International Journal of Information Technologies and Systems Approach
Volume 17 • Issue 1

# Estimation and Convergence Analysis of Traffic Structure Efficiency Based on an Undesirable Epsilon-Based Measure Model

Xudong Cao, School of Automotive and Transportation Engineering, Hefei University of Technology, China
Chenchen Chen, College of Civil Engineering, Anhui Jianzhu University, China
Lejia Zhang, Guilin University of Electronic Technology, China
Li Pan, UCSI University, Malaysia*

## ABSTRACT

To improve transportation efficiency, this paper analyzes the factors of transportation structure from the two levels of transportation—input and system output. An epsilon-based measure model of non-expected output is introduced, and the environmental benefits of transportation are considered. This model is used to analyze the regional transportation efficiency of 30 provinces and cities in China. Tobit regression and geographically weighted regression are applied to analyze the causes and spatial variation of differences in the efficiency of the transportation structure, and corresponding structural adjustment strategies are proposed. The results show that the regression coefficients of the share of secondary industry output in GDP, population density, and social fixed asset investment exert the most significant effects on transportation structure efficiency. The spatial distribution of sub-variable coefficients shows that spatial heterogeneity exists in the degree of influence of various socio-economic factors on the transportation structure efficiency in different regions.

## KEYWORDS

Epsilon-Based Measure, Environmental Benefit, Geographic Weighted Regression, Tobit Regression, Traffic Structure Efficiency

## INTRODUCTION

The construction and continuous development of the transportation industry greatly promotes the interconnection of regional industries, resources, and people and also enables sustainable economic growth (Prus & Sikora, 2021; Tian et al., 2020). According to the statistical bulletin on the development of the transportation industry published in 2022 by the Ministry of Transport of China, at the end of 2021, China invested 3,622 billion yuan annually into transportation fixed assets, which signifies an increase of 4.1% over the previous year (Ministry of Transport of the People's Republic of China, 2022). The railroad mileage has a length of 150,000 km, of which 40,000 km is suitable for high-speed rail. The national railroad network density is 156.7 km per 10,000 km,$^2$ and the total road mileage is

DOI: 10.4018/IJITSA.332798
*Corresponding Author

5,280,700 km, with a road density of 55.01 km per 100 km.[2] The total road mileage is 5,280,700 km, with a road density of 55.01 km per 100 km,[2] of which the highway mileage is 169,100 km, accounting for 3.2%. The inland waterways have a navigable mileage of 127,600 km and 20,867 berths for port production. Regarding transportation services, the annual operating passenger volume reached 8.303 billion people, which reflects a 14.1% decrease compared with the previous year. Passenger turnover is 1,975,815 million person-km, reflecting an increase of 2.6%. Operating freight volume is 52.160 billion tons, reflecting an increase of 12.3%, and completed cargo turnover is 21.8 trillion-ton km, reflecting an increase of 10.9%.

However, as a pillar industry for social and economic development, the transportation industry is characterized by concentrated investment, intensive energy consumption, and large pollution emissions. In 2020, the total energy consumption of China's transportation industry was about 413 million tons of standard coal, accounting for 8.29% of China's total energy consumption (Ministry of Transport of the People's Republic of China, 2022). In response to the surge of motor vehicle ownership, the total pollutant emissions from motor vehicles have also increased year by year. According to data released by the Chinese Ministry of Ecology and Environment, the total emission of four key pollutants from motor vehicles nationwide in 2020 was 15.930 million tons. Among them, carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides (NOx), and particulate matter (PM) emissions were 7.697 million tons, 1.902 million tons, 6.263 million tons, and 68,000 tons, respectively. Automobiles are the main contributor to total pollutant emissions, and their CO, HC, NOx, and PM emissions exceed 90% of the total motor vehicle emissions. The above data indicate that with increasing investment in the transportation industry and the continuous improvement of the transportation infrastructure, the demand for inter-regional passenger and cargo transportation is also increasing. At the same time, the scale of development of various modes of transportation and the amount of transportation undertaken vary greatly. The potentially resulting imbalances of the transportation structure will not only lead to an unreasonable distribution of social and economic resources, but also cause severe environmental pollution (Du et al., 2022; Li et al., 2021; Qiang et al., 2018). Therefore, in today's important period of transportation restructuring, measuring the efficiency of the transportation structure, continuously optimizing the transportation system, and realizing green, circular, and efficient transportation development have become key issues.

The transportation structure is a pattern of goods and people flow that formed under the specific conditions of spatial layout, population density, economic development, and social and natural environments. The transportation structure reflects the division of labor, organic combination, connectivity, and reasonable layout of the transportation complex. This complex is composed of various modes of transportation in the scope of socialized transportation and a unified transportation process according to technical and economic characteristics (Wei et al., 2021; Gao & Wang, 2021). As a result, the transportation structure can visually reflect the regional transportation strategy (Verma et al., 2021). To measure the rationality of the transportation structure system and the efficiency of coordination and cooperation between transportation modes, many scholars have studied the efficiency of the transportation structure from different perspectives (Lin & Wang, 2022; Zhai et al., 2019; Zeng & Wei, 2021; Kotikov & Kravchenko, 2020). Ganin et al. (2017) defined transportation system efficiency as the average annual delay time car commuters experience during peak hours. They developed and calibrated a model that uses link loads to evaluate traffic delays. Barnum et al. (2011) studied various technical efficiencies and resource allocation efficiencies between different transportation modes from a resource allocation perspective to assess the overall efficiency of urban public transportation. Guo et al. (2018) studied the passenger capacity that can be achieved by public transportation using the inputs of station density, station design, and years of operation to assess the efficiency of transit-oriented development (see also Salman et al., 2022). Existing research and practical experience clearly show that the reasonable allocation of capital investment, the promotion of the division of labor, collaboration among various modes of transportation, and the improvement of the efficiency of the integrated transportation system are urgently needed. These actions can achieve optimal allocation of transportation resources, improve the transportation structure system, and enhance the level of integrated transportation services. Therefore, in this paper, passenger and freight volumes as well as environmental pollutant emissions are selected to measure the efficiency of transportation structure under the premise of also considering the investment in the construction, operation, and maintenance of transportation facilities.

The methods commonly used to study the efficiency of transportation structures can be divided into three main categories: the comprehensive indicator evaluation method (Hu et al., 2022), the parametric method (Liu et al., 2021), and the non-parametric method (Wang et al., 2021; Tu et al., 2022). Among these methods, the comprehensive indicator evaluation method can compress multiple indicators into a few principal components, thus reducing repetition and redundancy among indicators, facilitating analysis and understanding. However, this type of method shows weak performance when dealing with nonlinear relationships among indicators and, further, certain weight settings are subjective (Liu et al., 2020; Abbott, 2014). While the main advantage of the parametric method is that it can separate purely random errors from non-efficiency values, its disadvantage is that it assumes that a specialized functional form connects inputs and outputs (Jofree et al., 2021). Inappropriate functional forms or assumptions about the distribution of error terms can potentially lead to the confusion of setting errors with efficiency estimates. Although the proposed hypothesis function can be verified via hypothesis testing, it can also add considerable unnecessary workload when the form of the proposed hypothesis function is unsatisfactory or when the functional relationship between inputs and outputs is unknown. In contrast, non-parametric methods, such as data envelopment analysis (DEA), are often used to analyze the efficiency (or performance) of individuals or units (Ma et al., 2021). The basic principle of DEA is to keep the input or output of the decision unit constant to determine the relative effectiveness of the production frontier. This form of analysis uses linear programming and statistical data and can judge the relative effectiveness by comparing the degree of deviation of the decision unit from the frontier. Therefore, the DEA method offers advantages when dealing with efficiency evaluation problems with multiple inputs and multiple outputs. This type of method only studies input and output data, without conducting any other data processing and without the need to understand certain expressive relationships between the data. There is no requirement for weights, and the optimal weights are only derived from the actual input-output data of the decision unit, which provides strong objectivity. Moreover, an epsilon-based measure model (EBM) is a hybrid distance function measurement model, which considers the coexistence characteristics of radial and non-radial relaxation variables and combines the advantages of the radial directional distance function model and the non-radial slack-based measure (SBM) model. It can not only calculate the improvement ratio between a target value and the actual value, but also calculate the non-radial values of various input-output values. Therefore, a more realistic efficiency of decision-making units (DMUs) can be calculated.

Based on the above analysis, in this paper, the regional transportation structure efficiency is measured by studying the transportation system structure of each Chinese province. Various transportation modes are considered such as railroad, road, and sea, and a super-efficient non-expectation EBM model with pollutants such as NOx and PM as non-expectation outputs is established. A Tobit regression and a geographically weighted regression (GWR) model are used to analyze the main factors and geospatial distribution characteristics that affect the efficiency of transportation structure. Moreover, transportation structure optimization strategies are proposed.

The main contributions of this paper include the following:

- This paper studies the structural efficiency of a transportation system that is jointly composed of multiple modes of transportation. A comprehensive assessment index system is established from three aspects: input, desired output, and non-desired output. This system considers not only the direct output of passenger and freight traffic, but also the environmental non-desirable output such as air pollutant emissions. The non-desired EBM model is applied to calculate the transportation efficiency of each region.

- The Tobit regression method and the GWR model are used to analyze the causes underlying the differences in the efficiency of transportation structures and the spatial heterogeneity of each factor in detail. Furthermore, based on the results of the analysis of 30 provinces and cities in China, the characteristics of regional transportation structure differences are fully considered and a targeted transportation structure adjustment strategy is proposed to achieve the most reasonable development direction of the transportation structure.

This paper is organized as follows: We first present a brief introduction to the background and significance of transportation structure research, the main methods with which traffic structure efficiency has been analyzed, and the structure and contribution of this paper. We follow with a literature review focused on the three main methods used to assess efficiency, a section on the transportation structure efficiency evaluation model and regression analysis method used in this paper, and a section introducing the composition of the evaluation index system and the selection of regression indicators. Next, the research data are analyzed and traffic efficiency values of each region are obtained through EBM model calculation; moreover, the main factors affecting the regional traffic efficiency and the distribution characteristics are analyzed, and corresponding improvement measures are proposed based on the results of the regression model. In the concluding section, the research results are summarized, including research contributions, policy recommendations, and future research directions.

