure, necessitates advanced tools for performance analysis and improvement. Benchmarking has come to the fore as an important performance assessment, best practice identification, and strategic decision-making tool to enhance efficiency and competitiveness in the sector. Benchmarking is a valuable tool for railway infrastructure managers to compare their performance against reference values, identify and address areas for improvement, and adopt best practices. This process is led by the Platform of Rail Infrastructure Managers in Europe (PRIME), which has created a catalogue of standardized key performance indicators (KPIs). The indicators encompass almost all the performance dimensions of the project, such as its contextual framework, safety, environment, performance, service delivery, finance, and growth. The complexity of railway systems, the numerous factors that influence them, and heterogeneity of data are the major challenges to traditional methods of performance analysis. With the use of standard KPIs and advanced analytical methods, benchmarking

is not only a comparison tool for performance, but also a decision-making tool for strategic infrastructure management and resource optimization.

This study is based on the application of KPIs as input data for the development of a hybrid benchmarking model for assessing the performance of railway infrastructure managers. The proposed model integrates three methods: Principal Component Analysis (PCA) for data dimensionality reduction, the Grey Best–Worst Method (G-BWM) for determining the weight coefficients of KPIs, and Assurance Region Data Envelopment Analysis (AR-DEA) for evaluating relative efficiency. This approach enables a precise and comprehensive performance assessment, considering the complexity and diversity of the railway infrastructure sector. PCA allows for dimensionality reduction, identifying key KPIs that best explain variability within the dataset. The G-BWM provides a reliable and flexible approach for determining weight coefficients in conditions of uncertainty and limited information. By leveraging expert assessments, the G-BWM identifies the most important (best) and least important (worst) KPIs and, through pairwise comparisons, calculates optimal weight coefficients. This ensures that weight coefficients accurately reflect the true significance of each KPI, enabling a fair performance evaluation. This process is particularly important for sectors with complex and heterogeneous operational environments, such as the railway industry. Traditional DEA analysis often has limitations, due to the possibility that certain inputs or outputs receive zero weight, which compromises result accuracy and disregards critical factors. The AR-DEA (Assurance Region DEA) method overcomes this issue by introducing weight constraints ("assurance regions") into the optimization process. These assurance regions define boundaries for input and output weight coefficients, ensuring that all key factors are included in the analysis. The AR-DEA method utilizes the weight coefficients derived through the G-BWM as initial values for setting assurance region limits. This establishes a link between subjective expert assessments and objective efficiency analysis, balancing model flexibility and accuracy. This integration enables the assessment of the relative efficiency of decision-making units (DMUs) while maintaining consistency in weight allocation. The application of the proposed model aims to enhance the benchmarking process by precisely identifying efficient and inefficient railway infrastructure managers. The model allows for a detailed comparison of different managers' performances, highlighting best practices that can serve as guidelines for improving less efficient units. Furthermore, the model provides an in-depth understanding of key factors influencing performance, ensuring that decision-makers receive accurate and reliable information for strategic planning, resource optimization, and evidence-based decision-making in railway infrastructure management. This study makes a significant contribution to the development of methodologies for monitoring and improving railway infrastructure performance, establishing a foundation for future research and innovation in resource management optimization and railway system sustainability.

This study proposes a hybrid benchmarking model that integrates PCA, the G-BWM, and AR-DEA to address key efficiency challenges in railway infrastructure management. By offering a quantitatively grounded and decision-maker-informed efficiency assessment framework, the model provides actionable insights for optimizing infrastructure operations, guiding investment strategies, and ensuring long-term sustainability in the sector. Accordingly, the rest of the paper is organized as follows: Section 2 provides a review of the relevant literature, focusing on benchmarking in the railway sector, efficiency concepts, and the application of advanced methodologies, including PCA, the G-BWM, and AR-DEA. Section 3 presents a detailed description of the hybrid methodology proposed in this study, outlining the specific steps for integrating PCA, the G-BWM, and AR-DEA. Section 4 applies the methodology to a selected sample of nine RIMs, demonstrating its practical implementation and results. Section 5, the discussion section, analyzes the obtained results,

exploring factors influencing efficiency, the advantages and limitations of the proposed methodology, and implications for both theory and practice. Finally, Section 6 concludes the study, summarizing the key findings and providing recommendations for future research.

2. Literature Review and Key Performance Indicators

This section presents a review of the relevant literature on benchmarking in the railway sector, with a focus on efficiency and the application of advanced methodological approaches. Special attention will be given to the use of methods such as Principal Component Analysis (PCA), the Grey Best–Worst Method (G-BWM), and Assurance Region Data Envelopment Analysis (AR-DEA), which are employed for performance analysis and improvement in railway transport. This review aims to provide a more profound understanding of existing theoretical and practical approaches that are relevant to this research.

2.1. Methodological Approaches in Benchmarking

For the development of this paper, the Systematic Literature Review (SLR) method was applied to evaluate research questions related to the application of hybrid models for benchmarking analysis of railway infrastructure managers. The systematic literature review was conducted through five phases. In the first phase, selection of scientific papers published in the last five years was performed to ensure the relevance and contemporaneity of the analysis. The selection of scientific papers published in the last five years ensured that the review would capture the most current and relevant research, aligning the study with contemporary advancements in hybrid models for benchmarking railway infrastructure managers. The second phase involved defining inclusion criteria, where key search terms were focused on rail transport, infrastructure managers, performance, efficiency, and benchmarking. Defining inclusion criteria with specific search terms helps to focus the review on key areas, such as rail transport, infrastructure managers, and performance benchmarking, ensuring that only relevant and targeted literature is included in the analysis. The third phase involved structuring the literature according to the applied methodologies (PCA, the BWM, DEA, the G-BWM, and AR-DEA). Structuring the literature by applied methodologies allowed for a systematic comparison of different approaches and their application to benchmarking analysis, facilitating the identification of trends and gaps in the current body of knowledge. This was followed by the fourth phase, in which the areas of application were analyzed to determine whether the combination of PCA, the G-BWM, and AR-DEA had been previously applied in railway transport or other sectors. The literature search was conducted using the following meta-query: (TITLE-ABS-KEY (“benchmarking” OR “performance” OR “efficiency”)) AND (TITLE-ABS-KEY (“railway infrastructure” OR “railway infrastructure managers” OR “rail transport” OR “traffic” OR “transport”)) AND (TITLE-ABS-KEY (“PCA” OR “principal component analysis” OR “G-BWM” OR “grey best-worst method” OR “BWM” OR “best-worst method” OR “AR DEA” OR “assurance region DEA” OR “DEA” OR “data envelopment analysis”)) AND ((LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “re”) OR LIMIT-TO (DOCTYPE, “cp”) OR LIMIT-TO (DOCTYPE, “cr”) OR LIMIT-TO (DOCTYPE, “bo”))) AND (LIMIT-TO (LANGUAGE, “English”)). Finally, the fifth phase of the research focused on identifying the objectives and purposes of the proposed models in the reviewed papers. Identifying the objectives and purposes of the proposed models in the reviewed papers contributes to a deeper understanding of the practical outcomes and goals of applying hybrid models in benchmarking, guiding future research and practical implementation strategies. The results of the systematic literature review include an analysis of 30 scientific papers. The review enabled the identification of key methodologies and approaches, highlighting PCA, the G-BWM, and AR-DEA as fundamental components for the hybrid model, integrating

dimensionality reduction, weight coefficient determination, and relative efficiency evaluation. Table 1 presents an overview of the analyzed studies, structured according to the defined inclusion criteria, including year of publication, key terms, applied methods, application domains, and objectives/purposes of the proposed models. Additionally, key challenges and emerging trends were identified, which will be further analyzed in the following sections of the study.

Table 1. Literature review.

Ref.	Crit. I—Year of Publication	Crit. II—Keywords	Criterion III—Applied Methods						Scope of Model		Purpose/Purpose of Model
			PCA	BWM	DEA	GBWM	AR DEA	Other Methods	Railway Traffic	Other Areas	
[19]	2024	Railway Carrying Capacity, Coordination Optimization	x					EFA, Regression, Scenario Analysis	x		Identification of key energy efficiency factors and optimization of railway systems.
[20]	2019	Sustainable transport, rail freight market, PCA	x						x		Analysis of impact of market liberalization on competitiveness of rail freight transport.
[21]	2021	PCA, railway maintenance	x					ANOVA, Kruskal–Wallis	x		Improvement of monitoring and maintenance of railway wheels.
[22]	2022	Life management, maintenance planning	x					Decision Trees, Bayesian Networks	x		Optimization of railway infrastructure maintenance.
[23]	2019	Maintenance; PCA; factor analysis	x					FA	x		Identification of key human factors that affect safety and efficiency of railway system maintenance.
[24]	2024	Rail network, BWM, decision-making		x				Lingo 17 softver	x		Identification and ranking of risk factors affecting rail transport safety.
[25]	2021	Urban rail, performance evaluation, BWM		x				Large-Scale Group Evaluation Technology, CS	x		Performance evaluation of PPP projects in urban rail transport.
[26]	2023	Rail transportation projects, Fuzzy MCDM		x				Stochastic VIKOR	x		Assessment and selection of railway development projects.
[27]	2022	Cross-regional railways, MCA, BWM		x				FAISM, DFS	x		Analysis of impact of changes in demand and supply on inter-regional rail connectivity in post-COVID-19 period.
[28]	2024	Performance evaluation, MCDM, G-BWM				x		RATMI, MCDM		x	Performance evaluation of retail warehouses.

Table 1. *Cont.*

Ref.	Crit. I—Year of Publication	Crit. II—Keywords	Criterion III—Applied Methods						Scope of Model		Purpose/Purpose of Model
			PCA	BWM	DEA	GBWM	AR DEA	Other Methods	Railway Traffic	Other Areas	
[29]	2023	intermodal terminal; Governance model, GBWM, GCOBRA				x		GCOBRA		x	Evaluation of governance models in intermodal terminal operations.
[30]	2024	Supply chain, circular economy, GBWM, GCOBRA				x		GCOBRA		x	Evaluation of agri-food supply chain risks.
[31]	2022	City logistics, Industry 4.0, GBWM, GCODAS				x		GCODAS		x	Evaluation of smart city logistics solutions.
[32]	2023	Railway transport, three-stage DEA,			x				x		Evaluation of efficiency of rail transport.
[33]	2019	railways, efficiency, benchmarking			x				x		Assessment of efficiency of railway operations.
[34]	2019	DEA, performance measurement			x			Two-stage clustering approach	x		Assessment of socio-technical performance of TCC staff.
[35]	2020	Railway undertaking, efficiency, DEA			x			FAHP	x		Assessment of efficiency of freight railway companies.
[36]	2019	Efficiency, Fuzzy EDAS, AR DEA					x	Fuzzy EDAS		x	Evaluation of efficiency of small inland intermodal terminals.
[37]	2023	Railway, efficiency measurement, NDEA			x			SBM, Network SBM	x		Analysis of efficiency of railway stations.
[38]	2023	DEA			x			SBM		x	Assessment of regional differences in vehicle environmental efficiency.
[39]	2019	Efficiency, rail systems, railways						Meta-analysis, Permutation Test	x		Analysis of factors affecting efficiency of railway systems.
[40]	2020	Railway indicators, HSR performance						SLR, expert consultation	x		Analysis of sustainability reporting practices of HSR operators.
[41]	2022	PCA–DEA, efficiency, transport	x		x			CRITIC, Entropy, MARCOS		x	Evaluation of efficiency of transport companies.
[42]	2020	Eco-efficiency, transport sector	x		x			MLR		x	Assessing eco-efficiency of transport in Nairobi and identifying factors for improving sustainability.

