Are Freight Production Models Transferable between Urban and Suburban Areas? Guiding Model Transfer in Geographically Sprawling Indian Cities

Bhavani Shankar Balla¹ · Prasanta K. Sahu¹ · Agnivesh Pani²

Received: 8 March 2021 / Accepted: 5 July 2021 / Published online: 15 July 2021
© The Institution of Engineers (India) 2021

Abstract Investigating the spatial transferability of freight generation (FG) models is an imperative research need to enable the usage of formerly estimated model parameters in new application contexts with or without the usage of local data. By understanding how to transfer models (and to what extent), planning agencies in large countries like India can save freight survey costs in regions where they lack the institutional capacity and resources. Due to geographically sprawling nature of most of the Indian cities, an important research question regarding transferability is whether the models developed for urban areas can provide accurate estimates of freight activity in the suburban areas or vice versa. This paper aims to provide two solutions to this problem: (i) compare the relative effectiveness of transferability depending up on the direction of transfer and (ii) assess which models can be transferred and which cannot. Data collected from seven cities are used for this study. A set of freight production (FP) models are developed using this data to understand the differential influence of geographical location and industry segment on the model coefficients. The estimated FP models show that suburban establishments exhibit significantly higher FP

rates as compared to other establishments. Subsequently, transferability direction and accuracy are determined using standard metrics such as transfer R^2 , relative aggregate transfer error and transfer index. The transferability findings will provide actionable insights into development of FP models in regions with data constraints, which is of great value in an era of declining budgets for travel surveys.

Keywords Spatial transferability · Model updating · Freight production · Regression · Establishment-based freight survey · Transfer R^2

Abbreviations

BTKM	Billion ton kilometer
CFS	Commodity flow survey
EBFS	Establishment-based freight survey
FG	Freight generation
FM	Freight movement
FP	Freight production
FTG	Freight trip generation
GDP	Gross domestic product
GFA	Gross floor area
GVA	Gross value added
ISIC	International Standard Industrial Classification
MPOs	Metropolitan planning organizations
NE	Number of employees
OLS	Ordinary least squares
RATE	Relative aggregate transfer error
RMSE	Root mean square error
TI	Transfer index
TR ²	Transfer R^2
WRMSE	Weighted root mean square error

✉ Prasanta K. Sahu
prasantsahu222@gmail.com

Bhavani Shankar Balla
ballabhavanishankar@gmail.com

Agnivesh Pani
agniveshpani@gmail.com

¹ Department of Civil Engineering, Birla Institute of Technology and Science Pilani, Hyderabad, Telangana 500078, India

² Department of Civil Engineering, The University of Memphis, Memphis, TN, USA

Introduction

Large-scale establishment-based freight data are crucial for forecasting freight demand. However, developing countries like India lack an established commodity flow survey (CFS) practice despite the presence of multiple megacities and pressing freight transport planning needs of logistics-driven economies. Freight demand forecasting models are thus of national and regional importance to plan for facilities that provide consumers with access to a wide variety of goods and services without traveling long distances [1, 2]. In particular, having solid estimates about the freight activity is critically important to understand why, how and where the freight activity takes place and what kind of infrastructure and policies need to be provided to respond effectively to the growing logistical requirements of businesses and households [1, 3]. Nonetheless, the freight model system is data intensive, and the planning agencies allocate a significant amount of resources in terms of cost and time on data collection programs [4, 5, 6].

The spatial transferability of the models could be one way for the metropolitan planning organizations (MPOs) to reduce the commitment of the significant portion of the financial resources for several data collection programs for developing freight modeling systems [7]. Investigations on model transferability, though not new, research is limited to a handful of freight demand studies [7–10]. Some notable studies addressing transferability of passenger trip generation and mode choice models are available [8–11]. Estimated model parameters can be used in an application context using the local dataset of relatively lesser sample size [11], and thereby, MPOs can plan facilities with time and budget constraints. The savings in cost and time are beneficial for a freight demand modeling system as it requires significant data from an establishment-based freight survey (EBFS) or CFS. In the absence of CFS and inaccessibility to freight data sources owing to ownership and privacy in a country like India and several other South-Asian countries, EBFS is a smarter option for obtaining data on freight activities [12]. However, with rigorous surveyor training, lower complete response rate, etc., EBFS becomes time overrun, incurring higher unit cost per data to meet the required sample size to represent the population [12, 13]. These issues highlight the importance of investigating the freight demand model transferability in an era where model estimation and diagnostic studies are evolving [1, 4, 14–16].

In the last two decades, due to rapid urbanization, freight movements in India have increased considerably. The nationwide freight movement (FM) was 2000 BTKM (Billion Ton Kilometer) during 2011–2012 and is estimated to grow at an annual rate of 9.7% till 2031–32,

which will lead FM to surpass 13,000 BTKM. At the same time, the overall passenger traffic is projected to rise by 15% approximately to reach 168,875 BPKM (billion passenger kilometer) in 2031–32 from 10,375 BPKM in 2011–12 [17]. The increase in truck activities more often conflicts to the commuter traffic, subsequently resulting to higher travel time or congestion experience in the urban road network. The conflicting phenomenon causes longer queues, network disruptions, environmental damage, poor air quality and degradation of overall quality of life in the urban areas.

As per the census 2011, the urban areas contribute to 31% of Indian population and gross domestic product (GDP) share of the cities is 63% of nation's GDP. Due to rapid urbanization, by 2030, it is predictable that the urban population increases to 40% and contributes to 75% of the national GDP [18]. Also, with the launch of 'Smart City Mission' [19] to improve the livability index, it is expected to notice accelerated economic activities and subsequent increase in truck trips. Also, Government of India has launched the 'Make in India' program to improve the ease of doing business index on a global level. This program influenced the manufacturing sector with a record growth of 7.9% in the gross value added (GVA) during 2016–17; GVA growth was 5.9% in 2017–18 and estimated to be 8.1% in 2018–19 [20]. The government's flexible approach in several policies for small and medium scale enterprises (SMEs) has been attracting investors from the local to national level to develop new industries both in peri-urban and suburban regions as there is a scarcity of land availability within the urban regions. The economic expansion programs initiated by the Government of India will boost the economy, nonetheless, it will also increase the trucking activities both at a firm and network level. Incidences like the ongoing pandemic COVID-19 have already added more truck-kilometers in all segments (first mile, long haul and last mile) of freight travel. The freight movement was reported to be 65.4 billion tons in April 2020 when nationwide lockdown was declared by the Government of India and increased to 110.110 billion tons in November 2020 [21]. The behavioral shift in consumer purchase preferences will keep on adding more truck-kilometers, and eventually, higher trucking will enhance the several negative externalities from trucking such as congestion, noise and air pollution in urban and suburban regions. To mitigate the external effects and augment the existing infrastructure, planning bodies require more representative freight demand models in both urban and suburban contexts. However, concerning the absence of CFS and EBFS in India makes it burdensome to planners and policy makers to develop a model system at the national level. Also, as the model system is data intensive, it is often

difficult to allocate financial resources for data collection program within budgetary constraints.

