Joint modeling of arrivals and parking durations for freight loading zones: Potential applications to improving urban logistics

Citation for the original published paper (version of record):
Kumar Kalahasthi, L., Sanchez-Diaz, I., Castrellon, J. et al (2022). Joint modeling of arrivals and parking durations for freight loading zones: Potential
applications to improving urban logistics. Transportation Research Part A: Policy and Practice, 166: 307-329. http://dx.doi.org/10.1016/j.tra.2022.11.003

(article starts on next page)

Transportation Research Part A 166 (2022) 307–329

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Joint modeling of arrivals and parking durations for freight loading zones: Potential applications to improving urban logistics

Lokesh Kumar Kalahasthi [a,*], Iván Sánchez-Díaz [a], Juan Pablo Castrellon [a,b], Jorge Gil [c], Michael Browne [d], Simon Hayes [e], Carles Sentís Ros [e]

[a] Department of Technology Management and Economics, Chalmers University of Technology, Vera Sandbergs Allé 8, Gothenburg 412 96, Sweden
[b] Department of Systems and Industrial Engineering, Universidad Nacional de Colombia, Bogotá 111321, Colombia
[c] Urban Design and Planning, Department of Architecture and Civil Engineering, Chalmers University of Technology, Sven Hultins gata 6 Gothenburg, 412 96, Sweden
[d] Urban Freight Transport and Logistics, Department of Business Administration, University of Gothenburg, Vasagatan 1, Gothenburg 411 24, Sweden
[e] PARKUNLOAD, Edifici NEC, Carretera de Ribes 4, 1° 08520 Les Franqueses del Vallès, Barcelona, Spain

ARTICLE INFO

Keywords:
Freight Parking Management
Truck Parking Zones
Urban Freight
Joint Modeling of Counts and Survival

ABSTRACT

This paper analyzes truck parking patterns in urban freight loading zones by jointly modeling the vehicle arrival rates and the parking durations. Three models were explored: 1) Count data (Negative Binomial) for vehicle arrivals, 2) Survival (Weibull) model for parking duration and 3) A joint model for arrivals and duration. The count data model estimates the parking demand i.e., the rate of truck arrival, while the survival model estimates the probability that a truck is parked for one more minute. The joint model is compared with separate models for predictability and performance. The dataset used in this research is obtained using a mobile phone parking application, at eight loading zones in the city Vic, Spain over an 18-month period from July 2018 to December 2019, comprised of vehicle parking durations, date, time of arrival and departure, professional activity, and vehicle type (weight). The parking activity data are complemented with built in environment variables of the loading zones, such as the number of establishments in a certain radius, the average walking distance to establishments, the presence of pedestrian pavement, the number of traffic lanes, among others. The joint model outperforms the models estimating the arrival rates and durations separately in goodness of fit and predictability. The model results showed that truck arrival rates vary significantly across days of the week, months, and arrival times. The parking durations are highly dependent on professional activity, vehicle type, and size. Tuesdays and Wednesdays have higher arrival rates compared to other days of a week (except Sundays). Among activities, the transport and parcels require longer parking durations. Among the vehicle types, trucks with gross weight larger than 3.5 tons park longer. This paper concludes by explaining the potential of these modeling approaches in improving urban freight operations, evaluation of various policy implications, limitations, and future research.

* Corresponding author.
E-mail addresses: klokesh@chalmers.se (L.K. Kalahasthi), ivan.sanchez@chalmers.se (I. Sánchez-Díaz), juanpabl@chalmers.se (J. Pablo Castrellon), jorge.gil@chalmers.se (J. Gil), michael.browne@gu.se (M. Browne), shayes@parkunload.com (S. Hayes), csentis@parkunload.com (C. Sentís Ros).

https://doi.org/10.1016/j.tra.2022.11.003

Available online 13 November 2022

1. Introduction

European cities concentrate nearly 80 % of the EU's population and generate 85 % of the gross domestic product (Bolay 2020). The rapid urbanization, the economic development, and the higher standards for quality of life have resulted in many transportation challenges for those cities. Urban planners and policy makers have to find a path to maintain the positive aspects of those developments while diminishing the negative aspects related to transportation, such as, traffic congestion, poor air quality, fossil energy consumption, noise and safety issues among other unsustainable impacts (Lindholm and Behrends 2012). Urban freight transport (UFT) has a major share in these externalities. For instance, UFT accounted for 15–25 % of total vehicle kilometers travelled and 26 % of total CO_2 emissions from the transportation sector in France (Schoemaker et al. 2006).

Planning and managing UFT are crucial to ensure a sustainable economic development in cities. Every day, thousands of freight vehicles drive into cities to supply businesses with goods that allow them to carry their economic activities and ensure city dwellers can fulfill their needs. Although it is undisputed that cities require those goods, there is a pressing need to design a system in which these goods can be delivered in a sustainable way. Several policy measures and strategies have been recommended by researchers to achieve sustainable UFT. Some of the latter focus on improving the supply side such as building new infrastructure, enhancing parking space and operation; some others focus on logistics interventions such as fostering consolidation centers; and others focus on freight demand management strategies such as pricing, off-hour deliveries and receiver-led consolidation programs (Allen et al. 2000, Muñuzuri et al. 2005, Holguín-Veras et al. 2015). Predominant UFT policies in Europe include low emissions zones which restrict poor environmental standard vehicles from accessing urban centers (Tögel and Špička 2014), urban consolidation centers aimed at mitigating freight traffic impacts by replacing partially loaded trucks with fully loaded ones (Allen et al. 2012), access restrictions to urban centers for certain vehicle sizes, time windows for deliveries (CLARS 2020), and land-use policies envisioned to incorporate freight operations in urban infrastructure development projects (Transmodal Limited 2012).

Freight parking management is a vital issue to consider while designing a land-use policy, that required further investigation into the possibilities to develop solutions to minimize negative externalities by UFT (Marsden 2006, Kladeftiras and Antoniou 2013). Despite the intense competition for space in urban environments, every municipal authority needs to provide public space for freight vehicles to park, load and unload their goods. Parking could be on-street along the curb or off-street reserved for freight vehicles. In addition, there are on-street loading zones (LZs) defined as the areas along the curbside of streets specifically reserved for loading and unloading of bulk cargo (Allen et al. 2007). Most LZs are designated using road marking and vertical signals, and lately a few started to include connectivity features mainly via smartphones apps. As documented in the literature, parking is a key issue for freight vehicles. A typical freight vehicle spends about 50 % of its daily operational time parked (Sanchez-Diaz et al. 2020), and the availability of a LZ is a key challenge for freight vehicles, which often double park and incur large parking fines (Jaller et al. 2013, Marcucci et al. 2015, Dey et al. 2019). The literature suggests different ways to approach this challenge. BESTUFS (2007) emphasizes the importance of developing connected electronic devices, and smart technologies to manage fleet operations at parking/loading zones. Dey et al. (2019) recommends increased enforcement at LZs to mitigate unauthorized parking and emphasizes the need for data-driven modeling approaches to estimate parking demand. The European Commission (2020) recognizes the importance of parking, and includes truck and commercial vehicles parking as one of the top priorities in their Intelligent Transport Systems (ITS) action plan and directive, with three major objectives: 1) truck drivers need to have appropriate information on safe and secure parking places, 2) improve capacity of truck parking areas, and 3) optimize existing parking capacity by digital information.

Although parking is recognized as one of the key challenges by drivers, transport managers and public authorities (Sanchez-Diaz et al. 2020), transport companies seldom consider the availability of LZ in their route plan, and public authorities plan LZs based on establishments' requests, on intuition or based on space availability. The main reason is that parking data are rarely available, and the development of parking demand predictive models to support decision-making is at a very early stage. The purpose of this paper is twofold, it seeks to shed light on the dynamics of urban freight parking demand, and it proposes a method to predict arrivals of vehicles and their parking durations at LZs. A novel model for joint estimation of arrival rates (count data) and parking duration (survival model) is explored and compared with the performance of separate models. To do so, the authors use historical parking data from a mobile phone app-based system (Parkunload 2020) to calibrate a set of statistical models (i.e., count data and duration models). This paper also proposes some practical implications that highlight how these data-driven approaches may lead to a smarter use of urban curb space and reduce the negative externalities of UFT.

The rest of the paper is organized as follows. Section 2 provides a comprehensive review of the relevant literature. Section 3 presents data description. Section 4 describes methodology. Section 5 presents the modeling results. Section 6 explains policy implications. Section 7 provides the concluding remarks.

2. Literature review

This section provides a comprehensive review of relevant literature. The literature on freight parking is rather limited. This review identified three subjects to classify the literature based on scope and methodology. The first subject is parking demand (arrival and duration) estimation, the second subject deals with studying parking behavior or simulation modeling, and the last category analyzes parking policies.

2.1. Parking demand and duration

Estimation of freight parking demand is studied predominantly using econometric approaches. For instance, Wang and Garber

Table 1
Summary of Parking Studies (Chronological).

No.	Publication	Dem-and	Dura-tion	Simu-lation	Data Source	Modeling Approach
1	Wang and Garber (2003)	X	–	–	Drivers/firms survey	Regression
2	Benenson et al. (2008)	X	–	X	Parking zone user survey	ABM
3	Chatterjee et al. (2008)	–	X	–	Stakeholders survey	Descriptive
4	Dieussaert et al. (2009)	–	–	X	Secondary sources	ABM
5	Jones et al. (2009)	X	–	–	Cordon and freight survey	Travel time study
6	Ibeas et al. (2011)	X	–	–	Parking payments	Regression
7	Habib et al. (2012)	–	X	–	Origin-Destination survey	Regression
8	Kladeftiras (2013)	–	–	X	Field visit	Microsimulation
9	André and João (2014)	X	–	–	Establishment survey	Regression
10	Nourinejad et al. (2014)	X	–	X	Driver survey	Binary logit
11	Marcucci et al. (2015)	X	–	–	Secondary sources	Multinomial logit
12	Vlahogianni et al. (2016)	–	X	–	Driver/carrier survey	Survival models
13	Roca-Riu et al. (2017)	–	–	X	NA	Microsimulation
14	Malik et. al (2017)	X	–	–	Cordon and freight survey	Descriptive
15	A. R. Alho et al. (2018)	X	–	X	Establishment survey	Microsimulation
16	Campbell et al. (2018)	X	–	–	Establishment survey	Regression models
17	Schmid et al. (2018)	–	X	–	Driver/carrier survey	Survival models
18	Dey et. al (2019)	X	–	–	LZ data	Descriptive analysis
19	Dalla Chiara et al. (2020)	X	–	X	Cordon and freight survey	Discrete choice model
20	Low and Cheah (2020)	–	X	–	Cordon and freight survey	Survival models
21	Sadek & Shaheen (2020)	X	–	–	Truck parking data	Fourier transformation
Current paper		**X**	**X**	**–**	**LZ data**	**Joint Model for Count Data and Survival**

Note: X = Check mark, NA = Not Applicable, - = Nil/Empty.

(2003) developed linear regression models to predict commercial vehicle parking demand on interstate highways, and concluded that the traffic volume and the distance between the parking zones are the major factors affecting demand. Ibeas et al. (2011) introduced geographically weighted regression models to better predict parking demand with results that outperform multilinear regression models. These findings gave insights about the importance of considering spatial heterogeneity in planning and predicting parking needs. Sadek et al. (2020) proposed a different tool from econometric models for predicting parking demand i.e., Fourier Transformations. Inputs for this model were one year trucking occupancy history for a dynamic forecasting model that showed up to 5 % estimation error. Campbell et al. (2018) proposed the use of Freight Trip Attraction (FTA) and Service Trip Attraction (STA) models to assess parking needs and added to the analysis Transport Demand Management (TDM) strategies to evaluate their impact on parking demand. The Campbell et al. (2018) study emphasized the need to conduct future research with quality data to achieve higher levels of accuracy. Alho and de Abreu e Silva (2014) found that employment and floor area are significant factors in estimating parking demand for retail establishments and, they expanded the analysis in Alho et al. (2018) with the use of these demand estimations to allocate loading/unloading bays. Butrina et al. (2017) developed a framework for analyzing the last 800 feet in freight delivery describing the delivery and facilities (location) characteristics that influence performance, and demand. Malik et al. (2017) analyzed reasons for imbalances between supply and demand in LZs. The study recommended the inclusion of factors such as vehicle characteristics, vehicle ownership, freight drivers paying parking fee or not, commodity type and, duration of goods delivery.

Parking duration for freight is the least studied among the subjects identified in this review. Different approaches have concluded the importance of considering travel activity-type to estimate duration and behavior. Vlahogianni et al. (2016) and Schmid et al. (2018) built survival models to identify the most significant factors that influence parking duration. Conclusions of this research suggest that vehicle and delivery characteristics are the most important aspects that determine parking duration. Low et al. (2020) found that in addition to Schmid et. al, parking location was also a significant factor influencing parking duration. These studies highlighted the need for future research efforts in understanding the parking duration patterns for commercial vehicles.

