MULTIVARIATE PLATOON DISPERSION MODELING AND SIGNAL COORDINATION WITH A PREDICTED PLATOON MAX-PRESSURE POLICY

A Dissertation

by

JIAN JIAO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yunlong Zhang
Committee Members,	Xiubin Wang
	Alireza Talebpour
	Sergiy Butenko
Head of Department,	Robin Autenrieth

May 2020

Major Subject: Civil Engineering

ABSTRACT

This research aims to describe the evolution of platoon vehicle speed distribution parameters and platoon characteristics change quantitatively, and utilize this information to develop advanced signal timing strategies for urban arterials that could handle platoon traffic effectively. Real platoon vehicle speed data was collected. Combined with data from simulation, the speed distributions for both homogeneous and heterogeneous traffic flow were studied. The speeds follow a truncated normal distribution for homogeneous flow and a Gaussian mixture model for heterogeneous flow. The factors that influence the parameters of the truncated normal distribution were identified and a multivariate distribution parameter model was built to describe the relationship between distribution parameters and influencing factors mathematically. A modified platoon dispersion model is developed based on the previous model. The multivariate model and the modified platoon dispersion model are proved to predict downstream speeds quite well, as model validation suggests. Based on the previous model development, a novel signal timing strategy, the Predicted Platoon Max-Pressure Policy is developed. The policy utilizes the information calculated from previous models and adjust signal timing for each approach of each intersection on the arterial in real time. The performance of the proposed policy is evaluated and compared with traditional methods in simulation. Results show that the proposed policy greatly improves all measures of effectiveness, from 15% to 50% under various scenarios. The practicality of the Predicted Platoon Max-Pressure Policy is further examined in different conditions. Simulation results prove that the policy has great potential in implementation in the real world.

DEDICATION

I would like to dedicate this dissertation to my father, Yongshun Jiao, and my mother, Jingchun Wu for their education and unconditional support, and being role models who led me grow up, and to my grandmother who took good care of me when I was little, and to my grandfather, who is exactly the kind of person I want to be. I would also like to dedicate this dissertation to my wife Xiduo Wang, for her love and support that help me get over the hard times.

ACKNOWLEDGMENTS

I would like to thank my advisor and the chair of my committee, Dr Zhang for his guidance on my academic studies in Texas A&M University. His advices and support will be one of the most precious things in my academic career and my future life. I would also like to thank my committee members, Dr Wang, Dr Talebpour and Dr Butenko for their support and valuable advices on my research. I am also grateful to my friends, faculties and staffs I met at Texas A&M University. Finally, I would like to express my thanks to all my family members who showed great support to my Ph.D. Study.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised and supported by a dissertation committee consisting of Dr Yunlong Zhang [advisor], Dr Xiubin Wang and Dr Talebpour of the Zachry Department of Civil & Environmental Engineering, and Dr Sergiy Butenko of the Department of Industrial & Systems Engineering.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a teaching fellowship from Texas A&M University.

There is no funding source for dissertation research

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

For major arterials in larger cities, urban intersections are often closely packed. It is beneficial to allow vehicles pass the intersections smoothly without having to stop. In order to reduce delay and the number of stops, signal coordination is often utilized on arterials to achieve optimal performance. Traditional signal coordination is based on fixed offsets, but a good knowledge of the characteristics of the upstream platoons my help improve the quality of signal coordination. Vehicle speed distribution is one of the methods that can best describe platoon dispersion characteristics and is often used in studies of platoon dispersion analysis as well as signal coordination. However, the traditional studies on vehicle speed distribution often assume that this distribution is fixed and do not change along the arterial. The parameters in the distribution model are selected as a fixed value. This may not be true since vehicle speeds change as time increases and the speeds are influenced by many factors along the arterial. Speed distribution parameters may change as influencing factors change, and even the distribution type may be different. Signal coordination may not be as effective at intersections further downstream. Hence, it is desirable to build a model that can describe the change in speeds over time along the arterial at different conditions, by describing the changes in the distribution parameters. The model can be further utilized in platoon dispersion model for arterial signal coordination. This study focuses on developing a flexible vehicle speed distribution model that describes the evolution of platoon speeds, by constructing mathematical relationships between distribution parameters and factors that may affect them, and incorporating the findings into a more reasonable platoon dispersion model. In addition, arterial signal timing strategies based on the developed model is also planned as an important part of the research effort. One of the novel scheduling methods is the implementation of Max-Pressure Policy developed from computer and electric engineering into the field of transportation engineering. In this study, an advanced Max-Pressure scheduling policy, the Predicted Platoon Max-Pressure Policy, is proposed to accommodate the evolution effect of the platoons and execute scheduling adjustments in advance, as well as taking into the switch-over delay into consideration.

1.1 Problem Statement

1.1.1 Traditional Platoon Dispersion Model

On urban arterials, vehicles proceed through the intersections in groups known as platoons. In the process of moving downstream, the platoon will spread out and causing vehicles arrive at the downstream intersection in a non-uniform way. This is because vehicles travel at different speeds. This phenomenon is called platoon dispersion. A non-negligible proportion of signal coordination strategies are based on platoon dispersion. The development of platoon dispersion theory can help improve the network performance both economically and environmentally.

To model platoon dispersion, three methods are most commonly used. They are Lighthill and Whitham's wave theory (1), Robertson's model (2), and Pacey's diffusion theory (3). Each of these methods has its limitations and problems. Lighthill and Whitham's theory, or the LWR model, is not very practical because its flow-density relationship equilibrium needs to be very accurate. Robertson's model is a recursive platoon dispersion model and has been used in the software TRANSYT. However, the model requires field data and is empirical in predicting vehicle behavior. The Pacey's model combines both theoretical and experimental work and has good accuracy in describing platoon behaviors. The main problem with Pacey's model is that it assumes the vehicle speeds follow a normal distribution. This assumption is often invalid, since the vehicle speeds cannot range from negative infinity to positive infinity as normal distribution does.

1.1.2 Possible Solutions

To address the problems with unrealistic high speeds and negative speeds in Pacey's model, more realistic distributions could be adopted to replace the normal distribution assumption by Pacey. These distributions include truncated distributions and mixed distributions. These distributions were proved to be able to represent vehicle speeds better than simple normal distribution, since the unrealistic speeds from normal distribution were omitted and heterogeneous flow conditions were considered. However, much more could be done besides adopting more complicated distributions to improve platoon dispersion model. Previous studies failed to consider the evolvement of the distribution over time and the influence of other factors on the speeds, resulting in the dispersion model to ineffectively describe the evolution of a platoon and lose its accuracy and effectiveness over time. In order to avoid such condition, the speed distribution model needs to be further modified.

1.1.3 Application in Arterial Signal Timing and Coordination

Speed distribution serves as an input in platoon dispersion model to calculate density at a certain location at a given time, and more importantly, the number of vehicles passed the location at the front of the platoon and the number of vehicles that have not passed the location at the rear of the platoon. This information could be utilized to optimize signal coordination along the arterial. If the speed distribution is not realistic and does not reflect the platoon characteristics well, signal coordination will not be effective and its performance will be greatly reduced. Some empirical distribution parameter values were presented by researchers before, but they may vary at arterials from different locations with different conditions. It is desirable to investigate how these conditions serve as influencing factors on the distribution parameters and develop corresponding distributions for different arterials. Signal timing and coordination could be modified based on platoon dispersion model with speed distribution parameters estimated according to the actual condition.

1.2 Objectives and Scopes

This research has two main objectives. The first one is to construct a multivariate model for speed distribution parameters so that platoon dispersion can be represented more accurately by implementing the multivariate model into the platoon dispersion model. The second is to utilize the new speed distribution characteristics and platoon dispersion model to develop a signal timing and coordination method with better performance.

In order to develop effective signal coordination for arterials, accurate upstream platoon information need to be provided. Understanding the speed distribution of incoming vehicles is essential for describing platoon dispersion characteristics. It is also desired to further investigate the change of vehicle speeds and the evolution of the speed distributions. To represent this gradual change in vehicle speed and speed distribution, a new model for speed distribution parameters that consider the change in these parameters based on different factors should be developed. Mathematical relationships between the distribution parameters and possible impacting factors should be built.

To develop such models, two issues need to be carefully addressed. The first issue is selecting the right factors that influence the changes in distribution parameters. A series of factors including time, traffic conditions and geometric conditions are considered and investigated. Statistical methods will be applied to test the significance of the factors and mathematical relationship should be built if possible. The second issue is to obtain enough required data for testing and building the distribution model. In order to have accurate and convincing data, real data collection should be conducted. Vehicle speed data are collected under different scenarios. The scenarios represent different values or levels of the considered factors. It should be noted that scenarios in real data collection are not able to cover all considered cases of the factors. Simulation study is used to aid data collection in cases which is hard to be obtained in real life. Vehicle speed distribution is then built on the combination of real and simulated data. The proposed model should be validated under different conditions. In addition, side friction factor should be considered.

With the speed distribution model, platoon dispersion characteristics could be described. A mathematical program to optimize arterial signal performance could be developed based on platoon dispersion characteristics. The developed signal timing and coordination strategy should be able to take the evolution of platoon characteristics into consideration. In addition, it is desirable to test the applicability of the proposed signal coordination strategy for different network conditions.

Therefore, the following specific objectives are to be accomplished:

1. Obtain real vehicle speed data aided by simulation data to investigate relationship between speed distribution parameters and possible factors.

2. Select factors which are significant in vehicle speed changes and build reliable model for speed distribution parameters, and develop platoon dispersion model based on the speed distribution parameter model.

3. Formulate signal timing or coordination strategies that is capable to take platoon dispersion and evolution of platoon characteristics into effect.

4. Test the signal coordination in various virtual network as well as simulated real networks.

1.3 Research Tasks and Dissertation Organization

The following task are completed in order to accomplish the objectives previously mentioned.

1. Literature review

2. Real and simulation data collection

3. Selection of influencing factors and statistical testing

4. Mathematical model development

5. Signal timing and coordination strategy development

6. Simulation and evaluation

7. Documentation

1.3.1 Task 1: Literature Review

This is the first task performed in order to accumulate enough background knowledge for the research topic. The literature review task also helped to determine the scope of the research so that new ideas were developed and meaningful and significant results were produced. A number of research topics were reviewed. These topics included traditional signal control and arterial signal coordination methods, advanced and newly-developed signal coordination strategies, vehicle speed distribution and platoon dispersion studies, and researches on multivariate model for distribution parameters as well as distribution studies from mathematical perspectives. The literature review was not limited to these topics. A number of other helpful research topics found during the during the study process were also included in the literature review.

1.3.2 Task 2: Real and Simulation Data Collection

Data collection is critical for the subsequent research process. Data with good quality are needed for distribution analysis. On one hand, simulation data alone may not be convincing enough to build a model. On the other hand, due to lack of data, real data are not sufficient for analysis on all possible cases considered. Therefore, a combination of real and simulated data was used for distribution analysis. Real data were used as the main data source while simulated data were used to cover cases which are hard to be accomplished in real life. The real data were collected during the research process. Simulated data were then collected based on building a same network as the real one in simulation software.

1.3.3 Task 3: Selection of Influencing Factors and Statistical Testing

A series of factors that may affect the vehicle speeds and platoon dispersion were selected before data collection and tested after data collection. Vehicle speed data were collected under different values of the factors, also referred to as cases. Each of these factors will be examined to determine if it has an effect on platoon speeds. Factors that has impact on the speeds went through a second test that examines the combined effect of the factors. This procedure decided the impacting factors to be considered in the speed distribution analysis.

1.3.4 Task 4: Distribution Fitting and Mathematical Model Development

In this task, the speeds were fitted to a number of distributions using statistical methods. Next, mathematical relationships were built between the vehicle speed distribution parameters and impacting factors. The mathematical relationship should be accurate in describing the distribution parameters, while being as simple as possible. This increases the applicability of the model. Part of the collected data were selected as the validation group to validate the developed model.

1.3.5 Task 5: Platoon Dispersion Model Development

To optimize the performance of arterial signal coordination, a platoon dispersion model is needed, in order to provide the vehicle density and calculate the number of vehicles in front and rear of the platoon. The platoon dispersion model utilizes the distribution parameter model to predict dispersion characteristics of the platoon at certain downstream locations.

1.3.6 Task 6: Signal Timing and Coordination Strategy Development

With the developed platoon dispersion model, follow-on studies on developing signal timing and coordination methods for arterials were conducted. Platoon density distribution and the number of vehicles in front and rear of the platoon were calculated as an input in optimizing arterial signal timing and coordination. Traditional mathematical optimization methods for signal coordination as well as advanced methods were developed and evaluated.

1.3.7 Task 7: Simulation and evaluation

The developed signal coordination strategy were evaluated using simulation software. On similar simulation network where the data were gathered, the proposed coordination strategies will be compared with some traditional methods to prove the advantage of the proposed method. In addition, the proposed coordination strategy was also tested under various scenarios for further evaluation on its application.

1.3.8 Task 8: Documentation

The research was documented in this final report which provides a complete description of the research process and results. The report include content such as problem statement, literature review, methodology, results, conclusion, recommendations etc. The gathering of the data, analysis on the data, and evaluation results were also documented in the final report.

The organization of the dissertation is as following. In Section 2, a comprehensive literature review on platoon dispersion, signal timing methods, and various transportation models using statistical methods is provided. Section 3 introduces the various methods used to collect and analyze the vehicle speed data, as well as the process to develop a multivariate distribution parameter model. Section 4 presents the results of the data analysis and evaluates the performance of the developed multivariate distribution parameter model. Section 5 utilizes to results of the previous sections to develop various signal timing strategies, and presents the evaluation results based on simulations. Section 6 presents some extensions of the proposed Predicted Platoon Max-Pressure Policy and discusses its practicality. The final sections provides conclusions from the research and gives recommendations on possible improvements and the directions of future researches.

2. LITERATURE REVIEW

A number of studies have been made on platoon dispersion in the past. Since the development of the original Robertson's theory (2) which was widely implemented in the traffic software for signal control like the "SATURN"(4), "TRAFLO"(5), "SCOOT"(6) and TRANSYT, and Pacey's diffusion theory (3), most of the later researches were made based on them. Seddon (7) further studied on Robertson's model and pointed out that the model was an equivalent of Pacey's model based on a shifted geometric distribution of travel time, while some studies (8), (9) showed that the vehicle travel time distribution is often normal, lognormal or gamma instead of a shifted geometric distribution. Pacey's theory was improved by Grace and Potts (10) by theoretically investigating Pacey's model in density aspect. Both Robertson's model and Pacey's model were further calibrated in the studies later on. The parameters of Robertson's model were calibrated based on link travel time (11), speed variability (12), number of lanes (13), turning vehicles in urban roads (14), and with high-resolution signal event data (15). The calibration on Pacey's model mainly focused on the speed distribution study, as proposed by Wei (16) and Badhrudeen (17).

Besides the early studies in platoon characteristics, platoon recognition is another interesting topic. The goal of this area is to identify the platoon in early stages and how it can help in signal control and traffic management. These studies include Gaur and Mirchandani's link density-based recognition algorithm (18), Chaudhary's field platoon identification and accommodation system at isolated intersections (19), (20), Krishnan and Polak's algorithm based on temporal flow density changes. Gaur and Mirchandani's method was further tested with real data and compared with cluster-based approach and image processing technique (21). Another study by Praveen and Ashalatha (22) focused on heterogeneous traffic flow with data collected from video graphic survey technique.

Recent studies on the platoon dispersion focused more on dynamic and stochastic side. The work by Li (23), (24) proposed to use Markov process for platoon dynamic behavior and state space model for individual vehicle behavior. He also investigated and derived statistical distribution models for various platoon characteristics such as platoon speed, platoon size, and within- and between-platoon headway. Li's studies concluded that the platoon speed follows a normal distribution, and a mixture of normal models will perform even better. Another research conducted by Shen (25) developed dynamic platoon dispersion model and studied factors that affect the model such as time, position of detectors, road section length, volume, and turning percentages. The simulation results showed that the dynamic versions outperform the original static version to a large extent. Other studies on platoon dispersion model includes stability and robustness of large platoons (26), platoon dispersion prediction with support vector regression (27), performance evaluation under midblock pedestrian and truck friction conditions (28), comparison with cell-transmission model (29), and applying platoon dispersion in analytical dynamic assignment process (30).

The above platoon dispersion and speed distribution studies are mostly based on homogeneous traffic flow with only passenger cars. It is also interesting to study the characteristics of heterogeneous flow which heavy vehicles like trucks and buses are involved. More complicated distribution will be required for mixed flows, such as mixture models. Dey (31) proposed that in addition to a unimodal distribution, a bimodal curve may also be realistic depending on the variation of speeds. Park, Zhang and Lord (32) proposed to use a finite mixture of normal distributions to capture speed heterogeneity based on real data. Ko and Guensler (33) also selected Gaussian mixture model in their effort to characterize traffic congestion. The Gaussian mixture model for speed distribution can be further applied to the derivation of platoon dispersion model, as performed by Wu (34), (35), if enough real traffic data could be obtained. Another study by Jiang (36) pointed out that the mixed simplified phase-type distribution is also capable of describing heterogeneous flow conditions and implementing into platoon dispersion model.

Many studies showed that the distribution type and parameter values used in the transportation engineering field may change as the influencing factors change. One of the first studies that goes into details is by Ye and Zhang (37). Their work showed that the headway distribution varies with different vehicle type and different vehicle compositions. Their study showed a good example on how the traffic volume and traffic compositions can affect the headway distributions. Though there are little researches that quantifies the evolution of the distribution characteristics, the work by Weng (38), (39) and Meng (40) on traffic safety was enlightening. In their researches, when historical statistical data was not available, they aim to build model using alternate datasets considering influencing factors. Their studies showed that the model based on the influencing factors perform as good as or even better in estimation and prediction than using historical data directly.

As mentioned previously, platoon dispersion model could be used for signal timing and coordination. A considerable number of researches have been made on signal timing and coordination, mainly focusing on one of the two objectives: maximizing bandwidth, or minimizing (optimizing) delay/stop or other types of performance measurement. Examples of the first type includes Brook's model (41), Messer's PASSER-II program (42), Little's MAXBAND program (43), Gartner's MULTI-BAND program (44), and Chaudhary's circular phasing optimization scheme (45). The second type that minimize delay/stop includes offline methods such as SIGOP-III (46), TRANSYT-7F (47), and online adaptive system such as SCOOT (48) and UTOPIA (49). There are however, studies aiming to achieve both objectives, such as the work by Skabardonis and May (50), Liu (51), and Lan and Messer (52). They either come to a bandwidth solution in the delay minimization program or solve the multiple objective function that minimize delay/stop and maximize bandwidth at the same time. The idea of multiple objectives is also applied in SYNCHRO (53). Some of the recent research work that optimize both objectives are by Zhou et al. (54), (55) that utilized uneven double cycling to reduce excessive delay at minor intersections. Some of the other recent researches on signal timing and coordination are on non-uniform flow involving heavy vehicles and connected vehicles, such as the person-based adaptive priority signal control with in a connected environment by Zeng et al. (56), arterial green-wave synchronous coordination model by Wei and Sun (57), simulation analysis of transit signal priority with arterial coordination by Mei et al. (58), and partition-enabled multi-mode band model for mixed traffic by Ma et al. (59).

3. RESEARCH METHODOLOGIES ON MULTIVARIATE DISTRIBUTION PARAMETER MODEL

3.1 Data Collection

Obtaining good quality vehicle speed data is critical to the analysis which include distribution fitting, significance tests and model construction. In terms of quality, the data should be able to represent real vehicle behavior. In terms of quantity, the data should have enough sample size for statistical analysis. In addition, the data should be collected from different conditions to represent the "cases" which indicates different values of the selected factors. Therefore, both real data and simulated data should be collected.

The real data are collected from Texas Avenue in College Station Texas. Texas Avenue is a major arterial that passes through the city of Bryan and College Station. The data were collected on weekdays from three links on Texas Avenue with different link lengths: the link between George Bush Drive and Harvey Road, the link between Harvey Road and Holleman Drive, and the link between Holleman Drive and Manuel Drive. The three links have different link lengths and their lengths are 1800 feet, 1250 feet and 100 feet respectively. This was to account for effects of different link lengths. There were about 3-5 data collection points with distances 300-500ft on each link to account for different travel times. The data were collected from three time periods during the day and later categorized into three groups according to volume level. The distances between two close objects and the times the vehicles reached the two objects were measured to calculate speed. Since the distance between the two objects were set to be small, the calculated speed could be assumed to be close to spot speed.

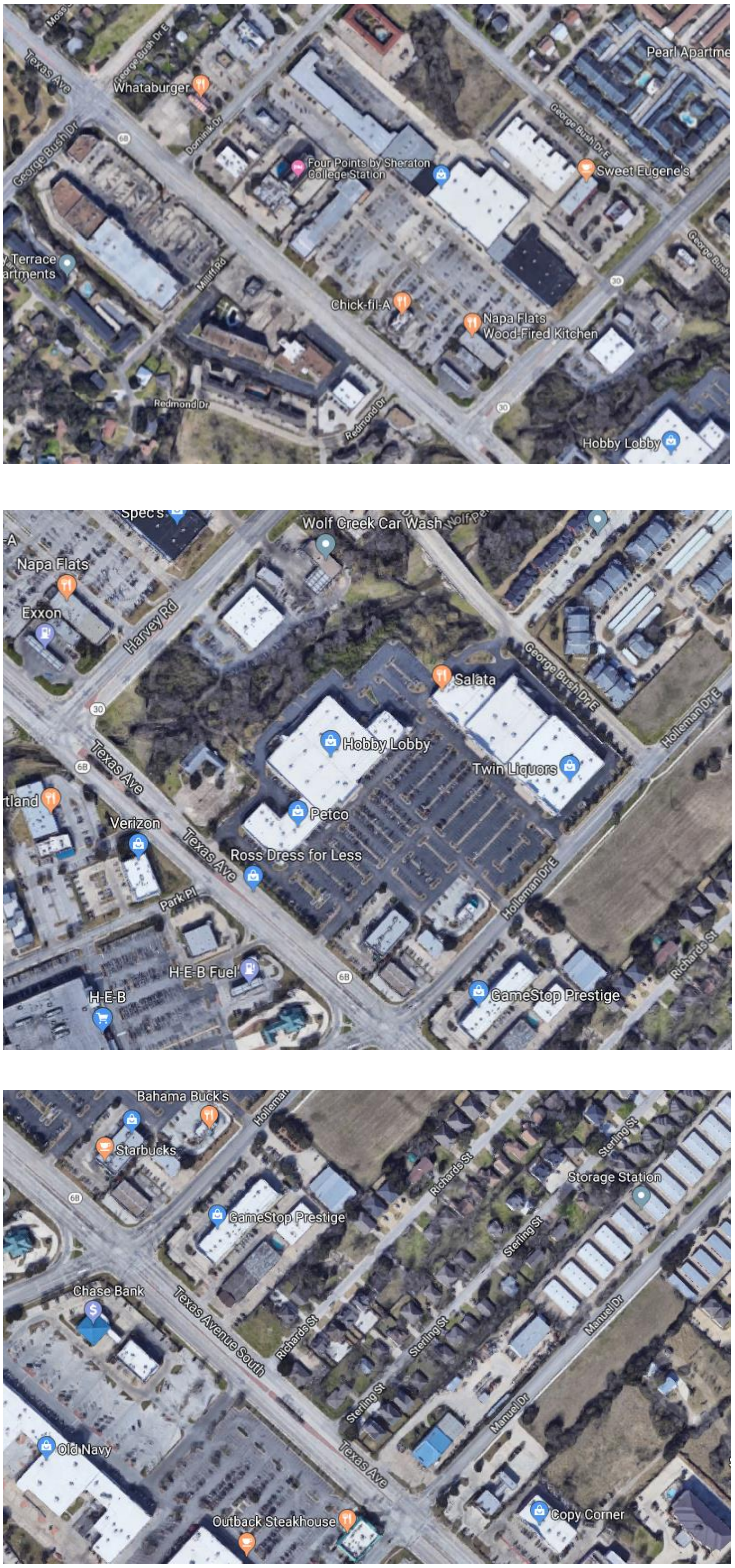


Figure 1. Three Links on Texas Avenue for Data Collection

To accommodate the needs of cases that cannot be obtained from real world data collection, the Texas Avenue was recreated in simulation software VISSIM. The geometric and traffic conditions were designed to be as close as the real arterial as possible. A layout of the VISSIM simulation network is presented below.

