11-16-2019

Traffic Density on Corridors Subject to Incidents: Models for Long-Term Congestion Management

Pedro Cesar Lopes Gerum
Cleveland State University, p.lopesgerum@csuohio.edu

Andrew Reed Benton

Melike Baykal-Gürsoy

Publisher's Statement

This version of the article has been accepted for publication, after peer review and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s13676-019-00149-2

Noname manuscript No.
(will be inserted by the editor)

Traffic Density on Corridors Subject to Incidents: Models for Long-Term Congestion Management

Pedro Cesar Lopes Gerum · Andrew Reed Benton · Melike Baykal-Gürsoy

the date of receipt and acceptance should be inserted later

Abstract The purpose of this research is to provide a faster and more efficient method to determine traffic density behavior for long-term congestion management using minimal statistical information. Applications include road work, road improvements, and route choice. To this end, this paper adapts and generalizes two analytical models (for non-peak and peak hours) for the probability mass function of traffic density for a major highway. It then validates the model against real data. The studied corridor has a total of 36 sensors, 18 in each direction, and the traffic experiences randomly occurring service deterioration due to accidents and inclement weather such as snow and thunderstorms. We base the models on queuing theory, and we compare the fundamental diagram with the data. This paper supports the validity of the models for each traffic condition under certain assumptions on the distributional properties of the associated random parameters. It discusses why these assumptions are needed and how they are determined. Furthermore, once the models are validated, different scenarios are presented to demonstrate traffic congestion behavior under various deterioration levels, as well as the estimation of traffic breakdown. These models, which

P. C. Lopes Gerum

Industrial and Systems Engineering, Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854

E-mail: pedro.gerum@rutgers.edu

A. R. Benton

Industrial and Systems Engineering, Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854

E-mail: andrew.benton@rutgers.edu

M. Baykal-Gürsoy

Industrial and Systems Engineering, CAIT and RUTCOR, Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854

E-mail: gursoy@soe.rutgers.edu

account for non-recurrent congestion, can improve decision-making without the need for extensive datasets or time-consuming simulations.

Keywords Random Queues · Traffic Density · Recurrent Congestion · Non-Recurrent Congestion · Traffic Breakdown

1 Introduction

In recent times, population growth, economic growth, and lifestyle changes have increased the demand for traffic infrastructure at an unprecedented rate, outpacing improvements to infrastructure. In turn, the rate of congestions and delays has increased, directly affecting millions of people worldwide. Even the introduction of autonomous vehicles may not curb congestion due to high demand unless policies are developed to encourage ridesharing or improve flow management. While traffic congestion and breakdown caused by excess cars during peak hours are common, other types of congestion have started to become more pervasive. This type of non-recurrent congestion and breakdown is a significant contributor to the total delay of vehicle travel time [Skabardonis et al., 2003, Kwon et al., 2006]. The two main factors that cause non-recurrent delays are weather deterioration and accidents that affect the capacity of the road. They account for well over half of the non-recurrent delays in urban areas, and nearly all non-recurrent delays in rural areas [Skabardonis et al., 1998].

This problem becomes increasingly ubiquitous as the worldwide population grows and concentrates [Thakur et al., 2012]. Unfortunately, transportation infrastructure investments are still lagging, putting pressure on those responsible for planning transportation systems [Schrank et al., 2015]. Severe travel delays are frequently the result of the inappropriate planning of transportation systems [Chiou, 2016]. Poor planning may cause insufficient provision of link capacity, particularly under uncertain travel demand and with the arrivals of non-recurrent incidents. Such issues need to be accounted for in roadway designs, even when it is infeasible to gather the appropriate data. Therefore, transportation investments must be carefully considered to help alleviate congestion and prevent future traffic problems.

This paper aims to address the critical need for analytical congestion assessment methods that are neither overly simplistic nor excessively complex. The proposed method uses simple, usually on hand parameters to describe the probability mass function of the traffic density analytically, while still accounting for non-recurrent incidents. The parameters are: a) mean traffic flow under normal and deteriorated con-

dition, b) mean speed under normal and deteriorated conditions, c) incident frequency, and d) clearance time, to estimate the probability distribution of traffic density.

Based on the previous literature [Baykal-Gürsoy and Xiao, 2004, Baykal-Gürsoy et al., 2009a,b], we extend their model and validate their approach against a real dataset. Although they derive the distribution of density in closed-form, its computation is nontrivial. Few properties of the distribution can be directly derived from its complex closed-form, limiting its application. However, in realistic scenarios, we can place reasonable bounds on the parameters, leading to a simpler closed-form distribution. The properties of the resulting distribution can then be directly computed and have an intuitive meaning, allowing decision-makers to attain a better understanding of the behavior of the system. Moreover, Baykal-Gürsoy et al. [2009b] validate the models against simulated data using Paramics®, and INTEGRATION; not against real data. This paper also addresses this gap by validating the model against data obtained from a Wisconsin highway traffic dataset.

This mathematical framework can answer questions such as *1. How much would a change in traffic behavior impact traffic congestion? 2. How can we estimate the probability of traffic flow breakdown on roadways when minimal data is available?* Moreover, this model explains how the probability of traffic breakdown changes according to the changes in traffic parameters, and how improvements on the road clearance time impact congestion. Tables 4 and 5 provide a few examples of how we can derive the probability from this model. This method, therefore, simplifies and improves the assessment of non-recurrent traffic density and its variability, leading to a reduction in congestion.

Despite some prior work present in literature, to the best of our knowledge, there is not an agreed-upon steady-state probability distribution for traffic density subject to non-recurrent congestion. Early work established in the literature largely depends on extensive simulation [Muñoz et al., 2003, Daganzo, 2006]. Recent research that includes non-recurrent congestion focuses on detecting incidents rather than computing the steady-state distribution [Anbaroglu et al., 2014, Chen et al., 2016, Laharotte et al., 2017]. Mathematically, they work on the transient portion of congestion behavior, thus focusing on real-time applications instead of planning and long-term decision-making. Hence, there is a gap in the literature for direct approaches to computing the steady-state distribution of traffic density that do not require simulation or extensive datasets and account for recurrent and non-recurrent congestion.

In section 2, we discuss current approaches to describe the behavior of traffic density from literature. In section 3, we present the data used in the validation of the traffic density distribution models. In section 4, we introduce an adaptation and an expansion of an existing model that demands minimal statistical

information that is usually known or can be easily approximated. The adaptation of the model provides a simple closed-form alternative for most scenarios during non-peak hours. The expansion provides solutions for a general case that includes recurrent and non-recurrent deterioration for both peak and non-peak hours. We perform a validation using data in section 5 to demonstrate the modeling performance. It uses the initial assumption that only the aggregate parameters, and not the data were initially known, then compares the predicted results with the actual data. Suggested applications are discussed in section 6.

2 Background

2.1 Congestion Forecasting

Several authors have derived solutions for the steady-state congestion problem. However, most current approaches face at least one of several shortcomings:

1. they require large datasets — and therefore are expensive or infeasible in most situations;
2. they lack a closed-form solution — and therefore require extensive computation;
3. they provide expected results, instead of full distributions — and therefore do not provide the full picture of the density behavior;
4. they do not account for non-recurrent congestion — and therefore lack robustness.

The traditional methodology for traffic modeling was introduced by Lighthill and Whitham [1955] and Richards [1956], which approximates traffic as a deterministic fluid governed by a conservation equation relating the flow, speed, and occupancy, the so-called the kinematic wave equation. While these initial models were powerful for modeling the emergent behavior seen in real traffic, they were mathematically cumbersome. Later models, like Newell [1993], made modifications that can accurately model traffic density with fewer technical complexities. However, these models are only capable of providing the expected flow parameters, not their distributions. Daganzo [1994] introduces the Cell Transmission Model (CTM), a numerical method to solve the kinematic wave equation. He also demonstrates that the CTM can analyze non-recurrent incidents behavior on a scenario-by-scenario basis by temporarily altering initial conditions or parameters. The Switching-Mode Model, proposed by Muñoz et al. [2003], describes an extension of the CTM with time-varying parameters. Their approach consists of multiple sets of CTM update parameters, each representing different congestion levels. These are then switched between, to model the effect of changing traffic conditions. The usefulness of this multiple-scenario approach is well demonstrated by the authors, but it does not provide the distributional information of the traffic density.

An alternative approach divides traffic into much smaller sub-units, usually at the scale of a single vehicle. Car-following models, like the Intelligent Driver Model proposed by Treiber et al. [2000], require extensive knowledge of driver characteristics. These data are often costly or infeasible to gather, preventing the modeling of large or multiple studies. This shortcoming is particularly relevant in the planning of new areas where infrastructure does not exist. Cellular Automata models like Nagel and Schreckenberg [1992] do not have the same data requirements, and Daganzo [2006] shows that they converge to the solutions of the original formulation by Lighthill and Whitham [1955]. The Cellular Automata models, as well as the CTM proposed by Daganzo [1994], can be stochastic, allowing traffic engineers to characterize the full distribution. Doing so, however, would require extensive simulation.

Part of the literature focuses on finding the probability of traffic breakdown. Traffic breakdown is triggered when a substantial speed decrease from the free flow speed is detected between two consecutive time intervals. This speed decrease drastically increases density, hence causing a sudden plunge in capacity. Kerner et al. [2002] adapt the original Cellular Automaton model to derive a theoretical probability for a spontaneous breakdown. With the further popularization of Cellular Automata models by Maerivoet and De Moor [2005], the concept of using simulation became a constant for the problem of finding the probability of traffic breakdown. With time, other variations have surfaced, such as the Monte-Carlo simulation model proposed by Dong and Mahmassani [2012]. Its novelty was the combination of a stochastic approach to macroscopic flow breakdown with a microscopic model of driver behaviors. However, these models face similar shortcomings — they require extensive computational resources and complex detailed data. Hence, probabilistic closed-form solutions for this problem have recently resurfaced in the literature [Arnesen and Hjelkrem, 2018, Han and Ahn, 2018]. Even then, the solutions proposed are limited for planning purposes, as they provide little direction for the decision-maker and still require complex data information. On the other hand, a full density distribution both allows for direct computation of the probability of traffic breakdown. It equips decision-makers with additional valuable information on the significance of different traffic parameters for breakdown and congestion.

Lately, authors have started employing new technologies to collect data for traffic density estimation. Zhang et al. [2017] develop a model to analyze traffic congestion through image processing or computer vision; Panichpapiboon and Leakkaw [2017] propose using cars as mobile sensors to generate data. This influx of data has been a boon for recent research in traffic congestion for dynamic systems, enabling machine learning approaches to arise. These models divide roadways into segments observed in discrete time units and either forecast traffic flow patterns [Celikoglu, 2014, Celikoglu and Silgu, 2016], or directly forecast the

traffic state [Polson and Sokolov, 2017] for some time horizon. Although most employ common variations of traditional traffic metrics (flow and speed), there is still no consensus in the literature regarding the most appropriate input for ML-based classifiers of traffic state. In addition, to the best of our knowledge, no literature has implemented machine-learning-based predictive models that directly use density as the output parameter. While these methods are promising for the future of traffic modeling, they are not well suited for planning applications. Typical solutions provide only expected values, and their accuracy decreases as the forecast horizon increases. While these advantages are compelling for real-time traffic monitoring, they do not transfer over to planning applications where these datasets are unavailable, and long-term distributional predictions are needed.

Finally, few papers have addressed non-recurrent incidents in their analysis of long-term behavior, even though they are a significant contributor to the total delay of vehicle travel time and traffic breakdown [Skabardonis et al., 2003, Kwon et al., 2006]. Baykal-Gürsoy and Xiao [2004] and Baykal-Gürsoy et al. [2009a,b] propose a model using queuing theory that accounts for non-recurrent congestion. They depict a segment of a roadway as two-state finite or infinite queues. The next section introduces stochastic queue-based traffic models.

