Cycle-by-cycle Delay Estimation at Signalized Intersections by using Machine Learning and Simulated Video Detection Data

2023 8th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) | 978-1-6654-5530-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/MT-ITS56129.2023.10241732

Ismet Goksad Erdagi
Graduate Student Researcher
University of Pittsburgh
Pittsburgh, PA, USA
ise6@pitt.edu

Nemanja Dobrota
Engineering Associate
Kittleson and Associates, Inc.
Washington, D.C., USA
ndobrota@kittelson.com

Slavica Gavric
Graduate Student Researcher
University of Pittsburgh
Pittsburgh, PA, USA
slg130@pitt.edu

Aleksandar Stevanovic
Associate Professor
University of Pittsburgh
Pittsburgh, PA, USA
stevanovic@pitt.edu

***Abstract*— Accurate estimation of delay is crucial for efficient traffic signal operations. Estimation of delay in the real-time manner using traditional loop detectors requires advanced detectors (in addition to stop-bar detection). In cases when this detection layout is not in place, delay estimates are approximated with a lower accuracy. Video detection is one of the most frequently deployed detection systems at signalized intersections in recent years. In most cases video detection operates in the same way as traditional inductive loops. However, when coupled with computer vision algorithms, video detection systems could be used to retrieve additional information (e.g., vehicular arrivals and departures) that cannot be taken out from the conventional systems (e.g., long stop-bar loop detectors). Although present for several decades, video detection data were not frequently examined for delay estimation purposes. In this study, we proposed a novel delay estimation model which can be developed with only data from stop-bar video detectors. Relevant data were collected from a simulation model of 11 signalized intersections at downtown Chattanooga, TN and processed to create needed inputs for model development. With the use of multigene genetic programming the authors developed a delay model that outperforms accuracy of multi regression model. Furthermore, authors evaluated the developed model by comparison with the other benchmark delay models, such as HCM and approach delay model. It was found that the developed MGGP delay model outperforms benchmark models for a wide range of traffic and signal operation conditions.**

Keywords— performance measures, delay, machine learning, traffic, video detection

I. Introduction

Urban areas are becoming increasingly affected by traffic congestion, which impacts the movements of billions of people, waste their time and energy, and significantly contributes to air pollution. The negative impact of traffic congestion at signalized intersections can quantitatively be measured by vehicle delay. Vehicle (control) delay is the additional travel time due to traffic signal control and is one of the most important traffic signal performance measures [1]. The importance of delay can be seen in its use to support real-time traffic signal performance monitoring [2], signal timing design [3], and signal timing optimization[4]. Thus, accurate estimation of control delay is crucial for efficient traffic signal operations.

Due to challenges of the field delay estimations (i.e., resources expensive, labor-intensive practice), analytical delay models have been widely used for estimation of delay [5]–[7]. Since early work by Clayton [8] and Webster [7], where the first analytical delay models were proposed, numerous studies were conducted to improve the accuracy of delay models [5], [6], [9]. Several delay models were developed recently by using machine learning (ML) techniques to overcome some limitations of traditional analytical models [10]–[13]. Although superior in their performance, when compared to the traditional models, the applicability of ML-based delay models depends on the specifics of locations where data for model development were collected. Interestingly, even though a fair number of traffic signals are coordinated, recently proposed ML models were developed with data from isolated intersections that do not depict progression effects. Thus, their applicability for coordinated movements may be considered questionable.

Recently, with the emergence of better sensor technologies, our industry proposed several delay estimation methods/models based on the data from the fixed sensors at intersections that allow for real-time delay estimation [14]. These models essentially use data from detectors (occupied, non-occupied) and traffic signal status (green, red, yellow), to generate arrival departure profiles and estimate delays. It was shown that the majority of these models [14] depend on the use of a specific detection layout. Particularly, advanced detectors (placed several hundred feet upstream from the stop bar) and stop-bar detectors are used to estimate delays. When such detector configurations do not exist, delay estimates are simplified. In particular, when using only stop-bar detection, the time difference between the first vehicle arrival time at stop bar detector during red and the beginning of green is used as a surrogate measure to delay, called approach delay [15]. Alternatively, many researchers use CV data to estimate delays. However, these datasets are still not widely available, and they have not been integrated into traffic signal performance measurement platforms.

