Monitoring specific travel behavior in floating car data

Zuzana Purkrábková, Pavel Hrubeš

***Abstract*— Modern telematics systems are based on data collected from real operation. The article works with a new source of traffic data, which is freely available to the public. Linking traffic data with spatio-temporal significance also allows us to examine data from the perspective of user behavior. From the principle of data collection, the data already have their specific assumed behaviors, which are described in the article. It is desirable to further investigate these behaviors, to what extent they are affected and whether they can be considered as parameters of the entire traffic flow.**

***Index Terms*— data collection, floating car data, geographic information systems, travel behavior, stakeholders**

I. Introduction

With increasing mobility, we also have to deal more with the optimization and management of traffic flow not only in urban areas but also in the suburbs.

The idea of Smart City and intelligent transport systems technology have a common building block, which is data collection. Floating car data (FCD) is one of the modern ways of collecting traffic data. Traffic monitoring on roads by connecting measuring points is already being used. However, the system costs for the operation of a system are proportionally high compared to the size of the monitored area. Data from floating vehicles do not need any additional installations around the road, which is a significant advantage of this technology.

An equally important part is the monitoring and analysis of user behavior, which further helps us to understand such a complex issue as transport.

The aim of this research is to extract the value of data collected by floating vehicles, especially from the perspective of the behavior of the user of the floating vehicle. The authors of the article have previously dealt with the calibration of the FCD model in the Czech Republic [9], but also with the possibilities of whether excesses, accidents and congestion can be observed on floating car data [10]. This article focuses on the specific behavior of floating vehicle users. The article describes the initial phase of research of alternative routes and analysis of driver behavior as such, and also suggests examining these drivers in the event of unexpected events.

Zuzana Purkrábková, Pavel Hrubeš, Czech Technical University in Prague, Faculty of Transportation Science, Department of Transport Telematics

978-0-7381-3158-0/21/$31.00 ©2021 IEEE

II. Literature review

Floating car data have come to the forefront of transport research groups in recent years. The article [1] examines FCD data and their application in road management. Authors work with the quality of the data sent and try to reduce the amount of data sent, which is often duplicated by several vehicles. The paper [2] discusses the possibilities of using data from taxi vehicles to track route selection in the city. The authors also point out the distortion of floating car data of taxi drivers and further discuss these parameters. In general, the use of similar data in urban areas is a common research element. These data have been used in a pilot project for example in Stockholm [3] or Beijing [4].

More research, such as [5] or [6], also examines travel behavior and user decisions. The second mentioned paper deals with rational decision-making, applied to transport issues.

Attention should also be drawn to the occasional confusion of floating car data terminology. For example, the authors [7] consider floating vehicles to be those where they detect the movement of the steering wheel, the depression of the brake pedal or the accelerator. Based on these signals, they monitor driver behavior.

In the Czech Republic, a study [8] was published containing recommendations and drawing attention to possible obstacles to the application of the data collection system from floating vehicles in the Czech Republic. The authors of the article, as a team of experts FTS CTU in Prague, participated in the calibration of the model created for data in the Czech Republic. The development and modification of this model is dealt with in already published reports, the latest [9].

III. Available data

A new source of traffic data from floating vehicles is available in the Czech Republic. The data source is freely available and state-guaranteed. These are not explicitly open data as we have known from other systems. For their collection, it is necessary to conclude a contract with the provider. It is not possible to download the data arbitrarily, it is necessary to have a contract with Directorate of Roads and Motorways, thanks to which the data is sent in the required interval to our server.

Data collection is only one part of similar systems. The Directorate of Roads and Motorways (hereinafter abbreviated as RSD) acts as a data provider, managing the data collection. Subsequent data collection and processing are also very important in floating vehicle systems. The data needs to be meaningfully passed on to other elements of the system who already want certain reliability of this data. That is why the company VARS Brno, a.s. is involved in the data processing in the role of model developer. An insignificant element is also the control and advisory authority from the academic sphere, which in this case is represented by the Faculty of Transportation Sciences, Czech Technical University in Prague (FTS CTU). A team of experts from FTS CTU worked on the calibration of the data model, which is used for the described data source.

Data are collected from commercial vehicles, such as business vehicles, work vehicles, taxis, public transport vehicles, and more. Typically, data is used by a company for its own needs, so the primary purpose of data collection is not to provide it to track traffic flows. Data are collected for the purpose of self-management or monitoring of the fleet, or for their management and optimization of operation. The data are only provided in a secondary view to monitor the general traffic flow. The advantage of such a principle is the fact that the data are collected with high quality (the primary use is for the company which is paying the collection). Data are from more, private and state entities and are therefore not unilaterally focused.