2.2 Stochastic Queuing Models

A theoretical framework used by some authors to predict congestion is queuing theory. Queuing analyses, together with deterministic (fluid-dynamics) models [May and Keller, 1967, Newell, 1971], are primarily used for performance evaluation purposes and the synchronization of traffic lights [Newell, 1965]. Early stochastic models assumed individual vehicle arrivals to follow a Poisson process [Zheng and Liu, 2017, Wang and Ahmed, 2017, Gazis, 2006, J.N. Darroch and Morris, 1964, Tanner, 1953], or as platoons of vehicles [Daganzo, 1994, Alfa and Neuts, 1995, Dunne, 1967, Lehoczky, 1972] to represent the behavior of cars moving between traffic signals.

Cheah and Smith [1994] and Jain and Smith [1997] studied stochastic queues to explore the usefulness of finite server queuing models with state-dependent travel speed for modeling both pedestrian and vehicle traffic flows. In this model, vehicles arrive according to a Poisson process, and the total time to traverse the corridor is assumed to follow a general distribution. If the roadway is at capacity, new arrivals must take an alternative path. Consider vehicles traveling on a corridor, as depicted in Figure 1. The space occupied by one individual vehicle represents a moving server. The service cycle initiates when a car arrives at the

corridor, and service (the act of traveling) is provided until the vehicle leaves the corridor. A server in this context is the moving vehicle-space, including the safe distance (space headway) to the car in front. The number of available servers is then the maximum number of vehicles that the corridor can physically accommodate. Consequently, when there is no space left for a vehicle to enter, i.e., all servers are full, the vehicle either is allowed to wait in a queue or is denied service [Cheah and Smith, 1994, Jain and Smith, 1997]. Other authors [Heidemann, 1996, Vandaele et al., 2000, Heidemann, 2001] studied a similar system but with a single server and infinite queuing capacity. Such a flow model is considered as a vertical queue that disregards the interdependence between vehicles within the same cell [Daganzo, 1994]. Although some of these models focus on congestion, none of them include the occurrence of non-recurrent incidents as part of the model.

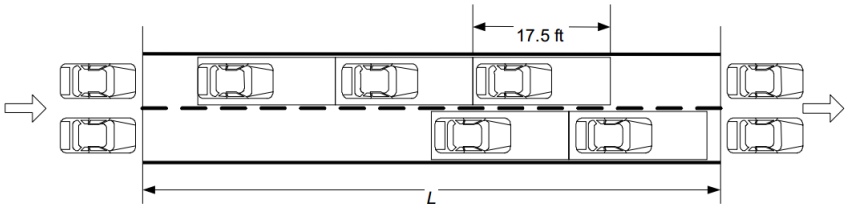

Fig. 1: Graphical depiction of a two-lane roadway segment.

The queuing model proposed by Baykal-Gürsoy and Xiao [2004] and Baykal-Gürsoy et al. [2009a,b] is the only one that considers non-recurrent congestion.

3 Measuring Traffic Data

In this section, we provide one example of how one may assess the congestion problem with real data. Data were collected from Wisconsin via 36 sensors in a 9-mile stretch, 18 on each side of the road (South-East and West-North Directions). Those sensors are inductive loop detectors embedded in the roadway. Figure 2 shows the roadway where data were obtained.

The sensors record three common indicators used in traffic models: speed, flow, and density. Density represents traffic congestion, counting the number of vehicles occupying a particular space unit at a specific time. Flow counts the number of vehicles passing through a certain point per unit time. Density, speed, and flow are closely related, as increasing flow also increases density initially. However, when more cars arrive than the highway can hold (surpassing the maximum flow), density keeps increasing while flow decreases. The limit of the relationship occurs when traffic is down to a complete stop, and flow is thus

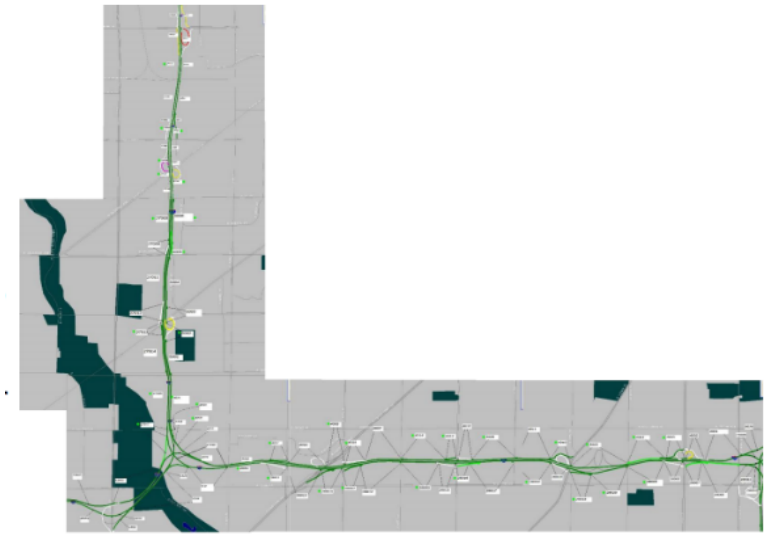

Fig. 2: Map from where data were collected.

zero, hence representing the complete breakdown of traffic [Daganzo, 1994]. Figure 3 represents a partial view of these relationships and showcases Milwaukee's data that this paper uses for validation. Most current data-gathering techniques cannot measure density directly but obtain speed, occupancy, and volume from which traffic density can be calculated [Kurzhanskiy, 2009]. Occupancy is a proxy measure for density. It gives the percent coverage of the sensor per unit time, while volume is a proxy measure for flow, when the period analyzed is divided into equally spaced time intervals. In this section, we describe the methods used to compute density given speed, occupancy, and volume.

The sensors analyzed provide occupancy measures for half-mile segments. In this paper, we consider an average car length plus headway of 22 ft, obtained through the procedure described in Dailey [1999]. We use ordinary least squares to find the best fit for the average car length plus headway. The number of cars, namely density, is then given by a simple ratio of occupancy times the length of the segment by the average length and headway of vehicles. However, occupancy data is generated in increments of 0.002 due to sensors' limitations. As a result, density generated from the data is rounded to specific values. Finally, although sensors also provided information on speed, the speed data was truncated at the speed limit (at 60 mph for some sensors, 65 mph for others). A 'reconstructed speed' data set is then constructed from the volume and occupancy data using a method proposed by Dailey [1999].

As seen in Figure 2, there are numerous merges and forks through which vehicles may leave and arrive. This setting introduces further error to volume and occupancy, therefore making the data harder to analyze. However, for most scenarios, the additional noise does not undermine the validity of the model.

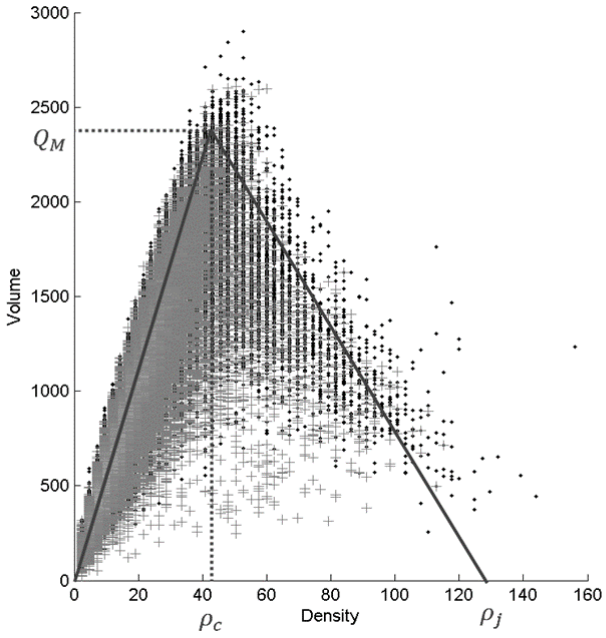

Fig. 3: The Fundamental Diagram with data from one sensor.

The most relevant traffic properties utilized in the validation are the frequency of incident occurrence, f, the duration until clearance and recovery, i.e., the so-called incident clearance time, $\frac{1}{r}$, and the severity of capacity reduction caused by an incident, α. We gather data from accidents and weather conditions to better understand incidents. We then compare the sensors' data with data from weather conditions in order to find relations between incidents or traffic deterioration and extreme weather events such as fog, snowstorms, thunderstorms, rain, or normal conditions. Precipitation data were obtained from the Climate Data Online system of the National Climatic Data Center of the National Oceanic and Atmospheric Administration [NOAA, Accesed in 2015-2016]. These data depict the hourly amount of precipitation in hundredths of inches recorded at the Milwaukee Mitchell International Airport weather station, for the same period as the traffic data. The data also include information on snow days and days with fog and thunderstorms. Our findings show that snowstorms and fog cause the highest impact on travel time.

After comparing the sensor's data with weather data and accident reports, we can split the speed, occupancy, and volume information into two situations: normal conditions (uptime), and adverse conditions (downtime). Downtime represents periods in which an accident or inclement weather condition deteriorates the traffic flow. Furthermore, data are separated into peak hour and non-peak hour and had weekends and late nights removed for accuracy. This separation allows us to analyze the behavior of each group accurately. Lastly, winter month's data is initially chosen to be the validation scenario during non-peak hours, because

extra congestion due to weather during wintertime tends to be predominant and extensive. Mean clearance time and the mean time to incident are calculated using only the times mentioned.

Since the previous assumption could be biased towards long stretches of down periods, i.e., snowfalls that last several days, we also considered the case in which the system may only be affected by incidents, with the occasional heavy rain. In this situation, we analyze the frequency of incident occurrence, as well as the duration until clearance for the summer months during non-peak hours. We show that the model works in both situations.

4 Analytical Model

The main objective of this paper is to analytically describe the probability mass function of the traffic density while accounting for non-recurrent incidents. We also show that simulation and extensive datasets may be unnecessary for long-term planning. In this section, we expand on previous research by creating a closed-form analytical model applicable to any roadway. A few of the assumptions used in the model are discussed and compared to a real-world dataset. The results include a simple approximation for the model proposed by Baykal-Gürsoy and Xiao [2004] and Baykal-Gürsoy et al. [2009b] during non-peak hours and a generalization for peak hours by Baykal-Gürsoy et al. [2009a]. Their models consider a segment of a road operating in a two-state environment process as a Markovian queuing system. The two environment states represent the situation of the roadway, which could be under normal or adverse conditions. The latter refers to incidents such as snowstorms or accidents. They cause a reduction in the road capacity because of closures or blockages, or because drivers tend to slow down and increase the distance to the car in front for increased safety. The time in each environment state is assumed to be random.

4.1 Model Assumptions

Our model relies on the following assumptions:

1. Times to incidents and clearance times are exponentially distributed.
2. Distributions for each segment can be generated independently.
3. Travel times are exponentially distributed, under normal and deteriorated conditions.
4. Arrivals follow a Poisson process, under normal and deteriorated conditions.

Assumption (1) is met. The histograms for the time to incident and clearance times are plotted together with a fitted exponential distribution in Figures 4 and 5 respectively. The distributions are fit with their

maximum likelihood estimators. r^2 values, the proportion of total variation in the outcomes explained by the model, are reported for each fit. They indicate that both times to incident and clearance times can be modeled as exponential random variables. Assumption (1) allows us to model the environment with a Markovian model, for which servers change state according to a continuous-time Markov chain (CTMC).

