Journal of Transport Geography 104 (2022) 103450

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Changes in warehouse spatial patterns and rental prices: Are they related? Exploring the case of US metropolitan areas

Renata Lúcia Magalhães de Oliveira [a,*], Laetitia Dablanc [b], Matthieu Schorung [b]

[a] *Department of Applied Social Sciences, Federal Center for Technological Education of Minas Gerais, Av. Amazonas 7675- Nova Gameleira, Belo Horizonte, MG, 30510-000, Brazil*

[b] *Université Gustave Eiffel, Laboratoire Ville Mobilité Transport, Unité commune ENPC et Univ Gustave Eiffel, 6-8 Avenue Blaise Pascal, 77420, Champs-sur-Marne, France*

A R T I C L E I N F O

Keywords:
Warehousing spatial patterns
US metropolitan areas
Logistics sprawl
Logistics real estate
Urban form

A B S T R A C T

In this paper, two hypotheses are explored linking urban characteristics to the spatial structure of warehouses: (i) the location of warehouses (central places or peripheral areas) is closely related to the land/rent values of these facilities; and (ii) logistics sprawl is higher in cities with a high differential between land/rent values in city centers and peripheral areas. For that, we have considered warehouse real estate and urban data for 46 United States metropolitan areas to analyze the urban spatial structure and the relationship among urban variables, warehouse location, and rental prices. We deliver an exploratory analysis among the 46 metropolitan areas. The main results are (i) it is essential to classify metropolitan portions of urban space into a typology to perform studies that consider different regions; (ii) warehouse location and rent prices are related to the concentration of urban activity; (iii) logistics sprawl is not significantly related to differential warehouse rental prices in the database that we explored.

1. Introduction

The rise of global supply chains, e-commerce consumption, and the outsourcing of logistics activities are among the factors that have driven the emergence of a dynamic metropolitan logistics real estate market around the world. Within metropolitan areas, urban land and floor space scarcity, economies of scale, and the need for extensive land parcels have relocated logistics facilities toward less dense and more peripheral areas of cities (Cidell, 2010a; Dablanc et al., 2014; Dablanc and Browne, 2019; Dablanc and Rakotonarivo, 2010; Giuliano and Kang, 2018; Heitz et al., 2017; Sakai et al., 2020). This process, known as 'logistics sprawl,' has compromised urban sustainability by adding many trucks and van freight trips within urban regions (Dablanc et al., 2014; Dablanc and Browne, 2019; Dablanc and Rakotonarivo, 2010; Heitz et al., 2017). In the past ten to twenty years, many case studies of logistics sprawl (LS) and warehouse locations in large urban areas have emerged (see references below). Investigations concerning urban form and function coordination, urban logistics, and real estate strategy have interested public and private actors. Eleven hypotheses were identified linking logistics sprawl and urban forms and characteristics (Dablanc et al., 2020), and some of them were missing sufficient data to be tested. This paper brings a methodological proposal using granular collaborative data and traditional geographic information to investigate two of these hypotheses. Through this research, we explore two hypotheses linking urban characteristics to the spatial structure of warehouses: (i) the location of warehouses is closely related to the land/rent values of these facilities; (ii) logistics sprawl is higher in cities with a high differential between land/rent values in city centers and peripheral areas.

Our starting point was measures of the changes in spatial patterns of warehouses (logistics sprawl measures, which concerns the decentralization of warehouses in the metropolitan areas) in different United States (US) metropolitan areas. These data were obtained from published previous studies and organized in a tidy dataset. We then gather and transform data concerning warehouse rental prices practiced by the real estate market and their urban structures. The goal is to understand the relationship between the evolution in the number and location of warehouses over time and the differential warehouse prices in activity hubs and peripheral activity zones. The primary sources of information about the spatial patterns of warehouses in the US were Kang (2020a) and (2020b). They were selected with additional publications and further work processing and structuring data (Dablanc et al., 2020), as detailed in Section 3 of the methodology.

* Corresponding author.
E-mail addresses: renataoliveira@cefetmg.br (R.L.M. Oliveira), laetitia.dablanc@univ-eiffel.fr (L. Dablanc), matthieu.schorung@univ-eiffel.fr (M. Schorung).

https://doi.org/10.1016/j.jtrangeo.2022.103450
Received 14 January 2022; Received in revised form 6 September 2022; Accepted 13 September 2022
Available online 25 September 2022

This paper's contribution is first methodological: we present an exploratory framework that can be applied to different economic and geographic contexts to understand urban attributes and their internal relationship. To represent urban centralities, we have transformed each variable into homogeneous spatial units (hexagon bins) and determined the proportion of higher activity hexagons. The disaggregated normalized data were then considered for exploring metropolitan regions concerning the spatial pattern of warehouses, warehouse real estate rental prices, and urban centralities. To our knowledge, no published research has explored the relationship between these variables.

Coordinating the aforementioned dimensions is essential to support urban logistics stakeholders' needs, cities' livability, and the real estate market. This research brings a framework for exploring different metropolitan areas within different geographic contexts based on spatial analyses and disaggregated data. This work's results are reproducible and can induce local and regional public authorities to develop a more effective public policy addressing logistics land use and transportation planning.

This paper is organized into five sections: (1) the introduction; (2) a literature review on the investigation of warehouse location and logistics real estate; (3) the presentation of the method; (4) results regarding the developments of this work; and (5) discussion and final considerations.

2. State of the art in warehouse spatial patterns and logistics real estate

In recent decades, several transformations in logistics facility systems have been observed (Sakai et al., 2020). Recent studies have analyzed the location of warehouses in metropolitan areas and the overtime evolution of this location. These studies have demonstrated a shift in the location of warehouses and logistics facilities to peri-urban areas (Allen and Browne, 2012; Bowen, 2008; Cidell, 2010a; Heitz et al., 2017; Heitz and Dablanc, 2015; Kang, 2020a). These changes have been identified as a "logistics sprawl" phenomenon that can be defined as "the tendency for warehouses to move from urban to suburban and exurban areas" (Dablanc and Ross, 2012, p. 434), resulting in a de-concentration of these facilities. In this work, we use "logistics sprawl" to refer to the de-concentration of warehouses in metropolitan areas over time (Dablanc and Rakotonarivo, 2010). Research has identified this phenomenon in different geographical contexts (Aljohani and Thompson, 2016; Cidell, 2010b; Dablanc et al., 2014; Dablanc and Rakotonarivo, 2010; Dablanc and Ross, 2012; Heitz and Dablanc, 2015).

Logistics warehouse location dynamics are based on several criteria and result in a complex supply chain cost structure. These criteria include transportation, accessibility, distribution activities, the regional economy structure, warehouse equipment, land use and real estate, and the organization of logistics flows in the last mile, among other factors (Dablanc and Rakotonarivo, 2010). In the case of North America, several works have analyzed case studies for Atlanta, Los Angeles, Seattle, and Toronto (Dablanc et al., 2014; Dablanc and Ross, 2012; Woudsma et al., 2016). More recently, a comparative analysis of Chicago and Phoenix regarding the location of logistics facilities was published (Dubie et al., 2020). Kang (2020a) has also identified warehouse spatial patterns for 64 US metropolitan areas.

The studies related to the location of logistics facilities seek to justify the local aspects that impacted the development or emergence of these facilities in a given territory. It is perceived in the presented investigations that accessibility to infrastructure is a factor that influences the location of logistics facilities (Sakai et al., 2020). Considering the effects of e-commerce on the logistics market, Janjevic and Winkenbach (2020) analyze urban last-mile distribution strategies in developed and emerging e-commerce markets (Janjevic and Winkenbach, 2020). *E*-commerce retailers and other firms participating in the supply chain are developing various strategies for last-mile e-commerce distribution in urban areas. These strategies must consider the local context that is particularly challenging in emergent markets and rarely discussed in the current literature.

Some studies have explored logistics facilities individually, considering the factors and characteristics influencing the location choice (Gingerich and Maoh, 2019; Giuliano and Kang, 2018; Woudsma et al., 2008). More recently, some papers have explored the determinants of the location of logistics facilities, such as the opportunity to access larger and cheaper vacant parcels in peripheral areas and proximity to highway networks and airports (Allen and Browne, 2010; Dablanc and Ross, 2012). This perspective is more in line with the observation of the local context stated before as a literature gap.

Among the studies that have addressed location determinants for warehouses, some have investigated: the growth of the logistics industry due to globalization and new production and distribution dynamics (Andreoli et al., 2010; Kang, 2020b; Sakai et al., 2020). The correlation of the dynamics of logistics establishments' location with economic dynamics at the national and regional levels for US metropolitan areas was explored by Bowen (2008), with data aggregated for each county within the metropolitan areas. Sakai et al. (2016) explored public regulatory tools regarding development permits and land use for logistics facilities (Sakai et al., 2016). Other authors have investigated transportation costs, even though they have become minor determinants over the past 30 years. The results indicate that the spatial distribution of logistics warehouses depends only marginally on transportation costs (Dablanc and Ross, 2012; Glaeser and Kohlhase, 2004), offering them "increased locational flexibility" (Rodrigue, 2004).

Hesse and Rodrigue (2004) explored the transformation of the logistics real estate sector, increasingly dominated by global firms whose activities are organized around multi-scalar distribution networks. Social and wage conditions can also play a role in the location of warehouses, such as the availability of a large and cheap labor force and the differential in terms of labor costs, as in the case of the Inland Empire in Southern California (De Lara, 2013). Sakai et al. (2020) have considered the influences of logistics facilities' location choices on the urban environment and vice-versa. They analyze local characteristics and specific logistics activity sectors to verify development attraction for logistics facilities in the Paris Region. In another study, the authors evaluate the site choice factors that influence warehouse locations in Istanbul, Turkey (Durmuş and Turk, 2014). They show that accessibility, industrial and commercial clusters, rent, distance to the city center, and customs significantly affect warehouse locations. Giuliano and Kang (2018) use data from warehouses in Los Angeles, USA, to evaluate the effects of demographic and accessibility variables on their locations. The results show that accessibility to interregional markets is essential for logistics facilities development. Gingerich and Maoh (2019) analyze warehouses in Toronto, Canada, and show that industrial land use, infrastructure accessibility (distance), level of urbanization, and land price have significantly influenced warehouse locations (Gingerich and Maoh, 2019).

Considering the intensification of globalization and outsourcing in the logistics real estate market and the changes occurring in the industrial and retail sectors, companies have opted for rental properties instead of owned ones to keep up with market changes (Wagner, 2010). Nevertheless, according to Lim and Park (2020), research on warehouse rental activity is still incipient despite extensive work regarding warehouses' location, except for industrial property value or rental prices (Lim and Park, 2020). In addition, these studies (Clark and Pennington-Cross, 2016; Lim et al., 2018; Lim and Park, 2020) focus on exploring one specific metropolitan area and not on an empirical and exploratory analysis among metropolitan regions, considering the relationship of warehouse rental prices, logistics sprawl and the concentration of urban activities.

Lim et al. (2018) considered the Bayesian spatial profile regression method to identify two warehouse rental submarkets in the Seoul Metropolitan Area (SMA) in South Korea: high-rent and low-rent groups. The high-rent group was strongly associated with proximity to the urban center in Seoul and Incheon Port, higher floor area ratio, relatively older building age, higher land price, transportation, and automated warehousing services. Lim and Park (2020) investigated the spatial dimension of warehouse rent determinants, identifying the regional specificities of the supply and demand of warehouse facilities and services. The spatial stationary and non-stationary relationship between warehouse rent and the transactional characteristics of the rental contracts, physical characteristics of the buildings, location factors, and various warehousing services were explored through spatial autoregressive regression (SAR) and mixed geographically weighted regression (MGWR) models, for the SMA.

Other studies addressed the rental price of logistics facilities and the variables that impact price formation, focusing on the physical characteristics of warehouses, and other attributes, such as the movement and changes in the logistics market (Baglio et al., 2019; Buttimer et al., 1997). For instance, Ma et al. (2018) used machine learning techniques to estimate warehouse rental pricing in China's Beijing area (Ma et al., 2018).

Finally, Heitz et al. (2017) faced different challenges regarding areal units while comparing Metropolitan areas in France and the Netherlands. The differences in size and population of the spatial units between the metropolitan areas did not allow these authors to compare population densities in absolute terms. They used quartile-based discretization for the studied variables generating a typology of spatial units to understand the pattern of relationships between logistical sprawl and population density. However, the authors state that "the diversity in the forms taken by logistics development is not simply due to the logistics and freight transport system, but also depends on the intrinsic characteristics of the regional spatial structure as well as local planning and land-use policies." (Heitz et al., 2017, p. 104).

To the best of our knowledge, only Kang (2020b) explored the spatial structure of warehouses in North American metropolitan regions, investigating the relationship between land prices, warehouse location, city size, and warehouse size for 48 US metropolitan areas (Kang, 2020b). However, in this last work, employment density was the proxy for land prices; and the level of internal disaggregation of spatial data was the ZIP code, whose areal unit is not uniform within each metropolitan region and among regions. The author also considered monocentric urban structures with a distinguished CBD. Further investigation is presented in this work through the characterization of multiple centralities (considering OSM POI and road length data as a proxy of urban activity) and exploring not only the warehouse rental prices but also the prices differential between central and suburban areas and its correlation with logistical sprawl.