Table 1. *Cont.*

Ref.	Crit. I—Year of Publication	Crit. II—Keywords	Criterion III—Applied Methods						Scope of Model		Purpose/Purpose of Model
			PCA	BWM	DEA	GBWM	AR DEA	Other Methods	Railway Traffic	Other Areas	
[43]	2020	DEA, BWM, road safety		x	x					x	Evaluation of efficiency of road traffic safety.
[44]	2020	DEA			x			AHP, VIKOR, MPI		x	Performance assessment of public road transport in India.
[45]	2023	Hospital waste management, DEA, PCA	x		x			ANP, CRS, Shannon's Entropy, CES		x	Assessment of efficiency of biomedical waste management in hospitals.
[46]	2023	BWM, DEA, PCA	x	x	x			SWOT		x	Assessing efficiency of private banks using CAMELS indicators.
[47]	2021	DEA			x			MLR, Graph-Theoretic Approach		x	Assessment of changes in efficiency and financial performance of companies in automotive industry.
[48]	2024	Hierarchical DEA			x					x	Assessment of logistics performance of countries.
[49]	2020	Logistics industry performance, SBM-DEA, PCA	x		x			Tobit Regression		x	Assessing efficiency of China's logistics industry.
[50]	2021	PCA; relative efficiency	x		x			VIKOR		x	Systematic assessment and ranking of suppliers in textile industry.
[51]	2022	DEA	x		x					x	Evaluation of supplier efficiency in food industry.
[52]	2020	PCA, DEA	x				x	Cluster Analysis, Tobit Regression		x	Energy system performance assessment of 125 countries.

Through the literature review, various applications of advanced analytical methods, including PCA, DEA, and the BWM, have been identified. PCA has been used for dimensionality reduction and the identification of key variables in studies on energy efficiency in railway systems [19], the market competitiveness of railway freight transport [20], optimization of railway wheel maintenance [21], key human factors [23], and railway infrastructure lifecycle management [22]. The BWM has been applied for determining weight coefficients of risk factors in railway safety analyses [24], evaluating public–private partnerships [25], analyzing the sustainability of railway projects [26], and assessing the impact of changes in demand and supply on inter-regional railway connectivity [27]. In addition to these methods, researchers have employed other techniques to enhance analytical frameworks. For example, meta-analysis has been used to aggregate results from multiple studies and identify key efficiency variables in the railway sector [39].

The DEA method, known for its ability to assess the relative efficiency of decision-making units, has found extensive application in the railway sector. It has been used

for evaluating the efficiency of railway transport systems in international and regional contexts [32,33]. Furthermore, DEA has been applied in socio-technical performance analysis [34] and in assessing the efficiency of freight railway companies [35], identifying key criteria for improving competitiveness. Additionally, network DEA models have enabled a deeper analysis of railway station efficiency [37], while specific DEA variants, such as the SBM and AR-DEA, have been employed to evaluate the performance of small intermodal terminals [36]. Although these methods have significantly contributed to understanding various aspects of the railway sector, their integration into a hybrid benchmarking model for railway infrastructure managers remains unexplored. The motivation for developing such a model stems from the application of these methods in other sectors. While the combination of PCA, the BWM, and DEA has been applied in the banking sector [46], it refers to their basic forms. An enhanced version that incorporates the G-BWM for improved flexibility and precision in weight determination under uncertainty, and AR-DEA, which introduces constraints on weight coefficients for a more realistic efficiency evaluation, represents a significant methodological advancement. This improved combination of methods has not been previously applied in the railway sector or other industries, underscoring the novelty and potential of the proposed hybrid model for enhancing performance evaluation and strategic decision-making.

The development of a hybrid model that combines PCA, the G-BWM, and AR-DEA represents a logical advancement in methodological frameworks, particularly within the railway sector. This model enables a more systematic analysis of complex interactions between key performance indicators (KPIs), where each method contributes to specific aspects of evaluation—from identifying key variables and ensuring reliable weighting of criteria, to accurately measuring relative efficiency. The model enhances performance evaluation by addressing DEA's limitations, incorporating decision-maker preferences through weight bounds, and using the BWM to capture those preferences in a grey environment, which allows for handling imprecise information. Additionally, PCA reduces dimensionality and improves discrimination power, ensuring more accurate, efficient, and realistic identification of efficient DMUs, making the model valuable for performance assessment across various complex, multi-criteria environments. The proposed approach offers a unique opportunity for developing dynamic and adaptable infrastructure management strategies, thereby contributing to the long-term sustainability and competitiveness of the railway sector. Its application opens new perspectives for enhancing analytical practices in transportation and beyond, providing a methodological framework that can be adapted to the specific needs and challenges of other industries.

2.2. Role of KPIs in Railway Benchmarking

Benchmarking in the railway sector plays a key role in performance assessment by identifying best practices and comparing them with reference values. According to research, benchmarking facilitates the understanding of the relative efficiency of different railway systems, and supports strategic decision-making aimed at improving the operational and financial aspects of infrastructure management [33]. Various studies have utilized benchmarking analyses to evaluate the efficiency of railway companies and identify best practices for enhancing operational performance. KPI-based benchmarking analysis is widely applied in the railway sector to optimize different operational aspects. For instance, benchmarking analysis is used for improving the energy efficiency of urban railway systems [53], assessing service quality in railway transport [54], enhancing national railway system safety [12], comparing performance measurement standards among different railway operators [55], evaluating railway system performance across Europe [56], assessing

high-speed rail network operational readiness [57], and evaluating the performance of infrastructure managers in the railway sector [58].

Benchmarking utilizing KPIs serves as a critical tool for assessing and improving railway infrastructure management, enabling systematic performance comparisons across different networks and identifying best practices for efficient management. The European Commission, through the PRIME (Platform for European Railway Infrastructure Managers) initiative, has established a catalogue of KPIs to standardize performance monitoring and enhance railway infrastructure management. This system covers a broad range of aspects, including contextual framework, safety, environment, performance, service delivery, finance, and growth, allowing for a structured approach to analyzing and comparing infrastructure managers' performance across Europe [59]. By utilizing KPIs, railway infrastructure managers can systematically monitor and analyze key parameters that impact the efficiency and safety of the railway system. The benchmarking process facilitates performance comparisons among different railway infrastructure managers, thereby identifying best practices. Using KPIs as a benchmarking foundation ensures precise evaluation and transparency, as clear metrics can be applied to assess success across various domains. The PRIME initiative not only provides a framework for internal evaluation, but also enables comparisons between different national and regional networks, fostering best practices in the railway transport sector. Through this methodology, performance comparisons can drive innovation and alignment with European standards, ultimately contributing to better conditions for passengers and more efficient railway system operations across Europe. Key challenges in benchmarking analysis within the railway sector include limited data availability, inconsistencies between different databases, and variations in regulatory frameworks across countries [60]. Consequently, methodologically robust approaches must be applied to ensure reliable and accurate performance comparisons. To ensure precise and reliable benchmarking, this paper employs KPIs of railway infrastructure managers. The benchmarking analysis is based on the application of PCA, G-BWM, and AR-DEA methods, which enable a comprehensive evaluation of railway infrastructure managers' performance.

3. Methodology—Hybrid PCA–G-BWM–AR-DEA Model

Benchmarking the performance of railway infrastructure managers is a key tool for identifying best practices and improving efficiency. This analysis requires a systematic approach that incorporates KPIs to ensure an objective and reliable evaluation. Given the complexity of infrastructure management, these indicators encompass various efficiency aspects, including operational, financial, technical, and environmental factors. However, evaluation challenges often arise, due to variable interdependencies, data multidimensionality, and subjectivity in determining the importance of individual criteria. To address these challenges, this paper proposes a hybrid model that integrates PCA, the G-BWM, and AR-DEA, leveraging their combined advantages to establish a comprehensive and robust framework for benchmarking railway infrastructure managers' performance based on KPIs. The proposed hybrid model offers a significant advancement in performance evaluation by addressing key limitations of the traditional DEA method. DEA's automatic weight assignment often leads to unrealistic evaluations, and the AR DEA model overcomes this by incorporating decision-maker preferences through weight bounds. However, defining these bounds requires a reliable method for capturing preferences, which is why the BWM is employed. The grey environment further enhances the model's robustness by enabling it to handle imprecise or incomplete information. Additionally, PCA is integrated to reduce dimensionality and improve discrimination power by focusing on the most critical KPIs, which enhances both computational efficiency and the accuracy of the evaluation. This

combination of methods ensures more flexible, accurate, and realistic identification of efficient DMUs, contributing to a more effective performance evaluation framework. This model's contributions to the field include improved discrimination power, computational efficiency, and applicability to complex, multi-criteria environments, making it a valuable tool for performance assessment across various disciplines.

PCA is employed to identify and extract the most relevant variables within the model. This method transforms the original dataset into a new set of variables—principal components—that maximize variance and reduce redundancy among indicators [61,62]. The first step in PCA involves data standardization, followed by the calculation of the covariance matrix, the determination of eigenvalues and eigenvectors, and the formation of feature vectors that define the principal components. The objective of this process is to eliminate collinearity among variables and enhance model stability. After dimensionality reduction, the G-BWM is applied to determine the weight coefficients of KPIs. The G-BWM is a Multi-Criteria Decision-Making (MCDM) method based on the grey system theory, allowing experts to quantify subjective preferences by identifying the best and worst criteria. The key steps in this method include defining the evaluation scale, selecting reference criteria, calculating grey vectors that describe the relative relationships among criteria, and estimating the initial and final weight coefficients. This approach enhances evaluation consistency under uncertainty and reduces subjective biases in weight assignment.

In the final stage, the AR-DEA method is used to assess the relative efficiency of railway infrastructure managers. DEA is a non-parametric method for measuring the efficiency of DMUs, employing linear programming to distinguish between efficient and inefficient units. In this study, AR-DEA extends the standard DEA model by introducing additional constraints that enhance the discriminatory power of the analysis, allowing for more precise ranking of units. The process involves data collection and normalization, the formation of input and output variable matrices, and the calculation of the relative efficiency of each DMU. The combination of these methods provides a comprehensive performance evaluation of railway infrastructure managers, integrating KPIs and expert assessments into the benchmarking process. The steps of the proposed model are described below and illustrated in Figure 1.

Step 1: Define the problem structure, i.e., the set of DMUs and the inputs and outputs for the efficiency evaluation.

Step 2: Apply PCA to analyze the relationships among input and output variables, reduce multicollinearity, and extract the most relevant components for further analysis [63].