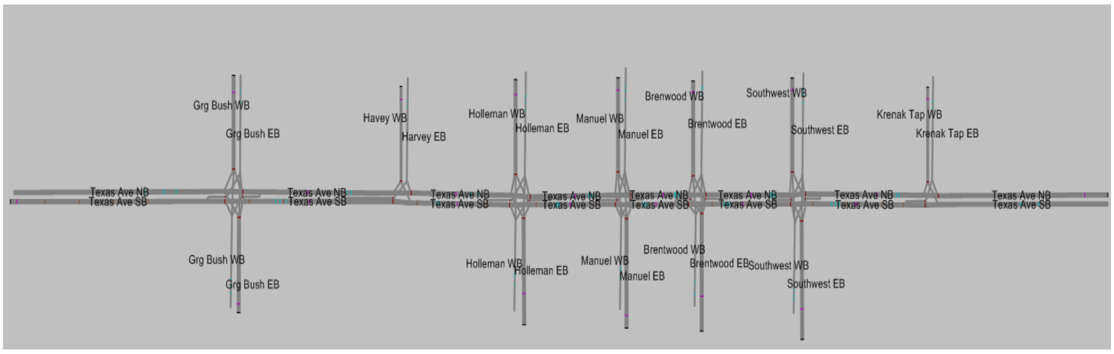


Figure 2. Layout of the Simulated Network for Texas Avenue

3.2 Distribution Fitting

This section explains the statistical methods adopted to fit the distribution of vehicle speeds and estimate the distribution parameters. In addition, since the preliminary results from Section 5 indicates that the vehicle speeds follow a truncated normal distribution, the characteristics of the truncated normal distribution are briefly explained in this section.

3.2.1 K-S Test

In the study, two types of statistical tests were conducted. The first test is the Kolmogorov-Smirnov test, or the K-S test. It is a non-parametric test for continuous distributions, and is based on the maximum distance between the theoretical distribution curve and the sample empirical distribution curve. Its test statistic can be defined as:

$$\mathrm{D} = \sup_x \left| F_{theoretical}(x) - F_{data}(x) \right| \quad (1)$$

where $\sup_x$ indicates the supermom of x, $F_{theoretical}(x)$ and $F_{data}(x)$ are the cumulative distribution function for the theoretical distribution and the sample data. The test statistic is then compared with the critical values in the K-S table to decide the acceptance or rejection of the null hypothesis, which should state that the sample is drawn from the theoretical distribution in the test. The K-S test is used to test the goodness of fit of various theoretical distribution on the vehicle speeds.

3.2.2 Akaike Information Criterion

The Akaike Information Criterion or the AIC provides a comparison of the quality of a group of statistical models for a given data sample, and is often used as a model selection criterion. The Akaike Information Criterion aims to balance the level of model fitting and the complexity of the model. The formula for the Akaike Information Criterion is:

$$\mathrm{AIC} = -2(\ln(L) - k)$$

Where k is the number of parameters in the model, and L is an indicator of the level of fit. A higher L indicates a better fit of the model. The AIC is used as a second type of statistical test in fitting theoretical distribution to the speed data. It is also used as a measure to examine the multiple distribution parameter models and select the best model among them.

3.2.3 Gaussian Mixture Model

The Gaussian mixture model is adopted to address the heterogeneity of the mixed traffic flow. It is a probabilistic model in statistics and can be used to represent subpopulation normal models within an overall distribution model. The model can be represented by the following:

$$p(x) = \sum_{i=1}^{K} \lambda_i N(x|\mu_i, \sigma_i) \quad (2)$$

$$N(x|\mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp(-\frac{(x-\mu_i)^2}{2\sigma_i^2}) \quad (3)$$

$$\sum_{i=1}^{K} \lambda_i = 1 \quad (4)$$

The Gaussian mixture model can have K components. Its ith component has a mean μ_i and variance σ_i. λ_i is the weight of the ith component of the mixture model. Equation (2) represents the general form of the overall model and Equation (3) represents the ith component with a normally distributed model. Equation (4) shows that the total probability normalizes to 1. The above equations are for a one-dimensional model and the Gaussian mixture model also has a multi-dimensional case.

3.2.4 EM Algorithm

The parameters of the Gaussian mixture model can be estimated based on the expectation-maximization (EM) algorithm (60), (61), (62). The EM algorithm is a statistical method that finds the maximum likelihood of the model parameters iteratively. The algorithm is best fit for models with equations that are hard to solve directly or involve latent variables. It is often used for mixture models where each data point from a certain component has a corresponding unobserved data point. There are two steps and the algorithm iteratively alternate between them.

The first step is the “expectation (E) step”, in which the log-likelihood expectation of the current step t is calculated as following:

$$Q(\theta|\theta^t) = E_{Z|X,\theta^t}(\log L(\theta; X, Z)) \quad (5)$$

where θ is the estimate of the parameter, X is the set of observed data, and Z is the set of unobserved data. The next step is the “maximization (M) step in which the following quantity is maximized:

$$\theta^{t+1} = argmax_{\theta} Q(\theta | \theta^t) \quad (6)$$

The parameters θ are first set to be some random initial values. Then the E step and the M step calculated iteratively until convergence.

3.2.5 Truncated Normal Distribution

From the preliminary analysis, the vehicle speeds were found to have truncated normal distribution. The truncated normal distribution is derived from a normal distribution with an upper and a lower bound. The truncated normal distribution is able to resolve the issue that normal distributions become less practical as it ranges from negative to positive infinity, when real values often do not. The truncated normal distribution avoids extreme values and hence is more suitable for speeds on arterials where vehicles are likely to be grouped together in platoons. The probability density function and the cumulative distribution function are as following:

$$f(x; \bar{\mu}, \bar{\sigma}, a, b) = \begin{cases} 0 & x \le a \\ \dfrac{\phi(x; \bar{\mu}, \bar{\sigma}^2)}{\Phi(b; \bar{\mu}, \bar{\sigma}^2) - \Phi(a; \bar{\mu}, \bar{\sigma}^2)} & a < x < b \\ 0 & x \ge b \end{cases} \quad (7)$$

$$F(x; \bar{\mu}, \bar{\sigma}, a, b) = \int_{-\infty}^{x} f(t; \bar{\mu}, \bar{\sigma}, a, b) dt = \frac{\Phi(x; \bar{\mu}, \bar{\sigma}^2) - \Phi(a; \bar{\mu}, \bar{\sigma}^2)}{\Phi(b; \bar{\mu}, \bar{\sigma}^2) - \Phi(x; \bar{\mu}, \bar{\sigma}^2)} \quad (8)$$

where ϕ and Φ are the pdf and cdf of the standard normal distribution. The truncated normal distribution has four parameters, the mean μ, the standard deviation σ, the lower bound value a and the upper bound value b. The parameter mean and variance for a two-sided truncation can be calculated as following, if letting $\alpha = \frac{a-\mu}{\sigma}$ and $\beta = \frac{b-\mu}{\sigma}$:

$$\mu = \bar{\mu} - \sigma \cdot \frac{\phi(\beta;0,1) - \phi(\alpha;0,1)}{\Phi(\beta;0,1) - \Phi(\alpha;0,1)} \quad (9)$$

$$\sigma = \bar{\sigma} \cdot (\sqrt{1 - \frac{\beta\phi(\beta;0,1) - \alpha\phi(\alpha;0,1)}{\Phi(\beta;0,1) - \Phi(\alpha;0,1)} - (\frac{\phi(\beta;0,1) - \phi(\alpha;0,1)}{\Phi(\beta;0,1) - \Phi(\alpha;0,1)})^2}) \quad (10)$$

3.3 Multivariate Model for Speed Distribution Parameters

The goal of this part of the study is to select the independent variables that contribute most significantly to the dependent variables, which in this case, are the parameters of the truncated normal distribution. The final model should account both accuracy and simplicity. Various methods could be applied in the model selection process. The simplest methods could be backward elimination and forward selection which is just a reversed backward selection method. A better method is combining the backward elimination and the forward selection to become the stepwise regression method. In the stepwise regression, various combinations of the independent variables will be tested, and one variable is removed or added to the model at each stage in the stepwise regression method. The best model can be identified with the largest R^2 value. Another type of model selection method is the criterion-based procedures, in which the best model is chosen based on certain criterion. The most common ones are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The BIC substitute the term “the number of parameters multiplied by 2” of the AIC with the number of parameters multiplied by the natural logarithm of the sample size. Both the AIC and the BIC aim to balance the level of fit with the model size, while BIC penalizes larger models more than the AIC. In the study, besides the stepwise regression methods, AIC is also adopted in selecting the best model for speed distribution parameters.

3.4 Platoon Dispersion Model

This section provides the platoon dispersion model development process based on a truncated normal distribution of vehicle speeds. Firstly, the probability density function for vehicle speeds could be re-written as:

$$f(v) = \begin{cases} c\dfrac{1}{\sqrt{2\pi}\sigma} e^{-0.5}(\dfrac{v-\mu}{\sigma})^2, & v_r \le v \le v_f \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

Where v_r and v_f are the minimum and maximum speeds of the platoon. Its cumulative distribution function is:

$$F(v) = \begin{cases} \int_{-\infty}^{v_r} f(v)dv = 0 & v < v_r \\ \int_{-\infty}^{v} f(v)dv = \int_{v_r}^{v} f(v)dv & v_r \le v \le v_f \\ \int_{v_r}^{v_f} f(v)dv = c[F(v) - F(v_r)] & v > v_f \end{cases} \quad (12)$$

For $v_r \le v \le v_f$, the equation can be further expanded as:

$$\int_{v_r}^{v} f(v)dv = c[\int_{-\infty}^{v} \frac{1}{\sqrt{2\pi}\sigma} e^{-0.5(\frac{v-\mu}{\sigma})^2} dv - \int_{-\infty}^{v_r} \frac{1}{\sqrt{2\pi}\sigma} e^{-0.5(\frac{v-\mu}{\sigma})^2} dv = c[F(v) - F(v_r)] \quad (13)$$

c is a constant that makes the probability density function integrates to 1 in its defined range.

Define the density at location x, time t to be k(x, t). Then at the time 0 where all vehicles are queued at the upstream intersection with queue length q the density is:

$$k(x,0)=\begin{cases}0 & x\geq 0\\ k_j & -q\leq x\leq 0\\ 0 & x\leq q\end{cases} \qquad (14)$$

The above equations assume that the initial head of queue position is at location "0" and the back of queue is at location "-q", Using a piecewise function to categorize the density calculation into four conditions, we have the density after some time t:

$$k(x,t)=\begin{cases}0 & x<v_r t-q, or \quad x>v_f t\\ k_j\int_{v_r}^{\frac{x}{t}+\frac{q}{t}} f(v)dv & v_r t-q\leq x<v_r t\\ k_j\int_{\frac{x}{t}}^{\frac{x}{t}+\frac{q}{t}} f(v)dv & v_r t\leq x<v_f t-q\\ k_j\int_{\frac{x}{t}}^{v_f} f(v)dv & v_f t-q\leq x<v_f t\end{cases} \qquad (15)$$

To avoid over-complication in the equations and calculations after substituting the probability density functions in the piecewise equations, we define $\frac{v-\mu}{\sigma}$ to be u and $\frac{\sigma}{\mu}$ to be α. Then for the probability density function of the truncated normal distribution with any lower and upper bound a and b, we have:

$$\int_a^b f(v)dv = c\int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}}e^{-0.5u^2}du = c\int_{\frac{at/\mu-t}{\alpha t}}^{\frac{bt/\mu-t}{\alpha t}} \frac{1}{\sqrt{2\pi}}e^{-0.5u^2}du = \frac{c}{2}[F(y)]_{y_1}^{y_2} \qquad (16)$$

where y_1 could be expressed as $\frac{at}{\sqrt{2}(\mu-t)\alpha t}$ and y_2 could be expressed as $\frac{bt}{\sqrt{2}(\mu-t)\alpha t}$, and $F[(y)]$ is the standard normal distribution as it is equal to $2\int_0^{\sqrt{2}y}\frac{1}{\sqrt{2\pi}}e^{-0.5u^2}du$. Substituting the simplified equation of $\int_a^b f(v)dv$ into the piecewise function above will give us:

$$
k(x,t)=\begin{cases} 0 & x<v_r t-q, or \quad x>v_f t \\ \frac{ck_j}{2}F[(y)]_{\frac{v_r t/\mu-t}{\sqrt{2\alpha t}}}^{\frac{x/\mu+q/\mu-t}{\sqrt{2\alpha t}}} & v_r t-q\le x<v_r t \\ \frac{ck_j}{2}F[(y)]_{\frac{x/\mu-t}{\sqrt{2\alpha t}}}^{\frac{x/\mu+q/\mu-t}{\sqrt{2\alpha t}}} & v_r t\le x<v_f t-q \\ \frac{ck_j}{2}F[(y)]_{\frac{x/\mu-t}{\sqrt{2\alpha t}}}^{\frac{v_f t/\mu-t}{\sqrt{2\alpha t}}} & v_f t-q\le x<v_f t \end{cases} \quad (17)
$$

The next is to calculated the number of vehicles at the front of the platoon that have passed the location x at time t A(x,t) and the number of vehicles at the rear of the platoon that have passed location x at time t B(x,t). The number of vehicles could be calculated by integrating the density over a specified distance under the following scenarios.

i) If $x>v_f t,\begin{cases} A(x,t)=0 \\ B(x,t)=\int_{v_r t-q}^{v_r t}k(x,t)dx+\int_{v_r t}^{v_f t-q}k(x,t)dx+\int_{v_f t-q}^{v_f t}k(x,t)dx \end{cases}$ (18)

ii) If $v_f t-q<x\le v_f t,\begin{cases} A(x,t)=\int_{x}^{v_f t}k(x,t)dx \\ B(x,t)=\int_{v_r t-q}^{v_r t}k(x,t)dx+\int_{v_r t}^{v_f t-q}k(x,t)dx+\int_{v_f t-q}^{x}k(x,t)dx \end{cases}$ (19)

iii) If $v_r t\le x\le v_f t-q,\begin{cases} A(x,t)=\int_{x}^{v_f t-q}k(x,t)dx+\int_{v_f t-q}^{v_f t}k(x,t)dx \\ B(x,t)=\int_{v_r t-q}^{v_r t}k(x,t)dx+\int_{v_r t}^{x}k(x,t)dx \end{cases}$

iv) If $v_r t-q\le x\le v_r t,\begin{cases} A(x,t)=\int_{x}^{v_r t}k(x,t)dx+\int_{v_r t}^{v_f t-q}k(x,t)dx+\int_{v_f t-q}^{v_f t}k(x,t)dx \\ B(x,t)=\int_{v_r t-q}^{x}k(x,t)dx \end{cases}$ (20)

v) If $x\le v_r t-q,\begin{cases} A(x,t)=\int_{v_r t-q}^{v_r t}k(x,t)dx+\int_{v_r t}^{v_f t-q}k(x,t)dx+\int_{v_f t-q}^{v_f t}k(x,t)dx \\ B(x,t)=0 \end{cases}$ (21)

For a segment x_1, x_2, the density could be calculated as:

$$\int_{x_1}^{x_2} k(x,t)dx = \frac{ck_j}{2}\int_{x_1}^{x_2}[F(y_2(x)) - F(y_1(x))]dx = \frac{\sqrt{2}ck_j\mu\alpha t}{2}[\int_{y_2(x_1)}^{y_2(x_2)} F(x)dx - \int_{y_1(x_1)}^{y_1(x_2)} F(x)dx]$$

(22)

To further simplify the above equation, we let $T(z) = \int F(z)dz = zF(z) + \frac{1}{\sqrt{\pi}}e^{-z^2}$,

then the above equation becomes:

$$\int_{x_1}^{x_2} k(x,t)dx = \frac{\sqrt{2}ck_j\mu\alpha t}{2}([G(z)]_{y_2(x_1)}^{y_2(x_2)} - [G(z)]_{y_1(x_1)}^{y_1(x_2)}) \quad (23)$$

Substituting the above equation into equations (18)-(22) will yield the vehicle number calculation result. Note that the parameters of the truncated normal distribution, the mean μ , the standard deviation σ , the maximum and minimum v_f and v_r should be calculated based on the multivariate model for vehicle speed distribution.

If the functions for calculating the mean, the standard deviation, the maximum and the minimum are $L_{mean}(w), L_{std}(w), L_{\max}(w), L_{\min}(w)$, then we have the following substitutions for equation (18)-(23):

$$v_f = L_{\max}(w) \quad (24)$$

$$v_r = L_{\min}(w) \quad (25)$$

$$\mu = L_{mean}(w) \quad (26)$$

$$a = \frac{L_{std}(w)}{L_{mean}(w)} \quad (27)$$

4. ANALYSIS ON MULTIVARIATE DISTRIBUTION PARAMETER MODEL AND PLATOON DISPERSION MODEL

This section provides analysis results on the platoon speeds. The first subsection introduces the distribution fitting process. The second subsection provides the decision process of selecting the most suitable model for the speed distribution parameters. The third subsection used real and simulated data to validate the proposed model.

To account for different levels for each of the possible impacting factor, the data were collected under various conditions. The factor "time" is the time the platoon have already traveled along the link after the head of the platoon has entered the current link. It has 3-4 levels, according to the length of the link. The factor volume has three levels, ranging from 1000-1800 vph. This is to ensure that the volume on the arterial do not reach over-saturated conditions, under which the effect of signal coordination is greatly decreased. The factor "link length" has three levels ranging from 1000 feet to 1800 feet, and is based on an approximation of real link lengths on Texas Avenue. In order to test more factors that may have an impact on the speed distribution, some conditions were simulated on a virtual network in VISSIM. This includes the truck percentage, the number of lanes, and the posted speed limit on the link. A table describing the levels of the factors is presented below. For the real data collected with factors "time", "volume" and "link length", theoretically there are total 5*3*3=45 scenarios accounting for volume levels of low, medium and high volume, link length levels of short, medium, and long link, and time intervals starting from 4 seconds to 20 seconds with 4 seconds increase. However, vehicles will travel less than 20 seconds to pass the short links. Hence there are actually 39 scenarios in total.

Table 1. List of Evaluating Scenarios

Possible Impacting Factors	Levels to be Examined
Time	4sec,8sec,12sec,16sec,20sec
Volume	1200vph,1400vph,1800vph
Link Length	1000ft, 1250ft,1800ft
Truck Percentage	5%-50%, on a 5% increase
Number of Lanes	1,2,3,4 lanes
Posted Speed Limit	35mph,40mph,45mph,50mph,55mph

4.1 Speed Distribution Fitting

4.1.1 Speed Distribution in Homogenous Traffic Flow

To understand platoon characteristics and develop platoon dispersion model, a suitable distribution needs to be selected and set up for the vehicle speeds. Some common distributions are tested in the statistical software R to examine the goodness of fit for the vehicle speeds. These distributions include Weibull, Gamma, lognormal and truncated normal distributions. The reason these distributions are selected is that they are the most commonly used distribution and most of the speed distributions fall into these categories. The distributions were fit for each of the 39 scenarios mentioned above. The fitting results show that the truncated normal distribution performs better than other distributions in all scenarios. The K-S test statistics and both AIC and BIC criterion displays better results for truncated normal distribution. Some examples are presented below.

Table 2. Goodness of Fit Test Results for Short Link, Low Volume at 20 Seconds.

	Weibull	Gamma	Lognorm	TrunNorm
K-S test	0.0462423	0.0306544	0.0332797	**0.0276881**
AIC	3922.693	3927.625	3932.748	**3918.764**
BIC	3932.509	3937.441	3942.564	**3928.579**

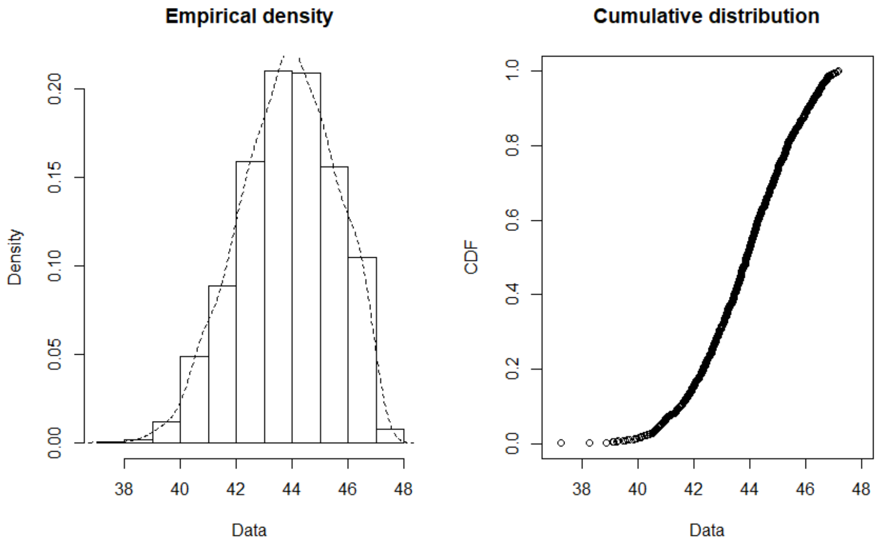


Figure 3. Histogram and Cumulative Distribution for Short Link, Low volume at 20 Seconds

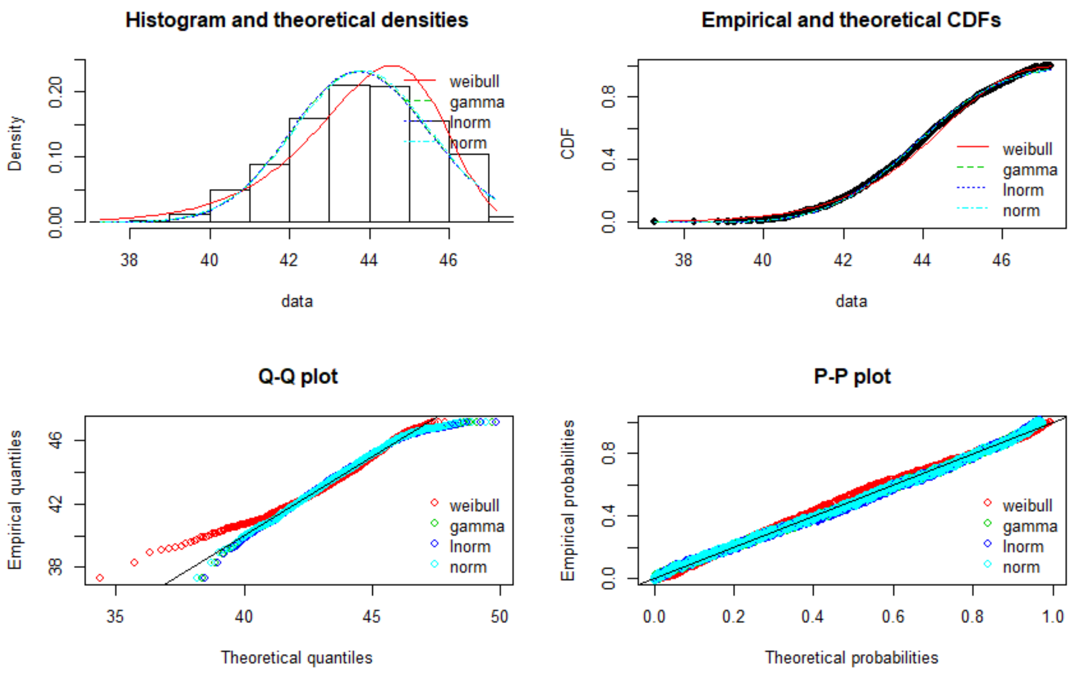


Figure 4. Diagnostic Plots for Short Link, Low Volume at 20 Seconds

Table 3. Speed Distribution Parameters for Short Link, Low Volume at 20 Seconds

	estimate	error
mean	44.02	0.05418
sd	1.733	0.03831
min	37.3	0.03768
max	47.2	0.04956

Table 4. Goodness of Fit Test Results for Medium Link, Medium Volume at 8 Seconds.

