



## DISTRIBUTION OF TRAFFIC SPEED IN DIFFERENT TRAFFIC CONDITIONS: AN EMPIRICAL STUDY IN BUDAPEST

Mohammad MAGHROUR ZEFREH,  T\*#

*Dept of Transport Technology and Economics, Budapest University of Technology and Economics,  
Budapest, Hungary*

Received 1 March 2018; revised 26 March 2019; accepted 4 July 2019

**Abstract.** Fundamental diagram, a graphical representation of the relationship among traffic flow, speed, and density, has been the foundation of traffic flow theory and transportation engineering for many years. Underlying a fundamental diagram is the relation between traffic speed and density, which serves as the basis to understand system dynamics. Empirical observations of the traffic speed versus traffic density show a wide-scattering of traffic speeds over a certain level of density, which would form a speed distribution over a certain level of density. The main aim of the current research is to study on the distribution of traffic speed in different traffic conditions in the urban roads since the distribution of traffic speed is necessary for many traffic engineering applications including generating traffic in micro-simulation systems. To do so, the traffic stream is videotaped at various locations in the city of Budapest (Hungary). The recorded videos were analysed by traffic engineering experts and different traffic conditions were extracted from these recorded videos based on the predefined scenarios. Then their relevant speeds in that time interval were estimated with the so-called “g-estimator method” using the outputs of the available loop detectors among the videotaped locations. Then different parametric candidate distributions have been fitted to the speeds by Maximum Likelihood Estimation (MLE) method. Having fitted different parametric distributions to speed data, they were compared by three goodness-of-fit tests along with two penalized criteria (Akaike Information Criterion – AIC and Bayesian Information Criterion – BIC) in order to overcome the over-fitting problems. The results showed that the speed of traffic flow follows exponential, normal, lognormal, gamma, beta and chi-square distribution in the condition that traffic flow followed over-saturated congestion, under saturated flow, free flow, congestion, accelerated flow and decelerated flow respectively.

**Keywords:** speed distribution, traffic condition, urban road traffic, traffic flow dynamics, speed–density relationship, interrupted traffic flow.

### Introduction

Fundamental diagram, a graphical representation of the relation among traffic flow, speed, and density, has been the foundation of traffic flow theory and transportation engineering for many years. For example, the analysis of traffic dynamics relies on input from this fundamental diagram to find when and where congestion builds up and how it dissipates; traffic engineers use a fundamental diagram to determine how well a road facility serves its users and how to plan for new facilities in case of capacity expansion. Underlying a fundamental diagram is the relation between traffic speed and density, which roughly corresponds to drivers’ speed choices under varying car-following distances. In fact, this relation serves as the basis to understand traffic system dynamics in various scientific areas including traffic flow (Wang *et al.* 2009). Empirical

observations show a wide-scattering of traffic speeds over a certain level of density. This scattering effect is due to the randomness of drivers’ speed choices. In fact, due to the stochastic nature of traffic flow, the observed speed may vary over a certain range, forming a distribution. These distributions in highways, or in general un-interrupted traffic flows, would follow the normal distribution over a certain level of density as shown in Figure 1 (Wang *et al.* 2009). Jun (2010) mentioned that the change or variability of speed distributions on a specified roadway during a certain period of time may explain the trends or patterns of how the characteristics of traffic on the roadway vary. The conditions that turn to a different speed distribution (rather than Normal distribution) are quite often realized in non-highway or urban roads, where, in general, the traffic

\*Corresponding author. E-mail: [torok.adam@mail.bme.hu](mailto:torok.adam@mail.bme.hu)

#Editor of the TRANSPORT – the manuscript was handled by one of the Associate Editors, who made all decisions related to the manuscript (including the choice of referees and the ultimate decision on the revision and publishing).

stream is much more complicated (Maghrour Zefreh *et al.* 2016). This assumption is supported by IMAGINE (2006), where different speed distributions are related to different traffic conditions. The development of mathematical tools focused on the modelling of the speed distribution in a traffic flow is widely reported in the scientific literature (Castro *et al.* 2008; Dey *et al.* 2006; Fitzpatrick *et al.* 2000; Trozzi *et al.* 1996). In general, speed distribution is necessary for many traffic engineering applications. For instance, an appropriate speed distribution model is the fundamental input for generating vehicles in traffic micro-simulation systems (Park, Schneeberger 2003; Llorca *et al.* 2015) and other applications such as activity-travel scheduling simulation (Liao *et al.* 2013; Liao 2016). Moreover, speed distribution can be utilized in theoretical analyses of traffic flow characteristics and to plan the appropriate traffic operational measures (Yu, Abdel-Aty 2014).

To sum up, in literature it is reported that, there is a normal distribution of traffic speed associated with each level of density in the so-called un-interrupted traffic flows (e.g. traffic flow in the highways) (Wang *et al.* 2013). Furthermore, different studies in the literature generally demonstrated the fact that distribution of traffic speed might deviate from normal distribution in interrupted traffic flows, e.g. urban road traffic flows (Maurya *et al.* 2015), however, to the best of the authors’ knowledge, there is no empirical study available in the literature that assigns the best fitted parametric distributions to the interrupted traffic speeds according to different traffic conditions. Therefore, the current research work attempts to fill this gap in the literature by studying on the distribution of the traffic speed in the so-called interrupted traffic flows where traffic is more complicated than highways due to the presence of traffic lights, intersections etc.

The remainder of this paper is organized as follows. The applied methodology in details including the extraction of desired traffic conditions from the videotaped traffic flow based on the predefined scenarios, speed estimation, distribution fitting and goodness-of-fit testing are

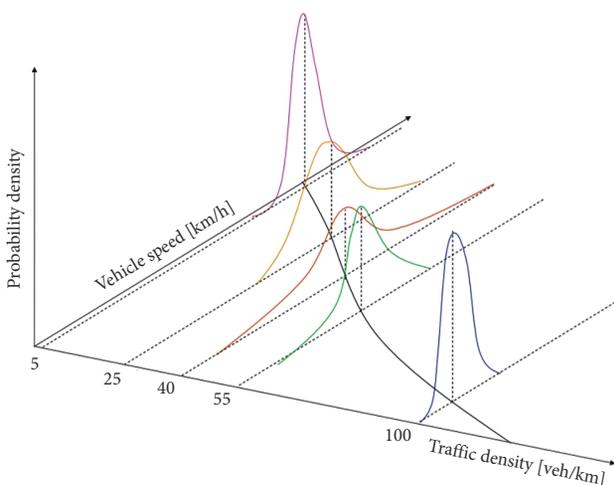


Figure 1. Three-dimensional representation of the speed–density relationship (Wang *et al.* 2009)

described in Section 1. A case study, at first, is considered in Section 2 where traffic flow followed acceleration process (for the illustration purpose) and then the obtained results for the other traffic conditions are presented and discussed respectively. Finally, the findings of the current research are concluded in the last section.

## 1. Methodology

In order to study on the distribution of the speed in the so-called interrupted traffic flow, a widespread traffic video tapping has been done in various locations in the city of Budapest (Hungary). Table 1 shows the investigated sites. Furthermore, Figure 2 shows the speed–density relationship of the traffic flow in the investigated area.

Table 1. Investigated sites for traffic flow analysis

Location	Video recording	Detector data	Number of lanes (in each direction)
Villanyi – Karolina Road (N)	√	√	1
Villanyi – Karolina Road (S)	√	√	1
Villanyi – Alsohegy Road (N)	√	√	1
Villanyi Road (710-4)	√	√	1
Villanyi Road (710-5)	√	√	1
Villanyi Road (710-6)	√	√	2
Villanyi Road (710-7)	√	√	2
Alsohegy – Villanyi Road (S)	√	√	1
Hamzsabegi Road	√	–	1
Szent Gellert Road	√	–	2
Budafoki Road	√	–	1
Moricz Zsigmond Road	√	–	1

### 1.1. Extracting different traffic conditions from the traffic flow

Having recorded the traffic flow, the recorded videos were analysed by traffic engineering experts and different traffic conditions (under saturated flow, free flow, congestion, over-saturated congestion, accelerated flow, decelerated flow) were extracted from these recorded videos based on the predefined scenarios.

#### 1.1.1. Defining different scenarios

The defined scenarios are theoretically explained here using a theoretical representation of a characteristic diagram for traffic moving at a traffic light when it turns to green shown in Figure 3.

The condition in which the traffic congestion would not be disappeared in the cycles demand (exceeds capacity for significant period) is considered as *over-saturated congestion*. This is the region  $x_j \leq x < x_L$  in Figure 3 where

