Research in Transportation Economics 108 (2024) 101491

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Research paper

eXplainable DEA approach for evaluating performance of public transport origin-destination pairs

Eun Hak Lee

Metropolitan & Statewide Planning Division, Texas A&M Transportation Institute, 1111 Rellis Pkwy, 77807, Bryan, TX, United States

ARTICLE INFO

Keywords:
Data envelopment analysis
Explainable artificial intelligence
Extreme gradient boosting
Shapley additive explanations
Smart card data
Public transportation

ABSTRACT

Understanding public transportation efficiency is crucial to enhancing urban mobility and economic growth. This study aims to evaluate the efficiency of the public transportation system using an explainable data envelopment analysis approach. Specifically, data envelopment analysis and explainable artificial intelligence techniques were incorporated to uncover the black-box relationship between the inputs and outputs of the efficiency score. The efficiency of the public transportation system was defined as travel times for subway, bus, and multimodal modes relative to the number of trips and travel distances. The efficiency score was evaluated by the origin-destination pair unit. As a result, the efficiency score was estimated to be 0.69 on average, indicating that a reduction of 31% in travel times is required to achieve a perfect score of 1.00. Among the 33,313 origin-destination pairs, 39 had a perfect score of 1.00. The results of the interpretation model showed the order of importance for features—buses, subways, and multimodal modes—with SHapley Additive Explanations values of 0.047, 0.029, and 0.022, respectively. These results suggest that focusing primarily on reducing bus travel times is effective for improving overall efficiency. In this manner, the explainable data envelopment analysis helps measure performance, understand results, and suggest improvement directions.

1. Introduction

With the growth of urban areas, there has been a rising interest in enhancing public transportation systems. According to the United Nations Economic Commission for Europe (UNECE, 2017), public transportation plays a crucial role in improving urban mobility and economic performance. For example, buses use road space four times more efficiently than motor vehicles. Railways are even more efficient, moving up to 50,000 passengers every hour without delays. Given these advantages, many large cities have extensively operated public transportation systems and encouraged people to use them. The mode shares of public transportation are 74.0% in Seoul, 74.1% in Tokyo, 60.3% in Singapore, 53.7% in Beijing, and 56.7% in London (The Seoul Research Data Service, 2023). These high usage rates were achieved because public transportation services became convenient and efficient through advanced systems and policies. For example, the public transportation system in Seoul has operated an automatic fare collection (AFC) system since 2004. This system allows users to pay fares based on the distance traveled, regardless of the mode of public transportation used (Lee et al., 2019a). Additionally, it enables convenient transfers between different modes of transportation without incurring additional base fares (Jang, 2010). Users pay for all transportation modes with a single card, simplifying the payment process and enhancing user convenience (Lee, et al., 2019b). Similarly, London, New York, and Tokyo have also adopted AFC systems to enhance the efficiency and convenience of their public transportation services. As a result, public transportation in major cities is regarded as a highly attractive option for citizens, with mode shares ranging from 50% to 75%.

The growing emphasis on public transportation has increased attention on evaluating its performance, i.e., efficiency, convenience, and connectivity (Azmoodeh et al., 2021; Bruun et al., 2018; Collins et al., 2023; Guo & Brakewood, 2024; Lee, 2022; Miller et al., 2016; Pan et al., 2024; Welch & Mishra, 2013; Yun et al., 2021). Specifically, efficiency is a critical concept from both the operator's and the user's perspectives. It refers to the optimal use of resources to maximize output. In the context of public transportation, efficiency involves providing the best possible service with the resources available. Various factors such as demand, travel time, operational costs, service satisfaction, and service quality are considered when defining efficiency (Guo et al., 2018; Nishiuchi et al., 2015; Yaliniz et al., 2011). Additionally, efficiency is often evaluated comprehensively through multi-criteria decision-making techniques, which allow for a balanced consideration

E-mail address: e-lee@tti.tamu.edu.

https://doi.org/10.1016/j.retrec.2024.101491
Received 3 February 2024; Received in revised form 29 October 2024; Accepted 5 November 2024
Available online 14 November 2024

of these different variables (Guo et al., 2018; Lee & Lee, 2024; Lee et al., 2019b; Wang et al., 2022; Wiegmans et al., 2018). For example, several studies have used data envelopment analysis (DEA) models to evaluate various issues in public transportation, such as connectivity, convenience, and fare allocation (Guo et al., 2018; Nishiuchi et al., 2015; Xie et al., 2022). Multiple inputs and outputs were used to calculate a single score, and the performance of public transportation was comprehensively evaluated. Similarly, analytic hierarchy process (AHP) and analytic network process (ANP) have been used to evaluate the overall performance of public transportation systems by determining the weights of different criteria through surveys (Daimi & Rebai, 2023). Moreover, SFA has been applied to measure the efficiency of public transportation systems by estimating a frontier function, allowing for the evaluation of performance relative to the most efficient observations (Chung & Chiou, 2023). This method is particularly useful in cases where stochastic variations need to be accounted for, offering a more robust efficiency evaluation. In addition, advanced models, such as fuzzy DEA and AHP-DEA, have been developed to evaluate public transportation efficiency. Fuzzy models better manage uncertainties in criteria weights and performance measures, leading to more accurate evaluations (Izadikhah et al., 2021). AHP-DEA combines the strengths of DEA and AHP, balancing objective efficiency analysis with subjective preferences on criteria (Li et al., 2016). Although previous studies have provided valuable insights into evaluating transportation efficiency, there are limitations in evaluating and interpreting complex public transportation systems. Specifically, traditional multi-criteria decision-making techniques struggle to clearly explain the relationship between inputs and outputs and often fail to consider multiple inputs and outputs simultaneously.

With these introductions, the main idea of this study was to evaluate public transportation efficiency by incorporating DEA and explainable artificial intelligence (XAI) techniques, such as explainable DEA (XDEA). Specifically, the decision-making unit (DMU) was set as the origin-destination (O-D) pair. Public transportation efficiency was defined as the number of trips and the O-D distance (outputs) relative to O-D travel times for subway, bus, and transfer modes (inputs). The XDEA model consisted of three models, i.e., DEA, eXtreme Gradient Boosting (XGBoost), and SHapley Additive exPlanations (SHAP). The DEA model was used to evaluate the public transportation efficiency of the O-D pair and the XAI model was to identify relationships between the efficiency score and multiple inputs. The results showed that the proposed model had advantages in providing an objective evaluation and understanding of the level of public transportation services for each O-D pair.

2. Methodology

This study aims to develop an XDEA model for evaluating the efficiency scores of O-D pairs in public transportation systems. Efficiency is defined by comparing the number of trips and O-D distances (outputs) to O-D travel times for subway, bus, and multimodal modes (inputs). O-D pairs were aggregated at the Dong level, the administrative census unit, to consider the public transportation network as a whole. This allows for an independent comparative analysis of the three modes of public transportation within different urban areas. This definition of efficiency allows for the independent evaluation of subway, bus, and multimodal (bus + subway) systems. The proposed approach includes three models: the DEA model, the prediction model, and the interpretation model. Firstly, the DEA model was developed to evaluate the efficiency of each O-D pair. The output variables and input variables were obtained from smart card data. Secondly, a prediction model was developed using XGBoost to model the relationships between the estimated efficiency score and input variables. XGBoost is one of the most powerful machine learning models. Finally, an interpretation model was created using SHAP to explore the impact of input variables on the efficiency score. SHAP excels at interpreting and explaining predictions from complex machine learning models. In summary, the DEA, XGBoost, and SHAP models were used to estimate the efficiency score, model the relationship between the efficiency score and input variables, and understand the impacts of the input variables on the efficiency score, respectively. The XDEA model framework for evaluating the efficiency score is shown in Fig. 1.

2.1. Data envelopment analysis

DEA is a non-parametric technique for measuring the efficiency or productivity of decision-making units (Charnes et al., 1978). The DEA model helps identify efficient units and enhance areas of inefficiency, especially when dealing with complex datasets that involve multiple inputs and outputs (Farrell, 1957). With these strengths, the DEA technique has been widely used to evaluate the performance of public transportation systems, i.e., efficiency, convenience, sustainability, and equity (Guo et al., 2018; Lee et al., 2019b; Nishiuchi et al., 2015).

