https://doi.org/10.51459/futajeet.2021.15.1.208

Journal of Engineering and Engineering Technology

ARRIVAL PATTERNS AND TRAFFIC FLOW CHARACTERISTICS AT SIGNALIZED INTERSECTIONS: CONSIDERING UPSTREAM SIGNAL INFLUENCE AND TIME RESOLUTIONS

[1*]Akinfala O. O., [2]Ege, E. E., and [3]Ogunwolu, L. F.

[1,2]*Department of Geography, Faculty of Social Science, University of Lagos, Nigeria.*
[3]*Department of Systems Engineering, Faculty of Engineering, University of Lagos, Nigeria*

ARTICLE HISTORY
Received: 21-12-19
Accepted: 18-08-20

KEYWORDS:
Traffic Signal, I-ratio, Arrival Patterns, Probability Distributions, Time Resolutions, Periodicities

ABSTRACT

Traffic arrivals at signal intersection approaches is inherently stochastic. This variability is typically reflected by I-ratio and there is a general consensus that the presence or absence of nearby upstream signal affects Variance to Mean Ratio (I-ratio). However, the effect of time resolution on arrival variability and the interaction effect between upstream signal and time resolution is yet to be examined in detail. This can lead to model misspecification and invariably, erroneous outcomes. This work examines the effect of time resolution and intersection type and their interaction on I-ratio and the resultant probability distributions. Traffic arrivals were measured at high time resolution- 10 seconds interval and then aggregated to lower time resolutions (30-150 seconds) at six intersections. Spectral density analysis showed statistically significant periodicity, specifically at 30 seconds interval with p-values < 0.0001 at all connected intersections while observations at isolated intersections lacked periodicity. Two-way ANOVA using I-ratio as the dependent variable and intersection type and time-resolution as the independent variables was performed. Statistically significant effect with F-value 8.606 at p-value < 0.0001 and R2 value 0.32 were observed. Intersection type, time resolution and the interaction between them were statistically significant, with p-values 0.002, < 0.0001 and 0.000 respectively. The combined effect of these factors led to a wide I-ratio range of 0.37-9.2. Negative Binomial, Poisson, and Binomial distributions represented 76.4, 20.4 and 4.2% of all I-ratios observed. Therefore, in contrast to literature which recommends Poisson, Negative Binomial may be a better suited probability distribution for traffic arrivals.

Correspondence:
fogunwolu@unilag.edu.ng

1.0 INTRODUCTION

The success of many decisions in traffic engineering, planning and policy making today are heavily dependent on reliable predictions. In many cases, due to either the cost of data collection, time limitation or simply unavailability of data, these predictions are based on limited amount of information. Consequently, the ability to characterize traffic observations into probability distributions accurately is crucial for traffic engineers to easily analyse the operational characteristics and performance of transport facilities with limited data. Characterizing vehicular arrivals into various statistical distributions has numerous applications in traffic engineering (Mauro and Branco, 2012; Ha *et al*, 2013). Traffic

https://doi.org/10.51459/futajeet.2021.15.1.208

arrival distributions are particularly useful in the following areas: designing turn bays, pedestrian systems, and traffic simulation (Gerlough and Barnes, 1971). Increasing congestion in many rapidly developing cities has also necessitated improvement in performance analysis for planning, design and operations. As a result of the highly stochastic nature of traffic, probability distributions offer a simple way to quickly and accurately model facility operations and performance. Furthermore, the probability distribution of arrival is a key parameter in the development of delay models at signalized intersections. The most widely used analytical delay models and formulas (Highway Capacity Manual (HCM), 1994-2010 and Ackelik, 1988) inherently assume a specific arrival process (step-wise or parabolic curve) at a signalized intersection (Cheng *et al*, 2015). Even though this is not specifically a probability distribution, it is a simplified way of representing an otherwise complex and uncertain arrival process. Notwithstanding, in early delay model development, (1940's - 1960's) the use of probability distributions was more direct, as they can even be categorized into various groups based on the type of arrival distribution upon which the model is predicated. For example, Beckman *et al*. (1956), Newell (1960) assumed Binomial arrival pattern, while Webster (1958), Miller (1963) and Newell (1965) assumed Poisson arrival pattern. The aforementioned models were heuristic and/or analytical in form. However, another group of researchers such as Olszewski (1990) and Heiderman (1994), who adopted Markov chains for a probabilistic approach towards queue formation and delay estimation also assumed a Poisson arrival process. Arrival processes are typically dependent of traffic conditions and facility characteristics. The Poisson arrival type with variance to mean ratio (I-ratio) of approximately 1 which is the most widely adopted in traffic signal operations analysis is typical of isolated intersections with random arrivals. Binomial arrivals with I-ratio < 1 are observed during congested periods and urban traffic where there is little room to maneuver, thus showing lower variance. Finally, Negative Binomial arrival with I-ratio > 1 is characterized by very high variance which is often as a result of cyclical disturbance from nearby upstream signals or when counting extends over both peak and off-peak periods (Gerlough and Huber, 1975).

The second quarter of the 20^{th} century has seen numerous works in the application of probability distributions to properly characterize traffic counts and arrivals. The works of Gerlough and Huber (1975) as well as May (1990) offer a comprehensive review of various distributions and their applications in traffic engineering. The earliest of the probability distribution models to be applied to traffic engineering is the Poisson model. Poisson is quite suitable for modelling of discrete random events, and it has been widely applied in traffic engineering because of its simplicity and wide applicability. Kinzer (1933) was probably the first to apply it to traffic. Poisson distribution was also adopted in the works of Adams (1936) and Greenshields et al (1947). This suggests they found I-ratios of 1 to be suitable for modelling traffic arrival in their works. On the other hand, Miller (1963), studied intersections in Birmingham and observed I-ratios ranging from 1 – 2, except in one case where I-ratio

https://doi.org/10.51459/futajeet.2021.15.1.208

was greater than 4. Newell, (1965) adopted a 60 sec counting interval and noted that variance to mean ratios of 1-1.5 are typical for 60 sec counting intervals. Although such long counting intervals have been known to obscure high variance (Gerlough and Huber, 1971), Newell's work was still able to show variability between mean and variance. Olszewski (1994) studied 5 intersections and found I-ratios to range between 0.77 and 1.74. Some studies have begun to identify that Poisson distribution may not necessarily be as good a fit for traffic arrivals at signalized intersections as literature presents. To this end, Ritchie (1983), studied several intersections in Melbourne, Australia during peak periods and I-ratios were between 1.1 and 5.4. The author compared the results of adopting Poisson and Negative Binomial models including two different methods of estimating parameter k namely (maximum likelihood and methods of moments) for Negative Binomial models. The author noted that at 10 sec counting interval, I-ratio increases nonlinearly to mean flow rate and Negative Binomial model was a better fit to observed data. Similarly, Baura *et al* (2015) adopted a simulation approach to investigate the suitability of Poisson or ARIMA time series model to predict arrival distribution and headway at three intersections in New Jersey, USA. The intersections were simulated in paramics software environment. Their results showed that ARIMA was a much better model for prediction of arrival pattern than Poisson.