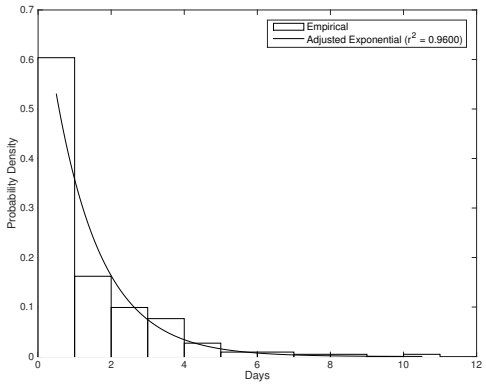

Fig. 4: Time to Incident Distribution.

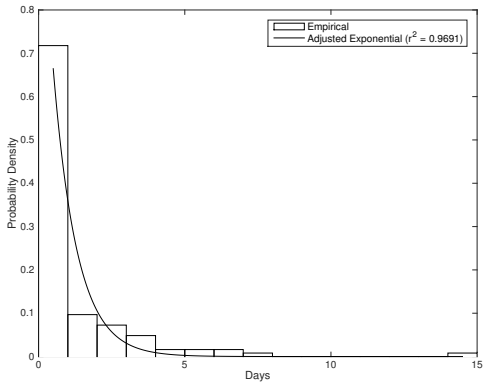

Fig. 5: Clearance Time Distribution.

Assumption (2) is also met. Although propagation effects will likely impact the parameters for upcoming segments, we expect that traffic engineers will use this model for individual segments in which they know or can estimate all the parameters. Section 4.2 lists these parameters, as well as their interpretation and methods of estimation. Capturing the correlation between sensors would only be necessary if the parameters were partially known, as depicted in figure 6. This paper's modeling approach generates parameters from independent datasets for each segment, thus accounting for possible propagation effects. Therefore, distributions can be generated independently despite a possible correlation between different segment's parameters.

Furthermore, the correlations between sensors are obtained indirectly by the model. The time of breakdowns will likely be correlated between segments, as weather events and car accidents will frequently impact adjacent segments. This correlation is captured by the change in arrival and service rates during a breakdown.

Assumption (3) is not met by our datasets, but we argue that this does not affect the final performance of our model. The histograms for travel times under normal and deteriorated conditions are plotted in Figures 7 and 8. They show that these random variables do not follow an exponential distribution. Nonetheless, we assume that these random variables do follow such distribution out of mathematical convenience (Baykal-Gürsoy et al. [2019] show that this assumption does not affect the expected density). We will see in Section 5

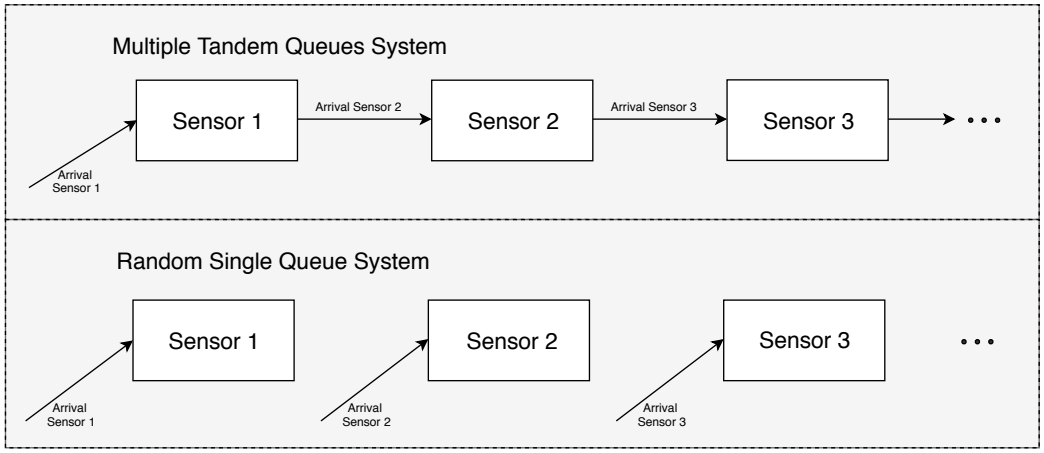

Fig. 6: Arrivals for the system proposed in this model are calculated independently of previous sensors. Every sensor has its arrival inputted separately. Propagation effects only need to be accounted for in Tandem Queue Systems, where the only arrival inputted is the one for Sensor 1.

that the model provides a good approximation for the probability mass function of traffic density, regardless of the validity of this assumption.

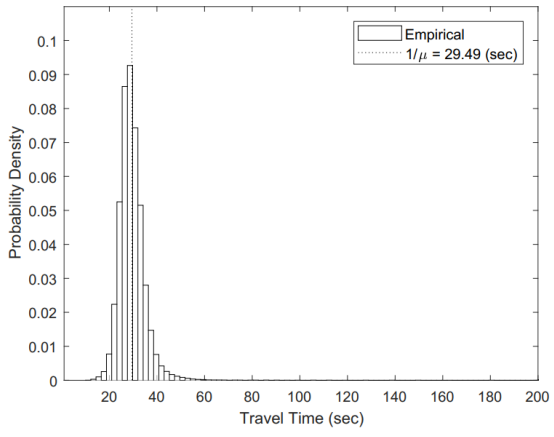

Fig. 7: Travel Time Distribution under normal conditions.

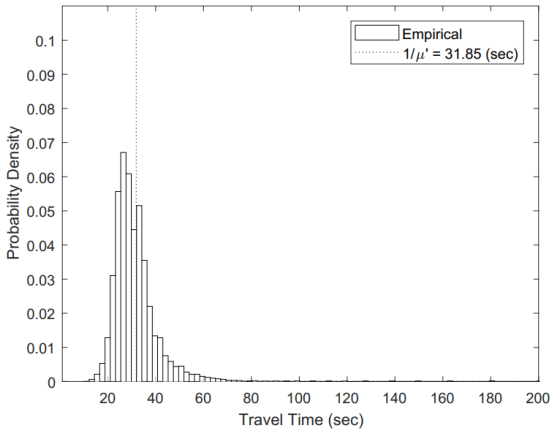

Fig. 8: Travel Time Distribution under adverse conditions - this distribution has a thicker tail than figure 7.

Assumption (4) is also not met by our dataset. However, this assumption is commonly used in literature [J.N. Darroch and Morris, 1964, Gazis, 2006, Zheng and Liu, 2017, Wang and Ahmed, 2017]. Although some methods in literature do not assume this, they introduce further complexity in the model, which then requires simulation for solutions. Numerical experiments indicate that many of these traffic simulations

produce nonhomogenous Poisson arrival processes. An example can be found in figure 9, which contains the histogram of interarrival times at an arbitrary cell in a Cellular Automata model. Similar results can be found for other microscopic traffic models, suggesting that even models which do not explicitly assume that arrivals are Poisson ultimately reproduce a (possibly nonhomogenous) Poisson process.

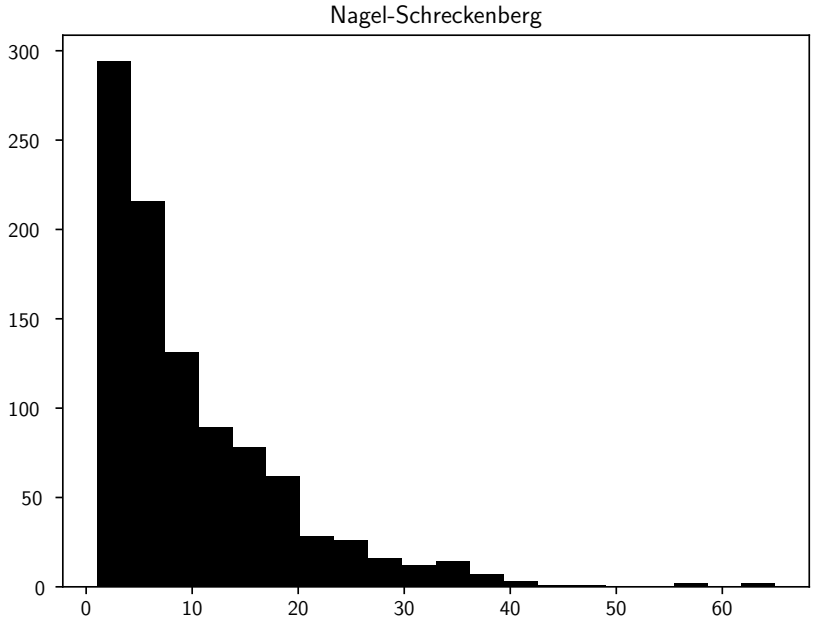

Fig. 9: Histogram of Interarrival Times from Nagel-Schreckenberg Simulation.

4.2 Model Structure and Parameters

4.2.1 Non-peak hours.

In the queue shown in Figure 10, each state of the Markov chain represents both the number of cars in the system and the condition, normal or adverse, of the system. Figure 10 also depicts the definitions of parameters.

Parameters λ and λ' represent the arrival rate of the system in normal and adverse conditions, respectively. It is typically true that $\lambda \geq \lambda'$, although this is not necessary. Under the assumptions of section 4.1, the expected time between arrivals is given by $\frac{1}{\lambda}$ and $\frac{1}{\lambda'}$. This gives a method for estimating λ and λ'

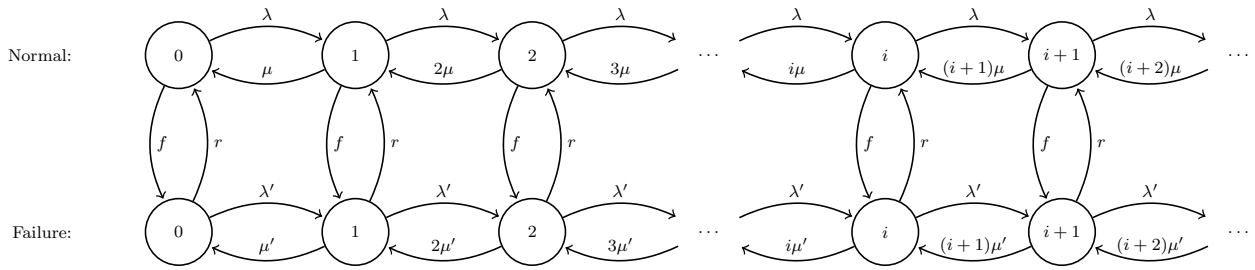

Fig. 10: State Transition Diagram for a Markov-modulated $M/M/\infty$ Queue.

directly from traffic flow data:

$$\lambda = \frac{1}{\mathbb{E}[\text{time between arrivals during normal conditions}]} = \mathbb{E}[\text{flow during normal conditions}],$$

$$\lambda' = \frac{1}{\mathbb{E}[\text{time between arrivals during adverse conditions}]} = \mathbb{E}[\text{flow during adverse conditions}].$$

Parameters μ and μ' represent the service rate of the system in normal and adverse conditions, respectively. The service rate is the instantaneous probability of a car leaving the segment of the road, and therefore, leaving the system. Because the model allocates one server per vehicle, there are infinitely many servers (as long as full capacity is not reached). Therefore, the service rate increases with the number of cars in the systems. This assumption applies to non-peak hours, given the unlikelihood of traffic breakdown during non-peak hours. We remove this assumption when discussing peak hours in section 4.2.2. The more cars there are in the system, the more likely one of them will leave. Thus, if there are n cars on the segment, then the total service rate becomes $n\mu$ during normal conditions, and similarly $n\mu'$ during adverse conditions. Additionally, $\mu > \mu'$ because the service rate must be higher during normal conditions.