	Weibull	Gamma	Lognorm	TrunNorm
K-S test	0.0599034	0.0294734	0.0320181	**0.0245284**
AIC	4048.366	3992.491	3994.373	**3981.401**
BIC	4058.181	4004.306	4008.188	**3994.217**

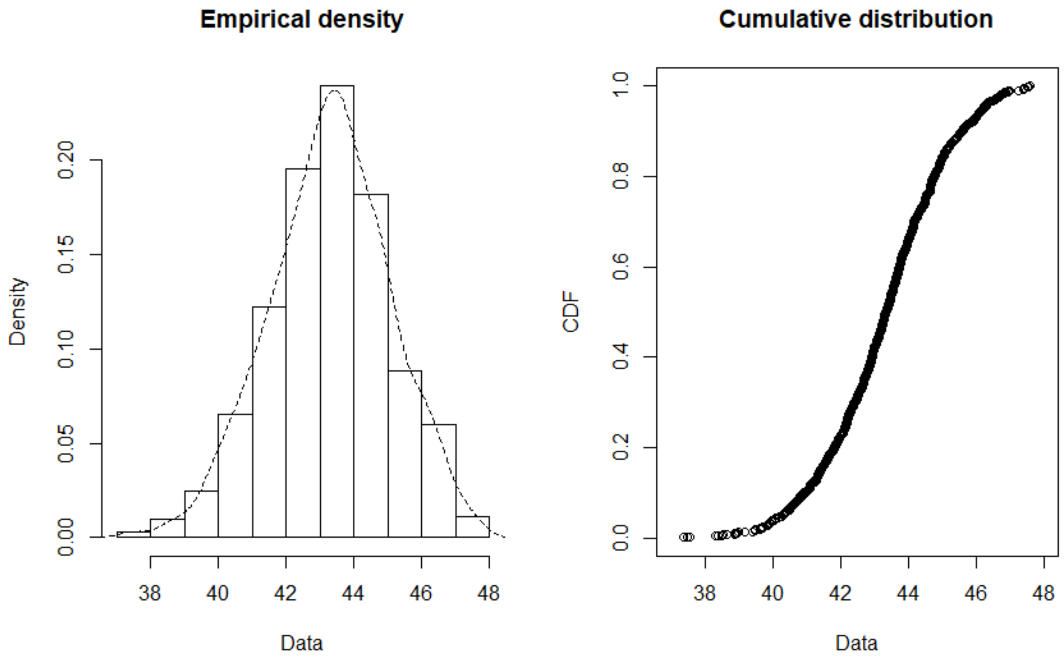


Figure 5. Histogram and Cumulative Distribution for Medium Link, Medium Volume at 8 Seconds.

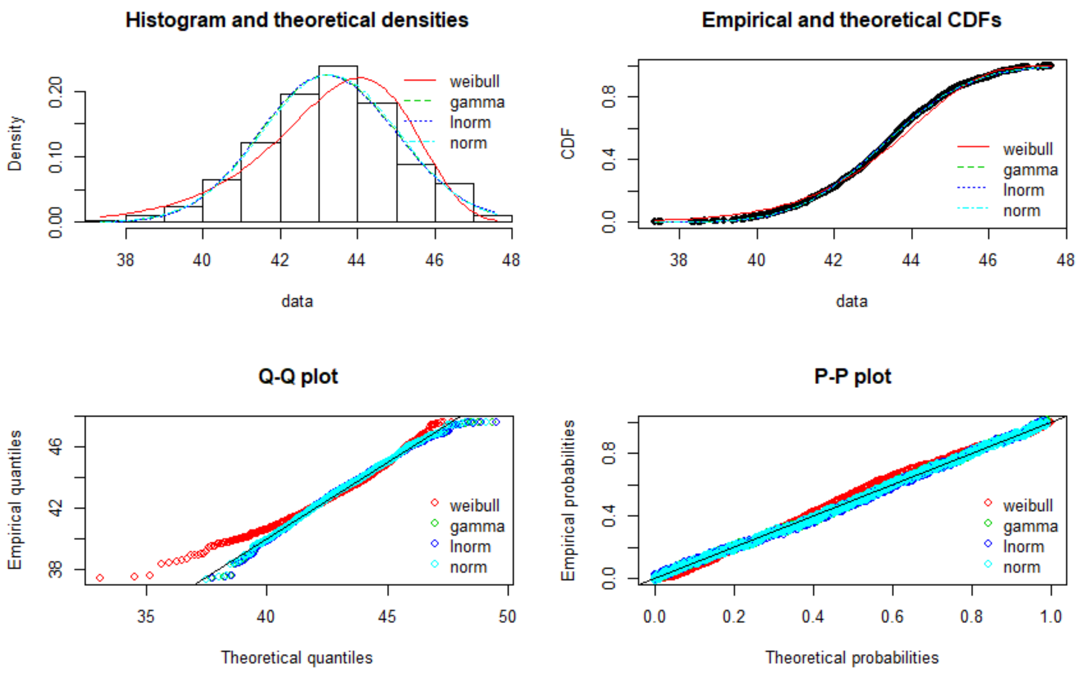


Figure 6. Diagnostic Plots for Medium Link, Medium Volume at 8 Seconds.

Table 5. Distribution Parameters for Medium Link, Medium Volume at 8 Seconds.

	estimate	error
mean	43.3	0.055988
sd	1.713	0.039589
min	37.4	0.042794
max	47.9	0.039921

Table 6. Goodness of Fit Test Results for Long Link, High Volume at 8 Seconds.

	Weibull	Gamma	Lognorm	TrunNorm
K-S test	0.0377407	0.0389533	0.0413466	**0.0339988**
AIC	3950.711	3954.753	3960.08	**3945.584**
BIC	3960.526	3964.568	3969.896	**3953.5**

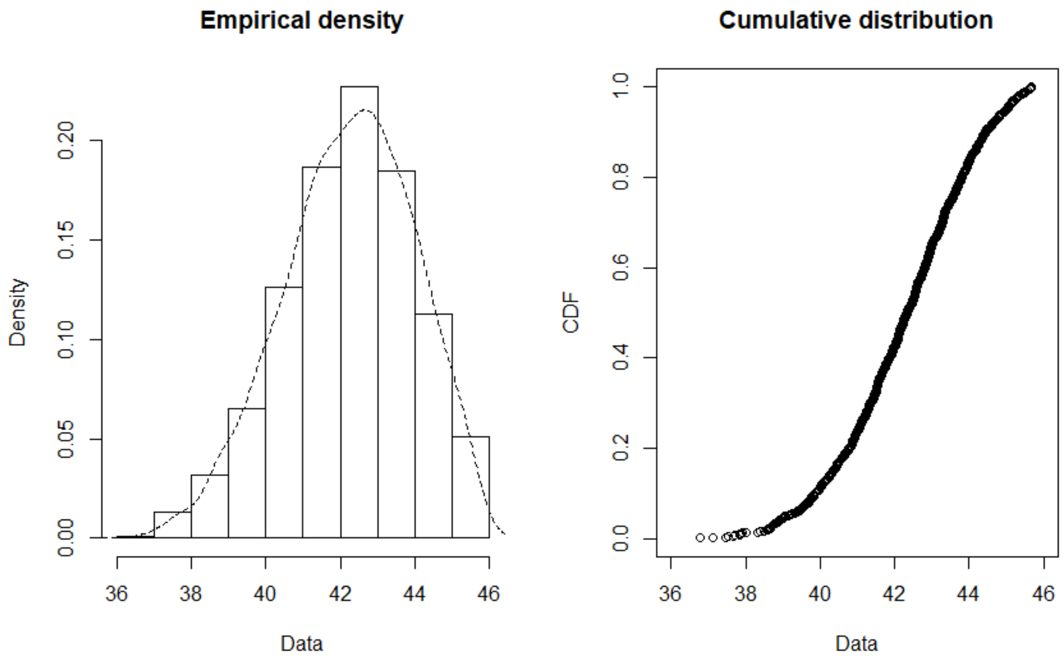


Figure 7. Histogram and Cumulative Distribution for Long Link, High Volume at 8 Seconds.

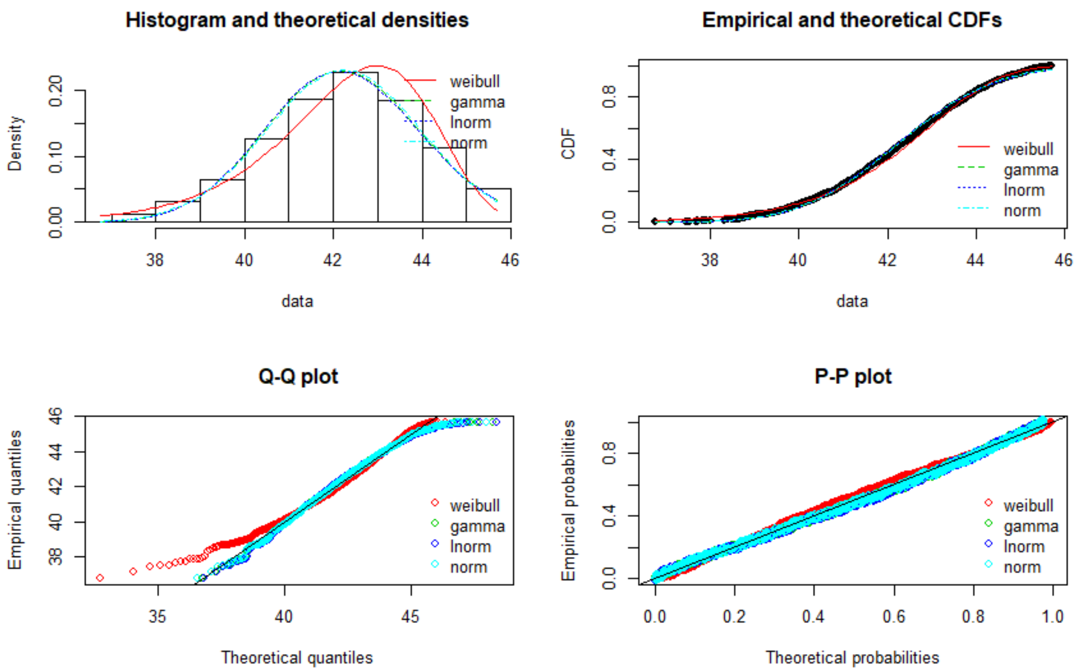


Figure 8. Diagnostic Plots for Long Link, High Volume at 8 Seconds.

Table 7. Distribution Parameters for Long Link, High Volume at 8 Seconds.

	estimate	error
mean	42.4	0.05491
sd	1.875	0.03883
min	40.1	0.04548
max	45.7	0.05977

Table 8. Summary of Distribution Parameters (4-8 seconds)

time=4sec				time=8sec			
low volume	short	medium	long	low volume	short	medium	long
mean	40.900	43.900	45.200	mean	43.400	45.500	46.700
sd	1.533	1.746	1.942	sd	1.615	1.861	2.058
min	36.900	37.400	37.200	min	38.300	38.100	37.700
max	44.200	47.800	49.500	max	46.700	49.300	51.000
medium volume	short	medium	long	medium volume	short	medium	long
mean	39.200	41.800	43.200	mean	41.000	43.300	44.800
sd	1.491	1.625	1.864	sd	1.582	1.713	1.959
min	35.300	38.400	39.700	min	38.400	38.800	39.200
max	42.100	44.500	46.600	max	45.600	47.900	48.200
high volume	short	medium	long	high volume	short	medium	long
mean	37.800	39.400	40.900	mean	39.200	41.000	42.400
sd	1.283	1.529	1.770	sd	1.420	1.644	1.825
min	35.200	36.400	36.200	min	37.100	38.200	39.300
max	40.400	42.300	44.200	max	42.000	44.800	45.700

Table 9. Summary of Distribution Parameters Continued (12-16 seconds)

time=12sec					time=16sec			
low volume	short	medium	long		low volume	short	medium	long
mean	44.000	46.900	48.100		mean	45.500	48.500	49.700
sd	1.733	1.982	2.163		sd	1.851	2.115	2.278
min	36.900	37.500	37.000		min	38.200	38.400	38.900
max	47.200	50.800	52.500		max	48.700	52.300	54.000
medium volume	short	medium	long		medium volume	short	medium	long
mean	42.400	44.800	46.300		mean	43.700	46.300	47.800
sd	1.667	1.846	2.074		sd	1.842	1.961	2.187
min	37.200	38.600	38.500		min	39.800	38.600	40.100
max	46.900	49.300	49.900		max	48.100	51.000	52.300
high volume	short	medium	long		high volume	short	medium	long
mean	40.600	42.300	43.900		mean	42.000	43.900	45.500
sd	1.533	1.758	1.990		sd	1.645	1.863	2.083
min	37.500	38.200	40.100		min	38.900	38.200	40.100
max	43.500	46.100	47.300		max	44.900	47.500	49.900

From the above summary tables, a clear increasing trend of the values of mean, standard deviation and maximum of the vehicle speeds can be observed. The values increase as the time increases, or as the link length increases. On the other hand, these values gradually decrease as the volume increases. The minimum value of the speeds however, does not have a clear pattern. It is likely that they do not change with the change of different impacting factors. Intuitively the increase of the mean, standard deviation and maximum indicates a larger dispersion as time and link length increases, or volume decreases. The range of the speeds (maximum-minimum value) also increases as time or link length increases, or as volume decreases. This may also be treated as an indicator of greater platoon dispersion. The quantitative analysis of the relationship of the parameters and the impacting factors are presented in the next section.

4.1.2 Speed Distribution in Heterogeneous Traffic Flow

This section presents analysis on the speed distribution of heterogeneous traffic flow with trucks involved. Since it is difficult to collect actual data, the truck speeds were simulated from the virtual networks of Texas Avenue. Truck-involved speeds displays a very clear bi-modal shape; hence a simple single distribution may not be suitable to describe the speed distribution characteristics. The Gaussian mixture model is adopted to accommodate the heterogeneity of the traffic flow.

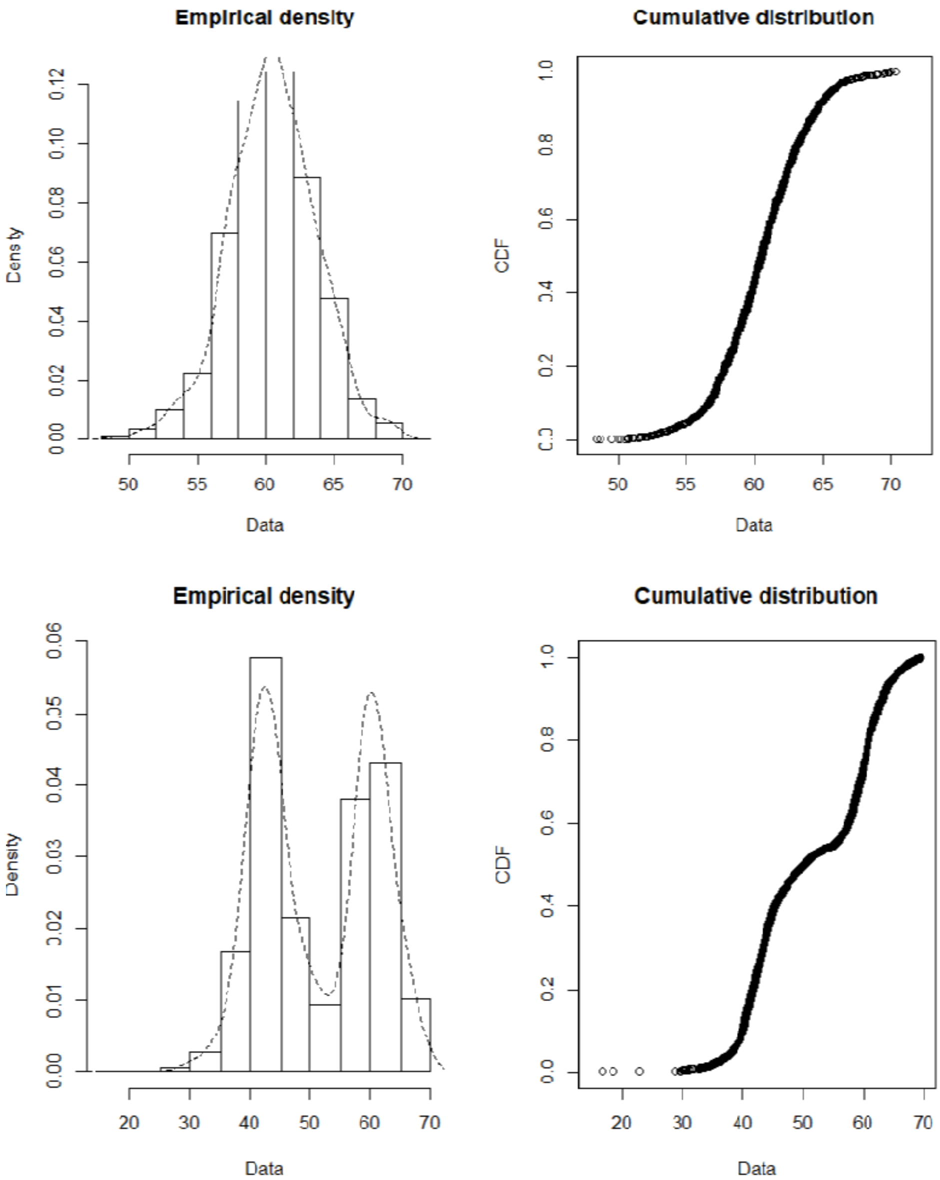


Figure 9. **Non-truck Involved Speed and Truck Involved Speed**

The parameters of the Gaussian mixture distribution are estimated using the EM algorithm. The Gaussian mixture model has two components, and each component has three parameters. The parameters are the weights of the first and second components λ_1 and λ_2, means of the first and second component μ_1 and μ_2, and variances of the first and second component σ_1 and σ_2. An example of the results of the parameter estimation is presented in the following tables, showing the parameters estimated from EM algorithm of different truck percentage categories at 20 seconds and at 15% of trucks, with speeds shown in ft/sec.

Table 10. **Summary of Gaussian Mixture Distribution Parameters at 20 Seconds.**

	Truck Percentages							
Parameters	5%	10%	15%	20%	25%	30%	35%	40%
λ_1	0.2008111	0.371839	0.53013	0.636909	0.73595	0.742507	0.8034252	0.829306
μ_1	44.05958	43.90233	43.00822	42.72704	42.38413	42.22718	42.073410	41.78459
σ_1	4.2671	3.988357	3.775228	3.667397	3.461127	3.36445	2.806854	2.749165
λ_2	0.799189	0.628161	0.46987	0.363091	0.26405	0.257494	0.1965748	0.170695
μ_2	63.22956	62.99763	62.80437	60.81698	60.73558	60.54013	60.14764	59.69984
σ_2	3.062812	3.216201	3.416746	3.440577	3.627539	3.745967	4.318205	5.376785

Table 11. Summary of Gaussian Mixture Distribution Parameters at 15% Trucks

Parameters	4 sec	8 sec	12 sec	16 sec	20 sec
λ_1	0.490849	0.531081	0.519861	0.53013	0.491746
μ_1	45.4821	44.63971	43.57135	43.00822	42.67205
σ_1	3.168855	3.325193	3.616188	3.775228	3.969188
λ_2	0.509151	0.46892	0.480139	0.46987	0.508254
μ_2	64.95777	64.33306	63.76431	62.80437	62.67844
σ_2	3.285173	3.398412	3.41605	3.416746	3.716746

Table 10 shows the relationship between Gaussian mixture distribution parameters and the truck percentages intuitively. The weight of the first component increases while the weight of the second component decreases. The increase in the amount of trucks in the traffic flow makes the platoon speeds decrease for both components. The variance of the first component also decreases while the variance of the second component increases as the truck percentage increases. The reason for the increase of λ_1 and decrease of λ_2 is simple, since more trucks are joining the traffic flow, making the weight of the first component increase and the second component decrease as there are less cars. The reason for decreasing variance of the first component may be that higher truck percentages confine the vehicles' speed selection range for the first component. The slower vehicles are more closely packed into a platoon as truck percentage increases.

Table 11 shows the relationship between Gaussian mixture distribution parameters and travel time when truck percentage is 15%. It can be seen that the means the two components decreases over time, while the variances increases over time. The increase of the variances conforms the previous experience that platoon diffuses over time. It should be noted that there seems no relationship between the weight λ and the travel time. This is reasonable since the truck percentages did not change. Fast and slow vehicles are unlikely to move from one component to another during this time period. Hence the weights are likely to remain at a certain value, in this 15% truck case, around 50%.

