Transportation Planning and Technology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gtpt20

Unveiling drivers of sustainability in Chinese transport: an approach based on principal component analysis and neural networks

Peter Fernandes Wanke, Amir Karbassi Yazdi, Thomas Hanne & Yong Tan

To cite this article: Peter Fernandes Wanke, Amir Karbassi Yazdi, Thomas Hanne & Yong Tan (2023) Unveiling drivers of sustainability in Chinese transport: an approach based on principal component analysis and neural networks, Transportation Planning and Technology, 46:5, 573-598, DOI: 10.1080/03081060.2023.2198517

To link to this article: https://doi.org/10.1080/03081060.2023.2198517

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 08 Apr 2023.

Submit your article to this journal

Article views: 734

View related articles

View Crossmark data

Citing articles: 4 View citing articles

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=gtpt20

TRANSPORTATION PLANNING AND TECHNOLOGY
2023, VOL. 46, NO. 5, 573–598
https://doi.org/10.1080/03081060.2023.2198517

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Unveiling drivers of sustainability in Chinese transport: an approach based on principal component analysis and neural networks

Peter Fernandes Wanke[a], Amir Karbassi Yazdi[b], Thomas Hanne[c] and Yong Tan[d]

[a]Center for Logistics Studies, COPPEAD – The Graduate School of Business Administration, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil; [b]School of Engineering, Universidad Católica del Norte, Coquimbo, Chile; [c]Institute for Information Systems, University of Applied Sciences and Arts Northwestern Switzerland, Windisch, Switzerland; [d]School of Management, University of Bradford, Bradford, West Yorkshire, UK

ABSTRACT
The paper analyzes the sustainability of the Chinese transportation sector by examining the relationship between energy consumption (and CO2 emissions), transportation modes, and macroeconomic variables. Principal Component Analysis (PCA) and Neural Networks (NN) are combined using monthly data from January 1999 to December 2017. Our goal is to propose a model that links China's transportation footprint to major macroeconomic factors while simultaneously controlling each mode of transportation. Inflation and credit policies exert relatively weak effects on the explained variable. In contrast, trade and fixed asset investments, as well as monetary and fiscal policies, show a positive and significant impact. The use of waterways and airways plays an imperative role in sustainable development compared to the use of roads.

ARTICLE HISTORY
Received 9 November 2022
Accepted 29 March 2023

KEYWORDS
China; transportation modes; sustainability; macro-economic variables; principal component analysis; neural networks

1. Introduction

As a result of the fastest global temperature rise in recorded history over the past five decades, climate change has adversely affected sustainable development worldwide. Over one hundred countries have adopted limits on global warming (Meinshausen et al. 2009). The transportation industry is a significant contributor to greenhouse gas emissions (GHG), which are the primary driver of climate change. According to World Bank statistics (IPCC 2015), transportation-related GHG emissions accounted for 14% of total global emissions, making it a crucial component in reducing emissions across countries (Heinold and Meisel 2018).

The situation in China is similar. Since 2005, carbon emissions have hit their peak. To achieve sustainable development, China plans to reduce carbon emissions by 60% to 65% by 2030 (Zhang, Yu, and Chen 2017). Due to the country's large population and rapid urbanization, there is a high demand for transportation in the industry

CONTACT Yong Tan y.tan9@bradford.ac.uk School of Management, University of Bradford, BD7 1DP, Bradford, West Yorkshire, UK

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

(Chen, Barros, and Yu 2017a; Chen et al. 2017b; Chen, Tzeremes, and Tzeremes 2018). In 2010, China consumed 40% of the total oil and emitted 20% of the carbon dioxide in the world. This poses a significant challenge for China's low-carbon development (Pan et al. 2018). The government is, therefore, putting in increased efforts to develop the sector sustainably, especially under strict environmental regulations, including raising automobile pollution standards and improving the efficiency of high-speed rail systems.

This study examines the sustainability of Chinese transportation modes using monthly data from January 1999 to December 2017. Previous studies have analyzed transportation modes using a time series approach (Lao et al. 2016; Kim, Hewings, and Amir 2017). However, this study develops an innovative approach to predicting China's transportation footprint by using Principal Component Analysis (PCA) and Neural Networks (NN) models. While several studies have utilized these methods, only How and Lam (2018) applied PCA and the Analytical Hierarchy Process (AHP) to evaluate the full spectrum of sustainability, including economic, environmental, and social domains.

In addition, this paper considers energy consumption and carbon emissions associated with both freight and passenger transportation footprints jointly. Previous works have examined either freight or passenger transportation (Zhu, Zhang, and Zhang 2018; Guimarães, Junior, and Silva 2018; Chao 2014). Combining freight and passenger transportation into one analysis may provide a better understanding of macroeconomic impacts.

Quantitative methods such as time-series analysis, multivariate regressions, and vector auto-regression are used in the literature. However, this paper uses a novel combination of PCA and NN to examine how macroeconomic variables influence China's transportation footprint. Neural networks are a common method for predicting data with better classification, generalization, and approximation ability and have been used in a wide range of functions for many years (Azoff 1994). PCA primarily uses variable dimension reduction to create new combinations of variables or indicators that are independent of one another and contain fewer variables or indicators (Dunteman 1989). These methods are flexible and computationally efficient, making them suitable for various applications, including transportation (Nagendra and Khare 2004; Liu et al. 2016; Liu et al. 2017).

This study aims to unveil the relationship between energy consumption (including CO2 emissions) of different transportation modes and significant macroeconomic variables by utilizing PCA and NN approaches. By using NN within the scope of PCA, we can explore the relationship between transportation footprint and macroeconomic factors while considering the impact of different transportation modes.

The study makes a valuable contribution to the existing literature by analyzing the feedback effects of macroeconomic factors on sustainability in the Chinese transportation industry. The paper is structured as follows: Section 2 provides some background information, while Section 3 summarizes the relevant literature. The methodology employed in this study is outlined in Section 4. Our empirical findings are presented in Section 5, followed by concluding remarks in Section 6.

2. Overview of the transportation system in China

The transportation system in China comprises railways, highways, air, and water transportation (Huang et al. 2018). As of 2015, the length of railways in China exceeded

120,000 kilometers. China operates approximately 67% of the world's high-speed railways, representing a significant growth in railway transportation in recent years (Wanke et al. 2018b). The main lines connect the major cities in China.

As China's economy and income grow, the civil aviation industry is also rapidly developing (Chen, Tzeremes, and Tzeremes 2018). Under the 'National Civil Airports Scheme', China plans to have 244 civil airports by 2020. In 2017, the civil aviation industry in China carried out transportation of 108.3 billion tons, a 12.5% increase from the previous year. IATA predicts that China's aviation market will grow by an average of 10% per year during 2017–2022, and the average yearly demand will be 0.5 flights per person, carrying 720 million passengers. The number of Chinese passengers is expected to reach 1.3 billion in 2035, surpassing that of the United States in 2024. Transportation in China is also facilitated by the road freight system. In 2017, approximately 130 million tons of freight were transported by road daily, making it China's most significant transport system. More than 78% of integrated freight is transported through roads. The Chinese expressway network includes eleven major lines, covering the country's economic centers. The inland rivers of China include the Yangtze, Pearl, Heilongjiang, Songhua, Huaihe, and Beijing-Hangzhou Grand Canal. The Yangtze River alone transports sixty percent of all freight (Wanke et al. 2018a.

According to the China Statistical Yearbook 2017, China's transportation industry consumed 383.18 million tons of standard coal in 2015. Diesel accounted for 29.13%, and gasoline accounted for 13.85%, respectively. The carbon emissions in China in 2017 were 291.142 million tons, accounting for 22.61% of total carbon emissions, as per the IPCC National Greenhouse Gas Inventories Guide. In 2010, gasoline accounted for 10% of carbon emissions. Diesel and gasoline remain the primary fuels used in transportation, which are associated with high carbon emission factors and severe pollution. These two primary energy sources urgently need to be addressed.

3. Review of literature

The existing studies on transportation freight and passenger time series are scarce. Khan et al. (2018) adopted econometric panel methods to investigate freight traffic in different transportation modes. The findings show that air-railways transportation can accelerate energy demand. Schodl et al. (2018) used a quantitative estimation technique to investigate the influence of transportation networks on the corresponding regions. Alizadeh (2013) examined the trading volume-price volatility relationship in the Forward Freight Agreement (FFA) market. Karlaftis and Vlahogianni (2009) developed fractionally integrated dual models and made relevant comparisons between the results with those from classic time series models under the traffic engineering context. Barros, Chen, and Wanke (2016) applied fractional models to Brazilian transportation modes, and the findings suggest that road transportation dominates over other modes.

In the Chinese context, several relevant studies have examined transportation with an emphasis on a few modes, including airlines, railways, and waterways (Bao, Hua, and Gu 2016; Hu, Chen, and Zheng 2018; Huo, Zhang, and Chen 2018; Wanke et al. 2018a; Antunes et al. 2023). However, the combined analysis of Chinese transportation modes – air, water, rail, and road – has not attracted much attention from transportation researchers so far (Feng and Wang 2018; Pan et al. 2018). Furthermore, existing works mainly focused on either passenger transportation or freight transportation, while there have been few efforts focusing on cross-impacts with macroeconomics (Halkos and Paizanos 2016; Zegras et al. 2013).

There have been limited studies investigating Chinese transportation modes using time series analysis. Wang and Lin (2019) examined the influence of urbanization, price, income, and structure on diesel and gasoline consumption in road transportation. Huang et al. (2018) investigated the dynamics of China's main transportation modes from monetary and energy perspectives. Existing studies were characterized by using low-frequency data, facilitated by traditional panel models, to evaluate causal relationships.

Regarding the environmental and economic impacts of transportation, Hadi-Vencheh et al. (2018), and Zhu, Zhang, and Zhang (2018) examined the impact of various modes of transport on monetary costs and environmental expenditures. Xia, Sun, and Feng (2022) investigated the factors determining the spatial inequalities of low-carbon development in China's transport section under a meta-frontier DEA-based decomposition approach. Liu and Feng (2020) proposed a temporal–spatial analysis to evaluate the relationship between transport energy-related carbon emission and economic development. For example, Du, Xie, and Ouyang (2017) adopted multiple regression and path analysis to assess the determinants of carbon emissions in China's transportation industry. The findings show that carbon emissions are significantly affected by traffic intensity, energy intensity, and economic levels.

