

Article

Environmental and Safety Performance of European Railways: An Integrated Efficiency Assessment

Arsen Benga ^{1,*}, María Jesús Delgado Rodríguez ² , Sonia de Lucas Santos ³  and Ghina El Mir ¹ 

¹ College of Business Administration, American University of the Middle East, Kuwait City 15453, Kuwait; ghina.el-mir@aum.edu.kw

² Universidad Rey Juan Carlos, 28032 Madrid, Spain; mariajesus.delgado@urjc.es

³ Universidad Autónoma de Madrid, 28049 Madrid, Spain

* Correspondence: arsen-benga@aum.edu.kw

Abstract

Railways play a pivotal role in advancing environmentally conscious and safe transportation systems, positioning them as a vital component of Europe's future mobility strategy. This study tackles the complex dimensions of sustainability in railway transport by combining environmental impacts and safety considerations within a single, integrated analytical framework. We extend the variable intermediate slack-based measure (VSBM) model to incorporate undesirable outputs—specifically accidents and emissions—allowing for a joint evaluation of safety and environmental performance. The revised model is applied to assess the operational efficiency of 14 European railway operators between 2010 and 2018. Compared to conventional efficiency models, our enhanced VSBM approach provides improved discriminatory power and reveals significant changes in relative efficiency rankings. By integrating safety and environmental dimensions, this study contributes a new perspective on sustainable railway performance measurement.

Keywords: climate change mitigation; sustainable mobility; green transportation; efficiency; measurement; railway decarbonization; integrated assessment models

1. Introduction

Railway transport systems are fundamental to the socioeconomic fabric of modern societies. Their ability to efficiently and safely move vast quantities of passengers and goods contributes significantly to national development strategies. Railways are particularly valued for their environmental and safety advantages. According to the European Commission [1], rail accounts for less than 0.5% of transport-related greenhouse gas emissions (excluding upstream emissions, vehicle manufacturing, maintenance, and end-of-life, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2528 (accessed on 16 December 2021)) within the European Union, positioning it as one of the most environmentally sustainable modes of transportation. Furthermore, rail travel has proven to be exceptionally safe, with an average of only 0.3 fatalities per billion passenger-kilometers, a stark contrast to the approximately seven fatalities observed in automobile transport for the same distance [2].

These inherent strengths in both environmental and safety domains are increasingly relevant as the European Union (EU) intensifies its commitment to sustainable transport. The European Green Deal aims to cut emissions from the transport sector by 90% by the year 2050, reflecting a bold commitment to decarbonization [3]. In this context, rail is



Academic Editor: Thyago Celso Cavalcante Nepomuceno

Received: 19 November 2025

Revised: 8 December 2025

Accepted: 17 December 2025

Published: 22 December 2025

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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

expected to play a central role in achieving these sustainability goals by mitigating negative externalities typically associated with mobility—such as emissions, accidents, and health hazards [4].

Although rail transport has notable sustainability advantages, the sector continues to encounter major obstacles in maintaining steady progress across the three core dimensions of sustainable development—environmental protection, economic sustainability, and social equity [5]. A comprehensive assessment that integrates all these dimensions would offer a clearer and more actionable view of sectoral performance. However, most existing studies focus exclusively on one domain—either environmental or safety—leaving an important gap in holistic evaluation frameworks [6,7].

Sustainable transport policymaking demands robust tools to evaluate efficiency and progress. Benchmarking environmental and safety performance not only supports regulatory oversight but also helps operators identify operational weaknesses and prioritize areas for improvement. While prior research has separately addressed environmental [8,9] and transport safety [10,11], only a few studies have attempted to examine both aspects concurrently [5,7].

In reality, the environmental and safety dimensions of railway performance are interdependent. For instance, train accidents involving hazardous materials can lead to environmental degradation beyond immediate physical damage and human casualties [12]. The International Civil Defense Organization has highlighted the ecological risks posed by railway incidents, particularly those involving dangerous cargo. Conversely, environmental policies—such as those promoting fuel-efficient, lightweight vehicles—may inadvertently reduce structural integrity, thereby affecting safety standards [13]. These interactions suggest a potential trade-off between environmental and safety priorities, reinforcing the need for an integrated performance assessment model.

The main aim of this research is to provide an integrated assessment of environmental and safety performance among railway operators in Europe. To achieve this, we introduce an extended version of the Variable Intermediate Slack-Based Measure (VSBM) model, originally developed by [14], and adapt it to incorporate undesirable outputs. This non-parametric methodology provides a rigorous approach to benchmarking, monitoring, and improving performance within a network structure, offering nuanced insights into the efficiency of rail operations with respect to both safety and environmental concerns.

Data Envelopment Analysis (DEA) continues to be a prominent tool in sustainability studies because it can evaluate multiple inputs and outputs simultaneously without the need to specify explicit functional forms [8]. While classical DEA models treat organizations as “black boxes,” more recent network DEA models consider the internal processes, enabling more refined evaluations of complex systems [15,16]. The Slack-Based Measure (SBM), introduced by Tone [17], is especially well-suited for evaluating systems with undesirable outputs, as it considers input and output slacks explicitly.

This study employs a modified Slack-Based Measure (SBM) integrated into a network DEA framework to evaluate the environmental and safety performance of 14 European railway operators during the 2010–2018 period. These operators are participants in the UIC-CER sustainability commitment [18], which aims to promote the best practices in environmentally and socially responsible transport. The chosen model, as shown in Figure 1, separates railway operations into two stages—production and service—thereby reflecting the actual organizational structure and allowing for efficiency measurement at each stage.

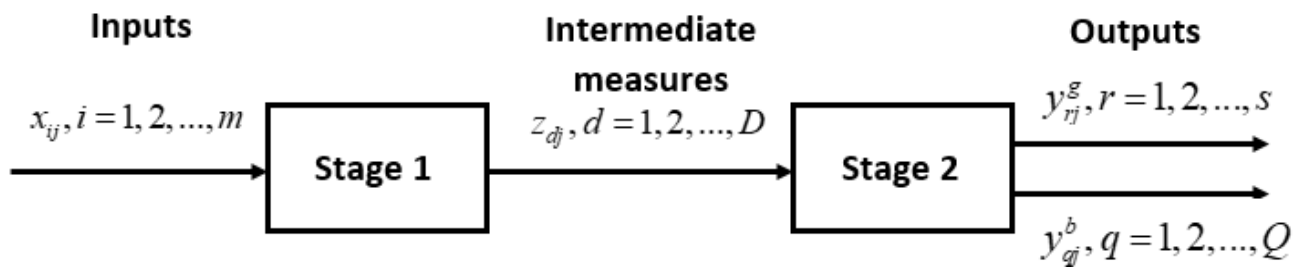


Figure 1. Schematic Representation of the Two-Stage Railway Production Framework.

To comprehensively assess sustainability in rail transport—particularly concerning safety and environmental impacts—we introduce an enhanced version of the Variable Intermediate Slack-Based Measure (VSBM) model. This approach builds upon the original framework of Chen et al. [19] and extends the dual-focus idea proposed by Wang [7], allowing for a unified evaluation of both safety and environmental efficiency. The model features several novel elements: it handles undesirable outputs, treats intermediate variables as flexible (free) elements within a two-stage production structure, and supports both input minimization and output maximization to optimize overall efficiency.

This study also makes several methodological contributions that differentiate it from prior research. Firstly, it is the first to implement a VSBM model incorporating undesirable outputs for the simultaneous evaluation of environmental and safety performance in the rail sector. Secondly, our framework explicitly accounts for potential trade-offs between these two dimensions by embedding them within a network DEA structure. Thirdly, we show that integrated and separate assessments (i.e., safety-only and environment-only models) can be complementary tools for diagnosing performance differentials. Fourthly, our ranking of railway operators using combined efficiency scores provides a practical benchmark for policymakers to compare and interpret performance across companies. Lastly, this joint framework offers actionable insights by enabling regulators and transport authorities to set evidence-based sustainability priorities, benchmark progress against top performers, and guide policy formulation using multidimensional efficiency measures.

The paper is organized to guide the reader through the study's key components. It begins with a review of relevant literature on sustainable transportation and DEA approaches (Section 2), followed by an explanation of the methodological framework and the extended VSBM model (Section 3). Section 4 outlines the dataset and defines the variables used in the analysis, while Section 5 discusses the findings from the efficiency assessments and comparative analyses. The final section (Section 6) concludes with a summary of the main insights and their implications for policy and practice.

2. Literature Review

This section reviews relevant literature concerning sustainable transportation and performance evaluation. It establishes the conceptual foundation for assessing environmental and safety efficiency in the railway sector and justifies the application of a modified DEA approach, specifically the Variable Intermediate Slack-Based Measure (VSBM) model with undesirable outputs.

2.1. Sustainable Mobility

The foundation of sustainable transport research in Europe traces back to the 1992 European Commission Green Paper, which emphasized the rising environmental and safety-related externalities of transportation and called for a cohesive community response [20]. Since then, the concept of sustainable mobility has evolved to encompass multiple dimen-

sions, including environmental protection, economic efficiency, social equity, technical innovation, and cultural adaptation [21].

In recent years, research on transportation sustainability has grown substantially, with comprehensive reviews by Zhao [22], Karjalainen & Juhola [23], Aloui et al. [24], and Kraus & Proff [25] documenting the evolution of the field. Environmental impacts, congestion, and safety risks have emerged as core themes, particularly in research related to policy development and system-level evaluation [2,26].

