Journal of Infrastructure, Policy and Development 2025, 9(2), 6564.
https://doi.org/10.24294/jipd6564

Article

Enhancing MRT purple line offset project through forecasting ridership using the direct demand ridership model

Kamonchanok Wangvittaya, Dahye Yang, Hyemin Park*

International School of Urban Sciences, University of Seoul, Seoul 02504, South Korea
*** Corresponding author:** Hyemin Park, minpark@uos.ac.kr

CITATION

Wangvittaya K, Yang D, Park H. (2025). Enhancing MRT purple line offset project through forecasting ridership using the direct demand ridership model. Journal of Infrastructure, Policy and Development. 9(2): 6564. https://doi.org/10.24294/jipd6564

ARTICLE INFO

Received: 22 May 2024
Accepted: 20 June 2024
Available online: 5 March 2025

Abstract: The Mass Rapid Transit (MRT) Purple Line project is part of the Thai government's energy- and transportation-related greenhouse gas reduction plan. The number of passengers estimated during the feasibility study period was used to calculate the greenhouse gas reduction effect of project implementation. Most of the estimated numbers exceed the actual number of passengers, resulting in errors in estimating greenhouse gas emissions. This study employed a direct demand ridership model (DDRM) to accurately predict MRT Purple Line ridership. The variables affecting the number of passengers were the population in the vicinity of stations, offices, and shopping malls, the number of bus lines that serve the area, and the length of the road. The DDRM accurately predicted the number of passengers within 10% of the observed change and, therefore, the project can help reduce greenhouse gas emissions by 1289 tCO_2 in 2023 and 2059 tCO_2 in 2030.

Keywords: climate change; greenhouse gas; MRT purple line; direct demand ridership model; Thailand; transportation plan; mitigation policy; mass rapid transit master plan; mass rapid transit authority; clean development mechanism; number of passengers

1. Introduction

Given the negative impacts of global warming on humanity, the 21st annual session of the Conference of the Parties (COP21) resolved to maintain the global surface temperature below 2 °C compared to pre-industrial revolution levels and cap a global increase at 1.5 °C. Consequently, the Paris Agreement, which was adopted by 196 parties at COP21 in 2016, requests that each country outline and communicate their post-2020 climate actions, known as their nationally determined contributions (NDCs).

The government of Thailand ratified their intended NDCs (INDCs), aiming to reduce greenhouse gas (GHG) emissions by 20%–25% by 2030 compared to the business as usual (BAU) scenario (555 $MtCO_2eq$). Thailand established its NDC Roadmap for 2030 in 2017, initiating a variety of national-level climate-change mitigation policies and actions to promote its transition to a low-carbon, resilient society. Of the sectors in Thailand's NDC Roadmap for 2030 reduction plan, the energy sector has the largest reduction target (113 $MtCO_2eq$); the second largest reduction target within the energy sector is the transportation sector (41 $MtCO_2eq$) (ONEP, 2018). Given that reducing GHG emissions in the transportation sector is critical to the NDC Roadmap, more projects and initiatives are necessary to assist in meeting targets.

One such initiative is the Transport Infrastructure Development Plan of Thailand, approved by the Thai government in 2015 (2015–2022). This plan proposes the promotion of rail mass transit, which contributes to GHG reduction in the transportation sector. The plan incorporates the Bangkok Metropolitan Region Mass Rapid Transit Master Plan (M-Map) and the Railway Master Plan (R-Map). According to the M-Map, under the Mass Rapid Transit Authority (MRTA) in Thailand, there are 12 rail lines (**Figure 1**), totaling 509 km. Approximately 17% of public rail transportation is operational, and the remainder is under construction.

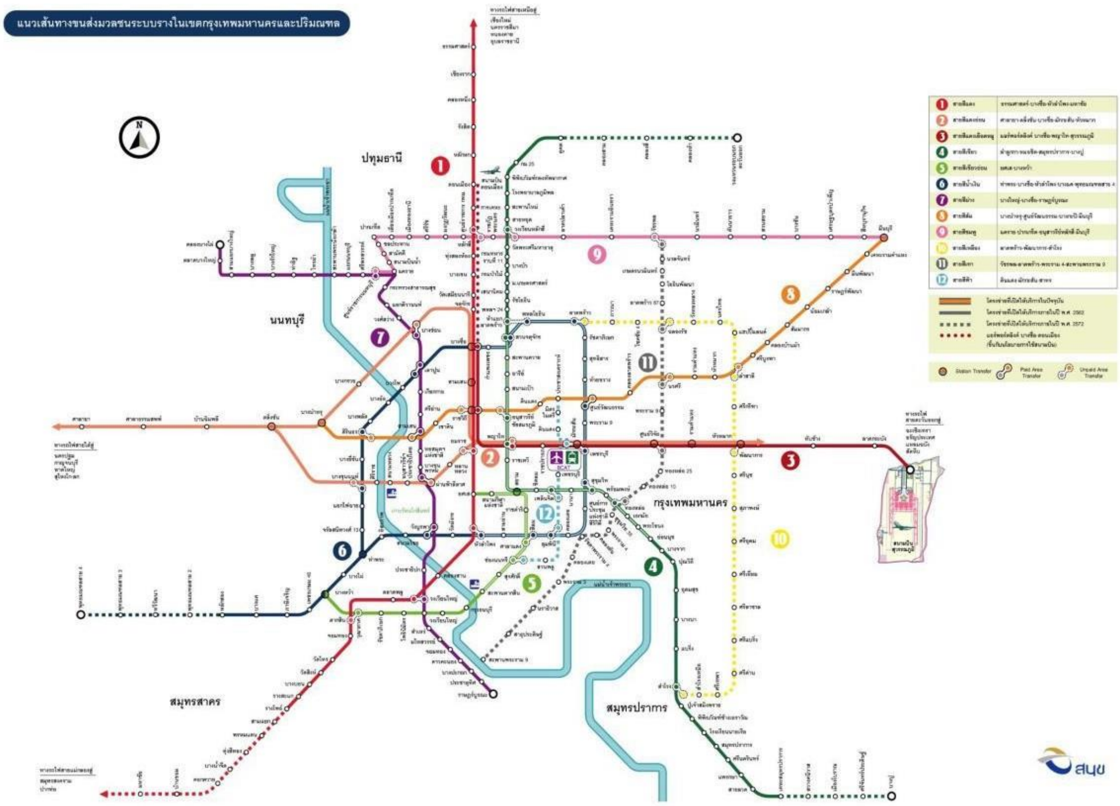

Figure 1. Mass rapid transit master plan (M-Map) in Bangkok metropolitan region (UNCRD, 2010).

