

## ESTIMATING DELAY AT A SIGNALIZED INTERSECTION USING QUEUING MODELS

Z. Mpanza<sup>1\*</sup> & H. Ncube<sup>2</sup>

<sup>1</sup>Department of Mechanical and Industrial Engineering  
University of Johannesburg, South Africa  
[zanelem@uj.ac.za](mailto:zanelem@uj.ac.za)

<sup>2</sup>Department of Quality and Operations Management  
University of Johannesburg, South Africa  
[henryn@uj.ac.za](mailto:henryn@uj.ac.za)

### ABSTRACT

Intersections have been identified as the most complex location in a traffic system. This is due to the number of movements that occur within an intersection. Delays have a negative impact on motorists as well as in the economy. Traffic signals are implemented to reduce this burden, but this is determined by the signal timing. The pre-timed traffic signal control assigns the right of way at an intersection according to predetermined schedule; and does not accommodate short-term fluctuations. The purpose of this study was to estimate the delay at an intersection using queuing models. The study uses field data collection.

---

<sup>1</sup> The author was employed as a Lecturer in the Department of Mechanical and Industrial Engineering, University of Johannesburg

<sup>2</sup> The author was employed as a Lecturer in the Department of Quality and Operations Management, University Johannesburg

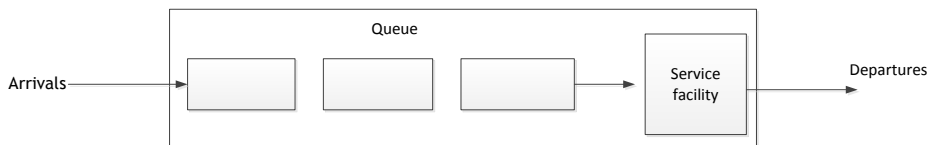
\*Corresponding author

## 1. INTRODUCTION

In urban areas, property developments are constantly increasing and this has caused an increase in the number of road users in the urban areas. With an increase in motorists, traffic congestions become inevitable. Traffic congestion contributes to economic, environmental and social issues in urban areas. Due to limited space, it becomes impossible to expand road networks and so traffic engineers have to come up with other mechanisms of making the road network effective. One of these mechanisms is to install traffic control systems.

Signalized intersections are basically points or nodes within a system of surface highways and streets; thus defining appropriate measures of effectiveness to describe the quality of operations within this system is somewhat difficult. Intersections are regarded as the most complex areas of traffic network [1]. This is due to a number of activities that take place within an intersection. A number of measures have been used in capacity analysis and simulation models, all of which quantify some aspect of the experience of traversing a signalized intersection in terms of what the driver comprehends. The most common factors include delay time, queue length and stopping time. Traffic control mechanisms like stop, yield and traffic lights are implemented at intersections to control traffic movement. The effectiveness of a traffic signal determines the overall time spent in a road network. Signal timing, signal synchronization, the arrival and departure rate of vehicles are just some of the influencing factors of the signalized intersection performance. Signal timing optimization is important to reduce traffic delays, relieve congestion, and improve the operation of an intersection. The performance of an intersection is measured by the delay which motorists experience within an intersection [1].

Vehicles approaching a traffic signal have the same characteristics of a queueing system. They have three parts, which are (1) the arrivals or inputs to the system, (2) the queue or waiting line itself, and (3) the service facility [2]. Vehicles arrive at a traffic signal, wait in the queue, receive a service (signal) and depart, as shown in Figure 1. The effectiveness of a traffic signal (service facility) determines the overall time spent in a road network.



**Figure 1: General Queueing system.**

At a signalized intersection, the level of service (LOS) or quality of traffic flow can be measured using various parameters [1]. Amongst them is vehicle delay; which is the most important parameter since it indicates the time loss by a vehicle while intersecting an intersection. Travel times are not easy to determine at an intersection and traffic signal control due to stochastic properties of traffic flow and stochastic arrivals and departures. This causes determining delay to be one of the complex task because it is influenced by many variables [1]. Also, delay is one of the activities that are regarded as non-value adding activities.

