Analysis of Time Headway Distribution on Korean Multilane Highway Using Loop Event Data

Jinhwan JANG  
Research Specialist  
Advanced Transport Research Division  
Korea Institute of Construction Technology  
1190, I Ilsan, Goyang, 411-712, South Korea  
Fax: +81-31-910-5694  
E-MAIL: jhjang@kict.re.kr

Byunghwa KIM  
Research Fellow  
Advanced Transport Research Division  
Korea Institute of Construction Technology  
1190, I Ilsan, Goyang, 411-712, South Korea  
Fax: +81-31-910-5694  
E-MAIL: bhim@kict.re.kr

Namkuk CHOI  
Administrative Official  
Ministry of Land, Transport, and Maritime Affairs  
18, Joong, Seoul, 100-120, South Korea  
Fax: +81-2-2125-2649  
E-MAIL: namkug@korea.kr

Byunghwa KIM  
Research Fellow  
Advanced Transport Research Division  
Korea Institute of Construction Technology  
1190, I Ilsan, Goyang, 411-712, South Korea  
Fax: +81-31-910-5694  
E-MAIL: bhim@kict.re.kr

Namchul BAIK  
Research Fellow  
Advanced Transport Research Division  
Korea Institute of Construction Technology  
1190, I Ilsan, Goyang, 411-712, South Korea  
Fax: +81-31-910-5694  
E-MAIL: nc100@kict.re.kr

Abstract: Vehicular time headway is of fundamental importance in traffic engineering. This research investigated theoretical headway models for a multilane highway in Korea using loop detector data. To analyze time headway by traffic flow states, the collected data were categorized into five flow rates and tested for randomness. Subsequent theoretical modeling was performed at the five flow rates. The Gamma distribution at low to intermediate flow rates, and the Pearson VI distribution at all flow rates were fitted by the Kolmogorov-Smirnov (K-S) statistic. Model parameters were obtained from conventional maximum likelihood estimation (MLE). Analyses were performed on residuals and parameters of the fitted models, as well as various headway statistics. Notably, parameters of the Pearson VI changed markedly during level of service (LOS) C, which presumably reflects the boundary between low and high traffic flows.

Key Words: headway distribution, multilane highway, Gamma distribution, Pearson VI distribution

1. INTRODUCTION

The distribution of time headway has been studied for many years. Time headway, defined as the elapsed time between successive vehicles in a single lane of traffic, affects safety, level of service (LOS), and capacity analysis. From the perspective of safety, the minimum time headway must be obeyed, in case the leading vehicle suddenly stops. Regarding the LOS, one index is the percentage of vehicles in platoon at headways less than 5 s (A. D. May, 1990). The time headway distribution determines the opportunities for merging and crossing at intersections or ramps. Capacity and saturation flow rate are reciprocals of the minimum time headway. Furthermore, accurate modeling of the time headway distribution is of fundamental importance for vehicle generation in traffic simulation models. A number of traffic simulation models have been developed to address different traffic problems. A key component of assessing the performance of simulation models is the generation of vehicle arrival times as input. Hence, simulation researchers have devoted considerable effort to mathematical models that accurately describe time headway.
Numerous theoretical models describing time headway distributions have been derived. Hoogendoorn et al. (1997 and 1998) proposed a new technique for parameter estimation of time headway models based on Fourier-series analysis, and estimated parameters for vehicle type-specific headway distributions with a Pearson-III-based generalized queuing model (GQM) using time headway data from two-lane rural highways. Luttinen (1999) described the principal statistical properties of the Cowan M3 time headway distribution model and evaluated the accuracy of moments and least-squares estimation methods with headway data from Finnish two-lane roads. Al-Ghamdi (2001) found that a gamma distribution for unsignalized urban arterials, and negative exponential, shifted exponential, or Erlang distributions for freeways provided good fits after analyzing substantial time headway data. He also defined the boundaries of traffic flow states by analyzing the parameters of the distributions. Pueboobpaphan et al. (2003) examined time headway distributions of probe vehicles with AVI tags on freeways in Houston. Arasan et al. (2003) performed an in-depth study on time headway distributions of mixed-traffic unaffected by nearby intersections and dominated by smaller vehicles such as motorized two-wheelers. Zhang et al. (2007) conducted a comprehensive study on time headway distribution models using headways observed on regular and HOV lanes of freeways in the Seattle area. Ha et al. (2010) modeled time headway distributions in different contexts, such as the type of lane, traffic flow, period of day, and change of cross-profile. They employed the composite time headway distribution, Generalized Queuing Model, and the Log-normal Model using event data from loop detectors on the A6 motorway in the south of Paris in France.