The service rate parameters come from the relationship between distance and speed. μ represents the rate of cars crossing a segment when the system is under normal conditions and is given by the ratio of the average speed and the segment length under normal conditions. Analogously, μ' represents the rate of cars crossing a segment when the system is under adverse conditions and is given by the ratio of the average speed and distance under adverse conditions. This interpretation allows us to estimate μ and μ' given the distance between sensors and the reconstructed speed (described in section 3):

$$\frac{1}{\mu} = \frac{\text{segment length}}{\mathbb{E}[\text{speed reconstructed during normal conditions}]} = \mathbb{E}[\text{travel time during normal conditions}],$$

$$\frac{1}{\mu'} = \frac{\text{segment length}}{\mathbb{E}[\text{speed reconstructed during adverse conditions}]} = \mathbb{E}[\text{travel time during adverse conditions}].$$

Parameters f and r represent the incident rate when the system is in normal condition, and the repair rate when the system is in an adverse condition. They are determined from the incident and weather reports by averaging the time until an incident (accident or adverse weather condition) occurs and the clearance time of each incident. The failure rate (the number of times the system goes from normal condition to adverse condition per unit time) is then given by one over the mean time to incidents, or the expected time that the system will remain in normal condition. Analogously, the clearance rate (the number of times the system goes from adverse condition to normal condition per unit time) is then given by one over the mean clearance time. For our dataset, the expected clearance time during summer is around 1 hour. The expected clearance time during winter is around 4.5 hours because snowfalls have a long lingering effect that takes longer to be cleared.

$$f = \frac{1}{\mathbb{E}[\text{time to incident}]} = \frac{1}{\mathbb{E}[\text{duration of normal condition}]},$$
$$r = \frac{1}{\mathbb{E}[\text{clearance time}]} = \frac{1}{\mathbb{E}[\text{duration of adverse condition}]}.$$

Keilson and Servi [1993] were the first to study such queues in a random environment. They derived the generating function of the stationary number of customers in the system in terms of Kummer functions. Baykal-Gürsoy and Xiao [2004] and Baykal-Gürsoy et al. [2009a,b] showed that the generating function reveals that the steady-state number of vehicles in the system is composed of two independent random variables. One represents the number of customers in an uninterrupted queue, and the other represents the customers accumulated during interruptions. In other words, the random number of cars on a corridor, X, is equal to $X = X_\phi + Y$ in steady-state, with X_ϕ representing the random number of vehicles accumulated during the normal condition, and Y representing the additional cars accumulated due to incidents. Moreover, X_ϕ and Y are independent of each other.

Furthermore, the complete distributions of X_ϕ and Y are derived. In equilibrium, X_ϕ follows a Poisson, while $Y = p \cdot Y_1 + (1-p) \cdot Y_2$ follows a mixture of two Poisson random variables Y_1 and Y_2 with random parameters coming from two different truncated Beta distributions, as detailed below:

$$X_\phi \sim \text{Poisson}\left(\frac{\lambda}{\mu}\right), \tag{1}$$

$$Y_1 \sim \text{Poisson}\big(B(a, b, -2\rho^*)\big), \tag{2}$$

$$Y_2 \sim \text{Poisson}\big(B(a+1, b+1, -2\rho^*)\big), \tag{3}$$

$$\text{with: } a = \frac{f}{\mu},\ b = \left(\frac{f}{\mu} + \frac{r}{\mu'}\right),\ \rho^* = \frac{1}{2} \cdot \left(\frac{\lambda}{\mu} - \frac{\lambda'}{\mu'}\right),\ p = \frac{\left(r + \frac{f\mu'}{\mu}\right)}{(r+f)},\ c = -2\rho^*.$$

Once these parameters are determined, the model yields the probability mass function for the traffic density. We compare it with the empirical traffic density obtained from occupancy data to assess how good the fit is per sensor.

As mentioned before, each sensor is treated individually for validation purposes. Nevertheless, the analytical models generated for each sensor produce correlated outputs because the input data is inherently correlated.

Model improvements

The sensor's data yield parameters a and b that are on the order of 10^{-5} for winter months, and a in the order of 10^{-5} and b in the order of 10^{-3} for summer months, meaning that incident and clearance rates, f and r, are considerably lower than μ and μ'. This relationship among these parameters depicts an ordinary situation since the rates of incidents and clearances tend to be notably lower than the rate of cars crossing the segment. This section proposes a simple approximation for this case. The simplified model has the accuracy of the full model but is simpler to calculate.

Another underlying assumption used in this model is that the addition of new cars will not affect the travel time – i.e., that the distribution for travel time is independent of the number of vehicles in the system. For non-peak hours, the primary source of congestion is not the arrival rate, but rather system deterioration caused by non-recurrent events, such as accidents or weather conditions. Section 4.2.2 also disregards this assumption, as it considers peak hours.

Applying these reasonable assumptions results in a single equation for the probability mass function of traffic density, which is later validated against the data. Moreover, we prove this equation to be just a mixture of two Poisson distributions when times to incidents and clearance times are considerably longer than travel times under normal and adverse conditions.

As we have discussed in this section, the probability mass function of traffic density is a convolution sum between a Poisson random variable and a mixture of two Poisson random variables with random parameters coming from truncated Beta distributions. It represents the probability mass function for the state of an infinite queue subject to two-server states.

$$X = X_\phi + Y, \tag{4}$$

$$Y = p \cdot Y_1 + (1 - p) \cdot Y_2. \tag{5}$$

The probability mass function of Y_1 given in equation 2 can be explicitly written as:

$$P\{Y_1 = k\} = \int_0^c e^{-\gamma} \cdot \frac{\gamma^k}{k!} \cdot \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \cdot \frac{\left(\frac{\gamma}{c}\right)^{a-1}\left(1-\frac{\gamma}{c}\right)^{b-a-1}}{c} d\gamma. \tag{6}$$

$P\{Y_2 = k\}$ can be computed through the same integral, but via substituting (a) to $(a+1)$ and (b) to $(b+1)$. Hence, derivations below can also be carried out for $P\{Y_2 = k\}$ following the same operations.

This integral is not robust for all ranges of a and b. Additionally, its lack of a closed-form prevents it from being readily applied. There are several effective methods for computing this, the most simple being solving the balance equations numerically. However, we will retain this closed-form as it allows for further simplifications.

Note that the exponential factor in equation 6 could be expanded by using Taylor series to rewrite it as a sum multiplied by the truncated beta function that is equal to $\frac{\Gamma(k+n+a)\Gamma(b-a)}{\Gamma(k+n+b)}$ (see Olver [2010]).

$$\begin{aligned}
P\{Y_1 = k\} &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^c \gamma^n \cdot \frac{\gamma^k}{k!} \cdot \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \cdot \frac{\left(\frac{\gamma}{c}\right)^{a-1}\left(1-\frac{\gamma}{c}\right)^{b-a-1}}{c} d\gamma \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{k+n}}{k!} \cdot \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \cdot \int_0^c \left(\frac{\gamma}{c}\right)^{k+n+a-1} \left(1-\frac{\gamma}{c}\right)^{b-a-1} d\frac{\gamma}{c} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{k+n}}{k!} \cdot \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \cdot \int_0^1 x^{k+n+a-1}(1-x)^{b-a-1} dx \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{k+n}}{k!} \cdot \frac{\Gamma(b)\Gamma(k+a+n)}{\Gamma(a)\Gamma(k+b+n)}.
\end{aligned}$$

The full equation can then be simplified by expanding the convolution sum to:

$$\begin{aligned}
X &= X_\phi + Y, \\
P(X = k) &= \sum_{q=0}^{k} P\{X_\phi = k-q\} P\{Y = q\} \\
P(X = k) &= \sum_{q=0}^{k} P\{X_\phi = k-q\}(p \cdot P\{Y_1 = q\} + (1-p) \cdot P\{Y_2 = q\}) \\
&= \sum_{q=0}^{k} e^{-\lambda/\mu} \frac{(\lambda/\mu)^{k-q}}{(k-q)!} \left[p \cdot \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{q+n}}{q!} \cdot \frac{\Gamma(b)\Gamma(q+a+n)}{\Gamma(a)\Gamma(q+b+n)} \right. \\
&\quad \left. + (1-p) \cdot \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{q+n}}{q!} \cdot \frac{\Gamma(b+1)\Gamma(q+a+1+n)}{\Gamma(a+1)\Gamma(q+b+1+n)} \right].
\end{aligned} \tag{7}$$

This new equation solves the issues the integral had for the extreme points.

In practice it is typically true that $f \ll \mu$ and $r \ll \mu'$. This has the interpretation that the time a vehicle spends traveling a segment is much shorter than the time it takes for traffic to accumulate or disperse from an incident or a clearance. Under this assumption, further simplifications are possible. Tricomi and Erdélyi [1951] prove the following asymptotic approximation for the quotient of gamma functions:

$$\frac{\Gamma(z+a)}{\Gamma(z+b)} = z^{a-b}\left[1 + \frac{(a-b)(a+b-1)}{2z} + O(|z|)^{-2}\right]. \tag{8}$$

This is approximately 1 when $|a-b| \ll z$, which is implied by $f \ll \mu$ and $r \ll \mu'$.

For $k = 0$:

$$\begin{aligned} P\{Y_1 = k\} &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot c^n \cdot \frac{\Gamma(b)\Gamma(a+n)}{\Gamma(a)\Gamma(b+n)} \\ &= \frac{\Gamma(b)\Gamma(a)}{\Gamma(a)\Gamma(b)} + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \cdot c^n \cdot \frac{\Gamma(b)\Gamma(a+n)}{\Gamma(a)\Gamma(b+n)} \\ &\approx 1 + \frac{\Gamma(b)}{\Gamma(a)}(e^{-c} - 1). \end{aligned} \tag{9}$$

For $k \geq 1$:

$$\begin{aligned} P\{Y_1 = k\} &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{k+n}}{k!} \cdot \frac{\Gamma(b)\Gamma(k+a+n)}{\Gamma(a)\Gamma(k+b+n)} \\ &\approx \frac{\Gamma(b)}{\Gamma(a)} \frac{e^{-c} \cdot c^k}{k!}, \end{aligned} \tag{10}$$

The final equation is the weighted Poisson distribution.

This approximation is imprecise, since the ratio $\frac{\Gamma(k+a+n)}{\Gamma(k+b+n)}$ slowly diverges from 1 as $k+n$ grows bigger. However, the increase in the factorial terms in the equations grows faster, thus offsetting such divergence. Numerical experiments indicate that this approximation has no apparent adverse effect on the final density function.

The derivations to determine the equations to solve for $P\{Y_1 = k\}$ can be used in a similar way for $P\{Y_2 = k\}$, under the same assumptions (small a's and b's). Since the only parameters changing are $a = a+1$ and $b = b+1$, it is easy to see that again $\frac{\Gamma(k+a+1+n)}{\Gamma(k+b+1+n)}$ can be approximated as 1 for all k's; besides, $\frac{\Gamma(b+1)}{\Gamma(a+1)}$ can also be approximated to 1, allowing for even further simplification.

Thus, the probability mass function of Y_2, for all k's is:

$$\begin{aligned} P\{Y_2 = k\} &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot \frac{c^{k+n}}{k!} \cdot \frac{\Gamma(b+1)\Gamma(k+a+1+n)}{\Gamma(a+1)\Gamma(k+b+1+n)} \\ &\approx \frac{\Gamma(b+1)}{\Gamma(a+1)} \frac{e^{-c} \cdot c^k}{k!} \\ &\approx \frac{e^{-c} \cdot c^k}{k!}, \end{aligned} \tag{11}$$

which is a Poisson distribution.

We can further derive the probability of traffic density as

$$X = \frac{r}{r+f} X_\phi + \frac{f}{r+f} X_b,$$

where X_ϕ is a random variable that follows a Poisson with parameter (λ/μ), and X_b is a random variable that follows a Poisson with parameter (λ'/μ'). Proof is postponed to the Appendix. With a and b being

small, the weight parameter of the mixture tends to $\frac{r}{r+f}$ for X_ϕ and $\frac{f}{r+f}$ for X_b. This is expected, because when incident and clearance rates are low, it is likely that each car will spend the whole time in the same state of the queue.