On the other hand, it is common to assess the sustainability or green technical efficiency of the transportation sector using DEA (Data Envelopment Analysis) (Chen et al. 2017b; Chen, Tzeremes, and Tzeremes 2018; Hadi-Vencheh et al. 2018; Wanke et al. 2018b) and SFA (Stochastic Frontier Analysis) (Chen, Barros, and Yu 2017a). Some researchers have also attempted to identify the factors that drive efficiency changes during a specific period, such as environmental regulation (Liu et al. 2016; Li, Meng, and Yao 2017; Pan et al. 2018) and transportation infrastructure (Xie, Fang, and Liu 2017). However, most of these studies have focused on a specific transportation mode within a country (Zuo et al. 2018). This research also contributes to this field of study since existing literature primarily relies on case studies and multivariate regression analyses.

Specifically, Huang et al. (2018) applied China's system diagram and time-series data to identify the monetary and environmental costs for different transportation modes. They concluded that private cars are the most costly method of mitigating environmental impact. Furthermore, Chang, Ma, and Wu (2018) used the STIRPAT model to analyze the effect of urbanization on the private and public sector energy consumption in China. They demonstrated that the urbanization level significantly affects the energy consumption of the public transportation system.

Moreover, some scholars have studied comparable topics internationally. Wang and Lin (2019) compared the energy use of road transportation in China and OECD countries and identified the factors that influence energy consumption in these countries. China's energy consumption, particularly its consumption of fossil fuels, is more affected by economic factors compared to other OECD countries. However, more research is needed to reveal the mechanisms and policy implications of improving sustainable development in China. Table 1 summarizes the reviewed studies above.

Transportation demand is derived from socioeconomic demand. This means that the level of economic activity and the behavior of consumers influence the demand for transportation services. As a result, transportation-related greenhouse gas emissions are related to macroeconomic variables such as income, population, and urbanization.

Regarding the concept of 'drivers of transport sustainability,' it is not surprising to consider macroeconomic variables as important factors in promoting sustainable transportation. By adopting green transportation strategies, it is possible to decouple transport-related greenhouse gas emissions from economic growth. For example, the promotion of railways and waterways, the use of clean energy vehicles, and the implementation of efficient logistics systems such as intermodal and multimodal structures can contribute to reducing transport emissions while maintaining economic growth.

In summary, it is possible to reduce transportation-related greenhouse gas emissions while promoting economic growth through the adoption of green transportation strategies (IEA 2017; UNEP 2019; EEA 2019). Macroeconomic variables can play a role in driving sustainable transportation, and it is important to consider them in developing policies and strategies to address climate change (ITF 2019; WB 2019; EC 2016).

4. Methodology

This section outlines the steps involved in developing the PCA-NN two-stage approach, based on the assumption that substantial macroeconomic variables in China are interconnected with transportation sustainability (energy consumption and CO2 emissions). The first methodological step is PCA, which is used to determine independent or uncorrelated factors (groups of variables). The transport footprint is then used as the dependent variable in an NN regression, using these extracted factors. The goal is to assess how significant macroeconomic variables and their relative importance to the transportation footprint influence the energy consumption of the transportation sector. These modeling steps are briefly described, and then the results are discussed. In the following paragraphs, we provide brief explanations of both PCA and NN and explain why they are used in combination.

4.1. Background on PCA

PCA is a method for studying the structure of data and determining patterns of covariance among variables. Therefore, PCA studies variance-covariance matrices. The purpose of PCA is to identify highly correlated variables or sets of variables. In addition to analyzing large datasets with multiple dimensions, PCA enhances interpretability, preserves information, and visualizes multidimensional data. PCA reduces the dimensionality of datasets. In some cases, a linear transformation of data can reduce the number of dimensions that cause a major part of variations. In two dimensions, data is clustered using the first two principal components. Principal component analysis is used, for instance, in microbiology, genetics, and atmospheric science.

Table 1. Review of previous related studies.

Author (year)	Aim	Data	method	Results	The advantages of our proposed method
Khan et al. (2018)	To investigate the impact of air transportation, railways transportation, and port container traffic on energy demand, customs duty, and economic growth in low-, middle-, and high-income countries.	Panel data from 92 countries over the period of 1995–2012.	Panel regression model	Air and railways transportation have a positive impact on economic growth, while port container traffic has a negative impact.	PCA-NN models can capture nonlinear and complex relationships between variables and can perform well even when there is heteroscedasticity in the data.
Schodl et al. (2018)	To develop a method for estimating regional impacts of innovative means of cargo transport.	Data on cargo flows, infrastructure, and transport modes in Austria.	Multi-criteria decision analysis (MCDA) and network simulation	The proposed method can estimate regional impacts of innovative means of cargo transport	PCA-NN models can identify the most important variables and their interactions and can provide insights into the structure of the data.
Alizadeh (2013)	To examine the relationship between trading volume and volatility in the shipping forward freight market.	Data on daily trading volumes and prices of forward freight agreements (FFAs) for the period of 2003–2008.	GARCH model and regression analysis	Trading volume has a positive impact on volatility in the shipping forward freight market, but the relationship is not linear.	PCA-NN models, can capture nonlinear and complex relationships, and can be applied to a wide range of data sets beyond financial time series.
Karlaftis and Vlahogianni (2009)	To investigate the memory properties and fractional integration in transportation time-series	Traffic volume data from Athens, Greece	Fractional integration analysis	transportation time-series exhibit long-term dependence and non-stationarity, and the memory properties differ across different types of transportation modes	PCA-NN models can identify the most important variables and their interactions and can provide insights into the structure of the data, even when the relationships are nonlinear and complex.
Barros, Chen, and Wanke (2016)	To evaluate the efficiency of Chinese seaports from 2002 to 2012	Data on Chinese seaports from the World Bank	Data envelopment analysis (DEA) and Malmquist productivity index (MPI)	the overall efficiency of Chinese seaports has improved during the period, and the major factors affecting efficiency are technology and scale efficiency	PCA-NN models can identify the most important variables and their interactions and can provide insights into the structure of the data, which can help to improve decision-making and increase efficiency.
Bao, Hua, and Gu (2016)	To examine the relationship between airport accessibility and airport competition	Data on airport accessibility and competition in China	Regression analysis	airport accessibility has a significant positive effect on airport competition, and the effect is	PCA-NN models can capture nonlinear and complex relationships between variables and can perform well even

(*Continued*)

Table 1. Continued.

Author (year)	Aim	Data	method	Results	The advantages of our proposed method
				stronger for domestic airports than for international airports	when there is heteroscedasticity in the data.
Hu, Chen, and Zheng (2018)	To analyze the impacts of congestion pricing on environmental costs at Guangzhou Baiyun International Airport	Data on airport operations and environmental emissions from Guangzhou Baiyun International Airport	Regression analysis	congestion pricing can effectively reduce airport congestion and environmental costs, and the optimal congestion pricing scheme depends on the level of environmental regulation and the sensitivity of passengers to congestion charges	PCA-NN models can capture nonlinear and complex relationships between variables and can perform well even when there is heteroscedasticity in the data.
Huo, Zhang, and Chen (2018)	Examine recent development of Chinese port cooperation strategies	Secondary data from reports, articles, and official documents	Literature review and case studies	Identified challenges and opportunities for Chinese port cooperation strategies and provided recommendations for future development	PCA and NN can handle large and complex datasets
Wanke et al. (2018a)	Measure Malmquist productivity indexes in Chinese ports	Data on 25 Chinese ports from 2008 to 2014	Fuzzy GMSS DEA approach	technological change and technical efficiency were the main drivers of productivity growth in Chinese ports	PCA and NN can capture non-linear relationships between variables
Feng and Wang (2018)	Analyze energy efficiency in China's transportation sector	Secondary data from reports and statistics	Data envelopment analysis and Tobit regression	Identified factors affecting energy efficiency in China's transportation sector and provided recommendations for policy makers	PCA and NN can identify hidden patterns and relationships between variables that may not be apparent with DEA and Tobit regression.
Pan et al. (2018)	Examine decarbonization of China's transportation sector	Secondary data from reports and statistics	Scenario analysis and modeling	Proposed a roadmap for decarbonizing China's transportation sector, including policy recommendations and technological solutions	PCA and NN can provide more accurate predictions and forecasts
Zegras et al. (2013)	To analyze the prospects of establishing metropolitan transportation authorities in Portugal within the context of fiscal federalism	Qualitative data from interviews with policymakers and public officials;	Case study approach	Establishment of metropolitan transportation authorities is feasible and may help improve transportation governance and service delivery	PCA and NN provide a more systematic and data-driven approach that can capture complex relationships between variables
Wang and Lin (2019)	To compare the fuel consumption and efficiency of road transport between China and OECD countries	Time-series data from 1980 to 2015	DEA and Malmquist productivity analysis	Fuel consumption and efficiency of road transport have improved in both China and OECD countries, but	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables

(Continued)

Table 1. Continued.

Author (year)	Aim	Data	method	Results	The advantages of our proposed method
				the gap between them remains significant	
Huang et al. (2018)	To assess the monetary, energy, and environmental costs of terrestrial transport modalities in China	Data from various sources	Life cycle assessment and cost-benefit analysis	Railway and waterway transport are more environmentally and economically sustainable than road transport in China	PCA and NN can handle both quantitative and qualitative data
Chen, Barros, and Yu (2017a)	To analyze the spatial distribution characteristics of Chinese airports	Data on 191 Chinese airports	Spatial cost function approach	The spatial distribution of airports in China is inefficient, with many airports located in areas with low demand and high costs	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Hadi-Vencheh et al. (2018)	To measure the sustainability of Chinese airlines with respect to CO_2 emissions	Data on 36 Chinese airlines over 2010–2014;	Modified slack-based measure model	Chinese airlines have low levels of sustainability	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Zhu, Zhang, and Zhang (2018)	To analyze the connectivity of intercity passenger transportation in China using a multimodal and network approach.	China Statistical Yearbook and Baidu Maps.	network analysis, clustering, and centrality analysis.	rail transportation plays a crucial role in intercity connectivity in China.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Chen et al. (2017b)	To evaluate the efficiency of Chinese airlines considering CO_2 emissions and flight delays	Data on 32 Chinese airlines over 2010–2014	Stochastic network DEA model	Most Chinese airlines are inefficient, and CO_2 emissions have a significant negative impact on their efficiency	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Chen, Tzeremes, and Tzeremes (2018)	To examine convergence in the Chinese airline industry using a Malmquist productivity analysis.	China's civil aviation statistics.	DEA and Malmquist productivity analysis.	technical efficiency and productivity have improved, but there is still a gap between state-owned and non-state-owned airlines.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Wanke et al. (2018b)	Investigate drivers of railway performance in selected Asian countries	Data from 14 Asian countries from 2008 to 2015	DEA, Tobit regression, and bootstrapping techniques	economic, social, and institutional factors, as well as technology, had significant impacts on railway performance	PCA and NN can handle missing and incomplete data. Additionally, PCA and NN can identify the most important variables that affect the outcome
Liu et al. (2016)	To propose an improved grey neural network model for predicting transportation disruptions.	Shanghai Metro System.	grey neural network modeling	proposed model had higher prediction accuracy than traditional models.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables

(*Continued*)

Table 1. Continued.

