Numerical Methods in Civil Engineering, 7-3 (2023) 66-77

Numerical Methods in Civil Engineering

Journal Homepage: https://nmce.kntu.ac.ir/

Freight production and attraction of industrial, agricultural and livestock, food, and fruit and vegetable commodities

Seyed Ali Ghaemi*, Mansour Hadji Hoseinlo**

ARTICLE INFO

APPLICABLE PAPER

Article history:
Received:
February 2022
Revised:
June 2022
Accepted:
August 2022

Keywords:
Freight transport models, Freight generation modelling, Sampling, Multiple linear regression

Abstract:

This paper analyzes freight production and attraction and their relationship with traffic analysis zone (TAZ) features. The effects of some parameters on the production and attraction of industrial, agricultural and livestock, food, and fruit and vegetable freight were evaluated using over 300 explanatory variables, i.e., land-use types, the numbers and areas of businesses, the characteristics of residents and employees, employment, land price, vehicle ownership per capita, and road network, and TAZ descriptors. The 2019 comprehensive master plan of Shiraz, Iran, in 325 TAZs was employed. A vehicle survey and roadside interviews were used to collect data in three cordons involving 143 stations. Vehicles of different types were counted in the vehicle survey, and the roadside interview forms included questions on the travel time, vehicle type, vehicle capacity, freight amount, freight type, and travel origin and destination. Then, the freight origin-destination (O-D) matrix was constructed. To evaluate the effects of variables on the production and attraction of industrial, agricultural and livestock, food, and fruit and vegetable freight using aggregated data in the TAZs of Shiraz, Iran, for different freight, modeling was performed based on minimizing the residual sum of squares, proposing a total of eight models. The adjusted coefficient of determinations (adjusted R2) was calculated to be larger than 0.55 for all eight models. In addition, the models were found to have root-mean-square errors (RMSE) below 130.

1. Introduction

Today, urban traffic and increased urban trips remain a major concern in metropolises. It is important to cope with increased freight transport traffic as it raises traffic congestion and causes traffic conflicts and car accidents. It is useful to predict freight travel demands at local and regional scales to estimate traffic flows in specific paths and plan for developing efficient infrastructures in the future. Since freight and passenger trips typically use roads with a right of way, travel demand growth in freight transportation increases the general travel cost for not only freight transportation but also passengers. The general travel cost can be minimized by using systematic planning and developing efficient transportation infrastructures. Furthermore, it should be mentioned that transportation is a fundamental requirement in society and plays a key role in the economy of a city.

Problem awareness and understanding are the first step in planning. To obtain awareness of future travel demands, a comprehensive master plan (CMP) can be implemented by urban transportation planning organizations. Financial resource allocation and investment decision-making in transportation projects are often carried out based on CMPs. Also, the transportation demand should be coordinated with the available transportation capacity. Travel demand analysis in four-step modelling, including trip generation, trip distribution, modal split, and traffic assignment, is a major component of a CMP. However, four-step freight transportation modelling efforts have been rare in comparison to passenger transportation modelling in developing countries, particularly Iran, probably due to limited information sources, limited public access to freight transportation databases, unknown (and difficultly identifiable) variables, commodity diversity, and urban policy priorities. Therefore, the present

* *Corresponding author: Department of Civil and Environmental Engineering, K. N. Toosi University of Technology, PHD Candidate, Tehran, Iran. Email: ghaemiseyedali@gmail.com*

***Department of Civil and Environmental Engineering, K. N. Toosi University of Technology, Associated Professor, Tehran, Iran.*

https://doi.org/10.52547/NMCE.2022.407

study seeks to model freight production and attraction to help quantitatively understand freight transport demand in Shiraz, Iran.

Freight transportation demand models are more complex than passenger travel demand models. Freight travel demand is dependent on numerous factors, and each freight type has different demand factors from other freight. Therefore, the separate modelling of each freight type is an important step in freight transportation demand modelling. The modelling of freight demand is a very complicated process due to a variety of unknown and difficultly identifiable and measurable variables, even after the classification of models. However, freight transportation importance for the efficient performance of an urban transportation system is gradually appreciated (Patil et al., 2021). The separate modelling of different freight types has not been studied in Iran.

It is necessary to determine the transported freight amount in each traffic analysis zone (TAZ) of an urban area in order to evaluate the effects of regional urban policies in terms of congestion and air pollution. Therefore, this study measures the effects of influential factors on the production and attraction of industrial, agricultural and livestock, food, and fruit and vegetable freight using aggregated data in TAZs of Shiraz for different freight types through multiple regression. In other words, this paper seeks to identify the relationship between urban features and freight production and attraction. The present study mainly aims to construct freight production and attraction models by considering freight differences.

Considering the differences between the present research and previous researches, these cases can be mentioned. Previous studies developed models with only one or two independent variables or small fit indices. The present study developed models based on 325 observations (many more than earlier works). In general, the present work employed many more measured quantitative explanatory variables to tackle the shortcomings of previous works that relied on descriptive datasets, few data points, or extrapolating data from complex quantitative models. Moreover, the proposed models utilized aggregated data of TAZs, unlike most previous studies. Also, a freight production and attraction model may estimate the freight demand based on the number of freight vehicles or freight weight (tons). The former is known as freight trip generation (FTG), while the latter is referred to as freight generation (FG). Since a systematic method of monitoring commodity flow is lacking in Iran, planners and policymakers can exploit the proposed FG model in urban transportation planning and transportation strategy development.

One of the results of this research, which is the estimate of the amount of freight production and attraction by the type of freight at the level of traffic analysis zone, can be used to plan the development of urban and infrastructure in the year of the plan horizon. Also, because currently, due to the problems and limitations mentioned in the previous paragraphs, city managers use estimation methods to estimate the amount of freight in the coming years. Therefore, the developed models of freight production and attraction in this article can help to accurately estimate the amount of freight in the coming years based on scientific forecasting and as a result, accurate planning for the development of the necessary infrastructure. Another result of this research is the knowledge of the influential variables and the degree of influence of each of them on the freight production and attraction, which can lead city managers to make better decisions regarding the allocation of budget in each of the sectors with accurate knowledge need to have. In this way, the waste of urban budget as a result of allocating budget to unnecessary things or not allocating budget to needed things is avoided so that urban development has a correct and fast process. This item is of great value considering the budget constraints.

It should be noted that this article has five parts. The first part deals with the introduction of the article and the second part reviews the literature of the subject area. The third part introduces the methodology, including the introduction of the study area, the method of gathering and data collection, the type of variables used in the model and the method of data processing, modelling method and statistical tests. The fourth part, by conducting statistical tests and checking the establishment of regression assumptions, presents the results of modelling and discussion regarding the outputs of the model. And in the fifth section, a summary has been made and the results of the current research have been compared with the results of other researches in the subject area.

2. Literature review

Since this article intends to examine the influential variables and their impact on freight production and attraction using modeling, it has been reviewed the research literature in several sections. The first part reviews the advantages and disadvantages of the classification model in terms of the type of freight and city zoning and refers to relevant researches. In the second part, in the second part, the difference between the models in terms of modeling of the estimation of the number of freight vehicles or freight weight is discussed. In the third part, many researches have been investigated in order to be aware of all the variables affecting the amount of production and load absorption, so that if possible, they can be used in this article. In the fourth part, the way of modeling in researches with the same purpose as the current research has been investigated.