An imperative study is required to investigate on the spatial transferability of freight generation models from one region to another to understand the degree of transferability between two different regions in a similar or different states. By understanding the extent of transferability, the planning agencies in developing countries like India can reduce the costs for the freight surveys in the regions where they lack institutional capacity and resources. Due to the geographically sprawling nature of the Indian cities, an important research question is about the extent of transferability when the estimated urban model is transferred to suburban region to estimate the freight activity or vice versa. This paper aims to provide two solutions to this problem: (i) compare the relative effectiveness of transferability depending up on the direction of transfer and (ii) assess which models can be transferred and which cannot. We developed models in an estimation context, separately for urban and suburban regions using data of 432 establishments collected through the EBFS in seven cities in Kerala — A southern coastal state in India. A ‘*a posteriori*’ segment classification scheme from the previous research [12] was used to estimate the classified models. The estimated models were assessed for transferability to determine the model transfer direction and accuracy levels. Next section discusses the most relevant literature, and after that, a brief outline of the current research is presented.

Past Research

Estimation Context

The modeling of freight generation is bi-conceptual which includes the concepts of freight generation (FG) and freight trip generation (FTG) [22]. FG focuses on weights, and FTG explains number of truck trips. An overview of previous studies reveals the use of various explanatory variables to explain FG/FTG, such as gross floor area [3, 23], number of employees [3, 11, 15], number of clients [24], sales [25], type of commodity [26]. Beagan et al. [27] developed several FG and FTG models based on employment, area of the building/establishment and type of land use by collecting the freight-related data from the USA. Sánchez-Díaz [3] developed a set of employment-based and area-based FG and FTG models using data collected from establishments in the City of Gothenburg, Sweden. Alho & de Abreu e Silva [16] used data collected from 604 retail establishments in the City of Lisbon, Portugal to develop a set of disaggregate FG models, using different modeling approaches.

In the context of developing economies, Pani et al. [15] developed city-specific disaggregate FG models for estimating FG and validated the explanatory power of years in business in addition to traditional business variables — gross floor area (GFA) and number of employees (NE) to model FG. Sahu & Pani [23] developed two sets of disaggregate FG models — (a) 52 practice-oriented FG models using ordinary least squares (OLS) regression technique, and (b) multiple classification analysis tables and nomograms, and examined the geographical disparities on FG at a regional level. The disaggregate FG models developed in aforementioned studies focus on forecasting FG at a regional level, and such models are useful for planning for inter-city freight transportation. Thus, there is a concerning lack of focus on the extent to which location of an establishment, i.e., whether the establishment is located in an urban or suburban area and acts as a modifier to the relationship between FG and establishment characteristics (as described by business variables), in the context of developing economies like India.

Application Context

In the age of constrained survey assets and ever-developing demand for disaggregated travel information, examining the spatial transferability of models has become a pillar of travel demand analysis [28]. An overview of previous studies [29–36] shows that various measures could be used to assess the transferability of the estimated models. For instance, Atherton & Ben-Akiva [37] used transferability test statistic to assess the transferability of work-trip modal-split model. In order to enhance the transferability, they also used the Bayesian update process. Taking forward their study, the updating techniques like transfer-scale approach and combined transfer estimator were developed to improve the accuracy of prediction of the transferred models [38]. Koppelman & Wilmot [39] used several measures such as transfer index (TI), transfer ρ^2 , root mean square error (RMSE), aggregate prediction statistic and relative aggregate transfer error (RATE) to assess transferability of disaggregate choice models.

The transfer ρ^2 is used to test the goodness of fit of the estimation context model when transferred to the application context [40]. To assess the transferability of ordered response model, Agyemang-Duah & Hall [10] employed weighted root mean square error (WRMSE), transfer pseudo R^2 , aggregate prediction statistic and root of sum of residual error. They further updated the model parameters of the estimated model by using scaling updating technique. McArthur et al. [41] used relative number of wrong predictions and standardized root mean square error to analyze the transferability of parameters in gravity models.

Transfer R^2 is used to assess the extent of transferability of linear regression models by Wilmot [42]. The key results obtained from some of the past studies are summarized in Table 1. From these studies, it can be interpreted that the sample size has a significant impact on model transferability. Models with small sample sizes have low TI values.

In the US, for different locales, transferability appraisal was done, and it was discovered that the reach out of transferability is more between the regions inside the state than over the states [9]. The spatial analysis of shopping trip generation in Toronto metropolitan area showed that a directly transferred ordered response model performed well in predicting the application context [10]. Some key past studies [8, 47, 48] suggested that models can be transferable between regions with contextual similarities in terms of socio-demographic factors such as household income, auto-ownership, household structure, key employment type/industries, employment pattern, commuting pattern and degree of urbanization [7, 49]. Institute of Transportation Engineers manual [50] uses the suburban passenger trip data for estimation context and transfer the trip generation model as an application context for the prediction of trips in the urban area. In one of the very few studies, the freight trip generation models were assessed for transferability using RMSE as a statistical measure [11]. The authors also used synthetic correction technique to improve the transferability. However, previous research efforts on FG models have neither estimated nor assessed the transferability by exploring the urban–suburban divide in FG by commercial establishments.

Current Research

The present research first develops separate sets of disaggregate freight production (FP) models for urban and suburban establishments and extends the evolving body of the literature on freight demand modeling for developing economies. These models are estimated by OLS regression technique using business size variables — Number of employees (NE) and gross floor area (GFA) — as

predictors. Model estimation and transferability assessment were done for two different cases: (i) for all urban and suburban establishments, (ii) ‘a posteriori’ segmentation [12] of urban and suburban establishments. Single variable models were estimated without intercept [23] for the above two cases. Initially, the FP models’ transferability assessment was carried out using the naïve transfer approach, and later for the FP models updated using combined transfer estimation. The metrics used for the transferability assessment of naïve and updated models were (i) transfer R^2 (TR^2), (ii) transfer index (TI), (iii) weighted root mean square error (WRMSE) and (iv) relative aggregate transfer error (RATE).