Author (year)	Aim	Data	method	Results	The advantages of our proposed method
Li, Meng, and Yao (2017)	To evaluate the sustainability performance of China's transportation industry under environmental regulation.	China Environment Yearbook and China Transportation Yearbook.	DEA and regression analysis.	environmental regulation has a positive impact on the sustainability performance of China's transportation industry.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Xie, Fang, and Liu (2017)	To investigate the effects of transportation infrastructure on urban carbon emissions in China.	China City Statistical Yearbook and China Energy Statistical Yearbook.	panel data analysis	the improvement of transportation infrastructure has a negative impact on urban carbon emissions.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Xia, Sun, and Feng (2022)	To decompose the causes of spatial inequalities in low-carbon development in China's transport sector using a meta-frontier DEA-based decomposition approach.	China Transportation Yearbook	meta-frontier DEA modeling.	the main factors contributing to spatial inequalities were differences in technology efficiency and input quality.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Liu and Feng (2020)	To analyze the temporal and spatial decoupling of transport CO2 emissions from China's economic expansion.	China Energy Statistical Yearbook and China City Statistical Yearbook.	decoupling analysis and spatial analysis.	some provinces have successfully decoupled transport CO2 emissions from economic growth, while others have not.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Zuo et al. (2018)	To determine whether road or rail transportation is the better solution for reducing carbon emissions related to the transportation of aggregates.	an aggregate transportation company in the UK.	life cycle assessment and cost analysis.	rail transportation is a more sustainable solution for the transportation of aggregates.	CA and NN can handle both quantitative and qualitative data
Chang, Ma, and Wu (2018)	To analyze the impact of urban development on residents' public transportation travel energy consumption in China, with a focus on hydrogen fuel cell vehicles	China Urban Construction Statistical Yearbook and China Transportation Yearbook.	regression analysis and scenario analysis.	the adoption of hydrogen fuel cell vehicles can significantly reduce energy consumption and carbon emissions in urban transportation.	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables
Antunes et al. (2023)	To examine the impact of R&D and innovation on Chinese road transportation sustainability performance using a novel trigonometric envelopment analysis.	the China Statistical Yearbook and China Transportation Yearbook.	DEA and trigonometric envelopment analysis.	R&D and innovation have a positive impact on the sustainability performance of Chinese road transportation	PCA and NN can handle high-dimensional data and capture nonlinear relationships between variables

In n-data analysis, a derived variable account for the majority of the variance in p variables. There is an assumption that its distribution is equal. After removing the first principal component, the second principal component explains the majority of variance. PCA reduces the number of highly correlated variables to an independent set by reducing the number of variables that are highly correlated. Predictive models are created using PCA. Data points can be reduced in dimensionality by projecting the first few principal components onto them. As a result, the first principal component maximizes variance. Summarizing principal components in orthogonal directions maximizes variance.

Every objective's eigenvector is its principal component. The eigen decomposition of data matrices is commonly used in principal component analysis. Eigenvector-based multivariate analysis, such as PCA, is more straightforward than factor analysis. In factor analysis, the eigenvectors of a different matrix are solved, and more domain-specific assumptions are taken into account. There is a connection between CCA (Canonical Correlation Analysis) and PCA. In CCA, coordinates are used to describe cross-covariance between two datasets, while in PCA, orthogonal coordinates are used to describe variance between datasets.

4.2. Background on neural networks

Artificial neural networks are a method in artificial intelligence (AI) used to solve problems by using simplified models of biological neural networks. The models use neurons as building blocks, which are connected using weights to express the strength of their connections. A positive connection is known as an excitatory connection, while a negative connection is known as an inhibitory connection. Weights are also applied to all inputs. The inputs of a neuron are summed up as a linear combination, multiplying each input with the respective weight. An activation function controls the output of a neuron.

Neural networks are nonlinear statistical data models or decision-making tools. They can be used to find patterns in data or model complex relationships between inputs and outputs. Adaptive control, predictive modeling, and data-driven applications can benefit from neural networks. Deductive reasoning can be used within networks to facilitate self-learning. Artificial neural networks are commonly used for the following tasks:

- Time series prediction and modeling using function approximation or regression analysis.
- Recognizing patterns and sequences, detecting novelty, and making sequential decisions.
- Filtering, clustering, blind signal separation, and compression.

4.3. Reasons for combining NN and PCA

Although training an NN model takes longer than training a linear PCA model, it was found that the former is more suitable for use in nonlinear systems, despite the longer training time. This model is also more comprehensive in terms of time domain features. Additionally, the proposed method improves the system's performance by using time domain features.

Readers should be aware that when using PCA and NN together, the number of weights related to each explanatory variable in the NN increases, and the amount of data required to accurately determine the network weights also increases (often quite rapidly), increasing the risk of overfitting. Dimensionality reduction reduces the network size, which, in turn, reduces the amount of training data required. A complete description of the error measurement function in terms of PCA can be obtained by learning from examples in layered feed-forward neural networks using optimization methods, such as back propagation.

4.4. The dataset used

Table 2 shows the variables used in the current study. All variables are determined on a monthly basis. We mainly have three different sources for collecting the data: 1) the yearbook of China's transportation; 2) the Wind database (www.wind.com.cn); 3) China's National Statistics Bureau. The average conversion factors for each transportation mode's efficiency and pollutant characteristics were used in the computation of CO2 emissions and energy levels (Wang, Ou, and Zhang 2017. Furthermore, an average weight of 70 kg was assigned to each passenger in order to establish a common basis for freight and passenger movements. The data period covered in the study ranges from January 1999 to December 2017.

Figure 1 displays the results of the returns of these time series after removing the trend and seasonal effects, with the exception of the loan rates, consumer price indexes, deposit rates, and exchange rates for which trend and seasonal effects have not been removed. To extract the random level component and seasonal effects from the original dataset, we apply a time series decomposition method. Subsequently, the observation at time t was divided by the previous observation at time t-1 to calculate the returns. These decomposition steps and return computations ensure the stationarity of the time series. Returns were calculated for each mode based on the sum of their respective shares, as well as for energy consumption and CO2 emissions.

Table 3 presents the descriptive statistics for the return time series. Table 4 shows the correlation coefficients, which also support the need to extract factors using PCA, since many values are relatively high (with an absolute value above 0.40). Table 3 shows that, except for macroeconomic variables with proportionally more significant standard deviations (SD), the recomputed time-series show mean returns fluctuating around 1.0 and the coefficient of variation (CV) is low. A consideration of minima, maxima, and means in Table 3 also reveals some asymmetry in the time series of transportation sustainability and macro-economic variables. It is apparent from Table 4 that some of these variables are strongly correlated, particularly within the context of each mode of transportation.

4.5. PCA

Following Hair et al. (1995), let the random vector, $X = [X_1, X_2, \ldots, X_p]$ (i.e. the criteria chosen to be aggregated), have the correlation matrix C, with eigenvalues, $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p \geqslant 0$, and normalized eigenvectors, $l_1, l_2, \ldots, l_p$. Consider the linear

Table 2. Descriptive statistics for the original dataset (SD = standard deviation, CV = coefficient of variation).

Variables	Min	Max	Mean	SD	CV
Total Turnover of Passenger Traffic (100 million person*kilometers)	557.320	3272.168	1922.790	706.304	0.367
Total Turnover of Freight Traffic (100 million tons*kilometers)	2767.800	17518.878	9430.423	4679.750	0.496
Total Energy Use (100 million Horsepower (HP))	2452217.050	16968737.164	8493750.660	4806812.482	0.566
Total Transportation CO_2 Emission (tons)	118105.025	1040037.946	489524.878	318947.832	0.652
Railway Turnover of Passenger Traffic (100 million person*kilometers)	159.160	1454.196	659.780	272.007	0.412
Railway Turnover of Freight Traffic (100 million tons*kilometers)	904.240	2592.156	1853.188	477.550	0.258
Railway Energy Use (100 million HP)	1237081.467	3539892.991	2532496.478	655340.256	0.259
Railway CO_2 Emission (tons)	31545.577	90267.271	64578.660	16711.177	0.259
Roadway Turnover of Passenger Traffic (100 million person*kilometers)	371.680	1780.888	942.625	326.934	0.347
Roadway Turnover of Freight Traffic (100 million tons*kilometers)	390.900	6404.753	2577.382	2178.130	0.845
Roadway Energy Use (100 million HP)	634161.471	9597547.660	3887302.150	3224837.145	0.830
Roadway CO_2 Emission (tons)	50022.657	757054.559	306630.394	254375.154	0.830
Waterway Turnover of Passenger Traffic (100 million person*kilometers)	2.880	13.330	6.496	1.448	0.223
Waterway Turnover of Freight Traffic (100 million tons*kilometers)	1438.790	9349.873	4977.091	2188.696	0.440
Waterway Energy Use (100 million HP)	287935.240	1870076.966	995509.186	437733.294	0.440
Waterway CO_2 Emission (tons)	28793.524	187007.697	99550.919	43773.329	0.440
Air Transportation Turnover of Passenger Traffic (100 million person*kilometers)	23.600	869.837	313.499	215.571	0.688
Air Transportation Turnover of Freight Traffic (100 million tons*kilometers)	2.310	23.034	10.408	5.515	0.530
Airway Energy Use (100 million HP)	175066.667	2693275.867	1078442.846	678809.654	0.629
Airway CO_2 Emission (tons)	3046.160	46863.000	18764.906	11811.288	0.629
Exchange Rate (RMB/U.S Dollar)	6.104	8.280	7.301	0.858	0.118
Loan Rate (%)	4.350	7.470	5.683	0.738	0.130
Deposit Rate (%)	1.500	4.140	2.486	0.689	0.277
Consumer Price Index (%)	−1.800	2.600	0.177	0.683	3.855
Total Trade (100 million U.S. Dollars)	196.990	4089.081	1964.353	1202.975	0.612
M2 Money Supply (100 million Yuan)	105500.000	1676768.540	632779.175	483581.809	0.764
Fiscal Expenditure (1 billion Yuan)	52.976	2701.591	688.320	610.871	0.887
Consumer Confidence Index	97.000	123.900	107.845	4.849	0.045
Fixed Assets Investment (1 million Yuan)	84519.000	63168396.000	11101454.122	14282428.935	1.287
Imported Crude Oil Price Index	88.470	948.796	443.219	229.345	0.517
Railway Share of Energy Use	0.170	0.589	0.369	0.127	0.345
Road Share of Energy Use	0.212	0.595	0.380	0.144	0.381
Waterway Share of Energy Use	0.086	0.200	0.128	0.027	0.215
Air Share of Energy Use	0.054	0.208	0.124	0.021	0.168
Railway Share of CO_2 Emission	0.068	0.319	0.182	0.079	0.432
Road Share of CO_2 Emission	0.344	0.741	0.536	0.147	0.275
Waterway Share of CO_2 Emission	0.139	0.389	0.242	0.072	0.295
Air Share of CO_2 Emission	0.020	0.064	0.040	0.007	0.184

