AN INVESTIGATION OF HEADWAY DISTRIBUTION MODELS ON TWO-LANE ROADS PASSING THROUGH PERI-URBAN AREA

PritamSaha and Rupali Roy
Indian Institute of Engineering Science and Technology, Shibpur
saha.pritam@gmail.com&rupaliroy.nit@gmail.com

Ashoke Kumar Sarkar
Birla Institute of Technology and Science Pilani
asarkarbits@gmail.com

ABSTRACT

This paper demonstrates a simplistic approach of modellingtime headways especially on roads that pass through peri-urban areas. Since, public transport facilities in those areas are not adequate, the gaps in such facilities are usually met by privately own vehicles and different types of para-transit modes of transportation including paddle tri-cycle. As a result, the prevailing traffic exhibits extreme heterogeneity in its mix which has an impact on vehicle-arrival characteristics and inter-arrival time or headways as a consequence.Frequent formation of platoons makes distribution of headways somewhat skewed and conventional negative exponential distribution cannot describe them. The current study, accordingly, aimed at developing a method of modelling headways and thereby, alleviating the implications on capacity and level-of-service on these roads. Since, it is difficult to calibrate 'mixed models' under mixed traffic, the paper proposed combined distribution as an alternative wherein shorter and longer headways are modelled separately. On the basis of field data collected on a two-lane highway passing through moderately developed peri-urban area, the study investigated the compatibility of combined distributions. Three distribution functions, namely Exponential, Lognormal and Gamma models were chosen and fitted to the observed headway data using maximum-likelihood estimation and K-S test technique with 5% level-of-significance.

KEYWORDS: Two-lane roads; peri-urban area; mixed traffic; time headways

INTRODUCTION

Most of the smaller states in India are capital-centric, even though they are towns or small cities. With the economic growth during the last few years they are growing quite fast and a large population from all over the state is settling down due to opportunities in employment, business and better medical and educational facilities. Unlike the large cities, usually these towns are having only one major road connectivity, typically a two-lane highway, from other important towns/ cities in the state and the development takes place along these roads since they provides easy accessibility to the city centre; sometimes it is referred as development of peri-urban areas. Systematic development plans of such towns/ cities are usually not present due to shortage of funds and lack of vision.

Peri-urban areas gradually start to generate considerable amount of traffic as the residents are dependent on the town/ city for almost all needs. In the absence of proper planning and lack of sufficient demand to sustain city bus services, the residents usually own bicycles, two-wheelers and cars and sometimes rely on para-transit modes, thereby, making the prevailing traffic mixed in nature and quite different from rural road sections. Substantial city traffic influence on such roads results in mobility reduction and subsequent deterioration of quality of service. Further, presence of a wide range of vehicle category in terms of static and dynamic characteristics makes it tricky to arrive at appropriate inputs in the process of developing simulation models. Since, time headway is a key component of determining the performance of a simulation model and also, plays a fundamental role in many traffic engineering applications, such as capacity and level of service analysis, it is imperative to develop a simplistic approach of modelling time headways which would exhibit aptness under mixed traffic as usually observed on such roads.

Statistically, the headway data can be described as exponential if the co-efficient of variation is unity. However, on the basis of empirical investigation, it was observed that the heterogeneity effect results in a deviation of co-efficient of variation and makes exponential model inappropriate in describing headways. Hypo-exponential model is considered suitable when co-efficient of variation is less than one whereas hyper-exponential model is considered statistically valid when it exceeds unity.Shifted erlang and shifted exponential models have been used as hypo-exponential as they meet all the traits. Mixture distribution typically represents hyper-exponential distribution. However, it involves complex mathematical approach for parameter calibration which aggravates more in the event of heterogeneous traffic. Thus, in the present study, combined distribution was introduced as an alternative of mixture distribution. Accordingly, a simplistic approach was developed using the concept of modeling shorter and longer headways separately. Clustering technique was adopted to group the data into shorter and longer headways. Moreover, Exponential, Lognormal and Gamma models were chosen and fitted to the headway data. The appropriate models of headway distributions were selected using a methodology based on K-S test with 5% level-of-significance and also examined visually on the basis of quantile plots.