This study aimed to calculate GHG reductions by predicting the number of passengers that will benefit from the Mass Rapid Transit (MRT) Purple Line project, which is one of the detailed lines of the M-Map. Prior to the service, the MRT Purple Line Project was part of the GHG operation under MRTA. That is, this study aims to provide a comprehensive analysis of the MRT Purple Line project by applying the Direct Demand Ridership Model (DDRM) to predict ridership and evaluating the GHG emissions associated with the proposed changes. These two aspects are explored separately to ensure clarity in methodology and results. Passenger numbers and energy consumption were estimated using the Feasibility Study Report of the MRT Purple Line by the MRTA, and those estimates were used to calculate the GHG reduction results from the project's implementation. However, the figures from the feasibility studies presented in the passenger section of the report are substantially inflated compared to the actual number of passengers using the MRTA in 2016. The performance of the railway project was either underestimated or overestimated due to multiple variables that may influence railway functioning.

Various previous studies have used models that consider variables to accurately predict the number of passengers in the railway sector. GHG emissions of urban rail transportation have been accurately measured based on passenger volume projections. It can help quantify the impact of urban rail transportation on reducing carbon emissions (Yuan et al., 2023). By identifying factors that attract the number of passengers, policies such as revitalization of public transportation and greenhouse gas reduction in the rail transportation sector were suggested. Liu et al. (2016) developed direct ridership models for railroad transfer stations in Maryland to predict the number of passengers and proposed policy implications to increase the number of public transportation users, but noted that this was limited by the state's specialization. Vergel-Tovar and Rodriguez (2018) investigated the association between BRT station-level demand and built environment attributes for 120 stations in seven Latin American cities. Using a direct ridership model, the study found that various land uses around the station, the location of the BRT station relative to the central business district, developable land around the station, and the integration of the station to the urban fabric determine BRT ridership. A study using a mathematical model incorporating an algorithm to predict CO_2 emissions in the transportation sector predicted accurate CO_2 emissions in the transportation sector in Canada and analysed changes in emissions in the transportation sector according to various energy resources. Accordingly, GHG reduction scenarios in Canada's transportation sector were evaluated (Javanmard et al., 2023). Previous studies have shown that accurate prediction of passenger demand using railways is related to GHG emissions.

No research has yet predicted passenger demand for the MRT Purple Line using a direct model applicable to other electric train projects. This study aims to accurately predict passenger numbers by analyzing proximity to stations, identifying factors motivating public transit use, and estimating GHG reductions. To accurately predict ridership and evaluate GHG emission reductions, it is crucial to consider various determinants that influence public transit usage and to employ robust methodologies for emission estimation. Recent studies emphasize the need for a comprehensive approach that incorporates socio-economic, environmental, and infrastructural variables (De Gruyter et al., 2018; Litman, 2020).

2. Predicting passengers using direct demand ridership models

A wide range of factors influence public transit ridership, including socio-economic characteristics, environmental concerns, and the quality of service. Eom et al. (2012) highlights that transit accessibility and connectivity significantly affect ridership levels, while socio-economic factors such as income and car ownership also play crucial roles. Previous studies typically rely on traditional regression models or simpler demand forecasting techniques (Kain, 1999; Litman, 2020). However, these methods often fail to account for the complex interplay of factors influencing ridership. DDRM addresses this gap by integrating variables such as population density, employment density, road length, and bus line connectivity. The detailed steps involved in the application of DDRM include data collection, variable selection, model calibration, and validation. This comprehensive approach has been shown to improve prediction accuracy in various contexts (Cervero et al., 2010; Duduta, 2013; Vergel-Tovar and Rodriguez, 2018).

DDRM was developed to predict, compare, and select transportation modes in urban areas in the United States. Numerous studies have used direct demand ridership models (DDRMs) to predict passenger numbers. These models provide precise results and make it easy to foresee the funding needed to support transport modes (Cervero et al., 2010). This model also considers the area surrounding the station and the service characteristics of mass transit systems when predicting passenger numbers.

The model can be applied in two primary ways: DDRM for train passengers and DDRM for express bus passengers. The difference between the two categories lies in the data used for analysis (Duduta, 2013).

Cervero et al. (2010) utilized DDRM to predict passenger counts for Bus Rapid Transit (BRT) in Los Angeles. They used 69 stations in three lines as the analysis sample, with an average of 744 passengers. These lines are connected to the orange and blue metro lines. This study aimed to examine the relationship between the built environment, transit services, and ridership.

We conducted an analysis to examine the number of passengers arriving at and departing from the stops and stations over time. The variables used can be categorized into three groups: one set of variables is related to service and two sets are related to characteristics of the station area; and all three groups of variables are associated with stop signs. Details are presented in **Table 1**.

Table 1. The variables used for analysis are classified by the characteristics of the variables (Cervero et al., 2010).

Variable class	Variable list
Availability of BRT service	The daily frequency of buses in both direction
	The number of service hours per day
	The number of feeder buses perpendicular to the daily station
	The number of bus line of feeder bus perpendicular of the daily station
	The number of connections to the subway per day
	The number of parallel train services per day
	The number of perpendicular train services per day
	Percentage of specific way for BRT
Location and attributes surrounding the station (0.5 miles away from the station)	Density of population around the station in 2000
	Employment density in 2000
	The total density between population and employment the station in 2000
	Street connectivity index
	Density of population around the station in 2000
	The distance to the nearest next station
	Destination station
Facilities within the station	The number of Park and Ride
	Capacity of Park and Ride
	Stopover while waiting for the bus
	Schedule for bus
	Passenger information system
	Bus waiting area with a roof
	Long-distance bus connection point
	The symbol has the BRT sign on it.

A multiple regression analysis with the least number of squares was performed. The results match **Table 2**, which shows that the number of people living within 0.5 miles of a bus stop is one of the factors that affects how a key station is cleared. Analysis results also showed that when more people live in an area, more passengers will use the bus stop. For example, if the number of people living within 0.5 miles of the stop sign increases from 5000 to 10,000, assuming all other variables remain the same, approximately 170 more people will ride the bus. Ridership increases as the distance between one station and the next station increases. The coefficient shows a value of 261.705 if there is a station 1.5 miles from the closest express bus station. Approximately 260 more people will use the bus than at the station half a mile away. The capacity of Park and Ride has a coefficient of 0.514, which is positive, indicating that when there is a bus stop with parking-lot capacity, it is likely that the bus will have more passengers.