## LITERATURE REVIEW

### Comprehensive Indicator Evaluation Method

Comprehensive indicator evaluation methods, such as fuzzy evaluation and hierarchical analysis, use multiple indicators to evaluate multiple participating units and are also known as multivariate comprehensive evaluation methods. The basic concept is the transformation of multiple indicators into one indicator that can reflect the comprehensive situation. Sahitya et al. (2021) analyzed the efficiency of traffic network structures in terms of road network connectivity, accessibility, and spatial morphology, using a neuro-fuzzy inference system with multiple linear regression and adaptive networks; they used this system to establish a computational model for evaluating network efficiency based on accessibility of road network structure parameters. Based on the hesitant fuzzy multi-attribute assessment method, Han et al. (2020) analyzed both the multimodal transportation cost and efficiency problems from the perspective of sea and rail intermodal transportation; they also identified the key factors affecting the development of multimodal transportation and provided suggestions such as incorporating the ecological environment into the evaluation system. Zhu and Li (2008) combined the traffic capacity of the urban road environment with traffic structure optimization and established an optimization model through regression analysis with traffic environment capacity as constraint and traffic flow and traffic efficiency as objective. Levinson (2003) considered transportation efficiency from the multiple perspectives of engineers, economists, managers, and planners; in this process, Levinson combined the subjective perspective of travelers with the objective perspective of professionals to construct a multidimensional transportation efficiency evaluation system in the four dimensions of mobility, utility, productivity, and accessibility. Although these methods are easy to operate, the selection of model parameters and weight coefficients is highly subjective, and the role of objective data is weakened.

### Parametric Method

The frontier analysis method is the most used efficiency assessment method in recent years. At its core, frontier analysis determines the external boundary of all possible inputs and outputs (i.e., the production frontier surface) based on known input-output observations. Consequently, all output values lie within the boundary and the distance between each observation and the boundary is the efficiency of that production point. Depending on whether the parameters included in the frontier production function need to be estimated, frontier analysis methods can be divided into parametric methods and nonparametric methods. Among them, parametric methods use an econometric method that utilizes multivariate statistical analysis techniques to determine unknown parameters in the frontier function, which are then used to calculate both the theoretical and actual values. According to different assumptions on the distribution of inefficient terms in the frontier function, the parametric method mainly includes the stochastic frontier approach, the distribution free approach, the thick frontier approach, and the recursive thick frontier approach.

The most widely used parametric method is the stochastic frontier approach (Amin et al., 2021). Liu et al. (2021) used a stochastic frontier approach to evaluate the total factor energy efficiency of China's road transport industry and proposed specific development strategies to increase road investments and develop advanced technologies based on the results of eight economic regions (Liu, et al., 2021; Bin & Sun, 2022). Wanke et al. (2020) studied the relationship between the $CO_2$ emission level of each transport mode and the associated passenger and freight turnover by building a robust Bayesian stochastic frontier analysis model; they then calculated the sustainability efficiency of each transport mode as a basis for the formulation and adjustment of transport policy. Sami et al. (2013) examined the actual operations of 64 road transport operators in 18 countries and used stochastic frontier analysis to identify a more significant degree of influence of operating profit, investment, and firm size on the level of firm efficiency. They found that larger firms with greater investment capacity were able to better exploit scale effects and were technically more efficient than smaller firms. De Borger et al. (2002) systematically summarized the application of stochastic frontier analysis in public transport performance assessments. Lin et al. (2010) included traffic accidents as a non-expected output of public transport systems and found a significant correlation between traffic accidents and public transport efficiency using a stochastic production frontier model. Holmgren (2013) used stochastic frontier analysis to show that much of the decline in the cost efficiency of public transport in Sweden stems from three factors: the diversionary effect of other transport modes, the decrease in relative passenger flows caused by an increase in the density of bus routes, and the increase in costs caused by an increase in service and environmental quality requirements. Other examples of the application of stochastic frontier analysis include analyses of the relationship between high-speed rail development and urban environmental efficiency and of the efficiency of public transport systems based on traffic emissions (Sun et al., 2020; Tamaki et al., 2019; Ankora, 2022; Zhang, 2021; Zhang et al., 2018). Parametric methods can consider the impact of random errors on the output and distinguish random errors from technical inefficiency. However, this type of method is based on assumption in the form of the relationship function between inputs, outputs, and external influencing factors that must be made in advance. If it is difficult to accurately characterize the production-input relationship of the DMU for the preset production function, the evaluation results will deviate substantially from the actual efficiency, and parameter calibration is also exceptionally difficult.

### Non-Parametric Method

Non-parametric methods do not impose strict assumptions on the underlying data distribution when data characteristics are unknown. Therefore, these methods can handle nonlinear relationships among variables more flexibly, provide a more accurate representation of complex data patterns, and avoid the need for explicit parameter estimation or hypothesis testing. In addition, these methods usually rely on the rank or order of data rather than the actual values, making them more robust to outliers and enabling them to resist hypothesis violations. Data envelopment analysis (DEA) is a non-parametric method that is widely used to assess the efficiency of decision-making units (DMUs). It compares the relative efficiency of multiple DMUs by evaluating their inputs and outputs. The main goal of DEA is to identify the most efficient units and suggest improvements for inefficient units.

Therefore, DEA is often used for traffic efficiency assessments. Khanh Van et al. (2022) studied the spatial efficiency of the transportation system using the Charnes-Cooper-Rhodes model and showed that the remaining capacity of parking lots in Sapporo City reached 27.81%, and the density of bus stops and roads exceeded demand by about 14%. Li et al. (2022) used a three-stage DEA model to assess the efficiency of the transportation industry and used a panel vector autoregressive model to analyze the interaction between traffic structure, transportation efficiency, and regional economic development. Karlaftis (2004) used DEA to demonstrate that the efficiency and effectiveness of a transportation system are positively correlated and that the optimal size of a transportation system varies with the evaluation metric. In further research, DEA was applied to assess the operational efficiency of airports and similar conclusions were obtained across various studies, namely that management and operation systems, system size, and technological change have a large impact on the operational efficiency of airports (Merkert & Hensher, 2011; Sukte et al., 2022; Lai et al., 2015; Henke et al., 2022). Hirschhausen and Cullman (2010) conducted a comparison of the technical efficiency of 179 public transport bus companies in Germany over a 15-year period, showing that economies of scale are not linearly related to production efficiency; moreover, production efficiency tends to decline when economies of scale are larger, thus indicating that the industrial structure can be adjusted through economies of scale. Deng and Yan (2019) applied the EBM model to optimize bus routes and departure frequencies. In other research, DEA was applied to evaluate the transfer efficiency of urban bus systems, urban rail systems, and both systems to uncover the influencing factors and improve the efficiency of transportation systems (Yao et al., 2019; Lee et al., 2019; Ma & Zhang, 2022; Taboada & Han, 2020).

A comparison of parametric and non-parametric methods shows that the parametric method is based on the assumption that the form of the relationship function between inputs, outputs, and external influencing factors is known in advance; therefore, if it is difficult for the predefined function to accurately describe the production-input relationship characteristics of the DMU, the efficiency evaluation results of the production function will show a large deviation from the actual production efficiency. It is also extremely difficult to calibrate the parameters in the function because the sample distribution characteristics are unknown. In contrast, nonparametric methods do not need to determine the form of the production function and parameter estimation, but rather identify the production frontier surface through linear programming. These methods can calculate the distance from sample points to the production frontier surface to determine the production efficiency. Although the underlying algorithm is more complex, it does not rely on subjective weights and can rely entirely on objective information to address the efficiency problem of multiple inputs and outputs. Nonparametric methods can also solve the relative efficiency problem with varying magnitudes.

In summary, most existing studies on transportation efficiency focus on single transportation modes, such as airports, railways, and buses, or on technical efficiency assessments of specific transportation enterprises. However, with the progress of transportation technology, integrated transportation systems have gradually formed. These represent the trend of future transportation restructuring, and there is still much to be explored by research on transportation structure efficiency assessment. In addition, with increasing public awareness of environmental protection, the concept of green transportation is also an important goal of transportation development; therefore, environmental benefits should be included in the evaluation system when measuring the efficiency of transportation systems.

## METHODS

### Super-EBM-Undesirable Model

DEA methods are usually divided into two categories: radial and non-radial methods. Radial methods, such as Charnes-Cooper-Rhodes and Banker-Charnes-Cooper (Kammoun, 2018), are based on Debreu-Farrell economic theory and determine the magnitude of efficiency by comparing the projected distance from an inefficient DMU to a relatively efficient production frontier. An inefficient DMU reaches the efficient frontier radially by reducing inputs or increasing outputs at the same proportion. In many practical cases, the gap between the DMU and the strongly efficient unit should be considered in addition to the equal proportional improvement component and the slack improvement component. To compensate for the shortcomings of the radial model, Tone (2001) proposed a non-radial DEA model based on slack variable measures rooted in Pareto-Koopmans economic theory, called slack-based measure (SBM). Domagała and Kadlubek (2023) established an SBM-undesirable model by incorporating the parameters related to non-desired outputs into the objective function. However, when assessing the efficiency of a system that has a large number of input and output indicators, multiple DMUs are usually effective. These effective DMUs have the same efficiency values, making further differentiation difficult. As a solution, Anderson and Petersen (1993) proposed a super-efficiency model to further evaluate the degree of effectiveness for effective or weakly effective DMUs (Andersen & Petersen, 1993; Jiang et al., 2020). However, the SBM model adjusts non-equal scaling for different inputs or outputs, and although it circumvents the assumption of the radial scaling of input factors, it loses the original scaling information of the projection value of the efficiency frontier. The optimal relaxation of the SBM model taking zero and positive values has a clearly different linear programming solution process. Based on the above problems, an EBM model that integrates radial and non-radial characteristics can effectively solve the problems existing in the SBM model to measure the efficiency fraction. Therefore, this paper adopts the non-expected output and non-oriented super-efficient EBM model to measure the efficiency of the traffic structure.

