COMPARISON OF TIME HEADWAY DISTRIBUTIONS IN DIFFERENT TRAFFIC CONTEXTS

Application on a French Motorway

Duy-Hung HA

Maurice ARON

Simon COHEN

The French National Institute for Transport and Safety Research, INRETS

This paper concerns the modeling of time headway distributions on the A6 Motorway in France. The aim of the work is to compare these distributions in different contexts such as type of lane, traffic flow, period of the day and change of cross-profile. In order to interpret and to distinguish the different contexts and behaviors, the composite time headway distribution Generalized Queuing Model and the best simple model, the Log-normal Model, are calibrated. In addition, the statistical analysis of a series of events is applied to understand the time headway process. The distribution of instantaneous speeds enriches such interpretations in several cases. The results show a significant goodness-of-fit of the composite time headway distributions with most samples. The Log-normal Model provides also good adjustments for many samples particularly on the middle and fast lanes. Consequently, the sensitivity of the parameters of these models is yielded according to the contexts.

INTRODUCTION AND OBJECTIVE

Time headway (TH) is a fundamental variable in traffic theory and has been studied over the past decades. This variable is the temporal clearance elapsed between two consecutive arrival instants at a designated point of measure. This notion is distinguished from time headway on board that is measured by the rapport of distance between two consecutive vehicles and the speed of the second vehicle. The TH in this paper is measured by inductive loops located on the road. The research fields concerned range from theoretical calculation to engineering operations. Statistical analysis showed the possibility of describing vehicular arrival pattern by using some particular probabilistic distributions. Many probabilistic distributions have been proposed from simple to composite ones. The simple model is a probabilistic law which TH could obey, for example exponential law or gamma law, etc. More complicated, the composite model is the result of a combination of simple distributions in which exponential distribution plays an important role to describe long TH tendency. A good TH model facilitates traffic safety analysis, traffic capacity calculation and microscopic simulation.

Motorways connect principal zones of a country. The network of this type of road gathers therefore a large amount of traffic, especially at the entrances of big cities. This is also the case of the A6 Motorway in this work which is located in the south of Paris. Previous reports on the A6 data were focused only on the global statistical properties such as travel time, mean speed, speed-density cartography, and level of service. However detailed microscopic studies which are closely linked to driver behavior remain an open issue that needs to be investigated further.

In previous works, TH distributions have been examined to compare the different probabilistic models and the different estimation methods (cf. Luttinen (1996), Hoogendoorn (1996, 1998) and Zhang *et al.* (2007)). The studies have focused on the goodness-of-fit of the models to the real data and the performance of the estimation techniques. In these researches, data have been generally collected according to the traffic flow. Besides, the main criterions used were based on the statistic test values.

In this paper, in addition to the comparison of TH distribution following the hourly traffic volume, the aim of the work is also at comparing the TH arrival patterns of vehicles in different traffic contexts by using the data collected on the urban branch A6a of the A6 Motorway. These comparisons are realized in the similar hourly traffic volume level in order to separate the factor of traffic demand. The contexts considered are period of the day, effect of reduction in cross-profile and type of lane. The latter factor has been rarely studied in the literature. There are three lanes on the A6 Motorway: the slow lane (SL), the middle lane (ML) and the fast lane (FL). The best TH simple model - the Log-normal Model (LNM) and one of the two best composite TH models - the Generalized Queuing Model (GQM) are calibrated by lane. In addition to the probabilistic modeling, statistical analysis of a series of events furnishes supplementary information in order to compare TH between contexts. Finally, instantaneous speed distribution contributes also additional points for the comparisons.

DATA COLLECTION

The data are collected on the A6a motorway in the south of Paris in France. The experimental sections are straight and far from interchanges. The average traffic volume is about 70 000 vehicles per day where 5 % represents heavy trucks. The legal maximal speed in the section studied is 110 km/h. In addition, cross-profile is reduced on a 2 kilometer-long section by marking lines on the motorway surface without any sign for drivers in advance. Normal lane width of 3.25 m was narrowed by 13 % on all lanes.

The periods measured were in the morning, in the afternoon and in the evening on week days in 1996. The stations of the SIREDO standard system recorded individual data in the experiment. Three key parameters are recorded for each vehicle: arrival instant (including time in hours, in minutes and in hundredths of second), instantaneous speed (in km/h) and vehicular length (in meters). Furthermore, the collected data were qualified by using fixed cameras observing the real events. This camera system allows traffic operators to refine data in the cases where the inductive loop system of the SIREDO station could lead to confusion. Moreover, data were separately recorded from the three different lanes of the A6 Motorway. In total, 25 hours data for each lane of the roadway were collected in different periods of 10 days in April and in May, from 8h to 15h, and also some from 18h to 21h.

Initially, the data in question were used to describe how the lateral position of vehicles is distributed, to observe the possible utilization of space between lanes by motorcyclists, to analyze the effect of lane reduction upon the capacity of the total roadway and on speed according to the period of the day. The time headway study was restrained in calculation of short TH in global circumstances (cf. Costilles). Such detailed individual data allows us to investigate in a different way the effects of different contexts on TH distribution. The two best TH distributions are calibrated and supplementary calculations of spectral analysis and of speed distribution were also carried out.

METHOD

Time Headway models

TH distribution is formulated by probability density function (PDF) $f(h)$ and cumulative density function (CDF) $F(h)$. The relation between the two functions is $dF(h) = f(h)dh$. Therefore, knowledge of function f is essential to understand the arrival pattern of vehicles in different contexts. Time headway h is a positive variable measured generally in seconds.

Previous work (Ha *et al.*) showed that the best goodness of fit obtained for a single model is the LNM while the GQM and the semi - Poisson (SPM) are the two best composite ones. Besides, calculation with the GQM is straightforward.

The LNM has been proposed for TH by Greenberg (cf. Greenberg) in 1966 with the following PDF:

$$f(h) = \frac{1}{(h-\tau)\sqrt{2\pi}\sigma} e^{-\frac{[\ln(h-\tau)-\mu]^2}{2\sigma^2}} \times 1_{\{h \geq \tau\}}(h) \qquad (1)$$

The LNM assumes that the logarithm of TH obeys the normal law of parameters μ and σ. Parameter τ is a positive translation of the whole distribution. The first and principal assumption of the four assumptions proposed by Greenberg in order to derive the log - normality of the TH pattern is:

$$H(t+dt) - H(t) = k(t) \times H(t) \qquad (2)$$

where $H(t)$ is the TH variable at instant t.

It means that the change of TH from present instant t is proportional to the actual TH value at t. The assumptions of Greenberg create also connections between the LNM and some models of car - following.

In this paper, the estimation of parameters of the LNM is implemented in parallel by two methods proposed by Cohen and Whitten (cf. Cohen *et al.*, Luttinen). They are the Local Maximum Likelihood Estimator (LMLE) and the Modified Maximum Likelihood Estimator (MMLE). Both methods are based on the solution of a system of three equations resulting from partially differentiating the Log-likelihood function of LNM with respect to the corresponding parameters of the LNM. The MMLM replaces the equation respecting τ in the LMLE by a simpler one: $F^{-1}\left(\frac{1}{n+1}\right) = h_{(1)}$ where n is the number of sampled TH and $h_{(1)}$ is the minimum TH value in the sample. In both methods, an equation of τ is then deduced and solved before being used to find the two remaining parameters μ and σ. The best result in terms of statistical tests such as chi-squared and Kolmogorov - Smirnov (K - S) is chosen between the LMLE and the MMLE.

The composite model GQM is a combination of an exponential tendency and another model. The aim is to calibrate the unimodal property of TH distribution and simultaneously to fit the exponential tail for long TH. To this end, global TH is broken down into two parts: the following TH (also called tracking TH or constrained TH) and the non-following TH (or free TH). The supposed relationship between these two TH categories will decide the final formulation of a composite model.

The GQM proposed by Branston and Cowan (cf. Branston, Cowan) supposes that free TH contains itself constrained TH and an exponential part. As a result, the PDF of GQM is:

$$f(h) = \theta g(h) + (1-\theta)\lambda e^{-\lambda h} \int_0^h g(u) e^{\lambda u} du \qquad (3)$$

where g indicates the PDF of the following TH. The gamma-distribution is the best choice for function g because of its unimodal form and also of its facilities in calculation (cf. Hoogendoorn *et al.*, Ha *et al.*). The parameter λ represents the intensity of the exponential tendency while the parameter θ is considered as the participating part of tracking vehicles. Gamma - function g is written with parameter α and β as below:

$$g(h) = \frac{\alpha^{\beta} (h-\tau)^{\beta-1}}{\Gamma(\beta)} e^{-\alpha(h-\tau)} \times 1_{\{h \geq \tau\}}(h) \qquad (4)$$

The previous work showed that estimation with $\tau = 0$ provided much better results particularly while using the Maximum Likelihood Estimator (MLE) (cf. Hoogendoorn *et al.*, Ha *et al.*). In this paper, this method is carried out by MATLAB 7.0 using the Toolbox optimization.

