The Issue of Analyzing Measurement Data of Driving Speed in Large Urban Areas

Emilia Skupień and Agnieszka Tubis

Abstract The measurement of speed is the subject of analysis and publication of many researchers around the world. This issue is analyzed from the point of view of movement engineering, urban planning, logistics, but also the city management. The researches, initiated by the authors in this paper, are focused on measuring the time of vehicle journeys on the selected section of one of the main streets of Wroclaw. The aim of the chapter is to present the issues related to the analysis of measurement data from the ITS, which are aimed at supporting decision-making processes of those responsible for organizing transport in the city. For this purpose, the authors presented the basic theoretical issues of the study area and the approximate principles of operation of ITS systems. Then on the basis of quantitative analysis, the authors characterized the difficulties that may be encountered while drawing conclusions from the analysis of the collected data.

Keywords Movement engineering · Data analyzing · ITS

1 Introduction

The most important attribute of the road traffic is the speed of vehicles. It is a measure of the quality of movement and its consequences, both for its individual participants, as well as the general public dimension. The speed determines the comfort, convenience, traffic safety and the economics of transportation and the rank of traffic impact on the environment [14]. In large urban areas the speed of vehicles in addition to the usual groups of factors discussed in the literature section,

E. Skupień (✉) · A. Tubis
Wroclaw University of Technology, 27 Wybrzeze Wyspianskiego Street, Wroclaw, Poland
e-mail: emilia.skupien@pwr.edu.pl

A. Tubis
e-mail: agnieszka.tubis@pwr.edu.pl

W. Zamojski et al. (eds.), *Dependability Engineering and Complex Systems*,
Advances in Intelligent Systems and Computing 470,
DOI 10.1007/978-3-319-39639-2_36

is also affected by the action of the Intelligent Transportation System, which, among other determines the applicable cycles of traffic lights.

The measurement of speed is the subject of analysis and publication of many researchers around the world. This issue is analyzed from the point of view of movement engineering, urban planning, logistics, but also the city management. In fact it is also a point of interest to institutions responsible for traffic management in urban areas, who use data on average speed to increase road safety and improve passenger comfort. A properly conducted measurement, and above all accurate analysis of the data obtained, will affect the correctness of the decisions made by the authorities.

The researches, initiated by the authors in this paper, are focused on measuring the time of vehicle journeys on the selected section of one of the main streets of Wroclaw. The aim of the chapter is to present the issues related to the analysis of measurement data from the ITS, which are aimed at supporting decision-making processes of those responsible for organizing transport in the city. For this purpose, the authors presented the basic theoretical issues of the study area and the approximate principles of operation of ITS systems. Then on the basis of quantitative analysis, the authors characterized the difficulties that may be encountered while drawing conclusions from the analysis of the collected data.

The research presented in the chapter is preliminary and is only the first stage initiated research.

2 Literature Review

The issue of traffic measurement is of interest to the researchers examining the various transportation systems [9, 12]. Traffic studies are conducted and used for very different short-term and long-term purposes. The literature emphasizes that the primary objectives of measurements and studies of traffic is to gather information on its condition. This will include the setting of such characteristics as [2]: traffic flow in vehicles per hour or per day; fluctuations in traffic during the day, week or year; traffic structure; traffic burden; traffic flows at junctions and intersections (the directional structure); speed of movement; intervals between the vehicles; loss of time when traveling along the roadway or intersection; the size of forming traffic jams; types of traffic (source, target, through, internal); road trips (route); pedestrian traffic.

Based on these data it is possible to prepare [3]:

- study of transport for planning urban transportation systems, together with the traffic forecasts;
- plans and projects of roads and crossroads;
- traffic organization, capacity assessment and traffic conditions;

- economic analyzes;
- accident analysis;
- arrangements for traffic noise levels, air pollution and other environmental impacts.

Traffic studies provide data for analysis, discussion and decisions in planning, design and management of transport solutions and controlling the traffic [6]. Their aim is to support the decision-making process targeting the best possible use of transport infrastructure and to ensure the safe and efficient movement of people and cargo.