Despite recent research developments discussed above, the relationships among the concentration of urban activity, logistics sprawl, and warehouse rental prices have not been explored yet. In this paper, a methodological approach, presented in the next section, is proposed and implemented to address this research gap based on unstructured and open-access data and spatial analysis.

3. Methodological approach

This work proposes a framework that can be implemented for different metropolitan areas in different contexts. It allows identifying the relative concentration of urban activity from open and collaborative spatial data and collecting current warehouse rental prices. The relationships among warehouse rental prices, relative urban activity concentration within the metropolitan areas, and logistics sprawl measures are explored considering homogeneous areal units to allow exploring different metropolitan areas and, therefore, address this gap in the literature.

In this paper, we explore two hypotheses linking urban characteristics to the spatial structure of warehouses: (i) the location of warehouses is closely related to the rental prices of these facilities; (ii) logistics sprawl is higher in cities with a high differential between land/rent values in city centers and peripheral areas. The methodological steps for this investigation are based on two primary analyses: (i) the analysis of the urban spatial structure in each metropolitan area and the relationship between urban variables, warehouse location, and real estate rent prices; and (ii) exploratory analysis of some US metropolitan areas, considering logistics sprawl measures in published studies (Dablanc et al., 2020). Fig. 1 presents the methodological steps performed in this paper. The following subsections present a detailed description of each step.

3.1. Data collection and processing

3.1.1. Logistics sprawl measures

To develop this research, we have considered data on 46 US metropolitan areas where logistics sprawl had been investigated and published in scientific journals. In this dataset (Dablanc et al., 2020), there was information regarding centrographic measures of logistics sprawl, timeframe and sources of data collection, the population in the respective timeframes, metropolitan administrative information, information on the spatial structure of the metropolitan areas, areas' importance as a gateway at a regional scale, and aggregated information on warehouse rental prices. For this paper, we have selected the 46 North American cases whose spatial structure changes regarding the location of warehouses were previously explored (Kang, 2020b). This data is here explored on a metropolitan scale and related to Centrographic measures (standard distance of warehouses to their barycenter) (Dablanc and Rakotonarivo, 2010).

3.1.2. Urban activity

We have collected open-access information (OpenStreetMap contributors, 2020) on urban activity and infrastructure and proposed an urban activity index (UAI). The UAI is a proxy for the concentration of urban activity derived from the location of points of interest and street network density, further detailed.

Data regarding geographic information, road infrastructure, and the location of points of interest were obtained from OpenStreetMap (OpenStreetMap contributors, 2020). OpenStreetMap (OSM) is an open database containing Voluntary Geographic Information (VGI). It is under the Open Database License and contains worldwide data, primarily collected and maintained by volunteers. Despite the differences in data quality within regions of the world (Vargas et al., 2021), the potential of OSM data to explore the concentration of urban activity is significant (Klinkhardt et al., 2021; OpenStreetMap contributors, 2020; Zhang and Pfoser, 2019).

OSM Points of interest (POI) represent different categories of urban activities, such as touristic points, archaeological sites, ATMs, banks, bakeries, bars, restaurants, grocery and other food retailers, beauty shops, book shops, general retailing, medical care centers, and other health e facilities, dentists, parks, fire stations, florists, urban build environments elements, like benches and fountains, graveyards, hotels, libraries, museums, theaters, playgrounds, post offices, schools, toilets, wastebaskets, and other urban facilities. We used this information as a proxy for urban activity intensity (Bakillah et al., 2014; OpenStreetMap contributors, 2020).

Zhang and Pfoser (2019) assessed the suitability of the VGI (OpenStreetMap POI data) to infer urban change. These authors state that little attention has been addressed to POI data, and most work has focused on road networks. One main conclusion in this work is that, from the method proposed, they consider that OSM POI data can be used to drive urban science research (Zhang and Pfoser, 2019).

Compared with traditional land-use data, POI data have the following advantages: (1) POI data makes it possible to address scale issues. We can transform point data into different scales, and (2) POIs may also represent people's preferences and interactions with the built environment (Wu et al., 2018).

Crowdsourced data, including POIs and collaborative data systems,

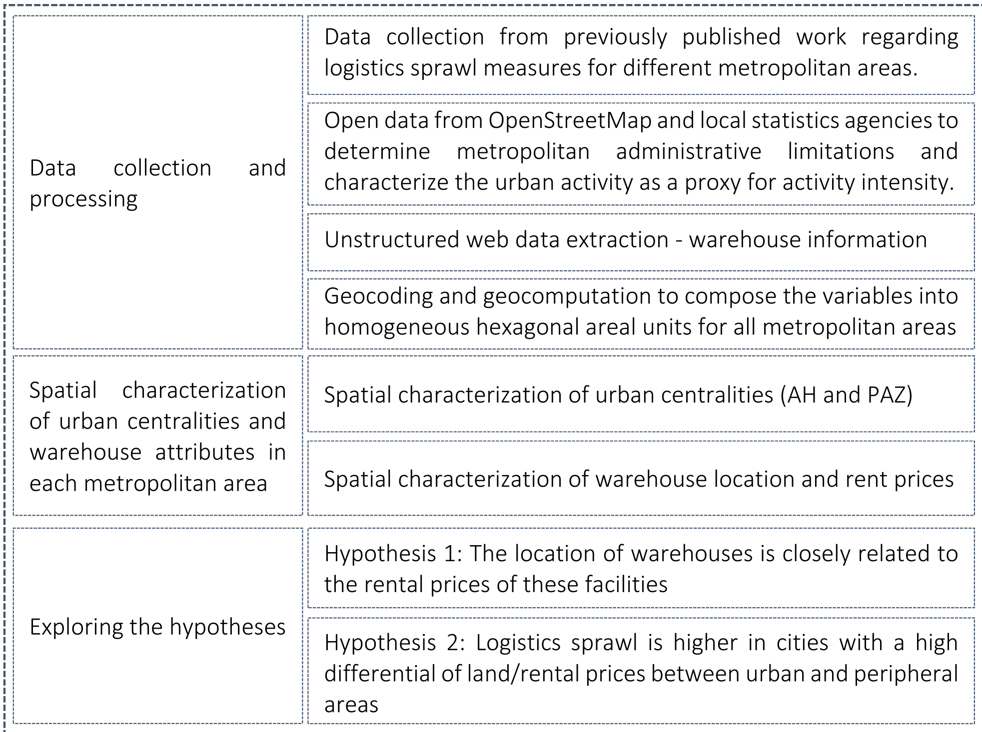

Fig. 1. Methodological steps.

are widely used in urban studies. These data sources have become disaggregated proxies for urban activity (Dong et al., 2020). Michael Batty (Batty, 2007; Whitehand et al., 1996) is a reference in studies considering cities as complex systems. He explores the potential of granular data provided by citizens and their contribution to urban research.

Niu and Silva (2020, p. 8) state that POIs data can connect human activities with the built environment, and, therefore, researchers can "understand purposeful urban activity and the function of land use." Accordingly, Miao et al. (2021) discuss how exploratory research can create new ideas by combining POI and road network data with spatial methods to understand the urban spatial structure and explore functional areas.

Similarly, Long and Liu (2015, p. 4) understand POIs as a means to "infer land-use intensity, function, and mixing at the parcel level." These authors mention human-oriented urbanization, emphasizing the granularity of research data and the human scale of cities. Therefore, we propose the use of crowdsourcing as a new research paradigm. Shen (2017) considered POIs and social media check-in data to represent land use regarding measures of functional centrality structures. This author state that VGI can be appropriately considered while analyzing urban function patterns (Shen, 2017).

Nevertheless, besides the positive aspects of VGI, when we consider the population of users and producers of crowdsourced data, open and collaborative data sources represents a small part of the population (Niu and Silva, 2020). So, we should not replace robust census data collection methods with crowdsourced data but include them in an exploratory investigation, which is the goal of this work.

At first, we had information regarding logistics sprawl for 48 metropolises. A preliminary exploratory step was understanding if the POI per capita was similar among these metropolitan areas. We found that Las Vegas and Seattle were outliers (interquartile range = 1.5) and decided to withdraw these metropolitan areas from the sample. Therefore, 46 metropolitan areas were considered regarding the proposed method.

We made sure that there were no facilities related to logistics activities within the POI dataset for the metropolitan areas, avoiding duplication of information when relating the urban attributes. Additional demographic and geographic information was collected in local public data from governmental agencies (Federal Register, 2010).

We performed different empirical analyses of the distribution of the urban activity index within metropolitan regions to decide on a proportion of hexagonal bins that could better represent spatial structure differentiation. We explored quartiles, quintiles, and deciles, but we adopted this distribution (5% + 95%) to evidence differences between activity hubs and peripheral activity zones.

The data was then spatialized, and additional data were collected from open-source and unstructured datasets regarding locational data on urban areas' spatial structure and logistics real estate information. The Urban Activity Index proposed for this work is detailed in Section 3.2.1.

3.1.3. Warehouse real estate data

We gathered warehouse real estate data (location and rental prices) for 46 US metropolitan areas to understand the relationships between this information and respective spatial structures. We explored different real estate websites for the US counties that composed the MSAs. However, most of them presented the information only by employing direct consultation with the real estate company. This approach could not be performed for different geographic contexts on the necessary scale. Some of the websites explored were Reonomy,[1] Prologis,[2] Loopnet,[3] and CBRE.[4] The LoopNet website had a significant amount of information with location (address), facility size (area), and monthly rental prices available for an immediate consultation and, therefore, was the chosen data source for this exploratory work (LoopNet, 2020). Data was then collected for the 46 US MSAs, resulting in the identification of 7370 warehouses.

[1] https://www.reonomy.com/properties/industrial-property/us/.
[2] https://www.prologis.com/property-search/.
[3] https://www.loopnet.com/search/listings/warehouses/usa/for-lease/.
[4] https://www.cbre.com/properties/properties-for-lease/industrial/.

In Fig. 2, we present the parameters for the automated search through the customized scrapping algorithm and the exhibited results for each consultation. The results were collected and structured on a county-wide dataset composing the MSAs under investigation.

The chosen real estate website presents unstructured data (e.g., no data frame with observations in rows and variables in columns) as an *Html* web page. Therefore, a customized algorithm was developed for an automated web data extraction method: web scraping. Individual scripts were developed to scrape the warehouse information from the *Html* file and assemble one dataset for each metropolitan area (Battisti et al., 2019; Ge et al., 2021; Luo and He, 2021; Pineda-Jaramillo and Pineda-Jaramillo, 2021). This process was mainly composed of seven steps for this paper: (i) find the URL where the data is published; (ii) inspect the webpage to find the data from its source code; (iii) build the prototype code, which was written in R language and using the *rvest* package and intended to extract and prepare the data; (iv) generalize the code considering functions, loops and debugging and run the code alternating among US metropolitan areas; (v) store the data as an organized data frame; (vi) check and clean the gathered data; (vii) geocode the information on each warehouse.

One limitation of this work is that no information on the real estate websites could differentiate warehouses from other logistics facilities, such as small manufacturing or business equipment, which are generally more expensive to rent. Therefore, we have considered the industrial facilities for rent as the search criteria and the warehouse as a keyword to enhance the search algorithm. Nevertheless, we could not control the type of logistics facilities that were gathered in the organized dataset. Since this is an exploratory study, we decided to consider this data on logistics facilities.

In order to bring more data consistency to the analysis, we performed an outlier identification and excluded extreme values. We have considered the multiplier 1.5 applied to the interquartile range to identify these extreme values. Figs. 3 and 4 present the warehouses' size and price before and after the removal of extreme values.

For example, the New York metropolitan area (MSA) comprises twenty-three counties. The number of identified warehouses is 860 facilities (LoopNet, 2020), and the number of observations after 724 with information on rental prices, address, and size. The number of warehouses identified in each county for NY is listed in Table 1. It is essential to highlight that the official database in the US that lists all establishments by their NAICS code is the County Business Patterns (CBP). For 2019, the CBP lists 993 logistics warehouses (under code 493 of the NAICS) for the New York Combined Statistical Area (Schorung, 2022). Despite being official, the CBP data considers all existing and "active" warehouses, excluding closed ones. The County Business Patterns helps count and localize new warehouses. Nevertheless, the official database also has limitations regarding all warehouses' stock.

The final dataset was then composed of logistics sprawl measures, the variables regarding warehouse rental price, location, and size gathered through the consultation on the real estate website; and open-access data from OpenStreetMap (Points of Interest and road network) to represent the classification of urban areas according to the activity spatial concentration. Despite the computational effort to collect publicly available logistics real estate data, it is essential to highlight that they were still incomplete regarding only warehouses available (industrial facilities) for rent. Therefore, it does not represent the warehouse stock in each metropolitan area, and the average rental prices only include the available facilities.

