Distilling actionable insights from big travel demand datasets for city planning

Alvin Chua, Serene Ow, Kevin Hsu, Yazhe Wang,
Michael Chirico and Zhongwen Huang

July 30, 2019

Distilling Actionable Insights from Big Travel Demand Datasets for City Planning

Alvin CHUA[a*], Serene OW[b], Kevin HSU[a], WANG Yazhe[b], Michael CHIRICO[b], HUANG Zhongwen[a]

[a] *Urban Redevelopment Authority, 45 Maxwell Road, Singapore 069118*
[b] *Grab, 9 Straits View, Marina One West Tower, #23-07/12, Singapore 018937*
[c] *Centre for Liveable Cities, 45 Maxwell Road, Singapore 069118*

*Corresponding author

JEL classification:
Data Estimation, Analytical Method, Urban Transportation System

Keywords:
Ride-hailing, Public transport, Land use transport interaction, Land use activities, Trip generation rates

ABSTRACT

Working towards a more data-informed land use, amenities and infrastructure planning process, the Singapore Urban Redevelopment Authority (URA) harnesses big data and spatial analytics to deepen its understanding of urban activity and mobility patterns. Big travel demand datasets from public transport and ride-hailing services enable planners to observe mobility patterns at a high level of detail for large numbers of users, trips, and trip types. Since August 2018, the URA has been working with leading technology company and ride-hailing operator Grab to understand how daily commute patterns vary between existing and new transport modes, and how the volume of activities in each area evolves across different times of day. This paper describes the novel dataset and analytical techniques utilised to study the relationship between urban activity and mobility. It will also report how spatiotemporal characteristics of the urban environment, such as land use mix, location accessibility, and peak-hour travel demand, influence commutes by different modes in each area. By studying mobility over a range of travel modes, this method of analysis will provide city planners with richer insights to better assess infrastructure requirements for new developments. The findings are also useful for emerging transport providers, who can improve service delivery across short- and medium-term time scales.

1. Introduction

Cities are centres of social and economic activities where multiple urban systems work in tandem, but each system has its own rhythm. To meet the needs of people and businesses, city planners must balance competing demands for limited urban space and provide the necessary infrastructure to support them. The interplay between land use and mobility systems plays a critical role in connecting people to opportunities and services, as well as in enabling the efficient flow of goods and resources. Demand for mobility services arises from this social and economic background, both at the origin where trips begin, and at the destinations where trips end. The level and pattern of demand for different types of mobility services differs across space and time, due to the type of land use activities, mix of uses, development density, accessibility, and myriad other possible factors.

Much effort has been made over the years to better understand the complex interactions between land use and mobility. Most research is focused on the mode choice, route choice, and implied time valuation of commuters, with less attention paid to developing a more comprehensive understanding of land use trip rates, and how these trip rates vary as land use mix, demographic profile and accessibility levels differ. Travel surveys offer valuable information on population preferences for mode and distance, and how these preferences vary by commuting purpose. The accuracy and completeness of travel surveys obviously have a critical impact on the quality of conclusions one can glean therefrom. For instance, it will be challenging to generalise the effects of land use on trip rates if there are insufficient observations for analysis, or if their representativeness is questionable or unknown. This has severely limited progress in modelling trip generation and attraction rates, which are critical estimates required to plan transport infrastructure and services. Even though cell phone-based interactive methods have improved the spatial and temporal granularity of travel surveys in recent years, the cost of obtaining sufficiently large population samples and the latency between survey iterations are still impediments to city planners closely monitoring changes in mobility patterns.

Big datasets like transport fare collection records provide granular population-scale observations of travel demand. The fusion of big travel demand data with land use and administrative data will provide city planners with an integrated view of urban dynamics, empowering analysts to characterise the land use and mobility patterns of each neighbourhood and distil the differences among them. More importantly, it can serve as a practical and cost-effective way to identify the factors that drive changes in mobility and understand the purposes of commutes. This will in turn help city planners develop a data informed assessment of infrastructure requirements and provide valuable insights for transport service operators to improve service standards.

In this paper, we describe a novel dataset assembled by the Singapore Urban Redevelopment Authority (URA) and leading technology company and ride-hailing operator Grab to study land use-transportation interactions in Singapore. We outline our analytical approach, which takes an activity-based perspective towards modelling land use and mobility interactions to evaluate how bus and rail mobility patterns differ from private hire cars and taxis. We will also show how characteristics of the urban environment such as land use mix, location accessibility, and peak-hour travel demand influence commute by the two modes. The findings from this case study demonstrate the value of a public-private collaboration to study land use mobility patterns, and the feasibility to obtain a set of land use and location sensitive space-time trip rates.

2. Urban Context

Singapore is a compact, high-density island city with land area of around 720 km^2; economic nodes are located across 4 gateways (i.e. Central Area, Eastern Gateway, Western

Gateway, Northern Gateway) to meet different business needs as well as to bring jobs closer to residential towns that are located throughout the city. Each day, approximately 6.1 million trips are made by rail and bus; the use of privately-owned cars comprises one third of the peak-hour transport mode share (Land Transport Authority, 2019). In planning for Singapore's land use, the goal is to create a sustainable Singapore that provides a quality living environment, offers plentiful growth opportunities and jobs for the people, and safeguards a clean and green landscape (https://www.ura.gov.sg/Corporate/Planning). To achieve this, Singapore adopts a long-term perspective in planning, balances economic, social and environmental considerations, and optimizes the use of limited land to ensure the current and future needs of the people are met. One of the key focus areas is to ensure that moving around Singapore continues to be convenient and pleasant. Singapore has adopted various strategies in the past to encourage the use of public transport, as well as manage car ownership and usage. In addition to these, Singapore continues to enhance connectivity across Singapore with expanded public transport and active mobility networks; plan for more jobs in the business nodes closer to homes where work and play will be more accessible; and harness new mobility technologies and operating models to achieve more efficient delivery of goods and a wider range of commuter options for residents (https://www.ura.gov.sg/Corporate/Planning/Draft-Master-Plan-19/Themes).