As noted in the introduction, the most important attribute in road traffic is the speed of vehicles [14]. Speed is an essential parameter in describing traffic conditions in the transport network. This parameter can be used to estimate the quality of travelling in the area. Moreover, with the development of Intelligent Transport Systems, vehicle speed and vehicle flow play an important role in advanced traffic management systems. The speed, along with other parameters, such as traffic flow, provide information about the current situation in the transport network [1]. Traffic speed data is one of the most important information sources for ITS, Advanced Traveler Information Systems (ATIS), and Advanced Traffic Management Systems (ATMS) [16].

Previous studies [3, 8, 15] indicate that the choice of speed by drivers in a free traffic is influenced by 6 sets of factors: (1) the characteristics of vehicles; (2) the characteristics of drivers; (3) the characteristics of the road and its environment; (4) weather conditions; (5) the legislation concerning speed limits and penalties for exceeding the speed limits; (6) the exact moment, the year, the season and time of the day.

In the case of researches carried out in city traffic, one should take into account additional factors that can affect the measured speed.

Among the different speed-related parameters known in the literature, average speed is very often used as measure for safe driving, mainly because elevated crash risk and severity have been related to an increase in average speed [4, 13]. It should also be noted that the formulas given in the literature to determine the average speed of vehicles, predominantly refer to vehicles in free traffic (e.g. Kempa's, Szczuraszek's and Kempa's equations) [6].

Some of the authors [5, 11] also emphasize that the use of aggregate statistics fails to recognize the probability distribution of the individual observed values. The apparent improvement in explaining the variation of the parameter is shown by the aggregation speed data, which may reduce the individual extreme values and overestimate the consistency and safety level. Therefore, individual speed profiles should be examined. Consequently, aggregate values also should be analyzed in order to find out if the difference between the individual and aggregate values is statistically significant.

3 The Use of Intelligent Transportation Systems in Traffic Management

Name of the Intelligent Transportation Systems means systems which represent a broad collection of diverse technologies (telecommunications, information technology, automatic and measuring), and management techniques for transport purposes in order to protect the lives of road users, improving the efficiency of the transportation and the protection of natural resources [17]. Traffic management using ITS technologies is divided into two subsystems: urban traffic management and traffic management on highways. Due to the scope of the study authors attention will be focused only on the first subsystem.

In urban traffic management most often used systems are [10]: (1) traffic management systems in the network of streets; (2) the automatic traffic surveillance; (3) the automatic charge collection.

These systems are used on a large scale in cities and urban areas. Also, the city, which is the subject of presented research is fitted with a system for ITS.

The main task of ITS is to improve traffic and safety of cars and public transport vehicles [18]. The main functional areas of the ITS system are [7]:

- the possibility of a dynamic and adaptive traffic management, responding to changing traffic conditions at any real-time;
- simultaneous traffic control at various levels: prioritizing public transport, liquefaction vehicular traffic, improving of openings of pedestrian crossings, crisis management;
- the ability to collect, process and distribute information to support the decision of the residents in the planning and implementation of the travel services and decisions on categorizing and prioritizing remedial action and maintenance of urban infrastructure.

These functions are realized by the system by continuously gathering information about the situation on the roads and intersections, by the infrastructure permanently embedded in the city. The data used by the ITS system to analyze the situation on the road, relate mainly to traffic at crossroads covered by its operation. Collected in the database values are obtained through detectors identifying and counting the vehicles, mounted next to the traffic lights at crossroads.

4 Specificity and Problems of Measurement Data Analysis

Customizing traffic lights by the Intelligent Transportation Systems is based on automatically collected and processed data. They therefore contain arising in such circumstances outliers and this can lead to undesirable settings. Knowing the way of collecting measurement data on the main roads of Polish cities, prior to their analysis, one can distinguish the expected reasons for outliers. These reasons include:

- taking into account by the system, emergency vehicles and other with a speed higher than allowed, whose traveling time will be significantly lower than the average values of the group,
- vehicles stop on the route (e.g. as a result of a collision or crash) or deviating from the route and returning after some time, their time of travel will be longer.

Analysis of traveling time of cars crossing the sections of roads in urban areas, poses many difficulties. Facilitation for inferences about a set data is the ability to match them to a known probability distribution. For this reason, the authors have made an attempt to find a universal method of outlier data rejecting, which allows fitting of the remaining data to a known probability distribution.