In recent years, video cameras, with the help of computer vision algorithms, have allowed us to enhance traditional stop-bar detection which records only times when detectors are occupied or not. Such conventional detection technologies impose several limitations; for instance, departure time and number of departures from the stop-bar

978-1-6654-5530-5/23/$31.00 ©2023 IEEE

detector, which is occupied by multiple vehicles, are not collected (unless loop stop-bar detectors are of a specific configuration or there is a separate exit detector). Similarly, by using traditional loop detectors, it may be challenging to track the arrival times of multiple vehicles and their number (it is usually not distinguished whether the detector is occupied by one or multiple vehicles). However, advanced video detection systems allow us to track each vehicle's entrance time in the detection zone (which length is also easily configurable), departure time from the detection zone, and type of vehicle detected [16], [17]. These emerging data sets are used for traffic detection by many transportation agencies around the US, primarily to count/classify vehicles. Although superior in terms of data reported, video-based loop detectors have not been extensively explored for reliable delay estimation.

This paper presents a machine learning-based estimation model for a vehicular delay that solely utilizes emerging video stop-bar detection data. The authors relied on the use of ML modeling because these models can track patterns from the data, which is very beneficial in this case (when partial information about arrivals is available). Data were collected from a microsimulation model of the real-world arterial located in downtown Chattanooga, Tennessee, US, and processed by multiple Python scripts since such an approach provides reliable ground truth delay values (instead of observing delay from traffic cameras for several hours). Results from the proposed models were validated against traditional analytical models (i.e., HCM6) and existing real-time delay estimation practice using only advanced detectors (i.e., approach delay [18].

The remainder of the paper is structured as follows. First, we provide a literature review on vehicle delay estimation models and other traffic estimation models developed with machine learning tools. Then, we present an overview of the methodology used in this research. In methodology, we explain the main differences between traditional loop stop-bar and emerging video detectors from the perspective of delay estimation and the main parameters considered to develop the delay model. Subsequently, data collection efforts and model development are presented. Later, the results are provided and accompanied by a detailed discussion and comparison of the developed model with benchmark delay models. Finally, the concluding remarks are presented with future research possibilities.

II. LITERATURE REVIEW

The first pioneering work on the development of delay models [8], [19] was done by Wardrop and Clayton. Shortly after, it was noticed that the delay estimates are not realistic or accurate due to the underlying model assumptions that the vehicular arrivals are uniformly distributed. The most famous stochastic delay, developed in 1958, is the Webster model [7]. However, delay models were questioned for accuracy in oversaturated conditions ever since. Several delay models were developed to account for both under and oversaturated conditions [6], [20] and show decent performance for delay estimations in so call time-dependent traffic profiles. In 2006, Incremental Queue Accumulation (IQA) model was proposed to overcome various limitations of the previous models [20].

Several research efforts were conducted in the past to estimate delay based on real-time signal and traditional loop detection high-resolution data. Sharma et al. proposed Input-Output delay estimation technique [14] and validated it against delays collected in the field. It was found that IO does not always provide good estimates compared to the field-collected values, mainly during peak-hour conditions when traffic may be saturated [14]. Day and Bullock also estimated analytical delay by using high-resolution data [21]. Estimated delays were not compared with ground truth values, and thus, the potential of such a method was not fully validated. Most studies rely on advanced detector data regardless of the method used.

To overcome the limitations of the previous estimation models several delay models were developed recently using machine learning (ML) techniques. Murat and Baskan developed an estimation model for average delay at an isolated signalized intersection by using an artificial neural network approach [22]. Ban et al. proposed to estimate traffic delay using travel time sampled from mobile sensors [23]. Korkmaz and Akgungor developed a differential evolution delay estimation model [24]. Bagdatli developed three different delay estimation models leveraging the gene expression programming method [25]. In a more recent study from 2021, Bagdatli and Dokuz developed delay estimation models by using four different machine learning methods (support vector regression (SVR), such as random forest (RF), k-nearest neighbor (kNN), and extreme gradient boosting (XGBT)) [26]. Dobrota and Stevanovic developed an MGGP model for the delay of protected/permitted left-turning vehicles [27].

To summarize literature findings: several studies have estimated vehicle delay by using the advanced detector data or in the connected vehicle environment, but none of the studies estimated delay in the cases when only data from stop-bar video detection are available from a coordinated corridor. The main contribution of this study is the development of a reliable delay estimation model for coordinated corridors based on limited data available from the stop-bar detectors.