Now we take each province as a decision unit and construct the optimal frontier surface of transportation inputs and outputs for different provinces. Suppose there are n decision units, denoted as $D_j$ $(j=1,2,\ldots,k,\ldots,n)$. Each decision unit has m inputs, $(x_1,x_2,\ldots,x_i,\ldots x_m)\in R$, there are p desired outputs, $(y_1,y_2,\ldots,y_r,\ldots y_p)$ $R$, and q undesired outputs, $(b_1,b_2,\ldots,b_t,\ldots b_q)$ $R$. Then the non-desired output, non-oriented super-efficient EBM model is as follows:

$$
\begin{aligned}
&\rho_k^{\ *} = \min \frac{\theta - \varepsilon_x \sum_{i=1}^{m} \frac{W_i^- S_i^-}{x_{ik}}}{\varphi + \varepsilon_y \sum_{r=1}^{p} \frac{W_r^+ S_r^+}{y_{rk}} + \varepsilon_b \sum_{t=1}^{q} \frac{W_t^{b-} S_t^{b-}}{b_{tk}}} \\
&s.t. \quad \sum_{\substack{j=1 \\ j\neq k}}^{n} x_{ij}\lambda_j + S_i^- = \theta x_{ik} \\
&\qquad \sum_{\substack{j=1 \\ j\neq k}}^{n} y_{rj}\lambda_j - S_r^+ = \varphi y_{rk} \\
&\qquad \sum_{\substack{j=1 \\ j\neq k}}^{n} b_{tj}\lambda_j + S_t^{b-} = \varphi b_{tk} \\
&\qquad \sum_{i=1}^{m} W_i^- = \sum_{r=1}^{p} W_r^+ = \sum_{t=1}^{q} W_t^{b-} = 1 \\
&\qquad \lambda_j \geq 0, S_i^-, S_r^+, S_t^{b-} \geq 0, \theta \leq 1, \varphi \geq 1 \\
&\qquad W_i^-, W_r^+, W_t^{b-} \geq 0
\end{aligned} \tag{1}
$$

where $\rho_k$* is the optimal efficiency considering the non-desired output case, $\theta$ is the efficiency value calculated by the radial model, $\varepsilon_x,\varepsilon_y,\varepsilon_b$ are the weight coefficients representing the non-radial part in the super-efficient EBM model, taking values in the range of [0,1], when $\varepsilon_x$, $\varepsilon_y$, $\varepsilon_b$=0, the EBM model is equivalent to the radial model, and when $\varepsilon_x,\varepsilon_y,\varepsilon_b$=1, the EBM model is equivalent to the weighted SBM model. $x_{ik},y_{rk},b_{tk}$ denote the $i$-th input, $r$-th desired output, and $t$-th undesired output of decision unit $k$, respectively; $S_i^-$, $S_r^+$, $S_t^{b-}$ are input slack variables, desired output slack variables, and undesired output slack variables, respectively; $W_i^-$, $W_r^+$, $W_t^{b-}$ are the weights of each input, desired output, and undesired output indicator, respectively. $\lambda_j$ denotes the linear combination coefficient of decision unit $j$, and $\varphi$ is the output expansion ratio.

Before calculating the optimal efficiency, the correlation matrix **P** of the input variables is constructed according to Tone and Tsutsui (Tone & Tsutsui, 2010):

$$P(i_1,i_2) = 1 + R(x_{i_1},x_{i_2})/2 \tag{2}$$

This denotes the Pearson correlation coefficient between the input index $x_{i_1}$ and $x_{i_2}$, $1 \le i_1,i_2 \le m$.

Then, the maximum eigenvalue of the correlation matrix **P** and the corresponding eigenvector $\mathbf{W}_i$=[$W_i$] are found and substituted into Equation (3) to obtain $\varepsilon_x$ and $W_i^-$:

$$\varepsilon_x = \begin{cases} \dfrac{m-\beta_x}{m-1} & (m>1) \\ 0 & (m=0) \end{cases}$$
$$W_i^- = \frac{W_i}{\sum_{i=1}^{m} W_i} \tag{3}$$

## Regression Analysis

Many factors affect the efficiency of transportation structures, and these factors are unevenly and discontinuously distributed in space. The traditional linear regression model is based on "smoothness assumption embedding," and it is difficult to estimate the spatial heterogeneity of socio-economic factors. Tobit regression and GWR models can effectively determine the strength of influencing factors, while detecting spatial non-stationarity and allowing local weight estimation to adjust the relationship between variables with spatial location changes, which is a more realistic approach.

First, this paper adopts the Tobit model with restricted dependent variables for regression, using the structural efficiency value as explained variable and influencing factors as explanatory variables to establish a Tobit regression analysis model. This model is then used to judge the direction and intensity of influencing factors on the structural efficiency of transportation through the coefficients of the explanatory variables and to guide the formulation of transportation restructuring strategies.

The Tobit regression model is shown in Equation (4):

$$\rho_j = \begin{cases} F_j\beta^T + \varepsilon_j, & F_j\beta^T + \varepsilon_j > 0 \\ 0\ , & F_j\beta^T + \varepsilon_j \le 0 \end{cases} \tag{4}$$

where $\rho_j$ is the efficiency value of the $j$-th decision unit, $\boldsymbol{F}_j$ is the vector of influencing factors of the $j$-th decision unit, $\boldsymbol{\beta}$ is the regression parameter vector, $\varepsilon_j$ is the error term perturbation of the $j$-th decision unit, and $\varepsilon_j \sim N(0, \sigma^2)$, $\sigma^2$ is the variance. In this paper, the regression parameter vector of the model is solved by the great likelihood method.

To address the problem of spatial non-stationarity and spatial dependence of factors influencing traffic efficiency, this paper adjusts the global regression coefficients with local parameter estimation by introducing a GWR model to fully consider spatial heterogeneity. This approach is more conducive to exploring inter-regional differences and correcting regression results. The GWR model is shown in Equation (5):

$$\rho_j = C_0(\mu_j, \nu_j) + \sum_{l=1}^{h} C_l(\mu_j, \nu_j) f_{jl} + \varepsilon_j \tag{5}$$

where $(\mu_j, \nu_j)$ is the geographic coordinates of the $j$-th decision unit, longitude $\mu_j$ and latitude $\nu_j$; $f_{jl}$ is the $l$-th influence factor of the $j$-th decision unit; $C_l\ (\mu_j, \nu_j)$ is the value of the regression coefficient of the geographic location regression function $C(\mu, \nu)$ on the $l$-th influence factor of the $j$-th decision unit, where $C_0\ (\mu_j, \nu_j)$ is the initial term of the geographic location regression function; and $h$ is the number of influence factors.

The regression coefficient of the $j$-th decision can be expressed by the geographic location regression function, as in Equation (6):

$$C(\mu_j, \nu_j) = (F^T W_j F^T)^{-1} F^T W_j \rho \tag{6}$$

where $\boldsymbol{C}(\mu_j, \nu_j)$ is the vector composed of the values of the geographic location regression function $C(\mu, \nu)$ at the $j$-th decision unit; $\boldsymbol{W}_j$ is the weight matrix of the $j$-th decision unit to describe the weights of other decision units affecting the $j$-th decision unit, as shown in Equation (7); $\boldsymbol{\rho}$ is the vector composed of the efficiency values of all decision units, $\boldsymbol{\rho}=[\rho_1, \rho_2, \ldots, \rho_n]$; $w_{jn}$ is the weight of the influence of decision unit $n$ on decision unit $j$; and $\boldsymbol{F}$ is the matrix composed of the influence factors:

$$W_j = \begin{bmatrix} w_{j1} & & & \\ & w_{j2} & & \\ & & \cdots & \\ & & & w_{jn} \end{bmatrix} \tag{7}$$

To avoid estimation errors caused by sparse data from neighboring samples of sample points, a Gaussian kernel function is used to determine the weights, as shown in Equation (8):

$$w_{jn} = \exp(-(d_{jn} / b)^2) \tag{8}$$

where $d_{jn}$ is the Euclidean distance between decision unit $j$ and decision unit $n$ and $b$ is the bandwidth, a non-negative decay parameter used to describe the functional relationship between the weights and the distance, determined using the minimization of the deficit pool information AICc method.

## INDICATOR DESIGN

The assessment subjects of the efficiency of the regional transportation structure can be divided into intra-city transportation system and inter-city passenger and freight transportation system. The intra-city transportation system is mainly used for passenger transportation and can be divided into two parts: road transportation and rail transportation. The intercity transportation system mainly includes five modes of transportation: water transportation, air transportation, pipeline transportation, road transportation, and railroad transportation. This paper focuses on road transportation and railroad transportation, as these have a large capacity and are widely used. Water transportation, air transportation, and pipeline transportation have the characteristics of strong specialization, poor universality, weak substitutability, and strong environmental dependence. Taking the transportation structure system of each Chinese province as the DMU, the input-output index system of regional transportation structure efficiency evaluation is established, as shown in Figure 1.

In terms of input, urban road network density $x_1$ and rail mileage $x_2$ characterize the accessibility and efficiency of the urban transportation skeleton network. Road mileage $x_3$ and rail mileage $x_4$ directly reflect the intercity passenger and freight communication capacity. The operation of facilities requires inputs of manpower, equipment, and energy; therefore, the number of public transportation vehicles per 10,000 people $x_5$, total energy consumption $x_6$, and total number of employees $x_7$ are selected as system operation cost input indicators.

On the output side, urban passenger volume $y_1$ and freight volume $y_2$ are quantitative indicators that reflect the transport industry's service to the national economy and people's life, as well as directional indicators that reflect the scale and speed of transport development. However, motor vehicle pollution has become an important source of air pollution in China. It is an important cause of haze and photochemical smog pollution, in which excessive emissions of NOx $b_1$ and PM $b_2$ seriously affect the health of residents and the ecosystem.