Step 2.1: For the initial data $X_t^{(i)}$, where $i = 1, 2, \ldots, N$, a $t = 1, 2, \ldots, T$, standardization is performed using the following formula:

$$\tilde{X}_t^{(i)} = \frac{\left(X_t^{(i)} - \tilde{X}^{(i)} \right)}{\sigma \left(X^{(i)} \right)} \tag{1}$$

where $\tilde{X}^{(i)}$ is the mean value, and $\sigma \left(X^{(i)} \right)$ is the standard deviation of the variable $X^{(i)}$.

Step 2.2: Let $\sum$ denote the covariance matrix of the standardized data $\tilde{X}$; then, it can be decomposed as follows:

$$\sum = \Gamma \Lambda \Gamma' \tag{2}$$

where Λ is a positive diagonal matrix containing the eigenvalues $\lambda_1, \lambda_1, \ldots, \lambda_N$, and the columns of the matrix Γ are the corresponding eigenvectors. It is assumed that the eigenvectors (principal components) are ordered in descending order of their eigenvalues,

where the first K eigenvectors represent the factor loadings $\left[q^{(1)}, q^{(2)}, \ldots, q^{(K)}\right]$. Then, K features $\left[F^{(1)}, F^{(2)}, \ldots \text{and } F^{(K)}\right]$ are obtained, which are defined as follows:

$$F_t^{(k)} = q^{(k)'} \widetilde{X}_t \tag{3}$$

where je $\widetilde{X}_t$ represents the N-dimensional vector of the standardized term structure at time t.

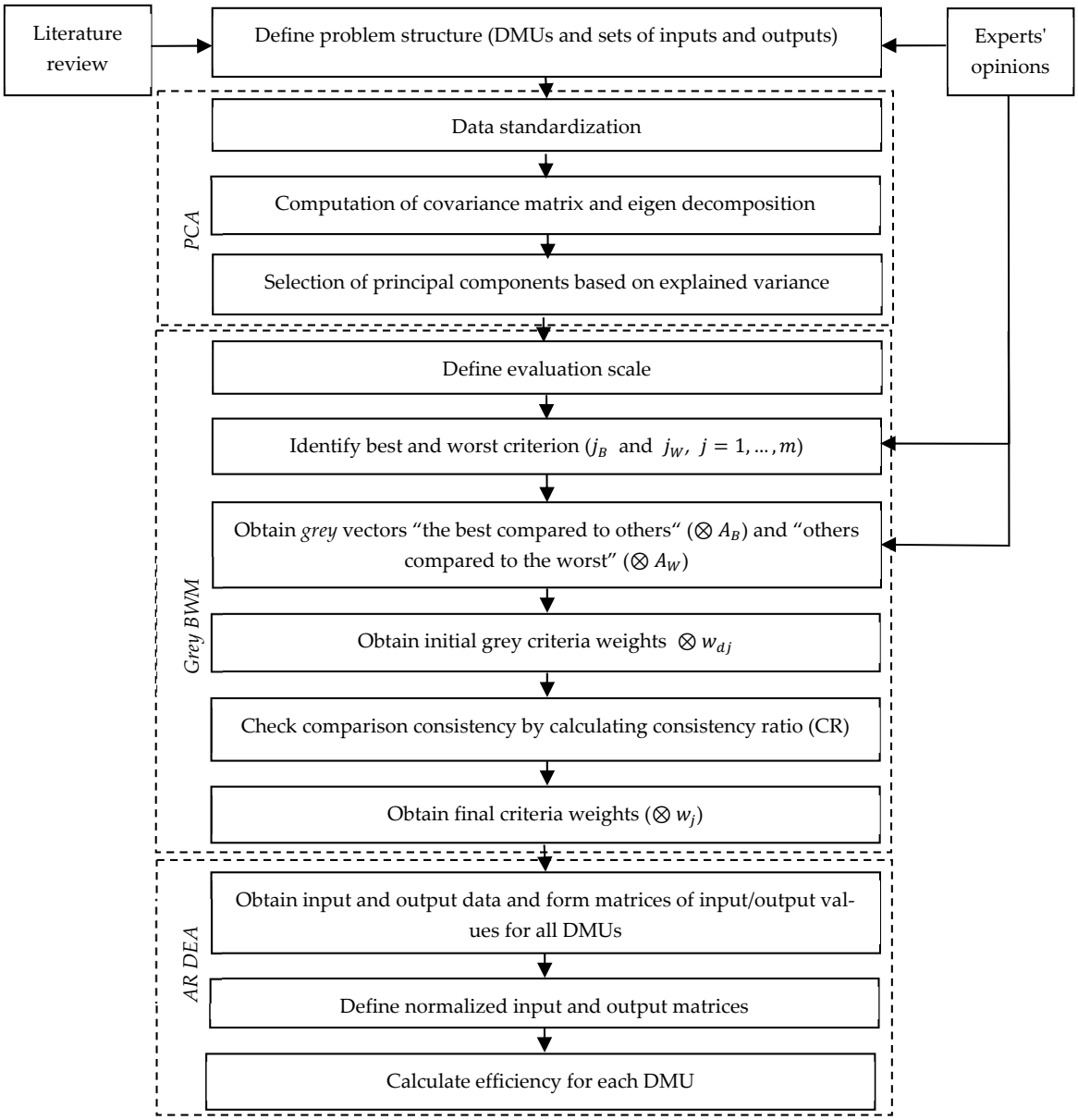


Figure 1. Structure of proposed model.

The covariance matrix captures the relationships between variables and helps to identify the directions of maximum variance, while the eigenvectors define the principal components along those directions, enabling dimensionality reduction and the transformation of data into a new uncorrelated feature space.

Step 2.3: To determine how well a principal component explains the data, the total variance and the proportion of variance of the m-th principal component are used, defined as follows:

$$\frac{\lambda_m}{\sum_{j=1}^{n} \lambda_j} \tag{4}$$

This proportion represents the explanatory power of each principal component. Accordingly, the principal component with the highest explanatory power becomes the first principal component, followed by the second with the second highest explanatory power, and so on. The principal components that capture the largest portion of the total variance (typically 80–90%) are selected, ensuring that the most important aspects of the data are preserved for further analysis.

Step 3: Define the input/output weight limits by applying the G-BWM (adapted from [64].

Step 3.1: Define the grey scale for the evaluations (Table 2).

Table 2. Grey evaluation scale.

Linguistic Term	Abbreviation	Grey Scale
"None"	"N"	[0, 2]
"Very Low"	"VL"	[1, 3]
"Low"	"L"	[2, 4]
"Fairly Low"	"FL"	[3, 5]
"Moderate"	"M"	[4, 6]
"Fairly High"	"FH"	[5, 7]
"High"	"H"	[6, 8]
"Very High"	"VH"	[7, 9]
"Extremely High"	"EH"	[8, 10]

Step 3.2: Decision-makers (DMs) d $(d = 1, \ldots, f)$ select the best and the worst input/output (j_B and j_W, $j = 1, \ldots, m$), and then evaluate the other inputs/outputs according to them using the scale from Table 1, thus obtaining grey vectors „the best compared to others"—$\otimes A_B = (\otimes a_{B1}, \otimes a_{B2}, \ldots, \otimes a_{Bm})$, and "others compared to the worst"—$\otimes A_W = (\otimes a_{1W}, \otimes a_{2W}, \ldots, \otimes a_{mW})$.

Step 3.3: Obtain the optimal grey input/output weight limits $\otimes l_{d1}, \otimes l_{d2}, \ldots, \otimes l_{dm}$, $\forall d = 1, \ldots, f$ by solving the following optimization problem:

$$\text{s.t.} \begin{cases} \min \otimes \xi, \\ \left| \frac{\otimes l_B}{\otimes l_{dj}} - \otimes a_{Bj} \right| \leq \otimes \xi, \\ \left| \frac{\otimes l_{dj}}{\otimes l_W} - \otimes a_{jW} \right| \leq \otimes \xi, \\ \sum_{j=1}^{m} L\left(\otimes l_{dj} \right) = 1 \\ \underline{l}_{dj} \leq \bar{l}_{dj} \\ \underline{l}_{dj} \geq 0 \\ j = 1, \ldots, m \end{cases} \tag{5}$$

where $\otimes \xi = \left[\underline{\xi}, \bar{\xi} \right]$, $\otimes l_B = \left[\underline{l}_B, \bar{l}_B \right]$ is the grey limit of the "best" criterion, $\otimes l_w = \left[\underline{l}_W, \bar{l}_W \right]$ is the grey limit of the "worst" input/output, $\otimes l_{dj} = \left[\underline{l}_{dj}, \bar{l}_{dj} \right]$ is the grey limit of the input/output j, $j = 1, \ldots, m$, $j \neq j_B, j_W$, $\otimes a_{Bj} = \left[\underline{a}_{Bj}, \bar{a}_{Bj} \right]$ is the grey preference of the "best" input/output over the input/output j, and $\otimes a_{jW} = \left[\underline{a}_{jW}, \bar{a}_{jW} \right]$ is the preference of

the input/output j over the "worst" input/output. $L(\otimes l_j)$ is the white value of the grey number $\otimes l_j$:

$$L(\otimes l_{dj}) = (\underline{l}_{dj} + \bar{l}_{dj})/2 \tag{6}$$

Step 3.4: Check the comparison consistency by calculating the consistency ratio (CR) with the following formula:

$$CR = R(\otimes\xi)/CI \tag{7}$$

where $R(\otimes\xi)$ is the white value of the grey number $\otimes\xi$, obtained by (6), and CI is the consistency index:

$$CI^2 - (1 + 2\bar{a}_{BW})CI + \left(\bar{a}_{BW}^2 - \bar{a}_{BW}\right) = 0 \tag{8}$$

where

$$\otimes a_{BW} = \max_j \{\otimes a_{Bj}, \otimes a_{jW}\} \tag{9}$$

The comparison is considered consistent if the CR value is close to 0.

Step 3.5: Obtain the final input/output limits, as follows:

$$\otimes l_j = \left[\underline{l}_j, \bar{l}_j\right], \ \forall j = 1, \ldots, m, \tag{10}$$

$$\underline{l}_j = \left(\prod_{d=1}^{l} \underline{l}_{dj}\right)^{1/d}, \tag{11}$$

$$\bar{l}_j = \left(\prod_{d=1}^{l} \bar{l}_{dj}\right)^{1/d}, \tag{12}$$

Step 4: Select the efficient DMUs by applying the AR DEA method (adapted from [65]).

Step 4.1: Define the input value matrices for all DMUs (X_{ij}), i.e., the output value matrices for all DMUs (Y_{is}), in the following manner:

$$X_{ij} = \left[x_{ij}\right]_{o \times e}, \tag{13}$$

$$Y_{is} = \left[y_{is}\right]_{o \times t} \tag{14}$$

where x_{ij} is the value of input j for DMU i, $i = 1, \ldots, o$, $j = 1, \ldots, e$, and y_{is} is the value of output s for DMU i, $s = 1, \ldots, t$. e and t represent the total number of inputs and outputs taken into consideration, respectively.