2.2. *Parking behavior and simulations*

Parking simulation integrates the demand estimation, with secondary data such as traffic counts, network data, to assist the evaluation of driver behavior, future infrastructure developments, and policy impacts. For instance, Benenson et al. (2008) simulated drivers' behavior under different spatial urban environments using agent-based modeling (ABM). Outputs from the model provided insights about the search time and distance to destination depending on spatial characteristics. Using the same paradigm (ABM), Dieussaert et al. (2009) proposed a simulation tool for analyzing the impacts of potential policy scenarios on the traffic generated while searching for a parking space. For analyzing infrastructure availability, Roca-Riu et al. (2017) designed a conceptual modeling approach for simulating dynamic delivery parking spots (DDPS). The DDPS model incorporates the variability of parking supply over time depending on traffic demand.

Regarding policy and impacts evaluation, Kladeftiras and Antoniou (2013) ran a microsimulation to evaluate the impact of double-parking on the environment. Results showed that effective enforcement strategies coupled with an appropriate driver behavior may lead to 44 % increase in the average speeds and significant reductions in air pollutant emissions. Nourinejad et al. (2014) developed a microsimulation model for urban parking operations, including freight and passengers, expanded the analysis to impact variables such

Fig. 1. Location of Loading Zones (LZ) in Vic, Spain.

as walking distance, congestion impacts, and parking search times. Alho et al. (2018) used simulation as a tool to represent and quantify double-parking effects in the evaluation of policies related to loading/unloading bays allocation. Dalla Chiara et al. (2020) proposed a combined parking-choice and simulation model to study parking choice behavior of commercial vehicle drivers. The simulation tested impacts of different parking management strategies and, found that reducing parking durations and providing incentives for light goods and service vehicles to park in larger carparks are effective in reducing double-parking and queues.

2.3. Parking policy analysis

Considering strategies to encourage parking zones' turnover and decrease vehicle-miles-traveled (VMT), the Cambridge Systematics Inc. (2007) deployed a set of measures that public agencies can implement based on experiences from different US cities. In a similar study, Marsden (2006) reviewed drivers response to different parking policies and confirmed how sensitive they are mainly to out-of-pocket costs (fees and walk time). Chatterjee et al. (2008) conducted a qualitative research of this nature with a special focus on freight parking, comparing initiatives implemented by local authorities at the city level. Stakeholders perceptions differ from those in Marsden (2006) and Cambridge Systematics Inc. (2007) with safety, traffic operational and planning issues found as the relevant factors in urban parking operations and policies. In terms of projects implementation for satisfying parking demand, Jones et al. (2009) quantified the positive impacts of various strategies on truck travel times such as reallocation of curb space, longer loading zones, metered loading zones and enhanced parking enforcement. Marcucci et al. (2015) complemented and extended previous research by analyzing some of the mentioned strategies from transport providers' perspective. Malik et al. (2017) confirmed previous approaches suggesting that successful parking policies implementations have to consider actions and factors related to street design, level of enforcement, parking fee, and stakeholders involved.

2.4. Summary of literature review

Table 1 summarizes the relevant literature on freight parking. As mentioned, parking demand is widely studied followed by simulation, and policy analysis. From Table 1, it is evident that there is very limited research on parking durations. Also, there is no study that combines demand estimation and survival modeling. The existing studies on parking duration are either based on driver or establishment surveys which are prone to have data collection errors as the actual time spent is not measured (Schmid et al. 2018) or suffer from missing data (Low et al. 2020).

To the best of the authors' knowledge Dey et al. (2019) is the only research that provided a descriptive analysis of historical parking data, the data was collected from LZs in Washington D.C., USA. The current research contributes to the literature by estimating econometric models using the historical parking data from LZs from Vic, Spain, measured by an electronic device and an cellphone app-based system that records every parking operation (Parkunload 2020). Also, there is no research on exploring a joint model in estimating freight parking demand and duration, which constitutes another key contribution of this paper.

3. Data and descriptive analysis

The modeling in this paper is based on the historical parking data for eight Loading Zones (LZs) in the city of Vic (Catalonia), Spain as shown in Fig. 1. Vic is the capital city of the comarca of Osona, in the Barcelona province of Spain with a population about 46,000, employment 15,000 and the area of 30 square kilometers. Major industry sectors are agriculture, construction and service (Statistical Institute of Catalonia 2020). The data were collected using a parking device developed for Vic Municipality by a technology company called ParkUnload as described at its corporate web-site (Parkunload 2020). Parkunload (2020) is a company providing technological solutions for urban goods distribution, last mile challenges, and a great source for big data in logistics. ParkUnload provides the truck drivers with a mobile phone-based app (or a Bluetooth button) which they turn on at the beginning of parking and turn off (check out) when leaving the LZs. The maximum allowable parking duration is 30 min, which was often violated prior to 2018 when the system was implemented. The ParkUnload solution enables the LZs to be monitored by the public agency for any parking violations. The distance between LZs vary from 40 m (LZ 2 to LZ 3) to 200 m (LZ 1 to LZ 2). Also, the LZs have different capacities and operating

Table 2
Seasonality Index of Arrival Rates (Vehicles per Hour).

Vairable		LZ1	LZ2	LZ3	LZ4	LZ5	LZ6	LZ7	LZ8	Total Sample
1. Arrival Time	06:00					0.39			0.47	0.57
	07:00					1.06			1.31	1.58
	08:00	2.10	1.91	1.79	1.76	1.76	1.75	1.70	1.80	1.75
	09:00	2.09	2.13	1.92	2.09	2.17	2.04	2.01	1.90	2.00
	10:00	1.73	1.95	1.83	1.97	1.95	1.85	1.74	1.76	1.81
	11:00	0.96	1.14	1.16	1.06	0.99	1.08	0.94	1.09	1.01
	12:00	0.42	0.42	0.47	0.42	0.37	0.46	0.39	0.47	0.41
	13:00	0.47	0.46	0.54	0.41	0.35	0.40	0.41	0.53	0.42
	14:00	0.78	0.77	0.86	0.66	0.80	0.75	0.81	0.92	0.77
	15:00	0.92	0.84	1.03	0.96	1.20	1.05	1.11	1.05	1.02
	16:00	0.66	0.72	0.81	0.90	0.98	0.89	1.03	0.95	0.88
	17:00	0.59	0.50	0.47	0.58	0.73	0.55	0.68	0.60	0.60
	18:00	0.27	0.16	0.13	0.19	0.25	0.18	0.18	0.15	0.19
Avg AR		1.27	1.70	1.07	3.14	3.67	1.58	2.94	2.60	2.31
2. Day of Week	Mon	0.69	0.98	0.91	0.87	0.91	0.87	0.85	0.84	0.87
	Tues	1.13	1.09	1.14	1.19	1.12	1.10	1.05	1.07	1.11
	Wed	1.14	1.04	1.01	1.00	1.03	1.07	1.04	1.04	1.04
	Thur	1.00	0.97	0.99	0.95	1.00	1.01	1.06	1.03	1.00
	Fri	1.03	0.92	0.94	1.00	0.93	0.94	1.00	1.03	0.97
	Sat	0.34	0.15	0.21	0.31	0.25	0.30	0.28	0.23	0.27
Avg AR		1.32	1.79	1.12	3.26	3.90	1.64	3.07	2.78	2.40
3. Month	Jan	0.97	1.13	1.16	1.19	1.01	1.22	1.13	1.13	1.11
	Feb	1.23	1.05	1.18	1.23	1.05	1.19	1.20	1.10	1.14
	Mar	1.24	1.11	1.12	1.24	1.07	1.12	1.10	1.10	1.13
	April	1.05	1.05	1.05	1.01	1.09	0.95	0.98	1.09	1.04
	May	1.01	1.02	0.99	1.07	1.05	1.03	0.97	1.17	1.05
	June	0.99	0.81	0.91	0.92	1.07	0.81	0.91	0.98	0.94
	July	0.71	0.75	0.75	0.72	0.90	0.77	0.82	0.73	0.79
	Sept	0.64	0.59	0.61	0.68	0.74	0.67	0.71	0.68	0.68
	Aug	0.90	0.94	0.92	0.85	0.97	0.89	0.96	0.90	0.92
	Oct	0.96	1.05	1.02	0.94	1.02	1.00	1.01	1.08	1.01
	Nov	1.20	1.21	1.17	1.07	1.03	1.20	1.16	1.07	1.11
	Dec	1.09	1.28	1.14	1.08	0.99	1.15	1.05	0.97	1.06
Avg AR		1.21	1.58	1.00	2.99	3.48	1.48	2.74	2.50	2.16

Note: The shaded cells represent *SI* greater than one.

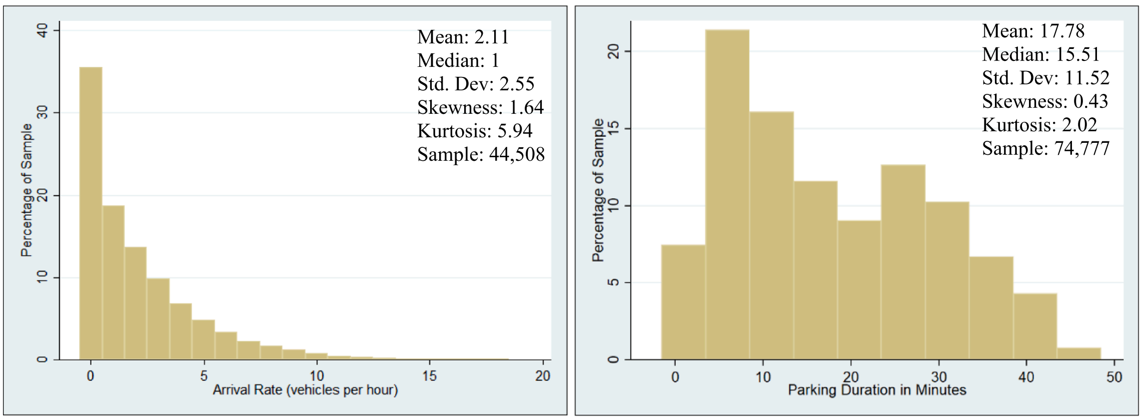

Fig. 2. Distribution of Arrival Rates and Duration Data.

timings can also vary from zone to zone. LZ 3 has the lowest capacity with two parking spots and 12 m, followed by LZ 1 and 6 with four spots and 20–23 m, LZ 2 has six spots and 46 m, and LZ 7 has 7 spots and 43 m in length. The LZ with the largest capacity are LZ 4, LZ 5 and LZ 8 that have 8 spots and 46–49 m in length. LZ 1 and 4 operate for 14 h from 6 AM to 8 PM and remaining six LZs operate for 12 h from 8 AM to 8 PM, a day. All LZs function six days a week, Monday through Saturday. The data were not collected on Sundays as the vehicles are free to park on Sundays for as long as they require.

The data were collected for all vehicles parked in these 8 LZs, for a total of 18 months: six months in 2018 from July to December (pilot) and the whole year of 2019 (first year of full system operation). The full (uncleaned) data comprises of 103,967 records (31,026 in 2018 and 72,941 in 2019) of vehicle parking including start time, end time, professional activity, and vehicle type. There are 6,412 vehicles with unique registration numbers parked in these 18 months (2,956 in 2018 and 5,233 in 2019). Around 15.6 % (4,857) in 2018 and 21.3 % (15,558) records have the device closed automatically as the drivers forgot to check out. The parking data are divided into two parts:

1. Vehicle Parking Duration Data (at vehicle level): This dataset comprises of all vehicles parked during the 18 months, with the predominant variables shown in Table 2. The observations with either the drivers overlooked to turn off the parking device or parked for less than a minute were not included in the analysis. This led to 74,777 observations as shown in Table 2. However, the count rate data explained below includes all vehicles.
2. Vehicle Arrival Rate Data (at LZ level): The vehicle duration dataset is aggregated for number of vehicles arriving in each LZ in each hour of the day. For example, if a vehicle arrived at 7:30 AM or 7:59 AM, it is assumed that the vehicle arrived between 7 and 8 AM. The vehicle durations are not captured in the dataset i.e., if the same vehicle stays beyond 8 AM, it is not included in the interval 8–9 AM. The dataset comprised of 44,508 observations as shown in Fig. 2.

The arrival rate data are merged with location data to account for these interactions of freight activity, urban infrastructure, and economic activity on parking arrival rates. The location data consists of a series of variables associated with each LZ that are calculated or extracted from different geospatial data sources. These location variables include:

1. LZ geometry details: The length and width of the LZ in meters, digitized manually over the 50 cm resolution orthophoto from the geodata services portal of the Spanish Spatial Data Infrastructure (IDEE in Spanish).
2. Infrastructure characteristics: This includes the road type (i.e. primary, secondary, slow, no access), number of lanes, presence of cycle lanes and presence of pedestrian pavement (binary). The variables were extracted from the OpenStreetMap (OSM) street network data and assigned to the nearest LZ and were complemented with observation of the orthophoto.
3. Economic activity: This includes the number of establishments (retail and hospitality) within a given walking distance from the LZ (50 m, 100 m, and 200 m), the average distance (m) to and total area (m^2) of those establishments, and the distance (m) to the nearest establishment. The location of establishments is extracted from OSM points of interest data and complemented by manually digitizing establishments from Google maps. The area of each establishment is calculated by dividing the area of the building footprint from OSM where the establishment is located by the number of establishments in the same building.

