

Research Article

Freight Demand Distribution in a Suburban Area: Calibration of an Acquisition Model with Floating Car Data

Giuseppe Battaglia,¹ Giuseppe Musolino ,² and Antonino Vitetta ²

¹Engineering Degree, Università Mediterranea di Reggio Calabria, Feo di Vito, Reggio Calabria 89122, Italy

²Università Mediterranea di Reggio Calabria, Feo di Vito, Reggio Calabria 89122, Italy

Correspondence should be addressed to Giuseppe Musolino; giuseppe.musolino@unirc.it

Received 23 September 2021; Revised 4 January 2022; Accepted 21 January 2022; Published 23 February 2022

Academic Editor: Jaeyoung Lee

Copyright © 2022 Giuseppe Battaglia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper deals with the interaction between freight flows, land use, and economy of a suburban area. In urban freight distribution, the process of freight acquisition of commercial establishments represents a crucial element. Therefore, it is necessary to plan city logistics in order to reduce the negative externalities generated by freight transport without depressing the economic and social vitality. In order to estimate and evaluate the performance and impacts generated by urban freight distribution, Transport System Models (TSMs) play an important role in reproducing and forecasting the process of distribution of final products to retailers and consumers. In recent years, Information and Communication Technologies (ICTs) have made new sources of data in great amounts available at lower costs than data provided from traditional surveys. Among the different data sources, Floating Car Data (FCD), obtained through GPS trackers installed on board the vehicles, offer partial insights into the examined process with lower cost and provide continuous information in space and time. In order to obtain reliable results, it is necessary to combine TSMs and ICTs as observed data could feed Transport System Models and could allow their calibration and therefore their application in real cases. In the above context, the paper deals with TSMs related to estimating urban freight demand flows. The work carried out concerns the specification-calibration-validation of a freight zonal acquisition model by means of FCD related to interzonal trips of commercial vehicles inside the Locride area (Reggio Calabria, Italy). The obtained models, estimated from FCD, could support the a priori evaluation of city logistics measures. The results are more important considering that they are obtained in an area with a lack of transport monitoring technologies. They could be adopted in other similar areas and confirm that adequate planning of city logistics determines advantages for environmental, social, and economic sustainability.

1. Introduction

The number of freight vehicles travelling inside urban areas has increased considerably in recent years due to changes in retailers' patterns (e.g., elimination of stocks, just-in-time) and in citizens' habits (e.g., e-commerce). This element generates impacts on the three pillars of sustainability [1]:

- (i) On the environmental side, there have been increasing levels of congestion air and noise pollution.
- (ii) On the economic side, there have been changes in logistics costs and, thus, in the price of products.
- (iii) On the social side, there have been increasing risks of accidents due to the presence of different types of vehicles on the roads.

On the other hand, several city logistics measures have been implemented and tested in many cities during the last two decades [1, 2]. In many cases, implemented measures have resulted to be unable to reach the expected sustainability goals because many of them were introduced without an ex ante quantitative evaluation. Transport System Models (TSMs) play an important role in estimating and evaluating the performance and impacts generated by city logistics measures. TSMs should be able to assess, forecast, and evaluate the process of last-mile distribution of final products [3–5].

The specification, calibration, and validation of freight TSMs are based on the observations obtained with two main approaches (more details are reported in Section 2). The first approach is related to the traditional disaggregate surveys on

operators' patterns, similar to the approaches adopted for the passengers' demand models. The second, more recent, is based on the use of the so-called big data provided by Information and Communication Technologies (ICTs) tools. In particular, the use of Floating Car Data (FCD), obtained through GPS installed on board, allows us to identify the space-time trajectories of freight vehicles during the last-mile delivery. When a great amount of data is required, the second approach is less expensive than the first approach. FCD is one of the less expensive systems to obtain traffic data.

The observed data are necessary to apply the specification-calibration-validation process of TSMs.

Further research on methods and models is necessary to meet the following requirements [6]:

- (i) The increase of the volume of transported goods on the roads and the reorganization of the last-mile logistic chain, imposed by the new patterns of retailers and citizens, require new approaches for modelling the city logistics process.
- (ii) The massive diffusion of emerging ICT tools made a great amount of data (e.g., FCD) available, which is able to support freight demand models' definition and requires new models and/or the validations of existing models.

The objective and the research contributions of this paper concern the test of the accuracy levels of the freight demand models calibrated using FCD, with a focus on the acquisition level of freight vehicles in zones where commercial establishments are present. The research presented in this paper aims to verify whether FCD can support the following:

- (i) Quantitative analyses of mobility of freight vehicles.
- (ii) Estimation of freight flow models in the current state and/or in future scenarios in areas characterized by the absence (or scarcity) of other transport monitoring technologies.

In many areas of the world, FCD are the main data available for supporting planning decisions about the transport system, traffic monitoring systems are not available, and there is a lack of transport monitoring technologies. TSMs need to be specified, calibrated, and validated with low-cost investments in these areas, and FCD seem to be suitable for achieving these results.

According to the recent trends in the use of FCD, this paper presents a descriptive model that has been specified, calibrated, and validated. The number of vehicles has been considered as the reference unit in the model.

This paper is structured as follows. Section 2 reports the state of the art related to freight demand models in urban areas. Section 3 proposes the specification of a model for the zonal acquisition of freight at suburban scale. Section 4 shows the results of the proposed freight demand model, which is calibrated for the Locride area (Reggio Calabria, Italy). Section 5 reports the main conclusions and the research perspectives.

2. State of the Art

In the field of TSMs for urban freight distribution, although relevant studies were conducted before (see, e.g., [7]), the existing literature has been largely developed since the end of the 90s, when the first city logistics experiences were carried out in some Northern European cities. The technicians and researchers felt the need to have a corpus of evaluation models to support the ex ante definition of city logistics measures before being implemented in the cities.

The literature covers different elements of the urban freight distribution process, called city logistics: from the representation of the last-mile distribution chain on the supply models' side to the simulation of stakeholders' behaviour and interactions on the demand models' side. Detailed literature reviews are reported by Taniguchi and Thompson [8], Holguín-Veras et al. [9], and Russo and Comi [10].

Freight demand models can be specified, calibrated, and validated by using different sources of data and statistical tools. Classical interviews on a sample of stakeholders can be used, and the data obtained can be processed using inferential statistics techniques. This approach provides a direct estimation of freight demand flows. In addition, other approaches can be used to estimate current demand and forecast future transport demand. One of these approaches is named aggregate, which uses aggregated and indirect information about stakeholders' behaviour. Aggregated information provided by traffic counts can be used not only for freight demand flows but also to calibrate the parameters of freight demand models. This approach can also be applied when FCD are available.

This section reports a state of the art, without being exhaustive, concerning urban freight demand models recently developed. Table 1 shows a classification of a selection of papers in relation to different criteria adopted. The selection of literature papers concerns the use of FCD in supporting the analysis of urban freight distribution process and the estimation of freight demand models.

The *model structure* of the freight demand models may be direct or partial (all the analyzed papers are partial considering the model structure). The *reference unit* of freight flow can be expressed in terms of vehicles travelling on the road network or in terms of quantities of freight transported (e.g., tons). The *distribution channel* or the movement of freight can be among the warehouses, the distribution centers, and the retailers (traditional channel) or between the warehouses of the virtual shop and the end-user (e-commerce channel). The *aggregation level* of the variables used classifies models into aggregated or disaggregated, in relation to whether the attributes (variables) present refer to disaggregated units such as the individual user or to aggregated units such as a category of users or an economic sector. The *model assumption* concerns the typology of the model representation that can be behavioural or descriptive, considering whether or not the stakeholders' choices derive from the analysis of the behaviour of decision-makers (producers, retailers, consumers, transport, and logistic companies). The *observed data* may be provided by

TABLE 1: Classification of a selection of papers using big data to estimate freight demand models.