Besides using open-access and unstructured data, this work's contribution is the disaggregated spatial structure exploration from open data from OpenStreetMap (OpenStreetMap contributors, 2020) to classify the spatial patterns of warehouse location, rent prices, and urban structure. This disaggregation and standardization of spatial units allow the exploratory study of metropolitan areas. Therefore, we have assembled all the data into a grid of hexagon cells with a short diagonal of five kilometers. Section road lengths were aggregated in each cell. As explained in the next section, the number of points of interest was computed for these spatial units to categorize urban space regarding the concentration of activities.

3.2. Spatial characterization of urban activity and warehouse attributes in each metropolitan area

A universal hexagon size was considered to standardize metropolitan space representation and eliminate the need to transform variables into area-related (density) ones. The ability to communicate and analyze the phenomena is more potent if the attributes are not calculated considering the area of the spatial units (such as density measures), especially in studies considering different contexts. Therefore, this study's chosen scale of analysis and spatial units (cells in grids) were hexagons with a dimension of 5 km in the minor diagonal (Ben-Joseph and Gordon, 2000; Birch et al., 2007; Crown et al., 2018).

To determine the territorial thresholds for each metropolitan area, we considered official geographical data sources (Federal Register, 2010) and excluded the portions of the natural area outside the administrative boundaries. Natural features within the territorial boundaries were not excluded as they impact the potential interactions in the regions.

The administrative information for the delimitation of the metropolitan regions was consulted in the documentation made available by the responsible institutions for providing statistical and geographic information in each region. We have considered the Metropolitan Statistical Areas (MSA) regarding the US metropolitan areas, and each MSA concerns at least one urbanized area with >50,000 inhabitants (Federal Register, 2010).

This methodological step aims at organizing derivate indicators to categorize city centers ("activity hubs") and peripheral areas ("peripheral activity zones"), warehouse locations, and respective rent prices.

3.2.1. Spatial characterization of city areas and urban activities

The urban variables transformed into cells were the number of points of interest (POI) and the sum of roads. The points of interest were not stratified according to activities' categories since the objective of collecting this information and the location of the road infrastructure was to identify the centralities of each metropolitan region. For this work, we define centralities as areas with higher intensity of activities and connectivity concerning the spatial distribution of these facilities (Sarkar et al., 2020).

The territory of each metropolitan region, divided into hexagonal bins, was overlaid with the spatial structure of POIs and highways. For each cell, the length of highways (Quinn, 2013) and the number of POI contained in that unit were then calculated to consolidate the spatial differentiation among metropolitan areas. We do not use population density in the method for two reasons: (i) population information is

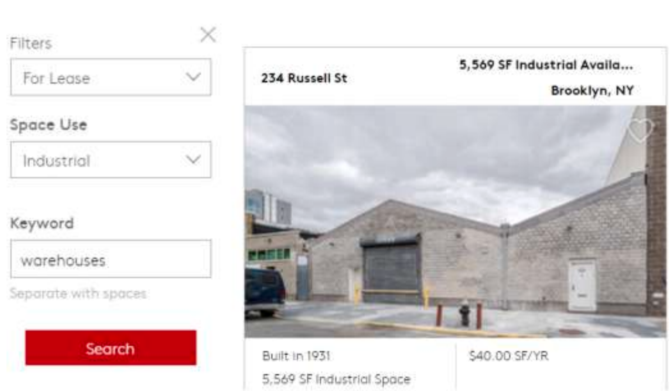

Fig. 2. Example of search parameters and warehouse rental information (LoopNet, 2020).

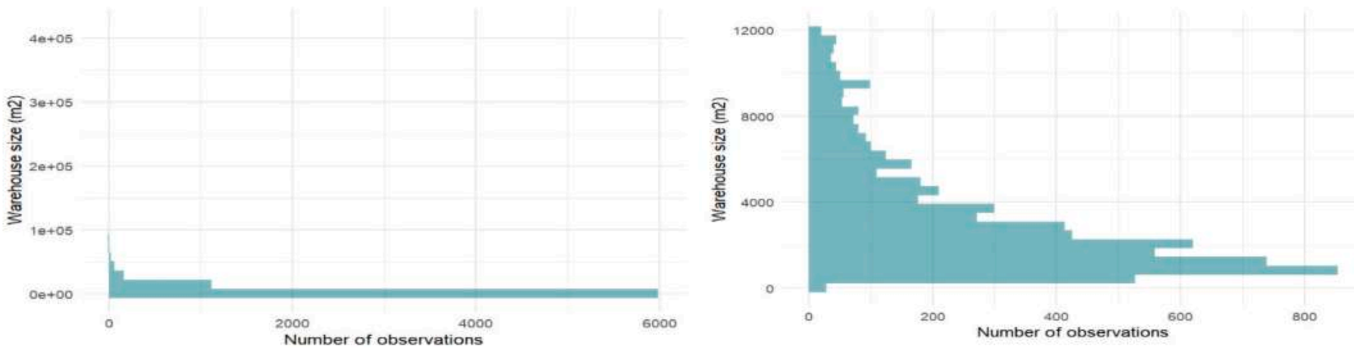

Fig. 3. Warehouse sizes before and after the removal of extreme values.

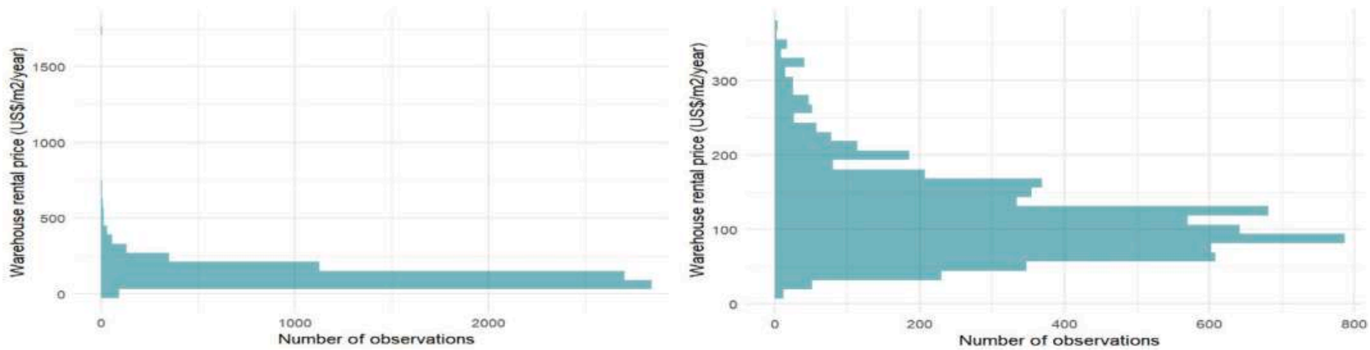

Fig. 4. Warehouse prices before and after the removal of extreme values.

Table 1
Number of warehouses identified for each county composing the New York MSA.

County	Number of warehouses	County	Number of warehouses
Bergen County	59	Ocean County	9
Bronx County	35	Orange County	15
Dutchess County	7	Passaic County	23
Essex County	34	Putnam County	4
Hudson County	19	Queens County	67
Hunterdon County	1	Rockland County	9
Kings County	99	Somerset County	14
Middlesex County	34	Suffolk County	105
Monmouth County	22	Sussex County	3
Morris County	29	Union County	29
Nassau County	62	Westchester County	45
New York County	0	**Total - warehouses**	**724**

aggregated in zones that are usually larger than the hexagons: we would have a different aggregation level of this information for each metropolitan area, leading to difficulties in analyzing different metropolitan areas in the same methodological framework; (ii) we would have to work with density variables and not absolute concentration. With homogeneous zoning, we could be more straightforward with the variables considered.

We then normalized the number of POI and road segments in each hexagon and composed a relative indicator of urban activity intensity, namely the Urban Activity Index (UAI). This index was composed of the sum of normalized road extension and the number of POIs for each hexagon λ. For the normalization of the variables, we considered a min-max approach. We have then considered outliers (3.0 hinge) to differentiate the metropolitan spatial structure, considering each cell's UAI to categorize areas within the metropolitan regions. The outliers in the lower bound and UAI lower than 95% of the overall information in each metropolitan area (percentile) were considered peripheral activity zones (PAZ). UAI values higher or equal to the top 5% (percentile) and upper outliers were considered activity hubs (AH). Table 2 shows this classification approach.

To explore the hypotheses designed for this study, the proportional difference between warehouse rental prices in activity hubs and peripheral activity zones of the metropolitan areas was also calculated from the classification of the hexagonal bins. This indicator was entitled Differential Warehouse Rent Prices (DWP) and concerns a continuous variable calculated from the ratio between the average warehouse rent price at activity hubs (AH) and peripheral activity zones (PAZ). This rationale is presented in Eq. (1).

$$DWP = \frac{Average\ warehouse\ rent\ price_{\mathbf{AH}}}{Average\ warehouse\ rent\ price_{\mathbf{PAZ}}} \tag{1}$$

Due to the lack of information for different timeframes concerning the urban activity, warehouse availability, and real estate information, these variables are static in time: we did not identify past DWPs.

Since the timeframes considered in previous logistics sprawl studies are different, we calculated the Yearly Logistics Sprawl (YLS). The YLS is the difference between the more recent average distance to the barycenter for logistics facilities and the previous measure (from the spatial structure of warehouses) divided by the period between these measures

Table 2
Classification of urban activity areas within metropolitan regions.

Classification (percentiles)	Category
Lower outlier	Peripheral activity zones
$< 95\%$	Peripheral activity zones
$\geq 95\%$	Activity hubs
Upper outlier	Activity hubs

(Eq. (2)).

$$YLS = \frac{MDB_{t1} - MDB_{t0}}{time\ between\ t0\ and\ t1} \quad (2)$$

where YLS is yearly logistics sprawl and MDB is the mean distance to the barycenter for each timeframe (the denominator in Eq. (2)).

3.2.2. Spatial characterization of warehouse location and rent prices

The same rationale adopted for the urban structure characterization was performed for warehouse location and rent price. The variables were: address and the number of warehouses for rental, the average warehouse rental price (US$/m^2/year) from the real estate site LoopNet (LoopNet, 2020). The addresses for warehouses available for rent were geocoded and spatialized.

The outliers and extreme values were identified for the warehouse concentration in the hexagons and the average prices. We have considered that values above Q3 + (1.5 x IQR) or below Q1 – (1.5 x IQR) were outliers. Values above Q3 + (3 x IQR) or below Q1 – (3 x IQR) were considered extreme outliers. We then categorized the hexagons with non-zero observations according to Table 3.

The consolidation of the indicators presented before resulted in the composition of two datasets at different scales: (i) a dataset of hexagonal bins for each metropolitan region and (ii) a dataset of summary indicators for all metropolitan regions. The results of this exploratory spatial and nonspatial analysis of the consolidated data are presented in Table 4. The data collection, cleaning, and processing were performed in Free Open-Source Software (FOSS): The R project 4.0.5 and GeoDa 1.14.0.

3.3. Exploring the hypotheses

3.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities

This first hypothesis does not increment the academic literature. Nevertheless, it was explored to understand the appropriateness of the method designed for this work. The anticipated results were in line with the classical bid-rent theory: warehouse rent is expected to be higher in the activity hubs like other real estate prices due to the scarcity of available land property in the urban centers. We then propose an analysis of this relationship through an intrametropolitan classification regarding the differential rent prices in central and peripherical areas.

To explore this first hypothesis, we have related three categorical variables: the number of warehouses, the average rent values, and the urban activity index. All three variables were computed for the hexagon bins for all metropolitan areas. We then visualize the relationship and perform Chi-square tests to understand if these variables are dependent on each other. It is essential to highlight that 46 metropolitan areas were considered, and only complete data, with no missing values, were included in the analysis. Pearson's Chi-squared test was done with simulated *p*-values (2000 replications) and a significance level of 5% to understand the dependence between urban classification and warehouse location/rent prices.

Table 3
Relative classification criteria for warehouse location and rent prices in each metropolitan region.

Classification (quantiles)	Category	
Low extreme values	Extremely Low	
Low values outliers	Very low	
< 25 %	Low	Range of values except for outliers
25 % - 75 %	Medium	
> 75 %	High	
High values outliers	Very high	
High extreme values	Extremely high	

Table 4
Results and analytical elements for the exploratory spatial and nonspatial analysis.

Item	Variables
A dataset containing all hex bins for all the metropolitan regions investigated	- Urban activity index (sum of normalized road length and POIs) - Categorical (binary) variable to determine activity hubs and peripheral activity zones for each metropolitan area - Number of warehouses in each hex bin - Average warehouse price in each hex bin - Classification of hex bins regarding warehouse prices.
A dataset containing the synthetic indicators for each metropolitan region	- Yearly logistics sprawl (YLS) from centrographic measures (ratio between variables from the previous studies) - Differential rent prices (DWP) (continuous variable calculated considering the category of activity hubs/ peripheral activity zones and average rent prices)

3.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential of land/rental prices between urban and peripheral areas

In this step, we develop different analyses to understand the relationship between urban structure and warehouse location and rent prices and address the second hypothesis proposed for this research: logistics sprawl is higher in cities with a high differential between city centers (more precisely, activity hubs, as defined above) and peripheral activity zones land/rent values.

