Scientific Journal of Silesian University of Technology. Series Transport
Zeszyty Naukowe Politechniki Śląskiej. Seria Transport

Volume 95 **2017**

p-ISSN: 0209-3324

e-ISSN: 2450-1549

DOI: https://doi.org/10.20858/sjsutst.2017.95.22

Journal homepage: http://sjsutst.polsl.pl

Article citation information:
Żochowska, R., Karoń, G., Sobota, A., Janecki, R. Selected aspects of the methodology of a household interview survey on an urban agglomeration scale with regard to its services. *Scientific Journal of Silesian University of Technology. Series Transport*. 2017, **95**, 239-249. ISSN: 0209-3324. DOI: https://doi.org/10.20858/sjsutst.2017.95.22.

Renata ŻOCHOWSKA[1], Grzegorz KAROŃ[2], Aleksander SOBOTA[3], Ryszard JANECKI[4]

SELECTED ASPECTS OF THE METHODOLOGY OF A HOUSEHOLD INTERVIEW SURVEY ON AN URBAN AGGLOMERATION SCALE WITH REGARD TO ITS SERVICES

Summary. The article presents the essential issues and algorithm of the methodology of a four-step transportation model, which was constructed in order to carrying out a household interview survey. The results of this research are source data for determining the travel behaviour of the users of transportation systems, including intelligent transport systems (ITS). The presented issues regarding the survey methodology also concern the specifics of the study area, an urban agglomeration area. The examples particularly relate to an urban agglomeration with the nature of a conurbation, namely, the Upper Silesian Agglomeration in Poland.

Keywords: four-step travel demand model; household interview survey; travel behaviour modelling; urban agglomeration; intelligent transport system (ITS).

[1] Faculty of Transport, The Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland. E-mail: renata.zochowska@polsl.pl
[2] Faculty of Transport, The Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland. E-mail: grzegorz.karon@polsl.pl
[3] Faculty of Transport, The Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland. Email: aleksander.sobota@polsl.pl
[4] Faculty of Economics, University of Economics, 1 May 50 Street, 40-287 Katowice, Poland. E-mail: ryszard.janecki@ue.katowice.pl

1. INTRODUCTION

This article presents the issues and algorithm for the collection of data during a household interview survey, as part of tasks in the algorithm of activities during comprehensive traffic and transportation studies (CTTS) on transportation modelling, along with an evaluation of an ITS.

A review of the state of the art revealed that discussion relating to household interview surveys in urban areas, as described in foreign literature, do not take into account, to a large extent, the determinants of the functioning of areas such as polycentric agglomerations (conurbations). In the national literature, there is a lack of comprehensive research; and, while there are many publications that deal with these issues, they do so in a fragmentary or superficial manner. Thus, the developed methodology, presented in this paper, comprising an algorithm for household surveys with regard to transportation modelling, along with an evaluation of an ITS project (which uses a systematic approach and a formalization of the key research questions) is an important scientific contribution.

Currently, the use of ITS services, in addition to influencing the distribution of traffic flows in the transportation network (by traffic control subsystems and VMS signs, for example) may also affect user decisions about certain types of travel. Users' decisions concerning their travel arrangements may be considered from at least three aspects:

- Executive aspect of the trip: realization vs. retreat
- Temporal aspect of the trip: during peak hours vs. during off-peak hours
- Spatial aspect of the trip: via the most popular paths (the shortest ones, but also the most charged by traffic flows) vs. via the least popular paths (the longest ones, but the less charged by traffic flows).

The construction of a transportation model requires gathering a wide variety of data, derived from different sources, i.e., surveys and traffic measurements [2], [3], [5], [9], [13]. The scope and level of detail of the data collected in a survey generate at least two outcomes. The first concerns the structure of the four-step transportation model and the basic scope of the necessary data. The second concerns the scope of the project (investment project, study, transport development strategy etc.), especially the assumptions, components and results of the project in terms of the impact on the transportation system and its users. An algorithm for survey sampling, included in the algorithm of activities during CTTS on transportation modelling, along with an evaluation of ITS project, is presented in Fig. 1.