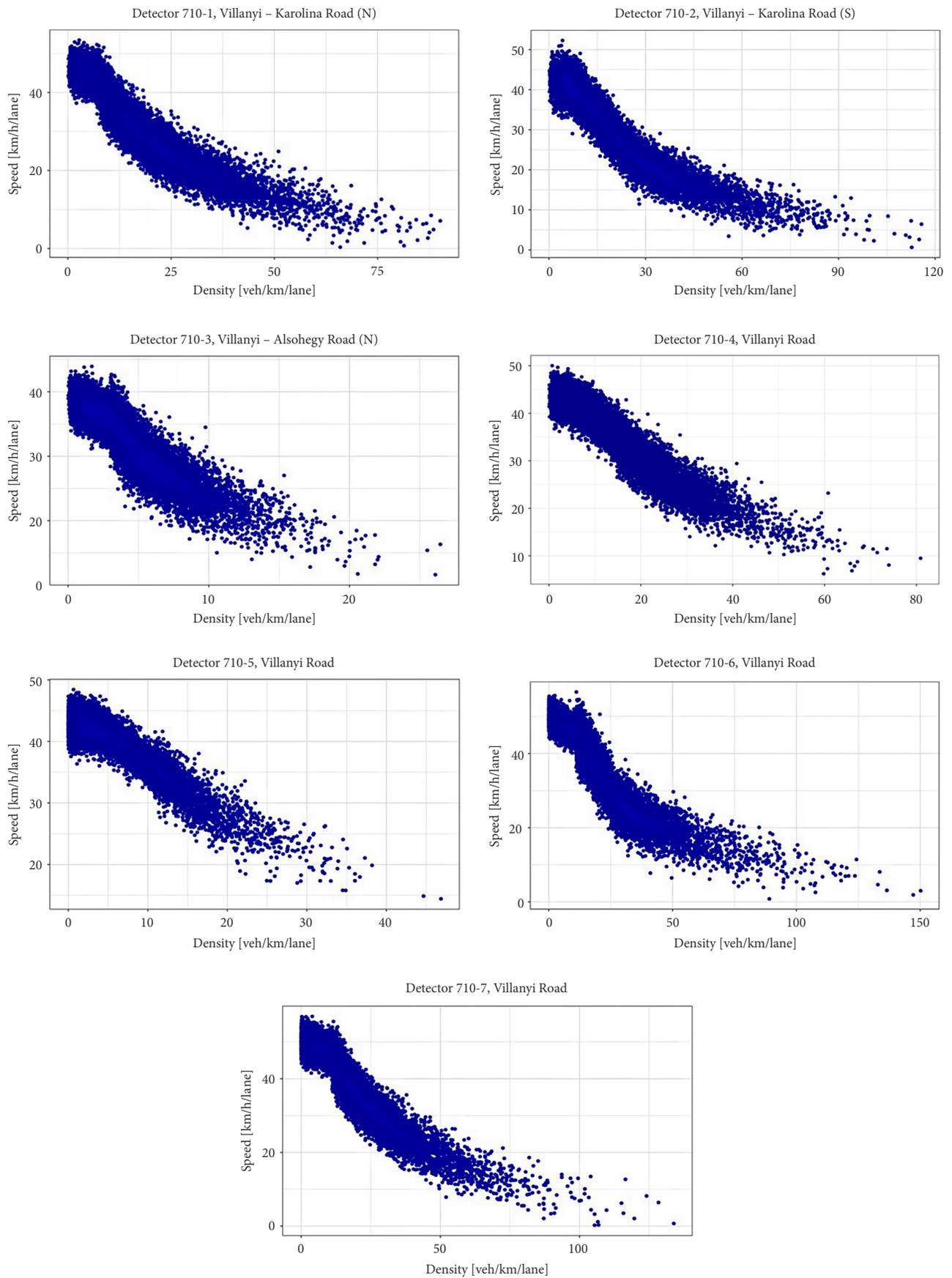


Figure 2. Speed–density relationship of the traffic flow in the study area (source: compiled by the authors)

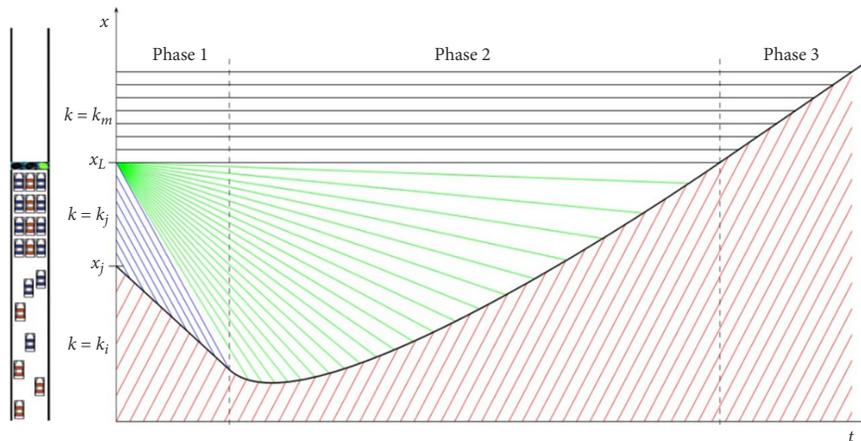


Figure 3. Characteristic diagram for traffic moving at a traffic light (Lustri 2010)

density is close to jam density  $k_j$ . The *under saturated* flow is considered when the traffic is close to the capacity (and will be discharged in a cycle) in the region  $x < x_j$ , where traffic is flowing with the density lower than optimal density  $k_i < k_m$ , as the flow is not unimpeded. The *free flow* traffic is considered when there are a few numbers of vehicles in the street (much lower than the capacity) in the region  $x < x_j$ , where traffic is flowing with the density lower than optimal density  $k_i < k_m$ , as the flow is not unimpeded. The *deceleration process* is considered in the condition where traffic flow is getting closer to the  $x_j$ , where the shockwave, black curve, travels backward through the traffic in Phase 1 shown in the Figure 3. When the traffic light turns green, vehicles are able to leave the light entirely unimpeded, so the density would be equal to optimal density  $k = k_m$  and obviously flow would be in its maximum. This condition is considered as the *acceleration process* in which the shock wave slows down and starts to move back towards the traffic light (Phase 2). The condition in which traffic light turns to green but the intersection is not completely empty (individual cycle failures) yet is considered as *congestion*.

### 1.1.2. Extracting different traffic conditions from the recorded videos

In the current research, traffic flow is videotaped at various intersections in the city of Budapest during the whole day. The video tapped traffic flow is further analysed by traffic engineering experts in order to extract different traffic conditions from these videos based on the already defined scenarios in Section 1.1.1. For instance, Figure 4a shows a series of the frames extracted from the recorded video when traffic flow followed the deceleration process in one of the videotaped locations (Hamzsabegi Road) during the day and Figure 4b shows a series of the frames extracted from the recorded video when traffic flow followed acceleration process in one of the videotaped locations (Szent Gellert Road) during the night.

It should be highlighted that literature has scientifically proved that night driving behaviour would influence on car-following conditions resulting in generating instabil-

ity in traffic flow (Jiang, Wu 2007; Bella *et al.* 2014; Bella, Calvi 2013). This instability in traffic flow is the basis of generating different traffic conditions that would be leaded to different speed profiles and distributions. Therefore taking traffic conditions as the basis of speed distribution analysis might seem to be a logical assumption in both day times and night times.

### 1.2. Speed estimation by g-estimator method from loop detector outputs

Data from loop detectors have been primary sources for traffic information, and single loops are the predominant loop detector type in many places. Unfortunately, the most common form of traffic detector, the single loop detector, is incapable of providing speed measurements. Therefore, traffic speed should be calculated based on the detector output, that is, traffic volume and occupancy time. Since the loop detector outputs may contain some incorrect/missed values due to equipment malfunctions and communication faults, before doing the speed estimation, the outputs of detectors have been validated and the incorrect/missed values were imputed based on the algorithms proposed by Maghrour Zefreh and Török (2018a). Having done the validation/imputation process, traffic speed is estimated by the so-called “g-estimator method” using the loop detector outputs as shown in Equation (1):

$$\bar{S}(i) = \frac{N(i)}{T \cdot O(i) \cdot g}, \quad (1)$$

where:  $i$  is the time interval index;  $\bar{S}$  is the speed for each time interval;  $N$  is the number of vehicles per interval (volume);  $O$  is the percentage of time in which loop is occupied per interval;  $T$  represents the hours per interval;  $g$  is the constant based on mean vehicle length and detector size.

It should be noted that, in this research, the parameter  $g$  in Equation (1) is calibrated based on Maghrour Zefreh *et al.* (2017). Having estimated the speeds, the desired time intervals (based on the time intervals of extracted traffic conditions from the recorded videos) were extracted for the further investigation (finding their distributions).

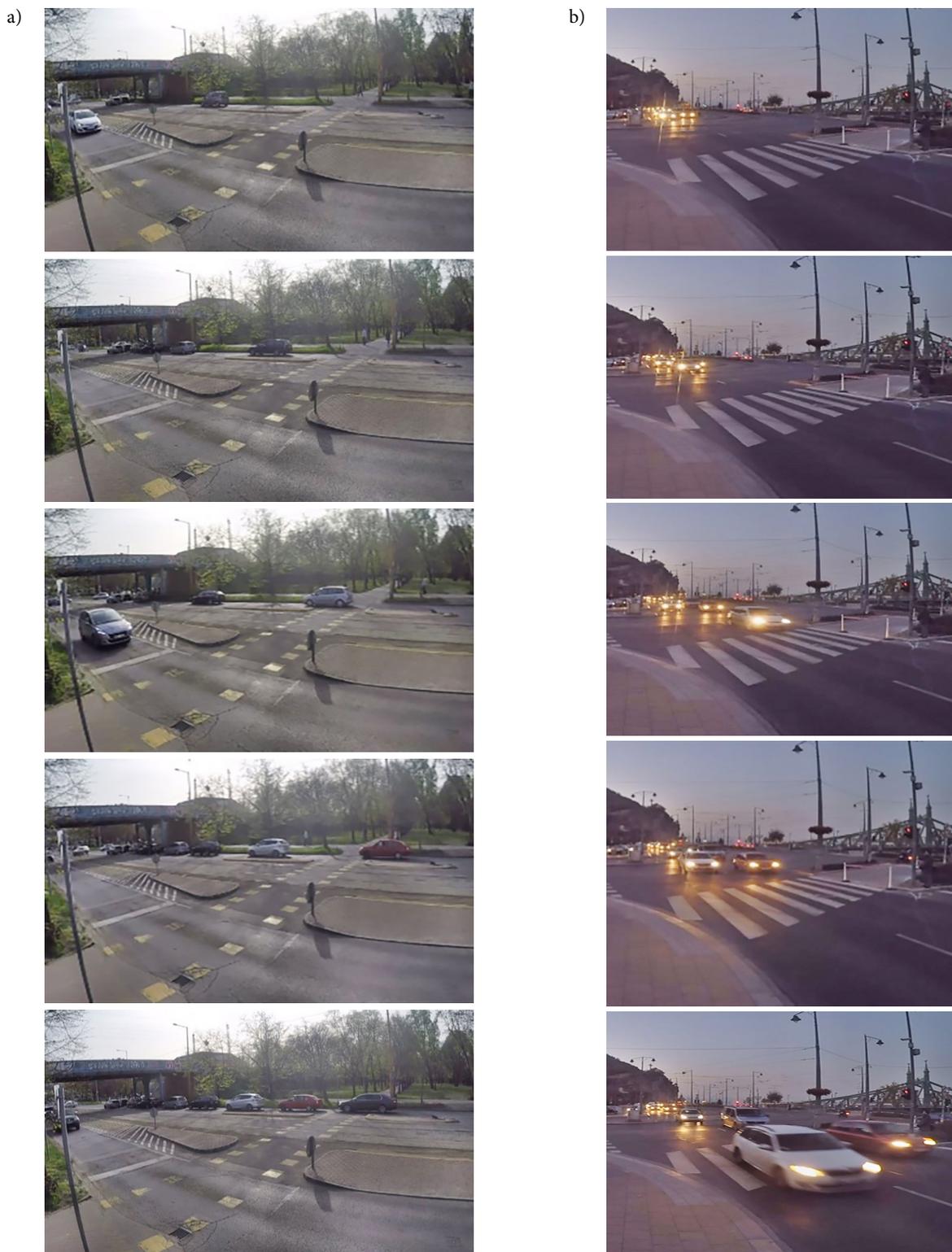


Figure 4. Queue: a – forming, representing deceleration in traffic speed (Hamzsabegi Road);  
 b – dissipating, representing acceleration in traffic speed (Szent Gellert Road)