3. Emergence of Shared Mobility Services

"On-demand" transit services have existed for decades, even without the support of today's advanced routing and scheduling technology (Kirby et al., 1974). In markets such as the United States, users can pre-schedule rides by telephone, where a driver will arrive at an agreed address and deliver the user to an agreed destination. These services include "paratransit" assistance for persons with disabilities, or "demand-response transit" services in low-density areas where fixed-route buses are not operationally appropriate or financially viable (Mageean and Nelson, 2003). In the 2000s, "vehicle sharing" emerged as a concept to take advantage of improvements in cellular technology. These services allow users to forgo owning a vehicle, by providing access to a range of mobility options like bicycles, scooters, and cars (Shaheen et al., 2009). Finally, the broad availability of smartphones and advanced routing technologies has enabled the integration of these trends, manifesting as next-generation shared mobility services, including ride-hailing, micro-transit, and fully-fledged transportation network companies (TNCs). This suite of emerging services, provided by a mix of public entities and private companies, has helped to rapidly reshape the urban transportation landscape.

Shared mobility services enable users to gain short-term access to a certain transportation mode, on an as-needed basis. They generally fall into one of two types: "core and incumbent services" that encompass traditional shared modes of transportation (e.g car rental, limos, public transit, taxis, etc.), and the newer "innovative services" that have arisen with the advent of mobile technology, or in some cases, to serve the workers of mobile technology companies (Cohen and Shaheen, 2018). Innovative services emerging in the past decade include app-based bike sharing, app-based car sharing, P2P vehicle sharing, and app-based on-demand ride services. This last category includes TNCs like Bolt, Didi, Grab, Lyft and Uber, as well as supplementary services in lieu of fixed-route bus and rail, such as high-tech companies' private shuttles and public-private micro-transit. King County Metro in Seattle, USA defines micro-transit as “IT-enabled, privately-operated, multi-passenger transportation service” that provides an “on-demand dynamic shuttle”.

Among these diverse modes, the most prominent are TNCs that operate much like on-demand taxi fleets. The term “transportation network company” was coined in 2013 by the California Public Utilities Commission to classify an emerging form of shared mobility that combines for-hire transportation with an online-enabled platform, enabling drivers to link with passengers (California Public Utilities Commission, n.d.). These companies began to appear as innovations in cellular and geolocation technology enabled riders to hail drivers in real-time via smartphone app, and then efficiently routed to a destination. These services can also be referred to as “ridesourcing” or “ride hailing,” or if carpooling is involved, “ridesharing” (Cohen and Shaheen, 2018). This form of shared mobility is currently dominated by companies like Uber and Lyft in the US, Didi Chuxing in China, Ola in India, Grab in Southeast Asia, and Bolt in some European and African markets. A notable achievement of TNCs is the speed with which they became a global phenomenon: The Natural Resource Defense council found that Uber, the largest of the ride-hailing companies, took only six years to exceed the market valuation of General Motors, a vehicle manufacturer founded in the early 20th century, and estimated that for every 10 transit trips taken in the United States, seven of them are taken in an Uber or Lyft vehicle.

3.1 Shared Mobility Networks as Sources of Data

With their extensive reach to wide user base, TNCs are becoming an increasingly important source from which data can be gathered, allowing city planners to study where and when trips occur, and how the supply and demand of mobility is met. Singapore is no exception to this global phenomenon. Ride hailing is an emerging mode of transport, with user penetration estimated at 32.2% in 2019 and expected to hit 39.9% in 2023 (Statista, 2019). In the past, public agencies were the major operator of mass transit services. Private companies had a share of the market but did not possess the means to reach a broader segment of the population. Either the government had the data, or no one had it. Now, with the emergence of TNCs and other shared mobility platforms, knowledge of travel demand is no longer the sole domain of the government. These emergent services yield plentiful data about user journeys, including origins, destinations, and time of use. Furthermore, the services also shape preferences and alter travel patterns, which changes the ecosystem in which mobility services operate. Thus, to develop optimal transport solutions for a city, neither the public nor private sector is able to operate alone. City planners must actively engage and cultivate the transportation ecosystem, by creating consortia, partnerships and potentially even physical spaces for collaboration.

Ride hailing operators like Grab have worked with city governments in South East Asia to address transportation problems such as helping to close transit gaps for commuters and public transport gaps in public infrastructure. As much as 13% of Grab rides in Malaysia connect to public transport, and 10% in Indonesia. Across the region, Grab aims to make it easier for commuters to get to their final destination regardless of which transportation service they use. Grab is also working on providing insights to help city planners make their city smarter and more liveable, for example through analysis of traffic data and travel patterns to provide insights on sustainable urban transport measures. One promising area of collaboration is for ride hailing operators and city planners to jointly study land use transport interactions to inform city and transport planning. By collaboratively investigating this topic area, they can draw on ride-hailing and public transport travel demand data, as well as land use information, to form a more complete picture of travel activity and behaviour. This collaboration in the context of Singapore constitutes the focus of this study.

4. Data

The land use data for our analysis comprises 24 variables derived from job estimation statistics, points of interest (POI) in the Singapore street directory, inventory of dwelling unit stock and population demographics. The data is aggregated to a hexagonal grid of equally-sized cells 200m per side, and 400m across. Travel demand data comprises public transport fare records and ride hailing transactions. Similar to land use variables, travel demand data is aggregated as origin-destination matrices summarising hexagon-to-hexagon travel demand at hourly intervals. Table 1 provides a comprehensive breakdown of the land use variables we consider in this study and the corresponding activities they primarily represent. We outline as secondary attributes potential factors that further shape our interpretation of results.

4.1.1 Job Estimation Statistics

For transport modelling purposes, URA estimates number of jobs per industrial sector at known locations, to calculate daily trips generated (see Table 1). While these trip generation estimates reliably capture the location and magnitude of work activity in Singapore, it is limited in several ways. First, statistics are estimated for traffic zones that are larger than the hexagonal cells considered for analysis. Weighted overlay with equal assigned weights was applied to smoothly distribute job estimates from each traffic zone to spatially intersecting hexagons (see Figure 1a-e). Furthermore, recreation and leisure activities are not adequately represented, so we turn to the place-of-interest (POI) data in the Singapore street directory as a supplementary data source (see Figure 1f).