The parking data could also be studied in the perspective of occupancy of LZs i.e., the empty parking spots in the LZs. This aspect is covered in the working paper Regal Ludowieg et al. (2020), which used various econometrics and machine learning techniques to estimate the probability that at least one parking spot is empty in a given LZ at a given time. Fig. 2 shows the distribution, mean, median, standard deviation (Std. Dev.), skewness and kurtosis values of both datasets. The arrival rates data are right skewed, varying from one to 19 vehicles per hour, with a high peak evident from the kurtosis being almost equal to six. The figure also shows the overdispersion (variance to mean ratio = 3.08) because of which the negative binomial model was found to provide a better fit than the Poisson distribution, since the latter does not allow the variance to be adjusted independent of the mean. More than 50 % of the sample has arrival rate less than two vehicle per hour, median being just one vehicle per hour. The duration data are flatter, varying from one minute to 45 min with a moderate skewness (0.43, less than one) and kurtosis (2.02, less than 3). The coefficient of variation is around 65 % shows more dispersion around the mean. Weibull distribution with an appropriate scale parameter is found to provide the better fit. Around 10 % of the observations violated the 30-minute limit for maximum allowable parking time.

Table 2 presents the seasonality indices (*SI*) for each LZ and the entire dataset (all LZs combined) with respect to three variables: 1) Arrival time in hours from 06.00 to 18.00, 2) Day of Week from Monday to Saturday, 3) Month from January to December. The *SI* is defined as the ratio between the average AR for each variable (arrival time, day of the week, month) and the aggregate average of ARs as shown in Equation (1) below. The average arrival rate (Avg AR) is also shown in the Table 2. The Avg. AR is defined as the average of ARs at all levels of a variable. For example, the Avg AR for the variable month is equal to the sum of ARs for each month divided by twelve (*N*, i.e., total number months in a year). Hence, the *SI* larger than one (highlighted in Table 2) represent the months, day of weeks that have larger ARs than the average.

$$SI = \frac{AR_i}{AvgAR} \forall i \in arrivaltime/dayofweek/month \tag{1}$$

Where,

Table 3
Summary Statistics of Parking Duration (in minutes).

Variable		2018 (6 months)			2019 (whole year)			Total Sample		
		Mean	Std. Dev	Obs	Mean	Std. Dev	Obs	Mean	Std. Dev	Obs
Loading Zones (LZ)	LZ 1	17.54	11.30	1,403	18.63	11.82	3,290	18.30	11.68	4,693
	LZ 2	17.52	11.31	2,149	17.43	11.37	4,314	17.46	11.35	6,463
	LZ 3	17.20	11.51	1,279	18.04	11.74	2,740	17.77	11.68	4,019
	LZ 4	16.69	11.08	3,535	17.50	11.26	8,862	17.27	11.22	12,397
	LZ 5	17.38	11.36	5,273	18.36	11.56	11,759	18.06	11.51	17,032
	LZ 6	18.19	11.60	1,915	17.91	11.65	4,259	18.00	11.64	6,174
	LZ 7	17.39	11.54	3,825	17.91	11.60	7,858	17.74	11.58	11,683
	LZ 8	17.51	11.70	3,955	17.93	11.62	8,361	17.80	11.65	12,316
Day of Week	Monday	17.39	11.51	4,361	18.07	11.72	8,517	17.84	11.65	12,878
	Tuesday	17.55	11.50	4,100	18.25	11.67	9,622	18.04	11.62	13,722
	Wednesday	17.32	11.51	4,571	17.81	11.51	10,093	17.66	11.51	14,664
	Thursday	17.56	11.26	4,819	18.24	11.44	11,426	18.04	11.39	16,245
	Friday	16.82	11.35	4,403	17.21	11.40	9,255	17.08	11.39	13,658
	Saturday	18.38	11.54	1,080	18.53	11.58	2,530	18.48	11.57	3,610
Arrival Time	6:00	17.63	12.06	304	18.49	11.29	685	18.23	11.53	989
	7:00	17.27	11.52	903	17.82	11.80	1,863	17.64	11.71	2,766
	8:00	16.54	11.34	3,689	17.19	11.46	8,367	16.99	11.42	12,056
	9:00	16.52	11.16	4,219	17.08	11.35	9,487	16.91	11.29	13,706
	10:00	16.95	11.37	3,710	17.44	11.41	8,292	17.29	11.40	12,002
	11:00	17.38	11.26	2,027	17.84	11.63	4,398	17.70	11.51	6,425
	12:00	19.75	11.83	769	20.05	11.77	1,624	19.96	11.79	2,393
	13:00	19.55	11.64	939	21.02	11.31	1,719	20.50	11.45	2,658
	14:00	18.00	11.54	1,702	19.18	11.70	3,428	18.79	11.66	5,130
	15:00	17.78	11.47	2,099	18.81	11.54	4,579	18.48	11.53	6,678
	16:00	18.20	11.34	1,746	18.79	11.63	4,004	18.61	11.54	5,750
	17:00	18.38	11.96	970	18.05	11.71	2,363	18.14	11.78	3,333
	18:00	17.82	11.04	257	16.77	11.26	634	17.08	11.21	891
Activity	Unspecified	22.18	11.59	4,217	22.39	11.57	10,075	22.33	11.58	14,292
	Install & Maintenance	14.12	10.32	6,296	15.01	10.68	11,912	14.71	10.57	18,208
	Transport & parcels	22.74	11.49	1,593	21.95	11.56	3,943	22.18	11.54	5,536
	Construction	18.94	11.25	1,972	19.08	11.71	4,885	19.04	11.58	6,857
	Local commerce	20.11	11.07	2,018	20.18	11.25	4,469	20.15	11.19	6,487
	Comercial Agent	14.72	10.32	4,330	14.92	10.45	9,022	14.86	10.40	13,352
	Food and Markets	13.23	9.99	107	17.50	10.70	252	16.23	10.66	359
	Automotive	16.17	11.87	1,836	15.93	11.57	4,393	16.00	11.65	6,229
	NA*	14.54	10.37	965	16.25	10.80	2,492	15.77	10.71	3,457
Vehicle type	Car	13.23	10.13	586	14.99	10.50	1,531	14.50	10.43	2,117
	Truck less than 3.5 T	17.05	11.42	14,405	17.48	11.54	31,402	17.34	11.51	45,807
	Truck 3.5 T to 12 T	15.43	10.24	1,401	16.83	10.73	3,086	16.39	10.60	4,487
	Truck greater than 12 T	15.45	9.63	216	20.00	10.24	399	18.40	10.25	615
	Van	18.92	11.63	6,726	19.44	11.68	15,025	19.28	11.67	21,751
Total Sample		**17.38**	**11.43**	**23,334**	**17.96**	**11.55**	**51,443**	**17.78**	**11.52**	**74,777**

Note: NA* = Not Applicable, Chi2 test for means vary significantly across all variable groups.

$$Avg.AR = \left(\sum_{i}^{N} AR_i\right) \Big/ N$$

LZ 4 and 5 have the highest Avg. ARs, while LZ 1 and 3 have the least (see Table 2). 8:00–10:00 is the peak hour with *SI* nearly-two. The arrivals rates vary drastically from morning to evening. Tuesdays and Wednesdays have higher ARs while Mondays and Saturdays have the least. Compared to arrival time and day of the week, the *SI* across months is not varied significantly, between 0.7 and 1.1. The busiest months are January, February, March, and November with *SI* about 1.1. July, September are the quieter months. Pearson Chi2 test for independence shows the strong variation in *SI* across arrival time and day of week with Cramer's V equals 0.19 and 0.14 respectively (Pearson 1900, Cramér 1946).

Table 3 shows the means, standard deviation, and sample size for duration in minutes across LZs, day of the week, professional activities, arrival time, and truck type by weight for 2018, 2019 and total sample. The chi-squared test (not shown in the table) for equality of means across different variable groups in Table 3 show significant variation across all the variables (James 1954, StataCorp 2007). LZ 1 has the highest duration, LZ 4 has the least. In days of week, Fridays have the least mean duration, and Saturdays have the highest. Across all LZs, year 2019 has slightly higher mean duration compared to 2018. Transport and parcels take longer than other sectors. The activity food and markets has the fastest deliveries. Vehicles arriving after noon park longer. Among the vehicle types, around 3 % of the sample is commercial cars which were parked for small deliveries and services. Other vehicle types are divided based on the 'N' category (N1, N2, and N3) as classified by the European commission (European Commission, 2020). The N1 category which has less than 3.5 tons of maximum mass is further divided into heavy-duty vehicles (Truck less than 3.5 T), and light-duty vehicles (Van). Trucks with gross weight between 3.5- and 12-tons park for lesser durations compared to other trucks and vans. Vans park longer than other vehicles. Overall, the table shows the potential variables that could affect the parking durations and are considered in the modeling in this paper.

4. Model specification

The commercial vehicle parking at eight LZs in Vic, Spain is analyzed for arrival rates and parking duration using the count data and the survival models respectively. Three models were estimated: 1) Arrival rates using count data models, 2) Parking duration using survival models, and 3) A novel model to jointly estimate the arrival rates and parking durations by combining the count data and duration models. The separate models for arrival rates and durations were compared with the joint model for the model performance. The model specifications are explained below.

4.1. Count data model

The arrival rates are modeled using the count data models. The panel data modeling approach grouped on LZs is adopted to account for the serial correlation among the arrival rates in the same LZ. During preliminary analysis, the overdispersion parameter for negative binomial (NB) model with random effects was found to be significant. Hence, NB model is selected over the Poisson model. After estimating different models and using Akaike Information Criterion (Akaike 1998), the fixed effects overdispersion NB model provided a better fit. The model specification for the fixed effects overdispersion NB model is shown in Equation (2). The count of vehicles arriving at an LZ (y) is Poisson distributed with parameter γ, and the parameter γ is gamma distributed with parameters λ and δ. In the fixed effects overdispersion NB model, λ is estimated as given in Equation (2) and the overdispersion parameter δ is assumed to be constant. The Equation (2) provides the probability of the count (number of vehicles arriving, y) in each group (LZ), for a given independent variable vector X, conditioned on vector of sum of vehicles arriving in each LZ ($\sum y$), where $\Gamma()$ represents gamma distribution. Hence, fixed effects overdispersion NB model does not have a overdispersion parameter δ in the model results (Hausman et al. 1984). Although there exists literature in sociology that recommends improvements to this model, the authors discovered that the model given in Equation (2) provided a better fit for parking analysis (Allison and Waterman 2002).

$$P(Y=y|X,\sum y)=\frac{\Gamma(\sum\lambda)\Gamma(\sum y+1)}{\Gamma(\sum\lambda+\sum y)}\prod\frac{\Gamma(\lambda+y)}{\Gamma(\lambda)\Gamma(y+1)} \tag{2}$$

Where, $\lambda = exp(-\beta X+\varepsilon)$ is a vector with each element for an LZ. β is the vector of coefficients (Coef.).ε is the error term.

4.2. Survival model

The parking duration is modeled using the parametric duration or survival models, which are used to estimate the probability until the occurrence of an event. The parking data are recorded from opening to closing time for each day in all LZs. Although the starting times are known, the observations where the truck driver had overlooked checking out from the parking device cannot be classified as left censored data. The end times for these vehicles are not only unknown but also it is more likely that a reasonable share of these vehicles could have parked beyond the end of the study period i.e., closing time of the LZ for the day. Therefore, the parking data considered in this study are a complete and uncensored data as both start and end times are available for all observations in the sample. This paper models the probability that the parking for a vehicle ends within a given duration. The Weibull model is found to provide better fit for the data. Unlike the count data models, the panel effects over the location (LZs) were not considered for survival models. Because the parking durations are modelled at the vehicle level whereas, the arrival rates are modelled at LZ level. Hence, the vehicle attributes such as professional activity, weights were included in the survival analysis but not in the count data model. It is a reasonable assumption, as the parking durations may not show serial correlation problem across LZs. However, the arrival rate models must consider the panel effects with respect to LZs, since the count data models estimate the number of vehicles arriving at an LZ. As mentioned in Section 3, the duration dataset is at vehicle level whereas, the arrival rate dataset is at LZ level. The survival function for Weibull model (accelerated failure time), defined as the cumulative probability that the parking duration is greater than or equal to the time t, is given in Equation (3) (Washington et al. 2009).

$$S(t)=P(T\geq t)=exp(-\lambda t^{p}) \tag{3}$$

$$\lambda = exp(-p\beta X+\varepsilon)$$

The λ in the above equation is estimated by using regression methods as a function of vector of relevant attributes or variables X. β is the vector of coefficients (Coef.). p is called the scale parameter that allows the probability to change over time. p greater than one indicates that the probability the parking would end increases monotonously with time. Similarly, p less than one indicates a monotonous decrease while p equals one is no change with time which becomes an exponential model.ε is the error term.

4.2.1. Model selection

The survival and count data models are selected based on likelihood Ratio (LR) test, Akaike information criterion (*AIC*) and Bayesian information criterion (*BIC*) as defined in Equations 4–6 (StataCorp 2007).

$$LRtest = -2(LL_{null}-LL_{model}) \tag{4}$$

Table 4
Survival Model Selection.