The implementation of ITS services in the specified functional and utility configuration should contribute to reducing congestion in the urban transportation network. It is therefore necessary to gain from the surveys data that will allow for mapping the impact of ITS services in a specified configuration on users' decisions about how they make a journey in terms of the following:

- Specified time
- Specified location
- Specified mode of travel
- Specified paths in the network

The presented transportation model is a tool that supports analyses of problems associated with the functioning and development of the transportation system for a specific urban area. This article addresses the fact that the area is an urban agglomeration (concentration of cities) and with the nature of the conurbation being studied, namely, the Upper Silesian

Agglomeration in Poland, and its capital city of Katowice. The presence of an urban agglomeration is specific to the study area. Therefore, the methodology of a travel behaviour survey on such a scale requires the appropriate conditions and limitations to be determined, as well as the assumption of the representativeness of the sample in terms of selected characteristics [2], [3], [4], [5], [9], [13], [15], [16], [17], [18].

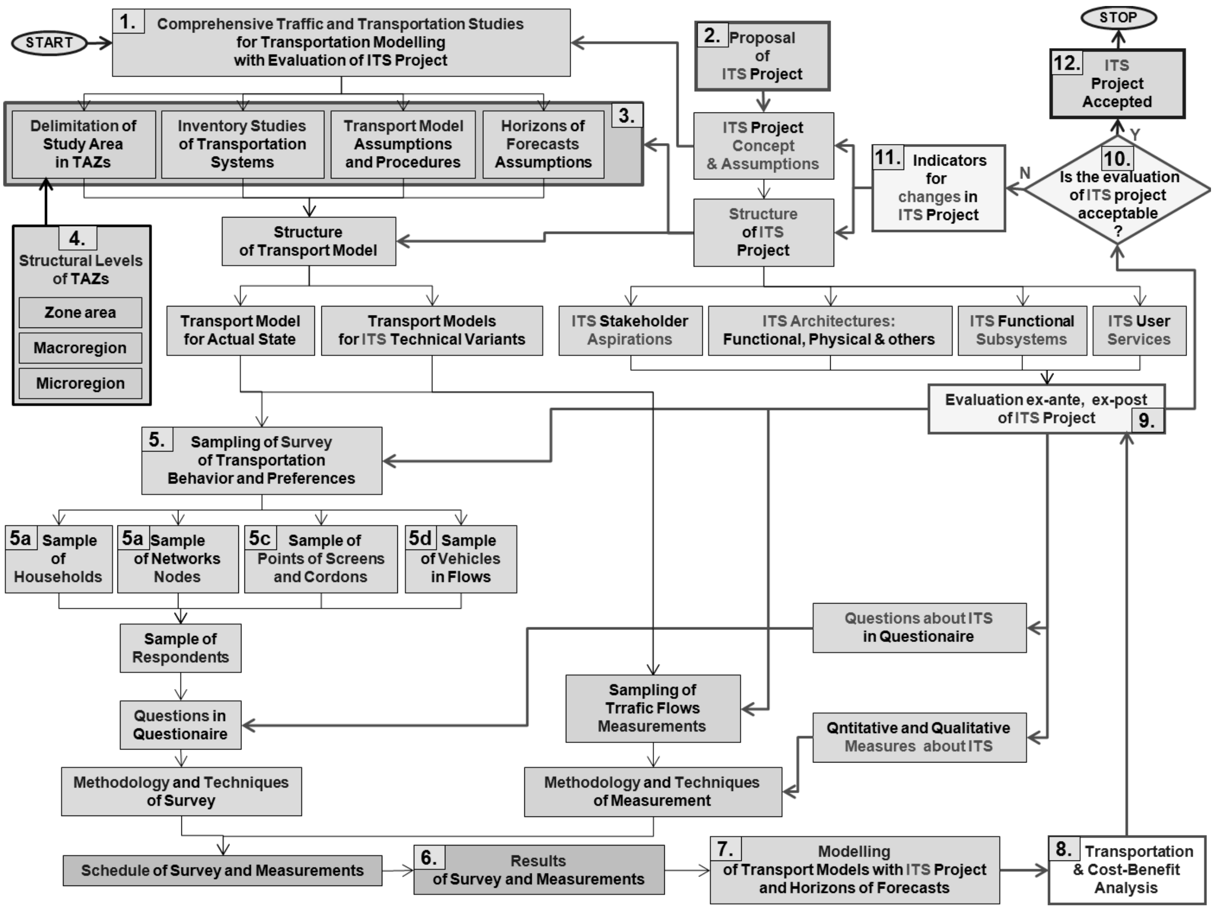

Fig. 1. Algorithm for survey sampling (blocks 5, 5a-5d), which is included in the algorithm of activities during CTTS on transportation modelling, along with an evaluation of an ITS Project (in the form of block scheme)

2. CONDITIONS AND METHODOLOGICAL ASSUMPTIONS OF HOUSEHOLD INTERVIEW SURVEY

When analysing the area on an urban agglomeration scale, it is necessary to determine its delimitation [2], [3], [4], [13], [17]. In the proposed methodology, it is assumed that the travel behaviour surveys are conducted in two zones (see Fig. 2):

– Zone 0, which is the internal area, covering municipalities belonging to the urban agglomeration
– Zone 1, which is the outer area, covering municipalities that directly impact on the study area

Fig. 2. Delimitation of the urban agglomeration area (Upper Silesian Agglomeration in Poland) for Zone 0 (a major area of study) and Zone 1 (direct surroundings).

2.1. The scope of the study

Surveys conducted in households are complex, providing basic sources of information on many characteristics that describe the travel behaviour of residents in the study area. In particular, information and characteristics should at least include:

- For examined households:
  - The number of traffic analysis zones (TAZs)
  - The number of persons in a household
  - The number of cars in a household
- For examined persons:
  - Anthropogenic features and characteristics of socio-economic activities (e.g., sex, age, occupation, income level)
  - A detailed description of the trip (e.g., origin and destination, purpose, start and end times, method of travel with the use of different transportation subsystems, duration, frequency)