4.1.2 Public Transport Data

Fares on the Singapore public transport network are almost exclusively paid with electronic, contactless smart cards. The card reader system records origin, destination, time, travel time, and a unique identifier which can be used to distinguish the transfer duration between successive bus or rail trips. As the public transport network follows a hub and spoke model, trip chaining (Alsger et al., 2015) is required to obtain a representative estimate of travel demand between locations, for activities besides transfer. Following parameters derived from an earlier study (Chua and Wang, 2018), sequential trips with transfer durations of at most 6.47 minutes toward rail and at most 11.7 minutes toward bus will be chained (i.e. merged and considered as a single trip in our analysis).

4.2 Public Transport Reach

To capture locational accessibility afforded by public transport, URA estimates the number of locations reachable island wide from a particular origin within 45 minutes' travel time. The estimates are calculated with a GIS network routing engine based on the Singapore bus, rail and pedestrian network. Travel times are rounded to cell centroids, considering only the shortest travel time though there may be multiple route options available. For the purposes of this study, accessibility scores are calculated for every cell. A spatial representation of reach by public transport is shown in Figure 2.

Table 1. Detailed breakdown of the variables considered in this study mapped to their respective information sources, as well as the corresponding activities and attributes.

Information Source	Variable	Primary Activity	Secondary Attributes
Jobs Estimation Statistics	Office Non-Manufacturing Manufacturing Public Utilities Military Port	Work	Not Applicable
	Airport	Flights	Partly representative of work activities
	Healthcare	Healthcare	
	Retail	Daily needs	
	Education	School	
	Hotel	Tourist accommodation	
Street Directory Place of Interest (POI)	POI Transport	Travel	Partly representative of activities related to work and daily needs
	POI Recreation	Recreation	
	POI Work	Work	Not Applicable
Dwelling Unit Stock	Landed	Accommodation	Low density affluent
	Non-landed		High density affluent
	Public 5 room & larger apartments		High density
	Public 4 room apartments		
	Public 3 room apartments		
	Public 1 & 2 room apartment		
Population	Population density		Crowded
	Senior density		Older development
	Children density		Younger development

5. Data Processing and Analytical Method

5.1 Public Transport Origin-Destination Demand Smoothing

One feature distinguishing ride-hailing from public transit trips is the availability of point-to-point service. If someone wants to go from their home to the park, they simply designate these endpoints in the Grab mobile application and (inasmuch as the origin and destination are both connected to the road network) their assigned driver will ferry them nearly “door to

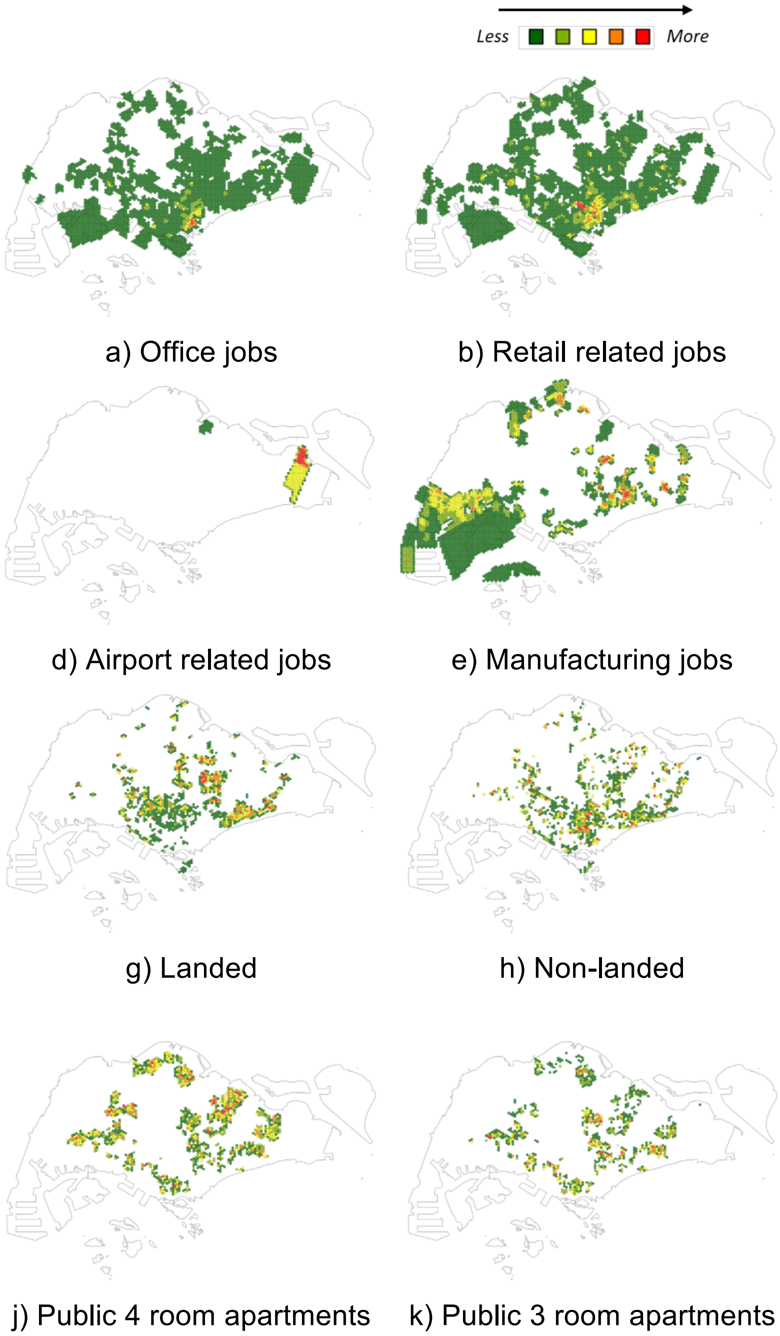


Figure 1. Spatial distribution of major land use variables. a-e) Job estimates by sector representing the whereabouts of work activities. f) Public transport hubs. g-k) Dwelling units by property type.

