TUAD, 2023, 6(2), 131–142

Trafik ve Ulaşım Araştırmaları Dergisi

Dergi Ana Sayfa: http://dergipark.org.tr/tuad

Safety Research Unit

Araştırma Makalesi

Establishment of Confidence Intervals for Average Vehicle Speeds

Mirhamid Baghirov[1*]

[1] Azerbaijan University of Architecture and Construction, Baku, Azerbaijan

Abstract

Accurate forecasting of the average speed of vehicles is important for early detection of traffic congestion and density, as well as for the correct reporting of the green wave mode. For this purpose, to evaluate the forecasting of traffic indicators using the artificial intelligence method a confidence interval was established on the example of the city of Baku, the capital of the Republic of Azerbaijan. However, this confidence interval includes forecasts for the next few days based on observations of existing roads, not of planned and reconstructed roads. During the preparation of the report, the first 21 days of April 2019 were selected and based on the obtained data, the objective numerical assessment of the quantitative anticipation for the number of hours in a day, standardized deviation's objective statistical approximation, the quantile of the Student's distribution, the lower limit, the upper limit and the values of the difference between them were determined, as well as the degree of freedom and the computed confidence interval. As a result of the reports, one confidence interval is given for each hour of the day, in which the indicator falls into this interval with 95% of the following days. It is also possible to determine the recommended speed limit, which is the most important part of the green wave mode, with the help of the mentioned prediction.

Keywords: confidence interval, average speed of vehicles, traffic flow, green wave, recommended speed of movement, road traffic management

Ortalama Araç Hızları İçin Güven Aralıklarının Oluşturulması

Öz

Araçların ortalama hızlarının doğru tahmin edilmesi trafik sıkışıklığının ve yoğunluğunun erken tespiti ve aynı zamanda yeşil dalga modunun doğru raporlanması için önemlidir. Bu amaçla yapay zeka yöntemi kullanılarak trafik göstergelerinin tahminini değerlendirmek için Azerbaycan Cumhuriyeti'nin başkenti Bakü şehri örneğinde bir güven aralığı oluşturulmuştur. Ancak bu güven aralığı, planlanan ve yeniden yapılan yollara değil, mevcut yollara ilişkin gözlemlere dayalı olarak önümüzdeki birkaç gün için tahminleri içerir. Raporun hazırlanması sırasında 2019 Nisan ayının ilk 21 günü seçilmiş ve elde edilen verilere göre günün saatlerine ilişkin matematiksel beklenti için yansız istatistiksel tahmin, standart sapma için yansız istatistiksel tahmin, Student dağılımı, alt sınır, üst sınır ve aralarındaki farkın değerleri belirlenmiş, ayrıca serbestlik derecesi ve güven aralığı hesaplanmıştır. Raporlamalar sonucunda günün her saati için bu güven aralığı verilir ve gösterge takip eden günlerin %95'i ile bu aralığa düşer. Bahsi geçen tahmin yardımıyla yeşil dalga modunun en önemli parçası olan tavsiye edilen hız limitini belirlemek de mümkün olmaktadır.

Anahtar Kelimeler: güven aralığı, araçların ortalama hızı, trafik akışı, yeşil dalga, önerilen hareket hızı, karayolu trafik yönetimi

[*] *İletişim / Contact: Mirhamid Baghirov, Azerbaijan University of Architecture and Construction, Baku, Azerbaijan. E-Posta / E-mail:* mirhemid.bagirov@gmail.com.
Gönderildiği tarihi / Date submitted: 10.02.2023, Kabul edildiği tarih / Date accepted: 03.06.2023
Alıntı / Citation: Baghirov, M. (2023). Establishment of confidence intervals for average vehicle speeds. *Trafik ve Ulaşım Araştırmaları Dergisi, 6*(2), 131–142. doi: 10.38002/tuad.1249767

Establishment of Confidence Intervals for Average Vehicle Speeds

For pedestrians to make judgments that would assure their safety when trying to cross the road, it is crucial that they have the capability to properly assess and anticipate the average vehicle speed (Sun et al., 2015).