Proposition 1 *The probability mass function of traffic density is the mixture of two Poisson random variables with rates* $\frac{\lambda}{\mu}$ *and* $\frac{\lambda'}{\mu'}$ *when the incident and clearance rates are considerably lower than* μ *and* μ'*. The mixing weight for the Poisson random variable with rate* $\frac{\lambda}{\mu}$ *is given by* $\frac{r}{r+f}$*.*

Although one may argue this result could follow from intuition, this proof formally demonstrates the result to be true for segments where the time to incident and clearance times are much longer than the average travel time to cross the segment. Furthermore, the proof shows that a more general formulation (equation 7) must be used for segments that do not meet these criteria. Note that this also suggests the probability mass function for the density depends on the clearance and incident rate, thus supporting the importance of considering non-recurrent incidents in the model. This closed-form solution also allows traffic engineers to compute higher moments when needed since they can be derived by weighing the moments from a Poisson distribution. The moment generating function (MGF) of a Poisson random variable Z with parameter ϕ is $M_Z(t) = e^{\phi(e^t-1)}$. The n^{th} factorial moment of the distribution can be computed by taking the n^{th} derivative of the MGF, then setting $t = 0$ [Ross, 1996]. The central moments of this Poisson random variable Z are $E[Z] = \phi$, $V[Z] = \phi$, Skewness$[Z] = \sqrt{\phi^{-1}}$, and Kurt$[Z] = \phi^{-1}$. Since the traffic density random variable X is approximated as a mixture of two independent Poisson random variables, we can immediately write its central moments as described in Table 1.

$E[X] = \frac{r}{r+f} \cdot \frac{\lambda}{\mu} + \frac{f}{r+f} \cdot \frac{\lambda'}{\mu'}$	$V[X] = \frac{r}{r+f} \cdot \frac{\lambda}{\mu} + \frac{f}{r+f} \cdot \frac{\lambda'}{\mu'}$
Skewness$[X] = \frac{r}{r+f} \cdot \sqrt{\frac{\mu}{\lambda}} + \frac{f}{r+f} \cdot \sqrt{\frac{\mu'}{\lambda'}}$	Kurt$[X] = \frac{r}{r+f} \cdot \frac{\mu}{\lambda} + \frac{f}{r+f} \cdot \frac{\mu'}{\lambda'}$

Table 1: Central Moments for the traffic density distribution under non-peak hours.

4.2.2 Peak hours.

When the system operates in peak hours, the higher arrival rate of cars is also a cause of congestion. In this case, the assumption that the travel time distribution is independent of the number of cars is not as consistent with real-life scenarios. Therefore, we need to account for another source of travel time deterioration: the current number of vehicles in the system.

For this case, we combine our deterioration model with the congested traffic model M/G/C/C discussed in Jain and Smith [1997]. Let us assume $\mathbf{a} = [\mathrm{a}_1, \mathrm{a}_2, \mathrm{a}_3, \ldots]$ is a vector where component a_n represents the deterioration coefficient caused by congestion when n cars are present. $\mathrm{a}_1 = 1$ because a single car can travel at free-flow speed. Moreover, $0 \leq \mathrm{a}_n \leq 1$, and a_n is monotonically decreasing as n grows, meaning that cars arriving can only maintain or worsen system conditions. As an initial suggestion, Jain and Smith [1997] propose function a_n to be linearly or exponentially decreasing in n. In this paper, we assume a linearly decreasing function for a_n, which avoids an overly fast deterioration caused by the additional cars.

Furthermore, since the probability of breakdown is not negligible in this scenario, the system capacity is truncated at a certain point C, i.e., we assume that no more cars arrive after C cars are in the system. The presence of C cars in the system represents a complete breakdown, where there is no space left for another car to arrive. The modeling thus follows an M/M/C/C queue in a random environment, represented in Figure 11.

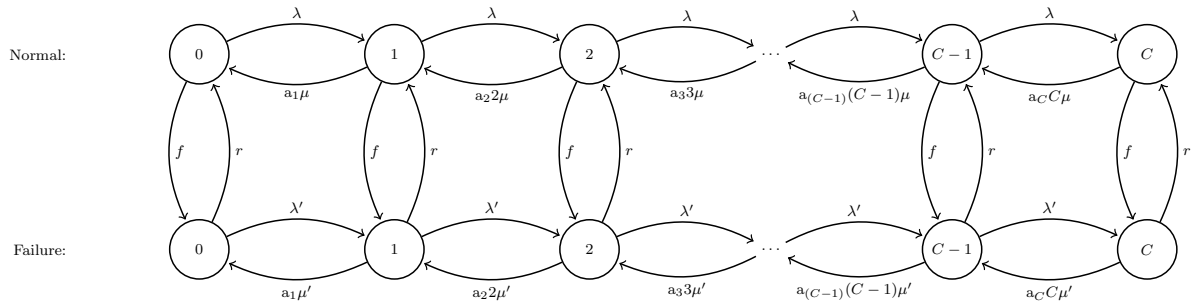

Fig. 11: State Transition Diagram for a Markov-modulated $M/M/C/C$ Queue. Note the inclusion of the parameters representing the extra congestion due to accumulation of cars, and the limited capacity C of the system.

The solution for such a queue is described below, and more details can be found in Baykal-Gürsoy et al. [2009a].

The balance equations are given by:

$$p_{iN}\left(\lambda + f + i\mu \mathrm{a}_i\right) = r p_{iF} + \left((i+1)\mu \mathrm{a}_{i+1}\right) p_{i+1,N} + \lambda p_{i-1,N}, \quad \text{for } i = 1, 2, \ldots, C-1,$$

$$p_{iF}\left(\lambda' + r + i\mu' \mathrm{a}_i\right) = f p_{iN} + \left((i+1)\mu' \mathrm{a}_{i+1}\right) p_{i+1,F} + \lambda' p_{i-1,F}, \quad \text{for } i = 1, 2, \ldots, C-1.$$

and the boundary equations are,

$$p_{0N}(\lambda + f) = rp_{0F} + \mu \mathrm{a}_1 p_{1N},$$

$$p_{0F}(\lambda' + r) = fp_{1N} + \mu' \mathrm{a}_1 p_{1F},$$

$$p_{CN}\left(C\mu \mathrm{a}_C + f\right) = rp_{CF} + \lambda p_{C-1,N},$$

$$p_{CF}\left(C\mu' \mathrm{a}_C + r\right) = fp_{CN} + \lambda' p_{C-1,F},$$

and the normalization equation is $\sum_{i=0}^{C}(p_{iN} + p_{iF}) = 1$.

Given a fixed value of C, determined as the maximum capacity of the road or as the number of cars for which the road is in a breakdown, we can solve for all p_{iN} and p_{iF}, as long as the values of the vector $\mathbf{a}$ are available.

Note that, unlike the non-peak hours' framework, the probability mass function for density during peak-hours does not have an intuitive closed form. However, we provide a straightforward and efficient framework to compute it numerically via a relatively small linear system of equations.

5 Validation

This section shows how the proposed model compares to the dataset. Here, we use a dataset to compute the aggregate parameters for the proposed model and compare it against the log-normal and Weibull distributions. These models also provide a reasonable fit to the data, but they require parameters that can only be found using observed density data. By comparison, the parameters for the proposed model are often known by traffic engineers or readily available with little data collection, allowing an effortless implementation of the model for different road segments.

5.1 Discussion and Findings

Results are obtained from the comparisons between the curve generated from the model and the data for each sensor, as depicted in figures 12, 13, and 14. The curve represents the cumulative distributions generated through the analytical model and from the density data computed from the occupancy data set. The main goal is to determine whether the curve from the analytical model is a good fit for the histogram and how it compares to other distributions used in literature. However, the dataset used is censored due to limitations in the sensors' precision. As a result, some points in the density distribution are misrepresented in the dataset. This limitation causes distribution tests, such as the Kolmogorov-Smirnov test, to fail for

both our model and other models used in literature. Therefore, we used the confidence interval obtained from the Kolmogorov-Smirnov test [Massey Jr, 1951] along with an added uncertainty level to account for sensors' lack of precision to determine the validity of the model. As the figures show, the empirical cumulative distribution is within the KS-test upper and lower bounds, and therefore matches the model's cumulative distribution. Lastly, we compare the analytical model to other distributions used in literature employing the Akaike Information Criterion.

5.1.1 Non-peak hours.

The best approach when using this analytical model is to separate specific periods in which these parameters behave according to the assumptions. The model assumes that time to incident and clearance times are exponentially distributed, as well as the interarrival times for up and down periods, and travel times during up and down periods. As an example, there could be a model for January and February, months known to be the snowiest, then separate the remaining months in two groups, dry season and rainy season, and have a model for each one of those. Such definitions depend on local factors and should be determined individually.

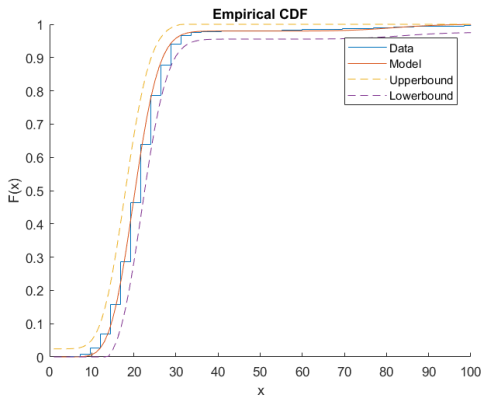

Fig. 12: Comparison between the CDF of the model and the empirical CDF from the data during non-peak hours for the month of January, accounting for both accident and snowfalls.

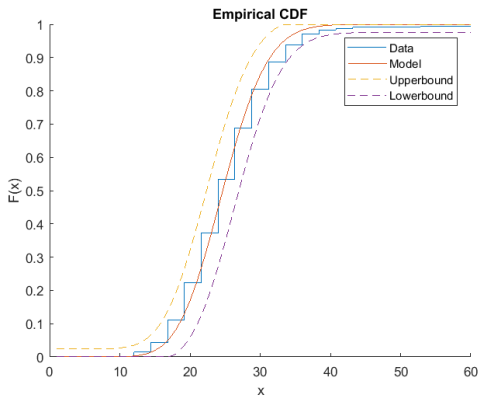

Fig. 13: Comparison between the CDF of the model and the empirical CDF from the data during non-peak hours for the months of June to August, accounting for accidents with occasional heavy rainfall.

Figures 12 and 13 depict one of the 36 sensors comparison for non-peak hours during winter and summer months. The upper and lower bounds represent the 5% significance level of the model distribution

[Massey Jr, 1951] when accounting for data uncertainty due to the censored dataset. Since the empirical distributions reside within the bounds, the model is valid for both sets of months. Non-peak hours data were gathered from 11 a.m. to 1 p.m., Tuesday to Thursdays. The main difference between the two figures is the slightly thicker tail in figure 12, mainly caused by snowstorms that tend to have a more lasting impact than the heavy rains and accidents in the summer.

	Model (Winter)	Lognormal (Winter)	Weibull (Winter)	Model (Summer)	Lognormal (Summer)	Weibull (Summer)
Mean	4931	4941	5042	1362	1173	1370
Best AIC (out of 36 sensors)	30	2	1	13	19	8

Table 2: Adapted AIC comparison between model and commonly used distributions during non-peak hours. The parameters are obtained from the data during the month of January (Winter), and August (Summer), Tuesday to Thursday, 11 a.m. to 1 p.m., for each sensor.