### 1.3. Fitting distributions to the estimated speeds

Speed distribution studies provide an insight into the aggregate flow of vehicles, which have important applications in lots of issues such as kinematical traffic simulation model, road design, speed limit evaluation, road traffic

noise prediction, traffic safety evaluation, bicycle performance evaluation, analysis of pedestrian walking, capacity estimation, Level of Service analysis, Safety analysis, bus operation analysis etc. (Iannone *et al.* 2013; Berry, Belmont 1951; Lin *et al.* 2008; Maghrour Zefreh, Török 2018b; Wang *et al.* 2015; Vadeby, Forsman 2016, 2017;

Maurya *et al.* 2015, 2016; Bassani *et al.* 2016; Hustim, Fujimoto 2012; Chandra, Bharti 2013; Du *et al.* 2017). The main aim of the current research is to investigate the variations of traffic speed in different traffic conditions in urban roads. This assumption is supported by Jun (2010) where the change between different speed distributions would show the pattern of traffic variations of a roadway system.

In this paper, distribution of the speeds within the desired time intervals (extracted traffic conditions from the recorded videos) is investigated by the Maximum Likelihood Estimation (MLE) method using “fitdistrplus” package in R programming language software (Delignette-Muller, Dutang 2015). The entire procedure is explained in details in forthcoming subsections.

### 1.3.1. Choice of candidate distributions

Before fitting one or more distributions to a data set, it is generally necessary to choose good candidates among a predefined set of distributions. The first attempt in choosing the candidate distributions for our set of data (vehicles speeds in different traffic conditions) was done by plotting the histogram and empirical distribution function of the speeds within the desired intervals based on Equation (2):

$$F_n(x) = \frac{1}{n} \cdot \sum_{i=1}^n I\{X_i \leq x\}, \quad (2)$$

where:  $X_1, \dots, X_n$  are independent and identically distributed random variables with Cumulative Distribution Function (CDF)  $F(x) = P(X_1 \leq x)$ ;  $I$  is the indicator function, namely,  $I\{X_i \leq x\}$  is 1 if  $X_i \leq x$  and 0 otherwise.

In addition to empirical plots, descriptive statistics of the data set (speeds) would help to choose candidate distributions among a set of parametric distributions particularly the skewness  $sk$  and kurtosis  $kr$  parameters linked to the third and fourth moments. The skewness and kurtosis from a sample  $(X_i)_i \stackrel{i.i.d.}{\sim} X$  with observations  $(x_i)_i$  is given by Casella and Berger (2002):

$$sk(X) = \frac{E\left(\left(X - E(X)\right)^3\right)}{\text{Var}(X)^{3/2}},$$

$$\hat{sk} = \frac{\sqrt{n \cdot (n-1)}}{n-2} \cdot \frac{m_3}{m_2^{3/2}}; \quad (3)$$

$$kr(X) = \frac{E\left(\left(X - E(X)\right)^4\right)}{\text{Var}(X)^2},$$

$$\hat{kr} = \frac{n-1}{(n-2) \cdot (n-3)} \times \left( (n+1) \cdot \frac{m_4}{m_2^2} - 3 \cdot (n-1) \right) + 3, \quad (4)$$

where:  $m_2, m_3, m_4$  denote empirical moments defined by  $m_k = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^k$  with  $x_i$  the  $n$  observations of variable  $x$  and  $\bar{x}$  their mean value.

The estimated skewness and kurtosis parameters of the empirical distribution are further investigated by a skewness–kurtosis plot to choose candidates in order to describe a distribution among a set of parametric distributions. The plot shows the locus of skewness–kurtosis pairs that the distribution can take by varying its parameter values. In fact, this plot shows the possible range of skewness–kurtosis combination a distribution can have. For instance, this combination can be a constant value (e.g. normal distribution with skewness of 0 and kurtosis of 3). It can also form a curve if the equation for estimating the skewness and kurtosis would be dependent on a single distribution parameter (e.g. gamma distribution). Skewness–kurtosis combination can further lie on a two-dimensional surface if the equation for estimating the skewness and kurtosis would be dependent on more than one distribution parameters (e.g. beta distribution). It should be noted that, for any distribution, kurtosis has to be greater or equal to the square of skewness plus one. The values lower than this threshold would be placed in the so-called impossible region where no distribution can fall in.

### 1.3.2. Fit of distributions by MLE method

Once selected the parametric distributions as the candidates, the distribution parameters  $\theta$  would be estimated by maximizing the likelihood function defined as:

$$L(\theta) = \prod_{i=1}^n f(x_i | \theta), \quad (5)$$

where:  $x_i$  is observed traffic speed;  $f(\cdot | \theta)$  is density function of the candidate parametric distribution.

The investigated parametric distributions fitted to the empirical distributions are as follows:

– normal:

$$f_n(x; \alpha, \mu) = \frac{\exp\left(-\frac{1}{2} \cdot \left(\frac{x - \mu}{\alpha}\right)^2\right)}{\alpha \cdot \sqrt{2 \cdot \pi}}; \quad (6)$$

– lognormal:

$$f_{ln}(x; \alpha, \mu) = \frac{1}{x \cdot \alpha \cdot \sqrt{2 \cdot \pi}} \times \exp\left(-\frac{(\ln(x) - \mu)^2}{2 \cdot \alpha^2}\right); \quad (7)$$

– exponential:

$$f_{ex}(x; \lambda) = \lambda \cdot \exp(-\lambda \cdot x); \quad (8)$$

– uniform:

$$f_u(x) = \frac{1}{B - A}; \quad (9)$$

– logistic:

$$f_{lo}(x; \alpha, \mu) = \frac{\exp\left(-\left(\frac{x - \mu}{\alpha}\right)\right)}{\alpha \cdot \left(1 + \exp\left(-\left(\frac{x - \mu}{\alpha}\right)\right)\right)^2}; \quad (10)$$

– beta:

$$f_b(x; k, \omega) = \frac{x^{k-1} \cdot (1-x)^{\omega-1}}{\eta(k, \omega)},$$

$$\eta(k, \omega) = \frac{\Gamma(k) \cdot \Gamma(\omega)}{\Gamma(k + \omega)}; \quad (11)$$

– gamma:

$$f_g(x; \alpha, k) = \frac{\alpha^k}{\Gamma(k)} \cdot x^{k-1} \cdot \exp(-\alpha \cdot x); \quad (12)$$

– Weibull:

$$f_w(x; \alpha, k) = \frac{k}{\alpha} \cdot \left(\frac{x}{\alpha}\right)^{k-1} \cdot \exp\left(-\left(\frac{x}{\alpha}\right)^k\right); \quad (13)$$

– chi-square:

$$f_{ch}(x; \tau) = \frac{1}{2^{\frac{\tau}{2}} \cdot \Gamma\left(\frac{\tau}{2}\right)} \cdot x^{\frac{\tau}{2}-1} \cdot \exp\left(-\frac{x}{2}\right), \quad (14)$$

where:  $\alpha$  is scale parameter;  $\mu$  is location parameter;  $\lambda$  is rate parameter;  $A$ : Min,  $B$ : Max,  $k$ : shape parameter,  $\omega$  is second shape parameter;  $\tau$  is degrees of freedom parameter;  $\eta(\cdot)$  is beta function;  $\Gamma(\cdot)$  is gamma function.

Having estimated candidate distributions parameters based on Equation (5), the candidate distributions would be fitted to the data set for the possible graphical comparison (goodness-of-fit plots) of the candidates with empirical distribution.

#### 1.4. Compare fitted distributions by goodness-of-fit test

Having estimated different candidate parametric distributions for traffic speed in different traffic conditions based on Equation (5), these different distributions were at first compared graphically and then compared to each other by goodness-of-fit tests in order to find best fitted speed distribution in each traffic condition. The goodness-of-fit statistics aims to measure the distance between the fitted parametric distribution and the empirical distribution (distance between CDFs). In the current research three goodness-of-fit tests (Kolmogorov–Smirnov, Cramér–Von Mises, and Anderson–Darling) are considered based on D’Agostino (2017):

– Kolmogorov–Smirnov (KS):

$$\sup |F_n(x) - F(x)|; \quad (15)$$

– Cramér–Von Mises (CvM):

$$\int_{-\infty}^{\infty} (F_n(x) - F(x))^2 dx; \quad (16)$$

– Anderson–Darling (AD):

$$n \cdot \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x) \cdot (1 - F(x))} dx, \quad (17)$$

where:  $F_n$  is empirical CDF of the vehicles speeds;  $F$  is fitted theoretical parametric CDF.

Apart from these statistics, two other classical penalized criteria based on the log-likelihood are further con-

sidered to tackle the over-fitting problems as follows:

– Akaike Information Criterion (AIC):

$$AIC = 2 \cdot k - 2 \cdot \ln\left(\hat{L}\right); \quad (18)$$

– Bayesian Information Criterion (BIC):

$$BIC = \ln(n) \cdot k - 2 \cdot \ln\left(\hat{L}\right), \quad (19)$$

where:  $k$  is number of estimated parameters in the model;  $\hat{L}$  is maximum value of the likelihood function for the model;  $n$  is number of observations.

## 2. Case study and results

The main aim of the current research was to study on the variation of the traffic speed in different traffic conditions. To do so, the traffic flow is disaggregated visually by the help of video recording based on predefined traffic conditions (see Section 1.1.1). These traffic conditions, in the current research, are as follows: under-saturated flow, free flow traffic, congestion, over-saturated congestion, acceleration process and deceleration process.