The tests were conducted for two groups of measuring data, collected for journey times on the same stretch of road in the same direction during morning and afternoon rush hours. A summary of the raw data containing all the measurements from mornings and afternoons rush hours are shown in Table 1 and Fig. 1.

Because of the predicted outliers and inability to adjust the raw data into any of the known statistical distribution, authors have made attempts to reject outliers in such a way as to obtain the universal method of data filtering.

Most of the known methods of rejecting outliers is used for normal distributions, but in the examined case measurement data did not show this pattern. Therefore, attempts have been made to find methods tailored to the individual specificities of the phenomenon.

Table 1 A summary of the raw data (time of traveling through the investigated section)

	Morning	Afternoon
Sample size	687	471
Average	1023	1102
Standard deviation	857	1412
Max value	9549	9645
Min value	360	374
Median	848	550

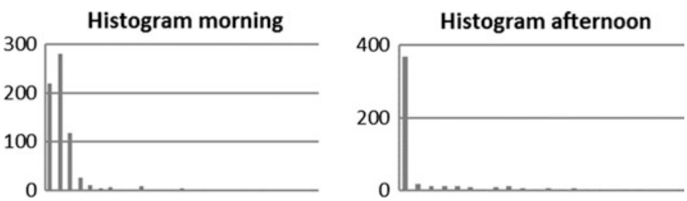

Fig. 1 Histograms transit times through the tested section at the morning and afternoon rush hours

The first attempt was to reject the maximum passes time until the median value was 0.95 times the average. Summary of the left data is shown in Table 2 and Fig. 2.

The histogram for the afternoon rush hours can be considered similar to the log-normal distribution, but the data obtained using the same method for the morning rush hours, do not allow adaptation to any known statistical distributions. Therefore, the result was not considered satisfactory.

A second attempt the rejection of outliers was based on the exclusion of times of travel at a speed of less than 0.33 of the maximum speed. The results are shown in Table 3 and Fig. 3.

The conclusions from the rejecting outliers are similar to previous ones. The results obtained in the afternoon rush hours differ significantly from the results of the morning. The histogram for the morning rush indicates the possibility of the coexistence of two groups of overlapping data. This phenomenon however has not been recorded for the afternoon rush hours.

Table 2 A summary of the data, for median 0.95 times the average

	Morning	Afternoon
Sample size	660	386
Average	363	141
Standard deviation	2684	1273
Max value	360	374
Min value	881	550
Median	837	524

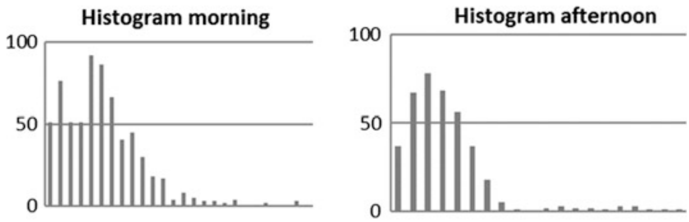

Fig. 2 Histograms transit times through the tested section at the morning and afternoon rush hours, for median 0.95 times the average

Table 3 A summary of the data, exclusion V < 0.33 Vmax

	Morning	Afternoon
Sample size	505	380
Average	727	540
Standard deviation	192	116
Max value	1079	1098
Min value	360	374
Median	848	550

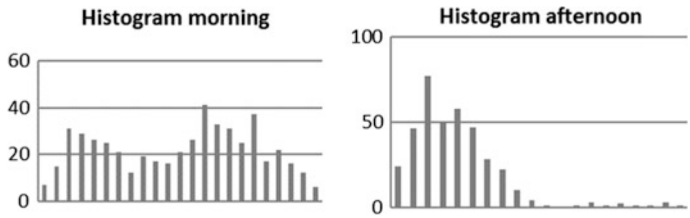

Fig. 3 Histograms transit times through the tested section at the morning and afternoon traffic rush hours, exclusion V < 0.33 Vmax

Due to the fact that the authors wanted the queried data rejection method to be universal, it cannot be regarded as satisfactory, so there have been no attempts of implementing for other data.

Another attempt differ from previous ones, this time rejected has been both the highest and lowest results. The authors rejected travel times for speeds less than 0.75 of average speed and above 1.5 average speed. The obtained data are shown in Table 4 and Fig. 4.