Reference	Model structure	Reference unit	Distribution channel	Aggregation level	Model assumptions	Observed data	Study area	Calibrated parameters
Russo and Comi [11]	P	V	R	A/D	B/D	T	U (Giarre, Italy)	Y
Ehmke et al. [12]	P	V	R	A	D	B	U (Stuttgart, Germany)	Y
Nuzzolo and Comi [13]	P	Q/V	R	A	D	B	U (Rome, Italy)	Y
Sharman [14]	P	V	R	A	D	B	R (Canada)	Y
Ben-Akiva et al. [15]	P	V	R	A	D	B/T	R (USA)	N
Hadavi et al. [16]	P	V	R/E	A	D	B	U (Brussels, Belgium)	N
Croce et al. [17]	P	Q/V	R	A/D	D	B/T	U (Locride, Italy)	Y
Diana et al. [18]	P	V	R	A	D	B/T	U (Turin, Italy)	N
Nuzzolo et al. [19]	P	V	R	A	D	B	R (Veneto, Italy)	N
Comi and Polimeni [3]	P	V	R	A	D	B	R (Veneto, Italy)	Y
Musolino et al. [20]	P	V	R	D	B	B	U (Locride, Italy)	Y
Wang et al. [21]	P	Q	E	D	D	B	E-commerce platform	Y
Russo and Comi [10]	P	V	R/E	A	D	B	Rome (Italy)	N

P, partial choice. V, vehicles; Q, quantity. R, one or more retailers; E, e-commerce. A, aggregate; D, disaggregate. B, behavioural; D, descriptive. T, traditional survey; B, big data. U, urban/suburban; R: regional. Y, presence of calibrated parameters; N, no calibrated parameters

traditional sources or by ICT sources (e.g., Floating Car Data (FCD)); in order to facilitate the estimation of freight demand flows and to build new models, several papers have been published regarding the role of big data (in many cases FCD) in supporting the definition of freight demand models. Recently, there has been much consideration to the movement of freight vehicles determined by e-commerce. The *study area* within the transport system is located, maybe urban, suburban, or regional. Finally, in relation to the *calibrated parameters*, the selected papers include models with calibrated parameters and without calibrated parameters.

The set of papers considered have the following common characteristics:

- (i) They have a partial model structure.
- (ii) The vehicle is considered a reference unit in the majority of the cases, given that observed data come from ICT tools.
- (iii) One or more retailers are almost always present in the distribution channel.
- (iv) An aggregate level and a descriptive assumption for the model are considered in the majority of the cases.
- (v) The models are calibrated by means of big data in almost all cases.

A brief description of each selected paper is presented in Table 1.

Russo and Comi [11] estimated freight flows in terms of both quantities of goods consumed and the number of freight vehicles. They were the first authors who proposed a multistage model based on a disaggregated approach for each decision-making level. Ehmke et al. [12] considered the influence of congestion on the road network in the freight distribution process. They also considered the time taken to complete a delivery round, in addition to the distance travelled. Three scenarios are simulated in relation to the freight trip origin destinations: movements originated in peripheral areas and with a destination in the city center, journeys from/to the city center, and journeys from/to points outside the city center. Nuzzolo and Comi [13] analyzed the actors involved in the movement of vehicles used for the transport and delivery of goods within a study area. They reported the scheme for distribution centers, wholesalers, retailers, and end buyers. In order to be able to model these types of movement, the authors proposed a model related to delivery cycles, generating an O/D matrix. The study provided calibrated data related to the study area of Rome. Sharman [14] conducted a study based on data collected via GPS devices. Three types of travel destinations have been identified: frequently visited destination, regular programming target, and unscheduled destination. In addition, the author used a discrete choice approach, using both Probit and Logit models. Since only trip data were available, it was not possible to associate data on the type of freight transported or on the type of service. Ben-Akiva et al. [15] used evaluation and planning models with

traditional surveys, applied to freight transport. Data analysis showed that drivers followed four behaviour styles. The study focused on the route choice in the presence of congestion on the road network. Hadavi et al. [16] considered the freight deliveries inside Brussels, acquiring GPS track data over a 28-day period. The data relating to the heavy freight vehicles used for the deliveries have been selected; each route has been associated with the data relating to the mass transported and the “EURO” regulation, to which the vehicle refers in order to estimate the environmental impact of the delivery. The authors used aggregated data. Croce et al. [17] integrated the data provided by traditional transport surveys with the big data provided by ICT for the construction of TSMs. The need to integrate these two data sources was born for improving the forecasting capacity of TSMs and for reducing the costs for specification, calibration, and validation process. Diana et al. [18] considered the routes, tracked by GPS on board, of the vans used for delivery within the city of Turin. The database generated associates to each route, vehicle, and fuel consumption data. The analysis aimed to study the freight vehicle delivery process by using an origin/destination matrix. Nuzzolo et al. [19] carried out a simulation model for the freight vehicles’ movements, considering the structure of the distribution chain and of the sales centers at a regional scale. The goal was the construction of an O/D matrix, where the origins are the places where the warehouses are located, and the destination is the places where the users are located (retailers). The authors used aggregated data to construct a multilevel model, based on the FCD database, considering more than 30.000 deliveries routes operated by 1579 vehicles within the Veneto region (Italy). The aim was to provide a transport model capable of simulating delivery cycles of freight vehicles, using aggregated data, including the socioeconomic attributes of zones and attributes about the existing logistic structures. Comi and Polimeni [3] investigated the potentialities offered by FCD to infer delivering activities. They calibrated a discrete trip-chain model for light goods vehicles, which describes delivery tours in terms of the number of stops/deliveries performed. The data used concern light goods vehicles operating in the Veneto region (Italy). Musolino et al. [20] proposed a model that integrates the path choice and vehicle routing levels through methodological and experimentation activities, the choice set was generated with a multicriteria approach, and a route choice model was calibrated. The research of Wang et al. [21] aimed to gain behavioural insights into online shopping in light of the Coronavirus disease by exploring the factors affecting the initial adoption and continuance intention of using delivery services by consumers. The calibrated model revealed that there are differences not only between the different identified types of consumers (i.e., the prior adopter, temporary adopter and permanent new adopter, and nonadopter) but also between the four types of products (i.e., grocery, food, home goods, and other packages) considered in the survey. Russo and Comi [10] proposed a unified formulation of the basic problem of fleets, like the Travelling Salesman Problem (TSP), which explicitly

includes the use of emerging ICTs (e-ICTs), pointing out the learning process of path costs in urban delivery. This research explored the opportunity to extend the path cost formation with a within-day and day-to-day learning process, including the specification of the attributes provided by e-ICTs.

The state of the art regarding the use of FCD in the analysis of urban freight distribution process showed that this source of data could provide great support in the estimation of the modelling components of freight demand, which deal mainly with the vehicle level. As a matter of fact, FCD provide information related to road vehicles, not to users’ behaviour. FCD provide elements about vehicle-based freight flows, while elements related to commodity-based levels are missing. In order to obtain insights into the entire supply chain, FCD could be integrated with specific information about transport operators and clients (e.g., type of delivered freight, explicit temporal and spatial constraints, and specific clients’ needs).

According to the above considerations, the significance of this research contribution concerns the calibration of an acquisition model, by means of FCD, of freight vehicles in zones where commercial establishments are present.

3. Methodology

In the context of specification-calibration-validation of freight demand models, this section deals with a model that describes part of the supplying process of freight by urban commercial establishments present in an urban area, which is called in the literature “*acquisition model*” (see [19] and the references included). In particular, the focus is on the evaluation of the number of freight vehicles having as destination zones where commercial establishments are present.

3.1. Definitions and Notations. The following definitions and notations are adopted for the model (see Figure 1).

The zones of the study area, in relation to freight flows, are defined as follows:

- (i) r , internal residential zone (residents), internal zone to the study area where the end-consumer consumes the goods and where the residential activities are located.
- (ii) c , internal commercial zone (retailers), internal zone to the study area where the end-consumer can purchase the goods which the retailers sell; the shops are located in these zones.
- (iii) w , internal warehouse zone (wholesalers), internal zone to the study area from where the retailer purchases goods to be sold in his/her shop.
- (iv) z , external warehouse/production zone (wholesalers), zone outside the study area where the end-consumer can purchase the goods or where the (internal) retailer can purchase goods to be sold in his/her shop.

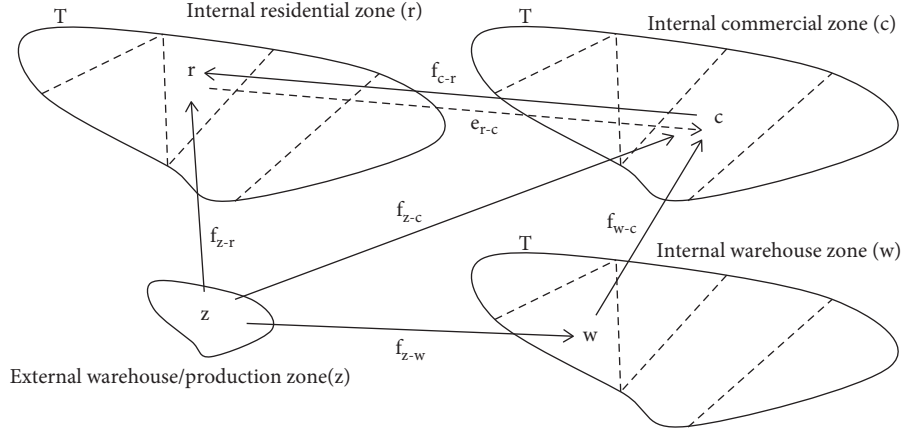


FIGURE 1: Zones, freight flows, and passenger flows relationships between zones of the study area.

