The possibility of using floating car data to monitor the occurrence of accidents

Initial analysis and results to verify whether it is possible to monitor the occurrence of accidents on the newly accessible data source

ZUZANA PURKRÁBKOVÁ

Czech Technical University in Prague Faculty of Transportation Sciences

PAVEL HRUBEŠ

Czech Technical University in Prague Faculty of Transportation Sciences

Today, traffic management is based on traffic data, which must be high quality and reliable. The article deals with a new data source that is available in the Czech Republic. Thanks to this dataset, it is possible to get an overview of a wide road network throughout the whole area of the country. The authors use a broader view of the road network to study the behavior of drivers of floating vehicles in an accident on the main road. The high data coverage promises the ability to monitor traffic spills in the event of an emergency. These data can be used in the future to improve the overview of the current state of traffic and road management.

CCS CONCEPTS • geographic information systems • floating car data • data processing

Additional Keywords and Phrases: accident detection, sustainable data source, traffic management

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, 2018. The Title of the Paper: NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

With the growing number of vehicles on the road network, there is a growing need to optimize and control traffic flow in real-time. This requires a considerable amount of information and traffic data, which serves as a basis for strategic decisions. In general, data is the golden grail for any traffic management. This article offers an insight into a new data source in the Czech Republic, which is linked to floating vehicles.

Floating car data (sometimes also known as floating probes) is time-stamped traffic data that contains time-stamped geolocation and velocity information. Data is collected directly from the road network without the need for additional installation of technology along the road. Each vehicle must have location equipment.

In the case of the Czech Republic, so-called GPS-based FCDs are used. This means data obtained from vehicles equipped with a GPS receiver. The GPS unit is activated automatically when the vehicle is started moving, collects data (GPS position, the direction of travel, speed, ...) while driving, and switches off again when the engine is switched off. The data is sent at a certain interval to the server, where it is processed and stored. Therefore, there is a need for hardware in the vehicle and a server for data storage.

This paper is structured as follows. After the literature overview, the dataset and its basic characteristics are described in section 3. Section 4 treats the question of whether a traffic jam or accident can be detected from the data source. A model and the result of the analysis are presented. Finally, a conclusion and an outlook on further research are given.

2 STATE OF THE ART

As mentioned in the introduction, the collection and use of traffic data is a key role in modern transport concepts. It is well known that traffic data is used not only for traffic monitoring, management, and prediction, but some research also deals with crisis management, such as [1], where data is undoubtedly also a key component.

In general, several researchers, but also entities from the private sector deal with the issue of data from floating vehicles. Very often, projects focus on data obtained from the taxi fleet, such as [2] or [3]. These researches focus on the use of data for urban areas. However, there are not many papers to obtain data from the suburbs. An example is [4], which deals with data from the motorway in Germany.

The initial analysis, which was carried out as part of this project, addresses, and controls the quality of the transmitted data. Several studies are examining how accurate traffic information is compared to the real situation on the roads. For example, article [5] examines traffic messages broadcast in Austria via a designated channel, where they have an average delay of about 10 minutes.

Paper [6] compares different methods of traffic data management and discusses the quality of traffic information. The basic idea is to understand quality as a level of similarity between the actual traffic situation and the transmitted data. One of the methods is based on tracking test vehicles, while another uses data from stationary detectors to reconstruct the traffic flow. These methods compare sources directly, without one of the sources being considered accurate.

Today, FCD-based technologies are used by almost all providers of traffic information such as TomTom, Google, Inrix, etc. As a result, they can provide real-time information. Article [7] Today, FCD-based technologies are used by almost all providers of traffic information such as TomTom, Google, Inrix, etc. As a result, they can provide real-time information. Article [8] works on the data from taxi vehicles in the urban area. Based on the model, they estimate the average speed and propose a method for identifying traffic congestion with an average speed.

After verifying and validating the model and its parameters, the described data could also be used for traffic modeling and management. Congestion prediction, which may be another output of these data in the future, is also addressed in the paper [9], which works on the data from taxi vehicles in the urban area. Based on the model, they estimate the average speed and propose a method for identifying traffic congestion with an average speed.

So far, there are not many papers dealing specifically with this data source. The authors [10] focused mainly on the initial design and implementation in the Czech Republic. Furthermore, the report [11] describes in more detail the calibration of the model as such and the debugging of errors. The last update and debugging of the computational algorithm were done by data from April 2021.