Analysis of serial events

The probabilistic method provides distributions or frequencies of all TH values. Nonetheless, this approach does not allow one to investigate the properties of independence or correlation between successive TH. The latter facts are however important to understand interactions between driver decisions. The calculations in this section are supposed to measure the correlation between vehicles and the distance between real characteristics of the specific arrival patterns compared with the renewal property. Notice that the last notion involves a series of random variables which are independent and identically distributed (i.i.d).

Let a series of n consecutive TH $h_1, h_2, \ldots, h_n$ be sampled in one hour. This arrival process contains n interval times h_i between arrival instants of a time series $X_0, X_1, \ldots, X_n$. So $h_i = X_i - X_{i-1}$. The correlations between the TH H_i and another TH H_{i+k} positioning at k interval distances far from the H_i form the *correlogram*. The value of k is integer ($k = 0,1,\ldots n-1$). Taking into account the standard deviation of TH, the correlation coefficient in k order is calculated in order to measure the level of correlation:

$$\begin{aligned} \rho_k &= \frac{\operatorname{cov}(H_i, H_{i+k})}{Var\left[H\right]} \\ &= Corr(H_i, H_{i+k}) \end{aligned} \qquad (5)$$

These calculated values are also called auto - correlation coefficients. The ρ_1 is used as well as the simplest method to test a renewal process because it reflects the interaction between two consecutive TH.

Another approach utilized is the spectral analysis when the series of ρ_k is known. This method is applied on the basis of the two following spectral functions, the power function $P(\omega)$ and the spectral density $f(\omega)$:

$$P(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \mathrm{cov}(H_i, H_{i+k}) \cos(k\omega) \qquad \text{With } -\pi \le \omega \le \pi \qquad (6)$$

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} Corr(H_i, H_{i+k}) \cos(k\omega) \qquad \text{With } -\pi \le \omega \le \pi \qquad (7)$$

This paper is focused on the estimations $I_n(\omega_p)$ of the power function $P(\omega)$ to establish the *periodogram*:

$$\begin{aligned} I_n(\omega_p) &= \frac{1}{2\pi} \sum_{k=-n+1}^{n-1} \left(1 - \frac{|i|}{n}\right) \tilde{C}_i \cos(i\omega_p) \\ &= \frac{1}{\pi} \left[\frac{C_0}{2} + \sum_{i=1}^{n-1} \left(1 - \frac{|i|}{n}\right) \tilde{C}_i \cos(i\omega_p) \right] \end{aligned} \qquad (8)$$

with $\omega_p = 2\pi p / n$ where $p = (1, 2, \ldots n/2]$ and $\tilde{C}_i$ is the estimated value of $\mathrm{cov}(H_i, H_{i+k})$.

Information represented in *correlogram* and in *periodogram* can be used to compare different arrival TH processes. The coefficient ρ_1 and the values of $I_n(\omega_p)$ could serve also to test the renewal hypothesis of samples examined. Notice that $2I_n / s^2$ is a constant of the value $1/\pi$ in any renewal process, where s is the estimated value of standard variation of the TH. Greater variations observed in both the *correlogram* and *periodogram* globally indicate higher correlations between vehicles. It means also that the arrival patterns are further from a renewal process. The number n is the number of events recorded in each sample. The estimated values of TH means and TH variances are recalculated in each computation of correlation coefficients in the order k, despite the fact that this operation is slightly affected when k is near to n (cf. Figure 1.).

In order to measure the homogeneity of variations of the $2I_n / s^2$, the simple method chosen is the study on means and standard deviations of $2I_n / s^2$. Notice that this statistic depends on the number n. Therefore, the comparison of two highly different traffic volumes samples using this statistic is not reliable.

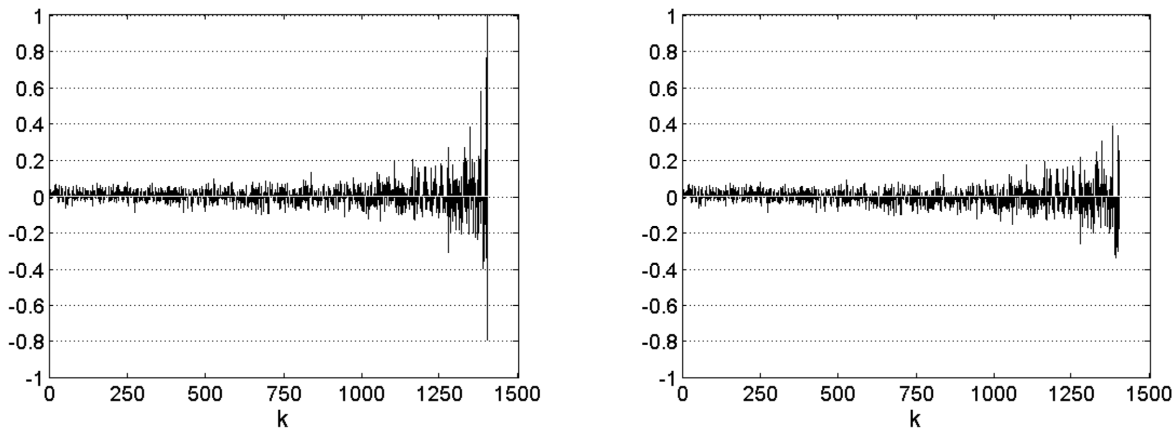


Figure 1. Estimation of ρ_k with estimated means and variances (Left) and with constant means and variances (Right)

RESULTS AND DISCUSSION

Traffic in different periods of the day

Traffic in the period (8h - 9h) was significantly heavier than traffic in periods from 10h to 15h. In the morning rush hour, drivers tended to utilize the FL the most, then the ML and the SL. Traffic on the SL was always the least in all cases of any hour examined. Nevertheless, more drivers have chosen the ML than the FL when traffic was intermediate (neither dense nor fluid). The utilization of the ML in such conditions seems to give drivers more flexibility either for speeding up or for slowing down only by changing the driving lane. In the above time periods, the slower the lane is, the more stable the traffic is maintained. Traffic of the FL was very scattered in the noon period that one could suppose that this lane was mainly used for transient purposes such as overtaking.

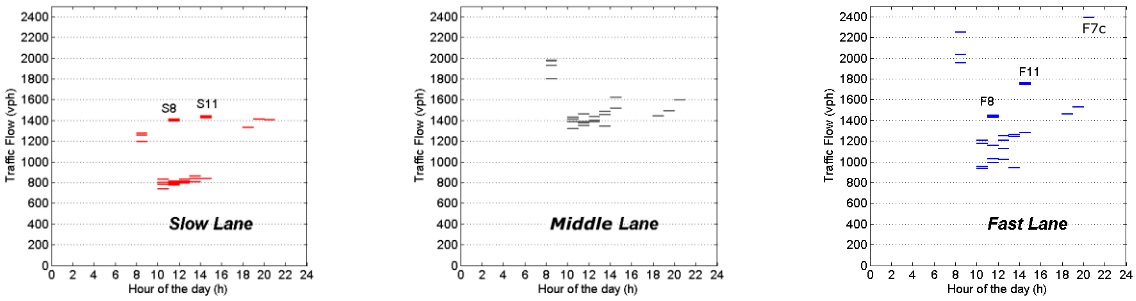


Figure 2. Traffic Volume per Lane of The Three Lanes

The traffic in the period (18h - 21h) stabilized at about the hourly traffic flow per lane of 1500 vph. The samples here were only collected in reduced cross-profile section corresponding to the periods where *traffic jams* were observed. Furthermore, traffic volumes in this condition were similar among the three lanes.

Five particular cases given in Figure 2 corresponding to the periods (11h - 12h) and (14h - 15h) on the SL and on the FL were S8, F8, S11, F11 and F7*c* (the first letter is the lane category, the succeeding number indicates the sample number, the letter *c* signifies the third hour of the sample N°7). All these cases coincide with the traffic jam phenomena observed both in normal and in reduced cross - profile sections.

The traffic flows studied were counted per lane by the number of vehicles passing the measure station regardless of vehicle type. In fact, the proportion of trucks whose length is longer than 7 m was particularly high on SL after the period (8h - 9h). It could reach 30 % at noon. On the contrary, a very small percentage of trucks were recorded on the ML and FL throughout the time of the experiment.

Statistical properties

Mean and standard deviation of motorway TH have close relationships with traffic flow. Figure 3 shows that the standard deviations decreased regularly following traffic flow on all lanes. The coefficients of variation (CV) were stable and less than 1 on the SL and on the ML. More variable on the FL, this coefficient was greater than 1 for intermediate traffic volume and a bit scattered with values less than 1 at high level of traffic. The CV greater than 1 indicates that TH is more grouped on the FL around the traffic flow of 1000 vph.

