RETAIL TRADE AREA ANALYSIS USING MULTIPLE VARIABLES MODELING AT RESIDENTIAL ZONE

Andrii GALKIN1*, Iryna BALANDINA2, Iryna POLCHANINOVA3, Olena GALKINA4
1Department of Transport System and Logistics, Transport System and Technologies Faculty, O. M. Beketov National University of Urban Economy in Kharkiv, 17, Marshala Baganova str., 61-002, Kharkiv, Ukraine
2, 3Department of Tourism and Hospitality, Management Faculty, O. M. Beketov National University of Urban Economy in Kharkiv, 17, Marshala Baganova str., 61-002, Khakiv, Ukraine
4Department of Water Supply, Sewerage and Water Treatment, Urban Ecology Faculty, O. M. Beketov National University of Urban Economy in Kharkiv, 17, Marshala Baganova str., 61-002, Khakiv, Ukraine
*E-mail: galkin tsl@gmail.com

Abstract. Purpose – the purpose of the paper is to set up a method of retail trade area analysis using multiple variables modeling at a residential zone.

Research methodology – system analysis; regression analysis; correlation analysis; simulating; urban characteristics analysis.

Findings – retail trade area analysis using multiple variables modeling at residential zone based on the proposed method is performed by directly processing and analysing data in a separate zone.

Research limitations – the obtained results can be used for data variation range of conducted experiment.

Practical implications – the proposed method makes some adjustments in estimating the limits of the trade area, specifying it with the help of the non-linearity factor and area slope, thereby changing the shape of the circle into a complex figure, depending on the geographical landscape and the structure of the roads. Implementation is made for one of the consumer’s zone in Kharkiv, Ukraine. The results allowed to adjust the trade zone, which in fact reduced it twice.

Originality/Value – the probabilities of visiting retailers have been calculated according to the developed model considering surrounding limitation. In such conditions, the analysis of the consumer market and identification of the trade zone of the retailers is one of the forms to improving the efficiency of the retailer functioning.

Keywords: consumers, logistics, zone, service, non-linearity factor, slope factor.

JEL Classification: M30, M31, M39, L66.

Conference topic: Contemporary Organizations Development Management.

Introduction

For today, the retail zone is often being part of the residence zone especially in urban area (Galkin, Dolia, & Davidich, 2018). Therefore, not related to measures implemented on a micro district scale, the developed models are not integrated with other components of urban mobility issues; end-consumers movement who purchases the goods in the same area they live.

To solve the problem of choosing a retailer store probability visiting the four main approaches are used: the checklist method (Colomé & Serra, 2000); analogue approach (Applebaum & Cohen, 1961); regression analysis (Lord & Lynds, 1981); models of revealed preferences (Huff, 1964). The mixed methods: fuzzy AHP, artificial neural network, etc., had been used for this issue (Kuo, Chi, & Kao, 2002). The checklist method implies an assessment by the end-consumer parameters that are determined by the consumer or given by the questionnaire (Hernandez & Biasiotto, 2001). The analogue approach is based on using the values of some characteristics of the store as initial data for demand forecasting (Applebaum & Cohen, 1961). Regression analysis was used to predict demand for various stores: food stores (Lusch, Serpkenci, & Orvis, 2015), various products and services (Birkin, G. Clarke, & M. Clarke, 2017), and grocery retailers (Oruc & Tihi, 2012). Baseline data
for the regression analysis can be obtained from standard checklists, statistics, and other surveys. Methods for identifying preferences include: Reilly (1931), Nakanishi and Cooper (1974), Louviere, Hensher, and Swait (2000), Fotheringham, Brunsdon, and Charlton (2000), Davis (2006). At the same time, such methods do not fully reflect modern market conditions. Certain parameters of existing models and methods do not provide a systematic approach to determining retailer visiting probability according to integration of end-consumer factors, retailer factors and environmental factors.