Step 4.2: Define the normalized matrices of inputs and outputs. The normalized matrices of inputs (X'_{ij}) and outputs (Y'_{is}) are formed by the normalized values of inputs and outputs, obtained in the following manner:

$$x'_{ij} = \frac{x_{ij}}{\max\limits_i x_{ij}}, \ \forall j = 1, \ldots, e, \tag{15}$$

$$y'_{is} = \frac{y_{is}}{\max\limits_i y_{is}}, \ \forall s = 1, \ldots, t \tag{16}$$

Step 4.3: Calculate the efficiencies for each DMU. In order to calculate the efficiencies for each DMU, it is necessary to solve the following linear optimization model:

$$max \sum_{s=1}^{t} u_s y_{is}, \tag{17}$$

subject to the following:

$$\sum_{j=1}^{e} v_j x_{ij} = 1, \tag{18}$$

$$\sum_{s=1}^{t} u_s y_{is} - \sum_{j=1}^{e} v_j x_{ij} \leq 0, \quad \forall i = 1, \ ..., \ o, \tag{19}$$

$$\underline{lv}_j \leq v_j \leq p\overline{lv}_j, \quad \forall j = 1, \ ..., \ e, \tag{20}$$

$$q\underline{lu}_s \leq u_s \leq q\overline{lu}_s, \quad \forall s = 1, \ ..., t, \tag{21}$$

$$p \geq 0, \ q \geq 0, \tag{22}$$

where v_j, $j = 1, \ldots, e$, and u_s, $s = 1, \ldots, t$, are the input and output weights, respectively; $\underline{lv}_j$ and $\overline{lv}_j$ are the lower and upper weight limits for the input j; and $\underline{lu}_s$ and $\overline{lu}_s$ are the lower and upper weight limits for the output s, calculated by applying Equations (11) and (12). p and q are the scaling factors, enabling the fulfilment of conditions for introducing the input and output weight bounds.

4. Application of Proposed Model for Benchmarking Analysis of RIMs

In this section, the hybrid model for benchmarking analysis of RIMs is applied, utilizing a combination of PCA, G-BWM, and AR-DEA methods. The main objective of this approach is to conduct a detailed evaluation of the performance of nine selected RIMs, using a set of 35 KPIs. These KPIs serve as the foundation for quantifying efficiency and identifying best practices in railway infrastructure management. This multidimensional approach enables a systematic analysis and comparison of results, taking into account various business aspects, including contextual framework, safety, environment, performance, service delivery, finance, and growth. In Step 1, the DMUs in this study consist of nine RIMs from different EU countries. To ensure data privacy and confidentiality, the identities of the RIMs will remain anonymous, with each manager being assigned an alphanumeric code (e.g., RIM1, RIM2, etc.). These units were carefully selected to ensure comparability, thereby enhancing the accuracy and validity of the results. For the selection of KPIs, this study relies on the set of indicators defined in the PRIME report of the European Commission, which covers key aspects of railway infrastructure performance. The selection of KPIs was based on their high priority and relevance for benchmarking analysis. The indicators were meticulously chosen to ensure measurability and comprehensiveness, covering all critical performance dimensions of railway infrastructure managers. Additionally, the KPIs used in this study were selected based on previous research [60], where the same indicators were applied to assess the performance of railway infrastructure managers. Data collection in this phase is of critical importance, as the data must be accurate, verified, and standardized to ensure reliable subsequent analysis. Ensuring high-quality data and aligning them with the required standards guarantees that the analysis remains valid and that the results are relevant for decision-making. All data used in this phase were collected from reliable sources, including relevant reports and the PRIME platform, ensuring consistency and analytical integrity [18]. In Step 2, the PCA method was applied to reduce the dimensionality of the dataset KPIs, thereby enhancing analytical efficiency and improving data interpretation. The PCA implementation was conducted using Minitab software version 17.1, which facilitated the necessary statistical computations and visualization of results, streamlining further analysis and data interpretation. Table 3 presents the results of the principal component analysis, including the eigenvalues, proportions, and cumulative proportions of variance for each principal component (PC1 to PC8).

Table 3. Eigenanalysis of correlation matrix.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalue	10.805	8.411	5.118	3.581	3.111	1.429	1.357	1.187
Proportion	0.309	0.240	0.146	0.102	0.089	0.041	0.039	0.034
Cumulative	0.309	0.549	0.695	0.798	0.887	0.927	0.966	1.000

The first principal component (PC1) has an eigenvalue of 10.805, explaining 30.9% of the total variance in the dataset. PC2 and PC3, with eigenvalues of 8.411 and 5.118, account for an additional 24% and 14.6% of the variance, respectively. By analyzing the eigenvalues and cumulative proportion of variance, it is evident that the first five components (PC1–PC5) collectively capture 88.7% of the total variance, which serves as an optimal threshold for retaining critical information while reducing dimensionality. The remaining components (PC6–PC8) contribute significantly less to the explained variance, and may be excluded from further analysis if the goal is to simplify the model. Figure 2 presents a scree plot of the extracted components, visually illustrating the distribution of eigenvalues and enabling the identification of key components that retain the highest variance. Table 4 shows the results, including factor loadings, which indicate the strength of the relationship between individual KPIs and principal components. These insights help to identify the most significant factors for decision-making and performance optimization in railway infrastructure management.

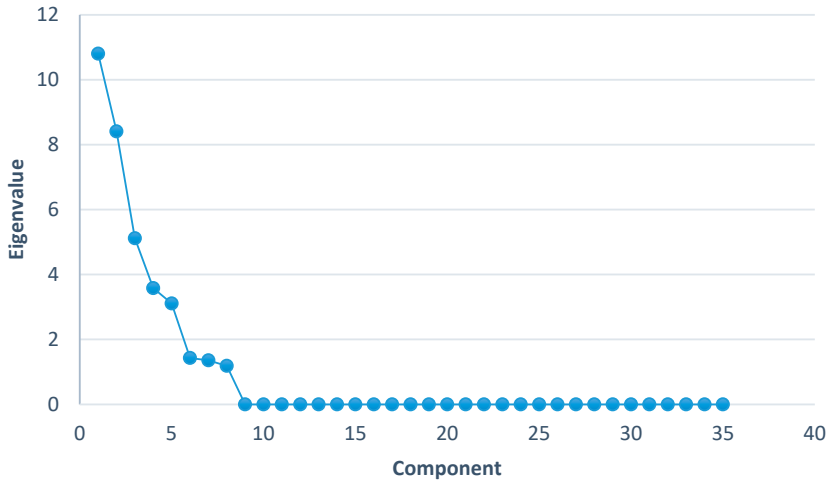


Figure 2. Scree plot of extracted components.

Following the application of PCA, selection of KPIs was performed, identifying those with the most significant impact on data variability within the first five principal components (PC1–PC5), which together accounted for 88.7% of the total variance. The selected variables were identified based on factor loadings, considering indicators with the highest absolute values, ensuring the retention of critical information necessary for further analysis. In the first principal component (PC1), the variables C1 (0.257), G5 (0.252), and D5 (−0.254) emerged as the most representative. These variables correspond to the degree of railway electrification, ATP (Automatic Train Protection) system coverage, and track failure frequency, highlighting their significant contribution to railway infrastructure management efficiency. The second principal component (PC2) was primarily influenced by the variable G6 (−0.315), which relates to the connectivity of the railway network with maritime ports. For the third component (PC3), the key indicators were C5 (−0.329), S3 (0.308), and S5 (0.306), which include the share of the national network managed by the RIM, safety indicators related to infrastructure management, and the share of electric trains in the total traffic. These indicators emphasize the relationship between safety aspects and energy efficiency within the railway system. Within the fourth component (PC4), the most significant indicator was C3 (0.404), which represents the modal share of rail transport in freight transport. This variable underscores the importance of freight transport as a critical efficiency factor in the railway network. The fifth component (PC5) identified S2 (0.359) and P4 (−0.53) as the most relevant KPIs. These variables are associated with the number of

serious accidents and the percentage of cancelled trains due to infrastructure-related issues, emphasizing the importance of safety and operational challenges in the railway sector.

Table 4. Factor loadings of KPIs.

Variables/KPIs	PC1	PC2	PC3	PC4	PC5
C1—Degree of electrification of total main track	0.257	0.012	0.143	0.121	0.196
C2—National modal share of railway in passenger transport	0.185	0.215	−0.032	0.018	0.118
C3—National modal share of railway in freight transport	0.048	0.183	0.084	0.404	0.08
C4—Length of national main lines	0.122	−0.17	−0.299	0.127	−0.049
C5—Share of national network managed % by RIM	0.07	−0.133	−0.329	0.175	0.004
C6—Length of high-speed lines	−0.201	−0.196	−0.015	−0.22	−0.133
C7—Active licences	0.074	−0.035	−0.32	0.234	−0.044
S1—Significant accidents	−0.078	−0.19	0.165	0.012	0.264
S2—Fatalities and serious injuries	−0.211	−0.005	0.04	0.068	0.359
S3—Infrastructure manager-related precursors to accidents	−0.045	−0.189	0.308	0.179	−0.046
S4—Share of diesel-powered trains	−0.207	0.011	−0.306	0.031	−0.031
S5—Share of electricity-powered trains	0.207	−0.011	0.306	−0.031	0.031
P1—Passenger train punctuality	0.188	0.097	−0.053	−0.311	−0.008
P2—Freight train punctuality	0.201	−0.067	0.206	−0.031	−0.079
P3—Delay minutes caused by infrastructure manager	−0.227	−0.014	0.021	−0.2	0.201
P4—Percentage of passenger train cancellations caused by RIM	−0.052	−0.066	0.059	−0.031	−0.53
D1—Asset failures in relation to network size	−0.221	0.2	−0.104	0.109	−0.004
D2—Signalling failures in relation to network size	−0.181	0.218	0.072	0.112	0.109
D3—Telecommunication failures in relation to network size	−0.227	0.219	−0.038	−0.027	−0.012
D4—Power supply failures in relation to network size	−0.235	0.076	0.116	0.175	0.023
D5—Track failures in relation to network size	−0.254	0.161	−0.06	−0.019	−0.028
D6—Tracks with temporary speed restrictions	0.191	−0.016	−0.134	0.267	−0.019
F1—Operational expenditures (OPEX) in relation to network size	0.065	0.299	−0.094	−0.125	−0.135
F2—Maintenance expenditures in relation to network size	−0.118	0.219	−0.125	−0.277	0.03
F3—Capital expenditures (CAPEX) in relation to network size	0.156	0.013	0.221	0.111	−0.274
F4—Proportion of Track Access Charges (TACs) in total revenue	0.134	−0.174	−0.166	−0.16	0.078
F5—Track Access Charge (TAC) revenue in relation to network size	0.177	−0.035	−0.301	−0.04	0.071
F6—Public funding for operational expenditures (OPEX)	−0.118	0.087	−0.028	0.21	−0.447
F7—Public funding for capital expenditures (CAPEX)	0.048	0.266	0.194	−0.15	−0.128
G1—Degree of network utilization—all trains	0.178	0.244	−0.05	−0.104	−0.066
G2—Degree of network utilization—passenger trains	0.171	0.226	−0.044	−0.186	−0.085
G3—Degree of network utilization—freight trains	0.134	0.216	−0.054	0.299	0.048
G4—ERTMS track-side deployment	0.15	0.269	0.097	0.05	0.136
G5—ATP (Automatic Train Protection) coverage	0.252	−0.01	−0.096	−0.163	0.118
G6—Core maritime port connection	0.039	−0.315	0.053	−0.066	−0.087

All the selected variables have high interpretative value, as they provide a more accurate assessment of the indicators influencing railway infrastructure efficiency. Furthermore, these indicators can serve as a foundation for optimizing existing operational strategies, and support informed decision-making in railway system management.