RESEARCH MOTIVATIONS

There have been several studies that investigated the time headway distribution of vehicles and suggested suitable models. Based on a study in North Carolina, Khasnabis and Heimbach studied headway distribution models for two-lane rural highway under varying traffic flow levels. They experimented with six headway distribution models, namely, Erlang, negative exponential, Pearson Type III, and Schuhl models, and their combinations and found Schuhl model appropriate for the headway distributions (Khasnabis and Heimbach 1980). Al-Ghamdi (2001) also observed headway distribution to vary with the flow rate and thereby, made an attempt to identify the boundaries where it changes. At low flow, where interaction among the vehicles are insignificant, negative exponential distribution fits well to the observed headways, whereas for medium flow shifted exponential and gamma distributions were found suitable. Erlang distribution displayed compatibility at heavy flow when car following interaction is frequent (Al-Ghamdi 2001). Luttinen (1994, 1999) analyzed headways on rural highways in Finland and concluded that the gamma distribution could be used for low-to-moderate traffic volumes which have low probability for short headways.The M3 distribution is a good headway model if accurate modelling of short headways is not required. He also suggested that although lognormal distribution is neither simple nor realistic enough, it can be considered as a model for the follower headway distribution. Couple of studies conducted on urban roads in India reported that while describing the headway characteristics under mixed traffic conditions both, hyperlang distribution (Chandra and Kumar 2001) and negative exponential distribution (Arasan and Khosy2003) could be applied. Abtahi et al. (2011) reported different headway distribution models in the passing and middle lanes in urban highways under heavy traffic condition. They proposed that lognormal and gamma models with 0.24 s and 0.69 s shifts are appropriate in passing and middle lane respectively. Bham and Ancha (2006) proposed lognormal and gamma distribution with a shift of 0.21 s and 0.26 s respectively for the modelling of preferred time headway and the time headway of drivers. In a study conducted in Dhaka, Hossain and Iqbal (1999) observed that the headways followed both lognormal and exponential distribution. Dey and Chandra (2009) proposed two continuous statistical distribution models, gamma and lognormal for desired time gap and time headway of drivers in a steady car-following state on two-lane roads under mixed traffic conditions. Mei and Bullen found that in a car-following situation time headways follow lognormal distribution and congregate to the shifted lognormal distribution with a shift of 0.3–0.4 s at higher traffic volumes (Mei and Bullen 1993). Kumar and Rao (1998) observed that negative exponential distribution adequately describes the headways at low to moderate flow levels. Hoogendoorn and Botma (1997) proposed Branston's generalized queuing model for time headway distributions. They proposed a new statistical procedure to estimate the parameters of a mixed-vehicle-type headway distribution model (Hoogendoorn and Bovy 1998). Rossi et al. (2013) analyzed the time headway data of a two-lane rural roads and considered gamma-generalized queuing model (gamma-GQM) to be an appropriate model in describing headway data. The model considered both, the endogenous traffic parameters such as traffic flow and composition and exogenous conditions such as weather and geometric features. Experience on urban roads in Bangkok, Thailand indicates that generalized extreme value (GEV) is quite effective in modelling headways and exponential is the least effective one (Panichpapiboon2014). Ye and Zhang (2009) collected headway data and identified differences between the vehicle type-specific headway distributions. He found that the truck-truck headway is the largest and the car-car headway is the smallest. Zhang et al. (2007) examined the performance of various headway models. The goodness-of-fit was checked by K-S test and visualized by Q-Q plot. The test result showed that the Double Displaced Negative Exponential Distribution model provided the best fit to the headways obtained at urban freeways and shifted lognormal distribution fits well to the headways obtained on other types of roads.

A fairly recent study on Indian traffic indicates that at car-following state under mixed traffic headway between two vehicles depends on the length of the lead vehicle. Typically, such traffic displays a wide range of vehicle type even under the same vehicle category, thus, making headway analysis under such traffic tricky (Penmetsa et al. 2015). In a study conducted in India under mixed traffic reveals that car-following interaction ceases beyond a headway value of 6 seconds (Saha et al. 2017).Further, the majority of the above studies investigated the appropriate model of time headways based on the data obtained at homogeneous traffic. Thus, the premise on which the present study is based considers a systematic analysis of headway data of heterogeneous traffic accounting for a wide range of flow scopes (volume to capacity ratio ranging from 0.2 to 1.0) and directional movements separately.

STUDY SITES AND FIELD DATA

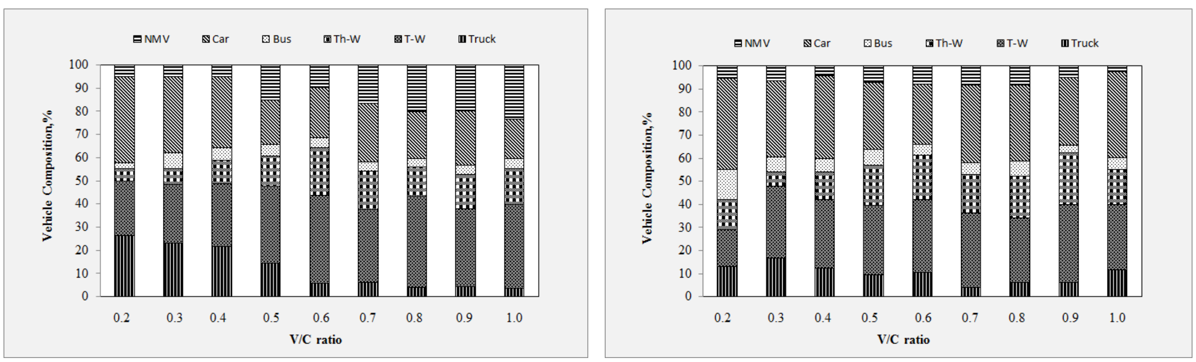
Field studies were carried out to observe the time headways at different flow scopeson a National Highway (two-lane highway) that passes through the city outskirts or peri-urban area of Agartala, the capital city of Tripura, India. A 20 km highway section that connects the city at its western end was selected and eight segments were identified on the section for conducting the study. Therefore, the directional movements of traffic were termed as west bound (city bound) and east bound traffic respectively. Study segments were selected in such a way so that they are free from the effect of intersection, curvature, ribbon development and also, pavement conditions were good and uniform (see FIGURE 1).

Video photographic survey technique was adopted to capture the traffic data appropriately. A reference line was marked on the pavement and two observation points were chosen for installing the video cameras; one in each direction, for recording the entry and exit of front and rear wheels of vehicles. The video files were then played on a computer to extract the traffic data (Saha et al. 2015). The necessary readings i.e. vehicle type, axle configurations and time were noted down independently for the two specified locations when a vehicle just crosses the reference line. The extracted traffic data was then processed and the time headway of vehicles was computed. The class interval for the observed data set wasassessed using Struge's rule (Sturges 1926; Scott 2009); determination of class interval is necessary for obtaining a smooth histogram.