For various Intelligent Transport Systems (ITS) applications, determining the average speed of vehicles is important. However, it is very difficult to determine the specified forecast on highways since the individual speed of the vehicle is influenced in a deterministic or stochastic way by many factors related to the vehicle driving and traffic safety (e.g., the state of the road traffic, kinds of motor vehicles plus psychophysiological characteristics of each driver). Machine learning makes it more accessible to determine average vehicle velocity forecast by investigating probable connections among average traffic speeds and other key factors in a traffic flow based on archival information on automobile infrastructure. This paper presents a new method influenced by data proposed for the long-term prediction of individual average vehicle speed along planned streets or avenues based on long and short-term memory algorithms using error back-propagation. In some sources, the use of the Pearson’s correlation coefficient increased the effectiveness of calculating and analyzing parameter correlation of the archived characteristics of the vehicle driving and traffic safety (vehicle, driver, road, traffic). Several accounts claim that the methodology for predicting the average speed of vehicles is further divided into parametric (constant) and non-parametric (multivariate) ways, as shown below (Lefèvre et al., 2014; Xiaolei et al., 2015; Yufang et al., 2019).

- Parametric approaches rely on models calculated using empirical data with predetermined parameters relying on a few theoretical premises;
- In non-parametric approaches, models are built based on historical data on traffic flow and vehicle management forecasting.

When it comes to large-scale transportation networks or long-term velocity forecasting of vehicles, parametric ways are typically useless. Recent vehicle velocity forecast studies are mostly focused on both short and long-term vehicle velocity forecast.

Long-term forecast of average vehicle speeds is of vital importance, which will lead to great achievements in improving traffic safety and management efficiency, as well as achieving optimal route programming. Thus, as can be seen from the above, Accurate long-term estimates of the average velocity of moving vehicles are particularly challenging in comparison with the short-term speed. However, the estimated velocity of vehicle is the outcome of a variety of aspects, namely psychophysiological state of the driver, the condition of the vehicle, road conditions, and other factors. Accurately obtaining these ratios will get more challenging the longer the predicting period. The creation and use of artificial sensory systems as well as other methods for machine learning to address the aforementioned issues has led to significant advancements in the fields of regression prediction (multivariate analysis), machine vision (computer interpretation of images), and human language technology (NLP) (Chen et al., 2018; Jiang & Learned-Miller, 2017; Kai et al., 2015; Moreno-Lopez & Kalita, 2017).

Given that forecasting vehicle’s velocity is also a problem of regression prediction, an adequate machine learning technique could improve the low precision and constrained forecasting scope brought on by the conventional "decision tree" model, Markov model, as well as other velocity forecasting methods to a certain degree (Jing et al., 2017).

In order to develop a vehicle speed prediction model, a number of studies have been conducted and international experiences in this field have been studied. So, the following sequence should be expected when setting the model parameters:

- Configuration of parameters of the “planned path error backpropagation method” model;
- Configuration of "long short-term memory" settings;
- Setting up parameter arrangements for the “support-vector machine” model.