No	Model	LL null	LL model	LR test	df	AIC	BIC	Corr*
1	Exponential	− 95,549.35	− 94,168.06	2,762.58	88	188,512.10	189,323.70	0.283
2	**Weibull**	**− 86,524.99**	**− 82,411.82**	**8,226.34**	**89**	**165,001.60**	**165,822.00**	**0.281**
3	Gompertz	− 85,674.35	− 82,616.81	6,115.08	89	165,411.60	166,232.30	0.281
4	Lognormal	− 91,500.20	− 88,300.00	6,400.40	89	176,778.00	177,598.80	0.283
5	Loglogistic	− 93,206.38	− 83,069.66	20,273.44	89	166,317.30	167,136.20	0.282

Note: LL = Loglikelihood, LR = Likelihood Ratio.

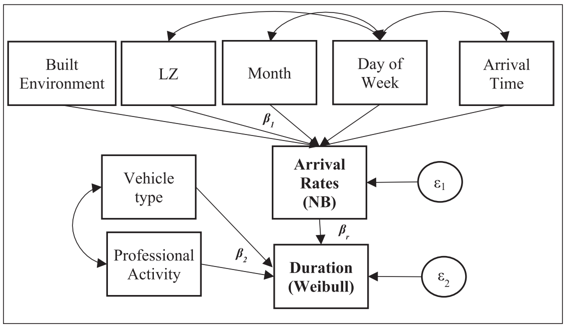

Fig. 3. Schematic of Joint Model for Arrival and Durations.

$$AIC = -2(LL_{model} - k) \tag{5}$$

$$BIC = -2(LL_{model} - k\ln(N)) \tag{6}$$

Where, LL_{null} is the loglikelihood for null model. The null model is the restricted model without including any parameters. LL_{model} is the loglikelihood of current model. Total number of parameters is given by k and N is the sample size.

In addition to the above three measures, the correlation coefficient (Corr) between the actual and estimated durations are also considered in the model selection. Table 4 shows the selection criteria for five duration model specifications namely, Exponential, Weibull, Gompertz, Lognormal, and Loglogistic. Since the Weibull model has the lowest Akaike Information Criteria (AIC), and Bayesian Information Criteria (BIC), with a Corr of 0.28, it is selected as the best model (bold in Table 4). The scale parameter (p) as explained in Equation (3) is significantly greater than one which supports the rejection of exponential model (Box-Steffensmeier et al. 2004). The Weibull model selection is further justified through Somers' D statistic as explained by Newson (2010), as this model produced the highest Somers' D (0.2) with respective p-value equal to zero.

Similarly for the arrival rates, the NB model is selected, as it produced the least AIC (134,669) and BIC (135,630) compared to the Poisson model that produced AIC and BIC of 137,898 and 138,081 respectively. The overdispersion parameter is significantly greater than zero led to the selection of NB model. Also, hurdle or zero-inflated count data models are not appropriate in this particular context as the zero values in the arrival rate in the dataset resulted from the actual lack of vehicles arriving in that time interval and not affected by any other parameter or user choice.

4.3. Joint model for arrivals and duration

Fig. 3 presents the schematic representation of the joint model that estimates the arrivals and duration of the vehicles parked at the LZs. The joint model combines a count data model for arrival rates and a survival model for parking duration. Joint estimation of counts and survival is seldom adopted in the medicinal sciences in modeling the number of incidences of an ailment and the probability of survival of a patient (Cowling et al. 2006). Recently, such joint models are used in the road safety analysis to estimate the number of crashes and time between the two consecutive incidents (Wu et al. 2022). The existing joint models assume the Poisson model for the counts to minimize the computation complexity. However, the model proposed in this paper overcomes this assumption by including a negative binomial (NB) model for the arrival rates to account for the overdispersion. The final joint model comprises of a mean-dispersion NB model (see Equation (7)) for the arrival rates and a Weibull model (see Equation (8)) for the parking durations. The model form for the arrival rates in the joint model (see Equation (7)) is different from that of the separate model (see Equation (2)) in the inclusion of the overdispersion parameter (δ). In essence, the former model results would have an overdispersion parameter (δ) while the latter model would not (since assumed as fixed). The Equation (7) estimates the probability of the number of vehicles arriving (y), for a given independent variables vector X, and Γ() represents gamma distribution.

Table 5
Count Data Model (NB).

Variables		Binary Terms		Day of Week (Binary Interactions)											
				Mon		Tues		Wed		Thur		Fri		Sat	
		Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z
Constant		−3.44	−3.11											0.21	3.27
LZ	LZ1	1.62	3.10	1.19	17.08					−0.07	−2.01			0.41	6.88
	LZ 2			1.63	25.14										
	LZ 3			1.56	22.61	0.09	2.30								
	LZ 4	0.76	3.70	1.54	24.25	0.15	5.53					0.09	2.94	0.46	9.51
	LZ5			1.48	23.95							−0.09	−3.16		
	LZ 6	2.78	4.28	1.48	22.61									0.36	6.38
	LZ 7*	0.61	4.90	1.46	23.01					0.06	2.03	0.05	1.67	0.30	6.04
	LZ 8			1.40	22.26										
Month	**Jan**					0.08	2.39					1.82	27.17		
	Feb	−0.07	−2.37			0.20	4.63	0.15	3.32			2.03	28.77	0.23	3.20
	Mar	0.08	4.35									1.79	27.01		
	Apr											1.61	22.28		
	May											1.65	24.82		
	June	−0.07	−3.03	−0.15	−2.78							1.70	23.89		
	July			−0.32	−10.35	−0.37	−12.59	−0.26	−9.06	−0.32	−9.75	1.40	21.21	−0.37	−5.74
	Aug			−0.42	−12.11	−0.42	−12.87	−0.43	−13.02	−0.48	−15.08	1.29	20.09	−0.35	−5.89
	Sept	−0.13	−8.31									1.75	27.24		
	Oct	−0.06	−4.26									1.62	24.89		
	Nov*	0.03	1.42			0.06	1.81	0.08	2.43			1.77	27.55	0.11	2.01
	Dec					0.06	2.03					1.73	27.15	0.21	4.18
Arrival Time	**7–7:59**	0.96	17.93	0.17	2.30	1.81	20.14	1.67	18.26	1.56	16.63			0.52	5.45
	8–8:59	1.47	66.73			1.83	29.24	1.67	26.39	1.64	25.88			0.19	4.34
	9–9:59	1.60	78.81			1.78	28.75	1.74	28.15	1.70	27.35				
	10–10:59	1.53	73.73			1.72	27.54	1.70	27.28	1.69	26.91				
	11–11:59	0.95	39.32			1.67	25.22	1.66	25.00	1.67	24.85				
	12–12:59			0.12	2.31	1.78	25.75	1.78	25.49	1.75	24.77				
	13–13:59			0.20	4.05	1.82	26.51	1.84	26.63	1.83	26.29			−0.66	−5.88
	14–14:59			0.89	24.13	2.42	38.46	2.43	38.33	2.37	37.02			−0.37	−3.70
	15–15:59	0.76	22.17	0.44	10.12	1.94	27.62	1.90	27.00	1.95	27.57			−0.76	−8.56
	16–16:59			1.00	27.88	2.51	40.32	2.58	41.40	2.56	40.72				
	17–17:59					2.09	31.71	2.13	32.19	2.22	33.91				
	18–18:59	−0.86	−17.22			1.72	17.10	1.77	17.61	1.77	17.28				

Built Environment Variables		
Emp. within 100*m*	0.0004	2.65
Pedestrian	0.76	4.61

Obs=44,508, Groups=8 (LZs); Wald chi2(1 1 8) = 35230.42; LL=-67215.622
LL_null=-84216.012; AIC=134,669; BIC=135,630; Prob > chi2 = 0
Pseudo R^2=0.202

Note: Empty cells = not significant at 10 % level (|z|≤1.64), * =significant at 10 % level, LL = Loglikelihood, LR = Likelihood Ratio.

Table 6
Survival Model (Weibull).

Variables		Binary Terms		Day of Week (Binary Interactions)											
				Mon		Tues		Wed		Thur		Fri		Sat	
		Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z
Constant		2.92	151					0.17	6.41	0.05	1.96	0.12	4.59		
Year 2019		0.02	4.42												
Loading Zones	**LZ1**			−0.06	−2.32	−0.07	−3.54								
	LZ 2	−0.04	−3.39	−0.06	−2.78	−0.05	−2.25								
	LZ 3	−0.09	−6.45					0.07	2.79	0.10	3.69				
	LZ 4	−0.06	−6.72							0.05	3.14				
	LZ5			−0.05	−3.12										
	LZ 6	−0.05	−4.95												
	LZ 7	−0.04	−5.39	−0.04	−2.35										
	LZ 8	−0.08	−8.09			0.04	2.48			0.05	3.19			0.10	3.36
Month	**Jan**											−0.10	−4.50		
	Mar							−0.06	−2.88			−0.05	−2.42		
	May					0.04	1.99								
	Aug							−0.05	−2.52						
	Sept											−0.04	−2.52		
	Nov									0.04	2.21				
Arrival Time	**6–7:59**			−0.07	−2.87	−0.08	−3.29	−0.10	−3.84	−0.08	−2.77				
	8–9:59	−0.07	−9.90			0.05	4.41								
	10–11:59	−0.05	−6.76												
	12–13:59	0.03	2.68			−0.05	−2.22								
	14–15:59							−0.03	−2.05						
	16–17:59							−0.06	−3.44			−0.04	−2.17		
	18–19:59	−0.10	−4.79												
Activity	**Unspecified**	0.19	10.8							0.06	3.11	0.04	2.20		
	Install & Maint…	−0.13	−7.68	−0.04	−2.72			−0.05	−3.13						
	Transport…	0.09	2.30	0.12	3.20	0.08	2.15	0.13	3.27	0.16	3.81	0.16	4.01		
	Construction	0.10	4.78	−0.08	−3.25	−0.09	−3.88			0.07	2.73	−0.10	−4.29		
	Local commerce	−0.13	−2.06	0.26	4.04	0.24	3.74	0.24	3.76	0.29	4.50	0.18	2.79		
	Comercial Agent	−0.16	−8.96							0.06	3.09				
	Food and Markets	−0.08	−2.07												
	Automotive	−0.09	−4.58							0.09	3.76			0.16	4.06
Vehicle Type	**Car**							−0.23	−5.73	−0.19	−4.65	−0.19	−5.03		
	Truck less than 3.5 T	0.14	6.16					−0.17	−6.51	−0.14	−5.30	−0.16	−6.15		
	Truck 3.5 T to 12 T	0.14	5.29					−0.18	−5.27	−0.09	−2.56	−0.10	−2.73		
	Truck greater than 12 T	0.14	5.03	0.15	2.02										
	Van			0.18	6.98	0.15	5.80							0.19	5.58
Shape Parameter (1/p)		0.62	162	**Obs. 74,777; LL = -82411.82; LR chi2(87) = 5295.71; Prob > chi2 = 0**											

Note: Empty cells = not significant at 10 % level ($\lvert z \rvert \leq 1.64$), LL = Loglikelihood, LR = Likelihood Ratio.

$$P(Y = y) = \frac{\Gamma((1/\delta) + y)}{\Gamma((1/\delta))y!}\left(\frac{(1/\delta)}{(1/\delta) + \lambda}\right)^{1/\delta}\left(\frac{\lambda}{(1/\delta) + \lambda}\right)^{y} \tag{7}$$

$$\lambda = exp(\beta_1 X_1 + \varepsilon_1)$$

The survival model for the parking duration is same as the Weibull model shown in Equation (3).

$$S(t) = P(T \geq t) = exp(-\alpha t^{p}) \tag{8}$$

$$\alpha = exp(-p\beta_r AR - p\beta_2 X_2 + \varepsilon_2)$$

$X_{1=}$ Variable vector affecting arrivals (built environment, LZ, month, day of the week and time).
$X_{2=}$ Variable vector affecting the parking duration (professional activity and vehicle type).
$\beta_{1,}\ \beta_{r,}\ \beta_2 =$ Coefficient vector for count data and survival models respectively.
$\varepsilon_{1,}\ \varepsilon_2 =$ Error terms for count data and survival models respectively.
$p =$ Shape parameter for Weibull model.

The arrival rate model has the built environment variables, LZ, month, day of the week and arrival time (in hours) as the observed variables (X_1). The duration model inputs the estimated arrival rates from Equation (7) (*AR*), along with the activity (business type) and the vehicle type. The model considers the interaction terms among these variables as depicted by the "arcs with arrows" in the Fig. 3. The correlation between the error terms ε_1 and ε_2, is not significant. Hence, the interaction between these error terms is not included. It is reasonable to assume that the arrival rates mostly depend upon the day, time, location of LZ and month while the durations highly depend on the type of activity, and size of the vehicle. For example, a LZ would most probably have high arrival rates in morning peak hours (say 6–9 AM) on a weekday compared to non-peak hour on a Saturday. Similarly, the bigger vehicles serving the transport and parcels are more likely to have longer parking durations compared to the smaller vehicles serving the local commerce.