The traffic composition consists of motorized two and three-wheelers, bus, truck, car and also non-motorized ones including paddle tri-cycle (Fig.2). A close look into the figure reveals that

Proportion of two-wheeler and car is significant; in the range of about 20-35 and 20-40 percent respectively. Also, city bound traffic consists significant proportion of non-motorized vehicles; in a way that they share around 20 percent of total traffic. An informal opinion poll of about 50 road users indicates that lack of adequate public transport facilities compel them to use para-transit modes like paddle tri-cycle or personalized transportation like car, motorized two-wheeler etc.

FIGURE 1A view of the study section (image by author)



(a) (b)

FIGURE 2Vehicle composition at the study sites: (a) West bound traffic and (b) East bound traffic

STUDY DESIGN

Modelling time headway distribution: Methodology

Development of an appropriate headway distribution model calls for an approach that would examine the statistical models in regards to the observed headways. Negative exponential distribution is normally used to describe the headway distribution pattern of highways. However, a number of researchers have reported the use of several other models in order to explain the headway distribution pattern more explicitly. Al-Ghamdi(2001) observed that Erlang and Gamma distributions provide a decent fit for a large range of flows whereas lognormal distribution was found to have a theoretical connection to the car-following models. Gamma distribution is widely used as headway model because of its flexibility and compatibility (Zhang et al. 2007)and Erlang distribution has wide applicability because of its relation to the exponential and gamma distributions.

Statistically, the headway data can be described as exponential if the co-efficient of variation is unity (Arasan and Khosy2003; Al-Ghamdi 2001). However, when it deviates from unity the headway distribution pattern could be described either by hypo-exponential if such value is less than one or by hyper-exponential type if it exceeds unity (Bolch et al. 2006). Typically, Erlang distribution represents hypo-exponential distribution function(Bolch et al. 2006)and hyper-exponential distribution can be well represented by mixed distributionfunction (Luttinen1996; Singh et al. 2007).As the scheme of the study was to develop headway distribution models, data corresponding to an extensive range of traffic flow were fitted to different types of distributions which meet the required traits and the following sections provide a detail insight of it.

An application of shifted headway distribution model

Hypo-exponential distribution appears to be appropriate if the co-efficient of variation of observed data is less than one.Erlang distribution is found to have similar kind of applicability like hypo-exponential distribution in stochastic process(Bolch et al. 2006). Application of single distribution models is easy and simple;however, they have limited capabilities in approximating shorter headways. The concept of shifted single distribution models is, therefore, applied in order to improve the accuracy of single distribution models(Zhang et al., 2007). *Eq. 1 & 2*demonstrate the cumulative density functions of shifted negative exponential and shifted erlang distributions indicating a shift of $'\tau'$seconds to the right. Investigations were accordingly made with these functions with a range of shifts (0 to 3 s)keeping a step of 0.015 s for determining the shift against which a function gives the best result.

Shifted negative exponential distribution,

$F(h) = 1 - exp\ (-\lambda(h - \tau))$(1)

Shifted erlang distribution,

$F(h) = \frac{\Gamma(h-\tau)/(m)}{\Gamma(m)})$(2)

Where: h = headways, λ= continuous inverse scale parameter (λ>0), m=shape parameter (positive integer), β= continuous scale parameter (Erlang distribution) (β> 0)

An approach of combined headway distribution model

On two-lane roads, the passing opportunities and number of widely spaced vehicles start diminishingwith the increase of flow rate. The situation further aggravates in the event of heterogeneity and sizable proportion of slower vehicles in traffic mix resulting in significantamount of shorter headways particularly at moderate and heavy flow. There have been a number of researchers who proposed the concept of mixed distribution models for describing headways at such situation: they are‘double displaced negative exponential distribution’ (Griffiths and Hunt 1991), ‘generalized queuing model’(Branston 1976), ‘Cowan M3 and M4 model’ (Zhang et al. 2007). *Eq. 3 to 5*demonstrate the mixed distribution functions that are used in traffic engineering studies:

Double Displaced Negative Exponential Distribution (DDNED),

$$f(h) = \begin{cases} \phi * \Upsilon_1 * \exp(-\Upsilon_1 * (h-d)) + (1-\Phi) * \Upsilon_2 * \exp(-\Upsilon_2 * (h-d)) & h \geq d \\ 0 & h < d \end{cases} \quad (3)$$

Generalized Queuing Model,

$$f(h) = \theta * g(h) + (1-\theta) * \Upsilon * \exp(-\Upsilon * h) * \int_{o}^{h} g(x) * \exp(\gamma * x)\, dx (4)$$

Cowan M3 distribution,

$$f(h) = (1-\Phi)\delta(h-\tau) + \Phi\lambda e^{-\lambda(h-\tau)}u(h-\tau) (5)$$

Where: *f (h)* is the probability density function; ϕ is a weighting factor constrained by $0 < \phi \leq 1$; $\Upsilon_1$andΥ_2 are constants associated with the flow status; and d is a displaced parameter; θ = proportion of vehicles tracking at the minimum headway; $\delta(\mathrm{h}-\tau)$ =unit impulse (Dirac delta) function

The premise on which the concept is based assumes that headway consists of two components: 'following'and 'free'. Accordingly, mixed headway distribution models are supposed to better capture headway dynamics. However, in the event of mixed traffic with slower vehicles, calibration and parameter estimation of such distribution model is somewhat difficult because of its complicated structure (Zhang et al. 2007).This is particularly true for mixed traffic situation where a variety of vehicles in terms of static and dynamic characteristics is observed.