Freight transportation production models are in their infancy in many developing countries, including Iran. A number of studies have been recently conducted in India [1] [2] [3], Ecuador [4], and Turkey [5], and the freight transportation structure differences of such countries from developed countries were mentioned [4] [6]. The presence of various industrial sectors and stakeholders and a shortage of quantitatively measurable criteria are explanations for the lower development of freight models than passenger models [7]. The development of a freight transportation model in several layers, each representing a particular type of industrial commodities or an industrial freight production sector, is an approach to tackle these challenges [8]. The modeling domain is restricted to industrial sectors to better understand demand production by particular sectors, their relationships, and planning areas [9]. Particular models of an industry can be separately employed using such layered approaches to investigate policies and the operation of the industry or to estimate the production and attraction of the freight in a city, province, or country. A number of freight transportation demand modeling studies reported that models developed based on industrial classification systems [5] outperformed those developed using land-use classes [1] [10] [11]. Moreover, some studies zoned the city for the spatial analysis of freight transportation demand [12].

As mentioned, freight production and attraction can be modeled by estimating freight weights (FG) or the number of freight vehicle trips (FTG) [13] [14] [15]. However, a number of researchers suggested that FG models were more efficient and effective than FTG models due to the direct dependence of business size variables (e.g., employment) on the commodity production and attraction scale, while FTG models are influenced by logistic strategies and decisions [13]. In light of using industry classification, FG models are more likely to be a better representative than FTG models for freight transportation description [16]. FTG models are generally utilized to indicate traffic consequences of transportation flows in a city, while FG models are employed to design and evaluate infrastructural transportation plans[5]. Moreover, FG models can be considered for realizing and measuring the transportation demands of major centers in terms of fleet vehicle types and capacities to transport the produced freight [17].

As mentioned, freight demand models typically use classification systems to classify institutions into assumed groups of homogenous patterns in freight production (FP) and freight trip production (FTP). Although such classification systems are an appealing and popular concept, the assumption of homogeneity in industrial classification has a reduction nature. Pani and Sahu evaluated this assumption and investigated the feasibility of data-based segmentation by measuring relationships between FP and FTP patterns and common classes. They created homogenous ensembles of posterior segmentations via aggregation. The alternative segmentation schemes were compared in the prediction of FP and FTP. Industrial classification systems (NAICS and ISIC) were found to significantly outperform product classification systems (ASICC). Furthermore, it was observed that a significant portion of diversity in FTP was independent of employment due to the fundamental effect of the freight size, and another segmentation scheme considering the freight size might be an effective middle ground to develop both FTP and FP models in transportation demand models [18]. As a result, researchers modeled only specific types of freight. Dhulipala and Patil modeled agricultural freight [19]. Diaz studied freight demand models in food and residential institutions [20]. Fliu and Pela analyzed the relationship between shopping trips and freight production in urban areas [21]. Furthermore, researchers focused on modeling retail malls freight in Singapore .[22]

A review of the literature indicates that previous studies investigated parameters such as the freight production and attraction center size using a variety of variables, e.g., the number of employees [3] [17] [23] [24], the number of customers [11], area [1] [23] [25] [26] [27], sales[28], business age and size [29], days of the week [25], industry type [30], the number of parking spaces [28], vehicle ownership, and economic variables (such as gross domestic product) [28]. A number of studies incorporated explanatory variables into FTG models, including the number of residents [31], employed population [2], and the number of trucks in each TAZ. Many researchers investigated several relationships between freight delivery and variables such as employment, location, and industry type (e.g., Ambrosini et al. [13] [23] [32]. Alho and Silva suggested that employment and area were the most important explanatory variables since their data could be used by planners and experts to further calibrate, estimate, and infer models [23]. However, the descriptive power of employment and area significantly varies, depending on whether an FG or an FTG model is implemented [32]. Jaller et al. reported employment-based FTG models to outperform area-based FTG models in the industrial sectors of New York [26]. This finding is consistent with Holgovin et al. [13]. They showed that land availability for facilities as a constraint acted as input for economic processes. Therefore, it restricts the ground floor net area to describe institution-generated freight. However, a comparative evaluation of employment and area in predictive power remains to be performed on FG models. Lavsen et al. [10] and Sanchez-Diaz et al. [17] studied other variables, including land use and location. McLeod et al. argued that commodity production and transportation patterns were not efficiently understood by planners and policymakers responsible for making complicated, strategic decisions on land use and transportation planning [33]. In the absence of detailed planning evidence, they may rely on little or descriptive data provided and extrapolated through complex quantitative models. Unfortunately, the outputs of predictive models may fail to completely agree with observations. Such models cannot predict complicated, long-term phenomena that may be generated alternatingly and disrupt commodity movement patterns. These researchers highlight a wider set of freight production factors, including demand factors such as land-use and corporate strategy changes [33]. Gonzalez-Feliu and Sánchez-Díaz studied the impacts of the aggregation level and category construction on the quality and relevance of FTG models [34]. They proposed a technique to compare constant generations and functional form models based on various classifications and estimating the mean average percentage error (MAPE). They evaluated functional forms using linear regression and compared them by the Pearson correlation coefficient. It was found that the aggregation level would not necessarily have a positive impact on the accuracy of an FTG model, and the optimal functional form improves model accuracy. Bakshi et al. investigated the effects of land use on freight transportation in New Delhi, India [35]. Boarnet et al. explored the effects of multi-center urban development patterns on freight transportation activity [36].

The ordinary least square (OLS) method is an efficient statistical approach to evaluate the contributions of independent variables, and models have been mostly developed using OLS based on trip estimation [13], [19], [23], [25], [37], [38], [39], [40]. Studies have applied OLS as one of the most efficient models [41]. OLS was employed in many vehicle-based models [42] [38] [43]. These models calculate the origin-destination (O-D) matrix directly based on commodity flows in different industrial sectors. Also, multiple classification analysis (MCA) is used in freight production and attraction models. Alho and de Abreu e Silva [23] employed MCA. However, earlier works mentioned that MCA does not impose a functional form and freely changes MCA parameters in independent variable intervals [10]. Adaptive analysis indicated that MCA had slightly higher performance than regression models in the estimation of freight trip attraction; however, there were differences, considering the overall error [32].Compared to recent interest in searching for more flexible models such as MCA [44] [45], it remains yet to be clarified whether MCA is as efficient as regression models to be estimated based on FG.