In this section, the proposed methodology is applied to the extracted speeds data (hereinafter sample time interval speeds) of a scenario in which traffic flow followed acceleration process (hereinafter called accelerated flow condition). Before fitting one or more distributions to our speeds data set, it is necessary to choose good candidates among a predefined set of distributions.

### 2.1. Choice of the candidate distributions for the sample time interval speeds in accelerated flow condition

The first attempt in choosing the candidate distributions was done by plotting the histogram and empirical distribution of the speeds based on Equation (2). Figure 5 shows the plotted histogram on the density scale together with the CDF of the sample time interval speeds, where traffic flow followed *acceleration process*.

In addition to empirical plots, the third and fourth moments of the empirical distribution of the sample time interval speeds were estimated based on Equations (3) and (4). Table 2 shows the summary statistics of the previously mentioned sample time interval speeds where traffic flow followed acceleration process.

Having estimated the skewness and kurtosis parameters, a skewness-kurtosis plot of the empirical distribution based on (Cullen, Frey 1999) is further investigated to choose candidates in order to describe a distribution among a set of parametric distributions according to the estimated skewness and kurtosis (last two rows of Table 2). It should be emphasized that the non-zero skewness reveals a lack of symmetry of the empirical distribution, while the kurtosis value quantifies the weight of tails in comparison to the normal distribution for which the kurtosis equals 3.

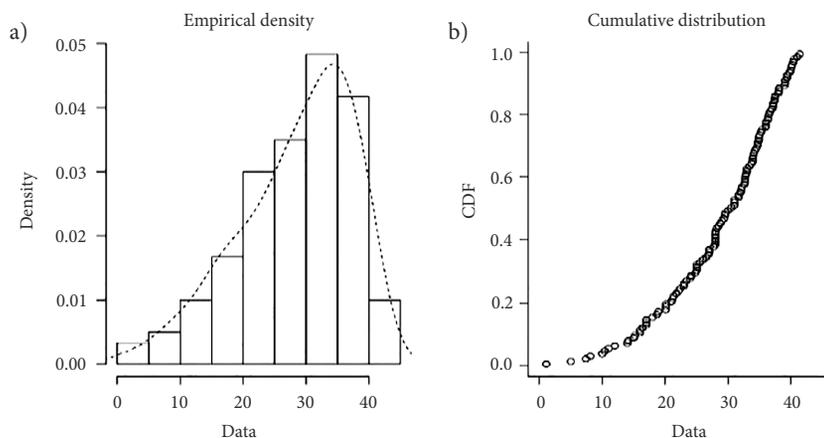


Figure 5. Histogram (a) and CDF (b) plots of an empirical distribution of the sample time interval speeds in accelerated flow condition

Table 2. Summary statistics of the sample time interval speeds in accelerated flow condition [km/h]

Minimum	Maximum	Median	Mean	Estimated SD	Estimated skewness	Estimated kurtosis
1.1	41.4	30.25	28.44799	9.013783	-0.7810547	2.993978

Figure 6 shows the plotted skewness–kurtosis graph of the speeds within a sample time interval (traffic flow followed acceleration) plotted by “fitdistrplus” package in R programming language software (Delignette-Muller, Dutang 2015). In order to take into account the uncertainty of the estimated values of kurtosis and skewness from the calculated speeds, a nonparametric bootstrap procedure is performed. Values of skewness and kurtosis are computed on bootstrap samples (1000 samples) and reported on the skewness–kurtosis plot as shown in Figure 6.

By taking a wide look at Figures 5 and 6 and considering descriptive statistics of the sample time interval speeds (Table 2), normal, lognormal, beta, gamma and Weibull distributions are considered as candidate distributions for further investigations (distribution fitting process) for the mentioned sample time interval speeds.

### 2.2. Fitting candidate distributions by MLE method to the sample time interval speeds in accelerated flow condition

Once selected the parametric candidate distributions (normal, lognormal, beta, gamma and Weibull distributions in this case), their distribution parameters were estimated by MLE method using Equation (5) and the related density functions in order to fit the candidates to the data set for the possible graphical comparison (goodness-of-fit plots) of the candidates with empirical distribution.

The estimated parameters of the candidate distributions for the previously mentioned sample time interval speeds are shown in Table 3 and their related goodness-of-fit plots (density plot, CDF plot, Q–Q plot<sup>1</sup>, P–P plot<sup>2</sup>)

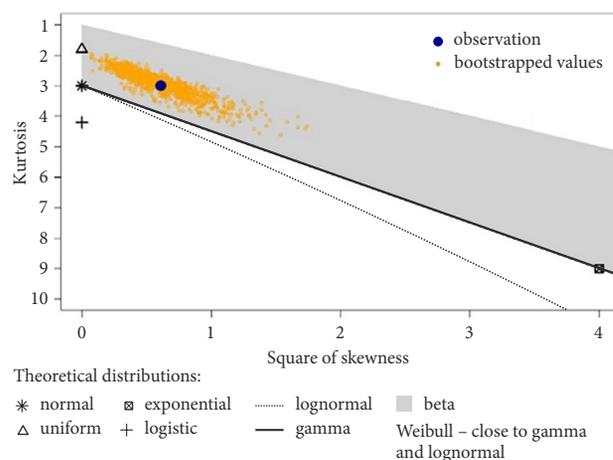


Figure 6. Skewness–kurtosis plot of the sample time interval speeds in accelerated flow condition

are presented in Figure 7. It should be highlighted that, since the beta distribution was among the candidate distributions in this time interval, the speeds were rescaled to (0–1) interval for distribution parameter estimation and distribution fitting.

It should be noted that all four plots of Figure 7 compare the candidate distributions by the empirical distribution of the sample time interval speeds in some aspects. For instance, the density plot represents the density function of the fitted distribution along with the histogram of the empirical distribution of the speeds. Apart from the two basic classical goodness-of-fit plots (density plot and CDF plot), the Q–Q plot emphasizes the lack-of-fit at the distribution tails while the P–P plot emphasizes the lack-of-fit at the distribution centre. Taking the Q–Q plot of the sample time interval speeds into account, one can simply find out that the beta and normal distributions describe the tails of empirical distribution better though the beta

<sup>1</sup> emphasizes the lack of fit at the distribution tails in candidate – empirical distributions comparison;

<sup>2</sup> emphasizes the lack of fit at the distribution centre in candidate – empirical distributions comparison.

Table 3. Estimated parameters for the sample time interval speeds in accelerated flow condition

<i>Weibull distribution</i>				
	Estimate	Std. error	Shape*	Scale*
Shape	3.6984370	0.28826273	1.0000000	0.2712543
Scale	0.7588289	0.01945914	0.2712543	1.0000000
Log-likelihood	11.40879	–	–	–
AIC	–18.81758	–	–	–
BIC	–13.24259	–	–	–
<i>Normal distribution</i>				
	Estimate	Std. error	Mean*	SD*
Mean	0.6871479	0.01979238	1	0
SD	0.2168146	0.01399398	0	1
Log-likelihood	13.17288	–	–	–
AIC	–22.34576	–	–	–
BIC	–16.77078	–	–	–
<i>Lognormal distribution</i>				
	Estimate	Std. error	Mean-log*	SD-log*
Mean-log	–0.4571250	0.04453794	1	0
SD-log	0.4878887	0.03149249	0	1
Log-likelihood	–29.29747	–	–	–
AIC	62.59494	–	–	–
BIC	68.16992	–	–	–
<i>Gamma distribution</i>				
	Estimate	Std. error	Shape*	Rate*
Shape	6.265413	0.7882665	1.0000000	0.9604612
Rate	9.117955	1.1943756	0.9604612	1.0000000
Log-likelihood	–8.500691	–	–	–
AIC	21.00138	–	–	–
BIC	26.57637	–	–	–
<i>Beta distribution</i>				
	Estimate	Std. error	Shape 1*	Shape 2*
Shape 1	2.236750	0.2906532	1.0000000	0.7177342
Shape 2	1.027871	0.1185741	0.7177342	1.0000000
Log-likelihood	28.81901	–	–	–
AIC	–53.63801	–	–	–
BIC	–48.06303	–	–	–

Note: \* – values represent the correlation matrix values.

distribution could be preferred for its better description of the empirical distribution centre considering the related P–P plot.

### 2.3. Goodness-of-fit test comparison of the sample time interval speeds in accelerated flow condition

Having compared the candidate distributions to the empirical distribution graphically, they were further compared to each other by three goodness-of-fit tests (KS, CvM, AD) and two penalized criteria (AIC and BIC) based on the Equations (15)–(19) to find the best possible fitted distribution to that traffic condition (acceleration process in this case). The computed values of these three goodness-of-fit statistics and two classical penalized criteria based on the log likelihood for the fitted distributions to the sample time interval speeds in accelerated flow condition are given in Table 4.

As previously mentioned, the main aim of the Goodness-of-fit tests is to measure the distance between the fit-

ted parametric distribution and the empirical distribution. Therefore the lower parameter in Table 4, the better fitted distribution. Taking the outputs of the goodness-of-fit statistics into account (Table 4), beta distribution would be considered as the best fitted distribution to the case study sample time interval speeds where traffic flow followed acceleration process. The situation in which traffic speed follows different distributions is quite recognized in the literature. For instance, Leong (1968) and McLean (1979) found that, for lightly trafficked two-lane roads where most vehicles are traveling freely, car speeds measured in time are approximately normally distributed with a coefficient of variation ranging from about 0:11...0:18. In addition, Minh *et al.* (2005) have studied that the speed distribution followed the normal distribution on the urban road. Wang *et al.* (2012) introduced truncated normal and lognormal distribution for modelling speeds and travel time. Zou (2013) proposed that skew-*t* distribution can reasonably take into account the heterogeneity in vehicle speed data.

