



A decision support framework to evaluate the sustainability performance of urban road transportation

Ardavan Babaei¹ · Majid Khedmati¹ · Mohammad Reza Akbari Jokar¹ · Erfan Babae Tirkolae^{2,3,4}

Received: 12 September 2023 / Accepted: 22 November 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

This study proposes a decision support framework (DSF) based on two data envelopment analysis (DEA) models in order to evaluate the urban road transportation of countries for sustainable performance management during different years. The first model considers different years independently while the second model, which is a type of network model, takes into account all the years integrated. A multi-objective programming model under two types of uncertainties is then developed to solve the proposed DEA models based on a revised multi-choice goal programming (GP) approach. The efficiency scores are measured based on the data related to several major European countries and the factors including the level of freight and passenger transportation, level of greenhouse gas emissions, level of energy consumption, and road accidents which are addressed as the main evaluation factors. Eventually, the two proposed models are compared in terms of interpretation and final achievements. The results reveal that the efficiency scores of countries are different under deterministic/uncertain conditions and according to the structure of the evaluation model. Furthermore, efficiency changes are not necessarily the same as productivity changes. The high interpretability (up to 99.6%) of the models demonstrates the reliability of DSF for decision-making stakeholders in the transport sector. Furthermore, a set of managerial analyses is conducted based on different parameters of the performance evaluation measures for these countries including the productivity changes during the period under consideration, resilience of the countries, detection of the benchmark countries, ranking of different countries, and detection of the patterns for improving the transportation system.

Keywords Urban road transportation · Sustainability performance · Energy consumption · Data envelopment analysis model · Revised multi-choice goal programming

Responsible Editor: Philippe Garrigues

✉ Mohammad Reza Akbari Jokar
Reza.Akbari@sharif.edu

Ardavan Babaei
Ardvan.Babaei@ie.sharif.edu

Majid Khedmati
Khedmati@sharif.edu

Erfan Babae Tirkolae
erfan.babae@istinye.edu.tr

¹ Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran

² Department of Industrial Engineering, Istinye University, Istanbul, Turkey

³ Department of Industrial Engineering and Management, Yuan Ze University, Taoyuan, Taiwan

⁴ Department of Industrial and Mechanical Engineering, Lebanese American University, Byblos, Lebanon

Introduction

Development of the transportation sector is one of the important factors that can increasingly contribute to the economic growth of countries. In spite of positive effects on economic growth, transportation development may have negative consequences for society and the environment. For example, despite the economic benefits of freight transportation for each country, carbon emissions are among the negative consequences of freight transportation (Stenico de Campos et al. 2019; Gandhi et al. 2022). Thus, developments in this sector need to be sustainable, in which sustainable development covers the two factors social and environmental factors in addition to the economic factors (Tian et al. 2020). The European Union (EU) is a signatory to the United Nations Framework Convention on Climate Change (UNFCCC) and

submits an annual report on its greenhouse gas inventory for the year $t-2$ within the area covered by its Member States. The report includes data on carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O), perfluorocarbons (PFCs), hydrofluorocarbons (HFCs), sulfur hexafluoride (SF₆), and nitrogen trifluoride (NF₃) (Statista Research Department 2022a). According to Tiseo (2023), the transportation sector was the primary source of greenhouse gas emissions in the United Kingdom in 2021. Additionally, the number of road fatalities in the European Union increased by about 5% between 2020 and 2021, with 1000 more deaths reported (Statista Research Department 2022b). In Italy, road transportation accounts for the largest share of energy consumption in the transportation sector (Statista Research Department 2022c). These statistics highlight the importance of considering sustainability in transportation planning and evaluation. The decision-making and evaluation instruments are required to determine the significance of economic, social, and environmental factors and, consequently, to define the roadmap of improvements for the transportation sector (Nag et al. 2018; Mahdinia et al. 2018). Safety is one of the important indices of sustainability which has received special attention in the literature. Transportation development leads to an increase in the rate of traffic crashes, and hence, traffic safety management is required as a basis for safety assessment in the transportation sector (Xie et al. 2019). The safety can also be assessed from the economic perspective where it affects the benefit–cost analyses through investment justification and prioritization of the projects (Daniels et al. 2019; Proost et al. 2014). Data envelopment analysis (DEA) is a decision-making tool which is utilized for assessment purposes in the transportation sector (Tovar and Wall 2017; Xie et al. 2018) and can also be employed for the assessment of sustainability and resilience (García-Palomares et al. 2018; Ji et al. 2016; Hahn et al. 2017; Twumasi-Boakye and Sobanjo 2018). This research focuses on evaluating sustainability in transportation at the country level using the widely used and useful tool of DEA. The first sub-section reviews the existing research literature, while the second sub-section identifies research gaps and highlights our innovations. Finally, the third sub-section explains the research framework of this article.

Literature review

Izadikhah et al. (2021) presented a DEA-based optimization model under uncertain conditions to evaluate the sustainability and resilience aspects of 21 public transportation providers in three Iranian megacities. Omrani et al. (2023) introduced a new efficiency score, sustainable efficiency, that considers all three aspects. Existing and new DEA models were applied to evaluate the technical, social, environmental, and sustainable efficiencies

of Iranian airlines. The Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS) approach was also used to integrate the results of the four models, providing a comprehensive ranking of the airlines. Babaei et al. (2022a) evaluated the distribution network configurations (which was subject to many transportation decisions) by developing a DEA model from the perspective of cost, transparency, service, and environmental criteria. Alper et al. (2015) used DEA to study the performance of local municipalities in traffic safety during 2004–2009. They applied three models for evaluation where the first model included 2 inputs and 14 outputs, the second model as a two-stage model included 2 inputs and 8 intermediate outputs in the first stage, and 8 intermediate inputs and 6 outputs in the second stage, and finally, the third model included 2 inputs and 8 outputs. They validated their models and concluded that a decrease in accidents leads to an increase in efficiency. Io Storto and Evangelista (2023) investigated infrastructure efficiency measurements for road and rail modes of transport in 28 EU countries from 2010 to 2017 through a DEA model. Their results showed that countries are hardly able to improve both operational and environmental performances at the same time. Nikolaou and Dimitriou (2018) used the output-oriented DEA model to assess 23 European countries in terms of road safety from 2005 to 2014 and showed that the conditions are associated with uncertainties during the study period. Considering road safety as a major challenge in the EU countries, Shen et al. (2013) applied the output-oriented DEA model for assessment of road safety in EU countries during 2001–2010. The population, passenger-kilometers, and passenger-cars are considered as the inputs while the number of fatalities is accounted for as the only output of the model. They employed the Malmquist index to account for the transportation conditions during the study period and indicated the progress in road safety management and fatality risk mitigation in the EU countries. Ganji et al. (2019) criticized the traditional DEA methods used in safety management. In this work, in addition to the efficient frontier, the anti-efficient frontier was also taken into account, and hence, the performance was investigated under both optimistic and pessimistic conditions. They implemented the input-oriented DEA model and implemented this model on a real case study in Iran; it was conducted that the cities of Ilam, Alborz, and Hormozgan are successful in road safety management. Saeedi et al. (2019) utilized a network DEA to evaluate the intermodal freight traffic. They employed a slack-based model to create discriminative power between different freight transportation chains. Egilmez and Mcavoy (2013) applied the output-oriented DEA model to minimize the number of fatal crashes and investigate road safety in 50 US states during 2002–2008, showing that the utility of

economic and social resources is still inefficient, despite the decline in the fatality trend.

Given the fact that freight delivery is one of the most important subjects in city logistics, Muñuzuri (2019) suggested a Charnes, Cooper, and Rhodes (CCR)-output-oriented model to address this problem. Dadashi and Mirbaha (2019) used the Monte Carlo simulation to address the uncertainty in DEA and proposed a model to allocate the budget to projects such that the safety improvements in highways are more effectively guaranteed. Determination of the exact level of carbon emissions is one of the other examples of uncertainty owing to its ambiguity, which was addressed by Ignatius et al. (2016) using a fuzzy DEA model. Zhao et al. (2019) offered a DEA model to evaluate 30 major Chinese cities from the viewpoint of sustainability. The proposed model consists of two subsystems in which, one subsystem was intended to evaluate the economic and environmental dimensions of efficiency, and the other subsystem was used to evaluate the social dimension of efficiency. They demonstrated that the environmental and economic dimensions of the western cities and the social dimension of the eastern cities are not favorable. Taking into account technology changes, Sueyoshi et al. (2019) introduced new DEA-related indicators to evaluate the CO₂ emissions from 10 sources in Chinese provinces. Kutty et al. (2022) assessed the sustainability performance of 35 European smart cities from 2015 to 2020 using a novel DEA model. The model considered undesirable factors in the technology set and used an integrated relative sustainability performance assessment model to determine the most efficient smart city under the six dimensions of sustainable development. Hussain (2022) examined the joint effect of economic and environmental factors on transport efficiency, including the role of climate change mitigation technology. Using data from 35 Organization for Economic Cooperation and Development countries between 2000 and 2020, two approaches were employed: DEA based on slack-based measure and cross-sectional dependence autoregressive distributed lag. Ding et al. (2019) proposed a hybrid model consisting of cross efficiency (CE) and Malmquist index to perform a dynamic assessment. In this research, 30 Chinese provinces were assessed in terms of carbon emission and the findings showed that the eastern provinces are more efficient than the western provinces. Kiani Mavi et al. (2019) took into account the concept of the ideal point in the Malmquist index to increase the discriminative power of their model and applied it to the Iranian freight transportation sector. The inputs of this model included the cargo tonnage, road length, fuel consumption, and public freight transportation while the outputs included safety education, R&D expenditure, CO₂ emission, freight turnover volume, and traffic accidents. The safety assessment has also been addressed in many other fields to improve efficiency. To

name a few, Evans et al. (2016) addressed safety management in the field of air traffic management, Nahangi et al. (2019) implemented it in the field of construction sites, and Djordjevi et al. (2018) and Nahangi et al. (2019) applied it to address the railroad safety. It should be noted that the DEA model is widely used not only in the field of transportation evaluation but also in the fields of transportation planning, supply chain management and supply chain network design (Babaei et al. 2023a). By utilizing DEA models, supply chain design and transportation planning can be optimized not only for efficiency but also for sustainability (Babaei et al. 2023c, d, e, f).

Recently, an entropy-based DEA model was offered by Stefaniec et al. (2021) to evaluate the social sustainability of European Union (EU) road transport. As one of the main findings, it was suggested that social sustainability consideration leads to the exclusion of the bias arising from the economic factors' involvement. Memari et al. (2022) suggested a combinatorial decision-making approach to evaluate the sustainability performance of renewable energy facilities. They took into account data reliability and uncertainty using Z-number-based DEA model combined with fuzzy decision-making trial and evaluation laboratory (DEMATEL) and fuzzy analytical network process (ANP) techniques. Babaei et al. (2023b) developed a leader–follower DEA model based on the number of vehicles, violations, and average speed of vehicles to evaluate freeways leading to Tehran (a city in Iran). There are also some other decision-making and artificial intelligence tools to assess the sustainability performance in the transportation and logistics sections such as human–computer interaction of technology innovation, machine learning (Chen et al. 2022), and multi-criteria decision-making (MCDM) methods (Nag et al. 2018; Tian et al. 2023).

Research gaps and our innovations

Based on a thorough literature review, some research gaps have been identified which have not been addressed previously. In the following, the research gaps and our contributions are described:

- (i) In the literature, the mathematical models of performance assessment are typically used to assess different years, independently. In this work, in addition to the model that examines the countries based on sustainability every year, a network model is developed to examine all the years for the countries at the same time. In this way, the sustainability performance of countries is assessed with the help of a comprehensive

- approach (both independent and simultaneous evaluation of years).
- (ii) In the mathematical models of performance assessment, no specific goals are determined for the best and the worst performance states. Accordingly, a multi-objective optimization model is developed to evaluate countries' performance not only based on DEA common objective functions but also based on best and worst performance as well as maximum deviations. Considering approaches based on the best and worst performance and maximum deviations increases the power of differentiation in calculating the efficiency scores of countries in the models developed in this research. Based on the developed models, countries are compared not only based on the efficient frontier but also based on the best and worst performances. In addition, an efficient method is offered to find the bounds of the goals related to the objective functions.
 - (iii) Despite considering the uncertainty conditions in the literature, the uncertainty has not been considered in the form of data variation combined with interval goals. Measurement errors, variation in input–output relationships, and preventing incorrect and biased results require that uncertainty be considered in DEA models (Deh-nokhalaji et al. 2022). In addition, it is difficult for the decision-maker or supply chain manager to determine the deterministic and accurate values for the goals related to the objective functions (Chang 2008; Ledwith et al. 2021). Since multi-objective optimization models are taken into account in this study, the uncertainty conditions are assessed in input data, output data, and determination of goal values.
 - (iv) Analyzing the performance of countries is not limited to the calculation of efficiency. Productivity analysis by considering efficiency scores in different time conditions, resilience analysis by examining the worst and best performances, determining best practices as a progress plan to improve the performance of countries, and verifying the efficiency scores of countries extracted from DEA models, including important and necessary components for analyzing countries, while the literature has not comprehensively covered all the mentioned types of analysis. However, this research provides a comprehensive analysis of countries' performance based on the aforementioned types of analysis.
 - (v) Although the assessment of sustainability at the country level has a great impact on city transportation policy, there is little research that has done an evaluation using sustainability indicators at the country level. Hence, this level of evaluation has been the focus of our research.