The model is programmed in STATA version 14 (Pitblado 2013). The outliers in the survival model are identified using the deviance residuals (greater than + 3 and less than −3). The outliers in the count data are detected using the threshold (less than 0.5) values defined by Enzmann (2015). For more explanation on outlier detection, readers can refer to Lambert and Royston (2009) for survival analysis and Enzmann (2015) for count data models. The next section presents and compares the results from the separate and the joint model for arrival and parking durations.

5. Results and analysis

The results, analysis, key findings, and cross-validation of the separate models for arrival rate (count data), parking duration (survival) and the joint models are presented below. The model performance of separate models is compared with that of the joint model with respect to the influencing factors and the predictability of the actual arrivals and parking duration.

5.1. Count data model

Table 5 shows the results for fixed effects overdispersion NB count data model with panel effects based on LZs. As discussed in Section 4, the fixed effects model does not estimate an overdispersion parameter separately. All coefficients including the binary and binary interaction terms, shown in Table 5 are significant at 5 % level ($|z| \geq 1.96$). The pseudo R^2 larger than *0.2* indicates a good fit of the model with the actual data. For the fixed effects overdispersion NB model, the marginal effects (the rate at which the arrivals change with respect to a variable) are equal to the model coefficients. Among the binary variables, the loading zones (LZs), month, arrival time, are significant and two built environment variables; number of employees within 100 m radius from LZ (Emp. within 100 m) and binary variables indicating the presence of pedestrian pavements (Pedestrian) are significant. The arrival times at one-hour interval are found to be significant factors affecting the arrival rates. The day of the week is significant when combined with either month or arrival times. The interaction terms of arrival time intervals with the day of the week are significant compared to that of month and loading zones. The arrival rates are lower from 12:00 to 15:00 (3 PM) and between 18:00–19:00 (7 PM). On an average Tuesdays, Wednesdays, and Thursdays have higher arrival rates highly dependent on the time of the day than the LZ. February, March, and November months have higher rates. April, May, and December have coefficients significantly not different from the base case. July and August months have lower arrival rates than other months across all times of the day. The existence of pedestrian pavements leads to an increase in the arrival rates. Surprisingly, the higher the number of establishments within 100 m from LZ, the lower the arrival rate. The goodness of fit indicator, Chi2 test shows that all the parameters in the model are statistically significant.

Table 7
Joint Model for Arrivals and Parking Duration (NB + Weibull).

Arrival Rates (Negative Binomial)															
Variables (X_1)		Binary Terms		Day of Week (Binary Interactions)											
				Mon		Tues		Wed		Thur		Fri		Sat	
		Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z	Coef	z
Constant		1.77	155.5			0.07	7.27					0.04	3.58	−0.79	−36.9
Loading Zones (β1)	**LZ 1**	−0.56	−54.7	−0.17	−5.94									0.34	6.69
	LZ 2	−0.30	−28.5	0.07	3.89					−0.08	−4.03	−0.09	−4.54	−0.11	−1.79
	LZ 3	−0.69	−57.6									−0.09	−3.50	0.24	3.41
	LZ 4	0.10	11.2	0.04	3.32	0.12	11.17					0.05	3.71	0.20	7.82
	LZ5	0.32	54.9									−0.10	−9.61		
	LZ 6	−0.43	−46.5									−0.09	−4.41	0.27	6.53
	LZ 7	0.05	7.9							0.07	6.60				
	LZ 8									0.05	4.31				
Month (β1)	**Jan**					0.08	5.98					0.09	5.43		
	Feb					0.11	8.91	0.04	2.70			0.20	13.68		
	Mar	0.07	9.6									0.04	2.69		
	Apr	0.06	8.3												
	May	0.05	6.1									−0.08	−4.38		
	July	−0.12	−16.5			−0.08	−5.75								
	Aug	−0.26	−36.6												
	Sept	−0.05	−7.81									0.06	4.33		
	Oct	−0.04	−6.14									0.09	6.09		
	Nov											0.05	4.13		
	Dec							0.05	4.38			0.04	2.69	0.18	6.19
Arrival Time (β1)	**6–6:59**	−0.83	−40.32												
	7–7:59	−0.21	−15.84			0.08	3.87			−0.12	−4.69				
	8–8:59	0.05	5.04	−0.10	−7.41	0.10	8.34	0.07	6.35					−0.10	−3.55
	9–9:59	0.22	24.51	−0.20	−16.92			0.06	5.59					−0.20	−7.11
	10–10:59	−0.93	−50.56	−0.05	−3.72	0.15	11.70	0.20	15.83	0.17	13.16				
	11–11:59	−0.59	−30.31	−0.27	−14.86	−0.21	−12.77	−0.13	−7.64	−0.10	−6.21			−0.27	−7.01
	12–12:59	−0.45	−26.97	−0.96	−26.28	−0.86	−30.67	−0.86	−27.79	−0.84	−26.40			−0.63	−8.70
	13–13:59	−0.55	−30.94					0.08	2.57	0.10	3.03				
	14–14:59	−0.72	−38.97	0.18	7.12	0.11	4.22	0.23	9.15	0.14	5.30			−0.17	−1.77
	15–15:59	−0.93	−36.26	0.21	9.71	0.13	5.96	0.15	6.88	0.18	8.63			−0.22	−2.91
	16–16:59			0.19	7.84	0.16	6.73	0.25	10.65	0.20	8.68				
	17–17:59					0.13	4.74	0.12	4.48	0.19	7.28				
	18–18:59									0.17	3.48				
Duration Model (Weibull)												**Model Stats**			
Variables (X_2)		**Binary**		**Truck less than 3.5T**		**Truck 3.5T to 12T**		**Truck greater than 12T**		**Van**					
		Coef	**z**	**Coef**	**z**	**Coef**	**z**	**Coef**	**z**	**Coef**	**z**				
Constant		2.82	198.4	0.15	5.70	0.33	8.29	0.28	5.06	−0.05	−2.74	Obs=74,777			
Activity (β2)	**Unspecified***	−0.80	−1.79	1.01	2.26	0.82	1.82	0.92	1.96	1.23	2.76				
	Install & Maint…			−0.15	−6.28	−0.15	−3.55					LL=-239427.63			
	Transport…	0.38	17.96	−0.17	−5.01	−0.43	−3.93					LL_null=-272080.85			
	Construction	−0.41	−4.87	0.46	5.22	0.30	3.11			0.77	8.97	df=136			
	Local Commerce	0.34	22.65	−0.24	−7.44	−0.83	−8.25	−0.74	−4.19			AIC=479166.7			
	Comercial Agent	0.11	6.10	−0.25	−8.26	−0.51	−11.46	−0.29	−4.61			BIC=480420.9			
	Food and Markets	0.21	3.30	−0.30	−3.78	−1.32	−3.55								
	Automotive	0.16	9.06	−0.19	−6.07	−0.68	−13.68								
Model Parameters															
Arrival Rate (βr)		−0.002	−2.6	**Dispersion (1/ln(δ))***				−18.0	−1.7	**Shape Parameter (1/p)**				0.46	157.7

<u>Note</u>: Empty cells = not significant at 10 % level, * =significant at 10 % level, LL = Loglikelihood, LR = Likelihood Ratio.

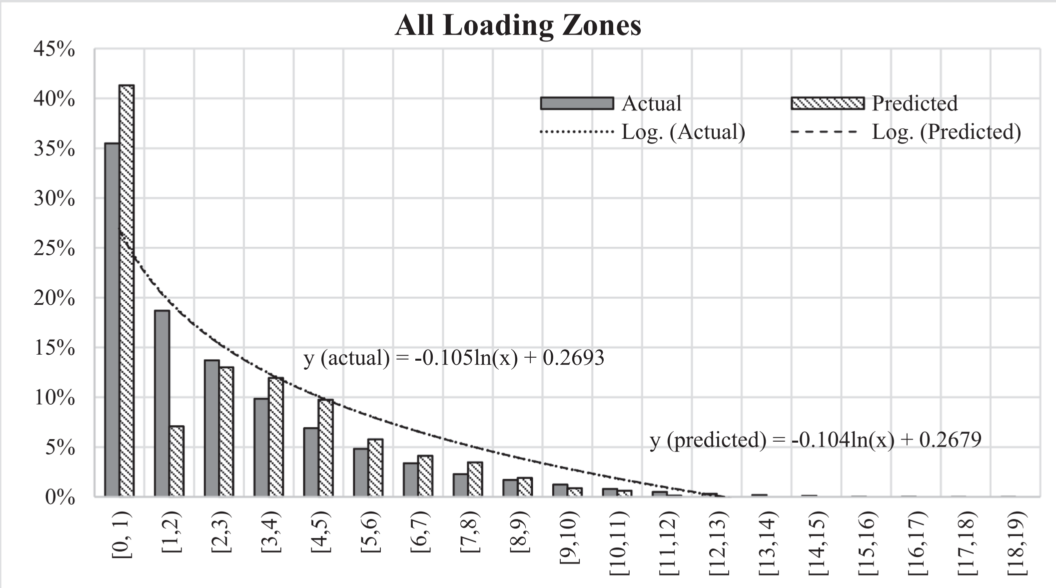

Fig. 4. Actual vs Predicted Arrival Rate Distribution.

5.2. Survival model

Table 6 shows the survival model results that estimate the cumulative probability that a vehicle is parked for one more minute at a given time *t*; as a function of year, LZs, day of week, arrival time, professional activity, and vehicle type. All coefficients are significant at 5 % level. The higher coefficient value the higher the value of survival rate, decreases exponentially with time. For example, at a given time *t*, the probability that a vehicle parks for one more minute is higher in 2019 compared to 2018.

The parking duration depends significantly on the professional activity and the vehicle type compared to the day of the week, month, and time. Six months (February, April, June, July, October, and December) do not affect the parking durations. LZ 1 and 5 have higher survival rates than other LZs. Compared to count data model in Table 5, aggregation of arrival time to two hours (instead of one hour) provided better fit. Wednesdays, Thursdays, and Fridays have higher survival rates. Transport and parcels have one the highest parking durations on all days of the week, while install and maintenance activity has the least. Vehicles arriving between noon and 17:00 (5 PM) have higher probability of longer parking durations. Trucks with gross weight greater than 12 tons park longer, which is in line with previous findings i.e., single unit trucks park longer than vans (Schmid et al. 2018). The Chi2 test indicates the model is significant and the inverse of the shape parameter ($1/p$) is greater than one, 0.62 and statistically significant.

5.3. Joint model for arrival rate and duration

Table 7 presents the results for the joint estimation of (see Equations (7) and (8)) mean-dispersion NB model for arrival rates and the Weibull model for parking duration. The mean-dispersion NB model contains an overdispersion parameter (δ, see Section 4). The separate model for the arrival rates (see Table 5), and the joint model (see Table 7) have same significant variables (time of the day and week) influencing the arrival rates similarly. For instance, Fridays and Saturdays have less arrival rates in both models. Among the months July and August have lesser arrival rates. There are two major differences in the joint model compared to the separate model for the arrival rates. Firstly, a greater number of variables (interaction terms between day of the week with month and time) are significant in the joint model. Second difference is related to the built environment variables such as employment, existence of pedestrian pavement which are not significant in the joint model. Similar to the separate model (see Table 6), the parking durations in the joint model significantly dependent on the professional activity and vehicle type (see Table 7). The negative coefficient for the arrival rate parameter in the duration model (β_r) shows that the probability of vehicles parking for longer durations decrease with increase in the arrival rates. Hence, the crowded LZs have lesser parking durations. In contrast to the separate model, the duration model in the Table 7 shows that the interaction terms of vehicle type and professional activity is highly significant. The over dispersion parameter (δ) for the arrival rates is larger than zero and significant at 10 % level. A higher shape parameter (p) in the joint model

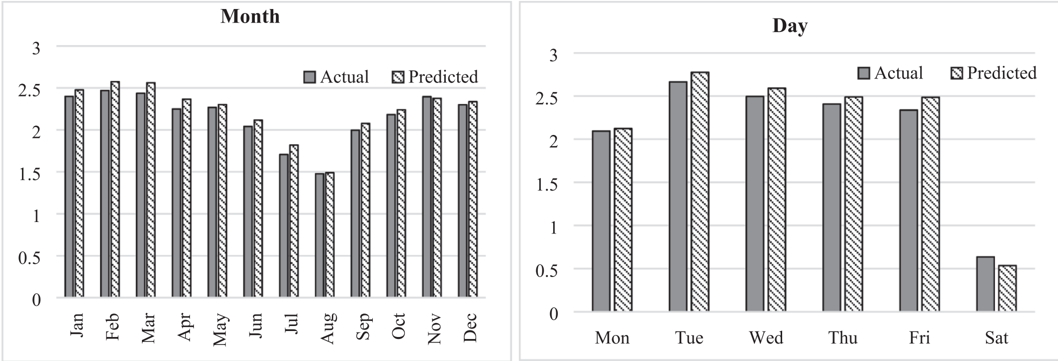

Fig. 5. Average Arrival Rate by Month and Day of Week.