Accordingly, to overcome this difficulty, combined distribution function was considered as an alternative of mixture distribution while describing headways at moderate and heavy flow. The simplistic analytical approach considers modelling shorter and longer headways separately.Partitional clustering technique exhibits its aptness in identifying the headway threshold. At car-following situationsLognormal and Gamma distributions seem to be appropriate because of their flexibility and compatibility(Zhang et al. 2007) andnegative exponential model describes well the widely spaced data. Therefore, Lognormal and Gammadistributions were applied on headway data for describing shorter headways and they all were applied for longer headways. The combined distribution, thus, represents 'following' and 'free 'components of headways separately on the basis of estimated headway threshold $'h'$ as displayed in *Eq. 6-9*.

Combined lognormal and gamma distributions,

$$F(h) = \int_{o}^{h} \Phi\left(\frac{\ln h-\mu}{\sigma}\right) dh + \int_{h}^{\alpha} \frac{\Gamma h/\,(\alpha)}{\Gamma(\alpha)} dh (6)$$

Combinedgamma and lognormal distributions,

$$F(h) = \int_{o}^{h} \frac{\Gamma h/\,(\alpha)}{\Gamma(\alpha)} dh + \int_{h}^{\alpha} \phi\left(\frac{\ln h-\mu}{\sigma}\right) dh (7)$$

Combined exponential and gamma distributions,

$$F(h) = \int_{o}^{h} (1-\exp(-\lambda h))dh + \int_{h}^{\alpha} \frac{\Gamma h/\,(\alpha)}{\Gamma(\alpha)} dh (8)$$ Combined exponential and lognormal distribution,

$$F(h) = \int_{o}^{h} (1-\exp(-\lambda h))dh + \int_{h}^{\alpha} \phi\left(\frac{\ln h-\mu}{\sigma}\right) dh (9)$$

Where: μ= continuous location parameter, σ= continuous scale parameter (Lognormal distribution), α= continuous shape parameter (α> 0)

Selection of appropriate headway model: goodness-of-fit test

The basis of selecting a particular distribution function is the extent it 'fits' the observed data; this compatibility is usually measured in terms of Goodness-of-fit tests. In traffic engineering problems, two such tests are quite common: Chi-square test and the Kolmogorov–Smirnov (K-S) test. However, the advantage that the K-S test offers over the chi-square test is that the K-S test can use data with a continuous distribution and there is no minimum frequency per test interval(Ye and Zhang 2009).Thus, the present study used K-S test while identifying the best fitted models. The test statistic is the difference between the cumulative percentage of the measured frequency and the cumulative percentage of the expected frequency. The test statistic 'D' is the largest of these differences over the entire measured population (Lilliefors 1967). Also,'quantile-quantile' plots helpinvisualizing the goodness-of-fit graphically by comparing their observed and estimated quantiles.

ANALYSIS AND RESULTS

Descriptive statistics of headway data

The possible capacity of the highway section was observed to be around 2300 pc/h (Saha et al. 2015). Headway data was collected considering directional segments separately and a wide range of traffic volume corresponding to volume to capacity ratio of 0.2-1.0 was covered. *Table 1*displays the descriptive statistics of the headway data. A careful examination of *Table 1* indicates thatat all flow scopes the median is less than mean and standard deviation decreases with the increase of flow. It signifies concentration of shorter headways and is attributed to high risk-ability of driver population which eventually lessens safety of traffic to a great extent.

Besides, the mean and standard deviation should result a 45^0plot i.e. co-efficient of variation should be unity in case of negative exponential distribution. Empirical investigation, however, indicates a deviation thereby, exhibiting inappropriateness of negative exponential model in describing headways. Moreover, co-efficient of variation was lessthan unity upto a flow level that corresponds to a volume to capacity ratio of 0.3 for both the directions and exceeds the value at moderate and heavy flow (volume to capacity ratio 0.4-1.0).

TABLE 1 Descriptive statistics of headways at different flow level

v/c ratio	West bound traffic			East bound traffic		
	Mean	Median	Standard deviation	Mean	Median	Standard deviation
0.2	14.87	10.50	12.65	10.03	4.50	9.22
0.3	11.57	7.50	10.64	10.10	7.50	9.18
0.4	8.12	4.50	8.52	8.75	4.50	9.02
0.5	6.70	4.50	6.81	7.27	4.50	8.22
0.6	5.70	4.50	6.11	6.00	1.50	7.05
0.7	4.86	1.50	5.54	4.69	1.50	5.05
0.8	4.49	1.50	4.92	4.50	1.50	5.17
0.9	3.97	1.50	4.21	3.60	1.50	5.12
1.0	3.66	1.50	4.10	3.55	1.50	4.78