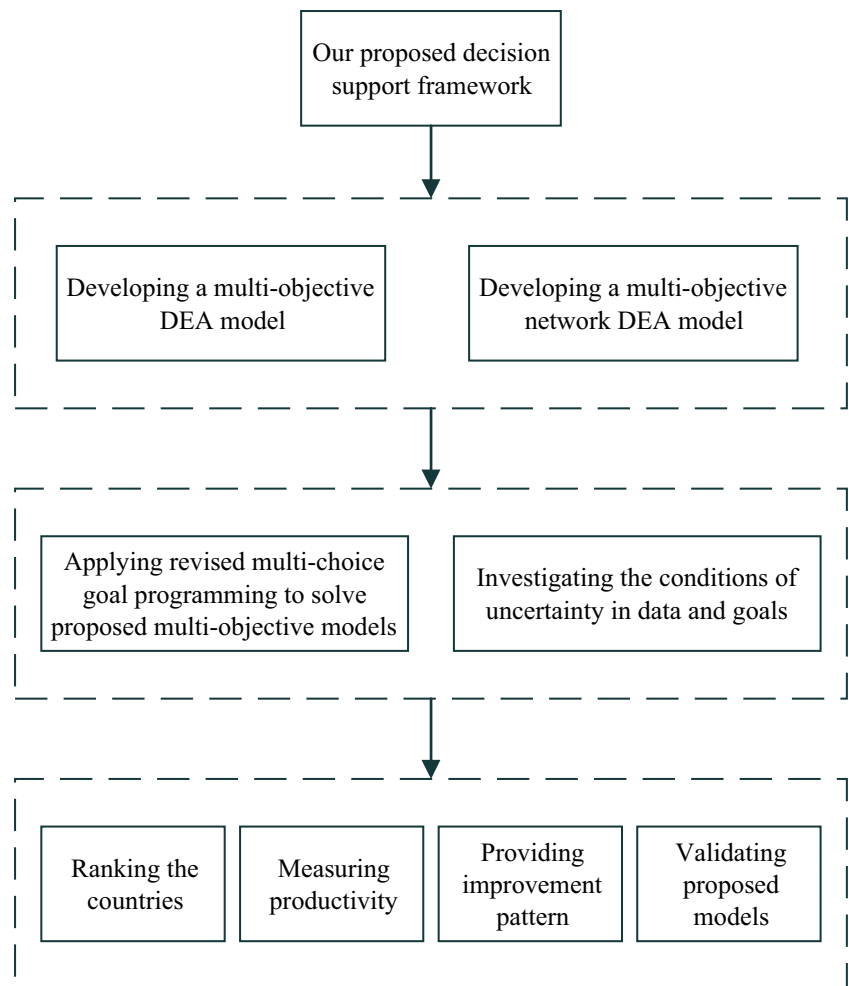
Based on identified research gaps and our innovative contributions, we have formulated the following research questions for our paper:

- I. How does the network model, which evaluates the sustainability performance of countries by considering all years simultaneously, compare to the traditional approach of assessing each year independently?
- II. How does the inclusion of best and worst performance states, as well as maximum deviations, affect the evaluation of countries' performance in the multi-objective optimization model?
- III. Why is it important to consider uncertainty in DEA models, particularly in terms of data variation and interval goals?
- IV. In what ways does the comprehensive analysis of countries' performance in this research, which includes efficiency analysis, productivity analysis, resilience analysis, identification of best practices, and verification of efficiency scores, differ from existing literature?
- V. How does this research contribute to the assessment of sustainability indicators at the country level?

Research framework

Considering research gaps and our contributions, our goal is to evaluate the performances of countries based on sustainability aspects through the development of multi-objective and network DEA models under uncertain conditions and provide comprehensive analysis regarding the performances. In this regard, a DSF as shown in Fig. 1, is presented in current research. To evaluate countries from the perspective of sustainability, two multi-objective optimization models are presented. One evaluates countries every year, as given by Formulas (6), (7), (8), (9), (10), (11), (12), and (13). The other is designed as a network model. This means that it evaluates the performance of each country in all years simultaneously (Formulas (14), (15), (16), (17), (18), (19), (20), and (21)). The revised multi-choice GP is then employed to solve the proposed multi-objective models, as shown by Formulas (22), (23), (24), (25), (26), and (27). To be closer to the real world, the uncertainty conditions in the developed models are considered both in the input and output data (Formulas (28), (29), (30), (31), (32), (33), (34), and (35)) and uncertainty conditions in determining the goal values (Formulas (36) and (37)). Then, the efficiency scores of countries are measured from the perspective of sustainability under deterministic and uncertain conditions based on the suggested multi-objective model and multi-objective network model (“**Results**”). Productivity measurement based on countries' efficiency scores under different years (“**Analysis of the productivity**”), examining of resilience based on the best and worst efficiency scores (“**Analysis of the resilience**”), analysis of countries' ranking based on resilience and productivity (“**Analysis of the ranks**”), identification of efficient patterns to provide the roadmap for improvement in low-performing countries (“**Analysis of the best practices**” and “**Analysis of the benchmarks**”), and

Fig. 1 Our proposed framework



validation of the proposed models through interpreting the efficiency scores (“[Analysis of the methods](#)”) are among the analyses provided by the proposed DSF.

The rest of the paper is organized as follows. In “Problem statement,” the suggested performance evaluation models are described in detail. The solution approach is provided in “Solution approach,” considering the multi-objective and uncertainty programming models. The extended performance evaluation models are implemented on the data of 12 European countries in “[Results](#).” Then, in “Managerial analysis,” the managerial analyses and some policy implications are presented. Finally, the discussion, concluding remarks, and suggestions for future research are given in “[Discussion](#)” and “[Conclusions and outlook](#).”

Problem statement

DEA is a methodology that can evaluate the performance of comparable decision-making units (DMUs) by forming an efficient frontier in which, there exist some inputs and outputs for each DMU. A linear mathematical

programming model is employed to compare the relative performance of all DMUs such that this model will be input-oriented if it maximizes the outputs by keeping the input level constant while it would be output-oriented if it minimizes the inputs by keeping the output level constant. The returns to scale (RTS) and orientation are two important characteristics of DEA models. In the constant RTS, a multiple of inputs produces the same multiple of outputs, while in variable RTS, each multiple of inputs can produce the same, greater or smaller multiple of outputs. The DEA models are able to provide benchmarks for inefficient units as reference sets in order to transform these units into efficient units. Charnes et al. (1978) presented a mathematical output-oriented model, abbreviated to CCR, with a constant RTS. This model is presented in Formulas (1), (2), (3), and (4) (Aldamak and Zolfaghari 2017):

$$\text{minimize } \sum_i v_i x_{io} \tag{1}$$

$$\text{subject to } \sum_r u_r y_{ro} = 1 \tag{2}$$

$$\sum_r u_r y_{r\tau} - \sum_i v_i x_{i\tau} \leq 0 \quad \forall \tau, \quad (3)$$

$$u_r, v_i \geq \varepsilon \quad \forall i, r, \quad (4)$$

where index $\tau = 1, \dots, n$ denotes the number of DMUs, $r = 1, \dots, s$ denotes the number of outputs and $i = 1, \dots, m$ represents the index of inputs, x_{ij} and y_{rj} stand for the input and output parameters, respectively, v_i and u_r represent the decision variables related to the input weight and output weight, respectively, and ε denotes a negligible value (Zhu 2009). In this model, Eq. (1) shows the objective function which minimizes the weighted sum of inputs of DMUs under consideration. Formulas (2), (3), and (4) express the constraints of the mathematical model where Eq. (2) sets output levels to 1, Constraint (3) states the ratio of output to input for all DMUs, and Constraint (4) expresses the positive variables greater than an equal to ε . The efficiency score of each DMU is determined by $\frac{1}{\sum_i v_i^* x_{ij}}$, after solving the output-oriented CCR model. In this equation, v_i^* is the optimal weight obtained from the mathematical model. Each time the mentioned model is solved, each DMU (here, each country) is evaluated with input and output criteria and its efficiency is measured.

In some cases, the inputs or outputs are undesirable (Liu et al. 2010). Generally, in DEA models, the inputs are undesirable and the outputs are desirable. Nevertheless, the outputs including carbon emissions, road accident fatalities and the quantity of energy (fuel) consumed in transportation are undesirable factors that should be transformed into desirable factors. Equation (5) indicates how undesirable factors can be transformed into desirable ones (Zhu 2009):

$$ny_{r\tau} = -y_{r\tau} + \max\{y_{r\tau}\} + 1, \quad (5)$$

where $ny_{r\tau}$ is the optimal output of the τ^{th} DMU and $\max\{y_{r\tau}\}$ represents the maximum output for all DMUs.

In this work, some indices are assigned to the input and output groups to evaluate the sustainability of road transportation in different countries wherein the outputs include the pollutant emission from transport (PET), the final energy consumption in road transport (FECRT), and the people killed in road accidents (PKRA). Since increasing the value of the outputs is considered undesirable, they can be transformed into desirable outputs based on Eq. (5). In addition, the inputs include passenger road transport on national territory (PRTNT) and total transported goods on national roads (TTGNR). The sustainability of the countries Germany (GE), Bulgaria (BU), France (FR), Croatia (CR), Italy (IT), Latvia (LA), Lithuania (LI), Poland (PO), Slovakia (SL), Finland (FI), Spain (SP), and Romania (RO) is assessed for years during 2011–2015, and it is attempted

to determine the values of PRTNT and TTGNR for which the optimized values of PET, FECRT, and PKRA can be obtained for each country. It is noteworthy that the criteria considered in this work are widely used in the literature of road transport sustainability performance evaluation (Babaei et al. 2022b; Omrani et al. 2020; Shen et al. 2012). For example, criteria related to road accidents, volume of freights, and passengers in Shen et al. (2013), Shen et al. (2015), and Wu et al. (2016) were considered in the topic of transport evaluation. Furthermore, other useful criteria in the aspects of sustainability that are related to energy consumption and emissions have been mentioned by some studies such as Chang et al. (2013), Ignatius et al. (2016), and Chen et al. (2020). All these research works inferred that the aforementioned criteria make the DEA model as a powerful evaluation tool in the field of transportation. Regarding the consideration of 5-year data (2011 to 2015), it is worth noting that the data is not collected and reported immediately after the end of each year. Data is collected over a longer period and then validated. In addition, the data from different countries are not released at the same time. Therefore, access to the data of countries regarding certain criteria becomes possible after the passage of time. Moreover, with the passage of time, the previous data are corrected by the collectors and reviewers, and as a result, they become more accurate and stable. Due to this, these years are taken into account to make the data of the countries available, accurate, and stable. It should be noted that our goal in this work is to come up with novel decision-making models based on DEA. Therefore, our models can be implemented on any data, and there are no restrictions in this regard for our models.

The discussions related to sustainability and transportation are topics that politicians, industries, people, and researchers are interested in. It is worth noting that evaluating the transportation of countries has a significant impact on cities in terms of policy and resources (Wątróbski et al. 2022). Countries with lower efficiency levels set regulations and standards that cities are required to implement. Furthermore, if countries with low-efficiency levels try to improve their transportation programs, cities will necessarily be involved in the improvement plan. Thus, it is necessary to anticipate and provide the necessary resources (e.g., financial resources for investment) required by cities.

In this study, two models are proposed to evaluate the sustainability performance of road transportation in different countries.

Model (1)

In this model, the efficiency score of each country is measured relative to other countries, in each iteration, with regard to sustainability. This model is represented in Formulas (6), (7),

(8), (9), (10), (11), (12), and (13). In this model, each country should be separately compared to other countries during each year. It is worth mentioning that the proposed multi-objective model considers wider dimensions compared to the research literature models. Our models not only minimize the weighted sum of the input of the DMU under investigation but also optimize the ideal and anti-ideal DMUs at the same time. In addition, considering the amount of deviations can create more differentiation in the calculation of the efficiency score of the DMUs.

$$\text{minimize } \sum_i v_i x_{io} \tag{6}$$

$$\text{minimize } M + d_o \tag{7}$$

$$\text{minimize } \sum_i v_i x_i^l \tag{8}$$

$$\text{minimize } \sum_i v_i x_i^u \tag{9}$$

$$\text{subject to } \sum_r u_r y_{ro} = 1 \tag{10}$$

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} + d_j = 0 \quad \forall j, \tag{11}$$

$$M \geq d_j \quad \forall j, \tag{12}$$

$$u_r, v_i \geq \epsilon, d_j \geq 0, M \text{ is a free variable } \forall i, r, \tag{13}$$

where $i, r,$ and j represent the inputs (PRTNT and TTGNR), the outputs (PET, FECRT, and PKRA) and the countries, respectively, d_j denotes the deviation variable for country j (inefficiency value), d_o represents the deviation variable for the country under consideration, x_i^l and x_i^u represent the minimum and maximum value of input $i,$ respectively, and M is the free variable of deviation. In this model, Eqs. (6), (7), (8), and (9) show the objective functions and Formulas (10), (11), (12), and (13) represent the constraints of the model. Objective Function (6) minimizes PRTNT and TTGNR for each of the countries under consideration. Objective Function (7) also tries to minimize the level of deviation for the country under consideration as well as the maximum deviation level for all countries. This objective function can bring the power of the mathematical model to create a differentiation between the efficiency scores of the DMUs. For more details, please see Ghasemi et al. (2014). Objective Function (8) considers the best state for the country under consideration because the outputs are obtained from the lowest levels of the inputs, and Objective Function (9) considers the worst state for the country under consideration

because the outputs are obtained from the highest levels of the inputs. Considering two Objective Functions (8) and (9) increases the discriminative power of the model to calculate the efficiency scores of the DMUs where, in addition to the normal state (conventional state that DEA uses to calculate the efficiency), the best and the worst states are also reported to the decision-maker. Equation (10) sets the output level to 1, Eq. (11) demonstrates the outputs to inputs ratio for each country, Constraint (12) displays the maximum deviation, and Constraint (13) indicates the non-negative and free variables. It should be noted that when the value of “ d_o ” is equal to zero, the DMU is fully efficient.

