Demiryolu Mühendisliği **Railway Engineering**
Temmuz 2024 July 2024
Sayı:20, Sayfa: 11-24 Issue:20, Page: 11-24
Araştırma Makalesi Research Article
doi: 10.47072/demiryolu.1407420
http://dergipark.org.tr/demiryolu
e-ISSN: 2687-2463, ISSN: 2149-1607

Evaluating the Performance of Railway Transportation Companies Using Multi-Criteria Decision-Making Methods

Çağdaş YÜKSEL*, Nuşin UNCU

Adana Alparslan Türkeş Science and Technology University, Faculty of Engineering, Industrial Engineering Department, Adana, Türkiye

**cagdasyuksel@tcdd.gov.tr*

(Alınış/Received: 20.12.2023, Kabul/Accepted: 29.02.2024, Yayımlama/Published: 31.07.2024)

Abstract: The purpose of performance evaluation is to generate measurable data on an organization's performance, with the goal of assisting managerial decision-making and enhancing overall performance. In this study, the key performance indicators (KPIs) for railway transportation companies are identified based on expert opinions and previous frameworks. The operational performance of various railway freight transport companies was evaluated using multi-criteria decision-making methods (MCDM). Among the MCDM approaches, the Evaluation Based on Distance from Average Solution (EDAS) method was applied as the main method. In addition to the EDAS method, alternative MCDM methods such as TOPSIS, PROMETHEE II, and COPRAS were used to highlight potential deviations when compared to the results obtained with the EDAS method. Based on the research findings, three out of the seven KPIs, namely safety, have the highest weight at 38%, followed by punctuality at 19%, and journey time at 12%. Subsequently, companies were ranked according to their performance based on all KPIs. Furthermore, a sensitivity analysis was conducted to demonstrate how changes in the relative weights of KPIs can affect the results.

Keywords: Key performance indicators, Multi-criteria decision making, Performance evaluation, Railway freight transport

Demiryolu Taşımacılığı Firmalarının Performanslarının Çok Kriterli Karar Verme Yöntemleri ile Değerlendirilmesi

Öz: Performans değerlendirmenin amacı, bir organizasyonun performansıyla ilgili ölçülebilir veriler üretmek, yönetimsel karar alma sürecine destek olmak ve genel performansı artırmaktır. Bu çalışmada, demiryolu taşımacılığı şirketleri için anahtar performans göstergeleri (APG'ler), uzman görüşleri ve önceki çerçevelere dayalı olarak belirlenmektedir. Çeşitli demiryolu yük taşıma şirketlerinin işletme performansı, çoklu kriterli karar verme yöntemleri (ÇKKV) kullanılarak değerlendirilmektedir. ÇKKV yaklaşımlarından biri olan Ortalama Çözüm Uzaklığına Dayalı Değerlendirme (EDAS) yöntemi, ana yöntem olarak uygulanmıştır. EDAS yöntemiyle elde edilen sonuçlarla karşılaştırıldığında potansiyel sapmaları göstermek için TOPSIS, PROMETHEE II ve COPRAS gibi alternatif ÇKKV yöntemleri de kullanılmıştır. Araştırma bulgularına göre, yedi APG'den üçü, yani emniyet %38 ile en yüksek ağırlığa sahipken, bunu %19 ile dakiklik ve %12 ile seyir süresi takip etmektedir. Daha sonra şirketler tüm KPI'lar baz alınarak performanslarına göre sıralanmıştır. Ayrıca, APG'lerin göreceli ağırlıklarındaki değişikliklerin sonuçları nasıl etkileyebileceğini göstermek için duyarlılık analizi yapılmıştır.

Anahtar kelimeler: Anahtar Performans Göstergeleri (APGler), Çok kriterli karar verme, Performans değerlendirme, Demiryolu yük taşımacılığı

1. Introduction

As an effect of transport networks, railways play a vital role in fostering economic growth, reducing road congestion, and limiting climate change [1]. In parallel with the advancement of global trade, railways have gained paramount importance not only within national borders but also between countries and even continents in the transportation of products and goods, particularly in terms of bulk and raw material transportation. This importance arises from their

Atıf için/Cite as: Ç. Yüksel, N. Uncu, "Evaluating the performance of railway transportation companies using Multi-Criteria Decision-Making Methods," *Demiryolu Mühendisliği*, sy. 20, ss. 11-24, Temmuz 2024. doi: 10.47072/demiryolu.1407420

attributes of safety, longevity, cost-effectiveness in construction, environmental friendliness, and independence from petroleum reliance. Social and economic growth has led to an increasing demand for more railway lines and services. Investments in the railway sector in Türkiye have increased in the last 19 years. Although Türkiye is located between the continents of Asia and Europe, it is not possible to say that it can sufficiently benefit from the geographical advantages created by its inter-regional position in terms of railway networks. It is aimed that Türkiye will reach the desired position with the completion of the ongoing investments and future investments [2].

Railways worldwide are facing increasing demand and limited infrastructure. As per the 2019 OECD report, under the baseline scenario, the annual investment in rail infrastructure is anticipated to rise to USD 315 billion (United States dollars) by the year 2050. This projection is established upon projects currently undergoing diverse stages of construction and planning [3]. Optimizing the use of the existing railway network and services is therefore becoming a central responsibility for all railway stakeholders (such as infrastructure managers, train operators, regulatory agencies, etc.), whether through upgrading infrastructure or developing improved management strategies to provide better services. Considering all of these, measuring railway performances is important for determining where and how to initiate these management efforts. In the context of railways, performance indicates the capacity of a railway system to fulfill its claims in terms of passenger and freight transportation [4]. UIC (International Union of Railways) [5] defines the performance of railways in terms of capacity in market conditions, infrastructure planning, program planning and capacity in operations.