The DEA model is divided into input-oriented and output-oriented models. The input-oriented model estimates efficiency by fixing output variables and minimizing input variables. Conversely, the output-oriented model measures efficiency by fixing input and maximizing output variables. Additionally, the DEA model is categorized into two representative models based on the assumption of returns to scale: the Charnes, Cooper, and Rhodes (CCR) model and the Banker, Charnes, and Cooper (BCC) model (Banker et al., 1984; Charnes et al., 1978). The CCR model is used under the assumption of Constant Returns to Scale (CRS), which assumes that the ratio of inputs to outputs remains constant, even as the scale of inputs and outputs increases. Conversely, the BCC model operates under the assumption of Variable Returns to Scale (VRS), which assumes that the ratio of inputs to outputs is not constant. By integrating these assumptions, four distinct models emerge: input-oriented CCR, output-oriented CCR, input-oriented BCC, and output-oriented BCC. Each model is used according to the analysis's objectives and the data's characteristics.

In this study, an input-oriented BCC model was used to evaluate the efficiency score regarding the public transportation systems. The DMU, which is the entity being evaluated for efficiency, was set as the O-D pair. The number of trips and average travel distance were set as output variables, and the travel times for subway, bus, and bus + subway were set as input variables. The direction was assumed to be input-oriented to improve efficiency by reducing travel time of public transportation modes compared to outputs such as given O-D pair characteristics, i.e., number of trips and average travel distance. For returns to scale, the VRS assumption was used since efficiency varies depending on the outputs or inputs scale of the O-D pair. The input-oriented BCC model used in this study is expressed as Equation (1):

$$\theta^{k^*} = \min_{\theta,\lambda,s^-,s^+} \left\{ \theta^k - \varepsilon \left(\sum_{m=1}^{M} s_m^- + \sum_{n=1}^{N} s_n^+ \right) \right\} \tag{1}$$

subject to

$$\theta^k x_m^k = \sum_{i=1}^{I} x_m^i \lambda^i + s_m^- \ (m = 1, 2, \ldots, M);$$

$$y_n^k = \sum_{i=1}^{I} y_n^i \lambda^i - s_n^+ \ (n = 1, 2, \ldots, N);$$

$$\sum_{i=1}^{I} \lambda^i = 1$$

$$\lambda^i, s_m^-, s_m^+ \geq 0, i = 1, 2, \ldots, I;\ m = 1, 2, \ldots, M;\ n = 1, 2, \ldots, N$$

where θ^{k^*} is the efficiency score of O-D pair k, x_m^i is the observed value of the m-th input for the O-D pair i, y_n^i is the observed value of the n-th output for the O-D pair i, i is the number of O-D pairs ($i = 1, \ldots, 33{,}313$),

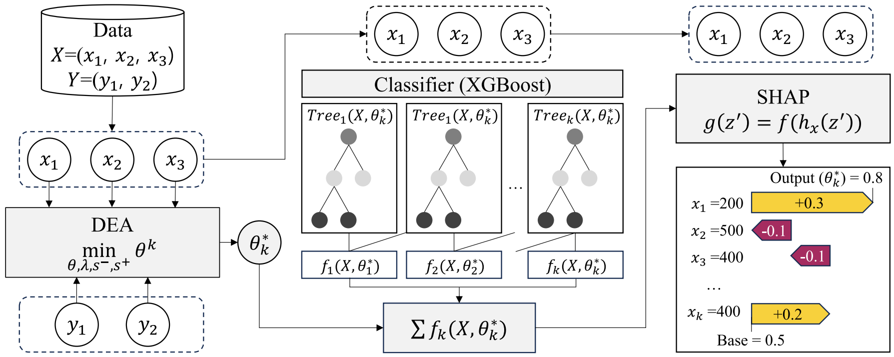

Fig. 1. eXplainable data envelopment analysis model framework.

n is the output variable ($n = 1$ and 2, $n = 1$: the number of O-D trips, $n = 2$: travel distance of O-D pair), m is the input variable ($m = 1, \ldots, 3$; $m = 1$: subway travel time, $m = 2$: bus travel time, $m = 3$: multimodal (bus + subway) travel time), λ^i is the weight assigned to O-D pair i, s_m^- is the slack variable for the m-th input, s_n^+ is the slack of the n-th output, and ε is a very small positive number.

2.2. eXtreme Gradient Boosting

XGBoost is an ensemble machine-learning technique that uses a series of decision trees (Chen & Guestrin, 2016; Lee, 2022). Its primary strengths include high predictive accuracy, rapid computation, and interpretability (Kim, 2021; Kwak & Lee, 2024; Lee, 2023). The core concept behind XGBoost is to refine the performance of preceding models by incorporating additional trees into the ensemble. As a result, XGBoost iteratively introduces trees and refines feature splits to better capture the residuals from previous predictions.

XGBoost was used to model the relationship between efficiency score and input variables, i.e., O-D travel times for subway, bus, and multimodal modes. The dataset consists of 33,313 samples, each representing an O-D pair. It includes independent variables x_i and a dependent variable $\theta_i^{k^*}$ which is the estimated efficiency score from the DEA model. The dataset is symbolized as $D = \left\{ \left(x_i, \theta_i^{k^*} \right) \right\}$, where $|D| = 33{,}313$. Each x_i has m variables ($m = 1:\text{subway}, 2:\text{bus}, 3:\text{multimodal}$), and these variables are associated with dependent variables $\theta_i^{k^*}$ ($x_i \in \mathscr{R}^m$; $\theta_i^{k^*} \in \mathscr{R}$). The tree ensemble model predicts the target value ($\hat{\theta}_i^{k^*}$) using Equation (2):

$$\hat{\theta}_i^{k^*} = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i) \tag{2}$$

where $\hat{\theta}_i^{k^*}$ is the predicted efficiency score for O-D pair i from the DEA model, x_i is the input vector of O-D pair i ($x_i \in \mathscr{R}^m$; $m = 1:\text{subway}, 2:\text{bus}, 3:\text{multimodal}$), w is the weight of the leaf node in the tree, and K is the number of tree functions. The set F is the space of trees the space of all possible trees that can be generated by the XGBoost model. It is defined as $F = \{f(x) = w_{q(x)}\}$, where $q : \mathscr{R}^m \rightarrow T$ describes the structure of the tree, mapping the input space x_i to the leaves T of the tree. The weights $w \in \mathscr{R}^T$ correspond to the values at the leaves of the tree, which are used to compute the final predictions. Each function $f_k(x_i)$ in the ensemble contributes to refining the prediction by adjusting these weights based on the structure of the tree.

The objective is to minimize $\mathscr{L}(\phi)$, and the mathematical expression is shown in Equation (3):

$$\mathscr{L}(\phi) = \sum_i l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*}\right) + \sum_k \Omega(f_k) \tag{3}$$

where, $\mathscr{L}(\phi)$ is the overall loss function, ϕ is the ensemble of tree functions, $l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*}\right)$ is the loss function that measures the difference between the actual efficiency score and the predicted efficiency score for O-D pair i, and $\Omega(f_k)$ is the regularization term that penalizes the complexity of the k-th tree function f_k. The mathematical expression of $\Omega(f_k)$ is expressed as in Equation (4):

$$\Omega(f_k) = \gamma T + \frac{1}{2}\alpha \|w_i\|^2 \tag{4}$$

where, γ is the regularization parameter that controls the penalty associated with the number of leaves in the tree, T is the number of leaves in the k-th tree, α is the regularization parameter that controls the penalty associated with the magnitude of the leaf weights, and $\|w_i\|^2$ is the squared L2 norm of the leaf weights w_i, which sums the squares of the weights assigned to the leaves of the tree.

The optimal (minimized) weight w_i^* of the leaf j is calculated as in Equation (5), and the corresponding optimal (minimized) value is estimated by Equations (6)–(8):

$$w_i^* = -\frac{\sum_{i \in I_j} \partial_{\hat{\theta}_i^{k^*-1}} l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*-1}\right)}{\sum_{i \in I_j} \partial^2_{\hat{\theta}_i^{k^*-1}} l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*-1}\right) + \alpha} \tag{5}$$

$$g_i = \partial_{\hat{\theta}_i^{k^*-1}} l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*-1}\right) \tag{6}$$

$$h_i = \partial^2_{\hat{\theta}_i^{k^*-1}} l\left(\theta_i^{k^*}, \hat{\theta}_i^{k^*-1}\right) \tag{7}$$

$$\tilde{\mathscr{L}}^t(q) = -\frac{1}{2}\sum_{j=1}^{T} \frac{\left(\sum_{i \in I_j} g_i\right)^2}{\sum_{i \in I_j} h_i + \alpha} + \gamma T \tag{8}$$

where, w_i^* is the optimal (minimized) weight of the leaf node for O-D pair i, g_i is the gradient of the loss function for the predicted efficiency score for O-D pair i, h_i is the Hessian (second derivative) of the loss function for the predicted efficiency score for O-D pair i, and $\tilde{\mathscr{L}}^t(q)$ is the objective function to be minimized for the tree structure q at iteration t.