In the Czech Republic, floating car data is collected using GPS units in the vehicle with a sending interval of one minute. The data contain a total of 15 parameters, on the basis of which it is possible to create analyzes of the road network. In addition to the time stamp, day, the designation of the TMC segment is mainly the number of vehicles, speed and travel times of the current and free traffic flow, the degree of data reliability, the mark for congestion, and estimation of its length and termination (see Table 1).

TABLE I
FLOATING CAR PARAMETERS

TMC segment	Travel time	Reaction time
Day	Delay	Traffic level
Time	Freeflow_speed	Congestion
Count	Freeflow_time	Congestion_from
Speed	Reliability	Congestion_length

More specifically, the collected data are described in [11]. To maintain the anonymity of the data, the first sections of the journey are not stored for floating vehicles. It is not possible to monitor the starting position of individual rides.

IV. DEFINITION OF THE KEY STAKEHOLDERS

Floating vehicles are new dynamic traffic information collected in real-time, which is why so much attention has been paid to them recently. According to the information above, it is clear which companies are directly involved in the operation of this system. However, if we want to further benefit the maximum from this data, it is appropriate to define who will use the data and also what will be expected from the system. Thanks to this, we will prevent possible discrepancies in the use of data by other systems and thus reduce the need for further pre-processing of data for new systems.

For this purpose, key stakeholders in the area were defined. These are the entities affected by the collection and subsequent processing and operation of data, as well as entities that continue to use FCD data. All subsequently defined participants participate in the operation of systems using FCD. Participants have their requirements for the system and their ideas about the organization and definition of data as such. A total of five main groups of key stakeholders (KS) affected by FCD issues have been identified:

- Entities providing inputs (RSD, VARS, companies that have fleet vehicles, rescue services, service operators (data operators, telecommunications services, GNNS), bus carriers, company trucks)
- Data processing entities (RSD (National traffic information center), VARS, FTS CTU in Prague)
- Entities using system outputs (RSD, traffic information service providers, producers and operators of telematics systems, producers and operators of navigation services)
- End users (drivers, passengers)
- Entities influencing legislation (Ministry of Transport, European Union)

The groups were designed to take into account the roles and goals they play concerning floating car data. The individual groups can be further described as a whole, as their attitudes and goals are usually the same. With regard to the areas of their activity, specific entities may be included in more than one group at the same time. The groups of key stakeholders and their influence are shown in the following diagram, see Fig. 1.

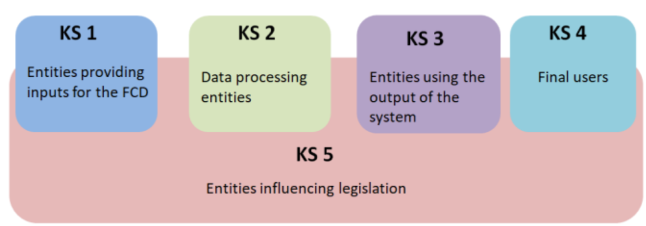

Fig. 1: Schema of the key stakeholders

Furthermore, the general objectives of individual groups of key players were defined on a very general scale. Table 2 is shown below. It is obvious that there is a partial intersection of the goals of individual entities in the use of data from floating vehicles. This also leads to the idea that the key stakeholders will try to find a compromise solution to get the system up and running quickly.

Table 2: The goals of the key stakeholders

GROUP OF KS	DEFINED GOALS OF THE GROUP
KS1	Supervision of vehicle movements in the network, smooth traffic flow
KS2	Reliable data, publishable data, calibrated model
KS3	Reliable data, satisfied customer, informed user
KS4	Smooth traffic flow, sufficiency of information, comfortable travel
KS5	Sustainable mobility and transport, legislative protection of the population

When building a new system like this, it is necessary to identify important links, representing the relationships between individual groups. On these links, then assess the common interfaces, ie determine the regularity of these links and, in the case of irregular links, find appropriate ways to tune them. Above all, it is necessary that there are regular links between key stakeholders providing data or information, so that, for example, data is not published in different formats, and the results are not mutually compatible.

Furthermore, the functions of individual groups of key stakeholders were defined, which are based on their distribution and use of data from floating vehicles. From the names of individual groups, their functions are noticeable.