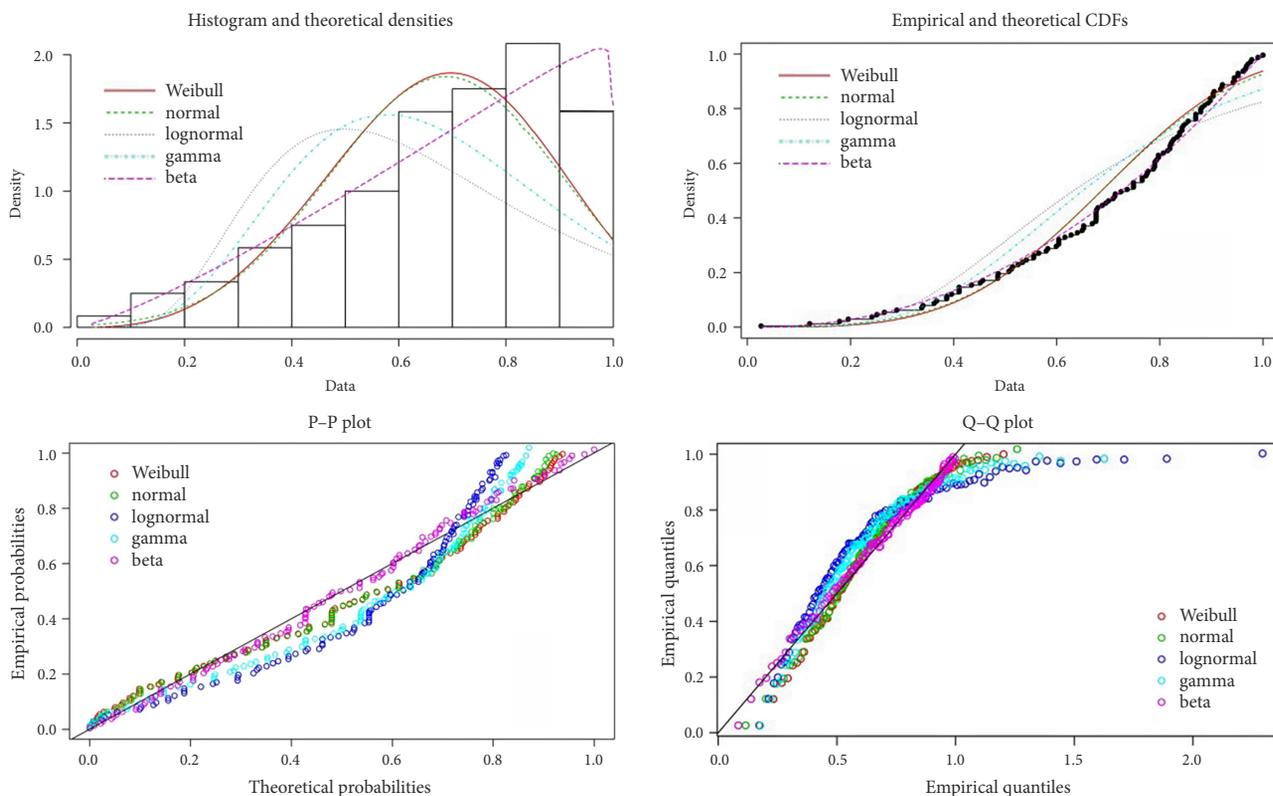


Figure 7. Goodness-of-fit plots of the candidate distributions fitted to the sample time interval speeds in accelerated flow condition

Table 4. Comparison of goodness-of-fit results of the sample time interval speeds in accelerated flow condition

Goodness-of-fit tests	Weibull	Normal	Lognormal	Gamma	Beta
KS	0.1110469	0.1081006	0.1743946	0.1537278	0.05633537
CvM	0.3430385	0.3299080	1.2681188	0.8381365	0.06010820
AD	2.1876732	2.0322768	7.2177582	4.8178281	0.46603657
AIC	-18.81758	-22.34576	62.59494	21.00138	-53.63801
BIC	-13.24259	-16.77078	68.16992	26.57637	-48.06303

Zou and Zhang (2011) said that a single normal distribution cannot accurately accommodate the excess kurtosis present in the speed distribution and they proposed skew-normal and skew-*t* distribution to fit speed data. Haight and Mosher (1962) considered that the speed data could be well represented by either a gamma or a lognormal distribution. Gerlough and Huber (1975) proposed the use of the lognormal distribution. This resembles the normal distribution but is skewed with a larger tail to the right. It offers the advantage that the same functional form is retained when the time speed distribution is transformed into a space-speed distribution and avoids the theoretical difficulty of the negative speeds given by the infinite tails of the normal distribution. This assumption is supported by IMAGINE (2006), where different speed distributions are related to different traffic conditions. Recently literature has remarked that there is a distribution of speed over each level of density in traffic flow, which might not necessarily be a normal distribution (Qu *et al.* 2017). The results

of the distribution fitting process of the current research in the urban road traffic (interrupted traffic flow) show the fact that traffic speed in the urban roads might follow different distributions taking different traffic conditions into account.

The best-fitted distributions to the traffic speed in different traffic conditions (results of the distribution fitting process for all of the defined traffic conditions) defined in Section 1.1.1 along with the statistical specifications of the traffic speed in different traffic conditions are shown in Table 5.

It should be noted that the minimum boundary of the acceleration process is considered as the situation in which the vehicles are almost stopped and they are ready to increase their speed. Moreover, the minimum boundary of the deceleration process is considered the situation in which the vehicles are decreasing their speed until the time that they are almost stopped (reaching the  $x_j$  in Figure 3).

Table 5. Speed variations in different traffic conditions

Traffic condition	Fitted speed distribution	Minimum speed [km/h]	Maximum speed [km/h]	Average speed [km/h]	SD of speed [km/h]
Over-saturated congestion	exponential	1	15.5	10.37	2.74
Under saturated flow	normal	25	55	38.14	6.46
Free flow	lognormal	35	55	47.18	5.5
Congestion	gamma	10	25	14.53	3.91
Accelerated flow	beta	1	45	27.41	13.15
Decelerated flow	chi-square	1	35	18.76	11.82

Furthermore, it is worth noting that in the current study, the possible presence of the HGVs in the urban roads has not been considered since the Heavy Goods Vehicles (HGVs) are prohibited to enter into Budapest since 2016 (neither day nor night). Hence, the share of HGVs has not been integrated into the approach.

#### 2.4. Sensitivity analysis

In this section, the sensitivity analysis in distribution fitting has been performed considering different levels of errors in speed estimation in under saturated flow, free flow, decelerated flow, accelerated flow, over-saturated congestion and congestion respectively. Tables 6–11 show the results of the sensitivity analysis considering different levels of error in speed estimation in different traffic conditions. By taking a wide look at the final results of the sensitivity analysis in under saturated flow condition (Table 6), it is evident that normal distribution is still the best-fitted distribution in under saturated traffic flow (the lower the values of goodness-of-fit tests the better fitted distribution). This is also the case for lognormal distribution in free flow condition (see the values in Table 7), beta distribution in accelerated flow condition (see the values in Table 9), exponential distribution in over-saturated congestion (see the values in Table 10) and gamma distribution in congestion (see the values in Table 11). It should be noted that the results of the sensitivity analysis in decelerated flow condition showed the fact that the lognormal distribution is better fitted than chi-square distribution considering +30% and –20% error level in speed estimation (pay attention to the values in +30% and –20% error rows in Table 8 and find the lowest value!) however the chi-square distribution is the best-fitted in the other levels.

#### Conclusions

Fundamental diagram, a graphical representation of the relationship among traffic flow, speed, and density, has been the foundation of traffic flow theory and transportation engineering for many years. Underlying a fundamental diagram is the relation between traffic speed and density, which roughly corresponds to drivers' speed choices under varying car-following distances. Empirical observations show a wide-scattering of traffic speeds over a certain level of density, which would form a distribution of speed over a certain level of density (see this scattering in Figure 2).

Literature often stated that these distributions in highways, where traffic flow is uninterrupted, would follow the normal distribution (Wang *et al.* 2013). The conditions that turn to a different speed distribution are quite often realized in urban roads, where, in general, the traffic stream is much more complicated.

The main aim of the current research was to investigate the distribution of the traffic speed in urban roads in different traffic conditions. To do so, the distribution of traffic speeds in various locations in city of Budapest (Hungary) has been examined using the recorded videos and the outputs of loop detectors in the investigation sites. It observed that the speed of the traffic flow followed exponential, normal, lognormal, gamma, beta and chi-square distribution in over-saturated congestion, under-saturated flow, free flow, congestion, accelerated flow and decelerated flow scenarios respectively.

Apart from distribution fitting analysis, the sensitivity analysis has been performed in the current study to investigate the effect of potential errors in speed estimation by loop detectors in the final proposed distributions. The results of the sensitivity analysis showed that, taking the +30% and –20% error level in speed estimation by loop detectors into account, the best-fitted distribution to the decelerated traffic flow would be changed from chi-square distribution to lognormal distribution (pay attention to the +30% and –20% error levels in Table 8 and find the lowest value!).

#### Acknowledgements

The research reported in this paper was supported by the Higher Education Excellence Program of the Ministry of Human Capacities in the frame of Artificial Intelligence research area of Budapest University of Technology and Economics (BME FIKP-MI/FM).

Moreover, the authors are grateful for the support of Hungarian Academy of Science (HAS) for providing the Janos BOLAYI scholarship.

The first author further gratefully acknowledges the support of *Stipendium Hungaricum* of Tempus Public Foundation.

#### Disclosure statement

Authors declare they do not have any competing financial, professional, or personal interests from other parties.