Key Performance Indicators (KPIs) measure the critical aspects vital to the success of a company [6]. The KPI concept will be used to assess the performance of railway freight transportation and train operating companies (TOCs). A two-year project, known as IMPROVERAIL, was initiated by the European Commission in 2001 and aimed at supporting the establishment of railway infrastructure management. As a result of the project, seven indicators were developed, which can be considered as KPIs, and are called “success dimensions”. These are efficiency, accessibility, financial effectiveness, asset utilization, innovation & growth, service quality & reliability, and safety [7]. Platform of Railway Infrastructure Managers in Europe (PRIME) is a forum that aims to enhance rail infrastructure managers' collaboration, support the implementation of European rail policy, and develop performance benchmarking [8]. In 2019, PRIME [9] identified twelve KPIs in five dimensions: safety, environment, performance, delivery, and financial. These KPIs reflect the needs and priorities of railway infrastructure managers and respond to the demands of stakeholders and customers.

In 2013, Law No. 6461 aimed to enhance transparency, cost-efficiency, neutrality, and competition in Türkiye's railway sector. As part of this law, the management of railway infrastructure and train operations were separated. The law has opened the Turkish rail transport sector to competition, and this means that TCDD Taşımacılık has to compete with the private train operating companies that would enter the sector [10]. Thus, in this competitive environment, performance has become important for these companies.

In 2017, TCDD was authorized as the railway infrastructure manager (IM) on the railway infrastructure and state-owned parts of the railway infrastructure in the national railway infrastructure network, and a new company, TCDD Taşımacılık was established as a train operating company (TOC) for passenger and freight transportation. The law opened the Turkish rail transport sector to competition, necessitating that TCDD Taşımacılık competes with newly-entered private train operating companies within the sector. Today, there are five certificated TOCs in the field of railway transportation. One of them has not started transportation activities yet [11].

In this study, the performance evaluation of TOCs is conducted using EDAS method, a relatively new method that facilitates the comparison and ranking of companies based on certain KPIs. The primary benefit of the EDAS method lies in its high efficiency and lower computational requirements when compared to alternative MCDM methods [12]. This method is highly effective when conflicting criteria exist [13]. The findings are reinforced through the application of various types of MCDM methods. This includes TOPSIS which is a distance-based approach similar to EDAS, PROMETHEE II, an outranking type and COPRAS, a utility-based approach.

The rest of the article is structured as follows: Section 2 provides a literature review. Section 3 explains the methodology of applying MCDM methods and related information on the performance KPIs of TOCs. Section 4 presents the results, and the findings are discussed in Section 5.

2. Literature Review

MCDM methods have been commonly applied for deciding between alternatives in the context of railway performance evaluation. Several case studies have been conducted to assess the performance of railway networks in various countries. Lu [4] proposed a framework for measuring British railway performance based on quality-of-service factors. Key performance indicators (KPIs) include accommodation, journey time, connectivity, punctuality, resilience, passenger comfort, energy usage, and resource usage. The Analytic Hierarchy Process (AHP) was employed to determine the order of importance of these KPIs. Additionally, the significant effects of quality-of-service factors on the KPIs were identified using Taguchi L27 design. The data were obtained from simulation runs conducted through the BRaVE program (Birmingham Railway Virtual Environment).

Frederico and Cavenaghi [14] developed a performance measurement system framework for railroad companies implementing within a major Brazilian railway company. They used semi-structured interviews and documentary data collection to obtain evidence on processes, environment, customers, and strategic objectives as perspectives needed to elaborate a performance measurement system based on balanced scorecards. Zhang et al. [15] assessed the China's railway transportation performance using the CRITIC (criteria importance through the intercriteria correlation method)-relative entropy method. The main criteria are safety in production, railway infrastructure, railway equipment, operation efficiency and green development.

Ranjan et al. [16] integrated MCDM methods DEMATEL and Vise Kriterijumska Optimizacija I Kompromisno Resenje (VIKOR) for evaluating the performance of Indian railway zones. Jose et al. [17] utilized the Hierarchical Fuzzy Axiomatic Design (H-FAD) method to assess the performance of sixteen zones within the Indian Railways. Petrovic et al. [18] assessed the performance of the Serbian railway system by employing the Entropy weight method to determine criteria weights. They utilized TOPSIS for ranking the performance across different years.

Other studies in the literature have been conducted to compare the railway performances of different countries or regions. Fraszczyk et al. [19] compared five European countries' railway systems on various passenger related parameters and revealed their performance levels compared to the European average using data analysis. Stoilova et al. [20] presented twenty-two infrastructural, economic, and technological criteria for comparing the performance of Trans-European Transport Network (TEN-T) Orient East Med Corridor railway network using SIMUS (Sequential Interactive Model for Urban Systems) method as an MCDM tool.

Stoilova [21] focused on evaluating the developmental levels of railway transport in various railway networks across the Balkan region. This assessment utilized the Shannon Entropy method and the Stepwise Weight Assessment Ratio Analysis (SWARA) method for criteria weights. Additionally, several MCDM methods, such as VIKOR, Weighted Aggregated Sum Product Assessment (WASPAS), and PROMETHEE, were applied to rank the railway performances of selected Balkan countries. Kara and Yalçın [22] utilized the CRITIC technique to determine criteria weights and the Range of Value (ROV) technique for ranking the performances of twenty-three European countries' railway networks. Bouraima et al. [23] revealed the importance of improvement suggestions as alternatives on the Sub-Saharan African railways performance based on six criteria using the weighted geometric Dombi Maclaurin Symmetric Mean (WGDMSM) operator within an intuitionistic fuzzy environment and combining the method with an interval rough PIvot Pairwise RElative Criteria Importance Assessment (PIPRECIA) operator.

