Keywords: driving cycle; electric buses; stochastic methods; computer simulation

Andrzej CZEREPICKI[1]*, Maciej KOZŁOWSKI[2]

APPLICATION OF THE BAYESIAN INFERENCE METHOD TO SYNTHESIZE URBAN DRIVING CYCLE SPEED SCHEDULES USING MEASURED DATA

Summary. Standardized driving cycles are widely used to simulate real-world vehicle operating conditions. They are used for estimating vehicle ranges, evaluating emissions, and designing powertrain characteristics. Recently, due to the intensive electrification of urban public transport fleets, estimating electric energy consumption has gained significant importance. This process requires up-to-date, validated driving cycles developed from real-world operating condition measurements. This paper presents an original method for synthesizing the driving cycle based on a stochastic approach using Bayesian inference. It first determines probability distributions of vehicle speed for the entire line. Next, all trips for the line are divided into segments of uninterrupted movement. A probability distribution of speed is also constructed for each of these segments. Then, quality fit indicators are calculated for each segment. Segments that support the hypothesis of the total distribution with maximum likelihood are then selected using Bayesian inference. A driving cycle for the selected line is synthesized by combining several segments with the highest degree of fit. The condition for applying the method is having a large dataset of measurements obtained during actual trips. The method can be effectively implemented computationally, ensuring high accuracy in reflecting the driving conditions in a selected area. The method has low computational requirements and a high capacity to adapt to changing operational conditions. This paper presents the results of applying this method to three bus lines in Warsaw and uses publicly available traffic data. The proposed method can be used for various applications, including estimating energy consumption for route travel, assessing the impact of traffic management changes on driving smoothness, and determining the load on electric bus engines and batteries.

1. INTRODUCTION

A driving cycle represents the typical driving pattern of a vehicle under specified conditions [1] and can be defined as a function of vehicle speed over time. Driving cycles simulate real-world vehicle operating conditions for diesel and electric vehicles [2]. They are used to estimate vehicle range under given conditions, fuel or electric energy consumption over a route, air pollution estimation, calculation of vehicle powertrain characteristics for matching to operating conditions, etc. Standardized driving cycles are developed by research institutions in various countries and may consider the geographical features of the location, climatic conditions, regional regulations, and types of vehicles (passenger cars, trucks, public transport). They may also pertain to urban, rural, or mixed traffic conditions. However, they do not always reflect specific real-world conditions [3].

[1] Warsaw University of Technology, Faculty of Transport; Koszykowa 75, 00-662 Warsaw, Poland; e-mail: andrzej.czerepicki@pw.edu.pl; orcid.org/0000-0002-8659-5695
[2] Warsaw University of Technology, Faculty of Transport; Koszykowa 75, 00-662 Warsaw, Poland; e-mail: maciej.kozlowski@pw.edu.pl; orcid.org/0000-0002-1068-8991
* Corresponding author. E-mail: andrzej.czerepicki@pw.edu.pl

Driving cycles can be constructed as a combination of motion phases: acceleration, cruising at a constant speed, deceleration, and stopping. A commonly used example is the New European Driving Cycle [4]. In practice, however, phase transitions occur much more rapidly than assumed in the cycle, and accelerations and decelerations in congested urban traffic are more aggressive. Therefore, in recent years, the New European Driving Cycle has been largely replaced by the Worldwide Harmonized Light Vehicles Test Cycle [5].

In the past decade, significant developments have occurred in electric-powered transportation, including in the realm of urban public transport, where a common trend is to replace diesel buses with their electric counterparts [6-8]. Zero-emission public transport is part of the EU's sustainable development strategy. The "Electromobility and Alternative Fuels Act" in Poland supports public transport vehicle electrification [9]. The first electric buses appeared on the streets of Warsaw in 2017.

Electric vehicles require the development of actual driving cycles due to the specific dynamics of electric bus propulsion. On the other hand, it is necessary to consider indirect factors such as the ability of the charging infrastructure to replenish energy consumed, as well as timetables setting time conditions [10]. Developing current driving cycles is an important research topic because accurately determining the driving cycle reduces electricity consumption and provides recommendations for drivers to follow [11]. Predicting energy consumption on a designated route facilitates better charger placement and power selection. Estimating energy flows in an electric vehicle's power system will also allow the vehicle's battery life to be assessed over extended use. On the other hand, not considering driving cycles that reflect current traffic conditions may result in operational problems [12].

Methods for constructing driving cycles employ advanced mathematical tools, including statistics, probability theory, and Markov chains, combined with computer modeling and neural networks. The approach presented in [13] is based on calculating a set of driving characteristics using a particular measuring device mounted on the vehicle. The trips were divided into segments according to the idle criterion. The need to equip the car with such a device limits the application of this method. In [14], a probabilistic approach focused on the multidimensionality of driving cycles, considering, for example, different weather or road conditions. Markov chains define the conditions and probabilities of transitioning to the next driving stage. Paper [15] introduced constraints to the evolution of Markov chains, which significantly accelerated, generating a representative driving cycle and applying self-adaptation mechanisms. Article [16] presented a prediction method based on a recursive, self-learning Markov chain. This further reduced the calculation time of the driving cycle by adjusting prediction parameters to achieve optimal performance.

Work [17] focused on the statistical analysis of the dependencies of driving cycle characteristics, including transition probabilities in the Markov matrix, based on travel times. In this case, the cycle model was limited to speed and acceleration. Paper [18], like this article, considered the specifics of urban bus driving cycles. It focused on analyzing the characteristics of driving cycles depending on the route features, using linear regression in the driving cycle prediction model. The macroscopic approach involves constructing a driving cycle based on route profile information obtained, for example, from geographic information systems. Article [19] applied the Monte Carlo method and the Markov chain to filter candidate driving cycles to select the most representative one. Among the remaining works, this article is characterized by an extensive set of measurement data collected over two years and a focus on electric vehicles. The microscopic approach relies on measurement data collected during trips on designated routes.

The method proposed in this paper for constructing a bus driving cycle uses probabilistic methods. The method uses a standard vehicle motion recorder, making it simple to collect a large amount of data from many vehicles simultaneously, as it does not require other specialized measuring devices. The novelty of the method lies in its application of Bayesian inference [20], [21], which, when combined with a large set of accurate measurement data, allows for the efficient implementation of the method in a computational form. This ensures high accuracy when reflecting actual traffic conditions for the entire selected area, as well as low computational requirements and a high capacity for adaptation to changing traffic conditions by incorporating real-time measurements.