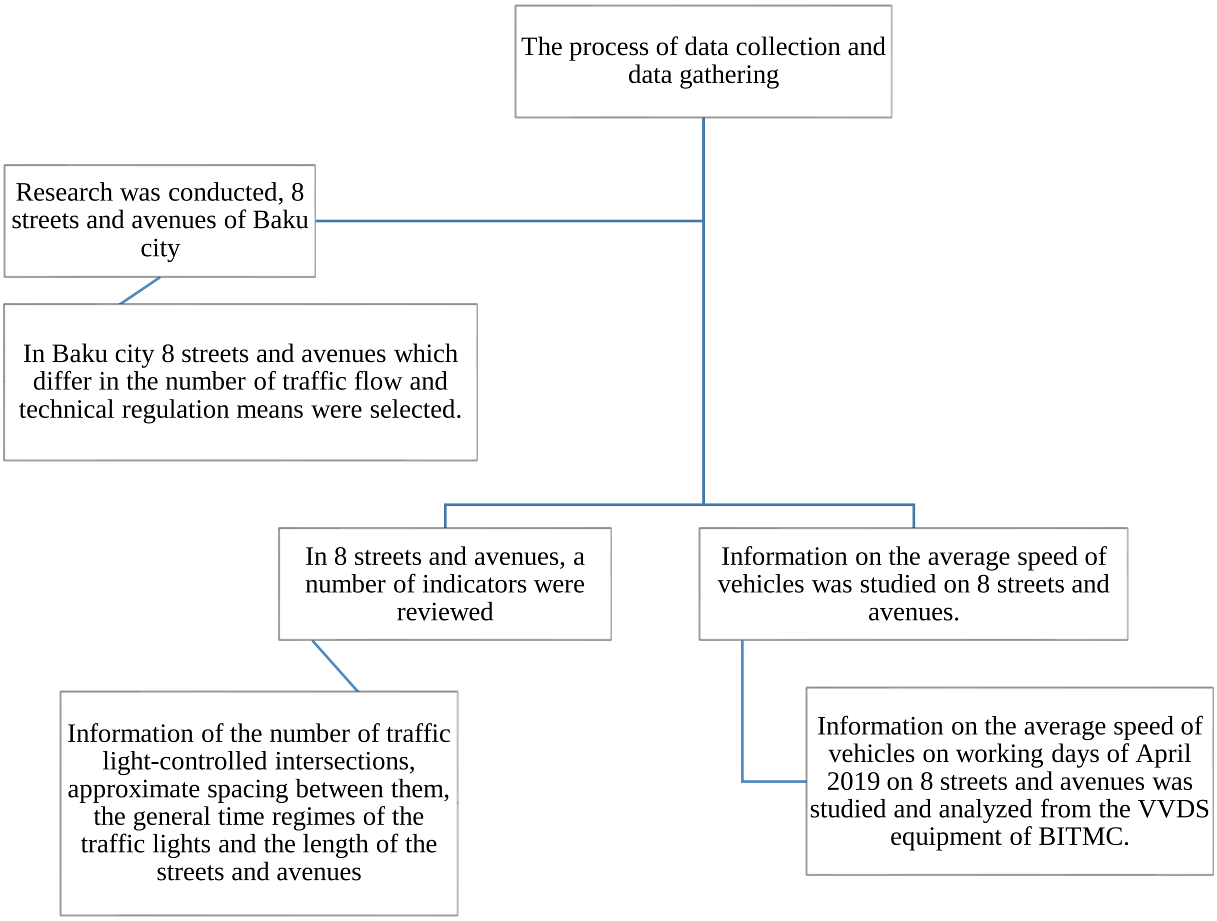
At the same time, in some sources, the method for predicting the average speed of a vehicle, which combines the algorithms of the “error backpropagation method” and “long short-term memory” on planned highways, the correspondence study involving the impact of vehicle on the "long-term velocity forecast" of the vehicle and the Pearson's correlation coefficient technique are used to emphasize the main determinant factors impacting the velocity of the vehicle, as a contribution to the long-term velocity forecasting of the vehicle model, where it offers higher forecasting precision and computational productivity. It has been discovered that the mathematical forecasting of each vehicle's speed in the long run has the ability to adapt to different roads and can choose a more suitable algorithm based on position and road information. Another thing to keep in mind is that the average forecasting of the vehicle’s velocity in the long run differs from the short-run velocity estimation demonstrated in the error features of the vehicle’s velocity for every point or time because unexpected and random elements that impair a vehicle's ability to corner are likely to have an impact on how foreseeable a vehicle's acceleration is over the long run. However, unlike the application of analyzing velocity in the short run, in optimized control at rapid or short-term power points, the energy spent on the intended route is much more indicative of the vehicle speed statistics, and it has been confirmed that the outcome of long-term forecasting of vehicle’s velocity are of practical significance for estimation of arrival and forecasted possibilities for energy usage.