Model (2)

In this model, in each iteration, the efficiency score of the country under consideration compared to the efficiency score of other countries during all the years is measured concerning sustainability. In this model, each DMU includes some divisions equal to the number of years, and each division has some inputs and outputs that are independent of the ones for other divisions (Kao 2014). The ratio of the number of runs in Model (2) to those in Model (1) is equal to $\frac{1}{NP}$, where NP denotes the number of years under study. This model is represented in Formulas (14), (15), (16), (17), (18), (19), (20), and (21):

$$\text{minimize } \sum_k \sum_i v_i x_{io}^k \tag{14}$$

$$\text{minimize } M + \sum_k d_o^k \tag{15}$$

$$\text{minimize } \sum_i v_i x_i^l \tag{16}$$

$$\text{minimize } \sum_i v_i x_i^u \tag{17}$$

$$\text{subject to } \sum_k \sum_r u_r y_{ro}^k = 1, \tag{18}$$

$$\sum_r u_r y_{rj}^k - \sum_i v_i x_{ij}^k + d_j^k = 0 \quad \forall j, k, \tag{19}$$

$$M \geq d_j^k \quad \forall j, k, \tag{20}$$

$$u_r, v_i \geq \epsilon, d_j^k \geq 0, M \text{ is a free variable } \forall i, r, k. \tag{21}$$

where $i, r,$ and j represent the inputs, outputs and countries, respectively. Index k represents each year under study, x_i^l and x_i^u represent the minimum and maximum input values for all the years. Objective function (14) minimizes the weighted sum of the country’s inputs during all years. Objective

function (15) minimizes the sum of deviations (inefficiencies related to all years) for the country under consideration and the maximum deviations (inefficiencies related to all countries and all years). Objective functions (16) and (17) minimize the best and worst states of the country under consideration (where there are the lowest weighted sum of inputs and the highest weighted sum of inputs, respectively). By considering both the worst and best-case scenarios, the resulting solutions will be more robust. Furthermore, the evaluation of countries will be more comprehensive. Therefore, a decision that takes into account not only normal conditions but also extreme situations in the real world ensures foresight and adaptability in different conditions. Constraint (18) conventionally sets the weighted sum of the outputs equal to one. Constraint (19) determines the inefficiency for each country and each year. Constraint (20) shows the maximum inefficiency. Constraint (21) represents decision variables.

Solution approach

In this section, two approaches are proposed to solve the models proposed in the previous section wherein the first approach is proposed to solve the multi-objective model while the second approach is proposed to address the uncertainty.

Multi-objective programming approach

Goal programming (GP) is one of the multi-objective programming methods which minimizes the deviations between each objective and the goal level of that objective. The goals are usually determined by the decision-maker. Depending on the significance of the objective functions, the significance of the objectives' deviations varies, too. In this regard, the deviations that are of higher significance are multiplied by a coefficient that is known as deviation weight. The objective has a higher significance if it has a greater coefficient. It should be noted that determining the goal level for each objective is very difficult and, in some cases, impossible. In some cases, considering multiple levels for each goal may result in more utility for the decision-maker than considering only a single level in which, multi-choice GP has been developed for these cases. However, these models contain some binary variables which makes it difficult to solve them. Accordingly, a revised multi-choice GP model with continuous variables is developed by Chang (2008) which is shown as follows:

$$\text{minimize } \sum_l [w_{o_l}(d_l^+ + d_l^-) + w_{d_l}(e_l^+ + e_l^-)] \quad (22)$$

$$\text{subject to } \sum_l f_l(x) - d_l^+ + d_l^- = p_l \quad \forall l, \quad (23)$$

$$p_l - e_l^+ + e_l^- = g_l^{\min} \quad \forall l, \quad (24)$$

$$g_l^{\min} \leq p_l \leq g_l^{\max} \quad \forall l, \quad (25)$$

$$H(x)(\leq = \geq) 0 \quad (26)$$

$$d_l^+, d_l^-, e_l^+, e_l^- \geq 0 \quad \forall l, \quad (27)$$

where l presents the index of each objective and d_l^+ and d_l^- denote the deviations of each objective function from the goal level (p_l). Furthermore, g_l^{\min} and g_l^{\max} represent the minimum and maximum goal level for each objective, e_l^+ and e_l^- represent the variables of deviation from the minimum goal level for each objective, w_{o_l} denotes the importance weight of the objectives, w_{d_l} denotes the importance weight related to p_l , and $H(x)$ stands for the other constraints of the model. In this model, Eq. (22) expresses the objective function which minimizes the weighted sum of deviations and Formulas (23), (24), (25), (26), and (27) are the constraints of the mathematical model in which, Eq. (23) defines the goal level of objective functions, Eq. (24) tries to bring the objective function close to its minimum level, Constraint (25) indicates the range of goal levels, Constraint (26) states the other constraints of the model, and Constraint (27) shows the non-negative variables of the model. Therefore, in order to solve the models in Formulas (6), (7), (8), and (9) and (14), (15), (16), and (17), they should be transformed into the model in Formulas (22), (23), (24), (25), (26), and (27). In this way, our proposed models transform into single-objective models. Accordingly, we can optimally solve these models through common solution methods such as simplex.

Uncertainty-handling approach

In this section, two types of uncertainties are taken into account for two models presented in "Problem statement." In the first type of uncertainty, DMU_o is investigated in the presence of data variation. This type of uncertainty falls into the category of stochastic models which leads to an effective sensitivity analysis (Cooper et al. 2004). Equations (28), (29), (30), and (31) display the changes in Model (1) under uncertain conditions:

$$\text{minimize } \sum_i v_i (x_{io} + \sigma_{io}^{in} \varphi^{-1}(\alpha)) \quad (28)$$

$$\text{subject to } \sum_r u_r (y_{ro} - \sigma_{ro}^{ou} \varphi^{-1}(\alpha)) = 1, \quad (29)$$

Table 1 Objective sampling

Objectives	$f_1(x)$	$f_2(x)$...	$f_n(x)$
$f_1(x)$	$f_{11}(x)$	$f_{12}(x)$...	$f_{1n}(x)$
$f_2(x)$	$f_{21}(x)$	$f_{22}(x)$...	$f_{2n}(x)$
...
$f_n(x)$	$f_{n1}(x)$	$f_{n2}(x)$...	$f_{nn}(x)$
Average	$\bar{f}_1(x)$	$\bar{f}_2(x)$...	$\bar{f}_n(x)$
Standard deviation	σ_1^f	σ_2^f	...	σ_n^f

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} + d_j = 0 \quad \forall j \neq o, \tag{30}$$

$$\sum_r u_r (y_{ro} - \sigma_{ro}^{ou} \varphi^{-1}(\alpha)) - \sum_i v_i (x_{io} + \sigma_{io}^{in} \varphi^{-1}(\alpha)) + d_j = 0 \quad \forall j = o, \tag{31}$$

where σ_{io}^{in} and σ_{ro}^{ou} denote the standard deviation of the inputs and outputs of the country under consideration, respectively, $\varphi^{-1}(\alpha)$ denotes the fractile function, and α represents a pre-determined number between 0 and 1. Equations (28), (29), (30), and (31) build the uncertain counterpart of Model (1).

Applying the first type of uncertainty for the second model, the adjusted Model (2) is presented in Eqs. (32), (33), (34), and (35). In other words, Eqs. (32), (33), (34), and (35) create the uncertain counterpart of Model (2). It should be noted that other components of Models (1) and (2) (which do not face the data of the country under consideration) remain unchanged.

$$\text{minimize } \sum_k \sum_i v_i (x_{io}^k + \sigma_{io}^{in} \varphi^{-1}(\alpha)) \tag{32}$$

$$\text{subject to } \sum_k \sum_r u_r (y_{ro}^k - \sigma_{ro}^{ou} \varphi^{-1}(\alpha)) = 1 \tag{33}$$

$$\sum_r u_r y_{rj}^k - \sum_i v_i x_{ij}^k + d_j^k = 0 \quad \forall j \neq o, k, \tag{34}$$

$$\sum_r u_r (y_{rj}^k - \sigma_{ro}^{ou} \varphi^{-1}(\alpha)) - \sum_i v_i (x_{ij}^k + \sigma_{io}^{in} \varphi^{-1}(\alpha)) + d_j^k = 0 \quad \forall j = o, k. \tag{35}$$

In the second type of uncertainty, the objective function of the problem is estimated by probability interval with the $(1 - \beta)$ percentage confidence in which, β denotes the risk level. For instance, this interval shows that if sampling is repeated 100 times, a total of $(1 - \beta)$ objective functions will fall within this confidence interval (Hespanhol et al. 2019). To do so, Table 1 is suggested for sampling.

At each iteration, the mathematical model with a goal corresponding to the first column of Table 1 is solved, and then, its result is used as a basis to obtain the values of other goals. For instance, $f_{11}(x)$ represents the value of the first objective function where the mathematical model is solved

based on the objective function $f_1(x)$. Here, $f_{12}(x)$ shows the value of the second objective function where the mathematical model is solved based on the objective function $f_1(x)$, and so on. Accordingly, a total of n samples are obtained for each objective function. The last two rows of Table 1 represent the average and standard deviation of the objective functions, respectively. The confidence interval needs the value of the probability based on the related distribution. In this regard, four conditions should be considered where (i) if the distribution function is normal, the standard normal distribution table can be utilized to obtain the probabilities, (ii) if the sample size is greater than 30, the central limit theorem can be used to estimate the probabilities based on the standard normal distribution, (iii) the t -distribution can be employed when the distribution function is assumed¹ to be normal but the sample size is small, and (iv) if the distribution is unknown and the sample size is small, the Jackknife and Bootstrap method can be used to estimate the mean and variance where, both Jackknife and Bootstrap methods are known as resampling methods in statistics (Afanador et al. 2014). Constraints (36) and (37) show the confidence intervals for a general objective function l , under Conditions (ii) and (iii):

$$\bar{f}_l(x) - \frac{\sigma_l^f}{\sqrt{n}} \left(z_{\frac{\beta}{2}} \right) \leq f_l(x) \leq \bar{f}_l(x) + \frac{\sigma_l^f}{\sqrt{n}} \left(z_{\frac{\beta}{2}} \right), \tag{36}$$

$$\bar{f}_l(x) - \frac{S_l^f}{\sqrt{n}} \left(t_{\frac{\beta}{2}, n-1} \right) \leq f_l(x) \leq \bar{f}_l(x) + \frac{S_l^f}{\sqrt{n}} \left(t_{\frac{\beta}{2}, n-1} \right), \tag{37}$$

where S denotes the sample standard deviation and $z_{\beta/2}$ and $t_{\beta/2, n-1}$ represent the probability values based on the standard normal and the t -student distributions, respectively. Moreover, σ_l^f represents the standard deviation of Objective function (1). These equations are also applicable to the revised multi-choice GP. In this case, p_l is replaced with $f_l(x)$ in Constraints (36) and (37), and g_l^{\min} and g_l^{\max} are defined differently. For example, according to Constraint (37), $g_l^{\min} = \bar{f}_l(x) - \frac{S_l^f}{\sqrt{n}} (t_{\beta/2, n-1})$ and $g_l^{\max} = \bar{f}_l(x) + \frac{S_l^f}{\sqrt{n}} (t_{\beta/2, n-1})$.

In summary, this section strengthened the DEA model through multi-objective programming and consideration of uncertainty conditions. The DEA model is one of the unique methods in decision-making that performs both data analysis and optimization. The development of a multi-objective DEA model strengthens the power of discrimination in the evaluation of DMUs. Moreover, the real world is full of uncertainty. For this reason, data variations in the DEA model are taken

¹ The t -distribution is used when the sample size is small and/or the population standard deviation is unknown, providing more accurate estimates and accounting for variability (Ahsanullah et al. 2014).

into account in order to upgrade the DEA model to a stochastic model. It is difficult for the decision-maker to solve the multi-objective model deterministically. The decision-maker has a big challenge in determining the deterministic value for the aspiration levels of the objective functions. Therefore, the aspiration levels are incorporated into the multi-objective DEA as intervals. Thus, the focus of this section was to strengthen the DEA model in order to consider uncertainty conditions.