Several studies have investigated environmental performance at the firm level. For instance, McMullen & Noh [27] analyzed the environmental efficiency of U.S. bus transit agencies, while Fukuyama et al. [9] benchmarked the ecological performance of Japanese railways against the aviation sector. Oum et al. [28] compared inter-city transportation efficiency between rail and air services in Japan. Chang et al. [29] extended the analysis to the global airline industry using a slack-based DEA model with environmental constraints.

Despite stringent regulations, maintaining rail safety remains challenging due to the high complexity and interdependence of infrastructure, rolling stock, stations, and signaling systems. Despite the growing attention to transport safety in academic and policy circles, safety indicators remain underrepresented in efficiency studies [30,31]. A few exceptions exist—Kim et al. [11], for example, incorporated safety metrics in evaluating ferry operators in Korea, and Yu & Fan [32] assessed accident-related costs in Taiwanese bus services.

Only a limited number of works integrate both environmental and safety considerations in a single assessment framework. Most efforts in this direction rely on composite indicator methods [33–35]. This study builds on this line of inquiry by applying a network DEA framework that simultaneously evaluates environmental and safety outcomes, offering a more comprehensive view of sustainable rail transport. Building on Wang's [7] conceptualization, we propose a model that treats intermediate variables as free and undesirable outputs as integral performance indicators.

2.2. Transport Efficiency Measurement

Transport efficiency can be interpreted as achieving either reduced resource usage or increased societal benefit, or both [36]. It spans several facets, such as technical, allocative, and overall efficiency [37]. Accurate efficiency measurements are essential tools in policymaking, infrastructure planning, and resource allocation, especially when evaluating sustainable transport systems [38].

Various approaches have been developed to evaluate the performance of transportation systems at multiple scales—from national to firm-level assessments [39]. These methods are broadly categorized into parametric and non-parametric models. Stochastic Frontier Analysis (SFA), a key parametric approach, assumes a predefined production function [40], whereas non-parametric models like Data Envelopment Analysis (DEA) do not impose such constraints [41].

DEA is particularly suited for evaluating systems with multiple inputs and outputs and has become a central method in transport efficiency research [42]. This study applies to the VSBM model proposed by Chen et al. [14], which extends traditional DEA frameworks by accounting for undesirable outputs and network processes. Unlike earlier works, our study is the first to employ this model to analyze combined safety and environmental efficiency in railway systems.

2.3. Methodology Background

DEA was initially introduced by [41] and formalized by [43], is a linear programming technique used to evaluate the relative efficiency of homogeneous decision-making

units (DMUs). DEA allows for performance assessment without predefined input-output relationships, making it especially useful in complex, real-world applications [7,44].

DEA is particularly advantageous because it can manage multiple inputs and outputs, operate effectively even when price information is unavailable, and pinpoint the sources of inefficiency within the system [45,46]. However, challenges such as data sensitivity, limited discrimination power with high-dimensional data, and dependency on sample size remain [47,48].

To address internal process complexity, DEA has evolved into network DEA (NDEA), which disaggregates systems into multiple stages [49]. In two-stage models, outputs from the first stage become intermediate inputs for the second [16]. This structure is especially applicable to transportation systems with distinct production and service components [39,50,51].

2.4. Undesirable Outputs

Efficiency evaluations in sustainability studies often involve a mix of desirable and undesirable outputs. While the goal is to maximize positive outputs, minimizing negative (undesirable) outputs—such as pollution or accidents—is equally critical. Conventional DEA models are not designed to manage undesirable outputs directly, necessitating methodological extensions [52].

Indirect methods reclassify undesirable outputs as inputs or transform them using inverse values [53–55]. In contrast, direct methods retain the original data structure while modifying the DEA model, as in the slack-based measure (SBM) [56], range-adjusted measure (RAM), and directional distance function (DDF) [57,58].

Among these, SBM-DEA is widely adopted in transport research for both environmental [59,60] and safety evaluations [11]. Lozano [61] pioneered the integration of both dimensions within a single SBM model, demonstrating improved discrimination and relevance in transport applications. However, most implementations aggregate all system components, neglecting sub-process intricacies—an issue noted by Tone & Tsutsui [62].

Chen & Cook [63] extended DEA into a network framework that accommodates undesirable outputs, although their model does not always satisfy stage-level efficiency properties [64]. Addressing these gaps, Chen et al. [14] proposed the Variable Intermediate Slack-Based Measure (VSBM), a two-stage DEA model that integrates undesirable outputs, supports non-proportional input/output changes, and treats intermediates as free variables. Our study builds on this framework and, for the first time, applies it to rail transport to jointly evaluate sustainability in terms of environmental impact and safety performance.

3. Methodology

Given the complexity of railway systems and their dual impact on the environment and safety, it is essential to adopt analytical tools capable of capturing both desirable and undesirable outputs. Previous literature has shown that treating rail transport as a multi-stage process offers a more accurate reflection of real-world operations [50,65–68]. In this study, we extend the variable intermediate Slack-Based Measure (VSBM) model by incorporating undesirable outputs such as pollution and accidents, making it more applicable to the assessment of railway transport sustainability. This represents a novel contribution, as no previous studies have employed a VSBM framework that integrates both environmental and safety indicators in a network DEA model.

Following Chen et al. [14], we assume there are n decision-making units (DMUs), where each DMU represents an individual railway operator included in the analysis. As illustrated in Figure 1, Stage 1 uses m inputs (x_{ij}), like track length and energy consumption, to produce D intermediate products (z_{ij}), like train-kilometers. These intermediates serve as inputs in Stage 2, which produces s desirable outputs (y_{ij}^s), like passenger-kilometers

and freight ton-kilometers, and Q undesirable outputs (y_{ij}^b), like the number of significant accidents and CO₂ emissions. Table 1 provides a detailed description and classification of variables utilized in the analysis. The slack variables in our model retain the same interpretation as in Färe [56] and Chen et al. [14]. A slack in input indicates that a railway operator is using more infrastructure or energy than necessary compared to the best-performing operators, meaning these inputs could be reduced without affecting service delivery. A slack in desirable output shows that the operator is not fully utilizing its available resources and therefore could increase the amount of passenger- or freight transport produced. A slack in undesirable outputs reflects that the operator experiences higher levels of accidents or emissions than the efficient benchmark and must reduce these negative outcomes to reach the frontier. In all cases, a positive slack clearly identifies an area where operational performance can be improved. The structure allows us to assess the system's overall efficiency as well as the performance of each stage individually.

Table 1. Description and Classification of Variables Utilized in the Analysis (Source: UIC).

Input/Output		Code	Description
Input	Length of lines	X_1	Total length of lines—end of year (km)
	Total energy consumption	X_2	The total energy used for traction, services, and facilities (GWh).
Intermediate Output	Train-Kilometers	Z	All train movements of the operator (Thousand train-kilometers)
Desirable Output	Passenger-Kilometers	Y_1^g	Passenger traffic of the railway operator (domestic + international) (Million passenger kilometer)
	Ton-Kilometers	Y_2^g	Freight traffic of the railway operator (domestic + international) (Million ton-kilometer)
Undesirable Output	Accidents	Y_2^b	Number of accidents—total (Number)
	Location-based emissions	Y_1^b	Total CO ₂ eq location-based emissions (tCO ₂ eq).

3.1. Slack-Based Measure Model

Färe [56] proposed the SBM model to address inefficiencies by incorporating slacks directly into the objective function. The model is non-radial and non-oriented, making it suitable for performance evaluation where both input excesses and output shortfalls exist. The efficiency score is calculated based on the mean proportion of slack variables relative to actual input and output values.

Model 1

$$\min \theta = \frac{1 - \frac{1}{m} \sum_{i=1}^m \frac{s_i^-}{x_{io}}}{1 + \frac{1}{s} \sum_{r=1}^s \frac{v_r^g}{y_{ro}}}$$

$$s.t.$$

$$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io}, i = 1, 2, \dots, m,$$

$$\sum_{j=1}^n \lambda_j y_{rj}^g - v_r^g = y_{ro}^g, r = 1, 2, \dots, s,$$

$$s_i^-, v_r^+ \geq 0, i = 1, 2, \dots, m; r = 1, 2, \dots, s,$$

$$\lambda \geq 0, j = 1, 2, \dots, n.$$

where s_i^- and v_r^g are slack variables for inputs and outputs, respectively, and λ is the intensity vector.

Tone [17] extended the SBM model to account for undesirable outputs. In this adaptation, outputs that are considered detrimental—such as emissions or accidents—are treated in such a way that higher values reduce the efficiency score. This modification allows the model to assess sustainability more effectively by penalizing bad performance indicators.

Model 2

$$\begin{aligned} \min \theta &= \frac{1 - \frac{1}{m} \sum_{i=1}^m \frac{s_i^-}{x_{io}}}{1 + \frac{1}{s+q} \left(\sum_{r=1}^s \frac{v_r^g}{y_{ro}} + \sum_{l=1}^q \frac{v_l^b}{b_{lo}} \right)} \\ \text{s.t.} \\ X\lambda + s_i^- &= x_o \\ Y\lambda - v_r^g &= y_o \\ B\lambda + v_l^b &= b_o \\ s_i^-, v_r^g, v_l^b &\geq 0, i = 1, 2, \dots, m; r = 1, 2, \dots, s, q = 1, 2, \dots, Q, \\ \lambda, \mu &\geq 0, j = 1, 2, \dots, n. \end{aligned}$$

where b denotes undesirable outputs, and their slacks s_b^- reflect their negative efficiency impact.