3 DATASET

For the third year, data from floating vehicles have been available in the Czech Republic through the interface of the Directorate of Roads and Motorways (hereinafter ŘSD). This data source is publicly funded and still in pilot operation and is suitable for further investigation. Sustainable transport is a big topic today, and this data source can contribute to more detailed or complementary information from road traffic.

This data is collected through a fleet of floating vehicles, which are mainly fleet vehicles of commercial and state entities, vehicles of public transport and other entities (such as emergency and safety services etc.) involved. In total, there are about 150,000 vehicles [12] on the road network in the Czech Republic.

This method requires a large penetration of vehicles in the network. Therefore, this method used to be prevalent in urban areas, where there is a high incidence of public transport vehicles and taxis. For motorways and highways, the floating vehicles must be detected at least 15 minutes intervals.

3.1 Spatial display of data on the network

The location on the road network is connected with the data. For this task, it is possible to use two networks, Global Network or TMC segment network (their comparison can be seen in Figure 1). The figure shows the coverage of both networks over the Czech Republic. However, for the purposes of analysis on motorways and 1st class roads, the network of TMC segments, which is shown on the right side of the figure, is sufficient. This network contains almost 30,000 sections of roads in the Czech Republic. In the case of TMC segments, these are defined sections The QGIS geographic information software is used for its sections due to its free version and wide user support.

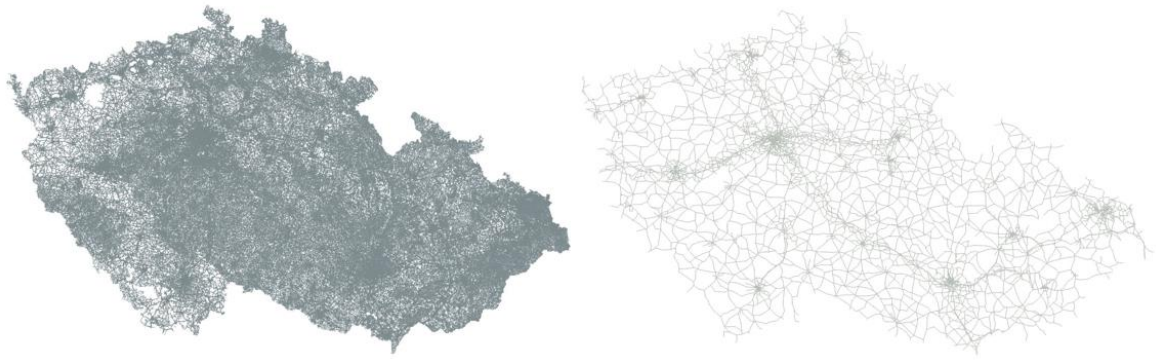

Figure 1: Comparison of Global Network (left part) and Network of TMC segment (right part)

When locating individual data on the spatial network, critical events will be searched. It is precisely these events that lead to the disruption of the free traffic flow. Already in the previous project, events leading to the complete closure of the main road were monitored and the diversion and change in behavior in this situation were monitored. Currently, the project deals with only a partial reduction of traffic flow. During these events, the diversion of the main road to detours will be further monitored. The user's behavior will also be monitored.

3.2 Monitored situations

For the analysis of the accident on the main road, the events listed in Table 1 were monitored. For these events, data from floating vehicles were examined, and especially the behavior of drivers in these events was examined. The events were selected based on subjective evaluation and general knowledge of the road network in the Czech Republic. In general, it should be recalled that during the period under review (the year 2020 and 2021), traffic was affected by the consequences of the COVID19 pandemics. Some accidents may have been distorted by a lower number of vehicles, including floating ones. At a critical time, the movement on the Czech roads decreased, so it is necessary to consider the possible distortion.

Table 1: List of monitored events

Date	Location	Excess	Note
12.1.2020	D5, 142.5km direction Rozvadov, 1pm	several traffic accidents	The motorway is closed at one time (approximately 2 hours)
26.6.2020	D1, 188km direction Brno, 5am	Traffic jam 6km long	Drivers take detours on road no. 602
9.11.2020	D8, 24km direction Ústí n.Labem, 1am	the motorway is closed at one time	Traffic jams also on detours
8.2.2021	D1, 80-100km both directions	Traffic jam several km long	Icing on the territory of the whole Czech Republic
17.2.2021	Whole country	Several traffic accidents	Icing
1.4.2021	D1, 11km direction Brno	Variable traffic jams	Easter, an extended weekend
24.8.2021	D1, 81-92km, both directions	Traffic jam up to 10km	Traffic is diverted to only one lane in each direction
8.9.2021	D8, 38km direction Prague, 6am	accident in the repaired section	Traffic jams also on detours
31.10.2021	D1, 209.5km, direction Brno,2pm	mass accident (6 vehicles)	Accidents also on detours, traffic jam till the evening on all roads Vyškov → Brno

3.3 Data pre-processing

The data needs to be pre-processed before examining the situations themselves. For this purpose, the work with the pgAdmin database together with the SQL language was used. After uploading the data, it was necessary to check how many TMC segments were detected in a given period, but also to check whether data errors were occurring. For example, it was checked whether the bidirectional sections showed two directions. Furthermore, a time base was created in the database. That could help to examine the data for a specific time across the entire road network in the Czech Republic.