Results

In this section, the proposed models are applied to the database² of Eurostat (2019) and the performance of the mentioned 12 European countries is evaluated in terms of transportation sustainability and safety in which, the indices PET, FECRT, PKRA, PRTNT, and TTGNR are considered as the inputs and outputs of the models.³ The deterministic conditions are considered based on Models (1) and (2) and the uncertainty conditions are determined using both types of uncertainty for Models (1) and (2). The deterministic model is solved based on the model of Chang (2008) given in Formulas (22), (23), (24), (25), (26), and (27) in which the goal levels for objectives are obtained in accordance with Table 1 and the uncertainty model is solved based on the second type of uncertainty. The values of α and $(1 - \beta)$ are set to 95%.

The standard deviation of inputs in the first type of uncertainty is assumed⁴ to be equal to the average standard deviation of all inputs for all years. The same procedure is followed for outputs. The ranking of the countries is obtained based on the proposed models and the results are presented in Table 2. Since our proposed models are all linear, the results reported in Table 2 are both feasible and globally optimal. The arithmetic mean is regarded to aggregate the efficiency scores of all the years obtained from Model (1) and to determine the final ranking of the countries, which is shown in the second and fourth columns of this table.

The results demonstrate that the countries LA, LI, SL, and CR have achieved the best ranks based on the arithmetic mean,

Table 2 Ranking of the countries

Country	Model (1)-deterministic	Model (2)-deterministic	Model (1)-uncertain	Model (2)-uncertain
GE	12	12	9	8
BU	5	4	6	6
FR	10	11	11	11
CR	3	3	3	3
IT	8	9	12	12
LA	1	2	1	2
LI	1	1	4	5
PO	11	10	8	9
SL	4	5	2	1
FI	7	7	7	7
SP	9	8	10	10
RO	6	6	5	4

although SL has more variability than CR. It is important to note that CR and FI achieved identical ranks in all the extended models. Therefore, the ranks of these countries are a good basis for comparison because they have stable results in all models and under different conditions. Figures 2 and 3 display the changes in the efficiency scores of the countries over a five-year period in which, Fig. 2 depicts the results obtained from Model (1) under deterministic conditions, and Fig. 3 shows the results obtained from Model (1) under the first type of uncertainty. The first, second, third, fourth and fifth years are marked in blue, orange, gray, yellow, and dark blue, respectively. According to these figures, the efficiency scores of LA and LI in the deterministic condition and the efficiency score of GE in the uncertainty condition have remained unchanged during different years.

The normalized efficiency values of countries over a 5-year period obtained from Model (2) under deterministic status and the first type of uncertainty are represented in Fig. 4. Based on the results, the countries SL, GE, PO, and RO have relatively better efficiency scores under the uncertainty condition.

Managerial analysis

This section is focused on the analysis of important issues such as productivity, resilience, rank, local best practice and benchmark. These issues will be effective in the decision-making of transportation and supply chain managers. In this regard, in “Analysis of the productivity,” the productivity of the countries is obtained from the aspect of sustainability by means of countries’ efficiency values in consecutive time periods. The countries’ efficiency scores are investigated considering disruptions in “Analysis of the resilience.” The ranking of the countries, considering the productivity and resilience indices, is analyzed in “Analysis of the ranks.” “Analysis of the best practices”

² For further information, please refer to (<https://ec.europa.eu/eurostat/web/transport/data/database>) and Babaei et al. (2022b).

³ Our indicators are derived from authoritative articles, such as Babaei et al. (2022b), Omrani et al. (2020), Shen et al. (2012, 2013, 2015), Wu et al. (2016), Chang et al. (2013), Ignatius et al. (2016), and Chen et al. (2020), which are widely used to evaluate the sustainability of the transportation sector.

⁴ Making this assumption simplifies modeling uncertain conditions and reduces computational complexity. This assumption also considers the performance of all years as an average, rather than focusing on a single year. Additionally, this assumption ensures a relative comparison between DMUs in DEA models (Thanassoulis et al. 1996), as all DMUs are measured under the same uncertainty criterion.

Fig. 2 Efficiency changes according to Model (1) under deterministic status

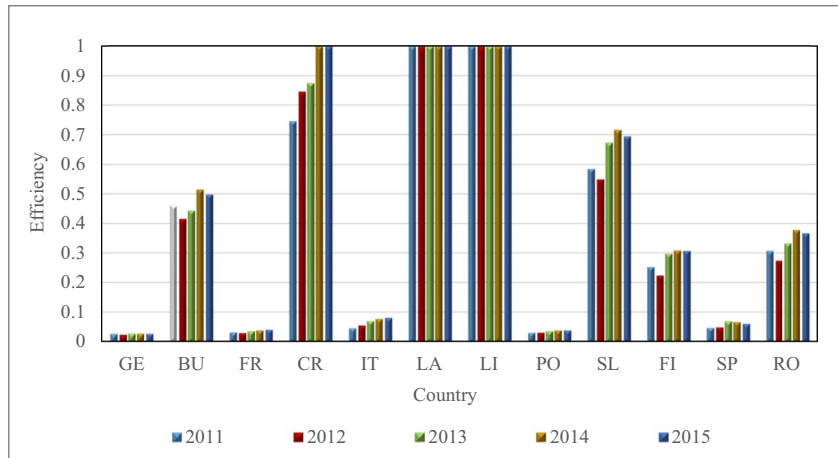


Fig. 3 Efficiency changes according to Model (1) and under the first type of uncertainty

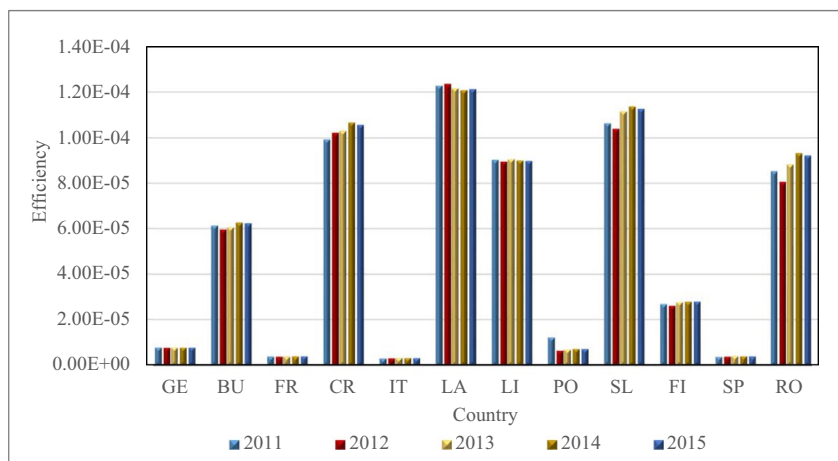
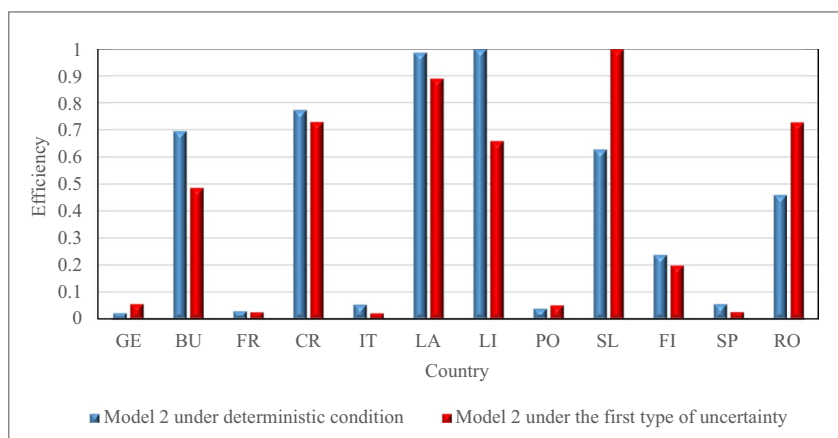


Fig. 4 Comparing the efficiency scores of Model (2) under the deterministic status and first type of uncertainty

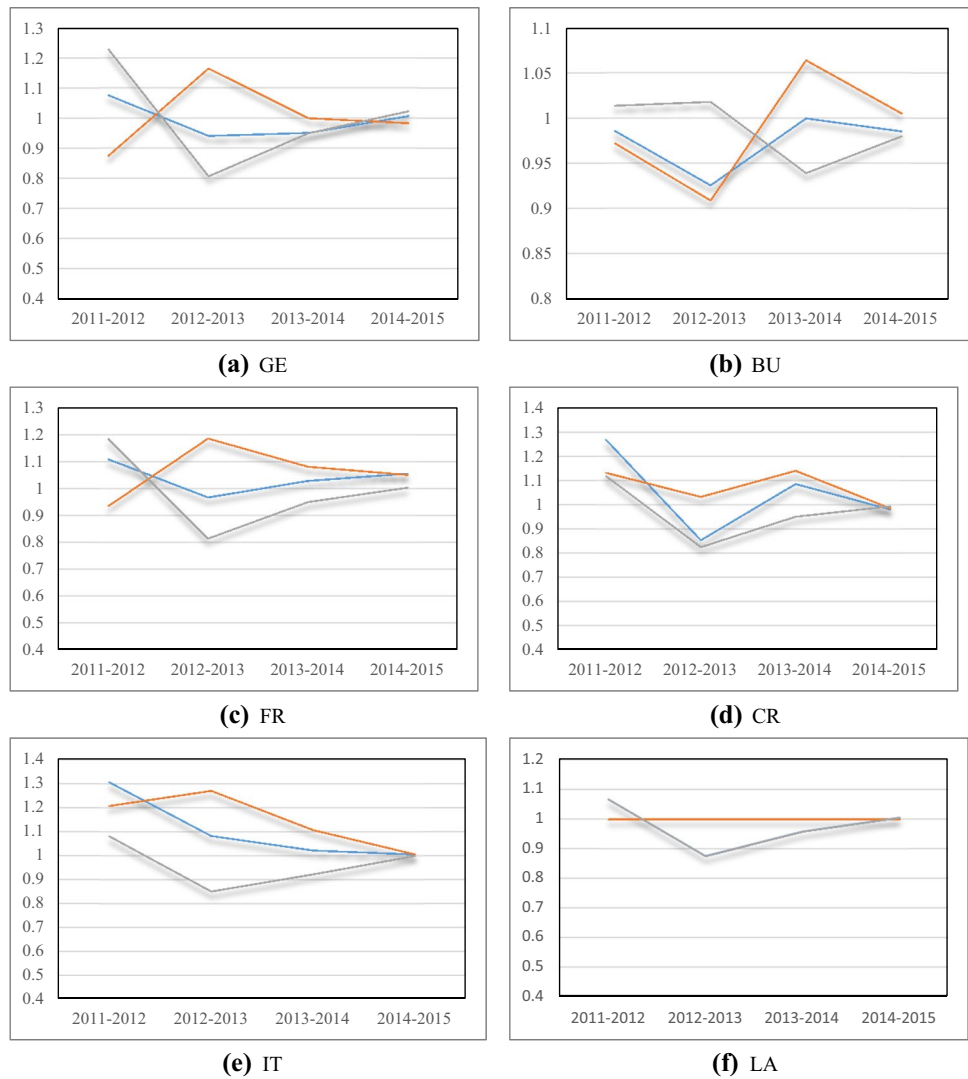


explains the best practices as a set of sustainable countries. “Analysis of the benchmarks” determines the countries which can be considered as benchmarks for other countries in order to improve their sustainability. Finally, “Analysis of the methods” compares two models which are introduced for modeling through the regression method.