In addition to the development of the transportation mode itself, the efficiency of the regional transportation structure is influenced by the related industrial structure and other factors within the socio-economic system. In this paper, seven indicators (chosen from demographic factors, economic level, industrial structure, and resource input) were finally selected as independent variables for the regression, as shown in Table 1.

## RESULT AND DISCUSSION

### Analysis of Traffic Structure Efficiency

In this paper, input-output data of transportation structures of 30 provinces in China are selected. An EBM-undesirable model with variable payoffs of scale for super-efficiency is established to comprehensively analyze the efficiency of regional transportation structures in each province, city, and autonomous region. The data were obtained from the 2018 *China Statistical Yearbook*, *China City Statistical Yearbook*, *China Urban Construction Statistical Yearbook*, *China Vehicle Environmental*

**Figure 1. Input and output index system**

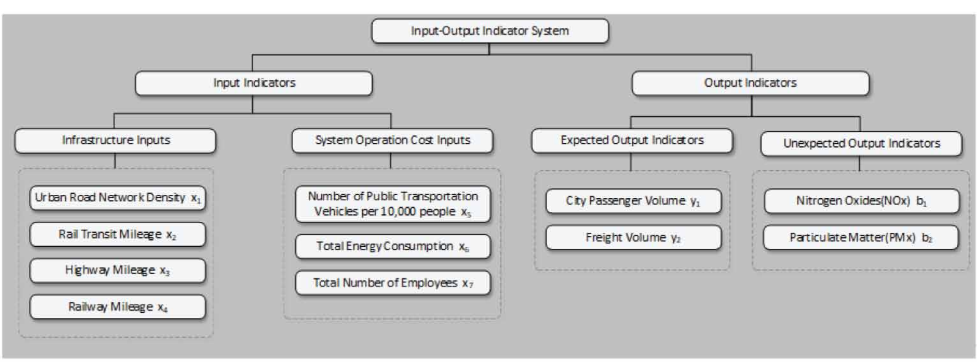


**Table 1. Definition of regression variables**

| Variable Type | Variable Name | Symbol |
|---|---|---|
| Dependent Variables | Transportation Structure Efficiency | TSE |
| Independent Variables | Density of Population / (persons /km$^2$) | DP |
| | Per capita Gross Domestic Product / yuan | PGDP |
| | Total Investment in Fixed Assets / billion yuan | TIFA |
| | Area of Land for Transportation / thousand hectares | ALT |
| | The Secondary Industry Proportion of Output Value to GDP /% | INP |
| | Total Investment in Environmental Pollution Control/ billion yuan | IEPC |
| | Number of Public Transportations per 10000 People /vehicle | TPB |

*Management Annual Report*, *China Energy Statistical Yearbook*, and national statistical yearbooks of provinces, cities, and departments. Using MaxDEA ultra7.0 software, the super-efficiency values of each province and regional traffic structure were calculated, and the results are shown in Tables 2 and 3.

Overall, a total of 22 provinces have an overall efficiency value of 1, reaching DEA validity and accounting for 73.33% of the total. This result indicates that most of the provinces in China have high transportation efficiency and basically meet the overall transportation needs of the society under the condition of limited resource input. The top five provinces are Shanghai (1.221), Guangdong (1.155), Yunnan (1.127), Jiangxi (1.083), and Anhui (1.074), while eight provinces (Qinghai, Henan, Hainan, Guangxi, Shanxi, Hubei, Xinjiang, and Jiangsu) have an overall efficiency value below 1 for their transportation structure. These provinces do not reach the DEA effective production frontier, and their transportation structure needs to be optimized.

In terms of regional distribution, geographical differences are more apparent in the overall efficiency of transportation. East China and Southwest China have a higher overall efficiency of transportation structure, while Northwest and Central China have a lower overall transportation efficiency. This result indicates that regional development achieves a certain convergence, and the synergistic and driving effects of transportation development in similar areas can be observed. The natural geographical barrier also results in substantial differences in traffic structures.

On the one hand, regions with a high degree of socio-economic development (such as Shanghai and Guangdong) have high population concentration and strong transportation demand. Their well-developed transportation infrastructures can bear the greater pressure of passenger and freight transportation, while system operation costs and negative environmental effects can be effectively controlled because of their relatively high management and technical level. On the other hand, Anhui, Jiangxi, Yunnan, and other fast-developing economic regions invest their limited transportation infrastructure construction funds in key areas with the greatest transportation benefits, such as coal transportation, agricultural products transportation, and tourism. Based on the more developed and reliable railroad transportation system, these regions strengthen the construction of road transportation corridors between key cities, maximize the characteristics of large capacity and stable operation, and use both existing and new transportation facilities to meet major passenger and freight needs. However, most of the provinces with less efficient transportation structures are still in the stage of incremental

**Table 2. Efficiency of traffic structure**

| DMU | EBM | Overall Efficiency | Pure Technical Efficiency | Scale Efficiency | Size Compensation | Rank |
|---|---|---|---|---|---|---|
| Beijing[1] | 1.065 | 1.000 | 1.000 | 1.000 | Unchanged | 7 |
| Tianjin[1] | 1.052 | 1.000 | 1.000 | 1.000 | Unchanged | 8 |
| Hebei[1] | 1.035 | 1.000 | 1.000 | 1.000 | Unchanged | 12 |
| Shanxi[1] | 0.940 | 0.940 | 1.000 | 0.940 | Incremental | 24 |
| Inner Mongolia[1] | 1.009 | 1.000 | 1.000 | 1.000 | Unchanged | 20 |
| Liaoning[2] | 1.013 | 1.000 | 1.000 | 1.000 | Unchanged | 17 |
| Jilin[2] | 1.023 | 1.000 | 1.000 | 1.000 | Unchanged | 14 |
| Heilongjiang[2] | 1.046 | 1.000 | 1.000 | 1.000 | Unchanged | 9 |
| Shanghai[3] | 1.221 | 1.000 | 1.000 | 1.000 | Unchanged | 1 |
| Jiangsu[3] | 0.845 | 0.845 | 0.849 | 0.995 | Decreased | 28 |
| Zhejiang[3] | 1.079 | 1.000 | 1.000 | 1.000 | Unchanged | 11 |
| Anhui[3] | 1.038 | 1.000 | 1.000 | 1.000 | Unchanged | 5 |
| Fujian[3] | 1.074 | 1.000 | 1.000 | 1.000 | Unchanged | 19 |
| Jiangxi[3] | 1.009 | 1.000 | 1.000 | 1.000 | Unchanged | 4 |
| Shandong[3] | 1.083 | 1.000 | 1.000 | 1.000 | Unchanged | 10 |
| Henan[4] | 0.860 | 0.860 | 1.000 | 0.860 | Incremental | 29 |
| Hubei[4] | 0.945 | 0.945 | 0.951 | 0.993 | Incremental | 23 |
| Hunan[4] | 1.012 | 1.000 | 1.000 | 1.000 | Unchanged | 18 |
| Guangdong[5] | 1.155 | 1.000 | 1.000 | 1.000 | Unchanged | 2 |
| Guangxi[5] | 0.863 | 0.863 | 1.000 | 0.863 | Incremental | 27 |
| Hainan[5] | 0.808 | 0.808 | 1.000 | 0.808 | Incremental | 28 |
| Chongqing[6] | 1.029 | 1.000 | 1.000 | 1.000 | Unchanged | 13 |
| Sichuan[6] | 1.021 | 1.000 | 1.000 | 1.000 | Unchanged | 15 |
| Guizhou[6] | 1.073 | 1.000 | 1.000 | 1.000 | Unchanged | 6 |
| Yunnan[6] | 1.127 | 1.000 | 1.000 | 1.000 | Unchanged | 3 |
| Shaanxi[7] | 1.020 | 1.000 | 1.000 | 1.000 | Unchanged | 16 |
| Gansu[7] | 1.000 | 1.000 | 1.000 | 1.000 | Unchanged | 22 |
| Qinghai[7] | 0.781 | 0.781 | 1.000 | 0.781 | Incremental | 30 |
| Ningxia[7] | 1.003 | 1.000 | 1.000 | 1.000 | Unchanged | 21 |
| Xinjiang[7] | 0.872 | 0.872 | 0.872 | 0.999 | Incremental | 25 |

Note: Superscripts 1 to 7 represent North China, Northeast China, East China, Central China, South China, Southwest China and Northwest China.

scale payoff, making it difficult to continuously improve transportation efficiency because of the scale of transportation infrastructure construction and the level of resource input.

For the above-mentioned eight provinces that did not reach the integrated efficiency, input-output projection analysis was used to calculate the input redundancy rate, output deficiency rate, and non-desired output redundancy rate of each province. The results are shown in Table 4.