Traffic delay is no different from any other delay experienced on a daily basis by customers, the delay can be in a processing plant or a service in a bank; they all have a negative impact on customers/users and they also indicate a poor performance of the system as a whole. Not only does it impact negatively on customers or users, it also impacts negatively on the performance of a business as customers end up choosing other alternatives to avoid delay. In traffic, users may choose to use other routes which are not designed for heavy traffic flow, which results in even more congestions. It is therefore important to identify the causes of delay and reduce it as much as possible.

Signal timing optimization is important to reduce traffic delays, relieve congestion, and improve the operation of an intersection. Signal timing is crucial during peak periods because it determines the number of vehicles that can cross a traffic light during green time. If there are fewer vehicles that can intersect, the vehicle queue length increases causing congestion and overflow. In queueing theory, this indicates that the service rate is less than the average arrival rate [2]. This problem is mostly caused by the pretimed traffic signals [1]. Traffic signals are said to be pretimed if they have a fixed signal timing, regardless of the queue length [1]. Pre-timed traffic signals are predetermined even during the peak period which does not solve the problem of queue length. One solution to this is to implement actuated traffic signals. In South Africa, most traffic signals are pretimed which results in traffic overflow on intersections. This study seeks to estimate delay at an intersection using a queueing model.

## 2. LITERATURE REVIEW

In this section, we look at some of the studies and models developed to estimate delay.

There are two basic categories of cost in a queueing situation: those associated with customers waiting for service; in this case motorists waiting to intersect at a traffic signal, and those associated with capacity [2]. The goal of queueing analysis is to balance the cost of providing service capacity with the cost of customers (motorists) waiting for service [2]. The goal of analysis is to identify the level of service (LOS) capacity that will minimize total cost. Much of the success of the analysis will depend on choosing an appropriate model. Model choice is affected by the characteristics of the system under investigation, the main characteristics are arrival and service patterns, queue discipline, number of channels and population source [2]. Waiting lines are a direct result of arrival and service variability. They occur because random, highly variable arrival and service patterns cause systems to be overloaded. Waiting lines are most likely to occur when arrivals are bunched or when service times are particularly lengthy. Usually vehicle arrivals at a traffic signal are unlimited. Vehicles arrive randomly since they cannot be predicted exactly, even though during the peak hours a prediction can be made.

Congestion is unavoidable in urban cities due to growth in traffic demand. This has led to the development of many traffic signal control systems which have been tested in a number of countries. In the 1960's, the early stage of systems was implemented in Munich, West London and Toronto and later the urban traffic control system project was completed in the United States [1]. The SCAT system, which has a multilevel hierarchical structure, was developed by Australia [1].

A number of studies have been done to estimate delay at a signalised intersection. Wu and Giuliani [3] estimate delay and capacity at a signalised intersection by measuring the cycleflow probability. The authors use the measured cycleflow probability to estimate delays and queueing lengths. Their results show that the cycleflow probability can estimate delays. Vajeeran and De Silva [4] conducted a delay analysis at a signalized T intersection at Katubedda. The authors used Vissim simulation software to model actual data of queue length. They found that signal timing changes and geometric changes can reduce delay. Christofa *et al.* [5] introduced a real-time signal control system that optimizes signal settings based on minimization of person delay on arterials. Tan *et al.* [6] uses a decentralized genetic algorithm (DGA) to minimize delay at signalized traffic network under an oversaturated condition. The results show that the developed DGA is able to reduce network delay by optimizing the traffic signal. Dabiri and Abbas [7] use a Particle Swarm Optimization (PSO) algorithm which is used as an optimizer for generating arterial traffic signal timing parameters. The results reveal that the proposed method is promising and outperforms for various traffic for traffic states. Lu and Yang [8] use a stochastic queue model that considers interdependence relations between adjacent intersections using probability-generating function to analyse delay in signalised networks under different congestion levels. Huang *et al.* [9] calculate intersection delay based on vehicle positioning data and analysed it with DBSCAN and Least Square Fit algorithms, without using the signal or traffic information.

Anusha *et al.* [10] use occupancy based method and the queue clearance based method to estimate queue and delay.