III. METHODOLOGY

In the following chapters, we describe specifics of emerging stop-bar video data, how the delay is estimated in real-time at a signalized intersection, details about modeling practice multigene genetic expression, and lastly formulation of the delay model based on available video data.

A. Emerging stop-bar video detection data

Before we introduce novel video detection data, let us discuss traditional loop detectors. Fig. 1, part *a),* illustrates cases of traditional loop detection capabilities. For instance, when vehicle 1 approaches the detector, occupancy status is equal to 0. Once vehicle 1 arrives at the detection zone, that time instance is recorded (t_{a_1}) and occupancy status changes from 0 to 1. In case when the second vehicle arrives at the detection zone that is occupied by vehicle 1, no changes will occur in occupancy (since the detector is already occupied); also, time instance when the second vehicle arrives is not recorded. In case when first two vehicles start departing the detection zone, and the third vehicle joins (and the detector is still being occupied), information about the third vehicle's arrival is not recorded, nor is the departure time of the first vehicle.

Let us observe the same case with novel video detection illustrated in Fig. 1, part *b)*. In case when the first vehicle approaches the detection zone, detector occupancy will change (from 0 to 1) and time instance (t_{a_1}) will be recorded. Further, when the second vehicle arrives at the stop-bar detector and first vehicle is fully stopped, unlike traditional loop detection, novel video detection will report the arrival time of the second vehicle (t_{a_2}) as well. When the first vehicle starts departing the detection zone and the third vehicle is arriving (occupancy is still 1), novel detection will report the departure time of the first vehicle as well as the arrival time of the third vehicle (t_{a_3}). This simple example shows how novel video detection data are superior to traditional loop detectors. The following subchapter provides a discussion about these two data sources' differences from the real-time delay estimation perspective.

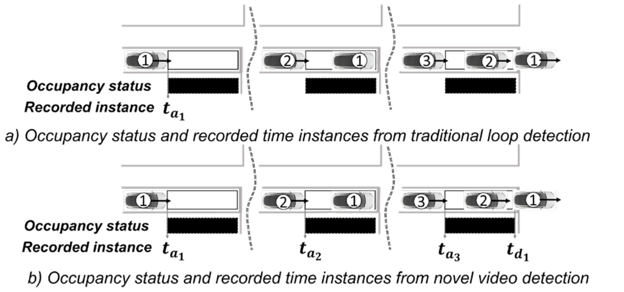


Fig. 1. Available data from traditional loop (a) and novel video (b) detection.

B. Delay estimation using real-time detection data

The most prominent method for delay estimation using real-time loop detection data is based on arrival and departure profiles estimated using advance detection and signal status data. The recorded arrival times estimated from advance detectors are used to estimate arrival profiles (i.e., queue formation) and for departure profiles (i.e., queue dissipation), assumed saturated headway is used (or measured if exit count detectors are available) to create delay polygon. The total delay is estimated based on constructed delay polygon. In cases when advanced detectors are not in place, our practice estimates the time to service, a surrogate measure for the delay, which is a time difference between the first vehicle arrival time at the intersection during red and the beginning of green.

As pointed out, to estimate total delay, it is necessary to estimate both the arrival profile and departure profiles. Therefore, we developed Fig. 2, 3, and 4 to illustrate differences in terms of available data from traditional loop detection and novel video detection data from a delay perspective. Fig. 2 illustrates available data from traditional and novel video detection data, on the left and right sides of the figure, respectively. Black dots in Fig 2. represent data that is possible to collect, whereas grey dots represent data that are not collectible. As it can be seen, loop detection allows us to track the moment when the first vehicle arrived at the detector zone (t_{a_1}) whereas for others ($t_{a_1} - t_{a_6}$), such information is not obtainable. The time difference between the beginning of green and the first arrival on red is also known as the time to service (*TTS*). With novel video data, one can notice that data about 4 arrivals ($t_{a_1} - t_{a_4}$) are collected (number of collected arrivals depends on configured detection zone length), whereas for others ($t_{a_5} - t_{a_6}$), beyond the detection zone, such information is not