  - Assessment of transport services, travel behaviour and preferences (e.g., concerning public or individual transport, application of selected ITS services both for public and for individual transport, accessibility of infrastructure for cycling in terms of the construction of Bike&Ride parking lots, accessibility to parking infrastructure in terms of the construction of Park&Ride parking lots)
  - Expectations in terms of improving travel experience by means of ITS services for public and individual transport

2.2. The basic assumptions

Any methodology of surveys on households requires:

- Defining the general population and random sample
- Determining the minimum sample size and formulating assumptions related to the sampling frame
- Developing the rules for assessing the accuracy of the results
- Developing the principles for the selection and sampling scheme

The proposed methodology assumes that the study population will comprise residents of the municipalities within the urban agglomeration (Zone 0), or zone 1, who are aged six years or more, regardless of their formal place of residence. Such a term also takes into account persons who do not have an actual address in the study area, but are temporarily staying within its boundaries and generating traffic. The age criterion has been formulated in such a way that, when analysing travel behaviour, one should take into consideration any movement associated with the need for all children of compulsory school age to get to a primary school. In the proposed methodology, the study population was divided into smaller subsets forming strata. Each stratum was then sampled as an independent subpopulation, comprising:

- For zone 0: residents of individual TAZs separated within the municipalities that belong to the urban agglomeration
- For zone 1: inhabitants of an individual municipality belonging to the outer area

The unit of observation is a household with one or more members. Therefore, the place where the household is located (and more specifically its address) has been assumed as a basic sampling unit. Thus, the sampling frame should be an address database containing the name of the municipality, the name of street and the number of the building. The frame should unambiguously map the general population and contain information relating to its subjective, territorial and temporal scope. The sampling frame has been divided into strata, which correspond to the TAZs in Zone 0 and individual municipalities in Zone 1. Such a structure ensures the most reliable and timely data on the socio-economic, economic, demographic and spatial characteristics. The household size is determined by the number of its members. The unit of study is a person who is a resident (aged six years or older) of the study area (municipalities belonging to Zone 0 or Zone 1), while a single household is treated as a group of units of the study (cluster) from the point of view of the travel behaviour of the persons belonging to the household. The interviewer should, therefore, strive to conduct surveys with as many members of a single household as possible.