Taking into account what was mentioned above, a confidence interval was established for the evaluation of the forecast on the example of a number of streets or avenues in Baku, the capital of the Republic of Azerbaijan (Valiyev, 2013). However, this confidence interval includes predictions for the next few days based on observations on existing roads, not on roads being designed and reconstructed. For the purpose of this forecast, one confidence interval is given for each hour of the day, in which the indicator falls to 95% for the following days (Sandercock, 2015). As can be seen from the percentage indicator, the interval has a very important indicator in determining the average movement speed for the application of the "green wave" mode in very large and large cities (Ahmadov & Baghirov, 2019; Lu et al., 2022; Zhang et al., 2020).

2. Method

2.1. Data Collection

In order to establish a confidence interval, research was conducted on several streets and avenues in Baku. For this purpose, first of all, average traffic speed indicators were investigated in 8 streets and avenues in Baku city. The aforementioned research was carried out with the help of the equipment of the Intelligent Transport Management Center operating in Baku. Thus, the Video Vehicle Detection System, which is one of the ITMC's equipment, collects data on the intensity of vehicles on the roads at 1-minute intervals (Scheme 1).



Scheme 1. Flowchart for Comprehending the Data Collection and Gathering Methodology

The marked streets and avenues are shown in red in the following figure (Figure 1). The number of traffic light-controlled intersections, approximate spacing between them, the general time regimes of the traffic lights and the length of the streets and avenues are shown in Table 1 below.

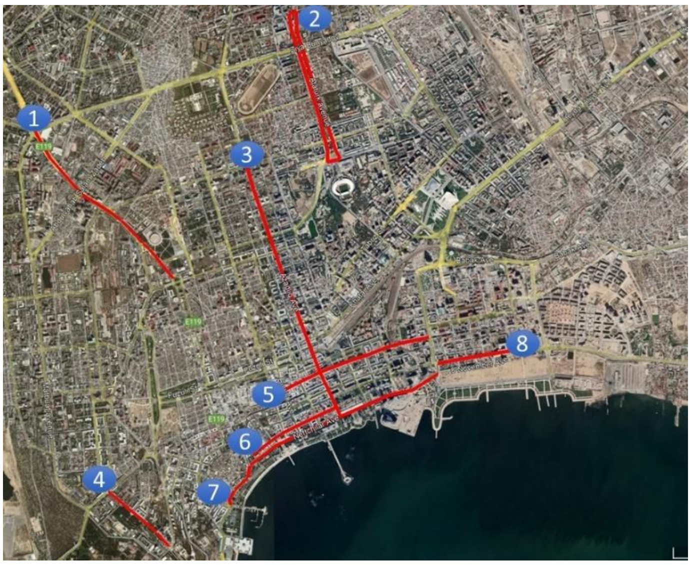


Figure 1. Research was conducted, 8 streets and avenues of Baku city

Table 1. The number of traffic light-controlled intersections, approximate spacing between them, the general time regimes of the traffic lights and the length of the streets and avenues

Street number	Number of traffic light-controlled intersections (number)	Approximate spacing of traffic light-controlled intersections (meters)	The length of the street (km)	Cycle in peak hours (second)	Cycle in non-peak hours (second)
1	6	400	2.51	90	70
2	4	410	2.06	90	70
3	16	100	3.30	90	70
4	1	490	0.98	90	70
5	6	230	1.90	90	70
6	3	290	1.18	90	70
7	2	1000	3.14	90	70
8	1	450	0.90	90	70

During the investigation of the first three weeks of April 2019, data were collected on the indicators of the average speed of motor vehicles from the first to the 21st of the month (including April 21) for all hours of the day (Table 2).

During the observations, taking into account that weekday drivers pay more attention to trips, for example, they often change the speed limit so as not to be late for work, and weekends are non-working days, the report was made only for working days. In the future, we consider it appropriate to conduct this type of report separately for non-working days.

2.2. Data Analysis

After appropriate analysis, a diagram of the dependence of the average speed of motor vehicles on the days of the month was constructed separately by hours of the day (Figure 2).