1. Analysis of recent researches and publications

There are different methods of identifying the area of service based on marketing approaches (Popa, 2015). Output parameter for which is the volume of consumption over a period. Each retailer has its own service area with individual parameters, which depends on consumer’s preferences, features of transport transit (Verroios, Efstathiou, & Delis, 2011) and geographic characteristics of the territory (Russo & Comi, 2012). Considering the influence of these parameters will allow adjusting the limits of the retailer’s service area, which will more accurately evaluating the demand.

The basis for identifying the limits of the service area in residential zones is the information about the trading range (Huff, 1964), or the trade area (Liliana, Gabriel, & Mircea, 2017), which are synonymous. In the trade area is considered a geographic sector that contains potential buyers of a member of the retail chain or store (Birkin, G. Clarke, & M. Clarke, 2017). The trade area can be divided into zones (districts). The trade area is been part of residential zones characterized by distance or time to walk to the retailer. Reilly (1931) identifies three zones, calling the primary zone of the neighbour, the secondary – the middle, tertiary – the far distance. The boundary of the near zone is determined by them no more than 2–4 km (10 minutes’ drive from the store). The average zone is located within the radius of 2–6 km (15–20 min.). The area which is located outside 6 km and can reach up to 25 km in large cities (see Table 1).

<table>
<thead>
<tr>
<th>1st level Micro district type</th>
<th>2nd level Sub district type</th>
<th>3rd level City type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction of the basic, urgent needs – in the goods and services of frequent demand and emergency demand. Demand is met by small stores, in which consumers can find an assortment, often bought, to provide themselves and household goods in the near future.</td>
<td>Satisfaction, pressing and periodically emerging needs, with greater variety and better quality of the offer. The assortment becomes wider and deeper: added products (services) of periodic demand, that is, new groups and positions in existing groups. The size of stores and shopping centers is increasing.</td>
<td>The district, urban type. Major and periodic demand for goods and services is accompanied by occasional and special demand satisfaction. In shops, consumers can find almost everything they need and do not even really need in cafes, restaurants and entertainment complexes – to get more impressions and more variety than the previous level.</td>
</tr>
</tbody>
</table>

The studies use a two-stage methodology for determining the relevant geographic retail market. In the first stage, the territory around the store is defined, within the boundaries of which consumers can reach the retailer at a certain time. In the second stage, the results of the first stage are specified taking into account local specific conditions of functioning of the retail market. But, as their analysis shows, such methods have a general and/or abstract nature, which does not allow assessing the real environment. Therefore, trade zone evaluation poor considers the geographical features of the terrain, such as (Galkin, 2017): the non-linearity factor slope coefficient, etc. The purpose of the paper is to set up method of retail trade area analysis using multiple variables modeling at the residential zone, which involved environmental factors, retailer factors and consumer’s factors. This approach will provide more flexible approach for evaluating the probability of end-consumer’s visiting of the retailer, defining sells volumes, etc.

2. Methodology

2.1. Attractiveness of selecting stores at residential zone

To determine the boundaries of the logistics service store’s area, an algorithm is proposed, that consists of four steps, depicted in Figure 1. According to offered algorithm procedure of experiment is conducted as follows:

- At the first stage, we identify the starting point, in order to construct for it a trade area of execution of the experiment.
In the second stage, we define the trading zone. According to the methodology (Halkin, et al., 2019), we identify the boundaries of the trading zone by plotting the radius. The resulting area is the primary trade area.

Adjustment of the trading zone is based on urban parameters. Researchers in their works agree that the radius is the first priority when calculating a particular trading zone (Halkin, 2018). Simultaneously, the influence of city parameters, such as the density of the street-road network (which affects the direction of movement) (Lobanov, 1990), natural deviations (Paroli & Maraschin, 2018), water and other natural obstacles and the railroad were considering (Crainic, Ricciardi, & Storchi, 2009) with it designing.