**Table 3. Regional traffic efficiency analysis**

| Area | Average Over-Efficiency Value | Regional Ranking |
|---|---|---|
| North China | 1.045 | 4 |
| Northeast China | 1.056 | 3 |
| East China | 1.12 | 2 |
| Central China | 0.906 | 6 |
| South China | 0.988 | 5 |
| Southwest China | 1.138 | 1 |
| Northwest China | 0.920 | 7 |

**Table 4. Adjustment rate of input and output**

| Projection | Metric | Shanxi | Jiangsu | Henan | Hubei | Guangxi | Hainan | Qinghai | Xinjiang |
|---|---|---|---|---|---|---|---|---|---|
| Input redundancy rate (%) | Urban Road Network Density | 25.28 | 55.86 | 3.33 | 47.75 | 22.75 | 91.41 | 74.24 | 36.00 |
| | Rail Transit Mileage | 0.00 | 6.14 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Highway Mileage | 0.00 | 11.69 | 12.53 | 24.85 | 0.00 | 0.00 | 39.38 | 0.00 |
| | Railway Mileage | 39.27 | 0.00 | 25.86 | 16.75 | 46.71 | 37.08 | 62.89 | 42.71 |
| | Number of Public Transportation Vehicles per 10,000 people | 0.00 | 24.38 | 0.00 | 0.00 | 0.00 | 85.77 | 79.08 | 24.58 |
| | Total Energy Consumption | 33.93 | 0.00 | 29.81 | 11.61 | 19.93 | 32.02 | 0.00 | 24.21 |
| | Total Number of Employees | 33.31 | 22.54 | 33.92 | 21.83 | 22.10 | 62.07 | 12.31 | 0.00 |
| Expected output shortfall rate (%) | City Passenger Volume | 0.00 | 13.24 | 5.70 | 3.97 | 42.58 | 22.55 | 23.80 | 12.45 |
| | Freight volume | 0.00 | 13.24 | 9.75 | 13.41 | 10.76 | 31.19 | 45.51 | 15.07 |
| Non-desired output redundancy rate (%) | Nitrogen oxides | 28.60 | 34.95 | 50.66 | 3.97 | 20.27 | 22.54 | 23.80 | 12.45 |
| | Particulate matter | 22.35 | 32.29 | 45.23 | 19.11 | 10.76 | 32.71 | 35.47 | 30.54 |

According to the results, in most provinces, there is a certain degree of input redundancy in the density of urban road networks, railroad mileage, and the number of employees. Specifically, urban road traffic and railroad transportation are characterized by high levels of frequency of use, capacity, economy, and practicality; therefore, they satisfy most of the passenger and freight demand, and the government's investment in roads and railroads increases year by year, resulting in a certain degree of input redundancy. At the same time, maintaining the normal operation of road and railroad systems will lead to the over-investment of human resources, resulting in an increase of operating costs and a decrease of the average efficiency of individuals. Therefore, from the perspective of input redundancy, these provinces and cities should appropriately reduce their large-scale investment in roads and railroads in non-major cities and should direct more social resources to the development of urban public transportation systems with high capacity and high efficiency, such as railways and buses. They should also further reduce resource wastage by streamlining organizations and personnel, increasing vocational skills training, and improving the efficiency evaluation system.

In addition, negative externalities of transportation can lead to an increase in overall transportation costs for society as well as a loss of environmental benefits. As shown in Table 3, all eight provinces that did not achieve effective overall efficiency have redundant non-desirable outputs. High non-desirable outputs weaken transportation efficiency and cause environmental pollution problems. Therefore, to control the negative environmental effects of transportation structure, provinces and cities should reasonably adjust the transportation energy structure, encourage the development of new energy sources and new technologies, reduce pollutant emissions, and improve transportation efficiency.

Different provinces and cities have very different transportation structures because of their different geographical and socio-economic environments; therefore, targeted structural adjustment strategies need to be proposed. For example, Shanxi Province is rich in mineral resources and is located in the Loess Plateau in western North China; its transportation structure is mainly based on the high-capacity railroad system, while the urban public transportation construction clearly suffers from under-investment. This has resulted in an imbalanced transportation structure. Therefore, Shanxi Province should adjust its energy structure at the right time, control the use and transportation of non-renewable resources, invest in the development of tourism and ecological economy, and increase the construction of special and convenient highways and urban public transportation systems. Jiangsu Province is located in the core area of the Yangtze River Delta and has a developed economy and superior transportation location; however, the analysis presented in Table 4 shows that the province fails to achieve the desired output with redundant inputs, and the non-desired output is not well controlled. Therefore, Jiangsu Province should reasonably plan its resource allocation, focus on improving the capacity and transportation efficiency of different transportation modes, and appropriately reduce bus lines and frequencies with low operational efficiency according to the traffic demand. Further, Jiangsu Province should design customized bus lines and assist the transportation industry in its gradual transformation from traditional energy to new energy.

## Regression Analysis of Traffic Structure

### *Tobit Regression*

In this paper, the seven indicators shown in Table 1 (such as population density and social fixed asset investment) are used as independent variables and the efficiency of transportation structure is used as dependent variable. Tobit regression analysis is used to determine the degree of influence of each factor on the regional transportation structure and the direction of adjustment of the transportation structure. To ensure that there is no redundancy of information in the independent variables that could affect the model regression results, a covariance test was run for the independent variables, and the model test results are shown in Tables 5 and 6.

The Pearson correlation coefficients among the indicators in Table 5 imply that the per capita gross regional product, the number of buses per 10,000 people, the total investment in environmental pollution control, and the area of land for transportation are correlated. Because the vitality of the regional economy is linked to an increase of the gross regional product, the government shares the benefits brought by socio-economic activities through taxation and other means; the government further invests in the construction of basic service facilities such as transportation, medical care, and environmental protection. The number of buses owned intuitively reflects the investment in transportation facilities and, therefore, there is a stronger correlation between the gross regional product and the development of the transportation system. The negative externalities of the transportation

**Table 5. Pearson correlation coefficient**

| Independent Variable | DP | PGDP | TIFA | ALT | INP | IEPC | TPB |
|---|---|---|---|---|---|---|---|
| DP | 1.00 | -0.23 | -0.06 | 0.03 | 0.04 | -0.22 | -0.25 |
| PGDP | -0.23 | 1.00 | 0.01 | -0.18 | -0.23 | 0.28 | 0.71② |
| TIFA | -0.06 | 0.01 | 1.00 | 0.57② | 0.25 | 0.48② | 0.09 |
| ALT | 0.03 | -0.18 | 0.57② | 1.00 | 0.38① | 0.63② | -0.17 |
| INP | 0.04 | -0.23 | 0.25 | 0.38① | 1.00 | 0.19 | -0.31 |
| IEPC | -0.22 | 0.28 | 0.48② | 0.63② | 0.19 | 1.00 | 0.36① |
| TPB | -0.25 | 0.71② | 0.09 | -0.17 | -0.31 | 0.36① | 1.00 |

Note: ▯ and ▯ indicate significance at the 5% and 1% levels.

**Table 6. Collinearity test**

| Independent Variable | DP | PGDP | TIFA | ALT | INP | IEPC | TPB |
|---|---|---|---|---|---|---|---|
| Average value | 2849 | 61575 | 1887 | 126 | 0.407 | 317 | 14.222 |
| Standard deviation | 1077 | 27747 | 1164 | 61 | 0.076 | 225 | 3.269 |
| Tolerance | 0.897 | 0.482 | 0.626 | 0.339 | 0.772 | 0.352 | 0.375 |
| VIF | 1.114 | 2.073 | 1.597 | 2.947 | 1.296 | 2.844 | 2.670 |

Note: VIF denotes variance inflation factor.

system are also evidenced by the large correlation between investment in environmental pollution and the area of land used for transportation. The correlation coefficients among most of the other indicators are less than 0.3, indicating that there is no clear correlation among indicators.

To further corroborate the existence of covariance or multiple covariance among independent variables, and to determine whether the correlation of specific indicators shown in Table 5 affects the precision of the regression results, the tolerance of variables and the variance inflation factor need to be calculated to verify the covariance of indicators. The variance inflation factor tests the regression model for severe multicollinearity problems by calculating the ratio of the variance of the regression coefficient estimates compared to the variance if no linear correlation is assumed between independent variables. When the tolerance of independent variables exceeds 0.1 and the variance inflation factor is less than 10, there is no problem of covariance between the independent variables. Otherwise, serious covariance exists between this independent variable and the other independent variables, and this variable should be eliminated. The results presented in Table 6 show that the standard deviations of each independent variable are relatively small, indicating that the distribution of variable values in each region is relatively concentrated. Among all indicators, that with the largest variance inflation factor is the area of land used for transportation, with a value of 2.947, which is far lower than 10. The tolerance of all indicators is greater than 0.3. Therefore, it can be assumed that there is no covariance problem between independent variables. Each independent variable can be used for Tobit regression analysis.

After the model passed the covariance test, Tobit regression analysis was conducted using Eviews 9.0. According to the regression results shown in Table 7, the four independent variables (i.e., population density, the proportion of the output value of secondary industry to GDP, the total investment in environmental pollution control, and the number of buses per 10,000 people) have

**Table 7. The regression results of Tobit**

| Independent Variable | DP | PGDP | TIFA | ALT | INP | IEPC | TPB |
|---|---|---|---|---|---|---|---|
| Regression coefficient | 0.4860② | 0.0140 | 0.1552 | 0.0059 | 0.9109③ | -0.0285① | 0.0271③ |
| Standard deviation | 0.2135 | 0.0120 | 0.2493 | 0.0062 | 0.2454 | 0.0162 | 0.0084 |
| Z-value | 2.2769 | 1.1699 | 0.6224 | 0.9516 | 3.7116 | -1.7645 | 3.2139 |
| P-value | 0.0228 | 0.2420 | 0.5337 | 0.3413 | 0.0002 | 0.0776 | 0.0013 |

Note: □, □ and □ indicate significance at the 10%, 5% and 1% levels, respectively.

high explanatory effects on the efficiency of the transportation structure. Among these independent variables, the proportion of secondary industry output value to GDP and the number of buses per 10,000 people passed the significance test at the 1% level and had a significant positive effect on the efficiency of the traffic structure. This result indicates that the layout of the regional industrial structure largely determines the regional passenger and freight demand as well as the choice of transportation modes, thus affecting both the efficiency and sustainability of transportation. The development of the secondary industry promotes the movement of materials and people between industries. The resulting increase in transportation demand and the pressure imposed by logistics costs further promote the adaptation of the transportation structure and the technological innovation of transportation facilities. Moreover, public transportation infrastructure is an important component of the urban transportation system. It can promote the transformation and optimization of the urban transportation structure, alleviate increasingly serious urban congestion, reduce logistics and transportation costs, improve urban transportation efficiency, and reduce air and noise pollution. The regression coefficient of the population density factor is 0.4860, which passes the significance test at the 5% level. This result indicates that the regional travel intensity in Chinese cities is positively influenced by population density. A higher population density reduces the average travel distance and increases the traffic intensity per unit area, which increases the effectivity of public transportation facilities, thus improving transportation efficiency and elevating the spatial adjustment of regional occupations to a balanced state.