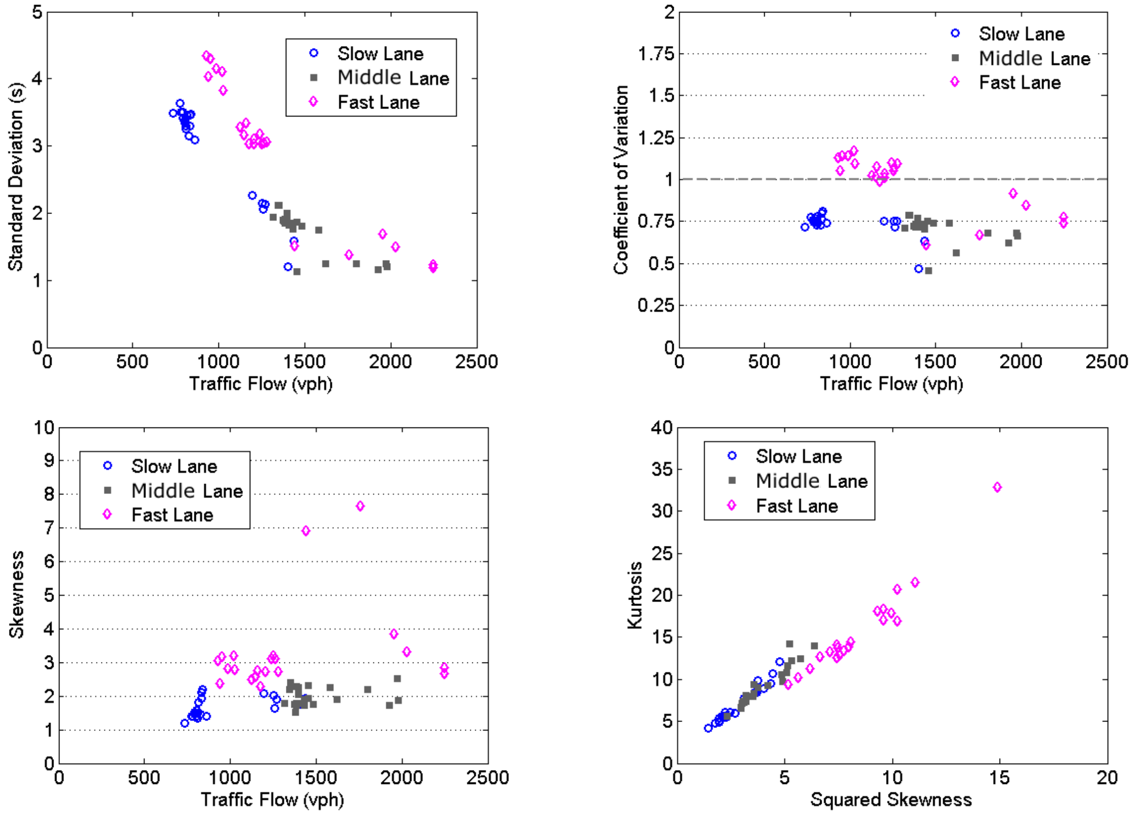


Figure 3. Statistical Properties of motorway Time Headway

The skewness values were at the same level for the data from the SL and the ML. They were higher with two extreme cases possible for the data from the FL. However, the relationship between squared skewness and kurtosis is always linear. In this case, statistical properties on SL and on ML are continuously close together and slightly different from those on the FL.

Generally, statistical characteristics of traffic in the same period of the day (from one hour to the next) were analogous namely on the SL and ML. Other properties are recapitulated in Table 1, where ρ_1 is the TH auto - correlation coefficient between one vehicle and the preceding one; proportion p is the percentage of long vehicles. Interaction between vehicles represented by ρ_1 was more significant in the rush hour period. Possible negative values could occur in the SL. This phenomenon on the SL corresponds to a high rate of heavy vehicles during the off-peak periods.

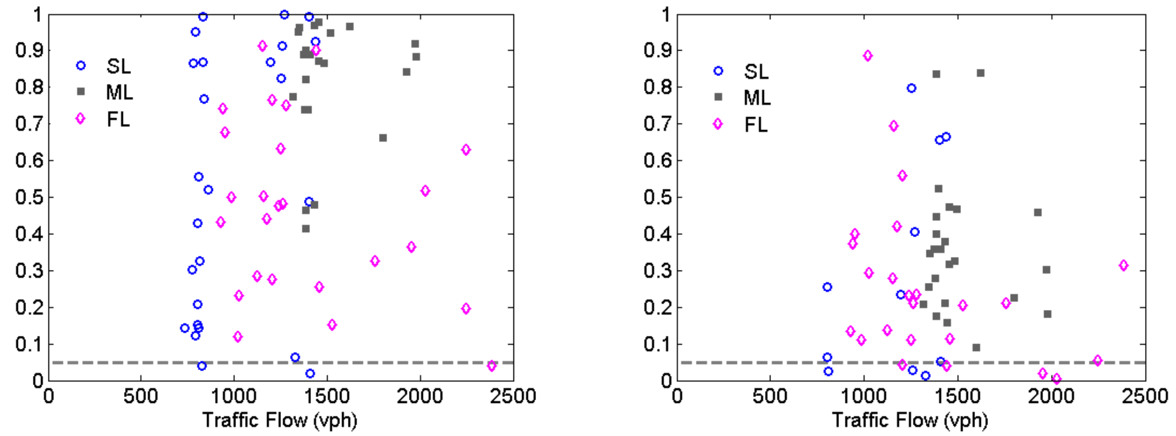


Figure 4. P - value of the GQM (Left) and the LNM (Right)

TH distribution and the traffic flow

The samples processed were modeled following the hourly traffic flow with the LNM, the GQM, the SPM and others. In the results, the GQM and the SPM provide substantially good fits for most cases with very high levels of P - values in terms of the both chi - squared test and K - S test. These two models have in fact equivalent levels of goodness of fit. Furthermore, variations and tendencies of the four parameters α , β , θ and λ were extremely similar between the GQM and the SPM. Consequently, the GQM is focused on in this section. Figure 4 illustrates also the performance of the LNM in many cases especially for samples of intermediate traffic volume from the ML and the FL.

Flow	Period	Mean	Std	Mode	CV	Skewness	Kurtosis	ρ_1	p %
					Slow Lane				
1255	8h - 9h	2.868	2.151	1.15	0.75	2.007	8.850	0.0838	9.88
1272	8h - 9h	2.827	2.134	1.15	0.75	1.902	8.387	0.1029	10.38
814	11h - 12h	4.419	3.452	1.05	0.781	1.821	7.931	-0.0002	22.60
810	12h - 13h	4.443	3.292	1.75	0.741	1.340	4.815	-0.0491	25.06
862	13h - 14h	4.176	3.083	1.45	0.738	1.398	5.313	-0.0662	23.55
					Middle Lane				
1974	8h - 9h	1.821	1.240	1.05	0.681	2.528	14.019	0.0824	1.52
1926	8h - 9h	1.869	1.160	1.05	0.621	1.733	7.125	0.0568	1.25
1372	11h - 12h	2.625	1.897	1.15	0.723	1.749	7.433	0.0268	1.24
1433	12h - 13h	2.511	1.767	1.05	0.704	1.894	8.368	0.0207	1.61
1485	13h - 14h	2.423	1.800	1.15	0.743	1.757	7.262	0.0611	0.81
					Fast Lane				
2247	8h - 9h	1.599	1.237	0.85	0.774	2.843	14.502	0.0923	0
2251	8h - 9h	1.598	1.185	0.75	0.742	2.664	13.235	0.1052	0
1158	11h - 12h	3.109	3.346	0.85	1.076	2.755	12.928	0.0877	0
1253	12h - 13h	2.875	3.027	0.75	1.053	3.203	20.720	0.0963	0
1243	13h - 14h	2.893	3.183	0.95	1.100	3.102	17.017	0.0295	0

Table1. Statistical properties of time headway in normal lane width section

We can see from the estimated values of the parameters of the GQM that (Figure 5):

1. Corresponding to the traffic volume of about 800 vph on the SL, the parameters α , β and θ vary widely (α and β from 2 to 6, θ from 0.05 to 0.25). In the higher traffic levels, the corresponding estimated parameters are higher but scattered on the SL and the ML. However, slight tendencies of the three parameters in question are observed for the FL data.

2. The parameter λ is sensitive to the variation of fluid traffic flow corresponding to off - peak hour periods. Indeed, parameter λ of data on the three lanes varies linearly following the traffic volume in the off-peak periods despite the fact that the traffic flow changes were small. However λ disperses more greatly on the SL and on the ML when traffic becomes denser. For the FL samples, parameter λ possibly maintains a linear trend.