Except for FM model estimations, the geographical analysis of parameters has the same importance since freight transportation is very sensitive to transportation costs [46]. Thus, the location features of freight production and attraction centers (e.g., closeness to seaports) lead to a change in business transactions and thus a change in FG patterns. As a result, variables that record locations, even if approximately, play a key role in freight transportation demand modeling [17]. For example, facilities that are located near a large interstate highway and can exploit geographical advantages to promote their business and have higher access to customers are likely to result in the generation of more transportation orders than an institution located far from a highway. The statistical importance analysis of such variables enables a method for evaluating the geographical stability of FG models [47]. The literature suggests that FG patterns differ in different zones, and a significant difference in FG often occurs in a state [44]. Therefore, national models or models constructed in an adjacent zone may induce significant errors in FG estimation. Apart from the main calculation methods such as multiple regression, other techniques have been adopted in freight transportation modeling, including Poisson regression models [48], zero-inflated negative binomial models and maximum likelihood estimator (MLE) [49], spatial econometric techniques such as the spatially lagged X (SLX) model, spatial Durbin model (SDM), and spatial Durbin error model (SDEM) [50], least squares techniques in two steps and instrumental variables, Monte Carlo method, and spatial autoregressive (SAR) models [51], artificial intelligence networks (ANNs) [52], generalized additive modeling (GAM) [21], conditional freight trip generation [53], time series [54], ordinal logit model [23], and machine learning (ML) [2].

The authors found that no FG model was available to evaluate transportation demands for different freight types in Iranian cities – or at least such models were so rare that the authors were not able to find one. Hence, it is crucial to model FG for different types of freight to obtain comprehensive insights into freight transportation in urban areas in Iran.

According to the literature review, in this article, the models will be made according to the type of freight and traffic zoning of the study area. The area of each of the conducted zoning units has a smaller size compared to other studies, and increasing the number of zoning units will increase the accuracy of the models. Also, according to the literature, the available information and the purpose of the research, in this article, the freight weight estimation model is used. Also, based on the recognition of effective variables from previous researches, necessary data and information will be collected and compiled. In each of the models, this article has used a larger number of significant variables compared to previous studies, so it has achieved a better fit. On the other hand, according to the literature review, the authors decided to use the multiple regression method due to the presentation of a functional model that can be used practically and to understand the extraction of the effect of each variable. Because some types of models are similar to a black box despite benefiting from a good fit and lack the value of understanding the relationship between variables. In this article, considering that four-step freight transportation modelling is rarely done in Iran, It was tried to strengthen the understanding of the relationships between the influencing variables by identifying more influential variables and using them in the multiple regression model in addition to the comprehensiveness of the study.

3. Methodology

Despite several transportation modeling approaches in the literature, it can be said that the selection of a modeling method is dependent on data availability and modeling objective. Considering the data quality of Shiraz, the present study adopted an FG model based on freight production and attraction for different TAZs and different freight types, including industrial, agricultural and livestock, fruit and vegetables, and food.

Shiraz is the capital city of Fars Province, Iran. It is located at 29.37° N-52.32°E, 919 km sought of Tehran. Shiraz occupies an area of 193 km^2 and has a population of 1,527,149 (8,160 per km^2) at an elevation of 1484 m in the Zagros Mountains. It has a moderate climate. Expansion limitations in the north and sought have led to physical expansion eastward and westward in Shiraz. The Shiraz Municipality has 11 districts.

The data of the dependent variables, including the production and attraction of industrial, agricultural and livestock, fruit and vegetable, and food freight, were obtained using vehicle surveying and roadside interviews in three cordons and five major freight production and attraction centers in the transportation CMP of Shiraz. The first cordon included portal information (a total of six portals), while the second and third cordons included 121 vehicle count stations and 43 interview stations in the incoming and outgoing lanes, respectively. The same procedure was implemented in the five major freight production and attraction centers of Shiraz, including Cargo Terminal, Shiraz Central Fruit and Vegetable Market, Shiraz Industrial Town, Shiraz Special Economic Zone, and Merchants Complex, as shown in Fig. 1.

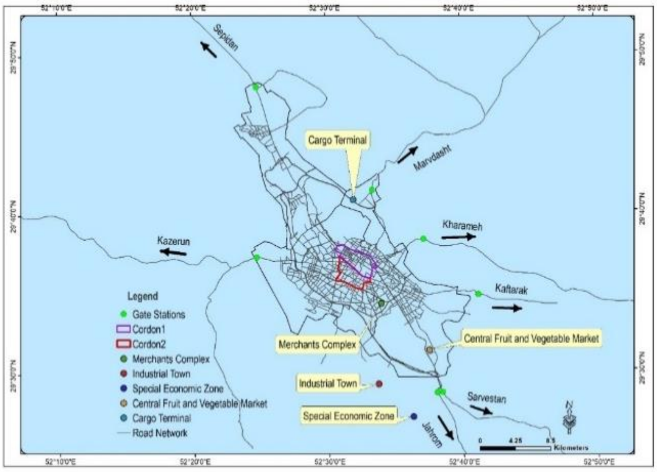

Fig. 1: Roadside survey stations

Different types of vehicles were separately counted in the vehicle survey. Moreover, the travel time, vehicle type, vehicle capacity, freight amount, freight type, travel origin, and travel destination, were inquired in the roadside interview forms. The data were collected in a 24-hour period from over 2,000 interviewees and traffic cameras of the Shiraz Municipality. The O-D matrix was built by aggregating the data. The travel data were introduced in the form of a matrix to EMME. Then, the O-D matrix was calibrated using screen lines and singular points, as shown in Fig. 2.

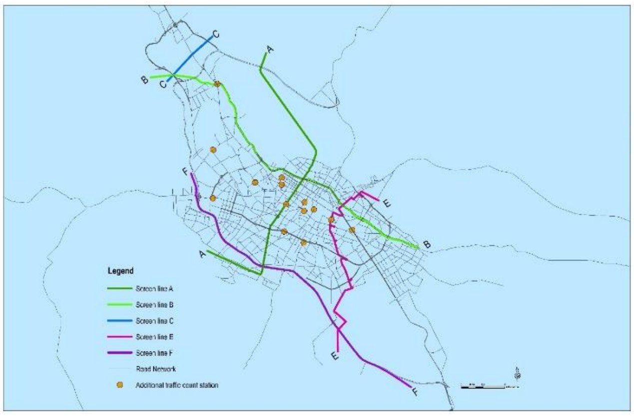

Fig. 2: Screen Lines

Additionally, the data of the explanatory (independent) variables were obtained using information from other organizations in Shiraz. It should be noted that such information were validated and digitized. Furthermore, since the model would be processed at the TAZ level, the data of the independent and dependent variables in 325 TAZs of Shiraz were aggregated to obtain 325 observations in the construction of each model. This phase was implemented by introducing the spatial data of all the variables in different layers and TAZs to ArcGIS Fig. 3 and Fig. 4 illustrates the aggregated data of some independent variables and Fig. 5 illustrates the Freight production and attraction of fruit and vegetable commodities.