Since it is difficult to compute all possible tree structures q. a greedy algorithm, which extends a single leaf to many branches iteratively, is used to estimate the optimal (minimized) value. This greedy algorithm is usually used to evaluate spilled candidates. The mathematical expression of the greedy algorithm is shown in Equation (9):

$$\mathscr{L}_s = -\frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \alpha} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \alpha} - \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \alpha}\right] - \gamma \tag{9}$$

where $I = I_L \cup I_R$, I_L is the instance set of left nodes after the split and I_R is the instance set of right nodes after the split, I_L represents the left node, which is the subset of data instances assigned to the left child node after a tree split, typically meeting a specific condition (e.g., feature value less than a threshold), I_R represents the right node, which is the subset of data instances assigned to the right child node after a tree split, typically not meeting the left node's condition (e.g., feature value greater than or equal to a threshold), and γ is a regularization term subtracted to penalize the addition of new leaf nodes.

2.3. SHapley Additive exPlanations

SHAP stands as a leading model-agnostic method in machine learning for explaining the output of black-box models such as machine learning models (Parsa et al., 2020). The core idea of SHAP is to allocate the contribution of each feature to a model's prediction, with the concept of cooperative game theory (Lundberg & Lee, 2017). Specifically, it calculates the average contribution of a feature value across all possible permutations of feature values. SHAP continually evaluates the impact of features for each estimation, updating the values through iterative questioning and answering (Lee, 2023). Finally, SHAP determines the contribution of each feature to the model's output based on marginal contributions. Given these advantages, SHAP values are extensively used to understand the relationships between input features and output features in black-box models. The SHAP is calculated as Equation (10) and d is determined by the additive feature function as Equation (11):

$$\psi_i(u) = \sum_{S\subseteq M\setminus\{m\}} \frac{|S|!(c-|S|-1)!}{c!}\left[u(v(S\cup\{m\})) - u(v(S))\right] \tag{10}$$

$$d(z') = f(p_x(z')) \tag{11}$$

where, $\psi_i(\xi)$ is the SHAP value of the m-th feature, u is the estimation function of v, M is the set of all input features $\{1,2,\ldots,m\}$, m is an index representing the specific feature (1: travel time for subway, 2: travel time for bus, 3: travel time for multimodal), S a subset of M, c is the total number of features, $v(S)$ is the vector for the subset S, $(v(S\cup\{m\}))-u(v(S))$ is the marginal contribution of i, referring to the difference in value generated when i is present or absent in the coalition S, d is the interpretation model, z' is the simplified input vector, and p_x is a mapping function from the simplified to the original input. Specifically, $p_x(z')$ represents what point in the original data space the coalition vector z' represents. The output of the prediction model u for this point should be the same as $d(z')$. This is one of the key principles of SHAP.

3. Application

3.1. Data description

The public transportation system in Seoul has been operated using an AFC system with smart cards since 2004. Smart card data provides 99% of trip information because the public transportation system operates entirely on an AFC system, requiring users to pay fares exclusively with smart cards. The daily ridership of public transportation in Seoul is approximately 8 million trips. Here, a trip refers to a single journey a passenger makes from origin to destination using public transportation. Smart card data was obtained from the Korea Transportation Safety Authority (KTSA), and data from May 17, 2017, was used to evaluate the efficiency of the OD-based public transportation system. Smart card data records details for each trip, including the trip ID, travel mode, origin station, destination station, transfer station, boarding time, alighting time, and transfer time. With these details, it is possible to obtain O-D-based public transportation attributes, i.e., the number of trips and travel time (Hwang et al., 2020).

Since smart card data provides trip information, preprocessing is required to obtain O-D-based information. Specifically, public transportation attributes for O-D are aggregated by Dong unit which corresponds to the census district unit in Seoul. Therefore, the number of trips and travel time for each mode are aggregated based on the Dong where the origin and destination stations are located. Moreover, cases where the origin and destination are reversed for a given O-D are also regarded as the same O-D. Since the O-D from o to d and d to o are regarded as the same, the number of trips for O-D pair i (y_1^i) is calculated by summing the number of trips from o to d (y^{od}) and the number of trips from d to o (y^{do}). Additionally, the number of trips for each public transportation mode m (m = 1: subway, m = 2: bus, m = 3: multimodal) is also considered. Equation (12) shows the mathematical expression for calculating the number of trips for O-D pair l.

$$y_1^i = \sum_{m=1}^{3}\left(y_{1,m}^{od} + y_{1,m}^{do}\right) \tag{12}$$

The distance between the origin and destination dis_i is calculated as the weighted average of travel distance for each mode (dis_m) based on the number of trips (y_1^i). This allows for a more accurate estimation of the average travel distance by considering the importance of each mode and the mathematical expression is shown in Equation (13):

$$dis_i = \frac{\sum_{m=1}^{3} y_{1,m}^i \times d_m}{y_1^i} \tag{13}$$

One of the advantages of smart card data is that it provides information about travel modes. The public transportation in Seoul is classified into three modes, i.e., subway, bus, and multimodal (bus + subway). Here, the multimodal refers to transfers between the subway and bus, or vice versa. Since the public transportation system in Seoul uses AFC, the fare is calculated based on the distance traveled, regardless of the mode of transportation used. As previously mentioned, the public transportation system in Seoul allows up to five free transfers between buses and subways, operating under a distance-based fare system. The multimodal option is significant for public transportation users, with multimodal trips accounting for about 20–30% of all public transportation trips. Regarding the three public transportation modes, travel times from origin to destination for each mode were obtained. For example, the subway travel time is calculated by dividing the total travel time of all subway trips for O-D pair i by the number of subway trips in O-D pair i. Similarly, the bus travel time and the multimodal travel time (bus + subway) are calculated in the same manner, i.e., dividing the total travel time of all bus trips for O-D pair i by the number of bus trips in O-D pair l, and dividing the total travel time of all multimodal trips for O-D pair i by the number of multimodal trips in O-D pair i.

The travel time for trip k is calculated by subtracting the tap-in time at the initial departure station from the tap-out time at the last destination station, as recorded in the smart card data. The travel time formulas for O-D pair i for the subway (x_1^i), bus (x_2^i), and multimodal (x_3^i) modes are shown in Equations (14)–(16).

$$x_1^i = \frac{\sum_{p=1}^{y_{1,sub}^i} x_{1,p}^i}{y_{1,sub}^i} \tag{14}$$

$$x_2^i = \frac{\sum_{p=1}^{y_{1,bus}^i} x_{1,p}^i}{y_{1,bus}^i} \tag{15}$$

$$x_3^i = \frac{\sum_{p=1}^{y_{1,tr}^i} x_{1,p}^i}{y_{1,tr}^i} \tag{16}$$

where, x_1^i is the average travel time for subway trips for O-D pair i, $y_{1,sub}^i$ is the total number of subway trips for O-D pair i, $x_{1,p}^i$ is the travel time for the p-th subway trip in O-D pair i, x_2^i is the average travel time for bus trips for O-D pair i, $y_{1,bus}^i$ is the total number of subway trips for O-D pair i, $x_{2,p}^i$ is the travel time for the p-th bus trip in O-D pair i, x_3^i is the average travel time for multimodal trips for O-D pair i, $y_{1,tr}^i$ is the total number of multimodal trips for O-D pair i, and $x_{3,p}^i$ is the travel time for the p-th multimodal trip in O-D pair i.

The preprocessed data revealed that 33,313 O-D pairs were used for public transportation trips on May 17, 2017. The average number of trips and the average travel distance for these 33,313 O-D pairs were found to be 139 trips and 18 km, respectively. The average travel times for the subway, bus, and multimodal (bus + subway) were 2,448, 3,375, and 2949 s, respectively. Fig. 2 shows the number of public transportation trips originating in Seoul.