door". By contrast, the same trip on public transit is often supplemented in the "last mile" by alternative transportation -- typically by foot, bike, scooter, etc. This presents a challenge to our analysis because travel demand on public transit modes tends to concentrate heavily at major transit nodes (i.e. those with subway stations and bus interchanges), and only peripherally at bus stops. Any hexagon without either a bus stop or a subway stop will therefore "generate" no demand in our data. This muddles the interpretation of our model, as

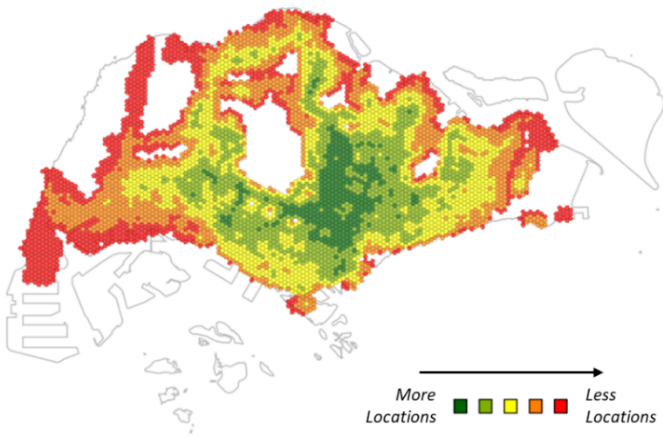


Figure 2. Spatial representation of reach by public transport estimated based on the shortest travel time from every hexagon. The scoring metric considers a combination travel by bus, rail and pedestrian modes.

public transit demand will mostly be driven by hexagons with major transit nodes, a small subset of all hexagons.

To overcome this, we use a simple agnostic spatial smoothing approach to better reflect the motivating example proposed above regarding “last mile” behaviour on public transit. Specifically, we use a geometrically weighted window at origin and destination to diffuse demand into the surrounding two rings of neighbouring hexagons (6 neighbours in the first ring, 12 neighbours in the second ring) for each observed hexagon of demand. That is, for each hexagon h with observed demand y_h, we redistribute this quantity among the neighborhood $N(h)$ of hexagons within the first 2 rings around h

$$y_{h'} = w_{h',h} y_h \quad \forall h' \in N(h)$$
$$w_{h',h} = r^{d(h',h)}$$
$$d(h',h) = \text{neighbor count of } h' \text{ vis-a-vis } h$$ [1][2]

r is a parameter which controls the assumed percentage of “organic” trips for the central hexagon, i.e., the percent of public transit trips originating in h where the passenger began their trip in h itself.

[1] More rigorously, if we consider our hexagon map as a graph where each hexagon is a node and two nodes are connected if the corresponding hexagons are adjacent, d(h',h) is the distance on this graph between h and h’.

[2] A subtle assumption here is that all neighbours in the second ring (where d(h’, h)=2) have equal weight, while for a hexagonal coordinate system this is not completely accurate, geometrically. All ring-1 hexagon centroids are distance s*sqrt(3) away from the “centre” hexagon (ring-0), where s is the hexagon’s side length, but half of the ring-2 hexagons are distance 3*s away while the rest are 2*s*sqrt(3) away (about 15% further). Lacking any data for validation, we adopt our approach for simplicity, but an alternative specification would be to impose a kernel discretised at the centroids, i.e., to measure the plane distance d from h to h’ and set weight as a function of d, e.g. for a normal kernel exp(-(l*d)^2), setting some distance cut-off for computational simplicity; an even more thorough accounting would set the kernel bandwidth according to a supplemental data set and then integrate the kernel over the hexagonal regions. Given the geometric weighting, our approach is most similar to using a Laplace kernel.

To be clear, this smoothing formulation assumes (1) all passengers arrive to a public transit node from the surrounding neighbourhood of about 860 meters; (2) the distribution of whence these passengers arrive is symmetric and follows a geometric pattern based on the hexagon ring structure; and (3) this relationship is constant throughout the region[3] and agnostic to whether the trips are from bus, subway, or both.

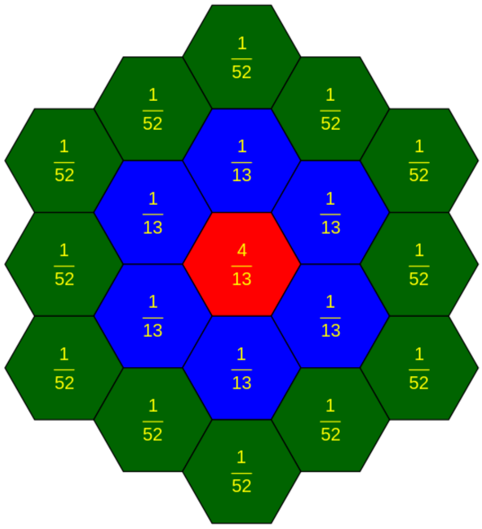


Figure 3. Smoothing kernel applied to redistribute public transit data into the surrounding neighbourhood of public transit nodes

We extend this to smoothing paired origin-destination demand $y_{o,d}$ in the natural way by smoothing demand both at origin and destination according to the geometric window, and multiplying the weight at origin and the weight at destination:

$$y_{o',d'} = w_{o',o} w_{d',d} y_{o,d} \quad \forall (o', d') \in N(o) \otimes N(d)$$

The additional assumptions implied by this extension are (1) accumulation at origin and dispersion at destination are independent, i.e., knowing the origin hexagon o' of a passenger travelling from o to d provides no information about their destination hexagon d'; (2) the dispersion pattern is the same at origin as at destination. Ultimately, the assumed latent demand ${y_{o,d}}^s$ for each hexagon pair, i.e., the outcome used in our regression model described below, is given by the smoothed demand:

$${y_{o,d}}^s = \sum_{(o',d') \in N(o) \otimes N(d)} w_{o,o'} w_{d,d'} y_{o',d'}$$