2. METHOD FOR DETERMINING A REPRESENTATIVE DRIVING CYCLE

Let the non-empty finite set $L = \{L_1, L_2, \dots L_n\}$ represent all the analyzed public transport lines serviced by buses, selected for constructing the representative driving cycle. Each line $L_i \in \{L\}$ is represented as a finite directed acyclic graph $S = \{S_1, S_2, \dots S_k\}$, where each node S_j represents a stop on the route and can be represented as an ordered pair $(name, next)$, where $name$ is the name of the stop, and $next$ is a pointer to the next stop. Stop S_1 is the beginning of the route, stop S_k is the end of the route, and $\forall m \in [1..k-1]$ the conditions $S_{m+1} = S_m.next$ and $S_k.next = null$ are satisfied.

The segment of the route L_i between two adjacent stops S_m and S_{m+1} $(1 \leq m \leq k-1)$ is denoted as D_m and represented by location functions dependent on time – specifically, $D_m(t) = \{lat(t), lon(t)\}$, where $lat(t)$ is latitude and $lon(t)$ is longitude. The stops may belong to more than one route *i.e., $L_a \cap L_b \geq \emptyset$ for $1 \leq a < b \leq n$). All route segments form the set $D = \{D_1, D_2, \dots D_p\}$, where segments may be common to two or more routes (i.e., $D_c \cap D_d \geq \emptyset$ for $1 \leq c < d \leq p$).

Speed is a discrete function, where each value $v_i = v(t_i)$ represents the current speed of the vehicle. Measurements of the vehicle's geographic position at times $t_i \in \{t_1, t_2, \dots t_r\}$ are taken at equal time intervals Δt, so $t_x = t_1 + (x-1)\Delta t$.

Fig. 1 shows the speed profile of one of the selected segments between stops.

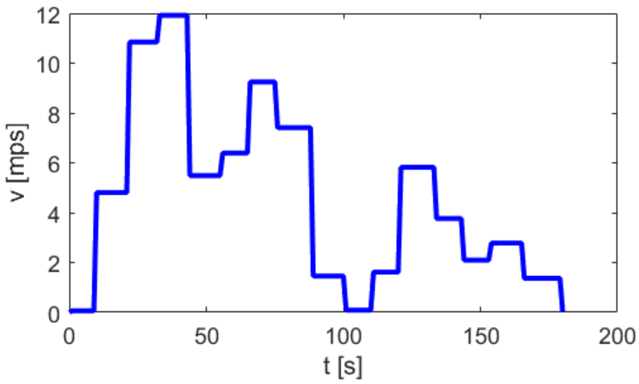

Fig. 1. Speed profile of one of the segments between stops

The drive of a segment can be characterized probabilistically, for example, by providing the probability distribution of speeds or by defining a stochastic matrix of conditional probabilities. In this case, the probability p_k that the speed v will fall within the interval $(v_k; v_{k+1})$ with an allowable error Δp_k is determined based on the relative travel time:

$$p_k = \frac{t_k}{t_n}, \ \Delta p_k = \sqrt{\frac{p_k(1-p_k)}{t_n}} \tag{1}$$

where t_k is the time spent driving at a speed within the specified segment, and t_n is the total travel time for the segment.

Fig. 2 shows an example of the speed probability distribution for a segment. The classification was conducted as follows: stop – no change in vehicle location for at least three seconds; speed class – discretization of speed in increments of 5 km/h within the 0–80 km/h range. Additionally, stops are distinguished between those occurring at designated stops and those occurring along the route.

Trips can also be characterized in terms of driving stages. Here, the phases can be distinguished as follows: 0 – stop, 1 – cruising at constant speed, 2 – acceleration, 3 – deceleration. The probability distribution associated with the analyzed segment is shown in Fig. 3.

The original method of synthesizing a representative driving cycle is intended to select a non-empty set of road segments whose chosen probabilistic characteristics best reflect the exact features obtained for all trips. The number of segments constituting the representative driving cycle can be selected depending on the requirements (e.g., the trip duration).

The practical implementation of the method requires an analysis of multiple journeys of road segments $D_i \in \{D\}$ belonging to different bus lines $L_i \in \{L\}$. For this purpose, a two-dimensional matrix T was created. The matrix rows specify ordinal numbers of road segment indices, while the matrix columns define categories of speed classes. A sample matrix fragment is shown in Table 1, where the classes are denoted as follows: 0 – stops, 5–40 speed classes in increments of 5 km/h (since speed classes 45, 50, ..., 80 are omitted in the table, the value of T_i may be greater than the sum of the rows of the table), Lp – ordinal number, k – class, and T_i – total travel time for segment number i.

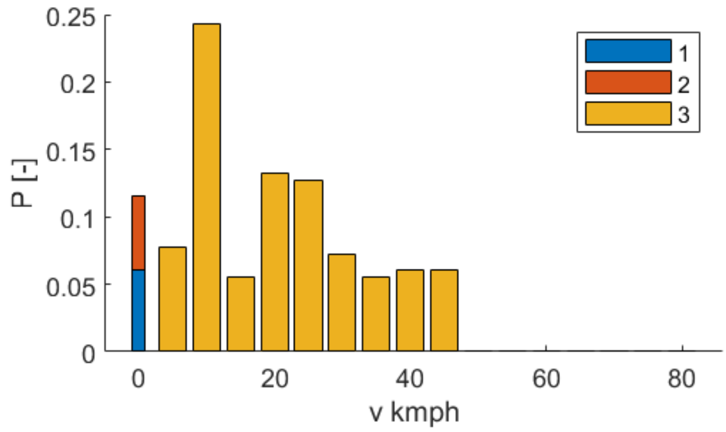

Fig. 2. Probability distribution of speeds on the selected route segment. The classes are numbered as follows: 1 – stops at bus stops, 2 – stops outside bus stops, and 3 – speed increments in steps of 5 km/h