Table 2. The variables used to estimate the number of passengers for 69 stations for BRT (Cervero et al., 2010).

Variables	**Cof.**	**Beta**	***T***	**Sig.**
Services of BRT				
Daily frequency of buses in both directions	5.103	0.176	3.771	0.000
The number of feeder buses perpendicular to the daily station	73.921	0.080	2.051	0.045
The number of bus lines of feeder buses perpendicular to the daily station	6.722	0.126	3.476	0.001
Location and attributes surrounding the station				
Population density around the station	0.017	0.134	4.303	0.000
Distance to the nearest next station	261.705	0.060	1.736	0.088
Interaction				
Specific lane for BRT and feeder buses (0–1). * Number of feeder buses perpendicular to the daily station	124.557	0.123	2.005	0.050
Specific lane for BRT and feeder buses (0–1). * Number of bus line of feeder buses perpendicular to the daily station	52.891	0.533	13.807	0.000
Specific lane for BRT and capacity of Park and Ride (0–1) * The number of Park and Ride stations	0.514	0.093	2.067	0.043
Specific lane for BRT and capacity of Park and Ride (0–1) * The total density of population and employment at the station	0.036	0.185	3.202	0.002
Constant	−541.164	-	−3.50	0.001

R square = 0.952; *F* Statistic (prob.) = 129.011 (0.000); 69 Samples.

Once completing the analysis, the number of passengers offered by the equation must be checked against the real number of passengers. Since the coefficient of determination is calculated at 0.952, it can be assumed that the estimated equations are close to the truth.

When using the DDRM to predict how many people will ride the BRT, some of the information considered is in line with this thesis. Therefore, this paper uses some variables to examine other data in greater detail (Cervero et al., 2010).

Duduta (2013) analysed a direct ridership model for the BRT and subways in Mexico City. However, BRT and subways exhibit different technological and service characteristics. Therefore, the model was divided into two types: the BRT model and the subway system model.

The analyzed data were divided into four categories:

1) Areas surrounding Mexico City station. We determined the distance that could

be covered between 10 and 15 min. This group of variables included the distance in meters to the closest station on the same metro line and the station density, which was considered to be the number of stations within a 5-km radius.

2) Network connectivity and accessibility to business centers are essential. For the network connectivity parameters, the connection of the BRT system to the subway was considered. The ease of accessibility to the business center variable considers how accessible the business center is from the station.

3) The station area density for the variables in this group considers the characteristics of the area surrounding the station, such as low building heights. In the absence of more precise data, density was measured in only three categories: Low, medium, and high (Duduta, 2013). Some areas around the station belonged to the low-density group; these areas have buildings that are no more than two stories and are not mixed-use spaces. This type of station is typically residential and located far from the city center. The second type has a medium density. In the medium density group, the entire area surrounding the station is developed. The area includes buildings between three and eight stories high that have mixed residential, commercial, and other uses. In the high-density group, the entire area around the station has been developed. It includes eight-story buildings and places of interest, such as shopping malls. This type of station is typically located in the city center.

4) Bus routes. For this category, the most important variable is the number of buses arriving at the station. More buses mean more passengers.

We employed multiple regression analysis to assess the model. The outcomes for the express buses and subways are presented in **Tables 3** and **4**, respectively. The results show that terminal stations, where passengers can conveniently wait before boarding the BRT, rather than stations on the way, have the greatest impact on BRT ridership. The density of the station area is important for both the BRT and subway; therefore, it is possible to determine which variables influence the number of passengers by analyzing the variables that affect them. According to this study, developing passenger terminals will allow more passengers to use BRT services. In addition, the proximity of a station to densely populated areas encourage more passengers to use the service.

Table 3. Model for forecasting the number of passengers (take-offs) for BRT (Duduta, 2013).

	Coefficient	***P***
Number of connecting bus routes	61.9	0.662
Density of the area surrounding the station (1 = low, 2 = medium, 3 = high).	3120.6	0.000
Connect to the subway (0 = no, 1 = yes).	2510.1	0.050
Terminal station (0 = no, 1 = yes).	8759.3	0.000
Area of interest (distance to the nearest station, meters)	5.2	0.008
Constant	−2395.1	0.234
Number of Samples	51	
R^2	0.51	
F (prob)	9.38 ($p < 0.001$)	

Table 4. Model for forecasting the number of passengers (take-offs) for metro (Duduta, 2013).

	Non-CBD model		**CBD model**	
	Coefficient	***P***	**Coefficient**	***P***
Area of interest (distance to the nearest station, meters)	7.2	0.178	-	-
Number of connecting bus routes	305	0.353	1224.8	0.002
Number of BRT and routes * Specific Lane (0–1)	981	0.034	-	-
Terminal at the station	17,586	0.000	-	-
Density of the area surrounding the station (1 = low, 2 = medium, 3 = high)	8909.4	0.000	16,161.5	0.000
Number of metro routes at the station	8628.5	0.049	-	-
-	9390	0.000	n/a	n/a
Constant	−16,347	0.05	−7309	0.349
Number of Samples	84		41	
R^2	0.54		0.51	
F (prob)	12.87 ($p < 0.001$)		12.89 ($p < 0.001$)	

The variables used in forecasting passenger numbers were unique to station characteristics. There is no consideration of the BRT and metro service systems, which are the factors that influence the use of both transportation systems.

Usvyat et al. (2009) estimated metro (heavy rail) passenger counts using data from 474 stations in representative cities, including Baltimore, Boston, Chicago, Cleveland, Los Angeles, Miami, New York, Philadelphia, San Francisco, and Washington, D.C., then conducted multiline linear regression analysis. Research indicates that the number of passengers utilizing a service depends on their demographic information.

The data used in this study were divided into four categories:

- Population data from the area surrounding the station.
- Specific details about each station.
- Information about the population along the route.
- Information about the population in metropolitan areas.

This research reveals that we can increase the number of variables as necessary and determine the relationships between them. Variables consistent with the research were included in the analysis along with other data.