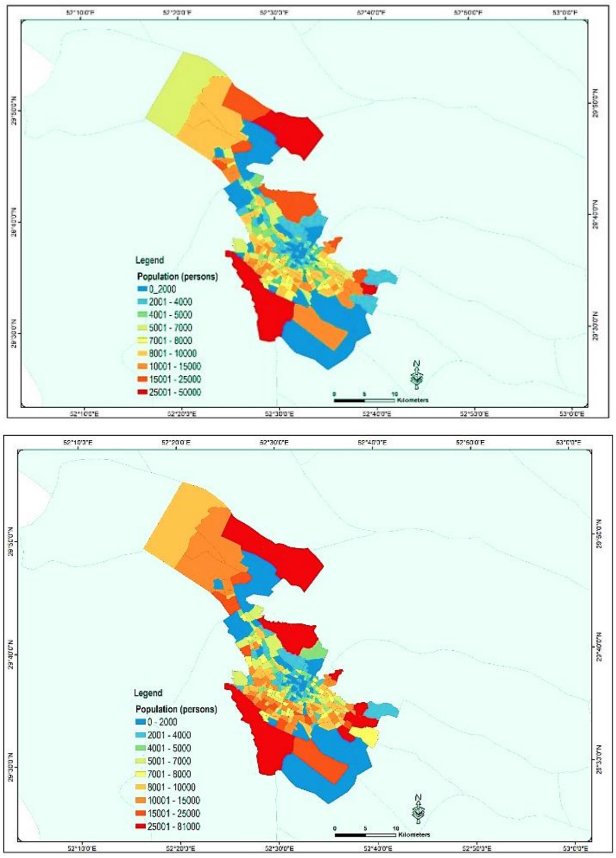

Fig. 3: Existing population and Estimated population

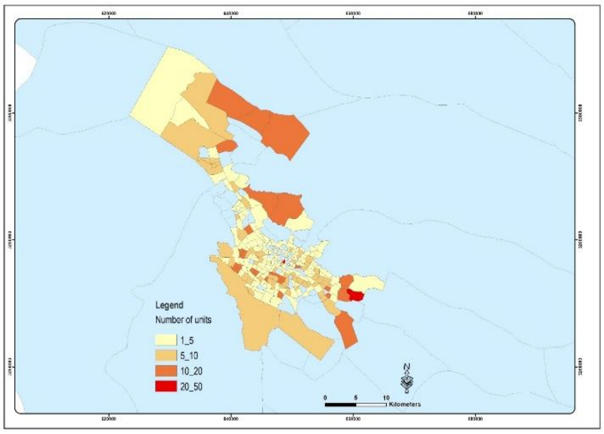

Fig. 4: Number of fruit and vegetable shops

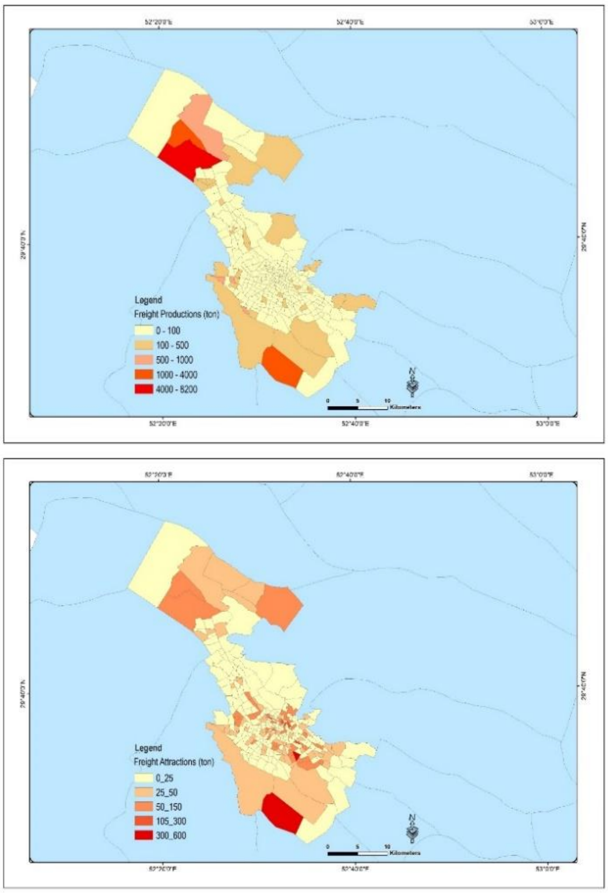

Fig. 5: Freight production and attraction of fruit and vegetable commodities

The model variables included the areas and peripheries of TAZs, the number and areas of warehouses, population, population pyramid, gender, employment, education, vehicle ownership, the number of home-based and non-home-based non-freight trips, the number of non-freight trips for different trip purposes, the number of car parking spaces, land price, location in the central business district, 36 land-use types, and 87 business groups in the 325 TAZs of Shiraz. As mentioned, the CMP of Shiraz divides the city into 325 TAZs, and the present study exploited the aggregated data of the TAZs. Fig. 6 depicts the TAZs.

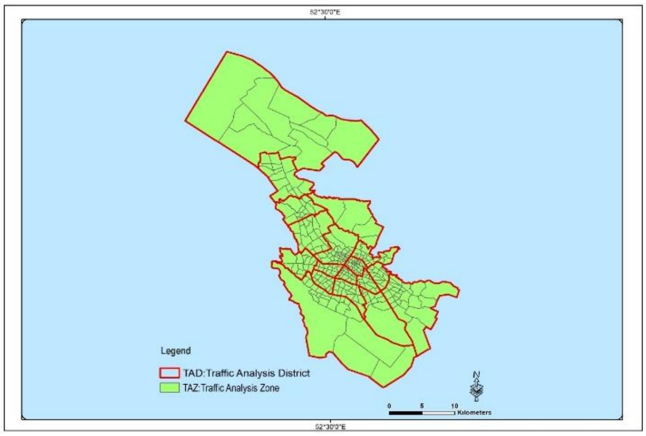

Fig. 6: Traffic analysis district and traffic analysis zone

This study developed freight production and attraction equations in the TAZs of Shiraz based on multivariate regression modeling using different databases. OLS methods were employed to process the model in RStudio.

First, the variables with reasonable causal relationships with the production and attraction of the freight types or discussed in previous studies were reviewed. To understand the extracted variables and their relationships with the dependent variables (estimating a suitable functional form), the scatter plot of the data was employed. Then, based on the functional form, a bivariate regression model of each variable and the dependent variable was developed. The insignificant variables at a confidence level of 95% would be excluded from modeling to avoid biased parameters and increased variance. Dummy variables were employed in the model to use descriptive and nominal data (e.g., TAZ location in the central business district) in modeling.

Once the significant independent variables had been identified for each dependent variable (freight production and freight attraction), linear multiple regression modeling was carried out. A large number of variables were tested in the construction of each model. The variables that were insignificant at a confidence level of 95% or had small effects on the dependent variable were excluded. Furthermore, independent variables that had good collinearity with the dependent variable but their combination with each other or with other variables led to insufficient fitness due to conceptual similarity or collinearity or led to insignificant were also excluded. Intensive collinearity reduces estimation accuracy. The correlation matrix was constructed and investigated to ensure that intensive collinearity would not occur. In each model, variables with correlations below 0.3 were employed.