3.2. Results of data envelopment analysis

Efficiency refers to how fast the public transportation system connects O-D pairs, considering the given number of O-D trips and O-D travel distance (Lee et al., 2019a; Nishiuchi et al., 2015; Guo et al., 2018). Regarding the efficiency concept, the efficiency score and improvement direction for each O-D pair were evaluated with the DEA model. The improvement direction stands for the reduction in travel time for public transportation services to enhance efficiency.

Fig. 3 shows the distribution of efficiency scores. As a result of the DEA model, the average efficiency score was estimated to be 0.69. The minimum, maximum, and standard deviation of efficiency scores were estimated to be 0.19, 1.00, and 0.10, respectively. Among 33,313 O-D pairs, 39 achieved an efficiency score of 1.00. These results indicated that the public transportation system evenly connected O-D pairs, and the efficiency score distribution closely resembled a normal distribution. From the perspective of policymakers, efficiency scores are useful for gaining insight into the overall public transportation system and identifying O-D pairs that require improvement.

Table 1 shows the result of the efficiency score. The results of the

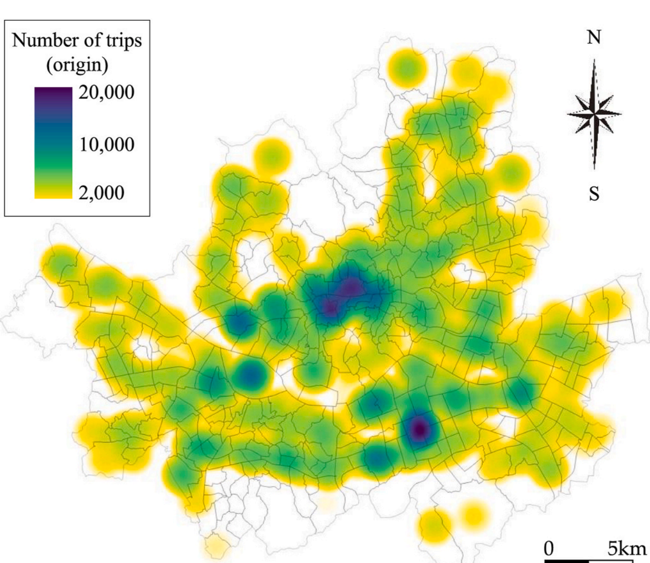

Fig. 2. Number of public transportation trips originating in Seoul.

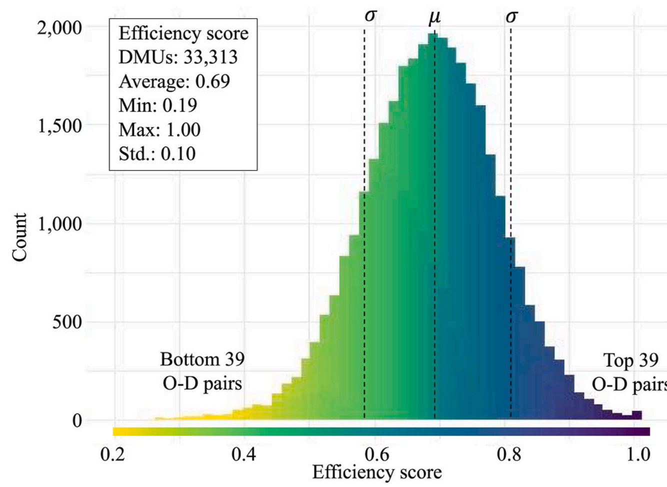

Fig. 3. Distribution of the efficiency score.

Table 1
Results of efficiency score.

		All 33313 O-D pairs	Top 39 O-D pairs	Bottom 39 O-D pairs
Efficiency score		0.69	1.00	0.27
Output	Number of trips	112	885	38
	Distance (km)	14.2	15.1	4.0
Input	Subway (sec)	2561	2128	3011
	Bus (sec)	3462	2463	2866
	Multimodal (sec)	2983	2336	3282
Optimal	Subway (sec)	1643	2128	1740
	Bus (sec)	1639	2463	1004
	Multimodal (sec)	1732	2336	1197

DEA model showed the efficiency score was estimated to be 0.69 on average. These results implied that, on average, a 31% travel time reduction was needed to achieve an efficiency score of 1.00. For the outputs, the number of trips and travel distance were 112 and 14.2 km, respectively. For the inputs, the subway, bus, and multimodal (bus + subway) travel times were 2,561, 3,462, and 2983 s, respectively.

Among 33,313 O-D pairs, 39 had an efficiency score of 1.00 and were estimated to be efficient. These top 39 O-D pairs were connected most efficiently by rapid public transportation modes, accommodating a high number of trips over extensive travel distances. Additionally, these efficient O-D pairs served as reference points for enhancing the efficiency of other, less efficient O-D pairs. The number of trips and travel distances for these efficient pairs were 885 and 15.1 km, respectively. These values were 7.94 and 1.06 times higher than the average values of the 33,313 O-D pairs. In terms of inputs, the travel times for the subway, bus, and multimodal modes were 2,128, 2,463, and 2336 s, respectively. These travel times were 0.83, 0.71, and 0.78 times shorter than the average values of the 33,313 O-D pairs for the respective travel modes. These results indicated that efficient O-D pairs transported a significant number of trips with shorter travel times.

The efficiency score for the bottom 39 O-D pairs was estimated to be 0.27, on average. These results implied that, on average, a travel time reduction of 73% (1.00–0.27) was needed to achieve an efficiency score of 1.00. The number of trips and travel distances for these O-D pairs were 38 and 4.0 km, respectively. These values were 0.34 and 0.28 times lower than the average values of the 33,313 O-D pairs. In terms of inputs, the travel times for the subway, bus, and multimodal modes were 3,977, 3,011, and 2866 s, respectively. The travel times for the subway and multimodal modes were much higher than the average values of their respective modes for the 33,313 O-D pairs, at 1.18 and 1.10 times, respectively. However, the bus travel time was only 0.83 times the average bus travel time for the 33,313 O-D pairs. The results suggested that these inefficient O-D pairs had disadvantages in the subway and multimodal modes but advantages in the bus mode.

To achieve an efficiency score of 1.00, there were specific requirements for the subway, bus, and multimodal travel times. The top 39 O-D pairs formed the efficiency frontier, representing the line that indicated the most efficient units at a given level of output. Inefficient O-D pairs referred to the efficiency frontier to determine the direction and requirements for improvement in each variable. The estimated optimal values for subway, bus, and multimodal travel times were, on average, 1,643, 1,639, and 1732 s, respectively. These represented about 0.64, 0.47, and 0.58 of the observed input values of the 33,313 O-D pairs, which were 2,561, 3,462, and 2983 s, respectively. This suggested that to achieve an efficiency score of 1.00, the travel times for each mode needed to be reduced by 36%, 53%, and 42%, respectively. As for the bottom 39 O-D pairs, the optimal travel times for subway, bus, and multimodal were estimated at 1,740, 1,004, and 1197 s, respectively. These values were about 0.58, 0.35, and 0.36 of the observed input values of the 33,313 O-D pairs, indicating a need for reductions of 42%, 65%, and 64%, respectively, to achieve an efficiency score of 1.00.

In summary, efficient O-D pairs offered 20–30% faster travel times on public transportation, despite longer travel distances and more trips compared to the average of 33,313 O-D pairs. These pairs were well-connected by major routes, such as subways or express buses. In contrast, inefficient O-D pairs resulted in longer travel times for subway and multimodal modes, despite shorter travel distances and fewer trips. These inefficient O-D pairs were primarily connected by feeder routes, indicating a need to reduce travel times to improve efficiency. From a transportation planning perspective, it is reasonable to reduce travel times for these inefficient O-D pairs by introducing express bus routes or improving transfer systems.

Fig. 4 shows the relationship between the travel times of various public transportation modes and the efficiency score. Efficient O-D pairs have shorter travel times compared to inefficient O-D pairs. Conversely, inefficient O-D pairs have longer travel times than efficient O-D pairs. To enhance efficiency, the travel times of inefficient O-D pairs need to be reduced to match those of the efficient O-D pairs. O-D pairs with an efficiency score of 1.0 form an efficiency frontier, represented by a purple-colored circle. O-D pairs outside of this frontier are considered inefficient and use the frontier as a benchmark to guide their improvement efforts. For example, the group with the lowest efficiency score, represented by a yellow-colored circle, is furthest from the frontier and requires a reduction in travel times to improve efficiency. As previously stated, O-D pairs with low efficiency scores have significantly longer travel times compared to efficient O-D pairs and require approximately a 45–65% reduction in travel times to achieve efficiency. These results provide guidance on how much travel times need to be reduced for each O-D pair to improve efficiency from the O-D perspective.