The freight flows between the zones are defined as follows:

- (v) f_{c-r} , freight attraction flow from commercial zone c to residential zone r .
- (vi) f_{w-c} , freight acquisition flow from internal warehouse zone w to internal commercial zone c .
- (vii) $f_{z-c(w,r)}$, freight acquisition flow from external warehouse zone z to internal commercial zone c (to internal warehouse zone w or to internal residential zone r).
- (viii) d_r , freight demand attracted in zone r , obtained as $d_r = \sum_w f_{w-r}$.
- (ix) d_c , freight demand acquisition in zone c , obtained as $d_c = \sum_w f_{w-c}$.

The freight demand acquired in the study area is

- (i) $d = \sum_c d_c$.

The freight flows generate passenger flows between the zones, defined as follows:

- (ii) e_{r-c} , passenger flow from internal residential zone r to internal commercial zone c .
- (iii) $e_c = \sum_r e_{r-c}$, passenger demand attracted in internal commercial zone c for purchase purpose.

3.2. Model Specification and Calibration. The proposed model is an aggregate discrete choice model of freight acquisition of an internal commercial zone c , d_c (see Figure 1).

Assuming that the total freight demand d acquired in the study area is given, the freight demand acquired in the internal commercial zone c , d_c , is obtained as

$$d_c = p_c(\boldsymbol{\beta}) \cdot d, \quad (1)$$

with

- (i) p_c , percentage of freight vehicles supplying commercial zone c , or probability that a wholesaler supplies commercial zone c , to be evaluated with a discrete choice model.
- (ii) $\boldsymbol{\beta}$, unknown vector of parameter, to be calibrated (or updated).

- (iii) d , total freight demand acquired in the study area.

According to the random utility theory [22, 24], a homogenous set of wholesaler i chooses to supply a commercial zone c , among the set of commercial zones of the study area T , $c \in T$, maximizing his/her associated perceived utility, U_c^i .

As the utility, $U_c^i(\boldsymbol{\beta})$, is a random variable, it is not possible to establish with certainty which zone will be chosen by the wholesaler i to supply. The probability that zone $c \in T$ will be chosen from the homogenous set of wholesaler i may be expressed as the probability that the perceived utility of commercial zone c , U_c^i , is greater than the perceived utility of every other commercial zone h , U_h^i :

$$p_c(\boldsymbol{\beta}) = \text{Prob}(U_c^i(\boldsymbol{\beta}) > U_h^i(\boldsymbol{\beta})) \forall h \neq c, c \in T, \quad (2)$$

where T is the choice set or the set of available alternatives (commercial zones) for the wholesaler.

It is worth noting that the apex i is omitted in the probability for simplicity sake. The model refers to a homogenous set of wholesaler i .

Assuming that the perceived utility, $U_c^i(\boldsymbol{\beta})$, is independently and identically distributed (i.i.d.) according to a Gumbel across the alternatives, equation (2) is specified with a multinomial Logit. Different model specifications are recently proposed, based on different distributions for the utility (i.e., Fuzzy or Quantum Utility models), for a different level of choice [23].

The probability of choosing the commercial zone c as a zone to be supplied, function of the unknown vector of parameters to be calibrated, $\boldsymbol{\beta}$, is

$$p_c(\boldsymbol{\beta}) = \frac{\exp(V_c(\boldsymbol{\beta}))}{\left(\sum_{h \in T} \exp(V_h(\boldsymbol{\beta}))\right)}, \quad (3)$$

where $V_c(\boldsymbol{\beta}) = \beta_x^T \mathbf{x}_c + \beta_y^T \mathbf{y}_c$ is the expected value of $U_c^i(\boldsymbol{\beta})$, considering that the parameter of the Gumbel probability distribution is included in the parameters; $\beta^T = [\beta_x^T, \beta_y^T]$ is the vector of the parameters to be calibrated (or updated); \mathbf{x}_c is the vector of the socioeconomic attributes of zone c ; \mathbf{y}_c is the vector of cost/(passive) accessibility attributes of zone c .

The vector of parameters, $\boldsymbol{\beta}$, is calibrated by means of an optimization model:

$$\beta^{\text{opt}} = \min_{\beta \in S_b} \phi(\beta), \quad (4)$$

where β^{opt} is the values of parameters belonging to vector β that minimize $\phi()$; $\phi()$ is the objective function, defined as a Nonlinear Generalized Least Square (NLGLS) estimator:

$$\phi(\beta) = k_\beta \cdot \sum_{k=1 \dots m} (\beta_k - \beta_k^*)^2 + k_d \cdot \sum_{d \in T} (p_c(\beta) - p_c^*)^2, \quad (5)$$

where β_k is the current value of parameter k belonging to vector β ; β_k^* is the initial value of parameter k belonging to vector β ; S_b is the feasibility set of the parameters; k_β is the weight associated with parameters; k_p is the weight associated with frequencies; $p_c^* = d^* / d^* [\%]$ is the observed frequencies of freight vehicles attracted in commercial zone c ; d_c^* is the observed freight flow expressed in vehicles/time attracted in commercial zone c (FCD attracted in c); d^* is the observed freight flow expressed in vehicles/time attracted in the study area.

The model specified with (3) is considered in this paper. The model has a Logit structure [24], and the expected value of utility depends on attractiveness (x_c) and cost attributes (y_c). The attributes of attractiveness are defined by means of socioeconomic variables; the cost attributes are defined by means of accessibility, or cost, variables.

The single attributes are weighted with the following parameters:

- (i) β_x is the vector of attractiveness attributes, which are greater than zero.
- (ii) β_y is the vector of cost attributes, and it is less than zero.

The probability evaluated with (3) increases with the increase of attractiveness attributes and with the decrease of cost attributes.

The parameters β_x and β_y have to be calibrated from observed data. The specification proposed refers to the classic NLGLS statistical theory. The model reported in (4), with the specification of (5), is adopted for the parameters' calibration. It minimizes the weighted (k_β, k_d) square error between the initial (β_k^*) and the estimated (β_k) parameters and between the estimated probability ($p_c(\beta)$) and the observed frequencies (p_c^*).

4. Experimentation

The proposed zonal acquisition model, specified in Section 3, has been specified, calibrated, and validated for the suburban area of Locride (Reggio Calabria, Italy).

4.1. Study Area and Data. The study area is the suburban area of Locride inside the Metropolitan City of Reggio Calabria, Southern Italy (Figure 2). A detailed description of the study area, the passenger trip patterns, and the freight distribution process in the study area is reported, respectively, in Malavenda et al. [25] and in Croce et al. [4]. The study area was subdivided into 17 zones. The zones are defined considering homogenous areas in terms of land use

and municipality boundaries. For the smaller municipalities (5) in terms of inhabitants, the zone's boundary is equal to the administrative boundary. For the bigger municipalities (6), the administrative area is subdivided into two zones in relation to the population, the land use and the dimension. An analysis of the existing transport infrastructures (roads) connecting the different zones was carried out. The road transport system is schematized by means of a graph approach, which describes the main space-time positions (nodes) and their connections (links) and relative cost, consistently with an objective to estimate freight and passenger vehicle flows affecting the study area.