The further evaluation is based on ten KPIs, categorized into input and output variables. The input variables include the degree of main railway line electrification (C1), coverage of the Automatic Train Protection (ATP) system (G5), share of the national network managed by the RIM (C5), connectivity with maritime ports (G6), modal share of rail freight transport (C3), and track failures relative to network size (D5). On the other hand, the output variables consist of infrastructure manager-related safety indicators (S3), the share of electric trains in total traffic (S5), the number of serious accidents (S2), and the percentage of cancelled trains due to infrastructure-related issues (P4). These indicators serve as the foundation for assessing factor significance using the G-BWM and evaluating efficiency through the AR DEA method. In Step 3, the G-BWM was applied to determine the weight boundaries for inputs and outputs. The process began with the definition of the grey scale for indicator evaluation (Table 2), utilizing interval values derived from grey number theory. Once the scale was defined, experts identified the most important (Best) and least important (Worst) input and output, which acted as reference values for evaluating the remaining indicators. Based on these reference values, the remaining input and output variables were assessed relative to the best and worst indicators, forming grey vectors. The first vector represents the preference of the best indicator over the others, while the second vector reflects the preference of all indicators concerning the worst indicator. The obtained scores were then used in an optimization model, where formulas (5) to (12) were applied to compute the lower and upper weight boundaries for each input and output. This process also included a consistency check to ensure that the expert assessments were logically coherent and mutually aligned, thereby increasing the reliability and robustness of the weighting process. Table 5 presents the input evaluations obtained using the G-BWM, where inputs are assessed relative to the most important (Best) and least important (Worst) criteria, while the final column displays the corresponding grey weights. Based on the results, the degree of main railway line electrification (C1) was identified as the most important input, having the highest weight range. In contrast, the modal share of rail freight transport (C3), the share of the national network managed by the RIM (C5), and ATP system coverage (G5) were assigned approximately equal importance. The lowest weight was assigned to track failures relative to network size (D5), indicating that while this indicator remains relevant for analysis, it has the least impact on overall infrastructure efficiency.

Table 5. Input evaluations and grey weights of inputs.

Input KPI	Evaluation Relative to C1 (Best Input)	Evaluation Relative to D5 (Worst Input)	Grey Weights of Inputs
C1—Degree of electrification of total main track (Best input)	BI	FH [5, 7]	[0.412, 0.491]
C3—National modal share of railway in freight transport	M [4, 6]	VL [1, 3]	[0.098, 0.101]
C5—Share of national network managed by RIM (%)	M [4, 6]	VL [1, 3]	[0.098, 0.101]
D5—Track failures in relation to network size (Worst input)	FH [5, 7]	WI	[0.054, 0.078]
G5—ATP (Automatic Train Protection) coverage	M [4, 6]	VL [1, 3]	[0.098, 0.101]
G6—Core maritime port connection	L [2, 4]	FL [3, 5]	[0.152, 0.216]

Table 6 presents the output evaluations, where the output variables are assessed relative to the best (Best) and worst (Worst) output. The infrastructure manager-related safety indicator (S3) was identified as the most significant output, exhibiting the highest weight, while the number of serious accidents (S2) was recognized as the least significant output, with the lowest weight range. The share of electric trains (S5) and the percentage of cancelled trains due to infrastructure issues (P4) demonstrated moderate importance in the performance evaluation, with P4 having slightly lower weights compared to S5. These results indicate that in assessing the efficiency of railway infrastructure managers, the greatest emphasis is placed on safety indicators associated with infrastructure management.

Table 6. Output evaluations and grey weights of outputs.

Output KPI	Evaluation Relative to S3 (Best Output)	Evaluation Relative to S2 (Worst Output)	Grey Weights of Outputs
S2—Fatalities and serious injuries (Worst output)	M [4, 6]	WO	[0.078, 0.125]
S3—Infrastructure manager-related precursors to accidents (Best output)	BO	M [4, 6]	[0.475, 0.596]
S5—Share of electricity-powered trains	L [2, 4]	L [2, 4]	[0.181, 0.250]
P4—Percentage of passenger train cancellations caused by RIM	FL [3, 5]	VL [1, 3]	[0.145, 0.150]

After determining the input and output weights using the G-BWM, Step 4 applied the AR-DEA method to assess the efficiency of railway infrastructure managers (RIMs). This analysis was based on ten key performance indicators (KPIs), where C1, C3, C5, D5, G5, and G6 were defined as inputs, while S2, S3, S5, and P4 were designated as outputs. Based on the previous step, each input and output were assigned weight limits and optimized through the G-BWM.

In the first step of the AR-DEA method, input and output matrices for all decision-making units (DMUs) were formed, as defined in formulas (13) and (14). These matrices enabled the evaluation of each unit's efficiency based on the available data. Subsequently, the data were normalized to ensure comparability across all inputs and outputs, following formulas (15) and (16). Once normalization was completed, efficiency scores for each DMU were computed by solving an optimization problem, where the weighted sum of outputs was maximized under the condition that the weighted sum of inputs equaled 1, as shown in formulas (17) and (18). The predefined input and output weight limits obtained from the G-BWM were incorporated as constraints in the model, as represented in formulas (19) and (20). Finally, scaling factors were introduced to ensure that the weight limits remained within the defined intervals, as described in formulas (21) and (22). By solving this optimization problem, efficiency scores were determined for each DMU. Units with an efficiency score of 1 are considered fully efficient, while those with lower efficiency scores indicate areas for potential improvement. Table 7 presents the input and output values for the analyzed railway infrastructure managers, along with their efficiency scores calculated using the AR-DEA method. The table displays the six selected inputs (C1, C3, C5, D5, G5, G6) and four outputs (S2, S3, S5, P4), which serve as the basis for determining efficiency coefficients for each RIM.

The results indicate that RIM8 is the only infrastructure manager with an efficiency score of 1.000, making it the reference unit in the analysis, and demonstrating optimal resource utilization. The other RIMs exhibit varying levels of efficiency, with the lowest efficiency scores recorded for RIM2 (0.190) and RIM4 (0.222). Among the non-optimal units, RIM6 achieves the highest efficiency score (0.341), suggesting that it is approaching optimal performance, but still has room for improvement. Analyzing the input and output values

reveals that units with lower efficiency are often characterized by high values of negative outputs (S2 and P4) or inefficient use of inputs, such as a high number of track failures (D5) and suboptimal electrification utilization (C1). Conversely, efficient RIMs maintain a better balance between inputs and outputs, indicating more effective infrastructure management. These results enable the identification of potential improvements for less efficient units through resource optimization and performance enhancement of key indicators.

Table 7. Input/output values and efficiency.

DMU	C1	C3	C5	D5	G5	G6	S2	S3	S5	P4	Effic.
RIM1	66	15.3	4138	110	19	100	0.09	2.7	85	50	0.295
RIM2	78	19	55,198	1	98	100	0.1	0.16	70	18	0.190
RIM3	76	10.7	3225	145	78	100	0.54	1.75	80	4	0.261
RIM4	73	22	27,301	198	6	100	0.61	2.8	82	8	0.222
RIM5	71	6.4	5972	172	100	100	0.07	0.5	89	24	0.301
RIM6	100	33.6	7891	35	100	0	0.15	0.2	98	3	0.341
RIM7	70	10.6	47,870	8	79	87	0.12	0.9	80	22	0.222
RIM8	46	22.8	11,473	720	12	0	0.45	0.01	56	23	1.000
RIM9	86	28.8	11,834	145	85	100	0.1	3.7	95	25	0.227

Sensitivity Analysis

The sensitivity analysis presented in Table 8 evaluates the stability of the obtained efficiency results across different scenarios of weight coefficient modifications. In the first four scenarios (Sc.1–Sc.4), the weight of the most important input was gradually reduced by 25%, 50%, 75%, and 90%, while in the following four scenarios (Sc.5–Sc.8), the same modifications were applied to the most important output. In Sc.9, the weights of all outputs were equalized, whereas in Sc.10, the weights of both inputs and outputs were balanced to test the model's stability under completely uniform criteria. To enhance the clarity of the sensitivity analysis, the results are visually represented in Figure 3, which facilitates the identification of efficiency trends across different scenarios.

The results indicate that changes in the weights of inputs and outputs had an expected impact on efficiency scores, but without drastic fluctuations, confirming the stability of the solution. The reduction in the most important input's weight (Sc.1–Sc.4) led to a gradual decrease in efficiency for all RIMs, with the least efficient units experiencing the most significant drop. In Sc.4, where the weight of the key input was reduced by 90%, the efficiency of RIMs such as RIM2 and RIM7 significantly decreased, confirming that this indicator had a strong influence on the initial results. Conversely, in scenarios Sc.5–Sc.8, where the weight of the most important output was reduced, an increase in efficiency was observed for all RIMs. As expected, the least efficient units experienced the highest efficiency gains, as reducing the significance of the key output lowered their limiting constraints in the model. In Sc.8, where the weight of the key output was reduced by 90%, some RIMs (RIM6, RIM5) achieved efficiency scores close to 1, suggesting that this output was the main restricting factor in the baseline analysis. Sc.9 and Sc.10, in which all output weights and both input and output weights were equalized, demonstrated that efficiency values remained within an acceptable range, without significant deviations. This further confirms the robustness of the model, as the results do not exhibit extreme fluctuations, even when all factors are treated equally. The sensitivity analysis validates the stability and reliability of the obtained results, showing that changes in key input and output weights affect efficiency in the expected direction. More efficient RIMs displayed lower sensitivity to weight modifications, while less efficient RIMs exhibited greater variations across different scenarios, indicating potential opportunities for optimizing their performance.

Table 8. Sensitivity analysis.