Although there has been a tremendous work done in this area, there are very few studies that are published in the South African context. According to White Paper on Roads Policy for South Africa [11], the roads industry in South Africa does not have a formalised system in place to guide the industry at the project implementation level which causes a challenge in implementation of best-practices for road authorities, road designers and builders. South African traffic engineers conduct these studies but they are seldom published and therefore this limits the availability of data.

### 3. METHODOLOGY

#### 3.1 Data collection site

Since the data was not readily available, data had to be collected manually for this study. The study was conducted in an intersection situated in Victory Park, which is a suburb in the North of Johannesburg. This site was selected because it is one of the busiest intersections as it connects to an arterial, but yet less attention has been given to improve it. Also, due to the new housing developments in the area, the traffic congestion has increased dramatically as experienced by motorists on a daily basis. The intersection experiences various traffic conditions from saturated to undersaturated with overflow at times. The intersection is a four-leg signalized intersection as seen in Figure 2. Table 1 gives a summary of the information about the intersection. The study was conducted on a Tuesday during peak time of the morning from 7:00 am to 8:00 am. Six cycles were recorded. According to Schroeder *et al.* [12] six cycles is adequate to get conclusive information. The weather condition was clear with no rain or thundershowers. The intersection is a straight paved road. Figure 2 shows the intersection under study. Only the through lane traffic was selected for surveying and therefore the right turn lane was excluded. In Figure 3, the red arrow shows the lane excluded from the survey.

**Table 1: Intersection information**

Intersection	Cycle length (seconds)	Direction	Green time(seconds)	Approach
Corner Tana Road and 3 <sup>rd</sup> Avenue	70	through	26.06	Eastbound



Figure 2: Intersection selected for study.



Figure 3: Showing the excluded right turning lane group [13].

### 3.2 Field survey

The data required for this study included queue length, signal timing (green, yellow and red), headways, etc. Before the actual collection of field data, preparation was done days prior to the study which involved observing the traffic flow behaviour, identifying which peak hour (afternoon or morning) was more critical. The morning peak was identified as more critical than the afternoon. The data was manually recorded using a stop watch and the designed recording sheets. Traffic video cameras were the preferred survey method but due to financial constraints these could not be obtained. Nevertheless, the surveyors were conveniently located where they could observe all movements and at a close proximity with the ability to maneuver.

One of the surveyors observed and counted stopped vehicles on an intersection approach at every red signal. This was to get the size of the queue length. A total of eleven sets of queue counts was recorded and these were added to get a total for a 15 minutes interval. The vehicles departing on the green signal were also counted and added up to get a total for a 15 minutes interval. A total of twelve sets of departure counts were recorded. Table 2 shows the total queue length and total departure count for 7:00 am - 8:00 am.

The signal timing was observed and recorded. This intersection has a three-phase signal cycle system with pretimed or fixed cycle time. Table 3 shows the observed signal times, which show green as 26.06 seconds, yellow with 2.94 seconds, red with 41.03 seconds and a cycle length of 70 seconds.

**Table 2: Departure count and queue length**

Time Period	Total Departure Count (veh)	Queue length (veh)
7:00 - 7:15 am	141	168
7:15 - 7:30 am	107	155
7:30 - 7:45 am	145	172
7:45 - 8:00 am	136	161
<b>TOTAL</b>	<b>529</b>	<b>656</b>

**Table 3: Signal times**

Signal	Time (seconds)
Green	26.06
Yellow	2.94
Red	41.03
Cycle time	70

When the green signal was initiated, vehicles crossing the curb line were observed and the time between the initiation of the green and the crossing of the first vehicle over the curb line was recorded. These times were recorded for eight successive headways. According to Schroeder *et al.* [12], eight is an ideal number since the headway starts levelling off after the fourth or fifth vehicle. This levelling off is defined as a saturation headway which was recorded as 2 seconds, and we use this to compute saturation flow rate which is the number of vehicles per hour of green time per lane [12]. The start-up lost time and clearance lost time was recorded as 1 second.