We have performed a Pearson correlation analysis (coefficient and test) to understand the relationship between differential warehouse rent prices and yearly logistics sprawl. Different representations of the variables were also performed, and the metropolitan areas were classified according to the sprawling phenomenon for logistics facilities and the differential warehouse prices.

4. Results

4.1. Data collection and treatment

The results of this methodological step are presented together with the following steps since, after collection and treatment, data were gathered into the proposed spatial units.

Fig. 5 presents the results of each investigation collected in published papers for US metropolitan areas (Dablanc et al., 2014; Dablanc and Ross, 2012; Giuliano and Kang, 2018; Kang, 2020a, 2020b). This information is the average distance to the gravity center in kilometers (logistics sprawl measure) for each metropolitan area's two timeframes considered in the studies.

Fig. 6 presents the average warehouse rental price for each metropolitan area. After the outliers are excluded from the dataset considering the methodological approach presented in the respective section, the metropolitan areas with the highest average warehouse rent prices are San Francisco, Los Angeles, San Diego, New York, and Miami. As a reference, the average warehouse rent price in San Francisco is 178.49 US$/m2/year.

4.2. Spatial characterization of urban activity and warehouse attributes in each metropolitan region

The main results of this methodological step are the organization of data into standard spatial areal units for all metropolitan regions, described in Section 3.2 of this paper. The consolidation of information is presented throughout this section represented through the generated

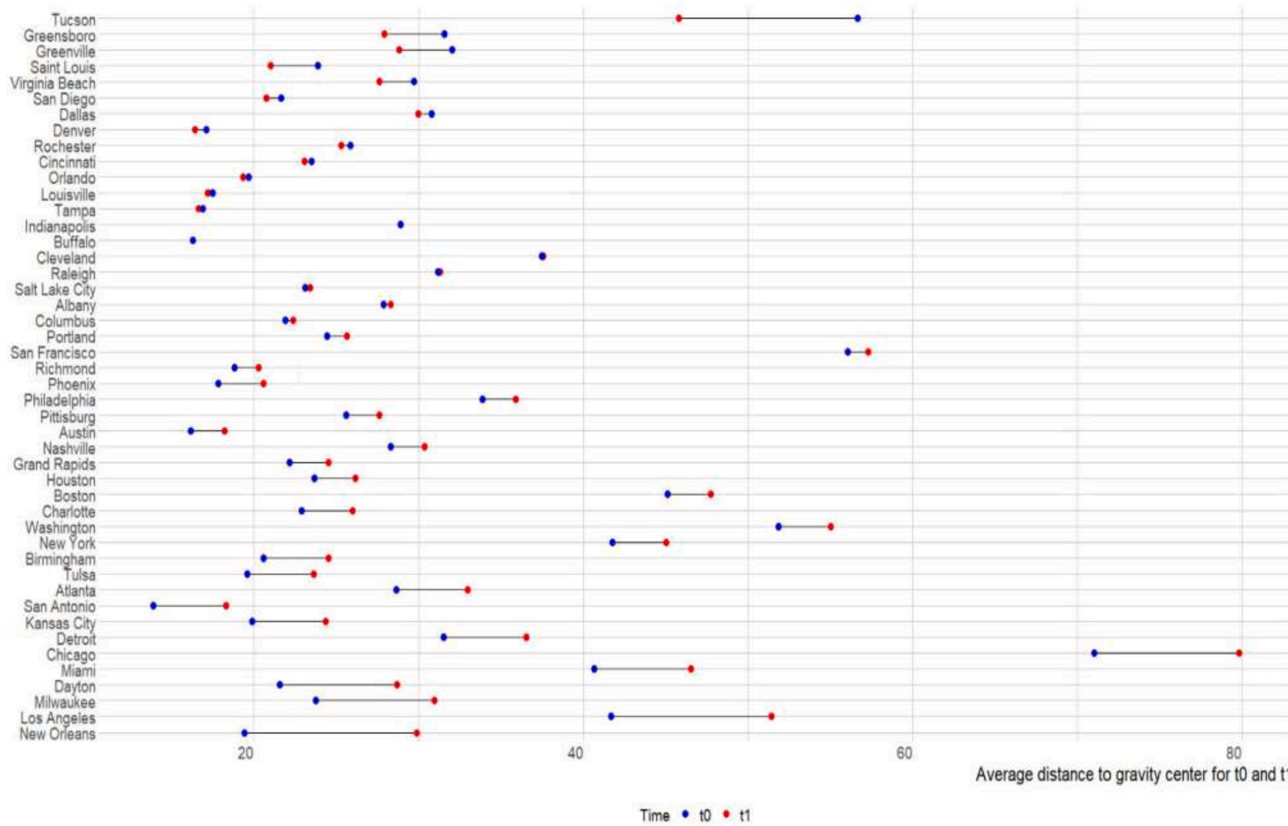

Fig. 5. Logistics sprawl measures.

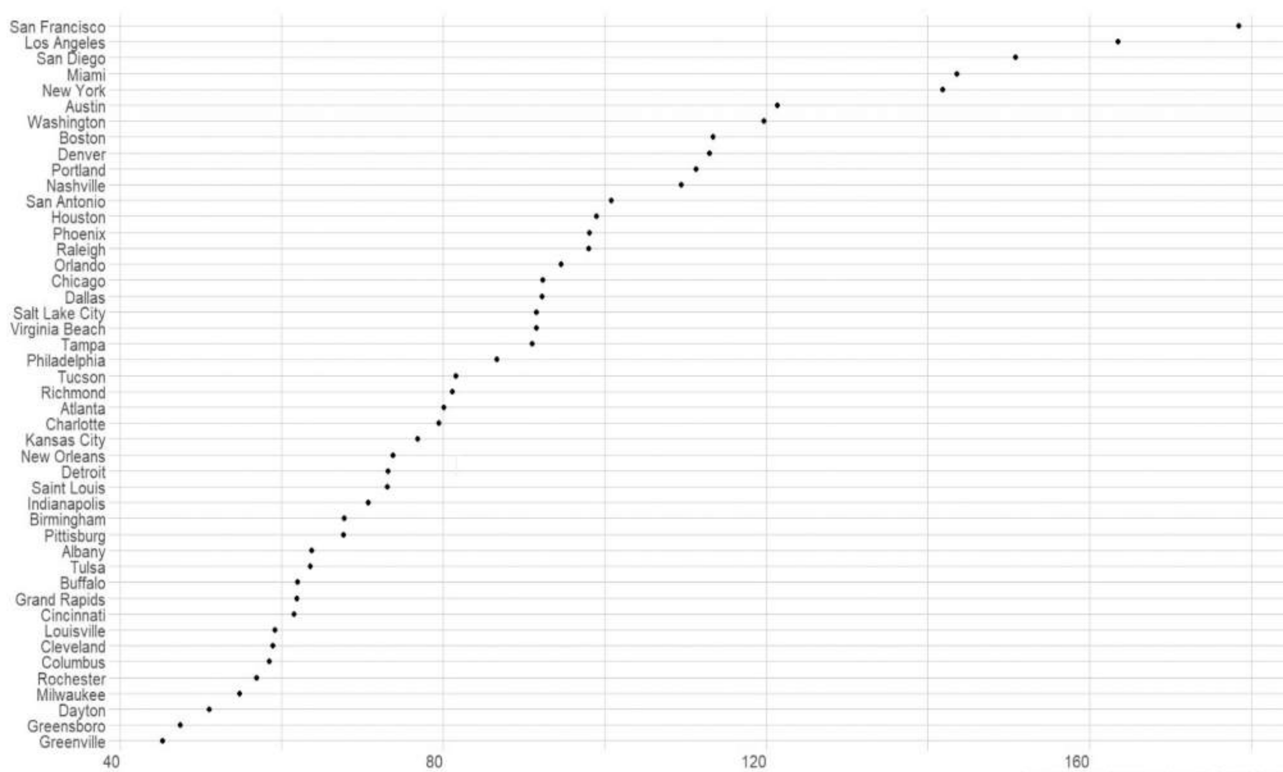

Fig. 6. Average warehouse rent prices (US$/m^2/year).

data:

i. urban classification (Activity Hubs, Peripheral Activity Zones) considering the Urban Activity Index (UAI);
ii. warehouse location and rent prices classification according to Table 2;
iii. differential rent prices related to the categories through the UAI;
iv. yearly logistics sprawl (Dablanc et al., 2020).

4.2.1. Spatial characterization of city areas and urban activities

Concerning the 46 metropolitan areas, we have transformed all the variables into hexagon bins as the sample for analyzing the relationship between the location of warehouses and the rent values of logistics (first hypothesis). There were 2486 hexagon bins with warehouses in them. After classifying the areas corresponding to Activity Hubs and Peripheral Activity Zones, these bins were considered for the analysis.

In Fig. 7, we present one example of the classification performed for the metropolitan region of New York. The classification method was also applied to all the US metropolitan areas under investigation.

4.2.2. Spatial characterization of warehouse location and rent prices

Figs. 8 and 9 represent the number of warehouses and the average warehouse rent prices, collected from the real estate website LoopNet (LoopNet, 2020), spatialized and classified according to the metropolitan area in AH and PAZ.

The pure examination of the spatial representation of the variables in Figs. 8 and 9 cannot drive any exploratory interpretation of the phenomenon. Therefore, Section 4.3 presents the statistical analysis to help understand the relationships between variables.

Fig. 10 presents AH and PAZ average warehouse rent prices for each metropolitan region. Fig. 11 represents the proportional warehouse differential (ratio between AH/PAZ) areas. The areas with no significant differential were <10% of the proportional difference in warehouse rent prices between central (AH) and suburban (PAZ) areas. Therefore, metropolitan areas with a ratio of <0.9 were classified as having higher PAZ and significant differences. The ones with a ratio >1.1 were categorized as metropolitan areas with higher warehouse prices in AH. The remaining regions were classified as having no significant differential. Three groups of cities were derived from this differential.

New Orleans, San Antonio, Albany, and Detroit present a counterintuitive result: warehouse rent prices are significantly higher in periphery activity zones. This result needs further exploration in incremental investigations, looking individually at each metropolitan region.

Finally, Fig. 12 presents the yearly logistics sprawl indicator. The three categories are (i) areas with negative indicators (no sprawl but recentralization over the years); (ii) areas with no significant yearly sprawl ($-1 < \text{YLS} < 1$) and (iii) areas with positive YLS.

4.3. Exploring the hypotheses

4.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities

For this analysis, we present the Chi-square test regarding the warehouse location (Urban Activity Index), warehouse concentration (number of warehouses), and average rent values for all hexagon bins, disregarding the metropolitan region to which they belong. This approach was designed to test hypothesis 1: Are warehouses' location and rent prices related to urban activity? For this analysis, we have considered only the hexagon bins with warehouses. Thus, 2486 hexagon bins were considered (all complete cases). Figs. 13 to 15 present the variables' relationships and the Chi-Square test results.

Considering the metropolitan regions (Figs. 13 and 14), 35% of the warehouses are AH and 65% PAZ. From the χ^2 we can reject the

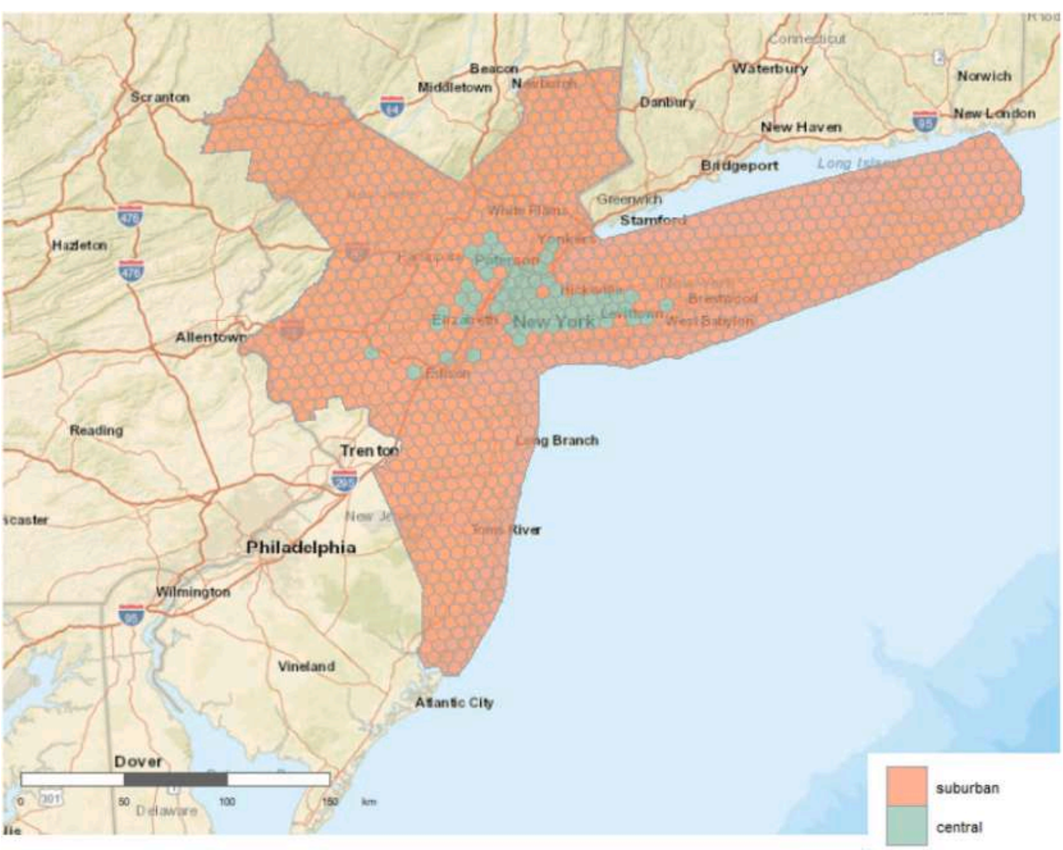

Fig. 7. Classification for activity hubs and peripheral activity zones from the urban activity index.