3. Methodology

3.1. Selecting the model's variables

The primary hypothesis of this study was that the variables influencing the annual number of passengers per station on the MRT Purple Line consist of factors related to the station's surrounding space utilization. The number of passengers per station per year is influenced by the characteristics and services of mass transit systems as well as the surrounding population. Using the number of passengers per station in 2017 and expected variables affecting the number of passengers, a model was developed to examine the relationship between the number of passengers and those variables in order to test the main hypothesis. State Railway of Thailand (SRT), Bangkok Mass

Transit System Public Company Limited (BTS) under the Bangkok Metropolitan Authority and Mass Rapid Transit Authority (MRTA) in Thailand provided passenger information for DDRM. Additionally, office space, shopping malls and residential areas within 500m of the station were referenced from the Department of city planning and urban development in Bangkok and the Department of Public works and town & country planning in Nonthaburi.

This work also contains sub-assumptions. The specifics are listed below:

1) The number of individuals surrounding the station influences the number of passengers. The greater the number of people surrounding the station, the more passengers there are.

2) On the journey with other systems, the terminal station and the terminal connecting station will have more passengers than the station along the way. Many connecting stations will have fewer passengers than the terminal station.

3) The use of space surrounding the station that influences the number of passengers includes office and company areas, hotel and shopping mall areas, and educational areas. If there are more of these areas near the station, the number of passengers will increase.

These variables influence the number of passengers at each station. The details are factors for space utilization around the station, population around the station, and service characteristics of the mass transit system.

The station's area utilization factor is examined in terms of thousands of square meters. Each variable contains the following information:

1) Residential areas surrounding the station are a variable that affects passenger volume. If there are more residential areas in the vicinity of the station, the population will increase. A literature review revealed that the residential area was not taken into account because it was highly correlated with the number of passengers around the station. Consequently, only considering the number of passengers within the station. However, residential areas were also considered in this study, as the population surrounding the stations used in the study was only a rough estimate. Consideration of additional living space could increase the model's efficiency.

2) The length of the roads around the station is a variable that affects the number of passengers. The length of a roads around the station can be considered in two ways. The longer the road, the easier it is to access the station by other means of transport. In other words, the more roads there are, the more passengers will use transportation. On the other hand, as the road gets longer, the number of passengers using the service decreases due to the availability of alternative modes of transport. The perimeter of a station is measured in km. The length of two types of roads, concrete roads and paved roads, was determined by measuring their lengths, taking into account the width of the road from one lane and up. According to the literature review, the length of roads surrounding the station was not considered because this variable is unlikely to affect the number of international passengers. In Thailand, however, the researcher anticipates that the length of the road surrounding the station will affect the number of passengers because the station's accessibility in Thailand has developed alternative modes of transport to the station. Therefore, the length of roads surrounding the station was assumed to mean the connectivity to the station and was considered as a variable. Because the characteristics of Thailand's transportation modes were taken into

consideration, this study assumed that this would have a positive impact on passengers.

3) Variables like offices, hotels, shopping malls, and public utilities define the employment-generating and service-using area around the station. As the number of employed individuals increases, the number of passengers will also rise.

The population surrounding the station is the group most likely to utilize the station's services. The greater the station's surrounding population, the greater the number of passengers. In this study, where the population distribution is a static distribution, the same assumption is still made. The greater the population of the station's surroundings, the greater the number of passengers.

The service characteristics of the mass transit system vary across different transportation systems. Key factors include park-and-ride capacity, the public transportation system in Bangkok, the number of bus routes to each station, and three station types. Here's a clearer breakdown:

1) Passenger count: The more people using the parking lot, the more likely they are to use public transportation due to its convenience.

2) Park-and-Ride Capacity: A literature review found a positive correlation between park-and-ride capacity and passenger numbers. Stations with larger parking spaces tend to have more passengers than those with limited or no park-and-ride facilities.

3) Bus routes: The number of bus routes at each station influences passenger numbers. In Thailand, buses can either transport passengers to the station (feeder buses) or serve as an alternative mode of transport. Feeder buses increase passenger numbers, while competing bus routes may decrease them. Overall, the number of bus routes positively correlates with the number of passengers if the buses act as feeder services.

4) Station types: Stations are categorized into terminal stations, interchange stations, and en-route stations. Terminal and interchange stations, which connect to other systems, tend to have more passengers compared to en-route stations. A literature review found a positive correlation between these station types and passenger numbers, leading to higher passenger use.

From the literature review, the primary variables influencing the annual passenger number per station can be summarized. **Table 5** provides a summary of the MRT Purple Line's study area's information-finding potential and suitability, based on which the study area was chosen.

Table 5. Consideration must be given to primary variables when conducting research.

Independent variable	Variable symbol	Expected coefficient
Road distance	Road	(+, −)
Parking lots	Parking	(+)
Buses line	Buses	(+, −)
Terminal station	Terminal	(+)
Interchange	Interchange	(+)
Populations	Population	(+)
Office and store	Office	(+)
Residential area	Residence	(+)

The variables to be analyzed with the least squares method of multiple regression are summarized in **Table 5**. In addition to specifying the symbol of the coefficient of the variable expected from the model, a literary review forecast specifies the coefficient of the model. The coefficients obtained from the model may differ from those found in the literature. The predicted outcomes of the initial variables for which the coefficient is anticipated to be positive are mixed area utilization, station population, parking capacity, station variables, residential, business office, hotel, shopping mall, and other industries. The length of the roads surrounding each station as well as the number of bus routes that serve each station are anticipated to have positive and negative coefficients.

In addition to analyzing the Purple Line Project, this study examined Bangkok's public transportation system. The BTS SkyTrain and Airport Rail Link are included. Consequently, the study was split into two models. The first model evaluated 26 stations and did not include the 7 BTS Sky Train stations because these stations are in urban centers. It is contradictory with the suburban location of the MRT Purple Line project, and the variables' characteristics are different, but the second model analyzed all 33 stations. The results are displayed in **Table 6**.

Table 6. The models that consider all stations and the models that exclude urban stations.