3.2. Variable Intermediate Slack-Based Measure Model

Chen et al. [14] introduced the VSBM model to evaluate two-stage production systems. Their approach treats intermediate outputs as flexible and free variables that link the stages. This flexibility allows the model to adapt to variations in intermediate usage and more accurately reflect the actual structure of complex systems like railway networks. The overall system efficiency and stage-level projections are obtained through this network-based framework.

Model 3

$$\begin{aligned} \min \theta &= \frac{1 - \frac{1}{m} \sum_{i=1}^m \frac{s_i^-}{x_{io}}}{1 + \frac{1}{s} \sum_{r=1}^s \frac{v_r^g}{y_{ro}}} \\ \text{s.t.} \\ \sum_{j=1}^n \lambda_j x_{ij} + s_i^- &= x_{io}, i = 1, 2, \dots, m, \\ \sum_{j=1}^n \lambda_j z_{dj} &= \tilde{z}_{do}, d = 1, 2, \dots, D, \\ \sum_{j=1}^n \mu_j z_{dj} &= \tilde{z}_{do}, d = 1, 2, \dots, D, \\ \sum_{j=1}^n \mu_j y_{rj}^g - v_r^g &= y_{ro}^g, r = 1, 2, \dots, s, \\ s_i^-, v_r^g &\geq 0, i = 1, 2, \dots, m; r = 1, 2, \dots, s, \\ \lambda, \mu &\geq 0, j = 1, 2, \dots, n. \end{aligned}$$

In models (1), (2), and (3), θ denotes the overall system efficiency. In DEA notation, the operator currently being evaluated is denoted as DMU_0 , while the remaining operators serve as reference units for benchmarking. The variable x_{ij} refers to input i of decision-making unit j (DMU_j), while x_{io} represents the corresponding input for the decision-making unit under evaluation (DMU_0). The term s_i^- captures the input slack for a DMU under evaluation (DMU_0). Similarly, z_{dj} is intermediate output d for DMU_j , and $\tilde{z}_{do}$ is the counterpart of DMU_0 . For outputs, y_{rj} denotes the desirable output r of DMU_j , y_{ro}

represents the output r of DMU_0 , and v_r^s represents the slack of the corresponding desirable output. The variables λ and μ correspond to the intensity parameters for each stage of the production process. A key feature of this model is that the intermediate variables $\tilde{z}_{do}$ are treated as free variables, following [14].

3.3. Variable Intermediate Slack-Based Measure with Undesirable Outputs

To broaden the applicability of the VSBM framework, this study proposes a refined model that explicitly accounts for undesirable outputs. Specifically, Stage 2 of the production process is modeled to generate both desirable outputs and undesirable byproducts, such as emissions and accident occurrences. It is assumed that, alongside the s desirable outputs $y_{rj}^s, r = 1, 2, \dots, s$, the second stage of the process also generates Q undesirable outputs $y_{qj}^b, q = 1, 2, \dots, Q$. Building upon the methodologies developed by Chen et al. [14], Färe [56], and Chen and Cook [63], the proposed VSBM model explicitly integrates undesirable outputs into the efficiency assessment. These undesirable outputs are incorporated into the denominator of the efficiency function, such that any increase in their magnitude proportionally decreases the overall efficiency score.

Model 4

$$\begin{aligned} \min \theta &= \frac{1 - \frac{1}{m} \sum_{i=1}^m \frac{s_i^-}{x_{i0}}}{1 + \frac{1}{s+Q} \left(\sum_{r=1}^s \frac{v_r^s}{y_{r0}^s} + \sum_{q=1}^Q \frac{v_q^b}{y_{q0}^b} \right)} \\ \text{s.t.} \\ \sum_{j=1}^n \lambda_j x_{ij} + s_i^- &= x_{i0}, i = 1, 2, \dots, m, \\ \sum_{j=1}^n \lambda_j z_{dj} &= \tilde{z}_{do}, d = 1, 2, \dots, D, \\ \sum_{j=1}^n \mu_j z_{dj} &= \tilde{z}_{do}, d = 1, 2, \dots, D, \\ \sum_{j=1}^n \mu_j y_{rj}^s - v_r^s &= y_{r0}^s, r = 1, 2, \dots, s, \\ \sum_{j=1}^n \mu_j y_{qj}^b + v_q^b &= y_{q0}^b, q = 1, 2, \dots, Q, \\ s_i^-, v_r^s, v_q^b &\geq 0, i = 1, 2, \dots, m; r = 1, 2, \dots, s, q = 1, 2, \dots, Q, \\ \lambda, \mu &\geq 0, j = 1, 2, \dots, n. \end{aligned}$$

The slack variables $-v_r^s$ and $+v_q^b$ associated with desirable and undesirable outputs are included in the denominator of the efficiency function. Their values are inversely related to the performance level, meaning that a decrease in desirable outputs or an increase in undesirable outputs leads to a lower efficiency score. In essence, higher slack in good outputs or greater generation of bad outputs results in a decline in overall efficiency [7].

Let,

$$\begin{aligned} t &= \frac{1}{1 + \frac{1}{s+Q} \left(\sum_{r=1}^s \frac{v_r^s}{y_{r0}^s} + \sum_{q=1}^Q \frac{v_q^b}{y_{q0}^b} \right)}, \\ S_i^- &= t s_i^-, V_r^s = t v_r^s, V_q^b = t v_q^b, \\ z'_{do} &= t \tilde{z}_{do}, \Lambda = t \lambda, M = t \mu \end{aligned}$$

Furthermore, following the transformation method proposed by Benga & Delgado Rodríguez [69], the non-linear model is reformulated as a linear programming problem to simplify the computational process. This transformation enables the model to minimize in-

put and undesirable outputs while maximizing desirable outputs, offering a comprehensive measure of transport sustainability within railway systems.

Model 5

$$\begin{aligned}
 \min \varphi &= t - \frac{1}{m} \sum_{i=1}^m \frac{S_i^-}{x_{io}} \\
 \text{s.t.} \\
 t + \frac{1}{s+Q} \left(\sum_{r=1}^s \frac{V_r^g}{y_{ro}^g} + \sum_{q=1}^Q \frac{V_q^b}{y_{qo}^b} \right) &= 1 \\
 \sum_{j=1}^n \Lambda_j x_{ij} + S_i^- &= t x_{io}, i = 1, 2, \dots, m, \\
 \sum_{j=1}^n \Lambda_j z_{dj} &= z'_{do}, d = 1, 2, \dots, D, \\
 \sum_{j=1}^n M_j z_{dj} &= z'_{do}, d = 1, 2, \dots, D, \\
 \sum_{j=1}^n M_j y_{rj}^g - V_r^g &= t y_{ro}^g, r = 1, 2, \dots, s, \\
 \sum_{j=1}^n M_j y_{qj}^b + V_q^b &= t y_{qo}^b, q = 1, 2, \dots, Q, \\
 S_i^-, V_r^g, V_q^b &\geq 0, i = 1, 2, \dots, m; r = 1, 2, \dots, s, q = 1, 2, \dots, Q, \\
 \Lambda, M &\geq 0, j = 1, 2, \dots, n.
 \end{aligned}$$

The variable t is implicitly constrained by the conditions outlined in Model (5) and therefore does not appear explicitly in the list of constraints. The optimal solution to the linear programming formulation (Model (5)) serves as a basis for solving the corresponding linear fractional program presented in Model (4).

$$\begin{aligned}
 \theta^* &= \varphi^*; \lambda_j^* = \frac{\Lambda_j^*}{t^*}, i = 1, 2, \dots, n; \mu_j^* = \frac{M_j^*}{t^*}, j = 1, 2, \dots, n; s_i^{-*} = \frac{S_i^{-*}}{t^*}, i = 1, 2, \dots, m; \\
 v_r^{g*} &= \frac{V_r^{g*}}{t^*}, r = 1, 2, \dots, s; v_q^{b*} = \frac{V_q^{b*}}{t^*}, q = 1, 2, \dots, Q; \tilde{z}_{do}^* = \frac{z'_{do}}{t^*}, d = 1, 2, \dots, D
 \end{aligned}$$

4. Data Description

The variables included in the railway operator efficiency model, as well as their measurement indicators, were determined by drawing on established studies and the operational features unique to railway systems. The primary data source for this study was the database provided by the International Union of Railways (UIC), accessible via <https://uic-stats.uic.org/> (accessed on 15 January 2020) for 14 European railway companies over the period 2010–2018. In addition, environmental performance data related to rail transport was obtained directly from the UIC's environmental performance database, with supplementary information provided by representatives of the UIC Environment Department through private email correspondence; this dataset is based on the voluntary reporting guideline developed by the UIC in 2011 for member railways that signed the UIC Declaration on Sustainable Mobility and Transport, ensuring harmonized and comparable sustainability data across operators (<https://uic.org/sustainability/> (accessed on 20 August 2025)).

Given the inherent network structure of railway operations—spanning from input resources to output transport services—this study follows prior work by Charnes and Cooper [68] in considering rail transport as a multi-stage production process. Consequently, the railway production and service operations are categorized into two distinct stages, in alignment with the modeling approach proposed by Zhou and Hu [67].