Figure 1. The method of the boundaries of the store’s service area setting

To describe the scheme of the separate territories, the non-linearity factor is used which evaluates the distance by air and roads on the equation (Galkin, 2017):  
\[
R^{ijz} = \frac{\sum_{j=1}^{J} l_{\text{road}}^{wijz}}{\sum_{j=1}^{J} l_{\text{air}}^{wijz}},
\]

where \( j \) – number of retailers; \( w \) – number of residential zone of consumers; \( z \) – number of trade areas in the city; \( l_{\text{road}}^{wijz} \) – the distance between points “roads”, km; \( l_{\text{air}}^{wijz} \) – the distance between points “by air”, km; \( n \) – number of measurements.

To take into consideration the natural slopes, you must take advantage of the topological map of the area. Changes in altitudes above sea level and natural slopes are measured using the c non-linearity factor and slope factor \( (k_{y}^{wijz}) \) (Paroli & Maraschin, 2018) by the equation:  
\[
k_{y}^{wijz} = 1 + \tan(A) = 1 - a / b,
\]

where \( a \) – catch (height difference from the starting point to the store), km; \( b \) – catch (horizontal distance from the starting point to the store), km.

Mark retailers on the map. Mark retailers on the corrected map of the corresponding trading profile: grocery, household goods, construction, electrical engineering and others.

Adjustment of the service area considered the inhabitants’ poll.

3. Results

3.1. Trade area identification

For instance, and calculations, the shopping area of the first level (residential type) was chosen. The study was conducted for a commercial facility in Kharkiv (Ukraine), but the method can be used for any other trading zone.

1. Mark the beginning of the experiment zone. Retail trade study in the consumer areas, allows to conduct a comparative analysis of their development in the territory of the settlement. Trading zone analysis determines the frequency of visits to retailers in this zone; the time of movement to the store and other parameters affecting the demand, the probability of a visit.
2. Mark the radius according to the selected area of service “by air”, we conduct a circle with a given radius – a trading zone (1.5 km). Use Google or other possible electronic charts for this (Figure 2).

3. Correction area is the service, taking into consideration the non-linearity factor and the slope factor. For non-linearity calculation the Google maps tool is used. The calculations are in the following sequence. We will route the routes from the starting point (home) to the end of the studied boundary (radius) “by road”, in 8 directions: north, north-east, east, south-east, south, south-west, west, northwest. Increasing the number of directions will show more accurate results. This will allow adjusting the boundaries of the pedestrian accessibility area of retail outlets to home users. Indirectness of the message considered natural obstacles (Figure 3). The presented distances are represented by criterion – the minimum distance.

Changes in altitudes and natural slopes are measured by the non-linearity factor \( R^{wij} \). To evaluate the topological map in this area. Increase of the value of factor slope results in increase of the time and the distance of movement to the retailers, thereby reducing the radius of the trading zone. In addition, to go “uphill” is more difficult and energy costly is higher than to go in a straight line. In doing so, it is necessary to consider the slope factor in the area \( k^{wij} \) on two sides of the consumer movement. Adjusted trading zone, taking into consideration the slopes of the area, is presented in Figure 4.

4. Mark the Retailers on the map. In the course of the research, all types of retailers selling consumer products were considered: supermarkets, kiosks, shops, etc. Retailers, which are included in the trading zone, taking into consideration corrections, entered in the research area (Figure 4).

Presented trading zone analysis (Figure 5) indicates a large cluster of retailers in it – 15 units. 21 retailers are not included in the trading area, “waived” after correction. The presented trading zone is the basis for gathering socio-economic information about the service area (Ibeas, Moura, Nuzzolo, & Comi, 2012). According to
Galkin, A.; Balandina, I.; Polchaninova, I.; Galkina, O. 2019. Retail trade area analysis using multiple variables modeling at residential zone

this, it is possible to determine: the area; quantity of inhabitants; consumer income and behaviour; distances from buildings and stations of urban passenger transport to retailer (places of gravitation) (Verroios, Efstathiou, & Delis, 2011); calculate the time of movement to each retailer.