combinations, where the superscript t represents the transpose operator,

$$X_{\mathrm{PC}_i} = l_i^t X = l_{1i}X_1 + l_{2i}X_2 + \ldots + l_{pi}X_p, \tag{1}$$

$$\mathrm{Var}(X_{\mathrm{PC}_i}) = l_i^t C l_i,\ i = 1, 2, \ldots, p, \tag{2}$$

and

$$\mathrm{Correlation}(X_{\mathrm{PC}_i}, X_{\mathrm{PC}_k}) = l_i^t C l_k,\ i = 1, 2, \ldots, p,\ k = 1, 2, \ldots, p \tag{3}$$

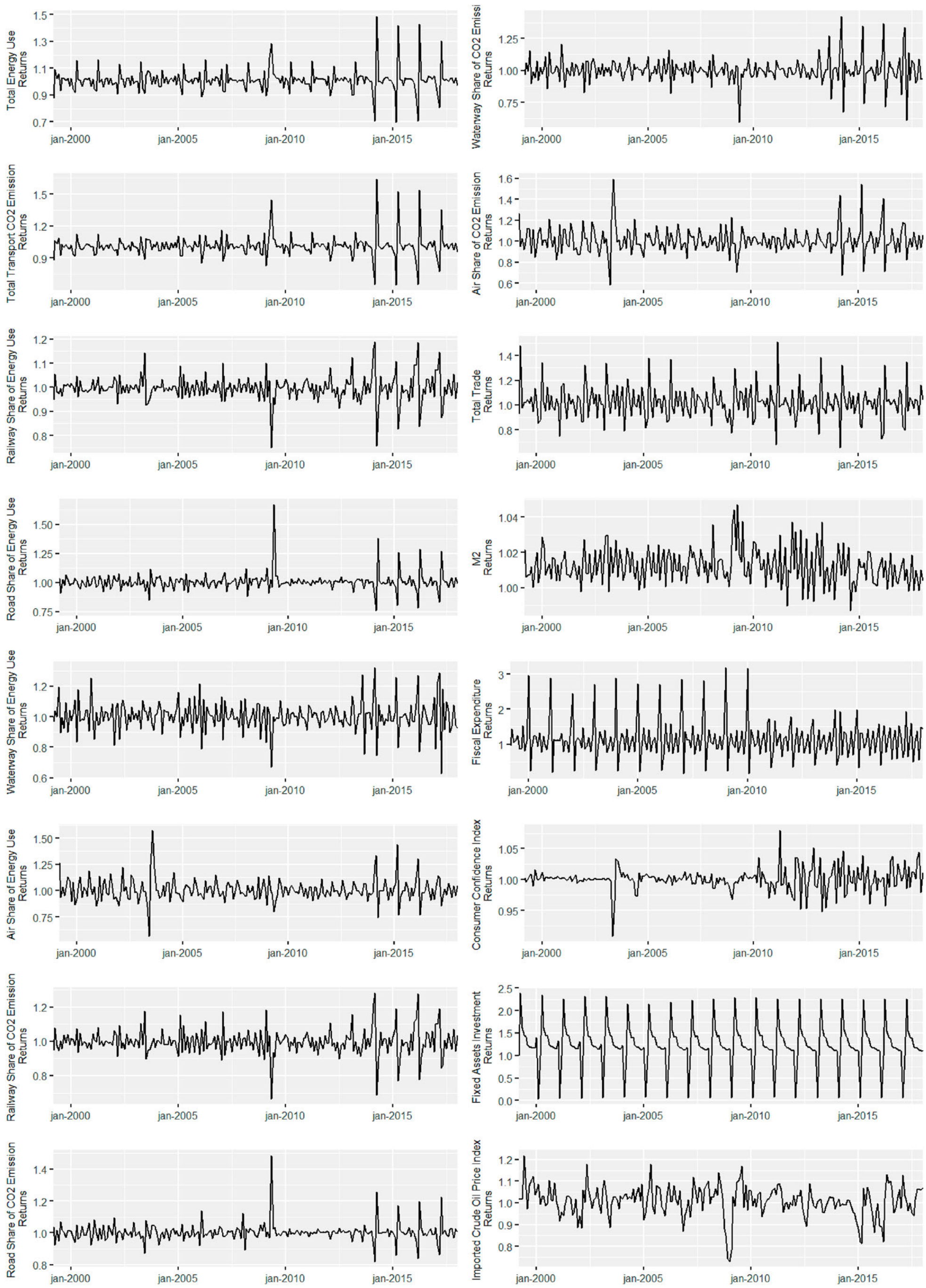

Figure 1. Time series of returns (x_t /x_{t-1}) after removal of seasonal and trend effects.

The principal components (PCs) are the uncorrelated linear combinations, $X_{(PC_i)}$, $X_{(PC_2)}$, . . . , $X_{(PC_p)}$, ranked by their variances in a descending order. The PCA used here is correlation-based and helps to identify the underlying transformed

Table 3. Descriptive statistics for the respective series of returns without seasonal and trend effects (*).

Variables	Min	Max	Mean	SD	CV
Total Energy Use (Returns)	0.6992	1.4891	1.0114	0.0871	0.0861
Total Transportation CO_2 Emission (Returns)	0.6515	1.6427	1.0138	0.1038	0.1024
Railway Share of Energy Use (Returns)	0.7505	1.1895	0.9970	0.0530	0.0532
Road Share of Energy Use (Returns)	0.7616	1.6734	1.0059	0.0780	0.0776
Waterway Share of Energy Use (Returns)	0.6308	1.3226	1.0035	0.0974	0.0970
Air Share of Energy Use (Returns)	0.5688	1.5768	1.0083	0.1044	0.1035
Railway Share of CO_2 Emission (Returns)	0.6661	1.2823	0.9974	0.0751	0.0753
Road Share of CO_2 Emission (Returns)	0.8210	1.4853	1.0038	0.0582	0.0580
Waterway Share of CO_2 Emission (Returns)	0.5965	1.4258	1.0027	0.0995	0.0992
Air Share of CO_2 Emission (Returns)	0.5845	1.5955	1.0092	0.1219	0.1208
Exchange Rate (RMB/U.S Dollar)	6.1043	8.2795	7.2968	0.8578	0.1176
Loan Rate (%)	4.3500	7.4700	5.6798	0.7377	0.1299
Deposit Rate (%)	1.5000	4.1400	2.4807	0.6852	0.2762
Consumer Price Index (%)	−1.8000	2.6000	0.1770	0.6843	3.8654
Total Trade (Returns)	0.6561	1.5129	1.0213	0.1306	0.1279
M2 Money Supply (Returns)	0.9873	1.0472	1.0123	0.0097	0.0096
Fiscal Expenditure (Returns)	0.1727	3.1913	1.1392	0.5306	0.4657
Consumer Confidence Index (Returns)	0.9088	1.0801	1.0008	0.0183	0.0182
Fixed Assets Investment (Returns)	0.0426	2.3933	1.2386	0.4709	0.3802
Imported Crude Oil Price Index (Returns)	0.7307	1.2181	1.0097	0.0702	0.0696

(*) Except for Loan Rate (%), Exchange Rate (RMB/U.S. Dollar), Loan Rate (%), Deposit Rate (%), and Consumer Price Index (%), which presented a stepwise behavior and, therefore, returns were computed directly, without removing trend and seasonal effects.

transportation and macroeconomic factors that are independent of each other. According to Tabachnick and Fidell (2007), only factor loads greater than 0.50 should be interpreted and taken into account.

Table 5 describes the four extracted factors for the macroeconomic variables: economic activity, credit policy, macroeconomic policy, and inflation. The economic activity factor encompasses the variables Total Trade and Fixed Assets Investment, reflecting the influence of the country's international trade and fixed-assets investment. Both global trade and fixed assets investment have played a crucial role in contributing to China's rapid economic growth, whereas the credit policy factor is based on the Loan Rate and Deposit Rate, reflecting the price of loans and deposits and the impacts of liquidity on the entire financial system. As the banking system dominates China's financial system, the interest spread is an important indicator reflecting the competitiveness and prosperity of the financial industry. The Monetary and Fiscal Policy factor is formed by M2 Money Supply and Fiscal Expenditure variables, indicating the potential effects of quantity-based monetary policy and government spending. Finally, the Inflationary Pressure factor, formed by the exchange rate, the Consumer Confidence Index, and the Imported Crude Oil Price Index, reflects the internal and external inflation pressure through foreign demands or the rising cost of bulk commodities. It covers almost all types of drivers of inflation.