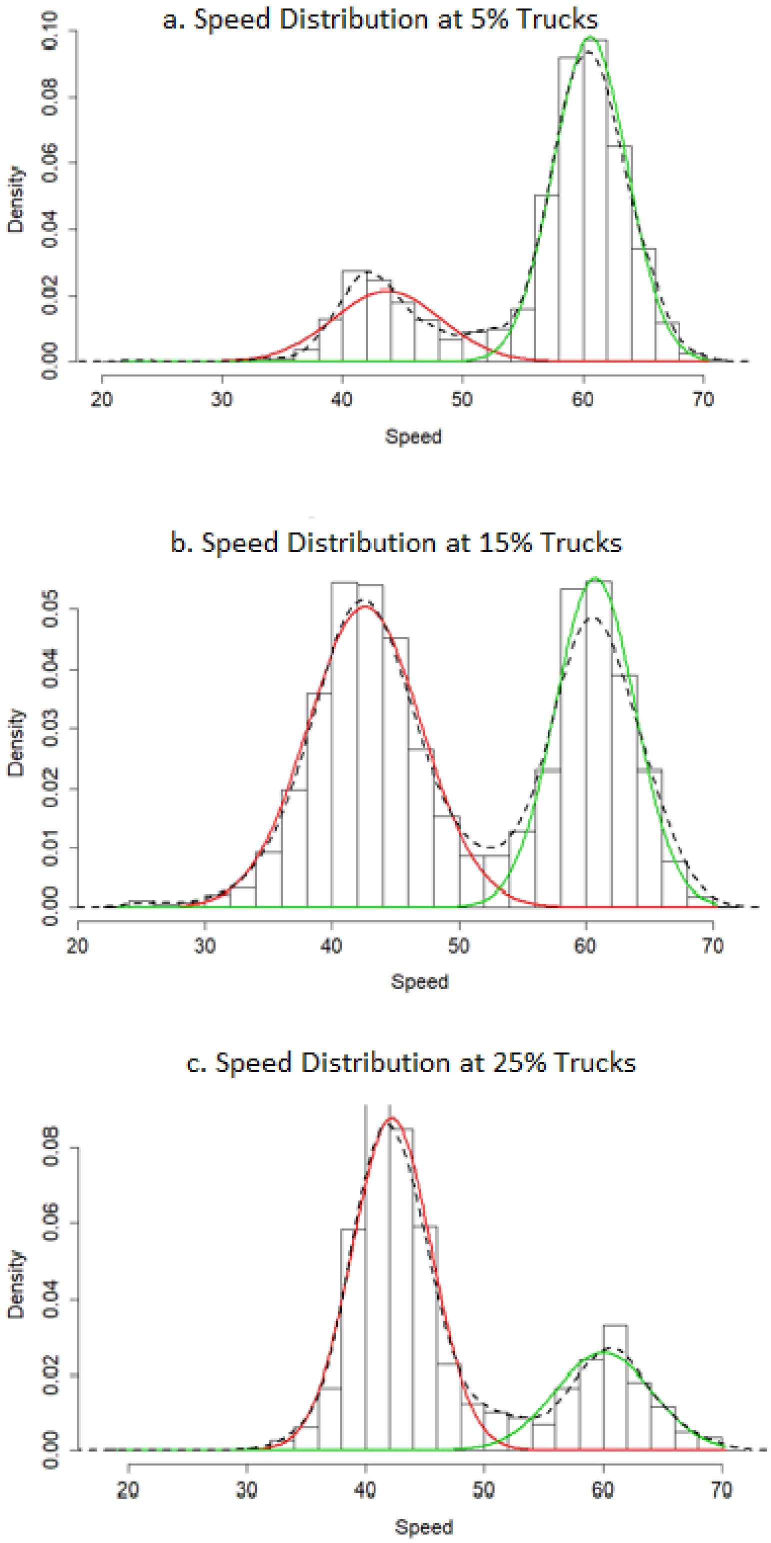


Figure 10. Gaussian Mixture Distribution Fitted Curves under Different Truck Percentages

4.2 Effect of Independent Variables and Model Selection

To start building the relationship for the distribution parameters and possible impacting factors, the individual effect of each of the factor on the parameters was examined first, and a stepwise regression was conducted. Statistical result show that the time, volume, link length have strong cause and effect relationship with the distribution parameters mean speed, standard deviation and maximum speed. The minimum speed on the other hand, is not influenced by these factors, as its p-value of the t-test is large enough to reject the null hypothesis. Another factor that affect the distribution parameters is the truck percentage, which not only change the parameter values but also change the distribution type. Due to the lack of real truck-involved data, the distribution parameter model can be formulated separately based on simulated data and is not included in the following model formulation. The effect of the number of lanes and the posted speed limit were also studied based on simulations. Results showed that when increasing the number of lanes from 1 to 2, the impact on the speed distributions is significant, but this impact diminishes quickly when increasing the number of lanes further from 2 to 4. This can be explained by the fact that overtaking is not allowed when there is only one lane. When there are multiple lanes and vehicles are able to overtake freely, the impact of number of lanes diminishes quickly. Since most arterials have multiple lanes, this factor is excluded from the model formulation. The posted speed limit does affect the mean speed and maximum speed, but does not seem to have a significant effect on the standard deviation, nor the distribution type. The main effect of increasing the posted speed limit is shifting the distribution rightwards and resulting in a change of the intercept. This indicates that the posted speed limit is likely to be one of the influencing factors and should be considered in the model. Since the data regarding the posted speed limit is also collected by simulation, the effect of posted speed limit could be excluded during regression and considered to modify the intercept.

The model was built starting with investigating the relationship between the parameters and each single factor. Results showed that the best fit relationship is a linear relationship, followed by a logarithm relationship. Both these relationships do not reject the null hypothesis. Polynomial and exponential relationship were also examined but has low significance levels to reject the null hypothesis. Hence, the multivariate model building starts with a linear relationship. The following tables show the estimates of each factor and the overall model performance.

Table 12. **Linear Multivariate Model for the Mean Speed**

	Estimate	Std error	P-Value
Intercept	41.72817	1.4396	< 2e-16
time	0.36382	0.06265	4.06E-06
volume	-0.46	0.07291	1.12E-06
link length	0.37176	0.06984	1.44E-05
Adjusted R-squared:	0.8538		
F-statistic p-value	5.91E-10		

Table 13. Linear Multivariate Model for the Standard Deviation of Speeds

	Estimate	Std error	P-Value
Intercept	1.24997	0.057114	< 2e-16
time	0.026786	0.002486	4.37E-11
volume	-0.02595	0.002893	1.93E-09
link length	0.048508	0.002771	6.59E-16
Adjusted R-squared:	0.9556		
F-statistic p-value	< 2.2E-16		

Table 14. Linear Multivariate Model for the Maximum Speed

	Estimate	Std error	P-Value
Intercept	45.34878	1.0504	< 2e-16
time	0.4406	0.04571	4.54E-10
volume	-0.585	0.0532	2.83E-11
link length	0.43821	0.05096	4.45E-09
Adjusted R-squared:	0.921		
F-statistic p-value	4.54E-15		

Though the adjusted R-squared values show that the models are feasible, the linear model could be further modified with adjusting the intercept. The current ones are 41 and 45 mph for mean and maximum, which do not show a strong relationship to any of the road conditions or vehicle characteristics. One way to solve this may be considering the posted speed limit, which indeed have an impact on the distribution parameters, 40mph as the pre-defined intercept. With this fixed intercept, the modified mean and maximum speed model is shown in the following table.

Table 15. **Modified Linear Multivariate Model for the Mean Speed**

Intercept 40	Estimate	Std error	P-Value
time	0.38034	0.06162	1.34E-06
volume	-0.39794	0.05183	2.94E-08
link length	0.37176	0.05755	7.51E-08
Adjusted R-squared:	0.9195		
F-statistic p-value	1.67E-15		

Table 16. **Modified Linear Multivariate Model for the Maximum Speed**

Intercept 40	Estimate	Std error	P-Value
time	0.49175	0.06185	1.51E-08
volume	-0.39291	0.05203	4.01E-08
link length	0.58771	0.05777	9.80E-11
Adjusted R-squared:	0.9716		
F-statistic p-value	< 2.2e-16		

The modified model has improved adjusted R-squared value, and much less F-statistic p-value, showing better fit and more significance of the variables than those not modified. Though the overall model is satisfying, it is still desirable to examine the nonlinear models for the parameters. A table comparing the adjusted R square value and the residual standard error (RSE) are presented below. Four polynomial models with time^2, link length^2, time^3, and volume^3, three exponentials models of time, volume and link length, and three logarithm models of time, volume and link lengths were tested. The logarithm models perform better than others, with the logarithm of link length very close to the linear model. The adjusted R square values are all smaller and the RSE are larger than the values of the linear model, indicating less fit and more error predictions of the nonlinear models. The AIC and BIC criterions also showed that the linear model has the smallest value. Hence, linear multivariate models were selected as the prediction model for the truncated normal distribution parameters to be used in the platoon dispersion model.

Table 17. **Nonlinear Model Comparisons**

nonlinear models	R-square	RSE	AIC	BIC
log transformation 1	0.7341	1.727	113.3	119.54
log transformation 2	0.6652	1.982	115.45	121.48
log transformation 3	0.9192	1.521	109.19	116.19
polynomial regression 1	0.4873	2.502	118.47	126.47
polynomial regression 2	0.4922	2.489	119.02	125.49
polynomial regression 3	0.4068	2.338	121.88	126.59
polynomial regression 4	0.2934	2.829	130.41	135.26
exponential 1	0.3789	2.531	123.59	128.32
exponential 2	0.4125	2.227	119.18	124.79
exponential 3	0.2478	2.909	133.92	136.65
linear	0.9195	1.315	106.72	113.72

It is also desirable to investigate the interactions among the independent variables in the regression model. Interaction effects in regression means that the effect of one independent variable depends on the effect of another independent variable. In this study, the interaction effect between time and volume, time and link length, volume and link length, and combination of all three were studied, though from relevant experience these three factors should not interact with each other The statistical test results are presented below. The p-values for these interaction terms indicate a rejected null-hypothesis and the adjusted R square values show that the overall model performance is bad. Hence, the interaction effect can be excluded from analysis.

Table 18. **Variable Interaction Results**

Interaction Model	p-Value	Adjusted R Square
time*vol	0.2653	0.3871
time*length	0.3206	0.4879
vol*length	0.2878	0.3727
time*vol*length	0.427	0.3036

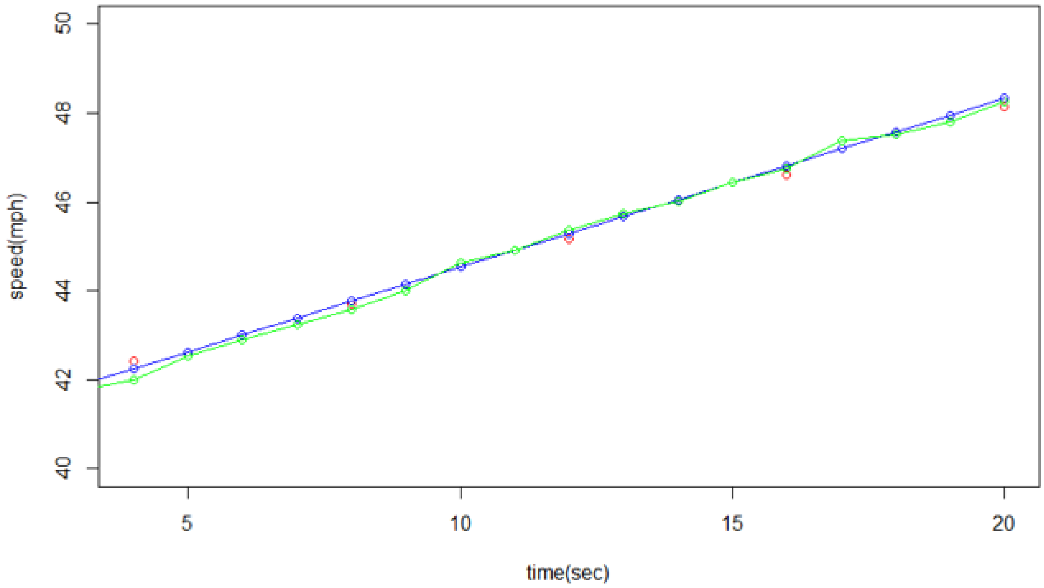
4.3 Model Validation and Platoon Density Distribution

This section is designed to validate the proposed model by examining the performance of the multivariate distribution parameter model and the modified platoon dispersion model. The performance of the multivariate distribution parameter model is evaluated by its prediction ability on speeds, and the performance of the modified platoon dispersion model is evaluated by its ability to estimate downstream densities. The validation begins by comparison with simulation data, and is followed by comparison with real data from Peachtree Street.

4.3.1 Model Validation with Simulation Data

4.3.1.1 Speed Prediction Evaluation

The first part of this section is to validate the prediction ability of the multivariate speed distribution parameter model based on simulation data. The model calculation results are compared with the test group of the real data and the simulated data, on an 850ft short link and a 1450ft long link, under two volume conditions. Some examples are presented below.



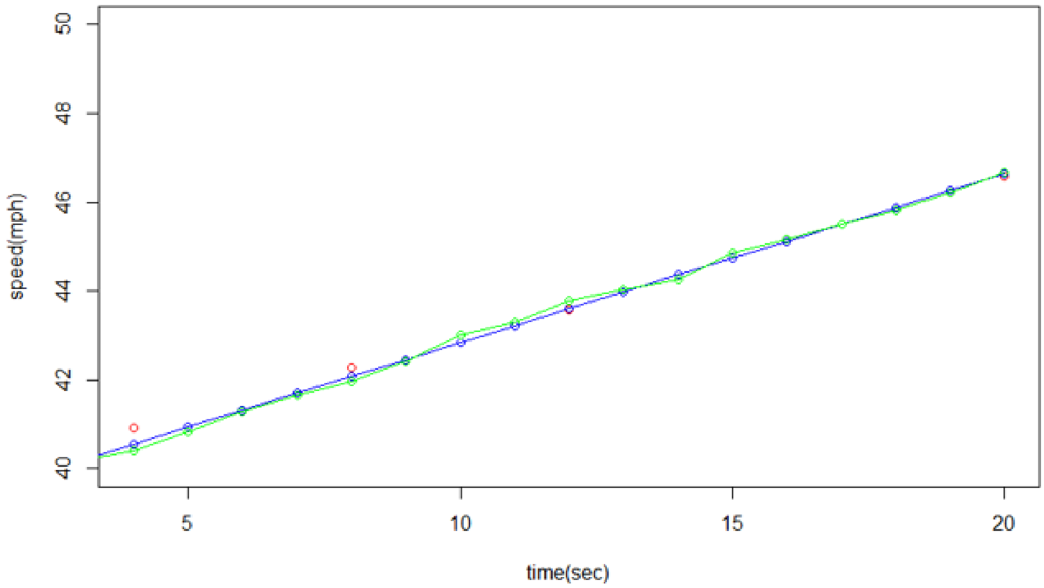


Figure 11. **Comparison between Calculated (blue), Simulated (green) and Real (red) Mean Speed under Low Volume (up) and High Volume (down) on Long Link**

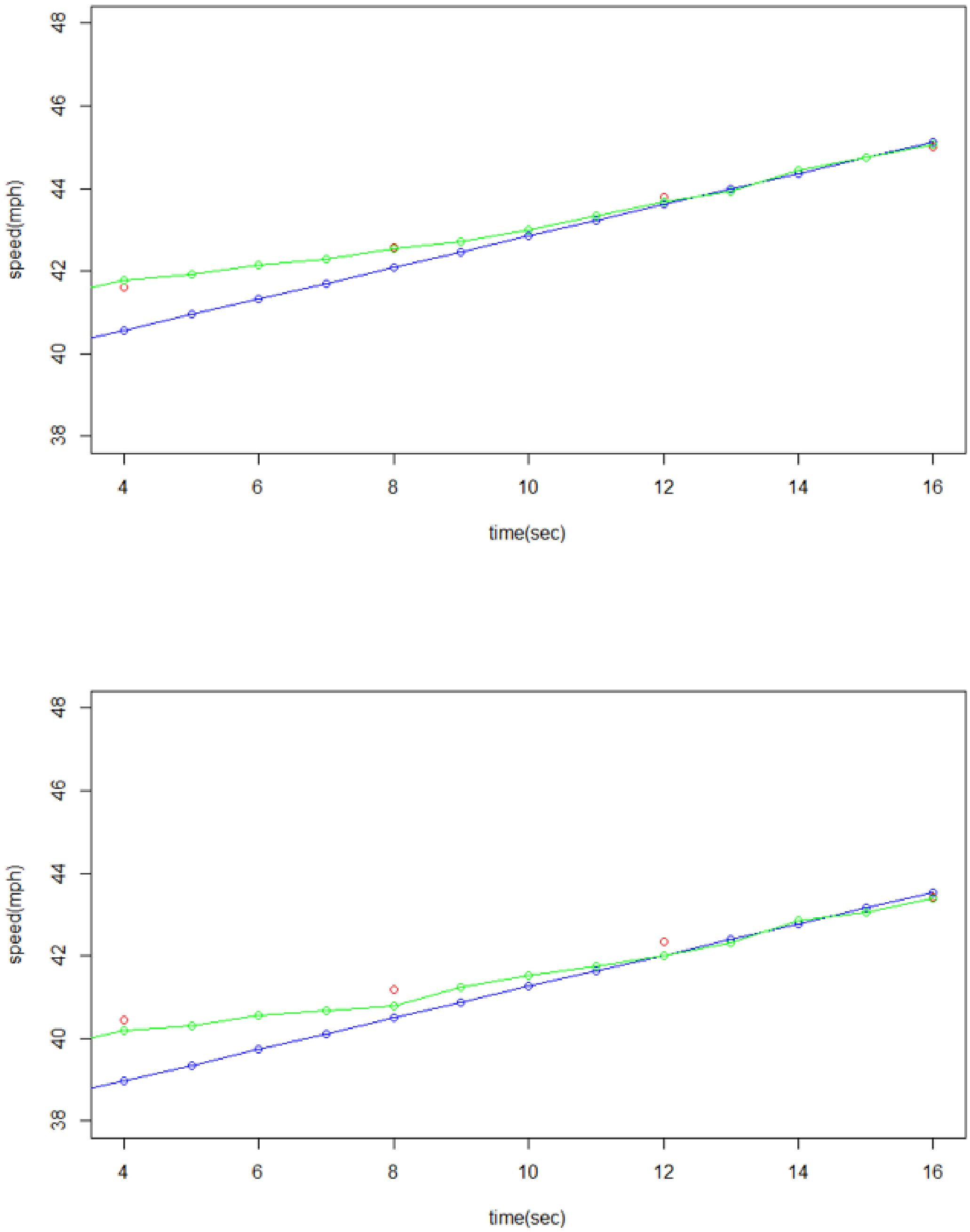


Figure 12. Comparison between Calculated (blue), Simulated (green) and Real (red) Mean Speed under Low Volume (up) and High Volume (down) on Short Link

The results showed that the model performs better when predicting distribution parameters for longer links. For shorter links, the prediction error is large at beginning, but gradually decreases as time increases. Similar patterns can also be observed for standard deviation and maximum speed prediction. This is acceptable because firstly, for most of the time the error is less than 1 mph. In addition, the speed distribution at downstream locations is more important for signal coordination, where the parameter prediction is rather accurate.

4.3.1.2 Density Prediction Evaluation

The model's ability to predict density at downstream is also examined. The performance of the improved platoon dispersion model is evaluated by comparing with Robertson's recursive model, on a virtual arterial constructed in VISSIM. Platoons are generated at an upstream location and the densities at a downstream segment of 100ft are detected. The simulated results are then compared with calculated values from the improved model and the traditional Robertson's model. The parameters of Robertson's model are calibrated to yield the optimal performance. The comparison results are presented second by second in 5 minutes for the simulation data. The results presented in Figure 13 indicates that the proposed modified platoon dispersion model has better performance in predicting downstream densities than Robertson's model, which sometimes has lower and late arrival predictions compared to the simulated or real data results. The proposed modified dispersion model on the other hand, performs better in predicting the maximum density of the segment and models the dispersion effect of the platoon better, especially for the front of the platoon. The RMSEs of the proposed modified model are much smaller than Robertson's model, indicating that the proposed modified model indeed performs better quantitatively.

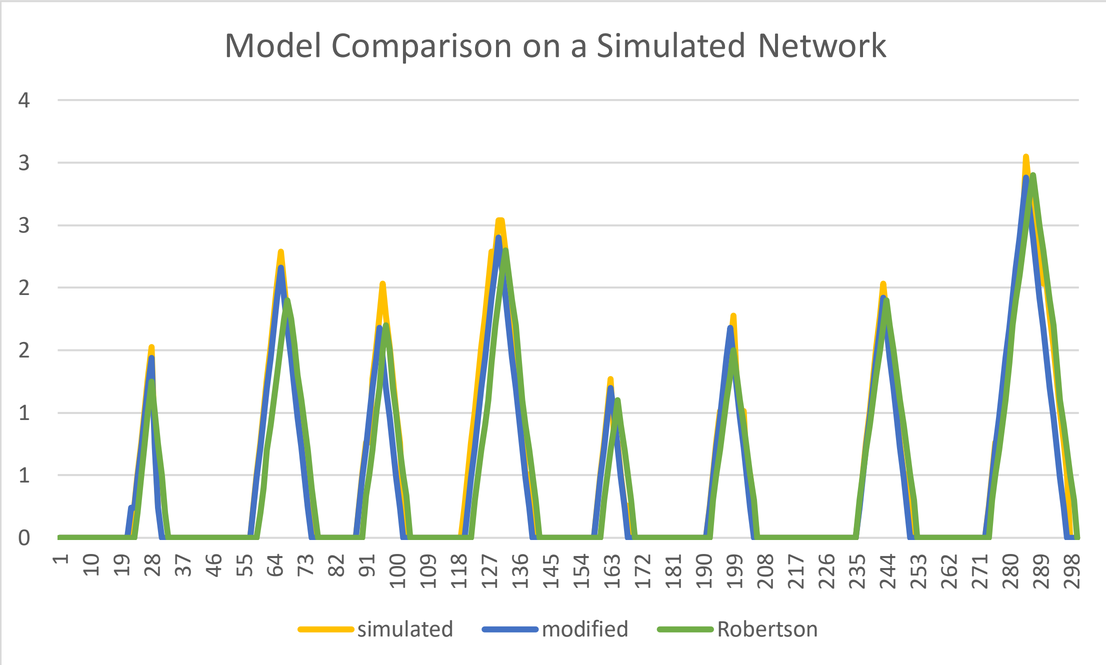


Figure 13. **Comparison of Densities (Simulated vs. Calculated) at A Downstream Segment**

4.3.2 Model Validation with Real Arterial Data

The proposed model is then validated with real data from Peachtree Street. The Peachtree Street Data is from the Next Generation Simulation program initiated by the US Department of Transportation. Peachtree Street is an urban arterial located in Atlanta, Georgia. The data collected are vehicle trajectory data collected by video cameras on buildings in two 15-minute periods. Similar to section 4.3.1, the performance of the multivariate distribution parameter model and the modified platoon dispersion model are evaluated. Speed data was extracted from two segments of Peachtree Street (between 13th and 14th Street, and between 11th and 12th Street).

