The Negative Externalities of the Transportation System in Megacities influenced by the Industrial and Commercial Establishments and the Urban Freights flows

Diego B. Tomasiello, Cláudia A. Soares Machado, Mariana A. Giannotti, Renato O. Arbex and José Alberto Quintanilha

Abstract

There are few activities taking place in a megacity (cities with over 10 million inhabitants) that do not require the movement of goods. However, these displacements compete for space and use the same road system of the transportation of people in private vehicles or mass transit. The São Paulo dynamics of urban freight distribution was analyzed based on 1,789,022 spatialized points provided by GPS receivers installed on freight vehicles. These points were spatially related to the road system to identify the main roads with a high concentration of cargo vehicles, and they were analyzed regarding the location of large companies of the food and drink sector, and also other kind of companies in general. Through techniques of kernel density estimator, Moran's index and the generation of outlier clusters, were identified several areas with high correlation between cargo vehicles, large companies that produce and/or distribute food products and other kind of companies.

D.B. Tomasiello (✉) · C.A.S. Machado · M.A. Giannotti · R.O. Arbex · J.A. Quintanilha
Geoprocessing Lab –Transportation Engineering Dept. – Polytechnic School of the University of São Paulo, Av. Prof. Almeida Prado, t.2, nº 83, SP, Brazil
E-mail: diegobt86@usp.br
C.A.S. Machado
E-mail: csmachado6@gmail.com
M.A. Giannotti
E-mail: mariana.giannotti@usp.br
R.O. Arbex
E-mail: renatoarbex@usp.br
J.A. Quintanilha
E-mail: jaquinta@usp.br

1. Introduction

The dynamics of urban freight distribution occurs predominantly in areas characterized by a concentration of industrial, commercial and service establishments, and causes negative consequences for the transportation system, which in most cases are more extreme than those of the passenger transportation system. Such effects are known as externalities of transport, and the impacts most strongly perceived by users are congestion, traffic accidents, pollution, vibration and pavement deterioration, and etc.

This paper aims to investigate, through spatial analysis, how the location of industrial, commercial and service establishments and the urban freight trips affect the transportation system regarding to the generation of externalities. The study area is the megacity of São Paulo, Brazil, which has a population of over 13 million inhabitants and one of the largest vehicle fleets in the world. The methodology correlates spatially the two sets of data (location of industrial, commercial and service establishments and flow of freight vehicles through the road network provided by GPS receivers) using spatial statistical tools (Kernel Density Estimator – KDE and Moran Index) to assess how these two variables are related.

The incorporation of monitoring and tracking vehicle data (GPS data) used for urban freight allows studying the spatiotemporal behavior of these dislocations and defines clusters of interest of the geographical distribution of urban freight and the socioeconomic characteristics of the study area for possible interventions and researches. Data, and particularly geo-spatial data (i.e. data that can be mapped), is fundamental to the success of urban freight transport studies. Such data, in particular in the context of megacities, tends to be 'big data' – i.e. data that includes millions of individual elements and requires careful, specialist processing techniques and management.

2. Theoretical Background

According to Crainic et al. (2004) urban freight is both an extremely important as disturbing activity for urban traffic. For the population ensures that there are adequate amounts of supplies in stores, as well as the delivery of goods at home. For companies established within the limits of cities, provides the vital link between suppliers and consumers. On the other hand, the urban freight causes a series of negative externalities for the urban transportation system. The transport negative externalities affect the performance and safety of urban traffic, and their consequences reduce population welfare, transportation system accessibility, urban mobility, and the value of urban land. To avoid the externalities, actions of planning, road safety and territorial organization are required (Crainic et al., 2004; Behrends et al., 2008; Russo and Comi, 2011).

The presence of the externalities is a strong evidence of market imperfection. It is necessary to improve the dynamics of urban freight distribution, pursue sustainable development and seek solutions to freight transport in urban centers that address the optimization of logistics activities, considering the social, environmental, and economic, impacts, and thus improving the efficiency of the local distribution system (Verhoef, 1994).

Filippi et al. (2010), in developing a methodology for urban areas, emphasize the need to implement measures aimed at mitigating the effects of the negative externalities of urban freight transport, mainly with respect to the reduction of urban accessibility and mobility and the emission of pollutants into the atmosphere. Similarly, Romero et al. (2014) propose a simulated model for the optimization of planning and managing urban freight transport, with the aim of minimizing the total costs of the system (operational costs of suppliers, costs borne by users of privately owned vehicles and those of public transport, etc.), thus reducing the impact on accessibility and mobility by means of reducing traffic jams, as well as the impact on the environment through the control and reduction of freight vehicle emissions.