Analysis of the productivity

There are various methods to quantify productivity over different time periods. One such measure, initially introduced in consumer theory and subsequently applied to productivity analysis, is the Malmquist index. This index offers several

Fig. 5 Malmquist index, efficiency change, and technical change in different countries



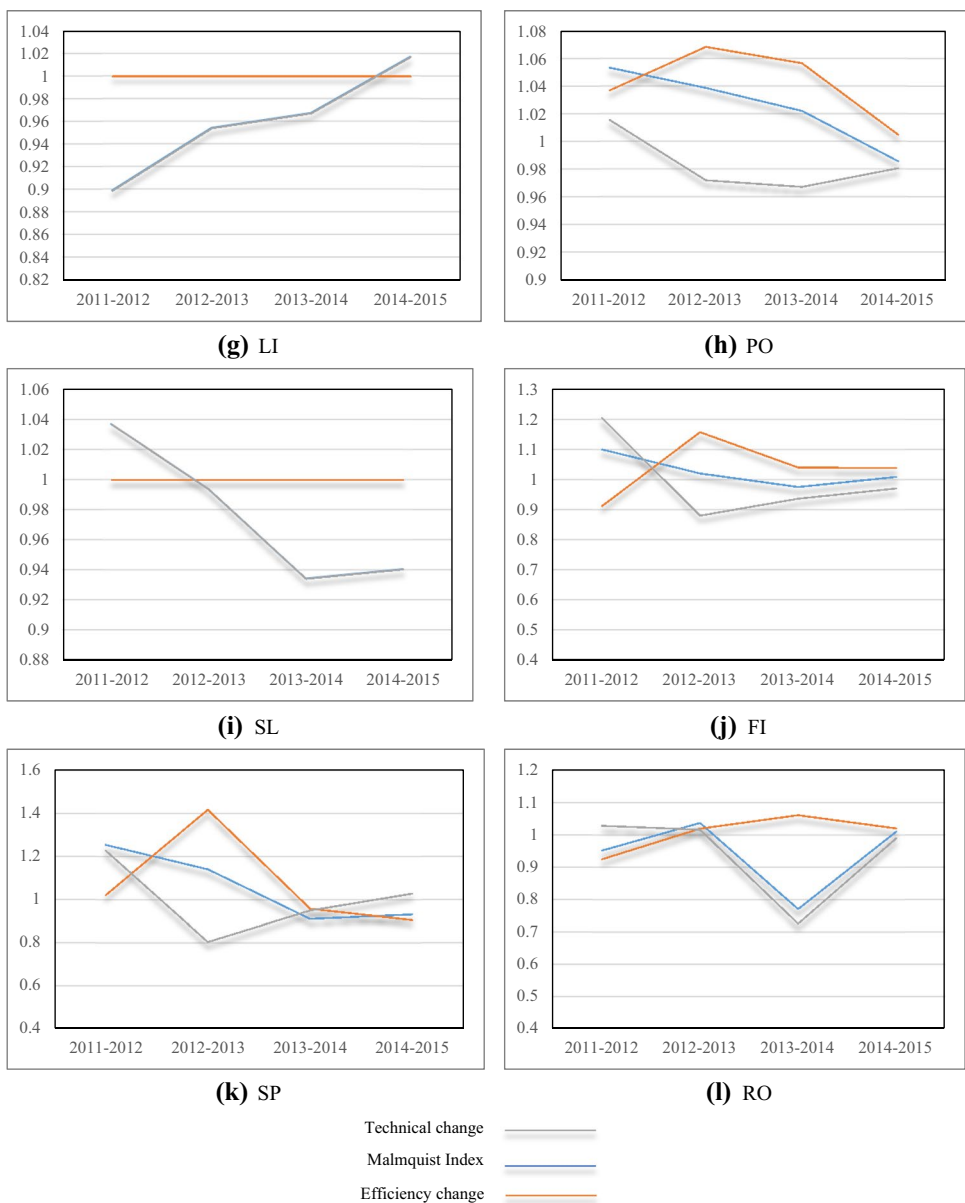
benefits, including the absence of any assumptions about the economic conduct of production units, such as cost minimization or revenue maximization (Odeck 2006). The Malmquist index is utilized to measure productivity changes over time (Zhu 2009). This index is presented in Eq. (38):

$$M_{o,t,t+1} = \left(\frac{\theta_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{\theta_o^t(x_o^t, y_o^t)} \right) \left(\frac{\theta_o^t(x_o^t, y_o^t)}{\theta_o^{t+1}(x_o^{t+1}, y_o^{t+1})} \times \frac{\theta_o^t(x_o^{t+1}, y_o^{t+1})}{\theta_o^{t+1}(x_o^t, y_o^t)} \right)^{\frac{1}{2}}, \tag{38}$$

where $M_{o,t,t+1}$ denotes the Malmquist index for the country input and output during $(t, t + 1)$. If the index is greater than one, it means that the productivity has been improved; if the index is equal to one, it means that the productivity has remained unchanged; and if the index is less than one, it means that the productivity has declined. The first and the second brackets show the efficiency changes (EC) and

technical changes (TC), respectively, θ_o^t stands for the efficiency score of country O in period t . The inputs and outputs of the periods t and $(t + 1)$ are used to calculate $\theta_o^{t+1}(x_o^t, y_o^t)$ and $\theta_o^t(x_o^{t+1}, y_o^{t+1})$, respectively. The efficiencies $\theta_o^{t+1}(x_o^t, y_o^t)$ and $\theta_o^t(x_o^{t+1}, y_o^{t+1})$ for other countries, however, are calculated by inputs and outputs of periods $(t + 1)$ and t , respectively. The Malmquist Index components are obtained from the model presented in Eqs. (1), (2), (3), and (4). Figure 5 represents the values of $M_{o,t,t+1}$, TC, and EC for different time periods and countries. The values of $M_{o,t,t+1}$ and TC are equal for countries SL, LA, and LI. Based on the results, the most negative and positive productivity changes have occurred in CR (2011–2013) and RO (2013–2015), respectively. The maximum positive and negative changes in EC are related to SP (2011–2013) and SP (2012–2014), respectively, and the maximum positive and negative changes in TC are related to RO (2013–2015) and SP (2011–2013).

Fig. 5 (continued)



Analysis of the resilience

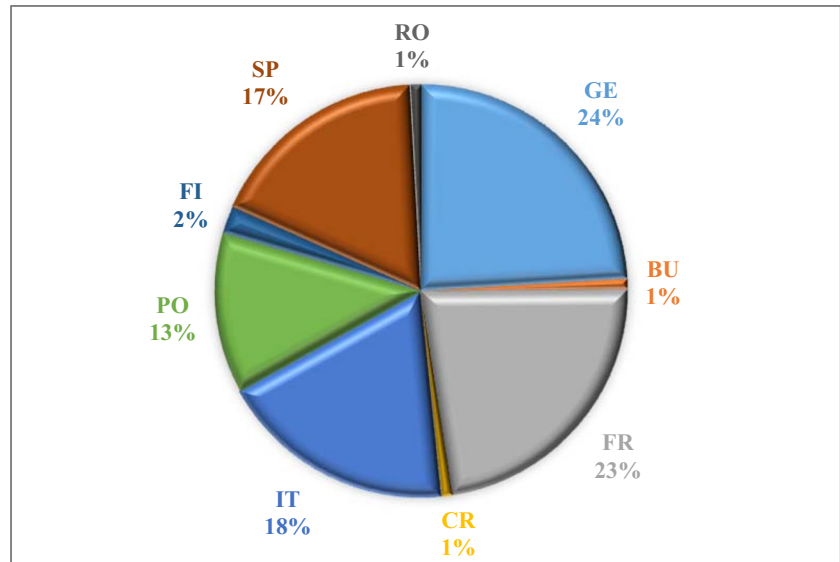
Resilience, in the transportation system, is defined as the ability to retain the performance of the system at the time of disruption. In this way, the retention ability is considered to be equal to the difference between the most and the least desirable performances in which, the greater values of the difference mean higher resilience of the system. The most desirable performance value is obtained when the country’s input and output have the highest output and the lowest input while, the other countries have the highest input and the lowest output (Hadi-Vencheh et al. 2015). The resilience level is obtained from the model presented in Formulas (1), (2), (3), and (4) based on the average of the input and output data in which, the results are displayed in Fig. 6. Based on the

results, the difference between the most and the least desirable performances of the three countries including LA, LI, and SL is zero and it means that these three countries are the most resilient. On the other hand, GE, FR, and IT cover the greatest surface area in the figure which shows a greater difference between the most and the least desirable performances, and thus, these countries are the least resilient ones.

Analysis of the ranks

Once the problem is solved and the results are analyzed, the decision-maker should prioritize the alternatives and select the best one. Simple additive weighting (SAW) is a decision-making method which compares some alternatives based on some criteria and assigns a score to each alternative.

Fig. 6 Resilience level of the countries



Considering the different dimensions of the criteria, the related scores should first be normalized. Then, the weighted sum of the score of each alternative in all criteria (with respect to the weight of the criteria) determines the final score of the alternatives (Seyedmohammadi et al. 2018). In this study, two criteria with 12 alternatives are considered for two decision-making problems. In the first decision-making problem, the first and second criteria are the ranking of the countries and the resilience, respectively, and the 12 countries are considered as alternatives. In the second decision-making problem, the first criterion is the ranking of the countries while the second criterion is obtained based on the Malmquist index.

The score assigned to each country on the basis of this criterion is determined by the number of times each country

has been productive. The second criterion in both decision-making problems is considered deterministic while, the first criterion, according to Table 1, is considered to be interval-based. In this case, an interval between the best and the worst rankings is considered for the first criterion as follows:

$$FS_j = \underline{R}_j + \gamma(\bar{R}_j - \underline{R}_j) \forall j, \tag{39}$$

$$0 \leq \gamma \leq 1, \tag{40}$$

where the lowest and highest ranks are denoted by \underline{R}_j and \bar{R}_j , respectively, and FS_j stands for the final score of each country in the first criterion. In this research, two decision-making problems mentioned above are solved for different values of γ including 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1. At each iteration, the weights of the first and second criteria are assumed⁵ to be (0.1, 0.9), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25), and (0.9, 1). In the next step, the arithmetic mean of the score of each country is calculated based on all γ s. The final ranking of the countries is determined based on this score for the first and second decision-making problems in which the results are reported in Tables 3 and 4, respectively.

The changes in the mean scores of the alternatives due to the changes in the weights are illustrated in Figs. 7 and 8 concerning the first and second decision-making problems, respectively. According to Fig. 7, a decrease in the significance of the second criterion resulted in a decrease in the final scores of all countries, except for GE. In addition, as

Table 3 Ranking in the first decision-making problem

Country	(γ_1, γ_2)				
	(0.1,0.9)	(0.25,0.75)	(0.5,0.5)	(0.75,0.25)	(0.9,0.1)
GE	12	12	12	11	11
BU	5	5	5	5	5
FR	11	11	11	12	12
CR	4	4	4	4	4
IT	10	10	10	10	10
LA	1	1	1	1	1
LI	2	2	2	2	2
PO	8	8	8	8	9
SL	2	2	2	2	2
FI	7	7	7	7	7
SP	9	9	9	9	8
RO	6	6	6	6	6

⁵ Based on this, a comprehensive investigation is conducted to determine the importance of each criterion in comparison to another criterion.

Table 4 Ranking in the second decision-making problem

Country	(γ_1, γ_2)				
	(0.1,0.9)	(0.25,0.75)	(0.5,0.5)	(0.75,0.2)	(0.9,0.1)
GE	9	4	12	12	11
BU	12	1	9	6	6
FR	4	8	10	11	12
CR	6	7	2	2	2
IT	1	12	4	8	10
LA	5	9	1	1	1
LI	10	2	6	3	3
PO	3	10	8	9	9
SL	10	2	6	3	3
FI	2	11	3	7	7
SP	8	5	11	10	8
RO	7	6	5	5	5

Fig. 8 depicts, any decrease in the significance of the second criterion leads to improvements in the final score of BU, CR, LA, LI, SL, and RO.

Analysis of the best practices

Recognition of the best practices (BPs) is one of the most important managerial problems. However, planning to achieve directly to the BPs is a difficult task. In such cases, partitioning can be used to detect the local BPs (LBPs) and global BPs (GBPs). In order to improve their status, LBPs can first achieve the higher-level (more efficient) LBPs and then GBPs. To this end, applying the model presented in Formulas (1), (2), (3), and (4) and taking into account the expected value of the data, GBPs and LBPs are achieved and presented in Fig. 9. The model is

Fig. 7 Changes in the scores of the first decision-making problem

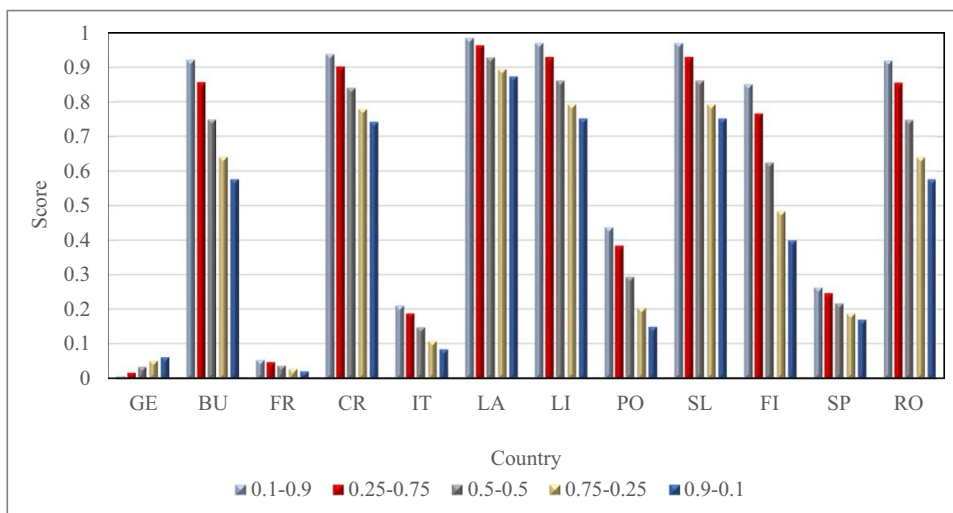
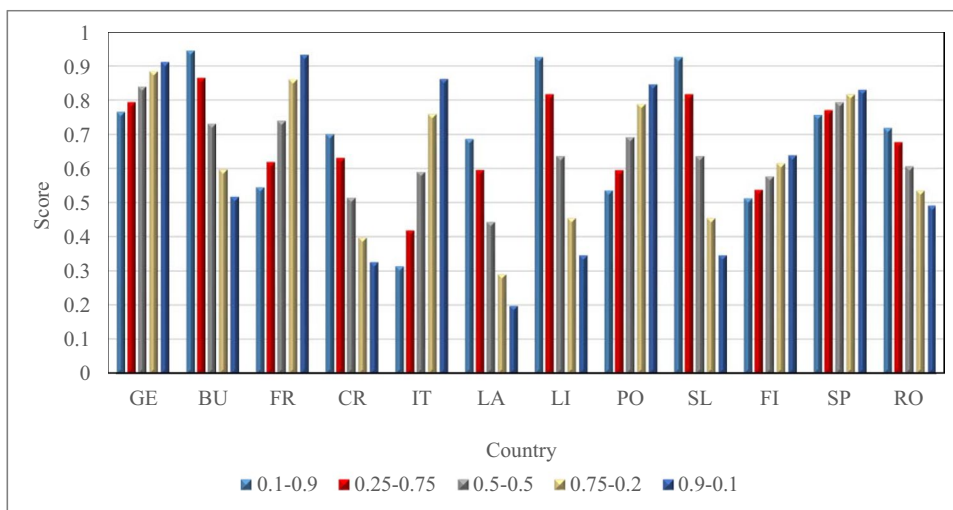