2.3. The sampling scheme

The sampling scheme is a concrete and coherent (in terms of being formal and substantive) procedure for drawing out sampling items from the general population, such that the condition of randomness is satisfied. The choice of optimal sampling scheme depends on the conditions and purpose of the research. To a significant extent, it determines the way in which specific parameters are estimated, as well as how certain parameters should be estimated. Household interview surveys require specific rules for sampling. The drawing out of sampling items requires a cluster sampling scheme, where each cluster is a single household. The assumptions concerning this drawing out process are as follows:

- The sampling frame must be an address database of municipalities in Zone 0 and Zone 1, which contains the name of the municipality, the name of the street and the number of the building. It is usually at the disposal of the state administration.
- The data contained in the sampling frame should be current (in the proposed methodology, it was established up to three years ago) and complete. In addition, the data should be arranged and contain information about assigning an address to the appropriate TAZ.
- The sampling frame should be divided into individual strata, corresponding to individual TAZs in Zone 0 and Zone 1.
- The sample size (number of respondents) in the individual municipalities in Zone 0 and Zone 1 is fixed on the basis of proportional allocation, according to the relation:

$$w_g = \frac{Lm_g}{Lm}, \qquad g \in \boldsymbol{G} \tag{1}$$

where:

w_g - the number of respondents in g-th municipality

G- the set of all municipalities in the study area (Zone 0 or 1)

Lm_g - the number of residents in g-th municipality

Lm - the total number of inhabitants in the study area (Zone 0 or 1)

- The sampling scheme in the programming phase of the survey (Stage I) includes:
  - Determining the initial number of addresses in relation to each stratum, which should take into account the surplus for the randomly selected addresses (e.g., 20%) concerning the situations where there is an absence of response or a refusal to carry out the interview.
  - Drawing out sampling items of a predetermined number from the sampling frame, which is an address database.
  - Grouping randomly selected addresses within each stratum. The essential criterion for each grouping is the location of addresses. It is worth combining the addresses located in close proximity into one set in order to reduce costs and improve the survey's outcomes.
  - Determining the number of surveys and identifying specific addresses relevant to the survey for each interviewer. In the proposed methodology, the number of surveys in a single TAZ of Zone 0 has been established on the basis of proportional allocation, according to the relation:

$$w_{rk(g)} = \frac{Lm_{rk(g)}}{Lm_g}, \qquad rk(g) \in \boldsymbol{RK}(g), \quad g \in \boldsymbol{G} \tag{2}$$

where:

$w_{rk(g)}$- the number of surveys in $rk(g)$-th TAZ in g-th municipality

RK(g) - the set of TAZs in g-th municipality

$Lm_{rk(g)}$- the number of residents in $rk(g)$-th TAZ in g-th municipality,

- The assumption about the proportional allocation of the number of surveys in the TAZs and the number of respondents in the municipalities of Zone 0 and Zone 1 satisfies the requirement concerning the equal probability of selection for all items constituting the general population.

These assumptions were used in the methodology for constructing the transportation model when preparing the tender documentation (including the specification of the essential terms of the contract as well as the description of the subject of the order) for two areas: the Upper Silesian Agglomeration (in Poland) and the territory of the central subregion of the Silesian Voivodeship.