Independent variable	**Model 1**			**Model 2**		
	Coef.	**T**	**sig**	**Coef.**	**T**	**sig**
Residential	−0.0002	−0.3155	0.7563	0.00183	2.7448	0.0113
Office and Store	0.0015	1.9129	0.0728	−0.00039	−0.5808	0.5668
Populations	0.0001	1.0031	0.3299	0.00155	1.9282	0.0657
Interchange station	−0.0073	−0.0691	0.9457	0.00010	1.4321	0.1650
Terminal	0.1890	1.4437	0.1670	0.27759	1.8165	0.0818
Buses	−0.0645	−2.9710	0.0086	0.28970	1.3684	0.1839
P&R	0.0000	0.2193	0.8290	−0.01359	−0.5832	0.5652
Road	0.0324	2.4035	0.0279	0.00004	0.2479	0.8063
R^2	0.804			0.600		
Adjust R^2	0.712			0.466		
Residual	0.2097			3.038		

In the evaluation, adjust R^2 and residual values were utilized. The Adjust R^2 value for Model 1 is closer to 1 than the value for Model 2. This indicates that Model 1 is a better fit for the data covered in this study. The residual statistic for the second model is greater than the residual statistic for the first model, indicating that the first model has a smaller squared estimation error. This relates to the initial hypothesis of the least squares approach for calculating the multiple regression equation's parameters. Considering both statistical figures, it is evident that Model 1 is more appropriate than Model 2 for analyzing the number of passengers per station annually. Therefore, it can be concluded that Model 1, or a model with limitless variables that does not account for the seven BTS Sky Train stations, is suitable for further study.

The relationships between the independent and source variables were determined prior to selecting the model variables. Semi-log regression modeling was used to estimate the variable coefficients, and **Table 7** illustrates which variables influence passenger counts. The stepwise addition of independent variables, or stepwise regression, was used to determine which variables should be included; this is a suitable method for determining the best predictor variable and the optimal number of variables. With stepwise analysis, an incoming predictor variable is tested each time a new variable is added to the equation, which means that if some predictor is entered into the equation, the equation is eliminated. If the predictive variable did not result in a statistically significant increase in the R-value, it was omitted.

Table 7. Variables selected for equation by stepwise analysis.

Independent variables	**B**	**t**	**sig**	**Standardized Coef. Beta**	R^2	**Adjusted** R^2
1 (Constant)	6.789	34.875	0.000	-	0.545	0.526
Buses	−0.092	−5.364	0.000	−0.738		
2 (Constant)	6.437	30.289	0.000	-	0.661	0.632
Buses	−0.075	−4.606	0.000	−0.602		
Population	7.041×10^{-5}	2.809	0.010	0.367		
3 (Constant)	5.803	16.585	0.000	-	0.722	0.684
Buses	−0.062	−3.790	0.001	−0.495		
Population	6.984×10^{-5}	3.007	0.006	0.364		
Road	0.030	2.192	0.039	0.269		
4 (Constant)	5.825	17.858	0.000	-	0.770	0.726
Buses	−0.066	−4.317	0.000	−0.530		
Population	5.341×10^{-5}	2.319	0.031	0.278		
Road	0.028	2.172	0.041	0.249		
Office and Store	0.001	2.083	0.050	0.233		

As shown in **Table 7**, the stepwise regression test revealed that the variables that affected the model were the length of the road surrounding the station, the number of bus routes at each station, the number of offices and shopping malls, and the population surrounding the station. The resulting Equation (1) is as follows:

$$Log10(Y) = 5.825 + 0.028(Road) - 0.066(Buses) + 0.001(office \; \& \; department) + 5.341 \times 10^{-5}(Population) \quad (1)$$

3.2. Evaluating GHG emissions reductions for the railway project

The GHG emissions analysis was conducted to evaluate the environmental impact of the proposed changes to the MRT Purple Line. This analysis involves calculating baseline emissions, project emissions, and the resultant emission reductions. The methodology follows established guidelines and integrates data on energy consumption, emission factors, and projected ridership Various methodologies are used to estimate GHG emission reductions from public transit projects. Boarnet et al. (2017) compared top-down and bottom-up approaches, finding that integrated models that account for both direct and indirect emissions provide the most accurate estimates. Additionally, advancements in data analytics and machine learning have improved the precision of these estimations. In the transportation sector can be achieved either through technological intervention on the fuels used by vehicles or through driver behaviour changes regarding vehicle activities.

The MRTA, the owner of the MRT Purple Line Project, submitted a letter of intent to develop the Clean Development Mechanism (CDM) project, under which potential emission reductions were assessed using relevant methods for the transportation sector. There are two estimation methods for CO_2 emissions, corresponding to two distinct mitigation strategies. These are based on fuel (top-down) and driving style (bottom-up). For the fuel-based methodology, emissions were determined based on previously aggregated fuel consumption data. The following method multiplies fuel consumption using the CO_2 emission factor for each fuel type. The fuel emission factor is calculated using the fuel's heat content, the proportion of carbon and hydrogen in the fuel that is oxidized, and the fuel's content coefficient. Data on fuel consumption can be obtained from various sources, including fuel receipts (although their accuracy remains questionable), financial records of fuel expenditure, or a direct measurement of fuel consumption. In the absence of specific information on fuel consumption, vehicle activity data (e.g., distance traveled) and fuel economy factors (e.g., km/L) can be used to calculate fuel consumption. In the absence of fuel consumption data, the distance-based method should be used. In this method, emissions are calculated using distance-based emission factors that vary based on driving patterns, which can be expressed in terms of vehicle km traveled (VKT) and passenger km, which are typically acquired using a variety of traffic engineering approaches that range from the conventional four-step model to cutting-edge traffic demand estimation models. Nevertheless, it appears that the existence of these errors does not satisfy the CDM's requirements for precise estimated emissions. All data used to calculate CO_2 emissions using this second methodology are difficult to measure precisely in situations where vehicles are not centrally controlled and are influenced by behavioral factors (Gojash et al., 2005). Because this method is primarily based on human-km emission parameters and compares project input emissions, there is a high risk of error, and it is difficult to determine carbon emissions when high-capacity rapid transit projects operate over long distances. Additionally, the rapid transit as an emission source may not capture the complex changes in overall transportation carbon emissions (Wang et al., 2022). Additionally, issues with parameter calibration and calculation accuracy have occurred (Chen et al., 2017). Because this technique cannot accurately predict emissions from the rapid transit projects as a source of emissions in a variety of environments, it will be difficult to calculate emissions based on ridership. However, these studies did not provide a reliable quantitative method for cities in developing nations that lack urban rail transit. There are numerous unanswered questions: What type of measurement is most suitable? Before and after the opening of a new rail transit line, residents' travel distances and modes of transportation will change (Zhang et al., 2020).