Table 6. Comparison of goodness-of-fit results of the sample speeds [km/h] in under saturated flow considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Normal	Lognormal	Gamma	Uniform	Error level
KS	0.05504001	0.04283302	0.05084550	0.04434948	0.1492870	+5%
CvM	0.07640241	0.02727461	0.04373036	0.02922211	0.8036464	
AD	0.54684425	0.18708341	0.31058764	0.20827640	inf	
AIC	777.7321	773.2594	775.5032	773.8661	n/a	
BIC	783.2565	778.7837	781.0275	779.3905	n/a	
KS	0.05499471	0.04283302	0.05084550	0.04440133	0.1492870	+10%
CvM	0.07625297	0.02727461	0.04373036	0.02929997	0.8036464	
AD	0.54639501	0.18708341	0.31058764	0.20862067	inf	
AIC	788.6178	784.1450	786.3888	784.7518	n/a	
BIC	794.1422	789.6694	791.9132	790.2762	n/a	
KS	0.05503531	0.04283302	0.05084550	0.04431115	0.1492870	+15%
CvM	0.07636038	0.02727461	0.04373036	0.02916014	0.8036464	
AD	0.54673476	0.18708341	0.31058764	0.20787495	inf	
AIC	799.0195	794.5468	796.7905	795.1535	n/a	
BIC	804.5439	800.0711	802.3149	800.6779	n/a	
KS	0.05500568	0.04283302	0.05084550	0.04437475	0.1492870	+20%
CvM	0.07626643	0.02727461	0.04373036	0.02925770	0.8036464	
AD	0.54645018	0.18708341	0.31058764	0.20836851	inf	
AIC	808.9785	804.5057	806.7495	805.1125	n/a	
BIC	814.5028	810.0301	812.2738	810.6368	n/a	
KS	0.05509182	0.04283302	0.05084550	0.04429315	0.1492870	-5%
CvM	0.07654561	0.02727461	0.04373036	0.02913593	0.8036464	
AD	0.54729315	0.18708341	0.31058764	0.20784474	inf	
AIC	754.3126	749.8398	752.0836	750.4466	n/a	
BIC	759.8370	755.3642	757.6080	755.9710	n/a	
KS	0.05492876	0.04283302	0.05084550	0.04432825	0.1492870	-10%
CvM	0.07614922	0.02727461	0.04373036	0.02918784	0.8036464	
AD	0.54601521	0.18708341	0.31058764	0.20805714	inf	
AIC	741.6609	737.1881	739.4319	737.7949	n/a	
BIC	747.1853	742.7124	744.9562	743.3192	n/a	
KS	0.05501837	0.04283302	0.05084550	0.04436164	0.1492870	-15%
CvM	0.07633856	0.02727461	0.04373036	0.02923681	0.8036464	
AD	0.54664741	0.18708341	0.31058764	0.20824154	inf	
AIC	728.2858	723.8130	726.0568	724.4198	n/a	
BIC	733.8102	729.3374	731.5812	729.9442	n/a	
KS	0.05509583	0.04283302	0.05084550	0.04432034	0.1492870	-20%
CvM	0.07654531	0.02727461	0.04373036	0.02917514	0.8036464	
AD	0.54730092	0.18708341	0.31058764	0.20797707	inf	
AIC	714.0996	709.6269	711.8707	710.2337	n/a	
BIC	719.6240	715.1512	717.3950	715.7580	n/a	

Table 7. Comparison of goodness-of-fit results of the sample speeds [km/h] in free flow considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Normal	Lognormal	Gamma	Uniform	Error level
KS	0.09705827	0.07834848	0.05738510	0.06447334	0.1660802	+5%
CvM	0.17364430	0.06128168	0.03939531	0.04246622	0.8325637	
AD	1.14189121	0.35862117	0.22840086	0.24244124	inf	
AIC	698.7769	688.6895	687.4441	687.4508	n/a	
BIC	704.1779	694.0905	692.8451	692.8517	n/a	
KS	0.09713114	0.07834848	0.05738510	0.06464948	0.1660802	+10%
CvM	0.17394637	0.06128168	0.03939531	0.04253815	0.8325637	
AD	1.14278860	0.35862117	0.22840086	0.24249715	inf	
AIC	709.0113	698.9239	697.6785	697.6852	n/a	
BIC	714.4123	704.3249	703.0795	703.0861	n/a	
KS	0.09717136	0.07834848	0.05738510	0.06456792	0.1660802	+15%
CvM	0.17420127	0.06128168	0.03939531	0.04254718	0.8325637	
AD	1.14356743	0.35862117	0.22840086	0.24276696	inf	
AIC	718.7907	708.7033	707.4579	707.4646	n/a	
BIC	724.1917	714.1042	712.8589	712.8655	n/a	
KS	0.09701301	0.07834848	0.05738510	0.06457777	0.1660802	+20%
CvM	0.17349244	0.06128168	0.03939531	0.04255825	0.8325637	
AD	1.14144892	0.35862117	0.22840086	0.24281917	inf	
AIC	728.1538	718.0664	716.821	716.8277	n/a	
BIC	733.5548	723.4674	722.222	722.2286	n/a	
KS	0.09711641	0.07834848	0.05738510	0.06452301	0.1660802	-5%
CvM	0.17387197	0.06128168	0.03939531	0.04251763	0.8325637	
AD	1.14256438	0.35862117	0.22840086	0.24267438	inf	
AIC	676.7586	666.6711	665.4258	665.4324	n/a	
BIC	682.1595	672.0721	670.8267	670.8334	n/a	
KS	0.0970845	0.07834848	0.05738510	0.06454240	0.1660802	-10%
CvM	0.1738569	0.06128168	0.03939531	0.04251456	0.8325637	
AD	1.1425470	0.35862117	0.22840086	0.24260415	inf	
AIC	664.8638	654.7763	653.5310	653.5376	n/a	
BIC	670.2647	660.1773	658.9319	658.9386	n/a	
KS	0.09769744	0.07834848	0.05738510	0.06457408	0.1660802	-15 %
CvM	0.17641785	0.06128168	0.03939531	0.04255090	0.8325637	
AD	1.15020960	0.35862117	0.22840086	0.24277725	inf	
AIC	652.2892	642.2015	640.9561	640.9628	n/a	
BIC	657.6901	647.6025	646.3571	646.3637	n/a	
KS	0.09707671	0.07834848	0.05738510	0.06453655	0.1660802	-20%
CvM	0.17369209	0.06128168	0.03939531	0.04251853	0.8325637	
AD	1.14202651	0.35862117	0.22840086	0.24264652	inf	
AIC	638.9515	628.8641	627.6187	627.6254	n/a	
BIC	644.3525	634.2650	633.0197	633.0263	n/a	

Table 8. Comparison of Goodness-of-fit results of the sample speeds [km/h] in decelerated flow considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Normal	Lognormal	Chi-square	Logistic	Error level
KS	0.07017497	0.07159481	0.04944783	0.03871958	0.05200988	+5%
CvM	0.08200117	0.08189015	0.05792295	0.02048399	0.05390933	
AD	0.52790972	0.52451279	0.43051457	0.17394709	0.46953517	
AIC	753.9579	754.6487	753.7931	748.2034	757.5318	
BIC	759.4822	760.1730	759.3175	750.9656	763.0561	
KS	0.07017885	0.07159481	0.04944783	0.04293905	0.05200988	+10%
CvM	0.08199925	0.08189015	0.05792295	0.02258951	0.05390933	
AD	0.52789862	0.52451279	0.43051457	0.20574382	0.46953517	
AIC	764.8435	765.5344	764.6788	759.1470	768.4175	
BIC	770.3679	771.0587	770.2032	761.9091	773.9418	
KS	0.07016279	0.07159481	0.04944783	0.04688528	0.05200988	+15%
CvM	0.08198816	0.08189015	0.05792295	0.02805678	0.05390933	
AD	0.52787995	0.52451279	0.43051457	0.27014133	0.46953517	
AIC	775.2453	775.9361	775.0805	769.8442	778.8192	
BIC	780.7696	781.4604	780.6049	772.6063	784.3435	
KS	0.07016066	0.07159481	0.04944783	0.05056152	0.05200988	+20%
CvM	0.08199170	0.08189015	0.05792295	0.03648743	0.05390933	
AD	0.52789471	0.52451279	0.43051457	0.36450384	0.46953517	
AIC	785.2042	785.8950	785.0395	780.3162	788.7781	
BIC	790.7286	791.4194	790.5638	783.0784	794.3025	
KS	0.07013587	0.07159481	0.04944783	0.05727033	0.05200988	+30%
CvM	0.08190228	0.08189015	0.05792295	0.06043900	0.05390933	
AD	0.52761898	0.52451279	0.43051457	0.63209731	0.46953517	
AIC	803.9342	804.6250	803.7695	800.6567	807.5081	
BIC	809.4585	810.1494	809.2938	803.4189	813.0325	
KS	0.07013940	0.07159481	0.04944783	0.03879238	0.05200988	-5%
CvM	0.08191911	0.08189015	0.05792295	0.02861514	0.05390933	
AD	0.52767222	0.52451279	0.43051457	0.22170578	0.46953517	
AIC	730.5383	731.2292	730.3736	725.4766	734.1123	
BIC	736.0627	736.7535	735.8979	728.2388	739.6366	
KS	0.07024338	0.07159481	0.04944783	0.04327660	0.05200988	-10%
CvM	0.08217810	0.08189015	0.05792295	0.04021206	0.05390933	
AD	0.52843258	0.52451279	0.43051457	0.30919795	0.46953517	
AIC	717.8866	718.5774	717.7219	713.6326	721.4605	
BIC	723.4110	724.1018	723.2462	716.3948	726.9849	
KS	0.07008660	0.07159481	0.04944783	0.04790153	0.05200988	-15%
CvM	0.08171569	0.08189015	0.05792295	0.05799990	0.05390933	
AD	0.52704242	0.52451279	0.43051457	0.44554787	0.46953517	
AIC	704.5115	705.2024	704.3468	701.4191	708.0855	
BIC	710.0359	710.7267	709.8712	704.1813	713.6098	
KS	0.07010143	0.07159481	0.04944783	0.05268503	0.05200988	-20%
CvM	0.08183545	0.08189015	0.05792295	0.08306093	0.05390933	
AD	0.52743270	0.52451279	0.43051457	0.63651677	0.46953517	
AIC	690.3254	691.0162	690.1606	688.7906	693.8993	
BIC	695.8497	696.5405	695.6850	691.5528	699.4237	