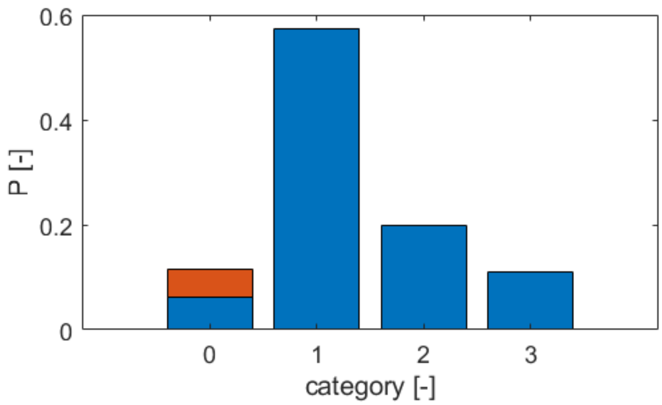

Fig. 3. The probability distribution of driving stages associated with the speed profile

Table 1

Matrix of segment travel times (fragment)

Lp/k	0	5	10	15	20	25	30	35	40	T_i
1	11	10	30	2	15	16	0	11	0	132
2	17	0	0	11	0	0	0	7	13	69
3	1	13	0	0	46	0	8	0	32	167
4	59	9	21	22	0	0	12	0	0	123
5	57	85	29	46	0	0	34	0	12	263
6	44	22	0	11	1	14	10	0	10	112
7	106	25	29	43	14	53	11	10	0	395
..	…	…	…	…	…	…	…	…	…	…

By summing the elements of matrix T along rows i, we obtain the travel time distributions for segments as:

$$T_i = \sum_{j=1}^{k} (\tau_i)_j \tag{2}$$

where: τ_{ij} is the travel time at speed belonging to class j on segment i (an element of the matrix shown in the table), and T_i is the travel time for segment i.

By summing the elements of matrix T along columns j, we obtain the speed count distributions. Dividing by the total travel time allows us to calculate the total probability distribution as:

$$P(V_j) = \frac{\sum_{i=1}^{n} (\tau_j)_i}{\sum_{i=1}^{n} T_i} \tag{3}$$

The resulting total probability distribution from all considered segments is shown in Fig. 4. Due to the large size of the analyzed data (the matrix describes ~20,000 segments), the percentage error of the probabilities is at most 2%. Fig. 5 illustrates the distributions of the following movement stages: stop, cruising at constant speed, acceleration, and deceleration.

Conditional probability distributions of achieving speed on selected segments $P\ (velocity\ |\ segment)$ are obtained by dividing the element of the matrix τ_{ij} by the travel time of the segment T_i:

$$P(V|i)_j = \frac{\tau_{ij}}{T_i}\ , j = 1, \ldots, k \tag{4}$$

where: $P(V|i)_j$ is the probability of achieving speed in class j on segment i. Based on this, the conditional probability distribution $P(segment\ |\ velocity)$ is determined by Bayes' theorem of conditional probability:

$$P(i|V)_j = P(V|i)_j * \omega_j \tag{5}$$

where: $P(i|V)_j$ is the probability that a trip took place on segment i if movement occurred at speed class j, and ω_i is the weight of probability.

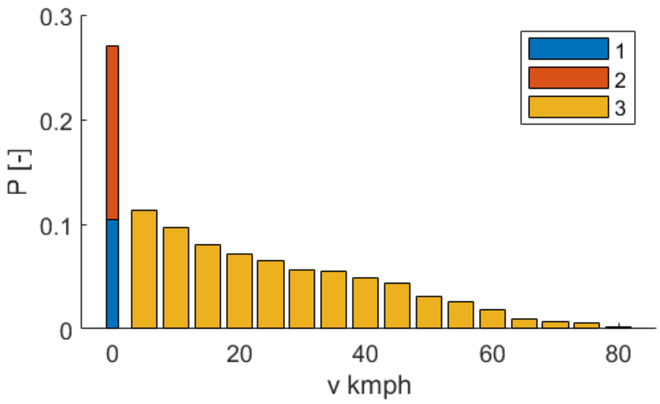

Fig. 4. The overall probability distribution of speed for a selected route (112)

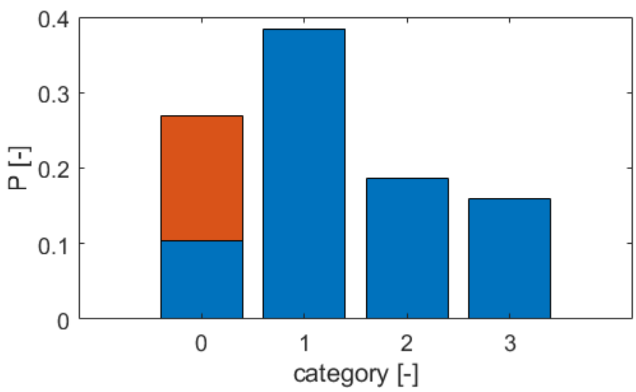

Fig. 5. The driving stages for the selected route (112)

In this case, the weight of probability determines the relative travel time of the segment:

$$\omega_i = \frac{T_i}{\sum_{i=1}^{n} T_i} \tag{6}$$

This principle can be extended to the total speed probability:

$$P(V) = P(V|j) \cdot OR^{-1} \tag{7}$$

However, currently specified probabilities will take on new meanings: OR – the odds ratio that there is a higher chance of the event occurring in the study group than in the reference group, $P(V|j)$ – a priori hypothesis that the total probability distribution can be described by the distribution $P(velocity \mid segment)$, and $P(V)$ – posterior distribution of total probability. This is a case in which Bayesian inference is applied based on credibility. The OR is the weight of Bayesian inference. A value of OR≈1 means that the chance of an event occurring in both groups is similar. If OR<1, there is a lower chance of the event occurring in the study group than in the reference group, whereas the opposite pattern is indicated by OR>1. Based on the last dependency:

$$OR = \frac{P(V|j)}{P(V)}; \quad \Delta OR = \sqrt{\left(\frac{P(V|j)\cdot\Delta P(V)}{\left(P(V)\right)^2}\right)^2 + \left(\frac{\Delta P(V|j)}{P(V)}\right)^2} \tag{8}$$

In this case, OR is the total speed probability distribution ratio to the distributions on individual segments. Table 2 shows the distribution of ORs describing the degree of support for the hypothesis $P(V)$ using the data $P(V|j = 7 \ldots 14)$ from Table 1.