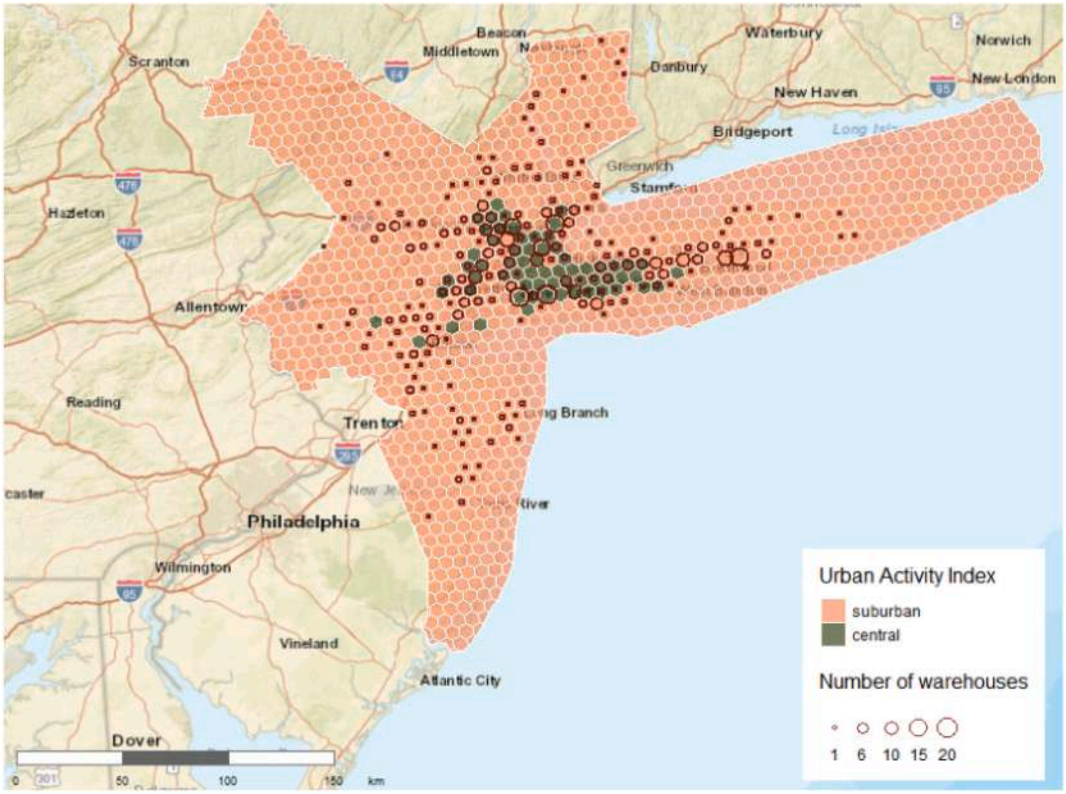

Fig. 8. Urban classification and spatialized warehouse concentration – New York metropolitan area.

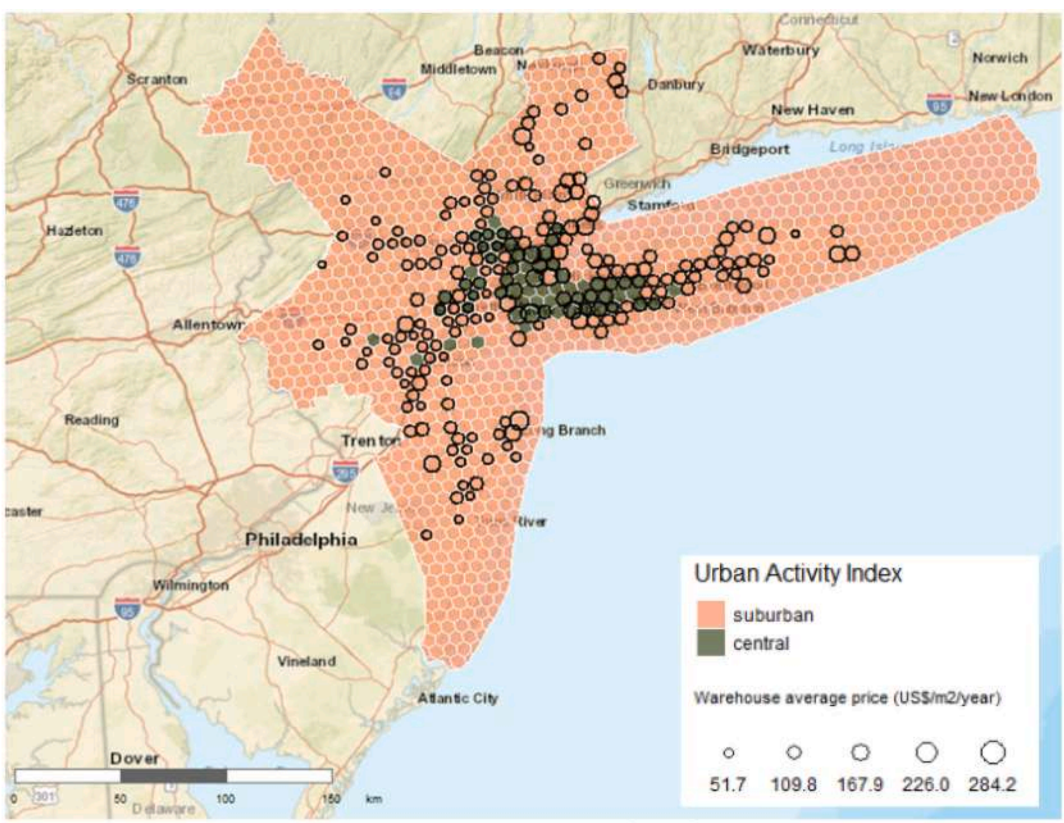

Fig. 9. Urban classification and spatialized average warehouse rent prices – New York metropolitan area.

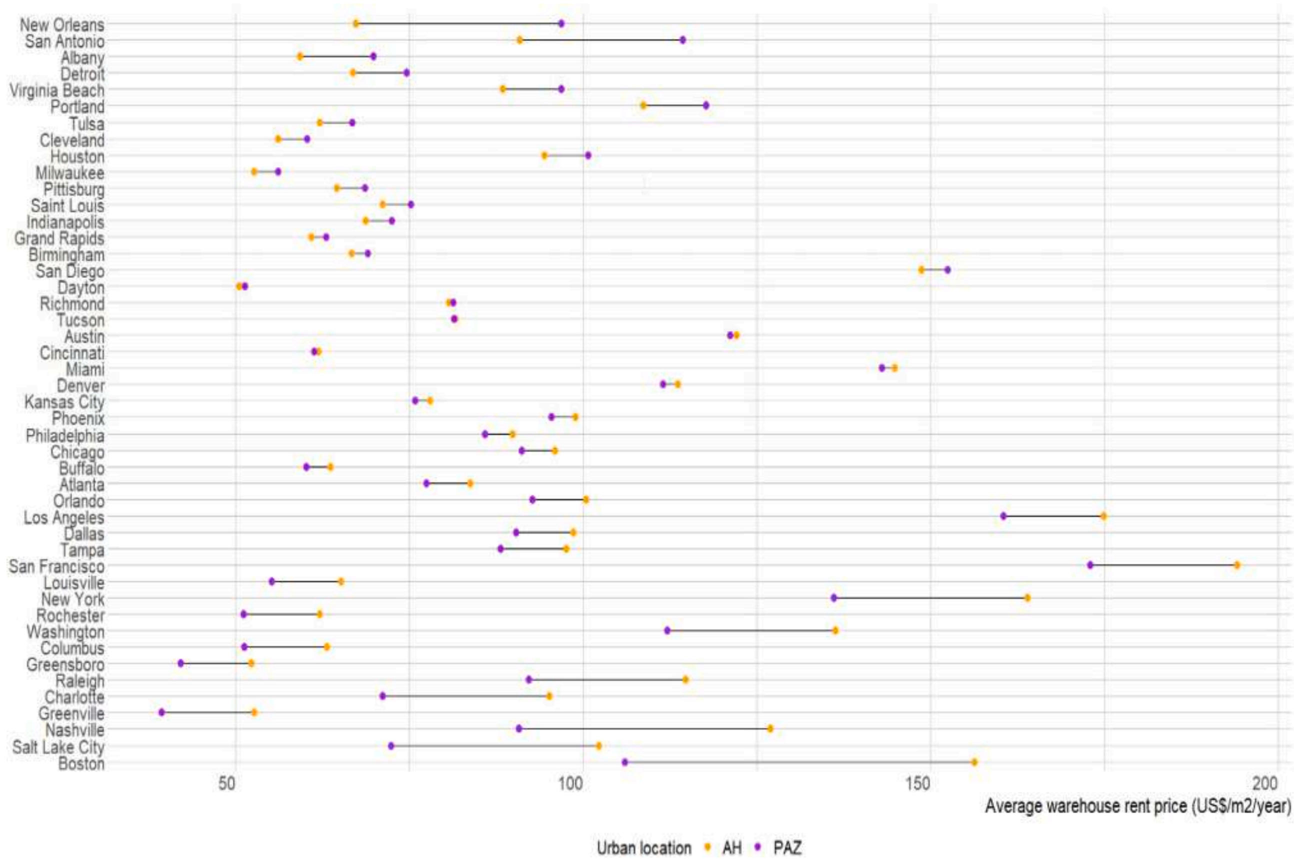

Fig. 10. Average warehouse rent prices classified by urban activity.

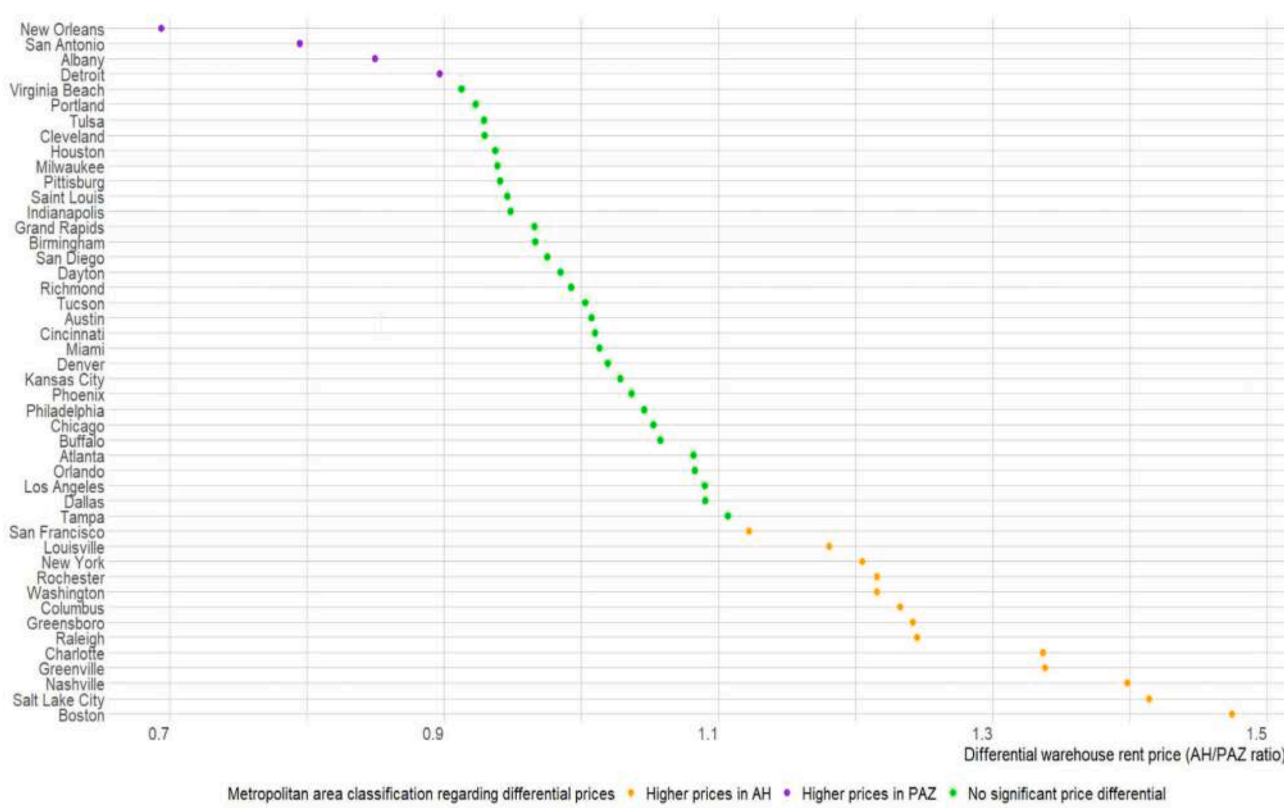

Fig. 11. Representation of rent price differential for warehouses for each metropolitan region.