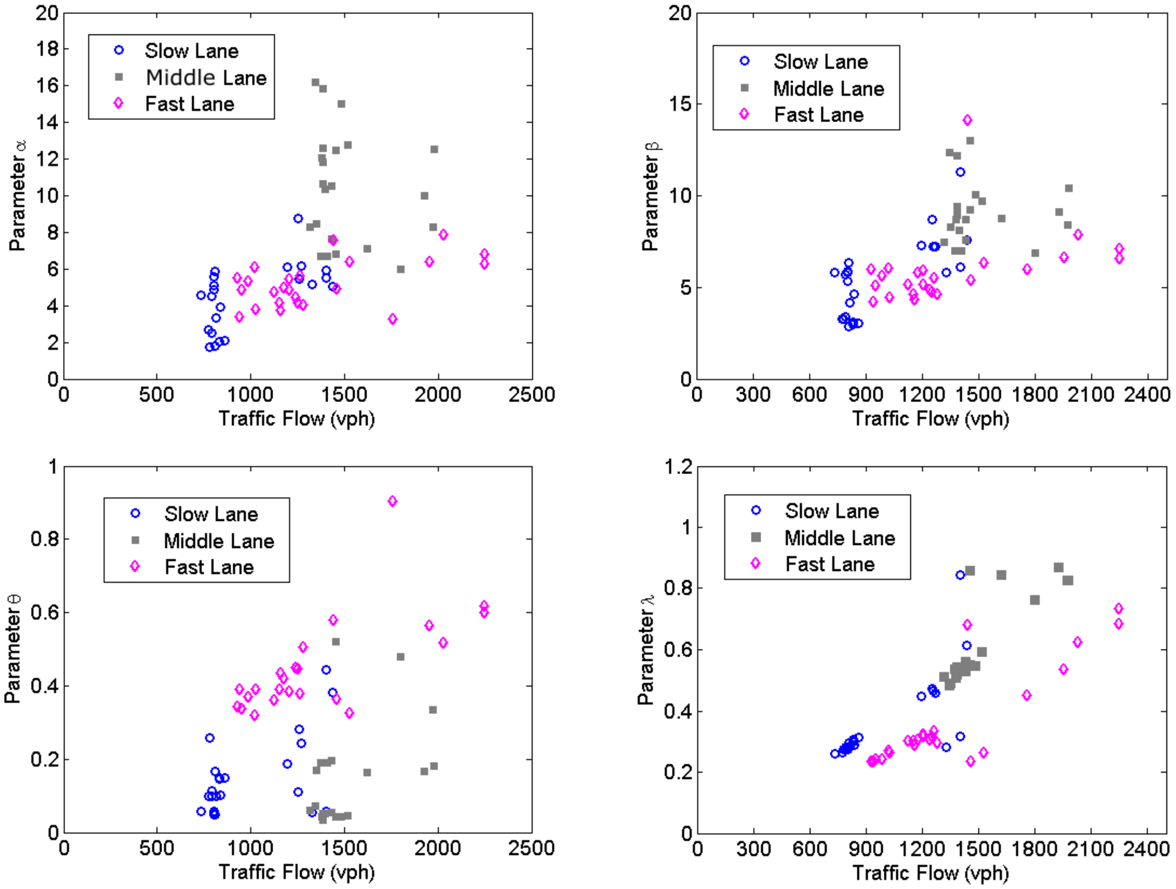


Figure 5. Variation of estimated values of four parameters of the GQM

3. A particular statistical property of the GQM for TH is that the relationship between the parameters α and β is linear (cf. Figure 6). This correlation was predictable by observing that the clustering of estimated values of these two parameters in Figure 5 has closely similar forms.

Q	Period	α	β	θ	λ	P_{χ^2}	$P_{K\text{-}S}$	A^2
814	11h - 12h	3.3249	4.1521	0.1001	0.2839	0.1531	0. 3253	0.7252
810	12h - 13h	1.7808	2.8602	0.1653	0.2943	0.2760	0.5575	0.5159

Table 2. GQM parameters of two samples (S3b and S3c) in the same period of the day

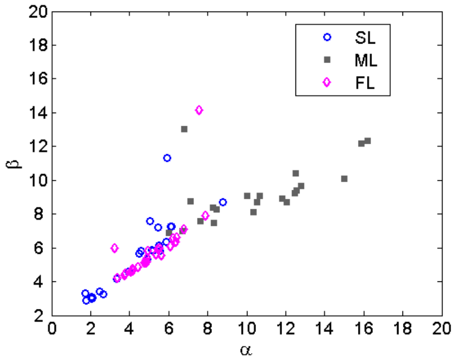


Figure 6. Relationship between the parameters α and β of the GQM

Figure 5 presents results in all lane width cases. Table 2 tabulates two periods sampled in similar traffic flow on the SL of the same section. In Figure 8, we notice that:

1. The GQM parameters are sensitive to the distribution of short TH less than 5 *s* (see the TH distribution in *histogram* of Figure 8). The difference is specifically reflected by the parameters α and β. The parameters θ and λ are more stable in a similar range of traffic volumes. By the way, the GQM furnishes always a high goodness of fit which is statistically demonstrated by high values of P - value.

2. In the *correlogram* of Figure 8, correlations between vehicles are small possibly because of low traffic flows. The estimated values in *periodogram* of the 814 vph vary a little more regularly around the renewal value than the 810 vph sample does. Indeed, the standard deviation of estimated values $2I_n / s^2$ in the first case (SL3b) is less than in the second case (SL3c) (cf. Table 8).

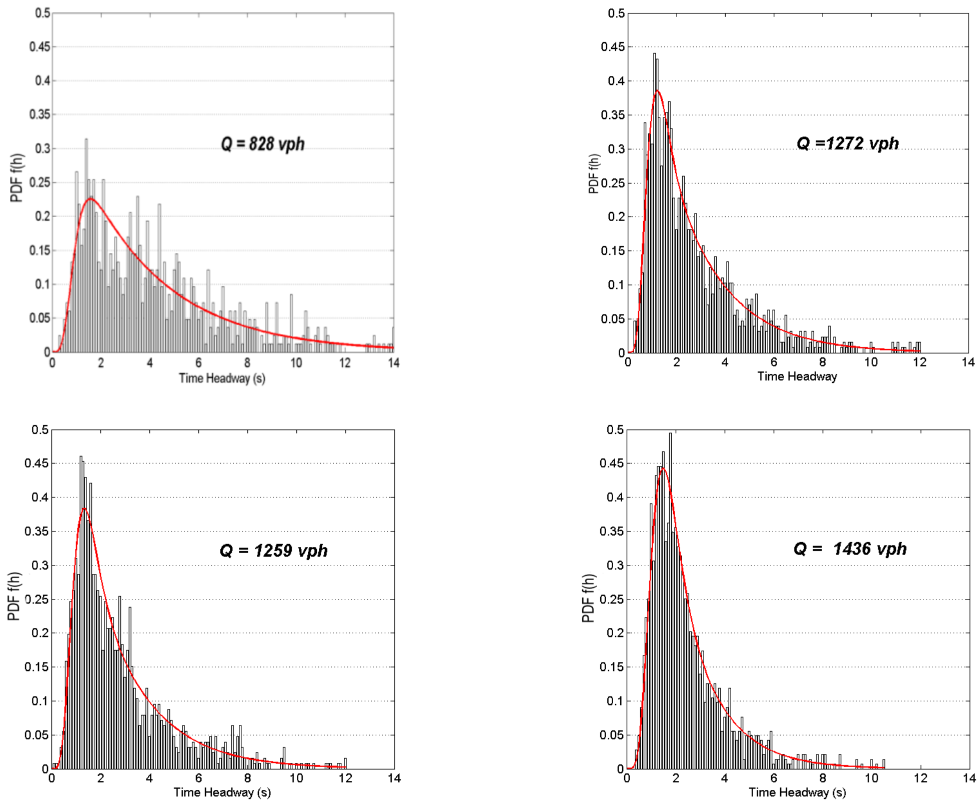


Figure 7. Samples with equivalent estimated values of parameters α and β

Q	Period	α	β	θ	λ	P^2_χ	$P_{K\text{-}S}$	A^2
828	14h - 15h	6.1477	6.9136	0.0601	0.2937	0.0352	0.0387	1.2829
1272	11h - 12h	6.1610	7.2461	0.2437	0.4580	0.7223	0.9984	0.1642
1259	10h - 11h	5.4728	7.2098	0.2825	0.4659	0.3506	0.9117	0.4753
1436	14h - 15h	5.0604	7.6027	0.3832	0.6153	0.8875	0.9238	0.1338

Table 3. GQM parameters of two couples of samples ((S6c,S2) and (S4, S11)) with similar values of parameters α and β

The second remark is enhanced when regarding the speed distributions. The average speeds are equivalent except that the larger diversion in the 810 vph case caused by some outlier values is observed. Between these two samples of similar traffic flow on the same lane, estimated values in the *correlogram* and in the *periodogram* are quite close together. Also, despite more stable estimated values of GQM parameters in the remaining SL samples in the noon period, the existence of such large variation of parameters α and β shows that the utilization of the notion of hourly traffic flow could possibly contain more than one different TH pattern corresponding to the same level of traffic.