Data Description

A total of 54,170 establishments exist across the seven cities, namely Cochin, Calicut, Malappuram, Kannur, Palakkad, Thrissur and Kottayam, based on the economic census [51]. For EBFS, the minimum required sample size [52] for the study area is found to be 382 on the assumption of confidence level of 95%, response distribution of 50% and margin of error of 5%. The EBFS survey using face-to-face interview was conducted on shippers (manufacturing units, assembling companies and raw material production sites) in Kerala, and a total of 432 completed responses from several industries were finally obtained consisting of information on weekly FP, NE and GFA of an establishment. The reliability of the data is cross-checked on item basis using publicly accessible information from various sources such as Commercial Tax Departments, and local search websites. Measurement tools in Google Earth are used for verifying the gross floor area of the establishment. NE and GFA were considered to model ‘Weekly FP.’ NE includes the total number of skilled and unskilled employees. GFA is the total covered area inside the building envelope of the establishment. The industrial sectors included in the survey were classified into multiple classes defined by the International Standard Industrial Classification (ISIC) system. We further used a posteriori

Table 1 Some key results from the literature

Reference	Model	Study area	Results
Wilmot [42]	Trip generation models	Different cities of South Africa	$TR^2 = -0.44$ to 0.80
Santoso and Tsunokawa [43]	Mode choice models	Urban and suburban areas of Ho Chi Minh City, Vietnam	TI = 0.45 to 0.99
Karasmaa [44]	Mode choice models	Helsinki Metropolitan Area and Turku region in Finland	TI = 0.80 to 0.95
Santoso and Tsunokawa [45]	Mode choice models	Ho Chi Minh city of Vietnam and Phnom Penh city of Cambodia	TI = 0.01 to 0.99
Sikder et al. [46]	Time-of-day choice models	Different counties in San Francisco Bay Area of California	TI = -0.21 to 0.94

segmentation scheme to classify the industries into three homogenous ‘a posteriori’ segments [12]. The details of ‘a posteriori’ segments are as follows: (i) Segment 1 (low value density industries) — ISIC 16: Wood, wood products, furniture and fixtures; ISIC 24–25: Basic metal, alloy, metal products; (ii) Segment 2 (medium value density industries) — ISIC 11: Beverages; ISIC 13: Textile mills; ISIC 17–18: Paper, paper products and printing; ISIC 22: Plastic and rubber products; ISIC 29–30: Transportation equipment; (iii) Segment 3 (high value density industries) — ISIC 10: Food products; ISIC 14: Wearing apparel; ISIC 20–21: Basic chemicals, chemical products and pharmaceuticals; ISIC 23: Non-metallic mineral products; ISIC 26–28: Machinery and equipment; ISIC 32: Other manufacturing industries.

Urban and Suburban Establishment Categorization

The establishments are classified as urban and suburban establishments based on their geographical location. In this study, the urban area is defined based on the definition given by Government of India while carrying out the census study 2011. According to the Census of India 2011, the urban areas are the places governed by a municipal corporation, cantonment board or notified town area committee. Furthermore, places with a minimum population of 5000, a population density more than 400 persons per km², and more than 75 percent of the men engaged in non-agricultural activities are regarded as urban areas [53]. The establishments in the urban area are grouped as urban establishments, and the remaining are grouped as suburban establishments. Out of 432 establishments, urban and suburban establishments are 327 and 105, respectively.

Preliminary Analysis

The descriptive statistics of FP and the establishment characteristics (GFA and NE) for both urban and suburban areas are presented in Table 2. Mean FP values of the establishments vary among the urban and suburban regions. It can be seen that the mean FP of the suburban establishments is more than that of the urban establishments. Also, the means of GFA and NE are higher in case of suburban establishments. In regard to the classified industries, the average GFA and average NE are the highest in the urban- (900.1 m²) and suburban- (35) establishments, respectively, and these establishments belong to segment 2. Whereas the highest average FP is noticed in segment 1 of the suburban establishments.

Pearson coefficients and the scatter plots indicate strong linear association between FP and business variables (GFA and NE) for urban and suburban establishments. Pearson correlation coefficient values for FP explanatory variables

vary between 0.767 and 0.849 which indicates a strong linear relationship. The scatter plots for urban and suburban establishments are shown in Fig. 1. The scatter plots between weekly FP and explanatory variables (GFA and NE) for urban and suburban establishments exhibit a very strong positive correlation. These findings suggested a linear model is appropriate to estimate the weekly FP.

Model Development

The linear regression is the most commonly employed statistical technique for the developing predictive models that quantify FG [1, 3, 15, 23]. The practitioners prefer regression-based models due to its ability to describe the causal relationship and statistical robustness [54]. In this study, the FP models are estimated using OLS regression technique with a single business size (GFA and NE). For conceptual reasons, the models are developed without intercept. The logical reason for not using intercept is that no economic activity is noticed when employment or gross floor area is zero [23]. One of the largest investigations on FG models is in agreement with this practice [1]. The model structure used in this study is given in Eq. (1).

$$FP_{X_i} = \beta_i(X_i) + \varepsilon_{X_i} \quad (1)$$

where FP_{X_i} = Freight production, in tons/week, by an establishment, predicted using business size variable $X_i \forall i$, $i = 1 \& 2$ (GFA & NE); β_i = OLS estimator for slope of the regression line; ε_{X_i} = stochastic error term such that $E(\varepsilon_{X_i}) = 0$.

The model estimation results are given in Tables 2 and 3. These tables constitute the sample size (n), coefficient of determination (R^2), standard error ($S.E$), standard deviation ($S.D$) and t -statistics. All the developed FP models of urban and suburban establishments are significant at 99.9% and the R^2 value lies between 0.588 and 0.721. On the other hand, FP models for the three industrial segments show that all the parameters are significant at 99.9% confidence interval, and goodness of fit (R^2) varies between 0.656 and 0.907.

The FP rates of the urban and suburban establishments can be explained from the FP models presented in Table 3. Area-based FP rates for the urban and suburban establishments are 1.87 tons/100 m² and 2.36 tons/100 m². On the other hand, the employment-based FP rates are 0.45 tons/employee and 0.61 tons/employee. From these FP rates, it can be inferred that the weekly FP rates in the suburban regions are more than that of urban regions. The more availability of land with affordable prices and employment of more people for lesser wages may be a logical explanation for higher FP rates in these regions. This explanation can be correlated with the theories of

Table 2 Summary statistics

Variables Unit	FP Tons/week Urban establishments	GFA m ²	NE No	Variables Unit Suburban establishment	FP Tons/week	GFA m ²	NE No
All establishments							
<i>n</i> = 327				<i>n</i> = 105			
Minimum	0.6	53.0	1	Minimum	1.9	65.3	3
Average	19.6	749.8	26	Average	26.8	848	31
Maximum	173.1	4333.5	200	Maximum	136.5	4599.4	185
C.V	1.1	0.9	1.1	C.V	1.0	0.9	1.0
Establishments classified under segment 1							
<i>n</i> = 77				<i>n</i> = 24			
Minimum	1.4	122.4	4	Minimum	8.7	260.5	4
Average	36.4	779.4	20	Average	52.6	865.9	26
Maximum	123.1	3888.5	76	Maximum	136.5	4599.4	83
C.V	0.7	0.7	0.7	C.V	0.6	1.0	0.7
Establishments classified under segment 2							
<i>n</i> = 70				<i>n</i> = 24			
Minimum	1.4	73.0	3	Minimum	2.7	65.3	5
Average	23.8	900.1	32	Average	29.2	863.8	35
Maximum	173.1	4333.5	180	Maximum	109.6	2660.3	160
C.V	1.1	1.0	1.0	C.V	1.0	0.9	1.0
Establishments classified under segment 3							
<i>n</i> = 180				<i>n</i> = 57			
Minimum	0.6	53.0	1	Minimum	1.9	159.2	3
Average	10.7	678.7	27	Average	14.9	833.9	32
Maximum	94.2	4333.2	200	Maximum	44.2	3681.3	185
C.V	1.0	0.9	1.2	C.V	0.8	0.9	1.0