A similar relationship can be used to determine a group of speed distributions on several – for example, three or four – different road segments, representing the total probability distributions most credibly. In this case, the OR distribution can be determined by the formula:

$$P(V|i)_{j=row} = \frac{\sum_{j=row}(\tau_i)_j}{\sum_{j=row}(T_i)_j}; \quad (OR_i)_{j=row} = \frac{P(V|i)_{j=row}}{P(V_i)} \tag{9}$$

Table 2

The OR distribution based on data from Table 1

Lp/k	0	5	10	15	20	25	30	35	40
1	0.80	0.67	2.36	0.19	1.58	1.85	0	1.52	0
2	1.54	0	0	1.99	0	0	0	1.86	3.85
3	0.06	0.69	0	0	3.82	0	0.84	0	3.92
4	3.30	0.65	1.77	2.24	0	0	1.72	0	0
5	1.32	2.86	1.14	2.19	0	0	2.28	0	0.93
6	3.10	1.74	0	1.23	0.12	1.91	1.57	0	1.83
7	1.63	0.56	0.76	1.36	0.49	2.05	0.49	0.46	0
…	…	…	…	…	…	…	….	…	…

For example, to calculate the OR for segments indexed 1, 3, 5, and 7 in the speed class of 10 km/h from Table 2, we determine the following: $(\tau_{10})_1 = 30$; $(\tau_{10})_3 = 0$; $(\tau_{10})_5 = 29$; $(\tau_{10})_7 = 29$. Thus, $\sum_{j=row}(\tau_i)_j = 88$. Similarly, $\sum_{j=row}(T_i)_j$=132+167+263+395=957. Therefore, $P(V|i = 10)_{j=1,3,5,7} = 0.0920$. Since $P(V_{i=10}) = 0.0964$, the OR based on the data from the selected four segments in the 10 km/h speed class is 0.953, which supports the hypothesis. The calculated ORs for segments in all speed classes are presented in Table 3.

The objective function should describe the condition that the odds of an event occurring in one group are close to the odds of it occurring in the other group. The evaluated condition can be defined as follows:

$$F_i = \left[(OR_i)_{j=row} - 1\right] < \alpha \tag{10}$$

where: α is the maximum deviation (e.g., α = 0.1), and F_i is a logical function for the checked speed classes. The α value characterizes the maximum permissible deviation of the OR from 1. In the example, the value of $\alpha = 0.1$ was chosen. Increasing α enhances the chances of a segment qualifying for cycle representation. If the value of α is too small, it can make it difficult to find segments that meet the hypothesis of similarity of distributions; if it is too large, it can reduce the accuracy of calculations.

Table 3

The OR coefficients for fitting the speed distribution of a randomly selected group of road segments and the evaluation index function

k	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**
OR(1,3,5,7)	0.99	1.21	0.95	1.17	1.06	1.08	0.95	0.39	0.92
F	1	0	1	0	1	1	1	0	1
k	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	
OR(1,3,5,7)	1.20	2.96	0.95	1.09	1.16	0	2.22	0	
F	0	0	1	1	0	0	0	0	

The quality indicator $F_{indicator}$ is the sum of the logical values:

$$F_{indicator} = \sum F_i \tag{11}$$

In the example presented above, its value is 8. The largest value of $F_{indicator}$ is the best, as it shows that the corresponding segment best reflects the entire route in all speed intervals. The proposed method uses $F_{indicator}$ to find segments that best represent driving cycles. If the calculations result in more groups being assessed identically, the quality of fitting the group can be further distinguished by calculating the sums of odds for individual speed classes.

3. ACQUIRING DATA FOR ANALYSIS

Public transport buses in Warsaw are equipped with motion recorders. These recorders save various vehicle movement characteristics (current vehicle speed, the geocoordinates of the vehicle, and the local time of the vehicle), measured using sensors. Measurement data is supplemented with course description data: the number of the serviced public transportation line, vehicle registration number, and the shifts of the drivers conducting the route. A data package from each bus is sent to a remote server via a modem at specified time intervals (every 10 seconds). The format of the data frame is shown in Table 4.

The Municipal Transport Authority information system (Fig. 6) contains vehicles and schedules databases. Vehicles record their driving parameters and send them to the server as a data package using the format shown in Table 4. Timetables are stored in the Schedules database. A registered user (or application) can access both databases free of charge via the Web Server using the application programming interface methods. The default format for data exchange is JavaScript Object Notation. A computer application was developed to analyze data from public transport vehicle journeys by collecting cyclical traffic data and storing it in the local Dataset. In this way, a journey database was created over two years. The data was then filtered to select journeys made on the indicated routes at a specific time of year and at particular times. Next, the data was analyzed using the MATLAB software environment.

The data obtained from the carrier's programming interface reflects the dynamics of the vehicle traffic process. Due to random occurrences, such as traffic jams, breakdowns, weather conditions, and route changes, these data may only partially accurately reflect the traffic process. Therefore, the measurement data was filtered according to the prevailing traffic distributions to construct a representative driving cycle. For each analyzed urban transport route, the driving schedules in effect on the day the measurement data was obtained were retrieved. The schedule data for each urban transport line included the names of the stops, their sequence, scheduled departure times from the stops, and the route length. The filtering process aimed to exclude atypical journeys (e.g., incomplete trips, significant service delays, skipped stops due to detours) whose inclusion in the standardized cycle could have affected their quality. The schedule data was obtained from the carrier's network server.

Table 4

The format of the data frame

Route	Vehicle	CrewNo	Latitude	Longitude	Speed, [km/h]	Time
523	N12345	4	52.24015	20.93415	35	2024-01-15 12:18:27
…	…	…	…	…	…	…