Table 2 indicates that the model closely matches the goodness of fit of log-normal and Weibull, two commonly used distributions (the complete AIC list for each sensor is given in Table 9, Appendix 2). Hence, our model is valid for roads in which incident and clearance rates (f and r) are much lower than μ and μ'. The values also suggest the validity of the model for both winter and summer months, although parameters may need adjustments for the analyzed period.

These results endorse the validity of this queuing theory approach, suggesting the simplified analytical model (a mixture of two Poisson distributions) is robust enough. It matches the performance of or performs better than other distributions used in literature. Furthermore, the results suggest that the assumptions made by the model are valid.

Moreover, a significant advantage of this analytical model is that it does not require detailed data, but rather aggregate parameters that can be easily estimated. Furthermore, minor errors in estimation are not overly harmful to the performance of the model. The derivatives of expected value and distribution of density with respect to the parameters $\lambda, \lambda', \mu, \mu', f, r$ allow for sensitivity analysis. Except for λ and λ', this model is robust to errors in estimation. For similar values to those seen on all sensors, a 10% error in μ or μ' results in a 1% error in expected density. The model is sensitive to changes in λ and λ', only when $\lambda \approx \lambda'$. This model is, therefore, appropriate as an initial gauge on how traffic density will behave in new

roadways, in roadways for which data are scarce, and in roadways that may have endured some change in behavior.

5.1.2 Peak hours.

Similar to non-peak hours, the results for peak hours also proved to be very robust. Peak hours are defined as the times between 8 a.m. and 10 a.m. for Southeast direction, and between 6 p.m. and 8 p.m. for Northwest direction, Tuesday to Thursday. These periods are chosen to match the commute from residential areas to work locations in the morning and back in the evening.

The values of **a** are calculated from a function of the number of cars in the system and the road segment capacity. They determine the deterioration in service level (travel time) caused by the presence of more cars, thus causing drivers to drive more carefully and slowly. Jain and Smith [1997] suggest that such function could be linear and follow $a_n = \frac{C+1-n}{C}$, where C is the capacity of the segment. In our dataset, segments are half-mile long and contain three lanes. Thus, we can obtain C as

$$C = \frac{\text{length segment} \cdot \text{number of lanes}}{\text{average length of a car}} = \frac{0.5 \cdot 3}{22/5280} = 360,$$

giving $a_n = \frac{361-n}{360}$.

An alternative approach is to generate **a** via travel time computed from the speed data. By comparing average speed data for each density point in each sensor, we can directly generate **a**. For each density point, we can observe every car's speed, thus creating an array of speed data points. For density points with fewer than three speed data points, we assume speed remained the same as the previous density data point to prevent outliers.

Although the two approaches yield good fits, the one that generates **a** from a linear function covers a little more of the variance of the data. The average r^2 for the linearly generated **a** is 0.776, and for the **a** coming from speed data is 0.762. This suggests that a closed-form function is a robust option even when data is not present. Hence, we choose to present the results obtained with the linear function for **a**.

Figure 14 presents the empirical CDF generated from the data along with the model generated CDF. The upper and lower bounds represent the 5% confidence interval of the model distribution [Massey Jr, 1951] when accounting for data uncertainty due to the gaps in the dataset. Since the empirical distribution resides within the bounds, the model is valid.

Table 3 compares the goodness of fit for our model with other commonly used distributions for traffic density (the complete AIC list for each sensor is described in Table 10, Appendix 2). Adapted AICs (fol-

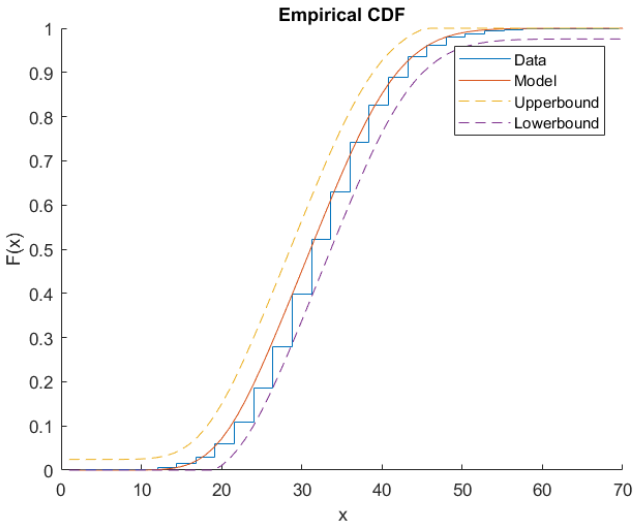

Fig. 14: Comparison between the CDF of the model and the empirical CDF from the data during peak hours for Summer months. Deterioration due to the number of cars, along with a similar arrival rate cause the difference between months to be minor during peak hours.

	Model (Winter)	Lognormal (Winter)	Weibull (Winter)	Model (Summer)	Lognormal (Summer)	Weibull (Summer)
Mean	5236	5377	5433	804	786	812
Best AIC (out of 36 sensors)	33 (1 tie)	3 (1 tie)	1	5 (1 tie)	23 (1 tie)	6

Table 3: Adapted AIC comparison between model and commonly used distributions during peak hours. The parameters are obtained from the data during the month of January (Winter), and August (Summer), Tuesday to Thursday, 8 a.m. to 10 a.m. for sensors in SE direction and 6 p.m. to 8 p.m. for sensors in the WN direction.

lowing the same algorithm described for non-peak hours) are calculated using the maximum log-likelihood function from the data for the model, lognormal and Weibull distributions. The MLEs calculated for the model are obtained using the Nelder-Mead method with five iterations [Nelder and Mead, 1965], which implies that certain improvements can still be achieved in the model's AIC if more iterations are performed. Results indicate that the analytical model performs slightly better than the other two fits during winter, and is consistent with the other two fits during summer months. They reiterate the importance of our model because it does not depend on large sets of data that can be expensive or infeasible to gather, differently from the other distributions.

6 Applications

6.1 Sensitivity on the deterioration level α

We consider the effect of the ratio between μ' and μ, which we call α. The lower the α is, the worse the system becomes when it deteriorates. A value of $\alpha = 0.80$ means service frequency drops 20% when the system deteriorates due to an incident. This parameter is one of the most difficult to change — it mainly reflects the overall infrastructure's resilience to incidents. A change in this parameter could represent the aging of a system or a significant shift in utilization. The model presented in this paper can help traffic engineers understand the effect a change in the deterioration level would have on congestion.

Tables 4 and 5 contain several points on the tail distribution for density during peak and non-peak hours. For example, we expect to see more than 11 cars with more than 30% probability when $\alpha = 0.8$ during non-peak hours and more than 18 with the same probability during peak hours. It is interesting to see that, as α decreases, the tail becomes thicker. The tables also show that the left side of the tail distribution is very similar for both non-peak and peak hours (as shown for $P\{X > x\} \geq 70\%$). In this part of the distribution, α has little impact. However, on the right side of the distribution (as shown for $P\{X > x\} \leq 30\%$) for peak hours, the addition of cars caused by a non-recurrent incident causes traffic to become worse when α is small, creating a cascading effect in traffic congestion.

$P\{X > x\}$	99%	70%	30%	1%	E[number of cars]
α=0.8	4	8	11	17	8.63
α=0.6	4	8	11	20	8.99
α=0.4	4	8	12	28	9.71
α=0.2	4	8	12	52	11.86

Table 4: Tail Probability for different levels of deterioration during non-peak hours when the arrival rate is 15 cars per minute, expected travel time is 1 minute under normal conditions, expected time to incident is 41 hours and expected clearance time is 28 hours. The expected travel time under adverse conditions vary with α as $\frac{1}{\alpha\mu}$, i.e., $\frac{1}{\alpha}$E[travel time under normal conditions].

More graphically, on Figures 15 and 16, we can see how the probability mass function behaves as α decreases for both non-peak and peak hours. At first, the tail starts to thicken. However, after some threshold, the system begins to become bimodal, with an evident separation between the normal condition

$P\{X > x\}$	99%	70%	30%	1%	E[number of cars]
α=0.8	8	14	18	27	15.01
α=0.6	8	14	18	33	15.69
α=0.4	8	14	19	49	17.14
α=0.2	8	14	19	112	22.81

Table 5: Tail Probability for different levels of deterioration during peak hours when the arrival rate is 25 cars per minute, expected travel time is 1 minute under normal conditions (before accounting for deterioration caused by congestion), and expected time to incident is 41 hours and expected clearance time is 28 hours. The expected travel time under adverse conditions (before accounting for deterioration caused by congestion) vary with α as $\frac{1}{\alpha\mu}$, i.e., $\frac{1}{\alpha}$E[travel time under normal conditions].

distribution and the adverse condition distribution. This bimodality is more prominent during peak hours than non-peak hours.

The probability of traffic breakdown directly follows from the resulting distribution. Traffic engineers may determine the number of cars that causes a traffic breakdown on this particular road, and then have the probability of breakdown determined by the model. This model is flexible enough to provide decision-makers with such a measure for different definitions of traffic breakdown since the literature has yet to agree on a specific definition.

6.2 Example on the usage of the model for planning

In this section, we provide an example of how this new model can be used by traffic engineers to help in their decision-making process. Suppose traffic engineers are planning to build a highway with different 0.5-mile sections. They are considering various investments, which alter multiple parameters in the model. In particular, they will look at the number of lanes to be built and the budget for nearby service vehicles and first responders. They assume all 0.5-mile sections within a region behave similarly. Such a study would usually entail collecting data for similar existing highways and simulating changes to them. However, the proposed model enables these analyses without such data or simulations.

In order to understand how the highway would behave throughout different periods, they separate the analysis into three period groups: low usage, which includes late nights, early mornings, weekends, and holidays; medium usage, which includes weekdays, but during times that avoid the main commute rush;

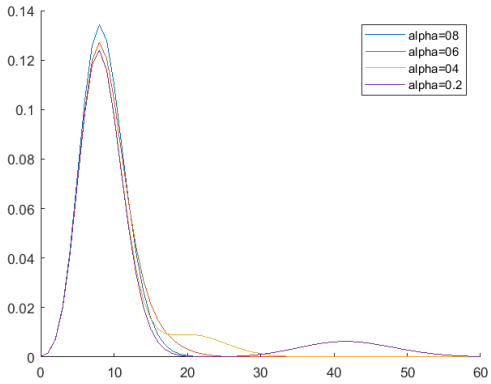

Fig. 15: PMF's for non-peak hours when the arrival rate is 15 cars per minute, expected travel time is 1 minute under normal conditions, and expected time to incident is 41 hours and expected clearance time is 28 hours.

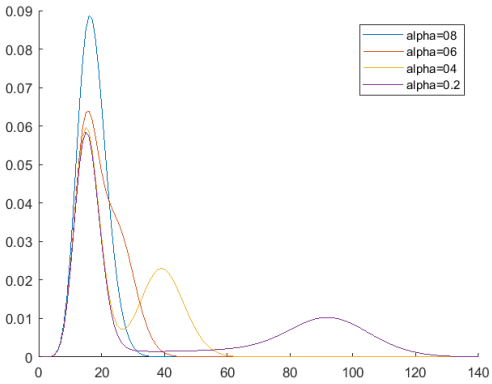

Fig. 16: PMF's for peak hours when the arrival rate is 25 cars per minute, expected travel time is 1 minute under normal conditions (before accounting for deterioration caused by congestion), and expected time to incident is 41 hours and expected clearance time is 28 hours.

high usage, which includes weekdays peak-hours. By comparing the region with other similar locations, the traffic engineers were able to estimate the following parameters:

	Low Usage	Medium Usage	High Usage
λ	160	650	1100
λ'	150	630	790
μ	60	21	18
μ'	40	14	12
f	0.00002	0.005	0.02
r	2	2	2

Table 6: Parameters estimated by the traffic engineers for the location in which the new highway will be built.