An important step was the calculation of penetration P_C, which describes the occupancy, resp. the speed of specific TMC segments during the period under review. The P_C can be obtained as follows

$$P_C = \frac{x}{\sum t_i}$$

where x represents the number of vehicles for a given section or the average vehicle speed for a given section. The sum then indicates the total number of time records.

It is also possible to calculate the penetration of vehicles only for peak periods when we can focus only on the periods of morning and afternoon traffic peaks. According to the knowledge of the area in our case we are talking about the period 7-11 am and 2-6 pm.

4 ANALYSIS AND EVALUATION

Two situations that occurred on the highway network of the Czech Republic were selected to demonstrate the results. The first event is from November 2020, when transport has not yet been hit hard by the pandemic COVID19. The second event was selected from the beginning of 2021, specifically from February 2021.

As the article describes the beginning of the research, the data were examined only in terms of penetration, coverage of TMC segments, and especially the speed and use of various routes for detours. The aim was to find out whether the selected accidents are recorded in the data, whether the traffic flow is obvious, and

therefore whether it would be possible to automatically detect or predict these accidents, congestion and other emergencies in the data in the future.

4.1 Situation A

The first situation follows a special event on the D8 motorway, which leads to the north of the republic. At the observed time, an accident occurred on the road. According to official traffic information (provided by the Police of the Czech Republic), this caused a slowdown in traffic and diversion of traffic to secondary roads.

The aim of this event was to monitor the behavior of drivers of floating vehicles, whether the reported behavior is transcribed into the examined data, and whether it is possible to monitor other peculiarities.

The observed event took place on Monday and the dates were compared with the surrounding days and the following Monday. The penetration P_C for the whole monitored period (November 2020) was on average 14% if we monitored only the peak period (7-11 am and 2-6 pm) we are talking about a value of 22.5%. Thus, the data can be considered relevant and it can be assumed that it should characterize the traffic flow.

The data from the floating cars are shown in red in figure 2. The rule is that the darker the color, the faster the traffic flow. If there is no red line on the road, data from floating vehicles was not captured for this road at the time (no floating vehicle was driving there). The pictures show when the accident occurred (picture on the left) and the course of the accident, when no vehicles drove along with the main traffic (picture on the right). It should be noted that, compared to similar times, the motorway is very sparsely occupied by floating vehicles, but relatively covered.

Figure 2: The course of the described situation A - creation of the accident (map on the left), the course of the accident (map on the right) on the D8 highway in the direction of Ústí nad Labem

It was also clear from the analysis that traffic limited the main road for about an hour (interruptions can be observed between 1:25 am and 2:45 am). This is different from the official reports reported by the police of the Czech Republic, which stated a longer closure of the highway and detours. However, this statement cannot be confirmed from the data.

This is the result for further discussion. The data raises the question of whether the given event can be monitored in the data (this was not confirmed by the initial analysis). However, it should be noted that the accident occurred at night when vehicle penetration is significantly lower than during the day. In this case, it would be appropriate to supplement the floating car data, such as data from toll gates, stationary detectors, or CCTV to validate this claim. In this case, the authors do not consider only data from floating vehicles as a relevant source.

4.2 Situation B

The second situation occurred in February 2021. Ice was detected throughout the whole country during the day, which significantly limited traffic locally. In this case, the project deals not only with a specific location, but in general with the entire road network and monitors whether the icing phenomenon is transcribed into data for the entire road network.

Locally, the project also focused on monitoring a specific locality, namely the D1 highway in both directions (highlighted in orange on figures below). This is the main highway that connects the three largest cities in the Czech Republic. The intensities and speeds on this road are traditionally higher than on other highways in the Czech Republic.

It is therefore a different situation than the first one, where one specific accident is not monitored. In this case, the behavior of drivers of floating vehicles was monitored for a defined area. Also, the penetration of vehicles was compared at the beginning, the penetration for specific days was solved, namely for the day measured by ice and the same day for the next week. The resulting total penetration was 17.3%, resp. 12.6%. In general, penetration values are lower than in the case of the November event. It can be assumed that the covid situation worsened the penetration.