Table 3 FP models based on geographical location

FP models for urban and suburban establishments												
	Area-based						Employment-based					
	<i>n</i>	GFA (β_1)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>	<i>n</i>	NE (β_2)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>
Urban	304	1.87	0.655	12.4	13.9	23.9***	302	0.45	0.588	13.2	13.4	20.7***
Suburban	97	2.36	0.721	15.9	20.5	15.7***	97	0.61	0.670	17.2	20.1	14.0***

***Indicates significant at 99.9%; GFA in 100 m²

production functions in neoclassical economics, which describes that the output produced (quantity of products) is dependent on a set of inputs (land, employment, capital) used in the production.

In Table 4, FP models for different industrial segments are presented. Area-based FP rates for the urban establishments range from 1.31 tons/100 m² (segment 3) and 3.52 tons/100 m² (segment 1). Whereas employment-based FP rates in these establishments vary from 0.32 tons/employee (segment 3) to 1.80 tons/employee (segment 1). The FP rates for the suburban establishments range from 1.53 tons/100 m² to 4.11 tons/100 m² and 0.4 tons/employee to

1.87 tons/employee for area-based and employment-based FP rates, respectively. When the industrial establishments are compared irrespective of their geographical location, the higher FP rates are noticed in the low value density industries (segment 1), which include ISIC 16: Wood, wood products, furniture and fixtures and ISIC 24–25: Basic metal, alloy, metal products. The possible reason behind the higher FP rates is that these industries save more money on overhead charges (handling, packing and transportation of freight), and the investment of these savings in production leads to more output (FP rates). However, from the developed models, it can be concluded that FP rates of

Table 4 FP Models for ‘a posteriori’ segmentation of urban and suburban establishments

Area-based FP models for establishments in three different segments												
Segment	FP models for urban establishments						FP models for suburban establishments					
	<i>N</i>	<i>GFA</i> (β_1)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>	<i>n</i>	<i>GFA</i> (β_1)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>
1	71	3.52	0.765	18.7	19.2	15.1***	23	4.11	0.746	30.7	31.7	8.0***
2	68	1.89	0.656	15.7	16.5	11.3***	23	2.85	0.786	17.1	25.5	9.0***
3	171	1.31	0.762	5.8	7.5	23.3***	52	1.53	0.849	6.8	11.3	16.9***
Employment-based FP models for establishments in three different segments												
Segment	FP models for urban establishments						FP models for suburban establishments					
	<i>n</i>	<i>NE</i> (β_2)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>	<i>n</i>	<i>NE</i> (β_2)	R^2	<i>S.E</i>	<i>S.D</i>	<i>t-stat</i>
1	72	1.80	0.803	17.2	24.2	20.4***	24	1.87	0.907	19.1	32.3	15.0***
2	66	0.58	0.822	10.9	16.1	17.3***	24	0.76	0.874	14.8	29.1	12.6***
3	166	0.32	0.810	5.1	7.1	26.6***	53	0.40	0.841	7.1	11.5	16.6***

***Indicates significant at 99.9%; GFA in 100 m²

the suburban establishments are more in the area-based and employment-based models on comparison with the urban FP models. Also, on comparing the models’ predictive ability based on the standard error, it is evident that the urban models are more suitable for predicting FP than that of suburban models.

Transferability Assessment

Transfer R^2 (TR^2), transfer index (TI), weighted root mean square error (WRMSE) and relative aggregate transfer error (RATE) are the statistical metrics which are used for transferability assessment. The first metric is TR^2 , which measures the proportion of variation of the dependent variable in application context data captured by the transferred model [39, 55]. The maximum possible value of TR^2 is 1, which indicates that the estimation context model (transferred model) is entirely transferable. A value equal to zero indicates that the transferred model has no explanatory power in the application context. The negative value of TR^2 indicates that the results from the transferred model are inferior to the mean of application context data [42]. The second metric, TI defined as TR^2 divided by R^2 , measures how well a transferred model in the application context data performs relatively to application context model, and it varies between 0 to 1 [42, 55]. Like TR^2 , the TI value of 1 specifies perfect transferability, and a value equal to zero explains that the transferred model describes nothing of application context. The negative value of TI is not suggested, as it indicates misleading outcomes [9, 42]. If TI is more than 1, it represents that the transferred model is superior to the local model [9]. The third metric is weighted RMSE (WRMSE), which is used as an index to

measure relative error of the model that is transferred from estimation context to application context using local data (application context data) [39]. The RATE is the fourth metric which quantifies the ratio between the transfer WRMSE and local WRMSE [39]. The equations of these metrics are as follows:

$$\text{Transfer } R^2 (TR^2) = 1 - \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y - \bar{Y})^2} \quad (2)$$

$$\text{Transfer Index (TI)} = \frac{TR^2}{R^2} \quad (3)$$

$$\begin{aligned} \text{Weighted Root Mean Square Error (WRMSE)} \\ = \sqrt{\frac{\sum_{i=1}^n \hat{Y}_i \times \text{REM}^2}{\sum_{i=1}^n \hat{Y}_i}} \end{aligned} \quad (4)$$

$$\text{Relative Aggregate Transfer Error (RATE)} = \frac{\text{WRMSE}_t}{\text{WRMSE}_a} \quad (5)$$

where, $\hat{Y}_i$ = predicted dependent variable values produced by transferred model (estimation context) operating on independent variable values in the application context.

data; $\bar{Y}$ = mean of dependent variable values in application context data; Y = observed dependent variable values in application context data; R^2 = coefficient of determination of linear regression model fitted to application context data; WRMSE_t and WRMSE_a represent the calculation of WRMSE of transferred model and application context model, respectively, using application context data; Relative error measurement, $\text{REM} = (\hat{Y}_i - \bar{Y})/\hat{Y}_i$; n = number of observations in the linear regression model of the application context.

Combined Transfer Estimation

There are various methods for updating the transferred model. In this study, for updating the coefficients, combined transfer estimation technique is used. This approach assumes transfer bias ($d = \beta_t - \beta_a$) to be non-zero [38, 56] which makes combined transfer estimation technique as a reliable approach.

$$\text{Combined Transfer Estimation Method, } \beta' = \frac{\frac{\beta_t}{\sigma_t^2 + d^2} + \frac{\beta_a}{\sigma_a^2}}{\frac{1}{\sigma_t^2 + d^2} + \frac{1}{\sigma_a^2}} \quad (6)$$

where β' is the updated model parameter; β_t and β_a are the parameters of the estimation context model and application model, respectively; σ_t and σ_a are corresponding standard deviations in estimation and application contexts; d is the difference of parameters of estimation context model and application context model.