A specific activity has been conducted to obtain the experimental data for model calibration, divided into different categories:

- (a) Socioeconomic data.
 - (b) Freight and passenger trip data (FCD).
 - (c) Transport cost/accessibility data (distances and travel times).
 - (d) Retailers' data.
- (a) Socioeconomic data are related to residential and economic activities within the examined study area. Table 2 shows data related to the study area and average and coefficient of variation of zonal data.
- (b) Freight and passenger trip data (Table 3) are obtained by processing available FCD concerning the study area (see [17]). The passenger vehicles data concern the number of passenger vehicles attracted in the zones of the study area during an average weekday, in the time period between 4:00 pm and 8:00 pm, on Saturdays and Sundays for two weeks in the winter period. The passengers' vehicles attracted in the zones during the above periods, or days, of the week are a proxy of the home-purchasing trips of the inhabitants of the study area. In particular, the number of passenger vehicles on a weekday (4:00 pm–8:00 pm) is used as an independent variable to estimate the potential of a zone to attract freight vehicle trips, as shown in the next sections. The freight vehicles data concern the number of freight vehicles attracted in the zones of the study area on an average working day. The FCD belong to commercial vans and small trucks of different companies. Data are distributed along the entire population. The number of average freight vehicles observed on a weekday is 249. The trips with both origin and destination external to the study area are not considered.
- (c) Transport cost (accessibility) data in terms of time and distance between each internal zone and external zones are related to interzonal travel times and distances along the minimum paths evaluated on the supply network. The average distances and times are calculated as follows. For the external southern and northern zones, the boundaries of the study area have been considered along the road "106-Jonica,"

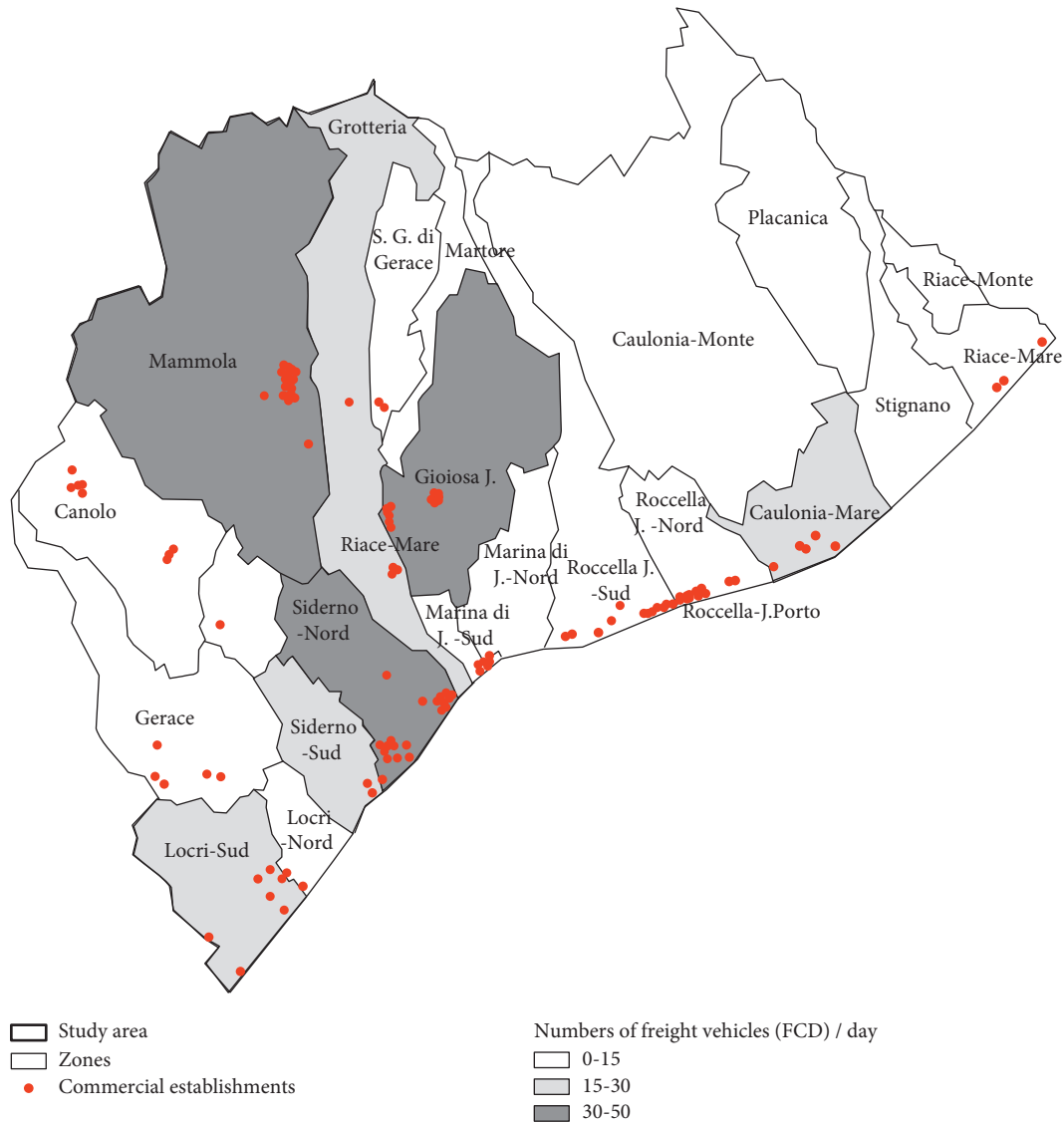


FIGURE 2: Study area, zoning, commercial establishments (red dots), and number of freight vehicles having a zone as destination per day.

which is the main road connecting the study area along the Jonian side of the Metropolitan City of Reggio Calabria. For the external western zones, the boundaries of the study area are considered along the road “682 Jonio-Tirreno,” connecting the Tyrrhenian side of the Metropolitan City of Reggio Calabria. For each internal zone, the average value of distance and time is calculated as an average between the distance and time for each couple. The transport costs are evaluated by means of a supply model with link cost functions previously estimated from FCD (see details in [26] and [5]). The user transport costs are common for all the users with the same origin, destination, and time of the day. Table 3 reports the average distance and time to reach each internal zone from the boundaries of the study area and the coefficient of variation between the zones.

(d) Data related to commercial establishments (Tables 4–6) concern the sale of food products,

subdivided into four groups: retail, hospitality, catering, and wholesale. The activities were further segmented in relation to the category of shop (ATECO coding), defined by the Italian National Institute of Statistics (ISTAT). The commercial establishments inside the study area have been identified one by one by consulting a public database on the web. Then, their dimensional characteristics, such as the surface of the shop, the number of shop windows, and the number of parking slots, have been estimated from the map assuming, on average, one slot every 20 m² of space for parking, including roads. Finally, each shop has been associated with each commercial zone. Concerning catering and hospitality activities, it was possible to identify further information by means of public web portals dedicated to clients of the facilities.

The number and the surface of commercial establishments of group 1 (retail) and of all groups (1, 2, 3, and 4)

TABLE 2: Aggregated data on socioeconomic attributes (for model calibration).

Attribute	Units of measurement	Study area	Zone	
			Average	Coefficient of variation
Number of commercial establishments	Number	129.00	7.59	0.95
Surface of commercial establishments	m ²	40 249.00	2 367.59	0.94
Number of households	Number	28 021.00	1 470.59	0.67
Population	Number	72 222.00	3 842.00	0.72
Population density	Number/km ²	105.32	105.32	0.96
Retail employees	Number	2 314.00	136.12	0.95
Wholesaler employees	Number	882.00	51.88	1.20
Total commercial employees	Number	2 415.00	218.29	1.13
Total employees	Number	14 271.00	813.12	1.10
Retail employment density	Number/km ²	12.78	12.78	1.56
Wholesale employment density	Number/km ²	4.70	4.70	1.59
Commercial employment density	Number/km ²	16.19	16.19	1.15
Total employment density	Number/km ²	64.34	64.34	1.30

TABLE 3: Average values of the cost attributes.

Attribute	Units of measurement	Average	Zone
			Coefficient of variation
Average distance	Km	26.83	0.12
Average time	Minutes	31.02	0.21

TABLE 4: Classification of commercial establishments.

Group	Type	Type	Product type	Car park	Surface
1	Retail	Store affiliated to LSR	Consumer multiproduct	No	Up to 400 m ²
		Store not affiliated to LSR	Consumer multiproduct	No	Up to 400 m ²
		Supermarket	Consumer multiproduct	Yes and No	Between 400 m ² –500 m ²
		Shop-emporium (bakery, butcher, fishmonger, greengrocer, and frozen food)	Consumer multiproduct belonging to a specific type	Yes and No	Not specified
2	Hospitality	Hotel, bed, and breakfast			
3	Catering	Bar/café			
		Restaurant, tavern, pizzeria, and takeaway			
4	Wholesale	Wholesale			

LSR, Large-Scale Retail.

TABLE 5: Retail data: number of activities.