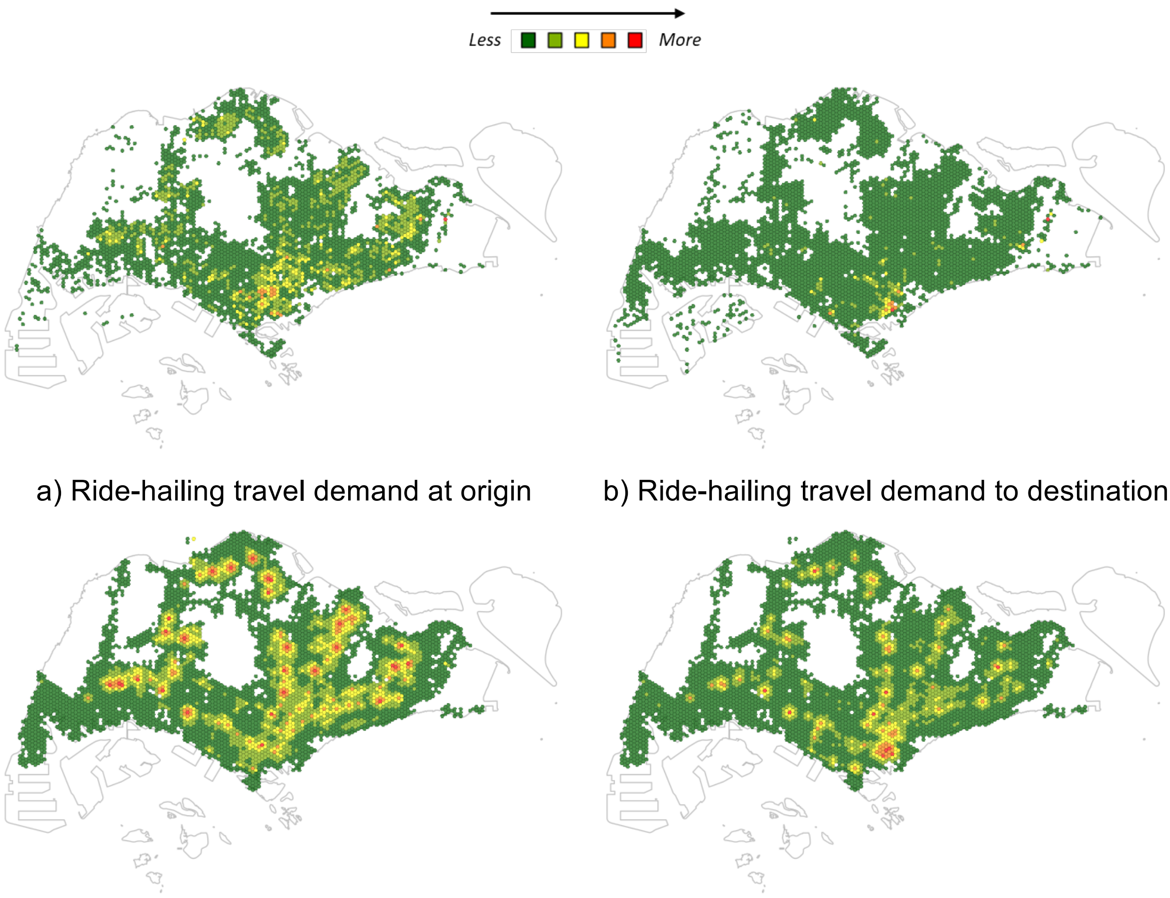
5.2 Regression Analysis

[3] We note here a further consideration -- for boundary hexagons (defined by hexagons which have as 2-level neighbours hexagons which produced neither public transport nor private ride hailing trips), we proportionally reassign any trips that would have been “diffused out” back into the “included” hexagons -- for example, if 5% of a centre hexagon h’s trips would be assigned to a hexagon in the ocean, that 5% is “distributed back” to the other neighbors of h in proportion to the original weights w(h, h’). This is equivalent to “dividing out” the weight as in a truncated probability distribution.

Linear regression is used extensively in practical applications (Yan and Xin, 2009) to explain variation in response variables that may be attributed to variation in the explanatory variables. We apply multivariate linear regression analysis to measure the relationship between land use variables and the travel demand of different modes. A multivariate linear regression model is a statistical linear model formulated as:

$$Y_j = \beta_0 + \beta_1 X_{j1} + \beta_2 X_{j2} + \ldots + \beta_p X_{jp} + \varepsilon_j$$

In the formula, $Y_j$is the j^{th} observation of the response variable (i.e. travel demand), j = 1, ... , n, X_{ji} is j^{th} observation of the i^{th} explanatory variable (i.e. land use variables), i = 1, 2, ..., p. ε_i is a random error and the β_i, i = 1, 2, ..., p are unknown regression coefficients to be estimated. We tabulate the average hourly travel demand from every origin destination pair of hexagons and combine these statistics with explanatory land use variables for model fitting. Distinct models are trained for public transportation and ride-hailing respectively for every hour of the day. All values are normalised to allow for direct comparison of regression coefficients.



a) Ride-hailing travel demand at origin

b) Ride-hailing travel demand to destination

c) Public transport travel demand at origin

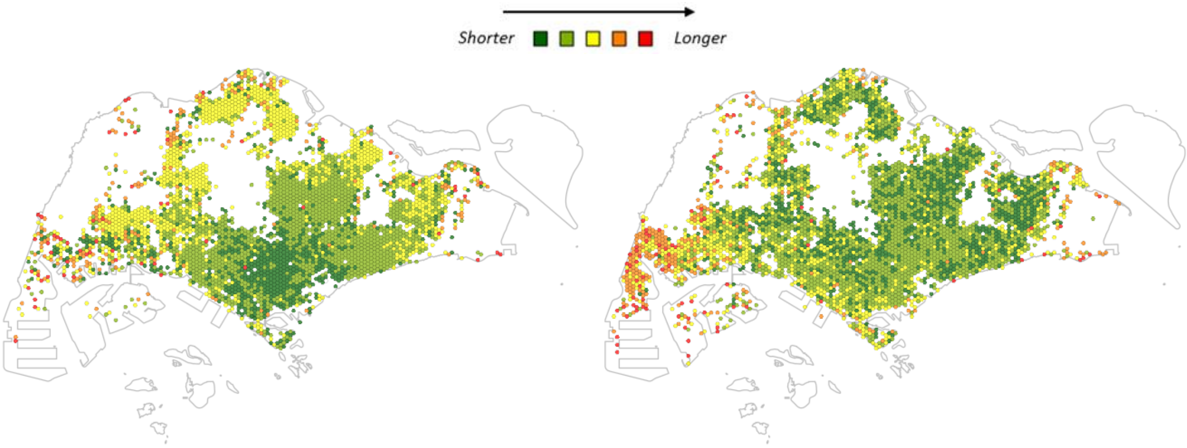
d) Public transport travel demand to destination

Figure 4. Average weekday travel demand by a-b) ride hailing and c-d) public transport at from 7 to 9 am.