Clearly, a wide variety of I-ratios have been found and probability distributions have been applied to traffic. However, (May, 1990) noted that no standard distribution can adequately describe traffic arrivals. Although a few studies have shown that Poisson may not necessarily be the most suitable distribution in all traffic scenarios, it is still widely used among researchers up till today. Viti and Van Zuylen (2009) adopted a Poisson distribution for arrivals as part of their model. However, they include some flexibility by allowing for modification in the event that another distribution is expected to be used. Harapap *et al* (2019) simulated the waiting time at an intersection using MATLAB. A Poisson process was assumed in their work. Farivar (2015) modelled the effect of a channelized right-turn on delay and capacity by developing a probabilistic model and comparing the results with VISSIM. A Poisson distribution was adopted in the work primarily because only isolated intersections were studied. Similarly, Xing *et al* (2016) developed optimal timing plans for coordinated signalized intersections by considering the effect of upstream arrivals at the intersection of interest. The downstream arrivals from upstream right turn movements or un-signalized intersections were modelled using a Poisson distribution. While the Poisson distribution may intuitively appear appropriate in both situations, Viti and Van Zuylen (2009) noted that this assumption may not necessarily be correct. This is because even if the traffic generation process is Poisson, if the distance between intersections are sufficiently long, vehicles generated by a Poisson process may eventually arrive downstream in platoons. Therefore, violating the Poisson assumption. Literature shows that studies which adopt the Poisson assumption abound even though it may be ill-suited.

A few studies have sought to investigate the impact of arrival distribution on the

https://doi.org/10.51459/futajeet.2021.15.1.208

operational performance of signals. To this end, Miller (1963) noted that delay is insensitive to the form of distribution of arriving traffic. This is perhaps why the author developed a general arrival model. In contrast, Olszweski (1994) and Leeuwarden (2006) both noted significant differences in terms of variability and increases in delay due to arrival distributions. Van Leeuwarden (2006) showed significant differences in terms of delay and performance between Poisson and geometric distribution. Olszweski (1994), demonstrated through simulation, that arrival distribution had a strong effect on delay. He showed that doubling I-ratio from 1 to 2 increased delay by up to 50%. While it is not the aim of this study to investigate the impact of I-ratio or a chosen probability on traffic performance, the results of these studies further underscore the need to understand and properly quantify the stochastic properties of arrival distribution. We currently lack information pertaining to arrival distributions and properties at signalized intersections in Lagos. Proper understanding of arrival properties based on field observation will enable engineers and planners develop more reliable predictive and operational performance models. Additionally, such information can serve as a key component for signal warrant analysis and also the type of control system to be installed (fixed, semi-actuated or fully actuated). There is a general consensus that arrivals characteristics at intersections are highly variable, and they are heavily dependent on factors such as type of intersection. There is also a growing body of work regarding time-resolution of traffic data collection and the impact on 1-ratio and variability. However, there is limited research considering these two factors together. In addition, we investigate whether Poisson is truly a general fit for traffic arrivals at intersections. This study characterizes traffic arrivals at different intersections under various operating conditions in Lagos, with specific interests in how various time resolutions and the presence of upstream signals affect arrivals patterns, I-ratio and invariably, the probability distribution of arrivals.

2.0 MATERIALS AND METHODS

A balanced experimental approach was adopted in this work. Two factors were considered. First, intersection type (isolated and connected) and second, time resolution (10, 30, 60, 90, 120, 150 seconds respectively). The data set used for this work was collected at six different sites in Lagos Metropolis over four weeks (June - July 2018). The screen lines (where traffic data is collected) at all the six selected intersections were situated beyond the point where the maximum queue typically reaches during peak periods. This is to ensure that true arrivals are captured and not the number of vehicles being served.

Three of the sites are isolated with no upstream signal within 800m of the approach under investigation and the remaining three were connected because they had traffic signals at least 800m upstream of the approach under investigation. Each site was studied for a duration of 6 hours each day for one working day with the six hours spreading over two hours in the morning (7am-9am), two hours in the afternoon (12pm-2pm) and two hours in the evening (5pm-7pm). This ensured we had a balanced design, since all observations were replicated an equal number of times. The

https://doi.org/10.51459/futajeet.2021.15.1.208

selection of three isolated and what we call connected (although, this does not mean they are coordinated or operating on the same cycle length), also enabled us examine the effect of upstream signal on arrival distribution (I-ratio/probability distribution) and the combined effect of intersection type and time resolution on I-ratio. The two major criteria for selecting connected intersections, in order to observe the effect of upstream nearby signals on arrival distribution were the distance of maximum of 800m and that the upstream signal functions properly and it is obeyed by motorist. The predominant adjacent land-use for all sites under investigation was commercial. An observer was positioned at a carefully selected screen-line to capture arrival of traffic. A computer program was used to collect traffic data at 10 seconds counting intervals for each two-hour period. Majority of previous research works have adopted 30 or 60 seconds or the cycle length. However, as mentioned in the previous section, longer intervals may obscure variability and thus yield inaccurate and unrepresentative arrival distributions. Moreover, high resolution time intervals may enable us better understand the implications of selecting one-time interval over the other, since from the high resolution (10 sec), we can still aggregate the data to lower resolutions (30 seconds, 60 seconds or the entire cycle length). Microsoft excel and its add-ins were used extensively in aggregating the various traffic data into the appropriate time resolutions and performing basic statistical analysis. The following analysis were performed: Two-way ANOVA (hereafter ANOVA) to examine the effect of time resolution and intersection type and their interaction effect on I-ratio. Spectral analysis was performed to examine if any strong periodicity exists due to intersection type. White noise tests such as Fisher's Kappa test and Barlett's Kolmogorov-Smirnov (BSK) test were performed to prove that the periodicities observed at various intersections were not random. Spectral analysis and white noise tests were performed using XLSTAT 2018 which is an add-in for advanced statistical analysis and modelling. A brief description of these methods is provided below.