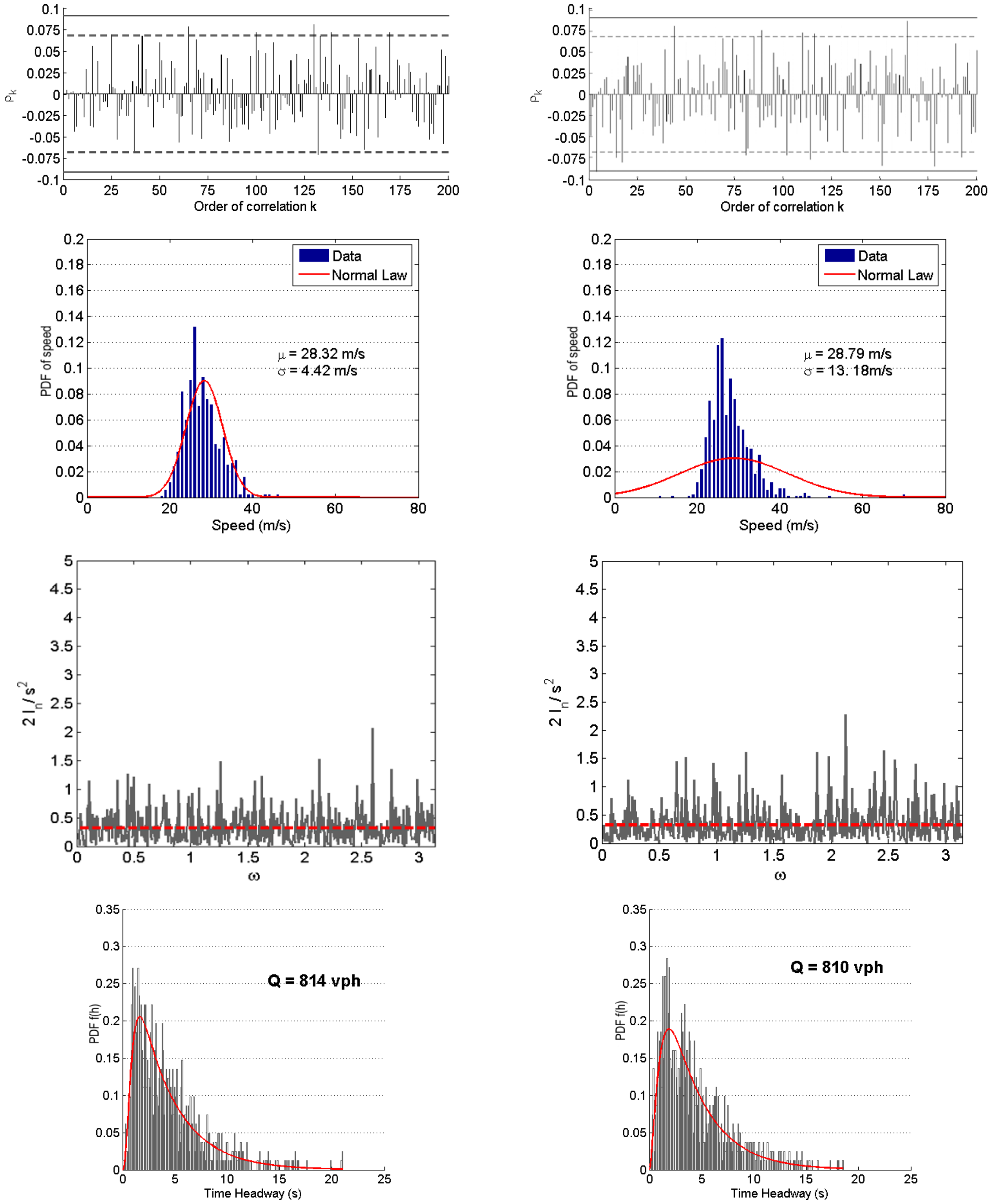


Figure 8. Two samples (S3b and S3c) with similar traffic flow

Table 3 lists the two couples of samples where parameters α and β are equivalent but where traffic volumes are completely different. The first couple with very near values α and β have two totally different distributions (cf. the upper row of Figure 7). This results from dissimilar values in parameters λ and θ. In general, in the vicinity of the Mode, both smaller and greater TH concentrate more around it in the higher volume case. Furthermore, the slope of decrease of the PDF curve after the Mode is much steeper in the heavier traffic case because of higher value of λ. The Mode in the heavier traffic flow is smaller and the corresponding mode frequency is higher.

To summarize, the GQM provides better fits for the heavier traffic with very high values of P - value in both squared - chi tests and K - S test. Smaller values of A^2 of the Anderson - Darling test are also obtained. In addition, the coincidences of parameters of the GQM could be investigated following some possibilities:

1 Hourly traffic flow is not a unique factor influencing TH distribution. This notion based on a long measure period perhaps contains more than one variation of arrival patterns. These variations therefore lead to the different distributions of TH.

2 The optimization procedure simultaneously for four parameters could be derived to numerical coincidences. This fact is maybe caused by correlations between parameters such as the relationship between α and β.

In all cases, the GQM presents a high flexibility and robustness that changing some parameters and maintaining others could also model well for two very different traffic contexts.

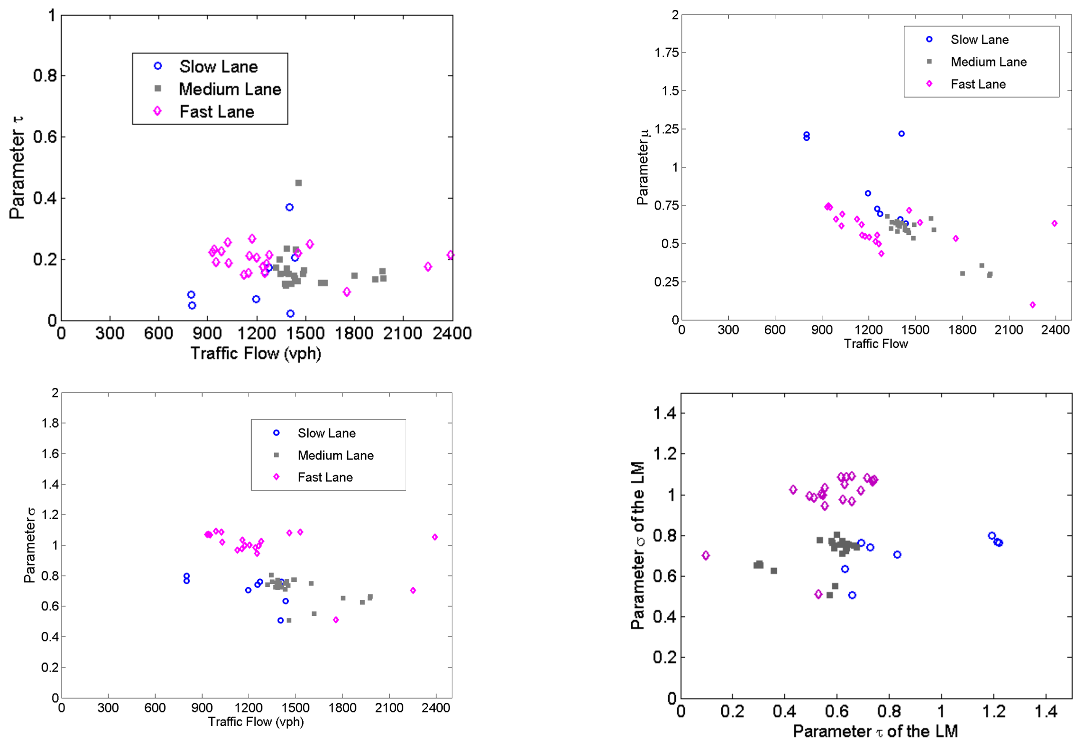


Figure 9. Variation of the parameters of the LNM

Following the traffic flow, the LNM fits most samples collected on the ML and on the FL. However, only some samples on the SL are successfully calibrated by the LNM.

Represented in Figure 9, the parameter of location τ is quite stable on both the ML and the FL. The estimated values of the parameter μ decrease following traffic flow. Linear tendency is obtained the most visually on the FL for traffic less than 1200 vph. The estimated values of σ are higher on the FL than other lanes. The LNM is less robust than the GQM and does not provide high goodness of fits for a wide sample range having different statistical characteristics which vary from fluid to high traffic demand. The relationship between τ and σ visualizes three separated areas of estimated points corresponding to the three lane types.