Type	Study area	Zone	
		Average	Coefficient of variation
Store affiliated to LSR	11	2.00	0.7
Store not affiliated to LSR	6	1.16	0.4
Supermarket	7	1.14	0.3
Shop-emporium (bakery, butcher, fishmonger, greengrocer, and frozen food)	22	1.55	0.7
Hotel, bed, and breakfast	13	4.10	0.2
Bar/café	41	2.80	1.2
Restaurant, tavern, pizzeria, and takeaway	28	1.00	2.7
Wholesale	1	0.00	

LSR, Large-Scale Retail.

TABLE 6: Retail data: surface (m²).

Type	Study area	Zone	
		Average	Coefficient of variation
Store affiliated to LSR	2 671	668	0.6
Store not affiliated to LSR	1 316	219	0.6
Supermarket	5 645	806	0.4
Shop-emporium (bakery, butcher, fishmonger, greengrocer, and frozen food)	4 188	322	1.1
Hotel, bed, and breakfast	7 376	820	0.7
Bar/café	13 550	1 355	0.6
Restaurant, tavern, pizzeria, and takeaway	4 803	483	1.2
Wholesale	670	670	0.0

LSR, Large-Scale Retail.

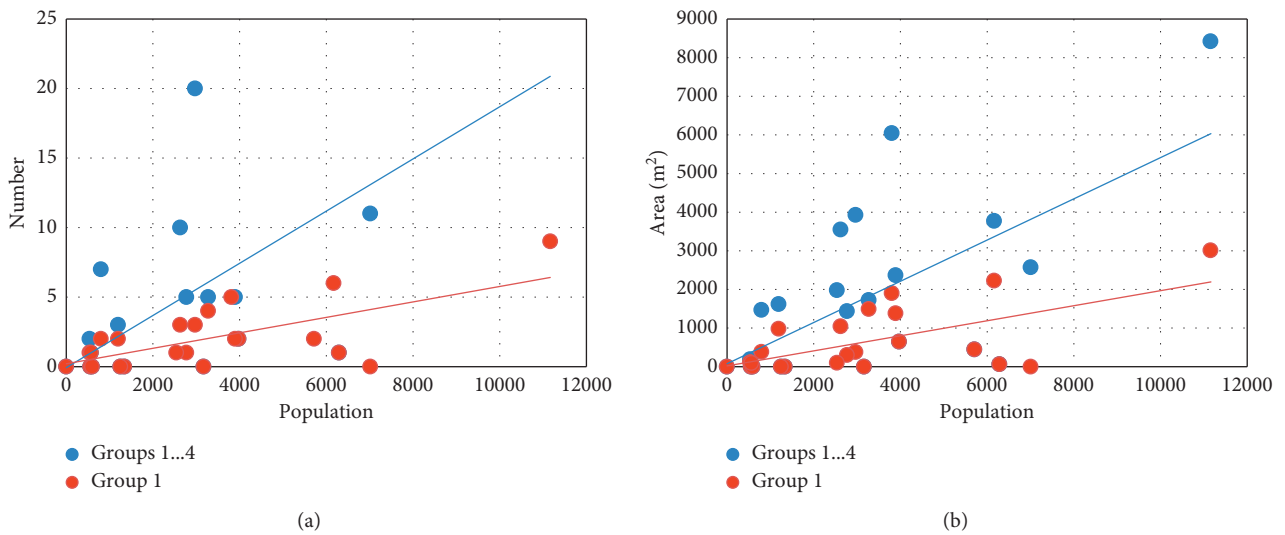


FIGURE 3: Number (a) and surface (b) of commercial activities versus population. Commercial activities of group 1 and of all groups 1, 2, 3, and 4.

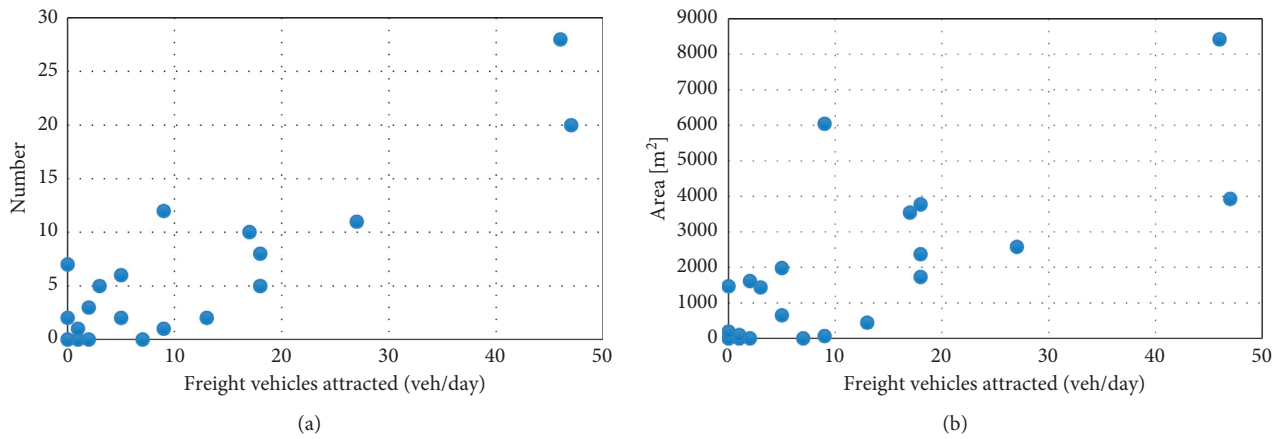


FIGURE 4: Number (a) and area (b) of commercial activities versus the number of attracted freight vehicles. Commercial activities of all groups.

TABLE 7: Values of calibrated parameters for equation (3): model specifications with direct variables.

ID	Number of commercial establishments (Num)	Activity area (m ²)	Number of households (Num)	Population >5000 (1/0)	Average accessibility Distance (km)	Time (minutes)	Number of passenger vehicles attracted (vehicles/day)	R ²
1	0.07279							0.67
2		0.00021						0.44
3			0.00052					0.46
4						-0.03902		0.07
5		0.00017		0.46961				0.49
6	0.07340						0.00012	0.68
7	0.06988					-0.01534		0.68
8		0.00019			-0.04339			0.45
9		0.00020				-0.01694		0.44
10				0.65946		-0.01643		0.10
11				0.60067	-0.04836			0.17

TABLE 8: Values of calibrated parameters for equation (3): model specifications with transformed variables.

ID	Ln (number of households)	Ln (average distance accessibility)	R ²
12	1.069804		0.47
13		-1.90510	0.08