As the overview above shows, there have been many studies on time headway distribution. However, there is little research on time headway distributions for various flow states, followed up by detailed analyses on features of the flow statistics and estimated parameters. Such analyses might derive meaningful conclusions on the features of traffic flow, and thus lead to the easy identification and understanding of certain characteristics of traffic flow. This study examines the characteristics of the time headway of a multilane uninterrupted highway. It proposes statistical distribution models for the highway from a considerable amount of accurate time headway data obtained with an inductive loop detector (ILD) at a sampling rate of 1/100 s. To analyze the time headway distribution for different flow conditions, the collected data were divided into five flow states. After exploring a large number of theoretical distribution models using Stat::Fit®, a widely used statistical package, a number of models fitted to each flow state were derived. Furthermore, the residuals, flow statistics, and estimated parameters of the derived models were investigated for different flow levels. As a result, meaningful conclusions on the boundary of low and high flow rates were drawn.

2. DATA COLLECTION

The Korea Ministry of Land, Transport, and Maritime Affairs (MLTM) deployed the Advanced Traveler Information System (ATIS) on a multilane uninterrupted highway named Jayuro, on which this study is based, to enhance traffic mobility and safety in 2009. The system was installed with loop detectors with 1.5-km spacing, variable message signs (VMSs), CCTVs, and road weather information systems (RWISs). The ATIS disseminates real-time information of travel time and inclement weather conditions via the VMSs, automatic response systems (ARSs), and the Internet; currently, a system is also being developed to provide this information through smart phones. In a data accuracy test of the detector using the baseline data source installed on the Korea detector test-bed (Jang et al., 2009), less than 2 % errors were observed for volume and speed data after a 3-day-long test.
An inductive loop detector (ILD) with a 100-Hz sampling rate, installed for gathering traffic data for the ATIS described above, was used to collect time headway data. After a 3-h long data verification test against volume data from video images, no errors were detected in the ILD data. Discrepancies were found only when straddling or lane-changing maneuvers occurred, which preferably should be removed from the time headway analysis. Assisted by the ILD, time headway data were collected for two consecutive days under clear weather conditions on the curb lane of a multilane uninterrupted highway in Gyeonggi province, South Korea, in May of 2010 (depicted in Figure 1).

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Flow rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60-300 v/h</td>
</tr>
<tr>
<td>Sample size</td>
<td>540</td>
</tr>
<tr>
<td>Minimum (s)</td>
<td>0.6</td>
</tr>
<tr>
<td>Maximum (s)</td>
<td>50.2</td>
</tr>
<tr>
<td>Mean (s)</td>
<td>12.7</td>
</tr>
<tr>
<td>Median (s)</td>
<td>9.8</td>
</tr>
<tr>
<td>Mode (s)</td>
<td>6.4</td>
</tr>
<tr>
<td>Std. deviation (s)</td>
<td>10.4</td>
</tr>
<tr>
<td>Coeff. of Var. (CV)</td>
<td>0.82</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.18</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.94</td>
</tr>
</tbody>
</table>

The study site is located about 1.5 km away from the nearest on-ramp, implying that no significant influence of ramp flow exists. The highway has four lanes (each 3.6 m in width) and a shoulder lane (2.5 m in width), level terrain, a design speed of 100 km/h, and a speed limit of 90 km/h. To examine the time headway distribution as a function of flow state, the data were collected at 5 v/m increments from 1 v/m (equivalent to 60 v/h) to 25 v/m (equivalent to 1,500 v/h). As shown in Table 1, the fundamental descriptive statistics of the collected data reasonably followed the general characteristics of time headway. The mode of the time headway is always less than the median, which in turn is always less than the mean, although these differences diminish with increasing traffic (A. D. May, 1990).