In the literature, although there are few studies outside of MCDM related to railway performance evaluation, they have been encountered. Sharma et al. [24] focused on measuring the performance of Indian Railways in terms of service efficiency. The study has an operational perspective using data envelopment analysis (DEA) methodology. The researchers incorporated quality-of-service dimensions in the performance measurement and benchmarking of the railway zones that constitute a national railway system under the public domain.

Through the literature review, it becomes evident that while numerous studies exist for the general evaluation of railway transportation systems, there is a scarcity of research focusing on the specific level of TOCs. Within the scope of the literature we reviewed, this study contributes by addressing TOCs' performance KPIs and serving as a case-specific investigation to evaluate the performance of railway transportation in Türkiye.

3. Methodology

In this section, we present the identification of KPIs for assessing railway performance. We provide insights into their components and offer a detailed account of the application of MCDM methods, particularly focusing on Buckley's Fuzzy AHP and EDAS.

3.1. Determination of KIPs

As a result of the focus group meeting with eight experts, seven unique KPIs were identified based on the reports of the railway transportation authorities, previous studies in the literature and the structure of the Turkish railway transportation sector [5, 6, 7, 9]. The qualifications of the experts are detailed in Table 1. Each of the experts has substantial knowledge about the railway industry and its operation and can identify the requirements of different stakeholders.

Table 1. The qualifications of experts

Expert	Title	Specialty	Experience
Expert 1	Department manager	Railway capacity management	36 years
Expert 2	Department assistant manager	Railway capacity management	35 years
Expert 3	Department manager	Railway traffic management	25 years
Expert 4	Department assistant manager	Railway traffic management	30 years
Expert 5	Department assistant manager	Railway traffic management	28 years
Expert 6	Department assistant manager	Railway traffic management	28 years
Expert 7	Department manager	Railway infrastructure maintenance	25 years
Expert 8	Department manager	Railway infrastructure modernization	24 years

The KIPs are determined as punctuality (C1), journey time (C2), resource usage (C3), profitability (C4), safety (C5), innovation & growth (C6) and environment (C7).

The punctuality of the trains is directly related to how well they operate according to the timetable. Commonly, punctuality is calculated by dividing the number of punctual trains in a specified period by the total number of trains and expressed as a percentage of punctual trains [25]. In this study, the ratio of the number of trains departing on scheduled time and prior to scheduled time divided by the total number of trains of a TOC is used as performance indicator of punctuality.

The journey time is the ratio of the total actual journey time to the total scheduled journey time within a specified time period.

Resource usage is determined as the ratio of the number of trains operated divided by the number of trains scheduled on the timetable in a given time period.

Profitability arises from a company's operational outcomes. Unlike profit, which is an absolute indicator, profitability is a relative measure that indicates the degree of a company's profit margins.

Safety is one of the most important and essential elements in the performance of a TOC. The number of railway accidents caused by TOC-related issues per million train kilometers is used as a measure of safety KPI.

Innovation & growth for a railway freight TOC involve engaging in innovative activities to enhance digitalization in operational processes, improve operational safety, optimize rolling stock maintenance operations, and modernize rolling stocks, among other initiatives.

The environment KPI, representing environmental friendliness, is measured by the ratio of total hybrid and electric locomotives to the overall locomotives in a TOC's rolling stock.

All KPIs are beneficial criteria, except for safety. As the safety ratio decreases, the performance increases; hence safety stands as a non-beneficial criterion.

3.2. Determining the weights of KPIs: Buckley's FAHP method

AHP, developed by Thomas L. Saaty, is a method for evaluating criteria within a hierarchical framework in MCDM [26]. When there is no available quantitative data about criteria, AHP is the one of the weighting method to be used [27]. On the other hand, fuzzy AHP (FAHP) is a synthetic approach developed as an extension of AHP, which emerges when it is considered that there is a certain fuzziness in the comments and evaluations of decision makers [28]. For these reasons, in this study the FAHP method was used to determine KPI weights.

Buckley [29] extended the pure AHP method by using fuzzy triangular numbers in comparison ratios. The steps of Buckley's fuzzy AHP approach are as follows [30].

In the first step, a fuzzy comparison matrix is created. If there is more than one decision maker, the preferences of each decision maker are combined by geometric mean method.

In the second step, the geometric mean of the fuzzy comparison values of each criterion is calculated by means of Equation 1. In the formula, $\tilde{r}_i$ is still a fuzzy number.

$$\tilde{r}_i = \left(\prod_{j=1}^{n} \tilde{d}_{ij} \right)^{1/n} \quad i = 1, 2, \dots, n \tag{1}$$

In the third step, by Equation 2, the triangular fuzzy weight of each criterion is calculated.

$$\widetilde{w}_i = \tilde{r}_i * (\tilde{r}_1 + \tilde{r}_2 + \dots + \tilde{r}_n)^{-1} = (lw_i, mw_i, uw_i) \tag{2}$$

In the fourth step, the obtained triangular fuzzy weights are defuzzified using the center of area method by Equation 3.

$$M_i = \frac{lw_i + mw_i + uw_i}{3} \tag{3}$$

In the last step, the obtained crisp values are normalized using Equation 4. The N_i values here are the relative weights of each criterion.

$$N_i = \frac{M_i}{\sum_{i=1}^{n} M_i} \tag{4}$$

3.3. Evaluation of performance: EDAS method

EDAS is a distance-based approach that uses positive and negative distances from the average solution [31]. In this method, the desirable alternative is related to distance from the average solution. There is no need to calculate an ideal solution in the EDAS method [13]. The application steps of the EDAS method are as follows [32]:

In the first step, a decision matrix with *n* alternative and *m* criteria, $D = \left[x_{ij}\right]_{n*m}$ is formed. Where x_{ij} is the performance value of i^{th} alternative on the j^{th} criterion.