2.1 Index of Dispersion/I-ratio

This index allows fitting count data into three main probability distributions. The results of I-ratio can be interpreted using Table 1 and expressed mathematically in (1). The probability distribution adopted has significant implication for data generation during modelling and simulation. Thus, appropriate distribution need to be selected for accurate modelling and simulation exercises.

$$I - ratio = \frac{\sigma^2}{\mu} \qquad (1)$$

where
σ^2 is variance of observations
μ is mean of observations

Table 1: I-ratio fit

I-ratio	Fit/ Distribution
< 1	Binomial
1	Poisson
> 1	Negative Binomial

2.2 ANOVA

ANOVA is a method for quantifying differences in the means of a set of observations by associating the observed mean differences with category/group/treatment effects, the interaction between the

https://doi.org/10.51459/futajeet.2021.15.1.208

categories/groups/treatments and random effects (Armstrong *et al,* 2002). To put it in another way, ANOVA examines the variation in the data to understand whether it is systematic (i.e. if it is due to the categories) or as a result of chance. The categories may have one or more levels such as low, medium or high which allow for the examination of two types of effects: main effects and interaction effects (Kim, 2014). The general form of ANOVA is (2)

$$x_{ij} = \mu + a_i + \varepsilon_{ij}$$

$$i = a, b \ldots\ldots k \text{ and } j = 1, 2, 3 \ldots.. n \qquad (2)$$

x_{ij} is the response or dependent variable (in this study, I-ratio), i refers to a group/category/treatment and j refers to replication or measurements within each category, μ is the overall mean of observations, a_i is the category effect, ε_{ij} is random error due to natural variation or measurement error. In this study, two categories were considered for ANOVA. The first was intersection type with two levels: isolated and connected. The second category considered was time resolution which had six levels (10, 30, 60, 90, 120, and 150 seconds). A total of two hundred and sixteen (216) I-ratio samples were used for the ANOVA analysis. For each of the six intersections surveyed, the hourly mean I-ratio was measured for 6 hours at each time resolution (6 hourly mean I-ratio x 6 time-resolutions x 6 intersections). Thirty-six (36) samples were used at each time resolution (one hundred and eight each for isolated and connected respectively). The judgment of the statistical significance or insignificance of the categories and their interaction effect was based on the results F-value, p-value and plots. F-values greater than the critical F statistic combined with a small p-value often lead to a rejection of the null hypothesis. P-value 0.05 was adopted for this analysis.

2.3 Spectral Analysis

In order to further establish a significant difference between the traffic arrival patterns at isolated and connected intersections, spectral analysis was performed to identify if any periodicity was present. Spectral analysis is a method for identifying any specific cyclical patterns or periodicities in a time series. By using a Fourier transform, any time series is decomposed into a sum of sine and cosines waves with varying wavelengths and amplitudes (SAS, 2018; Tian and Fernandez, 1998) Therefore, breaking down a complex time series into a set of sines and cosine components expressed in (3).

$$x_t = \sum_{k=1}^{m} (a_k cos\omega_k t + b_k sin\omega_k t) + e_t \qquad (3)$$

x_t is raw uniformly spaced time series

t is time subscript 1, 2, 3…… n-1

m is number of frequency from Fourier transform;

where $m = \frac{n}{2}$ if n is even and $m = \frac{n-1}{2}$ if n is odd

k is frequency subscript 1, 2, 3, …… m-1

a_k is coefficients of the cosine component specifying its amplitude

b_k is coefficients of the sine component specifying its amplitude

ω_k is Fourier frequencies,

where $\omega_k = \frac{2\pi k}{n}, subject\ to\ 0 \leq \omega_k \leq \pi$

e_t is random error

The periodogram which enables the detection of periodicities in a time series is based on the coefficients a_k and b_k. These coefficients obtained through Fourier transformation and with a little modification, become the value or ordinate or amplitude at a specific frequency k. The amplitude or value J_k of a periodogram is expressed in (4)

$$J_k = \frac{n}{2}(a_k^2 + b_k^2) \qquad (4)$$

A large J_k at a particular period or frequency suggests a strong periodic component. However, eyeballing of the plots alone might be misleading. As a result, a white noise test is performed to deduce whether high amplitudes observed at specific periods or frequencies are random or due to some inherent pattern.

Spectral analysis presented in this study was limited to 30 seconds time resolution only, since significant variability in I-ratio was observed at this time resolution.

2.3.1 White noise test for Periodicities

Two white noise tests were performed to determine whether the periodicities observed were simply random or due to some inherent patterns. Fisher Kappa's and Barlett Kolmogorov Smirnov tests were performed for this purpose. P-value 0.05 was adopted for both analyses. Therefore, since the null hypothesis for both tests are that the series is a white noise, p-values < 0.05 suggests the rejection of the null hypothesis and the acceptance of the alternative hypothesis; which is that the series has some periodicity embedded in it.

2.3.1.1 Fisher's Kappa (FK) test

FK test for white noise is predicated primarily on whether there is statistically significant difference between the peak and the average ordinate (or amplitude) of a periodogram, subject to a specific critical value (Moineddin *et al*, 2003). In a case where a time series is independently distributed i.e. autocorellation of zero, the periodogram value (J_k) is generally the same for all k. However, if a time series has an autocorrelation greater than zero, then there is significantly varying ordinate values across some of the frequencies (SAS, 2018).

2.3.1.2 Barlett Kolmogorov-Smirnov (BKS) Test

BKS tests whether there is a statistically significant difference between a normalized cumulative periodogram and a cumulative distribution function of a uniform (0,1) random variable (Moineddin *et al*, 2003; SAS, 2018). A normalized cumulative distribution is expressed in (5) below. The test calculates the absolute maximum deviation between the two distributions.

$$F_j = \frac{\sum_{k=1}^{j} J_k}{\sum_{k=1}^{m} J_k}, \quad j = 1,2,3 \ldots . m-1 \qquad (5)$$

where $m = \frac{n}{2}$ if n is even and $m = \frac{n-1}{2}$ if n is odd.

In both BKS test and FK test, low p-values (< 0.05) means the null hypothesis can be rejected and the periodicities observed are not random.

