J. Inst. Eng. India Ser. A
https://doi.org/10.1007/s40030-022-00698-2

ORIGINAL CONTRIBUTION

Introduction of a Simple Estimation Method for Lane-Based Queue Lengths with Lane-changing Movements

Shenura Jayatilleke[1] · Vasantha Wickramasinghe[2] · Niranga Amarasingha[1]

Received: 24 April 2022 / Accepted: 3 October 2022

Abstract Traffic congestions are increased globally due to rapid urbanization and expedited economic developments in many countries. Vehicle queue is a governing aspect of traffic congestion, studied over the past decades. Most of the existing queue estimation approaches are limited to homogeneous traffic conditions. However, the traffic conditions in many developing countries are heterogeneous and are heavily influenced by mixed vehicle composition, lane changing, and gap-filling behaviours. This study aims to estimate the queue length at signalized intersections having heterogeneous traffic conditions. The heterogeneity was assimilated with the consideration of Passenger Car Units (PCU) in the measurements of the traffic flow and the lane-changing movement within the considered road section. The influential factors of the queue length were contemplated with the arrival flow, discharge flow, outbound lane change, inbound lane change, and signal configuration. A Vector Auto Regression (VAR) model was developed to estimate queue length, with a lag time of 15 s for each variable. The results have indicated a higher accuracy in the queue estimation as well as the practical application for prediction, constituting the traffic characteristics of the formed vehicle queue. The R squared of the VAR model was 0.97, along with a Mean Absolute Percentage Error (MAPE) of 21.55%. The model estimation results of right turning lanes were well accurate with MAPE ranging from 15 to 17%, whilst for through movement lanes, accuracy was slightly low with MAPE in the range of 23–26%. The study manifests the functionality of the developed methodology for accurate queue estimations, asserting the practical applicability of VAR models in other locations constituting mixed traffic.

Keywords VAR model · Time series analysis · Mixed traffic · Signalized intersections · Queue prediction

Introduction

The investigation and evaluation of traffic behaviour help to demonstrate various perspectives of traffic predictions. Vehicle queue predictions are widely used in traffic management to predict future trends in the queue. The existing queue prediction approaches mainly scope within homogeneous traffic conditions, where car following and lane discipline are imminent. Thus, the use of such models in heterogeneous conditions reduces the accuracy of the results. The heterogeneous traffic flow is defined as a traffic stream containing various vehicles, either motorized or non-motorized [1]. Therefore, heterogeneous traffic belongs to different vehicle classes in terms of static characteristics such as shape and dimensions along with dynamic characteristics such as speed and acceleration. As mentioned by Verma et al. [2], heterogeneous traffic is dominated by different traffic behaviour and queuing behaviour. The sharing of the right-of-way (ROW) for the different vehicle classes, causing a difference in lateral and longitudinal directions, results in a weak lane discipline scenario. Therefore, weak lane discipline scenarios do not follow the First-In-First-Out (FIFO) or Last-In-First-Out (LIFO) queue disciplines. The car following models and other models based on lane discipline, cannot be rendered with such traffic behaviours.

The gap-filling behaviour is incurred mainly by smaller vehicles such as two-wheelers and three-wheelers, as they

✉ Shenura Jayatilleke
shenura.j@sliit.lk

1 Department of Civil Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

2 Department of Civil Engineering, University of Peradeniya, Peradeniya, Sri Lanka

Published online: 21 December 2022

percolate through the gaps between their larger neighbours (buses, trucks, etc.). As shown by Wickramasinghe et al. [3], smaller vehicles tend to approach the stop line faster and are served earlier than heavy vehicles. The smaller vehicles lead to the placement of vehicles at any given point, which largely ignores lane separation. In addition, the smaller vehicles may even create virtual lanes as they move forward and provide additional gaps to their own vehicles. Therefore, heterogeneous traffic conditions are very much different from homogeneous conditions, due to sudden lane-changing behaviours, weak lane discipline, and gap-filling behaviours. Consequently, FIFO queuing systems are inapplicable, and a novel approach with a queueing behaviour called Probably-First-In-Probably-First-Out (PFIPFO) should be incorporated. This was defined by Verma et al. [2], as a queuing behaviour in which a vehicle that has arrived before or after another given vehicle, may with some probability leave earlier than it.

The queue estimation methods have been mainly developed for signalized intersections. Seiran [4] classified the queue length estimation approaches as analytical queue estimation techniques and data-driven queue estimation techniques. The analytical techniques are the input–output technique and the shockwave-based approach. Whilst data-driven techniques were probabilistic probe vehicle-based positioning methods, Kalman-Filtering, Markovian Chains-Based models, direct video-based queue estimation methods, and neural network methods for queue prediction. According to Anusha et al. [5], the input–output method is fundamentally similar to the vehicles' conservation equation. Such that, the authors calculated the queue length of signalized intersections from the difference between the arrival vehicle number and the discharge vehicle number. The input–output methods have been used for more than seven decades as Newell [6] utilized it in the study for the queue approximation in a fixed cycle traffic light. In addition, Stephanopoulos & Michalopoulos [7], Sharma et al. [8] and Sieran [4] utilized a similar methodology of vehicle conservation by integrating the input–output method.