In the second step, the average solution (*AV*) is calculated for each criterion by Equation 5.

$$AV_j = \frac{\sum_{i=1}^{n} x_{ij}}{n} \tag{5}$$

In the third step, the positive distance from the average values (*PDA*) and the negative distance from the average values (*NDA*) are calculated by following Equations 6 and 7.

If j^{th} criterion is beneficial,

$$PDA_{ij} = \frac{\max\left(0, \left(x_{ij} - AV_j\right)\right)}{AV_j} \tag{6}$$

$$NDA_{ij} = \frac{\max\left(0, \left(AV_j - x_{ij}\right)\right)}{AV_j} \tag{7}$$

If j^{th} criterion is non-beneficial, Equation 8 and 9 are used.

$$PDA_{ij} = \frac{\max\left(0, \left(AV_j - x_{ij}\right)\right)}{AV_j} \tag{8}$$

$$NDA_{ij} = \frac{\max\left(0, \left(x_{ij} - AV_j\right)\right)}{AV_j} \quad (9)$$

In the fourth step, the weighted sum of the positive distance from the average solution values (SP_i) and the weighted sum of the negative distance from the average solution values (SN_i) for each alternative are calculated by Equations 10 and 11, respectively, where w_j is the weight of j^{th}criterion.

$$SP_i = \sum_{j=1}^{m} w_j * PDA_{ij} \quad (10)$$

$$SN_i = \sum_{j=1}^{m} w_j * NDA_{ij} \quad (11)$$

In the fifth step, SP_i and SN_i values are normalized for each alternative using Equations 12 and 13.

$$NSP_i = \frac{SP_i}{\max SP_i} \quad (12)$$

$$NSN_i = 1 - \frac{SN_i}{\max SN_i} \quad (13)$$

In the final step, the appraisal score (AS) is calculated for each alternative using Equation 14, where $0 \leq AS_i \leq 1$.

$$AS_i = \frac{1}{2} * (NSP_i + NSN_i) \quad (14)$$

3.4. Data preparation

We used operational data for 2021 from the TCDD Enterprise Resource Planning system to derive performance metrics for punctuality, journey time, resource usage, safety, and environment KPIs for the TOCs. However, due to their strategic significance to the companies, we could not directly obtain performance values for 'innovation and growth' and 'profitability' KPIs. Therefore, we employed the AHP technique to assess the performance values of these KPIs'. We conducted face-to-face interviews with the same group of experts during both the focus group process and the KPI weights determination process. Table 2 below shows all the performance metrics derived for the KPIs of the TOCs.

Table 2. KPIs performance values of TOCs

TOCs	Punctuality	Journey time	Resource usage	Profitability	Safety	Innovation & growth	Environment
TOC 1	0.4602	0.8688	0.7462	0,1314	2,9949	0.1829	0,2083
TOC 2	0.1511	0.6743	0.9059	0,4775	4,2538	0.4261	1,0000
TOC 3	0.6472	0.9265	0.7985	0,3911	1,2711	0.3910	0,5833

4. Results

4.1. Application of Buckley's FAHP method

Experts were asked to evaluate KPIs through the designed questionnaire; their linguistic expressions were converted into triangular fuzzy values using the scale developed by Zaki Mohamed Noor et al. [33]. The evaluations of KPIs by eight experts were combined into a single comparison matrix using the geometric mean method. This aggregated comparison matrix is presented in Table 3.

Table 3. The aggregated comparison matrix

	C1	C2	C3	C4	C5	C6	C7
	(l, m, u)	(l, m, u)	(l, m, u)	(l, m, u)	(l, m, u)	(l, m, u)	(l, m, u)
C1	(1, 1, 1)	(1.77, 2.24, 2.63)	(1.83, 2.3, 2.87)	(2.77, 3.69, 4.49)	(0.21, 0.25, 0.32)	(2.1, 2.8, 3.34)	(2.29, 2.95, 3.46)
C2	(0.38, 0.45, 0.57)	(1, 1, 1)	(0.87, 1.22, 1.51)	(1.77, 2.41, 3.02)	(0.24, 0.27, 0.37)	(1.49, 2.18, 2.77)	(1.09, 1.33, 1.57)
C3	(0.35, 0.43, 0.68)	(0.66, 0.82, 1.15)	(1, 1, 1)	(1.62, 2.13, 2.54)	(0.26, 0.3, 0.35)	(1.54, 1.87, 2.14)	(1.3, 1.51, 1.68)
C4	(0.22, 0.27, 0.36)	(0.33, 0.42, 0.57)	(0.39, 0.47, 0.62)	(1, 1, 1)	(0.17, 0.19, 0.25)	(0.65, 0.89, 1.11)	(0.73, 0.82, 0.95)
C5	(3.13, 4, 4.73)	(2.73, 3.69, 4.12)	(2.83, 3.38, 3.83)	(4.06, 5.27, 5.91)	(1, 1, 1)	(3.92, 5.11, 6.13)	(3.59, 4.12, 4.49)
C6	(0.3, 0.36, 0.48)	(0.36, 0.46, 0.67)	(0.47, 0.53, 0.65)	(0.9, 1.13, 1.54)	(0.16, 0.2, 0.26)	(1, 1, 1)	(0.87, 1, 1.15)
C7	(0.29, 0.34, 0.44)	(0.64, 0.75, 0.92)	(0.59, 0.66, 0.77)	(1.05, 1.21, 1.36)	(0.22, 0.24, 0.28)	(0.87, 1, 1.15)	(1, 1, 1)

For each KPI, the geometric mean ($\tilde{r}_i$) of the fuzzy comparison values was calculated by Equation 1. Triangular fuzzy weights ($\widetilde{w}_i$) of each KPI were calculated by Equation 2. Then, the triangular fuzzy weights were defuzzified by the center of area method using Equation 3 and the weights in crisp values (M_i) were obtained. These values are normalized using Equation 4 and N_i values, which were the relative weights of each KPI, were obtained. The results are given in Table 4 below.