Figure 2 shows that the first stage of production relies on two main inputs—total energy consumption and the length of operated lines—to produce a single intermediate output, measured in train-kilometers. Here, total energy consumption refers to the overall

“energy usage”, encompassing traction, services, and facility energy needs. The length of lines worked signifies the total track length operated within the reporting year, inclusive of the average extent of newly opened or closed lines and is used to reflect the infrastructural scale of the railway operator.

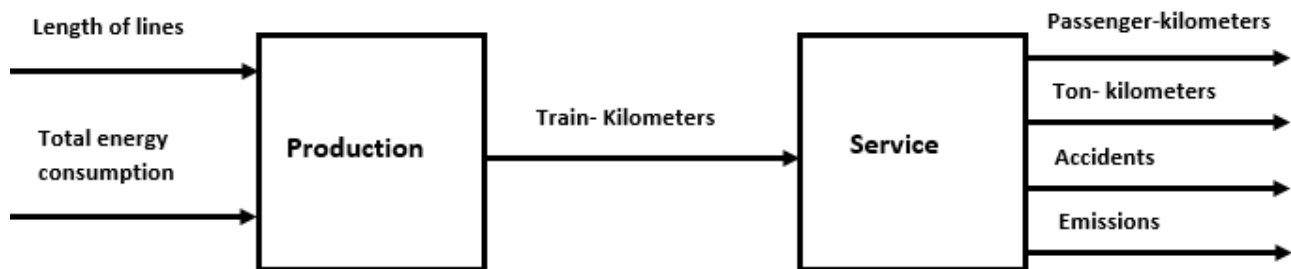


Figure 2. Structural Layout of the Railway Transport System in Two Stages.

The Train-Kilometers metric, encompassing all traction types, measures the movement of a train over a one-kilometer distance. Literature considers also other intermediary variables, like separating passenger-train-kilometers from freight-train-kilometers, or utilizing vehicle-kilometers, ton-kilometers of rolling stock, or capacity-kilometers. But these indicators aren’t always available for all operators, or they would break up the study in a way that makes it harder to compare. To keep the depiction of network use consistent, we use total train-kilometers as the intermediate output. For this same reason, only operators that offer both passenger and freight services are in the sample. Operators that only offer one type of service are not. This makes sure that all DMUs have a similar production structure and can be compared in a meaningful way using the same two-stage methodology.

In the subsequent stage (Stage 2—service delivery), this intermediate output serves as a sole input to produce the final outputs—transportation services for passengers and freight—as well as undesirable outputs such as accidents and emissions. Desirable outputs are represented by passenger and freight turnover. Passenger turnover, covering both domestic and international services, is obtained by multiplying the total number of passengers by their average travel distance. Similarly, freight turnover is calculated as the total volume of freight moved multiplied by the average distance transported per ton. Accidents, the sole safety-related undesirable variable, encompass collisions, derailments, injuries or fatalities, and level-crossing accidents. Regarding environmental impact, key pollutants generated by railway operations include nitrogen oxides (NO_x), particulate matter (PM), carbon monoxide (CO), and volatile organic compounds, as reported by the Community of European Railway and Infrastructure Companies (CER) via <https://www.cer.be> (accessed on 16 December 2025). For this study, location-based emissions (expressed as CO₂ equivalents) are used, reflecting the global warming potential (GWP) of the total greenhouse gas emissions in GWP/CO₂ units. All variables alongside their corresponding indicator definitions are presented in Table 1.

Several constraints affected the final composition of the dataset. First, the UIC data varied with respect to the extent of coverage across train operators and the availability period for each metric. Second, the analysis targeted railway operators offering both passenger and freight services to ensure a consistent basis for comparing DMUs. Lastly, this research concentrated on European railway operators participating in the UIC-CER sustainability initiative; however, only 18 of these companies submitted complete and reliable data, as reported by the UIC.

For consistency purposes, any railway operator with incomplete data for the chosen variables in a specific year was omitted from the sample. After this filtering process, the final dataset consisted of an unbalanced panel of 840 observations covering seven variables

for 14 European railway companies over the period 2010–2018. The final sample consisted of 14 railway operators: Germany (DB AG), Poland (PKP), Italy (FS), France (SNCF), Spain (RENFE), Austria (ÖBB), Belgium (SNCB), Bulgaria (BDZ), Switzerland (SBB), Czech Republic (CZ), Finland (VR), Portugal (CP), Romania (CFR), and Slovenia (SZ). Table 2 summarizes the descriptive statistics of all variables used in the analysis, providing the average value across the study period for each operator, with the corresponding standard deviation shown in parentheses to indicate the variability of performance over the years.

Table 2. Summary Statistics of Key Variables for the Railway Operators (Source: UIC).

	2010	2011	2012	2013	2014	2015	2016	2017	2018
Number of DMUs (count)	14	13	14	14	14	13	13	13	12
Length of lines (km)	11,439.7 (10,278.5)	11,978.5 (10,497.9)	11,341.1 (10,208.2)	11,419.8 (10,271.4)	11,421.4 (10,275.0)	11,868.9 (10,262.1)	11,898.3 (10,315.2)	10,790.7 (10,345.5)	11,270.3 (10,436.6)
Total energy consumption (GWh)	3257.4 (3681.9)	3380.2 (3691.4)	3035.7 (3537.7)	3045.4 (3533.6)	2946.9 (3368.4)	3088.5 (3283.2)	3038.5 (3151.1)	2835 (3195.1)	2999.6 (3150.1)
Train-Kilometers (Thousand train-kilometers)	196,663.4 (230,881.2)	206,733.3 (230,616.5)	188,214.5 (221,582.4)	185,710.3 (27,780.8)	183,374.8 (217,080.5)	188,859 (215,482.9)	194,486.3 (219,549.9)	187,050.7 (217,681.4)	199,791.3 (217,767.3)
Passenger-Kilometers (Million passenger kilometer)	21,746.2 (27,668.3)	23,032.4 (28,138.3)	21,131.6 (28,066.4)	21,330.7 (26,694.5)	21,395.7 (27,543.6)	22,752.5 (28,389.4)	23,591.6 (29,158)	23,888.8 (31,014.1)	25,609.7 (30,719.2)
Ton-Kilometers (Million ton-kilometer)	19,120.6 (26,713.6)	21,892.4 (29,329.7)	17,273.2 (20,394.3)	18,880.9 (27,780.8)	19,292 (26,308.5)	20,371.6 (25,604.1)	18,837.9 (24,446.1)	17,534.5 (24,300.6)	17,393.7 (23,627.7)
Accidents (count)	146.1 (209.4)	149.7 (222.3)	166.5 (227.9)	179.4 (236.3)	155 (193.4)	99.6 (100.7)	85 (92.6)	115 (117.7)	94.8 (103.1)
Emissions (tCO ₂ eq)	1146.3 (1557.5)	1243.1 (1674.4)	1107.9 (1595.1)	1043.5 (1551.1)	961.8 (1456.9)	997.2 (1375.6)	966.5 (1316.3)	819.3 (1204.4)	819.5 (1182.3)

Given the limited number of decision-making units (DMUs), using conventional cross-sectional DEA could result in inefficiencies. To mitigate this issue, the study employed a pooled approach in which each company-year observation is treated as an independent DMU, as suggested by Wang [7].

According to Kim et al. [11], ensuring isotonicity—meaning a non-decreasing relationship between inputs and outputs—is essential for obtaining reliable outcomes in DEA studies. This principle implies that a higher input level should lead to a corresponding increase in output during the production process. As indicated in Table 3, all selected input and output variables exhibit positive correlations that are statistically significant at the 1% level, confirming that the isotonicity assumption holds. A low correlation between accidents and inputs is expected, because accident levels are largely determined by safety practices and operational controls rather than system size. Therefore, instead of causing

instability, this variable strengthens the model by introducing additional discriminatory information on safety performance.

Table 3. Correlation Coefficients among Inputs and Outputs.

	Length of Lines	Train-km	Location-Based Emissions	Accidents	Passenger-km	Ton-km
Length of Lines	1	0.888 **	0.811 **	0.394 **	0.902 **	0.769 **
Total energy consumption	0.936 **	0.980 **	0.844 **	0.261 **	0.961 **	0.856 **
Train-km	0.888 **	1	0.882 **	0.257 **	0.920 **	0.905 **

** $p < 0.01$.

This study employs a comparable publicly accessible dataset to that of Benga et al. [70] and Charles et al. [71], yet it presents significant distinctions in the variables examined and the analytical framework utilized. The preceding research emphasizes operational indicators and demand conditions within a traditional single-stage DEA model. Conversely, the current study enhances the dataset by integrating accident statistics and CO₂ emissions, thereby facilitating the inclusion of negative outputs. The data are restructured into a two-stage network format, wherein operational activity (train-kilometers) is modeled independently from transport service delivery (passenger- and freight-kilometers). These extensions facilitate a comprehensive sustainability assessment and enable a more nuanced analysis of the trade-offs among safety, environmental performance, and operational efficiency.

5. Efficiency Evaluation of Railway Operations

5.1. Model Comparison and Validation

To begin, we compared the proposed Variable Slack-Based Measure (VSBM) model (4) with the conventional SBM model (2), which is commonly used in the literature for handling undesirable outputs. Table 4 presents the distribution of efficiency scores derived from both modeling approaches.

Table 4. Comparison of Efficiency Scores Between SBM and VSBM Models Incorporating Undesirable Outputs.