As for the factor extraction of the transportation variables, the results are presented in Table 6. Five factors were extracted, four of which offered a straightforward interpretation for each transportation mode (road, waterway, air, and railway) concerning their respective share of the transportation footprint. The fifth factor, transportation footprint, represents the CO2 emissions and total energy use for both freight and passenger transportation of all the transportation modes. Not only do Tables 5 and 6 offer

Table 4. Correlation matrix between transportation and macro-economic variables.

	Total Energy Use	Total Transportation CO_2 Emission	Railway Share of Energy Use	Road Share of Energy Use	Waterway Share of Energy Use	Air Share of Energy Use	Railway Share of CO_2 Emission	Road Share of CO_2 Emission	Waterway Share of CO_2 Emission	Air Share of CO_2 Emission	Exchange Rate	Loan Rate	Deposit Rate	Consumer Price Index	Total Trade	M2 Money supply	Fiscal Expenditure	Consumer Confidence Index	Fixed Assets Investment	Imported Crude Oil Price Index
Total Energy Use	1.000																			
Total Transportation CO_2 Emission	0.973	1.000																		
Railway Share of Energy Use	−0.555	−0.690	1.000																	
Road Share of Energy Use	0.600	0.716	−0.794	1.000																
Waterway Share of Energy Use	−0.499	−0.420	0.224	−0.544	1.000															
Air Share of Energy Use	−0.439	−0.520	0.331	−0.483	0.184	1.000														
Railway Share of CO_2 Emission	−0.537	−0.692	0.984	−0.791	0.140	0.431	1.000													
Road Share of CO_2 Emission	0.570	0.637	−0.665	0.961	−0.722	−0.377	−0.632	1.000												
Waterway Share of CO_2 Emission	−0.599	−0.565	0.427	−0.704	0.969	0.333	0.365	−0.827	1.000											
Air Share of CO_2 Emission	−0.470	−0.576	0.458	−0.560	0.152	0.986	0.561	−0.427	0.331	1.000										
Exchange Rate	−0.018	−0.033	0.030	−0.052	0.020	0.020	0.018	−0.048	0.009	0.012	1.000									
Loan Rate	−0.024	−0.024	−0.016	−0.013	0.012	−0.021	−0.022	−0.014	0.001	−0.025	0.039	1.000								
Deposit Rate	−0.016	−0.011	−0.027	−0.002	0.017	−0.014	−0.029	−0.007	0.009	−0.018	−0.261	0.929	1.000							
Consumer Price Index	−0.398	−0.361	0.137	−0.146	0.162	0.188	0.135	−0.141	0.189	0.188	−0.054	0.064	0.053	1.000						
Total Trade	0.620	0.557	−0.193	0.239	−0.328	−0.267	−0.190	0.248	−0.364	−0.265	0.059	0.001	−0.001	−0.422	1.000					
M2 Money Supply	0.133	0.137	−0.020	0.092	−0.082	−0.127	−0.032	0.083	−0.100	−0.121	0.059	0.049	0.021	0.010	0.234	1.000				
Fiscal Expenditure	0.277	0.320	−0.200	0.268	−0.072	−0.383	−0.256	0.208	−0.158	−0.399	0.040	−0.001	−0.014	−0.106	0.425	0.389	1.000			
Consumer Confidence Index	−0.044	−0.057	0.062	−0.071	0.094	0.222	0.080	−0.057	0.121	0.214	−0.009	−0.072	−0.065	0.076	−0.050	−0.083	−0.009	1.000		
Fixed Assets Investment	0.626	0.592	−0.309	0.282	−0.235	−0.289	−0.303	0.252	−0.291	−0.302	0.036	0.022	0.028	−0.555	0.684	0.008	0.411	−0.059	1.000	
Imported Crude Oil Price Index	0.164	0.166	−0.165	0.127	−0.033	−0.044	−0.170	0.110	−0.070	−0.073	0.159	0.065	0.035	−0.028	0.182	−0.069	−0.083	0.162	0.166	1.000

Table 5. Macro-economic variables factor extraction.

Macro-economy				
	Macro-Economic Factors			
Macro-Economic Variables	Economic Activity	Credit Policy	Monetary and Fiscal Policy	Inflationary Pressure
Exchange Rate	(0.009)	(0.147)	0.226	**0.597**
Loan Rate	(0.029)	**0.969**	0.054	0.034
Deposit Rate	0.000	**0.981**	(0.029)	(0.139)
Consumer Price Index (CPI)	(0.784)	0.087	0.124	0.102
Total Trade	**0.782**	0.026	0.350	0.107
M2 Money Supply	(0.062)	0.038	**0.847**	(0.032)
Fiscal Expenditure	0.340	(0.015)	**0.743**	(0.052)
Consumer Confidence Index	(0.108)	(0.060)	(0.098)	**0.521**
Fixed Assets Investment	**0.900**	0.044	0.118	0.040
Imported Crude Oil Price Index	0.209	0.153	(0.161)	**0.758**
Kaiser-Meyer-Olkin Measure of Sampling Adequacy				0.440
Bartlett's Test of Sphericity	Chi-square			1064.782
	p value			1.0971E-193
	Degrees of freedom			45

Table 6. Transportation variables factor extraction.

Transportation					
	Transportation Factors				
Transportation Variables	Road Share	Waterway Share	Air Share	Railway Share	Transportation Footprint
Total Energy Use	0.049	(0.322)	(0.211)	(0.252)	**0.886**
Total Transportation CO_2 Emission	0.133	(0.246)	(0.292)	(0.412)	**0.816**
Railway Share of Energy Use	(0.074)	0.155	0.142	**0.937**	(0.264)
Road Share of Energy Use	**0.511**	(0.464)	(0.284)	(0.621)	0.238
Waterway Share of Energy Use	(0.056)	**0.978**	0.028	0.012	(0.194)
Air Share of Energy Use	(0.059)	0.119	**0.965**	0.130	(0.182)
Railway Share of CO_2 Emission	(0.101)	0.070	0.262	**0.923**	(0.250)
Road Share of CO_2 Emission	**0.519**	(0.658)	(0.176)	(0.476)	0.197
Waterway Share of CO_2 Emission	(0.083)	**0.932**	0.154	0.206	(0.238)
Air Share of CO_2 Emission	(0.078)	0.082	**0.934**	0.274	(0.196)
Kaiser-Meyer-Olkin Measure of Sampling Adequacy					0.561
Bartlett's Test of Sphericity	Chi-square				7578.790
	p value				0
	Degrees of freedom				45

interpretabilities, but they also maximize the Bartlett's Test for Sphericity based on the number of factors observed.

Principal Component Analysis (PCA) and factor decomposition approaches are two common methods used in multivariate data analysis. Both techniques are used to identify underlying patterns in data and reduce the dimensionality of data. However, there are some advantages of PCA over the factor decomposition approach:

- Simplicity (Jolliffe 2002): PCA is a simpler and more straightforward method than the factor decomposition approach. It involves identifying the principal components of the data and using them to explain the maximum variation in the data. In contrast, the factor decomposition approach is a more complex process that involves identifying factors that underlie the observed variables.

- Independence of components (Abdi and Williams 2010): PCA assumes that the principal components are independent of each other, which makes it easier to interpret the results. In contrast, the factor decomposition approach assumes that the underlying factors are correlated with each other, making it harder to interpret the results.
- Robustness (Hubert, Rousseeuw, and Vanden Branden 2005): PCA is a more robust method than the factor decomposition approach. It can handle missing data and outliers more effectively, whereas the factor decomposition approach is more sensitive to these issues.
- Data reduction (Jackson 1991): PCA can be used to reduce the dimensionality of data by selecting the principal components that explain the most variation in the data. This makes it easier to visualize and analyze large datasets. The factor decomposition approach does not offer a straightforward method for dimensionality reduction.

4.6. Neural network

Neural networks are learning models that combine mathematical and artificial intelligence techniques based on machine learning. By experience, they are able to distinguish underlying patterns, just like biological neural networks. NNs can compete with regression, classification, and clustering in data analysis. Perceptrons are the simplest neural network model, consisting of a single neuron, an activation function, and weights linking inputs (predictors) to outputs (dependent variables). They are similar to multivariate statistical models (Faraway 2005). A NN learns the weights from required training data to maximize forecast accuracy by uncovering their hidden interactions, whereas statistical models estimate parameters or weights based on probability distributions. NNs have been successfully used to identify complex nonlinear relationships between predictors and dependent variables in transportation-related studies (Liu et al. 2016; Liu et al. 2017). Karlaftis and Vlahogianni provide a useful comparison of the merits and disadvantages of neural network and statistical methods in the transportation field. They provide guidance on how to choose the approach in different situations. Liu et al. (2017) designed a system to account for the number of passengers in public transportation based on a convolutional neural network. Liu et al. (2016) proposed a gray neural network to offer a more desirable prediction of market demand after transportation disruption.

Predictive models traditionally suffer from a trade-off between accuracy and interpretation complexity. Prediction accuracy must be prioritized, as a validated predictive model may be helpful even under challenging interpretations, but it is of little value without statistical validation. As long as a series of robustness checks are conducted to remove the forecasting bias, a predictive model can be validated. Forecasting bias can be mitigated by increasing or decreasing the number of hidden layers and perceptions. The model considered for the NNs can be specified as follows:

Transportation Footprint = f(economic activity, credit policy, macroeconomic policy, inflation, road share, air share, waterway share, railroad share). (4)

In other words, the main hypothesis is that China’s transportation footprint is directly affected by macroeconomic factors. These factors are, in turn, influenced by the relative shares of each mode of transportation regarding carbon emissions and energy consumption. For example, using a specific eco-efficient mode of transportation may mitigate the impact of macroeconomic factors on transportation.

5. Data analysis and discussion of results

5.1. Parameter setting of the neural network and cross-validation

Tables 4 and 5 present the variables used in the NN analysis, which were extracted using PCA. In addition to the transportation footprint, other variables were included as explanatory variables in the NN model (cf. Equation (4)). For the parameter settings of the NN, Figure 2 shows the apparent Root Mean Square Error (RMSE) as a function of the number of hidden layers in the NN, considering the differences between actual and predicted values. The RMSE can be used to evaluate each predictive technique across various hidden layers. The RMSE tends to decline with increasing hidden layers, which suggests a preference for a higher number of hidden layers. Furthermore, by using different bootstrapping and resampling methods, RMSE can distinguish model performance within the validation process. Kuhn and Johnson provide detailed information on this topic (2013).