Influence of the reduction on lane width

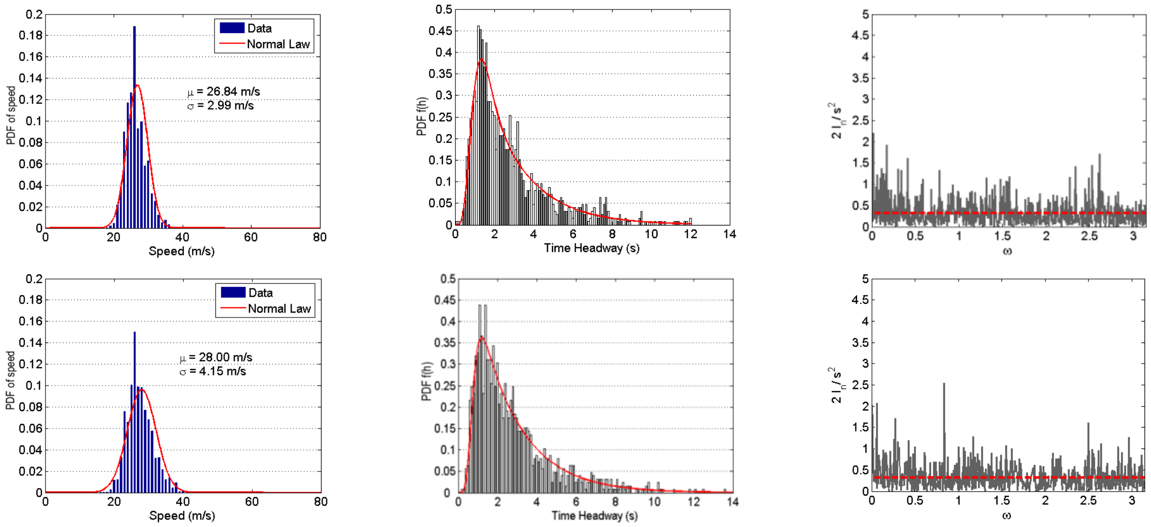
Table 4 recapitulates 4 couples of samples of similar traffic flow corresponding to the two cross-profile sections.

Lane	Q	Period	α	β	θ	λ	P_{χ^2}	$P_{\text{K-S}}$	A^2
Reduced Cross - Profile									
SL	1259	8h - 9h	5.4728	7.2098	0.2825	0.4659	0.3506	0.9117	0.4753
SL	828	12h - 13h	6.1477	6.9136	0.0601	0.2937	0.0352	0.0387	1.2829
ML	1384	10h - 11h	15.8654	12.1688	0.0332	0.5273	0.9978	0.9025	0.3361
FL	1204	12h - 13h	6.1206	6.0662	0.3196	0.2693	0.8451	0.1183	2.2148
Normal Cross - Profile									
SL	1255	8h - 9h	8.7727	8.6785	0.1099	0.4738	0.5208	0.8248	0.2189
SL	832	10h - 11h	2.0322	2.9788	0.1467	0.2982	0.6847	0.9932	0.2201
ML	1384	11h - 12h	10.6731	9.0676	0.0507	0.5421	0.9897	0.8224	0.2325
FL	1204	12h - 13h	4.8933	5.1749	0.3851	0.3182	0.3255	0.2765	0.7339

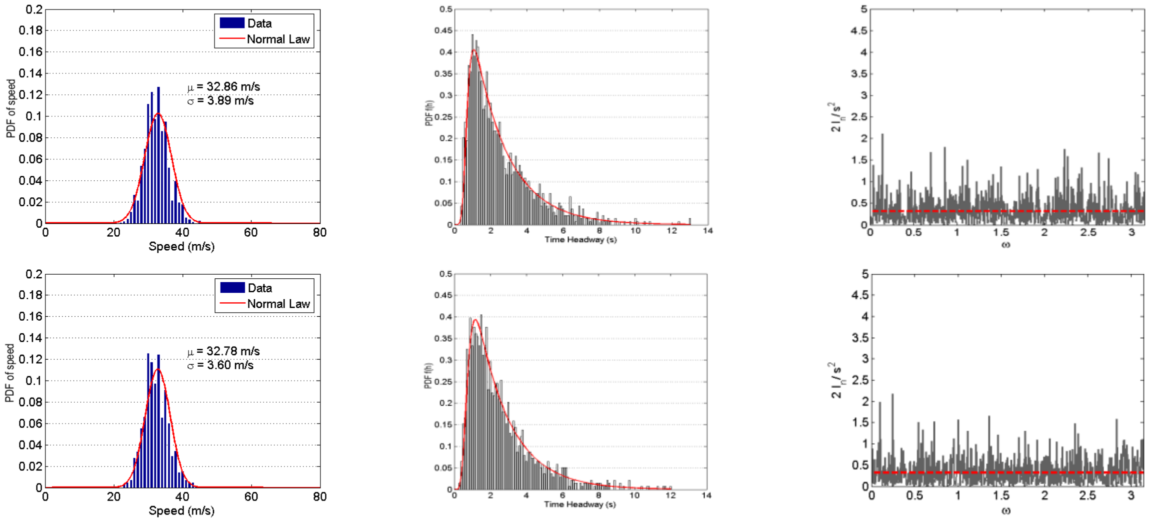
Table 4. GQM parameters of four couple samples of equivalent traffic flow in 2 sections

It is difficult to detect an influence caused by the reduction width through the *correlogram* and *periodogram* on the three lanes in peak hours. The standard deviation of estimated values of $2I_n / s^2$ changes more in off-peak hour periods both in the SL and on the FL (cf. Table 7), despite the fact that the directions of change are inverse when the reduction lane width occurred. The variation of standard deviations of $2I_n / s^2$ is minor on the ML in off-peak hour. One could investigate that heavy trucks make drivers in this lane pay more attention to the lane width because the heavier inertia of trucks must be taken into account. Also, drivers in the FL change possibly their overtaking behavior when lane width is modified. This effect seems only to be produced during off-peak hours when reduced lane marking is more visible. Comparing GQM parameters of each couple, the parameter λ does not change greatly. The faster the lane is, the greater the parameter λ changes. On all lanes in off-peak hours, the parameters α, β increase while θ decreases when lane width is reduced. In this case, parameter θ changes the most on the SL. On the SL in the noon period, to compensate for the almost stable values of λ, the three remaining parameters change greatly. Nevertheless, the tendencies of change in the parameters α, β and θ are inverse to those on the SL in peak - hour despite the fact that the distribution forms between two contexts are close together (cf. Figure 10 (S4 and S1)).

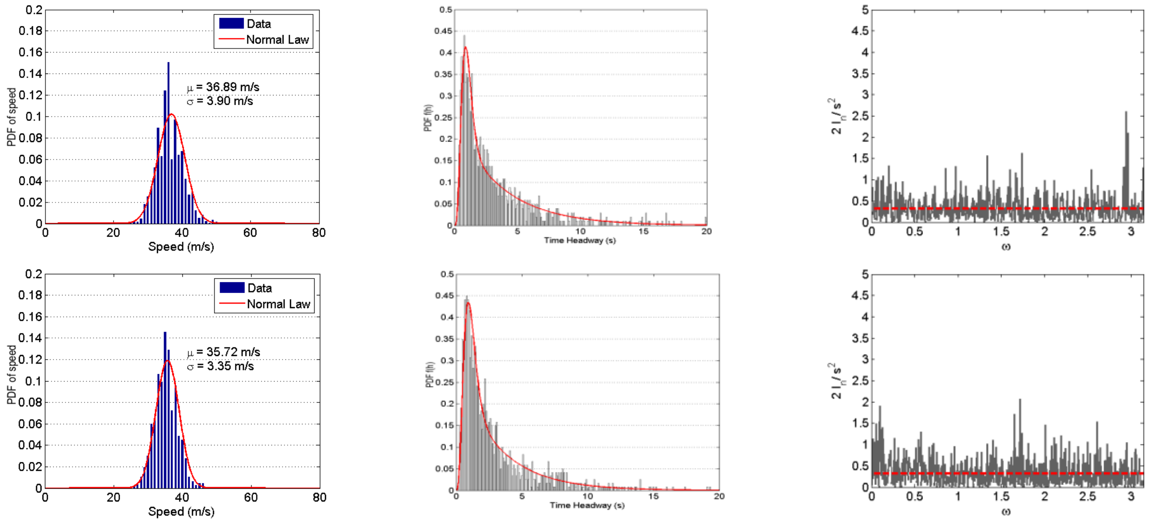
On the SL in the peak hour, the change in standard deviation of $2I_n / s^2$ is minor in modified cross-profile section (cf. Table 7). However, by comparing the speed distributions, we obtained the effect upon data on the SL where average speed and standard deviation are both greatly reduced. From this point of view, the ML is the least disturbed. These facts show that complementary utilization speed distribution is necessary (cf. Figure 10).



Two samples S4 and S1



Two samples M6a and M9b



Two samples F6c and F9c

Figure 10. Effect of reduction of cross - profile width

Comparison between lanes

In the samples considered, traffic volumes are always higher than 1200 vhp on the ML. Furthermore, traffic flow on this lane is steadiest with globally lower values of TH standard deviation and more constant values of TH skewness. The traffic flow on the FL especially in the noon period is very variable. In this latter case, higher values of TH standard deviation and CV greater than 1 are recorded. In general, the CV and skewness values of TH are similar on the ML and on the SL. These statistics are smaller than those of the same traffic volumes on the FL.