Headway distribution at low flow level

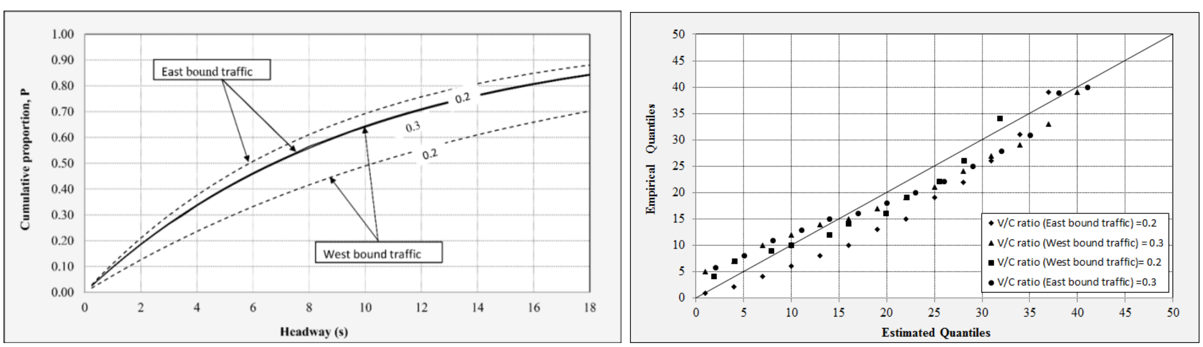
Shifted erlang and negative exponential distributions were applied to the headway data when the co-efficient of variation was less then unity. Attempts were made with shifts ranging from 0 to 3 seconds (with a step of 0.015 seconds) while examining the suitability of the selected distributions. A comparison of the K-S test statistic obtained with 5% level-of-significance for the selected models was made in order to examine the extent of fit to the observed data. The test statistic was observed to decrease with the increase in shifts, which however starts increasing beyond an optimal value. Accordingly, the trend discerns an optimal shift for the fitted models.*Table2*exhibits the estimated parameters of distribution models and K-S test results.Erlang distribution model with shifts of 0.675, 1.41 and 2.76 seconds and exponential distribution with 0.495 second shift were found to have acceptable statistical validity in terms of K-S test statistic.

TABLE2Goodness-of-fit details and estimated parameters of the fitted distributions

v/c ratio	K-S test (D-value)		Shift value(Second)		Fitted distribution	Estimated parameters
	Exponential	Erlang	Exponential	Erlang		
0.2*	0.**0942**	0.0954	**0.4950**	2.820	Shifted Exponential	λ =0.0673
0.3*	**0**.1328	**0.1312**	2.595	**1.410**	Shifted Erlang	m=1; β=9.8681
0.2@	0.1579	**0.1507**	2.955	**0.675**	Shifted Erlang	m=1; β=8.5335
0.3@	0.1735	**0.1624**	2.790	**2.760**	Shifted Erlang	m=1; β=9.5023

*West bound traffic; @East bound traffic; λ: continuous inverse scale parameter; m: shape parameter (positive integer); β: continuous scale parameter; Bold values indicate smallest D-values

Figure 3(a) displays the cumulative density function of the selected models and it indicates that proportion of shorter headways at low flow (volume to capacity ratio 0.2) is quite less for city bound traffic which, however, starts increasing when the flow increases. On the contrary, east bound traffic points to a diametric characteristic wherein proportion of shorter headways was significant at such flow level. This is attributed to the fact that, during morning city bound traffic was quite less and it starts increasing as the day progresses while trips to city outskirt was sizable during morning and thenceforth it decreases. *Figure 3(b)* displays that the data points clustered about a 45 degree plot signifying satisfactory agreement between theoretical model and the distribution of the observed data.



(a) (b)

FIGURE 3Shifted headway distributions at low flow:
(a) Cumulative density functions and (b) Q-Q plot

Headway distribution at moderate and heavy flow

Proportion of shorter headways was observed to be considerably high at moderate and high flow level. It was, therefore, important to describe shorter and longer headways separately. Accordingly, the field data collected at different flow scopes was grouped into cluster or subsets such that the data in each subset are similar to each other (Vimal et al. 2008).The similarity of data set is defined by a distance measure; this plays an important role in obtaining correct clusters (Vimal et al. 2008). The selection of distance measures, however, depends on the type of data. In context to the present study, the observed data set is of interval type as it contains a range of continuous values (Peng and T 2006). Therefore, Euclidean distance measure was adopted in the present study as the measure corresponds to interval data type. The number of clusters was selected in view of shorter and longer headways and then partitioning algorithm was used for identifying the boundaries. The mean of each cluster was used to determine its centroid. Accordingly, the k-mean algorithm was applied adopting Euclidean distance measure as an effective heuristic method of partitionalclustering(Kanungo et al. 2002). Subsequently, the dataset was grouped into two clusters and the limiting value of shorter headways was found to be 4 s.Similar value was also obtained by a study conducted in India (Dey et al. 2006).

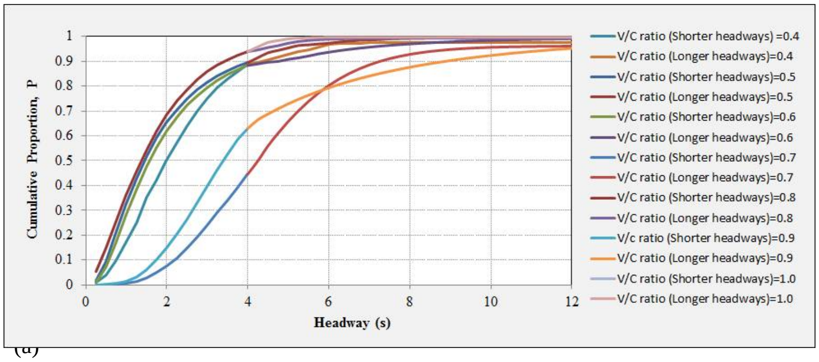
The present study introduces the concept of combined distribution wherein appropriate models for shorter and longer headways could be used to describe the headways. The present study considered lognormal and gamma distribution models to describe the shorter headways and lognormal, gamma and negative exponential distribution models to describe the larger headways. *Table 3*elucidates the goodness-of-fit details of the selected models and the calibrated model parameters. Since the limiting value of shorter headway was 4 seconds, shorter headways were modelled considering headway data up to 4 s and the remaining headways were modelled as longer headways separately. Best fitted models were selected based on K-S test results (see TABLE 3).