Table 4. Geometric means ($\tilde{r}_i$), triangular fuzzy weights ($\widetilde{w}_i$), crisp weights (M_i) and normalized crisp weights (relative weights, *Ni*)

Criteria	$\tilde{r}_i$	$\widetilde{w}_i$	M_i	N_i
	(l, m, u)	(l, m, u)		
C1	(1.3726, 1.6886, 1.9944)	(0.1337, 0.1925, 0.2742)	0.2001	0.1924
C2	(0.8110, 1.0044, 1.2230)	(0.0790, 0.1145, 0.1682)	0.1205	0.1159
C3	(0.7919, 0.9373, 1.1413)	(0.0771, 0.1069, 0.1569)	0.1136	0.1093
C4	(0.4207, 0.4955, 0.6132)	(0.0410, 0.0565, 0.0843)	0.0606	0.0583
C5	(2.8087, 3.4250, 3.8327)	(0.2735, 0.3905, 0.5270)	0.3970	0.3817
C6	(0.4871, 0.5693, 0.7129)	(0.0474, 0.0649, 0.0980)	0.0701	0.0674
C7	(0.5811, 0.6511, 0.7507)	(0.0566, 0.0742, 0.1032)	0.0780	0.0750

Within the specified framework, it has been revealed that the safety KPI is the criterion with the highest relative weight in terms of the performance of a freight TOC, while the profitability KPI stands as having the least relative weight. The relative weights of KPIs obtained with the fuzzy AHP method are shown in Figure 1.

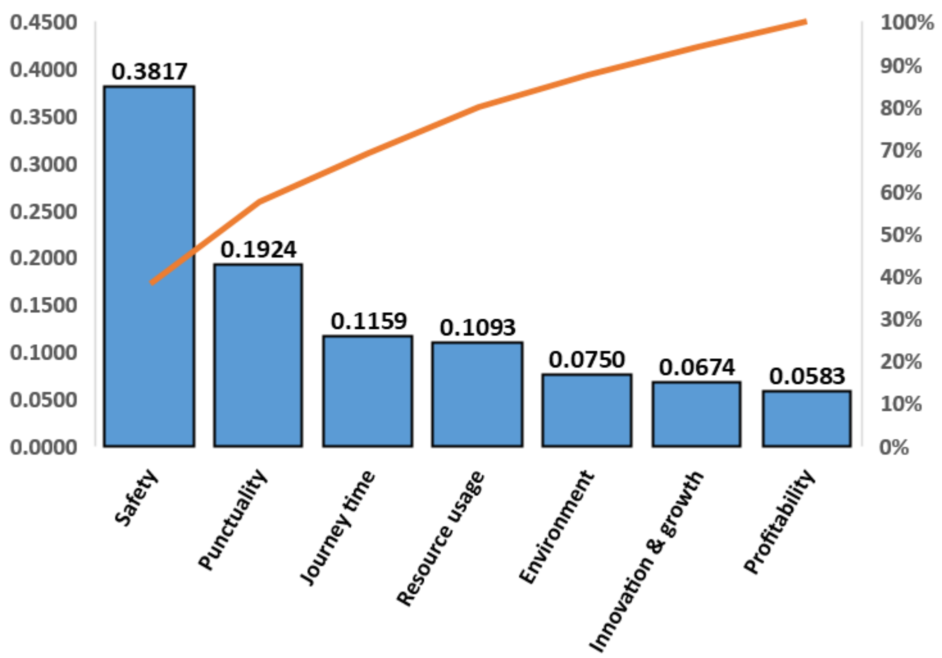

Figure 1. Relative weights (N_i) of KPIs obtained by FAHP method

4.2. Application of EDAS method

For three TOCs and seven KPIs, the decision matrix with the relative weights obtained by Buckley's FAHP method is shown in Table 5. Using Equation 5, the average solutions (AV_j) were calculated and presented for each TOC in Table 5.

Table 5. Decision matrix and average solutions

Weightage	0.1924	0.1159	0.1093	0.0583	0.3817	0.0674	0.0750
TOCs	Punctuality	Journey time	Resource usage	Profitability	Safety	Innovation & growth	Environment
TOC 1	0.4602	0.8688	0.7462	0.1314	2.9949	0.1829	0.2083
TOC 2	0.1511	0.6743	0.9059	0.4775	4.2538	0.4261	1.0000
TOC 3	0.6472	0.9265	0.7985	0.3911	1.2711	0.3910	0.5833
AV_j	0.4195	0.8232	0.8169	0.3333	2.8399	0.3333	0.5972

Since safety is considered as a non-beneficial criterion while other KPIs are considered as beneficial, Equations 6-9 appropriate for calculating PDA and NDA were used. The results are provided in Table 6.

Table 6. PDA and NDA values

TOC	Punctuality	Journey time	Resource usage	Profitability	Safety	Innovation & growth	Environment

	PDA	NDA	PDA	NDA	PDA	NDA	PDA	NDA	PDA	NDA	PDA	NDA	PDA	NDA
TOC 1	0.0187	0.0000	0.0064	0.0000	0.0000	0.0095	0.0000	0.0353	0.0000	0.0208	0.0000	0.0304	0.0000	0.0488
TOC 2	0.0000	0.1231	0.0000	0.0210	0.0119	0.0000	0.0252	0.0000	0.0000	0.1900	0.0188	0.0000	0.0506	0.0000
TOC 3	0.1045	0.0000	0.0145	0.0000	0.0000	0.0024	0.0101	0.0000	0.2109	0.0000	0.0117	0.0000	0.0000	0.0017

Using Equations 10 and 11, SP_i and SN_i values for each TOC were calculated. Using Equations 12 and 13, SP_i and SN_i were normalized, NSP_i and NSN_i values were obtained. Then in the final step, AS_i for each TOC were calculated using Equation 14. As the higher AS_i value indicates better performance, it's evident that TOC 3 exhibits the best performance, while TOC 2 demonstrates the least favorable performance as presented in Table 7.