The FP models of urban and suburban establishments within Kerala are assessed for the transferability. The assessment results of naïve and updated FP models with GFA and NE as explanatory variables are presented in Tables 4 and 5, respectively. In these tables, the naïve and updated models are compared to assess the transferability improvement in naïve models on parameter updating. The results suggest that the updated models are better than naïve models in all the cases. The previous studies are in accordance with this result [9]. As updated models perform better, most of the discussion in this study is about updated models.

When transferability assessment is carried from urban to suburban, it is evident that most of the area-based urban FP

models can be transferred to suburban context. Similar observations are seen when employment-based urban FP models are transferred to suburban context. Only in case of 'segment 1,' the urban model is not transferable to suburban context. TR^2 and TI in case of segment 1 are negative, suggesting that the transferred models perform worse than locally estimated models. This negative value is achieved when the transferred model predicts behavior contrary to that observed [42]. When the reverse assessment of transferability is done (suburban to urban), TI value of updated models ranges from -1.245 to 0.053 (area-based and employment-based together). This shows that even though the urban models are transferable to suburban context, suburban is not transferable to urban. It is evident that transferability is not symmetric between two regions within the state, and this finding is consistent with a previous study [57]. Also, it is noticed that the area-based models are more transferable compared to employment-based models. The possible reason is that compared to area, employment is not best representing freight activities due to possible variations in employment among various industries and the replacement of manpower with automation. In addition, it is noticed that some industrial segment models are not transferable. The possible reason for this is sample size used in developing these models is too small. The small data sets give poor transferability results [46].

In this study, WRMSE and RATE are used to assess the aggregate-level prediction of the transferred model. On updating the coefficients of models, it is noticed that for updated models, RATE values have improved. In general, RATE values of updated come closer to 1. The RATE value close to 1 suggests that the aggregate prediction error

Table 5 Transferability assessment of area-based FP models

S. No		Naïve models				Updated models			
		TR ²	TI	WRMSE	RATE	TR ²	TI	WRMSE	RATE
Urban (estimation context) to suburban (application context)									
1	All Establishments	0.450	0.625	1.326	1.331	0.414	0.625	1.196	1.200
	A posteriori segmentation								
2	Segment 1	− 0.380	− 0.509	1.199	1.252	− 0.416	− 0.509	1.125	1.175
3	Segment 2	0.529	0.673	1.532	1.650	0.498	0.673	1.281	1.379
4	Segment 3	0.338	0.398	0.708	1.247	0.296	0.398	0.655	1.153
Suburban (estimation context) to urban (application Context)									
5	All establishments	0.034	0.053	0.934	0.781	0.280	0.053	1.090	0.912
	A posteriori segmentation								
6	Segment 1	− 0.455	− 0.594	0.750	0.830	− 0.225	− 0.594	0.856	0.946
7	Segment 2	− 0.817	− 1.245	0.945	0.653	− 0.058	− 1.245	1.234	0.853
8	Segment 3	0.023	0.030	0.725	0.689	0.205	0.030	0.807	0.766

Fig. 1 Scatter plots for FP models of urban and suburban establishments

of transferred model in local data is equal to error of locally estimated model. When area-based updated models in urban context are transferred to suburban, RATE values range between 1.153 and 1.379 for area-based urban updated models in suburban context. These values indicate that the prediction error of area-based urban model in suburban context lies between 15.3% and 37.9%. When the employment-based models are transferred from urban to suburban, the error ranges between 3.7% and 28.2%. From the transferability assessment results, it is evident that the area-based urban models are more transferable than employment-based urban models. Also, when ‘a posteriori’ segmentation FP models are compared for transferability, it is noticed that all urban models are transferable except low value density industry (segment 1) models.

For better perspective of the transferability assessment results, TI and RATE values are considered. As TI is dependent on TR^2 and RATE is dependent on WRMSE, it is appropriate to consider these two metrics. From RATE values, the absolute transfer errors are calculated, and absolute transfer error is calculated as percentage of absolute value of $(1 - \text{RATE})$. In case RATE is equal to 1, the absolute transfer error is zero. If error is zero, then the transferability is said to be perfect. Considering TI and

absolute transfer error values, a graph is plotted among them to visualize the transferability. Figures 2 & 3 represent the graph to visualize various transferability assessments of all establishments and a posteriori segmentation, respectively. In the graphs, only the assessment metrics of updated models are considered as transferability results in these models are better than the naïve model. The summary of the transferability assessment results is given in Fig. 4.

Joint Context Estimation

Joint context estimation is a useful approach to assess transferability by estimating a joint model (base model). The base model is developed by combining the data from both the estimation context and application context [46]. In this approach, the ‘difference parameters’ are estimated to capture the differences in the parameters of estimation and application context. A simple t-test on these ‘difference parameters’ helps in understanding whether or not the base model is transferable [8, 58]. In this study, first, single variable base models (area-based and employment-based) are estimated using the data from the urban and suburban regions. Next, for each selected region, a dummy variable for that region is interacted with the variable in the base

Fig. 2 Transferability assessment of urban and suburban FP models

models to form ‘difference variables.’ From t-statistics results on these ‘difference variables,’ the following observations are made (i) the area-based base model is transferable to both urban and suburban contexts (ii) the employment-based base model is transferable only to the suburban context. From the t-test, we can determine whether the model is transferable but not the extent of transferability. These base models are further assessed using TI to find the extent of transferability in different regional contexts and different industrial segments. In addition, the TI values obtained from combined transfer estimation and joint context estimation are compared to understand if there is any improvement in transferability after pooling the data.

Fig. 3 Transferability assessment of a posteriori segmentation FP models

Tables 6, 7 represents the TI values obtained from the combined transfer estimation and joint context estimation. On comparing the base models and updated models, the extent of transferability of area-based base model (excluding industrial segments) to urban context has improved. In all other cases, the extent of transferability of updated models using combined transfer estimation technique is greater. The geographical dissimilarities of urban and suburban regions in terms of population density may be a possible reason for no improvement in TI values on transfer of base model. In this study, from the assessment result, it can be interpreted that the updated models are better transferable than the base models.