Fig. 5 shows the result of the efficiency score geographically for each O-D pair. It was found that the efficiency scores were higher for O-D pairs connecting areas with high commercial and business density, such as areas with a large floating population. On the other hand, the efficiency scores were lower for O-D pairs connecting residential-residential or residential-commercial areas. This implied that O-D pairs connecting commercial areas were well-served by public transportation. Conversely, for O-D pairs connecting residential-residential or residential-commercial areas, there was a need for improvement in public transportation services.

Another interesting observation was the significant difference in efficiency scores between the southern and northern regions of Seoul. Specifically, O-D pairs connected to areas in the south, such as Gangnam and Yeouido, were found to have relatively high efficiency scores. In contrast, O-D pairs connected to areas in the north, such as Jongno and Jungnang, had relatively low efficiency scores. Generally, the level of urban development and investment was higher in the southern region than in the northern region. This disparity was reflected in the efficiency scores, which offered valuable insights into the effectiveness of public transportation in different areas.

3.3. Results of prediction model

To validate the performance of the XGBoost model, two additional models were additionally developed using machine learning techniques, i.e., ANN and RF models. The performances of three models were evaluated using three performance measures, i.e., mean absolute percentage error (MAPE), root mean square error (RMSE), and mean absolute error (MAE). These metrics were widely recognized for highlighting the relationship between the precision of a prediction and its bias or error (Lee et al., 2021; Lee, 2023). The mathematical expressions for MAPE, MAE, and RMSE are shown in Equations (17)–(19):

$$MAPE = \frac{100}{n}\sqrt{\sum_{n}\frac{|x_{t+1}-\widetilde{x}_{t+1}|}{x_{t+1}}} \tag{17}$$

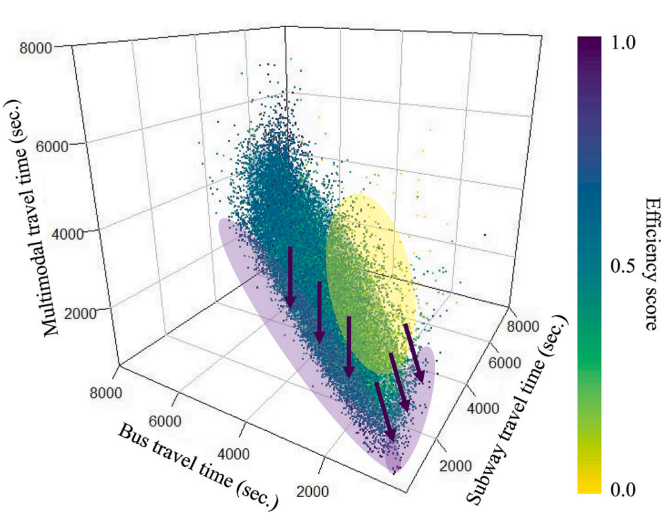

Fig. 4. Relationship between inputs and the efficiency score.

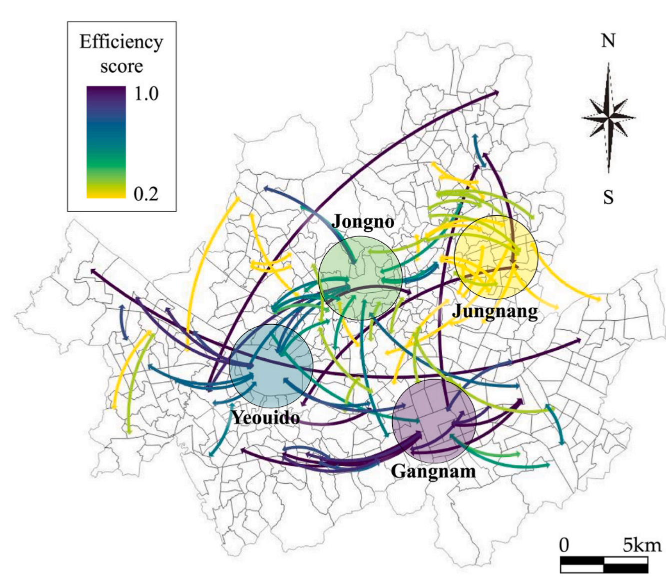

Fig. 5. Efficiency score with O-D pairs.

$$MAE = \frac{1}{n}\sum_{n}|x_{t+1} - \widetilde{x}_{t+1}| \tag{18}$$

$$RMSE = \sqrt{\frac{\sum_{n}(x_{t+1} - \widetilde{x}_{t+1})^2}{n}} \tag{19}$$

where n is the number of predicted efficiency scores, x is the actual efficiency score, and $\widetilde{x}$ is the predicted efficiency score.

The three models, i.e., ANN, RF, and XGBoost, were trained on 85% of the DMUs (O-D pairs) and tested on the remaining 15% of the DMUs. Both the training and test sets were selected randomly. The training set consisted of 28,316 out of 33,313 DMUs (O-D pairs), while the test data comprised the remaining 4997. The optimal hyperparameters for each model were determined using 4-fold cross-validation with a grid-search technique. The optimal hyperparameters were as follows: The ANN model had 5 neurons, 3 hidden layers, and a decay term of 0.1; the RF model used 8 trees, had a tree depth of 4, and used a learning rate of 0.1; similarly, the XGBoost model employed 8 trees, had a tree depth of 4, and used a learning rate of 0.1.

The performance results showed that the XGBoost model outperformed the others, with values of 0.04 for MAPE, 0.03 for MAE, and 0.04 for RMSE. These results aligned with previous studies where the XGBoost model often demonstrated superior performance (Kwak & Lee, 2024). Consequently, XGBoost was reasonable as the final prediction model to serve as the input for an interpretable model such as SHAP. The performances of the three efficiency prediction models are shown in Table 2.

3.4. Results of interpretation model

The feature importance with SHAP values of the three variables, i.e., travel times for subway, bus, and multimodal, from the XGBoost model is shown in Fig. 6. The variables were ordered based on their importance in estimating the efficiency score. The results indicated that the efficiency score was primarily influenced by bus travel time, with shorter bus travel times typically corresponding to lower efficiency scores. Despite this, the average SHAP value for bus travel time was found to be 0.047, suggesting that this feature generally had a positive impact on the efficiency score. This indicated that an increase in bus travel time led to a rise in the efficiency score. Subway travel time was identified as the second most influential variable, with an average SHAP value of 0.029, implying that it also positively affected the efficiency score. Lastly, the multimodal travel time was the third most significant variable. It had an average SHAP value of 0.022, indicating that it also had a positive impact on the efficiency score.

In summary, the features that had the greatest impact on the efficiency of public transportation are, in order, travel times for bus, subway, and multimodal transportation. This was because buses serve as feeder transportation and are often affected by road traffic and congestion, compared to subways and multimodal systems (Moon et al., 2021). Therefore, bus travel time was evaluated as the most sensitive factor affecting the efficiency of public transportation. Although buses had a lower direct connection function for O-D compared to other modes, they offered advantages in flexibility, accessibility, and route diversity (Chung & Chiou, 2023; Lee et al., 2019a). Recently, a policy to improve the efficiency of public transportation has been pursued by using express bus routes from the perspective of urban development. From this perspective, the efficiency score helps policymakers select O-D pairs where the connectivity of bus routes is weak and improves the level of services.

Table 2
Performance of the prediction model.

	ANN	RF	XGBoost
MAPE	0.11	0.08	0.04
MAE	0.07	0.05	0.03
RMSE	0.12	0.09	0.04

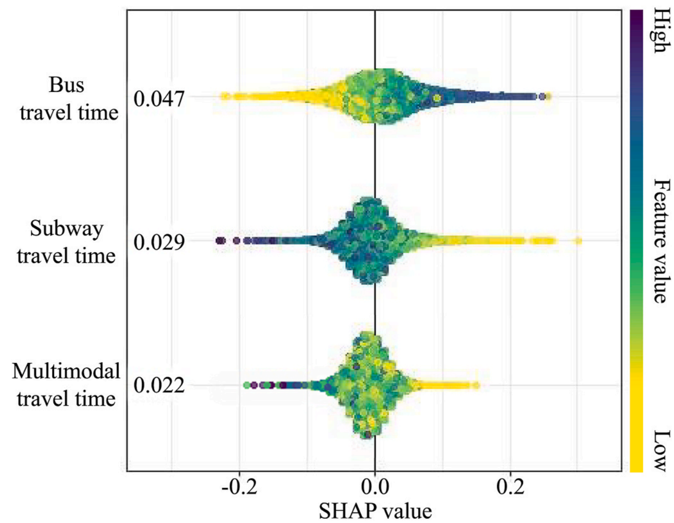

Fig. 6. Results of the feature importance with SHAP values.