<table>
<thead>
<tr>
<th>Independence tests</th>
<th>Flow rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60-300 v/h</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>Max. positive</td>
</tr>
<tr>
<td></td>
<td>Max. negative</td>
</tr>
<tr>
<td>Runs test (p-value</td>
<td>Median test</td>
</tr>
<tr>
<td>at sig.=0.05</td>
<td>Turning point test</td>
</tr>
</tbody>
</table>

All of the procedures to fit the collected data to certain theoretical distribution models assume that the observed data are independent and identically distributed (IID), that is, each point is independent of all the other data points and all data points are drawn from identical distributions. To verify the IID assumption of the collected time headway data, two tests were conducted: an autocorrelation test and a runs test. The autocorrelation, which varies between -1 (negative) and 1 (positive), was analyzed to examine the dependence of the collected time headway data. The analysis showed that little correlation existed between the data points in all flow states, as verified in Table 2. The runs test verifies the randomness of data by investigating the occurrence of an uninterrupted sequence of numbers with the same attribute. This test also revealed the randomness of the data for all flow states, in which all p-values
were higher than the predefined significance level of 0.05 (Table 2). Hence, the fitting procedures were subsequently performed for all five flow states.

3. DATA ANALYSIS

After the data were collected at the site as stated above, they were analyzed to determine the appropriate distribution model that reasonably describes the observations. After investigating broad-ranging statistical distribution models with Stat::Fit®, two single distributions were selected for the flow states (see Table 3). Compared to mixed distribution models such as the Generalized Queuing model (mathematically consistent with Cowan M4) and the Double Displaced Negative Exponential Distribution (DDNED), single distribution models are simple and easy to apply. Moreover, the single models derived in this study fitted the collected headway fairly well, as represented in Table 4 by the high p-values for the Kolmogorov-Smirnov (K-S) test. Unlike well-known models for time headway distribution, a rather unfamiliar distribution, the Pearson type VI, was derived for all of the flow rates. The gamma distribution, a somewhat more familiar model, was also accepted for the flow rates of 60-300, 360-600, and 660-900 v/h.

The Pearson type VI distribution, which is bounded on the low side, is referred to as the Beta distribution of the second kind due to the relationship between a random variable of the Pearson VI with that of the Beta distribution. Like the Gamma distribution, it has three distinct regions. For p=1, it resembles the Exponential distribution, starting at a finite value at the minimum x and decreasing monotonically thereafter. For p<1, it tends to infinity at the minimum x and decreases monotonically with increasing x. For p>1, a range which includes all of the optimized values in this study, it is zero at the minimum x, peaks at a value that depends on both p and q, and decreases monotonically thereafter (Johnson et al., 1994). The Gamma distribution, which is bounded at the lower side, also has three different regimes according to the range of the parameter (α), just like the Pearson VI distribution. When α is large, it can be used to approximate the Normal distribution while strictly maintaining a positive value of x. The peak of the distribution moves away from the minimum value as α increases, implying a broader distribution. The Gamma distribution has been commonly exploited to represent lifetimes, inter-arrival times, service times, etc. (Johnson et al., 1995). Time headway data can be regarded as a special case of inter-arrival or service times. The probability density functions, expected values, standard deviations, and parameter explanations of the two models are represented in Table 3.
Table 3 Equations and parameters of the fitted distribution models

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Equations</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>( f(x) = \frac{(x - \text{min})^{\alpha - 1}}{\beta^{\alpha}} \exp \left[ \frac{(x - \text{min})}{\beta} \right] )</td>
<td>( \text{min} = \text{minimum x} )  ( \alpha = \text{shape parameter &gt; 0} )  ( \beta = \text{scale parameter} )</td>
</tr>
<tr>
<td>Pearson type VI</td>
<td>( f(x) = \frac{\beta}{\beta + (x - \text{min})}^{p+q} ) ( B(p, q) )</td>
<td>( x &gt; \text{min}, \beta &gt; 0, B(p, q) = \text{Beta function} )  ( p )  ( q = \text{shape parameter &gt; 0} )</td>
</tr>
</tbody>
</table>

Statistics

\( \text{Mean} = \alpha \beta, \text{Std. dev.} = \sqrt{\alpha \beta^2} \)

\( \text{Mean} = \frac{\beta p}{q - 1}, \text{Std. dev.} = \sqrt{\frac{\beta^2 p(p + q - 1)}{(q - 1)^2(q - 2)}} \)

Table 4 Goodness of fit tests and estimated parameters for the fitted distribution