Comert [9] developed a stochastic model to estimate the cycle-to-cycle queue length from probe vehicle data and the results were within a $\pm 5\%$ error. However, the study did not predict future queue lengths. Hao et al. [10] developed seven cycle-by-cycle queue estimation models, using Bayesian network models. The data were collected from mobile traffic sensor data between the upstream and downstream of an intersection. The results show that the stochastic approach is more robust in low penetration rates, compared to deterministic approaches. However, the cycle-by-cycle approach in queue estimation is less applicable in the real world, as the queue variation is high within a cycle. Therefore, the model suffers from the lack of availability of actual ground truth data, as the queue is estimated only at a certain instant with this approach.

Zhan et al. [11] incorporated the Gaussian Process-based interpolation method and a car following model. The authors reconstructed the equivalent cumulative arrival—departure curve of each lane to estimate the queue length. The data collection was done using license plate data for each lane of the road. The Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) of the queue estimation was below 3.2 vehicles and 2.4 vehicles. However, the authors highlighted lane changing and inferring incorrect arrival times as the key limitations, as the data are from license plate capturing.

Zeng [12] predicted the queue length using a stochastic fluid theory. Therefore, two fluid theories were used considering the road traffic and congested traffic, to predict the queue length for single-lane and multi-lane scenarios. The average relative prediction error of the single-lane scenario was 24.7%, and for the multi-lane scenario was 38.2%. The higher prediction error percentage in the multi-lane scenario was justified by the authors with the lane-changing scenario. The higher error percentage for queue predictions in heterogeneous conditions was further highlighted by Li et al. [13]. They integrated the Lighthill–Whitham–Richards shockwave theory and Robertson's platoon dispersion model, to predict the arrival of vehicles, five seconds in advance for each lane and further integrated with the Kalman-Filter method to predict the queue length as the number of vehicles. The final queue prediction results of the maximum queue had an average RMSE of 2.33 vehicles, MAE of 1.82 vehicles, and Mean Absolute Percentage Error (MAPE) of 16.12%.

In contrast, Jayatilleke et al. [14] attempted to predict the queue length under heterogeneous conditions. The authors obtained a queue prediction model with an R squared of 0.724 and an MAE of 1.82 m. Even though the methodology quantified heterogeneous traffic as the mixed vehicle composition, lane changing was not incorporated. They concluded that lane changing is heavily attested in heterogeneous conditions, thus, should be dealt with accordingly. Additionally, they concluded that prediction accuracy for queue lengths above 50 m is reduced up to 40% due to lane-changing movements. Moreover, lane changing constitutes to violation of the FIFO phenomena of the input–output approach, consequently, increasing the percentage error of queue predictions. Thus, lane changing has a major influence on optimizing the accuracy of queue predictions under heterogeneous conditions.

Time series analysis has been used in econometrics to examine long-run and short-run dynamics among macroeconomic variables and stock exchanges. The studies were casual relationships among macroeconomic variables and stock prices in the Colombo Stock Exchange [15], the relation between the unemployment rate and stock prices in the

USA, China, and Japan [16], and the effect of macroeconomic variables on stock prices in Sri Lanka stock market [17], volatility models for the stock indices of Colombo Stock Exchange [18]. Therefore, time series analysis helps to identify short-run and long-run relationships, whilst the application of the time series method helps to assess the short-run and long-run relationships of queue length.

The present research study focuses on identifying the most influential characteristics of the vehicle queue. Moreover, the developed relationships between the vehicle queue length and other variables, help to estimate the queue accurately under heterogeneous conditions. Thus, the main objectives of the study were as follows,

- To quantify the governing factors of heterogeneous traffic conditions
- To develop an accurate lane-based queue estimation model under heterogeneous conditions
- To validate the developed queue estimation model, by applying it to a different location in contrast to the road geometry and traffic characteristics

The research methodology was developed by accounting for the heterogeneity of traffic in terms of lane changing and the mixed vehicle composition. The research was focused on developing a time series equation and its application, to predict the queue length at signalized intersections in heterogeneous traffic conditions. Time series analyses are mainly incorporated in econometrics where the variables rely on time. Therefore, this study explicates the queue length and the governing factors in the form of a generalized Vector Auto Regression (VAR) model and further predicts the vehicle queue at signalized intersections under heterogeneous traffic conditions.

Methodology

Data Collection

Study Area for Model Development

The model development study area was selected to represent the heavy heterogeneity of vehicle traffic. The study area (Location 1) for queue prediction is the Armour-Street Junction, which is a three-phased, four-legged junction as shown in Fig. 1. The considered leg (Maradana end) for the queue prediction was a three-lane road section, spanning a length of 120 m. The data were recorded from drone cameras on a weekend from 2.30 to 6.30 pm.

The geometric characteristics of the location are shown in Table 1.

Study Area for Queue Prediction

The developed model was applied to a different location in Thalawathugoda, Sri Lanka (Location 2), to further verify the model applicability as real-time queue predictions. Location 2 is a four-lane signalized junction, with one right turning lane, two through lanes, and one left turning lane as shown in Fig. 2. However, the left turning lane was disregarded, as it is not signal controlled. Two video cameras were used, to cover a total length of 100 m in the

Table 1 Geometric characteristics of location

Parameter	Lane 1	Lane 2	Lane 3
Location of lane	Lane near curb	Middle lane	Lane near centre median
Width of lane (m)	3.8	3.5	3.2
Turning movement	Through and left turn	Right turn	Right turn