Table 7. SP_i, SN_i, NSP_i, NSN_i and AS_i values

TOC	SP_i	SN_i	NSP_i	NSN_i	AS_i	Rank
TOC 1	0.0251	0.1449	0.0713	0.5665	0.3189	2
TOC 2	0.1065	0.3341	0.3028	0.0000	0.1514	3
TOC 3	0.3516	0.0042	1.0000	0.9874	0.9937	1

3.3. Sensitivity analysis

Sensitivity analysis, with the mainlines, can be defined as the exploration of how the outputs of a system are related to and are influenced by its inputs [34]. In mathematical models; sensitivity analysis plays a significant role in analyzing the impact and dependency of the inputs while determining the possible outputs of the model [35].

In this study, sensitivity analysis was carried out to reveal whether the alternatives of TOCs are sensitive to changes in KPI weights when fuzzy AHP application is performed with different decision makers.

Firstly, sensitivity analysis was performed on the EDAS method used in performance assessment by changing the KPI weights obtained with the fuzzy AHP method. Sensitivity analysis was carried out over seven different scenarios. In the scenarios, the weight of each KPI separately has been increased to 0.9, and the weights of the other KPIs have been evenly distributed. For example, In the scenario 1, it is assumed that the "punctuality" KPI is the most important criterion with 0.9 weight for performance evaluation of TOCs, and the weights of the other KPIs are evenly distributed. The KPI weights obtained using the fuzzy AHP method and the KPI weights assumed for scenarios of sensitivity analysis are shown in Table 8.

Table 8. KPI weights for sensitivity analysis

KPIs	FAHP	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7
Punctuality	0.192	0.900	0.017	0.017	0.017	0.017	0.017	0.017
Journey time	0.116	0.017	0.900	0.017	0.017	0.017	0.017	0.017
Resource usage	0.109	0.017	0.017	0.900	0.017	0.017	0.017	0.017
Profitability	0.058	0.017	0.017	0.017	0.900	0.017	0.017	0.017
Safety	0.382	0.017	0.017	0.017	0.017	0.900	0.017	0.017
Innovation & growth	0.067	0.017	0.017	0.017	0.017	0.017	0.900	0.017

Environment	0.075	0.017	0.017	0.017	0.017	0.017	0.017	0.900

The results obtained after the scenario implementation are outlined in Table 9. In the actual case, TOC 3 emerges as the best-performing company, while TOC 2 ranks as the lowest performing, as indicated in Table 9. However, upon sensitivity analysis, scenarios 3, 4, 6, and 7 demonstrate a change in rankings. Specifically, in these scenarios, TOC 2 emerges as the best-performing company, while TOC 1 falls to the lowest performing position. These findings suggest that TOC 1 wouldn't be the top-performing company, and TOC 3 wouldn't rank as the lowest performing company in the seven scenarios applied.

Table 9. Results of sensitivity analysis

Scenario	Ranking			Appraisal score		
	TOC 1	TOC 2	TOC 3	TOC 1	TOC 2	TOC 3
Real case	2	3	1	0.3189	0.1514	0.9937
Scenario 1	2	3	1	0.5610	0.0246	0.9994
Scenario 2	2	3	1	0.6030	0.0909	0.9979
Scenario 3	3	1	2	0.0105	0.8976	0.5120
Scenario 4	3	1	2	0.0031	0.9806	0.7196
Scenario 5	2	3	1	0.4169	0.0242	0.9992
Scenario 6	3	1	2	0.0047	0.9744	0.8294
Scenario 7	3	1	2	0.0021	0.9819	0.5035

The results of sensitivity analysis for each TOC based on the scenarios are illustrated in Figure 2. TOC 2 and TOC 3 appear to exhibit the best performance in half of the scenarios, while TOC 1 consistently performs the least favorably.

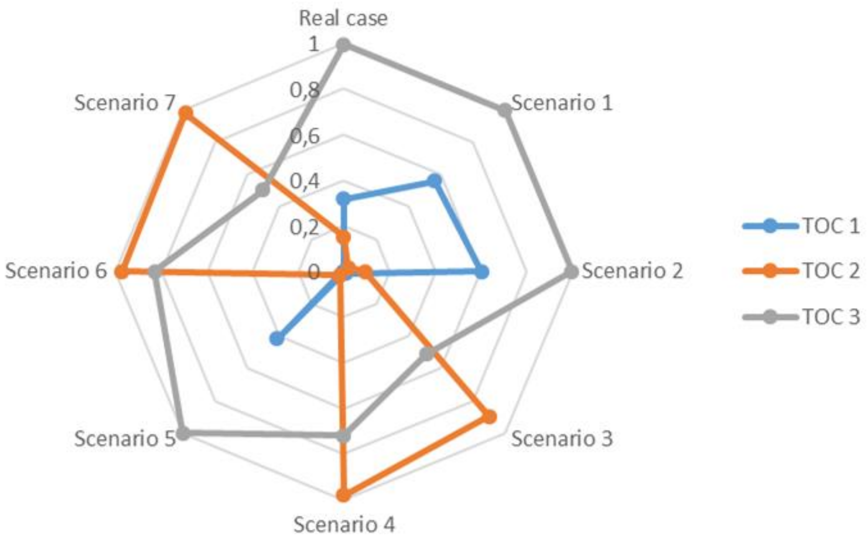

Figure 2. Sensitivity analysis for different scenarios

In addition, a sensitivity analysis was also performed to observe how the results change in case of evaluation with different types of MCDM methods. TOPSIS, COPRAS, and PROMETHEE II methods were applied to assess the performance of TOCs to support the EDAS results. The overall results are presented in Table 10.