Score	Normal SBM	Variable SBM
<0.500	79	107
0.500–0.599	16	3
0.600–0.699	4	1
0.700–0.799	4	2
0.800–0.899	3	2
0.900–0.999	1	3
1	13	2
Maximum	1	1
Minimum	0.150	0.101
Mean	0.475	0.319
Std. Deviation	0.240	0.194

A key benchmark for assessing model performance is the count of Decision-Making Units (DMUs) that achieve full efficiency, as highlighted by Wang et al. [7]. Table 4 illustrates

that the VSBM model—by accounting for intermediate measures within the two-stage structure and permitting them to vary without restriction—enhanced the differentiation among DMUs. While the SBM model yielded thirteen fully efficient DMUs (104, 20, 86, 99, 109, 111, 116, 85, 45, 96, 59, 72, and 17), the VSBM model recognized only two (DMUs 111 and 85), indicating stronger discriminatory power. This reduction in the number of units achieving the maximum score demonstrates the stronger discriminatory power of the VSBM model, as it reveals efficiency gaps that remain hidden under a single-stage specification. In practical terms, the VSBM provides a more realistic performance benchmark by distinguishing between operators that simply appear efficient in aggregate and those that genuinely operate efficiently across all stages of the production process.

Notably, DMUs 111 and 85 remained efficient across both frameworks. The efficiency scores from both models exhibited a strong positive correlation ($r = 0.88$), suggesting that units rated highly under SBM also tended to rank well under VSBM. However, the VSBM approach provides a more refined perspective by revealing efficiency gaps hidden in the single-stage model—particularly those related to emissions and accident reduction. Therefore, the VSBM model was selected for further analysis, as it offers enhanced discriminatory power and a more comprehensive representation of the multi-stage production process and sustainability dimensions of railway operations.

5.2. Performance Analysis Across DEA Configurations

This section focuses on evaluating the sustainability of railway transport by incorporating both environmental and safety-related undesirable outputs and comparing the efficiency scores derived from our proposed model with those generated by conventional models that either ignore or include only one undesirable factor (environment or safety). With a two-stage modeling framework, we aim to assess the consistency of efficiency outcomes across four distinct configurations: one without undesirable outputs, one considering only environmental factors, one including only safety-related factors, and one encompassing both environmental and safety aspects.

The application of the proposed VSBM model (Model 4) was tested under the four configurations: an unconstrained setting (no undesirable outputs), an environmentally constrained model, a safety-constrained model, and a model that integrates both environmental and safety constraints. Following the approach outlined by Wang [7], we examined the results across three analytical perspectives:

5.2.1. Comparative Summary of Operator Efficiencies

Table 5 provides a summary of efficiency scores for each railway operator in the first and last years of the study period. The percentage change between these two values serves as an index to assess overall changes in operational performance under each of the four scenarios: θ_{ES} —both constraints included; θ_E —environmental constraint; θ_S —safety constraint; and θ —no constraint.

The analysis revealed a clear alignment between the results of the full model (both constraints included) and those of the single-constraint models (environmental and safety), with all companies showing consistent change directions across these three. Out of the fourteen railway operators, eleven demonstrated improvements over time, while only two—BDZ and RENFE—experienced a decline. Notably, none of the companies achieved full efficiency throughout the entire timeframe.

Table 5. Changes in Efficiency Scores of 14 UIC-CER Member Railway Operators in Europe Under Different Model Specifications.

Company	N	θ_{ES}			θ_E			θ_S			θ		
		First	Last	Δ	First	Last	Δ	First	Last	D	First	Last	Δ
OBB	9	0.344	0.622	81.02%	0.361	0.616	70.5%	0.369	0.622	68.7%	0.371	0.503	35.7%
SNCB	9	0.255	0.279	9.13%	0.267	0.287	7.6%	0.286	0.298	4.2%	0.172	0.171	−0.3%
BDZ	5	0.148	0.138	−6.81%	0.167	0.157	−6.0%	0.160	0.143	−10.5%	0.162	0.153	−5.6%
SBB	9	0.778	1.000	28.62%	0.816	1.000	22.5%	0.558	0.624	11.8%	0.395	0.486	23.1%
CD	9	0.145	0.181	24.40%	0.160	0.197	23.1%	0.162	0.201	24.4%	0.177	0.201	13.9%
DB	9	0.326	0.361	10.55%	0.328	0.368	12.3%	0.373	0.406	8.8%	0.378	0.426	12.7%
RENFE	9	0.400	0.241	−39.66%	0.168	0.153	−8.9%	0.399	0.236	−40.9%	0.100	0.089	−11.4%
VR	9	0.219	0.344	57.06%	0.223	0.314	40.3%	0.244	0.361	47.6%	0.237	0.294	24.0%
SNCF	9	0.300	0.474	58.10%	0.246	0.313	27.1%	0.281	0.434	54.2%	0.114	0.128	12.4%
FS	9	0.288	0.331	14.73%	0.274	0.318	16.0%	0.321	0.363	13.0%	0.176	0.212	19.9%
PKP	7	0.239	0.196	−18.07%	0.270	0.216	−19.8%	0.252	0.217	−13.8%	0.298	0.225	−24.7%
CP	9	0.228	0.256	12.27%	0.249	0.282	13.4%	0.237	0.271	14.2%	0.157	0.182	15.8%
CFR	9	0.104	0.133	27.59%	0.117	0.151	28.8%	0.112	0.140	24.9%	0.112	0.134	19.7%
SZ	9	0.243	0.310	27.83%	0.259	0.280	8.0%	0.255	0.280	9.5%	0.215	0.203	−5.5%

However, some divergence was observed between the unconstrained model and the three constrained configurations. For 12 out of 14 operators, performance trajectories were consistent in direction between the unconstrained and constrained models. SNCB and SZ were the only exceptions, showing negative trends in the unconstrained model but improved performance when undesirable outputs were considered. This outcome suggests that despite limited operational performance, SNCB and SZ performed relatively well in terms of sustainability, reinforcing the value of including environmental and safety factors in efficiency evaluations.

5.2.2. Temporal Dynamics of Efficiency

Figure 3 illustrates the evolution of efficiency over time across the four models. CD, CFR, CP, DB, FS, SZ, and VR displayed steady patterns over time, whereas OBB, SBB, SNCB, and SNCF experienced more noticeable fluctuations. For most operators, performance under the full model is tracked closely with that of either the environmental or safety model, although it often diverged from the unconstrained model.

The data suggests different dominant drivers across companies: environmental factors appear to drive performance for DB and SBB; safety appears to dominate for FS, CD, RENFE, and VR; and both factors influence companies like CFR, CP, OBB, BDZ, PKP, and SZ. Conversely, SNCB and SNCF demonstrated erratic behavior that may be attributable to external managerial or economic influences. For example, SNCB's performance improvement post-2014, likely related to structural reforms that merged NMBS/SNCB-Holding into a single SNCB entity. These reforms may have enhanced customer service and streamlined operations, positively impacting efficiency.