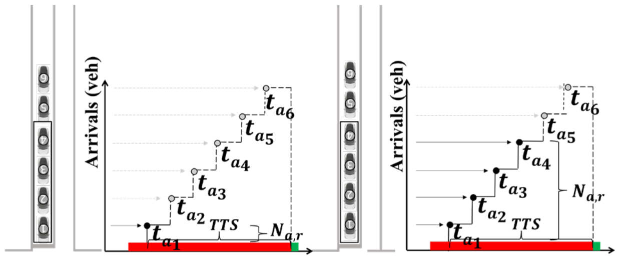


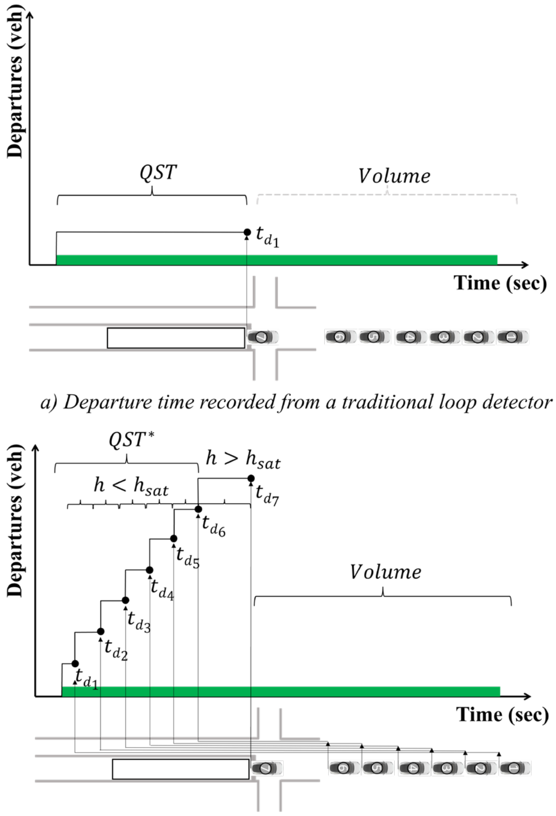
Fig. 2. Arrival profile at downstream intersection measured using tradition loop data (left) and novel video data (right).

collected (during red). Although video data do not provide a complete picture of all possible arrivals during red, partial information about queue length is obtainable. We label the number of detected arrivals on red, as $N_{a,r}$.

Another difference between traditional loop detectors and novel video data can be seen in the reporting capabilities during the queue discharge process. Fig. 3 shows data reported from traditional detection and novel video detection, parts *a)* and *b),* respectively. In the illustrated example, when 6 vehicles were queued during red, and subsequently released during the green, and the 7th vehicle arrived before the loop detector was unoccupied, the only reported departure is when the detector becomes unoccupied i.e., after the 7th vehicle. The time that passes from the beginning of green and the time when a continuous detector call is extinguished is called Queue Service Time (QST) [28]. In some cases, volume data can be retrieved from only stop-bar detector data if the technology is mature enough [16]. However, in cases when such loop detectors are not in place, exit detectors are used for the same purpose [29].

Novel video data is capable of reporting time instances when each vehicle departs the detection zone, regardless of occupancy status. One can notice in part b) of Figure 3, that for the same configuration of departures (illustrated in part a), one can reveal much more information about traffic operations. QST can be estimated based on saturation headway (h_{sat}), contrary to limited occupancy status. Thus, we label QST that can be derived based on saturation headway as QST^*, i.e., when saturated headway is higher than measured headway between two vehicles (h), we assume that the queue is discharging. It can be seen from this illustration that modified QST^* can improve the estimation of the number of queued vehicles. Additionally, volume data are fully reported from video detection.

Finally, one can approximate the queue polygon based on reported data from video detectors, as illustrated in Fig. 4 Such approximation is difficult to be performed analytically considering that the arrivals during red cannot be completely tracked down (as illustrated with dashed lines in Fig. 4). However, the authors of this study relied on machine learning techniques performing this task, considering that such a modeling practice, depending on a provided dataset (number of cycles with various delays, arriving, and discharging patterns), can provide a more realistic delay model. Considering computational efficiency and the ability to control model complexity during development, the authors of this study decided to use the MGGP technique to develop a delay model.



a) Departure time recorded from a traditional loop detector

b) Departure time recorded from a novel video detector

Fig. 3. Relationship between available data and total delay.

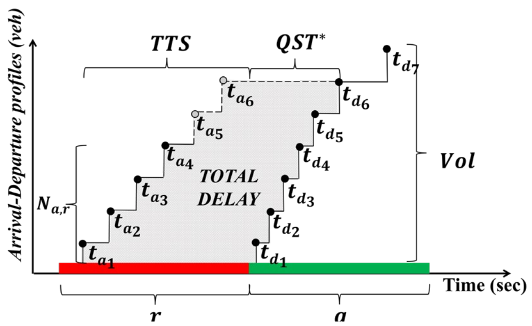


Fig. 4. Relationship between available data and delay.