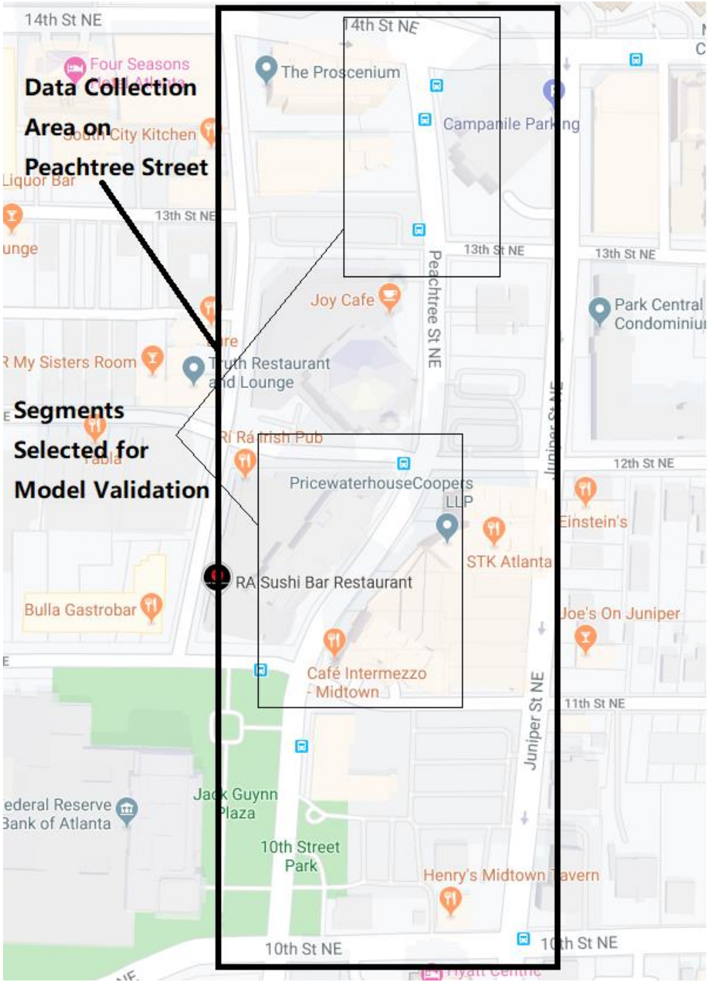
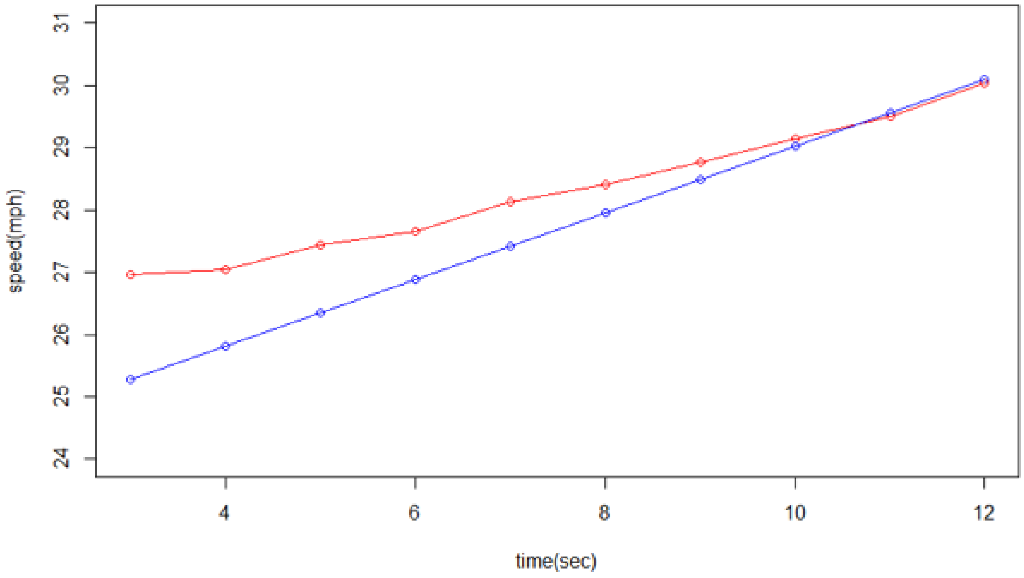


Figure 14. Segments of Peachtree Street

4.3.2.1 Speed Prediction Evaluation

The multivariate distribution parameter model demonstrates satisfactory ability to predict downstream speeds, as shown in Figure 1. Though the gap between the real and predicted value is larger at beginning, it is still in an acceptable range of 1-2 mph. More importantly, the model predicted the downstream speeds rather accurate.



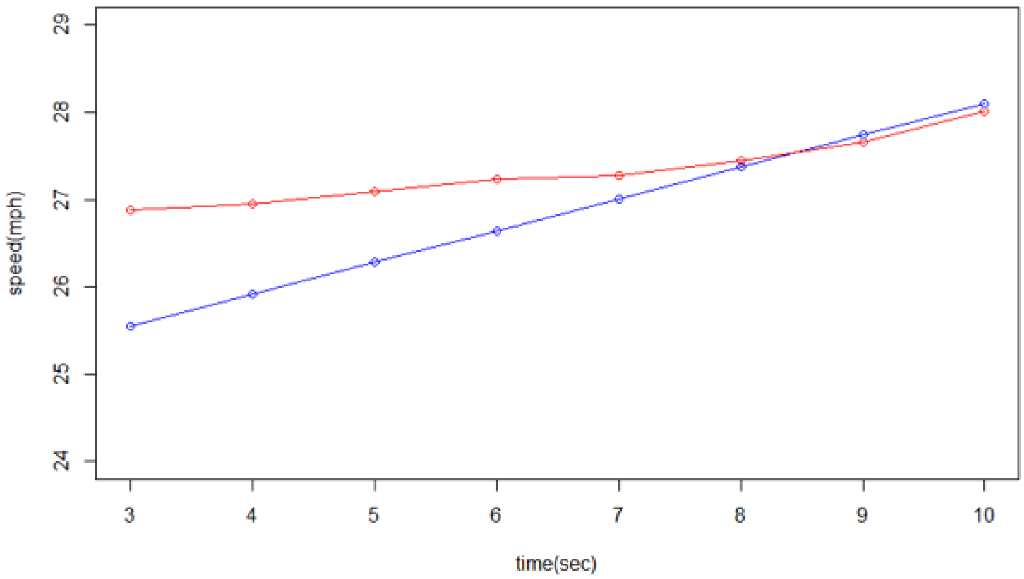


Figure 15. Comparison between Calculated (blue) and Real (red) Mean Speed at Two Sections on Peachtree Street

4.3.2.2 Density Prediction Evaluation

The density prediction of the modified platoon dispersion model is also examined in comparison with Robertson's model. The density at a downstream segment is calculated on a 2-second average for 10 minutes. The parameters of Robertson's model are calibrated again for best performance in density prediction. Similar trend can be observed, as the modified model has better ability in predicting platoon arrival and maximum density. The modified model also has a lower RMSE than Robertson's model.

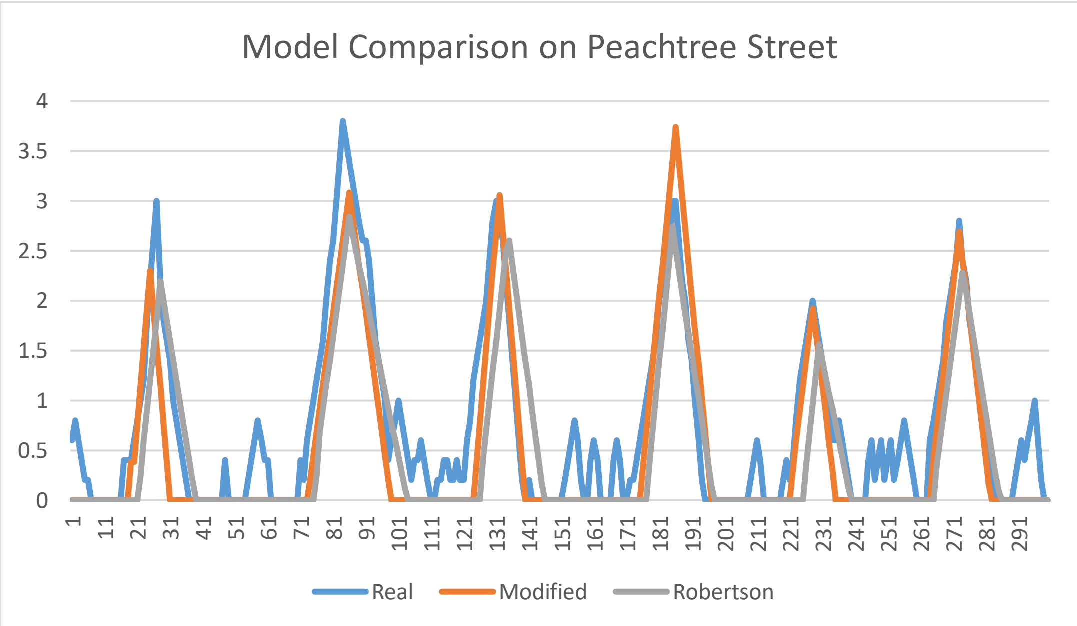


Figure 16. Comparison of Densities (Real vs. Calculated) at A Downstream Segment

The validation results in section 4.3.1 and 4.3.2 prove that the proposed multivariate distribution parameter model has strong ability in predicting downstream speeds and the modified platoon dispersion model is capable of providing a more accurate density prediction than Robertson's model. This result shows that the model has satisfactory performance and is capable of implementation in signal timing strategies.

4.3.3 Platoon Density Distribution

The third part of this section involves the study of the predicted platoon density distribution from the platoon dispersion model based on the multivariate speed distribution parameter model. From Figure 17, it can be observed that the density in the middle is higher than the front and rear of the platoon, since there are more vehicles travelling around the mean speed and the speed distribution determines the density distribution. As time increases, the platoon disperses more along the link, resulting a lower peak value and larger range. The curves are flatter. It can be further observed that there are three stages as vehicles spread between link section $\left(-q+v_r t, v_f t\right)$ at t after they depart at time 0. The three stages are as following:

a) $0 \le t \le \dfrac{\mu t_0}{v_f}$: The front of the platoon has not arrived at the downstream intersection, then the density at the location of the downstream intersection and after the downstream intersection is 0. This corresponds to the red curve in the graph

b) $\dfrac{\mu t_0}{v_f} \le t \le \dfrac{\mu t_0 + q}{vr}$: Some of the vehicles in the platoon have passed the downstream intersection, while the rest are still behind this location. This corresponds to the blue curve in the graph.

c) $\dfrac{\mu t_0 + q}{v_r} \le t$: The rear of the platoon has passed the downstream intersection, and the density at the downstream intersection and the location before it is 0. This corresponds to the green curve in the graph.

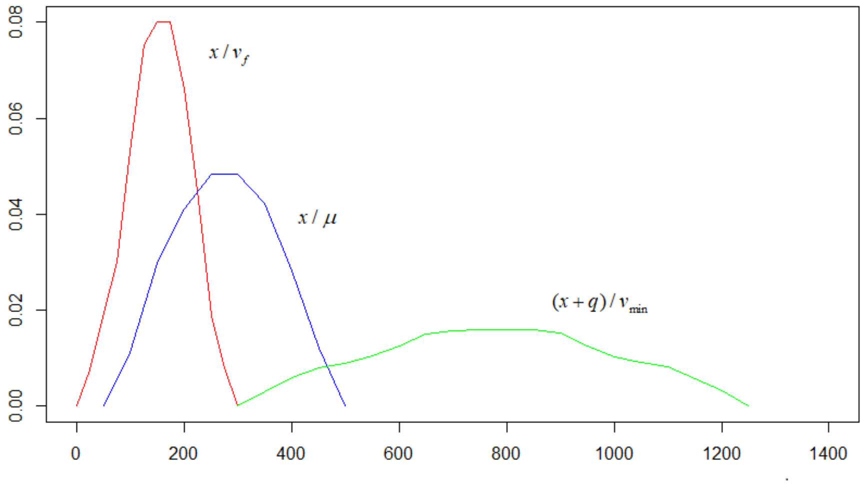


Figure 17. Platoon Density Distributions of Different Times

5. FORMULATION AND ANALYSIS ON PREDICTED PLATOON MAX-PRESSURE POLICY

In this section, arterial signal timing strategies are formulated and evaluated. The first subsection discusses the basic relationship between the offsets and the number of vehicles that are able to progress the downstream intersection. Some very simple experiments on the resulting delay and number of stops by adjusting the offset only is conducted. This experiment shows that just adjusting the offset on the arterial is not enough to accommodate the complicated nature of the arterial traffic. To address the issue, an advanced method based on the original Max-Pressure Policy, the Predicted Platoon Max-Pressure Policy (PPMP), is presented in the following subsection. This subsection introduces traditional Max-Pressure Policy first, and formally presents the PPMP Policy. The next subsection proves the optimality of the PPMP. The final subsection provides the evaluation of various PPMP measures and compares the policy with some traditional signal timing strategies.

5.1 Preliminary Analysis on Platoon Arrival and Offsets

The proposed model could be applied to aid the decision on signal timing and coordination strategy. A preliminary demonstration on the application is presented here by two simple experiments.

The first experiment is conducted on a virtual network constructed based on Texas Avenue. Part of the Texas Avenue consisting of three intersections was selected as preliminary testing network. The signal timing and offsets optimized from traffic software SYNCHRO was regarded as the base signal timing and base offsets for this corridor. The Texas Avenue main street through traffic was set to be about 1200 vph, for this section of the corridor. The turning volume and cross street volume were set close to the real volume of Texas Avenue. If the offset at the downstream intersection was set to be the value of x/μ , then vehicles at the front of the platoon will be stopped at the downstream intersection since the speeds of these vehicles are greater than the mean speed. Hence to reduce the number of stops and the delay caused, the offset should be adjusted earlier to allow vehicles at the front of the platoon pass the downstream intersection. On the other hand, if the offset was adjusted forward while the green time doesn't change, vehicles at the rear of the platoon will not be able to pass the downstream of the intersection. For example, if the downstream green ends at time $\frac{x+q}{\mu}$, then vehicles travelling at speeds lower than the mean speed will not be able to pass the downstream intersection. To consider this group of vehicles that travelling at lower speeds, a late offset is required. In order to find an optimum signal timing, three strategies were compared: the current offset, an early offset to minimize the number of vehicles stopped at the front of the platoon, and a late offset to minimize the number of vehicles stopped at the rear of the platoon. In this case, the number of vehicles stopped at the front is calculated as $A(x_l, t_s - o)$ where x_l is the link length, t_s is the green starting time of the downstream intersection, and o is the adjustment on the offset. The number of vehicles stopped at the rear can be calculated as $B(x_l, t_e + o)$ where t_e is the end time of the downstream green. The optimization results show that in order to minimize the vehicles stopped at the front of the platoon at the two downstream intersections, the offsets should be at least 8 and 6 seconds early. In order to minimize the vehicles stopped at the rear of the platoon, offsets with at least 13 and 10 seconds late is required. These offsets together with other offset values were set in VISSIM and the performance were compared with the original offset. The simulation results are as following.

Table 19. **Simulation Results of Different Offset Settings**

	Early Offset	Current Offset	Late Offset	4-second late offset
Average Delay	92.1	89.9	88.5	85.3
Average Number of Stops	2.01	2.13	1.98	1.92

The results show that both late and early offsets have better performance than the current offset in terms of the number of stops, especially the late ones. This may be caused by the fact that there are not as many vehicles in the front as in the rear, since the data collected has larger left tails. The distribution is truncated more on the right as there are not many high-speed vehicles. It should be noted that the total delay for early or late offset do not outperform the current one significantly. However, from the simulation, a 4-second late offset outperforms the best, in both delay and the number of stops. This indicates that the balance of delay between the front and rear of the platoon needs to be treated carefully and further investigation needs to be considered. However, either early or late offset does not improve the delay and stops much. The delay is at most 4 seconds better and the number of stops is only 0.2 better than the current value. For an arterial with multiple intersections, merely adjusting the offset may lead to worse results.

The second experiment is conducted on another virtual network. A short arterial with two intersections was constructed in VISSIM. The arterial is 1450 feet long and has three lanes in each direction. The major road has 1600 vph input volume and the cross streets have input volume ranging from 100vph to 200 vph. The signal timing for both arterials are optimized based on SYNCHRO. The offset of the downstream intersection was first set to be 22 seconds (the ratio of distance over platoon speed). Two measures of effectiveness, average delay and average number of stops for the through movement on the arterial, are then evaluated under different offset values second by second. From the Figure 18, it can be seen that the offset corresponding to the lowest average number of stops is 19 seconds and the offset corresponding to the lowest average delay value is 26 seconds. This indicates that the simple method to calculate offset is unsatisfactory. The results show that an earlier offset will reduce the number of stops and a late offset may reduce the average delay. The offset values which will yield the optimal delay and stops by the proposed modified model and Robertson's model are also calculated for comparison. The proposed modified model recommends 19 seconds for optimal number of stops and 27 seconds for optimal delay. Robertson's model recommends 20 seconds for optimal number of stops and 29 seconds for optimal delay. It can be observed that the offsets recommended by the proposed modified dispersion model is closer to the simulated optimal offset. This shows that the proposed model is able to model platoon dispersion better. However, it should be noted that the improvements in the MOEs is not too significant. In order to achieve a more desirable result, it is not enough by just simply adjusting the offset and a more complicated signal timing method is planned as the following research.

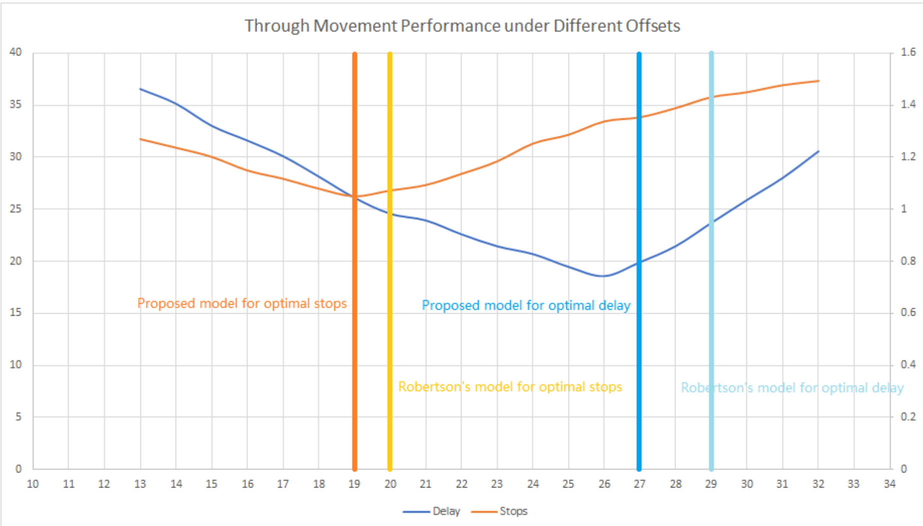


Figure 18. Average Number of Stops and Average Delay under Different Offsets Values

The above experiments show that just adjusting the offset is not enough. In order to improve the performance better, the following issues need to be addressed:

1) The platoon dispersion is not modeled accurately and in details as in Robertson's model. The evolution and changes of the platoon characteristics are not considered.

2) As traffic volume changes overtime, a fixed timing strategy may not be capable to deal with these volume changes. A dynamic signal timing strategy that responds to the various traffic conditions should be developed

3) The previous method considers fixed cycle and phase time for each intersection. To achieve optimal performance, the cycle time and phase split may be adjusted as well.

4) The previous method only considers main street through traffic. In reality, the turning vehicles and cross street vehicles are important as well.

In order to address the above issues, a novel signal timing strategy is desired and is further elaborated in the next section.

5.2 Max Pressure Policy and Predicted Platoon Max-Pressure Policy

The Max-Pressure policy first originated from the electrical and computer engineering field by Tassiulas and Ephremides (63) with implementation in computer networks. This policy was first applied in maximizing queueing system throughput in multi-hop networks. The early applications of the Max-Pressure Policy were introduced by Varaiya (64) and Wongpiromsarn et al. (65). A transportation network with multiple intersections can also be regarded as a multi-hop network which originates from the communication network area. In Varaiya's and Wongpiromsarn's works, the system throughput is optimized based on routing information and infinite queue capacity. The only required input for Max-Pressure Policy is the queues of adjacent intersections and the demand and saturation rates are not needed. Furthermore, the constraints on queue capacity and routing information is relaxed. (66) (67). The queue capacity constraint is resolved by assigning a finite but large enough queue capacity and the routing information constraint can be resolved by back-pressure-based control. Further studies on the implementation of Max-Pressure Policy in transportation area were made. Lioris et al. (68) tested the Max-Pressure Policy on a real network under various traffic conditions with uncongested and congested traffic. Anderson (69) identified Max-Pressure Policy's unpractical hardware problems and modified the policy by setting update times and allowable actuation domains for traffic signals with less green splits updating abilities.

Currently, traffic on arterials or networks are managed by traffic signals which switches between green and red periodically. When the traffic lights change from green to red which is also called a phase change, there will be a clearance interval due to safety considerations. The clearance interval contains to parts, a yellow change and a red clearance or the all-red interval. The former is to warn drivers of the upcoming change in the right-of-way, and the latter is to provide some additional time for vehicles already in the intersection to clear the intersection before the next phase. During this transition period, the intersection will have a throughput of almost zero. This will reduce the system performance. What makes the original Max-Pressure Policy even worse in its application in transportation is that the real vehicles need some time to accelerate or decelerate when a green phase starts or ends. The previously mentioned problems will result in a capacity loss and can be defined as the switch-over delay. To solve this issue on capacity loss, a Biased Max-Pressure Policy was proposed by Hsieh et al. (70). A bias factor is introduced to ensure that a phase switch will only happen if the maximum pressure is larger than the current phase pressure by a certain portion. This prevents the frequent phase switches and reduce the switch-over delay and capacity loss to a large extent.

The BMP was tested on a network with six intersections, and all intersections and links are treated equally. On arterials however, vehicles arrive in platoons on major road and cross street volumes are rather low. The arterial traffic may be treated with higher priority in this situation. The BMP only detects stationary queue length. When platoons arrive at the downstream intersection, they may be stopped by red light and experience undesired stops and delay. To allow platoons progress the downstream intersections without stops and further improve the system performance, a Platoon Predicted Max-Pressure Policy (PPMP) is proposed here. The policy takes into account of platoon dispersion and evolution of platoon characteristics to make advance decisions on signal timing allocations and phase switches.

We first define the pressure as:

$$W_{i,j}(t) = P_{i,j}(t) - \sum_{k:k \in D(j)} p_{j,k} * P_{j,k}(t) \quad (28)$$

Where i, j, k denotes links. i is the upstream link of j and j is the upstream link of k. $D(j)$ is the set of all downstream links of j, $p_{j,k}$ is the routing probability of vehicles entering from j choosing downstream link k. Note that the main difference here between the PPMP and the original BMP is the pressure based on queue length is adjusted to a predicted pressure $P_{i,j}(t)$ including the residual queue and the incoming platoon. There are several ways to define the adjusted pressure. In the current study, it's defined in the following methods and each definition was evaluated based on simulation. The evaluation results are presented in section 6.4.