Table 10. Results of MCDM methods applied

TOC	EDAS		TOPSIS		COPRAS		PROMETHEE II	
	Result	Rank	Result	Rank	Result	Rank	Result	Rank
TOC 1	0.3189	2	0.4437	2	56.46	2	-0.2160	2
TOC 2	0.1514	3	0.2147	3	54.90	3	-0.3066	3
TOC 3	0.9937	1	0.8944	1	100.00	1	0.5227	1

It is clearly seen from Table 10 that all the applied MCDM methods yielded the same ranking outcome despite the variations in performance scores. TOC 3 is identified as the best-performing company, while TOC 2 was identified as the lowest performing company.

5. Conclusion

In this study, the performances of three freight TOCs were evaluated using some MCDM methods. TOCs have a very important place in mass transportation in Türkiye as it has for the rest of the world. In Türkiye, with the Law No. 6461 on the Liberalization of Railway Transportation enacted in 2013, private railway TOCs were allowed to take part in the sector. As of 2022, three freight TOCs, two of which are privately owned and one of which is state-owned, are actively functioning in Türkiye. Considering that the number of private train operating companies will increase in the near future due to liberalization law, it will be more important for these companies to perform better in a more competitive environment.

The framework designed for the KPIs of TOCs was employed to assess company performance. Seven distinct KPIs were identified through a focus group meeting of eight experts, drawing on reports from railway transportation authorities, previous studies in the literature, and an understanding of the Turkish railway transportation sector. These determined KPIs include punctuality, journey time, resource usage, profitability, safety, innovation & growth, and environment. According to the Buckley's FAHP results, the findings indicate that safety has importance of 38%, punctuality 19%, journey time 12%, resource usage 11%, environment 8%, innovation & growth 7%, and profitability 6% on TOCs' performance.

The companies were ranked based on their performance according to KPIs using EDAS as main method. TOPSIS, PROMETHEE II and COPRAS methods are also applied to the problem to compare the results. Even though the performance scores calculated by MCDM methods are different, all four different methods give the same performance ranking. It was revealed that the company with the best performance is TOC 3, and the company with the worst performance is TOC 2. In order to present whether the alternatives of TOCs are sensitive to the changes in KPI weights in case of different decision makers in determining the KPI weights with the fuzzy AHP method, sensitivity analyzes were performed over seven different scenarios.

Railway freight and passenger transportation exhibit distinct structures and performance indicators. Notably, in Türkiye, public-owned railway companies primarily handle mainline passenger transportation, except for suburban and metro lines. However, if the landscape evolves in the future with an increase in the number of companies involved in passenger transportation, this study can be extended to incorporate passenger train operating companies, advancing its scope and relevance.

References

[1] M. Frost, S. G. Ison, and R. Watson, "UK rail transport: a review of demand and supply," *Proceedings of the ICE: Transport*, vol. 165, no. 3, pp. 225-234, 2012

[2] TCDD, "Turkish State Railways sectoral report," Turkey, 2020

[3] IEA, "The future of ail: Opportunities for energy and the environment", Paris, 2019

[4] M. Lu, "Evaluation of railway performance through quality of service," Ph.D. dissertation, University of Birmingham, England, 2016

[5] *UIC leaflet 406: capacity*, UIC International Union of Railways, France, 2004

[6] R. J. Anderson, R. Hirsch, M. Trompet, and W. Adeney, "Developing benchmarking methodologies for railway infrastructure management companies. Railway Technology Strategy Centre, Centre for Transport Studies, London, United Kingdom, 2003

[7] IMPROVERAIL, "Improved tools for railway capacity and access management: D2 benchmarking methodologies a uropa zationion of concepts in the railway sector," Competitive and Sustainable Growth Programme, European Commission, Belgium, 2003

[8] European Commission, "Infrastructure managers (PRIME)", 2023. [Online]. Available: https://transport.ec.europa.eu/transport-modes/rail/market/infrastructure-managers-prime_en [Accessed August 12, 2023]

[9] Key performance indicators for performance benchmarking, PRIME Catalogue, 2019

[10] A. Cebeci, H. Tüydeş-Yaman, and D. M. Z. Islam, "Spatial distribution of the rail freight demand in Turkey prior to railway reform," *Research in Transportation Business and Management*, vol. 44, 2022

[11] UAB-Ministry of Transport and Infrastructure, "Demiryolu tren işletmecisi yetki belgesi sahibi firmalar," 2020. [Online]. Available: https://uhdgm.uab.gov.tr/uploads/pages/demiryolu-tasimaciligi-yetki-belgesi-almis-olan fi/demiryolu-tren-isletmecisi-yetki-belgesi-sahibi-firmalar.pdf

[12] Y. He, F. Lei, G. Wei, R. Wang, J. Wu, and C. Wei, "EDAS method for multiple attribute group decision making with probabilistic uncertain linguistic information and its application to green supplier selection," International Journal of Computational Intelligence Systems, vol. 12, no. 2, pp. 1361-1370, 2019

[13] M. Keshavarz Ghorabaee, E. K. Zavadskas, L. Olfat, and Z. Turskis, "Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS)," *Informatica*, vol. 26, no. 3, pp. 435-451, 2015

[14] G. F. Frederico, and V. Cavenaghi, "Measuring performance in rail freight transportation companies," *International Business Research*, vol. 10, no. 11, pp. 117-128, 2017