Due to the spatial nature of the data involved in transport-related studies, the geospatial technologies, through the use of spatial analysis, provide a powerful analytical means to study urban freight systems. Jiang and Okabe (2014) affirm that thanks to the rapid development of geographic information science and its related technologies, ample range of transportation data have been collected in order to better understanding the transportation systems. Additionally, Theofilatos and Yannis (2014) emphasize the increasing availability of real-time traffic data gotten through geospatial technologies and how it has stimulated a proactive planning and safety management in transport networks. Unlike the conventional approaches, the geospatial methods analyze the spatial patterns of urban freight within the network space and are therefore not affected by the configuration of the street network or its distance (Çela et al., 2013). Bíl et al. (2013) claim that it has been facilitated by the application of geospatial technologies into transportation research, which has enabled the precise localization of cargo vehicles and the identification of spatial patterns involving the regions with high occurrence of freight trips.

In this article the adopted methodology of spatial analysis used the kernel density estimator (KDE) and Moran Index to assess how is the relationship between the geographical location of industrial, commercial and service establishments and the urban flow of freight vehicles in São Paulo City. The KDE calculates the probability density function of a distribution from which a sample has been observed, by centring a probability density function around each one of the observed events (Brunsdon, 1995). According to Flahaut et al. (2003) the kernel estimator is a non-parametric algorithm, which uses a density estimation method. Xie and Wu (2014) affirm that KDE is the most widely used nonparametric method in recent decades. The KDE describes the distribution of the location of an event occurrence and ignores its association with values. This distribution is characterized by the density of events that occur around a centroid, thus representing behavioral patterns of points or lines. In this study the events are the location (geocoded) of the business establishments (represented by points), and the KDE calculates the probability density function of each establishment location.

Kernel density analysis is performed by passing a moving window over the data, usually on a regular grid. The density of observations within a set radius is calculated for each event located on the grid, and the contribution of each observation is weighted by its proximity to the center of the moving window. The result of the KDE application is a density map (raster format). The value of each pixel represents the relationship between the concentrations of events per unit area. The KDE can be used to calculate the density of punctual events (i.e., the density of business establishments in a region) or linear events (i.e., the density of a road network in a zone). Silverman (1986) and Wand and Jones (1995) highlight the simplicity, satisfactory properties, and good results obtained from the KDE method.

Anselin and Lozano-Gracia (2008) define Moran's index as a measurement whose result indicates whether data are distributed randomly in space. Moran's I index only returns a numeric value, which can easily lead to interpretation mistakes (Prudente et al., 2014). An alternative for the Moran's index interpretation is to build Moran Scatter Plots, which is as a way to visualize overall indicators of spatial autocorrelation, revealing spatial lag of the aimed variable on vertical and horizontal axes (Montenegro and Betarelli Junior, 2008). The diagram is divided into four quadrants: High-High (HH), Low-Low (LL), High-Low (HL) and Low-High (LH). Moran Scatter Plots are four-quadrant graphics representing relationship between vector Z values and weighted means (Wz).

The HH quadrant represents positive values for Z and Wz, i.e. areas with high attribute values surrounded by high value areas. The LL quadrant represents negative values for Z and Wz, that is, areas with low attribute value next to low ones. The LH quadrant represents areas with negative values for Z and positive for Wz; these low value areas are near high value areas.

The HL quadrant, which represents positive values for deviation vector Z and negative for weighted means, express areas with high values for a given attribute close to low value neighbors (Prudente et al, 2014).

3. Methodology

3.1 Materials

3.1.1 Trucks Database

Analysis of trucks locations in the municipality of São Paulo were accomplished using data from GPS receivers installed on cargo trucks. The partnership with Maplink traffic data provider was essential for the realization of the studies. The database used as an input to the studies corresponds to a week of tracking, 06 to 10 October 2014, and has 1,789,022 points of trucks with latitude and longitude.

3.1.2 Streets Database

The street network database used in the analysis is the OpenStreetMap database. It was used to relate the information from trucks points to the road in the municipality of São Paulo. The database has 110,347 segments with several information, including classification by road type and streets name.

3.1.3 Location of Industries and Commerce

The location of industries and commerce was provided by the Metropolitan Studies Center (http://www.fflch.usp.br/centrodametropole/). The database has 452,816 records containing different sizes and segments of companies. The company's segment information was instrumental in the analysis of those that are more related to the high number of trucks on the road system.

3.1.4 Census Sectors

The census sectors were obtained from the Brazilian Institute of Geography and Statistics (IBGE) through the Metropolitan Studies Center. The census sector is a polygon corresponding to a suitable area for collecting information for census purposes. There is an amount of 18,953 census sectors in the municipality of São Paulo.

3.2 Procedures

The first procedure consisted in spatialize the trucks database and relate the information to the streets segments. For that, several treatments were performed in the tabular truck database, such as identification of duplicated records or any other type of inconsistency that could affect the processing or the results. The spatialization of the trucks database was performed using a GIS. Spatialized the tracked trucks points, they were compared to the streets vectors and evaluated the displacements between them. At the end of the analysis, it was observed that 80% of the spatialized points, or 1,141,441 points were located at a distance of 20 meters of the streets. The 20% of non-coincident points to the streets database were located mainly in large patios of distribution and logistics centers, which prevented matching them to specific streets segments. The spatialized points, with a distance up to 20 meters, were moved through a Snap tool to the corresponding street segment. Fig. 1 shows the result of the Snap process.