Once the candidate variables and functional forms had been selected, modeling was carried out using different combinations of the candidate variables. Several qualified models were processed to construct each freight production and attraction model. However, the best model of each group is reported. Concerning the functional forms, it should be noted that functional forms of higher complexity were found to have higher fitness, but they were not employed since the authors aimed to provide a simple, applicable, and comprehensible representation of the effect of each independent variable on the dependent variable. The present study employed linear, polynomial, logarithmic, and exponential functions to obtain the highest fitness. In the case of similar fitness values under different functional forms, the simpler function would be selected. However, the function of higher complexity was utilized whenever the simpler function led to almost the same fitness as the complex one but could not meet the third assumption.

4. Results and discussion

To realize the effects of influential variables on the production and attraction of industrial, agricultural and livestock, fruit and vegetable, and food freight using aggregated data of TAZs in Shiraz, mean squared error (MSE) minimization was used. To fit the model by minimizing MSE, six assumptions were evaluated, including (1) a linear model in terms of parameters, (2) randomness of samples, (3) a zero mean of residuals, (4) non-collinear independent variables, (5) homoscedasticity, and (6) the normality of error terms. It should be noted that the assumptions of the regression models using MSE minimization were evaluated before and after model processing. A number of assumptions could be evaluated before processing the model, while some others should be considered to be the case to process the model and validate the assumptions. The models that are found not to meet the assumptions during the process are excluded.

The first assumption is model linearity in parameters. The models were found to be linear in parameters, meeting the first assumption. As mentioned, the data were obtained from the transportation CMP of Shiraz, and sample randomness measures were implemented. Therefore, random sampling was performed, meeting the second assumption. The other assumptions were evaluated after model processing.

Concerning the third assumption, it should be mentioned that a zero conditional mean may be violated due to, for example, the non-incorporation of one or more necessary variables, the addition of irrelevant variables, the selection of an unsuitable functional form, variable measurement errors, and wrong statistical assumptions of the disturbance term (e.g., additive rather ran multiplicative). This study evaluated the suitable functional form. The present work tested the functional forms of the models using the Ramsey regression equation specification error test (RESET), as shown in Table 3. The third assumption is met in all the reported models.

The fourth assumption refers to the lack of complete collinearity between the independent variables. Many indices are available to detect multicollinearity. The present study adopted auxiliary regression to investigate the collinearity of a variable with a combination of other variables. The correlation coefficients of the auxiliary regressions were found to be small. Thus, it can be said that the models had no complete collinearity and met the fourth assumption.

The fifth assumption was homoscedasticity. Homoscedasticity was examined by subjective and objective tests. The graphical method is a common subjective test. It plots the squared estimated residuals versus the dependent variable. To meet homoscedasticity, the variation of the dependent variable is required to have no structured association with the squared estimated residuals. This was the case with the models. For example, Fig. 7 represents the graphical plot of the food freight production model.

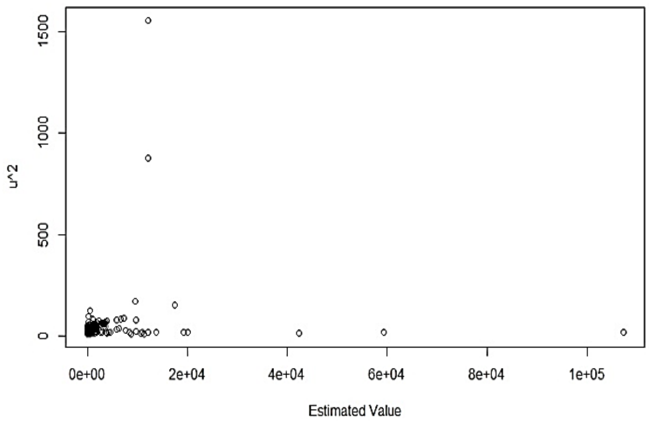

Fig. 7: The diagram of the square estimated residuals versus the estimated dependent variable

However, homoscedasticity cannot be firmly ensured using the graphical method. Thus, the Breusch-Pagan (BP) and White tests were utilized. The BP and LM values were compared to the critical value (77.93), and the null hypothesis (homoscedasticity) was not rejected. Therefore, it can be said that the fifth is met. However, even if this assumption was violated, the estimators would be still unbiased but would no longer be the best linear unbiased estimator (BLUE) and have the lowest variance.

Many tests could be used to investigate the sixth assumption, i.e., normality. As with the fifth assumption, the sixth assumption was examined in subjective and objective phases. The histogram of the residual is a subjective test. It can be compared to the histogram of the normal distribution. The histogram of the residual was plotted for all the models. It was observed that most histograms had slight differences from the normal distribution. The Q-Q plot is another subjective method. The distribution of residuals can be claimed to be normal when the residuals are situated around the bisector of the first quadrant. Both plots were developed for the eight models. Fig. 8 and Fig. 9 show the plots of the food freight production model. As can be seen, the two plot methods did not reject the assumption of a normal distribution of residuals for the models.

Moreover, the Shapiro-Wilk and Jarque-Bera tests were employed to evaluate the sixth assumption. The Shapiro-Wilk test has the highest detection power and is based on a regression relation or correlation analysis between explanatory variables and their expected values [55]. Also, the Shapiro-Wilk test is very sensitive to outliers. The null hypothesis of the Shapiro-Wilk test is normal population distribution. It is often used when data points are fewer than 5,000. The proposed model involved 325 observations. The p-value was calculated, revealing that the error terms had a normal distribution. Concerning the Jarque-Bera test, it should be noted that the null hypothesis is a normal distribution of the disturbance term, and disturbance term normality is proved when the null hypothesis is demonstrated not to be rejected.