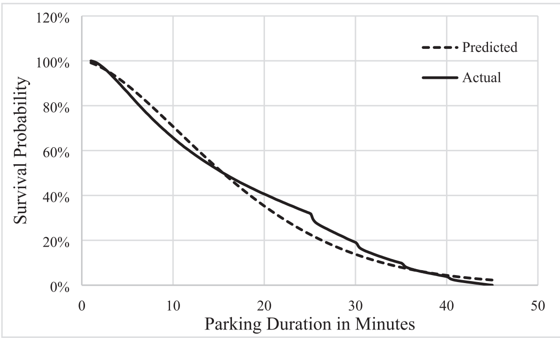

Fig. 6. Actual vs Predicted Survival Rates.

indicates that longer tail for the survival probability function i.e., better captures the vehicles that park for either very short or long durations compared to the mean.

The arrival rate and count data models (both separate and joint) are validated for overfitting and compared for model performance through k-fold cross-validation method (see Appendix A). The number of folds used is ten based on the optimal number of folds found in the previous studies (Abu-Mostafa et al. 2012, Marcot and Hanea 2021). So, the dataset is split randomly into ten folds of equal sample size. In each iteration, nine folds act as the training dataset and remaining onefold serve as the testing data. The cross-validation used Root Mean Squared Errors (RMSE) between observed and estimated arrivals and parking durations as the metric for evaluation. The lower the RMSE, the better is the model. The Appendix A shows that the models presented in Tables 5-7 do not show overfitting as all ten testing datasets provided almost similar RMSE for both arrival rates and parking durations. The goodness of fit for the models is evident from the lower standard deviation of RMSE across the testing datasets (Std. Dev), varying between 0.06 and 0.1. The joint model has lower RMSE compared to the separate model for both arrival rates and parking duration. However, the separate models have lesser Std. Dev for the duration model displaying better consistency in the estimated values. The join model provides the best fit for the arrival rates with respect to both, smaller RMSE and Std. Dev.

5.4. Predicted vs Actual values (Joint Model)

5.4.1. Arrival rates

Fig. 4 shows the distribution of the actual and predicted arrival rates for all LZs estimated using the joint model shown in section 5.3. The predicted arrival rates are obtained using a linear approximation for the expected value of NB model. Hence, the actual arrival is a discrete value while the predicted ones are continuous. For example, the column [0,1) shows the percentage of observations that have zero arrival rates for actual values but covers the values between zero and one for predicted values. The figure shows that the estimated counts have closer distribution as that of actual arrival rates, with slightly higher predictions for counts between zeros and ones, and lower for counts between one and two. There is no predicted arrival rate larger than thirteen as these values are found to be less probable in the dataset. Also, the log trendlines for both distributions match. The mean and standard deviations of the predicted distribution is 2.17 (2.11 for actual) and 2.15 (2.55 for actual) respectively. About 50 % of the overdispersion is accounted in the model as the ratio between the standard deviation and mean is 0.99 compared to 1.21 in actual data. The median of predicted arrival rates, 2 is higher than that of the actual distribution. Hence, the arrival rate models provide good fit to the dataset. In addition, the NB count

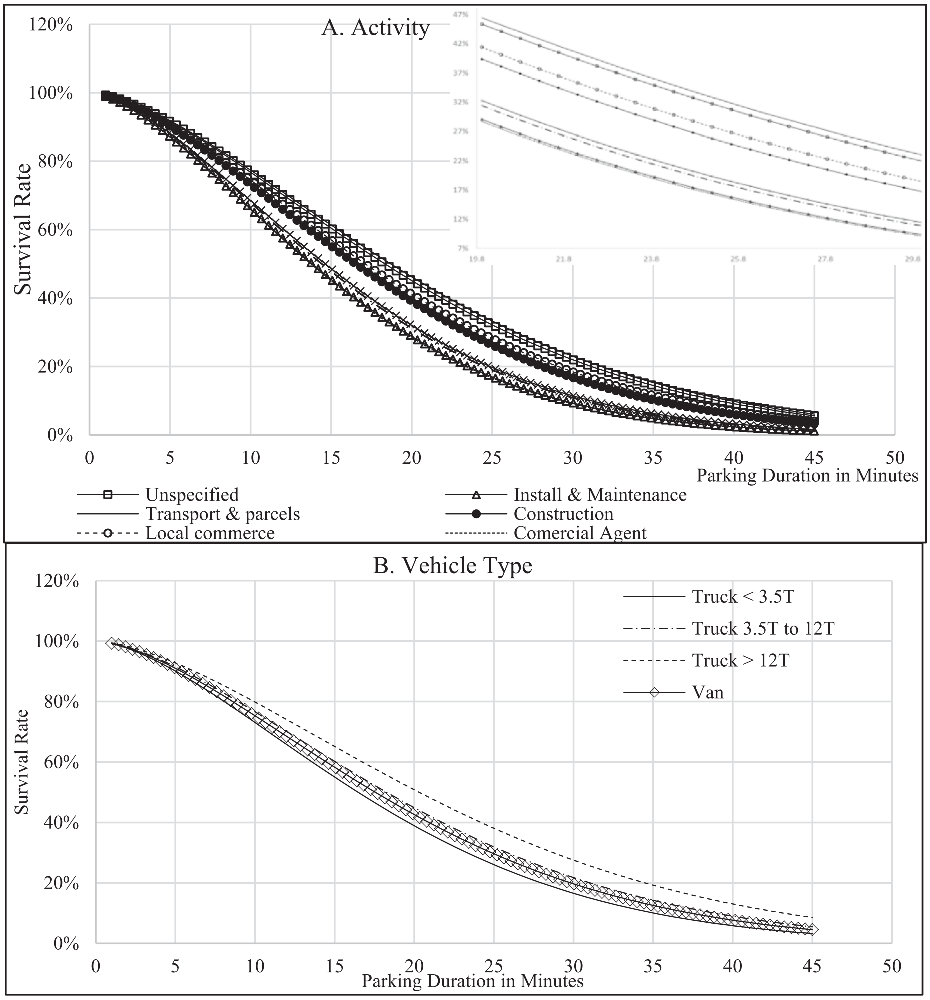

Fig. 7. Survival Rates for Professional Activity and Truck Type.

data model estimates the parking demand using as a probability of obtaining discrete demand compared to the approximation of the linear regression models which assume a continuous values of parking demand.

The count data model performance is further examined to capture the mean arrival rates by month and day of the week. Fig. 5 shows the actual and predicted mean arrival rates for all LZs by month and day of the week. The predictions are match the actual mean values. The months with better predictions are August, November, and December. Mondays and Thursdays have less error between actual and predicted arrival rates. Saturdays predicted mean arrivals are less than the actual mean. The distribution of predicted and actual arrival rates by LZ is shown in Appendix B. The predicted arrival rates by LZ show similar distribution as that of actual arrival rates. Therefore, the joint model is capable of replicating the actual arrival rates for different LZs, month, and day.

5.4.2. Duration

The joint model in Table 7 (predicted) is compared with Kaplan-Meier curve (actual), which is a non-parametric estimation method that estimates the cumulative probability as the ratio between the sample sizes as shown in Equation (9) below (Kaplan and Meier 1958). Fig. 6 shows that the predicted survival function better matches the actual survival rate.

$$S(t) = 1 - \frac{No\ of\ vehicles\ left\ LZ\ from\ start\ to\ time\ t}{No\ of\ vehicles\ parked\ from\ start\ to\ time\ t} \tag{9}$$

We can observe a monotonous decrease in the probability of the vehicle staying for one more minute at an LZ with increase in the time. There is a 30 % chance that, on average, a vehicle in an LZ is parked for less than 10 min. The probability increases to 65 % if the threshold duration increases to 20 min.

The results could be compared with findings from the limited literature available on parking durations. Schmid et al. (2018) found that in New York City (NYC), the mean parking duration is 15.66 min (17.78 in current paper). The distribution is right skewed with majority of vehicles parking for less than 10 min unlike the flatter distribution found in current research as shown in Fig. 2. Low et al.

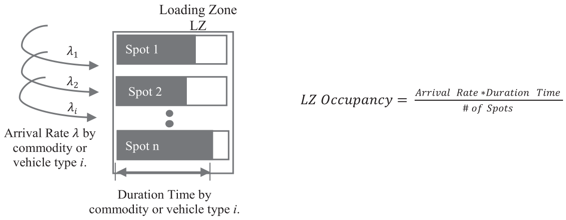

Fig. 8. Loading Zone Occupancy.

Table 8
Share of Vehicles vs Parking Duration Limits.

Share of Vehicles Activity	Parking Duration Limit (minutes)			
	25 %	50 %	75 %	90 %
Food and Markets	8	14	22	30
Transport & parcels	11	19	29	40
Construction	9	16	25	35
Local commerce	10	17	26	36
Automotive	8	14	23	31
Commercial Agent	8	13	21	30
Install and Maintenance	8	14	21	29

(2020) conducted a parking study on LZs in Singapore showed similar distribution as Schmid et al. (2018) with an average parking duration of 18 min for delivering the cargo and 17 min for picking up. The survival model estimated by Schmid et al. (2018) is also a Weibull model, but with steeper slope i.e., the probability that a vehicle parks for less than 20 min is 80 % in (Schmid et al. 2018) as compared to 65 % in the current research as shown in Fig. 6. The current study deals with bulk cargo deliveries at LZs; it is noted that the vehicles park (slightly) longer than the normal commercial deliveries at NYC. All the studies discussed above show a handful of parking violations in exceeding the maximum time limit.

Fig. 7 shows variation of survival function from the joint model in Table 7, by professional activities and vehicle types. The activities in the increasing order of the probabilities of parking longer, with 50 % chances of parking longer than time in the parenthesis are: Transport and parcels (~19.2 min), unspecified (~17.5 min), local commerce (~17 min), construction (~16.4 min), automotive (~16.1 min), food and markets (~14.5 min), install and maintenance (~14 min), and commercial agent (~13.4 min). Similarly, the vehicle types in the increasing order of the probabilities of parking longer are trucks with gross weight larger than 12 tons (~20 min), trucks between 3.5 and 12 tons (~18 min), vans (~17 min) and trucks with less than 3.5 tons (~16 min). The sector unspecified could transport a wide range of miscellaneous cargo that cannot be defined precisely. It is interesting to note that the vans park longer than the small or pickup trucks with less than 3.5 tons. This could be due to nature of the activity the vans belong to. The dataset comprises of 28 % of vans in unspecified activity, and 22 % in local commerce. Whereas around 32 % of the pickup trucks belong to one of the activities with the least parking duration, install and maintenance. Hence, it is important to include the interaction terms between the vehicle type and activity in the survival models.

For duration models, the marginal effects are defined as the change in the median duration with respect to a unit change in a given variable ($d\mu/dx$). A positive marginal effect implies that the median duration increases with the addition of a variable and a negative marginal effect implies a decrease in the median duration. The variables with the highest marginal effect are professional activity unspecified (3.06), vehicle type truck with weight larger than 12 tons (2.27), and truck with weight less than or equal to 3.5 tons (2.26). The variables with the least marginal effects are commercial agent (-2.46), install and maintenance (-2.12), and local commerce (-2.04). The marginal effects show that the median parking duration is more sensitive to professional activity compared to the vehicle type.

6. Practical implications

The methodology and modeling framework developed in this research has a wide range of applications in urban freight and parking policies as explained below. The appropriate model should be selected from the three models proposed in the research based on the policy requirements and data availability. For example, a joint model is better in case of data on professional activity, and vehicle type are available. Otherwise, the separate model for arrival rates has to be used.

6.1. Parking demand estimation

Parking demand estimations have had different approaches in literature. Establishment-based models are one of the most commonly used, as explained in literature review (Section 2) constrained by survey design and availability of databases from public or private entities. Since the current research estimates parking demand at LZ level, there is no need for aggregation which is required when the establishment-level estimates are used. The aggregation is required due to the fact that a vehicle can serve multiple establishments each time it parks. According to Marujo et al. (2018), one vehicle stop can make several freight deliveries if customers are up to 500 m apart. Also, due to the nature of vehicle arrivals, the count data models better capture the parking demand behavior than continuous regression models. Hence, the models in Table 5 and 7 help in improving the parking demand estimation without the need for disaggregate input data at establishment-level, by circumventing the issues with aggregation bias.

6.2. Urban infrastructure planning and space allocation

From the public sector perspective, joint estimation of arrivals and parking duration (Table 7) is a suitable tool for infrastructure planning and allocation e.g., decisions related to the location and number of LZs needed at an urban area, LZ's capacity requirement (number of parking spots and sizes per LZ), and optimal allocation of urban curbside space for commercial requirements. This combined approach is a step towards designing customized measures for freight transportation based on the needs of different city zones, commodities, industry sectors, and vehicle types, by incorporating parking availability, and occupancy durations at the LZs into the planning process (Castrellon et al., 2022). Fig. 8 illustrates the combined use of count and survival models. Public authorities or LZ managers could use this approach to evaluate occupancy rates i.e., the ratio between the occupied parking spots to total number of spots at an LZ, where they will be able to make decisions about expanding or reducing LZ capacity if occupancy rates are high (~100 %) or low (~50 %) respectively. These decisions may change during the year, month, day and even time of the day depending on LZ dynamics. For the specific case of Vic, city authorities could implement contingency plans for satisfying parking demand during peak periods e.g., winter months (from November to March) and considering alternative infrastructure usage for time-windows where LZs showed overcapacity e.g., July-August period.