3.2. Attraction of stores for consumers in the residential zone

The research was in visiting stores via consumers in the service area and monitoring the selected indicators. The experiment was performed using the Xiaomi Mi Band electronic clock, which calculated the quantity of meters passed, and the time for various elements of the purchasing process. The set of products was determined according to the consumer basket of the Law of Ukraine, as a set of food products, formed by the central executive body. Recent statistics show that in Ukraine more than 50% use this set of goods. In general, 24 people (women and men) aged 19–65 participated in the study. Parameters of the research area were selected and calculated using the method shown in the papers (Galkin, Dolia, & Davidich, 2018). The received data are presented in the Table 2.

Table 2. Output data for obtaining time changing models

<table>
<thead>
<tr>
<th>Num</th>
<th>Time movement, min. to store</th>
<th>Slope factor to store</th>
<th>Non-linearity factor by air</th>
<th>Distance, km by air</th>
<th>Time of trade service, km</th>
<th>Size of the store, m²</th>
<th>Total visiting time, m.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>11</td>
<td>0.990</td>
<td>1.010</td>
<td>1.23</td>
<td>0.57</td>
<td>0.7</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>96</td>
<td>6</td>
<td>4</td>
<td>1.100</td>
<td>0.900</td>
<td>1.00</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 3. Results of model evaluation

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value for energy cost model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Criterion: tabular</td>
<td>1.97</td>
</tr>
<tr>
<td>The ratio of the selected price for the consumer basket at the shop to the average price for a consumer basket in the area</td>
<td>20,3767</td>
</tr>
<tr>
<td>The ratio of the zone-retailer distance to average zone-retailer distance in the area</td>
<td>−9,6174</td>
</tr>
<tr>
<td>The ratio of the selected zone inhabitants quantity to average quantity of inhabitants in zones of the area</td>
<td>35,8947</td>
</tr>
<tr>
<td>The ratio of the selected non-linearity factor zone-retailer direction to average non-linearity factor in the area</td>
<td>−7,48068</td>
</tr>
<tr>
<td>The ratio of the selected slope factor zone-retailer direction to average slope factor in the area</td>
<td>−4,30166</td>
</tr>
<tr>
<td>The ratio of the current retailer size to average store size in the area</td>
<td>−2,32263</td>
</tr>
<tr>
<td>Fischer Criterion: tabular</td>
<td>3,88</td>
</tr>
<tr>
<td>calculated</td>
<td>1803,09</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>0,99</td>
</tr>
<tr>
<td>Determination coefficient</td>
<td>0,98</td>
</tr>
</tbody>
</table>

The processing of the research results was performed using regression and correlation analysis methods. The probability model of the choice of a store by end-consumer from \( w \) residential zone in the trade area has the made:

\[
P_{jwz} = 0.00785 \cdot \exp(1 - \frac{\Theta^j_{wz}}{\hat{\Theta}^j_{wz}}) - 0.0008 \exp\left(\frac{L^w_{wz}}{L^z}\right) + 0.00514 \left(\frac{N^j_{wz}}{N^w_{wz}}\right)^2 - 0.0892 \cdot \log\left(\frac{k^w_{wz}}{k^y_{wz}}\right) - 0.0044 \cdot \log\left(\frac{R^w_{wz}}{R^y_{wz}}\right) - 0.00055 \cdot \log\left(\frac{S^{shop}_{wz}}{S^{shop}_{wz}}\right),
\]

where: \( j \) – number of retailers; \( w \) – number of residential zone of consumers; \( z \) – number of trade areas in the city; \( \Theta^j_{wz} \) – the price for the consumer basket at the retailer, UAH; \( \hat{\Theta}^j_{wz} \) – average price for a consumer basket in the area, UAH; \( L^w_{wz} \) – distance from the household in zone to the retailer, km; \( k^w_{wz} \) – average distance from...
household in zone to retailer in the area, km; \( N_{iz} \) – quantity of inhabitants in the zone, consumers; \( \overline{N}_{iz} \) – average quantity of inhabitants in zones of the area, consumers; \( k_{ij}\gamma z \) – slope factor from zone to retailer; \( \overline{k}_{ij}\gamma z \) – average slope factor from zone to retailer in the area; \( R_{ij}\gamma z \) – none-linearity factor from zone to retailer; \( \overline{R}_{ij}\gamma z \) – average none-linearity factor from zone to retailer in the area; \( \text{shop}_{jz} \) – size of the retailer, m\(^2\); \( \overline{	ext{shop}}_{jz} \) – average size of the retailer in this area.