URBAN TO SUBURBAN TRANSFERABILITY Area-based FP models <ul style="list-style-type: none"> More transferable than employment-based FP models. Except segment 1 industries FP models, all other segment industries FP models are transferable. Employment-based FP models <ul style="list-style-type: none"> Except segment 1 industries FP models, all other segment industries FP models are transferable. 		Segment 1 industries (Low value density industries) ISIC 16: Wood, wood products, furniture, and fixtures ISIC 24-25: Basic metal, alloy, metal products Segment 2 industries (Medium value density industries) ISIC 11: Beverages ISIC 13: Textile mills ISIC 17-18: Paper, paper products, and printing ISIC 22: Plastic and rubber products ISIC 29-30: Transportation equipment Segment 3 industries (High value density industries) ISIC 10: Food products ISIC 14: Wearing apparel ISIC 20-21: Basic chemicals, chemical products and pharmaceuticals ISIC 23: Non-metallic mineral products ISIC 26-28: Machinery and equipment ISIC 32: Other manufacturing industries
SUBURBAN TO URBAN TRANSFERABILITY Area-based FP models <ul style="list-style-type: none"> No FP model is transferable Employment-based FP models <ul style="list-style-type: none"> No FP model is transferable 		

Fig. 4 Summary of transferability assessment

Table 6 Transferability assessment of employment-based FP models

S. No		Naïve models				Updated models			
		TR ²	TI	WRMSE	RATE	TR ²	TI	WRMSE	RATE
<i>Urban (estimation context) to Suburban (application context)</i>									
9	All Establishments	0.388	0.580	1.751	1.461	0.357	0.580	1.536	1.282
<i>A posteriori segmentation</i>									
10	Segment 1	− 0.082	− 0.090	0.503	1.059	− 0.103	− 0.090	0.492	1.037
11	Segment 2	0.377	0.510	0.918	1.477	0.327	0.510	0.825	1.326
12	Segment 3	0.271	0.322	0.963	1.407	0.221	0.322	0.866	1.265
<i>Suburban (estimation context) to urban (application context)</i>									
13	All Establishments	− 0.362	− 0.615	1.079	0.704	− 0.025	− 0.615	1.354	0.883
<i>A posteriori segmentation</i>									
14	Segment 1	− 0.054	− 0.068	0.575	0.953	− 0.013	− 0.068	0.592	0.982
15	Segment 2	− 0.799	− 0.972	0.644	0.745	− 0.210	− 0.972	0.791	0.915
16	Segment 3	− 0.782	− 0.964	0.686	0.755	− 0.311	− 0.964	0.828	0.912

Table 7 Transfer index – combined transfer estimation and joint context estimation

Transfer method	Area-based FP models			Employment-based FP models		
	Transferred to			Transferred to		
	Transferred from	Urban	Suburban	Transferred from	Urban	Suburban
<i>All establishments</i>						
Combined transfer estimation	Urban [^]	NA	0.625	Urban [^]	NA	0.580
	Suburban [^]	0.053	NA	Suburban [^]	– 0.615	NA
Joint context estimation	Base Model	0.163	0.454	Base Model	– 0.164	0.473
<i>Segment 1</i>						
Combined transfer estimation	Urban [^]	NA	– 0.509	Urban [^]	NA	– 0.090
	Suburban [^]	– 0.594	NA	Suburban [^]	– 0.068	NA
Joint context estimation	Base model	– 0.154	– 0.512	Base Model	– 0.5	– 0.667
<i>Segment 2</i>						
Combined transfer estimation	Urban [^]	NA	0.673	Urban [^]	NA	0.510
	Suburban [^]	– 1.245	NA	Suburban [^]	– 0.972	NA
Joint context estimation	Base Model	– 0.215	0.610	Base Model	0.035	0.476
<i>Segment 3</i>						
Combined transfer estimation	Urban [^]	NA	0.398	Urban [^]	NA	0.322
	Suburban [^]	0.03	NA	Suburban [^]	– 0.782	NA
Joint context estimation	Base Model	– 2.130	– 1.200	Base Model	– 2.964	– 0.893

NA = Not applicable, [^]represents updated models

Conclusions and Policy Implications

This paper contributes to existing knowledge of freight demand generation in India by establishing statistically significant relations between FP and business size variables. For this reason, this study uses data from 432 establishments in Kerala, India, which are obtained through an EBFS. The establishments are categorized to urban and

suburban based on the geographical location to analyze the freight demand model transferability assessment. Urban and suburban ‘a posteriori’ segmentation to interpret the influence of homogeneous industries. OLS regression method is used to develop single variable models without intercept. The models indicate that the FP rates in suburban regions are more significant than that of urban, irrespective of ‘a posteriori’ segmentation. Among the ‘a posteriori’

segments, irrespective of urban or suburban, higher FP rates are noticed in Segment 1 (ISIC 16: Wood, wood products, furniture and fixtures).

The financial burden on metropolitan planning organizations (MPOs) in freight data collection urges for spatial transferability of freight demand models. Transferability assessment is carried out using both urban and suburban models to analyze the model performance as well as the direction of transferability. Transferability assessment performance is carried out using TR^2 , TI, WRMSE and RATE. The models are directly transferred initially; later the naïve models are transferred by updating their coefficients using combined transfer estimation technique. The transferability assessment is done for naïve and updated models, and it can be seen that updated models offer better transferability performance. It is also noticed that most of the urban models are transferable in the suburban context. However, when the suburban models are transferred to the urban context, only a few models are transferable. These statements about transferability between urban and suburban conclude that transferability is asymmetric. The exclusion of spatial factors like population density, road density, road intersection, distance of establishment from seaport and city center, etc., in the models can be a possible reason for the asymmetric transferability between the urban and suburban regions. Also, the suburban models are not transferable because the sample size used for developing these models is small. Finally, among the urban models, area-based models are more transferable than employment-based models. The medium and high value density industry (segment 2 and 3) urban models are transferable to suburban context, and among them, medium value density industry (segment 2) models are more transferable. In suburban models, only the area-based model of high value density industry (segment 3) is transferable. Joint context estimation technique is also used to check whether or not the pooled data models give better transferability results. We did not notice higher TI values in this case as compared to the combined transfer estimation technique. In general, as there is rapid growth of freight traffic and shortage of freight data in suburban regions, the spatial transferability of urban models to suburban context can overcome the freight demand model availability to a certain extent in suburban regions.

From this study, a number of inferences have emerged with significant implications for planning and policy making. The first finding is that the degree of spatial transferability widely varies among different industrial classes and depends on the following factors: (i) business variables employed for measuring FP (ii) type of community (iii) commodity value density of industry sectors. The first factor suggests that the area-based models show better transferability than the employment-based models. The

second factor suggests that the transferability is possible from urban community to suburban community and not vice versa. The agencies looking forward to transfer FP models may do well if the transfer is done from regions with high population density to low population density. The last factor suggests that the transferability of models is better in medium and high value industries. In this study, the transferability of the FP models opens the possibility for transportation agencies, planners and practitioners to identify direction and the extent of transferability. The research findings are expected to be helpful in saving financial resources for data collection exercises and develop freight model system within budgetary constraints. However, like any other study, this study has certain limitations, which paves the way for further research.