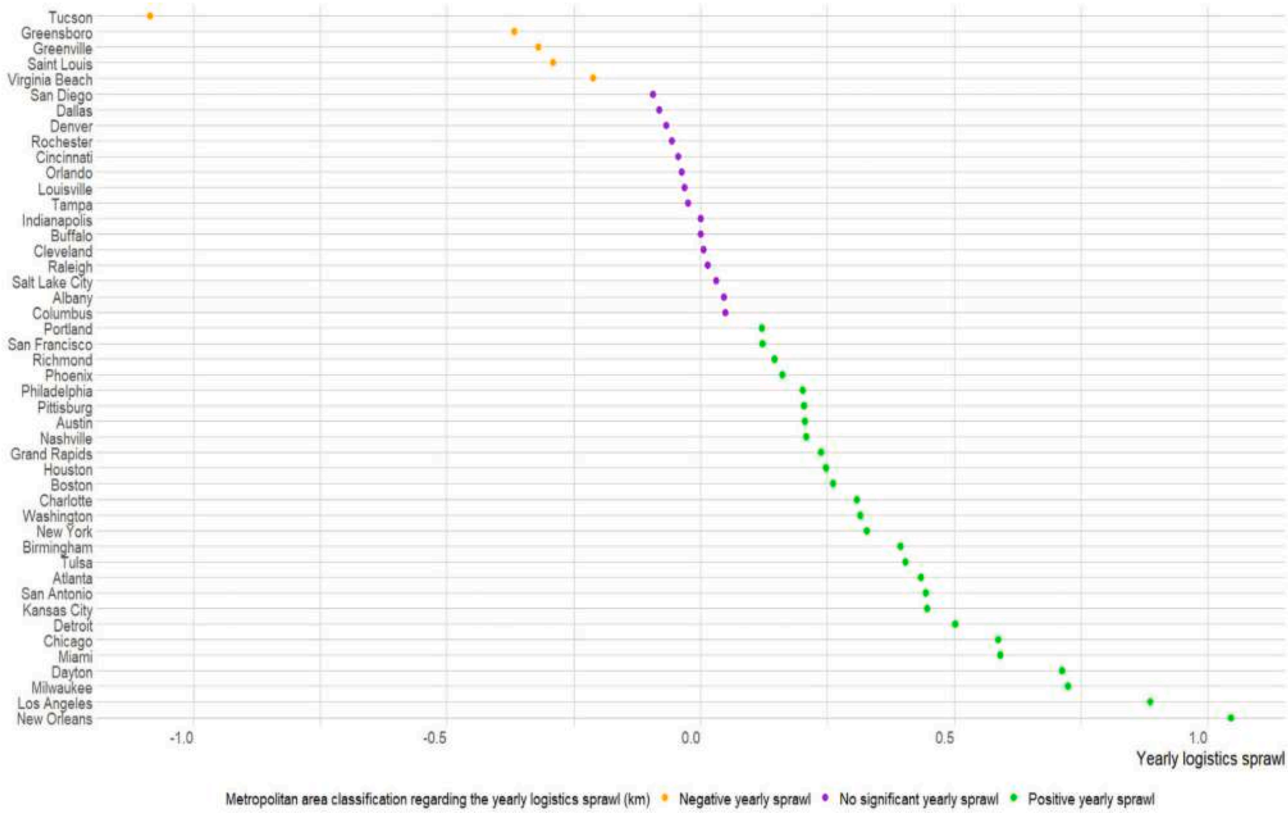

Fig. 12. Representation of yearly logistics sprawl.

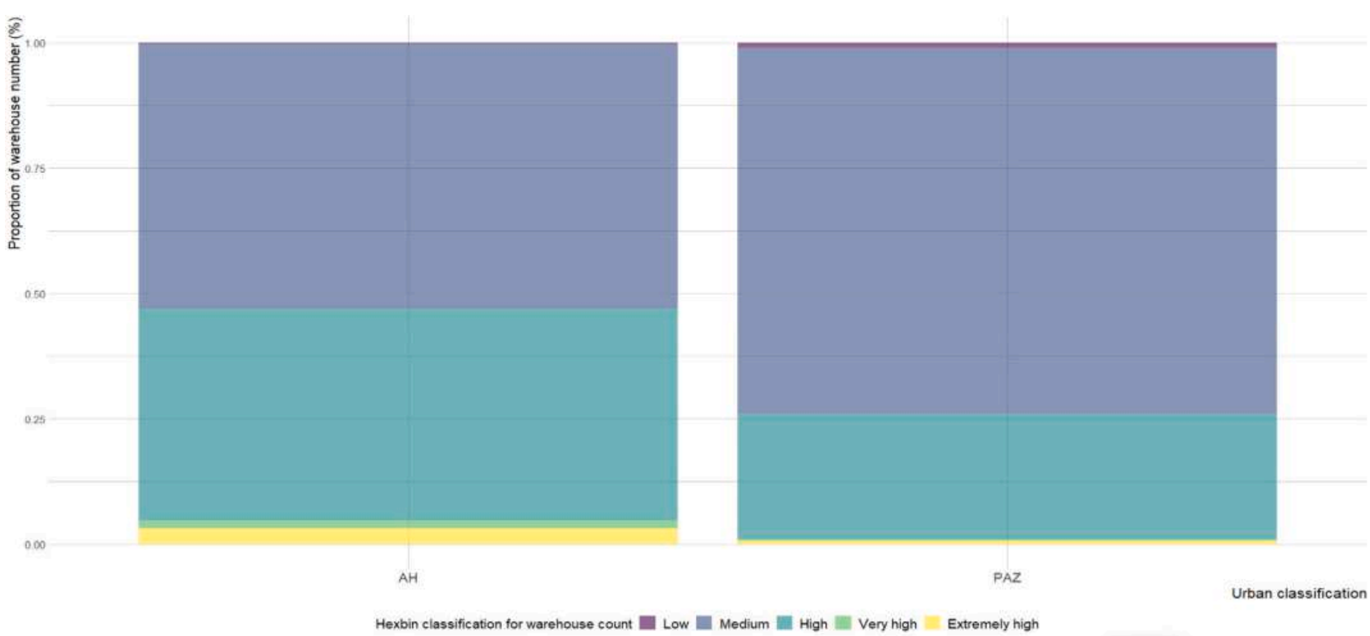

Fig. 13. Warehouse count x urban classification.

independence hypotheses and, therefore, assume that, in this case, the concentration of warehouses depends on the urban activity index. Also, AH areas have a higher proportion of hexagons with a high number of warehouses than PAZ.

Fig. 15 shows the relationship between average warehouse rent prices and the location of these facilities within the city. Also, considering all metropolitan areas, the average warehouse prices decrease as we move closer to PAZ, and we can reject the hypothesis that these variables are independent. In other words, considering a significance level of 5%, the warehouse rent prices depend on location in the metropolitan areas.

Synthetically, considering all metropolitan areas investigated, it is possible to state that the rent prices of the warehouses depend on the location (AH or PAZ) and spatial concentration of these logistics

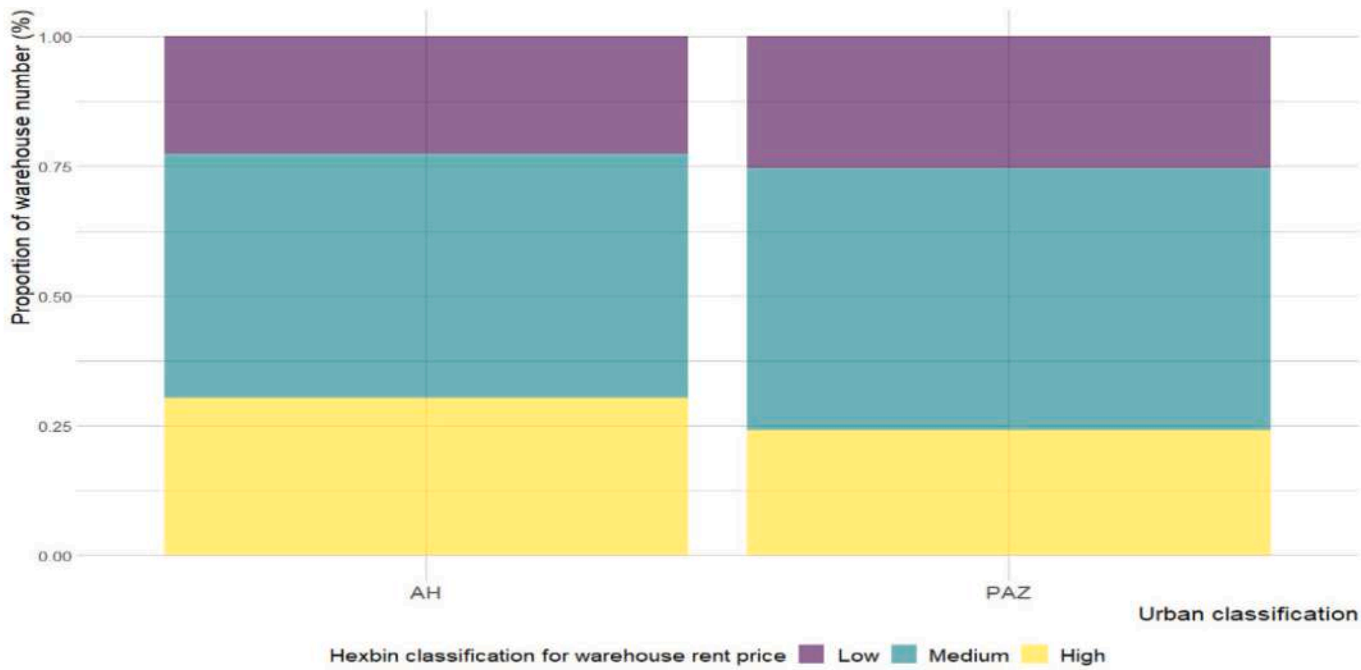

Fig. 14. Warehouse average prices x urban classification.

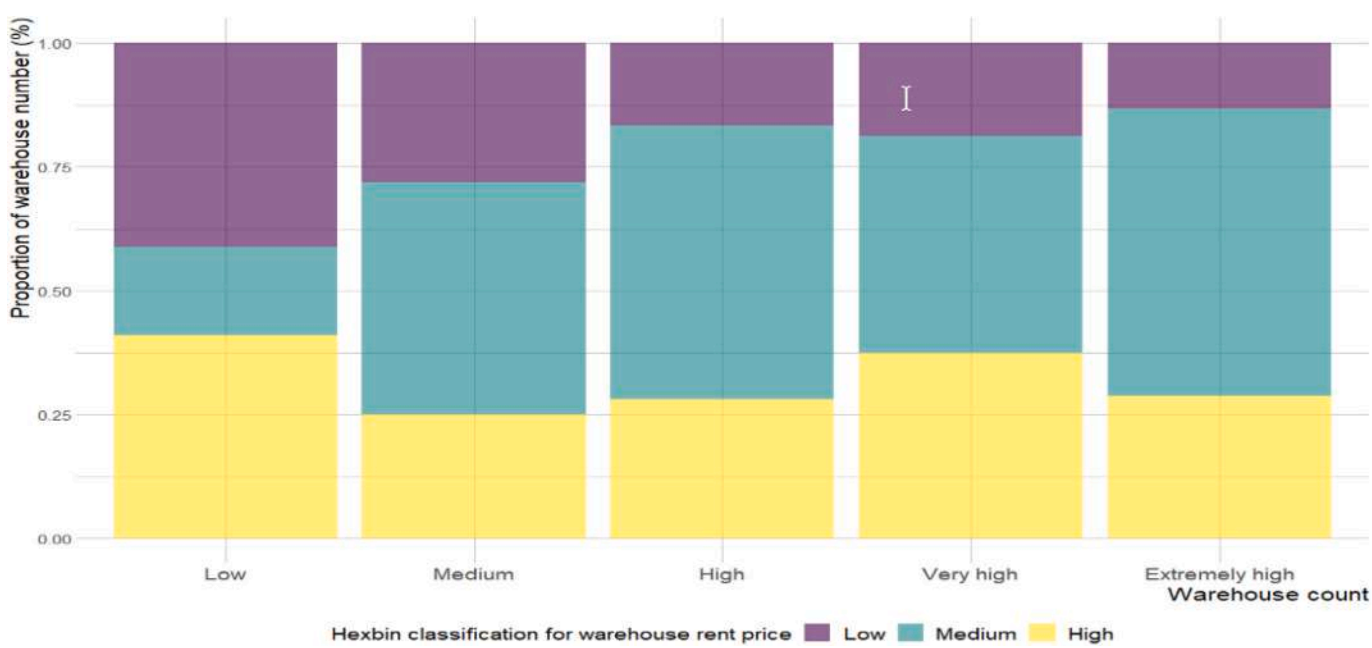

Fig. 15. Warehouse count and warehouse average rental prices.

facilities. This interpretation is related to a possible effect of agglomeration economy or locational decisions performed for similar reasons (Onstein et al., 2019). Therefore, this work's first hypothesis cannot be rejected from the tests performed considering the investigated metropolitan areas and methods.