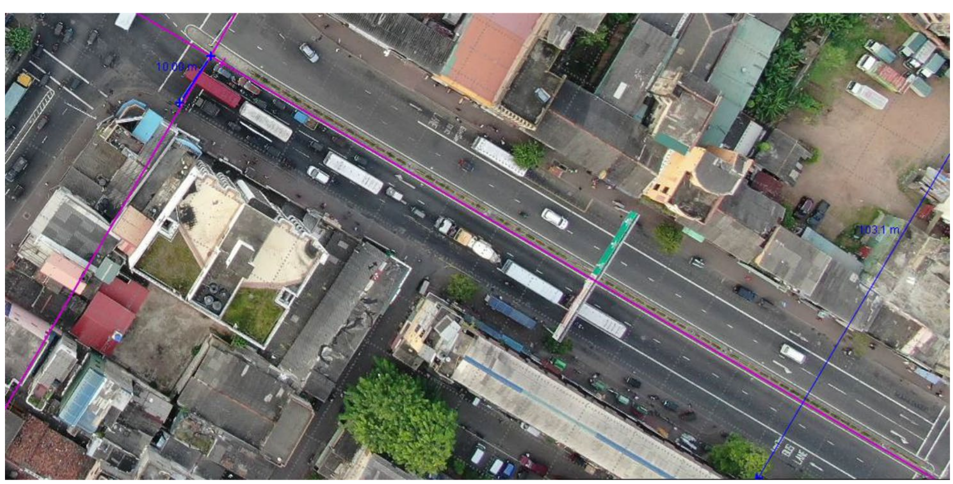


Fig. 1 Study area for model development (Location 1)

Fig. 2 Study area of real-time queue prediction (Location 2)

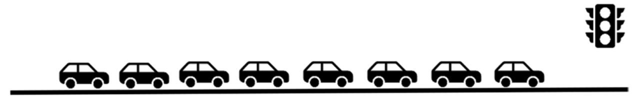


Fig. 3 Stopped queue

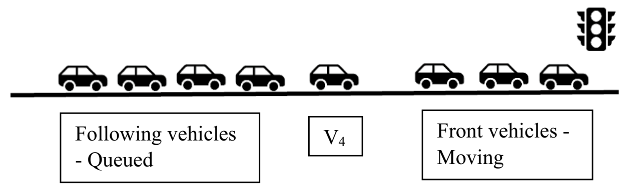


Fig. 4 Partially moving queue

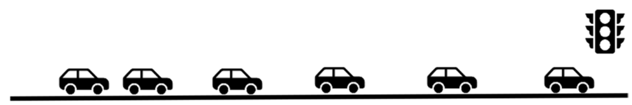


Fig. 5 Moving queue

Thalawathugoda–Pannipitiya leg of the Junction, capturing from 7.30 to 11.30 am on a weekday.

Data Extraction

The methodology was developed considering the four main aspects: vehicle queue, traffic flow, lane changing, and traffic control. Therefore, six data categories were extracted as queue length, arrival flow, discharge flow, inbound lane change, outbound lane change, and signal type. The data extraction was initially conducted for time sequences of 2 s and 5 s as shown below.

- Maximum queue length at the end of the time sequence
- Cumulative arrival and discharge flow for each time sequence
- Cumulative inbound and outbound lane changing for each time sequence
- Traffic control for each time sequence

Queue Length

The queue length was measured in metres, and the considered reference points were the start of the queue at the stop line to the back of the vehicle queue at the end of each time sequence. The starting point of the queue was considered as the stop line at the start, regardless of the vehicle's stopping position. This was considered to address the heterogeneity, as small vehicles tend to stop beyond the marked stop line. The vehicle queue was assessed as the maximum queue length of the vehicles at the time sequence, whilst the queues were further characterized according to the classification of Macioszek & Iwanowicz [19]

- Stopped—This was considered when all the vehicles in the queue were completely stopped due to being conditioned by a red signal. The stopped queue is illustrated in Fig. 3.
- Partially moving queue—This was considered when vehicles are moving very slowly or when there are vehicles whose drivers alternately start and stop manoeuvres, due to the oversaturation of the queue. The partially moving queue is illustrated in Fig. 4. In the vehicle queue, the first three vehicles have left the queue and the fourth vehicle (V_4) has started moving as the traffic control is released/ front vehicles are moving. However, the following vehicles after the fourth vehicle is still in the queue, as the maximum back of the queue is still formed. Therefore, in this scenario, the queue was considered from the stop line to the downstream until the back of the maximum queue formed. If the maximum back of the queue is moved forward, the queue was considered accordingly.
- Moving Queue—This was considered when the vehicles whose drivers have starting-up, manoeuvred to leave the intersection at the green signal duration, but are still influential at the back of the queue. The moving queue is illustrated in Fig. 5.
- The queue length phases for the queue transformation are characterized by three phases: queue formation, queue stagnation, and queue dissipation as defined in Table 2.

Arrival Flow and Discharge Flow

The arrival flow was determined at the upstream point of the road section, whilst the discharge flow was measured at the stop line. One of the core factors governing heterogeneous traffic conditions is the diversified vehicle composition. The

Table 2 Queue length phases of the queue transformation

Queue length phase	Definition
Queue formation	The queue length is gradually increased with respect to the increasing vehicle arrival flow at the red signal phase
Queue stagnation	The queue length is stagnating at a constant rate, due to the reduction/discontinuation of the traffic flow at the red signal phase
Queue dissipation	The queue length is gradually decreased with respect to the increasing vehicle discharge flow at the green signal phase

Table 3 PCU Factors for considered vehicle categories

Vehicle category	PCU factor
Motor bike	0.5
Three-wheeler	0.67
Car/Van/Jeep	1
Light goods vehicle (LGV)	1.25
Medium goods vehicle (MGV)	1.75
Bus/Heavy good vehicle (HGV)	2.25

Table 4 Lane-changing movements

Lane-changing type	Lane 1	Lane 2	Lane 3
Inbound lane change	2 to 1	1 to 2 3 to 2	2 to 3
Outbound lane change	1 to 2	2 to 1 2 to 3	3 to 2

use of arrival and discharge flows was designed to address this predominantly. The mixed traffic composition was dealt with considering the PCU of the vehicles [20]. Thereby, the considered PCU values for each vehicle category are shown in Table 3.