<table>
<thead>
<tr>
<th>Flow rate (v/h)</th>
<th>Fitted distribution at sig. = 0.05</th>
<th>Estimated parameters</th>
<th>Goodness-of-fit-test (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>K-S test</td>
</tr>
<tr>
<td>60-300</td>
<td>Gamma</td>
<td>Min=0.6, ( \alpha = 1.2, \beta = 9.8 )</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Pearson type VI</td>
<td>Min=0.6, ( \beta = 146, p = 1.3, q = 16 )</td>
<td>0.22</td>
</tr>
<tr>
<td>360-600</td>
<td>Gamma</td>
<td>Min=0.5, ( \alpha = 1.1, \beta = 5.7 )</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Pearson type VI</td>
<td>Min=0.5, ( \beta = 866, p = 1.0, q = 155 )</td>
<td>0.64</td>
</tr>
<tr>
<td>660-900</td>
<td>Gamma</td>
<td>Min=0.4, ( \alpha = 1.3, \beta = 3.0 )</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Pearson type VI</td>
<td>Min=0.4, ( \beta = 13186, p = 1.3, q = 4478 )</td>
<td>0.41</td>
</tr>
<tr>
<td>960-1200</td>
<td>Pearson type VI</td>
<td>Min=0.5, ( \beta = 10.2, p = 1.3, q = 5.8 )</td>
<td>0.13</td>
</tr>
<tr>
<td>1260-1500</td>
<td>Pearson type VI</td>
<td>Min=0.3, ( \beta = 13.3, p = 1.2, q = 8.7 )</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Conventional maximum likelihood estimates (MLEs), which are generally known as the best unbiased estimates for single distribution models examined here (G. Zhang, 2007), were used to estimate parameters. In addition, goodness-of-fit tests were performed for the proposed models with the K-S and chi-square (\( \chi^2 \)) statistics at the 5% significance level (see Table 4). The estimation results reveal relatively good fits with the K-S test in most cases, whereas most distributions were rejected by the \( \chi^2 \) test, except for the Pearson VI at 1260–1500 v/h. This finding verifies that the K-S test that fits a cumulative distribution to the observations point by point, is more conservative than the \( \chi^2 \) test. That is, it is least likely to reject the fit in error (Mei et al., 1981). The highest p-value at 360–600 v/h and the lowest p-value at 960–1200 v/h indicate that the derived distribution fits best at 360–600 v/h and vice versa at 960–1200 v/h.
Figure 2 Histograms and density plots using estimated parameters for (a) 60-300 v/h, (b) 360-600 v/h, (c) 660-900 v/h, (d) 960-1200 v/h, (e) 1260-1500 v/h, and (f) cumulative probability functions

Histograms of the observed time headway and the fitted distributions with the estimated parameters for the each flow level are represented in Figure 2 (a-e). To compare distributions across flow states, the cumulative probability density functions are also plotted in Figure 2 (f). The cumulative distribution for the 60–300 v/h diverges from the others, implying that many vehicles are moving in a platoon even though the traffic volume slightly increases.

Plots of the residuals for the K-S tests (see Figure 3) reveal significantly low values in most cases except for the residuals near 20 s at a flow rate of 60–300 v/h. The expectations and standard deviations of the Pearson VI are also similar to the measured values, as plotted in Figure 3 (f). Moreover, the minimum headways of the fitted models exactly reproduced the observed values. Interestingly, all residuals tend to show higher values under lower headways and diminish as the flow increases. This means that the theoretical time headway distributions do not fit well under car-following states with shorter headways than in freely moving states.
It can be of great value to investigate the relationships among variables to further understand the time headway characteristics. In this study, the mean time headway (i.e., the reciprocal of flow) is highly correlated with the standard deviation of the time headway, as depicted in Figure 4 (a). Given that the latter can be directly obtained from observing the flow, the relationship between the mean and standard deviation of headway may be highly regarded.
As in previous studies showing a concave relationship between the coefficient of variation (CV) of the time headway and the flow on uninterrupted facilities (Al-Ghamdi, 2001; Luttinen, 1992), the present study also demonstrates this concave relationship with a relatively low determination coefficient, as shown in Figure 4 (b). However, compared to the results of other studies (Luttinen, 1992; Breiman, 1977; Buckley, 1968) in which the CV is near or above 1 over a broad range of flow rates, the CV in this study ranges from 0.77 to 0.92. However, further examination with sufficient samples is recommended because of the low correlation and small sample size.