1) The predicted pressure is simply calculated by multiplying the predicted density and the platoon length when the platoon arrives at the downstream intersection.

$$P_{i,j}(t) = \sum d_{i,s} * seg_s + q_{i,j} \quad (29)$$

2) The predicted pressure is adjusted by dividing predicted incoming platoon to segments with different density and assigning a weight factor to each segment of the platoon. The weight factor can be assumed linearly:

$$P_{i,j}(t) = \sum d_{i,s} * seg_s * w_s + q_{i,j} \quad (30)$$

$$w_s = k w_{s_0} \quad (31)$$

Where $d_{i,s}$ is the density of platoon segment s on upstream link i, seg_s is the segment s of the platoon when it arrives at the downstream intersection, $q_{i,j}$ is the current queue at the downstream intersection, and w_s is the weight assigned to segment s. By choosing a base segment s_0 and assign a weight value to this segment, the weight of the rest of the segments can be calculated by equation (31).

3) The predicted pressure is adjusted by squaring the density of each segment and multiply the segment length of the predicted platoon at downstream intersection.

$$P_{i,j}(t) = \sum (d_{i,s})^2 * seg_s + q_{i,j} \quad (32)$$

4) The predicted pressure is adjusted by assigning a weight to the segment density.

$$P_{i,j}(t) = \sum d_{i,s} * seg_s * w_d + q_{i,j} \quad (33)$$

Where the density weight w_d is obtained from calibration. The most suitable value is close to the square of the ratio of segment density over overall density of the predicted platoon at downstream intersection.

It can be seen that method 1 is a simple prediction of the number of vehicles in the platoon when the platoon arrives at the downstream intersection. It should be noted due to platoon dispersion, vehicles spread unevenly along the link with fast vehicles in front and slow vehicles moving at rear. Therefore, it is not necessary to allow all vehicles in a platoon to progress the downstream intersection, as doing this may sacrifice the privilege of vehicles from other directions and resulting in worse performance. The platoon may be cut at a certain percentage value to achieve the optimal result. A sensitivity study on the cut percentage is presented in section 6.4. The problem with method 1 is that it treats all vehicles in the platoon equally. This may not yield the optimal result as vehicles distribute unevenly in the platoon. Some segments should be more favored than others. Method 2 can be used when favoring on the front or rear of the platoon is desired. Method 3 and 4 are different ways of favoring on the denser segments of the platoon.

The pressure calculated in the previous step can be implemented in the following algorithm for signal control.

1) Check the current queue at the intersection first. At time t_n, the length of the k-th superframe is calculated based on the sum of the queue length vector of each movement:

$$T_k = (\sum Q_{i,j}(t_k))^{\alpha}, \alpha \in (0,1) \quad (34)$$

$$t_{k+1} = t_k + T_k \quad (35)$$

2) Perform the step 2) with $W^*_{i,j}(t)$ calculated from stable queue stopped at the intersection first. The phase with maximum pressure for intersection u at time t is:

$$I_u(t) \in \arg\max_{I \in I_u} \sum \mu_{i,j} I_{i,j} W^*_{i,j}(t) \quad (36)$$

Where $\mu_{i,j}$ is the service rate of the movement i, j, $I_{i,j}$ is the indicator function of whether $Q_{i,j}$ is scheduled at a certain intersection. If $I_u(t)$ equals to $I_u(t-1)$, continue with $I_u(t)$.

Next, perform step 2) with $W^*_{i,j}(t)$ calculated from adjusted pressure as in equations (29)-(33). If $I_u(t)$ does not equal to $I_u(t-1)$, switch-over should be initiated for next T time slots, and $I_u(t)$ should then be applied for the next slot.

3) For the rest of the k-th superframe, find the phase $I_u(t)$ with maximum pressure. A switch can be made if the following condition is satisfied, that is, only if the maximum pressure is larger than the current phase pressure by a certain portion:

$$(1+B_u(t))\max(\sum \mu_{i,j} I_{i,j}(t-1) W^*_{i,j}(t), 0) < \max(\sum \mu_{i,j} I_{i,j}(t) W^*_{i,j}(t), 0) \quad (37)$$

Where $B_u(t)$ can be treated as a bias function and is calculated as:

$$B_u(t) = \zeta \min\{1, (\max(\sum W^*_{i,j}(t), 0)^{-\beta}\}, \beta \in (0,1), \varsigma > 0 \text{ (38)}$$

First check with $W^*_{i,j}(t)$ calculated from stable queue stopped at the intersection. Then calculate $W^*_{i,j}(t)$ again with adjusted pressure as in equations (29)-(33). Calculate the platoon arrival time based on the modified linear multivariate model for maximum speed presented in section 5.2. If the platoon arrives within the current superframe, initiate switch-over after $T_{arrival} - T_{clearance}$, where $T_{clearance}$ is the time to clear the current queue.

Otherwise, stay in the current phase.

4) Repeat step 3) and 4) until the end of the k-th superframe.

5) Repeat 1) to 4) at time $t = t_{k+1}$.

5.3 Proof of Throughput Optimality

At one superframe, the adjusted queue length is the difference between the arrived vehicles (arrival rate $A_{i,j}(t)$) and the vehicles that left the system (service rate $S_{i,j}(t)$):

$$Q_{i,j}(t) = \sum A_{i,j}(t) - \sum (\min\{S_{i,j}(t) * I_{i,j}(t) * X_{i,j}(t), Q_{i,j}(t)\}) \quad (39)$$

Where $X_{i,j}(t)$ is the indicator function of whether movement i, j is active at t.

The arrived vehicles can be re-written as the sum of all upstream movements M :

$$\sum A_{i,j}(t) = \sum \sum_{m,i \in M} \min\{S_{m,i}(t) * I_{m,i}(t) * X_{m.i}(t), Q_{m,i}(t)\} R_{i,j}(t) \quad (40)$$

Now define a Lyapunov drive over the studied superframe:

$$L(t) = 2Q(t)^T \Delta Q(t) + \Delta Q(t)^T \Delta Q(t) \quad (41)$$

Where $\Delta Q(t)$ is the difference between $Q(t)$ and $Q(t+1)$.

The conditional drift over one superframe is upper bounded as:

$$E\left[L(t) \middle| Q(t)\right] \le -2\varepsilon T_k \sum \max\{W_{i,j}(t), 0\} + C_1 \sum_{v \in V} M^v (\sum_{i,j \in M} \max\{W_{i,j}(t), 0\}) + C_2 \sum_{v \in V} \sum_{i,j \in M} \max\{W_{i,j}(t), 0\} + C_3 T_k^2 + C_4 T_k$$

(42)

Which is obtained by finding the upper bound for each of the terms in equation (41), and

$$C_1 = 2T_S(\varepsilon + S_{\max}) \quad (43)$$

$$C_2 = 2S_{\max}(\max CL) \quad (44)$$

$$C_3 = |M| U_{\max}^2 V_{\max}^2 + 2(\sum_{v \in V} (U_{\max} + 1) V_{\max} (\sum_{i,j \in M} \mu_{i,j})) \quad (45)$$

$$C_4 = 2(\sum_{i,j \in M} \mu_{i,j} S_{\max}) \quad (46)$$

Where V are the set of intersections v, CL is the fix-time control cycle length, $V_{\max}$ is the maximum of the arrival and service rate, and T_S is the length for switch over delay.

The pressure vector $W_{i,j}(t)$ has a lower bound based on the queue vector $Q_{i,j}(t)$ by applying the Perron-Frobenius theorem to the routing matrix and obtaining eigenvectors chosen to have strictly positive components:

$$\sum_{i,j \in M} \max \{ W_{i,j}(t), 0 \} \geq \delta \sum_{i,j \in M} Q_{i,j}(t) \quad (47)$$

Applying (47) to (42), the following inequality can be obtained based on inequality (42):

$$E\left[L(t) | Q(t) \right] \leq -2\varepsilon\delta T_k \sum_{i,j \in M} Q_{i,j}(t) + C_1 \sum_{v \in V} M^v \sum_{i,j \in M} Q_{i,j}(t) + C_2 \sum_{v \in V} \sum_{i,j \in M} Q_{i,j}(t) + C_3 T_k^2 + C_4 T_k \quad (48)$$

Since the time of the n^{th} switch-over at superframe k is lower bounded as:

$$T_n^k \geq C_5 B_v t_n^k \sum_{i,j \in M} \max \{ W_{i,j}(t_n^k), 0 \} \quad (49)$$

Where $C_5 = \mu_{\max}((2 + \zeta T_S)|M|^2 (U_{\max} + 1) V_{\max})^{-1}$.

The second term on the right side of inequality (42) can be written as:

$$M^v (\sum_{i,j \in M} \max(W_{i,j}(t), 0)) = \gamma (\sum_{i,j \in M} Q_{i,j}(t)^{1+\beta}) \quad (50)$$

Hence it can be observed from inequality (48) that the first term on the right side is the dominating term. Inequality (48) can be re-written with a positive constant C:

$$E\left[L(t) | Q(t) \right] \leq C - \varepsilon\delta (\sum_{i,j \in M} Q_{i,j}(t))^{1+\beta} \quad (51)$$

Let $G(t)=\sum_{0}^{T_{k-1}}\sum_{i,j\in M}Q_{i,j}(t+\Delta t)$, the following inequality can be obtained:

$$G(t)\le\sum_{0}^{T_k-1}\sum_{i,j}(Q_{i,j}(t)+\sum_{s=0}^{T_k-1}A_{i,j}(t+s))+\sum_{0}^{T_k-1}\sum_{i,j}Q_{i,j}(t)\ (52)$$

The conditional expectation of $G(t)$:

$$E\left[G(t)\middle|Q(t)\right]\le C_6(\sum_{i,j}Q_{i,j}(t))^{1+\beta}\ (53)$$

Re-writing inequality (51):

$$E\left[L(t)\middle|Q(t)\right]\le C-\frac{\varepsilon}{C_6}E\left[G(t)\middle|Q(t)\right]\ (54)$$

By looking at all the superframes, inequality (54) becomes:

$$\sum E\left[G(t)\middle|Q(t)\right]\le\sum(C-\frac{\varepsilon}{C_6}E\left[G(t)\middle|Q(t)\right])\ (55)$$

Since $L(0)<\infty$ and $\sum E\left[G(t)\middle|Q(t)\right]>-L(0)$, the following can be obtained:

$$\lim_{T\to\infty}\sup((\sum_{0}^{T-1}E[\sum_{i,j}Q_{i,j}(t)])/T)\le C_6*(C+L(0))/\varepsilon<\infty\ (56)$$

Since we have $\lim_{T\to\infty}\sup((\sum_{0}^{T-1}\sum_{i,j}E[Q_{i,j}(t)])/T)<\infty$, we can say that the multi-hop transportation system is strongly stable. Therefore, the throughput is optimal for tor the current system.

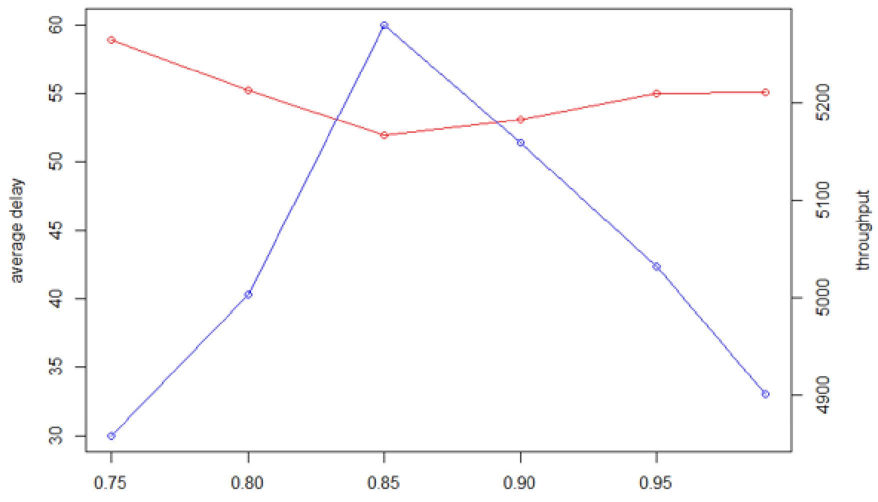
5.4 Performance Evaluation

5.4.1 Analysis on Platoon Percentage Cut for Upstream Pressure

In this section, the performance of various traditional and advanced signal timing strategies is tested on simulation networks and evaluated. To start, a study on the percentage value to cut the platoon to calculate upstream pressure is presented. As mentioned in section 6.2, it may not be necessary to include all platoon vehicles while calculating the upstream platoon predicted pressure. A percentage from the front to the rear of the platoon may be selected as the “active portion” of the upstream pressure calculation. To investigate the percentage value that yield the optimal result, a calibration on the platoon percentage cut was conducted. The platoon percentage cut starts from 75% percentile and increases by 5%, to 99% which approximately represents all platoon vehicles are counted as “active”. Table 20 shows the calibration results as the four measures of effectiveness of the system are set as the indicator. The simulation was conducted on a virtual network with 5 intersections and the network and simulation setting are the same as in section 6.4.2.

Table 20. Calibration on the Optimal Platoon Percentage Cut

	Average Delay		Throughput		Average Number of Stops		Average Stop Delay	
Density Percentage	1200vph	1600vph	1200vph	1600vph	1200vph	1600vph	1200vph	1600vph
75%	38.482	58.908	3790	4858	1.729	1.863	26.915	42.552
80%	36.54	55.225	3886	5003	1.655	1.811	24.271	40.08
85%	31.095	**51.962**	**4007**	**5280**	**1.582**	**1.764**	22.405	**36.981**
90%	**30.898**	53.121	3945	5159	1.607	1.878	**22.108**	37.796
95%	33.413	54.991	3902	5032	1.723	1.892	24.643	39.002
99%	35.102	55.066	3879	4901	1.76	1.904	25.437	39.873



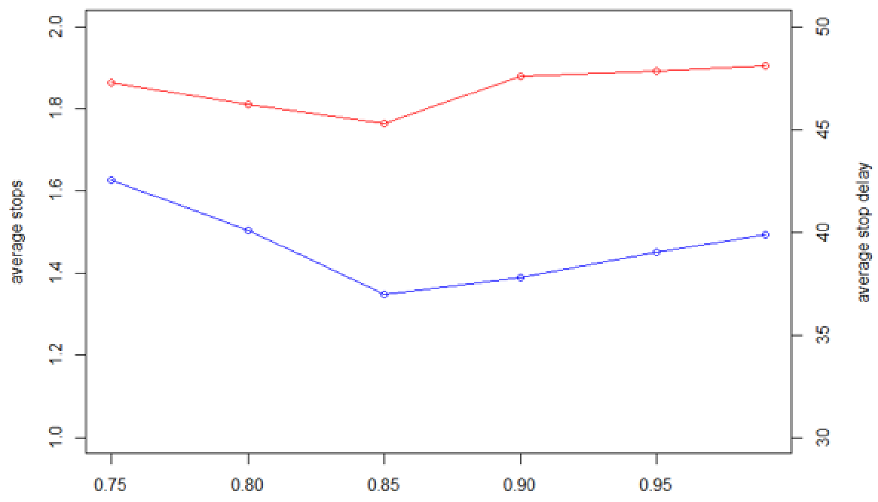


Figure 19. Calibration on the Optimal Platoon Percentage Cut with Average Delay and Throughput (up), and Average Number of Stops and Average Stop Delay (bottom)

From Table 20. and Figure 19, it can be seen that optimal platoon percentage cut value should be around 85% to 90% percentile, as all the average delay, throughput, average number of stops and average stop delay are the lowest in this interval. This indicates that the pressure calculation may not need to include all vehicles. The pressure calculation based on a certain percentage may yield better results.

5.4.2 Simulation Results and Evaluation

To evaluate the performance of various PPMP policies and compare with the baseline methods, simulations were conducted on two virtual networks constructed in microscopic traffic simulation software VISSIM. To obtain real-time information of the traffic information, an additional interface, VISSIM COM was used to establish communication with outside software MATLAB for computation and optimization purposes. The results from MATLAB can be used as a feedback to VISSIM through the COM interface and perform real-time signal timing control.

The simulation was first conducted on an artificial arterial with 5 intersections, and each link was constructed at an arbitrary length including short, medium and long link lengths. The main purpose is to test the strategies' applicability on a random arterial and investigate the performance of various strategies under different volume levels. A layout of the simulation arterial is presented in the Figure 20. Each direction of the main arterial has three lanes with a left-turn bay at the intersections. The cross streets have two lanes on each direction with volume randomly set between 100 vph and 200 vph. The main arterial has four volume levels, starting from 1200 vph and increasing at 400 vph to 2400 vph. The traffic composition is set to be passenger cars only. From section 5.1.2, it is known that including trucks may result in a different distribution type and distribution parameter model. Hence, simulation with heterogeneous flow may be regarded as one direction of the future research when enough real traffic data involving trucks is available. The simulation length was 1 hour and was run five times with different seeds. The final results were averaged on the five runs. Four most common measures of effectiveness were selected: the average delay, the throughput, the average number of stops, and the average stop delay. Eight signal timing strategies were tested for comparison. The results were presented in Table 21 and Table 22. A brief explanation for each of the timing strategy is listed below.

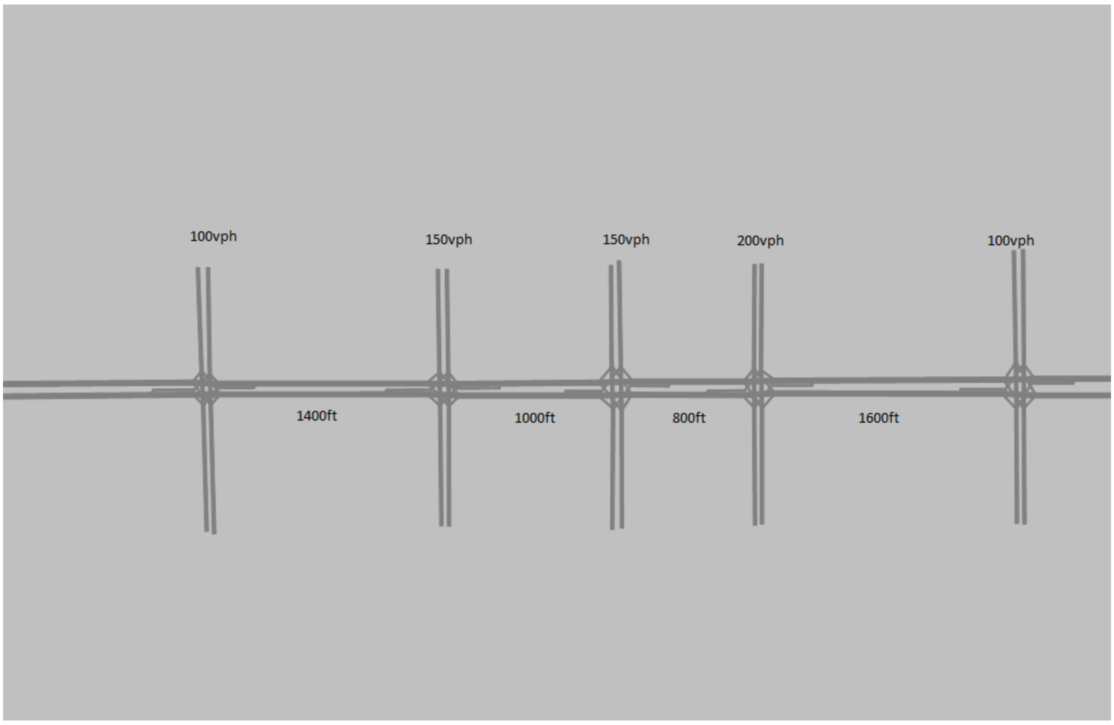


Figure 20. Layout of the Artificial Arterial in VISSIM

Fix-Time: Each intersection has fixed cycle time and fixed phase split. The timing for each intersection was optimized through traffic optimization software SYNCHRO and the offset of each intersection was optimized for the whole arterial. The optimized timing was used as input for VISSIM simulation.

Actuated: Each phase of the signal timing at an intersection is actuated and requires a detection on the traffic. The timing and phase splits are subject to real demand. The actuated timing was also optimized through SYNCHRO.

Max-Pressure: Original Max-Pressure Policy implemented directly in VISSIM simulation through COM interface.

Biased Max-Pressure: Biased Max-Pressure Policy that reduces the switch-over delay by assigning a bias factor while computing the pressure.

PPMP1: Platoon Predicted Max-Pressure that computes pressure by directly multiplying the density with length.

PPMP2: Platoon Predicted Max-Pressure that computes pressure by assigning weights to each segment. The weights were assigned linearly decreasing from the front to represent a strategy that favors the front of the platoon.

PPMP3: Platoon Predicted Max-Pressure that computes pressure by simply squaring the density value of each segment to favor platoon segments with larger density

PPMP4: Platoon Predicted Max-Pressure that computes pressure by assigning a weight factor according to the density for each segment. From calibration, the weight factor is most closely to the square root of the ratio of segment density over overall density.

Table 21. **Performance Evaluation of Various Signal Timing Strategies-Average Delay and Throughput**

	Average Delay				Throughput			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	40.916	61.724	87.588	145.229	3890	5028	5898	6558
Actuated	39.703	58.842	80.412	121.832	3904	5067	6004	6725
Max-Pressure	50.086	61.432	90.058	143.567	3742	4992	5372	6419
Biased Max-Pressure	30.422	54.626	71.772	81.877	3956	5109	6120	7104
PPMP1	30.785	56.43	73.002	92.08	3948	5120	5957	6613
PPMP2	30.085	51.862	67.731	75.449	4013	5265	6288	7286
PPMP3	35.677	58.412	75.564	94.297	3960	5209	6190	7099
PPMP4	**29.984**	**50.147**	**62.125**	**70.119**	**4129**	**5410**	**6402**	**7555**

Table 22. **Performance Evaluation of Various Signal Timing Strategies-Average Number of Stops and Average Stop Delay**

	Average Number of Stops				Average Stop Delay			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	2.465	2.82	4.187	9.576	29.451	49.102	62.595	86.565
Actuated	2.422	2.851	3.965	6.882	28.527	45.441	55.228	84.473
Max-Pressure	3.902	5.442	7.006	12.017	36.319	52.172	63.452	87.326
Biased Max-Pressure	2.039	2.389	2.925	3.487	24.913	41.74	49.582	56.155
PPMP1	1.644	2.397	2.558	3.034	23.998	39.802	48.653	55.477
PPMP2	**1.512**	**1.788**	**2.047**	**2.515**	**21.145**	**35.987**	**40.879**	**45.085**
PPMP3	1.749	2.135	3.067	3.471	25.4	43.157	48.556	56.478
PPMP4	1.611	1.882	2.219	2.996	22.458	38.39	43.504	49.721

From Table 21 and Table 22, Most PPMs provide better results than the traditional strategies. PPMP2 and PPMP4 yield the best results, viewing from two perspectives. PPMP4 has lower average delay and higher throughput, while PPMP2 performs better in terms of average number of stops and average stop delay. It can also be observed that all the PPMPs perform better at larger volume levels. The improvement in average delay increases from 10 seconds to nearly 20 seconds comparing 1200 vph and 2400 vph. The improvement in throughput also increases from 5% to about 15% comparing PPMP and actuated control. The PPMPs have significant advantage in reducing the number of stops. From Table 22, for traditional methods, the number of stops increases drastically at about three to four times as volume increases. The original Max-Pressure Policy performs even worse as frequent switch-overs take place in these situations. The PPMPs however, are not affected by the increase in volume, with only about 65% increase. Compared with the Biased Max-Pressure Policy which is also not sensitive to the increase in volume, PPMP2 outperforms by 20% at low volumes and bout 30% at high volumes. Although the two strategies outperform others to a large extent, the difference between them is rather small. Either these two strategies can be selected, and implemented depending on whether delay and throughput or the number of stops is regarded as more important. Figure 21-Figure 24 provides a graphic illustration on the performance of various strategies. The improvements under high volume conditions can be observed more intuitively.