The arrival flow and the discharge flow were measured for each time sequence accordingly. Therefore, the cumulative count of the PCU values within the considered time sequence was used.

Lane Changing

The lane changing was determined for each of the lanes by considering the complete lanes as the areal sections. The area of the road section was determined by the length and width of the considered lane. Where inbound lane changing was considered if the vehicles entered the lane section, within the given time sequence. Whilst the outbound lane change was considered if the vehicles exited the lane section. The considered lane-changing movements are summarized in Table 4.

Traffic Control

The traffic control in signalized intersections is directly attributed to the signal lights. The signal controls constituting the vehicle queues are either by red signals for queue formations, or green signals for queue dissipations. Therefore, the signal control variable was extracted as a categorical variable as shown below.

- Green signal control = 0
- Red signal control = 1

Thus, the sample dataset of the first 10 time sequences of the training dataset, accounting for 50 s is shown in Table 5.

Development of the time series model

Model development was done based on the time series analysis for time sequence 2 s and 5 s using the EVIEWS Student Version Lite 11 software. The analysis was conducted in six main phases: checking for stationarity, likelihood-ratio test, optimal lag length, checking for granger causality, checking for cointegration and model fitting & validation.

Stationarity is a common assumption in many time series techniques. The stationarity means there is no growth or decline in data. That is data fluctuate around a constant mean, independent of time and the variance of the fluctuations remains constant over time. On the contrary, non-stationary data have means, variances and covariances that change over time. Behaviour of non-stationary can be trends, cycles, random walks, or a combination of these three. Non-stationary data cannot be used to model or forecast. The stationarity was visually inspected to identify potential trends and any other seasonal patterns with the aid of the Auto Correlation Function (ACF) graph and the Partial Auto Correlation Function (PACF) graph of the correlogram. These tests were performed to identify the correlation between the data points at time t with datapoints at time t = 1.

Augmented Dickey-Fuller (ADF) is the augmented version of the popular unit root test called the Dickey-Fuller test. The unit root test was conducted with respect to the ADF test. The test was performed in two phases where the first phase was for intercept only, and the second phase was for intercept and trend. The ADF test statistics were

Table 5 Sample dataset

Time sequence (s)	Arrival flow (PCU/5 s)	Discharge flow (PCU/5 s)	Inbound lane change (PCU/5 s)	Outbound lane change (PCU/5 s)	Actual Queue (m)	Traffic control
5	0	0	0	0	86.1	1
10	0	0	0	0	86.1	1
15	0	0	0	0	86.1	1
20	1.25	0	0	0	86.1	1
25	0	0	0	0	86.1	1
30	0	0	0	0	86.1	1
35	0	0	0	0	86.1	1
40	0	0	0	0	86.1	1
45	0	0	0	0	86.1	1
50	0	0	0	0	86.1	1

checked by adhering to the t-statistics and the probability. The ADF t-statistics were checked to have a lower value than the t-statistics than the 5% level t-statistics. The probability was checked to be lower than 0.05. The undue variables were differenced as the first difference variables and the test was reperformed.

The Likelihood-ratio (LR) test was performed for exogenous variables, with dummy variables as input variables. Such that, the LR value of the signal variable was checked to be greater than the chi-squared value, to be acceptable in the model.

Granger causality test was performed as two regressions with hypothesis considerations for each x variable and the y variable as shown below.

First regression.

$H_0 = X$ does not Granger-cause Y Vs. $H_1 = X$ does Granger-cause Y.

Second regression.

$H_0 = Y$ does not Granger-cause X Vs. $H_1 = Y$ does Granger-cause X.

The cointegration test was performed to determine whether a group of non-stationary series are cointegrated or not. The unrestricted cointegration rank test was obtained as a trace test and maximum eigenvalue test, adhering to the linear deterministic trend assumption. The considered hypothesis of the Johannes-Cointegration test is shown below.

$H_0 =$ Cointegrating equation does not exist Vs. $H_1 =$ Cointegrating equation exists.

The variables were fitted as VAR models as unrestricted VAR as the variables were known to have no cointegrations. The model fitting was done for one location.

The model fitting was conducted for 80% of the overall dataset whilst the model was validated with 20% of the overall dataset to further confirm the reliability of the estimated queue [21]. Moreover, the analysis was conducted for the two considered time sequences (2 s and 5 s), until a highly accurate model was obtained.

Results and Discussion

Time Sequence Optimization

The time sequencing for the VAR model relies on the vehicle traffic flow, and for higher traffic flows, the time sequencing should be lower or vice versa. Table 6 highlights the time sequence selection based on the traffic flow and the accuracy of the final model by considering the R squared of the fitted model. The two different time sequences of 2 s and 5 s were selected as a trial and error, to compare the accuracy of the models. The time sequences of 2 s and 5 s were selected based on factors such as red signal timing and the length between the upstream entry point and the downstream exit point. Thus, it was noted that the proposed time sequence in the model should be lower than that of the red signal timing yet should be optimum to avoid the multiple capturing of the residual queue at the intersection. Additionally, the consideration of the length of the study area is imperative, to evade under/over capturing of the traffic flow. A total of six separate models were developed initially as shown in

Table 6 Time sequence selection based on R squared

	Lane 1	Lane 2	Lane 3
R squared	2 s–0.29	2 s–0.17	2 s–0.77
	5 s–0.49	5 s–0.97	5 s–0.88

Table 6, hence the model with the highest *R* squared was chosen.