The skewness and kurtosis of the observed headway were also analyzed according to flow rates. As shown in Table 1, the skewness lies in the range of 1.18–2.33, and the kurtosis lies between 0.94 and 7.56. As can be seen in Figure 5, the two statistics are highly correlated with correlation coefficient of 0.98. In general, the skewness of headway is known to decrease as traffic flow increases because the headway distribution approaches the Normal distribution as the flow increases (A. D. May, 1990). However, the skewness in this study shows an increasing pattern with growing traffic, which indicates that the headway distribution in this study diverges from the Normal distribution as traffic volume increases. On the other hand, the kurtosis at middle-range flows closely resembles that of the Normal distribution, with values around 3. As mentioned above, further investigations of these phenomena with more data are required because of the small sample size.
It would also be of interest to investigate the estimated parameters to examine the existence of any different features by flow states. In this study, the parameters of the Pearson VI distribution were thoroughly scrutinized. From Figure 6, \(\ln(\beta)\) and \(\ln(q)\), which are the scale and shape parameters of the Pearson VI, respectively, increase gradually until 900 v/h and then decrease markedly. This may reflect characteristic changes of the time headway in the range of 900–1200 v/h. As evidence supporting this assumption, the Gamma distribution fitted only up to the flow rate of 900 v/h (Table 4).

![Figure 6 Characteristics of the Pearson VI distribution parameters by flow states](image)

To examine this issue further, the volume-to-capacity ratio (v/c) of the present study site was calculated for each flow level. As noted above, the geometric condition of the site is ideal; hence, the only adjustment factor requiring application is a heavy vehicle factor. Using the observed traffic volume and the heavy vehicle percentage recorded in the Statistical Yearbook of Traffic Volume published by the Korean government (KICT, 2010), the v/c was computed according to the methodology presented in the Korea Highway Capacity Model (KHCM) (KOTI et al., 2001). Consequently, the parameters show a moderately increasing pattern, diminish substantially when v/c is in the range of 0.49–0.65 (approx. range of LOS C), and become stable thereafter. This presumably represents a change in traffic flow characteristics during the LOS C. In fact, the KHCM states that drivers are affected and at times restricted by other vehicles as from LOS C, unlike in free driving conditions below v/c 0.45 (i.e., the upper limit of LOS B).

### 4. CONCLUSIONS

This study proposed two single probabilistic distribution models fitted to time headway data collected on a multilane highway in Korea. The headway data, obtained with an ILD used for the ATIS, were categorized into five flow groups and were verified for their randomness by autocorrelation and runs tests. After exploring many statistical distributions with Stat::Fit®, the Gamma distribution at flow rates of 60–300, 360–600, and 660–900 v/h, and the Pearson type-VI distribution at all flow states including 60–300, 360–600, 660–900, 960–1200, and 1260–1500 v/h, were considered as the appropriate models.

Parameters for each model were estimated from MLE, generally known as the most unbiased estimation technique for single distributions. Goodness-of-fit tests were conducted by the K-S
and χ² tests. The K-S test accepted all proposed models with high p-values, whereas the χ² test rejected most of the models except for the Pearson VI at a flow rate of 1260–1500 v/h. This verifies that the K-S test, using a cumulative distribution for the fitting process, is a more forgiving goodness-of-fit test than the χ² test, which rejects a proposed model despite relatively high errors in classes with small sample sizes.

Key statistics of the collected headway data and the shape and scale parameters of the Pearson VI model were examined to understand further the features of the collected data. As a result, the mean time headway showed a strong linear correlation with the standard deviation of the time headway, and the flow versus CV showed a concave relationship analogous to findings from previous studies. In contrast to findings in the literature, however, the skewness of the observed headway increased with growing traffic, implying that the headway distribution diverges far from the Normal distribution as the flow increases. Given the small sample size in this paper, this aspect might be further analyzed with an adequate amount of data.

The investigation of ln (β) and ln (q), the scale and shape parameters of the Pearson VI, respectively, revealed that traffic flow characteristics change in the range of 900–1200 v/h (approx. the range of LOS C). In fact, in LOS C, driving conditions are affected or occasionally restricted by other vehicles according to the KHCM. Although this study was conducted with a substantial amount of accurate field data, the spatial transferability of the proposed distribution models and research findings should be verified with time headway data from other areas. With data from the RWISs, analyses on headway distributions under different meteorological conditions could also be highly recommended.

REFERENCES

D. J. Buckley, A semi-poison model of traffic flow (1968), Transportation Science, 2(2).
Luttinen, R. T., Statistical properties of vehicle time headways (1992), Transportation Research Record: Journal of the Transportation Research Board, No.1365, TRB, National Research Council, Washington D.C.