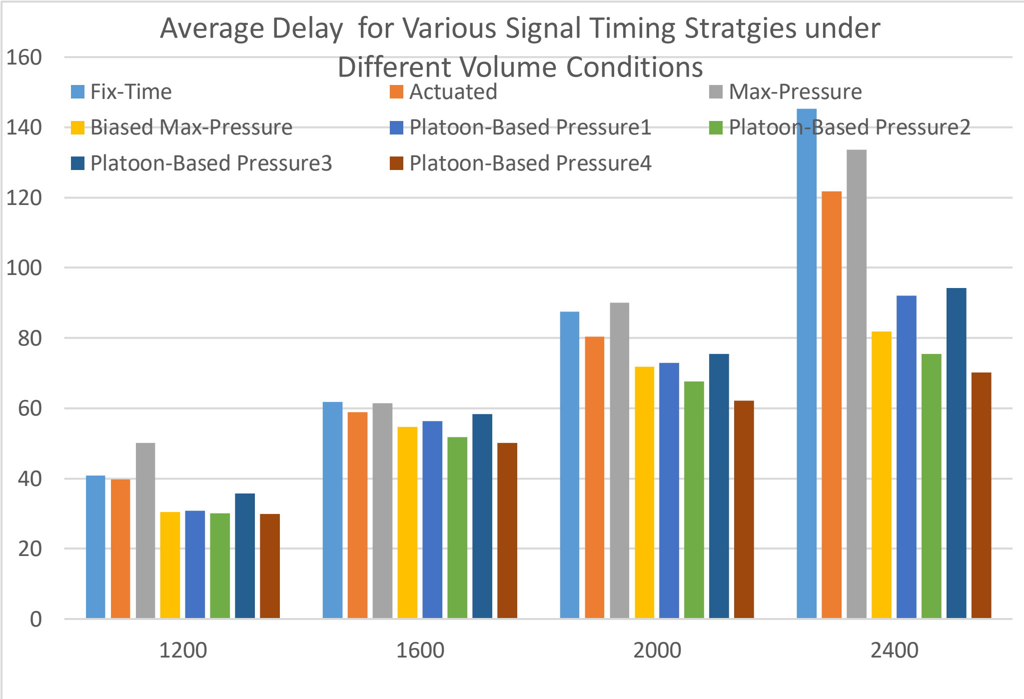


Figure 21. Average Delay under Different Volume Conditions

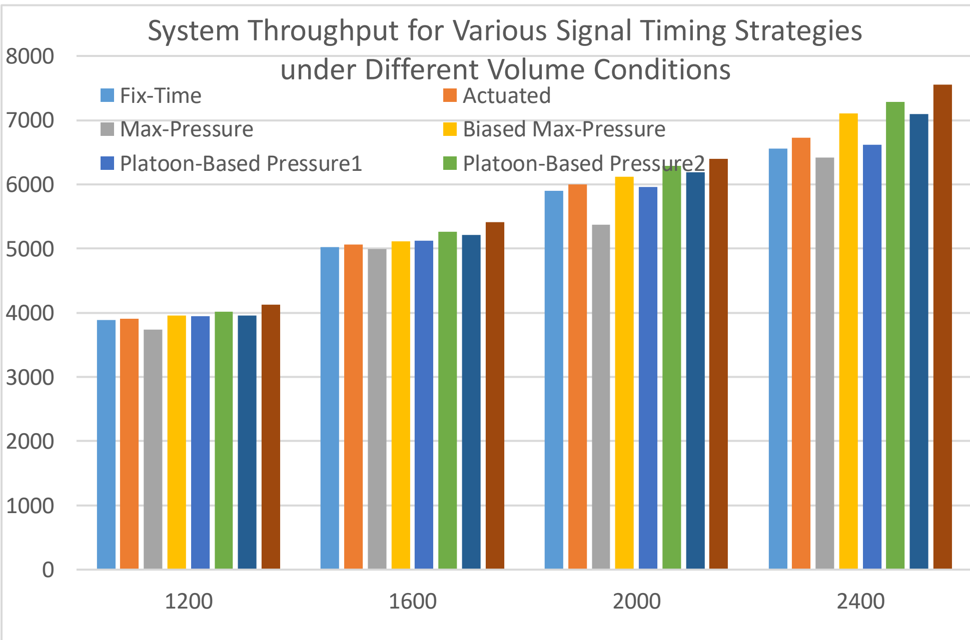


Figure 22. Throughput under Different Volume Conditions

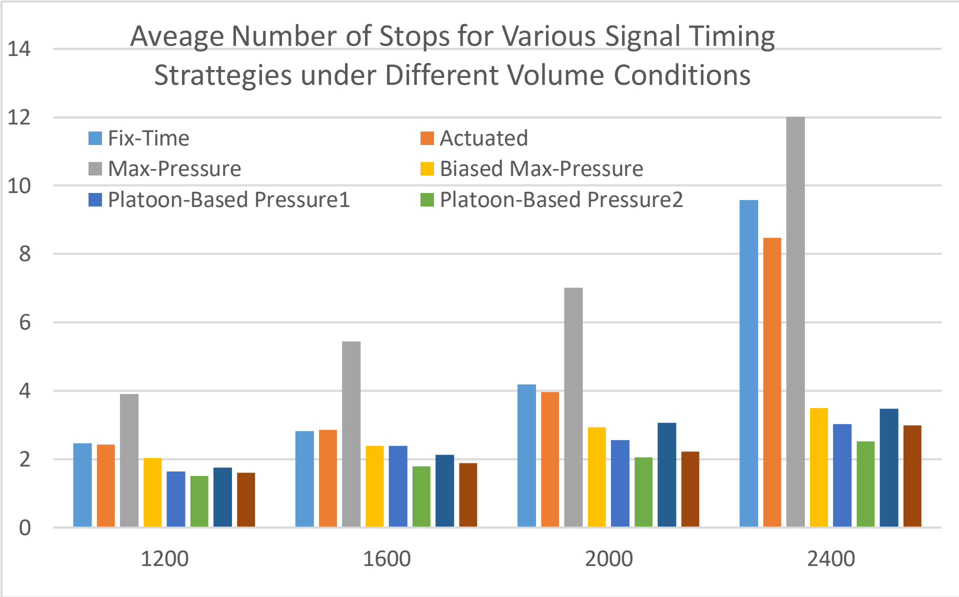
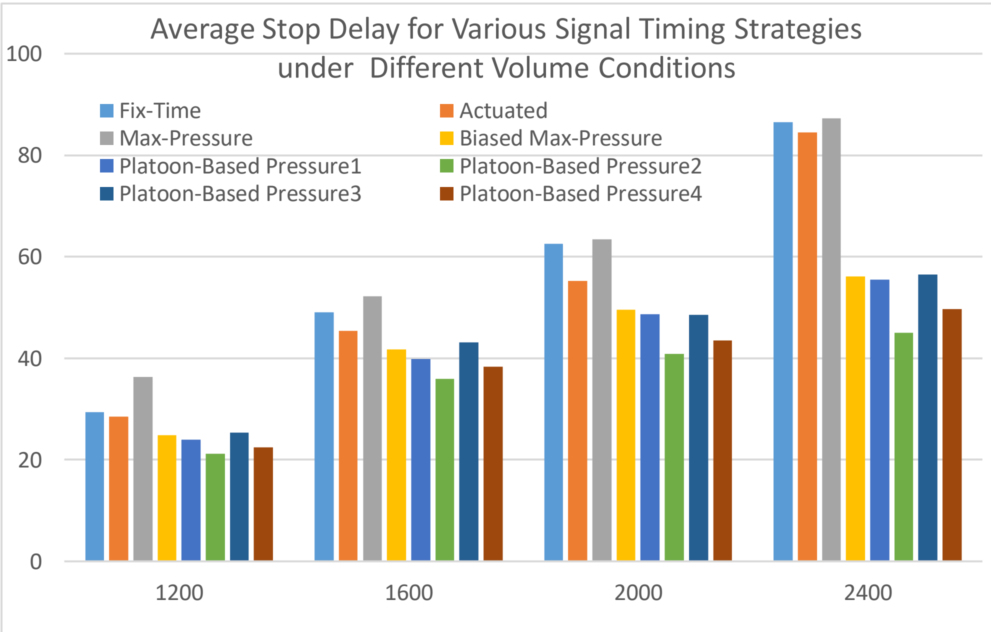


Figure 23. Average Number of Stops under Different Volume Conditions



3

Figure 24. Average Stop Delay under Different Volume Conditions

The simulation was also conducted on a virtual arterial constructed based on a section of the real arterial, Texas Avenue in College Station, Texas. The section starts from George Bush Drive and ends at Southwest Parkway. There are total six intersections, namely from left to right are George Bush Drive, Harvey Road, Holleman Drive, Manuel Drive, Brentwood Drive and Southwest Parkway. The input volume was based on real volume data for each link. The geometrical conditions were set to be as close to the real-world conditions as possible, as shown in Figure 25. The simulation results are presented in Table 23. It can be seen that on a simulated real-world arterial, similar results can be observed. PPMP2 and PPMP4 yield the best results.

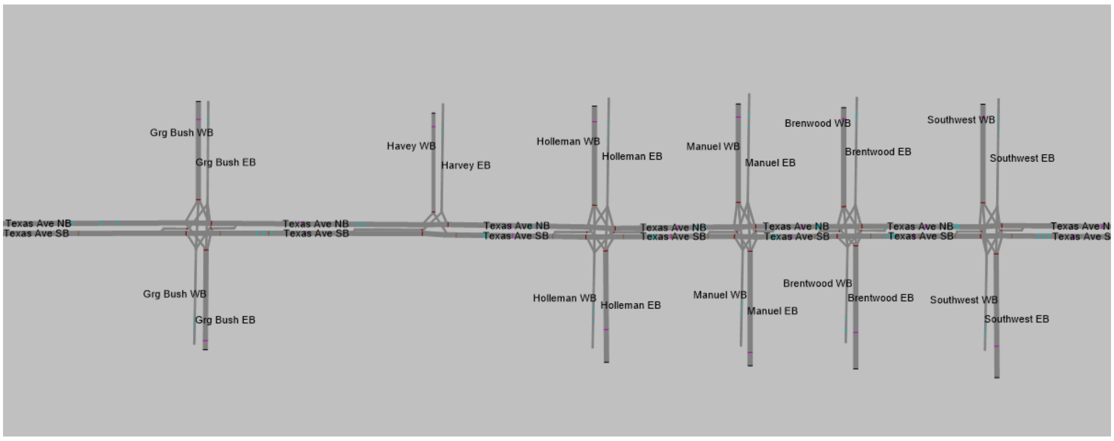


Figure 25. **Simulated Arterial based on Texas Avenue**

Table 23. Performance Evaluation on Texas Avenue

	Average Delay	Throughput	Average Number of Stops	Average Stop Delay
Fix-Time	70.458	4694	2.588	49.62
Actuated	69.404	4811	2.455	47.986
Max-Pressure	71.578	4782	5.889	52.774
Biased Max-Pressure	62.125	4898	2.229	43.883
PPMP1	64.086	4881	2.242	44.052
PPMP2	57.862	4969	**2.004**	**41.055**
PPMP3	63.412	4902	2.259	49.127
PPMP4	**56.147**	**5113**	2.142	43.004

5.4.3 Comparison with Dynamic Re-Grouping Strategy

In this section, a simple signal timing strategy that detects platoon information and dynamically re-group the dispersed platoon on a long arterial is presented. The strategy is based on the fact that platoons disperse over long distances. As the distance increases, the platoon dispersion becomes larger and the benefits of signal coordination will gradually decrease. In this situation, it may be more desirable to re-group the dispersed platoon vehicles to form a new platoon that is more compact, rather than merely adjusting the offset. The dynamic re-grouping strategy is planned to dynamically find an optimal downstream intersection to stop the platoon vehicles and re-group, based on real-time data detected upstream. A preliminary formulation of the dynamic re-grouping strategy is presented below. It serves as another comparison with the traditional methods and the max-pressure policies.

Assuming that the intersections are numbered as 1, 2, … i … n and are originally with fixed timing and coordination. Intersection 1 is regarded as the master intersection and each intersection has a fixed offset O_k^i (O_f^i is simply computed as link length over speed). The goal of the Dynamic Re-Grouping Strategy is to find the intersection r that has a different offset value O_k^r which stops the entire of or part of the platoon for re-grouping of the platoon vehicles. The intersections after intersection r will have adjusted offsets with intersection r as the master intersection. The strategy will find the optimal values of r and O_k^r that produce the maximum throughput.

1) Firstly, platoon with platoon size V_0 arriving at the arterial system at the k -th cycle is detected. k takes positive integer values. The platoon information was used to perform the following series of calculations starting from determining the time to arrive the downstream intersections by the multivariate model presented in section 5.2.

2) Secondly, the number of vehicles that have passed and not have passed the first downstream intersection with distance x_1 at time t_k^1 , $A_k^1(x_1, t_1^k)$ and $B_k^1(x_1, t_1^k)$, with function $H(x,t)$ which is a simplified representation for the function groups (18) -(23). $t_k^1 = s_k^1 + O_k^1$ where s_k^1 is the link travel time for intersection 1.

3) Next, assuming that the vehicles that have not passed the first intersection from the previous cycle, $B_{k-1}^1(x_1, t_{k-1}^1)$, joins the front of the platoon of the current cycle without disrupting the platoon of cycle k . Since they are in front of the platoon, all vehicles from the previous cycle can progress to the next intersection if assuming that the capacity is sufficient. The new platoon size V_1 is calculated as

$A_k^1(x_1,t_k^1)+B_{k-1}^1(x_1,t_k^1)$. With the new platoon size, Step 2) can be repeated to calculate $A_k^2(x_2,t_k^2)$ and $B_k^2(x_2,t_k^2)$ using function $H(x,t)$. The platoon size before the i-th intersection is $A_k^{i-1}(x_{i-1},t_k^{i-1})+B_{k-1}^{i-1}(x_{i-1},t_{k-1}^{i-1})$. Note that when $k=1$, the number of vehicles in the previous cycle are 0, and $t_k^i=t_k^{i-1}+s_k^i+O_k^i$.

4) Repeat step 3) iteratively to intersection r. For the first cycle, the number of vehicles that have not passed that downstream intersection is $B_1^r(x_r,t_1^r)$, with calculation of $t_k^r=t_k^{r-1}+s_k^r+O_k^r$. The stopped vehicles are only able to progress until the next cycle, together with the arriving vehicles at the second cycle $A_2^{r-1}(x_{r-1},t_2^{r-1})$. The total number of vehicles that are able to progress the next intersection will be the small values of $A_2^{r-1}(x_{r-1},t_2^{r-1})+B_1^r(x_r,t_1^r)$ and the maximum capacity C of the green phase of intersection r which is a known value. The smaller of the two will be the platoon size for the next intersection. For the rest of cycles, the platoon size will be $\max\left\{A_{k+1}^{r-1}(x_{r-1},t_{k+1}^{r-1})+B_k^r(x_r,t_k^r)-C,0\right\}+A_{k+2}^{r-1}(x_{r-1},t_{k+2}^{r-1})$.

5) For the rest of the intersections, intersection r will be regarded as the master intersection. Their new offsets will be based on intersection r and needs to be updated. Assuming they are denoted as O_k^j, $j>r$, and the time to travel to intersection $j=r+1$ is $t_{r+1}^k=t_{r-1}^k+s_r^k+R+s_{r+1}^k+O_k^{r+1}$. Repeat step 2) and 3) with $H(x,t)$ until the last intersection n of the arterial system. The number of vehicles that have passed the entire system for vehicles entering at the k-th cycle is $A_k^n(x_n,t_k^n)$, where $t_n^k=t_{n-1}^k+O_k^n$. The objective is to find values of r and O_k^r that maximizes $A_k^n(x_n,t_k^n)$. r takes values 2 to n and O_k^r takes values from 0 to cycle length.

The strategy is tested on a ten-intersection arterial in VISSIM under a low volume condition and a high-volume condition as shown in Figure 26. The settings are similar to the network in 6.4.2. The simulation results are presented in Table 24. From the table, it can be seen that the re-grouping strategy indeed have better performance compared with fix-time control, but performs much worse than the PPMPs. However, the Dynamic Re-Grouping Strategy that takes advantage of the evolution of the platoon characteristics and platoon dispersion patterns can still be a feasible solution in situations that PPMPs are not allowable due to constraints such as local geometric conditions, budgets and technology constraints.

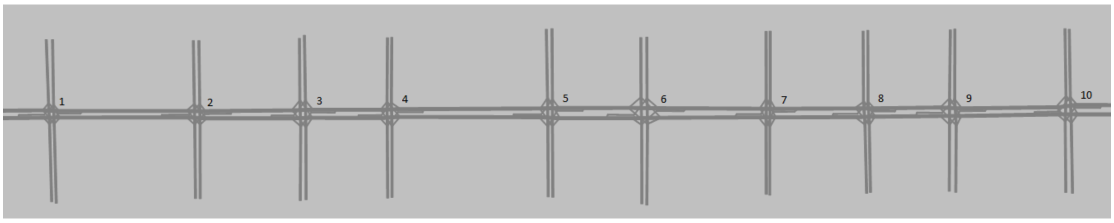


Figure 26. **Virtual Arterial with Ten Intersections**

Table 24. **Performance Evaluation for PPMP and Dynamic Re-grouping Method**

Comparison on a 10 Intersection Arterial								
	Delay		Throughput		Average No. Stops		Average Stop Delay	
Volume Level	1600	2400	1600	2400	1600	2400	1600	2400
Fix-Time	99.447	181.284	6170	7449	4.887	11.906	84.846	124.665
Dynamic Re-grouping	95.442	155.47	6224	8009	4.288	9.145	75.114	102.355
PPMP2	90.709	102.276	6399	8311	**3.004**	**6.753**	**66.592**	**76.774**
PPMP4	**88.492**	**99.401**	**6442**	**8545**	3.454	5.929	68.412	78.212

6. EXTENSIONS OF THE PREDICTED PLATOON MAX-PRESSURE POLICY

This section discusses the extensions of the Predicted Platoon Max-Pressure Policy by supplementing more evaluation scenarios and demonstrating possible applications in real world. Section 6.1 adds one more result evaluation on through movements while Section 6.2-6.4 discusses various ways of improving the PPMP's practicality in the real world.

6.1 Evaluating Through Movements on Major Arterials

Since the signal timing strategy is designed for urban arterials, the performance of the through traffic on the major arterial is often more important as the through vehicles on the major arterial take up a very large proportion of the total volume composition. Sometimes the cross-street performance may be compromised as the through performance is a stronger indicator to the effectiveness of the signal timing plan. Hence, simulations were conducted on the network as shown in Figure 20 to evaluate through movements on the major arterial. The results are presented in Table 25 and Table 26 with similar trends as in the evaluation for all movements. The results show that by looking at through movements only, the measures of effectiveness are also greatly improved. The delay and stops are improved at about 15-30% for low volume and 40%-50% for high volume conditions. The results show that though the PPMP like its family, the Max-Pressure Policies, are not regarded as a traditional coordination strategy that only favors through traffic, they are still capable of handling and giving priorities to the major direction of traffic. This is accomplished by the policy's nature of assigning green by "pressure", which is directly related to volume. In addition, for the PPMP, the favored direction can be adjusted in real time by changing the weights of different directions.

Table 25. Performance Evaluation for Through Movements of Various Signal Timing Strategies-Average Delay and Throughput

	Average Delay				Throughput			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	36.916	57.724	82.588	121.229	3024	4248	5005	5874
Actuated	40.058	60.884	72.756	114.074	3102	4467	5212	6039
Max-Pressure	62.899	65.453	91.46	111.523	3499	4683	5015	6209
PPMP2	25.439	48.076	62.598	70.017	3545	4786	5893	6886
PPMP4	**23.984**	**42.071**	**47.544**	**65.408**	**3888**	**5109**	**6099**	**7213**

Table 26. Performance Evaluation for Through Movements of Various Signal Timing Strategies-Average Number of Stops and Average Stop Delay

	Average Number of Stops				Average Stop Delay			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	2.165	2.52	3.899	8.842	25.071	50.432	60.531	84.223
Actuated	2.049	2.376	3.774	8.054	29.613	42.869	52.407	81.541
Max-Pressure	3.102	4.829	5.782	11.265	34.422	50.718	62.965	88.423
PPMP2	**1.502**	**1.878**	**2.065**	**2.664**	**19.012**	**33.447**	**40.084**	**48.387**
PPMP4	1.608	1.956	2.244	2.789	21.448	35.605	44.199	49.787

6.2 Implementation of Minimum Green and Pedestrian Crossing Time

To test the PPMP's practicality in the real world, a minimum green time should be added first. The minimum green time is an important parameter in traffic signal control. It not only help to provide enough time for pedestrians to cross the streets, but also ensures that the green time is long enough to meet the vehicle drivers' expectations. The PPMP is capable of taking a user-defined minimum green time parameter into its computation process. According to the HCM, one way to calculate the pedestrian crossing time is Min G=Street width/4fps + 7 – Inter-green time. While running simulations, the minimum green time of each intersection can be calculated by substituting local values for each term of the equations. The simulation results based on the network shown in Figure 20 is presented below. It can be seen that adding minimum green will result in a reduction on the PPMP's performance, especially at low volumes. The PPMP's performance improvement is rather small. At high volumes however, the negative impact of minimum green time is negligible. This proves that the PPMP has great capability in handling high volume traffic than traditional methods, and is capable of adapting scenarios with mixed traffic including pedestrians.