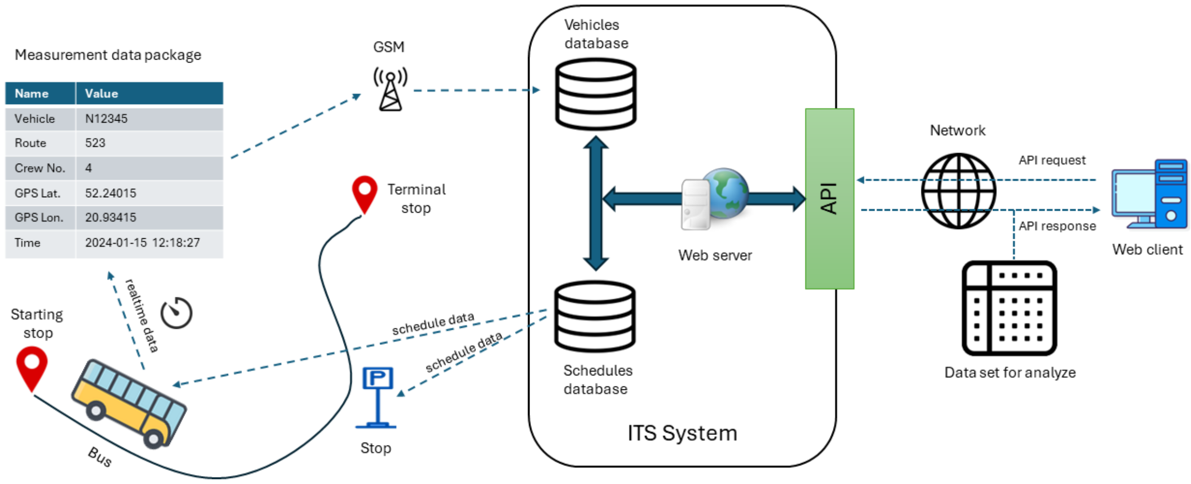

Fig. 6. The architecture of the information system used for collecting vehicle traffic data

The method for determining a representative urban cycle was presented using measurement data collected for three selected urban transport routes (Fig. 7). The routes were chosen to represent different districts of the city, including sections of highways, residential areas, bridges, streets with designated lanes for urban transport, and sections passing through the historical places of the Old Town, among others. Table 5 displays the essential characteristics of the selected routes.

The representative cycles were constructed as follows. First, the vehicle numbers servicing each route were selected. Then, each vehicle's trip from the initial to the final stop was extracted in the analyzed time interval. Each trip was then verified against the current schedules and discarded if its association with the line could not be determined. Complete trips were divided into segments based on bus station stops, resulting in uninterrupted driving segments. Each segment was characterized by specific features (distance, travel time, average speed, maximum speed, etc.). Next, speed intervals (groups) of 5 km/h were established, for which probability distributions of speed occurrence by group were constructed. Similarly to the segment data, speed probability distributions were built for the entire route. Then, using Bayesian inference, hypotheses were tested to match individual segments' speed distributions to the whole route's speed distributions. A specific number of candidate segments were identified based on the ranking of segments according to the selected objective function, whose combination formed a representative driving cycle for the selected area.

4. CALCULATION RESULTS

The first set of calculations determined the total probability distributions of speeds and driving stages for lines 112, 180, and 523. The results presented in Figs. 4 and 5 were based on calculations for line 112. We will show the results of identical calculations for the remaining lines. The speed distributions and stage probabilities for line 180 are shown in Figs. 8 and 9, respectively.

The speed distributions and driving stages for line 523 are shown in Figs. 10 and 11.

The presented total probability distributions show minimal differences. In the second set of calculations, the speed probability distributions were determined for groups of four randomly selected segments, and the ORs were calculated. In the next step, quality fit indicators were calculated. Based on these calculations, the data (in the form of sequences of selected four segments) that most strongly supported the hypothesis of the total distribution (with maximum reliability) were identified. The results are presented in the following figures.

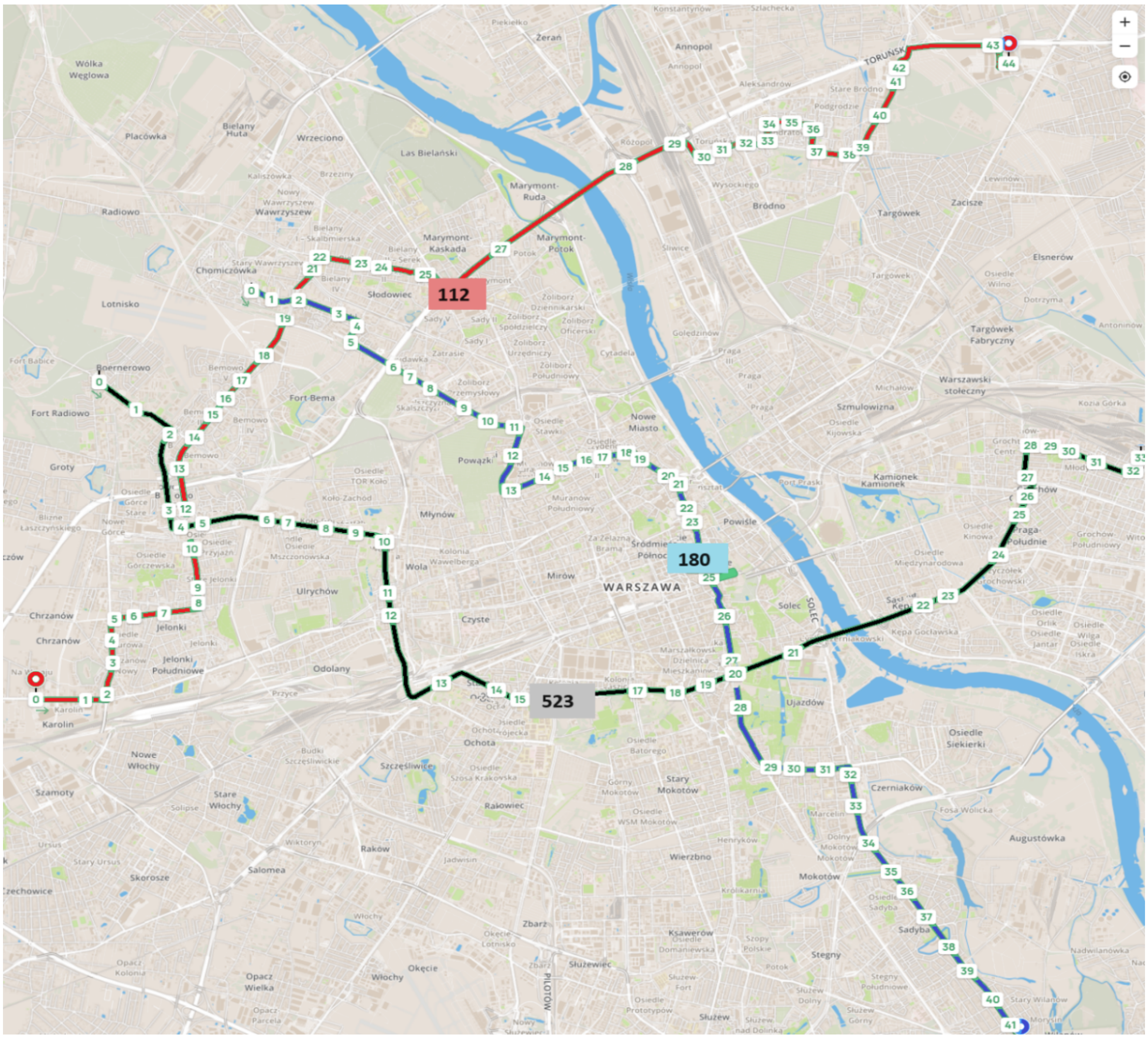

Fig. 7. Selected public transportation routes (source: [23])

