IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)
e-ISSN: 2278-1684,p-ISSN: 2320-334X, Volume 17, Issue 6 Ser. I (Nov. – Dec. 2020), PP 13-20
www.iosrjournals.org

Methodology to Calibrate the Saturation Flow in Traffic Light Urban Intersections

Fabio de Souza P. Borges[1], Adelayda P. Fonseca[1], Reinaldo C. Garcia[1]
[1]*(Faculdade de Tecnologia, Programa de Pós-Graduação em Transportes, Universidade de Brasília, Brazil)*

Abstract
The urban infrastructure has experienced great challenges due to the increase in the population and consequently growth of the vehicle fleet. The traffic light intersections are among the main factors to optimize the urban road systems aiming to achieve the efficiency of the infrastructure system. The flow saturation can be used to synchronize the traffic cycles applying detection and control traffic technologies taking into account the huge amount of the collected traffic data. This work applies then data mining techniques to calibrate the saturation flow and its robustness is shown being validated by the Shanteau[1] and the Highway Capacity Manual [2] methods.
Keywords*:Traffic Congestion, Macroscopic Fundamental Diagram, Saturation flow, Data Mining.*

Date of Submission: 01-11-2020 Date of Acceptance: 13-11-2020

I. Introduction

Traffic volumes increases at the same rate as the population growth due to the strong correlation between population and individual vehicles [3]. The expansion of the urban infrastructure has not been able to attend the advance in the demand of the vehicle fleet due to the lack of space and the scarcity of public resources [4][5].

The urban road system has become an essential social problem in the current world context requesting the need for improvements in the safety and the efficient use of the transport system [6]. In particular, the yearly estimated costs due to the traffic congestion in São Paulo was about US$ 8 billion in 2012 [7] showing the need to achieve efficiency in the urban transportation structure.

The signaled intersections have been considered important elements for the development of an efficient urban road system [8]. Furthermore, the saturation flow is an important parameter to determine the traffic light cycles being a measure of the operational level performance of these intersections showing then their capacity in an ideal situation of their operation. Therefore, the importance of the saturation flow to determine and update the traffic light plans applying even traditional methods, besides the work required to achieve efficiency through the collected field data has already been shown in previous studies [9][10]. These works show the need to detect not only the formation and the discharge of the queues but also the saturation flows during the traffic light cycles.

The saturation flow is a fundamental parameter to measure the service level, determining the capacity of the lanes being applied to optimize the synchronism of the traffic light cycles at intersections [11]. It has been extensively used to design, operate and control signalized intersections to better manage the urban road transport system. Moreover, the developed models need to consider the different factors influencing the saturation flow, including climatic conditions, lane width and driver behavior [11][12][13].

Recent research has been published applying new techniques including data mining to improve the efficiency of the transportation system [4][14][15][16]. These works through the collected real-time data set forecasted the traffic flow for close periods of time to better manage the traffic congestion [5][17][18].

Previous works have also intended to determine the different traffic features besides calibrating the required parameters ascertain the different flow regimes [15][18]. Techniques aiming to influence the traffic flow on real-time and methods to evaluate the saturation flow at traffic lights have also been implemented [6][8][12][19]. The technique of Artificial Intelligence (AI) has achieved good results in the already mentioned publication when analyzing the traffic flow parameters.

This article describes a methodology to calibrate the saturation flow in urban approaches applying mining techniques on real-time data collected by electronic surveillance equipment. The paper is organized as follows. The introduction presents a review of the literature of the already related published work. The second section describes the methodology to calibrate and validate the evaluated saturation flow. The third section presents the results when the developed methodology is applied and, the last section states our conclusions and recommendations.

II. Background

One of the most important tasks in traffic engineering is to estimate accurately the saturation flow [11]. Previous studies generally rely on expensive field research techniques without taking into account the dynamic changes in the saturation flow. Moreover, the published works in the last decade show the saturation flow relevance as a fundamental parameter to traffic flow emphasizing the need to modernize and automate its use in different traffic conditions.