3.0 RESULTS

The effect of time resolution, intersection type and their interaction on I-ratio and probability distributions

3.1 Time resolution

The results of time-resolution analysis in Table 2 show that the same traffic arrival data examined under different time resolutions yield distinct I-ratios. Although, the differences are not generally very large at

https://doi.org/10.51459/futajeet.2021.15.1.208

Table 2: Time resolution and I-ratio

TIME RESOLUTION (SECS)	10	30	60	90	120	150
INTERSECTION	I-RATIO	I-RATIO	I-RATIO	I-RATIO	I-RATIO	I-RATIO
ALAGOMEJI (Connected)	1.5	2.5	2.85	2.6	1.5	1.8
GBAGADA (Connected)	1	1.3	0.98	1.29	1.11	1.02
MKO (Connected)	4.9	9.2	4.6	1.9	3.4	2.99
ST FINBARRS (Isolated)	0.87	0.81	0.76	1	0.93	1.3
IYANA OWORO (Isolated)	1.14	1.19	1.25	1.06	1.3	0.86
FESTAC (Isolated)	1.41	1.34	1.21	1.17	1.34	1.44

isolated intersections, they appear much larger at connected intersections.

3.2 ANOVA

Based on the computed F-statistic (8.606) and p-value <0.0001 in Table 3, the model is clearly statistically significant at 5% alpha level.

Table 4 shows the different sources of variation for the I-ratio results, considering time resolutions, intersection type and the interaction effect. The three factors have a statistically significant effect on I-ratio, with p-value of < 0.05 and R^2 of 0.32. The INTERSECTION variable accounts for the largest source of variation.

The time resolution chosen to examine arrival characteristics also has a statistically significant impact on I-ratio with p-value 0.002. The results also show that a statistically significant interaction effect exist between time resolution and intersection type on I-ratio with p-value 0.000 (Figures 1a and b). Overall, the results show that the strongest influence on I-ratio is whether the intersection is connected or isolated. In Figure 1a and b, Intersection-1 and Intersection-2 near the x represent isolated and connected intersections respectively. The two graphs (Figure 1 and b) show that I-ratio at isolated intersections are very close to 1 at virtually all time resolutions. I-ratio also remains nearly parallel across all time resolutions. However, as it can be seen in 1a, there is a huge spike at 2 and appreciable spikes at 1 and 3. Similarly, in figure 1b, all points at 1 on the x axis clearly cluster at one point while the points at 2 are scattered. This further highlights the interaction effect of intersection type and time resolution. Clearly, there is a striking difference in the case of connected intersections with influence from upstream signals. The two graphs (Figure 1a and b) show that time has a strong effect on I-ratio particularly at 30 seconds interval which is labelled as 2 for both time and intersection. The remaining time resolutions, that is, 10, 60, 90, 120 and 150 seconds are labelled as 1, 3, 4, 5 and 6 under time respectively. Intersection labelled as 1 represents isolated intersections and 2 connected intersections. Finally, as shown in Table 6, the mean I-ratio which is essentially the variability of traffic arrivals, is higher at connected intersections in comparison to isolated intersections. The variability is particularly higher at 30 seconds intervals at the connected intersections.

3.3 Spectral Analysis of Arrivals at Isolated and Connected Intersections

The results show that isolated intersections have no periodic pattern, as the white noise tests failed. Connected intersections on the other hand showed strong periodic pattern.

https://doi.org/10.51459/futajeet.2021.15.1.208

Table 3: Anova: Two-Factor (type of intersection and time resolution) and their effect on I-ratio

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	11	56.666	5.151	8.606	**< 0.0001**
Error	204	122.105	0.599		
Corrected Total	215	178.771			

Table 4: Anova results for source of variation

Source	DF	Sum of squares	Mean squares	F	Pr > F
TIME	5	11.880	2.376	3.977	**0.002**
INTERSECTION	1	28.968	28.968	48.482	**< 0.0001**
TIME*INTERSECTION	5	15.615	3.123	5.227	**0.000**

Table 5: Mean I-ratio based on intersection type and time resolutions

INTERSECTION\TIME (Secs)	10	30	60	90	120	150
Isolated	1.065	1.087	1.171	1.126	1.174	1.221
Connected	1.883	2.869	2.068	1.497	1.548	1.396

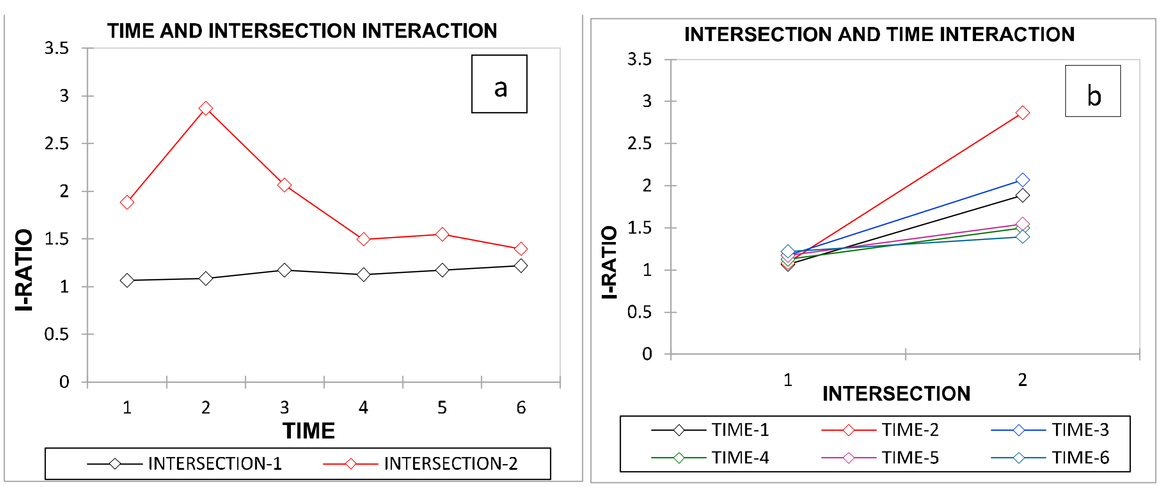

Figure 1a and b: Interaction effect between time resolution and intersection type

https://doi.org/10.51459/futajeet.2021.15.1.208

The results for both types of intersections are presented below.

3.3.1 Isolated intersections

Iyana Oworo

The white noise test (Table 6) for the data at this intersection shows the patterns in the data set are essentially random in nature as the p-value is greater than the significant level of 0.05. The data points in the spectral density graphs (Figure 2a and b) also show no distinct pattern. The presence of multiple spikes (Figure 2a) suggest very rapid oscillations across all frequencies. Although noticeable spikes are observed at frequency of approximately 2 and period of 3, the difference between these peak spectral densities or amplitudes and the spectral densities at any other period or frequency is not statistically significant. Therefore, no discernable cyclical pattern is present in the data.