The data collected assisted in the calculation of:

- The arrival rate
- Saturation flow rate

- Capacity of lane
- Effective green time
- Effective green time ratio
- Volume to capacity ratio
- Delay time

All the above parameters are calculated to aid in the estimation of delay. The formulas used to calculate these are listed in section 3.3.

### 3.3 Data Analysis

Webster [14] developed a model for estimating the delay experienced by motorists at signalised intersections using deterministic queuing analysis. The model is presented in Equation 3.1.

$$d = \frac{c(1-\lambda)^2}{2(1-\lambda X)} + \frac{X^2}{2v(1-X)} - 0.65 \left(\frac{c}{v^2}\right)^{1/3} [X^{2+5\lambda}] \quad (3.1)$$

where:

- d* = average overall delay per vehicle (seconds),
- λ* = proportion of the cycle that is effective green (*g*/*C*),
- C* = cycle length (seconds),
- v* = arrival rate (vehicles/hour),
- c* = capacity for lane group (vehicles /hour),
- X* = volume to capacity ratio of lane group,
- g* = effective green time (seconds).

The first term of equation 3.1 represents the average delay to the vehicles assuming uniform arrival. The additional delay, which is due to the randomness of vehicle arrival, is represented by the second term. To correct the delay estimates the third term of the equation was introduced as an adjustment factor.

We calculate delay using the Webster's queueing model [14]. First, we need to compute calculations for saturation flow rate, capacity of lane, effective green time and finally calculate the delay time. We use the following equations:

- 1) Saturation flow rate:

$$s = \frac{3600}{h} \quad (3.2)$$

where: *s* = saturation flow rate (vphgpl),  
*h* = saturation headway (seconds),  
 3600 = seconds/hour.

- 2) Effective green time:

$$g = G + Y - t \quad (3.3)$$

where:

- G* = actual green time
- Y* = sum of yellow plus all red time
- t<sub>L</sub>* = total lost time per phase

- 3) Capacity of each lane group:

$$c = s \frac{g}{C} \tag{3.4}$$

where:

$c$  = capacity of lanes serving movement, (vph or vphpl),  
 $s$  = saturation flow rate for movement (vphg or vphgpl),  
 $g$  = effective green time for movement (seconds),  
 $C$  = signal cycle length, (seconds).

4) Effective green time ratio:

$$\lambda = \frac{g}{C} \tag{3.5}$$

where:

$\lambda$  = proportion of the cycle that is effective green ( $g/C$ ),  
 $g$  = effective green time (seconds),  
 $C$  = cycle length (seconds).

5) Volume to capacity ratio:

$$X = \frac{v}{c} \tag{3.6}$$

where:

$X$  = volume to capacity ratio,  
 $v$  = arrival rate (vehicles/hour),  
 $c$  = capacity for lane (vehicles/hour).

#### 4. RESULTS AND DISCUSSION

When estimating delay at an intersection, a number of factors come into play. The basic characteristics used to model any intersection operation are applied the same way when modeling delay, that is, the manner in which vehicles depart, or discharge from the intersection when a GREEN indication is received. The following calculations are necessary to build the model for calculating delay at an intersection.

Figure 4 illustrates a group of vehicles at a signalised intersection, waiting for the GREEN indication.