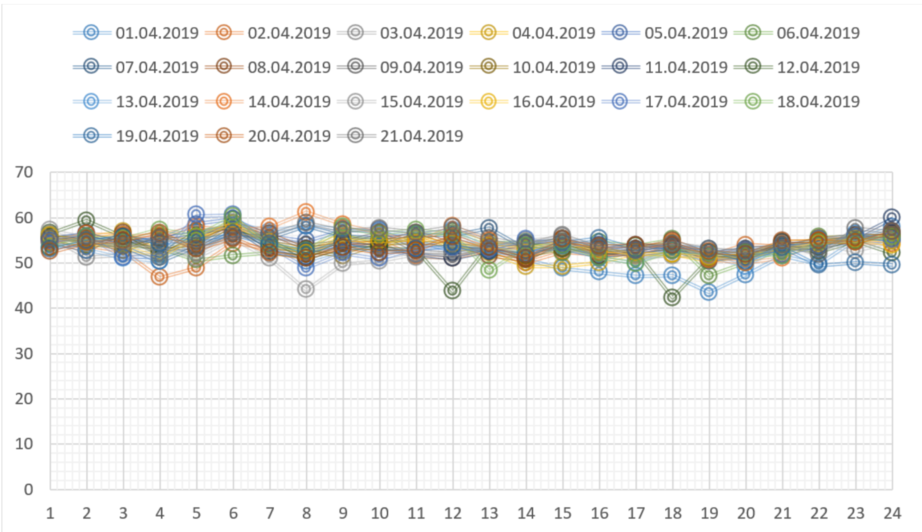


Figure 2. Comparison diagram of the average speed of vehicles by hours of the day for April 1-21,2019

Table 2. Average speed of motor vehicles for April 2019

Data	**Average speed of motor vehicles**																				
	For April 2019																				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
days of the week / hours per day	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun
1	53	56	56	56	56	56	56	55	55	56	55	57	54	54	57	56	55	53	53	52	55
2	54	54	51	56	57	54	55	53	56	55	56	59	53	57	56	54	56	56	55	54	55
3	56	53	52	57	54	54	54	57	55	53	55	56	51	56	54	54	51	55	56	55	54
4	54	47	53	53	55	53	50	52	53	56	55	52	50	56	53	51	51	57	54	57	55
5	52	49	52	55	59	51	54	55	53	55	57	57	57	58	51	56	60	56	55	53	54
6	57	54	57	60	60	52	57	59	58	58	58	58	59	57	55	59	61	60	56	55	56
7	54	53	54	55	56	52	55	56	55	54	52	55	57	58	51	54	54	54	54	52	56
8	50	52	52	53	55	51	58	53	54	53	51	51	58	61	44	52	49	53	53	51	59
9	53	53	55	56	57	56	56	54	55	52	54	53	57	59	50	55	52	58	54	53	57
10	54	55	53	53	57	55	55	54	57	54	54	53	57	57	50	55	51	56	52	52	58
11	55	51	51	56	54	55	56	57	53	55	52	53	57	54	52	55	53	57	53	53	57
12	53	51	51	53	55	57	56	54	52	54	51	44	57	58	52	54	53	56	54	56	58
13	53	53	53	52	53	54	58	52	53	53	53	54	53	55	53	54	53	48	53	53	55
14	51	50	54	49	53	54	53	51	51	52	54	54	52	53	55	53	55	54	52	51	54
15	49	52	55	49	54	52	53	55	54	53	55	55	54	55	56	55	54	54	54	55	56

Table 2. Continued

16	48	52	54	50	52	55	51	52	52	53	53	52	51	52	54	53	54	51	55	54	53
17	47	53	51	52	52	54	52	53	52	52	54	53	50	52	52	52	53	50	53	54	53
18	47	53	52	52	53	55	54	55	54	53	53	42	53	55	52	52	53	53	54	54	54
19	44	52	51	50	50	47	52	52	51	52	52	51	52	51	52	51	52	52	53	52	53
20	47	54	52	52	52	50	53	51	51	52	51	52	49	50	51	52	53	52	50	52	52
21	52	54	54	55	53	53	54	54	53	53	54	54	54	51	53	53	52	52	53	55	54
22	50	55	55	55	53	56	54	55	53	54	54	52	55	55	54	54	54	54	49	55	52
23	55	56	56	55	55	56	56	56	55	56	56	55	56	55	53	55	56	55	50	54	58
24	58	56	57	56	55	56	55	56	56	55	60	52	55	54	55	54	55	56	50	57	57

3. Results and Discussion

Based on the above-mentioned data, the unbiased statistical estimate of the mathematical expectation for the hours of the day, the unbiased statistical estimate of the standard deviation, the Student's distribution quantile, the lower and the upper limits criteria are obtained, and the values of the difference between them were determined and the df-degree of freedom and the confidence interval were calculated (Table 3) (Chen et al., 2020; Hurst, 2010; Shan et al., 2014; Thelwall & Fairclough, 2017; Zhang et al., 2021).