Table 6: White noise test

Statistic		Value	p-value
Fisher's kappa	0.155	6.369	
Bartlett's Kolmogorov-Smirnov		0.073	0.594

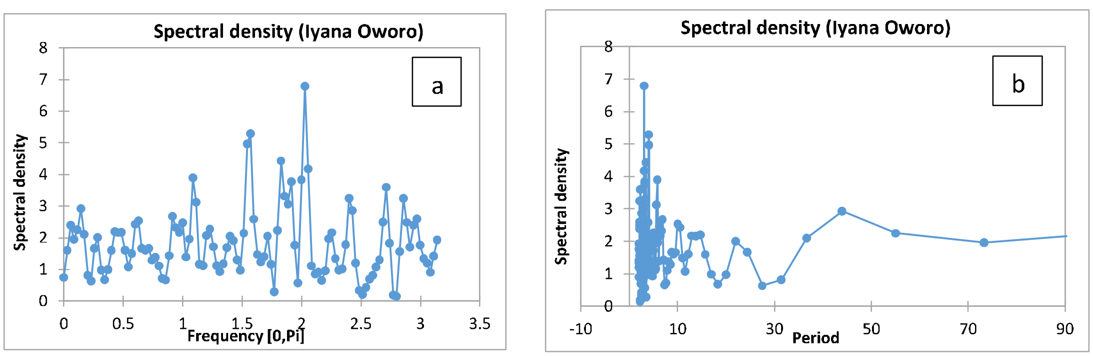

Figure 2a and b Spectral density chart (frequency and period) at Iyana Oworo

Festac

The white noise test (Table 7) for the data at this intersection shows the patterns in the data set are essentially random in nature as the p-value is > 0.05. While there is a noticeable spike close to 1 in Figure 3a and close to 10 in Figure 3b of respectively, the amplitude at those points are not statistically different from the amplitudes at other points.

Saint Finbarrs

The white noise test (Table 8) for the data at this intersection shows the patterns observed in the data-set are essentially random in nature, as the p-value is greater than the significant level of 0.05. Figure 4a shows clearly that multiple spikes are present in the data. However, the amplitudes at the periods and frequencies are not statistically different. Hence, traffic arrivals are simply sporadic at this intersection. Therefore, no distinct cyclical pattern is discernable.

https://doi.org/10.51459/futajeet.2021.15.1.208

Table 7: White noise test at Festac

Statistic	Value	p-value
Fisher's kappa	6.508	0.066
Bartlett's Kolmogorov-Smirnov	0.098	0.606

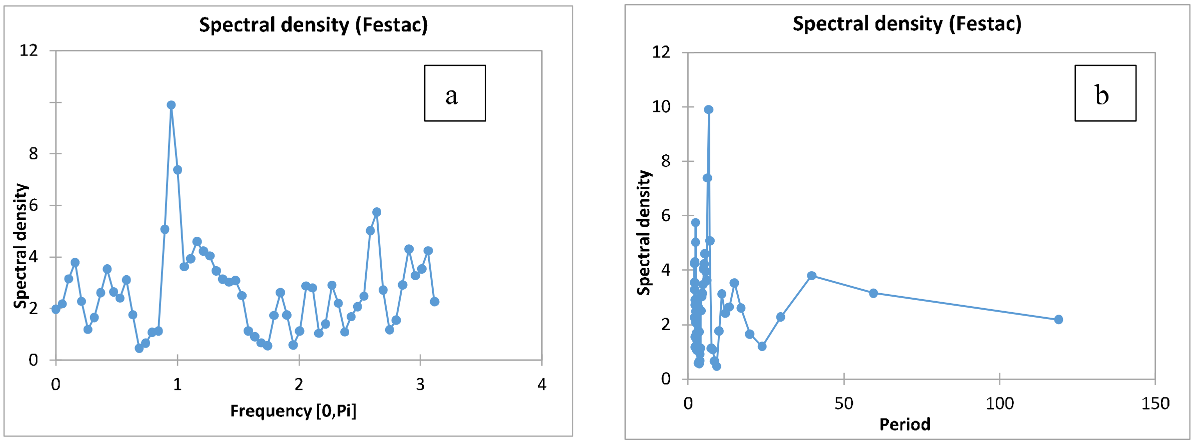

Figure 3a and b Spectral density chart (frequency and period) at Festac

Table 8: White noise test at Saint Finbarrs

Statistic	Value	p-value
Fisher's kappa	4.032	0.683
Bartlett's Kolmogorov-Smirnov	0.126	0.285

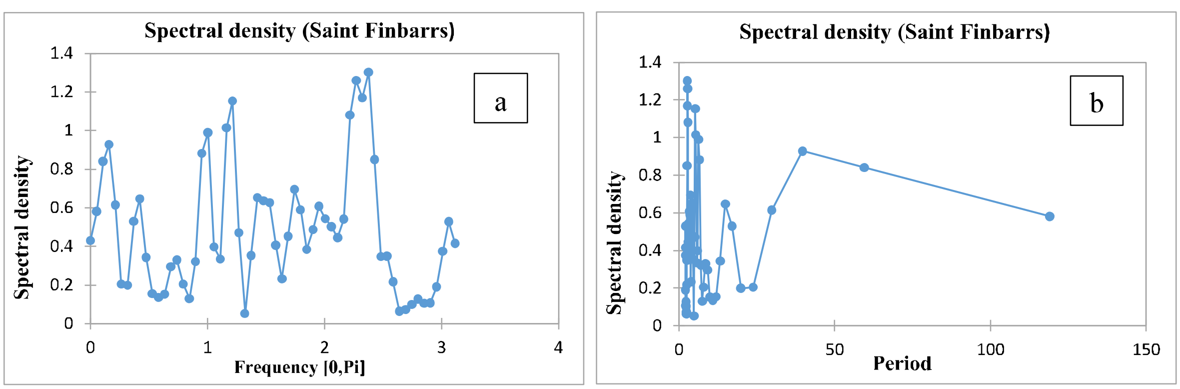

Figure 4a and b Spectral density chart (frequency and period) at Saint Finbarrs

3.3.2 Connected intersections

Gbagada

The results from the white noise test in table 9 shows some specific patterns are discernable from the data as p-value are significantly lower than 0.05 alpha level. This is further verified in the spectral density charts (Figure 5a and b). Clearly, a strong pattern can be observed with a very sharp peak in both charts with a peak amplitude of 37.7. The can be interpreted as a strong periodic pattern with traffic arrivals varying over regular cycles of 108 seconds.