	Sc.0	Sc.1	Sc.2	Sc.3	Sc.4	Sc.5	Sc.6	Sc.7	Sc.8	Sc.9	Sc.10
RIM1	0.295	0.287	0.274	0.256	0.174	0.360	0.461	0.643	0.843	0.549	0.565
RIM2	0.190	0.177	0.161	0.139	0.090	0.227	0.284	0.386	0.501	0.344	0.244
RIM3	0.261	0.253	0.242	0.224	0.156	0.318	0.408	0.571	0.758	0.469	0.460
RIM4	0.222	0.214	0.203	0.186	0.129	0.271	0.347	0.486	0.641	0.371	0.357
RIM5	0.301	0.290	0.273	0.249	0.170	0.366	0.466	0.646	0.848	0.575	0.538
RIM6	0.341	0.341	0.342	0.343	0.311	0.411	0.520	0.717	0.957	0.629	0.627
RIM7	0.222	0.211	0.195	0.173	0.118	0.270	0.345	0.480	0.628	0.402	0.333
RIM8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
RIM9	0.227	0.218	0.205	0.187	0.132	0.277	0.355	0.495	0.649	0.412	0.374

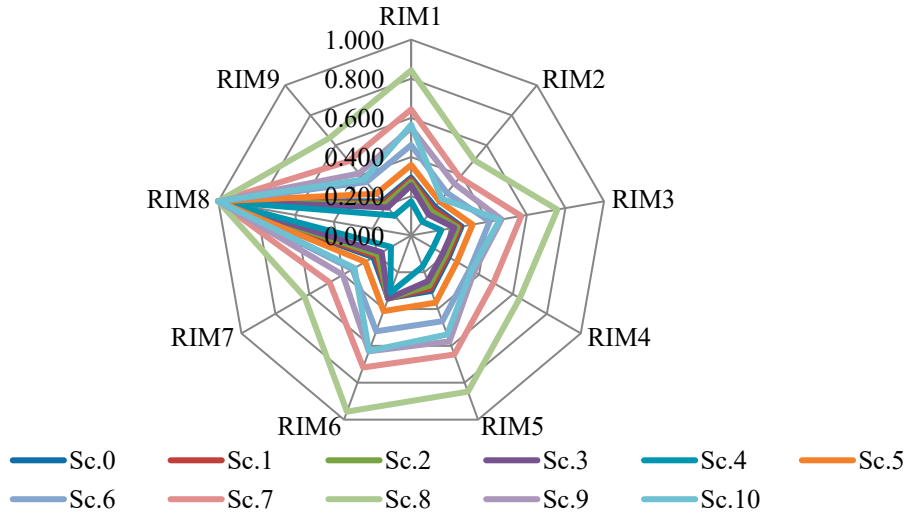


Figure 3. Sensitivity analysis.

5. Discussion

The application of a hybrid approach in this study, within the benchmarking analysis of RIMs, has enabled an objective assessment of their performance using a combination of PCA, G-BWM, and AR-DEA methods. This approach represents a significant advancement compared to traditional benchmarking methods, providing deeper insights into various aspects of railway infrastructure management and allowing for more precise strategic decision-making.

The proposed hybrid model is based on the DEA method, which identifies efficient DMUs using a set of KPIs as inputs and outputs. Compared to traditional performance evaluation approaches, DEA offers advantages such as handling multiple inputs and outputs simultaneously, not requiring predefined functional relationships, enabling direct comparisons between DMUs, and accommodating different units of measurement. However, conventional DEA has limitations, particularly in weight allocation, as it automatically assigns input and output weights without decision-maker (DM) influence. This can lead to extreme weighting scenarios in which a DMU appears efficient by emphasizing only one input and one output, while ignoring others. To address this, the AR DEA model introduces weight bounds, incorporating DM preferences and aligning the analysis with real-world conditions. Since defining these weight bounds relies on DM evaluations, a Multi-Criteria Decision-Making (MCDM) method is required. This study employs the Best–Worst Method (BWM), due to its simplicity, higher consistency, and reduced comparison requirements compared to alternatives like AHP and ANP. Additionally, the BWM is applied within a grey environment, which enhances decision-making in cases where DM

evaluations are imprecise or incomplete. The grey number structure, defined by upper and lower values, aligns well with the AR DEA model's need for weight constraints. Another drawback of conventional DEA is that the number of efficient DMUs tends to increase with more inputs and outputs, reducing discrimination power. To mitigate this, PCA is integrated into the model to identify the most critical KPIs, reducing dimensionality while preserving key information. PCA minimizes multicollinearity, enhances differentiation among DMUs, and improves computational efficiency. Thus, the hybrid model combines PCA for key input–output selection, the grey BWM for determining their priority and weight constraints, and AR DEA to derive truly efficient DMUs, ensuring a more realistic and effective evaluation framework.

In this specific case, a large number of KPIs are determined and monitored for the observed RIMs. Although at first glance, this seems like a positive thing, it makes it difficult to identify realistically effective RIMs. One of the reasons for this is that there is significant correlation and overlap between individual KPIs, and the other is that the existing and most commonly used methods for determining efficiency cannot adequately take into account a large number of inputs and outputs. In addition, there is no information on how important KPIs are in assessing the effectiveness of RIMs. The proposed model overcomes these problems by enabling the identification of key KPIs for the observed RIMs. Based on the factor loadings, which indicate the strength of the relationship between individual KPIs and principal components, PCA allowed a reduction in the problem size. At the same time, the AR DEA method identified effective RIMs, based on the limits (weight, importance) of the KPIs (input/output) obtained by applying the grey BWM for evaluation by DMs based on their experience and real-life circumstances. This made it possible to single out as efficient RIMs those that have high performance in the business segments that are truly important to the participants in these processes.

The results indicate that the most critical KPIs relate to the degree of electrification, ATP system coverage, and track failure rates. The degree of electrification directly influences energy efficiency, operational costs, and environmental sustainability, with higher electrification levels reducing reliance on fossil fuels and lowering carbon emissions. ATP system coverage enhances railway safety by preventing accidents caused by human error, ensuring compliance with speed limits, and reducing the risk of collisions, which is vital for maintaining operational reliability. Track failure rates serve as a key measure of infrastructure integrity, as frequent failures can lead to increased maintenance costs, service disruptions, and safety hazards, directly impacting network reliability. This underscores the essential role of infrastructure resources and safety measures in achieving effective railway infrastructure management. The results show that more efficient railway infrastructure managers exhibit a better balance between infrastructure resources and safety indicators, while less efficient RIMs are often burdened by a higher number of failures and lower coverage of advanced safety systems. A detailed analysis of the results reveals that RIM8 is the only fully efficient manager, with an efficiency score of 1.000, meaning it optimally utilizes available resources and can serve as a benchmark in the analysis. Other RIMs exhibit varying levels of efficiency, with RIM2 and RIM4 showing the lowest efficiencies of 0.190 and 0.222, respectively, indicating significant room for improvement. The efficiency results are primarily driven by the weighting of KPIs using the G-BWM and their values in the AR-DEA model. The proportion of the national network managed by a railway infrastructure manager affects strategic planning and investment efficiency, as larger networks require advanced asset management strategies to optimize performance and minimize operational risks. The degree of electrification of the total main track (C1) and infrastructure manager-related precursors to accidents (S3) had the highest weights, making them decisive factors. Precursors to accidents managed by infrastructure opera-

tors serve as early warning indicators, enabling proactive risk management and targeted interventions to enhance railway safety and prevent major disruptions. A higher share of electric trains in total traffic is critical for achieving sustainability goals, as electric traction significantly reduces air pollution, lowers energy costs, and enhances network efficiency. RIM2's inefficiency stems from its low value in infrastructure manager-related precursors to accidents (S3), which carries the highest weight. RIM8, despite having the highest track failures relative to network size (D5), achieved full efficiency due to strong performance in infrastructure manager-related precursors to accidents (S3), which had a greater influence in the AR-DEA model. Conversely, RIM4's inefficiency is driven by its high number of fatalities and serious injuries (S2) and infrastructure manager-related precursors to accidents (S3), both of which significantly impact efficiency due to their high weights. The number of fatalities and serious injuries is a fundamental safety KPI, as reducing these incidents is crucial for maintaining public trust, regulatory compliance, and the overall reliability of railway operations. High rates of passenger train cancellations due to infrastructure failures indicate operational inefficiencies and service reliability challenges, which can negatively impact customer satisfaction and financial performance. Strong connectivity with maritime ports is essential for intermodal transport efficiency, as it facilitates seamless freight transfers, reduces logistics costs, and enhances the competitiveness of rail freight networks. A higher modal share of rail freight transport indicates a well-utilized railway network, contributing to economic efficiency and sustainability by reducing road congestion, fuel consumption, and greenhouse gas emissions. These findings highlight that while infrastructure resources influence efficiency, safety, and operational stability, they are the most critical factors in performance evaluation.

The conducted sensitivity analysis confirmed the stability of the model, where the reduction in the weight of the key input was gradually reflected in lower efficiency scores, while the reduction in the most important output weight led to an increase in efficiency. Notably, less efficient RIMs were more sensitive to weight changes, suggesting that their performance is more dependent on specific infrastructure factors. In scenarios where input and output weights were equalized, efficiency scores remained within an acceptable range, further confirming the robustness of the approach.

The findings of this paper provide valuable insights that can assist infrastructure managers in making informed decisions, optimizing operational strategies, and defining future investments. Identified areas for improvement allow for targeted actions that can enhance the efficiency of the railway sector as a whole. Additionally, benchmarking analysis enables comparative evaluation, ensuring transparency and fostering competitiveness among infrastructure managers. Despite the significant advantages of the developed model, certain challenges and limitations remain. The reliability of results depends on the availability of accurate and up-to-date data, while subjectivity in expert judgement may affect the precision of weight coefficients in the G-BWM. Furthermore, the complexity of the AR-DEA method requires careful interpretation of results to ensure their practical applicability.

The hybrid model presented in this study could be adapted for other transport sectors, such as road and air transport, by incorporating sector-specific KPIs and adjusting the weighting process to reflect industry-specific efficiency drivers. For example, in road transport, key indicators could include congestion levels, vehicle emissions, and infrastructure maintenance costs, while in air transport, factors such as runway utilization, on-time performance, and the environmental impact of airport operations could be integrated. Additionally, the model's structured approach to efficiency evaluation could serve as a foundation for transport policy-making, offering data-driven insights to regulatory bodies for optimizing resource allocation, setting performance benchmarks, and promoting sustainable infrastructure development.

One of the key challenges in collecting and validating data for the evaluation of RIMs is the inaccessibility and inconsistency of data across different countries. The PRIME platform, an initiative of the European Commission for collecting data from RIMs, plays a crucial role in standardizing and making information available. However, data submission to the PRIME platform is not mandatory, leading to issues with data incompleteness. Many RIMs lack unified or standardized performance tracking systems, and some do not monitor all KPIs, which further complicates the benchmarking analysis. Additionally, a key limitation of this study is its exclusive focus on RIMs within the EU, where regulatory frameworks, operational standards, and data availability exhibit a relatively high degree of uniformity. Nevertheless, the proposed model and methodology can be adapted for application to RIMs outside Europe, provided that necessary modifications are made to accommodate distinct regulatory landscapes, operational complexities, and data accessibility constraints. As previously discussed, data unavailability poses a significant challenge even within the EU, and this issue becomes more pronounced in non-EU contexts, due to inconsistencies in data collection methodologies and KPI reporting standards. While the PRIME platform facilitates a certain level of data standardization among European RIMs, the non-mandatory nature of data submission results in gaps that impede direct comparability across different railway infrastructure networks.