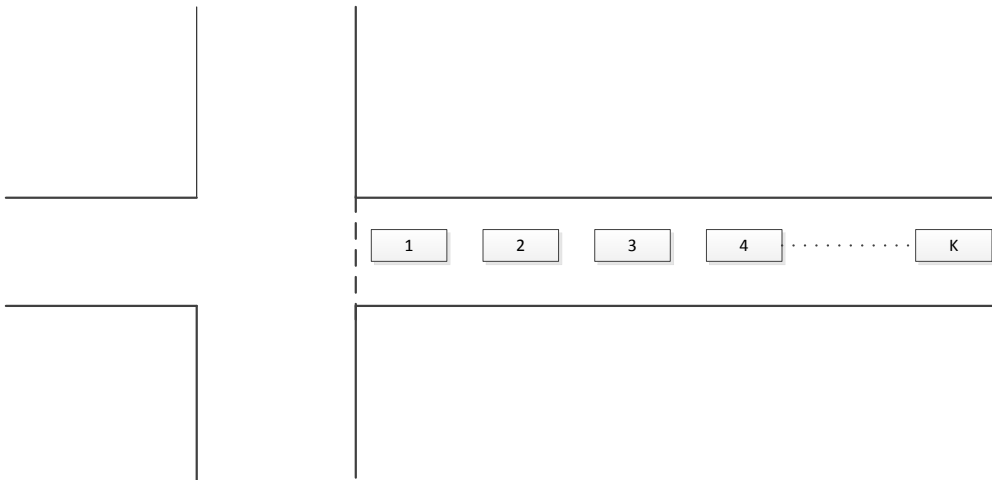
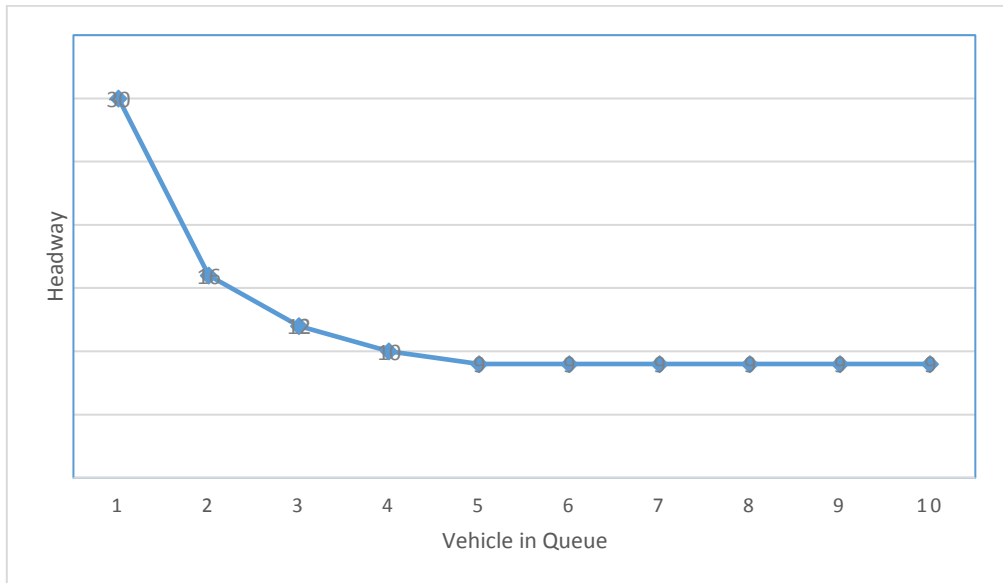


Figure 4: Vehicles waiting in queue for a green signal indication.



When GREEN is initiated, headways between departing vehicles is observed as vehicles cross the curb line. The first headway is the time between the initiation of GREEN and the crossing of the first vehicle over the curb line. The second headway is the time between the first and the second vehicles crossing the curb line, etc. As can be seen in the Figure 5, the first headway is longer as it includes the first driver's reaction time, and the time necessary to accelerate. The second headway is shorter because the second driver can overlap her reaction and acceleration time with the first driver's. Each successive headway gets a little smaller and the headways finally tend to level out. The level headway, or saturation headway can be seen from the fifth vehicle in the queue as shown in Figure 5.



**Figure 5: Headways departing signal.**

The leveling off is noticed with the fifth headway. From this illustration, we can now calculate the saturation flow rate using equation 3.2 which yields:

$$s = \frac{3600}{h} = \frac{3600}{2} = 1800 \text{ vphgpl}$$

This means that 1800 vehicles/hour/lane could enter the intersection if the signal were always green. If there is more than one lane, this number is multiplied by the number of lanes. In reality, the signal is not always green for any movement. Therefore, Miller [15], Akcelik [16] developed mechanism for dealing with the cyclic starting and stopping of a movement at a signal. This entails the calculation of the amount of green time available (effective green time,  $g$ ) to be used at a rate of one vehicle every 2 seconds (saturation headway). Calculating effective green time  $g$  using equation 3.3:

Using data listed in Table 1 of section 3.2 and equation 3.3, we get:

$$g = G + Y - t = 26.02 + 2.94 - (1 + 1) = 27 \text{ seconds}$$

Therefore, the amount of available green time to be used at a rate of one vehicle every 2 seconds is 27 seconds.