Still analyzing the first hypothesis (the location of warehouses is closely related to the land/rent values of logistics facilities), considering the classification of differential prices, we have then performed the Pearson Correlation test for all the hexagon datasets (46 US metropolitan areas) and each combination of categories, relating the variables (i) differential warehouse price; and (ii) urban activity index.

For all the metropolitan areas, the Pearson coefficient was 0.099172 (p-value = 1.4e-06). Since the p-value is lower than 0.05 (5% significance level), the null hypothesis that the relationship between the average warehouse rent prices and the urban activity index is insignificant can be rejected. Therefore, the greater the urban activity index, the higher the rent prices.

If we consider the dataset stratified for each category, relationships are still significant, except for the lower prices in activity hubs, which is counter-intuitive and opposes the literature. The metropolitan areas with lower rent prices in activity hubs should be further investigated to understand the urban structure in relation to warehouse rent prices.

- Higher prices in activity hubs: $r = 0.17766$ (p-value = 2.2e-06)
- Lower prices in activity hubs: $r = -0.079609$ (p-value = 0.39)
- No significant differential prices: $r = 0.13738$ (p-value = 1.8e-08)

4.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential of land/rental prices between urban and peripheral areas

This section explores the urban attributes, logistics facilities' spatial structure, and real estate practices in metropolitan areas. For that, two variables were used: (i) differential warehouse prices (DWP) and (ii) yearly logistics sprawl (YLS). DWP is the ratio between the average warehouse price in activity hubs and peripheral activity zones (in 2021, analysis time). YLS is the difference in average distance to the mean center between years in analysis divided by the number of years considered in logistics sprawl studies. Fig. 16 shows the classification of metropolitan areas according to the differential price category (three groups) and the relationship between YLS and DWP.

In activity hubs, there are four metropolitan areas with lower warehouse rent prices – Detroit, New Orleans, San Antonio, and Albany. These findings can be counter-intuitive. One hypothesis is that the logistics facilities in activity hubs are aging and obsolete and, therefore, have lower rent prices. On the other hand, warehouses in peripheral activity zones have been more recently built or remodeled and could present higher rental prices.

When calculating the correlation coefficient, it is assumed that at least one of the variables is normally distributed. For that, we have performed the Shapiro-Wilk normality test for DWP and YLS in each category of differential prices. We could only reject the null hypothesis for DWP, at a significance level of 5% ($p > 0.05$), for the metropolitan areas classified as having *no significant differential price,* considering the YLS. In other words, we can say that all metropolitan variables considering the differential price classification are normally distributed except for the combination of *no significant differential price and* YLS. Therefore, assessing the variable DWP and its relationship with the YLS for the metropolitan area sets is possible.

Regarding the second hypothesis (logistics sprawl is higher in cities with a high differential between land/rent values in Activity Hubs and Peripheral Activity Zones), we have computed the Pearson correlation index to investigate the relationship between YLS and DWP. Considering the DWP and YLS for each metropolitan area (dataset with 46 metropolitan areas), the $r = -0.23194$ (p-value = 0.12). For each category, we have:

- No significant price differential: $r = 0.088623$ (p-value = 0.65)
- Lower prices in Activity Hubs: $r = -0.7449$ (p-value = 0.26)
- Higher prices in Activity Hubs: $r = 0.1194$ (p-value = 0.7)

The null hypothesis that the relationship between the differential warehouse prices (DWP) and the yearly logistics sprawl (YLS) is not significant cannot be rejected for all the data combinations.

There is no evidence that differential warehouse rent prices are related to the yearly sprawl. Therefore, the second hypothesis is rejected, considering the method to relate yearly logistics sprawl with differential warehouse rent prices.

5. Discussion and final considerations

In this work, we present a disaggregated methodological approach to explore the hypotheses: (i) the location of warehouses is closely related to the land/rent values of logistics facilities, and (ii) logistics sprawl is higher in cities with a high differential between central and suburban land/rent values.

Regarding the first hypothesis, which aimed at investigating whether warehouse rent prices were higher in more central places within metropolitan areas, we can state that there is statistically significant evidence that the location of warehouses and average rent prices are not independent.

We explore the correlation between the differential rent values in central and suburban areas; and yearly logistics sprawl (second hypothesis). Considering all metropolitan regions, we identified that the null hypothesis (the relationship is not significant) could not be rejected. Nevertheless, since the metropolitan areas are significantly different in many dimensions, this statement cannot be generalized when we gather them into more similar groups.

Finally, we conclude that: (i) it is essential to classify portions of the urban space into a typology to allow the inclusion of different metropolitan areas, from different contexts, into one methodological framework; (ii) for the scope explored, warehouse location and prices are related to the density of urban activity; (iii) for the scope explored, logistics sprawl is not significantly related to the differential in warehouse rental prices in central and suburban areas.

Many limitations exist in this work, and some additional directions should be further explored in other research efforts. The price differential for renting warehouses is based on a sample of prices from specialized websites that are not comprehensive. Other websites exist and were not considered. We do not have a complete database that lists

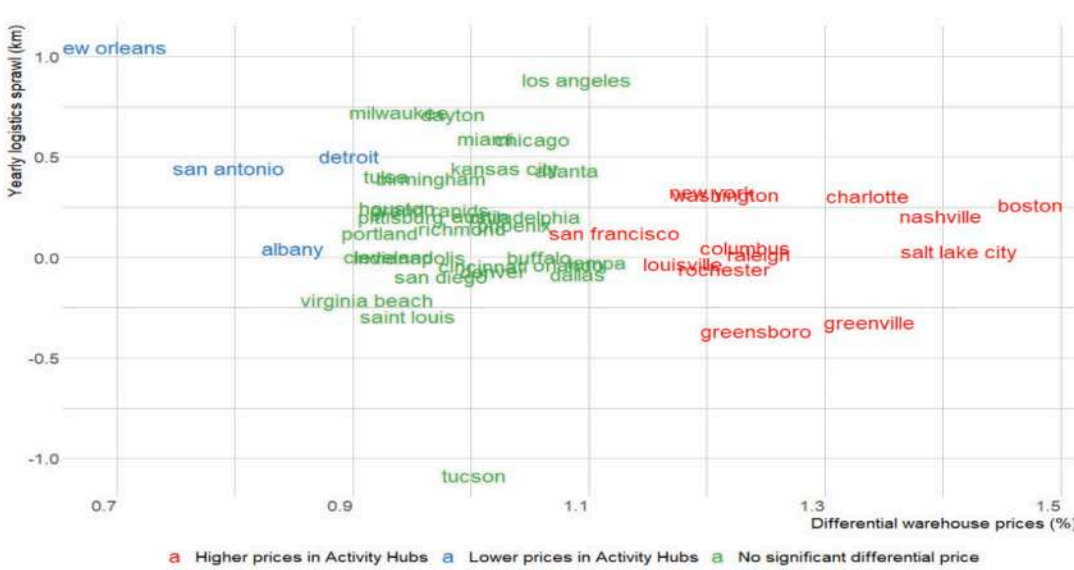

Fig. 16. Differential classes and a scatterplot for DWP and YLS.

all real estate and land values for warehouses in each geographic area and at different times. There is a risk that the warehouses' data collected (i) under-represent the whole number of facilities or (ii) result in an inadequate representation of the spatial distribution of warehouses in each metropolitan area. Further investigation needs to be performed to quantify this limitation. In further work, the County Business Patterns and scraping in real estate sites can be combined better to represent the stock of warehouses in each metropolitan area.

The proposed method concerns VGI, which results in incomplete intersectional data. Comprehensive data is even more scarce when research is held in underdeveloped countries. The completeness of VGI depends on citizens' access to technology (Bright et al., 2018). Despite having improved in recent years, this access is still inequitable. So, it is still important to clearly state the limitations of the works, even when of exploratory nature. We understand that the potential of using OSM data will require long-term efforts to use it in academic research and that our work brings one contribution but will require other research efforts to address intersectionality issues of VGI.

Besides location, other factors that influence warehouse rent prices should be included in further work. These factors concern the buildings' physical characteristics, transactional agreements, and accessibility to transportation infrastructure, for instance.

The real estate prices for logistics facilities were collected for one static timeframe, which compromises the analysis of the dynamic relationship between warehouse structure and rental prices. They were only representative of the warehousing rental market in 2020 when the data were collected. The time of our analysis is during the COVID-19 pandemic, which displayed an increase in warehouses' prices, as the market has shown high demand for warehouses and cost increases. We recommend a systematic collection of this information to consolidate a dataset that can reflect this dynamic phenomenon.

It is possible that areas with a relatively high density of urban activities were overlooked because of the methodological decision to classify this attribute considering 5% and 95% of the UAI distribution. Despite the exploratory nature of this work and having empirically explored different classification reasonings, we recommend that further research should make an effort to quantify the possible results considering other classes for differing urban regions.

Determining a typology for the metropolitan regions might be necessary if we include other areas and cities from different contexts. In this case, multivariate analysis can help the identification of similarities and dissimilarities among metropolitan regions.

This initial research, especially for the second hypothesis (links between logistics sprawl and differential in rental prices between urban and suburban areas), will require further study, particularly by integrating cases from other world regions. Asian cities (in particular Japanese, Chinese, and South Korean cities) and European cities would help obtain a valuable sample for international analysis.

We have not identified a statistical relationship between logistics sprawl and differential warehouse rent prices for central and peripheral areas, considering the method proposed for this work. This finding does not strictly oppose the results of Kang (2020b). Kang (2020b) state that the "results provide robust evidence that high land prices push large warehouses away from central locations." Nevertheless, we did not relate land prices to sprawl. We investigated the relationship between the *differential* warehouse rent prices between central and peripherical areas and yearly sprawl measures. These different outcomes need to be further explored, including other factors (e.g., economies of agglomeration concerning infrastructure and economies of scale) that need to be considered for locational decisions (Onstein, 2021; Onstein et al., 2019).

The number of warehouses for each metropolitan area is significantly different. This issue can result in bias while analyzing the areas comparatively, but we addressed this issue through homogeneous disaggregated spatial units and normalization of variables.

The definition of metropolitan areas for mapping and data analysis (number of warehouses, prices) required delineation choices. These delimitations can, in some cases, modify the cartographic rendering and the final analysis, depending on the size of the metropolitan area, for example. However, we believe our delimitations were overall reasonable.

In some metropolitan areas, peripheral warehouses are more expensive than warehouses located in central or pericentral areas. This result instigates further research, which explores data collection to understand possible causes, such as construction age and the number of concurrent facilities within the same region.

This work brings a methodological contribution since we present a framework for exploring metropolitan regions considering the spatial pattern of logistics facilities and urban characteristics through disaggregated open data. This method is reproducible for other approaches and other city scales.

It can help decision-making toward developing more effective public policy on logistics land use and transportation planning. Public policymakers can, for instance, consider these relationships regarding price and location to stimulate ruptures in the current real estate trend. For example, public stakeholders support and fund urban logistics projects in France, and this approach can be explored in different geographical contexts as an effort to promote more direct actions of public stakeholders. Also, the results of this work can promote incentives for warehouses to be located in more central areas, aiming to promote a trade-off between location price and distribution costs. In this case, more central warehouses, which are beginning to characterize the recent changes in the logistics real estate market, can result in less environmental impact through reduced traveled miles and the possibility of using non-motorized alternatives of last-mile distribution. Coordinating these dimensions is essential to support urban logistics stakeholders' needs, cities' livability, and the real estate market. Likewise, land use regulation can also be proposed to limit sprawl in the outskirts of metropolitan areas.

CRediT authorship contribution statement

Renata Lúcia Magalhães de Oliveira: Conceptualization, Methodology, Software, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Matthieu Schorung:** Supervision, Conceptualization, Methodology, Validation. **Laetitia Dablanc:** Supervision, Conceptualization, Methodology, Validation.

Data availability

Data will be made available on request.

Acknowledgments

We acknowledge the support received from the Logistics City Chair, the Université Gustave Eiffel (UPEM et IFSTTAR), Sogaris, Poste Immo, and Région Île-de-France to develop this research. We also thank the National Council for Scientific and Technological Development (CNPq) for the support (314045/2021-0).

References

Aljohani, K., Thompson, R.G., 2016. Impacts of logistics sprawl on the urban environment and logistics: taxonomy and review of literature. J. Transp. Geogr. 57, 255–263. https://doi.org/10.1016/j.jtrangeo.2016.08.009.