Based on this powerful technique, we propose the formulation of the delay model as a function of parameters that can be easily obtained from novel video detection data and were discussed in some detail in previous sections. The proposed formulation of delay is as follows:

$$Delay = f\left(N_{a,g}, N_{a,r}, TTS, QST^*, \frac{g}{C}, Volume, r, C\right) \quad (1)$$

where all variables were previously defined and their importance for delay estimation. In the following chapters, we discuss data collection efforts and the development and evaluation of the proposed delay model.

C. Data collection

For this study, we use PTV VISSIM, a widely used traffic microsimulation tool, to collect simulated traffic data on a cycle-by-cycle basis. One of the busiest roadways in downtown Chattanooga, MLK Boulevard, was modeled and used for data collection (see Fig. 5). MLK Boulevard network was previously described by Harris et al. in [30] where video analytics are detailed. In this paper, we mimicked the video detection zones in the simulation environment and analyzed the simulation output through Python scripting. By using simulation data, the authors assume that video detection performs flawlessly without any interruptions or false/missed calls.

Fig. 5. Subject network.

D. Data collection and model development

IV. Development of MGGP Model

In this section, a multi-gene genetic programming (MGGP) approach is employed. MGGP is a machine learning method used to create mathematical equations that can predict outcomes based on input variables. The MGGP creates a population of mathematical equations, which include input variables, and then evolves their forms through a process of natural selection similar to how living organisms evolve over time. The most successful equations, those that best predict the outcomes given by the input data, are selected to "breed" and create new equations in the next generation. This process is repeated until the algorithm finds the equation that best fits the input data. Once the equation is found, it can be used to make predictions on new data. The steps followed in this study to find the best MGGP delay models are as follows:

a. The input parameters impacting the delay are selected.
b. Simulation data for model development is prepared and organized.
c. The collected field data is divided into training, testing, and validation datasets. The training dataset is taken for the learning process. The testing dataset is used to evaluate the developed model after training. The validation dataset is used to further measure the performance of the developed model and compare it with benchmark delay models.
d. MGGP is run on the training data to find a mathematical model that connects the independent variables to delay.
e. The best MGGP model is chosen considering both its simplicity and the best performance on the training and testing data.
f. The model is validated by comparing it with benchmark delay models (HCM and Approach Delay).

The setting of the parameters significantly influences the generalization and accuracy of the final MGGP model. Therefore, a trial-and-error approach was adopted to find the optimal parameter setting. A range from 500 – 2000 was considered for the program's population size and number of generations. The maximum allowable number of genes and the maximum tree depth determine the complexity of the final models. Based on the values recommended in the previous studies [31], a practical range was considered for each parameter to find the best setting value. The training parameter settings are presented below.

V. RESULTS AND DISCUSSION

In this section, the results for the best performing MGGP delay model are shown. The population size, maximum number of genes, tournament size, and maximum tree depth for the best developed MGGP model are 500, 800, 80, 6, and 4, respectively. As mentioned in previous studies [31], the number of genes and the size of tree depth impact the size of search space. Therefore, to avoid overfitting and reducing the algorithm's speed the authors limited tree depth to 6.

A. MGGP delay model performance

The equation of the best obtained MGGP model is provided in (2). It should be noted that values from x_1, …, and x_8 stand for the independent input variables $(N_{a,g},\ N_{a,r}, TTS, QST^*, \frac{g}{C}, Volume, r, C)$.

$$Delay = x_1 x_2{}^2 x_3 x_4 (x_5 - x_6) \cdot 3.717 \cdot 10^{-5} + x_1 x_2 x_4{}^2 x_7 \cdot 1.326 \cdot 10^{-4} - 20.46 x_1 + 19.26 x_6 - \frac{19.26 x_3}{x_7} + \frac{0.2159 x_7}{x_5} - \frac{x_1 x_3{}^2 (x_5 - x_6) \cdot 1.832 \cdot 10^{-2}}{x_5 x_8} - 3.032 \qquad (2)$$

Pareto front of MGGP models is provided in Fig. 6 GPTIS provides the pareto front models based on the trade-off between accuracy and complexity. In Fig. 6, the best model, which has the highest R^2 value is presented as a blue dot with a red circle.