One of the first studies to suggest a structure to calibrate the occupation versus flow diagram in a road section was given by [18]. The study aimed to calibrate the model using a linear regression approach adjusted to the occupation and the flow data. The researchers used the capacity definition given in [2] to determine the capacity of the segments of the case study. They also applied an approximate quantile regression technique to create the supercritical occupation and flow adjustment to the collected traffic data in the Interstate 880 highway

The calibration of the speed and the density parameters using data mining techniques was addressed by [17]. A new hierarchical clustering algorithm was proposed based in the k-means method and an improvement in the grouping performance was achieved. The new approach was tested with the collected data in the Beijing's third ring road and the obtained results showed the robustness of their proposed algorithm achieving better results than previous published works.

Another method to generate and calibrate the density versus flow diagram was also presented by [20]. Firstly, according to the traffic time series fluctuations, the traffic data was partitioned into free flow and congested flow. A minimal principle technique was then used to identify the equilibrium states applying a mixed integer optimization technique to obtain linear adjustments to the density and the traffic flow. The optimization procedure goal was to adjust to the minimum absolute deviations the analyzed parameters. The methodology was implemented to three California highways and, the results showed the robustness of the developed approach.

The parameters calibration structures associated with the density and the traffic flow were also studied by [21]. They applied data mining techniques to calibrate the speed of the vehicles through a large volume of collected data improving the accuracy of the calibrated parameters. Furthermore, they implemented not only the agglomerative hierarchical grouping approach but also the method of the nearest k-neighbor to optimize the classifications obtaining new samples in the best adjusted groupings. The proposed algorithm performed well in the precision and efficiency of the parameter's estimation.

The application of clustering techniques when partitioning the traffic flow data in free and congested flow regimes has already been proposed by [16]. A clustering structure based on the fundamental variables of traffic flow including flow, speed and density is presented and, the obtained results are compared when applying the three grouping techniques of the Hierarchical, the k-means and the Gaussian Mixture Models. The results showed that clustering is an effective way to classify the flows of the traffic data creating then the fundamental traffic flow diagrams.

A method applying the Grubbs or Extreme Value Test to evaluate the saturation flow for traffic signal intersections in real time was presented by [13]. The proposed method through the analysis of the traffic flow theory and the composition of the queues eliminated the data outliers with higher probability excluding them from the measured headways. The attained results suggest that the traditional method to calculate the saturation flow underestimates its value since it is biased by the different types of vehicles, mainly the large ones. Their proposed algorithm was a more accurate one since it excluded the distortion due to the large and medium vehicles.

A two-stage grouping methodology based on big data to calibrate the fundamental traffic diagram in expressway networks was presented by [19]. The k-means algorithm jointly with the hierarchical and the modified hierarchical groupings technique were applied obtaining the fundamental diagrams to the Melbourne motorway network. A kriging method to model the saturation flow at traffic light intersections in India was developed by [8]. They compared their results with the ones obtained by the conventional and regression methods to assess the percentage of absolute error and their proposed methodology achieved better results than previous approaches.

A method to estimate the saturation flow in China freeways was presented by [11] where the collected data by video detectors were used to overcome the limitations of the field research. Analyzing the images, the headways were evaluated applying the linear regression least squares method and, afterwards they the saturation flow was evaluated by the average of the saturation headways using the Dickey-Fuller test.

The presented literature shows new applications being used to the traffic flow theory topic to make the transportation infrastructure more efficient. The integration of the described concepts can then be implemented to build a methodology taking advantage of the large amount of data collected by electronic vehicle inspection tools. These devices can be installed in the main traffic light intersections of the Brazilian cities and, applying modern tools including data mining, the collected data can be used to increase the performance of the traffic sector.

III. Methodology

The proposed methodology in this research is based on the k-means clustering technique to classify the data collected in two distinct phases, the free and the congested flows. The methodology obtains the transition point on the border between the two conditions of the traffic flow determining the saturation flow in the signalized intersections.

The collected data includes thirty-one months of vehicle counting by electronic vehicle inspection equipment installed in the approach of a traffic light intersection. Moreover, the applied procedures to validate the evaluated saturation flow are [1] and [2] ones.