Fig 1 Results of the snap process between trucks and streets

After moving the points, it was performed a Join operation of these points with the streets database. After the join operation, the points possess the information of the streets segments identification, allowing the tabular relationship and future analysis and identification of the corresponding segment.

To assess the amount of different trucks that passed through each street segment in the municipality of São Paulo, the trucks were counted only once when they passed through each segment for the day week under review. The differentiation of each truck was possible because of the existence of a unique vehicle identifier in the database.

Identified the amount of trucks per street segment for each day week, it was observed that there is a wide variation in the number of vehicles for each day. A comparison of the five-day week's numbers is described in Table 1.

Table 1 Comparison of the five weekdays and the number of trucks

Weekday	Number of Trucks
Monday (October/6th/2014)	133,044
Tuesday (October /7th/2014)	242,842
Wednesday (October /8th/2014)	253,977
Thursday (October /9th/2014)	255,779
Friday (October /10th/2014)	255,799

Due to this issue, the Wednesday was chosen as the most appropriate day for the rest of the analysis, since it has an average amount of cargo vehicles.

From the amount of trucks per street segment on Wednesday, it was performed a density analysis in order to identify which segments receives, in proportion to its length, the largest amount of cargo vehicles in the municipality. Fig. 2 shows the street classification by the trucks density.

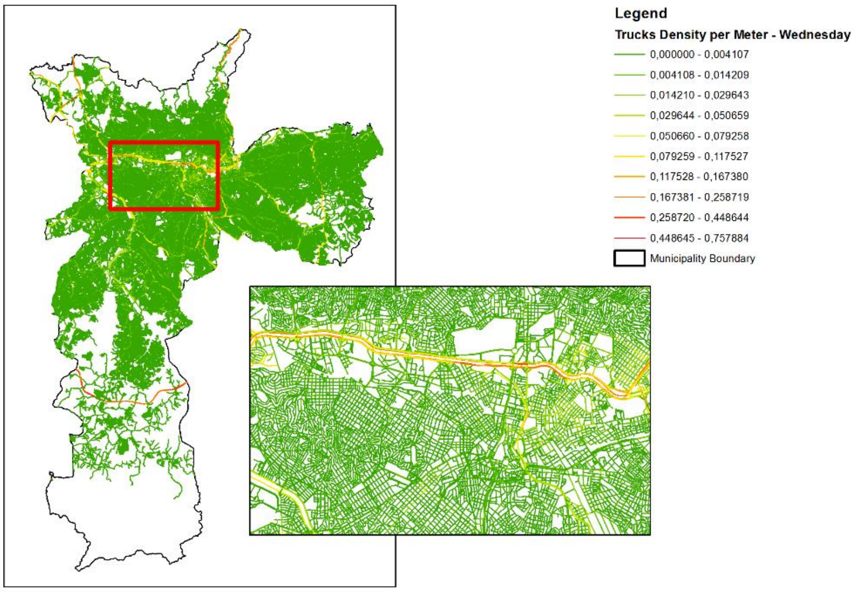


Fig. 2 Street classification by the trucks density.

Since the database containing the trucks data is now treated and spatialized, several analyses were performed regarding the companies' database. The database was stratified by the company size, which is measured by the number of employees, and by its market segment, which in the case of this study are the food and drinks companies. The largest companies, with more than 500 employees, were separated and a Kernel density estimator was performed in order to generate a continuous surface indicating the density distribution of the companies and trucks in the municipality of São Paulo. Fig. 3 indicates the results of the density distribution of the companies and trucks.