As of 01:00 every day of the first three weeks of April, the calculations were made in the following order:

- Unbiased statistical estimate for mathematical expectation (Poznyak, 2009):

$$v_{ME} = \frac{\sum_{i=1}^{15} v_{av,i}}{15} = 55.14\ km/h$$

here,

$v_{av,i}$ – respectively, the sum of the average speed indicators of vehicles on the i-th working day of April;

15− is the number of days the calculation is made.

- Unbiased statistical estimate for the standard deviation:

$$Y_{SD} = \sqrt{\frac{\sum_{i=1}^{15}\left(v_{av,i} - v_{ME}\right)^2}{15-1}}$$

here,

$v_{av,i}$ – respectively, the sum of the average speed indicators of vehicles on the i-th working day of April;

v_{ME} − unbiased statistical estimate for mathematical expectation

15 − is the number of days the calculation is made.

- Df-degree-of-freedom value:

$$df = 15 - 1 = 14$$

- Upper limit estimation:

$$v_{ul} = v_{ME} - \frac{2.14 \times Y_{SD}}{\sqrt{15}}$$

here,

2.14 − quantile of Student's distribution;

15 − is the number of days the calculation is made.

- Lower limit estimation:

$$v_{ll} = v_{ME} - \frac{2.14 \times Y_{SD}}{\sqrt{15}}$$

- Difference between upper and lower limit estimations:

$$v_{ul} - v_{ll}$$

Under the above-mentioned calculation rules, calculations were made for all hours of the day and these indicators are listed in table 3.

Table 3. Calculations for all hours of the day

Hour	An unbiased statistical estimate for the mathematical expectation	An unbiased statistical estimate for the standard deviation	DF-degree of freedom	The confidence interval	Quantile of Student's distribution	The lower limit	The upper limit	The difference
01:00	55.14	1.18	14	95.00%	2.14	54.49	55.79	1.31
02:00	55.14	1.85			2.14	54.12	56.16	2.04
03:00	54.38	1.76			2.14	53.40	55.35	1.95
04:00	53.15	2.52			2.14	51.76	54.54	2.79
05:00	54.79	2.95			2.14	53.15	56.42	3.27
06:00	57.98	1.77			2.14	57.00	58.96	1.96
07:00	54.11	1.35			2.14	53.36	54.86	1.50
08:00	51.67	2.53			2.14	50.27	53.07	2.80
09:00	54.13	2.10			2.14	52.97	55.29	2.32
10:00	54.00	1.93			2.14	52.93	55.07	2.13
11:00	53.85	1.95			2.14	52.77	54.92	2.16
12:00	52.53	2.90			2.14	50.92	54.13	3.21
13:00	52.69	1.39			2.14	51.92	53.47	1.54
14:00	52.44	1.85			2.14	51.42	53.47	2.05
15:00	53.59	2.05			2.14	52.45	54.73	2.27
16:00	52.33	1.80			2.14	51.33	53.33	2.00
17:00	51.80	1.66			2.14	50.88	52.72	1.84
18:00	51.81	3.16			2.14	50.06	53.56	3.50
19:00	50.97	2.17			2.14	49.77	52.17	2.41
20:00	51.45	1.44			2.14	50.66	52.24	1.59
21:00	53.19	0.97			2.14	52.66	53.73	1.07
22:00	53.59	1.80			2.14	52.60	54.59	1.99
23:00	54.79	1.54			2.14	53.94	55.65	1.71
24:00	55.39	2.38			2.14	54.07	56.70	2.63

As a result of the reports, we will get a comparative diagram of the average speed of vehicles in Baku, the capital of the Republic of Azerbaijan, during the first 21 days (15 working days) of April 2019, a result of the gap between the maximum and minimum hours of the day (Figure 3).