Data Mining: k-means technique

It is a classic data mining technique based on the interactive distance between data points. The technique consists in: (i) choose the number of clusters in which data is intend to be grouped and randomly generate k points as the center of its clusters; (ii) classify the remainder of data points according to their Euclidean distances to the nearest cluster center; (iii) recalculate the value of the centroides or means of each cluster; (iv) repeat the process with new cluster centers until definitive centroid stabilization is achieved or a stopping criterion has been reached [22][23].

Saturation Flow Calibration

The available data by the Federal Transit (DF) transit agency was composed by raw data containing the arrival times and the vehicle speed at the intersections. The data were then analysed to identify not only periods without records in the given time series but also daily periods where the flows were higher with speed reductions. This preliminary analysis guarantees a consistency in the data set and the existence of a congested flow [18].

The speed for each vehicle (i) is obtained from the database and, the density is evaluated by the number of vehicles contained in the last 1 km travelled by that same vehicle (i). Afterwards, the flow and then the fundamental traffic flow diagram are evaluated. The points in the fundamental traffic flow diagram are then classified in free and congested flows applying the k-means method. The boundary between the two states is evaluated and the points for each cluster are adjusted by linear regression. The intersection given by the two drawn regression lines represents the threshold between the two flow states, the free and the congested one, and therefore the corresponding saturation flow.

Validation of The Saturation Flow

The collected data by the electronic vehicle inspection equipment used were used to graphically check the traffic flow behaviour applying the already cited methods, the Shanteau method [1] and the Highway Capacity Manual one [2]. Therefore, the peak periods were previously identified using the Shanteau method and the HCM methods determining the saturation flow when approaching the intersection.

The Shanteau method [1] measured the saturation flow graphically by a discharge of the accumulation curve for the vehicle queue [24]. In their method, the axes of ordinates and abscissa represent respectively the quantities and accumulated times of each vehicle. If the rows dissipate under saturation, the graph shows a straight line whose slope given by the ratio of the change in the number of vehicles (ΔN) and the time interval (Δt) is the saturation flow.

The saturation flow was also determined by the Highway Capacity Manual [2] following the HCM´s guidelines. The first 4 vehicles of the retention line were then excluded to avoid the reaction´s responses of the first drivers due to their inertia and, at least 15 cycles having more than 8 vehicles were used to perform the calculations.

The graphs given in Figure 1 show the straight lines described by [1]. The sample was taken for three different days, Tuesday, Wednesday and Thursday from 6:45 am to 7:15 am. Moreover, there were 10 sets of data extracted for each day. Thus, the saturation flows were measured by the slope of the linear regression lines ($\Delta N / \Delta t$) and converted into vehicles per hour:

$$S=[3.600\times(\Delta N/\Delta t)] \qquad (1)$$

Finally, aiming to confirm the evaluated saturation flow calibration by the proposed methodology, a new field sampling was performed, independent of the data collected by the electronic vehicle inspection equipment. The field survey was carried out to assess the saturation flow when approaching the intersection and, the collected sample was made from November 21st to 23rd and, December 5th to 7th, 2017 from 6:45 amto 7:15 am. The research team included 10 analysts, grouped in pairs and positioned near the retention lines of each intersection route under the saturation regime. One researcher observed the last vehicle in the queue when the traffic light became green and, a second one recorded the times of each vehicle of the corresponding retention line when crossing the traffic light. The results were then analyzedand plotted, being in accordance withthe ones obtained by the Shanteau method [1].