The results of the feature dependency analysis are shown in Fig. 7. It illustrates the impact of each feature on the efficiency score and highlights the interactions with other features using SHAP values. Fig. 7(a) shows the impact of subway travel time on the efficiency score and its interaction with bus travel time. Specifically, in the initial phase (less than 1000 s), shorter subway travel times generally resulted in positive SHAP values, indicating an increase in efficiency. However, as subway travel time extended beyond 1000 s, these values shifted to negative, suggesting that longer subway travel times began to detract from efficiency. The relationship between subway travel time and bus travel time shows a positive correlation. Fig. 7(b) shows the impact of bus travel time on the efficiency score and its interaction with multimodal travel time. The efficiency score of an O-D pair tended to increase as bus travel time increased. Specifically, in the initial phase (less than 2000 s), shorter bus travel times generally resulted in negative SHAP values, indicating a decrease in efficiency. As bus travel time increased (beyond 2000 s), these values shifted to positive, suggesting that longer bus travel times started to enhance efficiency. This phenomenon was more pronounced in situations where buses served as long-distance feeder services. The relationship between bus travel time and multimodal travel time showed a positive correlation. Fig. 7(c) shows the impact of multimodal travel time on the efficiency score and its interaction with subway travel time. The efficiency score of an O-D pair tended to decrease as multimodal travel time increased. Specifically, in the initial phase (less than 2000 s), the impact on efficiency was mixed, with both positive and negative Shapley values. However, as travel time extended beyond 2000 s, the Shapley values increasingly shifted to the negative, suggesting that longer combined travel times tended to reduce efficiency, particularly when subway travel time was significant.

In summary, the efficiency score of an O-D pair decreased as subway and multimodal travel times increased. Conversely, the efficiency score of an O-D pair increased as bus travel time increased, specifically after 2000 s. This was due to the nature of each mode. Subways and multimodal modes are typically designed for rapid, direct transit across longer distances (Lee et al., 2022). Longer travel times for subways and multimodal modes indicated inefficiencies such as delays, congestion, or less direct routes. Conversely, buses often serve as feeder services for local routes (Moon et al., 2021). Longer bus travel times reflected their role in connecting more remote areas. Therefore, balancing bus and subway travel times is crucial for improving overall transportation efficiency. This could involve enhancing bus route planning to reduce travel times and making transfers between modes more seamless to

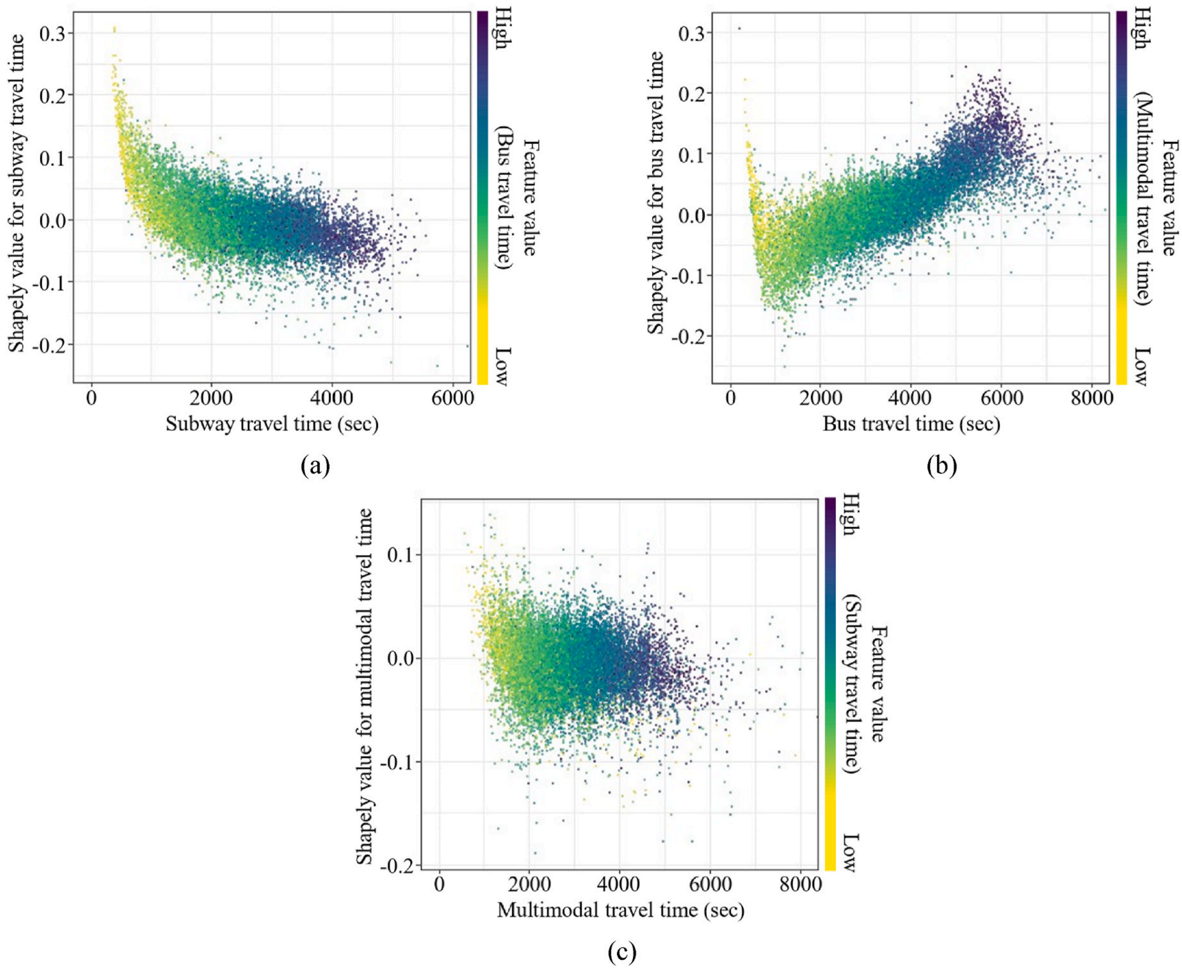

Fig. 7. Result of SHAP dependency analysis: (a) relationship between subway travel time and bus travel time, (b) relationship between bus travel time and multimodal travel time, (c) relationship between multimodal travel time and subway travel time.

minimize delays and inefficiencies in the system.

3.5. Discussion

The proposed XDEA model is also highly useful from a practical standpoint. Additional analysis was conducted to identify its practical implications. For this purpose, the efficiency scores, which were initially estimated based on O-D pairs, were re-aggregated on a per-origin Dong basis. For each origin Dong, the efficiency scores of all connected destination Dongs were averaged.

As a result, the efficiency score, connected to three of the bottom 39 O-D pairs, was estimated to be the lowest at 0.55. Fig. 8 shows the efficiency scores of the O-Ds connected to Sinnae. Specifically, Fig. 8(a) shows that the efficiency scores for the O-D pairs between Sinnae and areas within a 10 km radius (yellow circle), i.e., Jongno, Dongdaemun, Gwangjin, and Sanggye, were estimated to be relatively low. Conversely, the efficiency score for the O-D pair between Sinnae and Guro, the farthest travel distance, was evaluated as high at 0.88. Fig. 8(b) and (c) show the SHAP values of the efficiency scores for the O-D pairs between Sinnae and Dongdaemun, and Sinnae and Guro, respectively. The O-D with Dongdaemun as the destination had the most negative impact with a SHAP value of −0.179 for bus travel time. Conversely, the O-D with Guro as the destination had the most positive impact with a SHAP value of +0.121 for bus travel time. Sinnae area is a newly developing area and does not yet have sufficient public transportation infrastructure. Since there are only two local bus routes, it takes a long time to travel to surrounding areas. On the other hand, there is a direct express bus route operating with Guro, so the travel time is relatively short.