Table 27. Performance Evaluation with Minimum Green Time (Pedestrian Crossing Time)-Average Delay and Throughput

	Average Delay				Throughput			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	45.916	61.724	87.588	145.229	3890	5028	5898	6558
Actuated	39.703	58.842	80.412	121.832	3904	5067	6004	6725
Max-Pressure	50.086	61.432	90.058	133.567	3742	4992	5372	6419
PPMP2	39.085	60.288	77.952	97.808	3978	5202	6245	7160
PPMP4	**34.984**	**55.147**	**72.125**	**86.403**	**3992**	**5354**	**6328**	**7433**

Table 28. **Performance Evaluation with Minimum Green Time (Pedestrian Crossing Time)-Average Number of Stops and Stop Delay**

	Average Number of Stops				Average Stop Delay			
Volume	1200	1600	2000	2400	1200	1600	2000	2400
Fix-Time	2.565	2.82	4.187	9.576	29.451	49.102	62.595	86.565
Actuated	2.22	2.851	3.965	8.482	28.527	45.441	55.228	84.473
Max-Pressure	3.902	5.442	7.006	12.017	36.319	52.172	63.452	87.326
PPMP2	**1.518**	**1.885**	**2.247**	**2.815**	**27.145**	**41.987**	**47.879**	**54.085**
PPMP4	1.619	1.988	2.409	3.396	28.458	43.39	49.504	59.721

6.3 Implementation of Mixed Intersections

In the real world, not all intersections on an arterial can be equipped with PPMP controllers due to budget, local geometric conditions etc. This section demonstrates the implementation of the PPMP under two scenarios that only half of the intersections are equipped with PPMP controllers. For each scenario, three out of six intersections on Texas Avenue are randomly selected to be equipped with PPMP controllers. In the first scenario denoted as PPMP-semi1 in Table 29, intersections 1, 3, and 5 are selected, while in the second scenario denoted as PPMP-semi2, intersections 2, 3 and 6 are selected. The simulation results show that PPMP with mixed intersection control is still able to improve the average delay about 15%, while increasing the other measures of effectiveness to a satisfactory level as well.

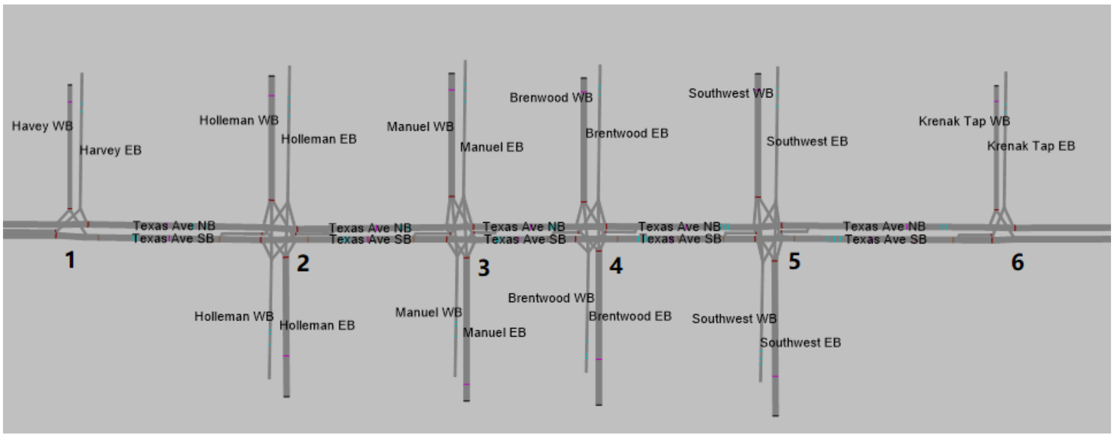


Figure 27. **Mixed Control for Six Intersections on Texas Avenue**

Table 29. Performance after Implementation of Mixed Control on Texas Avenue

	Average Delay	Throughput	Average Number of Stops	Average Stop Delay
Fix-Time	70.458	4694	2.588	49.62
Actuated	69.404	4811	2.455	47.986
Max-Pressure	71.578	4782	5.889	52.774
PPMP2	57.862	4969	2.004	41.055
PPMP2-semi1	65.089	4853	**2.259**	**44.155**
PPMP2-semi2	66.511	4845	**2.264**	**44.882**
PPMP4	56.147	5113	2.142	43.004
PPMP4-semi1	**60.987**	**4890**	2.405	47.094
PPMP4-semi2	**61.124**	**4883**	2.412	46.971

6.4 Time-Varying Traffic

The PPMP's practicality in the real world is further tested by applying time-varying traffic as input. There are cases when the traffic volume suddenly increases to a very high level due to some events which may cause temporary failures to traffic signals. This situation is worth to be studied and should be taken good care of. The time-varying volume input based on a 600-second (10 minutes) interval is presented in Figure 29. The simulation results are presented in Table 30. It can be concluded from the simulation results that the PPMP has great potential in handling time-varying traffic, by greatly improving all four measures of effectiveness in the one-hour period.

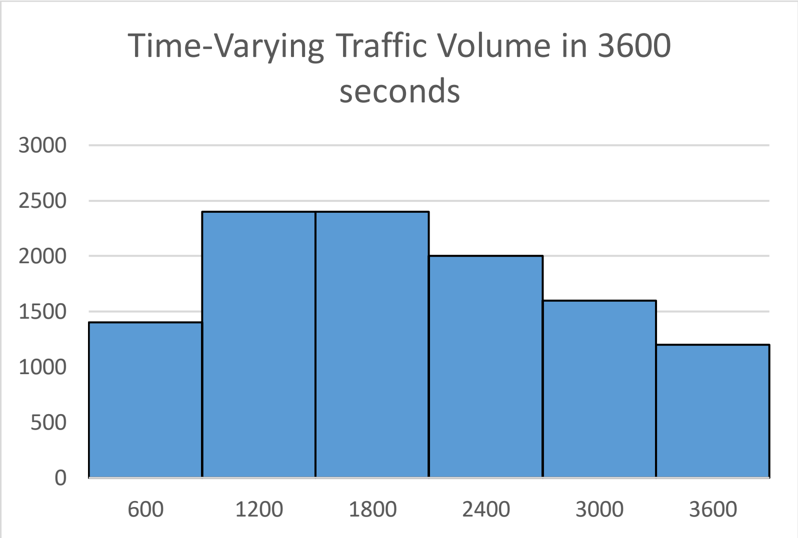


Figure 28. **Traffic Volume Changes in One-Hour Simulation Period**

Table 30. **Performance Comparison under Time-Varying Traffic Volume**

	Average Delay	Throughput	Average Number of Stops	Average Stop Delay
Fix-Time	85.08	5241	4.88	61.412
Actuated	79.54	5310	4.26	56.453
Max-Pressure	95.44	5107	6.84	59.443
PPMP2	60.47	5742	**2.019**	**37.281**
PPMP4	**57.56**	**5899**	2.196	42.054

7. CONCLUSIONS

7.1 Contributions

By performing real data collection, data analysis, model construction, and simulation, this research made the following contributions:

- Built platoon speed distribution based on real data combined with simulated data
- Investigated the impact of influencing factors on platoon speeds and platoon dispersion, for both homogeneous and heterogeneous traffic flow
- Constructed mathematical relationship for distribution parameters and variable factors that have an influence on the parameters
- Implemented the speed distribution parameter model to improve platoon dispersion model and better describe the actual platoon dispersion characteristics
- Developed signal timing strategies-the Predicted Platoon Max-Pressure Policy (PPMP) that can handle platoons with changing speed distribution patterns and control arterial signals in real time
- Evaluated the applicability of the proposed signal timing strategies in various network conditions and proved the PPMP has great practical potential in the real world.

7.2 Summary of Key Findings

This research has to major findings. First, by collecting real world vehicle speed data and performing various statistical analysis on these data, the distribution of the speeds is studied and key factors that influence the speed distribution parameters are identified. A multivariate model is constructed between the parameters and the influencing factors. The evolution of the speed distribution and its parameters, can be estimated based on the multivariate distribution parameter model. The model shows strong ability to predict downstream platoon speeds. A platoon dispersion model is modified based on the multivariate model to predict platoon densities and the vehicle in front and rear of the platoon. The modified dispersion model is proved to outperform the traditional Robertson's model. Second, the aforementioned model is implemented in signal timing strategy development. The Predicted Platoon Max-Pressure Policy (PPMP) that utilizes the previous results is developed. Simulation analysis show that the PPMP could outperform the traditional signal timing methods to a large extent and has great practicality under various real-world scenarios.

7.3 Future Research

Although the proposed multivariate distribution parameter model and the Predicted Platoon Max-Pressure Policy showed it great potential in application in the real world, there is more interesting aspects to be further studied. First, obtaining more real-world data under various scenarios will greatly help the future research and improve the model even better. The current data were collected on Texas Avenue and by simulation. Many of the conditions on Texas Avenues are fixed and there may be some influencing factors that could not be analyzed, such as the side friction. More real-world data may also help to further validate the proposed model. Second, the application of PPMP is worthy to tested under more scenarios if it is desired for a direct implementation. In addition, the adaption of connected vehicles may have an interesting impact on the topic. Connected vehicles are one of the most trending research topics. Traffic with mixed vehicles-connected and regular vehicles, may have different patterns and the platoon characteristics may change. It is desirable to study these scenarios and modify the proposed models accordingly. Last but not least, a preliminary analysis on speed distribution was performed for heterogeneous traffic flow with trucks. Due to the lack of data, the study on truck involved traffic is not thorough and in-depth. If enough truck involved data are present, it is desirable to construct mathematical models and develop signal timing plans for heterogeneous traffic flow.

REFERENCES

1. Lighthill, M. J., & Whitham, G. B. (1955). On Kinematic Waves II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *229*(1178), 317-345.

2. Robertson, D. I. (1969). TRANSYT-A Traffic Network Study Tool. Report 253. *London: TRRL.*

3. Pacey, G. M. (1956). The progress of a bunch of vehicles released from a traffic signal. *Road Research Laboratory Note RN/2665/GMP*, *1956*(1956), 36-50.
4. Hall, M., & Willumsen, L. (1980). SATURN-A Simulation-Assignment Model for the Evaluation of Traffic Management Schemes. *Traffic Engineering & Control*, *21*(4).

5. Lieberman, E., & Andrews, B. TRAFLO (1980). A New Tool to Evaluate Transportation Management Strategies, *Transportation Research Record 772, TRB.* Washington DC: National Research Council, 1980: 9-15.

6. Hunt, P. B., Robertson, D. I., Bretherton, R. D., & Winton, R. I. (1981). SCOOT-A Traffic Responsive Method of Coordinating Signals. *TRRL Report LR, 1014*, l93l.

7. Seddon, P. A. (1972). Another Look at Platoon Dispersion: 3. The Recurrence Relationship. *Traffic Engineering and Control*, *13*(10), 442-444.

8. Tracz, M. (1975). The Prediction of Platoon Dispersion Based on Rectangular Distribution of Journey Time. *Traffic Engineering & Control*, *16*(11).

9. Polus, A. (1979). A Study of Travel Time and Reliability on Arterial Routes. *Transportation*, *8*(2), 141-151.

10. Grace, M. J., & Potts, R. B. (1964). A Theory of The Diffusion of Traffic Platoons. *Operations Research*, *12*(2), 255-275.

11. Yu, L. (2000). Calibration of Platoon Dispersion Parameters on The Basis of Link Travel Time Statistics. *Transportation Research Record*, *1727*(1), 89-94.

12. Farzaneh, M. (2005). Modeling Traffic Dispersion (Doctoral dissertation, Virginia Tech).

13. Bie, Y., Liu, Z., Ma, D., & Wang, D. (2012). Calibration of Platoon Dispersion Parameter Considering the Impact of The Number of Lanes. Journal of Transportation Engineering, 139(2), 200-207.

14. Bie, Y., Liu, Z., Li, Y., & Pei, Y. (2014). Development of a New Platoon Dispersion Model Considering Turning Vehicles in Urban Road Environment. *Advances in Mechanical Engineering*, *6*, 279570.

15. Day, C. M., & Bullock, D. M. (2012). Calibration of Platoon Dispersion Model with High-Resolution Signal Event Data. *Transportation Research Record*, *2311*(1), 16-28.

16. Wei, M., Jin, W., & Shen, L. (2012). A Platoon Dispersion Model Based on A Truncated Normal Distribution of Speed. Journal of Applied Mathematics, 2012.

17. Badhrudeen, M., Thomas, H., Vanajakshi, L. D., & Sharma, A. (2017). Platoon Dispersion Analysis Based on Diffusion Theory. In *MATEC Web of Conferences* (Vol. 124, p. 01003). EDP Sciences.

18. Gaur, A., & Mirchandani, P. (2001). Method for Real-Time Recognition of Vehicle Platoons. *Transportation research record*, *1748*(1), 8-17.

19. Chaudhary, N. A., Abbas, M. M., & Charara, H. A. (2006). Development and Field Testing of Platoon Identification and Accommodation System. Transportation research record, 1978(1), 141-148.

20. Chaudhary, N. A., Charara, H. A., Longmire, R. R., & Johnson, J. (2009). Platoon Identification and Accommodation System Implementation in Brownwood and Caldwell, Texas (No. FHWA/TX-10/5-5507-01-1). Texas Transportation Institute.

21. Sudheer, S., Thomas, H., Sharma, A., & Vanajakshi, L. D. (2016, November). Automated Techniques for Real Time Platoon Detection and Identification. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1820-1825). IEEE.

22. Praveen, P. S., & Ashalatha, R. (2018). Identification of Platoon Dispersion Pattern under Heterogeneous Traffic Conditions. Case Studies on Transport Policy.

23. Li, B. (2017). Stochastic Modeling for Vehicle Platoons (I): Dynamic Grouping Behavior and Online Platoon Recognition. *Transportation Research Part B:*

Methodological, *95*, 364-377.

24. Li, B. (2017). Stochastic Modeling for Vehicle Platoons (II): Statistical Characteristics. *Transportation Research Part B: Methodological*, *95,* 378-393.

25. Shen, L., Liu, R., Yao, Z., Wu, W., & Yang, H. (2018). Development of Dynamic Platoon Dispersion Models for Predictive Traffic Signal Control. *IEEE Transactions on Intelligent Transportation Systems*, (99), 1-10.

26. Hao, H., & Barooah, P. (2013). Stability And Robustness of Large Platoons of Vehicles with Double ‐ Integrator Models and Nearest Neighbor Interaction. *International Journal of Robust and Nonlinear Control*, *23*(18), 2097-2122.

27. Lu S., & Liu. X. (2007). Platoon Dispersion Prediction under the Condition of Adjacent Cycle Traffic Flow Overlapping Based on Support Vector Regression. In *2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007)* (pp. 918-921). IEEE.

28. Forde, A., & Daniel, J. (2017). Performance Evaluation of The HCM 2010 Platoon Dispersion Model under Midblock Pedestrian and Truck Traffic Friction Conditions. *Journal of traffic and transportation engineering (English edition)*, *4*(6), 522-534.

29. Maher, M. (2011). A Comparison of The Use of The Cell Transmission and Platoon Dispersion Models in TRANSYT 13. *Transportation Planning and Technology*, *34*(1), 71-85.

30. Shin, M. S., Ran, B., He, R. R., & Choi, K. (2000). Introducing Platoon Dispersion into An Analytical Dynamic Assignment Process. *Transportation research record*, *1733*(1), 96-104.

31. Dey, P. P., Chandra, S., & Gangopadhaya, S. (2006). Speed Distribution Curves under Mixed Traffic Conditions. *Journal of transportation engineering*, *132*(6), 475-481.

32. Park, B. J., Zhang, Y., & Lord, D. (2010). Bayesian Mixture Modeling Approach to Account for Heterogeneity in Speed Data. *Transportation research part B: methodological*, *44*(5), 662-673.

33. Ko, J., & Guensler, R. L. (2005, January). Characterization of Congestion Based on Speed Distribution: A Statistical Approach Using Gaussian Mixture Model. In *Transportation Research Board Annual Meeting*.

34. Wu, W., Jin, W., & Shen, L. (2013). Mixed Platoon Flow Dispersion Model Based on Speed-Truncated Gaussian Mixture Distribution. *Journal of Applied Mathematics*, *2013*.

35. Wu, W., Shen, L., Jin, W., & Liu, R. (2015). Density-Based Mixed Platoon Dispersion Modelling with Truncated Mixed Gaussian Distribution of Speed. *Transportmetrica B: Transport Dynamics*, *3*(2), 114-130.

36. Jiang, Y., Yao, Z., Luo, X., Wu, W., Ding, X., & Khattak, A. (2016). Heterogeneous Platoon Flow Dispersion Model Based on Truncated Mixed Simplified Phase‐Type Distribution of Travel Speed. *Journal of Advanced Transportation*, *50*(8), 2160-2173.

37. Ye, F., & Zhang, Y. (2009). Vehicle Type–Specific Headway Analysis Using Freeway Traffic Data. *Transportation research record*, *2124*(1), 222-230.

38. Weng, J., Meng, Q., & Yan, X. (2014). Analysis of Work Zone Rear-End Crash Risk for Different Vehicle-Following Patterns. *Accident Analysis & Prevention*, *72*, 449-457.

39. Weng, J., & Meng, Q. (2011). Modeling Speed-Flow Relationship and Merging Behavior in Work Zone Merging Areas. *Transportation research part C: emerging technologies*, *19*(6), 985-996.

40. Meng, Q., Weng, J., & Qu, X. (2010). A Probabilistic Quantitative Risk Assessment Model for the Long-Term Work Zone Crashes. *Accident Analysis & Prevention*, *42*(6), 1866-1877.

41. Brooks, W. D. (1964). Vehicular Traffic Control–Designing Arterial Progressions Using A Digital Computer. *Data Processing IBM*.

42. Messer, C. J., Whitson, R. H., Dudek, C. L., & Romano, E. J. (1973). A variable-sequence multiphase progression optimization program. *Highway Research Record* No. 445, 24-33.

43. Little, J. D., Kelson, M. D., & Gartner, N. H. (1981). MAXBAND: A Versatile Program for Setting Signals on Arteries and Triangular Networks.

44. Gartner, N. H., Assman, S. F., Lasaga, F., & Hou, D. L. (1991). A Multi-Band Approach to Arterial Traffic Signal Optimization. *Transportation Research Part B: Methodological*, *25*(1), 55-74.

45. Chaudhary, N. A., Pinnoi, A., & Messer, C. J. (1991). Scheme to Optimize Circular Phasing Sequences. *Transportation Research Record*. 1324, 72-92.

46. Lieberman, E. B., Lai, J., & Ellington, R. E. (1983). SIGOP-III User's Manual. Federal Highway Administration, Office of Implementation.

47. Wallace, C. E., Courage, K. G., Reaves, D. P., Schoene, G. W., & Euler, G. W. (1984). TRANSYT-7F User's Manual (No. UF-TRC-U32 FP-06/07).

48. Hunt, P. B., Robertson, D. I., Bretherton, R. D., & Royle, M. C. (1982). The SCOOT On-Line Traffic Signal Optimization Technique. *Traffic Engineering & Control*, *23*(4).

49. Mauro, V., & Di Taranto, C. (1990). Utopia. *IFAC Proceedings Volumes*, *23*(2), 245-252.

50. Skabardonis, A., & May, A. D. (1985). Comparative Analysis of Computer Models for Arterial Signal Timing. *Transportation Research Record*. 1021, 45-52.

51. Liu, C. C. (1988). Bandwidth-Constrained Delay Optimization for Signal Systems. *ITE Journal*, *58*(12), 21-26.

52. Lan, C.J., Messer, C.J., Chaudhary, N.A., & Chang, E.C.P. (1992). Compromise Approach to Optimize Traffic Signal Coordination Problems during Unsaturated Conditions. *Transportation Research Record*, 1360, 112–120

53. Husch, D., & Albeck, J. (2007). Synchro Studio 7 User Manual. Trafficware, (50), 100.

54. Zhou, H., & Hawkins Jr, H. G. (2014). Arterial Signal Coordination with Uneven Double Cycling (SWUTC/15/600451-00024-1). Texas A&M Transportation Institute

55. Zhou, H., Hawkins Jr, H. G., & Zhang, Y. (2017). Arterial Signal Coordination with Uneven Double Cycling. *Transportation Research Part A: Policy and Practice, 103,* 409-429.

56. Zeng, X., Sun, X., Zhang, Y., and Quadrifoglio, L. (2015). Person-Based Adaptive Priority Signal Control with Connected-Vehicle Information. In: *Transportation Research Record*. 2487:78-87.

57. Wei, M., & Sun, B. (2017). An Arterial Green-Wave Synchronous Coordination Model for Bus and Non-Bus Lanes Based on Platoon Dispersion Theory. *Wireless Personal Communications*, *95*(3), 2091-2109.

58. Mei, Z., Tan, Z., Zhang, W., & Wang, D. (2019). Simulation Analysis of Traffic Signal Control and Transit Signal Priority Strategies under Arterial Coordination Conditions. *Simulation*, *95*(1), 51-64.

59. Ma, W., Zou, L., An, K., Gartner, N. H., & Wang, M. (2018). A Partition-Enabled Multi-Mode Band Approach to Arterial Traffic Signal Optimization. *IEEE Transactions on Intelligent Transportation Systems*, (99), 1-10.

60. Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1-22.

61. Benaglia, T., Chauveau, D., & Hunter, D. R. (2009). An EM-Like Algorithm for Semi-and Nonparametric Estimation in Multivariate Mixtures. *Journal of Computational and Graphical Statistics*, *18*(2), 505-526.

62. Hall, P., Neeman, A., Pakyari, R., & Elmore, R. (2005). Nonparametric Inference in Multivariate Mixtures. *Biometrika, 92*(3), 667-678.

63. Tassiulas, L., and Ephremides, A. Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks. In IEEE Wongpiromsarn, T. Transactions on Automatic Control 37, 12: 1936-1948, 1992

64. Varaiya, P. Max Pressure Control of a Network of Signalized Intersections. In Transportation Research Part C: Emerging Technologies. 36: 177-195, 2013.

65. Wongpiromsarn, T., Uthaicharoenpong, T., Wang, Y., Frazzoli, E., and Wang, D. Distributed Traffic Signal Control for Maximum Network Throughput. In *Proc. Of IEEE Conference on Intelligent Transportation Systems*, 2012.

66. Xiao, N., Frazzoli, E., Li, Y., Wang, Y., and Wang, D. Pressure Releasing Policy in Traffic Signal Control with Finite Queue Capacities. In *Proc. of IEEE Conference on Decision and Control*, 2014.

67. Gregoire, J., Frazzoli, E., De La Fortelle, A., and Wongpiromsarn, T. Back-Pressure Traffic Signal Control with Unknown Routing Rates. In *IFAC Proceedings* 47, 3:11332-11337, 2014.

68. Lioris, J., Kurzhanskiy, A., and Varaiya, P. Adaptive Max Pressure Control of Network of Signalized Intersections. In *IFAC-PapersOnline*, 49, 22: 19-24, 2016.

69. Anderson, L., Pumir, T., Triantafyllos, D., and Bayen, A. M. Stability and Implementation of a Cycle-Based Max Pressure Controller for Signalized Traffic Networks. In Networks and Heterogeneous Media, 13, 2: 241-260, 2018.

70. Hsieh, P. –C., Liu, X., Jiao, J., Hou, I., Zhang, Y., and Kumar, P. R. Throughput-Optimal Scheduling for Multi-Hop Networked Transportation Systems with Switch-Over Delay. In *Proc. of the 18th ACM International Symposium on Mobile ad Hoc Networking and Computing*, 2017.