Future research could refine benchmarking models by incorporating additional KPIs, such as infrastructure resilience metrics (e.g., track maintenance backlog, climate-related service disruptions) or digitalization indicators (e.g., level of automated train control, real-time monitoring adoption). Moreover, examining external variables like regulatory policy changes, economic shifts, and emerging technologies (e.g., AI-driven asset management) would enhance the adaptability and robustness of efficiency assessments in railway infrastructure management. The proposed hybrid model confirms its usefulness in the railway infrastructure sector, enabling managers to enhance their performance through objective and quantitative insights, thereby contributing to sustainable development and more efficient resource management.

6. Conclusions

The performance of RIMs is a key factor in ensuring the efficiency, reliability, and sustainability of the railway sector. As rail networks face increasing demands for capacity expansion and service optimization, ensuring efficient infrastructure management is essential to achieving long-term competitiveness and sustainability in the railway sector. Given the complexity of infrastructure operations and the necessity for data-driven decision-making, benchmarking provides a valuable framework for evaluating performance, identifying best practices, and guiding strategic improvements. Benchmarking analysis, as a methodological framework for comparative performance assessment, enables the identification of best practices, the determination of reference values, and the recognition of areas that require improvement. Through systematic performance comparison among different RIMs, benchmarking provides objective insights that support informed decision-making and the optimization of operational processes to enhance the sector's competitiveness and efficiency. This study has developed an integrated framework for evaluating the efficiency of railway infrastructure managers, combining PCA, the G-BWM, and AR-DEA. The proposed approach enables the identification of key performance factors, the determination of the relative importance of selected indicators, and an objective assessment of the efficiency of RIMs. The research results highlight that infrastructure and safety factors are crucial for the efficiency of railway infrastructure management. From an initial set of 35 KPIs, the PCA method identified six key inputs (C1, C3, C5, D5, G5, G6) and four outputs (S2, S3, S5, P4), covering the most relevant aspects of infrastructure,

safety, and operational performance. The application of the G-BWM identified the degree of electrification (C1) as the most important input, while safety indicators (S3) and the share of electric trains (S5) were recognized as the most significant outputs. The AR-DEA enabled the assessment of the relative efficiency of infrastructure managers, with RIM8 identified as the only fully efficient unit, while other RIMs demonstrated varying levels of efficiency. RIM2 and RIM4 recorded the lowest efficiency scores, indicating substantial room for improvement. The sensitivity analysis confirmed the robustness of the model, showing that less efficient RIMs were more sensitive to changes in input and output weights, whereas more efficient RIMs demonstrated greater resilience to variations in indicator weights.

The key contributions of the study are the following:

- Development of a Hybrid Benchmarking Framework: A novel methodological framework integrating PCA, the G-BWM, and AR-DEA for systematically assessing the performance of railway infrastructure managers (RIMs).
- Reliable Dimensionality Reduction with PCA: PCA effectively reduces data dimensionality, identifying the most relevant performance indicators while eliminating redundancy and multicollinearity.
- Objective Weight Determination Using the G-BWM: The Grey Best–Worst Method (G-BWM) ensures a structured and objective approach to weighting KPIs, even in conditions of uncertainty and incomplete information.
- Enhanced Efficiency Measurement with AR-DEA: The Assurance Region Data Envelopment Analysis (AR-DEA) method improves efficiency evaluation by incorporating realistic operational constraints and decision-maker preferences.
- Practical Guidelines for Performance Improvement: The study identifies key efficiency gaps and provides actionable recommendations for infrastructure managers to optimize resource allocation, improve maintenance strategies, and enhance overall network performance.
- Support for Strategic Decision-Making: The model offers a data-driven tool for policymakers and infrastructure managers to develop investment strategies, optimize operational performance, and ensure long-term railway sector sustainability.

The practical implications of this study provide actionable insights that RIMs can implement to enhance efficiency and optimize operations. For instance, RIMs with lower efficiency scores, such as RIM2 and RIM4, can prioritize targeted investments in infrastructure upgrades, particularly by increasing network electrification and enhancing accident prevention measures, both of which were identified as key efficiency drivers. Additionally, by leveraging the benchmarking results, underperforming RIMs can adopt best practices from high-performing counterparts, such as RIM8, which demonstrated optimal resource utilization and safety management. For maintenance optimization, infrastructure managers can use the findings to develop predictive maintenance programmes, reducing downtime and infrastructure-related delays. Moreover, regulatory bodies can utilize this model to allocate funding based on efficiency gaps, directing resources toward RIMs that require urgent improvements in operational reliability and safety. By integrating these insights into long-term strategic planning, RIMs can systematically improve network performance, enhance service quality, and contribute to a more sustainable and competitive railway sector.

Future research should explore the impact of various external factors that directly influence efficiency and performance, including economic, regulatory, and technological changes. For example, evolving EU regulations on railway liberalization, infrastructure funding, and safety standards can significantly alter how resources are allocated and performance is measured. Technological innovations—such as IoT-based monitoring, AI-driven asset management, and predictive maintenance—are also transforming operations by enabling more accurate and proactive decision-making. Moreover, the way in which

infrastructure managers are organized—whether under state control, private management, or hybrid models—can impact investment priorities and operational effectiveness. Considering these external drivers will help to ensure that future benchmarking models remain robust and practical for improving performance in an ever-changing industry. Additionally, the model could be expanded by incorporating additional criteria and adapting it to specific national and regional conditions, including the consideration of RIMs outside of the EU, where differences in regulatory frameworks and data availability may require further methodological adjustments. The proposed framework demonstrates high potential for practical application, providing reliable and objective insights that can serve as a foundation for decision-making aimed at improving the efficiency of the railway sector and achieving sustainable transport infrastructure development.

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

Funding: This research received no external funding.

Data Availability Statement: All data used in the study are presented within the paper.

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

References

1. Lema, N.M.; Price, A.D. Benchmarking: Performance Improvement Toward Competitive Advantage. *J. Manag. Eng.* **1995**, *11*, 28–37. [CrossRef]
2. Anand, G.; Kodali, R. Benchmarking the benchmarking models. *Benchmarking Int. J.* **2008**, *15*, 257–291. [CrossRef]
3. Ganguly, S.; Ray Chaudhury, S.; Mukherjee, A.; Si, A. Performance Analysis of Indian Railway Zones using MCDM Approaches. *Decis. Mak. Adv.* **2025**, *3*, 111–125. [CrossRef]
4. Lee, Y.; Zailani, S.; Soh, K.L. Understanding Factors for Benchmarking Adoption. *Benchmarking Int. J.* **2006**, *13*, 548–565.
5. Moffett, S.; Anderson-Gillespie, K.; McAdam, R. Benchmarking and Performance Measurement: A Statistical Analysis. *Benchmarking* **2008**, *15*, 368–381. [CrossRef]
6. Gus, M. Overcoming obstacles to developing effective performance measures. *Work. Study Emerald Group Publ. Ltd.* **1999**, *48*, 223–229. [CrossRef]
7. Yoshida, Y.; Fujimoto, H. Japanese-airport benchmarking with the DEA and endogenous-weight TFP methods: Testing the criticism of overinvestment in Japanese regional airports. *Transp. Res. Part E Logist. Transp. Rev.* **2004**, *40*, 533–546. [CrossRef]
8. Dias, J.; Azevedo, S.; Ferreira, J.J.; Palma, S. A comparative benchmarking analysis of main Iberian container terminals: A DEA approach. *Int. J. Shipp. Transp. Logist.* **2009**, *1*, 260–275. [CrossRef]
9. Egilmez, G.; McAvoy, D. Benchmarking road safety of U.S. states: A DEA-based Malmquist productivity index approach. *Accid. Anal. Prev.* **2013**, *53*, 55–64. [CrossRef] [PubMed]
10. Sharma, M.G.; Debnath, R.M.; Oloruntoba, R.; Sharma, S.M. Benchmarking of rail transport service performance through DEA for Indian railways. *Int. J. Logist. Manag.* **2016**, *27*, 629–649.
11. Hong, P.; Hong, S.; Roh, J.; Park, K. Evolving benchmarking practices: A review for research perspectives. *Benchmarking Int. J.* **2012**, *19*, 444–462. [CrossRef]
12. Blumenfeld, M.; Lin, C.Y.; Jack, A.; Abdurrahman, U.T.; Gerstein, T.; Barkan, C.P. Towards measuring national railways' safety through a benchmarking framework of transparency and published data. *Saf. Sci.* **2023**, *164*, 106188. [CrossRef]
13. Lin, C.Y.; Blumenfeld, M.; Gerstein, T.; Barkan, C.P.; Jack, A.; Abdurrahman, U.T. International benchmarking of railroad safety data systems and performance–a cross-continental case study. *J. Rail Transp. Plan. Manag.* **2023**, *26*, 100384. [CrossRef]
14. Rungskunroch, P.; Shen, Z.J.; Kaewunruen, S. Benchmarking environmental and economic impacts from the HSR networks considering life cycle perspectives. *Environ. Impact Assess. Rev.* **2021**, *90*, 106608. [CrossRef]
15. He, Q.; Wang, T.; Chan, A.P.; Xu, J. Developing a list of key performance indictors for benchmarking the success of construction megaprojects. *J. Constr. Eng. Manag.* **2021**, *147*, 04020164. [CrossRef]
16. Pugi, L.; Berzi, L.; Spedicato, M.; Cirillo, F. Hydrogen for railways: Design and simulation of an industrial benchmark study. *Int. J. Model. Identif. Control.* **2023**, *43*, 43–53. [CrossRef]