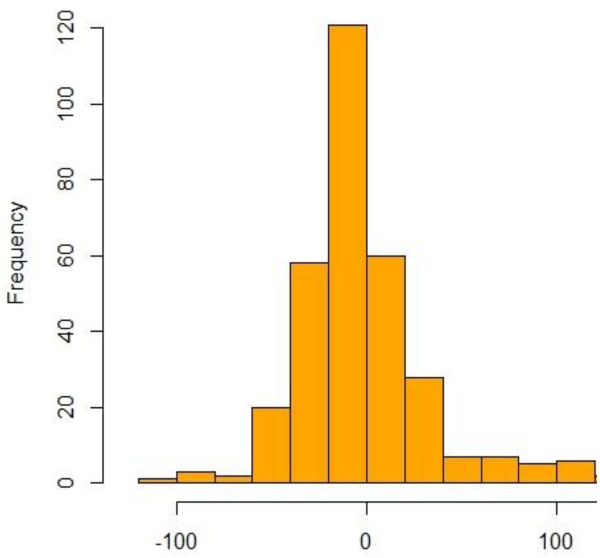

Fig. 8: Histogram of residual of food freight production model

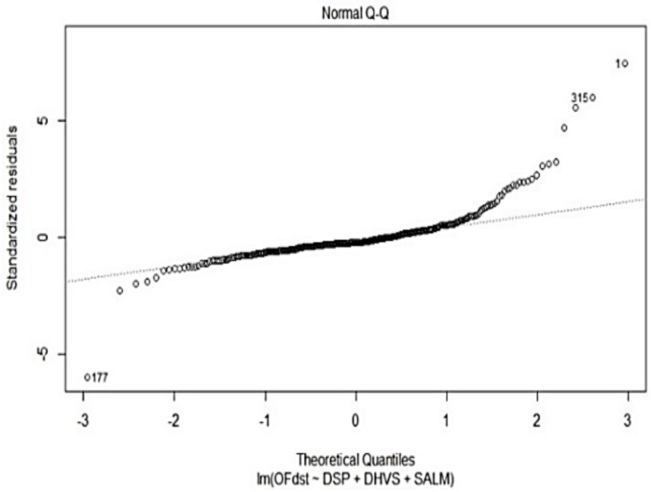

Fig. 9: Q-Q plot for food freight production model

The Jarque-Bera function has a chi-squared distribution with two degrees of freedom [41]. The critical Jarque-Bera value is 5.99, based on the number of observations in the processed models. The null hypothesis (i.e., a normal disturbance term distribution) is rejected if the Jarque-Bera value is larger than 5.99 [41]. As a result, a normal distribution of the disturbance term requires the Jarque-Bera value not to be significant. Based on the calculated chi-squared, the null hypothesis is not rejected, and it is inferred that the distribution of residuals is normal. This is consistent with the previous tests. The normal distribution of residuals (i.e., the sixth assumption) is not solid, and it is recommended that the maximum likelihood estimation (MLE) method be adopted for models with a quantitative normal distribution. Table 1 represents the assumption testing results. As can be seen, all six assumptions were found to be met in the models.

Table 1: The assumption testing results

Model	Linear in Parameters	Random Sampling	Zero Conditional Mean (Ramsey's RESET Test)	No Perfect Multi Collinearity (Axillary Regressions)	Homoscedasticity		Normality	
					Breusch -Pagan	White	Shapiro-Wilk	Jarque -Bera
Oindus	O.K.	O.K.	0.0	O.K.	2.7	1.4	0.02	1.6
Dindus	O.K.	O.K.	3.0	O.K.	15.2	10.1	0.04	5.6
Olivagr	O.K.	O.K.	8.8	O.K.	67.7	5.0	0.03	5.1
Dlivagr	O.K.	O.K.	2.0	O.K.	2.7	60.3	0.04	5.9
OFru	O.K.	O.K.	0.1	O.K.	0.0	0.0	0.03	5.1
DFru	O.K.	O.K.	0.1	O.K.	0.0	0.1	0.03	4.5
OFdst	O.K.	O.K.	6.0	O.K.	17.4	5.2	0.02	2.9
DFdst	O.K.	O.K.	4.8	O.K.	4.3	0.1	0.02	2.9

Table 2 shows the processed models of the production and attraction of industrial, agricultural and livestock, fruit and vegetable, and food freight. The numbers within parentheses represent the corresponding Student's t-value. The adjusted index of fit, F-value, and root-mean-square error (RMSE) are also reported.

Table 2: The estimated models

Model	Production	R^2_{adj}	F & RMSE	Attraction	R^2_{adj}	F & RMSE
indus	$9.8 + 2.1\sqrt{\text{Countgld}} + 127.2\text{TP}_{31}^2 + 791.8DIT$	0.55	F= 133.2	$30.2 + 1059.2TP_{23}^2 + 1.7\sqrt{Tol}$	0.71	F= 406.9
	(1.8) (3.9) (3.0) (19.1)		RMSE=41.0	(12.2) (28.1) (4.6)		RMSE=37.6
livagr	$1.6 + 0.09DHVS + 5.4NOHCS + 0.006TP_{37} + 2.64Tox$	0.64	F= 140.7	$4.6 + 6.0NOHCS + 0.29Merc + 109.5Sl + 5.2e - 17\,exp(NDHVS)$	0.59	F= 117.7
	(2.1) (14.9) (11.8) (4.5) (10.5)		RMSE=24.9	(7.1) (14.5) (8.9) (9.6) (9.8)		RMSE=11.4
Fru	v$8.2 + 0.1Fru$	0.98	F= 16000	$9.9 + 0.1Fru$	0.98	F= 72000
	(5.2) (126.8)		RMSE=28.1	(12.6) (268.3)		RMSE=14.2
Fdst	$8.3+0.34DHVS+0.018SALM +848.3DSP$	0.84	F= 547.7	$28.1+274Whsal^3+0.18NDET +547.0DSP$	0.60	F= 165.3
	(2.1) (13.0) (2.7) (21.2)		RMSE=127.8	(12.4) (7.4) (2.8) (20.8)		RMSE=36.8

Table 3: Defines the variables used in the freight production and attraction models.

variable	Variable definition	variable	Variable definition	variable	Variable definition
indus	industrial	livagr	agricultural and livestock	Fru	fruit and vegetable
countgld	the total number of businesses	DHVS	home-based van taxies in the destination	Fruv	the number of fruit and vegetable vendors
$\boldsymbol{TP_{31}}$	the area of combined residential, business, and workshop land-use	NOHCS	the number of non-home-based taxies in the origin,	Fdst	food freight
DIT	the district of industrial town-Dummy variable (TAZ 315=1 , etc=0)	$\boldsymbol{TP_{37}}$	the origin, river area	SALM	the number of in-site vendors and workmen
$\boldsymbol{TP_{23}}$	the area of industrial land-use	Tox	the number of fertilizer stores	DSP	dummy variable DSP (Shiraz Industrial Town and Merchants Complex)- (TAZ 117,315=1 , etc=0)
Tol	the number of ironmongery and tools stores	Merc	the number of slaughterhouses	Whsal	The number of food distributors and wholesalers
NDHVS	and the number of non-home-based van taxies in the origin	SL	the number of grain depots	NDET	the number of non-home-based trips for shopping

As can be seen, the total number of businesses, the area of combined residential, business, and workshop land-use, and the district of industrial town (DIT) were found to affect the production of industrial freight, and the adjusted-R^2 was calculated to be 0.55. However, different independent variables influence the attraction of industrial freight, including the area of industrial land-use and the number of ironmongery and tools stores. The coefficient of determination was estimated to be 0.71 for the industrial freight attraction model. Considering the t-values, the industrial freight production and attraction models are both significant at a confidence level of 99%.

It was found that the number of home-based van taxies in the destination, the number of non-home-based taxies in the origin, river area, and the number of fertilizer stores explained 64% of the average variation of agricultural and livestock freight production. Moreover, the number of non-home-based taxies in the travel origin, the number of slaughterhouses, the number of grain depots, and the number of non-home-based van taxies in the origin explained 59% of the average variation in agricultural and livestock freight attraction. The variables relating to the number of ride services are a proxy variable of the number of agricultural and livestock product workers (e.g., farmers).

The fruit and vegetable freight production and attraction models shared the independent variable of "the number of fruit and vegetable vendors," and both models had a fit index of 0.98. The fit indices of these two models are larger than those of the other six models partially as the distribution and sales centers of fruit and vegetables are more specialized than those of the other types of freight. In other words, fruit and vegetables are distributed in fruit and vegetable markets and sold to the end consumers in stores. The distribution and sales centers of industrial freight are very diverse.