Table 5

Essential characteristics of selected routes for analysis

Line	112 (regular line)		180 (regular, tourist line)		523 (express line)	
Terminal stops	Karolin – CH Marki	CH Marki - Karolin	Chomiczówka - Wilanów	Wilanów - Chomiczówka	Stare Bemowo – PKP Olszynka Grochowska	PKP Olszynka Grochowska - Stare Bemowo
Route length, [km]	23.76	25.78	19.60	19.80	20.67	20.45
Number of stops	45	48	42	45	35	33
Average number of trips per day	176		176		254	
Average duration of route operation (per day), [h]	17:34		19:26		17:18	

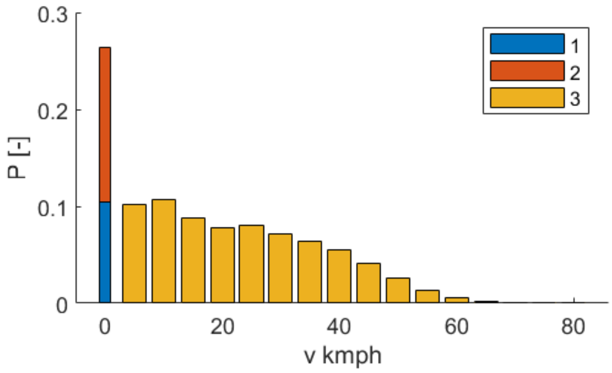

Fig. 8. Speed distributions for line 180

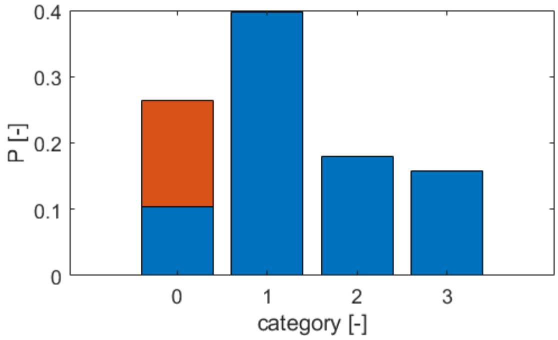

Fig. 9. Stage probability distributions for line 180

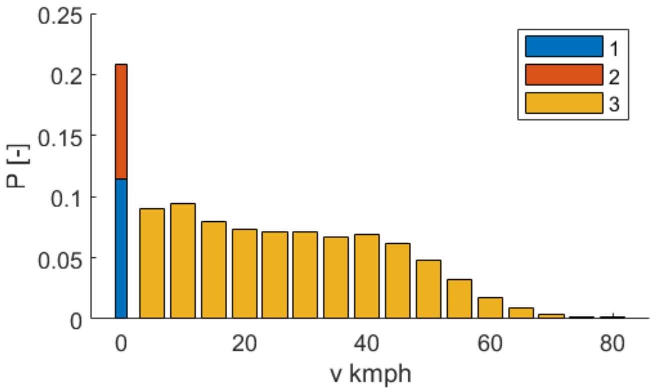

Fig. 10. Speed distributions for line 523

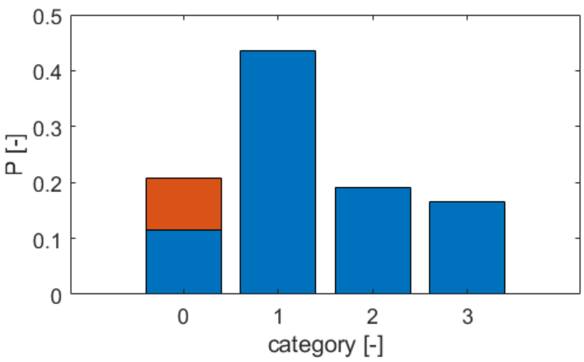

Fig. 11. Stage probability distributions for line 523

Fig. 12 shows the speed profiles of four segments of line 112, for which the quality evaluation indicator $F_{indicator} = 12$. Fig. 13 shows the histogram of the speed distribution for these segments.

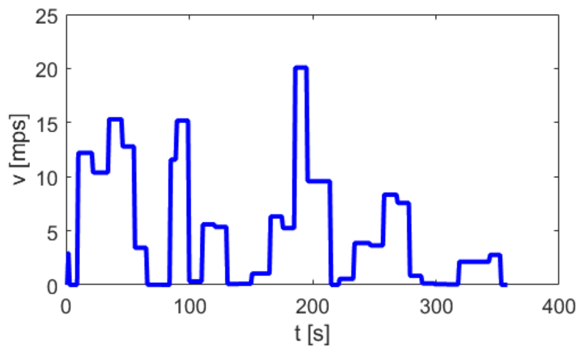

Fig. 12. Speed profiles of four segments of line 112 providing a representative driving cycle

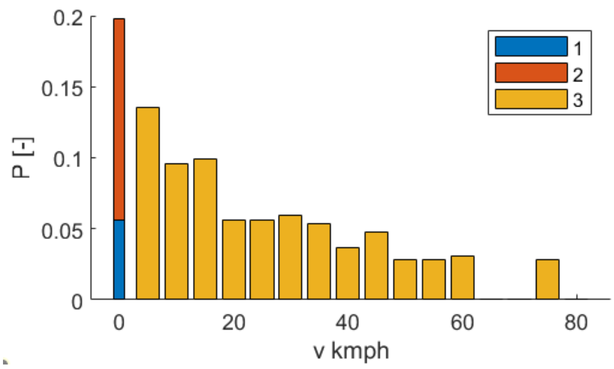

Fig. 13. Histogram of the speed distribution of four segments of line 112

Fig. 14 shows the speed profiles of four segments of line 523, for which the quality evaluation indicator $F_{indicator} = 13$. Fig. 15 shows the histogram of the speed distribution for these segments.