The following behaviors can be observed in documented images. Already during the night hours, a noticeable decrease in speed was observed on the main highway. This could have been caused by a drop in vehicle penetration at night. However, even from the previous analyzes, the highway network is successfully being filled with a sufficient number of floating vehicles after about 5 am. Here it is clear that even after 6:30 am the speed of floating vehicles is around 50kmph.

The situation is described in figure 3 and 4. In figure 3, it can be seen the situation in the morning. Normally, the highway is full at this time. But in the first map, there is no clear difference in speed on highways and other roads. This can be observed by all being the same shade of red. This situation is settling to around 11 am, which also corresponds to official reports from the Police of the Czech Republic. Until the normal traffic flow (ie to a state where the highway is recognizable in the data from other roads), the situation on the D1 highway arrives in the afternoon, specifically after 5 pm (see figure 4, where the difference is clear, the dark red is for speed high way). This can again be compared with other days of the month.

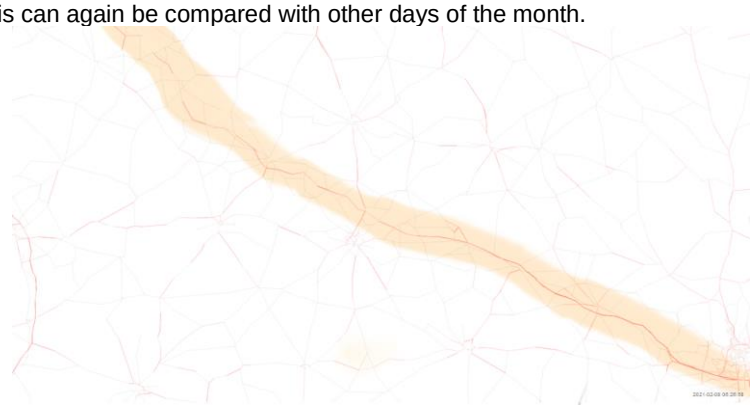

Figure 3: The situation on the D1 highway in the morning with congestion

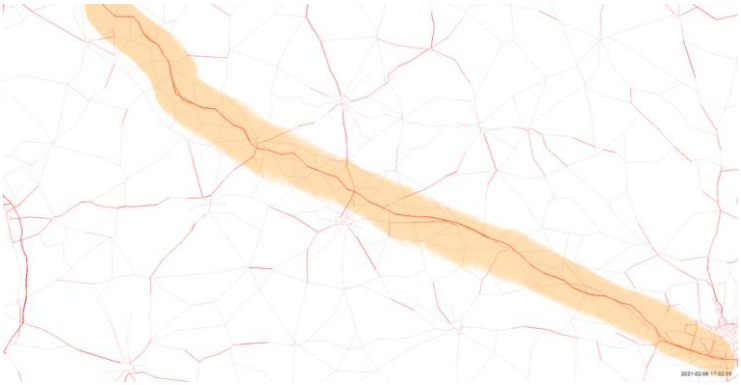

Figure 4: The situation on the D1 highway in late afternoon without the congestion

The event described above was also clearly reflected in the data. Surprisingly, not much traffic has moved to other surrounding roads. Official reports did not mention it either, but from the subjective knowledge of the territory, it was expected. This may be caused by meteorological conditions throughout the day when the icing formed on the territory of the entire Czech Republic. Another explanation may be the transfer of traffic to the second highway connecting these 3 cities, which runs north to our territory. To declare this, it is necessary to further examine the data from a global perspective.

5 CONCLUSION AND FUTURE WORK

The paper summarizes the initial processing and analysis of a new data source available in the Czech Republic. Since this is a free data source, guaranteed by the state, it is appropriate to further investigate it. The authors aim to verify whether the data can be used to clean up extraordinary events and driver behavior on detour routes. This has not been fully confirmed.

Data from 2020 and 2021 were processed, for which several events were selected. An initial analysis, including penetration values, was presented on the example of two events. Prior to the impact of the pandemic COVID19, the total vehicle penetration per month averaged over 20%, which is a quite good result. According to data, the value of the pandemic has fallen slightly.

In the future, it is necessary to further focus on whether the accommodation of the supplied data really occurred. It would also be appropriate to further link the data, for example, with meteorological data, which greatly affects traffic. If necessary, verify individual events based on data from stationary detectors so that results are conclusive and relevant.