For projects involving mass rapid transit, the ACM0016 methodology for rail-based urban mass rapid transit system (MRTS) is the only technique for measuring GHGs in Thai rail systems that has been certified by the Thailand Greenhouse Gas Management Organization (TGO) (UNCRD, 2010). A typical ACM0016 project involves extending an existing rail line or expanding the existing rail infrastructure (e.g., new rail lines), and GHG emissions mitigation involves the displacement of more GHG-intensive transport modes (e.g., an existing fleet of buses operating in mixed traffic conditions) with less GHG-intensive ones (e.g., newly developed rail-based systems or segregated bus lanes) (UNFCCC, 2021).

For the ACM0016 methodology, a top-down method was employed. The baseline emissions, project emissions, leakage emissions, and emission reductions were quantified separately. The term “emission reduction” describes the reduction in GHG emissions that results from the implementation of any strategic project to reduce emissions. Baseline, project, and leakage emissions must be considered when carrying out such projects. Equation (2) provides a formula for calculating emissions reduction by deducting the project’s emissions and leakage emissions from the baseline emissions.

$$\text{Emission Reduction} = (\text{Baseline Emission} - \text{Project Emission} - \text{Leakage Emission}) \tag{2}$$

3.3. Baseline emissions

The emissions that would have resulted from passenger transportation between the point of origin (O) and final destination (D) prior to the project activity were calculated. This calculation is differentiated based on the modes of transport (relevant vehicle categories) that passengers would have used if the project had not been constructed. These are the maximum possible emissions from an existing project that has not been subjected to any emissions reduction strategies, plans, or schemes.

Baseline emissions are the potential GHG emissions before the project is implemented. Equation (3) shows the product of the baseline emission factor and the amount of energy generated by the baseline project before emission reduction or green projects were added (UNFCCC/CCNUCC, 2008).

In the absence of fuel consumption data, a distance-based approach should be used to calculate fuel consumption. In this method, emissions were calculated using distance-based emission factors that vary based on driving patterns, which can be expressed in terms of VKT, passenger km, which are typically acquired by various traffic engineering approaches, and range from the conventional four-step model to state-of-the-art traffic demand estimation models. However, errors in these models do not appear to meet the CDM’s requirements for emission estimation precision (Gojash et al., 2005).

$$\text{Baseline Emission} = \text{Energy Emission Factor} \tag{3}$$

3.4. Project emission

The emissions from a project are the highest amounts that can be emitted as a result of the project’s energy consumption. Equation (4) represents this as a by-product of the project’s operational energy use and emission factor:

$$\text{Project Emission} = \text{Energy Consumed During Project Operation} \times \text{Emission Factor} \tag{4}$$

3.5. Leakage emission

The leakage emission is composed of:

- Emissions caused by change of load factor of taxis and buses in the baseline transportation system.

- Emissions from affected roadways that make traffic less crowded, but also cause cars to travel faster, giving it a rebound effect.

The impact of the project on traffic (increased trips) was calculated in the emissions calculation for the project, not the leakage calculation. If the program does not occur, the non-attending passenger's trip will be counted.

Total leakage is calculated from:

$$LE_y = LELFB_{,y} + LELFT_{,y} + LECON_{,y} \tag{5}$$

where:

LE_y: Volume of leakage in year y (tCO_2);

$LELFB_{,y}$ Leakage caused by a shift in the load factor of buses in a given year y (tCO_2);

$LELFT_{,y}$: The amount of leakage in year y caused by a change in the cab load factor (tCO_2).

$LECON_{,y}$: The amount of leakage caused by light traffic in y.

If the leakage emission $LE_y < 0$, it was not included in the calculation of emission reduction.

4. MRT purple line analysis

The predicted ridership of the MRT purple line was greater than the actual ridership. This will affect a variety of components and parties, including investment costs, the economy, and cash flow. The evaluation of GHG emissions from passengers showed that ridership was one of the most important operating factors. This section provides an equation for passenger forecasting using multiple linear regression as well as the relationship between independent variables that affect ridership growth. This equation can be applied to both future and current projects to increase and forecast ridership.

4.1. Model results

As shown in **Table 8**, the overall *R* square of the model was 0.770, indicating that the model evaluated 77% of the dependent variables that could be predicted by the four independent variables.

Table 8. Model summary.

Model	*R* Square	*F*	*P*-Value	Durbin–Watson
Entered	0.770	17.538	0.000[a]	2.405

Note: a. Predictors: (Constant), pop_2, Road, office_Department, buses.

Table 9 displays the coefficient of each variable. All of variables are significant: road distance, bus line, office and store, and populations.

Table 9. Model coefficients.

Variable	**Variable name**	**Coefficients**	***t***	***P*-Value**
Road Distance	Road	0.028	2.172	0.0414
Buses Line	buses	−0.066	−4.317	0.0003
Office and Store	Office_Department	0.001	2.083	0.0496
Populations	pop_2	5.341×10^{-5}	2.319	0.0305

4.2. Forecasting number of passengers for MRT purple line

From the analysis in Section 4.1, it is possible to forecast the number of passengers through 2030 and use this number to evaluate the GHG emission reduction for the MRT Purple Line Project. Our analysis revealed that DDRM outperformed traditional regression models in predicting ridership changes. Specifically, DDRM achieved a prediction accuracy within 10% of actual ridership figures. Additionally, DDRM's ability to incorporate detailed socio-economic and infrastructural data provided a more comprehensive understanding of the factors driving ridership. These results underscore the practical advantages of DDRM in urban transit planning.

According to the Nonthaburi Provincial Administrative Organization (Office of the National Economic and Social Development, Bangkok Thailand, 2017), the number of bus routes, office and store area, and road size from 2017 to 2018 will continue to be at the same scale in predicting passenger numbers in 2023 and 2030 as the provincial administrative policy continues. The organization indicated that there was no plan to alter the area to within 500 m and to include the number of bus lines that have not changed and the road expansion policy for the next ten years (2020–2030). As a result of comparing the actual number of passengers in 2019 and the value predicted by this model, the difference was within 10%, as shown in **Table 10**. The projected passenger numbers for 2023 and 2030 applying DDRM are shown in **Table 11**.

Table 10. Actual and model-predicted passenger numbers in 2019.

Station	**Actual**	**Model**	**% Different**
PP14	1,002,207	1,118,520.63	+10%
PP15	1,038,545	1,032,625.87	−1%
PP16	2,720,806	2,916,352.39	+7%
All stations	17,624,794	15,849,465	−10%

Table 11. Passenger projections for the MRT Purple Line through 2030.