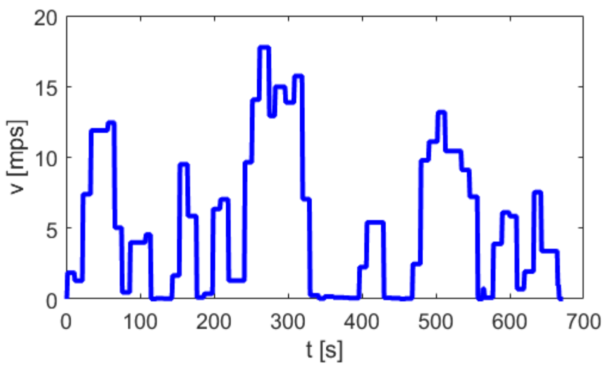

Fig. 14. Speed profiles of four segments of line 523 providing a representative driving cycle

The speed profiles shown in Figs. 12 and 14 are representative driving cycles for the selected sample routes. The method was demonstrated by randomly selecting four segments for each route whose summed quality equation indicator had the greatest significance. The order in which the selected sections were combined into a cycle was not important. The cycle length for each line was different. Route 523 (express) is characterized by longer distances between stops. Therefore, its representative driving cycle is longer. The proposed method makes it easy to obtain a cycle of the desired length. It is sufficient for each analyzed route to select the number of combined sections accordingly.

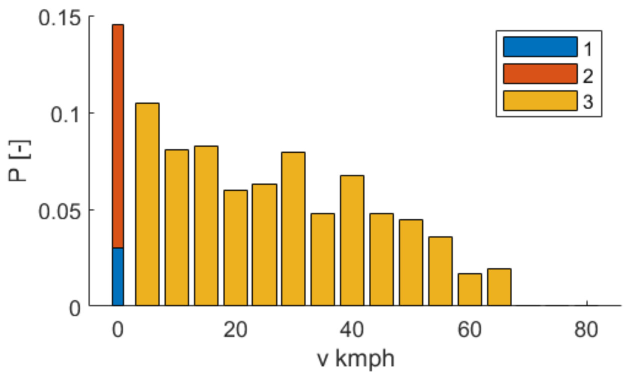

Fig.15. Histogram of the speed distribution for four segments of line 523

5. CONCLUSIONS

This article proposes a method for determining the driving cycles of city buses based on large sets of real-world measurement data utilizing Bayesian inference. This method allows a driving cycle of any duration (tailored to user needs) to be synthesized. The cycle consists of actual segments of inter-stop route sections whose selected probabilistic characteristics best reflect the overall movement on the analyzed routes.

The primary goal of developing the proprietary method was to create a Warsaw driving cycle for electric buses. This is primarily related to the intensive electrification of the public transport fleet. An analysis of standard driving cycles revealed that, due to their universality, they do not accurately reflect specific driving conditions. When the method is combined with an electric vehicle model adapted to the route profile [7], it provides a comprehensive solution for selecting the characteristics of an electric vehicle to service communication lines in a designated area. Furthermore, considering the approach presented in [10], the method serves as one of the pillars in planning the infrastructure for servicing electric vehicles at charging stations, including the selection of power and the number of chargers.

The advantage of this method is its high accuracy in reflecting the actual bus driving cycle. Measurement data are collected and updated continuously, allowing the constructed driving cycles to be used for traffic condition analysis at any time without the need for repeated complex and time-consuming calculations. The source data for demonstrating the method was obtained from files recorded for three selected bus lines over 30 days. The accumulated data set allowed for the creation of a matrix of characteristic travel features comprising a total of approximately 20,000 segments. The obtained data volume enabled the determination of speed development probabilities in specific classes with an error below 5%. It is also worth noting that the method does not use computationally complex mathematical models. The calculations can be performed, for example, on embedded computers placed on the vehicle [24], which significantly expands the application of the method and enables the prediction of the cycle on IoT-embedded devices [25].

The method's limitations include difficulties in its application to a new area without conducting real-world operational traffic characteristic measurements. Since the method is based on actual data as opposed to simulations, it is strongly tied to the characteristics of a specific area. Another issue is that sampling periods are not initially constant due to transmission delays. Not all measurement data are suitable for analysis due to human handling errors (e.g., turning off the transponder or transmission interruptions). Therefore, the development and further practical application of the method should include more accurate quality checks of the data incoming to the server by conducting control geolocation measurements [26] for selected trips.

Driving cycles calculated using the proposed method can easily be scaled by adding new segments and extending to new routes. Adding markers to the records collected in the database, such as weather conditions (temperature, season, road conditions, etc.), will allow for real-time cycle updates. Using obtained driving cycles in the operational mode will enable real-time vehicle range estimation on designated routes. Exploring data accumulated in the data warehouse from different periods and years in analytical mode will allow the detection of long-term dependencies in cycle characteristics from many input parameters. The results of such an analysis can be used in the design of new urban infrastructure, such as routes, schedules, and stops.

Further practical application of the method is planned to assess the impact of restrictions imposed by the urban traffic organization on the smoothness of road segment travel [27]. This situation occurs, for example, during renovations or reconstruction of communication routes (e.g., the tram to Wilanów and the introduction of bus lanes). The method is also planned to be used to determine engine and battery loads by representing the developed driving cycles in terms of instantaneous power consumption conditions. This will allow for the determination of continuous, hourly, and peak power to predict electric energy consumption [28] and the state of battery charge. It will also enable consideration of battery aging conditions.

In summary, the method – when combined with the simulation models of electric vehicles in public transport proposed and approved by the authors, the method for selecting drive system parameters and battery characteristics based on route features, and the method for estimating energy consumption for the route and its replenishment during stops – provides a comprehensive solution to the problem of analyzing the efficiency of electric vehicles in public transport and allows for the planning of its development. It should be noted that the proposed method is not limited to the driving cycles of electric vehicles and can be applied to vehicles with other propulsion systems.

Acknowledgments

This research was co-funded by the Scientific Council of the Discipline of Civil Engineering and Transport of Warsaw University of Technology and a subsidy for implementing research tasks within the statutory activity of the SIMT Department of the Faculty of Transport at the Warsaw University of Technology.