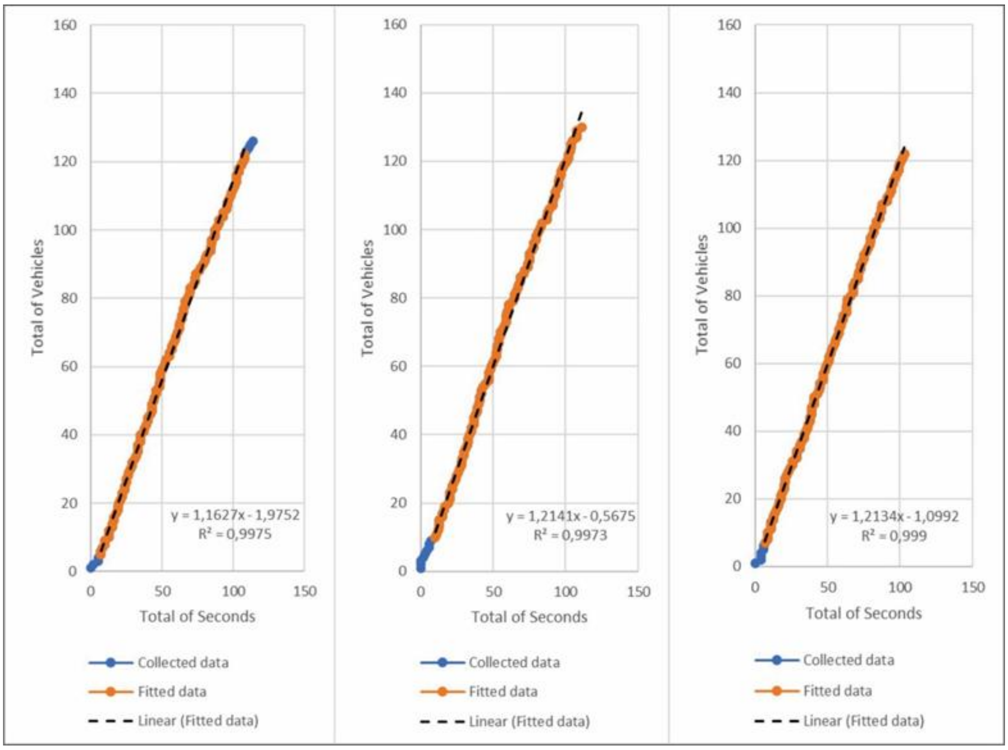

Figure 1: Graphical examples of discharges from the rows of accumulated vehicles represented in straight lines

IV. Case Study: Analysis of The Results

The data set already described was used to obtain the density and the flow of the studied intersection. The raw data and its classification in congested and free flows regimes are shown in Figure 2.

The applied methodology is then shown in Figure3 highlighting (a) the linear regressions adjusted to the clusters of each type of flow, the free and the congested one, and (b) the line passing by the intersection of these regressions from the origin of the axis system, defining the point S on theboundary of the two flows. Finally, the calibration of the saturation flow is concluded as it is shown in Figure 4, where the values of the saturation flow and the corresponding density, q=1,430 veh./h and k=36 veh./km, are obtained for the studied data.