6.3. Time restrictions

Besides infrastructure planning, public decision-makers can also consider measures to foster efficient use of LZs by implementing regulations with time-limit restrictions according to the vehicle type and commodity using parking models in Table 6 and 7. Results showed that the probability of exceeding the 30-minute parking threshold is nearly 30 % for trucks larger than 12.5 tons and more than 20 % for the rest of vehicle types. This may imply the need for reassessment in the decisions regarding parking enforcement or traffic restrictions for heavier vehicles. For professional activities, the probability of exceeding the time limit of 30 min is greater than 20 % for parcel deliveries, local commerce and construction sector, activities that could be included as exceptions in pilots to evaluate sector

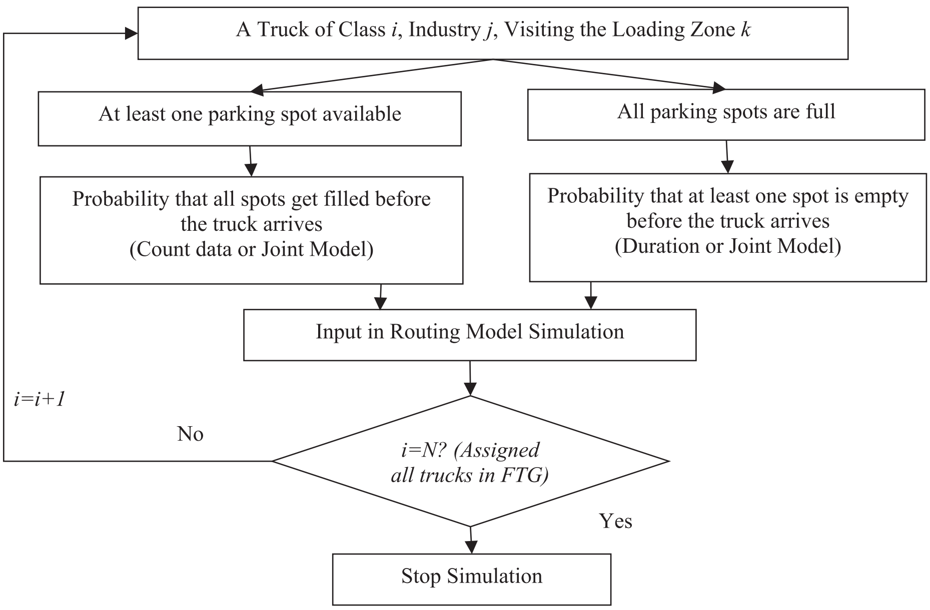

Fig. 9. Parking Simulation.

level acceptance of demand management strategies that minimize parking times.

One way to assess time limits is to calculate the percentage of vehicles that can perform their operation during a given time limit considering different vehicle configurations and professional activities. For the case of Vic, Table 8 shows the time limit that a public authority could define to meet a certain portion of vehicles that can operate without exceeding that threshold. For instance, Transport and parcels deliveries is one of the activities that needs wider time windows. According to the table, 90 % of the Transport & parcels operations can be done in a timeframe of 40 min, while other activities i.e., Food, Install and Maintenance and Commercial Agent, can have the same proportion of activities in 30 min or less. By implementing dynamic restrictions in urban loading zones depending on vehicle type, professional activity a successful strategy to make the most of the scarce urban public space will be possible. The design of dynamic restrictions along with technology-based innovative solutions to support policy enforcement are the key areas of future research on smart loading zones (Castrellon et al., 2022).

6.4. Truck routing simulation

From the private sector perspective, modeling parking demand and duration is a valuable input for designing vehicle routing plans that consider parking availability at different times of the day and days of the week. The joint modeling approach fosters informed decisions by allowing drivers to know beforehand parking availability. Logistics service providers can overcome empirical practices used to plan routes without parking information and avoid costly repercussions such as cruising or double-parking. For instance, truck delivery tour could include parking availability as constraints in obtaining the optimal route for urban deliveries. The analytical framework for management of LZs presented in this paper is a step towards helping the truck drivers, and logistics providers to better manage their deliveries using a mobile application, which incorporates parking constraints in the route plan. The latter will become increasingly important as routing decisions will become more systematic in autonomous freight vehicles that use information from connected infrastructure. Fig. 9 shows the schematic representation of the parking simulation model. Each truck trip in the city belonging to an industry type makes certain number of stops based on the delivery schedule. The truck tour with respect to parking at LZs, depends on two constraints. The first constraint is whether there is an empty spot available once reaching an LZ given by the count data model in Table 5 or Table 7. Second, if an LZ is fully occupied when the earliest departure of the existing truck given by the duration model in Table 6 or Table 7 (Kalahasthi et al. 2021).

6.5. Improved data collection

Last but not the least, the combined demand-duration modeling approach proposed in this paper would not have been possible without the data availability about vehicle arrivals and parking time for different commodities and vehicle types over an 18-month period. As a takeaway of this research, policymakers should look for technology solutions that allow capturing real-time data as an input for analysis such as those presented in this paper. In the specific case of Vic, data attributes included vehicle parking durations by day of the week, time of arrival/exit, emission standards, and industry sector. In this case, input data came from a mobile phone app-based solution that monitored LZ activity in real-time. However, other technologies can also cope with this purpose e.g., sensors, cameras, among others. Also, there is scope to include other potential variables such as industry sector, weight, or volume of cargo, and to collect more complete and precise location data on establishment type, location and size drawn from municipal commercial records. Developments of this kind would minimize data errors or missing data that normally occur in studies based on driver or establishment surveys while improving the freight operations at LZs.

7. Concluding remarks

Truck parking, especially the curb space management in cities is a key challenge faced by urban freight policy makers and transportation planners. The existing literature and modeling approaches on urban freight parking also have a huge scope for improvements. The past research mainly focused on estimating freight parking demand at establishment-level which could be harder to aggregate at loading zone levels as a single truck could make deliveries or pickups at multiple establishments. Also, parking durations play a vital role in demand management as the availability of parking spot at a given time at a loading zone depends on both arrival rate and parking duration. In addition to these limitations with modeling and estimation, data obtained from establishments are prone to have rounding errors as they were mostly based on surveys or interviews from the individuals. Because of these issues, the existing parking demand is either modeled as a continuous variable or discrete choice of using a loading zone. Therefore, this paper aimed to improve the existing urban commercial parking research by exploring a joint modeling approach to estimate the demand and durations of commercial parking at loading zones. The joint model performance is compared with the separate standard models for count data and survival analysis. The parking demand was modeled as a function of arrival rate using count data (negative binomial) models. The count data models are more appropriate for parking as the count data consists of non-negative integers and are widely used in similar situations such as vehicle arrivals at tollgates, ramps, etc. (Washington et al. 2009). The durations are modeled using survival (Weibull) models.

The dataset used in this research is robust and has a low risk of data collection errors as it was a historical real-time data collected from a cellphone-based parking application (Parkunload 2020) at eight loading zones in Vic, Spain for 18 months. Significant variables affecting the arrival rates are, arrival times, day of the week, and month. Duration was dependent on the professional activity and vehicle type. The potential policy implications of the models developed in this research include infrastructure planning such as space allocation for new or modified loading zones, land-use policies; estimation of flexible limits for parking time, vehicle size, and weight restrictions; assisting carriers in obtaining optimal truck routing plan incorporating the parking constraints, and improved freight data collection, to name a few. The applications of such research as the one presented in this paper, with further advancements as mentioned below, would lead to huge benefits to freight community in terms of reduction in emissions, energy consumption, and vehicle kilometers travelled.

In spite of the considerable advantages of the current research, there are several limitations worth mentioning that set foundations for future research on improving urban commercial parking. First limitation is the lack of some relevant variables in the dataset e.g., industry sector, weight or volume of cargo, establishments by sector, service vehicles, activity type (delivery or pick up). Though, the variable "professional activity" covered a wide range of industry sectors, there are many observations labelled as unspecified activity. The gross vehicle weights are included. However, the actual weight of cargo would be one of the deciding factors on time spent in parking. Freight or service vehicles were not differentiated although it is less likely that service vehicles which require higher parking durations could be parked on a loading zone with a maximum parking limit of 30 min (Holguín-Veras et al. 2021). Second limitation is on data and modeling, which could be disaggregated further. The count data and survival models could be improved by estimating the arrival rates and durations by industry sector and including the respective freight trip generation and establishment variables by sector. Final and vital limitation is related to the model application. The model developed in this paper is from a small city. Hence, future research efforts are required in assessing the geographic transferability of these models to other cities, especially the arrival rate distribution (Fig. 2) could be different in bigger cities. Also, it is important to explore spatial modeling approaches in the joint estimation of arrivals and durations to better examine the proximity effects of LZs. If data are available on the factors (e.g., access restrictions for certain vehicle types) influencing the zero arrival rates (if exists) at some LZs, the use of zero-inflated count data models could be investigated for effective capturing of the overdispersion in the arrival rates. Therefore, the modeling framework developed in the paper is a big step towards advancing urban freight research while opening up promising opportunities for future research. This research also sheds light on the effectiveness of using cellphone-based applications, not only for better parking management at loading zones but also for obtaining high quality data for research purposes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was made possible through the financial support from the Fordonsstrategisk Forskning och Innovation (FFI) from VINNOVA, the Urban Freight Platform from the Volvo Research and Education Foundation (VREF), and the Chalmers University of Technology. The authors also acknowledge the support and guidance from the ParkUnload team with the data and modeling efforts.

Appendix

A. Cross-Validation (10-fold) Results for Separate and Joint Models.
See Table 9.
B. Actual vs Predicted Arrival Rates by Loading Zone.
Se Fig. 10.

Table 9
Root Mean Squared Errors (RMSE) for Testing data.

Fold	RMSE			
	Separate Models		Joint Model	
	Arrival	Duration	Arrival	Duration
1	2.27	11.16	2.10	10.89
2	2.28	11.20	2.14	11.03
3	2.36	11.13	2.09	11.08
4	2.30	11.23	2.09	10.88
5	2.36	11.11	2.07	11.18
6	2.29	11.15	2.07	10.90
7	2.38	11.24	2.11	11.12
8	2.39	11.24	2.13	11.01
9	2.39	11.18	2.11	11.07
10	2.29	11.34	2.10	11.10
Mean	2.33	11.20	2.10	11.03
Std. Dev	0.05	0.06	0.02	0.10

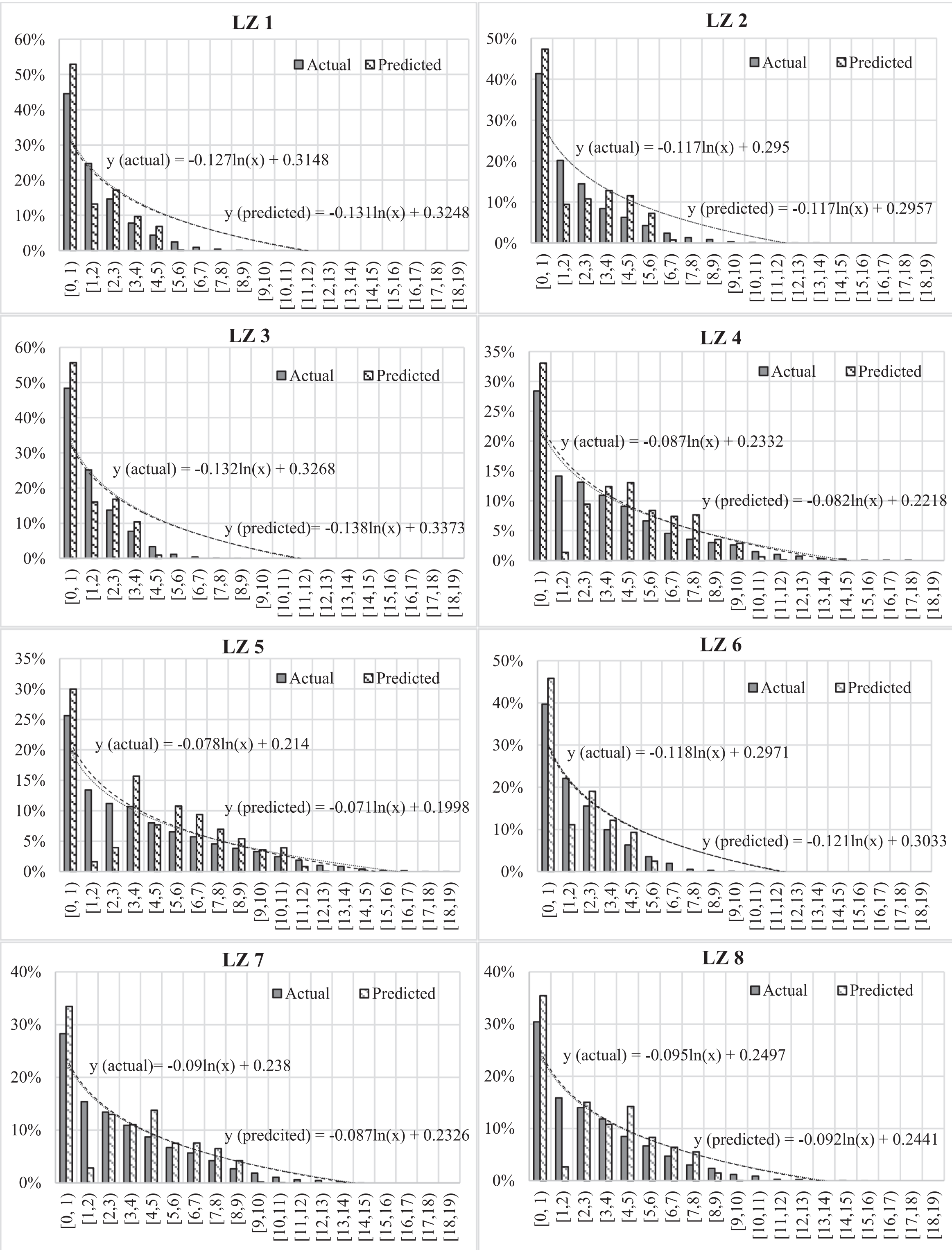

Fig. 10. Predicted vs Observed Arrival Rates by LZ.