The key stakeholders' group no. 3, which uses FCD data in its applications or models, has a great influence on solving the issue of user behavior on roads. The group can transform data and information into a passable and easy-to-understand form. This information can then be passed on to other key stakeholders, as it is already, for example, visualized or specified for individual situations. The main function of this group is thus the convenient processing of data into a publishable form that is easy to understand for end-users

The article will also deal with a group of end-users who belong to the key stakeholders' group no. 4, while their function in the system is primarily the end-use of the transmitted data. Most often, the data, thanks to this group, serve to speed up the decision-making of road users leading to smooth traffic, or, thanks to timely information, lead to the minimization of risks and accidents. This group will most influence the target effect of the whole system, as the whole transport system depends on its decision-making and appropriate use of traffic information. In other words, the driver, who decides based on current information about his further behavior, greatly influences the entire traffic flow.

V. Data specificity based on travel behavior

As follows from the previous paragraphs, the great advantage of this data source is a simple solution in terms of technology and installation. There is no need to install additional equipment in the surroundings of the infrastructure for data collection. From the user's point of view, this provides more information from the road network. Data from roads that are not equipped with detectors are also available.

However, the data addressed also have certain specifics, mainly due to the fleet providing transport data. It is desirable to present the data correctly to avoid inappropriate presentation of data in related systems. Therefore, the following part is an analysis of the specifics of the data, which is crucial for multiple key stakeholders in the system.

The resulting model is highly dependent on vehicle penetration, which is the distribution of floating cars on the network. With low penetration of vehicles, especially at night and in the morning, there are larger fluctuations in the measured parameters in the data. A similar problem occurs on weekends, when the traffic of company's vehicles, tucks and buses is smaller.

As floating vehicles are mostly company or business vehicles, their drivers also drive accordint to the traffic rules more than other drivers. This can be explained by the fact that drivers of floating vehicles very often feed their driving license and are aware of their possible offenses. This is evident, for example, in the distribution of speeds on the highway section of July last year (2020).

The graph (see Fig. 2) shows a speed histogram on the TMC segment number TS23990T25532, which is located on the D1 highway near Prague. The x-axis shows the individual speed values in km/h, the y-axis shows their frequencies. This is a section of highway that is not speed limited. It is obvious that drivers of floating vehicles very rarely exceed a speed of 130 km / h, which is the maximum speed limit for highways in the Czech Republic.

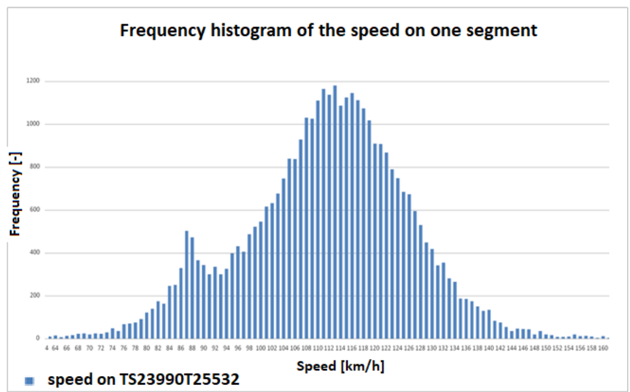

Fig. 2: Histogram of the speed

In the data, it is also possible to monitor the atypical behavior of the traffic flow on the main roads and the creation of the so-called asymmetric traffic flows. These are probably due to a good ratio of high density of roads in the Czech Republic to the area of the Czech Republic. This creates opportunities for drivers to use multiple roads and connections within the route of one day, multiple cities located on different routes.

Let's take a look at the problem on the example of the highway number D1. The number of vehicles traveling from Prague in the morning is not even equal to the number of vehicles traveling to Prague in the evening. This asymmetry can be observed throughout the month. Therefore, it cannot be said that the flows in the network are comparable in both directions during the month. The data for this example are shown in Fig. 3.

It is possible to observe a comparison of the count of floating vehicles on the section of highway D1 at the 17th km numbered from Prague from 8 to 20 July 2020. A whole month was intentionally not selected, which would be

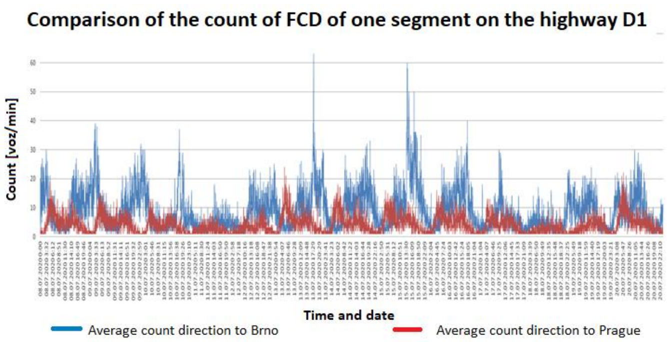

Fig. 3: Comparison of the count on the highway

distorted by public holidays at the beginning of the month and also by the beginning of the summer holidays.