6. Results

6.1 Ride-hailing Utilisation Patterns

To inform land use transport planning, city planners are interested to understand how commuters utilise ride-hailing, particularly in the morning on weekdays traditionally seen as peak travel times. Figure 4 compares the average ride hailing (Figures 4a and 4b) and public transport (Figure 4c and 4d) travel demand at origin and destination, from 7 to 9 am on an average weekday. From the maps, we can observe that ride-hailing and public transport travel demand differ substantially. While the Central Business District (CBD, in Singapore's south) attracts many commuters for both modes in general, public transport trip destinations are not solely concentrated in the CBD. As illustrated in Figure 4d, public transport trip destinations are evenly distributed throughout the island in the locality of transport hubs (see Figure 1f) where a mix of retail and hotel related work activities, as well as daily needs activities like shopping occur (see Figures 1b and 1c). The loci of public transport and ride-hailing trip origins also differ substantially. While a good share of ride-hailing trips originate near the CBD in the locality of affluent high density dwellings (see Figure 1h), public transport trip origins stem from public housing residential towns surrounding the CBD (see Figures 1i-k). Trips to and from the airport appears to be a distinctive characteristic of the way ride-hailing is used (see Figure 1d). Travel for flights and trips originating from the locality of hotels suggests active utilisation by tourists.



a) Mean travel distance from origin   b) Mean travel distance to destination

Figure 5. Average ride hailing travel distance from a) origin and the catchment at b) destinations for an average weekday between 7 - 9 am.

Closer analysis of ride-hailing travel distances reveal that trips originating from the central area tend to be shorter than those originating from the periphery (see Figure 5a), suggesting that travel occurs predominantly for office-related work activities and largely remains in the CBD. Cross-referencing travel demand patterns illustrated by Figures 4a and 4b suggests that ride-hailing is primarily utilised by an affluent segment of the population. Travel across longer distances to locations primarily associated with work-related manufacturing activities in Singapore's west is also apparent in Figure 5b, but this appears to be a relatively small subset of the overall travel demand.

Table 2. Summary of office and non-landed dwelling trip rates for public transport segmented by accessibility levels per hexagon of 104,000m^2. Most hexagons are mixed use and therefore can be substantial trip generators and attractors for accommodation and work activities respectively.

Public Transport Reach *(1 - more, 5 least)*		**7am**		**8am**		**9am**		**10am**		**11am**	
		Trip Gen.	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**
Office	**1**	0.076	0.149	0.059	0.099	0.036	0.028	0.03	0.027	0.032	0.027
	2	0.175	0.139	0.133	0.108	0.084	0.063	0.069	0.051	0.068	0.052
	3	0.069	0.054	0.062	0.047	0.039	0.031	0.031	0.026	0.03	0.026
	4	0.017	0.037	0.018	0.019	0.011	0.012	0.009	0.009	0.009	0.009
	5	0.027	0.037	0.034	0.042	0.022	0.027	0.017	0.021	0.019	0.022
Non-landed	**1**	2.953	5.812	2.296	3.857	1.423	1.502	1.188	1.04	1.26	1.044
	2	0.04	0.032	0.031	0.025	0.019	0.015	0.016	0.012	0.016	0.012
	3	0.044	0.034	0.039	0.03	0.025	0.02	0.019	0.016	0.019	0.017
	4	0.048	0.051	0.049	0.053	0.031	0.033	0.024	0.025	0.025	0.026
	5	0.129	0.172	0.159	0.195	0.103	0.126	0.081	0.099	0.087	0.103

Table 3. Summary of office and non-landed dwelling trip rates for ride-hailing segmented by accessibility levels per hexagon of 104,000m^2.

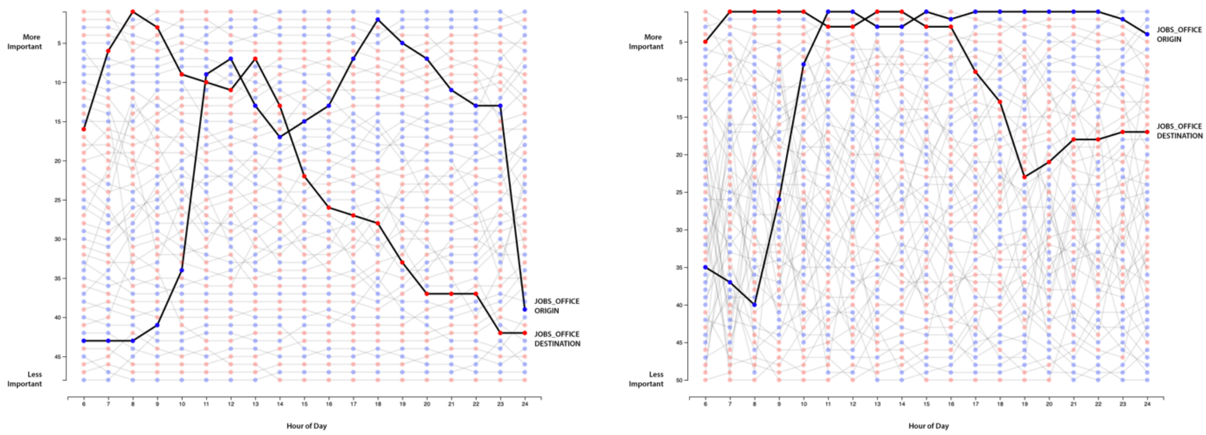
Public Transport Reach *(1 - more, 5 least)*		**7am**		**8am**		**9am**		**10am**		**11am**	
		Trip Gen.	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**	**Trip Gen.**	**Trip Attr.**
Office	**1**	0.104	0.185	0.105	0.181	0.103	0.188	0.106	0.182	0.106	0.182
	2	0.15	0.19	0.148	0.189	0.151	0.195	0.15	0.195	0.15	0.191
	3	0.03	0.029	0.03	0.029	0.003	0.029	0.03	0.029	0.031	0.029
	4	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	5	0.02	0.021	0.02	0.021	0.02	0.021	0.02	0.021	0.02	0.021
Non-landed	**1**	3.363	1.064	0.701	1.117	0.701	1.069	0.719	1.125	0.714	1.123
	2	0.033	0.039	0.033	0.039	0.033	0.04	0.033	0.039	0.033	0.039
	3	0.021	0.021	0.021	0.029	0.021	0.02	0.021	0.02	0.021	0.021
	4	0.028	0.029	0.028	0.02	0.028	0.029	0.028	0.03	0.028	0.029
	5	0.093	0.099	0.093	0.099	0.093	0.097	0.093	0.099	0.093	0.099