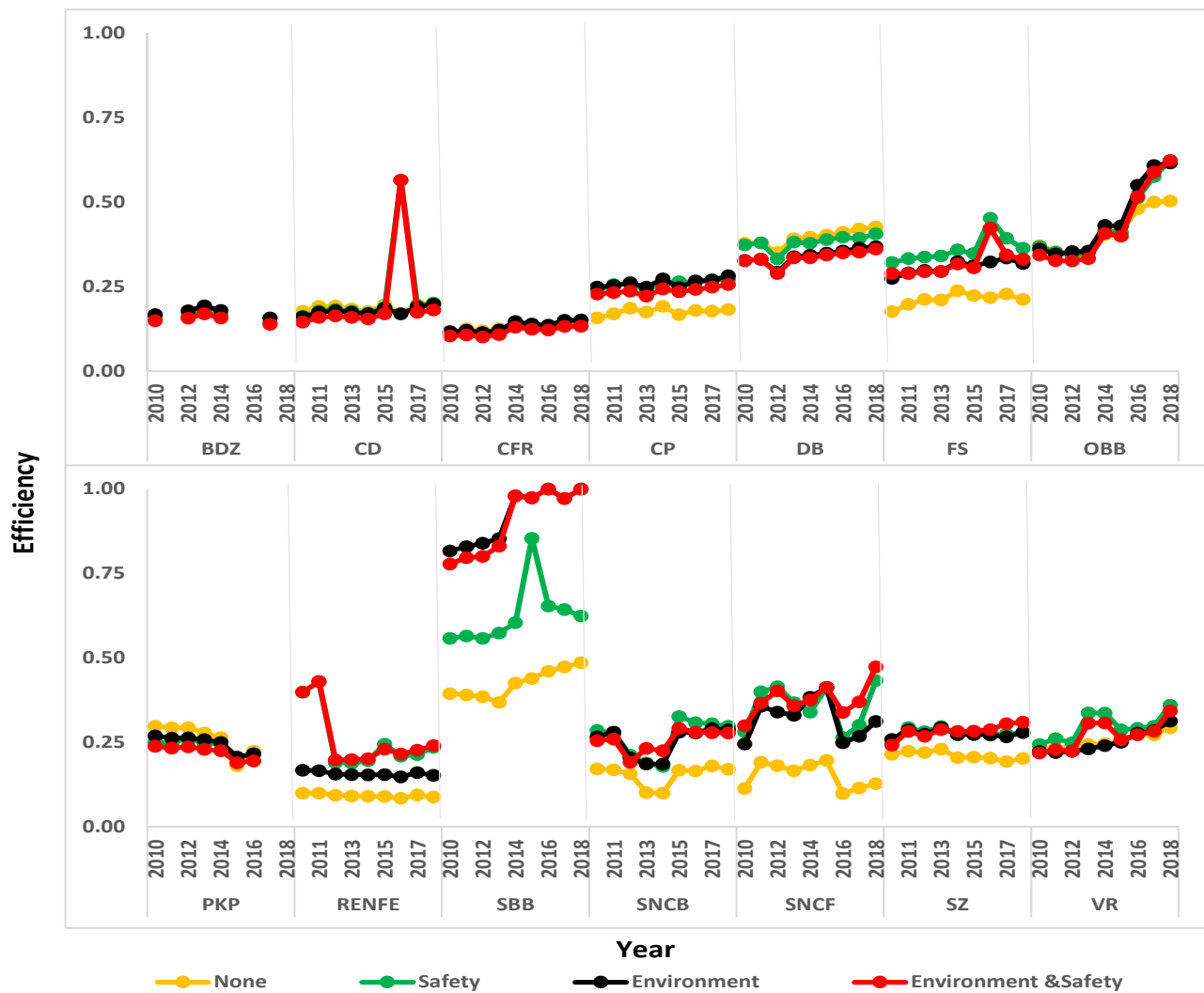


Figure 3. Efficiency Trends of 14 UIC-CER Member Railway Operators in Europe, 2010–2018.

The performance of CD significantly diverges when undesirable outputs are factored in, indicating that its operations continue to produce relatively greater environmental and safety externalities. In contrast, SBB continues to excel; but any variations in emissions or accident rates may moderately influence its integrated efficiency score due to its proximity to the frontier. These discrepancies validate that sustainability-focused efficiency evaluations reveal enhancement opportunities overlooked by traditional approaches.

To quantify changes, Table 6 reports the Euclidean distances between normalized time-series scores across the models. The Euclidean Distance quantifies the deviation of each operator's sustainability-adjusted efficiency score from its conventional score. A high score indicates that the incorporation of emissions and accident data markedly alters the assessment results, while a low value implies alignment between operational and sustainability performance. These results confirm greater dissimilarity between the unconstrained model and the combined model, confirming that the ranking of performance changes more substantially when environmental and safety factors are included in the analysis. SZ showed the highest level of deviation, with distances exceeding two across all comparisons.

Table 6. Euclidean Distance Measures Between Model Configurations.

DMU	θ_E vs. θ	θ_S vs. θ	θ_{ES} vs. θ
CD	0.037	4.377	4.956
CFR	0.423	0.107	1.275
CP	1.404	1.148	2.737
DB	0.299	0.431	0.736
FS	0.440	2.332	3.154
OBB	0.150	0.255	0.559
RENFE	0.234	1.641	1.871
SBB	2.524	0.121	1.380
SNCB	1.441	1.538	2.677
SNCF	1.755	2.870	3.336
SZ	2.973	3.116	4.829
VR	0.391	2.215	2.180

5.2.3. Cross-Model Ranking Analysis

This final part of the analysis examines how operator rankings varied across the four models. Table 7 presents Kendall's tau and Spearman's rho correlation coefficients for rank comparisons.

Table 7. Rank Correlation Statistics Between the Full and Single-Constraint Models by Year.

		2010	2011	2012	2013	2014	2015	2016	2017	2018
Kendall's Tau	θ_{ES} vs. θ_E	0.648 **	0.718 **	0.912 **	0.714 **	0.868 **	0.923 **	0.564 **	0.744 **	0.788 **
	θ_{ES} vs. θ_S	0.890 **	0.923 **	0.978 **	0.890 **	0.890 **	0.974 **	0.846 **	0.821 **	0.939 **
	θ_{ES} vs. θ	0.231	0.308	0.495 *	0.451 *	0.582 **	0.564 **	0.359	0.538 *	0.576 **
Spearman's Rho	θ_{ES} vs. θ_E	0.723 **	0.692 **	0.969 **	0.881 **	0.952 **	0.978 **	0.665 *	0.857 **	0.923 **
	θ_{ES} vs. θ_S	0.974 **	0.978 **	0.996 **	0.974 **	0.969 **	0.995 **	0.945 **	0.929 **	0.986 **
	θ_{ES} vs. θ	0.314	0.363	0.679 **	0.591 *	0.749 **	0.742 **	0.484	0.626 *	0.664 *

** Correlation is significant at the 0.01 level, * Correlation is significant at the 0.05 level.

The rank correlations between the comprehensive model and the individual constraint models were high and statistically significant at the 1% level, indicating a strong degree of consistency between them. However, the correlation between the unconstrained model and the full model was generally weaker and occasionally statistically insignificant. Correlation values between the unconstrained and full models consistently fell below 0.75, supporting the necessity of incorporating undesirable outputs to better assess railway sustainability.

Additionally, comparing the combined model with the individual environmental and safety models further emphasized the importance of joint evaluation. Kendall's tau ranged from 0.55 to 0.95, while Spearman's rho spanned 0.66 to 0.98—signifying moderate to strong but not perfect alignment. These discrepancies reinforce the notion that separate consideration of undesirable outputs may overlook critical interactions relevant to sustainability assessment.

5.3. Combined Environmental and Safety Performance

Following the comparative assessment of the four model configurations and the distinction between the results of the combined environmental- and safety-constrained efficiency model and the other three alternatives (namely, the unconstrained model, the model with only environmental constraints, and the model with only safety constraints), this subsection focuses on analyzing the performance trends captured by the proposed two-stage DEA model that integrates both environmental and safety factors.

As presented in Table 5 and illustrated in Figure 3, most railway operators exhibited low efficiency scores at the beginning of the period under the combined environmental and safety model. In the first year of the analysis, the mean efficiency score stood at 0.29, ranging from 0.104 at the lowest to 0.78 at the highest. By the end of the study period, the average score had risen modestly to 0.35, with values spanning from a minimum of 0.133 to a maximum of 1.

BDZ (Bulgaria), RENFE (Spain), and PKP (Poland) recorded a reduction in efficiency across the years. In contrast, the Swiss operator SBB consistently achieved the highest scores across the evaluation period, whereas CFR (Romania), BDZ, and CD (Czech Republic) remained among the least efficient performers. Notably, OBB (Austria), SNCF (France), and VR (Finland) recorded substantial improvements, each surpassing a 50% gain in efficiency during the period analyzed.

There are many structural and operational features that make SBB a top performer. Switzerland has one of the busiest rail networks in Europe, thanks to frequent service, efficient signaling systems, and a lot of public money spent on rail infrastructure. The operator also benefits from a mostly electrified network that gets most of its power from low-carbon sources. This means that the network has very low emissions per passenger-kilometer. SBB also has a great safety record, with very few major accidents. These characteristics work together to help SBB continuously attain high efficiency results in both operational and sustainability areas. OBB also gets good results because it has a modern fleet of rolling stock, a lot of electrification, and it keeps putting money into rail infrastructure and safety technology. Austria additionally supports modal shift by having good rail policies and connections across borders, which helps keep passenger and freight numbers high. These traits allow OBB to keep up a high level of performance, not just when looking at traditional measures of efficiency, but also when looking at accident and emission rates.

The general trend of increasing performance under the model that jointly considers both environmental and safety criteria was stronger than under the other configurations. This suggests that sustainability-oriented factors have a significant influence on the performance of railway operators. These findings underscore the relevance and necessity of incorporating environmental and safety dimensions into efficiency assessments to support the advancement of sustainable railway transport.

6. Conclusions and Implications for Policy and Research

This study introduced an innovative approach to assess the sustainability of railway transport systems by integrating both the network structure and the presence of undesirable outputs. Specifically, we implemented a novel Data Envelopment Analysis model—Variable Slack-Based Measure with intermediate measures—capable of incorporating undesirable outputs, accounting for non-proportional input–output variations and treating intermediate outputs as flexible variables. Recognizing the critical roles of environmental performance and safety in sustainable railway development, our model was applied in multiple configurations: one with no undesirable outputs, one with only environmental constraints, another with only safety constraints, and a fourth incorporating both.

The empirical investigation examined 14 European railway operators participating in the UIC-CER sustainability initiative, each providing both passenger and freight services between 2010 and 2018. Safety concerns were captured through accident-related metrics, including derailments, collisions, fatalities, and level-crossing incidents. Environmental impacts were quantified using CO₂-equivalent emissions, representing the contribution of each operator to greenhouse gas emissions based on global warming potential.

For comparative purposes, we also applied both the traditional Slack-Based Measure model and the enhanced version proposed in this research. The findings indicate that the enhanced model offered greater discriminatory power in distinguishing efficient from inefficient decision-making units. Although there was a positive correlation between the efficiency scores obtained from both models, the enhanced model provided more differentiated results. Moreover, we observed that efficiency scores tended to be higher when environmental and safety constraints were introduced, reinforcing the importance of including these dimensions in sustainability assessments. Notably, the outcomes from the combined evaluation of environmental and safety impacts were not identical to those derived from separate assessments, suggesting that a joint approach offers added analytical value.