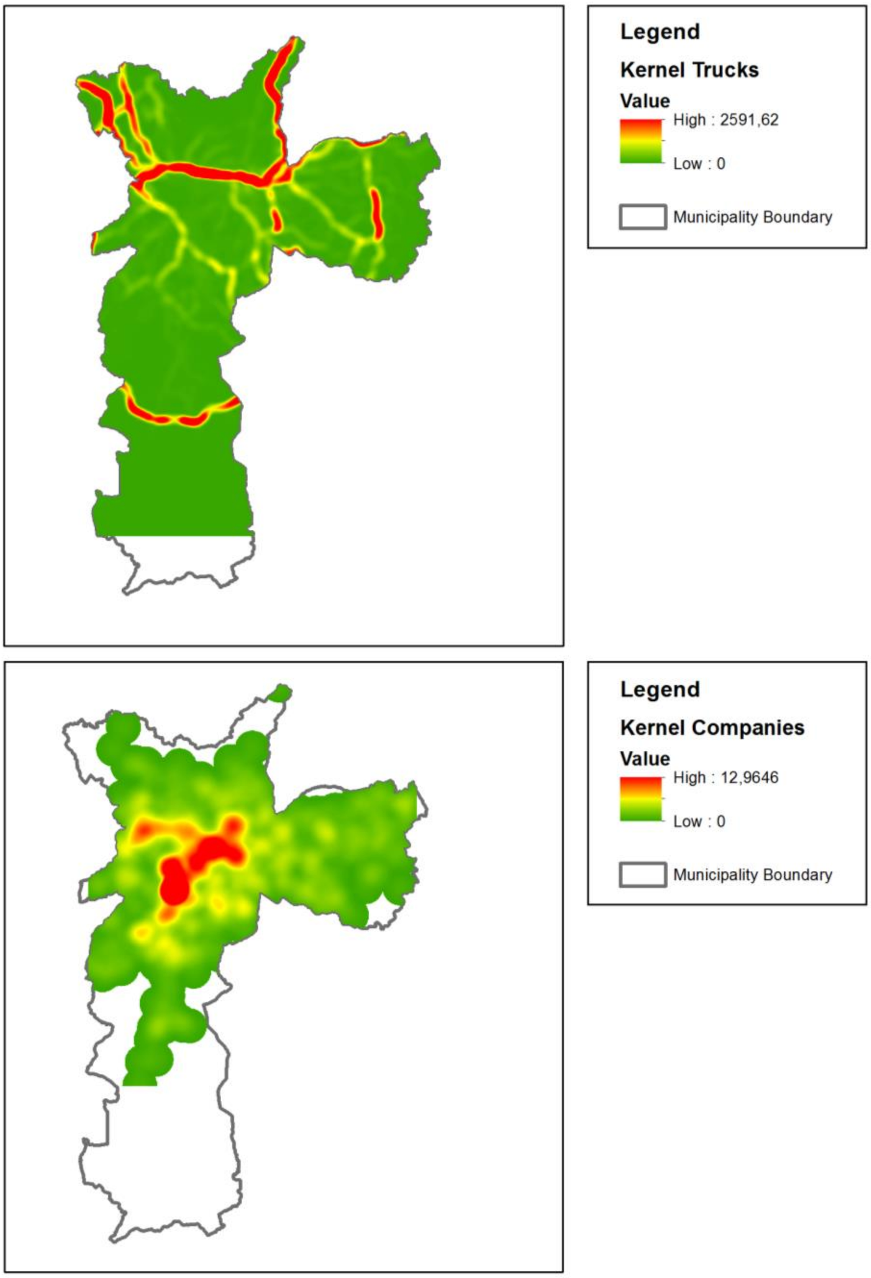


Fig. 3 Kernel density estimator results for trucks and companies.

From the continuous density surfaces of large food and drink companies and trucks, it was generated the Moran index and identified clusters with high concentration of the analyzed data. For the generation of the Moran index and clusters outlier, continuous surfaces of trucks density and large companies were transformed into point geometry and related to census sectors of the municipality of São Paulo. For each census sector polygon, the average density of companies and trucks were associated to the geometry extracted. From the associated values were performed the Moran index and identified the clusters of trucks and companies density. Fig. 4 shows the results.

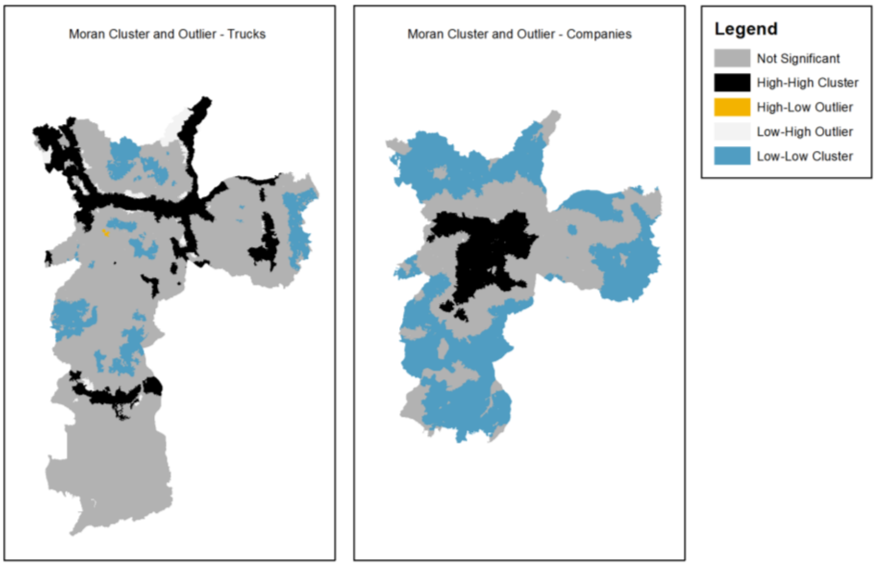


Fig. 4 Moran's index clusters for trucks and companies.

Due to time restriction for movement of cargo vehicles in the expanded center of São Paulo, it was performed the stratification of the Wednesday data with respect to time. The database was splitted into two zones, one with movement of cargo vehicles from 5:00 am to 9:00 pm and another from 9:00 pm to 5:00 am. From the data partition, there were accomplished new analyzes using the Kernel density estimator and generated the Moran index for the identification of clusters. Fig. 5 shows the results for the analysis performed.