Year	**Forecasting from model**
2023	18,410,904
2030	19,782,043

Note: 2016 is not included because the line was not open for the entire year.

4.3. Baseline emission from GHG evaluation

This section discusses the results of a GHG evaluation using the model's passenger projections from 2023 to 2030. Baseline emissions are the emissions potentially generated by the project's passengers in the absence of the project. This includes buses, private automobiles, motorcycles, and taxis. The base emission amount is equal to the total cost of a single-passenger trip. Various types of vehicles were observed beginning at their departure point.

The baseline emissions per passenger are computed based on the vehicle type, the distance travelled in each vehicle type, and the emission factor for each vehicle type. The calculation was broken down into two cycles to facilitate verification. The GHG emissions per person per km were first calculated and converted to $CTBL_{,y}$, as shown in **Table 12**. The $CTBL_{,y}$ value was then multiplied by the projected number of passengers for each year through 2030 to obtain the projected GHG emissions, as shown in **Table 13**.

Table 12. The baseline rate emissions of greenhouse gases per person per kilometer.

Type of Vehicles	BSPi (%)	$EFPKM_{,i,y}$ (gCO_2/Passengers-km)	$CTBL_{,y}$ (gCO_2/Passengers-km)
Buses	53.4	25.5	13.62
Private Car	23.9	99.7	23.83
Taxi	4.1	227.7	9.34
Motorcycle	5.4	36	1.94
Baseline Emission per person per kilometer			56.13

Table 13. The baseline emissions of greenhouse gases based on the predicted number of passengers.

Year	$CTBL_{,y}$ (gCO_2/Passengers-km)	Passengers from DDRM	Baseline emission per year
2023	56.13	18,410,904	10,334.04
2030		19,782,043	11,103.66

Only a single annual CT value is used to represent the absence of GHG emissions from a project.

4.4. Project emission from GHG evaluation

The project's emissions result from the emissions generated by the project's electrification. Project emissions are computed based on the energy used in the railway system. The calculation is divided into two cycles to facilitate verification. GHG emission per person per km were first calculated by multiplying the Purple Line's estimated electricity consumption (EC) by the GHG emission coefficient of the Purple Line, an electrified railway system (EF) and then converted to $CTPJ_{,y}$, (**Table 14**). The $CTPJ_{,y}$ value was then multiplied by the projected passenger numbers for each year through 2030 to obtain the projected GHG emissions if the program was not implemented.

Table 14. The project rate emissions of greenhouse gases per person per kilometer.

Year	$ECElc_{,y}$ (kWh/Passengers-km)	$EFEC_{,y}$ (tCO_2/MWh)	$CTPJ_{,y}$ (gCO_2/Passengers-km)
2023	0.1033	0.4758	49.1275
2030	0.0961	0.4758	45.7223

As shown in **Table 14**, the electricity consumption ($ECELc_{,y}$) of the electric train system in this study from the 2021 annual report was 19,009,690.51 kWh, which was also used in the GHG calculations for 2023 and 2030. This electricity consumption divided by ridership from the DDRM indicates that, in 2023 and 2030, one person will use less electricity to travel the same average distance as in 2030 (**Table 15**). As the number of passengers increases, individual electricity consumption decreases.

Table 15. The project emissions of greenhouse gases based on the predicted number of passengers.

Year	CTPJ,y (gCO_2/Passengers-km)	Passengers from DDRM	Project emission per year (tCO_2/year)
2023	49.1275	18,410,904	9044.81
2030	45.7223	19,782,043	9044.81

4.5. GHG emission reduction from MRT purple line

The results of the MRT Purple Line project's ability to reduce GHG emissions are shown in **Table 16**.

Table 16. The results of the evaluation of the MRT Purple Line Project's reduction of GHG emissions.

Year	Baseline emission per year (tCO_2/y)	Project emission per year (tCO_2/y)	The amount of reductions in greenhouse gas emissions (tCO_2/y)
2023	10,334.04	9044.81	1289
2030	11,103.66	9044.81	2059

5. Conclusion

The findings from this study have several practical implications for urban transit planning and implementation. First, the accurate prediction of ridership using the Direct Demand Ridership Model (DDRM) can significantly improve the planning and optimization of transit routes and schedules. By understanding the key factors that influence ridership, transit authorities can better allocate resources, enhance service efficiency, and meet passenger demand more effectively. For example, the positive correlation between population density and ridership suggests that expanding transit services in densely populated areas can maximize utilization and operational efficiency.

Furthermore, the detailed analysis of variables such as road length, bus lines, and surrounding land use provides actionable insights for infrastructure development. Planners can use these insights to design transit-oriented developments (TODs) that integrate residential, commercial, and recreational spaces with transit services, thereby encouraging higher ridership and reducing reliance on private vehicles. This approach not only improves accessibility but also promotes sustainable urban growth.

The main purpose of this study is to estimate the exact amount of greenhouse gas reduction by applying DDRM, which was developed by considering as many factors as possible based on previous studies, to predict the number of passengers for the Thailand MRT Purple Line project. In this study, variables that affect the number of passengers per station per year were considered through three factors: area utilization around the station, population around the station, and characteristics and services of the public transportation system. This research assumes the variables that affect the number of passengers on the MRT Purple Line Project. In the ridership prediction equation, the variables that increase ridership are the length of the road around the station, office and shopping mall areas, and the number of people around the station. However, transportation characteristics have resulted in a decrease in the number of passengers on the MRT Purple Line.

According to the evaluation of GHG reduction from the MRT Purple Line Project using the predicted number of passengers, the MRT Purple Line Project could reduce GHG emissions by 1289 tCO_2 in 2023 and 2059 tCO_2 in 2030.

Currently, the use of the area surrounding a station depends on its location. For instance, a city center station prioritizes the use of space for offices and commercial areas, whereas stations located outside the city prioritize the use of space for residential areas. The operation of areas surrounding various stations has varying effects on the number of passengers per station per year. However, this study presents variables from the direct demand ridership model that affect the number of passengers per station per year.

DDRM is appropriate for predicting and analyzing the number of passengers because it includes factors of space utilization around stations, such as offices, shopping malls, and road length. The population factors surrounding the stations, and the characteristics and services of the mass transit system became variables for the number of bus routes of each station, train system routes, and terminals and stations connecting buses to other systems.

The following recommendations can be derived from the model:

1) Maximize the use of land around stations by developing office spaces, companies, shopping malls, and accommodations to increase employment and area usage, thus boosting passenger numbers.