Fig. 8 Changes in the scores of the second decision-making problem



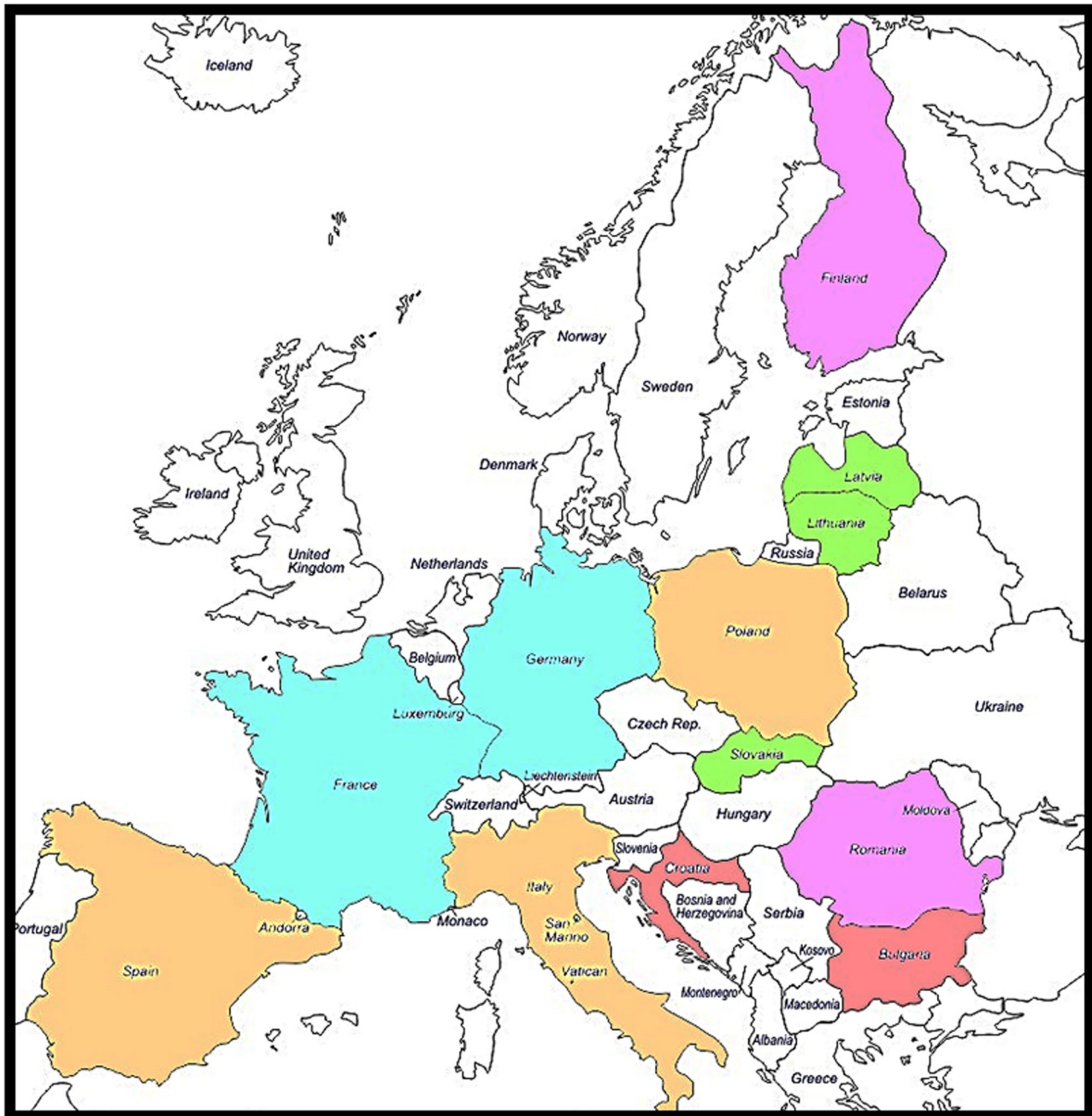


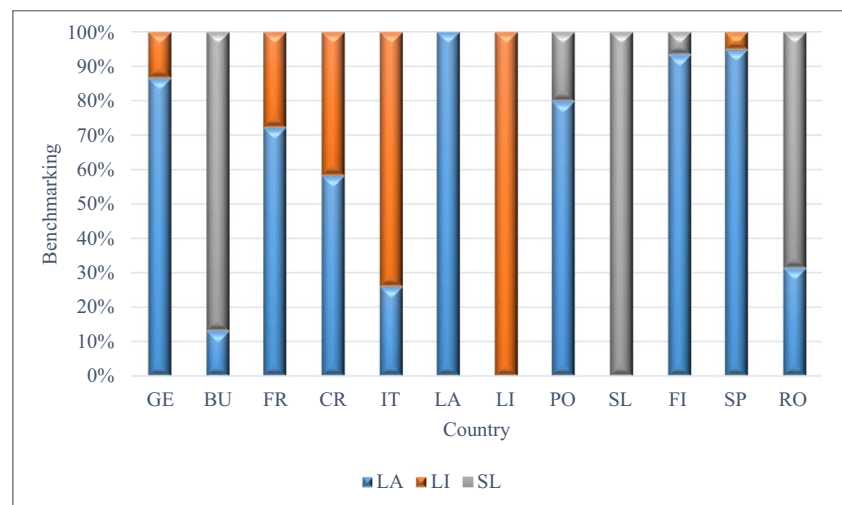
Fig. 9 Local and global best practices

first solved for all the countries and then, BPs are detected and omitted. Then, the model is solved for the remaining countries this procedure is repeated and accordingly, all BPs are detected at different stages (Seiford and Zhu 2003). In Fig. 9, the colors of green, red, purple, yellow and blue represent the different levels of BPs in which the green shows GBPs and the other colors represent LBPs at each stage. It must be noted that countries with lower

efficiency scores can adopt policies based on this subsection to improve their efficiency score in the field of transportation sustainability step by step.

Higher stages cover less efficient countries. Based on the results, the countries LI, SL and LA are GBPs, the first-level LBPs include BU and CR, the second-level LBPs include FI and RO, the third-level LBPs include PO, IT and SP, and finally, the fourth-level LBPs include FR and GE.

Fig. 10 Benchmarking percentage of different countries from LA, LI, and SL



Analysis of the benchmarks

Applying the dual model in DEA leads to the detection of benchmark units for inefficient units (Emrouznejad and Amin 2009). The average changes in the PET, FECRT, and PKRA indices which are required for countries to reach the efficient frontier with respect to the benchmark countries are presented in Fig. 10. This figure indicates the benchmarking level of different countries from LA, LI, and SL. Based on the results, LA is the country that is considered as the benchmark by most countries. For example, GE should consider LA and LI as benchmark countries while more attention should be paid to LA. Generally, LA, LI, and SL receive higher weights as the benchmark countries, respectively. This subsection encourages policymakers in the field of transportation who seek sustainable development to improve their efficiency score by taking patterns from sustainable countries.

Analysis of the methods

In this section, the linear regression technique is utilized to compare Model (1) and Model (2), under identical conditions. In the regression model, the efficiency score is considered as the dependent variable while each output of the DEA model is considered as the continuous independent variable, and then, the coefficient of determination is employed as an index to perform the comparisons. In this regard, five models including exponential (Exp), linear (Li), logarithmic (Lo), second-order polynomial (Pol), and power (Pow) are used to perform the regression modeling. The geometric mean of each method is presented with *Geo-Mean*. The results are summarized in Table 5 in which, the results demonstrate that the best estimation is obtained from the Pol method. The final geometric mean obtained from the best method is presented in the *Final Geo-Mean* row.

As Table 5 displays, the integrated model provides better interpretations. Furthermore, the results obtained from implementing the regression method on the efficiency scores of Model (1) and Model (2) are presented in Table 6. Depending on the year under study, Model (1) provides different results. Hence, in order to compare the results of Model (1) with those of Model (2), the average (Avg), maximum (Max), and minimum (Min) values of the results of Model (1) are, separately, taken into account. As Table 6 depicts, there is a very strong interpretation between the two efficiency scores. The best estimates were obtained from the Pow model. High interpretability shows that our models are trustworthy and transportation stakeholders can make decisions based on their results.

Discussion

The movement of goods and passengers is one of the important factors of prosperity in transportation activities. However, the growth in transportation activities results in emissions, road accidents, and energy consumption. Facilitation in the movement of goods and passengers leads to economic growth. Therefore, the transportation of passengers and freight is effective in the ability of businesses to access new markets, create jobs, reduce trade barriers, increase competition advantages, and plan cost reduction (regarding fuel, maintenance, and consolidation of shipments). For this reason, this research accounted for PRTNT and TTGNR as helpful criteria on the economic aspect. On the other hand, energy consumption and emissions affect the environmental aspect. Safety as a protector of individuals and communities has significant effects on the social aspect. The safety measure includes accidents that involve human lives. To do so, PET, FECRT, and PKRA were taken into account along with economic criteria. The mentioned criteria, as stated in “[Problem statement](#)” according to

Table 5 Regression results of the different methods

Year	Method	Model (1)					Model (2)				
		Exp	Li	Lo	Pol	Pow	Exp	Li	Lo	Pol	Pow
2011	Output 1	0.84	0.74	0.23	0.89	0.31	0.84	0.64	0.19	0.80	0.33
	Output 2	0.11	0.10	0.01	0.57	0.03	0.12	0.09	0.001	0.64	0.02
	Output 3	0.86	0.79	0.45	0.80	0.52	0.84	0.72	0.39	0.75	0.49
	Geo-Mean	0.42992	0.38811	0.10115	0.74037	0.16911	0.43912	0.34614	0.04200	0.72685	0.14788
2012	Output 1	0.81	0.63	0.19	0.80	0.31	0.82	0.62	0.19	0.80	0.32
	Output 2	0.14	0.10	0.004	0.52	0.04	0.15	0.11	0.008	0.52	0.05
	Output 3	0.82	0.72	0.42	0.74	0.47	0.82	0.71	0.42	0.73	0.48
	Geo-Mean	0.45305	0.35664	0.06834	0.67521	0.17996	0.46549	0.36449	0.08611	0.67216	0.19730
2013	Output 1	0.80	0.62	0.18	0.79	0.31	0.81	0.61	0.18	0.79	0.31
	Output 2	0.18	0.15	0.01	0.49	0.06	0.22	0.17	0.02	0.51	0.07
	Output 3	0.84	0.73	0.47	0.77	0.54	0.82	0.71	0.46	0.74	0.54
	Geo-Mean	0.49455	0.40795	0.09458	0.66799	0.21576	0.52671	0.41913	0.11831	0.66805	0.22714
2014	Output 1	0.79	0.63	0.19	0.80	0.32	0.81	0.61	0.18	0.79	0.31
	Output 2	0.20	0.15	0.16	0.43	0.07	0.23	0.19	0.02	0.46	0.08
	Output 3	0.83	0.73	0.50	0.76	0.64	0.82	0.71	0.49	0.73	0.65
	Geo-Mean	0.50806	0.41013	0.24771	0.63943	0.24293	0.53458	0.43496	0.12083	0.64254	0.25261
2015	Output 1	0.80	0.61	0.18	0.78	0.33	0.82	0.62	0.18	0.79	0.32
	Output 2	0.21	0.15	0.02	0.39	0.08	0.26	0.21	0.03	0.47	0.09
	Output 3	0.81	0.71	0.43	0.75	0.60	0.82	0.71	0.48	0.74	0.62
	Geo-Mean	0.51436	0.40200	0.11568	0.61105	0.25114	0.55916	0.45216	0.13737	0.65011	0.26137
Final Geo-Mean		0.66544					0.67131				

Values in bold indicate the best results

Table 6 Regression results for the efficiency

Model (1)	Model (2)				
	Exp	Li	Lo	Pol	Pow
Avg	0.894	0.993	0.884	0.994	0.998
Max	0.920	0.992	0.870	0.998	0.999
Min	0.850	0.984	0.900	0.986	0.994
Geo-mean	0.88753	0.98966	0.88458	0.99265	0.99699

Values in bold indicate the best results

the research literature, are effective on sustainability. Accordingly, it is necessary to evaluate the transportation systems of countries from the perspective of sustainability. In this way, policymakers identify and develop strategies to improve the performance of their country's transportation system.