In addition, the total investment in environmental pollution control is negatively correlated with traffic structure efficiency, passing the significance test at the 10% level. This result indicates that although the investment of government environmental protection funds is beneficial for the attenuation of the environmental impact of fixed and mobile pollution sources, input redundancy exists. Specific environmental protection measures, such as restricting industrial production, restricting traffic numbers, and adjusting parking fees, affect transportation costs and limit transportation demand to a certain extent, thus weakening the structural efficiency of transportation in the short term. From a long-term perspective, actions such as optimizing environmental investment, focusing on the development of transportation modes with low pollution and energy consumption, focusing on the output efficiency of government investment, and avoiding the waste of public resources will be more conducive to the future improvement of transportation efficiency. Therefore, optimizing urban function planning and industrial spatial layout, combined with adopting reasonable population guidance strategies to promote the transformation of transportation development towards a more refined mode are important measures to realize a green, efficient, and recyclable transportation structure system.

### *GWR Regression*

To ensure the validity of the GWR model calculation results, a spatial correlation analysis of the traffic structure efficiency was conducted. The global spatial autocorrelation results show that the global Moran's index of each province is 0.18, the Z-value is 2.7, which is higher than the threshold value of 1.96, and the P-value is 0.006 (i.e., significant at the 1% level). This result indicates that the traffic efficiency shows strong spatial autocorrelation.

At the same time, the current level of local indicators of spatial association was measured and the significant agglomeration map was drawn, as shown in Figure 2. The results indicate that the Beijing-Tianjin-Hebei region presents significant high-high aggregation characteristics, Jiangsu and Henan provinces are regional transportation efficiency depressions, and Guangdong province is the regional highland. Therefore, local spatial agglomerations and anomalies can be found in the distribution of traffic structure efficiency, and the spatial heterogeneity features are clear, which proves the feasibility of the constructed GWR model.

Using ArcGIS software, GWR model regression analysis was performed, and the regression coefficient estimation results are shown in Table 8.

The GWR regression results show that all indicators passed the significance test except for the transportation land use indicator. Density of Population, Per capita Gross Domestic Product, Total Investment in Fixed Assets, Total Investment in Environmental Pollution Control, and Number of Public Transportations (the five influencing factors) passed the 1% significance level test, which further corroborated the explanatory effect of the selected indicators on the efficiency of the transportation

Figure 2. Moran's I significance map of traffic structure efficiency

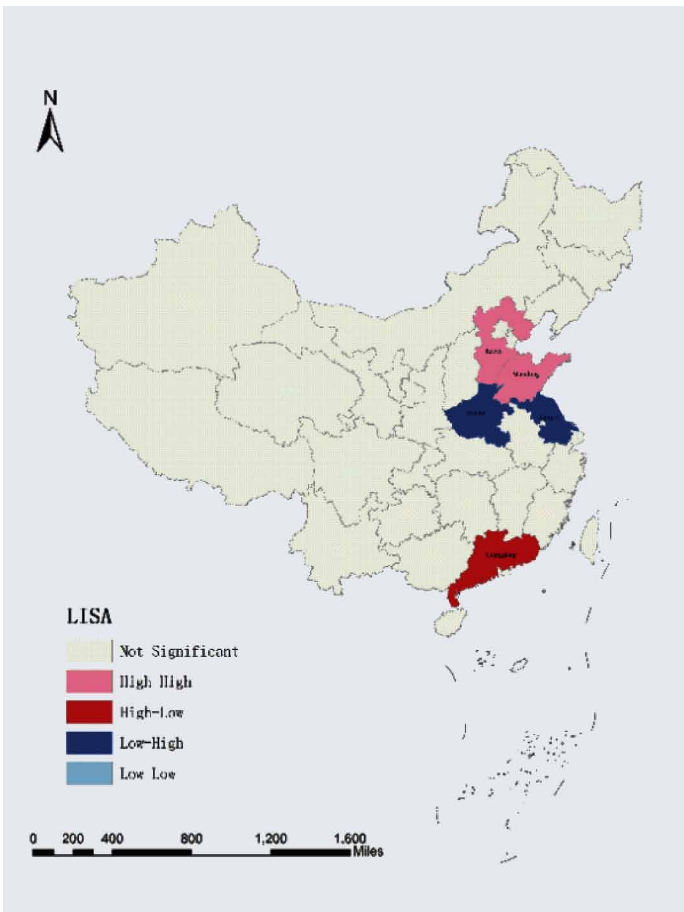


**Table 8. The regression results of GWR**

| Independent Variable | DP | PGDP | TIFA | ALT | INP | IEPC | TPB |
|---|---|---|---|---|---|---|---|
| Regression coefficient | 0.0872③ | 0.1558③ | 0.1031③ | -0.0321 | -0.0087① | -0.0466③ | 0.0534③ |
| t-value | 3.5617 | 4.8253 | 2.6122 | -0.2618 | -1.6170 | -3.5446 | 3.8144 |
| Maximum value | 0.0877 | 0.1562 | 0.1037 | -0.0319 | -0.0081 | -0.0465 | 0.0530 |
| Minimum value | 0.0874 | 0.1560 | 0.1034 | -0.0320 | -0.0084 | -0.0465 | 0.0531 |
| Sig. | 0.0000 | 0.0000 | 0.0083 | 0.7884 | 0.0769 | 0.0000 | 0.0000 |

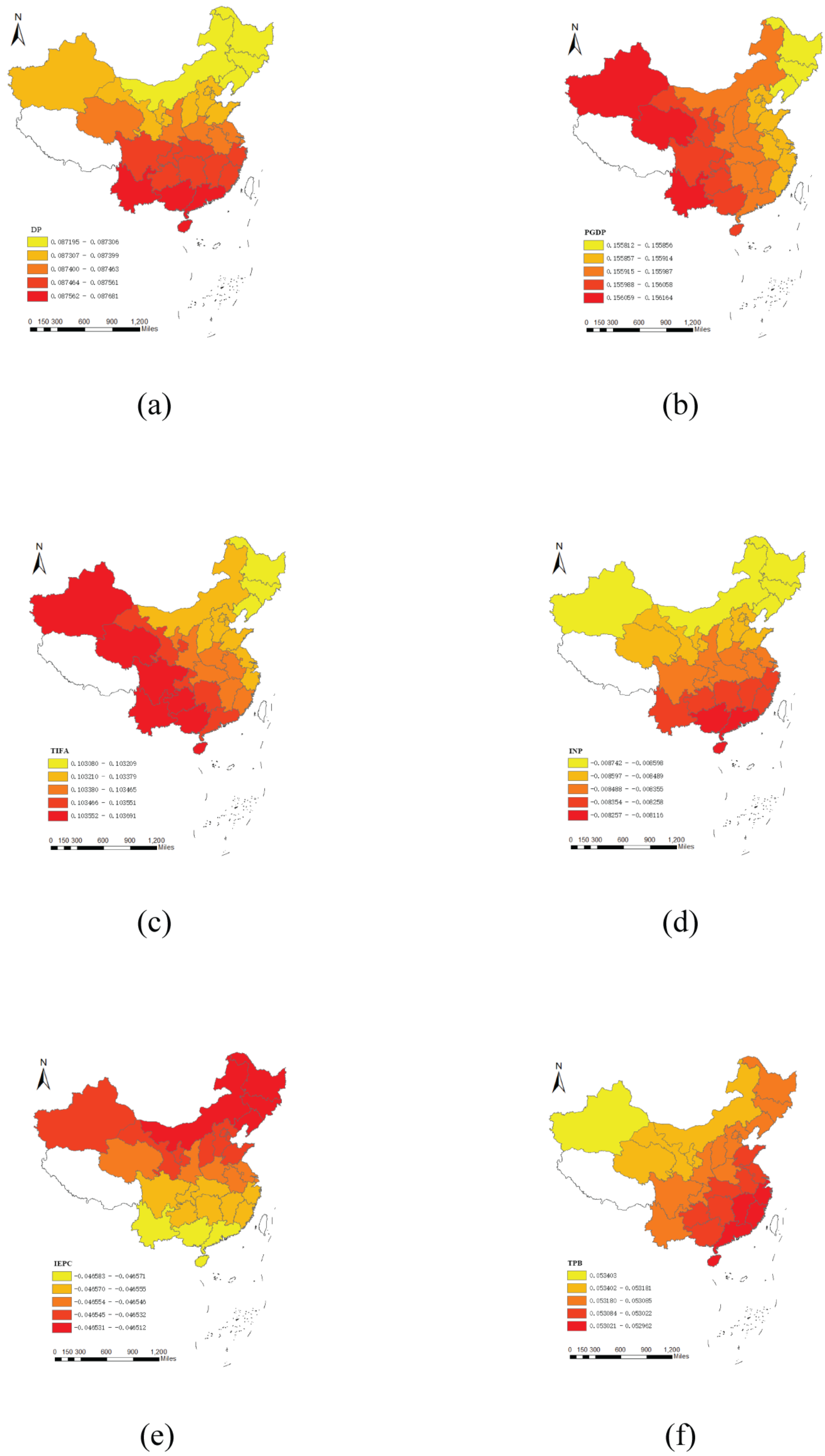
Note: ①, ②, and ③ denote significance at the 10%, 5%, and 1% levels, respectively.

structure. Specifically, density per capita, GDP per capita, investment in fixed assets of the whole society, and the number of buses per 10,000 people have positive effects on the transportation efficiency of traffic structure, which is consistent with the results of Tobit regression analysis. The proportion of the output value of the secondary industry to GDP and the total investment in environmental pollution control have negative effects. In particular, the indicator of the share of the output value of secondary industry only passes the 10% significance level and diverges from the direction of Tobit regression. Therefore, the spatial pattern visualization function of the GWR model parameter estimation is used to graphically represent the spatial regression coefficients of the six significance indicators. The spatial distribution characteristics of each indicator are shown in Figure 3.