Table 9. Comparison of Goodness-of-fit results of the sample speeds [km/h] in accelerated flow considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Normal	Lognormal	Gamma	Beta	Error level
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06305521	+5%
CvM	0.3430218	0.3299080	1.2681188	0.8381228	0.07394376	
AD	2.1875908	2.0322768	7.2177582	4.8177793	0.58546846	
AIC	-18.81705	-22.34524	62.59546	21.00191	-53.58349	
BIC	-13.24207	-16.77025	68.17044	26.57689	-48.00851	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06318114	+10%
CvM	0.3430218	0.3299080	1.2681188	0.8381227	0.07421670	
AD	2.1875907	2.0322768	7.2177582	4.8177790	0.58777253	
AIC	-18.81705	-22.34523	62.59546	21.00191	-53.58391	
BIC	-13.24207	-16.77025	68.17045	26.57689	-48.00892	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06333252	+15%
CvM	0.3430218	0.3299080	1.2681188	0.8381227	0.07457412	
AD	2.1875907	2.0322768	7.2177582	4.8177788	0.59062494	
AIC	-18.81705	-22.34523	62.59547	21.00191	-53.58435	
BIC	-13.24206	-16.77025	68.17045	26.57690	-48.00937	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06342779	+20%
CvM	0.3430217	0.3299080	1.2681188	0.8381226	0.07481655	
AD	2.1875907	2.0322768	7.2177582	4.8177786	0.59264767	
AIC	-18.81705	-22.34523	62.59547	21.00191	-53.58483	
BIC	-13.24206	-16.77025	68.17045	26.57690	-48.00984	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06277775	-5%
CvM	0.3430218	0.3299080	1.2681188	0.8381230	0.07326392	
AD	2.1875908	2.0322768	7.2177582	4.8177798	0.57992461	
AIC	-18.81706	-22.34524	62.59546	21.00190	-53.58280	
BIC	-13.24207	-16.77026	68.17044	26.57689	-48.00782	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06261044	-10%
CvM	0.3430218	0.3299080	1.2681188	0.8381231	0.07287825	
AD	2.1875908	2.0322768	7.2177582	4.8177801	0.57677902	
AIC	-18.81706	-22.34525	62.59545	21.00190	-53.58254	
BIC	-13.24208	-16.77026	68.17044	26.57688	-48.00755	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06246411	-15%
CvM	0.3430218	0.3299080	1.2681188	0.8381231	0.07255211	
AD	2.1875909	2.0322768	7.2177582	4.8177805	0.57401147	
AIC	-18.81707	-22.34525	62.59545	21.00189	-53.58234	
BIC	-13.24208	-16.77027	68.17043	26.57688	-48.00736	
KS	0.1110442	0.1081006	0.1743946	0.1537265	0.06225249	-20%
CvM	0.3430218	0.3299080	1.2681188	0.8381233	0.07209495	
AD	2.1875909	2.0322768	7.2177582	4.8177809	0.57027639	
AIC	-18.81707	-22.34525	62.59544	21.00189	-53.58223	
BIC	-13.24209	-16.77027	68.17043	26.57687	-48.00724	

Table 10. Comparison of goodness-of-fit results of the sample speeds [km/h] in over-saturated congestion considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Lognormal	Gamma	Exponential	Error level
KS	0.1191751	0.1352376	0.1227149	0.1076295	+5%
CvM	0.2231884	0.3314802	0.2325473	0.1913169	
AD	1.6539557	2.2459483	1.7130743	1.6000141	
AIC	481.8644	484.1806	481.7615	484.4919	
BIC	486.8417	489.1579	486.7388	486.9805	
KS	0.1191709	0.1352376	0.1227370	0.1076295	+10%
CvM	0.2232515	0.3314802	0.2325796	0.1913170	
AD	1.6543600	2.2459483	1.7133271	1.6000142	
AIC	490.1450	492.4612	490.0421	492.7725	
BIC	495.1222	497.4385	495.0193	495.2611	
KS	0.1192183	0.1352376	0.1227447	0.1076295	+15%
CvM	0.2233920	0.3314802	0.2325964	0.1913169	
AD	1.6552271	2.2459483	1.7134469	1.6000141	
AIC	498.0574	500.3736	497.9545	500.6849	
BIC	503.0347	505.3509	502.9317	503.1735	
KS	0.1191851	0.1352376	0.1227905	0.1076295	+20%
CvM	0.2232403	0.3314802	0.2326136	0.1913169	
AD	1.6542800	2.2459483	1.7136850	1.6000141	
AIC	505.6330	507.9492	505.5301	508.2605	
BIC	510.6103	512.9265	510.5074	510.7491	
KS	0.1192193	0.1352376	0.1227425	0.1076295	-5%
CvM	0.2233885	0.3314802	0.2324279	0.1913170	
AD	1.6552042	2.2459483	1.7124725	1.6000142	
AIC	464.0496	466.3658	463.9466	466.6770	
BIC	469.0268	471.3430	468.9239	469.1657	
KS	0.119168	0.1352376	0.1227722	0.1076295	-10%
CvM	0.223297	0.3314802	0.2326957	0.1913169	
AD	1.654652	2.2459483	1.7141004	1.6000141	
AIC	454.4256	456.7418	454.3227	457.0531	
BIC	459.4029	461.7191	459.2999	459.5417	
KS	0.1192008	0.1352376	0.1227363	0.1076295	-15%
CvM	0.2232869	0.3314802	0.2324706	0.1913170	
AD	1.6545681	2.2459483	1.7126990	1.6000141	
AIC	444.2514	446.5676	444.1485	446.8789	
BIC	449.2287	451.5449	449.1257	449.3675	
KS	0.1192119	0.1352376	0.1227209	0.1076295	-20%
CvM	0.2232823	0.3314802	0.2324812	0.1913170	
AD	1.6545321	2.2459483	1.7127131	1.6000141	
AIC	433.4602	435.7764	433.3573	436.0877	
BIC	438.4375	440.7537	438.3346	438.5763	

Table 11. Comparison of goodness-of-fit results of the sample speeds [km/h] in congestion considering different levels of error in speed estimation

Goodness-of-fit tests	Weibull	Normal	Lognormal	Gamma	Uniform	Error level
KS	0.08003632	0.07312253	0.07517967	0.06268930	0.2401795	+5%
CvM	0.10982133	0.08314372	0.07244533	0.05777517	1.7901477	
AD	0.86457138	0.61695308	0.44767556	0.38458526	inf	
AIC	618.4569	614.1415	609.0544	609.0750	n/a	
BIC	623.8760	619.5606	614.4734	614.4941	n/a	
KS	0.08020147	0.07312253	0.07517967	0.06270179	0.2401795	+10%
CvM	0.11038349	0.08314372	0.07244533	0.05776189	1.7901477	
AD	0.86640945	0.61695308	0.44767556	0.38448749	inf	
AIC	628.7843	624.469	619.3818	619.4024	n/a	
BIC	634.2034	629.888	624.8009	624.8215	n/a	
KS	0.08022457	0.07312253	0.07517967	0.06276694	0.2401795	+15%
CvM	0.11046431	0.08314372	0.07244533	0.05787347	1.7901477	
AD	0.86667554	0.61695308	0.44767556	0.38515794	inf	
AIC	638.6526	634.3373	629.2501	629.2707	n/a	
BIC	644.0717	639.7563	634.6692	634.6898	n/a	
KS	0.08008344	0.07312253	0.07517967	0.06276391	0.2401795	+20%
CvM	0.10998905	0.08314372	0.07244533	0.05782003	1.7901477	
AD	0.86512414	0.61695308	0.44767556	0.38481160	inf	
AIC	648.1009	643.7855	638.6984	638.719	n/a	
BIC	653.5199	649.2046	644.1174	644.138	n/a	
KS	0.08017335	0.07312253	0.07517967	0.06265826	0.2401795	-5%
CvM	0.11025152	0.08314372	0.07244533	0.05774832	1.7901477	
AD	0.86595356	0.61695308	0.44767556	0.38443757	inf	
AIC	596.2384	591.9230	586.8359	586.8565	n/a	
BIC	601.6574	597.3421	592.2549	592.2755	n/a	
KS	0.08009915	0.07312253	0.07517967	0.06276749	0.2401795	-10%
CvM	0.11008298	0.08314372	0.07244533	0.05791712	1.7901477	
AD	0.86545828	0.61695308	0.44767556	0.38544280	inf	
AIC	584.2354	579.9201	574.8329	574.8536	n/a	
BIC	589.6545	585.3392	580.2520	580.2726	n/a	
KS	0.08019808	0.07312253	0.07517967	0.06272397	0.2401795	-15%
CvM	0.11034252	0.08314372	0.07244533	0.05789188	1.7901477	
AD	0.86625584	0.61695308	0.44767556	0.38531653	inf	
AIC	571.5463	567.2309	562.1438	562.1644	n/a	
BIC	576.9653	572.6500	567.5628	567.5834	n/a	
KS	0.0801243	0.07312253	0.07517967	0.06269706	0.2401795	-20%
CvM	0.1101419	0.08314372	0.07244533	0.05784111	1.7901477	
AD	0.8656331	0.61695308	0.44767556	0.38500877	inf	
AIC	558.0876	553.7723	548.6851	548.7057	n/a	
BIC	563.5067	559.1913	554.1042	554.1248	n/a	