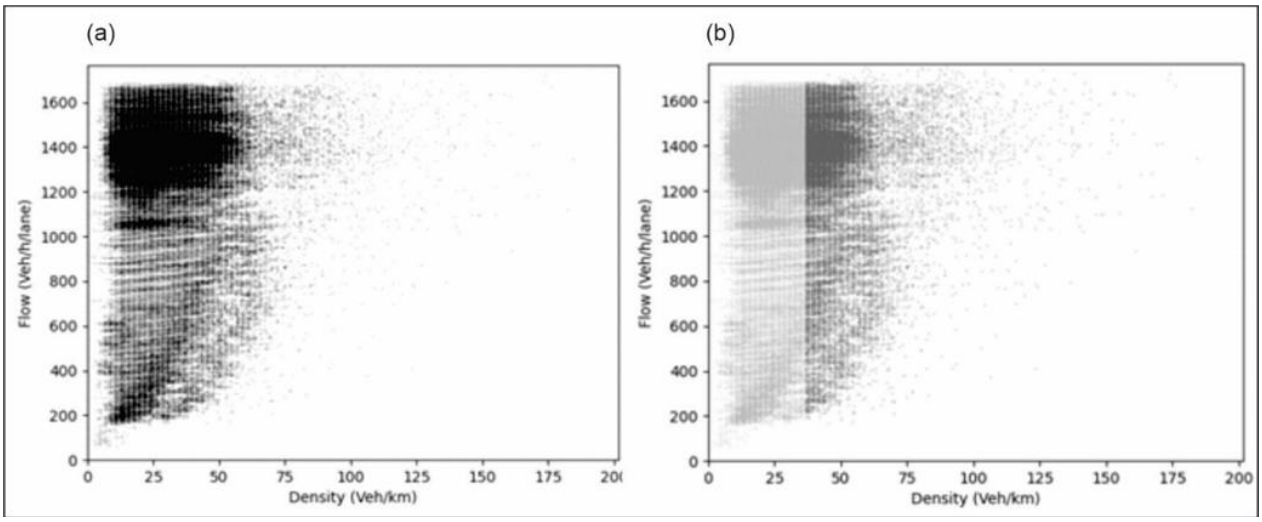

Figure 2: (a) Data presentation without clusters and (b) with clusters

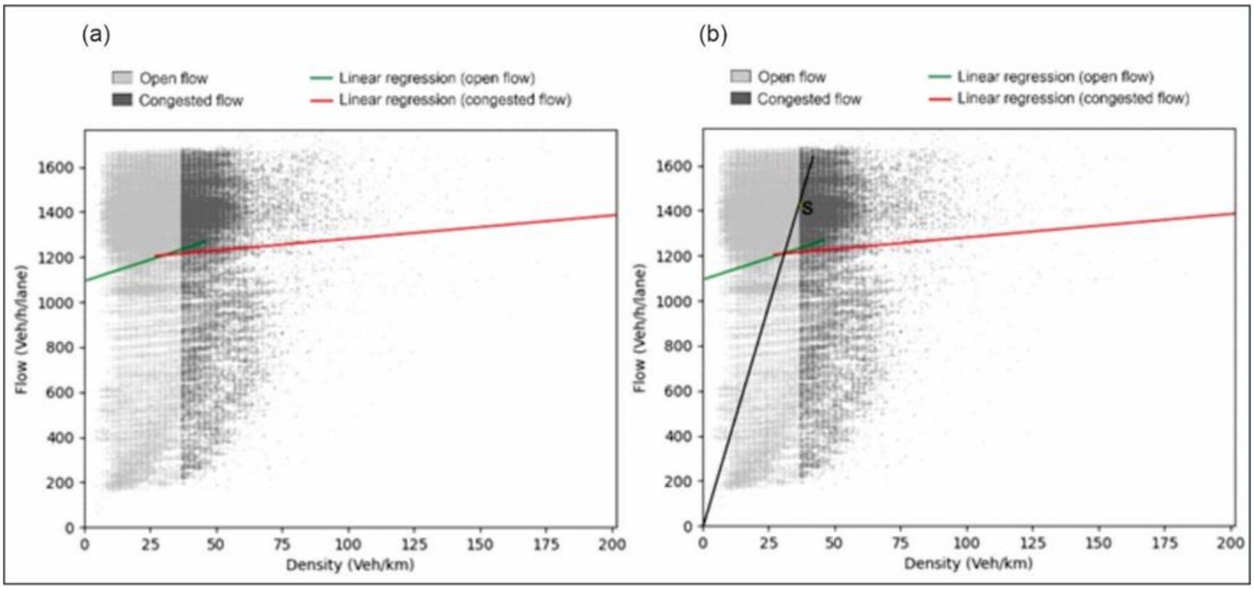

Figure 3: Geometric construction of point S

Regarding the validation of the presented results, Tableno 1 shows the measurements of the saturation flow using the Shanteau method [1] where the average of the evaluated saturation flow is 1,448 vehicles/hour/lane. According to the Kolmogorov-Smirnov test, the data adhere to a Normal Distribution (Figure 5) with 95% confidence and, considering the confidence interval of this distribution with 95% probability, the mean saturation flow measurement for this approach is about 1,452 vehicles/hour/lane excluding the value of 1,311 as it is outside the range with 95% of confidence.

The saturation flows measured according to the Highway Capacity Manual method [2] are given in Table no2 being then used to the second validation approach. The evaluated results by the HCM highlights the respective standard deviation and the sample mean. The measures inTable no2 were adjusted to a normal distribution as given in Figure 6. Therefore, it is evaluated that 68.3% of the saturation flows are between the two measures of 1,346 and 1,485 vehicles/hour/lane, resulting in the average saturation flow of 1,415 vehicles/hour/range.