The proposed method of discarding the lower and upper values, has made the two histograms difficult to interpret. For this reason, also this method could not be considered satisfactory.

Another proposed criterion for rejecting data become the longest and shortest forecasted travel time. The measurements of vehicles traveling above the permitted speed (50 km/h) and below 20 km/h (adding a possible waiting time at intersections

Table 4 A summary of the data, exclusion 0.75 V max < V < 1.5 V max

	Morning	Afternoon
Sample size	381	345
Average	791	532
Standard deviation	139	76
Max value	1042	764
Min value	523	411
Median	804	525

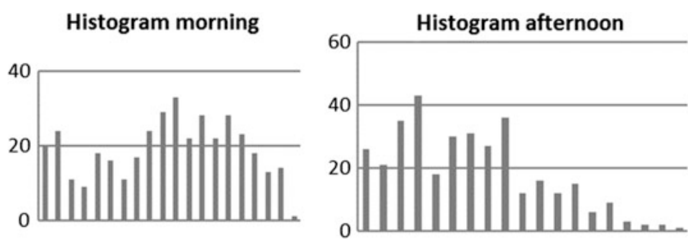

Fig. 4 Histograms transit times through the tested section at the morning and afternoon rush hours, exclusion 0.75 Vmax < V < 1.5 Vmax

Table 5 A summary of the data, exclusion 20 km/h < V < 50 km/h

	Morning	Afternoon
Sample size	597	360
Average	821	566
Standard deviation	250	150
Max value	1391	1323
Min value	415	416
Median	814	533

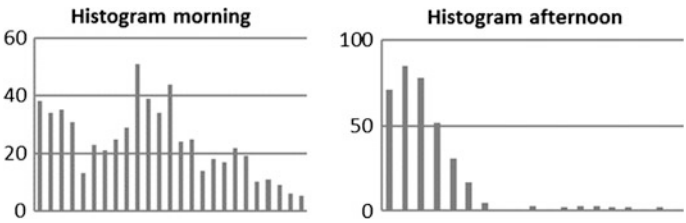

Fig. 5 Histograms transit times through the tested section at the morning and afternoon rush hours, exclusion 20 km/h < V < 50 km/h

with traffic lights by 30 s) has been rejected. Obtained in this way results are presented in Table 5 and Fig. 5.

These data confer similar to those occurring in the prior, difficulties of interpretation.

On the basis of this chapter the results of rejecting outliers tests, it can be stated that the analysis of data from measurements of travel times is a extremely complicated and demanding task and requires to take into account many factors. The very characteristics of the examined road sections undermines the credibility of some measurements. Registered traveling time, concerned road traffic on one of the main arteries of Wroclaw, connecting the western part of the city with the city center. On the investigated road is a provincial hospital, for this reason there are often recorded ambulance rides, whose speed exceeds the legal limit. This can affect an average speed of travel of all measured vehicles. However greater important can be attributable to the facilities, which are locate in the vicinity of the test route and may significantly understate measured travel time. These are—one of the largest shopping malls in Wroclaw and other shopping centers, business centers (collecting banks and other service companies), two large health clinics and the Municipal Social Centre. These objects are a place of temporary stop, often not exceeding 60 min. This causes the re-inclusion in the urban traffic of a vehicle in the studied time window, and the travel time is extended by the temporary stop.

The solution to this problem could be comparation each subsequent measurement time with the average travel time. This average travel time should be determined by the method of moving average with n values. The parameter n should be dependent on the frequency of the occurrence of the next recorded measurements (a large number of measurements recorded in the adopted unit of time will increase the value of n; a small number of recorded measurements will reduce the number of n parameters). However the proposed method is characterized by a large computational complexity. High complexity of the method requires a properly prepared calculation algorithm, supported by appropriate IT tool. The possibility of its use will be the subject of further study of the authors.

5 Conclusions

Large urban agglomerations invest in ITS in order to improve traffic flow and transport safety. Analyses of data collected by these systems are intended to support decision-making processes of people responsible for the organization of transport in cities. However, as shown in the paper: the historical data analyzed with a simple statistical methods, can lead to erroneous conclusions. It is necessary to carry out detailed analyzes that will allow to clean up the original data series. Only this way it is possible to select the appropriate data for further quantitative analysis. Additionally when given only historical data from the ITS system, the analyst is unable to carry out a credible inference. For a complete analysis one needs additional information, e.g.: weathering, random events (e.g. car accidents), events planned (e.g. a narrowing streets due to road works). Excluding the knowledge of conditions which accompany measurements, created plans for perfecting the urban traffic management may be based on false assumptions.