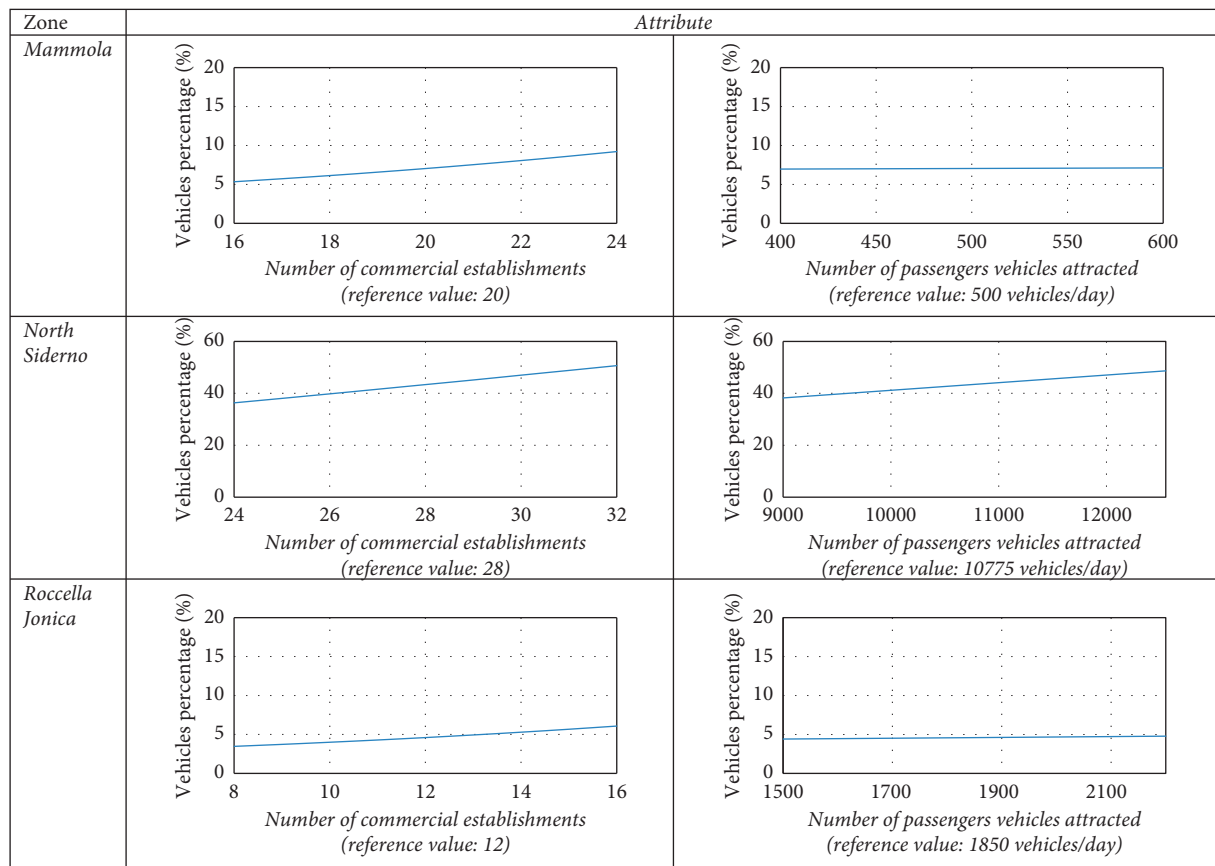


FIGURE 5: Probability of zonal acquisition (or percentage of freight vehicles having a zone as destination) estimated by model 6. Examples for zones of Mammola, Siderno, and Roccella Jonica.

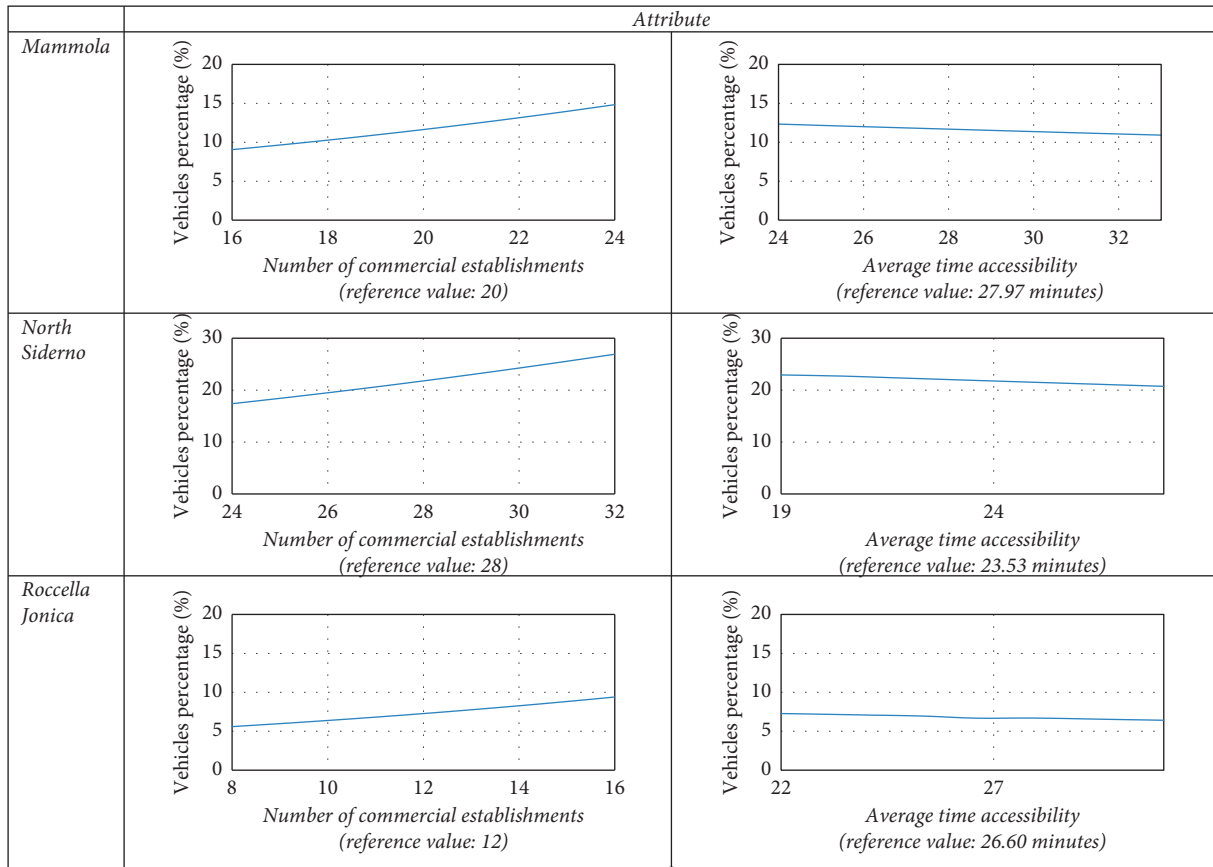


FIGURE 6: Probability of zonal acquisition (or percentage of freight vehicles having a zone as destination) estimated by model 7. Examples for zones of Mammola, Siderno, and Roccella Jonica.

TABLE 9: Maximum, minimum, and mean values of model attributes.

	Attributes		
	Number of commercial establishments	Average time accessibility (min)	Number of passenger vehicles attracted (veh/day)
Maximum	28.0	31.8	10725
Minimum	1.0	21.3	50
Mean	7.6	26.8	2843

TABLE 10: Values of model attributes for the selected zones.

Zone	Attributes		
	Number of commercial establishments	Average time accessibility (min)	Number of passenger vehicles attracted (veh/day)
Mammola	20	27.9	500
Northern Siderno	28	23.5	10725
Agnana Calabria	1	31.8	150
Roccella Jonica	12	26.6	1850
Southern M. di Gioiosa Jonica	6	21.3	5300

TABLE 11: Elasticity of models 7 and 8 versus increases of attribute values of +10% and +5%.

	Zone	Model 6		Model 7	
		Number of commercial establishments	Number of passenger vehicles attracted (veh/day)	Number of commercial establishments	Average time accessibility (min)
<i>Increment 10%</i>	Mammola	1.5	0.4	1.6	-0.1
	Northern siderno	2.2	0.4	2.3	-1.4
	Agnana Calabria	0.1	0.5	0.1	-0.02
	Roccella Jonica	0.9	0.4	0.9	-0.2
	Southern M. di Gioiosa Jonica	0.4	0.3	0.5	-0.7
<i>Increment 5%</i>	Mammola	0.0	0.0	1.5	-0.1
	Northern siderno	2.1	0.4	2.2	-1.3
	Agnana Calabria	0.1	0.5	0.1	-0.02
	Roccella Jonica	0.9	0.4	0.9	-0.2
	Southern M. di Gioiosa Jonica	0.4	0.3	0.4	-0.6

versus population of each zone are presented, respectively, in Figure 3. The figures show that both the number and the surface of the retail activities on average increase as the population of the zone increases. The number and surface of commercial activities of all groups (1, 2, 3, and 4) versus freight vehicles attracted in the zone are presented, respectively, in Figure 4.

4.2. Model Specification, Calibration, and Validation. Several models have been specified and calibrated using different combinations of the observed zonal attributes described above. Some specifications were rejected due to their low statistical significance considering informal and formal tests.

The section contains calibrated parameters and elasticity analyses of specified models that have provided higher statistical significance considering informal and formal tests. Table 7 shows the models selected specified with direct variables, based on the sign of the calibrated parameters and on the value of the coefficient of determination R^2 . Table 8 reports two values of the calibrated parameters of two specifications with transformed variables: the natural logarithm of the number of households living in the zone and the natural logarithm of the average distance accessibility.

4.3. Model Application. Some applications of the models calibrated in Table 7, useful for the validation of the results, are presented in the following. In particular, the following analyses are reported:

- The analysis of estimated probabilities versus attributes.
- The analysis of demand elasticity.
- The comparison between the observed frequencies and estimated probabilities.

The analysis is executed on models 6 and 7, which are selected for two reasons. The first concerns the presence of both a socioeconomic and a transport/mobility attribute in their specification; the second is related to their value of the coefficient of determination.

4.3.1. Estimated probabilities versus Attributes. The variation of the *probability of supplying one zone* versus each attribute of model 6 and of model 7 is plotted, respectively, in Figures 5 and 6. Considering that the two models have two attributes, for the attribute not considered, an average value is assumed, as reported in the figures. The probability is plotted considering three representative zones of the study area: Mammola, which is a hilly internal zone; Northern Siderno, which is a portion of a coastal populated town; and Roccella Jonica, where there is a touristic port.