Thus, when applying cross-validation techniques, a typical pattern can be observed in Figure 3, where the RMSE peaks at lower values of hidden layers, decreases, and then levels off within the range of 0.805 to 0.825 in some methods. The neural network with the highest accuracy (RMSE = 0.7997215) was achieved using the 10-fold cross-validation technique, with 19 hidden layers and a decay rate of 0.0090.

5.2. Relative importance and sensitivity analysis of the macro-economic factors

Regarding the NN results, Figure 4 shows the relative importance of macroeconomic factors in predicting the transportation footprint. The Chinese government's development plans and initiatives aimed at sustaining accelerated economic growth have made economic activity (trade and investment in fixed assets) and monetary and fiscal

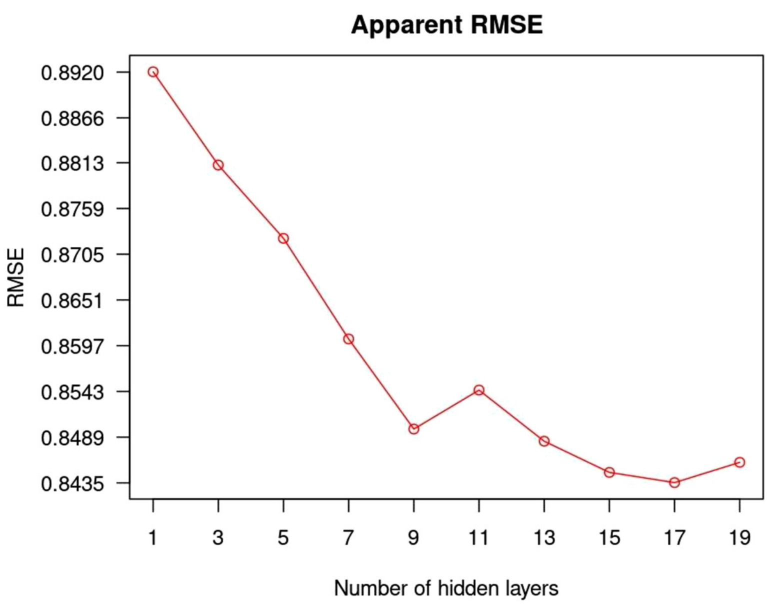

Figure 2. Apparent RMSE for different numbers of hidden layers.

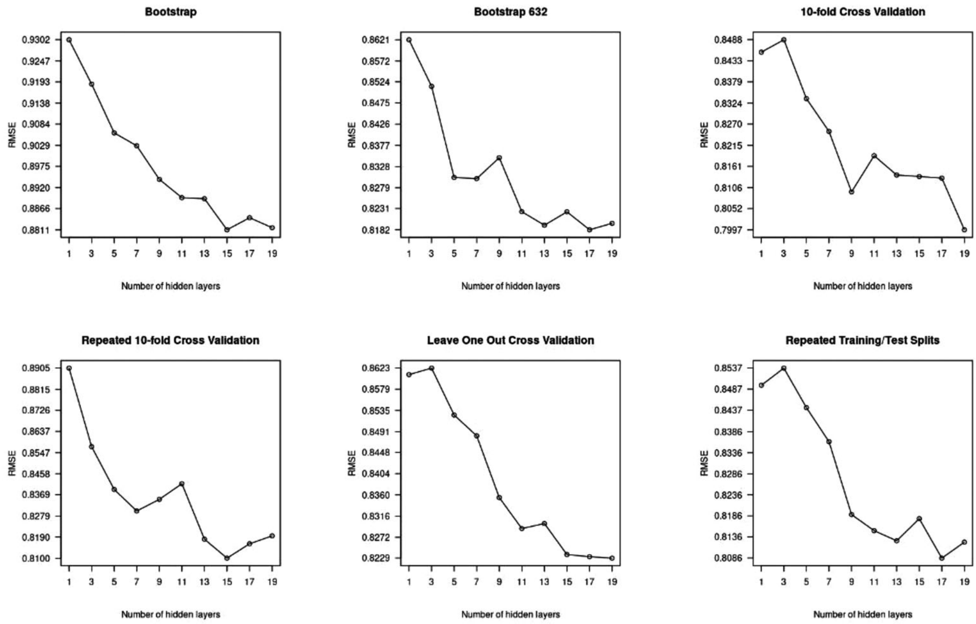

Figure 3. RMSE values for different numbers of hidden layers using various cross-validation procedures.

policies (M2 money supply and fiscal expenditure) the main factors affecting energy consumption and CO2 emissions in the transportation industry. Inflationary pressures (Consumer Confidence Index (CCI) imported oil price index, and exchange rate) and credit policy (loan rate and deposit rate) have a minor impact on transportation. Thus, trade opportunities, fixed investments, and government actions significantly influence CO2 emissions and energy consumption in goods and transportation. However, exchange rates and oil prices have a minor impact, possibly due to the

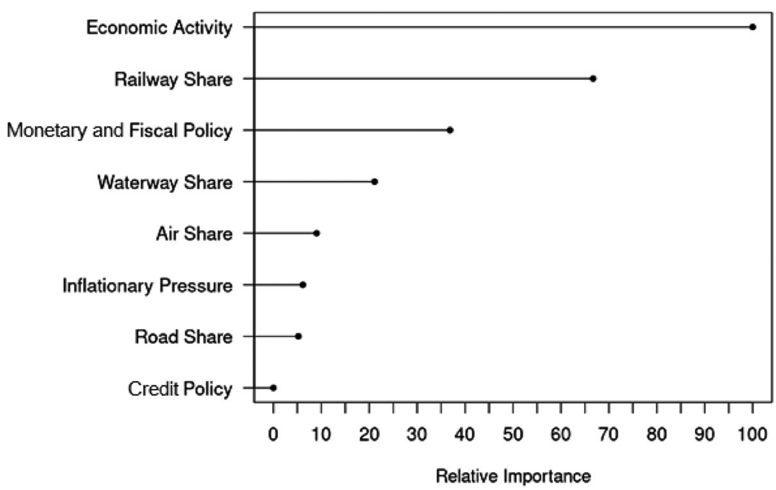

Figure 4. Relative importance of macro-economic factors and control variables.

Chinese economic development model, which prioritizes the production and supply of goods. The negative influence of increased economic activity and expansionary macroeconomic policy on the transportation footprint can be mitigated by using eco-efficient transportation modes, including waterways and railways, which can minimize CO2 emissions during the transportation of goods and people.

It is interesting to note that the railway share has greater importance than monetary and fiscal policy, which suggests the soundness of the government's efforts to encourage the use of eco-efficient transportation. We further argue that these plans can be easily accommodated in an expansionist economy driven by less restrictive monetary and fiscal policies. In light of this, the Chinese government should encourage the development of railway transportation. The country had 67,212 km of railways by the end of 2015, which ranked second in the world in terms of length. The country had 19,000 kilometers of high-speed rail (HSR), which ranked first in the world. Meanwhile, diesel locomotives have been almost entirely replaced with electric ones in recent decades, which have higher energy efficiency and emit less air pollution (Wanke et al. 2018a). Thus, it can be expected that the railway share plays a more significant role in the transportation footprint, in parallel with the dyad formed by economic activity and monetary and fiscal policies.

In the literature on transportation footprints, microeconomic factors have not been adequately considered as far as directional impacts are concerned (Wanke et al. 2018b). Figure 5 shows a sensitivity analysis of the directional impact of macroeconomic factors on the transportation footprint. In this sensitivity analysis, one explanatory variable is changed while the others are held constant at their mean values, as recommended by Faraway (2005). The contextual variables were standardized to have a mean of zero and a standard deviation of one. All factors, except inflationary pressure, had a positive impact on the transportation footprint. We should remind readers that, based on Equation (4), a positive impact on the transportation footprint means that increases in a given factor also contribute to increasing CO2 emissions and energy consumption in the Chinese transport system. Economic prices tend to increase in the short term, resulting in fewer goods and people being transported, while all other factors, including money supply, remain constant.

In contrast, the impact of different transportation modes on sustainability is well understood. Waterways, railroads, and airplanes typically emit more CO2 and consume more energy since they specialize in moving larger volumes of goods and passengers at lower speeds or smaller volumes of goods and passengers at higher speeds. However, contrary to preconceived expectations, the roadway share has the lowest coefficients of all transportation modes, as indicated in the sensitivity analysis presented in Figure 5. This reflects the fact that roadway transportation has much less room left to contribute to sustainable development in China. For the past few years, the Chinese government has implemented restrictive CO2 emission regulations. The sensitivity analysis indicates that all three other modes of transportation consume substantially more energy to meet considerably higher transportation demands. Moreover, as railway transportation shares increase, the impact of railway transportation on the transportation footprint decreases. Railway transportation will eventually diminish China's transportation footprint as its share of transportation increases. This may be explained by railway transportation's network scale effects (Xu et al. 2018). Therefore, an improved and increased railway infrastructure provides augmenting returns to scale in freight and people