6.2 Land Use Drivers of Travel Demand

Land use trip rates are essential for city planners to assess urban development proposals and land use policy implications. They are also fundamental material for mobility service operators to forecast travel demand. To illustrate, we show the land use trip rates from 7 to

11 am on weekdays where commute between non-landed dwelling units and office related work occurs. Summarised in tables 2 and 3 are the trip rates for public transport and ride-hailing respectively. These trips rates indicate the number of trips departing from (i.e. trip generation) or arriving at (i.e. trip attraction) a hexagon for each office job and non-landed dwelling unit present. City planners sometimes distinguish between peak and non-peak trip rates. We further segment trip rates by hour of day and locational accessibility to demonstrate how different urban environment characteristics influence the utilisation of both modes.

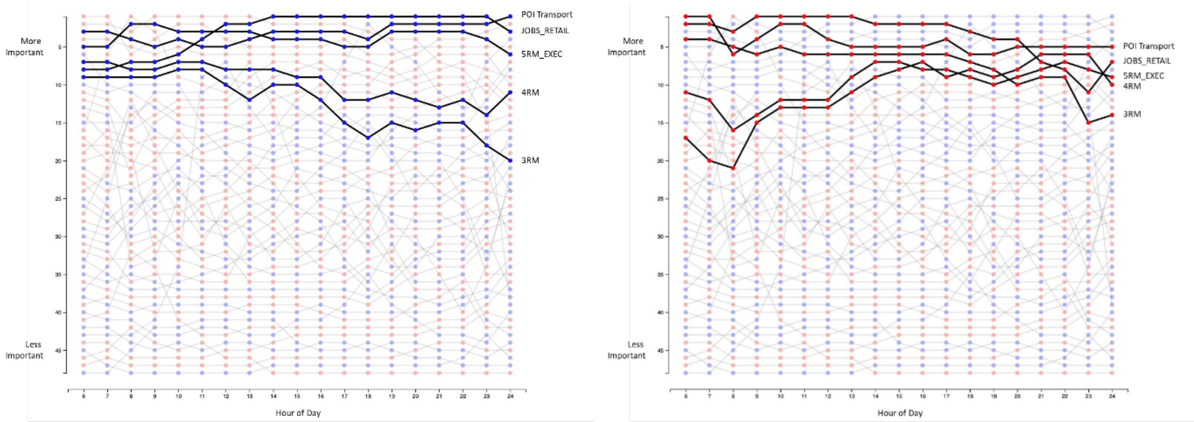
As can be observed in both tables, trip rates vary substantially by location and time of day. In general, public transport trip rates are more dynamic than ride-hailing. In table 2, we can observe that trip attraction rates to office jobs diminish as time progresses. However, the rate of change varies between locations with different levels of accessibility. As most commute to the CBD would have arrived at their destinations before 9 am, a sharp decline in trip attraction to office jobs located in hexagons scoring public transport reach of 1 and 2 can be observed between 8 and 9 am. In comparison, the rate of change is less distinctive for hexagons scoring public transport reach between 3 to 5 as these are located closer to residential centres. Ride-hailing trip attraction rates to office jobs are observably different from public transport. As shown in table 3, trip attraction remains relatively high and stable from 7 to 11 am in comparison to public transport, suggesting consistent demand for travel to office workplaces by ride hailing. There are some observable inconsistencies in both tables. Specifically, the trip rates for non-landed dwellings units in hexagons scoring public transport reach 5 is relatively large for both modes. We believe this to be primarily due to the small number of trips departing from or arriving at such locations. Similarly, the trip generation rate for office related work activities in hexagons scoring public transport reach 1 is comparably lower than hexagons scoring public transport reach 2. This is largely the result of planning initiatives to promote mixed use developments that incorporate accommodation and work related activities at highly accessible locations.



a) Public transport travel demand

b) Ride-hailing travel demand

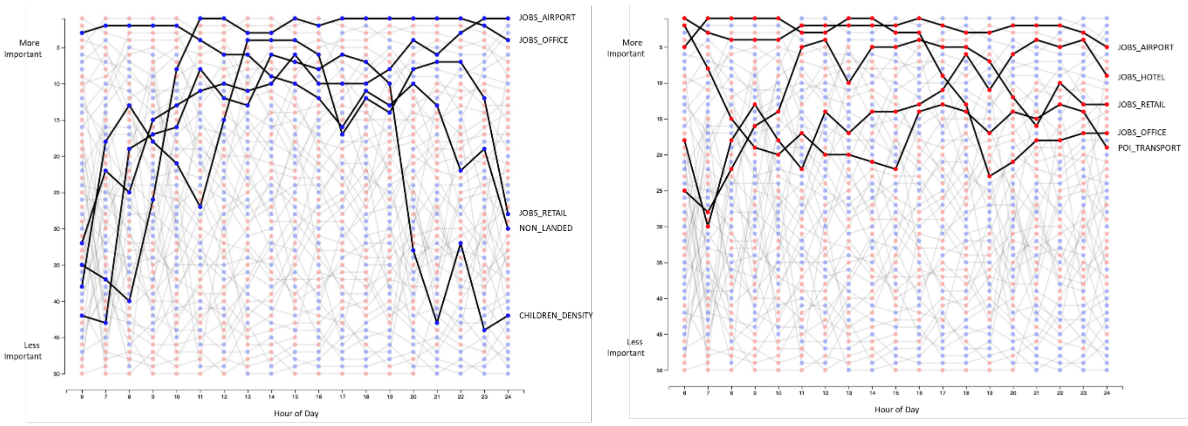
Figure 6. Relative importance of office work related activities as a driver of a) public transport and b) ride-hailing travel demand over the course of an average day. Blue dots indicate the importance of office work related activities at origin as trip generators, while red dots indicate its importance at destination as trip attractors.



a) Consistently important variables at origin

b) Consistently important variables at destination

Figure 7. Consistently important variables influencing public transport travel demand over the course of an average day at a) origin as trip generators and at b) destination as trip attractions.



a) Most important variables at origin

b) Most important variables at destination

Figure 8. Most important variables influencing ride hailing travel demand over the course of an average day at a) origin as trip generators and at b) destination as trip attractions.