[15] L. Zhang, Q. Cheng, and Q. Siyuan, "Evaluation of railway transportation performance based on CRITIC-Relative Entropy Method in China," *Journal of Advanced Transportation*, vol. 2023, 2023

[16] R. Ranjan, P. Chatterjee, and S. Chakraborty, "Performance evaluation of Indian Railway zones using DEMATEL and VIKOR methods," *Benchmarking: An International Journal*, vol. 23, no. 1, pp. 78-95, 2016

[17] E. Jose, P. Agarwal, J. Zhuang, and J. Swaminathan, "A multi-criteria decision making approach to evaluating the performance of Indian railway zones," *Annals of Operations Research*, vol. 325, no. 2, pp. 1133-1168, 2023

[18] N. Petrović, J. Mihajlović, V. Jovanović, D. Ćirić, and T. Živojinović, "Evaluating annual operation performance of Serbian railway system by using multiple criteria decision-making technique" *Acta Polytechnica Hungarica*, vol. 20, no. 1, 2023

[19] A. Fraszczyk, T. Lamb, and M. Marinov, "Are railways really that bad? An evaluation of rail systems performance in Europe with a focus on passenger rail," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 573–591, 2016

[20] S. Stoilova, N. Munier, M. Kendra, and T. Skrúcaný, "Multi-criteria evaluation of railway network performance in countries of the TEN-T orient-east med corridor," *Sustainability*, vol. 12, no. 4, pp. 1482, 2020

[21] S. D. Stoilova, "A multi-criteria assessment approach for the evaluation of railway transport in the Balkan region," *Promet-Traffic&Transportation*, vol. 31, no. 6, pp. 655-668, 2019

[22] K. Kara, and G. C. Yalçın, "Assessing railway transportation performance of European countries with CRITIC and ROV techniques" *Demiryolu Mühendisliği*, no. 17, pp. 93-106, 2023

[23] M. B. Bouraima, A. Saha, A. Stević, Z. Antucheviciene, J. Qiu, and P. Marton, "Assessment actions for improving railway sector performance using intuitionistic fuzzy-rough multi-criteria decision-making model," *Applied Soft Computing*, vol. 148, 2023

[24] M. G. Sharma, R. M. Debnath, R. Oloruntoba, and S. M. Sharma, “Benchmarking of rail transport service performance through DEA for Indian railways,” *The International Journal of Logistics Management*, vol. 27, no. 3, pp. 629-649, 2016

[25] N. O. Olsson, and H. Haugland, “Influencing factors on train punctuality - results from some Norwegian studies,” *Transport policy*, vol. 11, no. 4, pp. 387-397, 2004

[26] B. D. Dağıdır, and B. Özkan, “A comprehensive evaluation of a company performance using sustainability balanced scorecard based on picture fuzzy AHP,” Journal of Cleaner Production, vol. 435, no. 140519, 2024

[27] N. H. Zardari, K. Ahmed, S. M. Shirazi, and Z. B. Yusop, “Weighting methods and their effects on multi-criteria decision-making model outcomes in water resources management,” *Springer International Publishing*, Switzerland, pp. 1-5, 2014

[28] A. Özdağoğlu, and G. Özdağoğlu, “Comparison of AHP and fuzzy AHP for the multi-criteria decision making processes with linguistic evaluations,” *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 6, no. 11, pp. 65-85, 2007

[29] J. J. Buckley, “Fuzzy hierarchical analysis,” *Fuzzy sets and systems*, vol. 17, no. 3, pp. 233-247, 1985

[30] M. B. Ayhan, “A fuzzy AHP approach for supplier selection problem: A case study in a gear motor company,” *International Journal of Managing Value and Supply Chains*, vol. 4, pp. 11-23, 2013

[31] M. Mathew, and S. Sahu, “Comparison of new multi-criteria decision making methods for material handling equipment selection,” *Management Science Letters*, vol. 8, no. 3, pp. 139–150, 2018

[32] N. Yalçın, and N. Uncu, “Applying EDAS as an applicable MCDM method for industrial robot selection,” *Sigma Journal of Engineering and Natural Sciences*, vol. 37, no. 3, pp. 779-796, 2019

[33] A. Zaki Mohamed Noor et al., “Fusion of fuzzy AHP in selecting material for drinking water bottle based on customer needs,” *Journal of Engineering and Applied Sciences*, vol. 12, no. 14, pp. 4243-4249, 2017

[34] S. Razavi et al., “Future of sensitivity analysis: An essential discipline for systems modeling and policy support,” *Environmental Modelling and Software*, vol. 137, 2021

[35] B. Iooss, and A. Saltelli, “Introduction to sensitivity analysis,” in Handbook of Uncertainty Quantification, *Springer*, Cham, pp. 1-31, 2017

Resume

Çağdaş YÜKSEL
He was born in 1989, completed his bachelor's degree at Çukurova University, Department of Industrial Engineering, and his master's degree at Adana Alparslan Türkeş Science and Technology University, Department of Industrial Engineering. He works at TCDD 6th Regional Directorate, Department of Capacity Management.
Email: cagdasyuksel@tcdd.gov.tr

Nuşin UNCU
She was born in 1980, completed her bachelor's degree at Çankaya University Industrial Engineering (with scholarship). She works as a faculty member at Adana Alparslan Türkeş Science and Technology University. Her research areas are modelling of production and service processes and decision-making methods.
Email: nuncu@atu.edu.tr

Ethics Statement:

Authors confirm that the article is original, there is no plagiarism issue.

Contribution of authors: Çağdaş YÜKSEL: Conceptualization, Methodology, Software, Visualization, Investigation, Supervision. Nuşin UNCU: Resources, Validation, Writing-Original draft preparation, Writing-Reviewing and Editing.