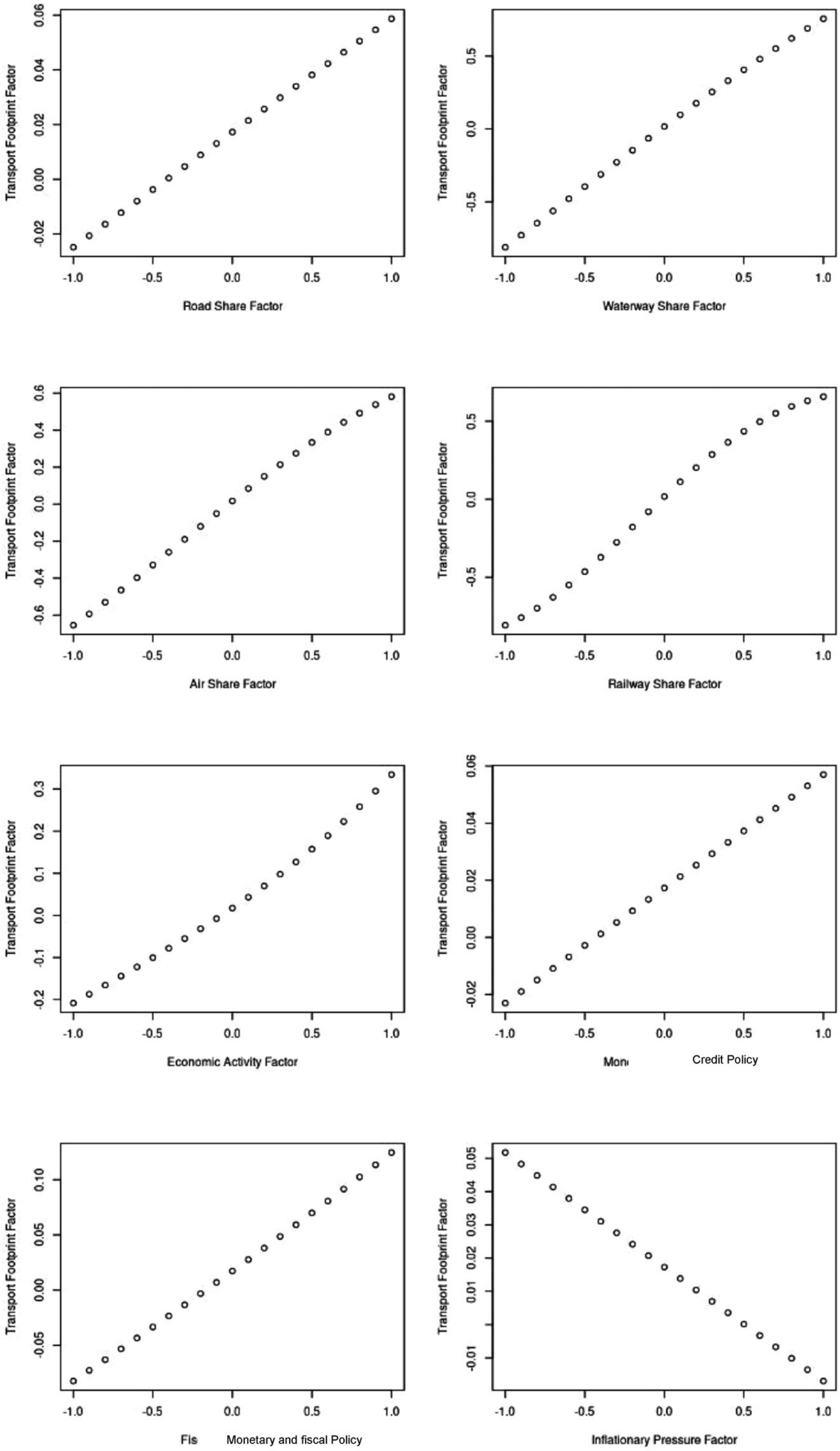

Figure 5. Sensitivity analysis on the directional impact of each factor on the transportation footprint.

transportation while enhancing the system's energy efficiency (Blair et al. 2017; Tsionas, Chen, and Wanke 2017; Darroch, Beecroft, and Nelson 2018). This finding is also consistent with the conclusion of Zuo et al. (2018), who analyzed the situation of the railway transportation mode in the U.K.

All these discussions have policy implications. Economic activity (trade and investment in fixed assets), as well as monetary and fiscal policies, should strongly focus on utilizing the higher efficiency of railways, airways, and waterways (rail, air, and water), which are most intensive in CO2 emissions and energy consumption. This is evident from the quasi-S shaped sensitivity curves presented in Figure 5. For the current state-of-the-art of technology, the transportation footprint may increase proportionally less as these shares increase. This quasi-S shape does not appear in the roadway mode sensitivity result, which rather shows a straight line. In combination with the low-scale transportation of goods and people, this suggests that automotive emission restrictions have so far produced the desired effect. Furthermore, multimodal network structures should be built to facilitate easy transshipment between modes to ensure widespread returns to scale for the Chinese transport system. It is cheaper to construct multimodal connections than to expand existing infrastructure. Consequently, these two alternative investments (multimodal connections vs. mode infrastructure) could be used countercyclically to mitigate the effects of inflationary pressures on economic activity. Investments in multimodal connections are lower when inflation rates are high, and monetary and fiscal policies are restrictive and higher when the economy is expanding with low inflationary pressure.

6. Conclusions

The relationship between sustainability in transportation and the macroeconomic environment in China was endogenous from 1999 to 2017, and analyses were conducted across different transportation modes. The results of this study not only fill gaps in the literature but also can aid the Chinese authorities in establishing effective policies and investments to reduce CO2 emissions, especially since China is ambitious in reducing carbon emissions and improving sustainable development. This study identified the most significant determinants affecting transportation footprints by employing PCA and NN. Besides investment conditions, economic activity, and macroeconomic policy, several factors influence the transportation footprint regarding energy consumption and CO2 emissions. To promote economic development, the government should pay attention to these two indicators. As discussed above, the study also highlights some policy implications. The first step in reducing emissions should be to clarify the objectives of policymakers.

Furthermore, it is necessary to decompose the target by the transportation sector. Railway transportation contributes significantly to carbon emissions in China as part of the transportation sector. Considering that the use of waterways and airways has a more substantial impact on energy consumption and CO2 emissions than the use of roads, the government should give them more attention than roads.

Compliance with ethical standards

There is no funding to report for this research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abdi, H., and L. J. Williams. 2010. “Principal Component Analysis.” *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (4): 433–459. doi:10.1002/wics.101.

Alizadeh, A. H. 2013. “Trading Volume and Volatility in the Shipping Forward Freight Market.” *Transportation Research Part E: Logistics and Transportation Review* 49 (1): 250–265. doi:10.1016/j.tre.2012.08.001.

Antunes, J., Y. Tan, P. Wanke, and C. J. C. Jabbour. 2023. “Impact of R&D and Innovation in Chinese Road Transportation Sustainability Performance: A Novel Trigonometric Envelopment Analysis for Ideal Solutions (TEA-IS).” *Socio-Economic Planning Sciences*, doi:10.1016/j.seps.2023.101544.

Azoff, E. M. 1994. *Neural Network Time Series Forecasting of Financial Markets*. New York: John Wiley & Sons, Inc.

Bao, D. W., S. Y. Hua, and J. Y. Gu. 2016. “Relevance of Airport Accessibility and Airport Competition.” *Journal of Air Transport Management* 55: 52–60. doi:10.1016/j.jairtraman.2016.04.009.

Barros, C. P., Z. Chen, and P. Wanke. 2016. “Efficiency in Chinese Seaports: 2002–2012.” *Maritime Economics & Logistics* 18 (3): 295–316. doi:10.1057/mel.2015.4.

Blair, John, Cielo Roldan, Sumita Ghosh, and Shih-Hsien Yung. 2017. “Greening Rail Infrastructure for Carbon Benefits.” *Procedia Engineering* 180: 1716–1724. doi:10.1016/j.proeng.2017.04.334.

Chang, X. Y., T. Ma, and R. Wu. 2018. “Impact of Urban Development on Residents’ Public Transportation Travel Energy Consumption in China: An Analysis of Hydrogen Fuel Cell Vehicles Alternatives.” *International Journal of Hydrogen Energy* 44 (30): 16015–16027. doi:10.1016/j.ijhydene.2018.09.099.

Chao, C. C. 2014. “Assessment of Carbon Emission Costs for Air Cargo Transportation.” *Transportation Research Part D: Transport and Environment* 33: 186–195. doi:10.1016/j.trd.2014.06.004.

Chen, Z., C. Barros, and Y. Yu. 2017a. “Spatial Distribution Characteristic of Chinese Airports: A Spatial Cost Function Approach.” *Journal of Air Transport Management* 59: 63–70. doi:10.1016/j.jairtraman.2016.11.011.

Chen, Z., P. Tzeremes, and N. G. Tzeremes. 2018. “Convergence in the Chinese Airline Industry: A Malmquist Productivity Analysis.” *Journal of Air Transport Management* 73: 77–86. doi:10.1016/j.jairtraman.2018.08.010.

Chen, Zhongfei, Peter Wanke, Jorge Antunes, and Ning Zhang. 2017b. “Chinese Airline Efficiency Under CO_2 Emissions and Flight Delays: A Stochastic Network DEA Model.” *Energy Economics* 68: 89–108. doi:10.1016/j.eneco.2017.09.015.

Darroch, N., M. Beecroft, and J. D. Nelson. 2018. “Going Underground: An Exploration of the Interfaces Between Underground Urban Transport Infrastructure and its Environment.” *Tunnelling and Underground Space Technology* 81: 450–462. doi:10.1016/j.tust.2018.08.027.

Du, K., C. Xie, and X. Ouyang. 2017. “A Comparison of Carbon Dioxide (CO2) Emission Trends among Provinces in China.” *Renewable and Sustainable Energy Reviews* 73: 19–25. doi:10.1016/j.rser.2017.01.102.

Dunteman, G. H. 1989. *Principal Components Analysis*. New York: Sage.

European Commission. 2016. EU Reference Scenario 2016: Energy, Transport and GHG Emissions - Trends to 2050. https://ec.europa.eu/energy/sites/default/files/documents/20160713%20EU%20Reference%20Scenario%202016.pdf.

European Environment Agency. 2019. Transport and Environment Reporting Mechanism (TERM) 2019: Tracking progress towards environmental targets in Europe. https://www.eea.europa.eu/publications/term-report-2019.

Faraway, J. J. 2005. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Florida: CRC Press.

Feng, C., and M. Wang. 2018. "Analysis of Energy Efficiency in China's Transportation Sector." *Renewable and Sustainable Energy Reviews* 94: 565–575. doi:10.1016/j.rser.2018.06.037.

Guimarães, V. A., I. C. L. Junior, and M. A. V. Silva. 2018. "Evaluating the Sustainability of Urban Passenger Transportation by Monte Carlo Simulation." *Renewable and Sustainable Energy Reviews* 93: 732–752. doi:10.1016/j.rser.2018.05.015.

Hadi-Vencheh, Abdollah, Peter Wanke, Ali Jamshidi, and Zhongfei Chen. 2018. "Sustainability of Chinese Airlines: A Modified Slack-based Measure Model for CO_2 Emissions." *Expert Systems* 37 (3): 1–11. doi:10.1111/exsy.12302.

Hair, Jr., Joseph F, Rolph E. Anderson, Ronald L Tatham, and Willam C Black. 1995. *Multivariate Data Analysis*. 5th ed. New Jersey: Prentice Hall.