The x-axis shows the time and day, the y-axis the average count of FCD vehicles, which was counted from minute intervals to the more suitable time period. The data of vehicles going to Prague are shown in red, and the vehicles going to Brno are shown in blue. It is obvious that in weaker traffic times the data are relatively comparable, but at peak times the number of vehicles going from Prague to Brno always predominates.

This may be due to the characteristics of data from floating vehicles, where these vehicles are often logistics service vehicles, where, for example, Prague, Brno, and Hradec Králové can be served by one vehicle in one day, which is reflected in the asymmetry of traffic flows.

Based on data from the Czech Republic, it was examined whether it is possible to detect and monitor accidents in a small area of the Czech Republic. This is followed by the creation of alternative paths and their selection by individual users.

VI. Travel behavior in the case of accident

The issue of congestion has been discussed for a long time. The White Paper on European Transport Policy for 2050 [12] has already defined that congestion costs in Europe are around 1% of GDP. In 2019, residents of major European cities spent about 140 hours in congestion according to the INRIX rankings [13]. There is therefore a clear need to address this issue.

It is obvious that the route requirements will be slightly different for urban and interurban areas. It is logical that over longer distances, people stop dealing with the number of kilometers traveled, but on the contrary, they are interested in minimizing their travel time.

If the assumption that data from floating vehicles can be perceived as concise characteristics of the traffic flow is confirmed, this data can be used in the future, for example, for the detection of traffic excesses. The following part is an introduction to the initial phase of this research, where the behavior of the user of floating vehicles is monitored, then it will be compared with other data sources that are not yet available.

The authors have already addressed the detection of traffic excess from a technical point of view in the article [10]. Two events on the main highway D1 between Prague and Brno were discussed, when the main traffic flow was completely closed. In one case a detour was recommended, in the other case, there wasn't any detours for the accident. At least for a while, it was possible to assume the transfer of traffic to the surrounding roads.

The behavior of users in the event of congestion or accident is one of the questions that the authors will address in the future. In particular, the discussion of the hypothesis of how unexpected situations affect the behavior and decisions of drivers and what is the impact of their decisions on the environment. To verify this hypothesis, it is appropriate to use a different type of apparatus than purely mathematical but to lean, for example, more towards economic or sociological theories. It is this nationwide source of data that should confirm our hypotheses.

The theory of rational choice should be used for this if we assume that users behave rationally in the economic sense. The assumption is that the user evaluates how it could benefit from individual routes (whether the route will be faster, more comfortable, or mentally more satisfying) or what benefit it will have if it stays on the main route.

The driver decides on the route that will provide him with the most subjectively evaluated benefit. This means that if he is offered a faster route, even if it is not more comfortable than an existing route, for example, he will probably use it. Only if the options are comparable and do not find a difference between them, it is likely to remain on the main route.

Rational decision-making includes a choice of all available options and the user is interested in what results arising from individual actions. In this case, actions are only a tool to achieve a result. The possibilities are expressed as a set of actions $A = \{a_1, a_2, ..., a_j\}$, where in our case there will be elements and individual alternative routes. It is assumed that the alternatives can be compared and the resulting order of preference determined.

In our case, however, there is a limit in the time for choice. This limits the distance of the current vehicle location from the last chance to change the route in the event of an accident. In that case, we are talking about the so-called limited rationality. The user then decides on a satisfactory solution rather than an optimal one. The driver does not have to do complex analyzes and his decision is feasible in practice.

We will take a closer look at the situation where a detour was recommended. The behavior of drivers was analyzed based on data from floating vehicles. Assumptions were made as to how drivers would behave. The most common assumption was:

- a_1 = Stay on the main road and wait for the situation to be mended
- a_2 = Follow a detour (II / 602)
- a_3 = Choose another alternative route in its sole discretion

The assumption of following dynamic navigation based on the alerts of other users is not stated, as the event took place late in the evening and no oversaturation of the detour route is expected.

From the currently available data, the day of the accident and also the other days of the month (July 2020) were analyzed. The behavior of FCD drivers and other characteristics that can be read from the data were monitored. It is obvious from the data that in the situation of an accident, the drivers in the absolute majority chose a detour (road II / 602).