2) Optimize bus line management by recognizing that an increased number of bus lines can lead to fewer passengers per line due to more available options, and consider the impact on GHG emissions from buses when planning routes.

3) Coordinate public and private sector efforts to support the development of various types of buses in the study area.

4) Governments should integrate multiple plans to develop a more effective public transportation system and ensure that services provided are beneficial and consistent across different modes of transport.

Author contributions: Conceptualization, KW and HP; methodology, KW and HP; software, KW; validation, KW, DY and HP, writing—original draft preparation, KW; writing—review and editing, DY and HP; funding acquisition, HP. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: This research was supported by the Carbon Neutrality, a specialized program of the Graduate School through the Korea Environmental Industry & Technology Institute (KEITI) funded by Ministry of Environment (MOE, Korea).

Conflict of interest: The authors declare no conflict of interest.

References

Bangkok Mass Transit System Public Company Limited (BTS). (2018) Annual Report 2017–2018. Available online:

https://btsapp1.bts.co.th/WebApplication//WareHouse/UploadPdf/201261132833th_eng_annual_report_2017-2018.pdf (accessed on 26 April 2024).

Boarnet, M. G., Wang, X., & Houston, D. (2017). Can new light rail reduce personal vehicle carbon emissions? a before-after, experimental-control evaluation in Los Angeles. Journal of Regional Science, 57(3), 523–539. https://doi.org/10.1111/jors.12275

Cervero, R., Murakami, J., & Miller, M. (2010). Direct Ridership Model of Bus Rapid Transit in Los Angeles County, California. Transportation Research Record: Journal of the Transportation Research Board, 2145(1), 1–7. https://doi.org/10.3141/2145-01

Chen, F., Shen, X., Wang, Z., et al. (2017). An Evaluation of the Low-Carbon Effects of Urban Rail Based on Mode Shifts. Sustainability, 9(3), 401. https://doi.org/10.3390/su9030401

De Gruyter, C., Truong, L. T., and Taylor, E. J. (2018). Can high quality public transport support reduced car parking requirements for new residential apartments? Journal of Transport Geography, 70, 165–174.

Duduta, N. (2013). Direct Ridership Models of Bus Rapid Transit and Metro Systems in Mexico City, Mexico

Energy Policy and Planning Office (EPPO). (2015). Thailand Power Development Plan 2015–2036. Available online: https://www.eppo.go.th/images/POLICY/ENG/PDP2015_Eng.pdf (accessed on 26 April 2024).

Eom, J., Schipper, L., & Thompson, L. (2012). We keep on truckin': Trends in freight energy use and carbon emissions in 11 IEA countries. Energy Policy, 45, 327–341. https://doi.org/10.1016/j.enpol.2012.02.040

Gojash, O., Fukuda, A. and Seki, Y. (2005). Preliminary study on estimation of ancillary benefits from potential Clean Development Mechanism (CDM) project in transport sector. Journal of the Eastern Asia Society for Transportation Studies, 6, 3225–3240.

Javanmard, M. E., Tang, Y., Wang, Z., and Tontiwachwuthikul, P. (2023). Forecast energy demand, CO_2 emissions and energy resource impacts for the transportation sector. Applied Energy, 338, 120830.

Kain, J. F. (1999). The urban transportation problem: A reexamination and update. In: Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer. Brookings Institution Press.

Litman, T. (2020). Evaluating Public Transportation Health Benefits. Victoria Transport Policy Institute.

Liu, C., Erdogan, S., Ma, T., and Ducca, F. W. (2016). How to increase rail ridership in Maryland: Direct ridership models for policy guidance. Journal of Urban Planning and Development, 142(4), 04016017.

Mass Rapid Transit Authority (MRTA). (2019). MAP Bangkok metro system updated 2019. Available online: https://www.mrta.co.th/en/number-of-passengers-of-the-mrta-project (accessed on 26 April 2024).

Office of Natural Resources and Environmental Policy and Planning (ONEP). (2018). Thailand's Third National Communication. Available online: https://unfccc.int/sites/default/files/resource/Thailand%20TNC.pdf (accessed on 26 April 2024).

Office of the National Economic and Social Development, Bangkok Thailand. (2017). 12th National Economic and Social Development Plan (2017–2021). Available online: https://www.nesdc.go.th/nesdb_en/download/article/article_20170705111314.pdf (accessed on 26 April 2024).

UNCRD. (2010). Environmentally Sustainable Transport Towards Low Carbon Society: THAILAND. Available online: https://uncrd.un.org/sites/uncrd.un.org//files/5th-est_b1g4_thailand.pdf (accessed on 27 April 2024)

UNFCCC. (2008). Kyoto Protocol Reference Manual. Available online: https://unfccc.int/resource/docs/publications/08_unfccc_kp_ref_manual.pdf (accessed on 26 April 2024).

UNFCCC. (2021). Decision-/CMA.3 Rules, modalities and procedures for the mechanism established by Article 6, paragraph 4, of the Paris Agreement. Available online: https://unfccc.int/sites/default/files/resource/cma3_auv_12b_PA_6.4.pdf (accessed on 26 April 2024).

Usvyat, L., Meckel, L., DiCarlantonio, M. and Lane, C. (2009). Sketch model to forecast heavy-rail ridership. In: Proceedings of the 88th Annual Meeting of the Transportation Research Board.

Vergel-Tovar, C. E., & Rodriguez, D. A. (2018). The ridership performance of the built environment for BRT systems: Evidence from Latin America. Journal of Transport Geography, 73, 172–184. https://doi.org/10.1016/j.jtrangeo.2018.06.018

Wang, Z., Ding, J., Wang, L., et al. (2022). Ex-ante and ex-post approaches of evaluating carbon emission reduction in urban rail transit. Mitigation and Adaptation Strategies for Global Change, 27(7). https://doi.org/10.1007/s11027-022-10020-x

Yuan, Z., Yuan, X., Yang, Y., et al. (2023). Greenhouse gas emission analysis and measurement for urban rail transit: A review of research progress and prospects. Digital Transportation and Safety, 2(1), 36–51. https://doi.org/10.48130/dts-2023-0004

Zhang, N., Wang, Z., Chen, F., et al. (2020). Low-Carbon Impact of Urban Rail Transit Based on Passenger Demand Forecast in

Baoji. Energies, 13(4), 782. https://doi.org/10.3390/en13040782