https://doi.org/10.51459/futajeet.2021.15.1.208

Table 9: White noise test for Gbagada

Statistic	Value	p-value
Fisher's kappa	14.123	< 0.0001
Bartlett's Kolmogorov-Smirnov	0.217	0.006

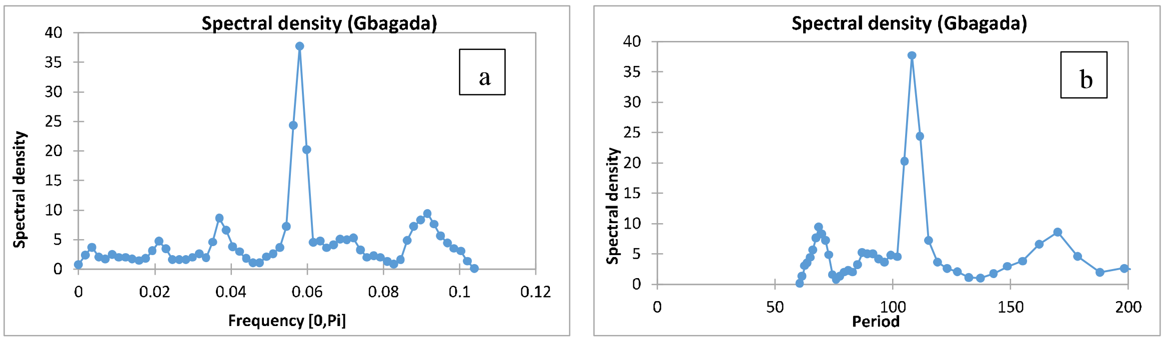

Figure 5a and b Spectral density chart (frequency and period) at Gbagada

MKO Gardens

The results of the white noise test in table 10, show there is a pattern that can be extracted from the data, as the p-values are both < 0.0001. The spectral density charts (Figure 6) show a periodic pattern with peak amplitude of 68.7. The very sharp distinct peaks in both charts can be interpreted as a strong periodic pattern with traffic arrivals varying quite regularly at approximately 93 seconds.

Table 10: White noise test for MKO Gardens

Statistic	Value	p-value
Fisher's kappa	18.884	< 0.0001
Bartlett's Kolmogorov-Smirnov	0.320	< 0.0001

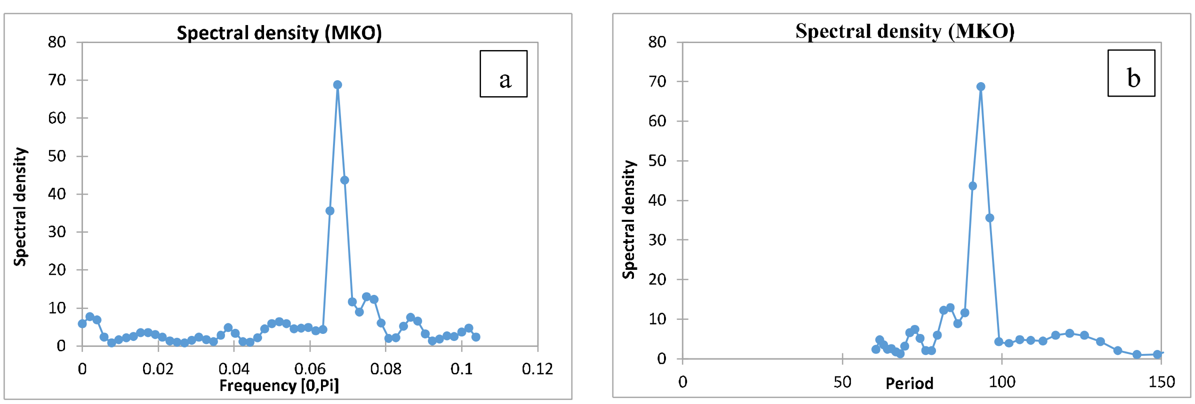

Figure 6a and b Spectral density chart (frequency and period) at MKO Gardens

https://doi.org/10.51459/futajeet.2021.15.1.208

The data from this intersection also shows periodicity, although it may not be as conclusive as those of the other two intersections, primarily because one of the white noise tests failed. The white noise test in table 11 also shows that the spikes with a peak amplitude of 36.7 observed in Figure 7a and b are statistically significant based on the FK test. However, the BKS test failed the test with p value 0.053. This may suggest that the periodicity at this intersection is not as strong as those of other connected intersections. Nonetheless, Figure 7a and b still show strong peaks which suggests a pattern of periodicity which may not due to randomness. This observation can be interpreted as a periodic pattern with traffic arrivals varying regularly at approximately 133 seconds.

Table 11: White noise test for Alagomeji

Statistic	Value	p-value
Fisher's kappa	13.937	< 0.0001
Bartlett's Kolmogorov-Smirnov	0.172	0.053

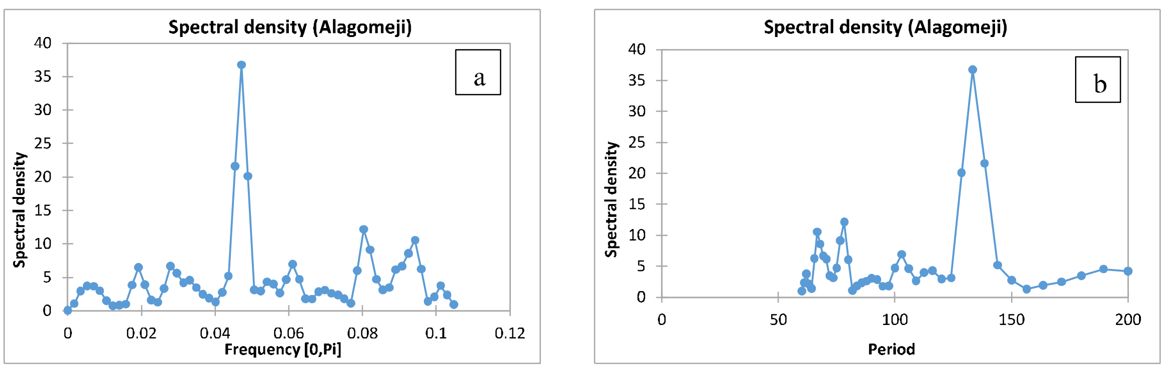

Figure 7a and b Spectral density chart (frequency and period) at Alagomeji

4.0 DISCUSSION

This study has highlighted some striking differences with regards to the characteristics of traffic arrivals and probability distributions under various conditions. Probability distributions were predominantly negatively Binomial (76.4%), followed by Binomial and Poisson which represented 20.4% and 3.2% respectively. A further breakdown of the distributions based on intersection type still showed that negative Binomial was the most prevalent fit for arrivals. At connected intersections, Negative Binomial, Binomial and Poisson were observed in 88.9, 10.2 and 1.2% of cases respectively. While at isolated intersections, negative Binomial, Binomial and Poisson were observed in 63.9, 30.6, 4.6% of cases respectively. Therefore, a key finding in this work is that in contrast to the vast majority of literature which suggests a Poisson distribution for traffic arrivals, the negative Binomial appears to be superior. The negative Binomial distribution appeared to be the better fit in for the vast majority of cases and this finding agrees with (Ritchie, 1983). The results of this study are also in agreement with (Mauro and Branco, 2013) which showed that Negative Binomial was a better