As shown in Figure 3, the regression coefficients of the six influencing factors have different spatial distribution trends, indicating that spatial heterogeneity exists in the degree of influence each socioeconomic factor exerts on the efficiency of the transportation structure in different regions. Specifically, southern China is mainly dominated by labor-intensive high-tech industries and service industries, which actively engage in population importation and industrial spillover and have high economic activity; however, the northeastern region has a deep industrial foundation and, although the population and industries are sparsely distributed, the industries are large in scale and generally have a high per capita output value. Therefore, population density plays an important role in the economic development and transportation activity of southern China. In the western region, with its relatively backward economic development, the regional transportation infrastructure is still in a period of vigorous construction and improvement. Investment and construction efforts of its transportation system are mainly oriented by policy planning and are gradually carried out in conjunction with the level of industrial and economic development of the region. Therefore, the per capita output value as well as social fixed asset investment become dominant factors affecting the efficiency of transportation structure in many western provinces. The increase of social fixed asset investment plays an important role in the promotion of the efficiency of comprehensive transportation in western regions.

In addition, because the secondary industry, mainly manufacturing, often requires large land resources and the production process is accompanied by large energy consumption and pollutant emissions, this industry clearly has a negative impact on the densely populated and land-stressed southern regions of China. In contrast, the characteristics of vast land and abundant natural resources in the northern region can be fully utilized to increase the effectiveness of the transportation structure and to alleviate the negative effects of industrial production. At the same time, the negative impact of environmental pollution control investments on the efficiency of the transportation structure is characterized by a decrease from north to south. This characteristic corresponds to differences in the regional area, industrial distribution, and population distribution between China's north and south.

Figure 3. The spatial distribution of sub-variables' coefficient



## CONCLUSION

In traditional transportation efficiency evaluations, a single transportation mode is associated with problems as it only considers carbon emissions and neglects other non-desirable outputs. To alleviate these problems, this paper establishes a super-efficient-energy-based-undesirable model that includes both urban and intercity transportation systems and considers non-desirable outputs. The transportation structure efficiency of 30 provinces is analyzed, and the influencing factors and spatial distribution differences of transportation structure efficiency are examined using Tobit and GWR regression. The following conclusions are obtained:

- More than 70% of China's provinces have high transportation efficiency, which basically meets the overall transportation needs of society. In terms of regional distribution, clear geographical differences were found in the overall efficiency of transportation. East China and Southwest China have a transportation structure with higher overall efficiency, while Northwest China and Central China have a transportation structure with lower overall efficiency. Synergy and driving effects of transportation development in similar areas can be observed, and a grouping trend is apparent. Natural geographical barriers also cause significant differences in traffic structure.
- Population density and the number of buses per 10,000 people significantly contribute to the efficiency of the transportation structure. At the same time, over-investment in environmental management can inhibit the effectiveness of transportation.
- The regression results of the GWR model illustrate the characteristics of regional heterogeneity in the influence of variables (such as population density and gross regional product per capita) on traffic structure. The relationship between variables is spatially non-stationary with changes in geographic location and population distribution. The accuracy of the results of the Tobit model can be corroborated through a comparison of regression coefficients and significance levels.

This study is subject to limitations, which provide avenues for future research. Because of the problem of data availability, the analytical index system used in this paper is not sufficiently comprehensive and should be expanded to also include indicators such as geographic structural characteristics, road network density, and bus route duplication rate. Inclusion of these indicators can comprehensively reflect the actual operational efficiency of the regional transportation structure and the factors affecting it while, at the same time, optimization strategies of the regional transportation structure can be formulated according to local conditions. In the future, the analysis of the impact of traffic restructuring strategies on traffic operation efficiency should be ramped up. Moreover, a complete library of restructuring strategies should be established by tracking traffic development policies and changes in the inputs and outputs of the traffic system of each region over an extended period of time. The mapping relationship between the efficiency analysis and the restructuring strategies should be established using methods such as decision trees. By employing these strategies, transportation structure optimization can be guided, the overall transportation efficiency can be improved, resource waste and environmental pollution can both be reduced, and an effective balance between transportation demand and transportation supply can be achieved.

## ACKNOWLEDGMENT

The authors are grateful to support given by Hefei University of Technology. The authors would like to thank the anonymous reviewers for their invaluable comments.

## DATA AVAILABILITY

The data used to support the findings of this study are included within the article.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## FUNDING STATEMENT

This work was supported by China Postdoctoral Science Foundation (Grant Nos. 2022M720980), Anhui Provincial Key Research and Development Project (Grant Nos. 2022k07020007) and Young Teacher Research Innovation Special Initiation Project of Hefei University of Technology (Grant Nos. JZ2022HGQB0210).

## AUTHOR CONTRIBUTIONS

Conceptualization, X.C. and C.C.; funding acquisition, X.C.; methodology, X.C; validation, L.Z. and C.C.; writing—original draft, X.C. writing—review and editing, L.Z., X.C. and C.C. All authors have read and agreed to the published version of the manuscript.

## AUTHOR NOTE

Correspondence concerning this article should be addressed to Li Pan, Faculty of Engineering, Technology and Built Environment, UCSI University, Cheras, Kuala Lumpur 56000, Malaysia. Email: zzdxsdsf@zzut.edu.cn.

## REFERENCES

Abbott, D. (2014). *Applied predictive analytics: Principles and techniques for the professional data analyst*. Wiley.

Amin, C., Mulyati, H., Anggraini, E., & Kusumastanto, T. (2021). Impact of maritime logistics on archipelagic economic development in eastern Indonesia. *Asian Journal of Shipping and Logistics*, *37*(2), 157–164. Advance online publication. doi:10.1016/j.ajsl.2021.01.004

Andersen, P., & Petersen, N. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, *39*(10), 1179–1297. doi:10.1287/mnsc.39.10.1261

Ankora, C., & D, A. (2022). Integrating user stories in the design of augmented reality application. *International Journal of Information Technologies and Systems Approach*, *15*(1), 1–19. doi:10.4018/IJITSA.304809

Barnum, D., Karlaftis, M., & Tandon, S. (2011). Improving the efficiency of metropolitan area transit by joint analysis of its multiple providers. *Transportation Research Part E, Logistics and Transportation Review*, *47*(6), 1160–1176. Advance online publication. doi:10.1016/j.tre.2011.04.006

Bin, S., & Sun, G. (2022). Research on the influence maximization problem in social networks based on the multi-functional complex networks model. *Journal of Organizational and End User Computing*, *34*(3), 1–17. doi:10.4018/JOEUC.302662

De Borger, B., Kerstens, K., & Costa, Á. (2002). Public transit performance: What does one learn from frontier studies? *Transport Reviews*, *22*(1), 1–38. doi:10.1080/01441640010020313

Deng, Y., & Yan, Y. (2019). Evaluating route and frequency design of bus lines based on data envelopment analysis with network epsilon-based measures. *Journal of Advanced Transportation*, *2019*, 1–12. Advance online publication. doi:10.1155/2019/5024253

Domagała, J., & Kadłubek, M. (2023). Economic, energy and environmental efficiency of road freight transportation sector in the EU. *Energies*, *16*(1), 461. Advance online publication. doi:10.3390/en16010461

Du, Q., Lu, C., Zou, P., Li, Y., Li, J., & Cui, X. (2021). Estimating transportation carbon efficiency (TCE) across the Belt & Road Initiative countries: An integrated approach of modified three-stage epsilon-based measurement model. *Environmental Impact Assessment Review*, *90*, 106634. doi:10.1016/j.eiar.2021.106634

Ganin, A., Kitsak, M., Marchese, D., Keisler, J., Seager, T., & Linkov, I. (2017). Resilience and efficiency in transportation networks. *Science Advances*, *3*(12), e1701079. Advance online publication. doi:10.1126/sciadv.1701079 PMID:29291243

Gao, Y., & Wang, J. (2021). A resilience assessment framework for urban transportation systems. *International Journal of Production Research*, *59*(7), 2177–2192. doi:10.1080/00207543.2020.1847339

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*, 1–10. doi:10.1155/2018/6701484

Han, B., Wan, M., Zhou, Y., & Su, Y. (2020). Evaluation of multimodal transport in China based on hesitation fuzzy multiattribute decision-making. *Mathematical Problems in Engineering*, *2020*, 1–9. Advance online publication. doi:10.1155/2020/1823068

Henke, I., Esposito, M., Corte, V., Del Gaudio, G., & Pagliara, F. (2022). Airport efficiency analysis in europe including user satisfaction: A non-parametric analysis with dea approach. *Sustainability (Basel)*, *14*(1), 283. Advance online publication. doi:10.3390/su14010283

Hirschhausen, C., & Cullmann, A. (2010). A nonparametric efficiency analysis of German public transport companies. *Transportation Research Part E, Logistics and Transportation Review*, *46*(3), 436–445. doi:10.1016/j.tre.2009.11.005

Holmgren, J. (2013). The efficiency of public transport operations—An evaluation using stochastic frontier analysis. *Research in Transportation Economics*, *39*(1), 50–57. doi:10.1016/j.retrec.2012.05.023

Hu, X., Chen, X., Zhao, J., Yu, K., Long, B., & Dai, G. (2022). Comprehensive service quality evaluation of public transit based on extension cloud model. *Archives of Transport*, *61*(1), 103–115. doi:10.5604/01.3001.0015.8198

Jiang, X., Ma, J., Zhu, H., Guo, X., & Huang, Z. (2020). Evaluating the carbon emissions efficiency of the logistics industry based on a super-SBM model and the Malmquist index from a strong transportation strategy perspective in China. *International Journal of Environmental Research and Public Health*, *17*(22), 8459. doi:10.3390/ijerph17228459 PMID:33203168

Jofree, A., Hashim, N., Fahmy-Abdullah, M., Sieng, L., & Ridzuan, S. (2021). Efficiency assessment of transport manufacturing firms using a stochastic frontier analysis approach. *Sains Malaysiana*, *50*(8), 2445–2453. doi:10.17576/jsm-2021-5008-25

Kammoun, R. (2018). The technical efficiency of Tunisian ports: Comparing data envelopment analysis and stochastic frontier analysis scores. *Logistics and Sustainable Transport*, *9*(2), 73–84. doi:10.2478/jlst-2018-0011

Karlaftis, M. (2004). A DEA approach for evaluating the efficiency and effectiveness of urban transit systems. *European Journal of Operational Research*, *152*(2), 354–364. doi:10.1016/S0377-2217(03)00029-8