Halkos, G. E., and E. A. Paizanos. 2016. "The Effects of Fiscal Policy on CO2 Emissions: Evidence from the USA." *Energy Policy* 88: 317–328. doi:10.1016/j.enpol.2015.10.035.

Heinold, A., and F. Meisel. 2018. "Emission Rates of Intermodal Rail/Road and Road-Only Transportation in Europe: A Comprehensive Simulation Study." *Transportation Research Part D: Transport and Environment* 65: 421–437. doi:10.1016/j.trd.2018.09.003.

How, B. S., and H. L. Lam. 2018. "Sustainability Evaluation for Biomass Supply Chain Synthesis: Novel Principal Component Analysis(PCA) Aided Optimisation Approach." *Journal of Cleaner Production* 189: 941–961. doi:10.1016/j.jclepro.2018.03.104.

Hu, R., L. Chen, and L. J. Zheng. 2018. "Congestion Pricing and Environmental Cost at Guangzhou Baiyun International Airport." *Journal of Air Transport Management* 70: 126–132. doi:10.1016/j.jairtraman.2018.04.016.

Huang, Shupei, Haizhong An, Silvio Viglia, Gabriella Fiorentino, Fabiana Corcelli, Wei Fang, and Sergio Ulgiati. 2018. "Terrestrial Transport Modalities in China Concerning Monetary, Energy and Environmental Costs." *Energy Policy* 122: 129–141. doi:10.1016/j.enpol.2018.06.047.

Hubert, M., P. J. Rousseeuw, and K. Vanden Branden. 2005. "ROBPCA: A New Approach to Robust Principal Component Analysis." *Technometrics* 47 (1): 64–79. doi:10.1198/004017004000000648.

Huo, W. W., W. Zhang, and P. S. Chen. 2018. "Recent Development of Chinese Port Cooperation Strategies." *Research in Transportation Business & Management* 26: 67–75. doi:10.1016/j.rtbm.2018.01.002.

International Energy Agency. 2017. Global EV Outlook 2017: Two Million and Counting. https://www.iea.org/reports/global-ev-outlook-2017.

International Transport Forum. 2019. Transport Innovation for Sustainable Development. https://www.itf-oecd.org/sites/default/files/docs/19cpb_transport_innovation_for_sustainable_development.pdf.

IPCC (Intergovernmental Panel on Climate Change). 2015. *Climate Change 2014: Mitigation of Climate Change (Vol. 3)*. New York: Cambridge University Press.

Jackson, J. E. 1991. *A User's Guide to Principal Components*. Wiley Online Library. doi:10.1002/9780470316801.

Jolliffe, I. T. 2002. *Principal Component Analysis*. Wiley Online Library. doi:10.1002/9781118445112.stat02052.

Karlaftis, M. G., and E. I. Vlahogianni. 2009. "Memory Properties and Fractional Integration in Transportation Time-Series." *Transportation Research Part C: Emerging Technologies* 17 (4): 444–453. doi:10.1016/j.trc.2009.03.001.

Khan, H. U., M. Siddique, K. Zaman, S. U. Yousaf, A. M. Shoukry, S. Gani, A. Khan, S. S. Hishan, and H. Saleem. 2018. "The Impact of air Transportation, Railways Transportation, and Port

Container Traffic on Energy Demand, Customs Duty, and Economic Growth: Evidence from a Panel of low-, Middle-, and High-Income Countries." *Journal of Air Transport Management* 70: 18–35. doi:10.1016/j.jairtraman.2018.04.013.

Kim, E., G. Hewings, and H. Amir. 2017. "Economic Evaluation Projects: Am Application of Financial Computable General Equilibrium Model." *Research in Transportation Economics* 61: 44–55. doi:10.1016/j.retrec.2016.09.002.

Lao, Xin, Xiaoling Zhang, Tiyan Shen, and Martin Skitmore. 2016. "Comparing China's City Transportation and Economic Networks." *Cities* 53: 43–50. doi:10.1016/j.cities.2016.01.006.

Li, Z., N. Meng, and X. Yao. 2017. "Sustainability Performance for China's Transportation Industry Under the Environmental Regulation." *Journal of Cleaner Production* 142: 688–696. doi:10.1016/j.jclepro.2016.09.041.

Liu, Y., and C. Feng. 2020. "Decouple Transport CO2 Emissions from China's Economic Expansion: A Temporal-Spatial Analysis." *Transportation Research Part D: Transport and Environment* 79: 102225. doi:10.1016/j.trd.2020.102225.

Liu, Chunxia, Tong Shu, Shou Chen, Shouyang Wang, Kin Keung Lai, and Lu Gan. 2016. "An Improved Grey Neural Network Modal for Predicting Transportation Disruptions." *Expert Systems with Applications* 45: 332–340. doi:10.1016/j.eswa.2015.09.052.

Liu, Guojin, Yin Zhenzhi, Jia Yunjian, and Xie Yulai. 2017. "Passenger Flow Estimation Based on Convolutional Neural Network in Public Transportation System." *Knowledge-Based Systems* 123: 102–115. doi:10.1016/j.knosys.2017.02.016.

Meinshausen, Malte, Nicolai Meinshausen, William Hare, Sarah C. B. Raper, Katja Frieler, Reto Knutti, David J. Frame, and Myles R. Allen. 2009. "Greenhouse-gas Emission Targets for Limiting Global Warming to 2 C." *Nature* 458 (7242): 1158–1162. doi:10.1038/nature08017.

Nagendra, S. S., and M. Khare. 2004. "Artificial Neural Network Based Line Source Models for Vehicular Exhaust Emission Predictions of an Urban Roadway." *Transportation Research Part D: Transport and Environment* 9 (3): 199–208. doi:10.1016/j.trd.2004.01.002.

Pan, Xunzhang, Hailin Wang, Lining Wang, and Wenyi Chen. 2018. "Decarbonization of China's Transportation Sector: In Light of National Mitigation Toward the Paris Agreement Goals." *Energy* 155: 853–864. doi:10.1016/j.energy.2018.04.144.

Schodl, R., S. Eitler, B. Ennser, A. Breinbauer, B. Hu, K. Markvica, M. Prandtstetter, et al. 2018. "Innovative Means of Cargo Transport: A Scalable Method for Estimating Regional Impacts." *Transportation Research Procedia* 30: 342–349. doi:10.1016/j.trpro.2018.09.037.

Tabachnick, B. G., and L. S. Fidell. 2007. *Using Multivariate Statistics*. Boston: Allyn and Bacon.

Tsionas, M. G., Z. Chen, and P. Wanke. 2017. "A Structural Vector Autoregressive Model of Technical Efficiency and Delays with an Application to Chinese Airlines." *Transportation Research Part A: Policy and Practice* 101: 1–10. doi:10.1016/j.tra.2017.05.003.

United Nations Environment Programme. 2019. Emissions Gap Report 2019. https://wedocs.unep.org/bitstream/handle/20.500.11822/30797/EGR2019.pdf.

Wang, T., and B. Lin. 2019. "Fuel Consumption in Road Transport: A Comparative Study of China and OECD Countries." *Journal of Cleaner Production* 206: 156–170. doi:10.1016/j.jclepro.2018.09.092.

Wang, H. L., X. Ou, and X. L. Zhang. 2017. "Technology, Energy Consumption, and Resulting CO2 Emissions in China's Transport Sector up to 2050." *Energy Policy* 109: 719–733. doi:10.1016/j.enpol.2017.07.010.

Wanke, Peter, Chen Zhongfei, Antunes Jorge, and Carlos Barros. 2018a. "Malmquist Productivity Indexes in Chinese Ports: A Fuzzy GMSS DEA Approach." *International Journal of Shipping and Transport Logistics* 10 (2): 202–236. doi:10.1504/IJSTL.2018.090094.

Wanke, Peter, Chen Zhongfei, Wankun Liu, Jorge Antunes, and Md. Abul Kalam Azad. 2018b. "Investigating the Drivers of Railway Performance: Evidence from Selected Asian Countries." *Habitat International* 80: 49–69. doi:10.1016/j.habitatint.2018.08.004.

World Bank. 2019. Connecting to Compete 2018: Trade Logistics in the Global Economy. https://openknowledge.worldbank.org/handle/10986/31330.

Xia, Y.-S., L.-X. Sun, and C. Feng. 2022. "What Causes Spatial Inequalities of low-Carbon Development in China's Transport Sector? A Newly Proposed Meta-Frontier DEA-Based Decomposition Approach." *Socio-Economic Planning Sciences* 80: 101151. doi:10.1016/j.seps.2021.101151.

Xie, R., J. Y. Fang, and C. J. Liu. 2017. "The Effects of Transportation Infrastructure on Urban Carbon Emissions." *Applied Energy* 196: 199–207. doi:10.1016/j.apenergy.2017.01.020.

Xu, Wangtu, Jiangping Zhou, Linchuang Yang, and Ling Li. 2018. "The Implications of High-Speed Rail for Chinese Cities: Connectivity and Accessibility." *Transportation Research Part A: Policy and Practice* 116: 308–326. doi:10.1016/j.tra.2018.06.023.

Zegras, Christopher, Joshua Nelson, Rosario Macário, and Christopjer Grillo. 2013. "Fiscal Federalism and Prospects for Metropolitan Transportation Authorities in Portugal." *Transport Policy* 29: 1–12. doi:10.1016/j.tranpol.2013.03.003.

Zhang, N., K. Yu, and Z. Chen. 2017. "How Does Urbanization Affect Carbon Dioxide Emissions? A Cross-Country Panel Data Analysis." *Energy Policy* 107: 678–687. doi:10.1016/j.enpol.2017.03.072.

Zhu, Z. R., A. M. Zhang, and Y. H. Zhang. 2018. "Connectivity of Intercity Passenger Transportation in China: A Multimodal and Network Approach." *Journal of Transport Geography* 71: 263–276. doi:10.1016/j.jtrangeo.2017.05.009.

Zuo, Chengchoa, Mark Birkin, Graham Clarke, Fiona McEvoy, and Andrew Bloodworth. 2018. "Reducing Carbon Emissions Related to the Transportation of Aggregates: Is Road or Rail the Solution?" *Transportation Research Part A: Policy and Practice* 117: 26–38. doi:10.1016/j.tra.2018.08.006.