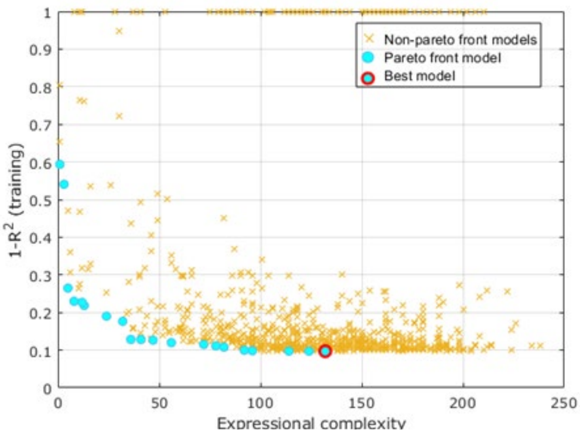


Fig. 6. Pareto front of MGGP models.

In Fig. 7, the authors show the performance of the best MGGP model on training and testing datasets. Models show high accuracy on both datasets.

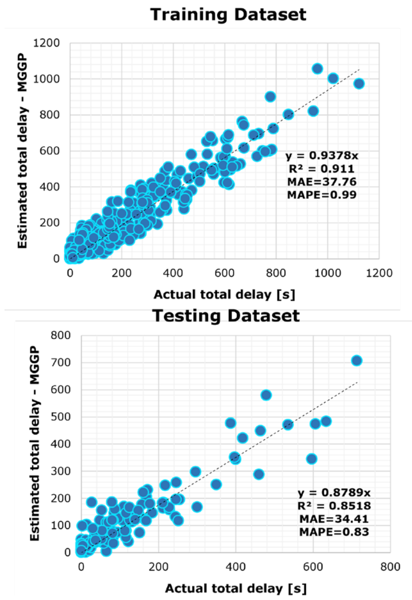


Fig. 7. Performance of best MGGP model on training and testing datasets.

B. Comparison of HCM, Approach delay and MGGP delay model

To compare the developed model with widely known models (HCM and approach delay) we exposed all three models to a completely new unseen dataset (validation) containing 150 cycles. Fig. 8 represents comparison results and the accuracy of each model.

The proposed model performs significantly better than the two others as can be noticed by observing various statistical parameters. One should remember that HCM uses a Percent on Green (POG) in its estimation. Thus, the HCM cannot provide accurate delay estimates for corridors that are equipped with video detection systems (where percent on green is an unknown parameter). Also, approach delay strongly relies on time to service (*TTS*) and in most cases, its accuracy is questionable, especially when analyzing coordinate networks where, for example, one vehicle arrives on red from the upstream intersection's side street and causes high *TTS* whereas all other vehicles arrive as a platoon and do not experience delay due to coordination favoring major through movements. Thus, it can be noticed that the assumption that all vehicles departing from an intersection wait on average, for a time equal to *TTS* is a hard assumption and can be only accurate in certain situations.

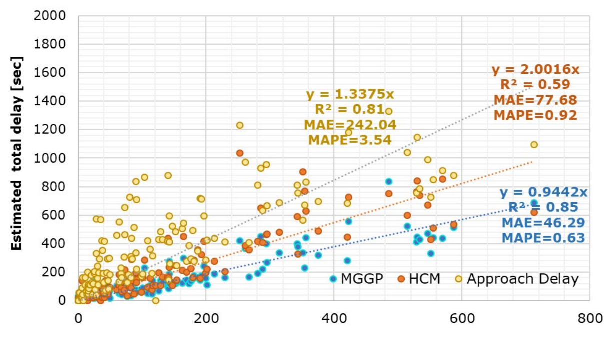


Fig. 8. Accuracy comparison of MGGP model, HCM and Approach delay using the validation dataset.

To get a better understanding of the scatter plot in Fig. 9 we arbitrarily selected 50 cycles and calculated the average delay (total delay divided by volume) to show the estimation accuracy of MGGP, HCM, and the Approach Delay models and presented this information in Fig. 9. Per Fig. 9, MGGP is performing significantly better than HCM and Approach delay.

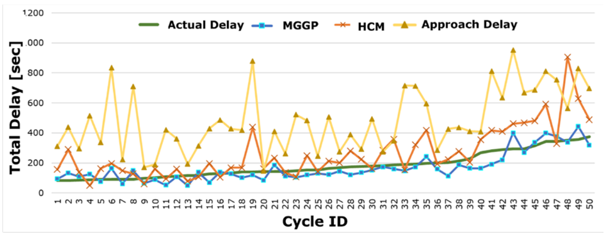


Fig. 9. Accuracy comaprison of MGGP model, HCM, and Approach delay using the validation dataset.