The policy implications of this research are multifaceted. First, the proposed model serves as a valuable tool for benchmarking the sustainability performance of railway operators while reflecting the real-world complexity of their operations. Efficiency rankings can inform policymakers about how each operator is performing relative to peers, thereby guiding strategic decisions aimed at system-wide improvements. Second, by analyzing the divergence in performance under environmental versus safety constraints, decision-makers can identify specific areas in need of targeted interventions. For instance, if an operator performs well in environmental metrics but poorly in safety, this highlights a clear area for focused improvement. Third, a combined assessment of both undesirable outputs provides a comprehensive perspective for policymakers interested in promoting holistic and sustainable rail operations.

It is worth noting that the concept of sustainable transportation typically encompasses environmental, social, and economic dimensions. A limitation of this study is its exclusion of direct measures of economic impact, accessibility, and affordability. Although these aspects may be indirectly reflected through transport volume indicators, future research could benefit from incorporating explicit indicators for these dimensions to strengthen the comprehensiveness of sustainability evaluations. Additionally, the scope of this study was limited to a relatively small subset of European railway operators that participated in the UIC-CER initiative and operated both freight and passenger services. Including a larger number of operators in the sample would improve the applicability of the results and allow for stronger, more comprehensive conclusions across the European railway industry.

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

Funding: This research does not receive any external funding.

Data Availability Statement: Data were obtained from the UIC Railway Statistics (<https://uic-stats.uic.org/>) (accessed on 15 January 2025) portal and via confidential correspondence with the International Union of Railways (UIC) and, due to provider restrictions, are not publicly available; interested researchers may request access directly from UIC, and the authors can share variable lists and processing details upon reasonable request subject to UIC's approval.

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

References

1. European Commission. *The Transport and Mobility Sector*; European Commission: Brussels, Belgium, 2020; pp. 3–5.
2. Catalano, G.; Daraio, C.; Diana, M.; Gregori, M.; Matteucci, G. Efficiency, Effectiveness, and Impacts Assessment in the Rail Transport Sector: A State-of-the-Art Critical Analysis of Current Research. *Int. Trans. Oper. Res.* **2018**, *26*, 5–40. [\[CrossRef\]](#)
3. European Commission. *The Journey Begins—2021 Is the European Year of Rail!* European Commission: Brussels, Belgium, 2021; pp. 2020–2021.
4. Djordjevic, B.; Sadashiv, A.; Krmac, E. Research in Transportation Business & Management Analysis of Dependency and Importance of Key Indicators for Railway Sustainability Monitoring: A New Integrated Approach with DEA and Pearson Correlation. *Res. Transp. Bus. Manag.* **2021**, *41*, 100650.
5. Stefaniec, A.; Hosseini, K.; Xie, J.; Li, Y. Sustainability Assessment of Inland Transportation in China: A Triple Bottom Line-Based Network DEA Approach. *Transp. Res. Part D Transp. Environ.* **2020**, *80*, 102258. [\[CrossRef\]](#)
6. Mahmoudi, R.; Emrouznejad, A.; Shetab-Boushehri, S.N.; Hejazi, S.R. The Origins, Development and Future Directions of Data Envelopment Analysis Approach in Transportation Systems. *Socio-Econ. Plan. Sci.* **2020**, *69*, 100672. [\[CrossRef\]](#)
7. Wang, Q.; Wu, Z.; Chen, X. Decomposition Weights and Overall Efficiency in a Two-Stage DEA Model with Shared Resources. *Comput. Ind. Eng.* **2019**, *136*, 135–148. [\[CrossRef\]](#)
8. Zhou, H.; Yang, Y.; Chen, Y.; Zhu, J. Data Envelopment Analysis Application in Sustainability: The Origins, Development and Future Directions. *Eur. J. Oper. Res.* **2018**, *264*, 1–16. [\[CrossRef\]](#)
9. Fukuyama, H.; Yoshida, Y.; Managi, S. Modal Choice between Air and Rail: A Social Efficiency Benchmarking Analysis That Considers CO₂ Emissions. *Environ. Econ. Policy Stud.* **2011**, *13*, 89–102. [\[CrossRef\]](#)
10. Djordjevic, B.; Krmac, E.; Mlinarić, T.J. Non-Radial DEA Model: A New Approach to Evaluation of Safety at Railway Level Crossings. *Saf. Sci.* **2018**, *103*, 234–246. [\[CrossRef\]](#)
11. Kim, J.; Lee, G.; Kim, H. Analysis of Operational Efficiency Considering Safety Factors as an Undesirable Output for Coastal Ferry Operators in Korea. *J. Mar. Sci. Eng.* **2020**, *8*, 367. [\[CrossRef\]](#)
12. Evans, A.W. Accidental Fatalities in Transport. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2003**, *166*, 253–260. [\[CrossRef\]](#)
13. Jacobsen, M.R. Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior. *Am. Econ. J. Appl. Econ.* **2013**, *5*, 1–26. [\[CrossRef\]](#)
14. Chen, Y.; Li, Y.; Liang, L.; Salo, A.; Wu, H. Frontier Projection and Efficiency Decomposition in Two-Stage Processes with Slacks-Based Measures. *Eur. J. Oper. Res.* **2015**, *250*, 543–554. [\[CrossRef\]](#)
15. Kao, C. A Classification of Slacks-Based Efficiency Measures in Network Data Envelopment Analysis with an Analysis of the Properties Possessed. *Eur. J. Oper. Res.* **2018**, *270*, 1109–1121. [\[CrossRef\]](#)
16. Cook, W.D.; Zhu, J.; Bi, G.; Yang, F. Network DEA: Additive Efficiency Decomposition. *Eur. J. Oper. Res.* **2010**, *207*, 1122–1129. [\[CrossRef\]](#)
17. Tone, K. Dealing with Undesirable Outputs in DEA: A Slacks-Based Measure (SBM) Approach. *GRIPS Discuss. Pap.* **2004**, *1*, 44–45.
18. UIC; CER. *Moving Towards Sustainable Mobility*; CER: Bruxelles, Belgium, 2012.
19. Lin, W.; Chen, B.; Xie, L.; Pan, H. Estimating Energy Consumption of Transport Modes in China Using DEA. *Sustainability* **2015**, *7*, 4225–4239. [\[CrossRef\]](#)
20. EC. *A Community Strategy for ‘Sustainable Mobility’*; EU Commission: Brussels, Belgium, 1992.
21. Holden, E.; Gilpin, G.; Banister, D. Sustainable Mobility at Thirty. *Sustainability* **2019**, *11*, 1965. [\[CrossRef\]](#)
22. Zhao, X.; Ke, Y.; Zuo, J.; Xiong, W.; Wu, P. Evaluation of Sustainable Transport Research in 2000–2019. *J. Clean. Prod.* **2020**, *256*, 120404. [\[CrossRef\]](#)
23. Karjalainen, L.E.; Juhola, S. Urban Transportation Sustainability Assessments: A Systematic Review of Literature. *Transp. Rev.* **2021**, *41*, 659–684. [\[CrossRef\]](#)
24. Aloui, A.; Hamani, N.; Derrouiche, R.; Delahoche, L. Systematic Literature Review on Collaborative Sustainable Transportation: Overview, Analysis and Perspectives. *Transp. Res. Interdiscip. Perspect.* **2021**, *9*, 100291. [\[CrossRef\]](#)
25. Kraus, L.; Proff, H. Sustainable Urban Transportation Criteria and Measurement—A Systematic Literature Review. *Sustainability* **2021**, *13*, 7113. [\[CrossRef\]](#)
26. Bakker, S.; Contreras, K.D.; Kappiantari, M.; Tuan, N.A.; Guillen, M.D.; Gunthawong, G.; Zuidgeest, M.; Liefverink, D.; van Maarseveen, M. Low-Carbon Transport Policy in Four ASEAN Countries: Developments in Indonesia, the Philippines, Thailand and Vietnam. *Sustainability* **2017**, *9*, 1217. [\[CrossRef\]](#)
27. Starr McMullen, B.; Noh, D.W. Accounting for Emissions in the Measurement of Transit Agency Efficiency: A Directional Distance Function Approach. *Transp. Res. Part D Transp. Environ.* **2007**, *12*, 1–9. [\[CrossRef\]](#)
28. Oum, T.H.; Pathomsiri, S.; Yoshida, Y. Limitations of DEA-Based Approach and Alternative Methods in the Measurement and Comparison of Social Efficiency across Firms in Different Transport Modes: An Empirical Study in Japan. *Transp. Res. Part E Logist. Transp. Rev.* **2013**, *57*, 16–26. [\[CrossRef\]](#)