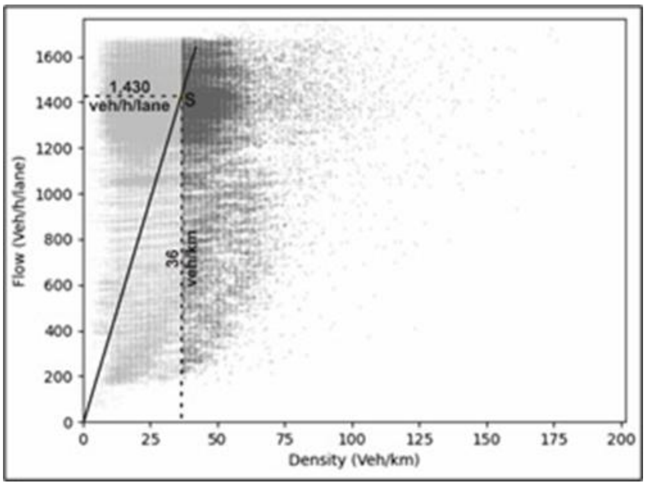

Figure 4: Saturation flow calibration

The saturation flow obtained by the method of Shanteau[1] applied to the base data - 1,452 vehicles / hour / lane - has a 70.23% probability according to the accumulated probability density curve derived from the fieldsurvey.The results for the saturation flows applying the Shanteau and the HCM methods are quite close when using collected data by the vehicle inspection devices or by the field survey. Furthermore, having determined thesaturation flows quite accurately, it is then possible to simulate the traffic flows through the intersections achieving a more valuable use of the transport infrastructure system.

Table 1: Saturation flow measurements from the data set.

No.	Flow	No.	Flow	No.	Flow
1	1,470	11	1,457	21	1,513
2	1,542	12	1,349	22	1,549
3	1,424	13	1,460	23	1,527
4	1,395	14	1,472	24	1,471
5	1,560	15	1,468	25	1,311
6	1,431	16	1,427	26	1,435
7	1,414	17	1,415	27	1,507
8	1,443	18	1,419	28	1,408
9	1,418	19	1,552	29	1,349
10	1,428	20	1,340	30	1,479

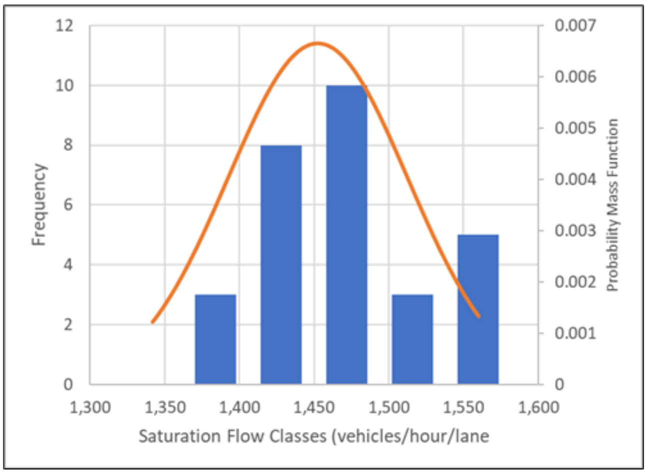

Figure 5: Histogram of saturation flow measurements being fitted by the normal distribution - Data set obtained by the electronic surveillance equipment

Table 2: Saturation flows measured (vehicles per hour per lane).

No	Flow	No	Flow	No	Flow	No	Flow	No	Flow	No	Flow
1	1.440	11	1.482	21	1.440	31	1.309	41	1.385	51	1.394
2	1.500	12	1.520	22	1.335	32	1.391	42	1.380	52	1.396
3	1.457	13	1.543	23	1.423	33	1.271	43	1.425	53	1.440
4	1.423	14	1.520	24	1.315	34	1.575	44	1.335	54	1.400
5	1.482	15	1.274	25	1.333	35	1.350	45	1.350	55	1.440
6	1.440	16	1.391	26	1.357	36	1.389	46	1.440	56	1.396
7	1.440	17	1.482	27	1.467	37	1.309	47	1.467	57	1.389
8	1.425	18	1.333	28	1.467	38	1.396	48	1.482	58	1.425
9	1.286	19	1.500	29	1.440	39	1.385	49	1.520	59	1.385
10	1.385	20	1.400	30	1.440	40	1.350	50	1.553	60	1.440
Meansaturationflow:				1.415,12 veh./hour/lane							
Stdofsaturationflow:				69,43 veh./hour/lane							

V. Conclusion and Recommendations

This research showed that the results of the survey of the saturation flow in the field and the procedure based on the method of Shanteau[1] are consistent and compatible with each other.