From the results, we can assume that the system is unlikely to reach full capacity under Medium Usage, and we, therefore, use the non-peak hours' method to calculate the density distribution. For the High Usage time frame, we use the peak-hours model instead.

We first explore the effect of lane count, assuming an average clearance time of 30 minutes. Increasing the lane count is equivalent to increasing the capacity of the system. Table 7 shows the probabilities of

seeing more than 10% of the road occupied and the probability of having less than 90% of the road occupied under Medium and High Usage given the number (1, 2 or 3) of lanes. We immediately see the trade-off between under-utilization in medium usage periods and over-utilization in high usage periods. The correct selection of lanes depends on the constraints faced by the decision-maker. For the sake of discussion, we select two lanes, which will be over-utilized 25% of the time during peak hours.

	Medium Usage		High Usage	
Lanes	$P\{X > 0.1 \cdot C\}$	$P\{X < 0.9 \cdot C\}$	$P\{X > 0.1 \cdot C\}$	$P\{X < 0.9 \cdot C\}$
1	0.9999	1	1	0.2306
2	0.8799	1	1	0.7583
3	0.1608	1	1	0.9998

Table 7: Probability of under-utilization and over-utilization for different lane counts. C represents the capacity of the system, which is a function of the number of lanes.

With the number of lanes selected, we consider changes to the clearance time. It can be affected by other investments, such as the response time of first responders and service vehicles. Again we see a trade-off that must be reconciled by the decision-maker — over-utilization can be reduced, but only by making significant reductions in clearance time. Another consideration is that increasing the clearance rate (i.e., decreasing expected clearance time) has a stronger impact on lowering over-utilization than under-utilization.

		Medium Usage		High Usage	
E[Clearance Time]	r	$P\{X > 24\}$	$P\{X < 216\}$	$P\{X > 24\}$	$P\{X < 216\}$
8.5 min	7	0.8796	1	1	0.9092
15 min	4	0.8797	1	1	0.8587
30 min	2	0.8799	1	1	0.7583
5 h	0.2	0.8825	1	1	0.2146
50 h	0.02	0.9036	1	1	0.0002

Table 8: Probability of under-utilization and over-utilization for different clearance times, with 2 lanes.

7 Conclusion

The analytical queuing model for traffic density proposed by Baykal-Gürsoy and Xiao [2004], Baykal-Gürsoy et al. [2009b] proved valid under certain conditions. Its original form has been algebraically changed for

more direct computational results, making its application more natural. When incident and clearance frequencies have low variance, and traffic conditions are steady, the model further simplifies into a mixture of two Poisson random variables. Moreover, the model has now been expanded to account for peak hour congestion in addition to non-peak hour congestion estimation. Results are comparable to other models used in literature. The simplicity of this model provides a direct application that makes it easily adaptable to different conditions. The full distribution output also allows for efficient performance measurement, which includes estimating the probability of breakdown or analyzing what impact changes in the infrastructure may cause. However, because the proposed model assumes independence between segments, it may not be suitable for planning in long stretches of road where the aggregate parameters vary across space. Tandem queueing networks allow for local changes in the incident and clearance rates, as well as the inclusion of the effect of propagation across segments. Further research into the distributions of Markov-modulated tandem queueing networks and their comparison with the presented model is planned. Nonetheless, the proposed model has the potential to impact long-term congestion planning by providing decision-makers with the full probabilistic behavior of congestion emerging from their decisions, even when only minimal statistical information is available.

Acknowledgements

The authors are grateful to Peter J. Jin for providing the data and to Marcelo Figueroa-Candia for helping with data preprocessing. The first author would like to thank CNPq (Brazilian National Council for Scientific and Technological Development) for funding this research.

References

A. Alfa and M. Neuts. Modelling vehicular traffic using the discrete time Markovian arrival process. *Transportation Science*, 29:109–117, 1995.

B. Anbaroglu, B. Heydecker, and T. Cheng. Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. *Transportation Research Part C: Emerging Technologies*, 48:47–65, 2014.

P. Arnesen and O. A. Hjelkrem. An estimator for traffic breakdown probability based on classification of transitional breakdown events. *Transportation Science*, 52(3):593–602, 2018. doi: 10.1287/trsc.2017.0776. URL `https://doi.org/10.1287/trsc.2017.0776`.

M. Baykal-Gürsoy and W. Xiao. Stochastic decomposition in M/M/∞ queues with Markov-modulated service rates. *Queueing Systems*, 48:75–88, 2004.

M. Baykal-Gürsoy, Z. Duan, and H. Xu. Stochastic models of traffic flow interrupted by incidents. *IFAC Proceedings Volumes*, 42(15):442–449, 2009a.

M. Baykal-Gürsoy, W. Xiao, and K. Ozbay. Modeling traffic flow interrupted by incidents. *European Journal of Operational Research*, 195(1):127–138, 2009b.

M. Baykal-Gürsoy, M. Figueroa Candia, and Z. Duan. Completion times on jobs on two-state service processes and their asymptotic behavior. *under review*, 2019.

H. B. Celikoglu. Dynamic classification of traffic flow patterns simulated by a switching multimode discrete cell transmission model. *IEEE Transactions on Intelligent Transportation Systems*, 15(6):2539–2550, 2014.

H. B. Celikoglu and M. A. Silgu. Extension of traffic flow pattern dynamic classification by a macroscopic model using multivariate clustering. *Transportation Science*, 50(3):966–981, 2016.

J. Cheah and J. Smith. Generalized M/G/C/C state dependent queuing models and pedestrian traffic flows. *Queueing Systems*, 15:365–385, 1994.

Z. Chen, X. C. Liu, and G. Zhang. Non-recurrent congestion analysis using data-driven spatiotemporal approach for information construction. *Transportation Research Part C: Emerging Technologies*, 71:19–31, 2016.

S.-W. Chiou. A robust urban traffic network design with signal settings. *Information Sciences*, 334:144–160, 2016.

C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B, Methodological*, 28(4):269–287, 1994.

C. F. Daganzo. In traffic flow, cellular automata= kinematic waves. *Transportation Research Part B: Methodological*, 40(5):396–403, 2006.

D. J. Dailey. A statistical algorithm for estimating speed from single loop volume and occupancy measurements. *Transportation Research Part B: Methodological*, 33(5):313–322, 1999.

J. Dong and H. S. Mahmassani. Stochastic modeling of traffic flow breakdown phenomenon: Application to predicting travel time reliability. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1803–1809, 2012.

M. C. Dunne. Traffic delays at a signalized intersection with binomial arrivals. *Transportation Science*, 1: 24–31, 1967.

D. C. Gazis. *Traffic theory*, volume 50. Springer Science & Business Media, 2006.

Y. Han and S. Ahn. Stochastic modeling of breakdown at freeway merge bottleneck and traffic control method using connected automated vehicle. *Transportation research part B: methodological*, 107:146–166, 2018.

D. Heidemann. A queueing theory approach to speed-flow-density relationships. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, France, July 1996.

D. Heidemann. A queueing theory model of nonstationary traffic flow. *Transportation Science*, 35:405–412, 2001.

R. Jain and J. Smith. Modeling vehicular traffic flow using M/G/C/C state dependent queueing models. *Transportation Science*, 31:324–336, 1997.

G. N. J.N. Darroch and R. Morris. Queues for vehicle-actuated traffic light. *Operations Research*, 12:882–895, 1964.

J. Keilson and L. Servi. The matrix M/M/∞ system: Retrial models and Markov modulated sources. *Advances in Applied Probability*, 25:453–471, 1993.

B. S. Kerner, S. L. Klenov, and D. E. Wolf. Cellular automata approach to three-phase traffic theory. *Journal of Physics A: Mathematical and General*, 35(47):9971, 2002.

A. A. Kurzhanskiy. Set-valued estimation of freeway traffic density. *IFAC Proceedings Volumes*, 42(15): 271–277, 2009.

J. Kwon, M. Mauch, and P. Varaiya. Components of congestion: Delay from incidents, special events, lane closures, weather, potential ramp metering gain, and excess demand. *Transportation Research Record: Journal of the Transportation Research Board*, 1959(1959):84–91, 2006. ISSN 0361-1981. doi: 10.3141/1959-10. URL `http://trrjournalonline.trb.org/doi/10.3141/1959-10`.

P.-A. Laharotte, R. Billot, and N.-E. El Faouzi. Detection of non-recurrent road traffic events based on clustering indicators. In *ESANN*, 2017.

J. Lehoczky. Traffic intersection control and zero-switch queues. *J. of Applied Probability*, 9:382–395, 1972.

M. Lighthill and G. Whitham. On kinematic waves: II. a theory of traffic on long crowded roads. In *Proc. Roy. Soc. London Ser. A 229*, pages 317–345, 1955.

S. Maerivoet and B. De Moor. Cellular automata models of road traffic. *Physics reports*, 419(1):1–64, 2005.

F. J. Massey Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.

A. May and H. Keller. Non-integer car-following models. *Highway Res. Rec.*, 199:19–32, 1967.

L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez. Traffic density estimation with the cell transmission model. In *Proceedings of the 2003 American Control Conference, 2003.*, volume 5, pages 3750–3755. IEEE, 2003.

K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I*, 2 (12):2221–2229, 1992.

J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4): 308–313, 1965.

G. Newell. Approximation methods for queues with application to the fixed-cycle traffic light. *SIAM Rev.*, 7:2:223–240, 1965.

G. Newell. *Applications of Queueing Theory.* Chapman and Hall, London, 1971.

G. F. Newell. A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, 27(4):281–287, 1993.

NOAA. National climatic data center, Accesed in 2015-2016.

F. W. Olver. *NIST Handbook of Mathematical Functions Hardback and CD-ROM.* Cambridge University Press, 2010.

S. Panichpapiboon and P. Leakkaw. Traffic density estimation: A mobile sensing approach. *IEEE Communications Magazine*, 55(12):126–131, 2017.

N. Polson and V. Sokolov. Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, 79:1–17, 2017.

P. Richards. Shock waves on the highway. *Operations Research*, 4:42–51, 1956.

S. Ross. *Stochastic Processes*, volume 2. Wiley New York, 1996.

D. Schrank, B. Eisele, T. Lomax, and J. Bak. Urban mobility scorecard. *College Station: Texas A&M Transportation Institute and INRIX*, (August), 2015.

A. Skabardonis, K. Petty, P. Varaiya, and R. Bertini. Evaluation of the Freeway Service Patrol (FSP) in Los Angeles, ucb-its-prr-98-31. Technical report, California PATH Research Report, Institute of Transportation Studies, University of California, Berkeley, 1998.

A. Skabardonis, K. Petty, and P. Varaiya. Measuring recurrent and non-recurrent traffic congestion. In *Proc. of the 82th Annual Meeting of the Transportation Research Board*, 2003. volume CD-ROM, Washington D. C.

J. Tanner. A problem of interface between two queues. *Biometrica*, 40:58–69, 1953.

G. S. Thakur, P. Huiz, and A. Helmy. Modeling and characterization of urban vehicular mobility using web cameras. In *Proceedings of the 2012 IEEE Conference on Computer Communications Workshop*, pages

262–267. IEEE, 2012.

M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2):1805, 2000.

F. G. Tricomi and A. Erdélyi. The asymptotic expansion of a ratio of gamma functions. *Pacific Journal of Mathematics*, 1(1):133–142, 1951.

N. Vandaele, T. VanWoensel, and N. Verbruggen. A queueing based traffic flow model. *Transportation Research-D: Transportation and Environment*, 5:121–135, 2000.