Model 6 is composed of two attributes: the number of commercial establishments and the number of passenger vehicles attracted. By considering the zone of Mammola as an example, the reference value of the number of commercial establishments is 20, and the range of variation is set between 16 and 24.

Model 7 is composed of two attributes: the number of commercial establishments and the average time accessibility. By considering the zone Mammola as an example, the reference value of the number of commercial establishments is again 20 (as in model 6), the range of variation is set between 16 and 24, the reference average time accessibility is 27.97 minutes, and the range of variation is set between 24 and 33 minutes.

The effects of the model are shown by comparing the plots among the three examined zones, where the reference value is different from one zone to another. Consequently, the range of variation of estimated probability is different. Given a model and an attribute, assuming a value of the percentage variation of a specific attribute, the effect in the probability increases with the absolute value of the attribute. For this reason, the elasticity indicator has to be evaluated for integrating this analysis.

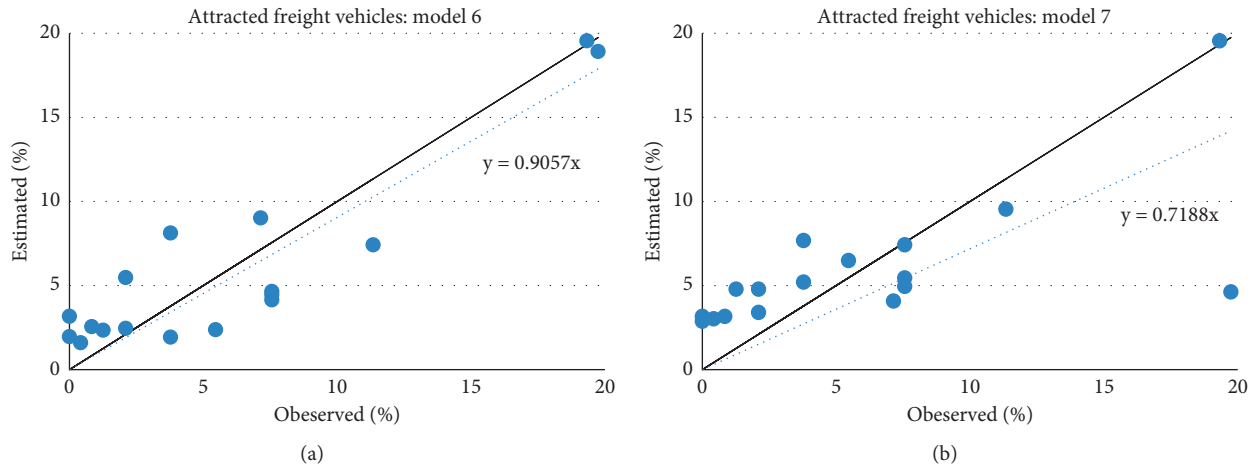


FIGURE 7: Comparison between observed and estimated values of attracted freight vehicles. Scatterplots for model 6 (a) and model 7 (b).

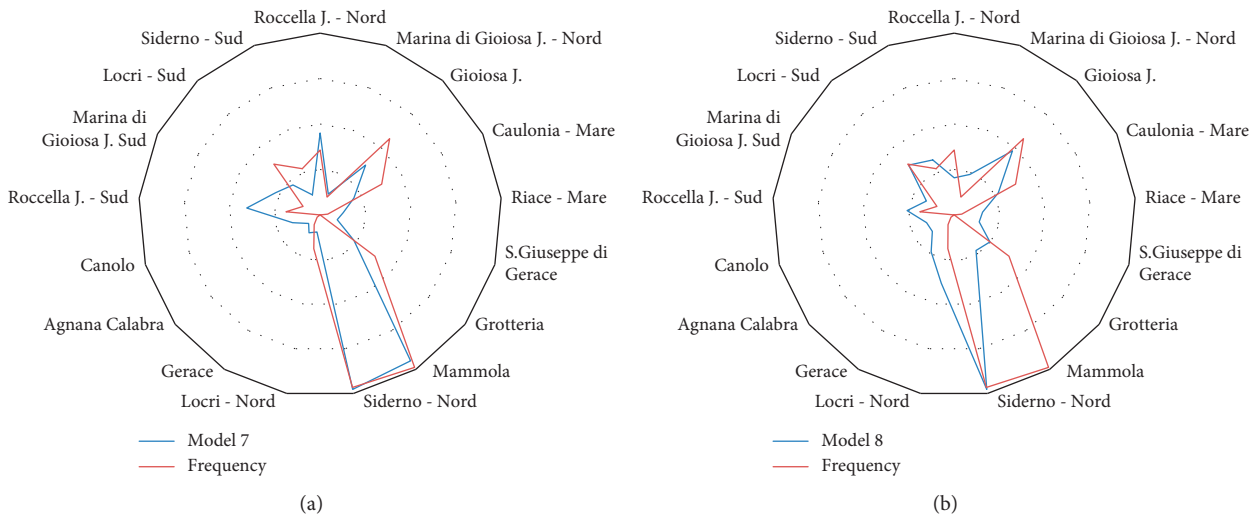


FIGURE 8: Comparison between observed and estimated values of attracted freight vehicles. Radar diagrams for model 6 (a) and model 7 (b).

4.3.2. *Elasticity.* Elasticity provides information on the percentage change in demand with respect to the percentage change of one attribute. The calculation of direct point elasticity related to an attribute x_k is executed according to the following equation:

$$\epsilon_{k,c} = \frac{\partial p_c / \partial x_{k,c}}{x_{k,c} / p_c}, \quad (6)$$

where $x_{k,c}$ is the value of the attribute k for the alternative c ; p_c is the value of probability associated with alternative c .

In the case of the Logit model with linear specification of the utility function, the direct elasticity has the following specification:

$$\epsilon_{k,c} = (1 - p_c) \beta_k x_{k,c}, \quad (7)$$

where $\beta_{k,c}$ is the parameter of the attribute $x_{k,c}$ (including the parameter of the Gumbel distribution).

Models 6 and 7 are considered for the calculation of elasticity by means of equation (6) or (7) for five zones considering an increase of 10% and 5% ($100 \Delta x_k / x_k$) for each attribute. The five zones are selected according to the average, maximum, and minimum values (see Tables 9 and 10) of the attributes of the models. Table 11 presents the values of the estimated direct elasticities. The following elements emerge. Both two models are elastic ($|\epsilon_{k,d}| > 1$) with respect to the attribute “number of commercial establishments” for the zones of Mammola and Northern Siderno. As far as concerns the attribute “average time accessibility,” model 8 is elastic ($|\epsilon_{k,d}| > 1$) for the zone of Northern Siderno. In all the other zones, the models are rigid ($|\epsilon_{k,d}| < 1$).

4.3.3. *Observed versus Estimated Values.* The comparison between observed and estimated values, with models 7 and 8, of attracted freight vehicles are reported, respectively, in Figure 7. The thicker solid lines in the scatterplot diagrams

are the bisectors of the quadrant represented by the axes of the observed and estimated values, while the dashed lines represent the regression line of the estimated values. The regression has no significance in terms of demand estimation. It gives information about over- or underestimation using the calibrated model. The following elements may be highlighted. In both models, the regression line lies below the bisector line. In the case of model 7, the angular coefficient of the regression line is 0.905; this means that the model underestimates the estimated values by 10% on average. In the case of model 8, the angular coefficient of the regression line is 0.718; that is, the model on average underestimates the estimated values by almost 30%.

The same above comparison, via radar diagrams, between observed and estimated values of attracted freight vehicles for each zone of the study area is presented in Figure 8.

5. Conclusions

The paper is part of a wider research topic concerning the influence of freight transport on the economy inside a subregional area. Since freight transport represents a fundamental component in the logistic process of commercial establishments, it requires adequate planning of supply services operated by road vehicles.

In recent years, the increase in the freight volumes transported on the roads has asked for new technologies (emerging ICTs) to monitor the freight movements and for new models to simulate the last-mile freight distribution process. The use of FCD, obtained through GPS trackers present on board the vehicle, offers the advantage of providing some characteristics of the phenomenon examined at lower costs.

The research objective of the paper is to verify whether and how FCD support quantitative analyses about freight vehicle mobility to obtain demand flow estimates in the current state and/or in future scenarios in a territorial area characterized by the lack of transport monitoring technologies.