It was found that food freight production was dependent on the number of home-based van taxies in the destination, the number of in-site vendors and workmen, and dummy variable DSP (Shiraz Industrial Town and Merchants Complex), with an adjusted R^2 of 0.84. The number of food distributors and wholesalers, the number of non-home-based trips for shopping, and DSP were observed to influence food freight attraction, and the model was calculated to have an adjusted R^2 of 0.60. The coefficient of dummy variable DSP indicates that the two TAZs involving the Shiraz Industrial Town and Merchants Complex had much larger contributions than the other zones to food freight production and attraction due to the concentration of numerous food production and packing centers in the Shiraz Industrial Town and large food stores in the Merchants Complex. To explain the larger fit index of food freight production than that of food freight attraction, it can be said that food freight is produced in wholesale centers, which are mostly found in a few zones of Shiraz, while food freight is attracted to many more small retail stores across Shiraz

All eight models had small constants compared to the mean of the dependent variables. The f-values and p-values of all the models are significant (at a confidence level of 99%). Also, all the models were calculated to have small RMSE values, suggesting the satisfactory fit of the models.

Furthermore, the models can be claimed to have satisfactory adjusted coefficients of determination since freight production and attraction models have small coefficients of determination due to the nature, dynamism, and diversity of transported commodities in urban areas (Patil et al., 2021([1]. However, case study results may differ in different cities, depending on data availability.

5. Conclusion

The present study exploited data from the transportation CMP of Shiraz to construct eight linear multivariate regression models. The factors impacting the production and attraction of industrial, agricultural and livestock, fruit and vegetable, and food freight were explored using aggregated data of TAZs using multiple regression in Shiraz. The proposed eight models introduced new factors impacting freight production and attraction.

Previous studies developed models with only one or two independent variables or small fit indices. The present study developed models based on 325 observations (many more than earlier works). In general, the present work employed many more measured quantitative explanatory variables to tackle the shortcomings of previous works that relied on descriptive datasets, few data points, or extrapolating data from complex quantitative models.

Since freight transportation is dependent on transportation costs, the use of location-representing variables was expected to improve FG modeling. For example, a factor in the TAZ containing the Shiraz Industrial Town would have different behavior in freight production and attraction due to higher customer access and business promotion than an institution outside this zone. Therefore, the present study used dummy variables of the Shiraz Industrial Town and Merchants Complex zones.

Moreover, the proposed models utilized aggregated data of TAZs, unlike most previous studies. This might have been due to the unavailability and monitoring of spatial data in developing countries. The present study exploited data of over 300 variables in 325 TAZs.

Considering a lack of consensus on the functional form of explanatory variables and their effects on freight production and attraction, the present study examined many functional forms of each variable to identify the best functional form of each variable based on the available data, association with the other variables, OLS assumptions, and model fitness.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

[1] Pani, A., Sahu, P. K., Patil, G. R., & Sarkar, A. K. (2018). Modelling Urban Freight Generation: A Case Study of Seven Cities in Kerala, India. Transport Policy, 69, 49-64.

[2] Venkadavarahan, M., & Marisamynathan, S. (2022). Development of urban freight trip generation models concerning establishment classification process for a developing country. International Journal of Transportation Science and Technology, 11(3), 603-618.

[3] Pani, A., Sahu, P. K., & Bhat, F. A. (2021). Assessing the Spatial Transferability of Freight (Trip) Generation Models Across and Within States of India: Empirical Evidence and Implications for Benefit Transfer. Networks and Spatial Economics, 21(2), 465-493.

[4] Puente-Mejia, B., Palacios-Argüello, L., Suárez-Núñez, C., & Gonzalez-Feliu, J. (2020). Freight Trip Generation Modeling and Data Collection Processes in Latin American cities. Modeling framework for Quito and Generalization Issues. Transportation Research Part A: Policy and Practice, 132, 226-241.

[5] The National Academies Press. (2012). Transportation Research Board and National Academies of Sciences, Engineering, and Medicine, Freight Trip Generation and Land Use (No. NCHRP Project 08-80, Project NCFRP-25).

[6] Dhonde, B., & Patel, C. R. (2020). Characterization of Freight Trip from Textile Powerloom Units–A Case Study of Surat, India. Transportation Research Procedia, 48, 428-438.

[7] Sahu, P.K., & Pani, A. (2020). Freight Generation and Geographical Effects: Modelling Freight Needs of Establishments in Developing Economies and Analyzing Their Geographical Disparities. Transportation, 47(6), 2873–2902.

[8] Souleyrette, R., Maze, T. H., Strauss, T., Preissig, D., & Smadi, A. G. (1998). Freight Planning Typology. Transportation Research Record, 1613(1), 12-19.

[9] Ottemöller, O., & Friedrich, H. (2016). Opportunities of Sectoral Freight Transport Demand Modelling. Case Studies on Transport Policy, 4(1), 9-12..

[10] Lawson, C. T., Holguín-Veras, J., Sánchez-Díaz, I., Jaller, M., Campbell, S., & Powers, E. L. (2012). Estimated Generation of Freight Trips Based on Land Use. Transportation Research Record, 2269(1), 65-72.

[11] Ha, D. H., & Combes, F. (2015, July). Building a Model of Freight Generation with a Commodity Flow Survey. In Commercial Transport: The 2nd Interdiciplinary Conference on Production Logistics and Traffic 2015.

[12] Pani, A., Sahu, P. K., Chandra, A., & Sarkar, A. K. (2019). Assessing the Extent of Modifiable Areal Unit Problem in Modelling Freight (trip) Generation: Relationship Between Zone Design and Model Estimation Results. Journal of Transport Geography, 80, e102524.

[13] Holguín-Veras, J., Jaller, M., Destro, L., Ban, X., Lawson, C., & Levinson, H. S. (2011). Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates. Transportation Research Record, 2224(1), 68-81.

[14] Madar, G., Maoh, H., & Gingerich, K. (2021). Modeling Commercial Vehicle Trip Generation at The Business-Establishment Level. Canadian Journal of Civil Engineering, 48(6), 669-680.

[15] Takyi, I.K. (2021). TRIP GENERATION ANALYSIS IN A DEVELOPING COUNTRY CONTEXT. Transp. Res. Rec., (1285).

[16] Holguín-Veras, J., & Thorson, E. (2000). Trip Length Distributions in Commodity-Based and Trip-Based Freight Demand Modeling: Investigation of Relationships," Transp. Res. Rec. J. Transp. Res. Board, 1707(1),37–48.

[17] Sánchez-Díaz, I., Holguín-Veras, J., & Wang, X. (2016). An Exploratory Analysis of Spatial Effects on Freight Trip Attraction. Transportation, 43, 177-196.

[18] Pani, A., & Sahu, P.K. (2019). Comparative Assessment of Industrial Classification Systems for Modeling Freight Production and Freight Trip Production. Transp. Res. Rec. J. Transp. Res. Board, 2673(3), 210–224.

[19] Dhulipala, S., & Patil, G. R. (2020). Freight Production of Agricultural Commodities in India Using Multiple Linear Regression and Generalized Additive Modelling. Transport Policy, 97, 245-258.