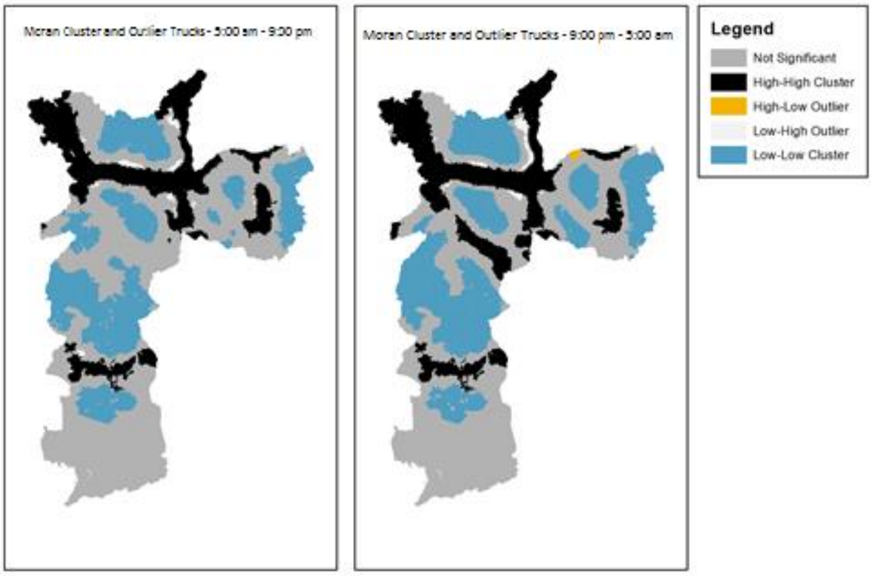


Fig. 5 Moran's index clusters for time restriction (5:00 am – 9:00 pm) and no time restriction (9:00 pm – 5:00 am)

Further investigation was the identification of cargo trucks with speed up to 4 km/h and relationship of these vehicles with the businesses near their locations. From this analysis, it was possible to infer that low speed or stopped vehicles were near of what kind of businesses.

4. Results

Through the trucks, large food and drink companies density data, were performed a Boolean AND analysis and identified regions where there are overlaps between a density higher than 10 trucks per square kilometer and more than 1 large company per square kilometer. Fig. 6 shows the overlapping areas.

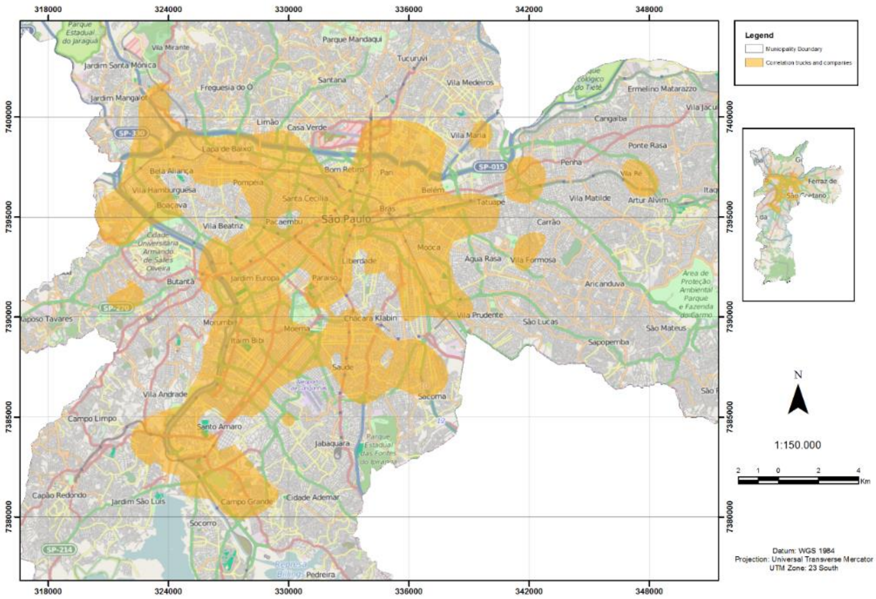


Fig. 6 Results of the overlapping analysis between trucks density and companies' density.

The Moran's index identified clusters of census sectors that have a high-density index of trucks and large companies of food and drinks. After the identification of these clusters of high density of trucks and companies, a Join operation was performed and both clusters with high concentration (trucks and companies) were separated. Fig. 7 shows the results of the operation.