Khanh Van, H., Ha, T., Asada, T., & Arimura, M. (2022). Assessing transportation system efficiency in its relati onship with urban housing: A data envelopment analysis. *Asian Transport Studies*, *8*, 302–309. doi:10.1016/j.eastsj.2022.100065

Kotikov, J., & Kravchenko, P. (2020). Assessment of combined transportation energy efficiency based on Bartini's LT-table entities. *Transportation Research Procedia*, *50*, 302–309. Advance online publication. doi:10.1016/j.trpro.2020.10.036

Lai, P., Potter, A., Beynon, M., & Beresford, A. (2015). Evaluating the efficiency performance of airports using an integrated AHP/DEA-AR technique. *Transport Policy*, *42*, 75–85. doi:10.1016/j.tranpol.2015.04.008

Lee, E., Lee, H., Kho, S., & Kim, D. (2019). Evaluation of transfer efficiency between bus and subway based on data envelopment analysis using Smart Card data. *KSCE Journal of Civil Engineering*, *23*(2), 788–799. doi:10.1007/s12205-018-0218-0

Levinson, D. (2003). Perspectives on efficiency in transportation. *International Journal of Transport Management*, *1*(3), 145–155. doi:10.1016/j.ijtm.2004.01.002

Li, Z., Hao, T., & Zheng, R. (2022). The relationship between transportation industry efficiency, transportation structure, and regional sustainability development in China: Based on DEA and PVAR models. *Sustainability (Basel)*, *14*(16), 10267. doi:10.3390/su141610267

Lin, E., Lan, L., & Chiu, A. (2010). Measuring transport efficiency with adjustment of accidents: Case of Taipei bus transit. *Transportmetrica*, *6*(2), 79–96. doi:10.1080/18128600802630265

Lin, S., & Wang, J. (2022). Carbon emission reduction effect of transportation structure adjustment in China: An approach on multi-objective optimization model. *Environmental Science and Pollution Research International*, *29*(4), 6166–6183. doi:10.1007/s11356-021-16108-2 PMID:34448144

Liu, H., Yang, R., Wu, J., & Chu, J. (2021). Total-factor energy efficiency change of the road transportation industry in China: A stochastic frontier approach. *Energy*, *219*, 119612. Advance online publication. doi:10.1016/j.energy.2020.119612

Liu, Y., Eckert, C. M., & Earl, C. (2020). A review of fuzzy AHP methods for decision-making with subjective judgements. *Expert Systems with Applications*, *161*, 113738. doi:10.1016/j.eswa.2020.113738

Ma, Q., Jia, P., & Kuang, H. (2021). Green efficiency changes of comprehensive transportation in China: Technological change or technical efficiency change? *Journal of Cleaner Production*, *304*, 127115. doi:10.1016/j.jclepro.2021.127115

Ma, S., Wang, Y., & Zhang, S. (2022). Improved artificial bee colony algorithm for multimodal optimization based on crowding method. *Journal of Organizational and End User Computing*, *34*(3), 1–18. doi:10.4018/JOEUC.302661

Merkert, R., & Hensher, D. (2011). The impact of strategic management and fleet planning on airline efficiency—A random effects Tobit model based on DEA efficiency scores. *Transportation Research Part A, Policy and Practice*, *45*(7), 686–695. doi:10.1016/j.tra.2011.04.015

Ministry of Transport of the People's Republic of China. (2022). Statistical bulletin on the development of the transport industry in 2021. Ministry of Transport of the People's Republic of China. Index No. 000019713O04/2022-00119.

Prus, P., & Sikora, M. (2021). The impact of transport infrastructure on the sustainable development of the region—Case study. *Agriculture*, *11*(4), 279. doi:10.3390/agriculture11040279

Qiang, Y., Tian, G., Liu, Y., & Li, Z. (2018). Energy-efficiency models of sustainable urban transportation structure optimization. *IEEE Access : Practical Innovations, Open Solutions*, *6*, 18192–18199. doi:10.1109/ACCESS.2018.2818738

Sahitya, K., & Prasad, C. (2021). GIS-based urban road network accessibility modeling using MLR, ANN & ANFIS methods. *Transport & Telecommunication*, *22*(1), 15–28. doi:10.2478/ttj-2021-0002

Salman, R. A., Myeongbae, L., Jonghyun, L., Cho, Y., & Changsun, S. (2022). A comparative study of energy big data analysis for product management in a smart factory. *Journal of Organizational and End User Computing*, *34*(2), 1–17. doi:10.4018/JOEUC.291559

Sami, J., Pascal, F., & Younes, B. (2013). Public road transport efficiency: A stochastic frontier analysis. *Journal of Transportation Systems Engineering and Information Technology*, *13*(5), 64–71. Advance online publication. doi:10.1016/S1570-6672(13)60123-3

Sukte, C. D., Mark, E., & Deshmukh, R. R. (2022). Efficient cryptographic protocol design for secure sharing of personal health records in the cloud. *International Journal of Information Technologies and Systems Approach*, *15*(1), 1–16. doi:10.4018/IJITSA.304810

Sun, X., Yan, S., Liu, T., & Wu, J. (2020). High-speed rail development and urban environmental efficiency in China: A city-level examination. *Transportation Research Part D, Transport and Environment*, *86*, 102456. doi:10.1016/j.trd.2020.102456

Taboada, G., & Han, L. (2020). Exploratory data analysis & data envelopment analysis of urban rail transit. *Electronics (Basel)*, *9*(8), 1270. doi:10.3390/electronics9081270

Tamaki, T., Nakamura, H., Fujii, H., & Managi, S. (2019). Efficiency and emissions from urban transport: Application to world city-level public transportation. *Economic Analysis and Policy*, *61*, 55–63. doi:10.1016/j.eap.2016.09.001

Tian, N., Tang, S., Che, A., & Wu, P. (2020). Measuring regional transport sustainability using super-efficiency SBM-DEA with weighting preference. *Journal of Cleaner Production*, *242*, 118474. doi:10.1016/j.jclepro.2019.118474

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, *130*(3), 498–509. doi:10.1016/S0377-2217(99)00407-5

Tone, K., & Tsutsui, M. (2010). An epsilon-based measure of efficiency in DEA – A third pole of technical efficiency. *European Journal of Operational Research*, *207*(3), 1554–1563. doi:10.1016/j.ejor.2010.07.014

Tu, Z., Qiao, L., Nowak, R., Lv, H., & Lv, Z. (2022). Digital twins-based automated pilot for energy-efficiency assessment of intelligent transportation infrastructure. *IEEE Transactions on Intelligent Transportation Systems*, *23*(11), 22320–22330. doi:10.1109/TITS.2022.3166585

Verma, A., Harsha, V., & Subramanian, G. (2021). Evolution of urban transportation policies in India: A review & analysis. *Transportation in Developing Economies*, *7*(2), 25. Advance online publication. doi:10.1007/s40890-021-00136-1

Wang, X., Zhen, W., Shuang, Z., & Haoyue, W. (2021). Transportation structural optimization based on the evaluation of comprehensive transportation efficiency. *Journal of Computational Methods in Sciences and Engineering*, *21*(3), 685–702. doi:10.3233/JCM-215109

Wanke, P., Chen, Z., Zheng, X., & Antunes, J. (2020). Sustainability efficiency and carbon inequality of the Chinese transportation system: A Robust Bayesian Stochastic Frontier Analysis. *Journal of Environmental Management*, *260*, 110163. doi:10.1016/j.jenvman.2020.110163 PMID:32090849

Wei, F., Zhang, X., Chu, J., Yang, F., & Yuan, Z. (2021). Energy and environmental efficiency of China's transportation sectors considering CO2 emission uncertainty. *Transportation Research Part D, Transport and Environment*, *97*, 102955. doi:10.1016/j.trd.2021.102955

Yao, D., Xu, L., & Li, J. (2019). Evaluating the performance of public transit systems: A case study of eleven cities in China. *Sustainability (Basel)*, *11*(13), 3555. doi:10.3390/su11133555

Zeng, P., & Wei, X. (2021). Measurement and convergence of transportation industry total factor energy efficiency in China. *Alexandria Engineering Journal*, *60*(5), 4267–4274. doi:10.1016/j.aej.2021.03.032

Zhai, W., Bai, X., Peng, Z., & Gu, C. (2019). A bottom-up transportation network efficiency measuring approach: A case study of taxi efficiency in New York City. *Journal of Transport Geography*, *80*, 102502. doi:10.1016/j.jtrangeo.2019.102502

Zhang, F. (2021). Evaluating public organization performance under extreme weather events: Does organizational adaptive capacity matter? *Journal of Environmental Management*, *296*, 113388. doi:10.1016/j.jenvman.2021.113388 PMID:34328864

Zhang, K., Xu, Y., & Sun, D. (2018). A mixed frontier model for urban bus performance evaluation. *Transport*, *171*(2), 65–74. doi:10.1680/jtran.15.00123

Zhu, Z., & Li, T. (2008). Environmental traffic capacity and traffic structure optimization on urban road. *11th International IEEE Conference on Intelligent Transportation Systems*, 329–333. doi:10.1109/ITSC.2008.4732524

*Xudong Cao received the Ph.D. degree in transportation engineering from Chang'an University, Xi'an, China. in 2021. He is currently a postdoctoral fellow at Hefei University of Technology. His research interests include traffic management and intelligent traffic control.*

*Chenchen Chen received the M.E. degree in transportation engineering from Chang'an University, Xi'an, China. in 2017. She is currently a lecturer in the College of Civil Engineering at Anhui Jianzhu University. Her research interests include transportation economics and management.*

*Lejia Zhang received the B.E. degree in transportation engineering from Hebei University of Engineering, Handan China. in 2021. She is currently pursuing the M.E. degree with the Guilin University of Electronic Technology, Guilin China. Her research interests include traffic planning and management.*

*Li Pan has obtained a master's degree and is currently pursuing a doctoral degree in UCSI University, Malaysia. She is currently a professor at Zhengzhou Institute of Engineering and Technology. Her main research direction is computer network and artificial intelligence.*