https://doi.org/10.51459/futajeet.2021.15.1.208

approximation for most of the traffic data collected. In this study, negative Binomial, at least using I-ratio, was a better fit for the data and this was the case irrespective of type of intersection or time resolution selected

While isolated intersections can be assumed to operate under near-Poisson distribution, as by and large, many of the observations appeared to have I-ratios close to 1. However, a conundrum still exists because only 4.6% of observations at isolated intersections actually fitted the Poisson distribution (i.e. I-ratio of 1). In many of the instances, I-ratios were slightly less or greater than 1, for example, 0.95 or 1.1. Strictly speaking, these should be categorized as Binomial and Negative Binomial distributions respectively. However, these values are quite close to 1 and more research is needed to understand the threshold of I-ratio deviations that may result in significantly different performance during modelling or simulation.

Given the randomness of arrivals at isolated intersections, it might appear more operationally efficient to adopt some form of adaptive signal control to deal with the variability. The random nature of arrivals at isolated intersections was further verified by performing a spectral density analysis. The results of the spectral analysis showed no periodic arrival pattern and the oscillating observations in the chart were primarily as a result of white noise. The results from Fisher's test and Kappa statistic for all isolated intersections had p-value greater than 0.05 which suggest arrival patterns are random or due to white noise. On the other hand, connected intersections with upstream signal influence showed strong periodicity based on the spectral density analysis performed. Clear periodic patterns were observed at various times ranging from 93 seconds at MKO, 108 seconds at Gbagada and 133 seconds at Alagomeji respectively. The periodic patterns observed are probably as a result of the platooning and filtering effect from upstream signals (Rouphail et al, 1999). Since, the number of vehicles arriving at a downstream signal are typically dependent on the maximum number of vehicles which are released from upstream and the duration of the red time at the upstream intersection. Furthermore, the results from white noise test using Fisher's Kappa test statistic showed the observations were non-random with p-values for all connected intersections less than 0.0001. This suggests that these intersections will operate more optimally if there is some form of coordination with the immediate upstream signal.

The variability of probability distributions is a key factor which may affect overall congestion (Law, 2013; Olszewski, 1994). The I-ratios observed in this study were wide, ranging from 0.37-9.2. This further underscores the high traffic variability from intersection to intersection and the need to examine arrival patterns carefully before and during any modelling effort. The results of this study may have a somewhat larger impact on traffic simulation rather than analytical traffic signal performance models such as the HCM and Ackelik models. This is primarily because simulation efforts require measured traffic data input in addition to the specification of their probability distributions. On the other hand, the need for probability distribution specification in analytical models such as those mentioned previously are obviated as

https://doi.org/10.51459/futajeet.2021.15.1.208

they almost exclusively require measured traffic data input for computation. While measured traffic data are required, in traffic simulation modelling, probability distributions are necessary to generate data for the hundreds or thousands of trial runs for the model. Therefore, misspecification of the probability distribution leads to wrong data being generated and invariably inaccurate results and higher output error (Law, 2013).

Time resolution of traffic data has a significant impact on the statistical fit of the data. The results showed that using the different time resolutions on the same one-hour traffic data yield dramatically different I-ratio and hence statistical fits. For example, for the same one-hour traffic data, MKO intersection had I-ratios of 4.9, 9.2, 4.6, 1.9 at 10, 30, 60 and 90 seconds respectively. While Saint Finbarr's had I-ratios of 0.87, 0.81, 0.76 and 1 at 10, 30, 60, and 90 seconds respectively. This can have a significant effect on modelling efforts particularly simulation models that typically involve traffic data generation at selected time intervals. Higher and lower time resolutions should both be considered when developing timing plans or during optimization efforts. For example, a shorter time/higher time resolution such as 10-30 seconds may be more ideal in some actuated systems were green splits and cycle times are short and where parameters such as green extension, and gap time are critical. While in the case of higher time resolutions such as 60-150 seconds may be quite important for fixed time control when there is need to understand how arrivals during the red time of an approach at an intersection will vary from cycle to cycle over the course of an hour for the purpose of determining the green split.

Two-way ANOVA was performed and it showed a strong interaction effect between time resolution chosen and type of intersection (isolated or connected). These both had a huge impact on variability of traffic arrivals at the studied sites, with intersection type being the most important factor. The results showed that most I-ratios for isolated intersections were very close to 1 with very little difference in I-ratio irrespective of time resolution selected. In contrast, connected intersections showed strong variation in I-ratio under different time resolutions. The strong interaction effect observed particularly at connected intersections suggests, that care should particularly be taken when carrying out planning, modelling and timing or optimization efforts at connected intersections. The purpose of the analysis being performed or project objective should guide the time resolution selected.

5.0 CONCLUSION

It is imperative that counts and analysis of traffic for signal timing, modeling and performance analysis should be based on the objective of the projects, as adopting the appropriate time resolution is necessary to achieve the desired results. It has been established in previous research dating back many decades, that isolated and connected intersections possess quite different arrival patterns. This work has taken it a step further by comparing these patterns under different time resolutions. The results show that the traditional approach to categorizing arrival patterns are too broad. Moreover, as opposed to conventional wisdom, the Negative Binomial distribution represents majority of

https://doi.org/10.51459/futajeet.2021.15.1.208

the data collected in Lagos regardless of intersection type or time resolution, at least if the I-ratio approach is adopted.

More importantly, this work has shown that arrivals at intersections are highly variable and this is dependent on the type of intersection and time resolutions. Additionally, the interaction between the two variables can have a significant impact on arrival patterns at an approach of an intersection. Variability is the bane of most control systems and traffic signals are not excluded. Incompatibility between arrivals and timing plans can lead to severe underperformance through either excessive or inadequate green time allocation. This ultimately leads to delays and long queues which contribute to road user stress, frustration, discomfort, increased cost of vehicle maintenance, loss of productive hours and increased emissions. It is therefore imperative that the purpose of any analysis or traffic operation project objective be considered carefully, before data collection even begins, since we have shown that higher time resolutions can be aggregated to lower time resolutions.