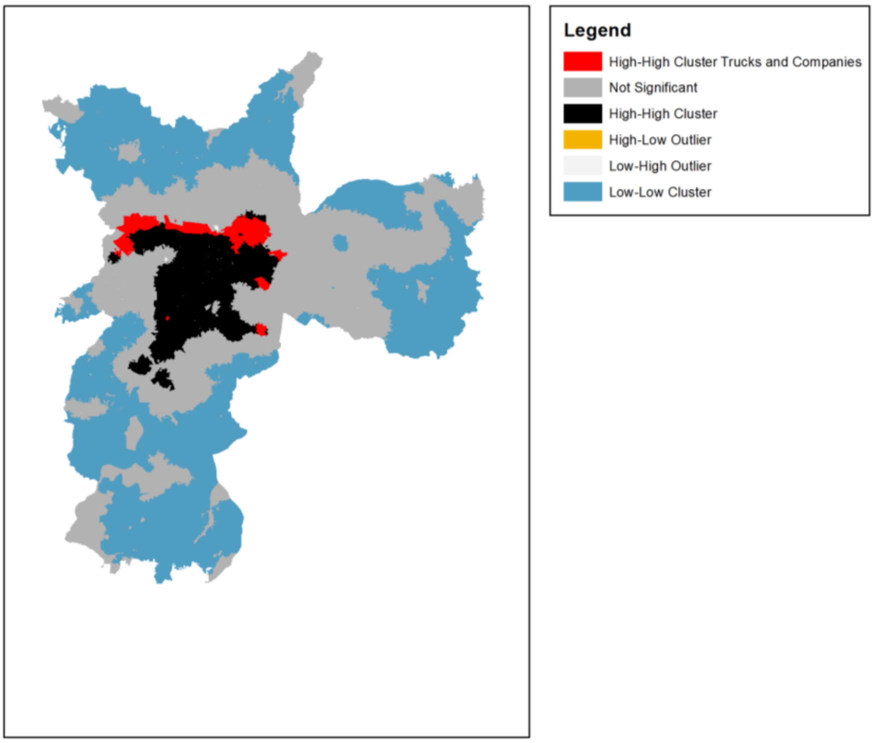


Fig. 7 Clusters with high concentration of trucks and large food and drinks companies.

Due to the differentiation of the time, according to the restrictions on movement of cargo vehicles in the expanded center of São Paulo, different clusters, according to the analysis period, were identified. The first analysis was regarding the no cargo vehicles restriction period, from 9:00 pm to 5:00 am. During this specific period, the correlation between the large food and drinks companies and the cargo vehicles movement is higher than the whole Wednesday. This result is clear analyzing the correlation between the high-high clusters of trucks and companies, represented in the Fig. 8.

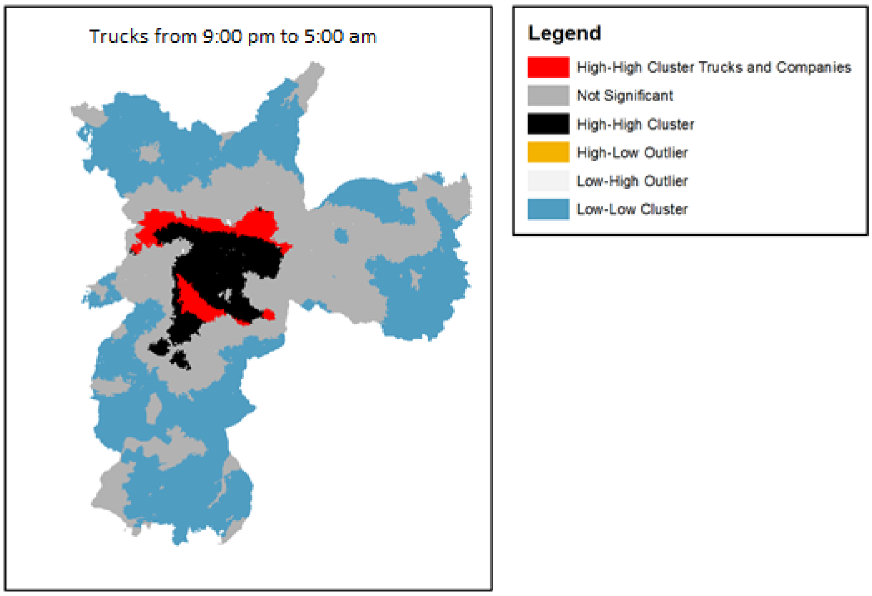


Fig. 8 Correlations of the large food and drinks companies and the cargo vehicles movement from 9:00 pm to 5:00 am.

The analysis of the correlation between the large food and drinks companies and trucks for the restriction period was similar to the analysis of the whole Wednesday. Due to the high amount of trucks movement in the restriction time, the result for the whole day was influenced by the restriction period. Fig. 9 shows the results of the correlation between the large food and drinks companies and the trucks.