Lane	Q	Period	α	β	θ	λ	P_{χ}^2	$P_{\text{K-S}}$	A^2
SL	1272	8h - 9h	6.1610	7.2461	0.2437	0.4580	0.7223	0.9984	0.1642
ML	1926	8h - 9h	10.00	9.1031	0.1659	0.8697	0.4370	0.8422	0.1748
FL	2251	8h - 9h	6.2671	6.5915	0.5998	0.7331	0.4788	0.1953	0.6534
SL	831	10h - 11h	2.0322	2.9788	0.1467	0.2982	0.6847	0.9932	0.2201
ML	1411	10h - 11h	6.7236	6.9742	0.1901	0.5363	0.0736	0.8891	0.3587
FL	1205	10h - 11h	5.4496	5.9030	0.3839	0.3235	0.8805	0.7668	0.3557

Table 5. Samples of the three lanes in the same period

We notice that negative values of the first seven orders of ρ_k appeared mainly on the SL. The majority of samples on the FL have positive values of the first three correlation coefficients ρ_1, ρ_2 and ρ_3.

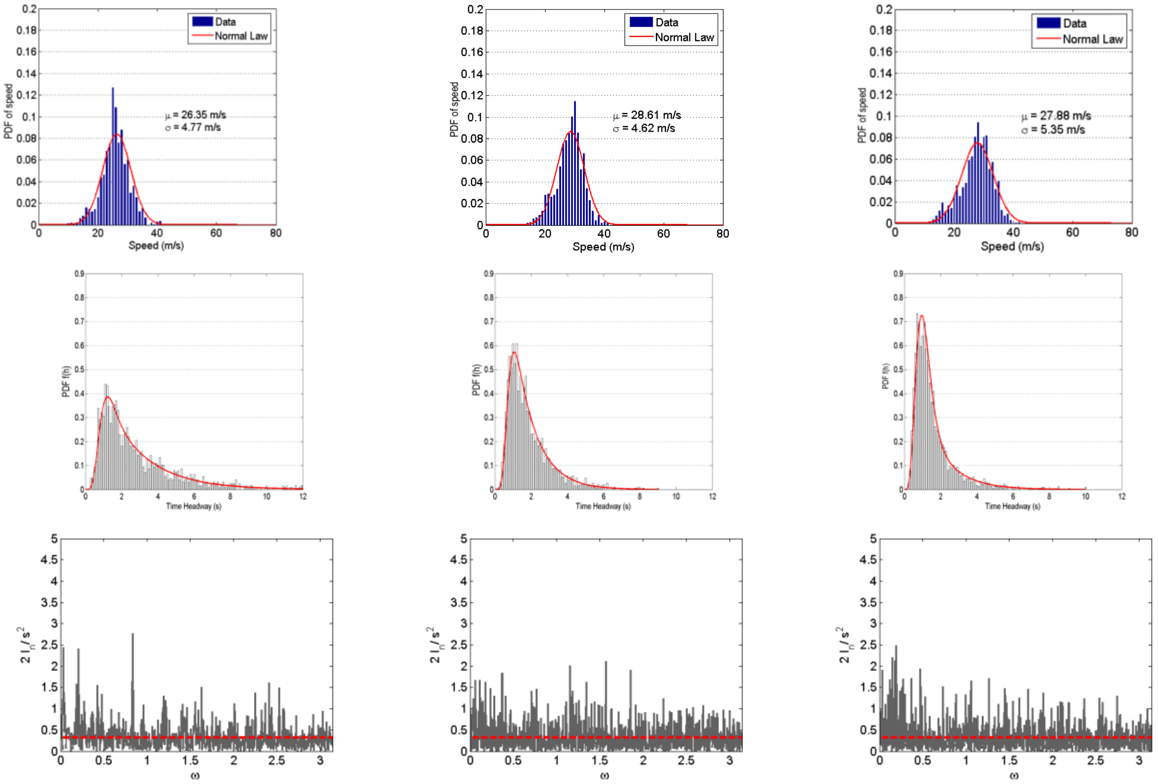
In the same period studied, comparative research of the first ten coefficients ρ_k presents a complicated relationship between lanes, and no remarkable result is observed. The *periodograms* in Figure 11 and Table 8 do not show clearly which lane is the most different. Notice that comparing the standard deviations of $2I_n / s^2$ between the samples of different event numbers n is not reliable.

Lane	Q	Period	α	β	θ	λ	P_{χ}^2	$P_{\text{K-S}}$	A^2
SL	1255	8h - 9h	8.7727	8.6785	0.1099	0.4738	0.5208	0.8248	0.2189
FL	1253	12h - 13h	4.1709	4.7568	0.4480	0.3183	0.0634	0.6321	0.5997

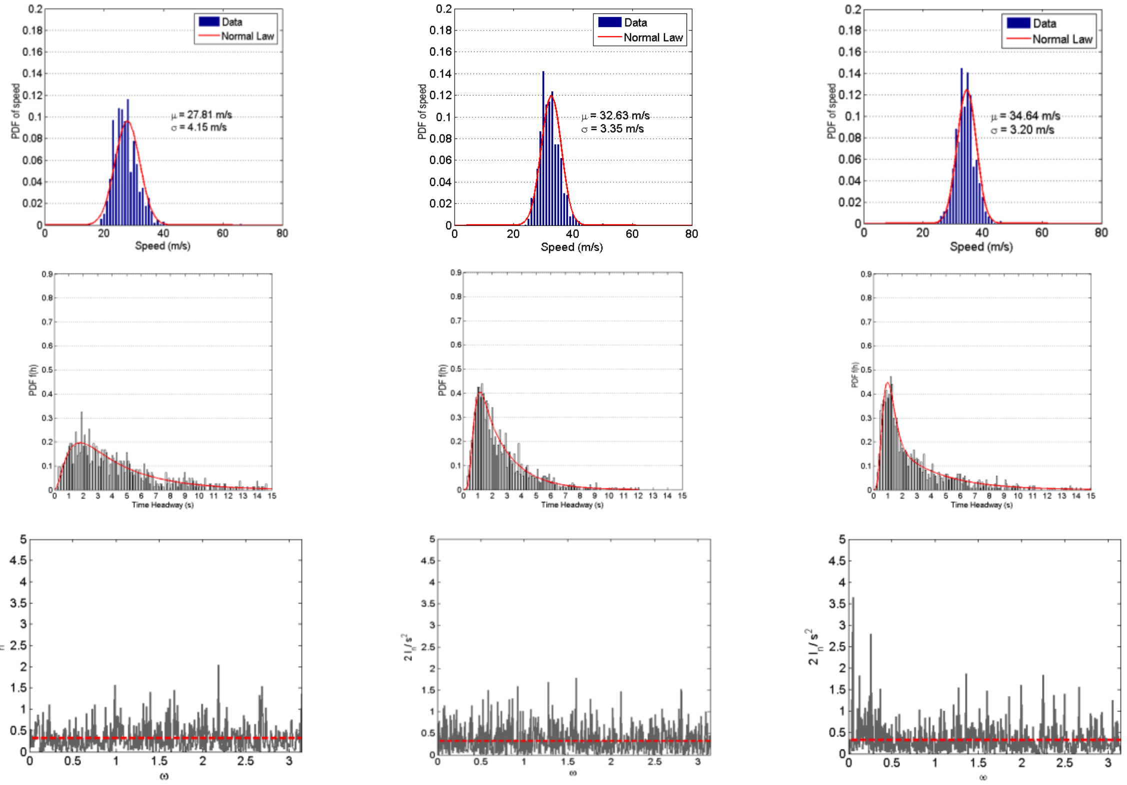
Table 6. Two samples (S1 and F3c) of similar flow rates in different lanes

The difference between TH distribution among three lanes is typical in peak hour in which α and β are high (normally greater than 5). Parameter λ is the smallest on the SL and the greatest on the ML. In the peak hour period on the FL, P - value of test K - S is the smallest, possibly because in the middle TH range, shorter values concentrate near the Mode than those on the SL and the ML. As a result, the CDF curve of the FL samples fits with more difficulty the data in the critical area of the K-S test that is principally located in the middle TH values. The Modes decrease respectively according to the SL, the ML and the FL. The Mode frequencies increase in other ways going from the SL to the FL.

In the same peak hour periods, the highest average speed and the lowest speed variation are obtained on the ML (cf. Figure 11). This phenomenon is handed over to the FL in off-peak hours. The lowest average speed is found on the SL. The fit of the speed distribution to a normal law depends on the time period examined, traffic flow and operating lane.