VI. CONCLUSIONS

In this study, the authors proposed a novel delay model, which serves to overcome the limitations of existing models that currently cannot estimate total delay based only on stop-bar video detection data. For this purpose, multigene genetic programming (MGGP) was utilized, and the most important findings are as follows:

- Total delay per cycle can be reliably estimated based on the video detection stop-bar data.
- When compared to the HCM and approach delay models, the novel model performs significantly better which results in a higher delay estimation accuracy.
- MGGP provides transparent and structured equations which can be used in the future on various corridors to estimate total delay.
- The major advantage of the developed MGGP model when compared to the benchmark models is that it does not require percent (of arrivals) on green. For example, the HCM's model requires information about the percentage of arrivals on green. However, such data are not always available.

The developed delay model fills the gap in the current body of knowledge by reliably estimating the total delay per cycle at the corridors where emerging video detection systems are employed, without advanced detectors. The authors acknowledge the limitations of this approach, and potential risks and uncertainties associated with relying solely on simulated data and suggest that future research investigates how the proposed model performs on various corridors with real video footage. Future research should also investigate different possibilities of utilizing powerful machine learning tools on emerging video detection data. Finally, future research should investigate development of optimal MGGP model and integration of constraints pertaining to combinations of arithmetic operations between input variables during model training. This approach would enhance the accuracy and effectiveness of the MGGP method in practical applications.

ACKNOWLEDGMENT

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Award Number DE-EE0009208.

REFERENCE

[1] F. Dion, H. Rakha, and Y.-S. Kang, "Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections," *Transportation Research Part B: Methodological*, vol. 38, no. 2, pp. 99–122, Feb. 2004, doi: 10.1016/S0191-2615(03)00003-1.

[2] P. J. Jin, T. Zhang, T. M. Brennan Jr, and M. Jalayer, "Real-Time Signal Performance Measurement (RT-SPM)," Art. no. FHWA NJ-2019-002, Jul. 2019, Accessed: Jul. 08, 2022. [Online]. Available: https://trid.trb.org/view/1681169

[3] T. Urbanik *et al.*, *Signal Timing Manual - Second Edition*. Washington, D.C.: Transportation Research Board, 2015, p. 22097. doi: 10.17226/22097.

[4] J. Stevanovic, A. Stevanovic, P. T. Martin, and T. Bauer, "Stochastic optimization of traffic control and transit priority settings in VISSIM," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 3, pp. 332–349, Jun. 2008, doi: 10.1016/j.trc.2008.01.002.

[5] HCM, *Highway Capacity Manual*. Washington, DC: Transportation Research Board, 1985.

[6] Akcelik, *Traffic Signals: Capacity and Timing Analysis*. 1981.

[7] F. V. Webster, *Traffic signal settings*. London: H.M.S.O., 1958.

[8] Arthur John Hamblin Clayton, *Road Traffic Calculations*. Journal of the Institution of Civil Engineers, 1941.

[9] R. M. Kimber and P. N. Daly, "Time-dependent queueing at road junctions: Observation and prediction," *Transportation Research Part B: Methodological*, vol. 20, no. 3, pp. 187–203, Jun. 1986, doi: 10.1016/0191-2615(86)90016-0.

[10] A. P. Akgüngör and E. Korkmaz, "Bezier Search Differential Evolution algorithm based estimation models of delay parameter k for signalized intersections," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 13, p. e6931, 2022.

[11] N. Dobrota and A. Stevanovic, "Modelling of Delay for Protected/Permitted Left Turning Vehicles usingMultigene Genetic Programming," Nov. 2021. doi: 10.11159/iccefa21.134.

[12] M. E. C. Bagdatli and A. S. Dokuz, "Vehicle Delay Estimation at Signalized Intersections Using Machine Learning Algorithms," *Transportation Research Record*, vol. 2675, no. 9, pp. 110–126, Sep. 2021.

[13] M. E. C. Bagdatli, "Vehicle Delay Modeling at Signalized Intersections with Gene-Expression Programming," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 9, p. 04020107, Sep. 2020, doi: 10.1061/JTEPBS.0000423.