[20] Sánchez-Díaz, I. (2018). Potential of Implementing Urban Freight Strategies in The Accommodation and Food Services Sector. Transportation Research Record, 2672(9), 194-203.

[21] Gonzalez-Feliu, J., & Peris-Pla, C. (2017). Impacts of Retailing Attractiveness on Freight and Shopping Trip Attraction Rates. Research in Transportation Business & Management, 24, 49-58.

[22] Cheah, L., Mepparambath, R. M., & Surribas, G. M. R. (2021). Freight Trips Generated at Retail Malls in Dense Urban Areas. Transportation Research Part A: Policy and Practice, 145, 118-131.

[23] Alho, A.R., & de Abreu e Silva, J. (2017). Modeling Retail Establishments' Freight Trip Generation: A Comparison of Methodologies to Predict Total Weekly Deliveries. Transportation, 44(5), 1195–1212.

[24] Rashidi, T. H., & Roorda, M. J. (2018). A Business Establishment Fleet Ownership and Composition Model. Transportation, 45, 971-987.

[25] De Oliveira, L.K., De Albuquerque Nóbrega, R.A., Ebias, D.G., & Souza Corrêa, B.G. (2017). Analysis of Freight Trip Generation Model for Food and Beverage in Belo Horizonte (Brazil). REGION, 4(1), 17–30.

[26] Jaller, M., Sanchez-Diaz, I., Holguín-Veras, J., & Lawson, C.T. (2021). Area Based Freight Trip Generation Models.

[27] Jaller, M., Sánchez-Díaz, I., & Holguín-Veras, J. (2015). Identifying Freight Intermediaries: Implications for Modeling of Freight Trip Generation. Transportation Research Record, 2478(1), 48-56.

[28] Asuncion, J. S. (2014). The Geographic Adaptive Potential of Freight Transportation and Production System in The Context of Fuel and Emission Constraints.

[29] Pani, A., Bhat, F. A., & Sahu, P. K. (2020). Effects of Business Age and Size on Freight Demand: Decomposition Analysis of Indian Establishments. Transportation Research Record, 2674(2), 112-126.

[30] Kulpa, T. (2014). Freight Truck Trip Generation Modelling at Regional Level. Procedia-Social and Behavioral Sciences, 111, 197-202.

[31] Gonzalez-Feliu, J., Palacios-Argüello, L., & Suarez-Nanez, C. (2020). Links Between Freight Trip Generation Rates, Accessibility and Socio-Demographic Variables in Urban Zones. Archives of Transport, 53.

[32] Holguín-Veras, J., Sánchez-Díaz, I., Lawson, C. T., Jaller, M., Campbell, S., Levinson, H. S., & Shin, H. S. (2013). Transferability of Freight Trip Generation Models. Transportation Research Record, 2379(1), 1-8.

[33] Allen, J., & Farber, S. (2019). Sizing Up Transport Poverty: A National Scale Accounting of Low-Income Households Suffering from Inaccessibility in Canada, and What to Do About it. Transport policy, 74, 214-223.

[34] Gonzalez-Feliu, J., & Sánchez-Díaz, I. (2019). The Influence of Aggregation Level and Category Construction on Estimation Quality for Freight Trip Generation Models. Transportation Research Part E: Logistics and Transportation Review, 121, 134-148.

[35] De Bakshi, N., Tiwari, G., & Bolia, N. B. (2020). Influence of Urban Form on Urban Freight Trip Generation. Case Studies on Transport Policy, 8(1), 229-235.

[36] Boarnet, M. G., Hong, A., & Santiago-Bartolomei, R. (2017). Urban Spatial Structure, Employment Subcenters, and Freight Travel. Journal of Transport Geography, 60, 267-276.

[37] Beagan, D.F., Fischer, M.J., & Kuppam, A.R. (2021). Quick Response Freight Manual II. e FHWA-HOP-08-010.

[38] Yang, Y. (2015). Development of The Regional Freight Transportation Demand Prediction Models Based on The Regression Analysis Methods. Neurocomputing, 158, 42-47.

[39] Vagadia, D.H. (2018). Review of Freight Demand Models. Res. Appl. Sci. Eng. Technol.,6(6), 1706–1712.

[40] Zaveri, P., & Sood, S. K. Analysis and Generation of Multiple Linear Regression for Residential Area of Vesu, Surat. 5(11) ,7.

[41] Gujarati, D.N,. & Porter, D.C. (2009). Basic Econometrics. Boston: McGraw-Hill Irwin.

[42] Council, B. M. (2002). Development of Truck Models.

[43] List, G. F., Konieczny, L. A., Durnford, C. L., & Papayanoulis, V. (2002). Best-Practice Truck-Flow Estimation Model for The New York City Region. Transportation Research Record, 1790(1), 97-103.

[44] Holguín-Veras, J. (2021). Using Commodity Flow Survey Microdata and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook. NCFRP Res. (No. 37, Art. no. Project NCFRP-25(01)).

[45] De Jong, G. (2014). Freight Service Valuation and Elasticities. In Modelling Freight Transport, 201-227.

[46] Rodríguez-Pose, A., Tselios, V., Winkler, D., & Farole, T. (2013). Geography and The Determinants of Firm Exports in Indonesia. World Development, 44, 225-240.

[47] Sikder, S., Pinjari, A. R., Srinivasan, S., & Nowrouzian, R. (2013). Spatial Transferability of Travel Forecasting Models: a Review and Synthesis. International Journal of Advances in Engineering Sciences and Applied Mathematics, 5, 104-128.

[48] Patil, G. R., Thadoju, S., Sahu, P. K., Mondal, A., & Bajpai, V. (2021). Data Collection and Modeling of Restaurants' Freight Trip Generation for Indian Cities. Transportation in Developing Economies, 7, 1-11.

[49] Middela, M. S., & Ramadurai, G. (2021). Incorporating Spatial Interactions in Zero-Inflated Negative Binomial Models for Freight Trip Generation. Transportation, 48(5), 2335-2356.

[50] Venkadavarahan, M,. & Marisamynathan, S. (2021). Exploring Spatial Interaction Effects in Freight Trip Generation Model for Intermediate and Pure Receiver Establishment. Transportation Policy.

[51] Krisztin, T. (2018). Semi-Parametric Spatial Autoregressive Models in Freight Generation Modeling. Transportation Research Part E: Logistics and Transportation Review, 114, 121-143.

[52] Al-Deek, H. M. (2001). Comparison of Two Approaches for Modeling Freight Movement at Seaports. Journal of Computing in Civil Engineering, 15(4), 284-291.

[53] Günay, G., Ergün, G., & Gökaşar, I. (2016). Conditional Freight Trip Generation Modelling. Journal of Transport Geography, 54, 102-111.

[54] Patil, G.R,. & Sahu, P.K. (2016). Estimation of Freight Demand at Mumbai Port Using Regression and Time Series Models. KSCE J. Civ. Eng., 20(5), 2022–2032.

[55] Çelebi, D. (2021). Chapter 1 - Issues in Freight Transport Modeling for Emerging Economies: Insights from Turkey.in Freight Transport Modeling in Emerging Countries.Elsevier,1–16.