REFERENCES

Adams, W. F. (1936). Road traffic considered as a random series. *J. Inst. Civ. Engrs, 4*, 121-130. http://dx.doi.org/10.1680/ijoti.1936.14802

Akcelik, R. (1988). The highway capacity manual delay formula for signalized intersections. ITE journal,58(3), 23–27

Armstrong, R. A., Eperjesi, F., and Gilmartin. B. (2002) The application of analysis of variance (ANOVA) to different experimental designs in optometry. Ophthal. Physiol. Opt. 22: 248–256

Barua, S., Das, A., and Roy, C.K (2015). Estimation of Traffic Arrival Pattern at Signalized Intersection Using ARIMA Model. International Journal of Computer Applications (0975 – 8887). DOI: 10.13140/RG.2.1.5029.9363

Beckmann, M. J., McGuire, C. B., and Winsten, C. B. (1956). Studies in the Economics of Transportation. Yale University Press.

Cheng Cheng, Yuchuan Du, Lijun Sun & Yuxiong Ji (2015): Review on Theoretical Delay Estimation Model for Signalized Intersections, Transport Reviews. DOI:10.1080/01441647.2015.1091048

Farivar, S. (2015). Modeling Capacity and Delay at Signalized Intersections with Channelized Right-turn Lanes Considering the Impact of Blockage. PhD Thesis. University of Nevada, Reno, Department of Civil and Environmental Engineering,

Francesco Viti and Henk J. Van Zuylen (2009) The Dynamics and the Uncertainty of Queues at Fixed and Actuated Controls: A Probabilistic Approach, Journal of Intelligent Transportation Systems, 13:1, 39-51 http://dx.doi.org/10.1080/15472450802644470

Gerlough D. L., and Barnes, F. C. Poisson and Other Distributions in Traffic. Eno Foundation for Transportation,

https://doi.org/10.51459/futajeet.2021.15.1.208

Saugatuck, CT, 1971

Gerlough D. L., and Huber, M. J. (1975) Traffic Flow Theory. TRB, Special Rept. 165, 1975

Greenshields, B. D., Shapiro, D., and Ericksen, E. L. (1947). Traffic Performance at Urban Street Intersections. New Haven, CT: Tech. Rep. No. 1, Yale Bureau of Highway Traffic

Ha, D. H., Aron, M., and Cohen, S. (2012). Time Headway variable and probabilistic modeling. *Transportation Research Part C*, *25*, 181-201. http://dx.doi.org/10.1016/j.trc.2012.06.002

Harahap, E., Darmawan, D., Fajar, Y., Ceha, R., Rachmiatie, A. (2019) Modeling and simulation of queue waiting time at traffic light intersection. Journal of Physics: 1188(2019) 012001 doi:10.1088/1742-6596/1188/1/012001

HCM. (1985-2010). Highway capacity manual. Washington, DC: Transportation Research Board.

Heidemann, D. (1994) Queue Length and Delay Distributions at Traffic Signals. Transpn. Res.-B. Vol. 28B, No. 5, pp. 377-389

Kim, H (2014) Statistical notes for clinical researchers: Two-way analysis of variance (ANOVA)-exploring possible interaction between factors. Open lecture on statistics http://dx.doi.org/10.5395/rde.2014.39.2.143

Kinzer, J. P. (1933). Application of the Theory of Probability to Problems of Highway Traffic. *Inst. Traff. Engrs.,5*, 118-124

Law, A.M. (2013). A Tutorial On How To Select Simulation Input Probability Distributions. *Proceedings of the 2013 Winter Simulation Conference*

Leeuwaarden, van, J. S. H. (2006). Delay analysis for the fixed-cycle traffic-light queue. (Report Eurandom; Vol. 2006014). Eindhoven: Eurandom.

Mauro, R., and Branco, F. (2013) Update on the Statistical Analysis of Traffic Counting on Two-Lane Rural Highways. Modern Applied Science; Vol. 7, No. 6; 2013. doi:10.5539/mas.v7n6p6

May, A. D. (1990). Traffic flow fundamentals. New Jersey, NY: Prentice Hall

Miller, A. J. (1963). Settings for fixed-cycle traffic signals. OR, 373–386.

Moineddin, R., Upshur, R.E.G., Crighton, E., and Mamdani, M. (2003). Autoregression as a means of assessing the strength of seasonality in a time series. *Population Health Metrics* 2003, 1:10 http://www.pophealthmetrics.com/content/1/1/10

Newell, G. F. (1960). Queues for a fixed-cycle traffic light. The Annals of Mathematical Statistics, 31(3), 589–597.

Newell, G. F. (1965). Approximation methods for queues with application to the fixed-cycle traffic light. SIAM Review, 7, 223–240.

Olszewski, P. (1990). Modeling of Queue Probability Distribution at Traffic Signals. In Proceedings of the 11th International Symposium on

Transportation and Traffic Theory. M. Koshi. Amsterdam, Elsevier: 549–588

Olszewski, P., S. (1994) Modeling probability distribution of delay at signalized intersections. Journal of advanced transportation. Vol 28, No. 3, pp 253-274. DOI: 10.1002/atr.5670280306

Ritchie, S. G. (1983) Application of Counting Distribution for High Variance Urban Traffic Counts. Transportation Research Record 905

Rouphail, N., Tarko, A., and Li, J. (1999). Traffic flow at signalized intersections. Transportation Research Board Special Report 165 (Traffic flow theory), Chapter 9. Washington, DC.

SAS Institute Inc. (2018). User's Guide The SPECTRA Procedure

Tian, J., and Fernandez, G. C. (1998). Seasonal Trend Analysis of Monthly Water Quality Data (June 1, 1998). University of Alberta School of Business Research Paper No. 2013-1203,The Western Users of SAS Software, 1999, Available at SSRN: https://ssrn.com/abstract=2284439

Webster, F. V. (1958). Traffic Signal Settings. Road Research Technical Paper, 39.

Xing, Y., Gao, Z., Qu, Z., and Hu, H. (2016) Study on Vehicle Delay Based on the Vehicle Arriving Distribution at Entrance Lanes of Intersection. Procedia Engineering 137, 599 – 608. doi: 10.1016/j.proeng.2016.01.297