These results highlight the impact of individual public transportation modes on efficiency. In low-efficiency areas, such as connections within a 10 km radius of Sinnae, increasing local bus frequency or adding multimodal links helps reduce travel times. Conversely, for longer distances, as in the O-D pair between Sinnae and Guro, implementing direct express bus services significantly improves travel times.

As such, the XDEA model supports decision-making for enhancing public transportation efficiency. Specifically, it is useful for identifying O-D pairs that need improvement and for prioritizing upgrades to these routes. By enhancing public transportation services, it is possible to increase efficiency scores. Recently, demand-responsive transit (DRT) services have emerged as an option to improve the public transportation system, offering flexible scheduling on specific routes and times. DRT services provide the benefits of both main and feeder routes through effective route design.

Regarding the policy applicability of the XDEA model, O-D pairs with low efficiency scores were selected, and economic and efficiency evaluations were performed with the introduction of DRT services. Specifically, 16 O-D pairs connecting the Yong-am and Joonghwa areas were selected. These areas had the lowest efficiency score when taking the weighted average of efficiency scores based on the number of O-D pairs in each administrative area. This approach allows the model to focus on the least efficient connections, targeting improvements where they are most needed.

The network optimization assumptions are as follows: 1) Bus operation time was calculated by summing link travel time and dwell time.

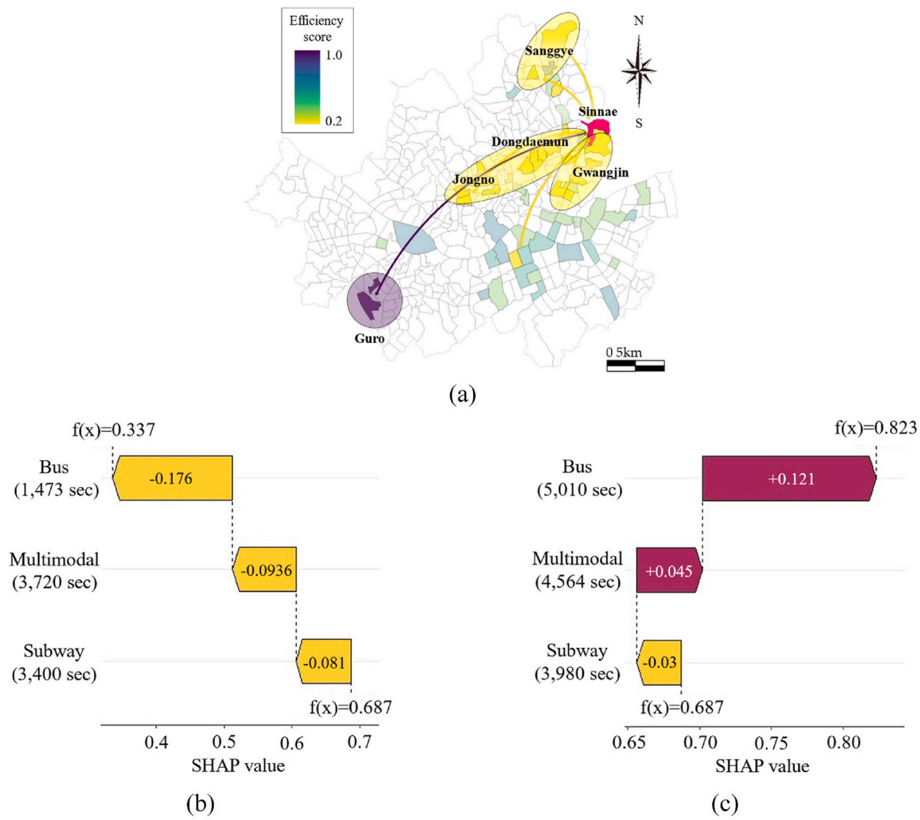

Fig. 8. Result of efficiency in Sinnae: (a) Efficiency score of Sinnae, (b) Efficiency score from Sinnae to Dongdaemun, (c) Efficiency score from Sinnae to Guro.

2) The passenger value of time is set at 9000 KRW per hour (Conversion: 1 USD ≈ 1300 KRW) based on Ministry of Land, Infrastructure and Transport data (2017). 3) Bus operation costs are set at 150,000 KRW. 4) Bus capacity is fixed at 45 passengers. 5) The travel demand at 7 a.m. during the morning peak for the 16 O-D pairs was 48 passengers, with the assumption that all board the vehicle.

The results show that the efficiency score of the 16 O-D pairs improved from 0.55 to 0.73, an increase of approximately 18%. Average travel time decreased by about 26 min, from 121 min to 95 min. The estimated benefit was 175,500 KRW (calculated as 26 min/60 min * 9000 KRW * 50 passengers), with a benefit-cost ratio of 1.17 (175,500 KRW/150,000 KRW). Assuming a fare of 3000 KRW, revenue was projected at 135,000 KRW, resulting in a revenue-cost ratio of 0.9. Given that Seoul's public transportation has a revenue-to-cost ratio of 0.55 (1046/1904 KRW), these results indicate high economic viability. Like this, flexible and cost-effective options like DRT provide more direct and efficient transit solutions. This aligns with previous research indicating that introducing express and local bus routes tailored to the characteristics of O-D pairs significantly enhances public transportation efficiency (Hwang et al., 2020; Moon et al., 2021). Although the implementation of DRT was part of a short-term plan, there is potential to further improve efficiency in the medium term by strengthening the transfer system and, in the long term, by introducing express rail lines. Fig. 9 shows the optimal DRT route planned based on efficiency scores from the XDEA model.

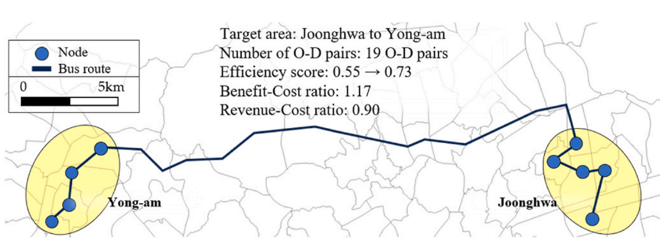

Fig. 9. New DRT route planning with XDEA model.

These findings suggest that the XDEA model is useful for identifying O-D pairs where travel times have potential for improvement, as it considers demand, distance, and travel time. Specifically, grouping low-efficiency O-D pairs and introducing new direct routes between them ensures considerable economic feasibility. Additionally, it provides direction for enhancing O-D pair efficiency by analyzing the relationship between travel times for each mode and efficiency scores. The model helps determine whether improvements are needed for major or feeder routes. Such insights are also related to economic feasibility, supporting the formulation of efficient policy-making.

4. Conclusions

Public transportation is crucial for strengthening economic and social connections in urban areas. It significantly influences a city's spatial planning, shaping its development and growth. Many major cities have emphasized and advanced efficient public transit for their transportation competitiveness. This study aims to evaluate public transportation efficiency using the XDEA technique, integrating DEA with the XAI technique. Specifically, the DMU was defined as the O-D pair, estimating efficiency based on trip numbers and O-D distance as outputs, compared to travel times for subway, bus, and transfer modes as inputs. The XDEA framework incorporated three advanced methods: a DEA model for evaluating O-D pair efficiency, a prediction model for learning relationships between efficiency scores and inputs, and an interpretation model for understanding relationships between efficiency scores and inputs. The findings highlighted the model provides an impartial evaluation and insights into public transportation service levels for individual O-D pairs.

The proposed model indicated that the efficiency score for the O-D pairs stood at 0.69, on average. This implied that an approximate reduction of 31% in travel times was required to achieve a perfect efficiency score of 1.00. Among the 33,313 O-D pairs, 39 had optimal efficiency with a score of 1.0. These top 39 O-D pairs were distinguished by their expedient public transportation services relative to their respective number of trips and travel distances. They served as references, providing benchmarks for enhancing the efficiency scores of other O-D pairs. Conversely, among the 39 O-D pairs with the lowest efficiency, the average efficiency score was 0.27. This highlighted that there needed to be a considerable reduction of about 73% in travel times to achieve a perfect efficiency score of 1.0. The SHAP analysis showed the order of importance for features such as buses, subways, and multimodal modes. The SHAP values for buses, subways, and multimodal travel were 0.047, 0.029, and 0.022, respectively. Specifically, O-D pairs with higher efficiency scores were linked with shorter travel times in subways and multimodal transport. Conversely, long bus travel times were associated with high efficiency scores. Considering these results, it was essential to focus on reducing bus travel times to enhance the overall efficiency of the O-D public transportation system.