Three samples S2, M2 and F2



Three samples S3a, M3a and F3a

Figure 11. Samples on three lanes in the same periods

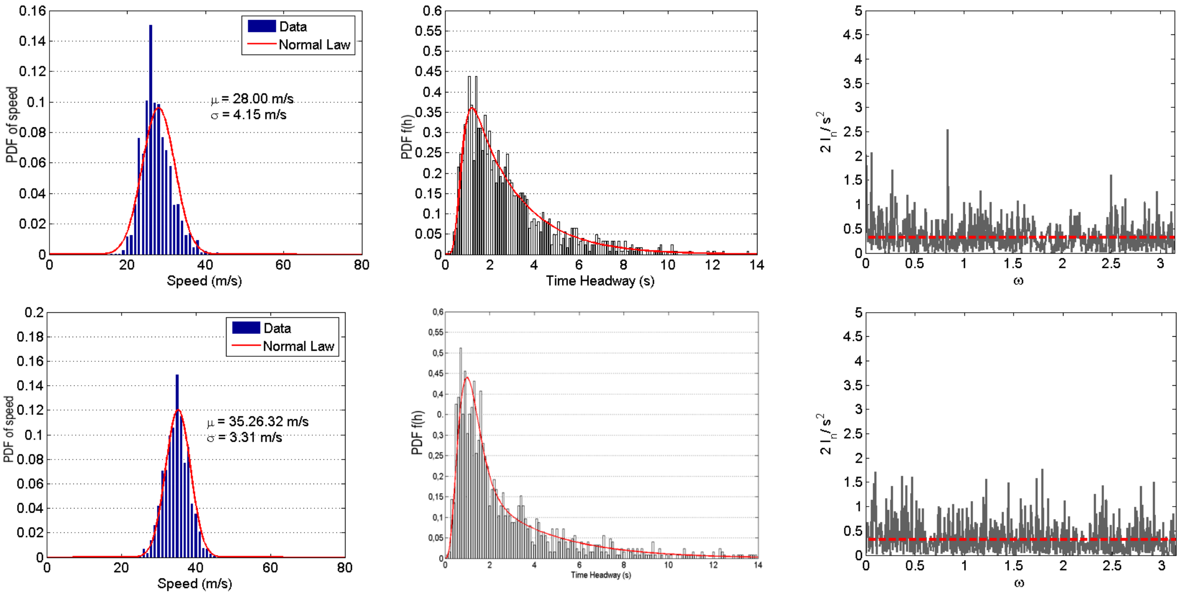


Figure 12. Two similar traffic demand samples in different lanes

Two samples in different lanes in Figure 12 are at the same level of traffic flow. In Table 6, a slight change in value of traffic volume makes a great difference in parameter λ. The high value of parameter θ on the FL makes the distribution form change more roughly from Mode to exponential decrease phase. The standard deviation of $2I_n / s^2$ is smaller in the SL case (SL1) than in the FL case (FL3c) despite the fact that the period of the SL sample is in peak hour.

Standard Deviation of $2\,I_n/s^2$				
	Peak hour	Off-peak hour		
Samples	SL4 & SL1	SL6c & SL3a	ML6a & ML3a	FL6c & FL9c
Narrowed	0.3078	0.3414	0.3113	0.3043
Normal	0.3056	0.3102	0.3146	0.3163

Table 7. Standard deviation of $2I_n / s^2$ in normal and narrowed lane width

CONCLUSIONS

Traffic flow depends first on the period of the day. Following the lane type in the same period, hourly traffic flow is either stable or scattered from one hour to the next. These traffic flow variations are smaller on the slower lane or in traffic jams. This fact is confirmed by the statistics of standard deviation, CV and skewness values on the three lanes. Furthermore, the statistical properties of the SL and the ML are close together and slightly different from those on the FL.

With four parameters, the GQM is flexible and capable of adapting to many traffic contexts. Each parameter of the GQM is sensitive to different traffic characteristics. More visible tendencies are observed for all the parameters on the FL according to the traffic flow. We suggest that such trends possibly could be obtained on other lanes by using traffic flow measured in periods shorter than one hour which are more homogeneous. On the other hand, the relationship shown by the graph between the parameters α and β suggests reducing the GQM parameters for TH modeling.

With three parameters of which one is the parameter of location, the LNM is less flexible than the GQM and seems to be simple to investigate precisely traffic phenomena. Indeed, the parameters of LNM are mainly linked to mean and standard deviation of logarithms of TH. However, research on this model with shorter measuring traffic flow is also necessary. The LNM model provides good fits particularly for samples on the ML and the FL.

The lane reduction effect could be detected by *periodogram* and by speed distribution. Variation on parameters of the GQM responds complicatedly to this lane narrowing. Changing lane width by modifying marking lines leads to a more significant response of the drivers on the SL. However, no obvious change is noticed on the ML in off-peak hour. Comparing lanes in the same period examined, TH distribution is typically different between lanes in the peak hour.

The hourly traffic flow is a vague macroscopic variable which could not characterize traffic properties well, although the literature has considered this flow as the unique principal factor influencing TH distribution. We noticed that the same hourly traffic volumes had different arrival patterns because of dissimilar driver behavior. One hour seems to be a long measure period for sampling data.

Studying the *correlogram* of ρ_k is not practical in order to compare possible vehicular correlations in different situations. Calculating with estimated values of *periodogram* should be tested in other traffic contexts and locations. Because this analysis depends on the number of events, the homogeneity of standard deviations of $2I_n / s^2$ has drawbacks in comparing two samples of different traffic flows. However, the analysis of instantaneous speed properties is vital for distinguishing separated contexts. Consequently, studying the association between the TH and the instantaneous speed presents an open issue for research in the near future. The variable for example h / v or the relationship between two distributions could lead to more interesting results. Moreover, the new variable h / v concerns the inverse of acceleration of vehicles which plays an important role in traffic safety research.

REFERENCES

Costilles, and Rouille. (1996). Une expérimentation de réduction du profil en travers sur voie rapide urbaine. Technical.report, RGRA/INRETS.

Ha, D. H., M. Aron, and S. Cohen. (2009). Modélisation du temps intervéhiculaire sur l'autoroute de France. Actes N°124, JDD SPI-STIC.

Greenberg, I. (1966). The log-normal distribution of headways. Australian Road Research, Vol.2, pp. 14-20.

Cohen, A., and B. Whitten. (1980). Estimation in the three – parameter lognormal distribution. Journal of the American Statistical Association, Vol.75, No.310, pp. 399-404.

Luttinen, R. (1996). Statistical Analysis of Vehicular Time Headways. PhD thesis, University of Technology Lahti Center (Neopoli, Lahti, Finland).

Branston, D. (1976). Models of single lane time headway distributions. Transportation Science, Vol. 10, No.12, pp. 125-148.

Cowan, R. (1975). Useful headway models. Transportation Research, Vol.9, pp. 371-375.

Hoogendoorn, S., and H. Botma. (1996). The estimation of parameters in headways distributions. Technical report, TRAIL Research School Delft.

Zhang, G., Wang, Y., Wei, H. and Chen, Y. (2007) Examining headway distribution models with urban freeway loop event data. Transportation Research Record, Vol. 1999, pp. 141–149.

Hoogendoorn, S., and Bovy. (1998). New estimation technique for vehicle – type – specific headway distributions. Transportation Research Record, Vol. 1646, pp. 18–28.

Statistic $2\,I_n/s^2$						
	Mean			Standard Deviation		
Samples	SL	ML	FL	SL	ML	FL
1	0.3189	0.3183	0.3184	0.3056	0.3178	0.3328
2	0.3176	0.3184	0.3170	0.3249	0.3182	0.3253
3a	0.3167	0.3179	0.3182	0.3102	0.3021	0.3554
3b	0.3178	0.3179	0.3164	0.2912	0.3270	0.3493
3c	0.3185	0.3179	0.3183	0.3287	0.3185	0.3265
3d	0.3167	0.3180	0.3181	0.3090	0.3078	0.3289
3e	0.3182	0.3179	0.3184	0.3116	0.3119	0.3439
4	0.3181	0.3181	0.3184	0.3078	0.3159	0.3125
5	0.3181	0.3180	0.3183	0.3246	0.3007	0.3328
6a	0.3186	0.3174	0.3187	0.3226	0.3113	0.3199
6b	0.3184	0.3183	0.3183	0.3284	0.3111	0.3318
6c	0.3175	0.3180	0.3155	0.3414	0.2901	0.3043
6d	0.3169	0.3175	0.3149	0.2875	0.2925	0.3255
7a	0.3185	0.3173	0.3182	0.3149	0.3269	0.3197
7b	0.3180	0.3181	0.3168	0.3276	0.3017	0.3102
7c	0.3174	0.3178	0.3164	0.3019	0.3092	0.3483
8	0.3170	0.3175	0.3176	0.3224	0.3078	0.3038
9a	0.3171	0.3181	0.3181	0.3256	0.3199	0.3257
9b	0.3167	0.3174	0.3175	0.3189	0.3146	0.3098
9c	0.3166	0.3177	0.3182	0.3087	0.3281	0.3163
9d	0.3184	0.3165	0.3174	0.2928	0.3165	0.3100
11	0.3181	0.3166	0.3185	0.3272	0.3455	0.3273
12a	0.3179	0.3173	0.3179	0.3180	0.2911	0.3181
12b	0.3155	0.3182	0.3171	0.2869	0.3188	0.3119
12c	0.3167	0.3182	0.3177	0.3122	0.2992	0.3364

Table 8. Mean and Standard Deviation of $2I_n / s^2$