S. Wang and N. Ahmed. Dynamic model of urban traffic and optimum management of its flow and congestion. *Dynamic Systems and Applications*, 26(3-4):575–587, 2017.

S. Zhang, G. Wu, J. P. Costeira, and J. M. Moura. Understanding traffic density from large-scale web camera data. *arXiv preprint arXiv:1703.05868*, 2017.

J. Zheng and H. X. Liu. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transportation Research Part C: Emerging Technologies*, 79:347–362, 2017.

APPENDIX 1 – Proof of Proposition 1

We will derive the probability mass function of the random number of cars on a segment from equations 1, 4, 5, 9, 10, and 11.

$$
\begin{aligned}
P\{X=k\} =& P\{X_\phi + Y = k\} = \sum_{q=0}^{k} P\{X_\phi = k-q\} \cdot P\{Y=q\} \\
=& P\{X_\phi = k\} \cdot P\{Y=0\} + \sum_{q=1}^{k} P\{X_\phi = k-q\} \cdot (p \cdot P\{Y_1 = q\} + (1-p) \cdot P\{Y_2 = q\}) \\
=& \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^k}{k!} \cdot \left[p \cdot \left(1 + \frac{\Gamma(b)}{\Gamma(a)} (e^{-c} - 1) \right) + (1-p) e^{-c} \right] \\
&+ \sum_{q=1}^{k} \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^{k-q}}{(k-q)!} \left[p \cdot \left(\frac{\Gamma(b)}{\Gamma(a)} \frac{(e^{-c} c^q)}{q!} \right) + (1-p) \frac{(e^{-c} c^q)}{q!} \right] \\
=& \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^k}{k!} \cdot \left[e^{-c} \left(1 + p \left(\frac{\Gamma(b)}{\Gamma(a)} - 1 \right) \right) + p \left(1 - \frac{\Gamma(b)}{\Gamma(a)} \right) \right] \\
&+ \left(1 + p \left(\frac{\Gamma(b)}{\Gamma(a)} - 1 \right) \right) \sum_{q=1}^{k} \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^{k-q}}{(k-q)!} \cdot \frac{e^{-c} c^q}{q!} \\
=& \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^k}{k!} \cdot \left[e^{-c} \left(1 + p \left(\frac{\Gamma(b)}{\Gamma(a)} - 1 \right) \right) + p \left(1 - \frac{\Gamma(b)}{\Gamma(a)} \right) \right] \\
&+ \left(1 + p \left(\frac{\Gamma(b)}{\Gamma(a)} - 1 \right) \right) \sum_{q=1}^{k} \frac{e^{-\lambda/\mu} \cdot (\lambda/\mu)^{k-q}}{(k-q)!} \cdot \frac{e^{-c} c^q}{q!}.
\end{aligned}
$$

Substituting $\left(1-p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)\right)=m$, we have

$$\begin{aligned}P\{X=k\}=&\frac{e^{-\lambda/\mu}\cdot(\lambda/\mu)^k}{k!}\cdot\left[e^{-c}m+(1-m)\right]+m\cdot\frac{e^{-(\lambda/\mu+c)}[(\lambda/\mu+c)^k-(\lambda/\mu)^k]}{k!}\\=&(1-m)\cdot\frac{e^{-\lambda/\mu}\cdot(\lambda/\mu)^k}{k!}+m\left(\frac{e^{-(\lambda/\mu+c)}[(\lambda/\mu+c)^k-(\lambda/\mu)^k]}{k!}+\frac{e^{-\lambda/\mu+c}\cdot(\lambda/\mu)^k}{k!}\right)\\=&(1-m)\cdot\frac{e^{-\lambda/\mu}\cdot(\lambda/\mu)^k}{k!}+m\left(\frac{e^{-(\lambda/\mu+c)}(\lambda/\mu+c)^k}{k!}\right).\end{aligned}$$

Therefore

$$P\{X=k\}=p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)X_g+\left(1-p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)\right)X_b.$$

Now we will show that $p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)$ is approximately equal to $\frac{r}{r+f}$ for small a and b.

Claim: $p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)\approx\frac{r}{r+f}$ for small a and b.

For $x>0$,

$$\Gamma(x)=\frac{\Gamma(x+1)}{x}.$$

The above relation for $x=a=\frac{f}{\mu}$ and $x=b=\frac{f}{\mu}+\frac{r}{\mu'}$, implies

$$\frac{\Gamma(b)}{\Gamma(a)}=\frac{\frac{f}{\mu}}{\frac{f}{\mu}+\frac{r}{\mu'}}\cdot\frac{\Gamma(b+1)}{\Gamma(a+1)}.$$

Using the approximation (8) for $z=1$, since $a\ll 1$, $b\ll 1$, and $b-a\ll 1$, one can deduce that $\frac{\Gamma(b+1)}{\Gamma(a+1)}$ is approximately equal to 1. Thus,

$$\frac{\Gamma(b)}{\Gamma(a)}\approx\frac{\frac{f}{\mu}}{\frac{f}{\mu}+\frac{r}{\mu'}}.$$

Because $p=\frac{r+f\frac{\mu'}{\mu}}{r+f}$, the result follows

$$\begin{aligned}p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)&\approx\left(\frac{r+f\frac{\mu'}{\mu}}{r+f}\right)\left(1-\frac{\frac{f}{\mu}}{\frac{f}{\mu}+\frac{r}{\mu'}}\right)\\&=\left(\frac{r\mu+f\mu'}{\mu(r+f)}\right)\left(\frac{\frac{r}{\mu'}}{\frac{f}{\mu}+\frac{r}{\mu'}}\right)=\frac{r}{r+f}.\end{aligned}$$

As a result, the probability mass function of density tends to the mixture $P\{X=k\}=\frac{r}{r+f}X_g+\frac{f}{r+f}X_b$, where X_g follow a Poisson with parameter (λ/μ) and X_b follow a Poisson with parameter (λ'/μ').

The error of the approximation can be derived similarly to the previous proof:

$$\begin{aligned}Error&=p\left(1-\frac{\Gamma(b)}{\Gamma(a)}\right)-\frac{r}{r+f}\\&=\left(\frac{r\mu+f\mu'}{\mu(r+f)}\right)\left(1-\frac{\frac{f}{\mu}}{\frac{f}{\mu}+\frac{r}{\mu'}}\cdot\frac{\Gamma(b+1)}{\Gamma(a+1)}\right)-\frac{r}{r+f}\\&=\left(\frac{r\mu+f\mu'}{\mu(r+f)}\right)-\left(\frac{f\mu'}{\mu(r+f)}\cdot\frac{\Gamma(b+1)}{\Gamma(a+1)}\right)-\frac{\mu r}{\mu(r+f)}\\&=\frac{f}{r+f}\cdot\frac{\mu'}{\mu}\cdot\left(1-\frac{\Gamma(b+1)}{\Gamma(a+1)}\right).\end{aligned}$$

One can see that when f and r are smaller in comparison to μ and μ', respectively, this approximation of the weight of the mixture holds. Figure 17 shows the error for different ratios between f and μ and r and μ'. We can see that when they are around 1% of the μ and μ', the error is on the order of 10^{-3}, and it decreases to lower than 10^{-5} when the ratios drop to 0.01%.

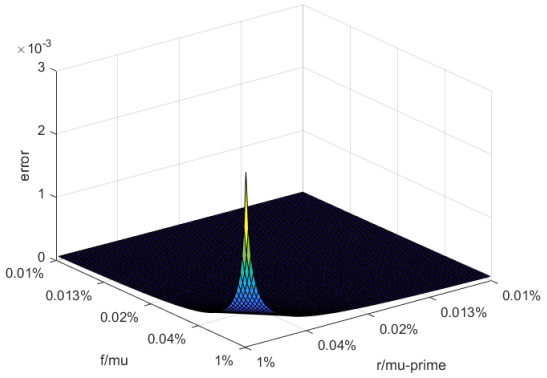

Fig. 17: Error between the algebraically found weight and $\frac{r}{r+f}$ for different ratios of f and μ, and r and μ'.

APPENDIX 2 – Tables containing AIC for each sensor

Sensor (SE) then (WN)	Model (Winter)	Lognormal (Winter)	Weibull (Winter)	Model (Summer)	Lognormal (Summer)	Weibull (Summer)
1	6057	5763	5607	783	784	804
2	4775	4795	5058	747	749	814
3	6427	5867	5732	817	814	840
4	4965	4985	5453	767	791	882
5	4555	4794	5307	738	769	862
6	4740	4849	5250	789	787	800
7	4948	4970	5405	776	779	778
8	4752	4819	4792	1303	1233	1344
9	4636	4708	4664	787	792	786
10	4865	4866	4938	771	776	774
11	4956	5007	4995	822	830	820
12	4689	4754	4721	820	832	831
13	4078	4157	4263	700	700	705
14	4762	4814	4874	800	803	802
15	4762	4838	4791	816	816	818
16	4670	4740	4706	802	803	805
17	5009	5070	5021	804	804	804
18	5983	5337	5361	830	826	831
1	4822	4862	4869	833	829	810
2	5334	5319	5507	891	893	882
3	4724	4818	5251	806	807	802
4	4724	4818	5251	806	807	802
5	4867	4889	4899	988	887	1070
6	5003	5026	5018	1034	907	1073
7	4795	4832	4825	993	892	1102
8	3990	4125	4007	1210	1062	1364
9	4740	4802	4770	1058	929	957
10	4904	4942	4937	1328	1111	1147
11	5106	5131	5127	1633	1271	1280
12	5319	5317	5287	1727	1632	1589
13	5642	5567	5549	1717	1588	1623
14	4900	4959	4914	1872	1972	2099
15	4722	4757	4947	4854	3187	N/A
16	4480	4620	4884	6314	4013	N/A
17	4890	4956	5196	3606	2302	9400
18	4941	5014	5328	2699	2437	5473

Table 9: Full list of the adapted AIC comparison between model and commonly used distributions during non-peak hours.

Sensor (SE) then (WN)	Model (Winter)	Lognormal (Winter)	Weibull (Winter)	Model (Summer)	Lognormal (Summer)	Weibull (Summer)
1	5505	6026	5767	774	772	780
2	4938	5132	5149	746	743	746
3	5663	6055	5775	800	796	798
4	5112	5302	5358	775	768	775
5	4777	5182	5341	736	728	734
6	5020	5147	5366	762	747	746
7	5255	5357	5626	771	776	767
8	N/A	5830	5990	874	831	883
9	5000	5257	4999	790	774	795
10	5230	5426	5262	777	786	831
11	5245	5494	5353	867	812	868
12	5163	5351	5247	829	809	854
13	4577	4982	5141	789	687	721
14	5343	5468	5626	822	798	807
15	5385	5678	5922	825	798	863
16	5643	5683	5943	747	797	867
17	6003	6050	6263	886	853	886
18	5955	6062	6262	926	846	907
1	5391	5395	5510	823	809	852
2	5765	5770	5847	882	876	920
3	5221	5245	5258	797	788	796
4	5221	5245	5258	797	788	796
5	5286	5310	5323	784	778	780
6	5345	5375	5374	806	801	795
7	5212	5241	5249	827	761	767
8	4505	4533	4507	747	747	753
9	4929	4961	4952	781	775	772
10	5169	5195	5198	792	787	783
11	5138	5152	5199	886	827	817
12	5351	5506	5476	809	772	788
13	5554	5734	5689	828	809	818
14	5039	5121	5180	779	758	806
15	4851	4864	4945	756	763	828
16	4814	4818	4972	738	746	821
17	5236	5228	5482	816	787	837
18	5402	5402	5784	806	803	871

Table 10: Full list of the adapted AIC comparison between model and commonly used distributions during peak hours.