In the real world, various institutions evaluate transportation from a sustainability perspective. For example, the International Transport Forum (ITF 2023) and United States Department of Transportation (DOT 2023) are aimed at creating a deep understanding of the role of transportation (both in freight and passengers) in strengthening the pillars of sustainability (social, environmental and economic). In this regard, this research presented an efficient DSF for assessing the transportation of

countries from the aspect of sustainability. This proposed DSF contains two developed DEA models to proceed with robust evaluations in terms of both data analysis and optimization. In order to strengthen the discriminating power of the DEA model, multi-objective DEA models were utilized. In addition, this study developed a model that evaluates countries simultaneously in all years to create an integrated assessment for transportation systems. uncertain conditions are considered in the data concerning the transportation to approach the real world. Human error (in the collection, interpretation and recording), lack or defect of information, bias in data collection, equipment and method used in data collection, environmental factors (i.e., weather, traffic, and roadblocks), and time limit (hasty data collection) are among the causes of uncertainty that have a significant impact on the data of inputs and outputs of road transport. It is important to note that there are significant reasons supporting the structure of our proposed approach. One weakness of the DEA model is its inability to distinguish between DMUs effectively. Therefore, our approach is based on multi-objective models that increase the power of differentiation in evaluating DMUs. Additionally, our proposed approach evaluates all years effectively, not only independently but also in an integrated manner. Since the real world is full of uncertainty, the data used in our proposed approach is considered uncertain. This uncertainty includes both the data of the DEA model (inputs and outputs) and the

solution of the multi-objective model. It is worth noting that our proposed approach is based on sustainability assessment, which is an important issue in both research and the real world.

The research framework (cf. Figure 1) illustrates our approach, which involves developing multi-objective and network evaluation models that can account for uncertainty. We have assessed uncertainty in both the input and output data of the evaluation models, as well as in the determination of the targets of the objective functions (cf. Table 1). Using a case study of several European countries, we ranked the countries under multi-objective and network models, both under conditions of uncertainty and certainty (cf. Table 2). We then conducted a sensitivity analysis to examine the performance changes according to the presented models and uncertainty conditions (as shown in Figs. 2, 3, and 4). Our article includes various applied analyses, such as productivity analysis based on efficiency in different periods (cf. Figure 5), resilience analysis by examining the highest and lowest limits of efficiency (cf. Figure 6), sensitivity analysis of ranks from the perspective of productivity and resilience (cf. Tables 3 and 4 and Figs. 7 and 8), introducing improvement benchmarks and models for countries with lower efficiency, and modeling countries with lower efficiency based on countries with better efficiency (as shown in Figs. 9 and 10). Finally, we also included the interpretability of models through regression (cf. Tables 5 and 6), which are important for effective decision-making.

The proposed DSF was implemented on the data of 12 European countries. This research framework has the ability to provide many analytical opportunities regarding ranking, productivity, benchmarking and resilience. According to Table 2, LI has the best rank in the multi-objective and network model under deterministic conditions. However, when uncertainty is included in the models, LA and SL get the best ranks. Therefore, it is effective to consider uncertainty and types of evaluation models (multi-objective model or network model) in the ranking result. Comparison through efficiency scores is not enough to evaluate countries. For example, in Fig. 5a, between the two periods of (2011–2012) and (2012–2013), efficiency change has increased, but productivity change has decreased. In Fig. 5l, in the same periods, efficiency and productivity changes were in the same direction (both have increased). In Fig. 5g, in the same periods, efficiency change has remained constant, while productivity change has occurred. Therefore, examining the transportation systems of countries only through efficiency may mislead policymakers and they may adopt policies that are not productive.

According to Fig. 6, LA, LI, and SL have the same efficiency scores in the worst and best cases. Therefore, it is appropriate for unsustainable countries to follow these countries in terms of performance, because their performance scores are resilient under the worst and best conditions. In this regard, as shown in Fig. 10, inefficient countries can adopt policies based on efficient countries. Modeling inefficient countries from countries with a solid foundation in the field of transportation can provide a robust improvement plan for inefficient countries. For

example, as outlined in Tables 3 and 4, LA, RO, and CR have the least variations under various importance weights (regarding rank, resilience, and productivity). It should be noted that the developed DSF provides reliable results to policymakers in the field of transportation because the proposed DSF performs the evaluation process through two models. This prevents the error of relying on a single model. In addition, if these two model results support each other, then the evaluation results obtained from the DSF become more reliable. In this regard, both Model (1) and Model (2) were implemented on the same data and both models aim to evaluate the transportation of countries from a sustainability perspective. The high interpretability of the efficiency scores obtained from these two models (based on Table 6) demonstrates that their obtained results support and validate each other. Transportation policymakers should be aware that despite the many benefits of this proposed DSF, the evaluation process in this DSF, as well as other research, faces challenges. All in all, the important challenges for assessing the transportation of countries from a sustainability perspective are the availability and measurement of data.

Conclusions and outlook

Transportation is one of the most important factors in the economic development of cities. Moreover, urban road transportation accounts for a considerable share of transportation in all countries, and therefore, it plays a vital role in logistics planning. In this regard, the environmental and safety factors are the major challenges of the countries that are affected by transportation. Therefore, evaluating the performance of the countries in terms of transportation sustainability is essential. In a word, an efficient DSF based on two DEA models was proposed to evaluate the sustainability of different countries. The first model dealt with the performance of countries in different years while the second model generally addressed the performance of countries during all the years as a network model. Since both models were multi-objective, a multi-objective programming model based on a revised multi-choice GP approach was developed to treat the models under uncertain conditions. In addition to uncertainty conditions in multi-objective programming, uncertainty in data was also considered in the proposed models. The proposed models were applied to the real data set related to some European countries. In this regard, the models were then solved and the ranking of the countries was obtained in which, the results under deterministic conditions revealed that the best rank belongs to LI while the worst rank belongs to GE. After LI, the countries of LA, CR, and SL achieved the best ranks, respectively. Furthermore, the obtained results varied based on uncertain conditions. Therefore, considering uncertain conditions along with deterministic conditions increased the decision-making accuracy of transport decision-makers. A set of managerial analyses were

also performed to assess the performance of the countries. The Malmquist index was used to analyze the productivity, EC, and TC of countries over the years under consideration. The resilience of the countries was also measured with regard to the efficiency changes under the best and worst conditions. Taking into account the ranking of the countries obtained from the proposed models and the potential effects of resilience and productivity, a sensitivity analysis was carried out on the ranking of the countries. The benchmarks for inefficient countries were determined based on the efficient countries. On the other hand, local and global best practices were determined for countries in order to enable them to improve their transportation system, step by step. Finally, the two models were compared to reveal their data interpretation potentials. The findings from managerial analyses revealed that most changes belonged to RO within the 2013–2015 time period in order to improve the productivity while most changes belonged to CR within the 2011–2013 time period to reduce the productivity. Moreover, LA, LI, and SL were introduced as the most resilient countries and as global best practices. Finally, the effectiveness of the proposed models was evaluated by comparing their efficiencies through the regression method. Sometimes determining policies for inefficient countries to reach the most efficient countries requires a lot of resource consumption. To avoid excessive consumption of resources, our DSF also provided step-by-step improvement of countries' efficiency. As shown in Fig. 9, there is no need for GE to follow LI, but in the first stage, it can reach the efficiency of countries like IT.

This work evaluated the transportation sustainability of countries based on five criteria, taking into account limitations in resources and data. It focused on European countries that have available and published data in predetermined periods. However, policymakers, such as the European Union, could benefit from more comprehensive results by considering additional effective criteria and a larger number of countries. While accounting for uncertainty and multi-objective models ensures the realism and discrimination power of the presented models, it also increases the complexity of data collection and mathematical problem-solving.

The inputs of a given period may affect not only the outputs of this period but also the outputs of other periods. Hence, evaluation of the potential effects of different periods on each other is recommended for future research. Analysis of the ranking of the countries using indices other than resilience and productivity such as agility and robustness can be taken into account to extend the current study. Moreover, big data analytics (BDA) can be used to gather required data and solve the models more efficiently. On the other hand, sustainability is a global issue, it is recommended that researchers evaluate Middle East countries from a sustainability perspective. Other multi-objective solution methods (e.g., augmented ϵ -constraint method) as well as other uncertainty-handling techniques (e.g., robust optimization) may be implemented and compared to the proposed ones in this work. Finally, advanced optimization algorithms

(e.g., hybrid heuristics and metaheuristics, adaptive algorithms and self-adaptive algorithms) can be utilized to treat the problem complexity efficiently (Dulebenets 2021, 2023; Pasha et al. 2022; Singh and Pillay 2022; Singh et al. 2022; Chen and Tan 2023).

Author contribution Data curation, conceptualization, investigation, software, methodology, and writing—original draft were done by A.B.; resources, supervision, and writing—review and editing were done by M.K.; conceptualization, supervision, and writing—review and editing were performed by M.R.A.J.; supervision, investigation, and writing—review and editing were done by E.B.T. Every author reviewed and gave their approval to the final manuscript.

Data availability Data was adopted from <https://ec.europa.eu/eurostat/data/database>. (Dataset, extracted on): (Pollutant emissions from transport [T2020_RK300], 21/09/2019); (Final energy consumption in road transport by type of fuel [TEN00127], 21/09/2019); National annual road transport by group of goods and type of transport [road_go_na_tgtt], 21/09/2019); (People killed in road accidents [SDG_11_40], 21/09/2019); (Passenger road transport on national territory [road_pa_mov], 21/09/2019).

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

- Afanador NL, Tran TN, Buydens LMC (2014) An assessment of the jackknife and bootstrap procedures on uncertainty estimation in the variable importance in the projection metric. *Chemom Intell Lab Syst* 137:162–172
- Ahsanullah M, Kibria BG, Shakil M (2014) Normal and Student's *t* distributions and their applications, vol 4. Atlantis Press, Paris
- Aldamak A, Zolfaghari S (2017) Review of efficiency ranking methods in data envelopment analysis. *Measurement: J Int Meas Confed* 106:161–172
- Alper D, Sinuany-Stern Z, Shinar D (2015) Evaluating the efficiency of local municipalities in providing traffic safety using the data envelopment analysis. *Accid Anal Prev* 78:39–50
- Babaei A, Khedmati M, Jokar MRA, Babaei Tirkolae E (2022a) Performance evaluation of omni-channel distribution network configurations considering green and transparent criteria under uncertainty. *Sustainability* 14(19):12607
- Babaei A, Khedmati M, Jokar MRA (2022b) A new model for evaluation of the passenger and freight transportation planning based on the sustainability and safety dimensions: a case study. *Process Integr Optim Sustain* 6(4):1201–1229
- Babaei A, Khedmati M, Jokar MRA, Tirkolae EB (2023a) Sustainable transportation planning with traffic congestion and uncertain conditions. *Expert Syst Appl* 227:119792
- Babaei A, Khedmati M, Jokar MRA, Tirkolae EB (2023b) An integrated decision support system to achieve sustainable development in transportation routes with traffic flow. *Environ Sci Pollut Res* 30(21):60367–60382
- Babaei A, Khedmati M, Jokar MRA (2023c) A branch and efficiency algorithm to design a sustainable two-echelon supply chain