When comparing other days, however, it is clear that the bypass road is not used as much on average as the road II / 395, which is used more by floating vehicles in the monitored period. This can be seen in the pictures from the model below (Fig. 4), where the speed is shown in the color scheme, where the lightest green means low speed, the darkest speed over 100 km / h. The figure above shows an accident situation where the occupation of road II / 602 is obvious and, conversely, there is not a single floating vehicle on road D1 (highlighted in light red). The picture below shows the situation on another day, when the occupation of other roads is obvious, very often II / 395.

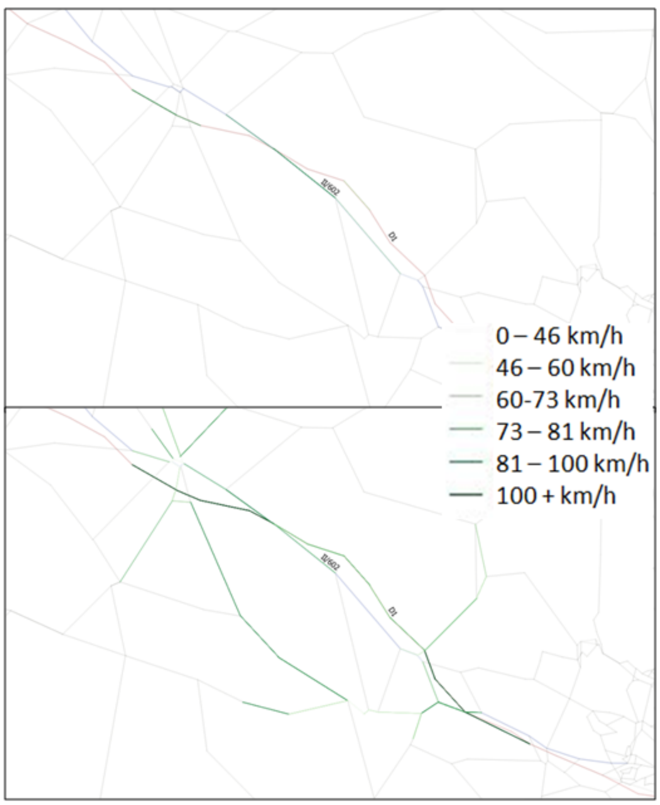

Fig. 4: Selection of communications by FCD drivers

Thus, other roads than the recommended detour route are commonly used. In this case, a detour route could be used for time reasons, when the situation was resolved in the evening and there was no capacity problem on other roads.

It is appropriate to further monitor the use of other roads, even lower classes, as shown in the following figure (Fig. 5), which shows the common situation of the monitored area in the early morning (5 a.m.). It is obvious that abbreviations between large cities are used in the area at a time when there is not a lot of traffic on the roads. In the picture, you can see the traffic leading off the highway, especially between the towns of Rosice and Velká Bíteš, while both towns are located directly at the exit of the highway.

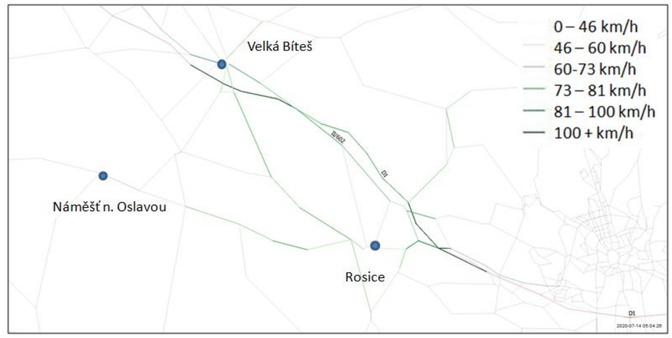

Fig. 5: Morning operation of FCD vehicles

It is obvious that with the help of this data it is also possible to monitor the selection of the communication route due to the spatiotemporal characteristics of the data. QGIS software was used to display and understand the significance of this data.

In the future, the user's behavior in unexpected situations will be examined in more detail, as well as the impact of their choice on the surroundings of the road.

VII. Conclusion and Future Work

The collection and processing of traffic data is now a much-discussed topic. For the modernization of transport systems, data is an elementary element that needs to be examined in more detail. The article describes the initial research of the data source, which is newly available in the Czech Republic, free and state-guaranteed.