4. Conclusion

Based on the obtained results, it will be more convenient to build a coordinated scheduling graph ("green wave" graph) on streets or avenues where it is not possible to determine the average speed of motor vehicles without detectors or video detection cameras. It will be possible to obtain the recommended speed indicators for all hours of the day on all streets or avenues where the "green wave" mode has been applied based on the confidence interval construction. Also, on streets or avenues that have implemented the "green wave" mode, the average speed of motor vehicles will be provided at the recommended speed according to the real mode for the hours of the day, which will not violate the coordinated scheduling graph, reduce traffic congestions and, most importantly, the organization of traffic will lead to increased efficiency and safety.

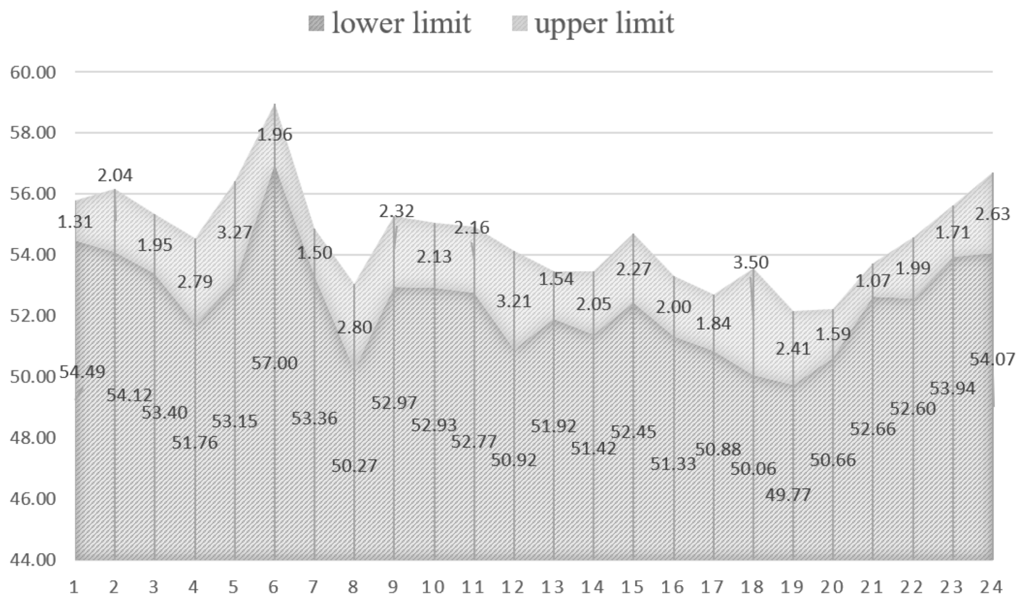


Figure 3. Comparative diagram of the average speed of motor vehicles during the first 21 days of April 2019, resulting from the gap between the maximum and minimum hours of the day

In the future, by increasing the number of investigated streets and avenues, as well as by more precisely studying other means of technical regulation of traffic on streets and avenues, it is possible to establish a more accurate confidence interval.

In addition, we inform you that according to the analyses, considering that weekday drivers pay more attention to trips, for instance, they often change the speed limit for not to be late for work, and weekends are non-working days, the report was made according to working days only.

Ethics Committee Approval Statement

Ethical committee approval is not required, as the study did not collect data from human or animal participants.

References

Ahmadov, G. M., & Baghirov, M. I. (2019). Application of Coordinated Regulatory System on Matbuat Avenue During Off-Peak Hours of the Day. *Scientific Works, 4*, 4.