Regression analysis yields coefficients that describe the influence of each land use variable on travel demand. By comparing coefficients for each hour of the day, it becomes possible to ascertain the role of each variable over time. Figure 6 depicts the relative importance of office work related activities for ride-hailing and public transport over the course of the day. As may be observed in figure 6a, office work related activities attract a large number of public transport trips in the morning, ranking as the top driver of travel demand at 8 am. While its influence as trip generator gradually diminishes in the later half of the day, its role as a trip generator grows and eventually peaks at 6 pm when most commuters begin to travel home. Similar patterns may be observed for ride-hailing (see figure 6b) in general, but the influence of office work related activities as trip generators remain high beyond 11pm as a result of the reduction in public transport services. Following this approach, we identify the top 5 drivers of travel demand for both modes. Compared to ride-hailing, the land use variables influencing public transport travel demand are far more consistent (see figures 7 and 8). Accommodation related activities, as well as shopping for daily needs ranks high as public transport trip generators and attractors at all hours of the day. In comparison, flights, tourist accommodation, office related work activities, high density affluent housing and shopping for daily needs appear as the most important variables driving ride-hailing travel demand.

7. Discussion and Conclusion

We have described a novel dataset that combines land use and travel demand data to study land use-transportation interactions in Singapore. Taking an activity-based analytical approach, we have shown how characteristics of the urban environment like land use mix, location accessibility, and peak-hour travel demand influence commute by the ride-hailing and public transport. The findings from this case study demonstrate the value of a public-private collaboration to study land use mobility patterns, and the feasibility to obtain a set of spatiotemporally sensitive land use trip rates that serves as a common understanding of land and mobility patterns in Singapore.

Joining land use information to trip origins and destinations allow city planners to infer details about prevailing travel patterns, by identifying the purpose of trips and the subset of commuters who might engage in them. Such information is useful to city planners who must characterize demand for services and infrastructure, and match them with sufficient supply. Accordingly, travel demand is better accounted for when considering both public transport and ride-hailing data in tandem, rather than either one alone. Public transport data has much larger volumes in Singapore, but does not capture true origin and final destination of trips. Ride-hailing data is comparatively smaller in volume, but more representative of actual travel demand at different locations island wide. Without visibility of how people travel the first and last mile, we are not able to fully ascertain how land use influences mobility. *POI_Transport* consistently surfaces in our analysis because the public transport network follows a hub and spoke model. Travel to transport hubs can occur for many purposes including work activities, daily needs retail activities or simply to make another trip to a destination elsewhere.

Our analytical approach provides city planners with a practical way to measure the influence of land use variables like zoning, number of jobs, dwelling units and demography on hourly travel demand. This can help city planers evaluate more thoroughly the impact of designating new or change existing land uses. More explicitly, such data can be translated into more precise trip generation tables. Because daily and hourly data are available, it is now feasible to estimate more precise trip rates, that are representative of travel demand for a given time of day and location. Different behaviours were revealed that did not seem apparent when analysing public transport data in isolation. Specific population sub-groups exhibit mobility patterns that did not follow the general trend observed in public transport like post-work leisure and entertainment. This points to the possibility that a unique subset of the population, who are more affluent, may be utilising ride-hailing services more frequently; or could indicate new lifestyle patterns and recreational activities that were previously not captured by public transport data, or are newly-enabled by ride-hailing today. This suggests that TNCs are not just meeting travel demand but also altering commuter mode choice, preferences, and travel patterns.

Acknowledgements

The authors extend their thanks to Mr Lucien W. Ong for his contribution to framing shared mobility and insights on experiments and evaluations of micro transit and other ride-hailing modes in the United States.

References

Alsger, A., Mesbah, M., Ferreira, L., & Safi, H. (2015). Public transport origin-destination estimation using smart card fare data. In *Transportation research board 94th annual meeting* (No. 15-0801).

California Public Utilities Commission, USA. *Transportation Network Companies* (2019). https://www.cpuc.ca.gov/tncinfo/ Accessed 15.07.19.

Cohen, A., & Shaheen, S. (2018). *Planning for shared mobility*.

Chua, A., & Wang, S. Y., *Journey Estimation with Smartcard Data for Land Use Planning* (2018). http://www.2018.csdm-asia.net/IMG/pdf/Journey_Estimation_with_Smartcard_Data_for_Land_Use_Planning.pdf Accessed 15.07.19

Kirby, R. F., Bhatt, K. U., Kemp, M. A., McGillivary, R. G., & Wohl, M. (1974). *Para Transit: Neglected Options for Urban Mobility* (No. UMTA-CA06-0045-74-2).

Mageean, J., & Nelson, J. D. (2003). The evaluation of demand responsive transport services in Europe. *Journal of Transport Geography*, *11*(4), 255-270.

Land Transport Authority, Public Consultation on the Land Transport Master Plan 2040 (2019). https://www.mot.gov.sg/ltmp2040/ Accessed 15.07.19.

Shaheen, S. A., Cohen, A. P., & Chung, M. S. (2009). North American carsharing: 10-year retrospective. *Transportation Research Record*, *2110*(1), 35-44.

Statista. Ride Hailing in Singapore (2019). https://www.statista.com/outlook/368/124/ride-hailing/singapore Accessed 15.07.19.

Yan, X., & Su, X. (2009). Linear regression analysis: theory and computing. World Scientific.