[14] A. Sharma, D. M. Bullock, and J. A. Bonneson, "Input–Output and Hybrid Techniques for Real-Time Prediction of Delay and Maximum Queue Length at Signalized Intersections," *Transportation Research Record*, vol. 2035, no. 1, pp. 69–80, Jan. 2007, doi: 10.3141/2035-08.

[15] M. Taylor, "UDOT ATSPM Train-the-trainer Workshop," *Dropbox*, 2017.

[16] M. S. Shirazi and B. Morris, "Vision-based vehicle queue analysis at junctions," in *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Karlsruhe, Germany, Aug. 2015, pp. 1–6. doi: 10.1109/AVSS.2015.7301732.

[17] G. Zhang and Y. Wang, "Machine Learning and Computer Vision-Enabled Traffic Sensing Data Analysis and Quality Enhancement," in *Data-Driven Solutions to Transportation Problems*, Elsevier, 2019, pp. 51–79. doi: 10.1016/B978-0-12-817026-7.00003-5.

[18] ATKINS, "Automated Traffic Signal Performance Measures Reporting Details," p. 65, 2020.

[19] J. G. Wardrop, "Road paper. some theoretical aspects of road traffic research.," *Proceedings of the Institution of Civil Engineers*, vol. 1, no. 3, pp. 325–362, May 1952, doi: 10.1680/ipeds.1952.11259.

[20] N. M. Rouphail, K. G. Courage, and D. W. Strong, "New Calculation Method for Existing and Extended Highway Capacity Manual Delay Estimation Procedures," 2006. Accessed: Jul. 08, 2022. [Online]. Available: https://trid.trb.org/view/776181

[21] C. M. Day and D. M. Bullock, "Computational Efficiency of Alternative Algorithms for Arterial Offset Optimization,"

Transportation Research Record, vol. 2259, no. 1, pp. 37–47, Jan. 2011, doi: 10.3141/2259-04.

[22] O. Baskan and Y. Murat, "Modeling vehicle delays at signalized junctions: Artificial neural networks approach," *Journal of scientific and industrial research*, vol. 65, pp. 558–564, Jul. 2006.

[23] P. Hao and X. (Jeff) Ban, "Long queue estimation for signalized intersections using mobile data," *Transportation Research Part B: Methodological*, vol. 82, pp. 54–73, Dec. 2015.

[24] E. Korkmaz and A. P. Akgüngör, "Delay estimation models for signalized intersections using differential evolution algorithm," *Journal of Engineering Research*, vol. 5, no. 3, Art. no. 3, Nov. 2017, Accessed: Jul. 08, 2022. [Online]. Available: https://kuwaitjournals.org/jer/index.php/JER/article/view/1566

[25] M. E. C. Bagdatli, "Vehicle Delay Modeling at Signalized Intersections with Gene-Expression Programming," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 9, p. 04020107, Sep. 2020, doi: 10.1061/JTEPBS.0000423.

[26] M. E. C. Bagdatli and A. S. Dokuz, "Vehicle Delay Estimation at Signalized Intersections Using Machine Learning Algorithms," *Transportation Research Record*, vol. 2675, no. 9, pp. 110–126, Sep. 2021, doi: 10.1177/03611981211036874.

[27] N. Dobrota and A. Stevanovic, "Modelling of Delay for Protected/Permitted Left Turning Vehicles usingMultigene Genetic Programming," Nov. 2021. doi: 10.11159/iccefa21.134.

[28] K. N. Balke, H. Charara, and R. Parker, "Development of a Traffic Signal Performance Measurement System (TSPMS)," Art. no. FHWA/TX-05/0-4422-2, May 2005, Accessed: Jul. 15, 2022. [Online]. Available: https://trid.trb.org/view/759658

[29] S. F. Smith, G. Barlow, X.-F. Xie, and Z. B. Rubinstein, "Smart Urban Signal Networks: Initial Application of the SURTRAC Adaptive Traffic Signal Control System," in *ICAPS*, 2013.

[30] A. Harris, J. Stovall, and M. Sartipi, "MLK Smart Corridor: An Urban Testbed for Smart City Applications," in *2019 IEEE International Conference on Big Data (Big Data)*, Dec. 2019, pp. 3506–3511. doi: 10.1109/BigData47090.2019.9006382.

[31] J. R. Koza, "Genetic programming as a means for programming computers by natural selection," *Stat Comput*, vol. 4, no. 2, pp. 87–112, Jun. 1994, doi: 10.1007/BF00175355.