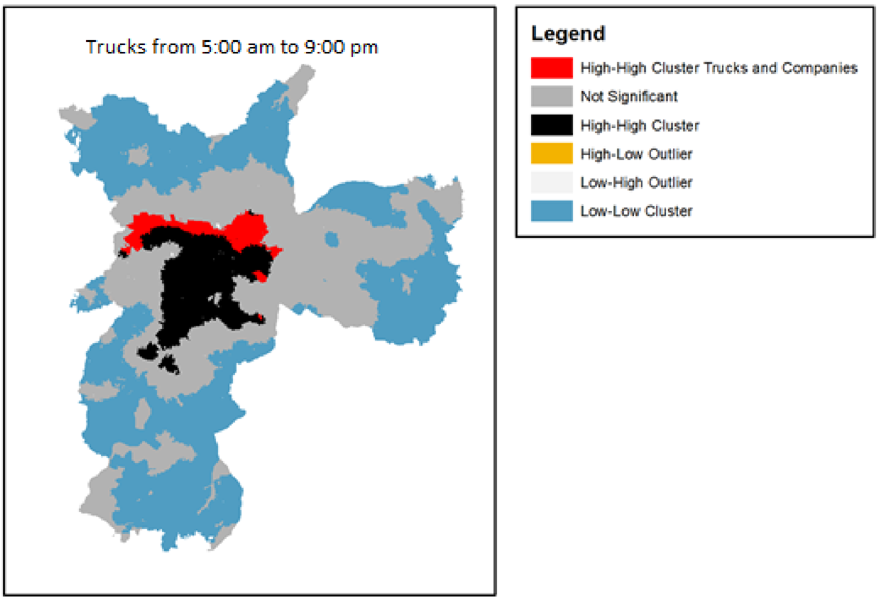
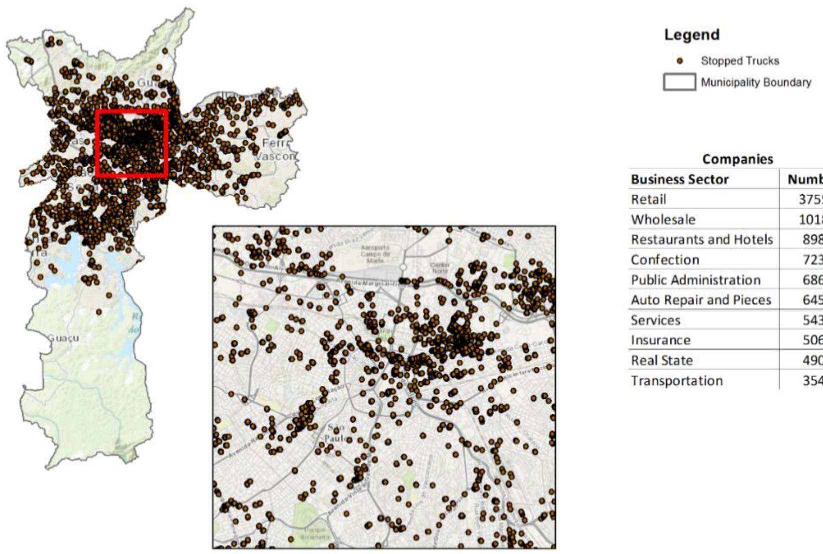


Fig. 9 Correlations of the large food and drinks companies and the cargo vehicles movement from 5:00 am to 9:00 pm.

The spatial distribution of truck points with travel speed up to 4 km/h, linked to the companies' points, allowed the identification of some companies related to specific business segments that are more associated to the amount of cargo transportation. It was performed a buffer analysis to all trucks points on Wednesday with travel speed up to 4 km/h and select the companies in a distance of 20 meters around the points. The selected companies were analyzed regarding the market segment and it was created a table containing the ten most related market segments to the low speed trucks. Fig. 10 shows the distribution of the low speed trucks in the municipality of São Paulo and the table containing the ten most related market segments to the low speed trucks.



Companies

Business Sector	Number
Retail	3755
Wholesale	1018
Restaurants and Hotels	898
Confection	723
Public Administration	686
Auto Repair and Pieces	645
Services	543
Insurance	506
Real State	490
Transportation	354

Fig. 10 Distribution of the low speed trucks in the municipality of São Paulo and the table containing the ten most related market segments to the low speed trucks.

5. Conclusion

Through the analysis, we concluded that there is a spatial relationship between the flow of trucks and the location of large companies of food and drinks in the municipality of São Paulo. Most companies are located in the expanded center of the city, which requires special authorization for cargo vehicles or contracts with Urban Freight Vehicles (VUCs), which are vehicles that have a special authorization to flow in the restriction zone. Companies established in the restriction zone boundaries, have higher correlation to the analyzed truck fleet by being close to major traffic routes and have no time restriction.

Due to time restrictions for the movement of cargo vehicles in the expanded center of São Paulo, analyses were performed at the same time of the time restriction, and when there is no time restriction. The outlier clusters of each analyzed period shows that for the period of no time restriction, the high cluster concentration of large food and drink companies are more related to the truck flow.

In addition to food and drinks companies, has been identified that much of the cargo vehicles at low speed or stopped, when corresponding to the road, were close to retail businesses, wholesale, restaurants and hotels, clothing, public administration companies, centers automotive repair, services, insurance companies, real estate or cargo transportation. Many companies identified in this analysis are branches that traditionally use cargo vehicles for the supply and distribution of products and goods.