TABLE3 Goodness-of-fit details and estimated parameters of the selected distributions

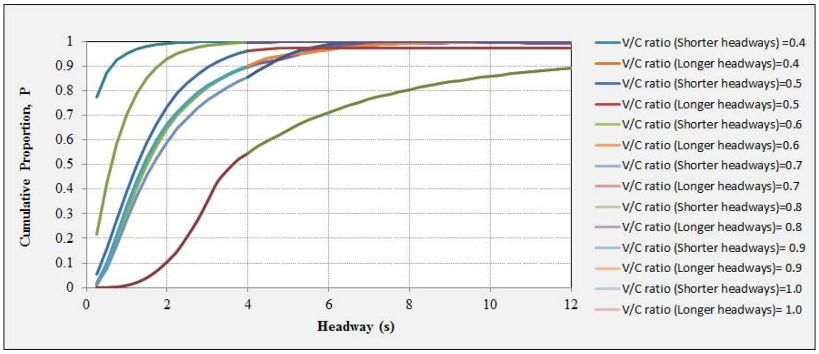
	v/c ratio	Shorter headways		Longer headways			Estimated parameters	
		Lognormal	Gamma	Lognormal	Gamma	Exponential	Shorter headways	Longer headways
West bound traffic	0.4	0.19345	**0.17950**	0.1428	**0.1383**	0.4540	α= 2.6448; β= 0.8611	α= 7.4232; β= 0.3789
	0.5	**0.23205**	0.24084	**0.1172**	0.1187	0.4556	μ= 0.3695; σ= 0.8114	μ= 0.9602; σ= 0.4194
	0.6	**0.20605**	0.20675	**0.1427**	0.1432	0.4577	μ= 0.4549; σ= 0.7817	μ= 0.5423; σ= 0.8168
	0.7	0.26508	**0.25685**	0.1632	**0.1622**	0.4667	α= 4.8001; β= 0 .9544	α= 6.026; β= 0.7138
	0.8	0.27781	**0.26696**	0.1327	**0.1289**	0.4666	α=1.7022; β= 0.9918	α= 6.832; β= 0.3789
	0.9	0.26713	**0.26529**	**0.1443**	0.1475	0.4749	α=4.8351; β= 0.7539	μ= 1.0985; σ= 0.8996
	1.0	0.27597	**0.25811**	**0.1730**	0.1751	0.4803	α=1.6505; β= 1.0135	μ= 1.1174;σ= 0.1985
East bound traffic	0.4	**0.21253**	0.21348	0.1422	**0.1407**	0.4462	μ= 0.4026; σ= 0.7822	α=6..531;β= 0.4603
	0.5	**0.23211**	0.23748	**0.1481**	0.1482	0.4721	μ= 0.3553;σ= 0.8130	μ= 0.9055; σ= 0.4837
	0.6	**0.20523**	0.22213	0.2497	**0.2256**	0.5080	μ= 0.5018;σ= 0.8401	α= 7.0215; β= 0.4232
	0.7	0.42970	**0.40975**	**0.1388**	0.1411	0.4746	α= 4.9778;β= 0.8171	μ= 1.2035; σ= 0.8761
	0.8	0.27278	**0.25856**	**0.1340**	0.1366	0.4550	α= 0.2585; β= 0.7893	μ= 0.8653; σ= 0.3621
	0.9	0.28101	**0.26112**	0.2217	**0.2192**	0.4597	α= 1.8156;β= 0.8442	α= 8.6817;β= 0.2419
	1.0	0.26135	**0.23247**	0.1956	**0.1472**	0.4632	α= 1.2463; β= 0.6542	α= 6.6846; β= 0.3683

[Bold values indicate smallest D-values]

Selected best fitted distributions for shorter (following component) and longer (free component) headways were accordingly combined (see Eq. 6-9) and applied for describing the data set. *Fig 4* illustrates the cumulative density function of combined models and indicates that the probability of shorter headways is quite high when compared to longer headways. This attributes to the fact that at moderate and high flow, platoon formation is frequent and car following interaction is quite high. The quantile plots of the combined models are shown in *Fig 5.* It is observed that the data points are close to the straight line signifying satisfactory agreement between theoretical model and the distribution of the observed data.

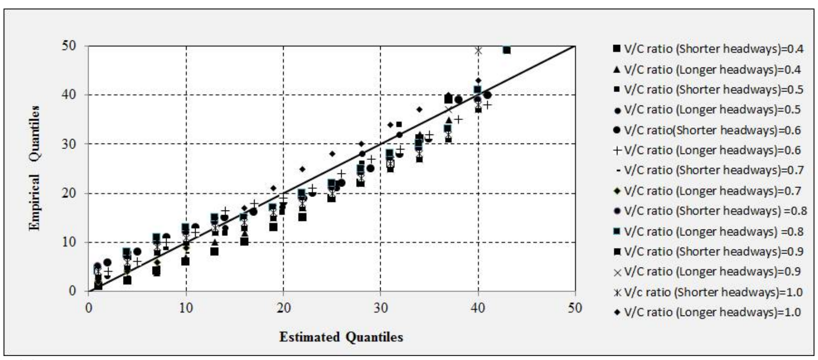


(a)