In the future, it would be the combination of telematics systems and data from floating vehicles that it would be possible to monitor, for example, the quality of the road surface or even plan their partial closure in less frequented hours. Thanks to data from floating vehicles, it is also possible to solve telematics systems that warn drivers of an obstacle on the road. The premise of the idea is that telecommunications are equipped with variable traffic signs or the use of cooperative systems.

In general, it can be argued that the penetration of floating vehicles should increase in the future. This is also connected with the expected development of connected vehicles, ie vehicles that will have a GPS transmitter and a SIM card or other data connection, for example to the car manufacturer. Then it will be possible to use

this data as well. However, there is still a need to investigate the sufficient vehicle penetration that may have been affected by the COVID19 pandemic in terms of future driver behavior.

ACKNOWLEDGMENTS

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] Růžička Jiří and Navrátilová Kristýna, Crisis management as the part of smart traffic control in cities. In: 2020 Smart City Symposium Prague. New York: IEEE Press, 2020. ISBN 978-1-7281-6821-0.

[2] Erik Jenelius and Haris N. Koutsopoulos, Travel time estimation for urban road networks using low frequency probe vehicle data. In: Transportation Research Part B: Methodological, vol.53, pp 64-81. 2013. ISSN 0191-2615. https://doi.org/10.1016/j.trb.2013.03.008

[3] Tiedong W. and Jingjing H., Applying floating car data in traffic monitoring. In: 2014 IEEE International Conference on Control Science and Systems Engineering, Yantai, vol. 322, pp 96-99. 2014. http://dx.doi.org/10.1002/andp.19053221004

[4] Rempe, Felix, Franeck, Philipp, Fastenrath, Ulrich and Bogenberger, Klaus. Online Freeway Traffic Estimation with Real Floating Car Data. 2016. DOI: 10.1109/ITSC.2016.7795854.

[5] Bernd Rainer and Martin Müllner, 2011. , TMCplus - Improving the TMC information chain. Bundesanstalt für Straßenwesen (BASt), WIRTSCHAFTSVERLAG NW. ISSN 0943-93075.

[6] Klaus Bogenberger and S. Weikl, Quality Management Methods for Real-Time Traffic Information. In Procedia - Social and Behavioral Sciences 54, pp 936-945. 2012. ISSN 1877-0428. https://doi.org/10.1016/j.sbspro.2012.09.809

[7] Kim Seoungbum and Benjamin Coifman, Comparing INRIX speed data against concurrent loop detector stations over several months. In Transportation Research Part C: Emerging Technologies vol. 49, 2014. https://doi.org/10.1016/j.trc.2014.10.002

[8] Chang Ande, Jiang Guiyan and Niu, Shifeng, Traffic Congestion Identification Method Based on GPS Equipped Floating Car. In: 2010 International Conference on Intelligent Computation Technology and Automation vol. 3, pp 1069-1071. 2010. DOI: 10.1109/ICICTA.2010.105

[9] Ayesha Ata, Muhammad Adnan Khan, Sagheer Abbas, Gulzar Ahmad and Areej Fatima, Modelling smart road traffic congestion control system using machine learning techniques. In: Neural Network World, 2019. DOI: 10.14311/NNW.2019.29.008

[10] Pavel Hrubeš, Martin Langr, Přemysl Derbek, Dušan Saiko a Martin Volný, Zmapování služeb a dat v oblasti FCD (Floating Car Data) pro využití v rámci informačních systémů ŘSD. Research report LSS 384/10, CTU in Prague Faculty of Transportation Sciences, 2010.

[11] Pavel Hrubeš, Zuzana Purkrábková, Michal Kovaljov, Kateřina Kovaljová a Josef Hampl, Kontrolní a expertní činnost při zavádění a následném provozu FCD – kalibrace výpočetního modelu III. Research report LSS 97/2019c, CTU in Prague Faculty of Transportation Sciences, 2021.

[12] Milan Frydryšek, Standardizovaná data z plovoucích vozidel jsou zdarma k dispozici široké veřejnosti. In: Dopravní noviny [online], 2020 [cited 17.06.2021]. https://www.dnoviny.cz/informatika/standardizovana-data-z-plovoucich-vozidel-jsou-zdarma-k-dispozici-siroke-verejnosti

Authors' background

Your Name	Position*	Email	Research Field
Zuzana Purkrábková	PhD candidate	purkrzuz@fd.cvut.cz	ITS, transport, floating vehicles, GIS
Pavel Hrubeš	Dean of faculty, associate professor	hrubes@fd.cvut.cz	ITS, transport, GIS, optimalization, driver – car interaction