Through the performed analysis may be observed that there is a high concentration of large food and drinks companies in the central regions of São Paulo. This concentration, which is associated with the high flow of cargo vehicles, ends up favoring the appearance of the negative effects of externalities, which are mainly to decreased mobility and urban accessibility and increasing pollution in the central portion of the city of São Paulo. The methodology applied in this study can be used by public authorities to make public policy actions that seek to balance the land use occupation of places that suffer from the effects of urban cargo transportation externalities.

In summary, our work presented an important contribution for planners and decision-makers regarding the understanding of the impacts of large food and drink companies and other kind of establishments to the freight vehicle flow.

Acknowledgments The authors would like to thank the company Maplink for the dataset of freight vehicle flow provided, and the Research Institute Centro de Estudos da Metrópole (Metropolitan Studies Center) for the data of the business establishments in São Paulo City. The authors also thank the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - National Council for Scientific and Technological Development), for granting scholarships to the researchers, and the Laboratory of Geoprocessing – Department of Transportation Engineering – Polytechnic School of the University of São Paulo for providing the infrastructure for this study.

References

Anselin, L.; Lozano-Gracia, N. Errors in variables and spatial effects in hedonic house price models of ambient air quality. Empirical Economics, Heidelberg, vol. 34, nº 1, 2008, p. 5-34.

Behrends, S.; Lindholm, M.; Woxenius, J. The impact of urban freight transport: A definition of sustainability from an actor’s perspective. Transportation Planning and Technology, vol. 31, nº 6, 2008, p. 693-713.

Bíl, M.; Andrásik, R.; Janoska, Z. Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. Accident Analysis and Prevention, vol. 55, 2013, p. 265–273.

Brunsdon, C. Estimating probability surfaces for geographical point data: An adaptive kernel algorithm. Computers & Geosciences, vol. 21, nº 7, 1995, p. 877-894.

Çela, L.; Shiode, S.; Lipovac, K. Integrating GIS and spatial analytical techniques in an analysis of road traffic accidents in Serbia. International Journal for Traffic and Transport Engineering, vol. 3, nº 1, 2013, p. 1 – 15.

Crainic, T. G.; Ricciardi, N.; Storchi, G. Advanced freight transportation systems for congested urban areas. Transportation Research Part C: Emerging Technologies, vol. 12, nº 2, 2004, p. 119-137.

Filippi, F.; Nuzzolo, A.; Comi, A.; Delle Site, P. Ex-ante assessment of urban freight transport policies. Procedia Social and Behavioral Sciences, vol. 2, 2010, p. 6332-6342.

Flahaut, B.; Mouchart, M.; San Martin, E.; Thomas, I. The local spatial autocorrelation and the kernel method for identifying black zones - A comparative approach. Accident Analysis and Prevention, vol. 35, 2003, p. 991–1004.

Jiang, B.; Okabe, A. Different Ways of Thinking about Street Networks and Spatial Analysis. Geographical Analysis, 2014, p. 1–4.

Montenegro, R.L.G.; Betarelli Junior, A.A. Analysis and investigation of the determinants of innovation in the municipalities of São Paulo. Regional and Urban Studies, vol. 2, nº 2, 2008, p. 7-31.

Prudente, V. H. R.; Souza, C. H. W.; Mercante, E.; Johann, J. A.; Uribe-Opazo, M. A. Spatial statistics applied to soybean production data from Paraná State for 2003-04 to 2009-10 crop-years. Engenharia Agrícola, vol. 34, nº 4, 2014, p. 755-769.

Romero, J. P.; Benavente, J.; Moura, J. L.; Ibeas, A.; Alonso, B. Na environmental approach to optimize urban freight transport systems. Computer-based Modelling and Optimization in Transportation. Advances in Intelligent Systems and Computing, vol. 262, SOUZA, J. F.; ROSSI, R. (eds), Switzerland: Springer International Publishing, 2014, p. 127-136.

Russo, F.; Comi, A. Measures for sustainable freight transport at urban scale: Expected goals and tested results in Europe. Journal of Urban Planning and Development, vol. 137, 2011, p. 142-152.

Silverman, B. W. Density estimation for statistics and data analysis - Monographs on Statistics and Applied Probability. London: Chapman and Hall, 1st edition, 1986, 200p.

Theofilatos, A.; Yannis, G. A review of the effect of traffic and weather characteristics on road safety. Accident Analysis and Prevention, vol. 72, 2014, p. 244–256.

Verhoef, E. External effects and social costs of road transport. Transportation Research A: Policy and Practice, vol. 28A, nº 4, 1994, p. 273-287.

Wand, M. P; Jones, M. C. Kernel smoothing - Monographs on Statistics and Applied Probability. London: Chapman and Hall, 1st edition, 1995, 212p.

Xie, X.; Wu, J. Some Improvement on Convergence Rates of Kernel Density Estimator. Applied Mathematics, vol. 5, 2014, p. 1684-1696.