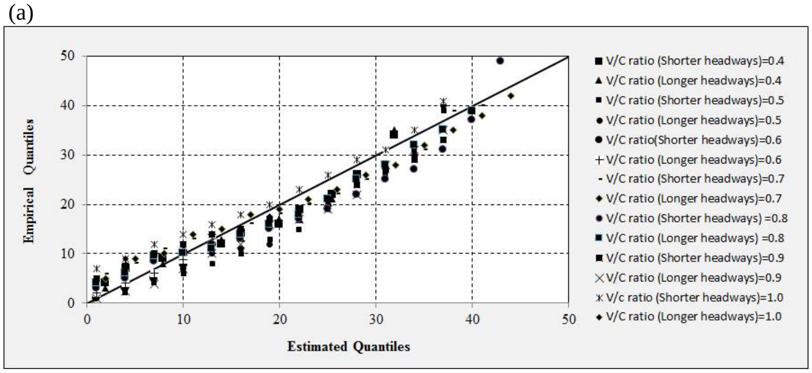


(b)

FIGURE 4 Cumulative distributions of combined models: (a) West bound and (b) East bound traffic



(a)



(b)

FIGURE 5Quantile-Quantile plots of combined models: (a) West bound and (b) East bound traffic

CONCLUSIONS

The present study was performed to investigate the time headway distributions ona two-lane highwaythat passes through moderately developed peri-urban area and exhibit extreme heterogeneity in its trafficmix. The driver's behaviours in choosing the headway were analysed for a wide range of traffic flow that corresponds to volume to capacity ratio of 0.2 to 1.0. It was observed that at car-following conditions, a large number of drivers adopt headways which are less than safe headway of 2 s (Luttinen 1992). Statistical parameters were computed and observed that at all flow scopes median is less than mean; it reveals the concentration of shorter headways. This is because of a high risk-ability of driver population which results in safety reduction. Further, empirical investigation reveals that co-efficient of variation deviates owing to the heterogeneity effect and makes exponential model inappropriate in describing headways. Hypo-exponential model is considered suitable when co-efficient of variation is less than one whereas hyper-exponential model is considered statistically valid when it exceeds unity.

The present study used shifted erlang distribution as hypo-exponential and combined distribution as hyper-exponential. Erlang distribution with shifts of 0.675, 1.41 and 2.76 seconds and exponential distribution with 0.495 second shift exhibit acceptable statistical validity in terms of K-S test statistic. Quantile plot of observed and estimated headways also explicates a satisfactory agreement. Combined distribution function was found as an alternative of mixture distribution while describing headway data at moderate and heavy flow; shorter and longer headway were modelled separately by simplistic approach. Clustering technique was adopted to group the data into shorter and longer headways. K-mean algorithm with Euclidean distance measure was selected for headway grouping. Lognormal and gamma distribution models were selected to describe the shorter headways whereas lognormal, gamma and negative exponential models were selected to describe the longer headways to identify the best fitted models. The Kolmogorov Smirnov (K-S) test with 5% level-of significance was used to find the appropriate models of headway data and further examined by the quantile plots as well. It has been observed that both shorter and longer headways can be described well using lognormal distribution up to moderate flow level. However, in case of high and congested flow Gamma distribution was found to have best fit with the observed headway data; this could be attributed to its flexibility and compatibility.

While shifted erlang distribution indicates insignificant proportion of shorter headway at low flow, cumulative density function of combined models confirms that at moderate and heavy flow probability of shorter headways is quite high when compared to longer headways. This attributes to the fact that at such flow levels, platoon formation is frequent and car following interaction is quite high. The present study, thus, creates a starting point of further initiativesaimed at establishing a robust method of modelling headways on two-lane highways that pass through city outskirts or peri-urban areas and exhibit extreme heterogeneity in its traffic mixbased on comprehensive fielddata.

REFERENCES

1. Abtahi, S. M.; Tamannaei, M.; Haghshenash, H. 2011. Analysis and modelling time headway distributions under heavy traffic flow conditions in the urban highways: case of Isfahan. *Transport* 26(4):375-382.doi: 10.3846/16484142.2011.635694
2. Al-Ghamdi, A. 2001. Analysis of time headways on urban roads: Case study from Riyadh, *Journal of Transportation Transportation Engineering* 127(4): 289-294.http://dx.doi.org/10.1061/(ASCE)0733-947X(2001)127:4(289)
3. Arasan, V. T.; Koshy, R. Z. 2003. Headway distribution of heterogeneous traffic on urban arterials, *Journal of Institution of Engineers(India):Civil Enigineering Division* 84: 210-215.
   a. http://cat.inist.fr/?aModele=afficheN&cpsidt=15411224
4. Bham, G. H..; Ancha, S. R. P. 2006. Statistical models for preferred time headway and time headway of drivers in steady state car following, in *Proceedings of the Ninth International Conference*, Chicago, Illinois, USA, Vol. 213: 344-349.
5. Bolch, G.; Greiner S.; De, M. H.; Trivedi, K. S.2006.*Queuing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications* .Wiley-Blackwell.
6. Branston, D. 1976. Models of Single Lane Time Headway Distributions.*Transportation Science* 10(2) :125-148.
7. Chandra, S.; Kumar, R. 2001. Headway modelling under mixed traffic on urban roads. *Road Transport Research* 10 (1): 61-