A proper understanding of the meaning of data is required for the application of data and its full extraction. Therefore, in the first part, the involved subjects were described, for which this step is basic in the implementation of the system in the future. From the introduction, it is also clear which vehicles become moving probes as floating vehicles and who is involved in the implementation of the entire model within the country.

Furthermore, the article shows the specifics of these data and raises the question of whether in the future these data can be used as a characteristic of traffic flow not only in urban areas but also in suburban areas and main highways. The specifics of compliance with the rules of the road, the average speed of the fleet of floating vehicles, and last but not least, the specifics of the road network were shown.

Furthermore, the authors propose an initial investigation of the behavior of FCD drivers. Accidents on the main Czech highway D1 were monitored and, based on floating car data, the choice of drivers' routes was retrospectively evaluated. It is obvious that the choice of communication can be well monitored on the basis of the solved data. In the future, this step will be linked to the technical background, where, based on hypotheses, data from floating vehicles will be examined for the creation of excesses, the creation and development of alternative routes, and other options. These solutions should then be supplemented by an analysis of the behavior of drivers and the impact of their choice and behavior on the surroundings of the road, ie especially the village.

Acknowledgement

Data from the Transport Portal of the Czech Republic were used for this project (data source: ŘSD ČR – JSDI – www.dopravniinfo.cz). This research was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS SGS21/080/OHK2/1T/16 - Detection of detours of an accident on main roads based on floating car data.

References

[1] B. S. Kerner, C. Demir, R. G. Herrtwich, S. L. Klenov, H. Rehborn, M. Aleksic, A. Haug , "Traffic state detection with floating car data in road networks", In Proceedings *2005 IEEE Intelligent Transportation Systems*, Vienna, 2005, pp. 44-49, doi: 10.1109/ITSC.2005.1520133

[2] M. Jürgen, K. Thiessenhusen, P. Wagner, "Using FCD-data for measuring route-choice probabilities", In Proceedings *Behavioural Responses to IT*, 2003.

[3] M. Rahmani, E. Jenelius, H. N. Koutsopoulos, "Nonparametric estimation of route travel time distributions from low-frequency floating car data", In: *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 343–362, 2015.

[4] L. Zhu, J. Sun, H. Wen, F. Chen, Y. Gao, "Experimental Research of Real-time Floating Car System in Beijing", 2008 11th International IEEE Conference on Intelligent Transportation Systems, Beijing, China, 2008, pp. 1173-1176, doi: 10.1109/ITSC.2008.4732614.

[5] E. D'Andrea and F. Marcelloni, "Detection of traffic congestion and incidents from GPS trace analysis", In: *Expert Systems with Applications*, vol. 73, pp. 43–56, 2017.

[6] S. Rasouli, H. J. P. Timmermans. *Bounded Rational Choice Behaviour: Applications in Transport* (eds. 2015). Bingley: Emerald.

[7] F. Meng-yin, S. Zhong-hong, Y. Yi, W. Mei-ling, "The recognition and applications of floating cars' driving behaviors," *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, Jilin, China, 2011, pp. 680-683, doi: 10.1109/MEC.2011.6025557.

[8] P. Hrubeš, M. Langr, P. Derbek, D. Saiko, M. Volný, "Zmapování služeb a dat v oblasti FCD (Floating Car Data) pro využití v rámci informačních systémů ŘSD*",* Prague: CTU in Prague, Faculty of Transportation Sciences, Department of control technic and telematics, 2010. Research report no. LSS 384/10.

[9] P. Hrubeš, Z. Purkrábková, M. Kovaljov, K. Kovaljová, J. Hampl, "Kontrolní a expertní činnost při zavádění a následném provozu FCD – kalibrace výpočetního modelu III", 2020, CTU in Prague, Faculty of Transportation Sciences, Department of transport telematics, Research report no. 97/2019b.

[10] Z. Purkrábková, P. Hrubeš, "Analysis of driver choice at the closure of the main road using floating car data", In: GARI winter multidisciplinary symposium 2020. Global academic research institute. 2020. ISBN 978-955-7153-04-9.

[11] Z. Purkrábková, P. Hrubeš, "The utilization of foating car data for traffic flow analysis", In: Young Transportation Engineers Conference 2020. Praha: CTU. Faculty of Transportation Sciences, 2020. p. 119-125. ISBN 978-80-01-06793-2.

[12] WHITE PAPER: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system, European Commision, 2011, Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011DC0144&from=EN

[13] INRIX 2019 Global traffic storecard: Interactive Ranking & City Dashboards, online [cited 12.02.2021], 2021, Available at: https://inrix.com/scorecard/?