We then compute capacity of the lane serving movement, using equation 3.4:

$$c = s \frac{g}{c} = 1800 \frac{27}{70} = 694 \text{ vphpl}$$

Thus, 694 vehicles per hour can transverse in each lane, given that the available green time is 27 seconds.

The effective green ratio is computed using equation 3.5:

$$\lambda = \frac{g}{c} = \frac{27}{70} = 0.38 \text{ seconds}$$

The arrival volume is computed by adding the departure volume and the net change in the size of the queue during the counting period as shown in Table 4. For period 1, which is 7:00 - 7:15 am, the arrival volume is calculated by subtracting the total departure from the queue length recorded at the beginning of the study.

**Table 4: Arrival Volumes**

<i>Time Period</i>	<i>Queue Length (veh)</i>	<i>Total Departure Count (veh)</i>	<i>Arrival Volume</i>
7:00 - 7:15 am	168	141	168-141 = 27
7:15 - 7:30 am	155	107	155+107-141 = 121
7:30 - 7:45 am	172	145	172+145-107 = 210
7:45 - 8:00 am	161	136	161+136-145 = 152
<b>TOTAL</b>		529	510

As can be seen in Table 4, between 7:00 - 7:15 am, 168 vehicles were in the queue at the start of the study period. Out of the 168 vehicles, 141 departed on green signal indication. The arrival volume is then computed by subtracting the departure from the queue length. Between 7:30 - 7:45 am queue length was at maximum with 172 vehicles. This also shows a peak departure flow at 145 veh/15 min with a peak arrival volume of 210 veh/15 min. The results in Table 4 tells us that the traffic flow is not too heavy from 7:00 - 7:15 am. It then starts to increase from 7:15 - 7:30 am, reaching its peak from 7:30 - 7:45 am. The total arrival volume is 510 vehicles per hour. It can be deduced from the results that more vehicles traverse between 7:15 - 8:00 am. Therefore, signal timing can be adjusted to accommodate the traffic conditions during this period to ease congestion.

Using equation 3.6 we calculate volume to capacity ratio:

$$X = \frac{v}{c} = \frac{510}{694} = 0.73$$

Finally, calculate delay using equation 3.1:

$$\begin{aligned} d &= \frac{c(1-\lambda)^2}{2(1-\lambda X)} + \frac{X^2}{2v(1-X)} - 0.65 \left(\frac{c}{v^2}\right)^{1/3} [X^{2+5\lambda}] \\ &= \frac{70(1-0.38)^2}{2(1-0.38 * 0.73)} + \frac{(0.73)^2}{(2)(510)(1-0.73)} - 0.65 \left(\frac{694}{510^2}\right)^{1/3} [(0.73)^{2+5(0.38)}] \\ &= 18.6 \text{ sec/veh} \end{aligned}$$

Thus, each vehicle is expected to experience 18.6 seconds of delay traversing in this intersection.

## 5. CONCLUSION

In this study, we were able to calculate the saturation flow rate, which is one of the elements that are used to time signals and estimate intersection capacity. Most agencies use standard constant values for saturation flow, but they vary between intersections and may vary significantly in the South African traffic network. More studies need to be conducted at several sites in South Africa to establish saturation rates. The calculations show that the arrival volume reaches peak between 7:15 - 8:00 am.

Delay was calculated to be 18.6 sec/veh, which means that every vehicle joining the intersection delays with 18.6 seconds in the intersection before traversing. This figure is not particularly large but can be eliminated by adjusting signal timing during peak hours. This may entail the use of technology that is able to detect or count the number of cars in the intersection and allowing sufficient green time for vehicles to pass (traffic actuated signal). This can improve the level of service in this intersection thereby reducing travel time for motorists. Reduction in delay may indirectly improve productivity as people will spend less time on the road and more time in the workplace.