2.4. Representativeness of the surveys

It is assumed that household interview surveys are conducted using the representative method, based on a random sample, which offers the possibility to generalize, within a specified margin of error, the survey results for all units of the general population. The concept of representativeness of the sample refers to the specific features of the study unit, which means that the structure of the random sample must correspond to the structure of the general population from the point of view of these features. In other words, a representative sample is a microcosm of the population due to the selected features, while it is assumed that the representativeness of the sample essentially depends, to a certain extent, on two factors:

- The choice of the respectively numerous sample
- The method of sampling and fulfilling the conditions of homogeneity, randomness and independence

A sample that is representative of one particular feature may not be representative in terms of other features. Accordingly, the problem regarding the proper selection of characteristics, for which representativeness is expected, seems to be very important. It is worth remembering that providing a representative sample for multiple features is simultaneously difficult to implement, while the complexity of the problem increases with the number of variables. Therefore, when designing the surveys, which are then used to analyse travel behaviour analyses, the minimum number of variables, which are closely related to the study variables, is usually selected in order to assess the representativeness of the sample. In this case, the choice of variables, on the presumption of their representativeness, is burdened by a certain degree of subjectivism.

In the case of drawing out large samples (at least tens of elements), one should expect that the probability of selecting samples of the structure, which are significantly different from the structure of the population, is relatively small. This obviously does not exclude the occurrence of extreme situations, in which surveys are only subjected to certain groups of respondents (e.g., due to the absence of working people, interviews only conducted with non-working elderly people), which may significantly distort the estimation of the real value of the observed feature. Therefore, to increase the accuracy of surveys, it is assumed that structure of the sample may differ by no more than 25% from the population structure in terms of sex and age with regard to inhabitants in each municipality of Zone 0 and Zone 1.

Putting together a sample of predetermined structure requires the efficient organization, high precision and constant control of the results of the survey at the various stages of its implementation. When conducting a survey, the current verification of the structure of the sample with the above assumptions and the potential correction are advisable. In the proposed methodology, it is assumed that the control of the sample structure, at the stage of carrying out the survey, is conducted at least once a week. This approach also allows for the verification of intensity in respect of the inflow of surveys and the corresponding adjustment regarding the choice of the sample items in order to obtain the desired data structure for the construction of a transportation model.

The result of the draw is a random sample, which is the basis of statistical inferences of unknown parameter values of the general population from which the sample has been selected. To assess the values of the parameter estimates, their estimators are used. They are the relevant functions (statistics) related to the sample, which correspond to the sequence of value of a random variable with a specified distribution, which is called the distribution from the sample. It is assumed that the applied estimators must at least be:

- Compatible, which means that sequences of assessments, which are obtained with their help, are stochastically convergent with the real value of the estimated parameters, that is, they are subjected to the law of large numbers. In practice, the effect of the compatibility of the estimator is obtained in a situation where, with increasing sample size, the variance of the estimator heads towards zero
- Unencumbered, which means that the expected values (mathematical expectations) of estimators correspond exactly to the real values of the estimated parameters.

To assess the precision of estimates in household interview surveys, four parameters of the general population (the characteristics that are the most relevant to the aim of the study) are chosen:

- The number of trips made for a certain purpose:

$$Lp(mp)_{rk(g)} = \sum_{i_{rk(g)} \in \boldsymbol{I}_{rk(g)}} Lp(mp)_{i_{rk(g)}}, \qquad rk(g) \in \boldsymbol{RK}(g), \quad g \in \boldsymbol{G} \tag{3}$$

where:

$Lp(mp)_{rk(g)}$ - the number of trips made for mp-th purpose in $rk(g)$-th TAZ in g-th municipality

$\boldsymbol{I}_{rk(g)}$ - the set of residents of $rk(g)$-th TAZ in g-th municipality

$Lp(mp)_{i_{rk(g)}}$ - the number of trips made for mp-th purpose by $i_{rk(g)}$-th resident of $rk(g)$-th TAZ in g-th municipality,

- The share of trips made in a certain mode:

$$Up(sp)_{rk(g)} = \frac{\sum_{i_{rk(g)} \in \boldsymbol{I}_{rk(g)}} Lp(sp)_{i_{rk(g)}}}{\sum_{i_{rk(g)} \in \boldsymbol{I}_{rk(g)}} Lp_{i_{rk(g)}}}, \qquad rk(g) \in \boldsymbol{RK}(g), \quad g \in \boldsymbol{G} \tag{4}$$

where:

$Up(sp)_{rk(g)}$ - the share of trips made in sp-th mode in $rk(g)$-th TAZ in g-th municipality

$Lp(sp)_{i_{rk(g)}}$ - the number of trips made in sp-th mode by $i_{rk(g)}$-th resident of $rk(g)$-th TAZ in g-th municipality

$Lp_{i_{rk(g)}}$ - the number of trips made by $i_{rk(g)}$-th resident of $rk(g)$-th TAZ in g-th municipality

– The average duration of the trip:

$$\overline{Tp}_{rk(g)} = \frac{\sum\limits_{i_{rk(g)} \in \boldsymbol{I}_{rk(g)}} \sum\limits_{j_{i_{rk(g)}} \in \boldsymbol{J}_{i_{rk(g)}}} Tp_{j_{i_{rk(g)}}}}{\sum\limits_{i_{rk(g)} \in \boldsymbol{I}_{rk(g)}} Lp_{i_{rk(g)}}}, \qquad rk(g) \in \boldsymbol{RK}(g), \quad g \in \boldsymbol{G} \tag{5}$$

where:
$\overline{Tp}_{rk(g)}$ - the average duration of trips in $rk(g)$-th TAZ in g-th municipality
$\boldsymbol{J}_{i_{rk(g)}}$ - the set of trips made by $i_{rk(g)}$-th resident of $rk(g)$-th TAZ in g-th municipality
$Tp_{j_{i_{rk(g)}}}$ - the duration of $j_{i_{rk(g)}}$-th trip made by $i_{rk(g)}$-th resident of $rk(g)$-th TAZ in g-th municipality

– The average mobility rate of respondents:

$$\overline{Lp}_g = \frac{\sum\limits_{i_g \in \boldsymbol{I}_g} Lp_{i_g}}{Lm_g}, \qquad g \in \boldsymbol{G} \tag{6}$$

where:
$\overline{Lp}_g$ - the average number of trips in g-th municipality
$\boldsymbol{I}_g$ - the set of residents in g-th municipality
Lp_{i_g} - the number of trips made by i_g-th resident of g-th municipality

Among the selected characteristics, only mobility rate, which is understood as the number of trips made during the day, is determined at the municipal level. Other characteristics should be analysed in the individual strata (TAZs).

3. CONCLUSIONS

The results of travel behaviour analyses are necessary when constructing a four-step transportation model, which reflects the existing state of the transportation system [12], [14], [19], and assessing the directions and trends from the perspective of changes in the system. A transportation model is also a tool for analysing the problems associated with the functioning and development of the transportation system. This feature enables a broad spectrum of activities under the conditions of the implementation of ITS services in urban areas.

Developing the methodology for a household interview survey on an urban agglomeration scale with regard to ITS services is related to the following decision problems:
– Defining the population and the sample
– Selecting the characteristics of the sample unit, which are important in terms of the representativeness of the survey
– Defining the units of observation and the units of study (basic and group units)
– Determining the strata and the sampling frame
– Selecting the sampling scheme
– Determining the process for controlling the implementation of the survey in terms of the representativeness of the data collected

Moreover, these decision problems are conditioned by the methodology of delimitation for the study area, by the methodology of description for the transportation systems, and by the methodology of identification for traffic flows in transportation systems.

Verification of the assumptions about the described methodology was made during the preparation of the tender documentation (including the specification of the essential terms of the contract, as well as the description of the subject of the order) for two areas: the Upper Silesian Agglomeration (in Poland) and the territory of the central subregion of the Silesian Voivodeship.