The findings of this study may not be generalized as the data pertain to only one state in India. However, the methodology is generic and can be utilized to investigate the freight model transferability direction in cities in other Indian states. The findings of this research are much useful for the planning agencies in Kerala or in a similar coastal state in India or elsewhere. The results of this research will assist them in data collection programs, and subsequently useful for developing comprehensive mobility plans. This research is limited to intrastate model transferability assessment and requires further investigation on interstate model transferability. It may also be noted that small sample size is the limitation of this study. Models built with small data sets produce poor or moderate transferability findings, suggesting the significance of using sufficient samples in building models. The generalization of this study requires data with larger sample size from several cities across different states in India. Also, it is important to identify the factors causing the transferability to be asymmetric and to know the reasons for transferability of only some industrial segment models.

Acknowledgements The study was financially supported by Research Initiation Grant (RIG Head 06/03/302). Birla Institute of Technology and Science (BITS) Pilani, Hyderabad, India. The authors acknowledge the Institute's financial support for this research effort.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. J. Holguín-Veras, C. Lawson, C. Wang et al., *NCFRP Research Report 37: Using Commodity Flow Survey Microdata and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook* (The National Academies Press, Washington DC, 2016)

2. A. Comi, R. Donnelly, F. Russo, Urban freight models, in *Modelling Freight Transport*. ed. by L. Tavasszy, G. de Jong (Elsevier Inc., New York, 2013), pp. 163–200
3. I. Sánchez-Díaz, Modeling urban freight generation: A study of commercial establishments' freight needs. *Transp Res Part A Policy Pract* (2017). <https://doi.org/10.1016/j.tra.2016.06.035>
4. A. Pani, P.K. Sahu, Planning, designing and conducting establishment-based freight surveys: A synthesis of the literature, case-study examples and recommendations for best practices in future surveys. *Transp Policy* **78**, 58–75 (2019). <https://doi.org/10.1016/j.tranpol.2019.04.006>
5. F. Toillier, M. Serouge, J.L. Routhier, et al (2016) How can urban goods movements be surveyed in a megacity? The case of the Paris region, in: *Transportation Research Procedia*. Elsevier B.V., pp 570–583
6. A. Pani, P.K. Sahu, Modelling non-response in establishment-based freight surveys: a sampling tool for statewide freight data collection in middle-income countries. *Transp Policy*. (2019). <https://doi.org/10.1016/j.tranpol.2019.10.011>
7. Z. Wafa, C.R. Bhat, R.M. Pendyala, V.M. Garikapati, Latent-segmentation-based approach to investigating spatial transferability of activity-travel models. *Transp Res Rec* **2493**, 136–144 (2015). <https://doi.org/10.3141/2493-15>
8. S. Sikder, A.R. Pinjari, S. Srinivasan, R. Nowrouzian, Spatial transferability of travel forecasting models: a review and synthesis. *Int J Adv Eng Sci Appl Math* **5**, 104–128 (2013). <https://doi.org/10.1007/s12572-013-0090-6>
9. S. Sikder, A.R. Pinjari, Spatial transferability of person-level daily activity generation and time use models. *Transp Res Rec* **2343**, 95–104 (2013). <https://doi.org/10.3141/2343-12>
10. K. Agyemang-Duah, F.L. Hall, Spatial transferability of an ordered response model of trip generation. *Transp Res Part A Policy Pract* **31**, 389–402 (1997). [https://doi.org/10.1016/S0965-8564\(96\)00035-3](https://doi.org/10.1016/S0965-8564(96)00035-3)
11. J. Holguín-Veras, I. Sánchez-Díaz, C. Lawson et al., Transferability of freight trip generation models. *Transp Res Rec* (2013). <https://doi.org/10.3141/2379-01>
12. A. Pani, P.K. Sahu, Comparative assessment of industrial classification systems for modeling freight production and freight trip production. *Transp Res Rec* **2673**, 210–224 (2019). <https://doi.org/10.1177/0361198119834300>
13. A. Samimi, A. Mohammadian, K. Kawamura, A nationwide web-based freight data collection. *Can J Civ Eng* **40**, 114–120 (2013). <https://doi.org/10.1139/cjee-2011-0555>
14. G. Giuliano, P. Gordon, Q. Pan et al., Estimating freight flows for metropolitan area highway networks using secondary data sources. *Netw Spat Econ* **10**, 73–91 (2010). <https://doi.org/10.1007/s11067-007-9024-9>
15. A. Pani, P.K. Sahu, G.R. Patil, A.K. Sarkar, Modelling urban freight generation: a case study of seven cities in Kerala, India. *Transp Policy* **69**, 49–64 (2018). <https://doi.org/10.1016/j.tranpol.2018.05.013>
16. A.R. Alho, de Abreu e Silva J, , Modeling retail establishments' freight trip generation: a comparison of methodologies to predict total weekly deliveries. *Transportation (Amst)* **44**, 1195–1212 (2017). <https://doi.org/10.1007/s11116-016-9700-z>
17. National Transport Development Policy Committee (2014) India Transport Report. Vol. 1: Executive Summary. New Delhi
18. Ministry of Urban Development (2015) Smart City: Mission Statement & Guidelines
19. Ministry of Housing and Urban Affairs (2016) Smart Cities Mission. <http://smartcities.gov.in/>. Accessed 2 Jan 2020
20. Ministry of Heavy Industries and Public Enterprises (2018) Annual Report 2018–19. New Delhi
21. (2020) CEIC. <https://www.ceicdata.com/en/india/railway-passenger-and-freight-traffic-monthly/freight-traffic-freight-originating>. Accessed 1 Jan 2020
22. J. Holguín-Veras, E. Thorson, Trip length distributions in commodity-based and trip-based freight demand modeling: Investigation of relationships. *Transp Res Rec* (2000). <https://doi.org/10.3141/1707-05>
23. P.K. Sahu, A. Pani, Freight generation and geographical effects: modelling freight needs of establishments in developing economies and analyzing their geographical disparities. *Transportation (Amst)* (2019). <https://doi.org/10.1007/s11116-019-09995-5>
24. Ha D-H, Combes F (2015) Building a model of freight generation with a commodity flow survey. ed. by Clausen U, Friedrich H, Thaller C, Geiger C (eds) *Commercial Transport, Proceedings of the 2nd Interdisciplinary Conference on Production Logistics and Traffic*. Springer, Cham, pp 23–37
25. A.R. Alho, J. de Silva, Freight-trip generation model. *Transp Res Rec J Transp Res Board* **2411**, 45–54 (2014). <https://doi.org/10.3141/2411-06>
26. C. Bastida, J. Holguín-Veras, Freight generation models: Comparative analysis of regression models and multiple classification analysis. *Transp Res Rec* (2009). <https://doi.org/10.3141/2097-07>
27. D. Beagan, M.J. Fischer, A. Kuppam, *Quick Response Freight Manual II - FHWA-HOP-08-010* (US Department of Transportation Modelling Freight Transport, Washington DC, 2007)
28. P. Agnivesh, P.K. Sahu, F.A. Bhat, Assessing the spatial transferability of freight (Trip) generation models across and within states of India: empirical evidence and implications for benefit transfer. *Net Spat Econ* **21**(2), 465–493 (2021). <https://doi.org/10.1007/s11067-021-09530-z>
29. P. Agnivesh, P.K. Sahu, J. Holguín-Veras, Examining the determinants of freight transport emissions using a fleet segmentation approach. *Transp Res Part D: Transp Environ* **92**, 102726 (2021). <https://doi.org/10.1016/j.trd.2021.102726>
30. A. Chandra, M.N. Sharath, A. Pani, P.K. Sahu, A multi-objective genetic algorithm approach to design optimal zoning systems for freight transportation planning. *J Transp Geogr* **92**, 103037 (2021). <https://doi.org/10.1016/j.jtrangeo.2021.103037>
31. G.R. Patil, S. Thadoju, P.K. Sahu, A. Mondal, V. Bajpai, Data collection and modeling of restaurants' freight trip generation for Indian cities. *Transp Develop Econ* **7**, 9 (2021). <https://doi.org/10.1007/s40890-021-00114-7>
32. P.K. Sahu, A. Chandra, A. Pani, B.B. Majumdar, Designing freight traffic analysis zones for metropolitan areas: identification of optimal scale for macro-level freight travel analysis. *Transp Plan Tech* **43**(6), 620–637 (2020). <https://doi.org/10.1080/03081060.2020.1780711>
33. A. Pani, F.A. Bhat, P.K. Sahu, Effects of business age and size on freight demand: decomposition analysis of indian establishments. *Transp Res Rec J Transp Res Board* **2674**(2), 112–126 (2020). <https://doi.org/10.1177/0361198120902432>
34. A. Pani, P.K. Sahu, B.B. Majumdar, Expenditure-based segmentation of freight travel markets: Identifying the determinants of freight transport expenditure for developing marketing strategies. *Res Transp Bus Manag* **33**, 100437 (2019). <https://doi.org/10.1016/j.rtbm.2020.100437>
35. A. Chandra, A. Pani, P.K. Sahu, Designing zoning systems for freight transportation planning: a GIS-based approach for automated zone design using public data sources. *Transp Res Procedia* **48**, 605–619 (2020). <https://doi.org/10.1016/j.trpro.2020.08.063>
36. A. Pani, P.K. Sahu, A. Chandra, A.K. Sarkar, Assessing the extent of modifiable areal unit problem in modelling freight (trip) generation: relationship between zone design and model estimation results. *J Transp Geogr* **80**, 102524 (2019). <https://doi.org/10.1016/j.jtrangeo.2019.102524>