More intersection delay studies need to be conducted so that we can compare the results. Studies similar to this one need to be conducted periodically so that we have enough traffic data for South African road network. Other studies may include delay caused by traffic signal failure in an intersection and the impact of traffic signal synchronization. These type of studies may assist in traffic control as well as transport planning.

## REFERENCES

- [1] Pande, A. and Wolshon, B. (Ed). 2016. *Traffic Engineering Handbook*, 7<sup>th</sup> Edition, John Wiley & Sons Inc.
- [2] Render, B., Stair R., Hanna, M. and Hale, T. 2015. *Quantitative Analysis for Management*, 12<sup>th</sup> Edition, Pearson.
- [3] Wu, N and Giuliani, S. 2016. Capacity and Delay Estimatin at Signalized Intersections under unsatutaed flow condition based on cycle overflow probability, *International Symposium on Enhancing Highway Performance*, 15(), pp 63-74
- [4] Vajeeran, A. and De Silva, G.L.D.I. 2017. Delay analysis at a signaized T intersection. *2017 Moratuwa enginerng Conference (MERCon)*, pp 325-330.
- [5] Christofa, E., Ampountolas, K. and Skabardonis A. 2016. Arterial traffic signal optimization: A person-based approach, *Transportation Research Part C* 66(), pp 27-47.
- [6] Tan. 1964. *Statistical Inference*, 3<sup>rd</sup> Edition, McGraw-Hill.
- [7] Dabiri, S and Abbas, M. 2016. Arterial traffic signal optimization using Particle Swarm optimization in an integrated VISSIM-MATLAB simulation environment, *201 IEE 19<sup>TH</sup> International Conference on Intelligent Transportation Systems (ITSC)*, Windsor Oceanico Hotel, Rio-de Janeiro, Brazil, November 1-4, 2016.
- [8] Lu, Y. and Yang, X. 2014. Estimating dynamic queue distribution in a signalized network through a probability generating model, *IEEE Transactions on Intelligent Transportation Systems*, 15(1), pp 334-344.
- [9] Huang, J., Li, G., Wang, Q. and Yu, H. 2013. Real time delay estimation for signalized intersection using transit vehicle positioning data, *13<sup>th</sup> International Conference on ITS Telecommunications (ITST)*.
- [10] Anusha, S., Vanajakshi, S. and Sharma, A. 2015. Performance comparison of two model based schemes for estimation of queue and delay at signalized intersections, *IEEE Intelligenet Vehicles Symposium (VI)*, June 28 - July 1, 2015, COEX, Seoul, Korea.
- [11] Department of Transport. 2017. *White Paper on National Transport Policy*, Pretoria, South Africa.

- [12] Schroeder, B., Cunningham, C., Findley, D., Hummer, J. and Foyle, R. 2010. *Manual of Transportation Engineering Studies*, 2<sup>nd</sup> Edition, Washington DC: ITE.
- [13] Google Maps, available at: <https://www.google.com/maps/place>, [accessed 31 July 2018].
- [14] Webster, F. 1958. Traffic signal settings, *Road Research Technical Paper No. 39, Road Research Laboratory, London, UK*.
- [15] Miller, A. 1963. Setting for fixed cycle traffic signals, *Operations Research Quarterly*, 14(4), pp 373-386.
- [16] Akcelik, R. 1988. The Highway Capacity Manual delay formula for signalized intersection, *ITE Journal*, 58(3), pp 23-27.

### Revisions Made

Revision location	Revision made
Page 3, to this implement	Inserted "is to implement"
Page 3, a Vissim	Removed "a"
Page 3, is used	Removed "is used"
Page 3, introduce	Changed to "introduced"
Page 3, uses decentralized	Inserted "a" (a decentralized)
Page 3, under a oversaturated	Inserted "an" (an oversaturated)
Page 5, change Figure 3	Changed to a Google maps picture
Page 7, caacity	Changed to "capacity"
Page 9, Figure 5 heading	Removed heading
Page 10, effecetive	Changed to "effective"
Page 10, subtrcting	Changed to "subtracting"
Page 10, deducted	Changed to "deduced"