The traffic density was evaluated in the last kilometer when approaching the intersection aiming to mitigate any deficiency in the collected data by the local traffic agency. Therefore, a methodology was developed to calibrate the saturation flow in an accurate way to count the vehicles arriving at the intersection. The approach can be implemented in real time situations, mitigating the inconsistency inherent in the opinions and decisions of operators and traffic technicians.

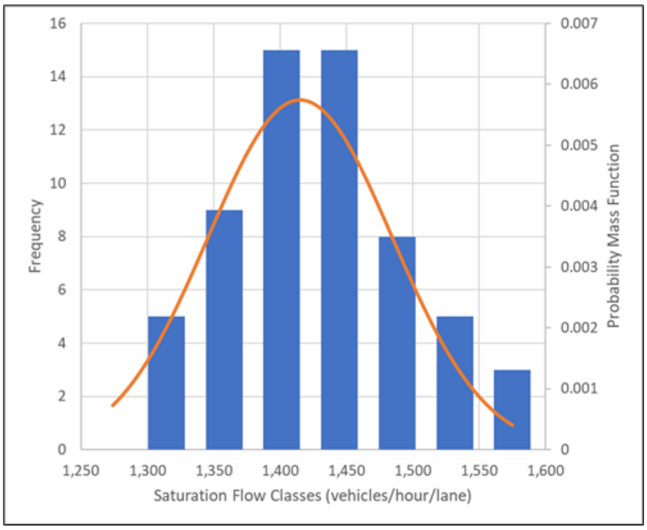

Figure 6: Histogram of saturation flow measurements versus fitted normal distribution for the sample of data collected by field researchers.

It is important to emphasize that the obtained results for the studied traffic light systems were validated in accordance not only by the observed traffic data but also by the HCM and the Shanteau methods [1][2]. The calibrated saturation flows have 58.49% of accumulated probability according to the realized field survey [2]. It represents also 79.4% of the maximum capacity of the arterial roads and the urban collectors of the analyzed traffic system.

It must be observed that the graphical results by the fundamental traffic flow diagram allow one to evaluate either the flow, or the density and the speed. A simulation can then be applied to the studied intersections based on the proposed methodology.

Finally, applying the graphical procedures to obtain the speed of the saturation flow and consequently the density and the flow, when collecting the real-time data, the procedure can be implemented dynamically. Since these approaches have been validated, a robust process has been presented where one can use for future calibration of the traffic light cycles.

References

[1]. Shanteau, R. M. 1988. Using Cumulative Curves ToMensure Saturation Flow and Lost Time. *ITE Journal*. 1988, Vol. 58, 10, pp. 27-31 *Manual, High Capacity*. Highway Capacity Manual. Transportation Research Board, Washington, DC, USA, 2000.

[2]. Manual, H. C. (2000). Highway capacity manual. *Washington, DC*, *2*.

[3]. McKenney, D., & White, T. (2013). Distributed and adaptive traffic signal control within a realistic traffic simulation. *Engineering Applications of Artificial Intelligence*, *26*(1), 574-583.

[4]. Garcia-Nieto, J., Olivera, A. C., & Alba, E. (2013). Optimal cycle program of traffic lights with particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, *17*(6), 823-839.

[5]. Tan, H., Wu, Y., Shen, B., Jin, P. J., & Ran, B. (2016). Short-term traffic prediction based on dynamic tensor completion. *IEEE Transactions on Intelligent Transportation Systems*, *17*(8), 2123-2133.

[6]. Verma, R., Paygude, P., Chaudhary, S., &Idate, S. (2018, February). Real Time Traffic Control Using Big Data Analytics. In *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)* (pp. 637-641). IEEE.

[7]. Souza, A. M., & Villas, L. A. (2015). A new solution based on inter-vehicle communication to reduce traffic jam in highway environment. *IEEE Latin America Transactions*, *13*(3), 721-726.