The paper reports the specification-calibration-validation of a freight supply-acquisition model with FCD. The simulated quantities are expressed in terms of the number of freight vehicles having a zone as a destination, which depend on socioeconomic attributes (e.g., commercial employees), accessibility (e.g., average travel time), and mobility (e.g., number of passenger vehicles attracted in a zone for purchasing purposes).

The calibrated model allows us to define some policy indications about logistics planning towards the suburban Locride area. From the elasticity analysis of the analyzed models, the following elements emerge. Accessibility plays no particularly relevant role considering that only in some cases the elasticity has values lower than 1. The presence of a high number of commercial establishments in a zone generates a high percentage of freight vehicles having the zone as a destination, considering that, in these zones, elasticity presents values higher than 1. This means that the location of commercial activities weighs more than

accessibility; therefore, from a logistic point of view, particular attention needs the location of new commercial establishments or the displacement of existing ones. This is probably also due to the small size of the examined study area.

The goodness of the indicators obtained testifies the quality of the obtained models for the reproduction of the observed phenomenon of the freight vehicle that reaches zones where commercial establishments are present. FCD are able to support some quantitative analyses of the freight vehicles' mobility in a territorial area characterized by the absence of transport monitoring technologies.

The models proposed have some limitations that must be considered and possibly improved in the future. The models are based only on FCD and do not consider flow data observed on the links and relative to the freights category. The calibrated parameters could be improved by means of aggregate calibrations from flows and also considering different freight categories in terms of vehicles and quantities. The model is calibrated in a specific territorial area; the temporal variation of the calibrated parameters and the transferability of the model to other contexts need to be verified.

It is desirable that further studies and simulations should be carried out in the future, starting from the structure of the model implemented, to improve the parameters of interest with larger databases and to allow the study of other segments of freight demand. Finally, the models' transferability to other areas with similar characteristics has to be tested.

Data Availability

The data are available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Giuseppe Battaglia contributed to formal analysis, data curation, and investigation; Giuseppe Musolino contributed to conceptualization, methodology, formal analysis, investigation, and reviewing and editing of the paper; Antonino Vitetta contributed to supervision, conceptualization, methodology, formal analysis, validation, and reviewing and editing of the paper.

Acknowledgments

This research was partially supported by the Dipartimento di ingegneria dell'Informazione, delle Infrastrutture e dell'Energia Sostenibile, Università Mediterranea di Reggio Calabria, and by the project "La Mobilità per i Passeggeri come Servizio-MyPasS", Fondi PON R&I 2014-2020 e FSC "Avviso per la presentazione di Progetti di Ricerca Industriale e Sviluppo Sperimentale nelle 12 aree di Specializzazione individuate dal PNR 2015-2020", codice identificativo ARS01_01100.

References

- [1] F. Russo and A. Comi, "Investigating the effects of city logistics measures on the economy of the city," *Sustainability*, vol. 12, no. 4, pp. 1–11, 2020.
- [2] F. Russo, C. Rindone, and P. Panuccio, "European plans for the smart city: from theories and rules to logistics test case," *European Planning Studies*, vol. 24, no. 9, pp. 1709–1726, 2016.
- [3] A. Comi and A. Polimeni, "Forecasting delivery pattern through floating car data: empirical evidence," *Future Transportation*, vol. 1, no. 3, pp. 707–719, 2021.
- [4] F. Tolentino-Zondervan, E. Bogers, and L. van de Sande, "A managerial and behavioral approach in aligning stakeholder goals in sustainable last mile logistics: a case study in The Netherlands," *Sustainability*, vol. 13, no. 8, p. 4434, 2021.
- [5] X. Wang, W. Kim, J. Holguín-Veras, and J. Schmid, "Adoption of delivery services in light of the COVID pandemic: who and how long?" *Transportation Research Part A: Policy and Practice*, vol. 154, pp. 270–286, 2021.
- [6] National Academies of Sciences, "Engineering, and Medicine," *Using Commodity Flow Survey Microdata and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook*, The National Academies Press, Washington, DC, 2016.
- [7] K. W. Ogden, *Urban Freights Movement*, Ashgate, Hants, England, 1992.
- [8] E. Taniguchi and R. G. Thompson, *City Logistics 3 – towards Sustainability and Liveable Cities*, ISTE Ltd, John Wiley and sons, London, UK, 2018.
- [9] J. Holguín-Veras, J. Amaya Leal, I. Sánchez-Díaz, M. Browne, and J. Wojtowicz, "State of the art and practice of urban freight management," *Transportation Research Part A: Policy and Practice*, vol. 137, no. C, pp. 360–382, 2020.
- [10] F. Russo and A. Comi, "Sustainable urban delivery: the learning process of path costs enhanced by information and communication technologies," *Sustainability*, vol. 13, no. 23, Article ID 13103, 2021.
- [11] F. Russo and A. Comi, "A modelling system to simulate goods movements at an urban scale," *Transportation*, vol. 37, no. 6, pp. 987–1009, 2010.
- [12] J. F. Ehmke, S. Meisel, and D. C. Mattfeld, "Floating car based travel times for city logistics," *Transportation Research Part C: Emerging Technologies*, vol. 21, no. 1, pp. 338–352, 2012.
- [13] A. Nuzzolo and A. Comi, "Urban freight demand forecasting: a mixed quantity/delivery/vehicle-based model," *Transportation Research Part E: Logistics and Transportation Review*, vol. 65, pp. 84–98, 2014.
- [14] B. W. Sharman, *Behavioural Modelling of Urban Freight Transportation: Activity and Inter-arrival Duration Models Estimated Using GPS Data*, Ph.D. thesis, Graduate Department of Civil Engineering, University of Toronto, Toronto, Canada, 2014.
- [15] M. E. Ben-Akiva, T. Toledo, J. Santos et al., "Freight data collection using GPS and web-based surveys: insights from US truck drivers' survey and perspectives for urban freight," *Case Studies on Transport Policy*, vol. 4, pp. 38–44, 2015.
- [16] S. Hadavi, S. Verlinde, W. Verbeke, C. MacHaris, and T. Guns, "Monitoring urban-freight transport based on GPS trajectories of heavy-goods vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3747–3758, 2019.
- [17] A. I. Croce, G. Musolino, C. Rindone, and A. Vitetta, "Transport system models and big data: zoning and graph building with traditional surveys, FCD and GIS," *ISPRS International Journal of Geo-Information*, vol. 8, no. 4, p. 187, 2019.
- [18] M. Diana, M. Pirra, and A. Woodcock, "Freight distribution in urban areas: a method to select the most important loading and unloading areas and a survey tool to investigate related demand patterns," *European Transport Research Review*, vol. 12, 2020.
- [19] A. Nuzzolo, A. Comi, and A. Polimeni, "Urban freight vehicle flows: an analysis of freight delivery patterns through floating car data," *Transportation Research Procedia*, vol. 47, pp. 409–416, 2020.
- [20] G. Musolino, C. Rindone, and A. Vitetta, "A modelling framework to simulate paths and routes choices of freight vehicles in suburban areas," in *Proceedings of the 7th International Conference on Models and Technologies for Intelligent Transportation Systems*, Heraklion, Greece, June 2021.
- [21] A. I. Croce, G. Musolino, C. Rindone, and A. Vitetta, "Route and path choices of freight vehicles: a case study with floating car data," *Sustainability*, vol. 12, no. 20, p. 8557, 2020.
- [22] E. Cascetta, *Transportation Systems Analysis: Models and Applications*, Springer, NY, USA, 2009.
- [23] A. Vitetta, "A quantum utility model for route choice in transport systems," *Travel Behaviour and Society*, vol. 3, pp. 29–37, 2016.
- [24] M. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press, Cambridge, Massachusetts, U.S.A., 1985.
- [25] G. A. Malavenda, G. Musolino, C. Rindone, and A. Vitetta, "Residential location, mobility, and travel time: a pilot study in a small-size Italian metropolitan area," *Journal of Advanced Transportation*, vol. 2020, Article ID 8827466, 11 pages, 2020.
- [26] B. Alonso, Á. I. Pòrtilla, G. Musolino, C. Rindone, and A. Vitetta, "Network Fundamental Diagram (NFD) and traffic signal control: first empirical evidences from the city of Santander," *Transportation Research Procedia*, vol. 27, pp. 27–34, 2017.
- [27] A. I. Croce, G. Musolino, C. Rindone, and A. Vitetta, "Traffic and energy consumption modelling of electric vehicles: parameter updating from floating and probe vehicle data," *Energies*, vol. 15, no. 1, p. 82, 2021.