Allen, J., Browne, M., 2010. Sustainability strategies for city logistics. In: McKinnon, A., Cullinane, S., Browne, M., Whiteing, A. (Eds.), Kogan Page, London.

Allen, J., Browne, M., 2012. Sustainability strategies for city logistics. In: McKinnon, A., Browne, M., Whiteing, A. (Eds.), Kogan Page, London, pp. 289–311.

Andreoli, D., Goodchild, A., Vitasek, K., 2010. The rise of mega distribution centers and the impact on logistical uncertainty. Transp. Lett. 2, 75–88. https://doi.org/10.3328/tl.2010.02.02.75-88.

Baglio, M., Perotti, S., Dallari, F., Garagiola, E.R., 2019. Benchmarking logistics facilities: a rating model to assess building quality and functionality. Benchmarking Int. J. 27, 1239–1260 doi:10/gm5gn8.

Bakillah, M., Liang, S., Mobasheri, A., Jokar Arsanjani, J., Zipf, A., 2014. Fine-resolution population mapping using OpenStreetMap points-of-interest. Int. J. Geogr. Inf. Sci. 28, 1940–1963 doi:10/gfxbbr.

Battisti, E., Shams, S.M.R., Sakka, G., Miglietta, N., 2019. Big data and risk management in business processes: implications for corporate real estate. Bus. Process. Manag. J. 26, 1141–1155 doi:10/ghqxxj.

Batty, M., 2007. Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals, 1. Paperback Ed. MIT, Cambridge, Mass, London.

Ben-Joseph, E., Gordon, D., 2000. Hexagonal planning in theory and practice. J. Urban Des. 5, 237–265. https://doi.org/10.1080/713683965.

Birch, C.P.D., Oom, S.P., Beecham, J.A., 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecol. Model. 206, 347–359. https://doi.org/10.1016/j.ecolmodel.2007.03.041.

Bowen, J.T., 2008. Moving places: the geography of warehousing in the US. J. Transp. Geogr. 16, 379–387. https://doi.org/10.1016/j.jtrangeo.2008.03.001.

Bright, J., De Sabbata, S., Lee, S., 2018. Geodemographic biases in crowdsourced knowledge websites: do neighbours fill in the blanks? GeoJournal 83, 427–440. https://doi.org/10.1007/s10708-017-9778-7.

Buttimer, R.J., Rutherford, R.C., Witten, R., 1997. Industrial warehouse rent determinants in the dallas/fort worth area. J. Real Estate Res. 47–56 doi:10/gndc2z.

Cidell, J., 2010a. Concentration and decentralization: the new geography of freight distribution in US metropolitan areas. J. Transp. Geogr. 18, 363–371. https://doi.org/10.1016/j.jtrangeo.2009.06.017.

Cidell, J., 2010b. Concentration and decentralization: the new geography of freight distribution in US metropolitan areas. J. Transp. Geogr. 18, 363–371. https://doi.org/10.1016/j.jtrangeo.2009.06.017.

Clark, D., Pennington-Cross, A., 2016. Determinants of industrial property rents in the Chicago metropolitan area. Reg. Sci. Urban Econ. 56, 34–45. https://doi.org/10.1016/j.regsciurbeco.2015.10.003.

Crown, W.D., Caldwell, J.H., Rosenstein, J.G., 2018. Standard 7 — geometry and spatial sense. N. J. Math. Curric. Framew. 209–249.

Dablanc, L., Browne, M., 2019. Introduction to special section on logistics sprawl. J. Transp. Geogr. 0–1 https://doi.org/10.1016/j.jtrangeo.2019.01.010.

Dablanc, L., Rakotonarivo, D., 2010. The impacts of logistics sprawl: how does the location of parcel transport terminals affect the energy efficiency of goods' movements in Paris and what can we do about it? Procedia Soc. Behav. Sci. 2, 6087–6096. https://doi.org/10.1016/j.sbspro.2010.04.021.

Dablanc, L., Ross, C., 2012. Atlanta: a mega logistics center in the piedmont atlantic megaregion (PAM). J. Transp. Geogr. 24, 432–442. https://doi.org/10.1016/j.jtrangeo.2012.05.001.

Dablanc, L., Ogilvie, S., Goodchild, A., 2014. Logistics sprawl. Transp. Res. Rec. 2410, 105–112. https://doi.org/10.3141/2410-12.

Dablanc, L., Palacios-Argüello, L., Oliveira, L., 2020. Logistics City Chair: Logistics Sprawl Characterization. Chair Logistics City, Paris.

De Lara, J., 2013. Warehouse Work: Path to the Middle Class or Road to Economic Insecurity. USC Program Environ. Reg. Equity.

Dong, L., Huang, Z., Zhang, J., Liu, Y., 2020. Understanding the mesoscopic scaling patterns within cities. Sci. Rep. 10, 21201. https://doi.org/10.1038/s41598-020-78135-2.

Dubie, M., Kuo, K.C., Giron-Valderrama, G., Goodchild, A., 2020. An evaluation of logistics sprawl in Chicago and Phoenix. J. Transp. Geogr. 88, 102298 https://doi.org/10.1016/j.jtrangeo.2018.08.008.

Durmuş, A., Turk, S.S., 2014. Factors influencing location selection of warehouses at the intra-urban level: Istanbul case. Eur. Plan. Stud. 22, 268–292 doi:10/gndc23.

Federal Register, 2010. Office of management and budget. Off. Manag. Budg. 75, 8.

Ge, L., Sarhani, M., Voß, S., Xie, L., 2021. Review of transit data sources: potentials, challenges and complementarity. Sustainability 13, 11450 doi:10/gnhwr8.

Gingerich, K., Maoh, H., 2019. The role of airport proximity on warehouse location and associated truck trips: evidence from Toronto, Ontario. J. Transp. Geogr. 74, 97–109 doi:10/gfz92q.

Giuliano, G., Kang, S., 2018. Spatial dynamics of the logistics industry: evidence from California. J. Transp. Geogr. 66, 248–258. https://doi.org/10.1016/j.jtrangeo.2017.11.013.

Glaeser, E.L., Kohlhase, J.E., 2004. Cities, regions and the decline of transport costs. In: Florax, R.J.G.M., Plane, D.A. (Eds.), Fifty Years of Regional Science, Advances in Spatial Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 197–228. https://doi.org/10.1007/978-3-662-07223-3_9.

Heitz, A., Dablanc, L., 2015. Logistics spatial patterns in Paris rise of Paris basin as logistics megaregion. Transp. Res. Rec. 2477, 76–84. https://doi.org/10.3141/2477-09.

Heitz, A., Dablanc, L., Tavasszy, L.A., 2017. Logistics sprawl in monocentric and polycentric metropolitan areas: the cases of Paris, France, and the Randstad, the Netherlands. Region 4, 93–107. https://doi.org/10.18335/region.v4i1.158.

Hesse, M., Rodrigue, J.-P., 2004. The transport geography of logistics and freight distribution. J. Transp. Geogr. 12, 171–184. https://doi.org/10.1016/j.jtrangeo.2003.12.004.

Janjevic, M., Winkenbach, M., 2020. Characterizing urban last-mile distribution strategies in mature and emerging e-commerce markets. Transp. Res. Part Policy Pract. 133, 164–196 doi:10/gm5gnr.

Kang, S., 2020a. Relative logistics sprawl: measuring changes in the relative distribution from warehouses to logistics businesses and the general population. J. Transp. Geogr. 83, 102636 https://doi.org/10.1016/j.jtrangeo.2020.102636.

Kang, S., 2020b. Why do warehouses decentralize more in certain metropolitan areas? J. Transp. Geogr. 88, 102330 doi:10/gnt2s3.

Klinkhardt, C., Woerle, T., Briem, L., Heilig, M., Kagerbauer, M., Vortisch, P., 2021. Using OpenStreetMap as a data source for attractiveness in travel demand models. Transp. Res. Rec. J. Transp. Res. Board. https://doi.org/10.1177/0361198121997415, 036119812199741.

Lim, H., Park, M., 2020. Modeling the spatial dimensions of warehouse rent determinants: a case study of Seoul metropolitan area, South Korea. Sustainability 12, 259. https://doi.org/10.3390/su12010259.

Lim, H., Yoo, E.-H., Park, M., 2018. Warehouse rental market segmentation using spatial profile regression. J. Transp. Geogr. 73, 64–74. https://doi.org/10.1016/j.jtrangeo.2018.10.007.

Long, Y., Liu, L., 2015. Big/Open Data in Chinese Urban Studies and Planning: A Brief Review, 13.

LoopNet, 2020. WWW Document. In: Wareh. Lease. URL. https://www.loopnet.com/warehouses-for-lease/.

Luo, S., He, S.Y., 2021. Understanding gender difference in perceptions toward transit services across space and time: a social media mining approach. Transp. Policy 111, 63–73 doi:10/gm4gsk.

Ma, Y., Zhang, Z., Ihler, A., Pan, B., 2018. Estimating warehouse rental price using machine learning techniques. Int. J. Comput. Commun. Control 13, 235–250. https://doi.org/10.15837/ijccc.2018.2.3034.

Miao, R., Wang, Y., Li, S., 2021. Analyzing urban spatial patterns and functional zones using Sina Weibo POI data: a case study of Beijing. Sustainability 13, 647. https://doi.org/10.3390/su13020647.

Niu, H., Silva, E.A., 2020. Crowdsourced data mining for urban activity: review of data sources, applications, and methods. J. Urban Plan. Dev. 146 https://doi.org/10.1061/(asce)up.1943-5444.0000566, 04020007.

Onstein, A., 2021. Factors Influencing Physical Distribution Structure Design (Master of Science in Urban and Regional Planning). Delft University of Technology, Delft, Netherlands.

Onstein, A.T.C., Tavasszy, L.A., van Damme, D.A., 2019. Factors determining distribution structure decisions in logistics: a literature review and research agenda. Transp. Rev. 39, 243–260. https://doi.org/10.1080/01441647.2018.1459929.

OpenStreetMap contributors, 2020. Planet Dump. retrieved from. https://planet.osm.org.

Pineda-Jaramillo, J., Pineda-Jaramillo, D., 2021. Analysing travel satisfaction of tourists towards a metro system from unstructured data. Res. Transp. Bus. Manag. 100746 doi:10/gnqvv9.

Quinn, P., 2013. Road density as a proxy for population density in regional-scale risk modeling. Nat. Hazards 65, 1227–1248 doi:10/f4jv28.

Rodrigue, J.-P., 2004. Freight, gateways and mega-urban regions: the logistical integration of the Bostwash corridor. Tijdschr. Voor Econ. En Soc. Geogr. 95, 147–161. https://doi.org/10.1111/j.0040-747X.2004.t01-1-00297.x.

Sakai, T., Beziat, A., Heitz, A., 2020. Location factors for logistics facilities: location choice modeling considering activity categories. J. Transp. Geogr. 85, 102710 https://doi.org/10.1016/j.jtrangeo.2020.102710.

Sakai, T., Kawamura, K., Hyodo, T., 2016. Location choice models of urban logistics facilities and the impact of zoning on their spatial distribution and efficiency. 95th Annual Meeting of the Transportation Research Board, Washington DC.

Sarkar, S., Wu, H., Levinson, D.M., 2020. Measuring polycentricity via network flows, spatial interaction and percolation. Urban Stud. 57, 2402–2422. https://doi.org/10.1177/0042098019832517.

Schorung, M., 2022. Atlas on Warehouse Geography in the US (Ebook Version). Logistics City Chair. University Gustave Eiffel, Paris.

Shen, Y., 2017. Understanding Functional Urban Centrality Spatio-Functional Interaction and its Socio-Economic Impact in Central Shanghai.

Vargas, J., Srivastava, S., Tuia, D., Falcao, A., 2021. OpenStreetMap: challenges and opportunities in machine learning and remote sensing. IEEE Geosci. Remote Sens. Mag. 9, 184–199. https://doi.org/10.1109/mgrs.2020.2994107.

Wagner, T., 2010. Regional traffic impacts of logistics-related land use. Transp. Policy 17, 224–229 doi:10/cnhp6k.

Whitehand, J.W.R., Batty, M., Longley, P., 1996. Fractal cities: a geometry of form and function. Geogr. J. 162, 113. https://doi.org/10.2307/3060277.

Woudsma, C., Jensen, J.F., Kanaroglou, P., Maoh, H., 2008. Logistics land use and the city: a spatial–temporal modeling approach. Transp. Res. E Logist. Transp. Rev. 44, 277–297 doi:10/cht5th.

Woudsma, C., Jakubicek, P., Dablanc, L., 2016. Logistics sprawl in north america: methodological issues and a case study in Toronto. Transp. Res. Procedia 12, 474–488 doi:10/gm5gns.

Wu, C., Ye, X., Ren, F., Du, Q., 2018. Check-in behaviour and spatio-temporal vibrancy: an exploratory analysis in Shenzhen, China. Cities 77, 104–116. https://doi.org/10.1016/j.cities.2018.01.017.

Zhang, L., Pfoser, D., 2019. Using OpenStreetMap point-of-interest data to model urban change—a feasibility study. PLoS One 14, e0212606. https://doi.org/10.1371/journal.pone.0212606.