37. T. Atherton, M. Ben-Akiva, Transferability and updating of disaggregate travel demand models. *Transp Res Rec* **610**, 12–18 (1976)
38. M. Ben-Akiva, D. Bolduc (1987) Approaches to model transferability and updating: the combined transfer estimator. *Transp Res Rec* 1–7
39. F.S. Koppelman, C.G. Wilmot, Transferability analysis of disaggregate choice models. *Adv Agron* **58**, 151–199 (1982). [https://doi.org/10.1016/S0065-2113\(08\)60255-2](https://doi.org/10.1016/S0065-2113(08)60255-2)
40. R.A. Galbraith, D.A. Hensher, Intra-metropolitan transferability of mode choice models. *J Transp Econ Policy* **16**, 7–29 (1982)
41. D.P. McArthur, G. Kleppe, I. Thorsen, J. Ubøe, The spatial transferability of parameters in a gravity model of commuting flows. *J Transp Geogr* **19**, 596–605 (2011). <https://doi.org/10.1016/j.jtrangeo.2010.06.014>
42. C.G. Wilmot, Evidence on transferability of trip-generation models. *J Transp Eng* **121**, 405–410 (1995). [https://doi.org/10.1061/\(ASCE\)0733-947X\(1995\)121:5\(405\)](https://doi.org/10.1061/(ASCE)0733-947X(1995)121:5(405))
43. D.S. Santoso, K. Tsunokawa, Spatial transferability and updating analysis of mode choice models in developing countries. *Transp Plan Technol* **28**, 341–358 (2005). <https://doi.org/10.1080/03081060500319694>
44. N. Karasmaa, Evaluation of transfer methods for spatial travel demand models. *Transp Res Part A Policy Pract* **41**, 411–427 (2007). <https://doi.org/10.1016/j.tra.2006.09.009>
45. C.-W. Kuo, M.-L. Tang, Survey and empirical evaluation of nonhomogeneous arrival process models with taxi data. *J Adv Transp* **47**, 512–525 (2011). <https://doi.org/10.1002/atr>
46. S. Sikder, B. Augustin, A.R. Pinjari, N. Eluru, Spatial transferability of tour-based time-of-day choice models: Empirical assessment. *Transp Res Rec* **2429**, 99–109 (2014). <https://doi.org/10.3141/2429-11>
47. H. Fu, C.G. Wilmot, E.J. Baker, Sequential logit dynamic travel demand model and its transferability. *Transp Res Rec* (2006). <https://doi.org/10.3141/1977-05>
48. R.J. Johnston, Choice experiments, site similarity and benefits transfer. *Environ Resour Econ* **38**, 331–351 (2007). <https://doi.org/10.1007/s10640-006-9073-4>
49. R.G. Schiffer, *NCFRP Report 735: Long-Distance and Rural Travel Transferable Parameters for Statewide Travel Forecasting Models* (The National Academies Press, Washington DC, 2012)
50. G. Tian, K. Park, R. Ewing, Trip and parking generation rates for different housing types: effects of compact development. *Urban Stud* **56**, 1554–1575 (2019). <https://doi.org/10.1177/0042098018770075>
51. Department of Economics & Statistics (2014) Provisional Report on 6th Economic Census. Thiruvananthapuram, Kerala
52. (2004) Roasoft Inc. <http://www.raosoft.com/samplesize.html>. Accessed 3 Jan 2021
53. Office of the Registrar General & Census Commissioner (2011) Provisional Population Totals Paper 2 of 2011: India (Vol II). Ministry of Home Affairs, Government of India
54. D.N. Gujarati, D. C. Porter, D. Gunasekar, *Basic Econometrics* (2004)
55. E. Kawamoto, Transferability of standardized regression model applied to person-based trip generation. *Transp Plan Technol* **26**, 331–359 (2003). <https://doi.org/10.1080/03081060310001635896>
56. M. Ben-Akiva, D. Bolduc, K. Pene, The combined estimator approach to model transferability and updating. *Empir Econ* **20**, 167–175 (1995). <https://doi.org/10.1007/BF01235164>
57. R. Nowrouzian, S. Srinivasan, Modeling the effect of land use on person miles traveled by using geographically weighted regression. *Transp Res Rec* (2013). <https://doi.org/10.3141/2397-13>
58. J. L. Bowman, M. Bradley, J. Castiglione, S. L. Yoder, Making advanced travel forecasting models affordable through model transferability, in *Transportation Research Board 93rd Annual Meeting*. Washington, D.C. (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.