29. Chang, Y.T.; Park, H.S.; Jeong, J.B.; Lee, J.W. Evaluating Economic and Environmental Efficiency of Global Airlines: A SBM-DEA Approach. *Transp. Res. Part D Transp. Environ.* **2014**, *27*, 46–50. [\[CrossRef\]](#)
30. Noroozzadeh, A.; Sadjadi, S.J. A New Approach to Evaluate Railways Efficiency Considering Safety Measures. *Decis. Sci. Lett.* **2013**, *2*, 71–80. [\[CrossRef\]](#)
31. Zhang, Q.; Zhuang, Y.; Wei, Y.; Jiang, H.; Yang, H. Railway Safety Risk Assessment and Control Optimization Method Based on FTA-FPN: A Case Study of Chinese High-Speed Railway Station. *J. Adv. Transp.* **2020**, *2020*, 3158468. [\[CrossRef\]](#)
32. Yu, M.-M.; Fan, C.-K. Measuring the Cost Effectiveness of Multimode Bus Transit in the Presence of Accident Risks. *Transp. Plan. Technol.* **2006**, *29*, 383–407. [\[CrossRef\]](#)
33. Castillo, H.; Pitfield, D.E. ELASTIC—A Methodological Framework for Identifying and Selecting Sustainable Transport Indicators. *Transp. Res. Part D Transp. Environ.* **2010**, *15*, 179–188. [\[CrossRef\]](#)
34. Miller, P.; de Barros, A.G.; Kattan, L.; Wirasinghe, S.C. Analyzing the Sustainability Performance of Public Transit. *Transp. Res. Part D Transp. Environ.* **2016**, *44*, 177–198. [\[CrossRef\]](#)
35. Sdoukopoulos, A.; Pitsiava-Latinopoulou, M.; Basbas, S.; Papaioannou, P. Measuring Progress towards Transport Sustainability through Indicators: Analysis and Metrics of the Main Indicator Initiatives. *Transp. Res. Part D Transp. Environ.* **2019**, *67*, 316–333. [\[CrossRef\]](#)
36. ITF. *Transport Infrastructure Investment*; ITF: London, UK, 2008. [\[CrossRef\]](#)
37. Choi, N.H.; Jung, K. Measuring Efficiency and Effectiveness of Highway Management in Sustainability. *Sustainability* **2017**, *9*, 1347. [\[CrossRef\]](#)
38. Daraio, C.; Diana, M.; Di Costa, F.; Leporelli, C.; Matteucci, G.; Nastasi, A. Efficiency and Effectiveness in the Urban Public Transport Sector: A Critical Review with Directions for Future Research. *Eur. J. Oper. Res.* **2016**, *248*, 1–20. [\[CrossRef\]](#)
39. Aigner, D.; Lovell, C.A.K.; Schmidt, P. Formulation and Estimation of Stochastic Frontier Production Function Models. *J. Econom.* **1977**, *6*, 21–37. [\[CrossRef\]](#)
40. Farrell, M.J. The Measurement of Productive Efficiency. *J. R. Stat. Soc. A* **1957**, *120*, 253–290. [\[CrossRef\]](#)
41. Costa, Á.; Markellos, R.N. Evaluating Public Transport Efficiency with Neural Network Models. *Transp. Res. Part C Emerg. Technol.* **1997**, *5*, 301–312. [\[CrossRef\]](#)
42. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the Efficiency of Decision Making Units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [\[CrossRef\]](#)
43. Seiford, L.M.; Thrall, R.M. Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis. *J. Econom.* **1990**, *46*, 7–38. [\[CrossRef\]](#)
44. Cooper, W.W.; Seiford, L.; Tone, K. *Introduction to Data Envelopment Analysis and Its Uses*; Springer: Berlin/Heidelberg, Germany, 2006.
45. Jordá, P.; Cascajo, R.; Monzón, A. Analysis of the Technical Efficiency of Urban Bus Services in Spain Based on SBM Models. *ISRN Civ. Eng.* **2012**, *2012*, 984758. [\[CrossRef\]](#)
46. Easton, L.; Murphy, D.J.; Pearson, J.N. Purchasing Performance Evaluation: With Data Envelopment Analysis. *Eur. J. Purch. Supply Manag.* **2002**, *8*, 123–134. [\[CrossRef\]](#)
47. Smith, P.C.; Goddard, M. Performance Management and Operational Research: A Marriage Made in Heaven? *J. Oper. Res. Soc.* **2002**, *53*, 247–255. [\[CrossRef\]](#)
48. Chen, L.; Lai, F.; Wang, Y.M.; Huang, Y.; Wu, F.M. A Two-Stage Network Data Envelopment Analysis Approach for Measuring and Decomposing Environmental Efficiency. *Comput. Ind. Eng.* **2018**, *119*, 388–403. [\[CrossRef\]](#)
49. Saeedi, H.; Behdani, B.; Wiegman, B.; Zuidwijk, R. Assessing the Technical Efficiency of Intermodal Freight Transport Chains Using a Modified Network DEA Approach. *Transp. Res. Part E Logist. Transp. Rev.* **2019**, *126*, 66–86. [\[CrossRef\]](#)
50. Marchetti, D.; Wanke, P. Brazil's Rail Freight Transport: Efficiency Analysis Using Two-Stage DEA and Cluster-Driven Public Policies. *Socio-Econ. Plan. Sci.* **2017**, *59*, 26–42. [\[CrossRef\]](#)
51. Seiford, L.M.; Zhu, J. Modeling Undesirable Factors in Efficiency Evaluation. *Eur. J. Oper. Res.* **2002**, *142*, 16–20. [\[CrossRef\]](#)
52. Berg, S.A.; Førsund, F.R.; Jansen, E.S. Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980–1989. *Scand. J. Econ.* **1992**, *94*, S211–S228. [\[CrossRef\]](#)
53. Lovell, C.; Pastor, J.; Turner, J.A. Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries. *Eur. J. Oper. Res.* **1995**, *87*, 507–518. [\[CrossRef\]](#)
54. Tyteca, D. Linear Programming Models for the Measurement of Environmental Performance of Firms—Concepts and Empirical Results. *J. Product. Anal.* **1997**, *8*, 183–197. [\[CrossRef\]](#)
55. Tone, K. A Slacks-Based Measure of Efficiency in Data Envelopment Analysis. *Eur. J. Oper. Res.* **2001**, *130*, 498–509. [\[CrossRef\]](#)
56. Färe, R.; Grosskopf, S. Directional Distance Functions and Slacks-Based Measures of Efficiency: Some Clarifications. *Eur. J. Oper. Res.* **2010**, *206*, 702. [\[CrossRef\]](#)
57. Picazo-Tadeo, A.J.; Reig-Martínez, E.; Hernández-Sancho, F. Directional Distance Functions and Environmental Regulation. *Resour. Energy Econ.* **2005**, *27*, 131–142. [\[CrossRef\]](#)

58. Kang, C.-C.; Khan, H.A.; Feng, C.-M.; Wu, C.-C. Efficiency Evaluation of Bus Transit Firms with and without Consideration of Environmental Air-Pollution Emissions. *Transp. Res. Part D Transp. Environ.* **2017**, *50*, 505–519. [[CrossRef](#)]
59. Park, Y.S.; Lim, S.H.; Egilmez, G.; Szmerekovsky, J. Environmental Efficiency Assessment of U.S. Transport Sector: A Slack-Based Data Envelopment Analysis Approach. *Transp. Res. Part D Transp. Environ.* **2018**, *61*, 152–164. [[CrossRef](#)]
60. Wang, D.D. Assessing Road Transport Sustainability by Combining Environmental Impacts and Safety Concerns. *Transp. Res. Part D Transp. Environ.* **2019**, *77*, 212–223. [[CrossRef](#)]
61. Lozano, S.; Gutiérrez, E. A Slacks-Based Network DEA Efficiency Analysis of European Airlines. *Transp. Plan. Technol.* **2014**, *37*, 623–637. [[CrossRef](#)]
62. Tone, K.; Tsutsui, M. Network DEA: A Slacks-Based Measure Approach. *Eur. J. Oper. Res.* **2009**, *197*, 243–252. [[CrossRef](#)]
63. Chen, Y.; Cook, W.D.; Kao, C.; Zhu, J. Network DEA Pitfalls: Divisional Efficiency and Frontier Projection under General Network Structures. *Eur. J. Oper. Res.* **2013**, *226*, 507–515. [[CrossRef](#)]
64. Tavassoli, M.; Faramarzi, G.R.; Saen, R.F. A Joint Measurement of Efficiency and Effectiveness Using Network Data Envelopment Analysis Approach in the Presence of Shared Input. *Opsearch* **2015**, *52*, 490–504. [[CrossRef](#)]
65. Yu, M.M. Assessing the Technical Efficiency, Service Effectiveness, and Technical Effectiveness of the World's Railways through NDEA Analysis. *Transp. Res. Part A Policy Pract.* **2008**, *42*, 1283–1294. [[CrossRef](#)]
66. Yu, M.M.; Lin, E.T.J. Efficiency and Effectiveness in Railway Performance Using a Multi-Activity Network DEA Model. *Omega* **2008**, *36*, 1005–1017. [[CrossRef](#)]
67. Zhou, H.; Hu, H. Sustainability Evaluation of Railways in China Using a Two-Stage Network DEA Model with Undesirable Outputs and Shared Resources. *Sustainability* **2017**, *9*, 150. [[CrossRef](#)]
68. Charnes, A.; Cooper, W.W. Programming with Linear Fractional Functionals. *Nav. Res. Logist. Q.* **1962**, *9*, 181–186. [[CrossRef](#)]
69. Benga, A.; Delgado Rodríguez, M.J.; de Lucas Santos, S.; Zeneli, G. Efficiency Analysis of European Railway Companies and the Effect of Demand Reduction. *J. Rail Transp. Plan. Manag.* **2025**, *34*, 100516. [[CrossRef](#)]
70. Benga, A.; Zeneli (Foto), G.; Delgado-Rodríguez, M.J.; de Lucas Santos, S. *Company Efforts and Environmental Efficiency: Evidence from European Railways Considering Market-Based Emissions*; Springer: Dordrecht, The Netherlands, 2023. [[CrossRef](#)]
71. Charles, V.; Aparicio, J.; Zhu, J. The Curse of Dimensionality of Decision-Making Units: A Simple Approach to Increase the Discriminatory Power of Data Envelopment Analysis. *Eur. J. Oper. Res.* **2019**, *279*, 929–940. [[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.