Thus, the time sequence of 2 s has the lowest model fitting with low R squared values of 0.29, 0.17, and 0.77 for lane 1, lane 2, and lane 3, respectively. Yet, the model with the time sequence of 5 s has the highest model fitting with the R squared of 0.49, 0.97, and 0.88 for lane 1, lane 2, and lane 3, respectively. Therefore, the lane 2 model with a 5 s time sequence was selected as the best fitting model, due to its high R squared in the developed model, and the high accuracy when applied to other locations for queue predictions.

The analysis results of parameter significance obtained for the VAR model estimates are shown in Table 7. The significant variables of the VAR model are queue length, arrival flow discharge flow, inbound lane change, and outbound lane change, whilst the lag is 3. The R squared value of the VAR model is substantially high, as it is 0.974 which is higher than 0.7 [21]. The high R squared of 0.974 implies the high model accuracy, as the developed VAR model predicts the actual data with an accuracy of 97.4%. The obtained Durbin–Watson value for the VAR model is 1.953, close to 2.0 implying that there is no serial correlation in the developed model.

The variables of queue length, arrival flow and discharge flow, inbound lane changing, and outbound lane changing are stationary at level. However, the signal variable was unacceptable as an exogenous variable in the model due to the failure in the LR test, as the LR value was smaller than the Chi-squared value of the model. The obtained VAR model for the queue length (at time t) is shown in Eq. 1. As per the coefficients of the model, it is evident that queue length arrival flow and inbound lane change at the nearest time sequences contribute positively to the queue. However, the furthest time sequence of queue length is inversely proportional to the queue length at present time. This is due to the residual queue dissipation in the furthest time sequences. On the other hand, discharge, and outbound lane change at nearest time sequences contribute negatively to the queue length. The impact from the furthest time sequences is comparatively lower overall.

$$\begin{aligned} QueueLength =& 1.049 + 1.3629 Queue_{t-1} - 0.295 Queue_{t-2} - 0.118 Queue_{t-3} \\ & + 0.377 Arrival_{t-1} + 0.136 Arrival_{t-2} + 1.714 Arrival_{t-3} \\ & - 0.985 Discharge_{t-1} - 0.985 Discharge_{t-2} + 0.242 Discharge_{t-3} \\ & + 0.266 Inboundlanechange_{t-1} - 0.691 Inboundlanechange_{t-2} \\ & + 2.999 Inboundlanechange_{t-3} - 0.873 Outboundlanechange_{t-1} \\ & - 0.411 Outboundlanechange_{t-2} + 0.266 Outboundlanechange_{t-3} \end{aligned} \quad (1)$$

where, (t − 1): time lag by one time sequence (− 5 s), (t − 2): time lag by two time sequences (− 10 s), (t − 3): time lag by three time sequences (− 15 s).

Table 7 VAR Estimate Significance

VAR estimates	Significance
$Queue_{t-1}$	0.000
$Queue_{t-2}$	0.033
$Queue_{t-3}$	0.015
$Arrival_{t-1}$	0.007
$Arrival_{t-2}$	0.018
$Arrival_{t-3}$	0.022
$Discharge_{t-1}$	0.002
$Discharge_{t-2}$	0.012
$Discharge_{t-3}$	0.051
$Inboundlanechange_{t-1}$	0.008
$Inboundlanechange_{t-2}$	0.027
$Inboundlanechange_{t-3}$	0.043
$Outboundlanechange_{t-1}$	0.005
$Outboundlanechange_{t-2}$	0.017
$Outboundlanechange_{t-3}$	0.029

VAR Model Training Results

The graphically illustrated VAR estimation for the first 800 s in Fig. 6, shows high accuracy in the predicted values. However, a lower model fit was identified in zero queues, causing a minor variation between the predicted and the actual queue. The higher queue length is incurred with the high vehicle density at the end of the red signal phase. Contrastingly, the zero queue is incurred with the low vehicle density in the green signal phase. However, the estimated values from the fitted VAR model at zero queues, have a higher percentage error due to the unpredictable nature of the model inputs. The prediction error can be contemplated as minimal, considering the minor variation of the absolute values less than 15 m [14]. The MAPE of the estimation was obtained as 21.55%. However, the error is minimized to a greater extent with the incorporation of the signal type as a categorical variable for traffic control.

VAR Model Validation Results

The trained VAR estimation model was applied to lanes 1 and 3 of location 1, as a step of model validation. The VAR prediction results of lane 1 and lane 3 with the actual field data are presented in Fig. 7 and Fig. 8, respectively. The model fit in lane 1 was high in the queue formation and dissipation phases, whilst it was low on two occasions: immediate transformation from queue formation to queue dissipation and when the queue length is zero. The MAPE for the two scenarios of lane 1 and lane 3 were 15.09% and 25.99%. The queue prediction behaviour of lane 1 is different from