## References

- Bassani, M.; Catani, L.; Cirillo, C.; Mutani, G. 2016. Night-time and daytime operating speed distribution in urban arterials, *Transportation Research Part F: Traffic Psychology and Behaviour* 42: 56–69. <https://doi.org/10.1016/j.trf.2016.06.020>
- Bella, F.; Calvi, A. 2013. Effects of simulated day and night driving on the speed differential in tangent–curve transition: a pilot study using driving simulator, *Traffic Injury Prevention* 14(4): 413–423. <https://doi.org/10.1080/15389588.2012.716880>
- Bella, F.; Calvi, A.; D'Amico, F. 2014. Analysis of driver speeds under night driving conditions using a driving simulator, *Journal of Safety Research* 49: 45–52. <https://doi.org/10.1016/j.jsr.2014.02.007>
- Berry, D. S.; Belmont, D. M. 1951. Distribution of vehicle speeds and travel times, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 31 July – 12 August 1950, University of California, CA, US, 589–602.
- Casella, G.; Berger, R. L. 2002. *Statistical Inference*. Cengage. 688 p.
- Castro, M.; Sánchez, J. A.; Vaquero, C. M.; Iglesias, L.; Rodríguez-Solano, R. 2008. Automated GIS-based system for speed estimation and highway safety evaluation, *Journal of Computing in Civil Engineering* 22(5): 325–331. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2008\)22:5\(325\)](https://doi.org/10.1061/(ASCE)0887-3801(2008)22:5(325))
- Chandra, S.; Bharti, A. K. 2013. Speed distribution curves for pedestrians during walking and crossing, *Procedia – Social and Behavioral Sciences* 104: 660–667. <https://doi.org/10.1016/j.sbspro.2013.11.160>
- Cullen, A. C.; Frey, H. C. 1999. *Probabilistic Techniques in Exposure Assessment: a Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Springer. 336 p.
- D'Agostino, R. B. 2017. *Goodness-of-Fit-Techniques*. CRC Press. 576 p.
- Delignette-Muller, M. L.; Dutang, C. 2015. Fitdistrplus: an R package for fitting distributions, *Journal of Statistical Software* 64(4): 1–34. <https://doi.org/10.18637/jss.v064.i04>
- Dey, P. P.; Chandra, S.; Gangopadhaya, S. 2006. Speed distribution curves under mixed traffic conditions, *Journal of Transportation Engineering* 132(6): 475–481. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2006\)132:6\(475\)](https://doi.org/10.1061/(ASCE)0733-947X(2006)132:6(475))
- Du, Y.; Deng, F.; Liao, F.; Ji, Y. 2017. Understanding the distribution characteristics of bus speed based on geocoded data, *Transportation Research Part C: Emerging Technologies* 82: 337–357. <https://doi.org/10.1016/j.trc.2017.07.004>
- Fitzpatrick, K.; Carlson, P. J.; Wooldridge, M. D.; Brewer, M. A. 2000. *Design Factors that Affect Driver Speed on Suburban Arterials*. Report No FHWA/TX-00/1769-3. Texas Department of Transportation, Austin, TX, US. 164 p.
- Gerlough, D. L.; Huber, M. J. 1975. *Traffic Flow Theory: a Monograph*. Transportation Research Board. 222 p.
- Haight, F. A.; Mosher, W. W. 1962. A practical method for improving the accuracy of vehicular speed distribution measurements, *Highway Research Board Bulletin* 341: 92–116.
- Hustim, M.; Fujimoto, K. 2012. Road traffic noise under heterogeneous traffic condition in Makassar city, Indonesia, *Journal of Habitat Engineering and Design* 4(1): 109–118.
- Iannone, G.; Guarnaccia, C.; Quartieri, J. 2013. Speed distribution influence in road traffic noise prediction, *Environmental Engineering and Management Journal* 12(3): 493–501. <https://doi.org/10.30638/eeemj.2013.061>
- IMAGINE. 2006. *Improved Methods for the Assessment of the Generic Impact of Noise in the Environment (IMAGINE)*. Funded under: FP6-Policies 2003–2006. Available from Internet: <https://cordis.europa.eu/project/rcn/73857/factsheet/en>
- Jiang, R.; Wu, Q.-S. 2007. The night driving behavior in a car-following model, *Physica A: Statistical Mechanics and its Applications* 375(1): 297–306. <https://doi.org/10.1016/j.physa.2006.09.011>
- Jun, J. 2010. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic, *Transportation Research Part C: Emerging Technologies* 18(4): 599–610. <https://doi.org/10.1016/j.trc.2009.12.005>
- Leong, H. J. 1968. The distribution and trend of free speeds on two lane two way rural highways in New South Wales, in *4th Australian Road Research Board (ARRB) Conference*, 1968, Melbourne, Australia, 4(1): 791–814.
- Liao, F. 2016. Modeling duration choice in space–time multi-state supernetworks for individual activity–travel scheduling, *Transportation Research Part C: Emerging Technologies* 69: 16–35. <https://doi.org/10.1016/j.trc.2016.05.011>
- Liao, F.; Arentze, T.; Timmermans, H. 2013. Incorporating space–time constraints and activity–travel time profiles in a multi-state supernetwork approach to individual activity–travel scheduling, *Transportation Research Part B: Methodological* 55: 41–58. <https://doi.org/10.1016/j.trb.2013.05.002>
- Lin, S.; He, M.; Tan, Y.; He, M. 2008. Comparison study on operating speeds of electric bicycles and bicycles: experience from field investigation in Kunming, China, *Transportation Research Record: Journal of the Transportation Research Board* 2048: 52–59. <https://doi.org/10.3141/2048-07>
- Llorca, C.; Moreno, A. T.; Lenorzer, A.; Casas, J.; Garcia, A. 2015. Development of a new microscopic passing maneuver model for two-lane rural roads, *Transportation Research Part C: Emerging Technologies* 52: 157–172. <https://doi.org/10.1016/j.trc.2014.06.001>
- Lustri, C. 2010. *Continuum Modelling of Traffic Flow*. University of Sydney, Australia. 15 p. Available from Internet: [http://wp.maths.usyd.edu.au/igs/wp-content/uploads/2014/03/Lustri\\_2010\\_TrafficFlow.pdf](http://wp.maths.usyd.edu.au/igs/wp-content/uploads/2014/03/Lustri_2010_TrafficFlow.pdf)
- Maghrouh Zefreh, M.; Baranyai, D.; Török, Á. 2016. Assessing the possibility of presenting a semi-stochastic speed–density function, *MATEC Web Conference* 81: 04002. <https://doi.org/10.1051/mateconf/20168104002>
- Maghrouh Zefreh, M.; Török, Á. 2018a. Single loop detector data validation and imputation of missing data, *Measurement* 116: 193–198. <https://doi.org/10.1016/j.measurement.2017.10.066>
- Maghrouh Zefreh, M.; Török, Á. 2018b. Theoretical comparison of the effects of different traffic conditions on urban road traffic noise, *Journal of Advanced Transportation* 2018: 7949574. <https://doi.org/10.1155/2018/7949574>
- Maghrouh Zefreh, M.; Török, Á.; Mészáros, F. 2017. Average vehicles length in two-lane urban roads: a case study in Budapest, *Periodica Polytechnica Transportation Engineering* 45(4): 218–222. <https://doi.org/10.3311/PPtr.10744>
- Maurya, A. K.; Dey, S.; Das, S. 2015. Speed and time headway distribution under mixed traffic condition, *Journal of the Eastern Asia Society for Transportation Studies* 11: 1774–1792. <https://doi.org/10.11175/easts.11.1774>
- Maurya, A. K.; Das, S.; Dey, S.; Nama, S. 2016. Study on speed and time-headway distributions on two-lane bidirectional road in heterogeneous traffic condition, *Transportation Research Procedia* 17: 428–437. <https://doi.org/10.1016/j.trpro.2016.11.084>
- McLean, J. R. 1979. Observed speed distributions and rural road traffic operations, in *9th Australian Road Research Board (ARRB) Conference*, 21–25 August 1978, Brisbane, Australia, 9(5): 235–244.

- Minh, C. C.; Sano, K.; Matsumoto, S. 2005. The speed, flow and headway analyses of motorcycle traffic, *Journal of the Eastern Asia Society for Transportation Studies* 6: 1496–1508. <https://doi.org/10.11175/easts.6.1496>
- Park, B.; Schneeberger, J. D. 2003. Microscopic simulation model calibration and validation: case study of vissim simulation model for a coordinated actuated signal system, *Transportation Research Record: Journal of the Transportation Research Board* 1856: 185–192. <https://doi.org/10.3141/1856-20>
- Qu, X.; Zhang, J.; Wang, S. 2017. On the stochastic fundamental diagram for freeway traffic: model development, analytical properties, validation, and extensive applications, *Transportation Research Part B: Methodological* 104: 256–271. <https://doi.org/10.1016/j.trb.2017.07.003>
- Trozzi, C.; Vaccaro, R.; Crocetti, S. 1996. Speed frequency distribution in air pollutants' emissions estimate from road traffic, *Science of the Total Environment* 189–190: 181–185. [https://doi.org/10.1016/0048-9697\(96\)05208-4](https://doi.org/10.1016/0048-9697(96)05208-4)
- Vadeby, A.; Forsman, Å. 2017. Changes in speed distribution: applying aggregated safety effect models to individual vehicle speeds, *Accident Analysis & Prevention* 103: 20–28. <https://doi.org/10.1016/j.aap.2017.03.012>
- Vadeby, A.; Forsman, Å. 2016. Speed distribution and traffic safety measures, in G. Yannis, S. Cohen (Eds.). *Traffic Safety* 4: 161–176. <https://doi.org/10.1002/9781119307853.ch11>
- Wang, D.; Zhou, D.; Jin, S.; Ma, D. 2015. Characteristics of mixed bicycle traffic flow on the conventional bicycle path, in *TRB 94th Annual Meeting Compendium of Papers*, 11–15 January 2015, Washington, DC, US, 1–14.
- Wang, H.; Li, J.; Chen, Q.; Ni, D. 2009. Speed–density relationship: from deterministic to stochastic, in *TRB 88th Annual Meeting Compendium of Papers DVD*, 11–15 January 2009, Washington, DC, US, 1–20.
- Wang, H.; Ni, D.; Chen, Q.-Y.; Li, J. 2013. Stochastic modeling of the equilibrium speed–density relationship, *Journal of Advanced Transportation* 47(1): 126–150. <https://doi.org/10.1002/atr.172>
- Wang, Y.; Dong, W.; Zhang, L.; Chin, D.; Papageorgiou, M.; Rose, G.; Young, W. 2012. Speed modeling and travel time estimation based on truncated normal and lognormal distributions, *Transportation Research Record: Journal of the Transportation Research Board* 2315: 66–72. <https://doi.org/10.3141/2315-07>
- Yu, R.; Abdel-Aty, M. 2014. An optimal variable speed limits system to ameliorate traffic safety risk, *Transportation Research Part C: Emerging Technologies* 46: 235–246. <https://doi.org/10.1016/j.trc.2014.05.016>
- Zou, Y. 2013. *A Multivariate Analysis of Freeway Speed and Headway Data*. PhD Dissertation. Texas A&M University, TX, US. 122 p. Available from Internet: <https://oaktrust.library.tamu.edu/handle/1969.1/151809>
- Zou, Y.; Zhang, Y. 2011. Use of skew-normal and skew-*t* distributions for mixture modeling of freeway speed data, *Transportation Research Record: Journal of the Transportation Research Board* 2260: 67–75. <https://doi.org/10.3141/2260-08>