Chen, L., Xing, Y., Zhang, J., & Na, X. (2018). Levenberg–Marquardt backpropagation training of multilayer neural networks for state estimation of a safety-critical cyber-physical system. *IEEE Transactions on Industrial Informatics, 14*(8), 3436–3446. doi: 10.1109/TII.2017.2777460

Chen, Z., Wang, B., & Gorban A. N. (2020). Multivariate Gaussian and Student-*t* process regression for multi-output prediction. *Neural Computing and Applications*, 32(8), 3005–3028. doi: 10.1007/s00521-019-04687-8

Hurst, S. (2010). *The Characteristic Function of the Student t-Distribution* (Financial Mathematics Research Report No. 95/6). Canberra: Centre for Mathematics and its Applications, School of Mathematical Sciences, ANU. https://www.worldcat.org/title/characteristic-function-of-the-student-t-distribution/oclc/37065789

Jiang, H., & Learned-Miller, E. (2017). Face detection with the faster R-CNN *Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition*, 650–657. doi: 10.1109/FG.2017.82

Jing, J., Filev, D., Kurt, A., Ozatay, E., Michelini, J., & Ozguner U. (2017). Vehicle speed prediction using a cooperative method of fuzzy Markov model and autoregressive model. *Proceedings of the IEEE Intelligent Vehicles Symposium*, 881–886.

Kai, C., Yi, Z., & Fangyan D. (2015) A LSTM-based method for stock returns prediction: a case study of China stock market. *Proceedings of the IEEE International Conference on Big Data (Big Data)*, 2823–2824. doi: 10.1109/BigData.2015.7364089

Lefèvre, S., Sun, Ch., Bajcsy, R., & Laugier, C. (2014). Comparison of parametric and non-parametric approaches for vehicle speed prediction. *Proceedings of the American Control Conference*, 3494-3499. doi: 10.1109/ACC.2014.6858871

Lu, K., Xin, T., Shuyan, J., Jianmin, X., & Yinhai, W. (2022). Optimization model for regional green wave coordinated control based on ring-and-barrier structure. *Journal of Intelligent Transportation Systems*, *26*(1), 68-80. doi:10.1080/15472450.2020.1795847

Moreno-Lopez, M., & Kalita, J. (2017). Deep learning applied to NLP (arXiv:1703.03091vl). doi: 10.48550/arXiv.1703.03091

Poznyak, A. S. (2009). 3-Mathematical Expectation, Advanced Mathematical Tools for Automatic Control Engineers: Stochastic Techniques. Elsevier.

Sandercock, P.A. (2015). Short History of Confidence Intervals. *Stroke, 46*(8), 184-187. doi: 10.1161/STROKEAHA.115.007750

Shah, A., Wilson, A. G., & Ghahramani Z. (2014). Student-*t* processes as alternatives to Gaussian processes. *Proceedings of the* 17th *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 877–885.

Sun, R., Zhuang, X., Wu, Ch., Zhao, G., & Zhang K. (2015). The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *30*, 97–106. doi: 10.1016/j.trf.2015.02.002

Thelwall, M., & Fairclough, R. (2017). The accuracy of confidence intervals for field normalized indicators. *Journal of Informatics*, *11*(2), 530-540. doi: 10.1016/j.joi.2017.03.004

Valiyev, A. (2013). Baku. *Cities*, *31*, 625-640. doi: 10.1016/j.cities.2012.11.004

Xiaolei, M., Zhimin, T., Wang, Y., & Yunpeng, W. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C, 54*, 187–197. doi: 10.1016/j.trc.2015.03.014

Yufang, L., Mingnuo, Ch., & Wanzhong, Zh. (2019). Investigating long-term vehicle speed prediction based on BP-LSTM algorithms. *IET Intelligent Transport Systems*, *13*(8), 1281-1290. doi: 10.1049/iet-its.2018.5593

Zhang, J., Shang, H., Li, X., & Yao, Y. (2020). An integrated arterial coordinated control model considering green wave on branch roads and pedestrian crossing time at intersections. *Journal of Management Science and Engineering*, *5*(4), 303-317. doi:10.1016/j.jmse.2020.09.004

Zhang, D., Min, L., Lan, Q., Zhang, Y., Jingcheng, L., & Jun, L. (2021). Analytical modeling of piezoelectric 6-degree-of-freedom accelerometer about cross-coupling degree. *Measurement*, *181*. doi: 10.1016/j.measurement.2021.109630