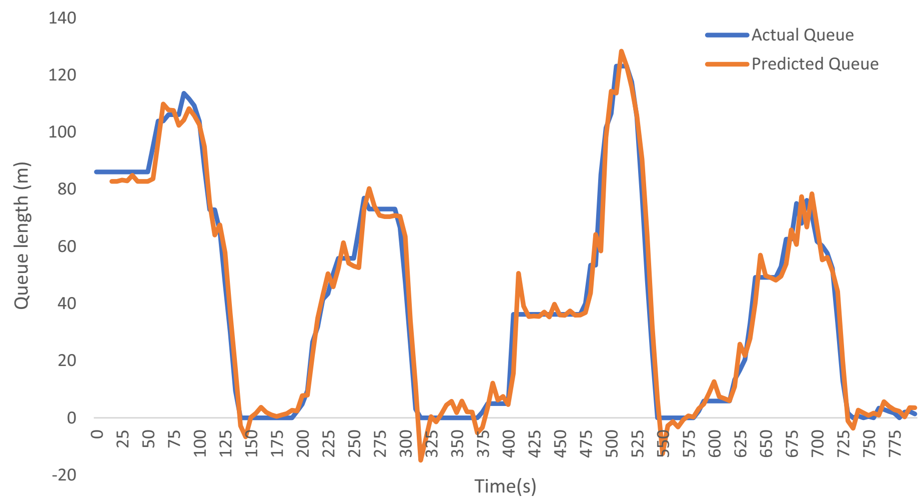


Fig. 6 VAR Estimation and actual field queue data of location 1—Lane 2

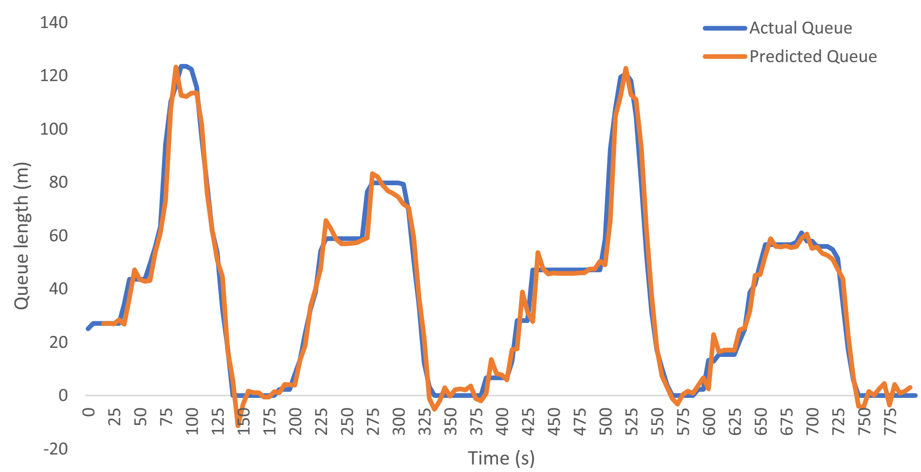


Fig. 7 VAR Prediction and actual field queue data of location 1—Lane 1

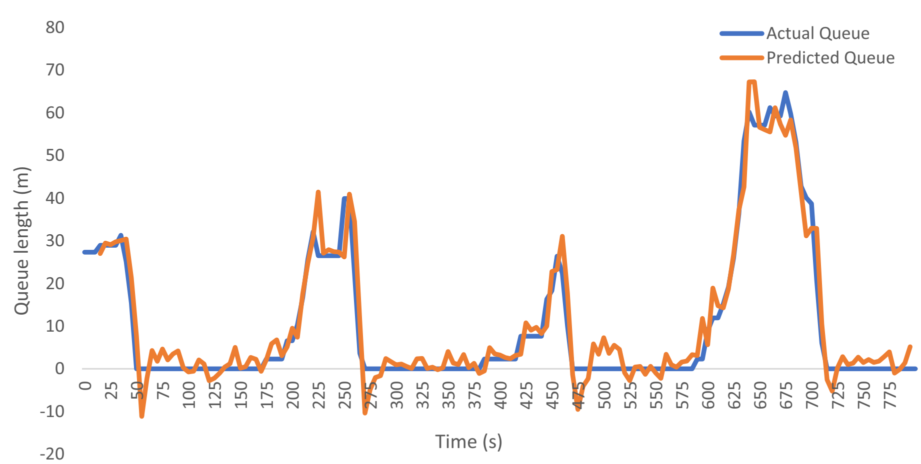


Fig. 8 VAR Prediction and actual field queue data of location 1—Lane 3

lanes 2 and 3, due to the changes in the traffic characteristics. Such that, lane 1 has turning movements of through and left turning with a comparatively reduced red signal timing of 72 s. Therefore, the green signal timing is high within the cycle time, and it alternately results in more zero queue lengths, which reduces the model fit of the queue prediction.