Based on the received calculations, using the software “Stat Graphics” the parameters of the model were analysed, (see Table 3).

The model adequately reflects the fact that with increasing retailer area, energy costs are also increasing during a purchase.

4. Discussion

Retail trade has a major role in the distribution system of most of the goods. The growth of large retail chains, the unification of wholesale and retail trade links, consumer orientation, are objective processes up-to-day (Krikke, Bloemhof-Ruwaard, & Van Wassenhove, 2003). This leads to the logistics system to adapt for changing demand conditions and effectively distribute the progress of material flows. The endless development of the retail market in Ukraine and in the world set up the issue of sustainable improve the methods and models of the operation of logistics systems considering the parameters of the demand of end consumers.

For the first time, the model of the probability of visiting retailers was presented, which the total expenses of consumers were used. Depending represented the price and non-price factors of the purchasing process, which allows more accurately identifying the behaviour of consumers and, as a consequence, demanding in the logistics system. The proposed model allows to move away from the traditional approach in determining the probability and allows to consider end-consumers maximally (revenue, cost structure, etc.) and, as a result, determine the working conditions of the logistics system. There have been constructed graphs of the influence of changes in the parameters of the service area on the probability of visiting a retailer (Figures 6–11).

Modeling allows to make the following conclusions:

1) Analysing the received timetable for changing the probability of visit a retailer, depending on the increase in the cost of purchasing goods (Figure 6), it can be concluded that with the increase in the cost of purchasing goods, the probability of visiting the store will decrease as the total cost of the buyer will increase. In this case, the important role is played by the price of each product from the buyer’s consumer basket. The higher the price, the greater will be the cost of buying goods, and, consequently, will reduce the overall attractiveness of the retailer and the probability of its visit.

2) Analysing the received timetable for changing the probability of a retailer visit, depending on the increase in the distance to the store (Figure 7), one can conclude that the larger the distance from the household to the store, the less probability of its visit, since in this case the rates of such factors as the value of the buyer’s time and hence the total costs of the buyer at the retailer.
3) With the increase in the quantity of inhabitants in the district, the probability of visiting the retailer will increase (Figure 8), since this indicator affects the quantity of potential end-consumers of certain goods of a certain type, which positively affects the probability of its visit.

4) The higher the coefficient of non-directional traffic, the greater the probability of a retailer visit (Figure 9) is due to the fact that most buyers are moving on their own cars and for them this factor is not essential, since it is likely to be the least uncomfortable in terms of the complexity of the way shop can provide comfortable conditions for customer service.

5) The greater the slope, the greater the probability of visiting the retailer (Figure 10). This is explained by the fact that it is easier for people to move along the slopes, and this affects the reduction of overall fatigue and the reduction of the total cost of the buyer.

6) With the increase in the area of the retailer, the probability of visiting the shop will decrease (Figure 11). This is due to the fact that when shopping in a large store, people spend more time and energy on finding and purchasing goods, which increases the cost of the buyer’s time when visiting the shop.

5. Limitations

Obtained models can be used for finding the market probability for set up variation range, presented in Table 4. Presented is variation range values are cover data all-over Ukraine and most east Europe countries. Obtained data allows to define limitations of the trade zone, volumes of retailers visiting and been origin demand: average check, retail network sales volume and other parameters.
The customary approach considers the trade area as concentric circles. The proposed method makes some adjustments in estimating the limits of the trade area, specifying it with the help of the non-linearity factor and area slope, thereby changing the shape of the circle into a complex figure, depending on the geographical landscape and the structure of the roads. The results allowed adjusting the trade area, which in fact reduced it twice.