References

Abu-Mostafa, Y., M. Magdon-Ismail and H. Lin (2012). “Learning from Data Vol. 4: AMLBook New York.”.

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. Selected Papers of Hirotugu Akaike. E. Parzen, K. Tanabe and G. Kitagawa. New York, NY, Springer New York: 199-213.

Alho, A. and J. de Abreu e Silva (2014). “Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics.” Transportation Research Record: Journal of the Transportation Research Board(2411): 45-54.

Alho, A. R., J. de Abreu e Silva, J. P. de Sousa and E. Blanco (2018). “Improving Mobility by Optimizing the Number, Location and Usage of Loading/Unloading Bays for Urban Freight Vehicles.” Transportation Research Part D: Transport and Environment **61**: 3-18.

Allen, J., Anderson, S., Browne, M., Jones, P., 2000. A Framework for Considering Policies to Encourage Sustainable Urban Freight Traffic and Goods/Service Flows. University of Westminster, London, Transport Studies Group.

Allen, J., Browne, M., Woodburn, A., Leonardi, J., 2012. The Role of Urban Consolidation Centres in Sustainable Freight Transport. Transport Reviews 32 (4), 473–490.

Allen, J., G. Thorne and M. Browne (2007). “BESTUFS Good Practice Guide on Urban Freight Transport.”.

Allison, P.D., Waterman, R.P., 2002. Fixed–Effects Negative Binomial Regression Models. Sociological Methodology 32 (1), 247–265.

Benenson, I., K. Martens, S. J. C. Birfir, Environment and U. Systems (2008). “PARKAGENT: An Agent-Based Model of Parking in the City.” **32**(6): 431-439.

BESTUFS, I. (2007). Good Practice Guide on Urban Freight Transport, BESTUFS, University of Westminster London, United Kingdom.

Bolay, J.C., 2020. Urban Facts. Urbanization: A Global Trend. Springer International Publishing:, Cham, pp. 7–55.

Box-Steffensmeier, J.M., Jones, B.S., Box-Steffensmeier, J.M., 2004. Event History Modeling: A Guide for Social Scientists. Cambridge University Press.

Butrina, P., G. D. C. Girón-Valderrama, J. L. Machado-León, A. Goodchild and P. C. Ayyalasomayajula (2017). “From the Last Mile to the Last 800 ft:Key Factors in Urban Pickup and Delivery of Goods.” **2609**(1): 85-92.

Cambridge Systematics Inc. (2007). Increase Cost of Parking in the Manhattan Central Business District (CBD). Technical Memorandum prepared for New York City Economic Development Corporation and the New York City Department of Transportation. New York: 1-58.

Campbell, S., Holguín-Veras, J., Ramirez-Rios, D., Gonzalez-Calderon, C., Kalahasthi, L., Wojtowicz, J., 2018. Freight and Service Parking Needs and the Role of Demand Management. European Transport Research Review 10 (47), 1–13.

Castrellon, J.P., Sanchez-Diaz, I., Gil, J., Smart loading zones. a data analytics approach for infrastructure network design in Freight Curbside Management. Available at SSRN: https://ssrn.com/abstract=4021963.

Chatterjee, A., Varma, A., Fischer, A., Swenson, J., 2008. Curbside Delivery of Freight by Trucks in Downtowns of Small- and Medium-Sized Urban Areas. Institute of Transportation Engineers. ITE Journal 78 (1), 32–42.

CLARS. (2020). “Urban Access Regulations in Europe.” Retrieved November 9, 2020, from https://urbanaccessregulations.eu/.

Cowling, B.J., Hutton, J.L., Shaw, J.E.H., 2006. Joint Modelling of Event Counts and Survival Times. Journal of the Royal Statistical Society. Series C (Applied Statistics) 55 (1), 31–39.

Cramér, H., 1946. A contribution to the theory of statistical estimation. Scandinavian Actuarial Journal 1946 (1), 85–94.

Dalla Chiara, G., Cheah, L., Lima Azevedo, C., Ben-Akiva, M.E., 2020. A Policy-Sensitive Model of Parking Choice for Commercial Vehicles in Urban Areas. 54 (3), 606–630.

Dey, S. S., B. O. Pérez, L. Richards, A. Pochowski, M. Sanders, M. Darst, E. C. Sanchez and S. Dock (2019). “Demystifying Urban Curbside Freight Management: Strategic Incremental Approach from Washington, D.C.” **2673**(12): 312-326.

Dieussaert, K., K. S. Aerts, T. and S. S. Maerivoet, K. (2009). SUSTAPARK: An Agent-based Model for Simulating Parking Search. 12th AGILE International Conference on Geographic Information Science. Leibniz, Germany.

Enzmann, D. (2015). NB_ADJUST: Stata Module to Identify and Adjust Outliers of a Variable Assumed to Follow a Negative Binomial Distribution.

European Commission. (2020). “Vehicle Categories.” Internal Market, Industry, Entrepreneurship and SMEs Retrieved Novemeber 30, 2020, from https://ec.europa.eu/growth/sectors/automotive/vehicle-categories/.

Hausman, J., Hall, B., Griliches, Z., 1984. Econometric Models for Count Data with an Application to the Patents-R&D Relationship. National Bureau of Economic Research Inc.

Holguín-Veras, J., Amaya-Leal, J., Wojtowicz, J., Jaller, M., González-Calderón, C., Sánchez-Díaz, I., Wang, X., Haake, D., Rhodes, S., Hodge, S.D., Frazier, R.J., Nick, M.K., Dack, J., Casinelli, L., Browne, M., 2015. Improving Freight System Performance in Metropolitan Areas. National Cooperative Freight Research Program. Washington, D.C., Transportation Research Board, pp. 1–212.

Holguín-Veras, J., Kalahasthi, L., Ramirez-Rios, D.G., 2021. Service Trip Attraction in Commercial Establishments. Transportation Research Part E: Logistics and Transportation Review 149 (102301).

Ibeas, Á., Cordera, R., dell'Olio, L., Moura, J.L., 2011. Modelling Demand in Restricted Parking Zones. Transportation Research Part A: Policy and Practice 45 (6), 485–498.

Jaller, M., Holguín-Veras, J., Hodge, S.D., 2013. Parking in The City: Challenges For Freight Traffic. Transportation Research Record 2379, 46–56.

James, G.S., 1954. Tests of Linear Hypotheses in Univeriate and Multivariate Analysis when the Ratios of the Population Variances are Unknown. Biometrika 41 (1–2), 19–43.

Jones, E., Chatterjee, A., Marsili Jr, R.L., 2009. A Collaborative Plan for Curbside Freight Delivery in Washington, DC, USA. ITE Journal (Institute of Transportation Engineers) 79 (5), 22–25.

Kalahasthi, L., I. Sanchez-Diaz, J. P. Castrellon, G. Jorge, M. Browne, B. Kulcsár, S. Hayes and C. Sentis Ros (2021). Commercial Vehicle Routing Model with Parking Chance Constraints (working paper).

Kaplan, E.L., Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 53 (282), 457–481.

Kladeftiras, M., Antoniou, C., 2013. Simulation-Based Assessment of Double-Parking Impacts on Traffic and Environmental Conditions. Transportation Research Record 2390 (1), 121–130.

Lambert, P.C., Royston, P., 2009. Further Development of Flexible Parametric Models for Survival Analysis. The Stata Journal 9 (2), 265–290.

Lindholm, M., Behrends, S., 2012. Challenges in Urban Freight Transport Planning – A Review in the Baltic Sea Region. Journal of Transport Geography 22, 129–136.

Low, R., Tekler, Z.D., Cheah, L., 2020. Predicting Commercial Vehicle Parking Duration using Generative Adversarial Multiple Imputation Networks. 2674 (9), 820–831.

Malik, L., Sánchez-Díaz, I., Tiwari, G., Woxenius, J., 2017. Urban Freight-Parking Practices: The Cases of Gothenburg (Sweden) and Delhi (India). Research in Transportation Business & Management 24, 37–48.

Marcot, B.G., Hanea, A.M., 2021. What is an Optimal Value of K in K-Fold Cross-Validation in Discrete Bayesian Network Analysis? Computational Statistics 36 (3), 2009–2031.

Marcucci, E., Gatta, V., Scaccia, L., 2015. Urban Freight, Parking and Pricing Policies: An Evaluation from a Transport Provider's Perspective. Transportation Research Part A: Policy and Practice 74, 239–249.

Marsden, G., 2006. The Evidence Base for Parking Policies-A Review. Transport Policy 13 (6), 447–457.

Marujo, L.G., Goes, G.V., D'Agosto, M.A., Ferreira, A.F., Winkenbach, M., Bandeira, R.A.M., 2018. Assessing the Sustainability of Mobile Depots: The Case of Urban Freight Distribution in Rio de Janeiro. Transportation Research Part D: Transport and Environment 62, 256–267.

Muñuzuri, J., Larrañeta, J., Onieva, L., Cortés, P., 2005. Solutions Applicable by Local Administrations for Urban Logistics Improvement. Cities 22 (1), 15–28.

Newson, R.B., 2010. Comparing the Predictive Powers of Survival Models using Harrell's C or Somers' D. The Stata Journal 10 (3), 339–358.

Nourinejad, M., Wenneman, A., Nurul Habib, K., Roorda, M.J., 2014. Truck Parking in Urban Areas: Application of Choice Modelling within Traffic Microsimulation. Transportation Research Part A 64, 54–64.

Parkunload. (2020). “Parkunload.” Retrieved November 11, 2020, from https://www.parkunload.com/.

Pearson, K., 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50 (302), 157–175.

Pitblado, J., 2013. Generalizing SEM in Stata. 2013 Stata Conference, Stata Users Group.

Regal Ludowieg, A., I. Sánchez-Díaz and L. Kalahasthi (2020). "Using Machine Learning to Predict Freight Vehicles Demand for Loading Zones in Urban Environments." Transportation Research Record: Journal of Transportation Research Board **in review**.

Roca-Riu, M., Cao, J., Dakic, I., Menendez, M., 2017. Designing Dynamic Delivery Parking Spots in Urban Areas by Minimizing Traffic Disruptions - Analytical and Simulation Approach. Ascona, Switzerland.

Sadek, B.A., Martin, E.W., Shaheen, S.A., 2020. Forecasting Truck Parking Using Fourier Transformations. 146 (8), 05020006.

Sanchez-Diaz, I., Palacios-Argüello, L., Levandi, A., Mardberg, J., Basso, R., 2020. A Time-Efficiency Study of Medium-Duty Trucks Delivering in Urban Environments. Sustainability 12 (1), 425.

Schmid, J., Wang, X., Conway, A., 2018. Commercial Vehicle Parking Duration in New York City and its Implications for Planning. Transportation Research Part A: Policy and Practice 116, 580–590.

Schoemaker, J., Allen, J., Huschebek, M., Monigl, J., 2006. Quantification of Urban Freight Transport Effects I. BESTUFS Consortium, NEA, University of Westminster, PTV. Transman. 76.

StataCorp, L., 2007. Stata Multivariate Statistics: Reference Manual. Stata Press Publication.

Statistical Institute of Catalonia, 2020. The Municipality in Figures: Vic. Retrieved November 11, 2020, from. https://www.idescat.cat/emex/?lang=en&id=082981#h3ff8000000000**.

Tögel, M., Špička, L., 2014. Low-Emission Zones in European Countries. Transactions on Transport Sciences 7 (3), 97–108.

Transmodal Limited, M. D. S. (2012). Study on Urban Freight Transport. DG MOVE European Commission. Chester, UK, Centro di ricerca per il Trasporto e la Logistica (CTL),: 154.

Vlahogianni, E.I., Kepaptsoglou, K., Tsetsos, V., Karlaftis, M.G., 2016. A Real-Time Parking Prediction System for Smart Cities. Journal of Intelligent Transportation Systems 20 (2), 192–204.

Wang, H. and N. J. Garber (2003). "Estimation of the Demand for Commercial Truck Parking on Interstate Highways in Virginia.".

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2009. Statistical and Econometric Methods for Transportation Data Analysis. Chapman & Hall/CRC.

Wu, L., Meng, Y., Kong, X., Zou, Y., 2022. Incorporating Survival Analysis into the Safety Effectiveness Evaluation of Treatments: Jointly Modeling Crash Counts and Time Intervals between Crashes. Journal of Transportation Safety & Security 14 (2), 338–358.