Fig. 9 VAR Prediction and actual field queue data of location 2—Lane 1

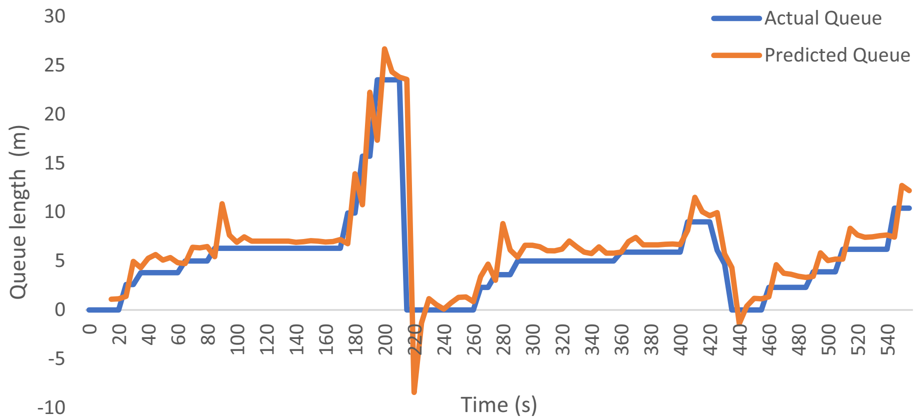


Fig. 10 VAR Prediction and actual field queue data of location 2—Lane 2

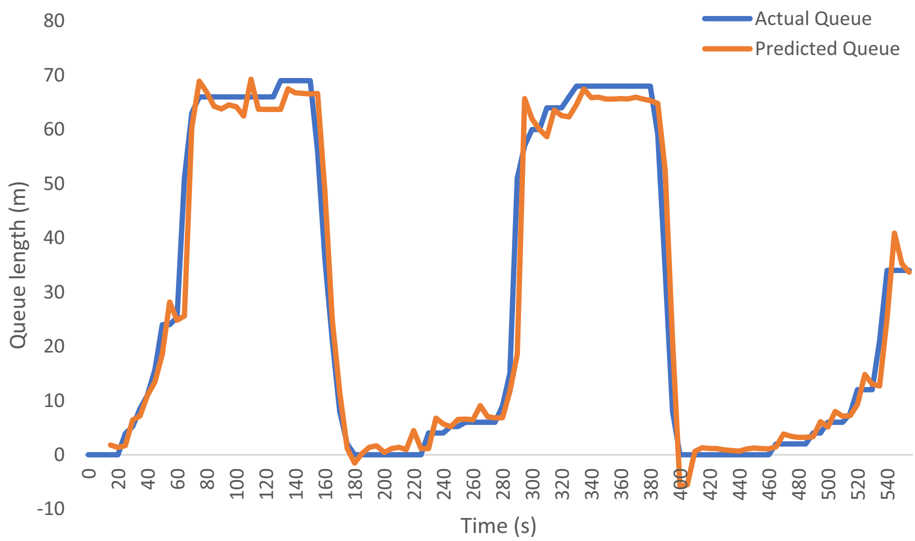
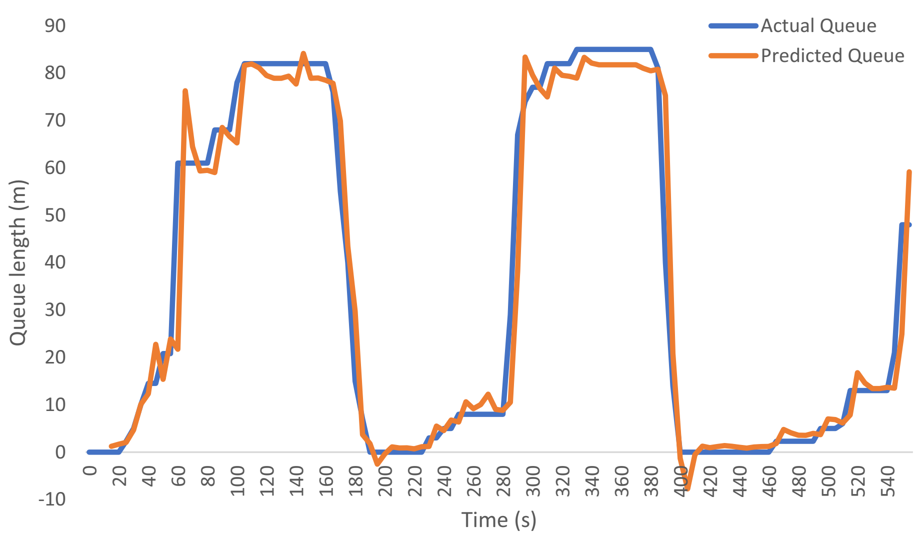


Fig. 11 VAR Prediction and actual field queue data of location 2—Lane 3



Real-Time Queue Prediction

The model was applied to the right turning lane (Lane 1), through lane near the centre median (Lane 2), and through lane near the curb (lane 3) as shown in Fig. 9, Fig. 10, and Fig. 11, respectively. The results of the location 2 model application demonstrated a similar model fit as location 1, where the right turning lane has a comparatively lower model fit in the queue stagnation phase. However, this error is low, considering the absolute difference between the observed and the predicted values. The MAPE for lane 1, lane 2, and lane 3 were 23.77%, 15.87%, and 16.99%. The MAPE variation of location 2 was similar to location 1 as the MAPE of the through movement lane was high, due to the zero queue length that occurred with the increased green signal timing. The results further showed that the model performs best in lanes with a high queue stagnation phase.

The prediction results show that a negative queue was obtained in two conditions: instantaneous queue drops to zero from a high queue length after the queue dissipation phase and continuous zero queue length in the queue stagnation phase. As pointed out by Sanchez et al. [22], negative queue estimates are erroneous and reduce accuracy. However, considering the absolute value of the error, the variation is negligible.

Conclusions

The study was conducted to estimate the queue length with adhering to the time series analysis of developing an unrestricted VAR model for 5 s time sequence data. Furthermore, the developed VAR model was utilized to predict the queue length of a four-lane road section with similar traffic conditions. The study assessed the relationships of influential factors for queue length. Moreover, the study considered the heterogeneous traffic conditions, specifically common among transportation networks in urban areas of developing countries. Therefore, the impact of heterogeneous conditions was accounted for with the quantification of mixed vehicle composition in the traffic flow and the lane-changing assessment. The results reveal that arrival flow, discharge flow, inbound lane changing, and outbound lane changing are significant towards the queue length. The dependency of the variables on the queue length of the present time (t) depends on the lag time of the previous three time sequences (t – 3). The results further demonstrate the variation in the accuracy of the predicted models by the traffic characteristics of turning movements. Therefore, MAPE of queue estimations in through movements lanes were ranging between 15 and 17%, whereas MAPE of right turning lanes were between 23 and 26%. The research methodology was limited in certain aspects. Thus, it can be further developed, adhering to the following future directions.