### Table 4. Range of data variation for the probability of store function estimation

<table>
<thead>
<tr>
<th>Num</th>
<th>Indicator</th>
<th>Units of measurement</th>
<th>Minimum value</th>
<th>Maximum value</th>
<th>Average value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The ratio of the selected slope factor zone-store direction to the average slope factor in the area</td>
<td>–</td>
<td>0.985</td>
<td>1.015</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>The ratio of the selected none-linearity factor zone-store direction to average none-linearity factor in the area</td>
<td>–</td>
<td>0.7</td>
<td>1.90</td>
<td>1.30</td>
</tr>
<tr>
<td>3</td>
<td>The ratio of the zone-store distance to average zone-store distance in the area</td>
<td>km</td>
<td>0.3</td>
<td>2.1</td>
<td>1.2</td>
</tr>
<tr>
<td>4</td>
<td>The ratio of the current store size to average store size in the area</td>
<td>m²</td>
<td>0.2</td>
<td>2.9</td>
<td>1.55</td>
</tr>
<tr>
<td>5</td>
<td>The ratio of the selected price for the consumer basket in the shop to average price for a consumer basket in the area</td>
<td>UAH</td>
<td>0.85</td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>The ratio of the selected zone inhabitants quantity to average quantity of inhabitants in zones of the area</td>
<td>UAH</td>
<td>0.15</td>
<td>1.95</td>
<td>1.00</td>
</tr>
</tbody>
</table>

An important topic for research is the combined consumer’s trips (Russo & Comi, 2006) and impulse purchases that are associated with them. Their essence lies in the fact, that a person can visit several stores on the same trip either on the way to or from work. Impulse purchases are caused by random factors. Consumer has an opportunity to buy a cup of coffee or tea on the way to the main destination point. The influence of the most significant parameters on the probability of choosing a trading point is reflected in the construction charts. The obtained results allow identifying the primary service area of customers inside of current trade area. This information can be initial data for further research in consumer’s marketing issue.

### Conclusions

In the market, competition is growing more and more, therefore, enterprises pay great attention to studying their clients, minimizing their losses, creating the most favourable conditions for purchasing and determining the attractiveness of retailer. Based on the received data and using the dependence (3), it is possible to determine the demand for each individual store.

The streams associated with purchases and the external costs of a society that arise from this are an integral process of integrating all processes. Increasing the level of motorization, lack of parking spaces, poor development of transport infrastructure leads to increased time and energy spent on purchases, reducing traffic jams and reducing the range of store services, market volume, etc. (Russo & Comi, 2012). In addition, taxes that is an integral part of sales change the attractiveness of domestic or international goods. Changing the size of taxes and fees leads to a redistribution of demand and the likelihood of a consumer’s choice of the retailer (Galkin, Dolia, & Davidich, 2018).

Therefore, using this approach, you can dynamically figure out how stocks, inflation, price changes, end-consumer income, a set of consumable products, and other phenomena will affect demand redistribution and, accordingly, choose a logistics supply and storage strategy for each logistics system.

Consumers in the examined area have different incomes, behaviour, etc., it leads to promotion by e-mixed and omni-channel ways in each case (Figure 2). The consumer’s influences the redistribution of material flows have high influence. The research confirms the modern theory of consumer preferences, which consists in the ability of the buyer to manage their own costs, and the ability of stores to manage the costs of their customers. The buyer, as a reasonable individual, decides to choose a store based on his own attributes, which depends on his behaviour, store’s factors, environmental factors.

Retail trade area analysis using multiple variables modeling at residential zone based on the proposed method is performed by directly processing and analysing data in a separate zone. In addition to primary data, when the market share is being studied for enlarged assortment positions, the secondary data obtained from Ukrainian and international statistical compilations, surveys or interviews can be used. Obviously, using different methods for determining market share will most likely give different results. In this case, it is necessary either to give preference to the most reliable method or to use the average value of the market share indicator for estimates obtained by different methods.
References