- network considering traffic congestion and uncertainty. *Environ Sci Pollut Res* 30(10):28274–28304
- Babaei A, Khedmati M, Akbari Jokar MR (2023d) A new branch and efficiency algorithm for an optimal design of the supply chain network in view of resilience, inequity and traffic congestion. *Ann Oper Res* 321(1–2):49–78
- Babaei A, Khedmati M, Jokar MRA (2023e) A comprehensive blockchain-enabled supply chain network design: an iterative model versus an integrated model. *Ann Oper Res*. <https://doi.org/10.1007/s10479-023-05562-5>
- Babaei A, Khedmati M, Jokar MRA (2023f) A new model for production and distribution planning based on data envelopment analysis with respect to traffic congestion, blockchain technology and uncertain conditions. *Ann Oper Res*. <https://doi.org/10.1007/s10479-023-05349-8>
- Chang CT (2008) Revised multi-choice goal programming. *Appl Math Model* 32(12):2587–2595
- Chang YT, Zhang N, Danao D, Zhang N (2013) Environmental efficiency analysis of transportation system in China: a non-radial DEA approach. *Energy Policy* 58:277–283
- Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Oper Res* 2(6):429–444
- Chen Z, Wang W, Li F, Zhao W (2020) Congestion assessment for the Belt and Road countries considering carbon emission reduction. *J Clean Prod* 242:118405
- Chen H, Zhou R, Chen H, Lau A (2022) Static and dynamic resilience assessment for sustainable urban transportation systems: a case study of Xi'an, China. *J Clean Prod* 368:133237
- Chen M, Tan Y (2023) SF-FWA: a self-adaptive fast fireworks algorithm for effective large-scale optimization. *Swarm Evol Comput* 80:101314
- Cooper WW, Deng H, Huang Z, Li SX (2004) Chance constrained programming approaches to congestion in stochastic data envelopment analysis. *Eur J Oper Res* 155(2):487–501
- Dadashi A, Mirbaha B (2019) Prioritizing highway safety improvement projects: a Monte-Carlo based data envelopment analysis approach. *Accid Anal Prev* 123:387–395
- Daniels S, Martensen H, Schoeters A, Van den Berghe W, Papadimitriou E, Ziakopoulos A, ... Perez OM (2019) A systematic cost-benefit analysis of 29 road safety measures. *Accid Anal Prev* 133:105292
- Dehnohalaji A, Khezri S, Emrouznejad A (2022) A box-uncertainty in DEA: a robust performance measurement framework. *Expert Syst Appl* 187:115855
- Ding L, Yang Y, Wang W, Calin AC (2019) Regional carbon emission efficiency and its dynamic evolution in China: a novel cross efficiency-malmquist productivity index. *J Clean Prod* 241:118260
- Djordjevi B, Krmac E, Josip T (2018) Non-radial DEA model: a new approach to evaluation of safety at railway level crossings. *Saf Sci* 103:234–246
- Dulebenets MA (2021) An adaptive polyploid memetic algorithm for scheduling trucks at a cross-docking terminal. *Inf Sci* 565:390–421
- Dulebenets MA (2023) A diffused memetic optimizer for reactive berth allocation and scheduling at marine container terminals in response to disruptions. *Swarm Evol Comput* 80:101334
- Egilmez G, Mcavoy D (2013) Benchmarking road safety of U. S. states: a DEA-based Malmquist productivity index approach. *Accid Anal Prev* 53:55–64
- Emrouznejad A, Amin GR (2009) DEA models for ratio data: convexity consideration. *Appl Math Model* 33:486–498
- Evans A, Vaze V, Barnhart C (2016) Airline-driven performance-based air traffic management: game theoretic models and multicriteria evaluation. *Transp Sci* 50(1):180–203
- Gandhi N, Kant R, Thakkar J (2022) A systematic scientometric review of sustainable rail freight transportation. *Environ Sci Pollut Res* 29(47):70746–70771
- Ganji SS, Rassafi AA, Xu DL (2019) A double frontier DEA cross efficiency method aggregated by evidential reasoning approach for measuring road safety performance. *Meas: J Int Meas Confed* 136:668–688
- García-Palomares JC, Gutiérrez J, Martín JC, Moya-Gómez B (2018) An analysis of the Spanish high capacity road network criticality. *Transportation* 45(4):1139–1159
- Ghasemi MR, Ignatius J, Emrouznejad A (2014) A bi-objective weighted model for improving the discrimination power in MCDEA. *Eur J Oper Res* 233(3):640–650
- Hadi-Vencheh A, Hatami-Marbini A, Ghelej Beigi Z, Gholami K (2015) An inverse optimization model for imprecise data envelopment analysis. *Optimization* 64(11):2441–2454
- Hahn JS, Kho SY, Choi K, Kim DK (2017) Sustainability evaluation of rapid routes for buses with a network DEA model. *Int J Sustain Transp* 11(9):659–669
- Hespanhol L, Vallio CS, Costa LM, Saragiotto BT (2019) Understanding and interpreting confidence and credible intervals around effect estimates. *Braz J Phys Ther* 23(4):290–301
- Hussain Z (2022) Environmental and economic-oriented transport efficiency: the role of climate change mitigation technology. *Environ Sci Pollut Res* 29(19):29165–29182
- Ignatius J, Ghasemi MR, Zhang F, Emrouznejad A, Hatami-Marbini A (2016) Carbon efficiency evaluation: an analytical framework using fuzzy DEA. *Eur J Oper Res* 253(2):428–440
- lo Storto C, Evangelista P (2023) Infrastructure efficiency, logistics quality and environmental impact of land logistics systems in the EU: a DEA-based dynamic mapping. *Res Transp Bus Manag* 46:100814
- Izadikhah M, Azadi M, Toloo M, Hussain FK (2021) Sustainably resilient supply chains evaluation in public transport: a fuzzy chance-constrained two-stage DEA approach. *Appl Soft Comput* 113:107879
- Ji X, Wu J, Zhu Q (2016) Eco-design of transportation in sustainable supply chain management: a DEA-like method. *Transp Res Part D: Transp Environ* 48:451–459
- Kao C (2014) Network data envelopment analysis: a review. *Eur J Oper Res* 239(1):1–16
- Kiani Mavi R, Fathi A, Farzipoor Saen R, Kiani Mavi N (2019) Eco-innovation in transportation industry: a double frontier common weights analysis with ideal point method for Malmquist productivity index. *Resour Conserv Recycl* 147:39–48
- Kutty AA, Kucukvar M, Abdella GM, Bulak ME, Onat NC (2022) Sustainability performance of European smart cities: a novel DEA approach with double frontiers. *Sustain Cities Soc* 81:103777
- Ledwith MC, Hufstetler BJ, Gallagher MA (2021) Stochastic preemptive goal programming to balance goal achievements under uncertainty. *J Multi-Criteria Decis Anal* 28(1–2):85–98
- Liu WB, Meng W, Li XX, Zhang DQ (2010) DEA models with undesirable inputs and outputs. *Ann Oper Res* 173:177–194
- Mahdinia I, Habibian M, Hatamzadeh Y, Gudmundsson H (2018) An indicator-based algorithm to measure transportation sustainability: a case study of the U.S. states. *Ecol Ind* 89:738–754
- Memari P, Mohammadi SS, Jolai F, Ghaderi SF (2022) Sustainability assessment of renewable energy site location using a combinatorial decision-making model under uncertainty and data reliability. *Int J Syst Sci: Oper Logist* 10(1):2121624
- Muñuzuri J (2019) Use of DEA to identify URBAN geographical zones with special di ff icult y for freight deliveries. *J Transp Geogr* 79:102490
- Nag D, Paul SK, Saha S, Goswami AK (2018) Sustainability assessment for the transportation environment of Darjeeling, India. *J Environ Manag* 213:489–502
- Nahangi M, Chen Y, McCabe B (2019) Safety-based efficiency evaluation of construction sites using data envelopment analysis (DEA). *Saf Sci* 113:382–388

- Nikolaou P, Dimitriou L (2018) Evaluation of road safety policies performance across Europe: results from benchmark analysis for a decade. *Transp Res A Policy Pract* 116:232–246
- Odeck J (2006) Identifying traffic safety best practice: an application of DEA and Malmquist indices. *Omega* 34(1):28–40
- Omrani H, Amini M, Alizadeh A (2020) An integrated group best-worst method–Data envelopment analysis approach for evaluating road safety: a case of Iran. *Measurement* 152:107330
- Omrani H, Shamsi M, Emrouznejad A (2023) Evaluating sustainable efficiency of decision-making units considering undesirable outputs: an application to airline using integrated multi-objective DEA-TOPSIS. *Environ Dev Sustain* 25(7):5899–5930
- Pasha J, Nwodu AL, Fathollahi-Fard AM, Tian G, Li Z, Wang H, Dulebenets MA (2022) Exact and metaheuristic algorithms for the vehicle routing problem with a factory-in-a-box in multi-objective settings. *Adv Eng Inform* 52:101623
- Proost S, Dunkerley F, Van der Loo S, Adler N, Bröcker J, Korzhenevych A (2014) Do the selected Trans European transport investments pass the cost benefit test? *Transportation* 41(1):107–132
- Saeedi H, Behdani B, Wiegman B, Zuidwijk R (2019) Assessing the technical efficiency of intermodal freight transport chains using a modified network DEA approach. *Transport Res E: Logist Trans Rev* 126:66–86
- Seiford LM, Zhu J (2003) Context-dependent data envelopment analysis - measuring attractiveness and progress. *Omega* 31(5):397–408
- Seyedmohammadi J, Sarmadian F, Jafarzadeh AA, Ghorbani MA, Shahbazi F (2018) Application of SAW, TOPSIS and fuzzy TOPSIS models in cultivation priority planning for maize, rapeseed and soybean crops. *Geoderma* 310:178–190
- Shen Y, Hermans E, Brijs T, Wets G, Vanhoof K (2012) Road safety risk evaluation and target setting using data envelopment analysis and its extensions. *Accid Anal Prev* 48:430–441
- Shen Y, Hermans E, Bao Q, Brijs T, Wets G (2013) Road safety development in Europe: a decade of changes (2001 – 2010). *Accid Anal Prev* 60:85–94
- Shen Y, Hermans E, Bao Q, Brijs T, Wets G (2015) Serious injuries: an additional indicator to fatalities for road safety benchmarking. *Traffic Inj Prev* 16(3):246–253
- Singh E, Pillay N (2022) A study of ant-based pheromone spaces for generation constructive hyper-heuristics. *Swarm Evol Comput* 72:101095
- Singh P, Pasha J, Moses R, Sobanjo J, Ozguven EE, Dulebenets MA (2022) Development of exact and heuristic optimization methods for safety improvement projects at level crossings under conflicting objectives. *Reliab Eng Syst Saf* 220:108296
- Stefaniac A, Hosseini K, Assani S, Hosseini SM, Li Y (2021) Social sustainability of regional transportation: an assessment framework with application to EU road transport. *Socioecon Plann Sci* 78:101088
- Stenico de Campos R, Tadeu Simon A, de Campos Martins F (2019) Assessing the impacts of road freight transport on sustainability: a case study in the sugar-energy sector. *J Clean Prod* 220:995–1004
- Sueyoshi T, Li A, Liu X (2019) Exploring sources of China's CO2 emission: decomposition analysis under different technology changes. *Eur J Oper Res* 279(3):984–995
- Thanassoulis E, Boussofiiane A, Dyson RG (1996) A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega* 24(3):229–244
- Tian N, Tang S, Che A, Wu P (2020) Measuring regional transport sustainability using super-efficiency SBM-DEA with weighting preference. *J Clean Prod* 242:118474. <https://doi.org/10.1016/j.jclepro.2019.118474>
- Tian G, Lu W, Zhang X, Zhan M, Dulebenets MA, Aleksandrov A, ..., Ivanov M (2023) A survey of multi-criteria decision-making techniques for green logistics and low-carbon transportation systems. *Environ Sci Pollut Res* 30(20):57279–57301
- Tovar B, Wall A (2017) Dynamic cost efficiency in port infrastructure using a directional distance function: accounting for the adjustment of quasi-fixed inputs over time. *Transp Sci* 51(1):296–304
- Twumasi-Boakye R, Sobanjo J (2018) Civil infrastructure resilience: state-of-the-art on transportation network systems. *Transp A: Transp Sci* 0(0):1–38
- Wątróbski J, Bączkiewicz A, Ziemia E, Sałabun W (2022) Sustainable cities and communities assessment using the DARIA-TOPSIS method. *Sustain Cities Soc* 83:103926
- Wu J, Zhu Q, Chu J, Liu H, Liang L (2016) Measuring energy and environmental efficiency of transportation systems in China based on a parallel DEA approach. *Transp Res Part D: Transp Environ* 48:460–472
- Xie C, Bai M, Wang X (2018) Accessing provincial energy efficiencies in China ' s transport sector. *Energy Policy* 123(June):525–532
- Xie B, An Z, Zheng Y, Li Z (2019) Incorporating transportation safety into land use planning: pre-assessment of land use conversion effects on severe crashes in urban China. *Appl Geogr* 103:1–11
- Zhao L, Zha Y, Zhuang Y, Liang L (2019) Data envelopment analysis for sustainability evaluation in China: tackling the economic, environmental, and social dimensions. *Eur J Oper Res* 275(3):1083–1095
- Zhu J (2009) Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets, vol 2. Springer, New York

Web References

- DOT (2023) <https://www.transportation.gov/about>. Accessed 25 May 2023
- Eurostat (2019) <https://ec.europa.eu/eurostat/data/database>. (Dataset, extracted on): (Pollutant emissions from transport [T2020_RK300], 21/09/2019); (Final energy consumption in road transport by type of fuel [TEN00127], 21/09/2019); National annual road transport by group of goods and type of transport [road_go_na_tggt], 21/09/2019); (People killed in road accidents [SDG_11_40]; 21/09/2019); (Passenger road transport on national territory [road_pa_mov], 21/09/2019)
- ITF (2023) <https://www.itf-oecd.org/about-itf>. Accessed 25 May 2023
- Statista Research Department (2022a) Greenhouse gas emissions of the transport sector in Germany 2011–2020. <https://www.statista.com/statistics/411863/annual-greenhouse-gas-emissions-of-the-transport-sector-in-germany/>. Accessed 4 Oct 2023
- Statista Research Department (2022b) Road deaths in European countries 2020–2021. <https://www.statista.com/statistics/323869/international-and-uk-road-deaths/>. Accessed 4 Oct 2023
- Statista Research Department (2022c) Final energy consumption by road transportation in Italy 1990–2020. <https://www.statista.com/statistics/1332886/final-energy-consumption-by-road-transport-italy/>. Accessed 4 Oct 2023
- Tiseo I (2023) Greenhouse gas emissions in the United Kingdom 1990–2021, by sector, <https://www.statista.com/statistics/417948/uk-greenhouse-gas-emissions-by-sector/>. Accessed 4 Oct 2023

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.