71. https://trid.trb.org/view.aspx?id=732752

8. Dey, P. P; Chandra, S.2009. Desired Time Gap and Time Headway in Steady-State Car-Following on Two Lane Roads. *Journal of Transportation Engineering*, Vol. 135, No. 10, 2009, pp. 687-693.
   a. doi:10.1061/(ASCE)0733-947X_(2009)135:10(687)
9. Dey, P. P.; Chandra, S.; Gangopadhaya, S.2006. Speed distribution curves under mixed traffic conditions. *Journal of transportation engineering*132(6):475-481.doi: 10.1061/_ASCE_0733-947X_2006_132:6_475_
10. Griffiths, J. D., and Hunt, J. 1991. Vehicle headways in urban areas. *Traffic Eng. Control* 32(10): 458-462.
    a. http://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds%5B%5D=citjournalarticle_63466_34
11. Hoogendoorn, S.; Bovy, P.1998. New estimation technique for vehicle-type-specific headway distributions. *In Transportation Research Record: Journal of the Transportation Research Board*1646: 18-28.http://trrjournalonline.trb.org/doi/abs/10.3141/1646-03
12. Hoogendoorn, S.; Hein, B.1997. Modelling and estimation of headway distributions. In *Transportation Research Record: Journal of the Transportation Research Board*1591*:*14-22.
13. Hossain, M.; Iqbal, G. A. 1999. Vehicular headway distribution and free speed characteristics on two-lane two-way highways of Bangladesh. *Journal of Institution of Engineers (India)* 80: 77-80..
    a. http://cat.inist.fr/?aModele=afficheN&cpsidt=1970688
14. Kanungo, T.; Mount, D. M.; Netanyahu, N.S.; Piatko, C.D.; Silverman, R.; Wu, A.Y.2002.An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence 24*(7): 881-892.doi: 10.11109/TPAMI.2002.1017616
15. Khasnabis, S..;Heimbach, C. L .1980. R Headway distribution models for two-lane rural highways, *Transportation Research Record*. 772, *Transportation Research Record* 772: 44-51. Available from Internet:
16. Kumar, M.; Rao. S. K.1998. Headway and speed studies on two-lane highways. *Indian highways: Indian Roads Congress* 5: 23-36.
    a. http://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds%5B%5D=citjournalarticle_239754_38
17. Lilliefors, H.W.1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown,*Journal of the American Statistical Association* 62 (318): 399-402.
18. Luttinen, R. T.1992. Statistical properties of vehicle time headways. In *Transportation Research Record: Journal of the Transportation Research Board* 1365: 92-107.http://onlinepubs.trb.org/Onlinepubs/trr/1992/1365/1365-010.pdf
19. Luttinen, R. T.1994. Identification and estimation of headway distributions. *Proceedings of the second International Symposium on Highway Capacit*y, Sydney, Australia.
20. Luttinen, R. T. 1996. Statistical Analysis of Vehicle Time Headways: D.TechThesis.Helsinki University of Technology, Finland. Available from Internet:http://lib.tkk.fi/Diss/199X/isbn951228474X/isbn951228474X.pdf
21. Luttinen, R.T. 1999. Properties of Cowans M3 headway distribution. In *Transportation Research Record: Journal of the Transportation Research Board* 1678: 189–196.http://trrjournalonline.trb.org/doi/abs/10.3141/1678-23
22. Mei, M.; Bullen, G.1993. Lognormal distribution for high traffic flows.In*Transportation Research Record: Journal of the Transportation Research Board.*1398: 125-128.https://trid.trb.org/view.aspx?id=383214
23. Panichpapiboon, S.2014. Time-Headway Distributions on an Expressway: Case of Bangkok. *Journal of Transportation Engineering* 141(1):1-8.
24. Peng, W.; Li, T.2006. Interval Data Clustering with Applications. *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence* 355-362.
25. Penmetsa, P.; Ghosh, I.; Chandra, S.2015. Evaluation of Performance measures for Two-Lane Intercity Highways under Mixed Traffic Conditions.*Journal of Transportation Engineering* 141(10):1-7.http://ascelibrary.org/doi/abs/10.1061/(ASCE)TE.1943-5436.0000787
26. Rossi, R.;Gastaldi, M.; Pascucci, F.2013. Gamma-GQM time headway model: endogenous effects in rural two-lane two-way roads. *Procedia - Social and Behavioral Sciences*, *Elsevier* 859-868.
27. Saha, P.; Sarkar, A. K.; Pal, M. 2015. Evaluation of speed-flow characteristics on two-lane highways with mixed traffic. *Transport- Taylor & Francis*1-9. doi: 10.3846/16484142.2015.1004369
28. Saha, P.; Roy, R.;Sarkar, A. K.; Pal, M. 2017.Preferred time headway of drivers on two-lane highways with heterogeneous traffic.*TransportationLetters:The International Journal of Transportation Research* http://dx.doi.org/10.1080/19427867.2017.1312859
29. Scott, D.W.2009. Sturges' rule, *Wiley Interdisciplinary Reviews: Computational Statisti*cs.1(3):303–306. http://onlinelibrary.wiley.com/doi/10.1002/wics.35/pdf
30. Singh, L. N.; Dattatreya, G. R. 2007. Estimation of the Hyperexponential Density with Applications in Sensor Networks. *International Journal of Distributed Sensor Networks*3(3): 311-330. doi:10.1080/15501320701259925
31. Sturges, H. A. 1926. The Choice of a Class Interval.*Journal of the American Statistical Associationn*21(153): 65–66. http://dx.doi.org/10.1080/01621459.1926.10502161
32. Vimal, A.; Valluri, S. R.; Karlapalem, K. 2008.An Experiment with Distance Measures for Clustering. *International Conference on Management of Data COMAD*,Mumbai, India.
33. Ye, Fan.; Zhang, Y. 2009. Vehicle type-Specific headway analysis using freeway traffic data, *Transportation Research Record: Journal of the Transportation Research Board* 2124: 222-230. http://dx.doi.org/10.3141/2124-22
34. Zhang, G.; Wang, Y.; Wei, H.; Chen, Y.2007. Examining headway distribution models with urban freeway loop event data, *Transportation Research Record: Journal of the Transportation Research Board* 1999: 141–149. http://dx.doi.org/10.3141/1999-15