17. Zintel, M.; Pluchet, J.; Hensler, A.; Bamberger, V.; Baron, R.; Watson, S. *RAIL 2040 Scenarios and CEO Agenda: Report 2023*; Arthur D. Little: Brussels, Belgium, 2023. Available online: https://www.adlittle.com/sites/default/files/reports/ADL_Rail_2040_2023.pdf (accessed on 24 February 2025).
18. Platform of Rail Infrastructure Managers in Europe (PRIME). *2021 PRIME Benchmarking Report*; European Commission: Brussels, Belgium, 2021. Available online: https://wikis.ec.europa.eu/download/attachments/44167372/PRIME%20External%20Report%202021.pdf?version=2&modificationDate=1687506392630&api=v2 (accessed on 24 February 2025).
19. Yi, F.; Zhijiang, L. Factor Analysis of Railway Carrying Capacity Coordination Optimization Considering Energy Consumption. *J. Electr. Syst.* **2024**, *20*, 676–683.
20. Duleba, S.; Farkas, B. Principal Component Analysis of the Potential for Increased Rail Competitiveness in East-Central Europe. *Sustainability* **2019**, *11*, 4181. [CrossRef]
21. Braga, J.A.P.; Andrade, A.R. Multivariate statistical aggregation and dimensionality reduction techniques to improve monitoring and maintenance in railways: The wheelset component. *Reliab. Eng. Syst. Saf.* **2021**, *216*, 107932. [CrossRef]
22. Lajevardi, S.; Lourenco, P.; Sousa, H.; Matos, J. Railway reinforced concrete infrastructure life management and sustainability index. *Acta Polytech. CTU Proc.* **2022**, *33*, 316–321. [CrossRef]
23. Panja, S.; Patra, S.; Gangopadhyaya, M.; Poddar, S. Human factor analysis for railway coach and bogie maintenance using principal component analysis and factor analysis. *Int. J. Hum. Factors Ergon.* **2019**, *6*, 35. [CrossRef]
24. Kamran, G.; Dragan, P.; Sarbast, M.; Parastou, B.; Domokos, E.K.; Iraj, M. Decoding rail derailments: Unraveling the weighted factors influencing safety and sustainability using the best-worst method. *Results Eng.* **2024**, *23*, 102539. [CrossRef]
25. Wang, Y.; Liang, Y.; Li, C.; Zhang, X. Operation Performance Evaluation of Urban Rail Transit PPP Projects: Based on Best Worst Method and Large-Scale Group Evaluation Technology. *Adv. Civ. Eng.* **2021**, *2021*, 4318869. [CrossRef]
26. Noruzi, M.; Naderan, A.; Zakeri, J.A.; Rahimov, K. A Novel Decision-Making Framework to Evaluate Rail Transport Development Projects Considering Sustainability under Uncertainty. *Sustainability* **2023**, *15*, 13086. [CrossRef]
27. Binxin, Y.; Guangdong, W. Multi-criteria analysis of cross-regional railways interconnection under the post COVID-19 pandemic crisis: A hybrid BWM-FAISM-DFS evaluation framework. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *179*, 103316. [CrossRef]
28. Barasin, A.M.; Alqahtani, A.Y.; Makki, A.A. Performance Evaluation of Retail Warehouses: A Combined MCDM Approach Using G-BWM and RATMI. *Logistics* **2024**, *8*, 10. [CrossRef]
29. Krstić, M.; Tadić, S.; Agnusdei, L. Evaluating governance models in intermodal terminal operations: A hybrid grey MCDM approach. *J. Intell. Manag. Decis.* **2023**, *2*, 179–191. [CrossRef]
30. Krstić, M.; Elia, V.; Agnusdei, G.P.; De Leo, F.; Tadić, S.; Miglietta, P.P. Evaluation of the agri-food supply chain risks: The circular economy context. *Brit. Food J.* **2024**, *126*, 113–133. [CrossRef]
31. Tadić, S.; Krstić, M.; Kovač, M.; Brnjac, N. Evaluation of smart city logistics solutions. *Promet-Traffic Transp.* **2022**, *34*, 725–738. [CrossRef]
32. Niu, Y.; Li, X.; Zhang, J.; Deng, X.; Chang, Y. Efficiency of railway transport: A comparative analysis for 16 countries. *Transp. Policy* **2023**, *141*, 42–53. [CrossRef]
33. Li, W.; Hilmola, O.-P. Belt and Road Initiative and Railway Sector Efficiency—Application of Networked Benchmarking Analysis. *Sustainability* **2019**, *11*, 2070. [CrossRef]
34. Topcu, T.G.; Triantis, K.; Roets, B. Estimation of the workload boundary in socio-technical infrastructure management systems: The case of Belgian railroads. *Eur. J. Oper. Res.* **2019**, *278*, 314–329. [CrossRef]
35. Blagojević, A.; Vesković, S.; Kasalica, S.; Gojić, A.; Allamani, A. The application of the fuzzy AHP and DEA for measuring the efficiency of freight transport railway undertakings. *Oper. Res. Eng. Sci. Theory Appl.* **2020**, *3*, 1–23. [CrossRef]
36. Tadić, S.; Krstić, M.; Brnjac, N. Selection of efficient types of inland intermodal terminals. *J. Transp. Geogr.* **2019**, *78*, 170–180. [CrossRef]
37. Foroutan, S.L.F.; Bamdad, S. Efficiency measurement of railway passenger stations through network data envelopment analysis. *Res. Transp. Bus. Manag.* **2023**, *46*, 100767. [CrossRef]
38. Tang, G.; Zhang, M.; Bu, F. Vehicle Environmental Efficiency Evaluation in Different Regions in China: A Combination of the Life Cycle Analysis (LCA) and Two-Stage Data Envelopment Analysis (DEA) Methods. *Sustainability* **2023**, *15*, 11984. [CrossRef]
39. Marchetti, D.; Wanke, P.F. Efficiency in rail transport: Evaluation of the main drivers through meta-analysis with resampling. *Transp. Res. Part A Policy Pract.* **2019**, *120*, 83–100. [CrossRef]
40. Labib, A.; Anson, J. Benchmarking the sustainability reporting of high-speed railways (HSRs): Towards a state-of-the-art benchmarking and reporting framework for HSRs. *J. Clean. Prod.* **2020**, *250*, 119505. [CrossRef]
41. Stević, Ž.; Miškić, S.; Vojinović, D.; Huskanović, E.; Stanković, M.; Pamučar, D. Development of a Model for Evaluating the Efficiency of Transport Companies: PCA–DEA–MCDM Model. *Axioms* **2020**, *11*, 140. [CrossRef]
42. Muge, A.K.; Jia, Q.; Wang, Y.; Li, F.T. Exploring the Eco-Efficiency of Transport-Related Particulate Matter Pollution in Nairobi City, Kenya. *Open Access Libr. J.* **2020**, *7*, e7067. [CrossRef]

43. Omrani, H.; Amini, M.; Alizadeh, A. An integrated group best-worst method—Data envelopment analysis approach for evaluating road safety: A case of Iran. Meas. *J. Int. Meas. Confed.* **2020**, *152*, 107330. [CrossRef]
44. Anand, K.; Gurmeet, S.; Omkarprasad, S.V. A Comparative Evaluation of Public Road Transportation Systems in India Using Multicriteria Decision-Making Techniques. *J. Adv. Transp.* **2020**, *2020*, 8827186. [CrossRef]
45. Ramaganesh, M.; Rebekka, S.; Watson, V.; Subramanian, A.; Sankaranarayanan, B. Measurement of HOSPITAL efficiency in biomedical waste management. *Mater. Today Proc.* **2021**. [CrossRef]
46. Taghipour, F.; Hamid, M.; Aghakarimi, E.; Rabbani, M. An integrated framework to evaluate and improve the performance of emergency departments during the COVID-19 pandemic: A mathematical programing approach. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **2023**, *237*, 683–705. [CrossRef]
47. Brandenburg, M.; Hahn, G.J. *Financial Performance and Firm Efficiency of Automotive Manufacturers and Their Suppliers: A Longitudinal Data Envelopment Analysis*; Logistics Research; Bundesvereinigung Logistik (BVL): Bremen, Germany, 2021; Volume 14, pp. 1–26, ISSN 1865-0368. [CrossRef]
48. Kok, F.S.; Yongli, G.; Ming, M.Y. Enhancing logistics performance measurement: An effectiveness-based hierarchical data envelopment analysis approach. *INFOR Inf. Syst. Oper. Res.* **2024**, *62*, 449–479. [CrossRef]
49. Fumin, D.; Lin, X.; Yuan, F.; Qunxi, G.; Zhi, L. PCA-DEA-tobit regression assessment with carbon emission constraints of China's logistics industry. *J. Clean. Prod.* **2020**, *271*, 122548. [CrossRef]
50. Karami, S.; Ghasemy, Y.R.; Mousazadegan, F. Supplier selection and evaluation in the garment supply chain: An integrated DEA–PCA–VIKOR approach. *J. Text. Inst.* **2021**, *112*, 578–595. [CrossRef]
51. Adel, H.M.; Siavash, H.; Per, J.A. A strategy-based framework for supplier selection: A grey PCA-DEA approach. *Oper. Res.* **2020**, *22*, 263–297.
52. Wu, T.-H.; Chung, Y.-F.; Huang, S.-W. Evaluating global energy security performances using an integrated PCA/DEA-AR technique. *Sustain. Energy Technol. Assess.* **2021**, *45*, 101041. [CrossRef]
53. González-Gil, A.; Palacin, R.; Batty, P. Optimal energy management of urban rail systems: Key performance indicators. *Energy Convers. Manag.* **2015**, *90*, 282–291. [CrossRef]
54. Menglei, L. *Evaluation of Railway Performance Through Quality of Service*; College of Engineering and Physical Sciences—University of Birmingham: Birmingham, UK, 2016.
55. Siripong, P.; Wasaporn, T.; Waranon, D. A comparative study of performance measurement standards of railway operator. In Proceedings of the MATEC Web of Conferences, Sibiu, Romania, 7–9 June 2017; Volume 138, pp. 1–10. [CrossRef]
56. Fraszczyk, A.; Lamb, T.; Marinov, M. Are railways really that bad? An evaluation of rail systems performance in Europe with a focus on passenger rail. *Transp. Res. Part A Policy Pract.* **2016**, *94*, 573–591. [CrossRef]
57. Rungskunroch, P. Benchmarking Operation Readiness of the High-Speed Rail (Hsr) Network. Ph.D. Thesis, University of Birmingham, Birmingham, UK, 2022.
58. Kalem, A.; Tadić, S.; Krstić, M.; Čabrić, N.; Branković, N. Performance Evaluation of Railway Infrastructure Managers: A Novel Hybrid Fuzzy MCDM Model. *Mathematics* **2024**, *12*, 1590. [CrossRef]
59. Platform of Rail Infrastructure Managers in Europe. *Key Performance Indicators for Performance Benchmarking*; European Commission: Brussels, Belgium, 2022.
60. Hansen, I.A.; Wiggenraad, P.B.L.; Wolff, J.W. Benchmark Analysis of Railway Networks and Undertakings. In Proceedings of the International Seminar on Railway Operations Modelling and Analysis, Copenhagen, Denmark, 13–15 May 2013; pp. 1–20.
61. Cao, L.J.; Chua, K.S.; Chong, W.K.; Lee, H.; Gu, Q.M. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing* **2003**, *55*, 321–336. [CrossRef]
62. Brint, A.; Genovese, A.; Piccolo, C.; Taboada-Perez, G.J. Reducing Data Requirements When Selecting Key Performance Indicators for Supply Chain Management: The Case of a Multinational Automotive Component Manufacturer. *Int. J. Prod. Econ.* **2021**, *233*, 107967. [CrossRef]
63. Chae, S.C.; Choi, S.-Y. Analysis of the Term Structure of Major Currencies Using Principal Component Analysis and Autoencoders. *Axioms* **2022**, *11*, 135. [CrossRef]
64. Krstić, M.; Tadić, S.; Miglietta, P.P.; Porrini, D. Biodiversity Protection Practices in Supply Chain Management: A Novel Hybrid Grey Best–Worst Method/Axial Distance-Based Aggregated Measurement Multi-Criteria Decision-Making Model. *Appl. Sci.* **2025**, *15*, 1354. [CrossRef]
65. Podinovski, V.V. Side effects of absolute weight bounds in DEA models. *Eur. J. Oper. Res.* **1999**, *115*, 583–595. [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.