References

1. Aaron Agbenyegah Agbo, Wenfeng Li, Lanbo Zheng, Yanwei Zhang, Charles Atombo 2017. "Optimisation of intermodal freight transport network". *European Transport/Transporti Europei* 63(1). ISSN 1825-3997
2. Boyce David, Huw Williams. 2015. *Forecasting Urban Travel. Past, Present and Future*. Cheltenham, UK; Northampton, MA: Edward Elgar Publishing. ISBN 978-1-84844-960-2.
3. Cambridge Systematics, Inc. 1996. *Travel Survey Manual*. Washington, DC: Department of Transportation and Local Government.
4. Department of Economic and Social Affairs. 2005. *Designing Household Survey Samples: Practical Guidelines*. *Studies in Methods*. New York, NY: United Nations.
5. Department of Transport. 2016. *National Travel Survey: England 2015*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/551437/national-travel-survey-2015.pdf.
6. Drisya Manghat, Aswathy KV, Krishnamurthy Karuppanagounder. 2017. "Automobile level of service criteria for two-line undivided heterogeneous urban corridors". *European Transport/Trasporti Europei* 63(9). ISSN: 1825-3997.
7. Erfan Hassannayebi, Seyed Hessameddin Zegordi, Mohammad Reza Amin-Naseri, Masoud Yaghini. 2017. "Optimizing headways for urban rail transit services using adaptive particle swarm algorithms". *Public Transport*. DOI:10.1007/s12469-016-0147-6.
8. Jacyna-Gołda Ilona, Jolanta Żak, Piotr Gołębiowski. 2014. "Models of traffic flow distribution for various scenarios of the development of proecological transport system". *Archives of Transport* 32 (4): 17-28. ISSN 0866-9546. DOI: 10.5604/08669546.1146994.
9. Kulpa Tomasz, Andrzej Szarata. 2016. "Analysis of household survey sample size in trip modelling process." *Transportation Research Procedia* 14: 1753-1761.
10. Nosal Katarzyna, Wiesław Starowicz. 2015. "Evaluation of influence of mobility management instruments implemented in separated areas of the city on the changes in modal split". *Archives of Transport* 35(3): 41-52. ISSN 0866-9546. DOI: 10.5604/08669546.1185186.
11. Pallav Kumar, Krinshnanunni, Shriniwas Arkatkar, Gaurang Joshi, 2017. "Examining microscopic and macroscopic traffic flow parameters at diverging section on multilane urban roads in India". *European Transport/Trasporti Europei* 63(2). ISSN: 1825-3997.
12. Pyza Dariusz. 2011. "Multi-criteria evaluation of transportation systems in supply chains". *Archives of Transport* 23(1): 47-65. ISSN 0866-9546.

13. Richardson Anthony J., Elizabeth S. Ampt, Arnim H. Meyburg. 1995. *Survey Methods for Transport Planning*. Melbourne: Eucalyptus Press.
14. Sierpiński Grzegorz. 2017. "Distance and frequency of travels made with selected means of transport. A case study for the Upper Silesian conurbation (Poland). In: G. Sierpiński (ed.). *Intelligent Transport Systems and Travel Behavior. Advances in Intelligent Systems and Computing*, 75-85. Cham, Switzerland: Springer International Publishing. ISBN 978-3-319-43990-7. DOI: https://doi.org/10.1007/978-3-319-43991-4_7.
15. Statistics Canada. 2010. *Survey Methods and Practices*. *Catalogue no. 12-578-X*. Statistics Canada: Ottawa. ISBN 978-1-100-16410-6.
16. Taylor Michael A.P., William Young, Marcus Wigan. 1992. "Designing a large-scale travel demand survey: new challenges and new opportunities." *Transportation Research Part A: Policy and Practice* 26A(3): 247-261. ISSN: 0965-8564. DOI: 10.1016/0965-8564(92)90035-6.
17. Transportation Research Board. 2007. *Metropolitan Travel Forecasting. Current Practice and Future Direction. Special Report 288*. Washington, DC: Transportation Research Board.
18. Transportation Research Board's Travel Survey Methods Committee. *The On-line Travel Survey Manual: A Dynamic Document for Transportation Professionals*. Available at: http://www.travelsurveymanual.org.
19. Wasiak Mariusz. 2016. "Vehicle selection model with respect to economic order quantity." *Archives of Transport* 40(4): 77-85. ISSN 0866-9546. DOI:10.5604/08669546.1225471.

Received 03.02.2017; accepted in revised form 30.04.2017