In the beginning of this chapter it was pointed out that the presented results are introducing the studies initiated by the authors. The purpose defined at the outset there was a presentation of the model analysis of the speed of vehicles based on historical data. The intention of the authors was to present the complexity and difficulty of quantitative analyzes based on primary data obtained from the ITS. For this reason, as the summary it should be stated that: (1) it is impossible to give a general statistical distribution that matches the analyzed historical data; (2) preliminary analysis of data should take into account the process of cleaning the raw data of erroneous or incomplete measurements; (3) the correctness of the performed inference should take into account the conditions associated with the measurement; (4) the person responsible for the analysis should be familiar with the characteristics of the assessed routes and common behavior patterns of drivers on the examined sections.

Therefore, the next stage of the researches is to develop a model procedure for treatment of primary data, enabling to obtain a satisfactory level of reliability of the data analysing for supporting traffic management in the city.

References

1. Birr, K., Jamroz, K., Kustra, W.: Travel time of public transport vehicles estimation. Transportation Research Procedia **3**, 359–365 (2014)
2. Datka, S., Tracz, M.: Przewodnik do ćwiczeń z inżynierii ruchu. Politechnika Krakowska, Kraków (1974)
3. Datka, S., Suchorzewski, W., Tracz, M.: Inżynieria ruchu. WKŁ, Warszawa (1999)
4. European Commission: MASTER Managing speeds of traffic on European roads (No. RO-96-SC.202). European Commission, Finland (1999)
5. Freedman, D.A.: Ecological inference and the ecological fallacy. Int. Encycl. Soc. Behav. Sci. **6**, 4027–4030 (1999)
6. Gaca, S., Suchorzewski, W., Tracz, M.: Inżynieria ruchu drogowego: teoria i praktyka. WKŁ, Warszawa (2014)
7. Hadas, A.: Zastosowanie zintegrowanego systemu transportu w aglomeracji miejskiej. Przegląd Komunikacyjny **4**, 8–13 (2015)
8. Kempa, J., Szczuraszek, T.: Design speed, actual speed and road designing. In: 2nd International Symposium on Highway Geometric Design. Transportation Research Board, Road and Transportation Research Association, Mainz, German (2000)
9. Kierzkowski, A., Kisiel, T.: The simulation model of logistic support for functioning ground handling agent, taking into account the probabilistic time of aircrafts arrival. In: Carpathian Logistics Congress—Congress Proceedings, pp: 463–469 (2013)
10. Oskarbski, J., Jamroz, K., Litwin, M.: Inteligentne systemy transportu– zaawansowane systemy zarządzania ruchem. www.pkd.org.pl/pliki/referaty/oskarbski,_jamroz,_litwin.pdf. Accessed January 2016
11. Park, Y., Saccomanno, F.: Evaluating speed consistency between successive elements of a two-lane rural highway. Transp. Res. Part A: Policy Pract. **40**, 375–385 (2006)
12. Restel, F.J.: Train punctuality model for a selected part of railway transportation system. In: Safety, Reliability and Risk Analysis: Beyond the Horizon—Proceedings of the European Safety and Reliability Conference (2014)
13. Shinar, D.: Traffic Safety and Human Behavior. Elsevier, Oxford (2007)
14. Szczuraszek, T.: Prędkość pojazdów w warunkach drogowego ruchu swobodnego. Komitet Inżynierii Lądowej i Wodnej PAN, Warszawa (2008)
15. Watanatada, T.: A Model for Predicting Free-flow Speed Based on Probabilistic Limiting Velocity Concepts: Theory and Estimation. Transportation Research Record, vol. 1091 (1986)
16. Zhang, J., Wang, F.Y., Wang, K., Lin, W.H., Xu, X., Chen, C.: Data-driven intelligent transportation systems: a survey. IEEE Trans. Intel. Transp. Syst. **12**(4), 1624–1639 (2011)
17. www.itspolska.pl. Accessed January 2016
18. http://zdium.wroc.pl/view/document/185. Accessed November 2015