[8]. Biswas, S., Chakraborty, S., Ghosh, I., & Chandra, S. (2018). Saturation flow model for signalized intersection under mixed traffic condition. *TransportationResearch Record*, *2672*(15), 55-65.

[9]. Luna, M. D. S. D. (2003). *SOBRE O FLUXO DE SATURAÇÃO: CONCEITUAÇÃO, APLICAÇÃO, DETERMINAÇÃO E VARIAÇÃO* (Masters dissertation, UNIVERSIDADE FEDERAL DO CEARÁ)

[10]. Oliveira Neto, F. M., Ary, M. B., de Oliveira, M. V. T., & Loureiro, C. F. G. (2006). Atualização de Planos de Tempo Fixo com Base em Dados Históricos de Sistemas Centralizados de Controle de Tráfego. In *Anais do XX Congresso de Pesquisa e Ensino em Transportes, ANPET*.

[11]. Wang, L., Wang, Y., &Bie, Y. (2018). Automatic estimation method for intersection saturation flow rate based on video detector data. *Journal of Advanced Transportation*, *2018*.

[12]. Chand, S., Gupta, N. J., & Velmurugan, S. (2017). Development of saturation flow model at signalized intersection for heterogeneous traffic. *Transportation Research Procedia*, *25*, 1662-1671.

[13]. Luo, W., Wu, Y., Yuan, J., & Lu, W. (2016, November). The calculation method with grubbs test for real-time saturation flow rate at signalized intersection. In *International Conference on Intelligent Transportation* (pp. 129-136). Springer, Singapore.

[14]. Stathopoulos, A., &Karlaftis, M. G. (2003). A multivariate state space approach for urban traffic flow modeling and prediction. *Transportation Research Part C: Emerging Technologies*, *11*(2), 121-135.

[15]. Sun, L., & Zhou, J. (2005). Development of multiregime speed–density relationships by cluster analysis. *Transportation Research Record*, *1934*(1), 64-71.

[16]. Kianfar, J., &Edara, P. (2013). A data mining approach to creating fundamental traffic flow diagram. *Procedia-Social and Behavioral Sciences*, *104*(0), 430-439.

[17]. Jiang, Z., & Huang, Y. X. (2009). Parametric calibration of speed–density relationships in mesoscopic traffic simulator with data mining. *Information Sciences*, *179*(12), 2002-2013.
[18]. Dervisoglu, G., Gomes, G., Kwon, J., Horowitz, R., &Varaiya, P. (2009, January). Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Transportation Research Board 88th Annual Meeting* (Vol. 15).
[19]. Gu, Z., Saberi, M., Sarvi, M., & Liu, Z. (2017). A big data approach for clustering and calibration of link fundamental diagrams for large-scale network simulation applications. *Transportation research procedia*, *23*, 901-921.
[20]. Li, J., & Zhang, H. M. (2011). Fundamental diagram of traffic flow: new identification scheme and further evidence from empirical data. *Transportation research record*, *2260*(1), 50-59.
[21]. Jiang, Z., Shubin, L. I., & Xiaoqing, L. I. U. (2012). Parameters calibration of traffic simulation model based on data mining. *Journal of Transportation Systems Engineering and Information Technology*, *12*(6), 28-33.
[22]. Ian, H. Witten, Eibe Frank and Mark A. (2011). Hall Data Mining: Practical Machine Learning Tools and Techniques. (3rd Edition). Morgan Kaufmann, 664.
[23]. Tan, P.; Steinbach, M., Karpatne, A. e Kumar, V. (2018) Introduction to data mining. (2ndEdition). Pearson, London, 864.
[24]. Santos, C. L. (2007). *FLUXO DE SATURAÇÃO DE INTERSEÇÕES COMPLEXAS CONTROLADAS POR SEMÁFOROS* (Doctoraldissertation, UNIVERSIDADE FEDERAL DO RIO DE JANEIRO).

Fabio de Souza P. Borges, et. al. "Methodology to Calibrate the Saturation Flow in Traffic Light Urban Intersections." *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, 17(6), 2020, pp. 13-20.