References

1. Das P.K. et al. (Eds.) *Fuel Cells for Transportation*. Elsevier. 2023. ISBN 978-0-323-99485-9.
2. Jeong, N.T. & Yang, S.M. & Kim, K.S. & Wang, M.S. & Kim, H.S. & Suh, M.W. Urban driving cycle for performance evaluation of electric vehicles. *Int J Automot Technol*. 2016. Vol. 17(1). P. 145-151.
3. Ho, S.H. & Wong, Y.D. & Chang, V.W.C. Developing Singapore Driving Cycle for passenger cars to estimate fuel consumption and vehicular emissions. *Atmos Environ*. 2014. Vol. 97. P. 353-362.
4. Nyberg, P. & Frisk, E. & Nielsen, L. Driving Cycle Equivalence and Transformation. *IEEE Trans Veh Technol*. 2017. Vol. 66(3). P. 1963-1974.
5. Giakoumis, E. & Zachiotis, A. Investigation of a diesel-engined vehicle's performance and emissions during the WLTC driving cycle – comparison with the NEDC. *Energies*. 2017. Vol. 10(2). No. 240.
6. Masłowski, D. & Kulińska, E. & Krzewicki, Ł. Alternative methods of replacing electric batteries in public transport vehicles. *Energies*. 2023. Vol. 16(15). No. 5828.
7. Kozłowski, M. & Czerepicki, A. Quick electrical drive selection method for bus retrofitting. *Sustainability*. 2023. Vol. 15(13). No. 10484.
8. Jagiełło, A. & Wołek, M. & Bizon, W. Comparison of tender criteria for electric and diesel buses in poland—has the ongoing revolution in urban transport been overlooked? *Energies*. 2023. Vol. 16(11). No. 4280.
9. *Electromobility and Alternative Fuels Act*. Available at: https://sip.lex.pl/akty-prawne/dzu-dziennik-ustaw/elektromobilnosc-i-paliwa-alternatywne-18683445.
10. Czerepicki, A. & Choromański, W. & Kozłowski, M. & Kazinski, A. Analysis of the problem of electric buses charging in urban transport. *Sci Tech*. 2020. Vol. 19(4). P. 349-355.

11. Torabi, S. & Bellone, M. & Wahde, M. Energy minimization for an electric bus using a genetic algorithm. *Eur Transp Res Rev*. 2020. Vol. 12(1). P. 2.
12. Brzeziński, J. *Gdansk bought electric buses for PLN 62 million. These, however, are not suitable for operation...* Available at: https://moto.pl/MotoPL/7,88389,30398966,gdansk-kupil-za-62-mln-zl-elektryczne-autobusy-te-jednak-nie.html.
13. Jing, Z. & Wang, T. & Zhang, S. & Wang, G. Development method for the driving cycle of electric vehicles. *Energies*. 2022. Vol. 15(22). No. 8715.
14. Esser, A. & Zeller, M. & Foulard, S. & Rinderknecht, S. Stochastic synthesis of representative and multidimensional driving cycles. *SAE Int J Altern Powertrains*. 2018. Vol. 7(3). P. 2018-01–0095.
15. Zhang, M. & Shi, S. & Cheng, W. & Shen, Y. Self-adaptive hyper-heuristic Markov chain evolution for generating vehicle multi-parameter driving cycles. *IEEE Trans Veh Technol*. 2020. Vol. 69(6). P. 6041-52.
16. Li, L. &, Sun, H. & Tao, F. & Fu, Z. Driving cycle prediction based on Markov chain combined with driving information mining. *Proc Inst Mech Eng Part D J Automob Eng*. 2023. P. 09544070231171741.
17. Shi, S. & Lin, N. & Zhang, Y. & Cheng, J. & Huang, C. & Liu, L. & et al. Research on Markov property analysis of driving cycles and its application. *Transp Res Part D Transp Environ*. 2016. Vol. 47. P 171-181.
18. Huang, D. & Xie, H. & Ma, H. & Sun, Q. Driving cycle prediction model based on bus route features. *Transp Res Part D Transp Environ*. 2017. Vol. 54. P. 99-113.
19. Zhao, X. & Ye, Y. & Ma, J. & Shi, P. & Chen, H. Construction of electric vehicle driving cycle for studying electric vehicle energy consumption and equivalent emissions. *Environ Sci Pollut Res*. 2020. Vol. 27(30). P. 37395-37409.
20. Chen, X. & Cheng, Z. & Jin, J.G. & Trépanier, M. & Sun, L. Probabilistic forecasting of bus travel time with a Bayesian gaussian mixture model. *Transp Sci*. 2023 Vol. 57(6).
21. Tang. J & Heinimann, H. & Han, K. & Luo, H. & Zhong, B. Evaluating resilience in urban transportation systems for sustainability: A systems-based Bayesian network model. *Transp Res Part C Emerg Technol*. 2020. Vol. 121. No. 102840.
22. Warsaw City Administration. *Open data*. Available at: https://api.um.warszawa.pl/.
23. *OpenStreet Map.* Available at: https://www.openstreetmap.org/.
24. Subramaniam, A & Ibrahim, N.A. & Jabar, S.N. & Rahman, S.A. Driving cycle tracking device big data storing and management. *Int J Electr Comput Eng*. 2022. Vol. 12(2). No. 1402.
25. Valera, J.J. & Heriz, B. & Lux, G. & Caus, J. & Bader, B. Driving cycle and road grade on-board predictions for the optimal energy management in EV-PHEVs. In: *2013 World Electric Vehicle Symposium and Exhibition (EVS27*). IEEE. 2013. P. 1-10.
26. Jaskowski, P. & Tomczuk, P. & Chrzanowicz, M. Construction of a measurement system with GPS RTK for operational control of street lighting. *Energies*. 2022. Vol. 15(23). No. 9106.
27. Liu, Y. & Feng, X. & Wang, Q. & Zhang, H. & Wang, X. Prediction of urban road congestion using a Bayesian network approach. *Procedia - Soc Behav Sci*. 2014. Vol. 138. P. 671-678.
28. Den, Z. & Xu, L & Liu, H. & Hu, X. & Duan, Z. & Xu, Y. Prognostics of battery capacity based on charging data and data-driven methods for on-road vehicles. *Appl Energy*. 2023. Vol. 339. No. 120954.

Received 16.02.2023; accepted in revised form 31.08.2024