1. Time sequence: This paper developed an estimation method for maximum queue length and predicted the short-term queue length for 5 s. Thus, predicting the queue length for higher time sequences is helpful to provide effective traffic management solutions.
2. Consideration of Vehicle Delay: This paper's main aim was to develop a methodology for queue estimation and prediction, evaluating heterogeneous traffic. However, the consideration of vehicle delay is a promising research direction.

Funding This research was supported by the Accelerating Higher Education Expansion and Development (AHEAD) Operation of the Ministry of Higher Education, Sri Lanka, funded by the World Bank Group, Non-State DOR. 003, Niranga Amarasingha.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

1. S. Jayatilleke, V. Wickramasinghe, H. Madushani, S. Dissanayake, Estimating the Delay to the Mainstream Traffic due to Jaywalking Pedestrians on Urban Roads, in Proceedings of International Conference on Transportation and Developmen*t* (ASCE, New York, 2021), pp. 73–83
2. A. Verma, G. Nagaraja, C. Anusha, S. Mayakuntla, Traffic signal timing optimization for heterogeneous traffic conditions using modified Webster's Delay model. Trans. Dev. Econ. **4**, 13 (2018)
3. V. Wickramasinghe, W. Nanayakkara, S. Jayatilleke, Loss Time Estimation at Signalized Intersections Due to Heterogeneous Traffic Conditions, in Proceedings of 8th International Conference on Transportation Systems Engineering and Management, Calicut, India, pp. 26–27 August 2021
4. H. Seiran, *Real-time Queue Length Estimation on Freeway Off-ramps Using Case Based Reasoning Combined with Kalman Filter* (University of Calgary, Calgary, 2015)
5. S. Anusha, L. Vanajakshi, A. Sharma, A simple method for estimation of queue length. Int. J. Innov. Res. Sci. Eng. Tech. **2**(1), 1–5 (2013)
6. G. Newell, Approximation methods for queues with application to the fixed-cycle traffic light. SIAM Rev. **7**(2), 223–240 (1965)
7. G. Stephanopoulos, P. Michalapoulos, Real-time control of critical intersections based on queue dynamics, In Proceedings of 17th Symposium on Adaptive Processes (IEEE, 1978)
8. A. Sharma, D. Bullock, J. Bonneson, Input-output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersections. J. Trans. Res. Rec. **2035**, 69–80 (2007)
9. G. Comert, Queue length estimation from probe vehicles at isolated intersections: estimators for primary parameters. Eur. J. Operat. Res. **252**(2), 502–521 (2016)

10. P. Hao, X. Ban, D. Guo, Q. Ji, Cycle-by-cycle intersection queue length distribution estimation using sample travel times. Trans. Res. Part B Method. **68**, 185–204 (2014)
11. X. Zhan, R. Li, S. Ukkusuri, Lane-based real-time queue length estimation using license plate recognition data. Trans. Res. Part C Emerg. Tech. **57**, 85–102 (2015)
12. X. Zeng, J. Zhan, L. Yang, Q. Xiong, Research on the queue length prediction model with consideration for stochastic fluid, In International Symposium for Intelligent Transportation and Smart City, (Springer, 2017), pp. 51–63
13. B. Li, W. Cheng, L. Li, Real-time prediction of lane-based queue lengths for signalized intersections. J. Adv. Trans. **2018**, 1–18 (2018)
14. S. Jayatilleke, V. Wickramasinghe, N. Amarasingha, K. Liyanage, M. Lakmali, Application of time-series analysis to predict vehicle queue length at signalized intersections with heterogeneous traffic conditions. Adv. Trans. Stud. Int. J. **57**, 17–30 (2022)
15. G. Wickremesinghe, *Macroeconomic Forces and Stock Prices: some Empirical Evidence from an Emerging Stock Market* (University of Wollongong, Wollongong, 2006)
16. F. Faiso, S. Fazel, The stock market and unemployment relationship in USA, China and Japan. Int. J. Econ. Fin. **5**, 24–26 (2013)
17. L. Menike, The effect of macroeconomic variables on stock prices in Sri Lankan stock market. Sab. Uni. J. **6**(1), 50–67 (2015)
18. H. Dharmarathne, D. Samayawardena, C. Tilakaratne, Volatility models for world stock indices and behaviour of All Share Price Index, In Proceedings of 8th International Research Conference, KDU (2015), pp. 175–182
19. E. Macioszek, D. Iwanowicz, A back of queue model of a signal-controlled intersection approach developed based on analysis of vehicle driver behaviour. Energies **14**(4), 1204 (2021)
20. RDA, Design Standards of Highway, Sri Lanka, Japan International Cooperation Agency, Colombo (2008)
21. D. Moore, N. William, M. Flinger (eds.), *The Best Practice of Statistics*, 6th edn. (W.H. Freeman and Co., New York, 2013)
22. R. Sanchez, R. Horowitz, P. Varaiya, Analysis of queue estimation methods using wireless magnetic sensors. Trans. Res. Rec. **2229**, 34–45 (2011)