The proposed model offered three significant contributions. Firstly, an XDEA model that integrated both DEA and XAI approaches was introduced to evaluate and understand the efficiency of the public transportation system. Secondly, evaluations were conducted based on O-D pairs, expanding upon the conventional analysis resolutions, i.e., station and area units, of efficiency evaluation using the DEA model. Thirdly, specific requirements to enhance the efficiency scores were pinpointed for each O-D pair. In conclusion, the XDEA model stood out for its ability to provide a unified metric for evaluating the efficiency of the integrated public transportation system and offered insights into its underlying factors. Also, the policy applicability of the proposed model was verified by economically introducing new public transportation routes.

Although the proposed model demonstrated notable performance and provided valuable insights, several issues still require further investigation. For example, this study determined the required travel times for subway, bus, and multimodal modes based on the number of trips and travel distances. However, travel demand could increase as the public transportation system for O-D pairs improves, potentially impacting efficiency. Specifically, considering a possible shift in demand from automobiles to public transportation, the scope of the study could be expanded by incorporating this aspect into the model. These considerations are crucial not only for improving the efficiency of public transportation systems but also for fostering a more interconnected and unified society.

Declaration of competing interest

The author declares no conflict of interest.

References

Azmoodeh, M., Haghighi, F., & Motieyan, H. (2021). Proposing an integrated accessibility-based measure to evaluate spatial equity among different social classes. *Environment and Planning B: Urban Analytics and City Science, 48*(9), 2790–2807.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.

Bruun, E., Allen, D., & Givoni, M. (2018). Choosing the right public transport solution based on performance of components. *Transport, 33*(4), 1017–1029.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794).

Chung, Y. S., & Chiou, Y. C. (2023). On the efficiency of subsidized bus services in rural areas: A stochastic metafrontier approach. *Research in Transportation Business & Management, 46*, Article 100811.

Collins, K., Der Wartanian, R., Reed, P., Chea, H., Hou, Y., & Zhang, Y. (2023). Social equity and public transit in the inland empire: Introducing a transit equity analysis model. *Transportation Research Interdisciplinary Perspectives, 21*, Article 100870.

Daimi, S., & Rebai, S. (2023). Sustainability performance assessment of Tunisian public transport companies: AHP and ANP approaches. *Socio-Economic Planning Sciences, 89*, Article 101680.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A, 120*(3), 253–281.

Guo, J., & Brakewood, C. (2024). Analysis of spatiotemporal transit accessibility and transit inequity of essential services in low-density cities, a case study of Nashville, TN. *Transportation Research Part A: Policy and Practice, 179*, Article 103931.

Guo, J., Nakamura, F., Li, Q., & Zhou, Y. (2018). Efficiency assessment of transit-oriented development by data envelopment analysis: Case study on the den-en toshi line in Japan. *Journal of Advanced Transportation, 2018*(1), Article 6701484.

Hwang, H., Cho, S. H., Kim, D. K., & Kho, S. Y. (2020). Development of a model for evaluating the coverage area of transit center using smart card data. *Journal of Advanced Transportation, 2020*(1), Article 8819791.

Izadikhah, M., Azadi, M., Toloo, M., & Hussain, F. K. (2021). Sustainably resilient supply chains evaluation in public transport: A fuzzy chance-constrained two-stage DEA approach. *Applied Soft Computing, 113*, Article 107879.

Jang, W. (2010). Travel time and transfer analysis using transit smart card data. *Transportation Research Record, 2144*(1), 142–149.

Lee, E. H. (2022). Exploring transit use during COVID-19 based on XGB and SHAP using smart card data. *Journal of Advanced Transportation, 2022*(1), Article 6458371.

Lee, E. H. (2023). *Traffic Speed Prediction of Urban Road Network based on High Importance Links using XGBoost and Shapley Additive Explanation*. IEEE Access.

Lee, E. H., Kho, S. Y., Kim, D. K., & Cho, S. H. (2021). Travel time prediction using gated recurrent unit and spatio-temporal algorithm. In *, Vol. 174. Proceedings of the institution of civil engineers-municipal engineer* (pp. 88–96). Thomas Telford Ltd. No. 2.

Kim, E. J. (2021). Analysis of travel mode choice in Seoul using an interpretable machine learning approach. *Journal of Advanced Transportation, 2021*(1), 6685004.

Kwak, K., & Lee, E. H. (2024). Impact of road transport system on groundwater quality inferred from explainable artificial intelligence (XAI). *Science of the Total Environment, 917*, 170388.

Lee, C., & Lee, E. H. (2024). Evaluation of urban nightlife attractiveness for Millennials and Generation Z. *Cities, 149*, Article 104934.

Lee, E. H., Kim, K., Kho, S. Y., Kim, D. K., & Cho, S. H. (2022). Exploring for route preferences of subway passengers using smart card and train log data. *Journal of Advanced Transportation, 2022*(1), Article 6657486.

Lee, E. H., Lee, H., Kho, S. Y., & Kim, D. K. (2019a). Evaluation of transfer efficiency between bus and subway based on data envelopment analysis using smart card data. *KSCE Journal of Civil Engineering, 23*, 788–799.

Lee, E. H., Shin, H., Cho, S. H., Kho, S. Y., & Kim, D. K. (2019b). Evaluating the efficiency of transit-oriented development using network slacks-based data envelopment analysis. *Energies, 12*(19), 3609.

Li, X., Liu, Y., Wang, Y., & Gao, Z. (2016). Evaluating transit operator efficiency: An enhanced DEA model with constrained fuzzy-AHP cones. *Journal of Traffic and Transportation Engineering (English Edition), 3*(3), 215–225.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*.

Miller, P., de Barros, A. G., Kattan, L., & Wirasinghe, S. C. (2016). Analyzing the sustainability performance of public transit. *Transportation Research Part D: Transport and Environment, 44*, 177–198.

Moon, S., Cho, S. H., & Kim, D. K. (2021). Designing multiple short-turn routes to mitigate the crowding on a bus network. *Transportation Research Record, 2675*(11), 23–33.

Nishiuchi, H., Todoroki, T., & Kishi, Y. (2015). A fundamental study on evaluation of public transport transfer nodes by data envelop analysis approach using smart card data. *Transportation Research Procedia, 6*, 391–401.

Pan, M. M., Wong, S., Tainter, F., Woelfel, S., & Ryan, A. (2024). Integrating equity in transportation scenario planning: A systematic review. *Transport Policy, 145*, 85–95.

Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis & Prevention, 136*, Article 105405.

The Seoul Research Data Service. (2023). https://data.si.re.kr/data/ Accessed: December 1. [Online]. Available:.

United Nations Economic Commission for Europe (UNECE). (2017). Sustainable urban mobility and public transport. https://www.citiesforum.org/news/the-changing-paradigm-of-public-transport/.

Wang, B., Liu, C., & Zhang, H. (2022). Where are equity and service effectiveness? A tale from public transport in shanghai. *Journal of Transport Geography, 98*, Article 103275.

Welch, T. F., & Mishra, S. (2013). A measure of equity for public transit connectivity. *Journal of Transport Geography, 33*, 29–41.

Wiegmans, B., Champagne-Gelinas, A., Duchesne, S., Slack, B., & Witte, P. (2018). Rail and road freight transport network efficiency of Canada, member states of the EU, and the USA. *Research in Transportation Business & Management, 28*, 54–65.

Xie, Q., Wu, X., Dai, Q., Zheng, X., & Wang, F. Y. (2022). An integrated data envelopment analysis and non-cooperative game approach for public transportation incentive subsidy allocation. *IEEE Transactions on Intelligent Transportation Systems, 23*(11), 21515–21530.

Yaliniz, P., Bilgic, S., Vitosoglu, Y., & Turan, C. (2011). Evaluation of urban public transportation efficiency in Kutahya, Turkey. *Procedia-Social and Behavioral Sciences, 20*, 885–895.

Yun, H., Lee, E. H., Kim, D. K., & Cho, S. H. (2021). Development of estimating methodology for transit accessibility using smart card data. *Transportation Research Record, 2675*(11), 159–171.